Culture

August 2016

Volume 36 Number 2 | ISSN 09650989

Evolving A National Culture Collection To Meet Current Challenges In Microbiology

Julie E. Russell
Head of Public Health England's Culture Collections, Porton Down, Salisbury, SP4 0JG, UK.

For nearly a century the UK National Collection of Type Cultures (NCTC), a Culture Collection of Public Health England, has been a trustworthy source of authentic bacteria for use in scientific studies. Founded in 1920, it is the most long-established collection in the world that was created with the specific purpose of providing microorganisms for other scientists to use. Funded initially by the UK Medical Research Committee, which later became the Medical Research Council (MRC), NCTC is one of four collections of biological resources operated by Public Health England, the other three being the National Collection of Pathogenic Fungi (1947), the European Collection of Authenticated Cell Cultures (1985) and the National Collection of Pathogenic Viruses (2001). The biological resources available from these collections are used worldwide in biomedical research and diagnostic testing. The food and environmental industry also uses NCTC bacteria as controls for tests for pathogens and indicator organisms.

In its early days the collection had a relatively broad remit and included fungi as well as bacteria. However, in 1947 the NCTC Curator, Dr Samuel T Cowan, decided that the collection should focus solely on bacteria of medical and veterinary interest. Non-medical bacterial and fungal cultures were transferred to other institutes with different expertise, leaving approximately 3000 cultures remaining in NCTC. Although the name of the collection refers specifically to 'Type Cultures', which are also known as type strains, i.e. the strains on which the description of a species is based, the collection also includes many reference strains. The collection in 2016 comprises more than 5000 bacteria of clinical importance, including over 900 type strains that are readily available and listed in NCTC's online catalogue, together with many thousands more preserved bacterial strains that have been donated from private collections and will be made available in the future (Figure 1).

The importance of reference strains

Reproducibility is essential in microbiology, so access to a reliable source of reference strains is critical. Reference strains provide a benchmark for results obtained from detecting, characterizing, enumerating, testing and sequencing the genome of other bacteria. They enable harmonization and repeatability within and between

Figure 1 ~ Phylogenetic tree diagram representing bacterial species available from NCTC.

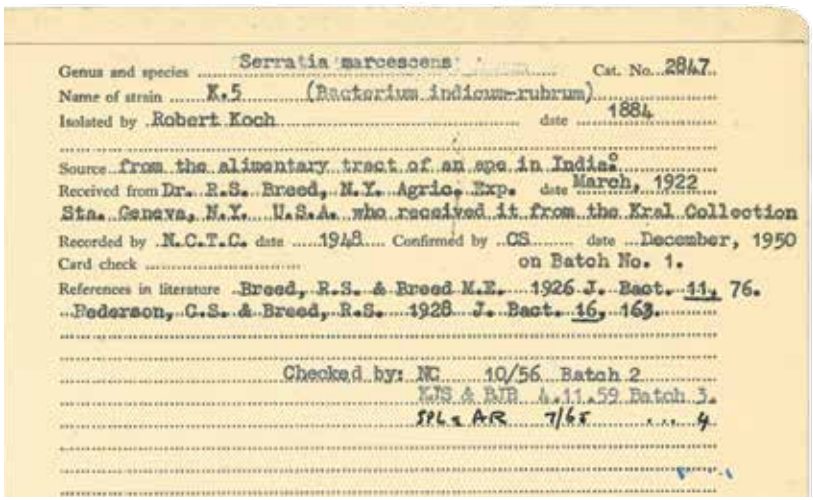

Genus and species Serratia marcescens Cat. No. 2847
Name of strain K.5 (Bacterium indicum-rubrum)
Isolated by Robert Koch date 1884
Source from the alimentary tract of an ape in India.
Received from Dr. R.S. Breed, N.Y. Agric. Exp. date March, 1922
Sta. Geneva, N.Y. U.S.A. who received it from the Kral Collection
Recorded by N.C.T.C. date 1948 Confirmed by CS date December, 1950
Card check on Batch No. 1.
References in literature Breed, R.S. & Breed M.E. 1926 J. Bact. 11, 76.
Pederson, C.S. & Breed, R.S. 1928 J. Bact. 16, 163.
Checked by: NC 10/56 Batch 2
KJS & RJR 1.11.59 Batch 3.
SPL & AR 7/65 4

Figure 2 ~ Quality Control report of an NCTC strain accession in 1948, originally isolated by Robert Koch.

microbiology laboratories on a day-to-day basis and over extended time periods. The reference strain of choice for a particular process may be stipulated in internationally accepted methods. For example, NCTC strains are listed in the European Committee on Antimicrobial Susceptibility Testing (EUCAST) quality control tables that provide guidance on the technical aspects of phenotypic antimicrobial susceptibility testing.

Some laboratories use NCTC reference strains to make their own stocks of control strains which can be cost effective. However, this requires careful management and good attention to detail, particularly with recording data such as the date when the NCTC authenticated strain was initially reconstituted, when the stock cultures were made and how many times the strain has been sub-cultured (passaged). Stock control strains should not be more than four sub-cultures from the original NCTC authenticated strain because serially sub-culturing bacteria can cause genetic variation which can impact on phenotypic characteristics. Published evidence suggests that this advice is not always followed[1].

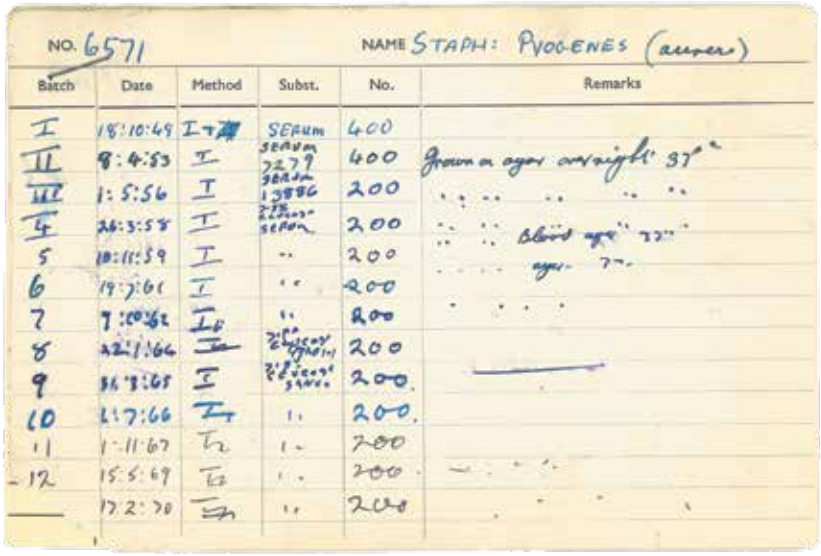
Developing a robust documentation and characterization process

In 1947, Dr Cowan requested a review of the collection's documentation, data and the characterization processes for the cultures. A large scale freeze-drying program to preserve the bacteria more effectively over extended time periods began in 1949 and a parallel work stream designed to improve characterization began in the same year with the introduction of a manual data recording system using individual hand-written cards prepared for every strain to record colony morphology, biochemical test results and freeze-drying records (Figure 2). This initiative resulted, according to a report in our archives, in NCTC being recognized as the foremost collection in the world with a vision to "encourage uniformity in the scientific work of microbiologists of all nations".

Dr Cowan's contributions to the first International Code of Nomenclature of Bacteria eventually led to the publication of Cowan and Steel's widely used '*Manual for Identification of Medical Bacteria*' (1965) which continues to have an essential role in bacteriology laboratories. The earlier editions focused on classical biochemical identification tests but more recent editions of the manual include sections on rapid and mechanized test methods, reflecting progressive developments within NCTC, and also emphasize the importance of laboratory quality control. Our present strategy, which requires continuous review to improve the relevance and quality of the collection strains, their associated data and the characterization and authentication procedures, is an approach that is compatible with that adopted throughout the life of NCTC.

It is essential that NCTC bacteria remain unchanged from the time they are accepted for accessioning into the collection until they are received for use in a requesting laboratory, a period that may span more than 90 years. During that time multiple consecutive freeze-dried batches of the strain may be prepared to ensure there are sufficient stocks to meet the requirements of the biomedical microbiology community (Figure 3). Characterization and authentication of NCTC strains on receipt, and after restocking, initially involved only traditional microbiology techniques such as microscopy, culture and biochemical tests; those methods continue to be standard practice in the NCTC laboratory today. Biochemical tests are now used in conjunction with serological testing, antigen detection, nucleic acid-based assays (tests that detect the DNA or RNA from the microorganism), mass spectrometry and, most recently, whole genome sequencing.

In the late 1960s increasing importance was attributed to the base compositions (G+C ratios) of deoxyribonucleic acid (DNA), as an adjunct to taxonomic studies. This is because although the DNA composition varies from one genus to another, there is less



NO. 6571 NAME STAPH: PYOGENES (aureus)

Batch	Date	Method	Subst.	No.	Remarks
I	18:10:49	I+	SERUM	400	
II	8:4:53	I	SERUM 2279	400	Grown on agar overnight 37°
III	1:5:56	I	SERUM 13886	200	" " " " "
IV	26:3:58	I	SERUM	200	" " Blood agar 37°
5	10:11:59	I	"	200	" " agar 7-
6	19:7:61	I	"	200	" " " "
7	7:10:62	I+	"	200	" " " "
8	22:1:66	I+	SERUM 2 years' storage	200	
9	16:8:68	I	SERUM 2 years' storage	200	
10	2:7:66	I+	"	200	
11	1:11:67	I+	"	200	
12	15:5:69	I+	"	200	
	17:2:70	I+	"	200	

Figure 3 ~ Early freeze-drying record of *Staphylococcus aureus* NCTC 6571 showing batches made from 1949 to 1970.

variance between species in the same genus, and the DNA from any single strain is usually homogeneous. Dr L R Hill (Curator 1978-1994) introduced tests to compare G+C contents of NCTC strains, DNA-DNA hybridization tests and restriction fragment length polymorphism (RFLP) analyses during the 1980s. 16S ribosomal RNA (rRNA) gene sequence analysis was introduced in the 1990s and NCTC strains continue to be routinely identified and authenticated using this method. 16S rRNA gene sequence analysis can often discriminate far more finely between strains of bacteria than is possible with traditional phenotypic methods and can allow more precise identification of poorly described, rarely isolated or phenotypically aberrant strains of a particular species[2]. 16S rRNA sequencing is still considered by many taxonomists to be the gold standard in bacterial identification and classification because 16S gene sequences are present in almost all bacteria, the function of the 16S gene appears to remain constant over time and the gene is sufficiently large for bioinformatic applications. However, the method is known to deliver low discrimination between some closely related species[3].

In the past few years Public Health England's specialist and reference microbiology laboratories that deliver the National Infection Service have embraced the use of mass spectrometry-based methods for more detailed bacterial identification. This technology detects and compares microbial proteins. All the structural and functional proteins present in an organism are known collectively as the proteome, the study of which is referred to as proteomics. Analysis of a bacterial proteome can supplement genomic analysis in furthering understanding of pathogenicity. Not all genes are expressed as proteins, one gene can code for more than one protein, the microbial proteome can change as a result of post translational modifications and gene expression can be dependent on factors such as environmental conditions. The most practical application of proteomics to date is the analysis of target proteins, as opposed to entire proteomes. This type of proteomics, referred to as functional proteomics, is always driven by a specific biological question and this approach is being used in on-going NCTC projects to study microbial proteins in several species including, *Bordetella pertussis*, where multiple strains, with apparently identical genomes, are exhibiting different degrees of pathogenicity.

An invaluable public resource

NCTC's combination of historical and current strains represents an invaluable public resource for the study of important bacterial pathogen groups and the emergence and evolution of pathogenicity factors and antimicrobial resistance. Our strategy focuses on a commitment to ensure continued scientific relevance and includes a requirement to develop improved electronic resources for increasing user access to information about the bacteria. In 2013, in partnership with the Wellcome Trust Sanger Institute (WTSI), NCTC was awarded a Wellcome Trust grant that has enabled us to progress this objective. The funding supports the 'NCTC 3000 project' which aims to provide annotated and assembled genomes for 3000 NCTC bacteria. A publically accessible web-based e-resource will integrate strain accession metadata and taxonomic and authentication information with publications, genome sequences, comparative analysis databases, and other resources available on the European Molecular Biology Laboratory (EMBL) and the National Center for Biotechnology Information (NCBI) websites.

The NCTC 3000 project is particularly important because whole genome sequencing (WGS) is emerging as a potential tool for routine use in clinical microbiology laboratories in high income countries. It can already tell us much about past events, and is

used for determining bacterial lineages, investigating outbreaks of infection and identifying biomarkers. Knowing the sequence of an entire genome, rather than that of only a few fragments, has been shown to increase the precision of molecular epidemiology and contact tracing for infections caused by bacteria such as *Mycobacterium tuberculosis*[5]. If NCTC is to remain scientifically important, it is essential that we embrace WGS technology and provide accurate reference genome sequences.

The strains to be included in the NCTC 3000 project prioritize type strains and strains of public health or historical importance, wherever possible. The sequences are derived using long read technology which has higher accuracy than shorter read technologies, delivers closed genomes and can identify DNA-base modifications such as methylations which are believed to play a role in pathogenicity. More than 1300 sequences are already publically accessible from the European Nucleotide Archive (ENA)/ EMBL-bank or via the NCTC or Sanger websites[6, 7].

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. In 2015 NCTC was awarded accreditation by the United Kingdom Accreditation Service (UKAS) to ISO 17025:2005 for a range of tests undertaken to quality control the NCTC strains in freeze-dried format. The NCTC 3000 project presents challenges about the quality and reproducibility of WGS methods; the reference genomes generated from the project are already being used by third parties. As the methods become more widely adopted, the quality of DNA extractions, genome sequences and assembled data will be challenged for accuracy and reproducibility. It is therefore paramount to stipulate the limitations of the methods, for example, in the NCTC 3000 project the fact that the sequences of small plasmids in certain bacteria will be lost using long read technology, and we have procedures in place to resolve this in the longer term by supplementing the data using other sequencing platforms.

Insight into antimicrobial resistance

The first important study resulting from the NCTC 3000 project was published in the Lancet in November 2014[8]. This was particularly significant because it included WGS data from NCTC 1, a strain of *Shigella flexneri* that was not only the very first bacterial strain to be included in the collection, but was isolated from a British soldier, Private Ernest Cable, who died of dysentery in World War 1. The study revealed that NCTC 1 belonged to a specific lineage (2a) of *S. flexneri*, with which it shares common characteristics and a large core genome. NCTC 1 was resistant to penicillin and erythromycin, even though it was isolated before the antimicrobial era, and it contained a complement of chromosomal antimicrobial resistance genes similar to that of more recent isolates. Comparison of the NCTC 1 genome with those of more recent strains indicates that genomic islands gained in the *S. flexneri* 2a lineage over time are predominately associated with additional antimicrobial resistances, virulence and serotype conversion. This demonstrates how bacterial populations can exploit multiple sources of resistance genes, and means of horizontal gene transmission, to develop an array of mechanisms of resistance in response to different antibiotics being introduced into clinical or agricultural practice.

One of the most interesting of NCTC's private collections is the Murray collection of several hundred strains of Enterobacteriaceae, collected from the pre-antibiotic era and relatively well-characterized. The strains were amassed by Canadian scientist, Professor E. D. G. Murray. After his death in 1964 the collection passed to his son, Robert E. G. Murray, also an eminent microbiologist, who in the early 1980s, in collaboration with NCTC microbiologists, transferred the Murray collection from the University of Western Ontario to the NCTC laboratory in Colindale, London. These strains are now publically available on request[9].

A recent study to investigate the global phylogeography and evolutionary history of *Shigella dysenteriae* type 1 compared the WGSs of 331 strains collected between 1915 and 2011, including 14 isolates from NCTC's Murray collection[10]. This analysis provided an interesting insight into the historical spread of the pathogen and the publication gained widespread publicity due to the implication that infection with *S. dysenteriae*, which remains a continued scourge in Africa and Asia, probably originated in Europe. In common with the NCTC 1 study outlined above, this research charted the development of a pathogen's resistance to antibiotics.

WGS has also contributed to the understanding of the more recent challenge of hospital-adapted pathogens that are resistant to antibiotics. Raven *et al.* undertook whole genome sequencing of *Enterococcus faecalis* strains associated with bloodstream infections that occurred over more than a decade in the United Kingdom and Ireland to determine the population structure and genetic associations with hospital adaptation[11]. They identified three lineages that predominated in the population; two of which (L1 and L2) were nationally distributed and L3 was geographically restricted. Genome comparison with a global collection that included 14 NCTC strains identified that L1 and L3 were also present in the United States but were genetically distinct. Each of the three main lineages contained a mixture of vancomycin-resistant and –susceptible *E. faecalis* strains which has important implications for infection control and antibiotic stewardship. As noted previously, antimicrobial resistance can result from horizontal gene transfer and also from the use of antibiotics in humans and animals. The overuse, or improper use, of antibiotics makes the development and spread of resistance much more likely.

Looking to the future

These studies not only reveal the value of WGS in improving our understanding of infectious diseases but also impact on the accessioning strategy for NCTC. The genomic data, such as that

presented by Raven *et al.*, can reveal gaps in the range of the current collection, such as identifying where strains are missing from specific clinically important lineages so we can procure them accordingly. The WGS information benefits our proteomic scientists who access the data to help to interpret the profiles generated by mass spectrometry. A benefit of belonging to Public Health England is the close relationship between NCTC and the scientists in the National Infection Service's Reference and Research laboratories who recognize the immense value of the collection and will deposit emerging and historical strains with phenotypic or genomic characteristics of biomedical significance. However, input of strains to the collection by the wider clinical and veterinary microbiology community is essential if the collection is to realize its aim of providing trustworthy source of microorganisms for the biomedical community. Depositing strains in NCTC is free of charge, allows the depositors free access to ampoules of their strains on request and provides the opportunity to raise the profile of depositors' publications relating to the NCTC strains.

The NCTC microbiologists are committed to ensuring that the collection of strains that span more than 100 years of bacterial infections, and was established by scientists with incredible foresight in recognizing the need for trustworthy biological resources, remains scientifically relevant for the emerging challenges of the 21st century. The most pressing of those is how to combat antimicrobial resistance, which the Chief Medical Officer for England, Dame Sally Davies, has described as a "catastrophic threat", and there is no doubt that NCTC strains have an important role to play[12].

References

1. Cross LJ, Russell JE, Desai M. (2011) Examining the genetic variation of reference microbial cultures used within food and environmental laboratories using fluorescent amplified fragment length polymorphism analysis. FEMS Microbiol Lett Aug;321(2):100-6. doi: 10.1111/j.1574-6968.2011.02320.x.
2. Clarridge JE. (2004) Impact of 16S rRNA Gene Sequence Analysis for Identification of Bacteria on Clinical Microbiology and Infectious Diseases. Clin Microbiol Rev. Oct; 17(4): 840–862.
3. Janda JM, Abbott SL. (2007) 16S rRNA Gene Sequencing for Bacterial Identification in the Diagnostic Laboratory: Pluses, Perils, and Pitfalls. Clin Microbiol. Sep; 45(9): 2761–2764.
4. GravesPR, Haystead TAJ. (2002) Molecular Biologists Guide to Proteomics Microbiology and Molecular Biology Reviews, Mar, p. 39–63 Vol. 66, No. 1.
5. Guerra-Assunção JA et al. (2015) Large-scale whole genome sequencing of M. tuberculosis provides insights into transmission in a high prevalence area. eLife 2015;4:e05166 DOI: http://dx.doi.org/10.7554/eLife.05166
6. Public Health England. National Collection of Type Cultures (NCTC). http://www.phe-culturecollections.org.uk/collections/nctc.aspx. Accessed 13/06/16.
7. Wellcome Trust Sanger Institute. Public Health England reference collections. http://www.sanger.ac.uk/resources/downloads/bacteria/nctc/ Accessed 130616.
8. Baker KB et al. (2014). The extant World War 1 dysentery bacillus NCTC 1: a genomic analysis. The Lancet. Volume 384 Number 9955 p1641-1720 e52-e56.
9. Baker KB et al. (2015). The Murray collection of pre-antibiotic era Enterobacteriacae: a unique research resource. Genome Medicine 7:97. DOI 10.1186/s13073-015-0222-7
10. Njamkepo et al. (2016) Global phylogeography and evolutionary history of Shigella dysenteriae type 1. Nature Microbiology 1, Article number: 16027 (2016) doi:10.1038/nmicrobiol.2016.27
11. Raven et al. (2015) Genome-based characterization of hospital-adapted Enterococcus faecalis lineages. Nature Microbiology 1. Article number: 15033 (2016) doi:10.1038/nmicrobiol.2015.33
12. Davies, S. (2016). Dame Sally Davies' speech at the Discovery Awards launch. Longitude Prize. https://longitudeprize.org/blog-post/dame-sally-davies-speech-discovery-awards-launch. Accessed 13/06/16.

Acknowledgements

Author: Julie E. Russell, Head of Public Health England's Culture Collections.

The author thanks all the scientists who compiled and maintained the NCTC archives over the past 96 years. This article was commissioned by Thermo Fisher Scientific. All views expressed are those of the author(s).

ThermoFisher SCIENTIFIC

Contact Information:
microbiology@thermofisher.com
USA +1 800 255 6730
International +44 (0) 1256 841144

LT2299A
991-500
August 2016