HOW TO SEQUENCE 10,000 BACTERIAL GENOMES AND RETAIN YOUR SANITY

an accessible, efficient and global approach

Blanca M. Perez-Sepulveda[1], Darren Heavens[2], Caisey V. Pulford[1], Alexander V. Predeus[1], Ross Low[2], Hermione Webster[1], Gregory Dykes[1], Christian Schudoma[2], Will Rowe[1,6], James Lipscombe[2], Chris Watkins[2], Benjamin Kumwenda[7], Neil Shearer[2], Karl Costigan[1], Kate S. Baker[1], Nicholas A. Feasey[3,4], Jay C. D. Hinton[1], Neil Hall[2,5], and the 10KSG consortium‡

[1]Institute of Infection, Veterinary & Ecological Sciences, University of Liverpool, Liverpool, UK, [2]Earlham Institute, Norwich Research Park, Norwich, UK, [3]Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, UK, [4]Malawi-Liverpool-Wellcome Programme, Blantyre, Malawi, [5]School of Biological Sciences, University of East Anglia, Norwich, UK, [6]University of Birmingham, Birmingham, UK, [7]College of Medicine, University of Malawi, Blantyre, Malawi

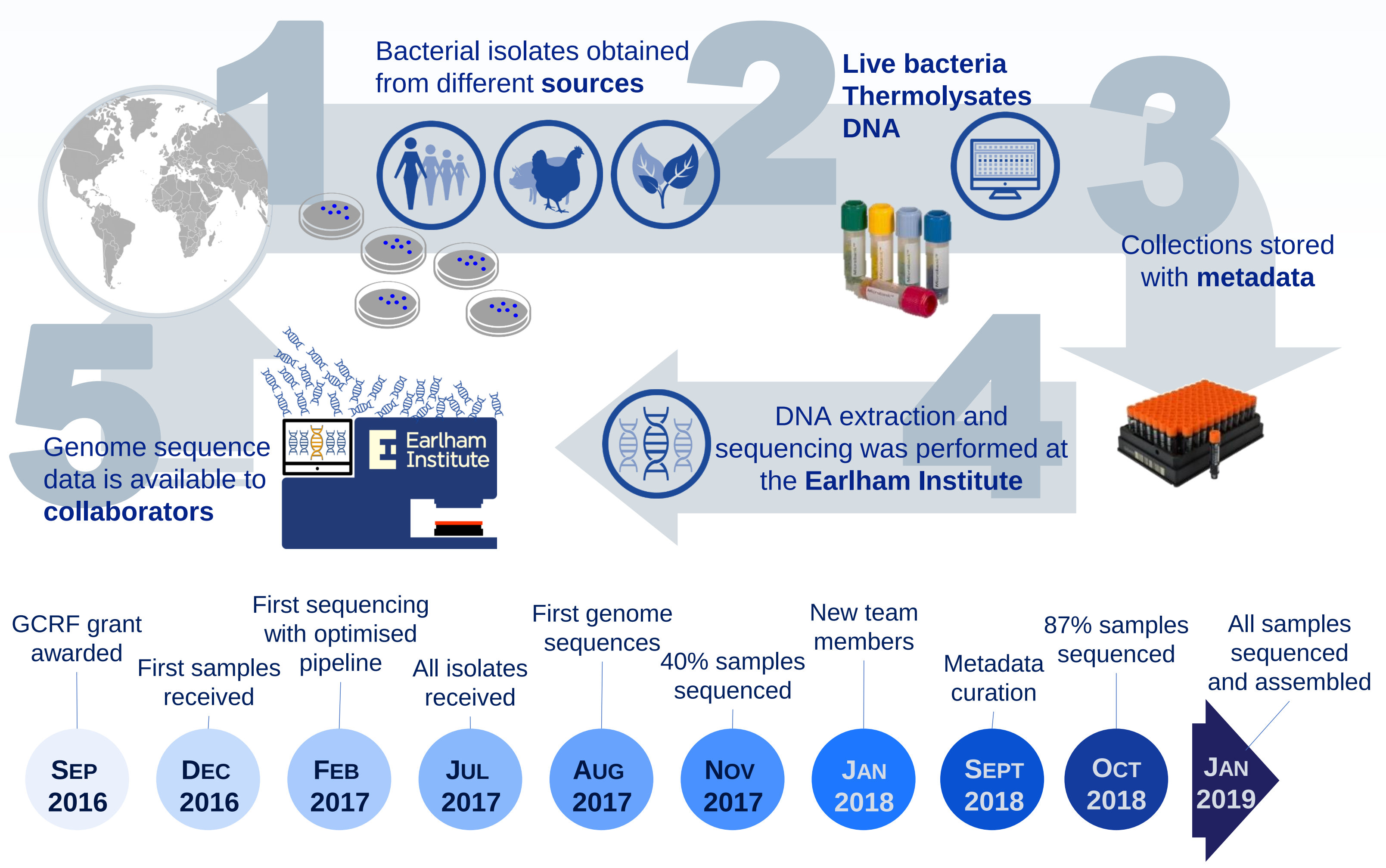
ABSTRACT

Non-typhoidal *Salmonella* (NTS) are typically associated with enterocolitis and linked to the industrialisation of food production. In recent years, NTS has been associated with invasive disease (iNTS disease) causing an estimated 77,000 deaths each year worldwide; 80% of mortality occurs in sub-Saharan Africa. New clades of *S.* Typhimurium and *S.* Enteritidis have been identified, which are characterised by genomic degradation, altered prophage repertoires and novel multidrug resistant plasmids. To understand how these clades are contributing to the burden and severity of iNTS disease, it is crucial to expand genome-based surveillance to cover more countries, and incorporate historical isolates to generate an evolutionary timeline of the development of iNTS.

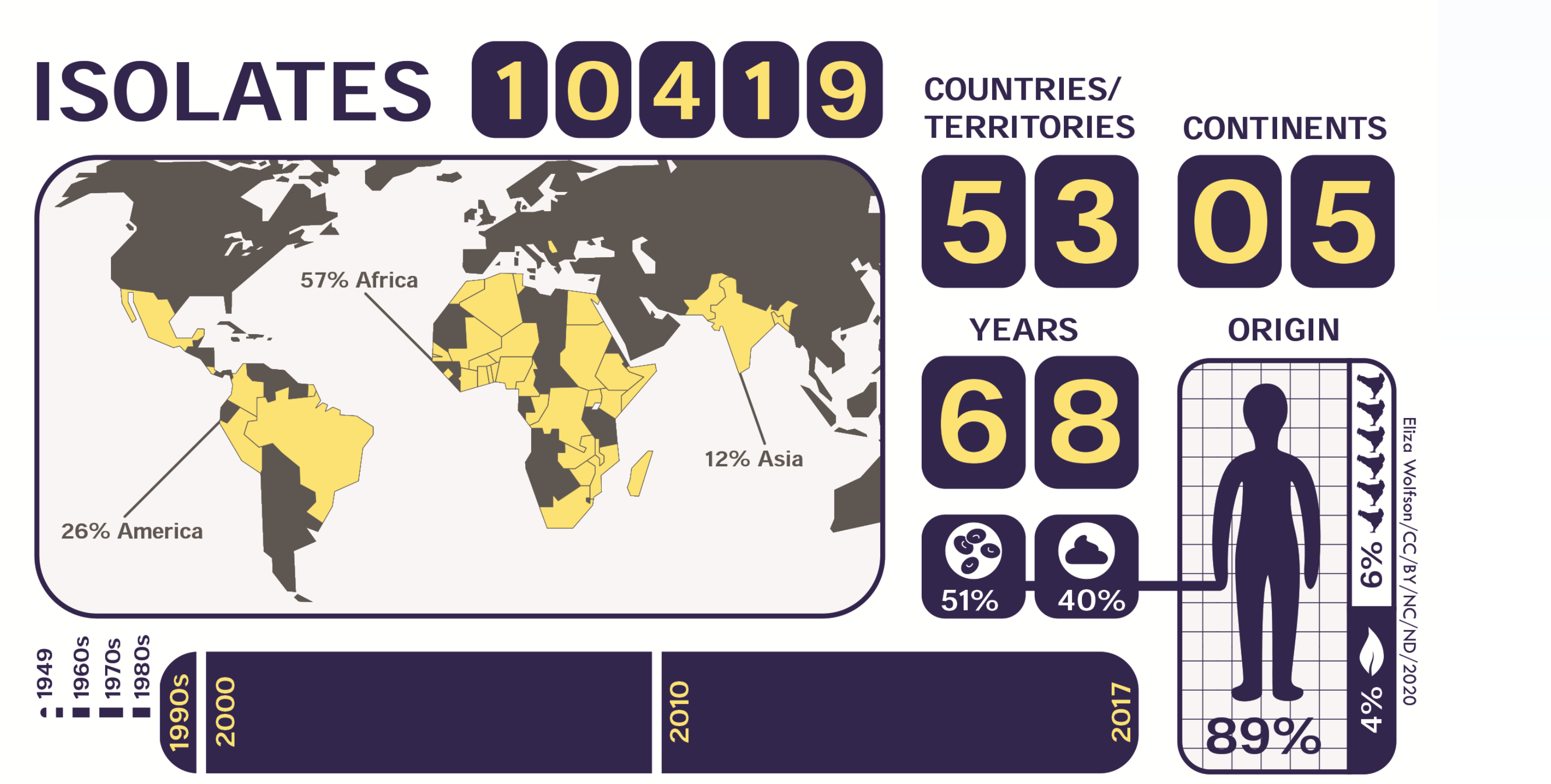
We developed and validated a robust and inexpensive method for large-scale collection and sequencing of bacterial genomes. The "10,000 *Salmonella* genomes" project established a worldwide research collaboration to generate information relevant to the epidemiology, drug resistance and virulence factors of *Salmonella* using a whole-genome sequencing approach. By streamlining collection of isolates and developing an efficient logistics pipeline, we gathered 10,419 clinical and environmental isolates from collections in low and middle-income countries within six months. Genome sequences are now available for isolates from 51 countries/territories dating from 1949 to 2017, with ~80% representing African and Latin-American datasets. Our method can be applied to other large sample collections that require maximisation of resources within a limited timeframe. Detailed genome analyses are in progress, and it is hoped that the resulting data will contribute to public health control strategies in low and middle-income countries.

Read it here or scan this QR code

COLLECTING THE ISOLATES

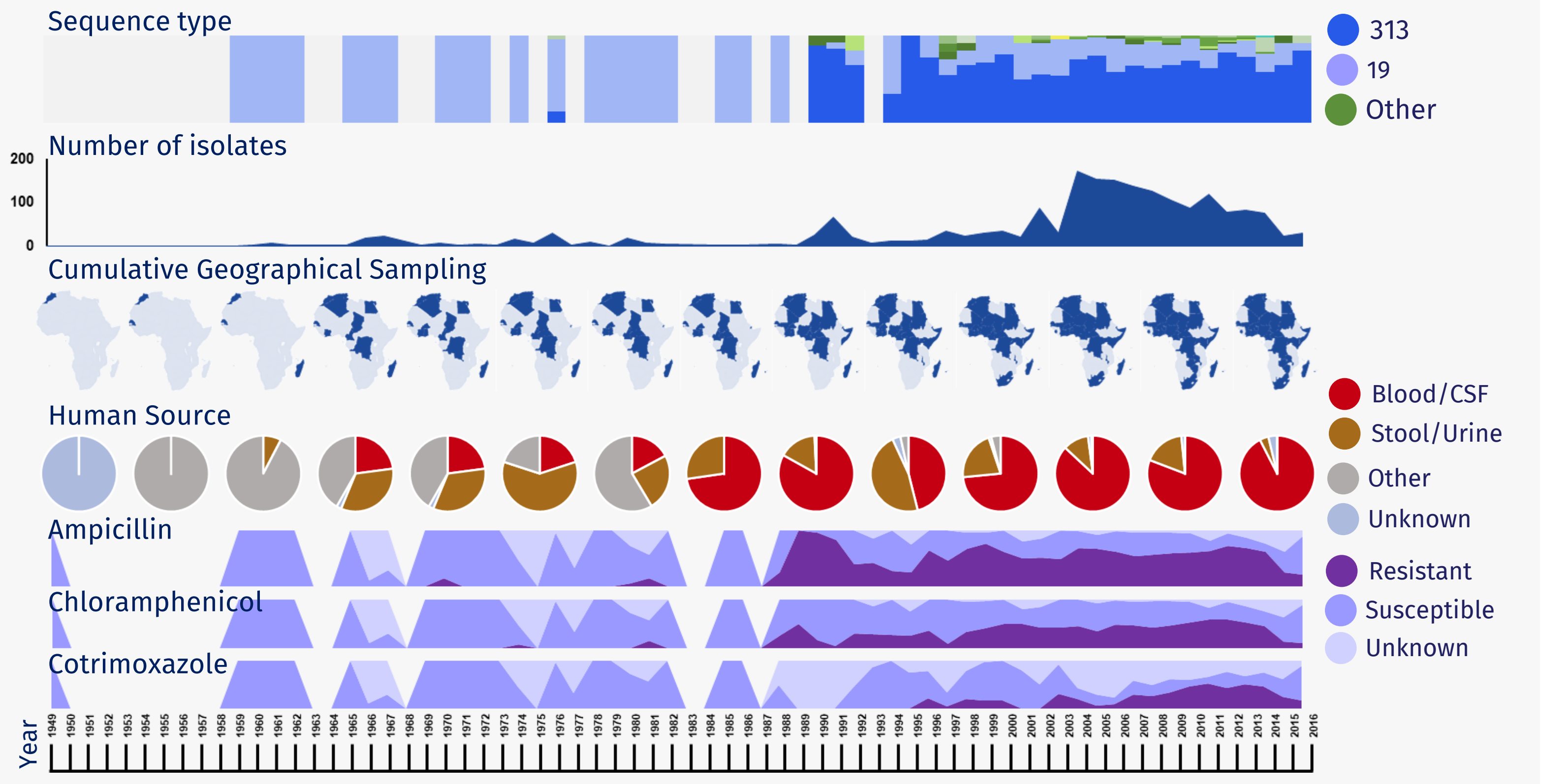


THE DATASET



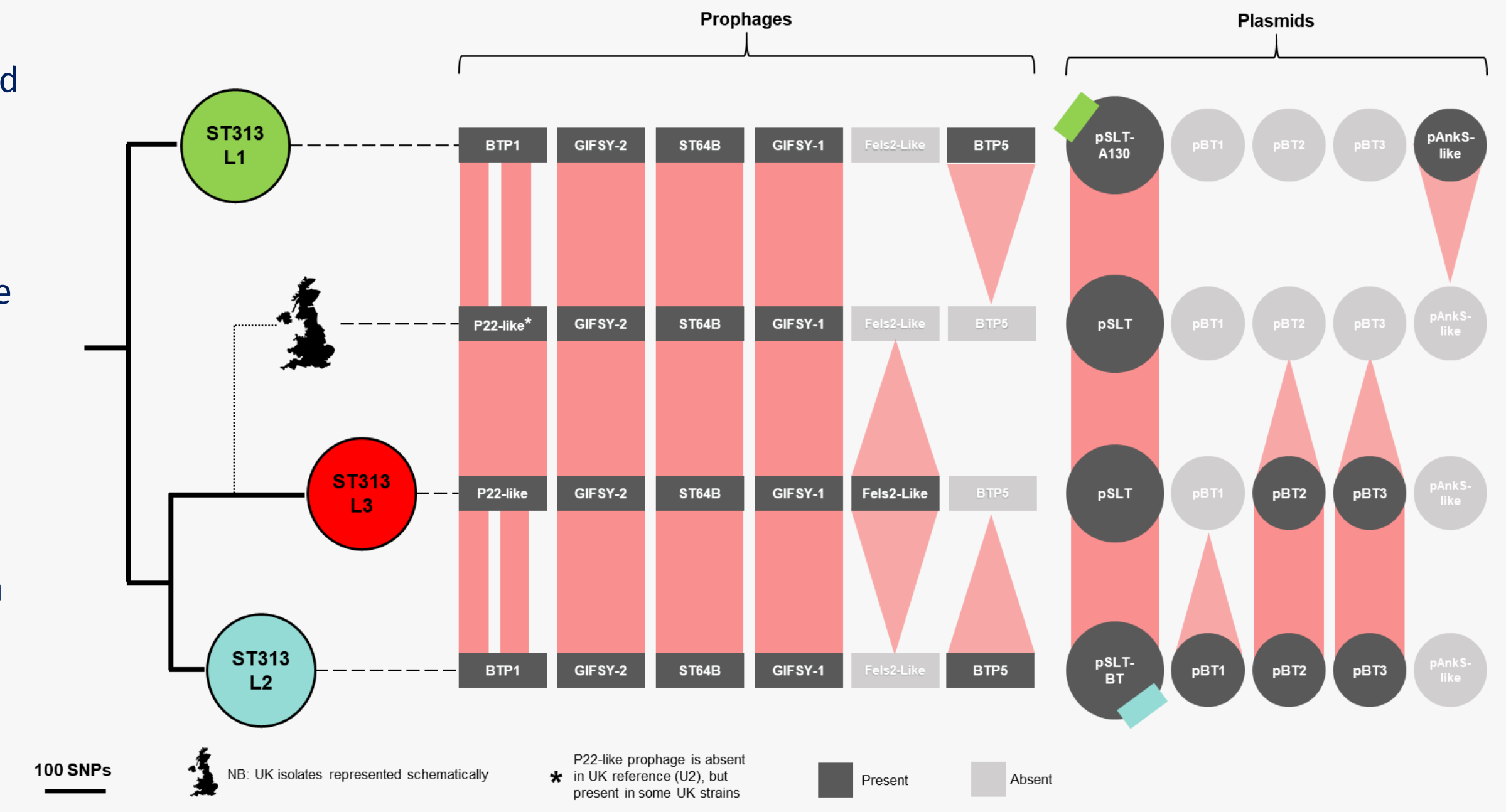
SOME OUTPUTS

S. TYPHIMURIUM GENOMES



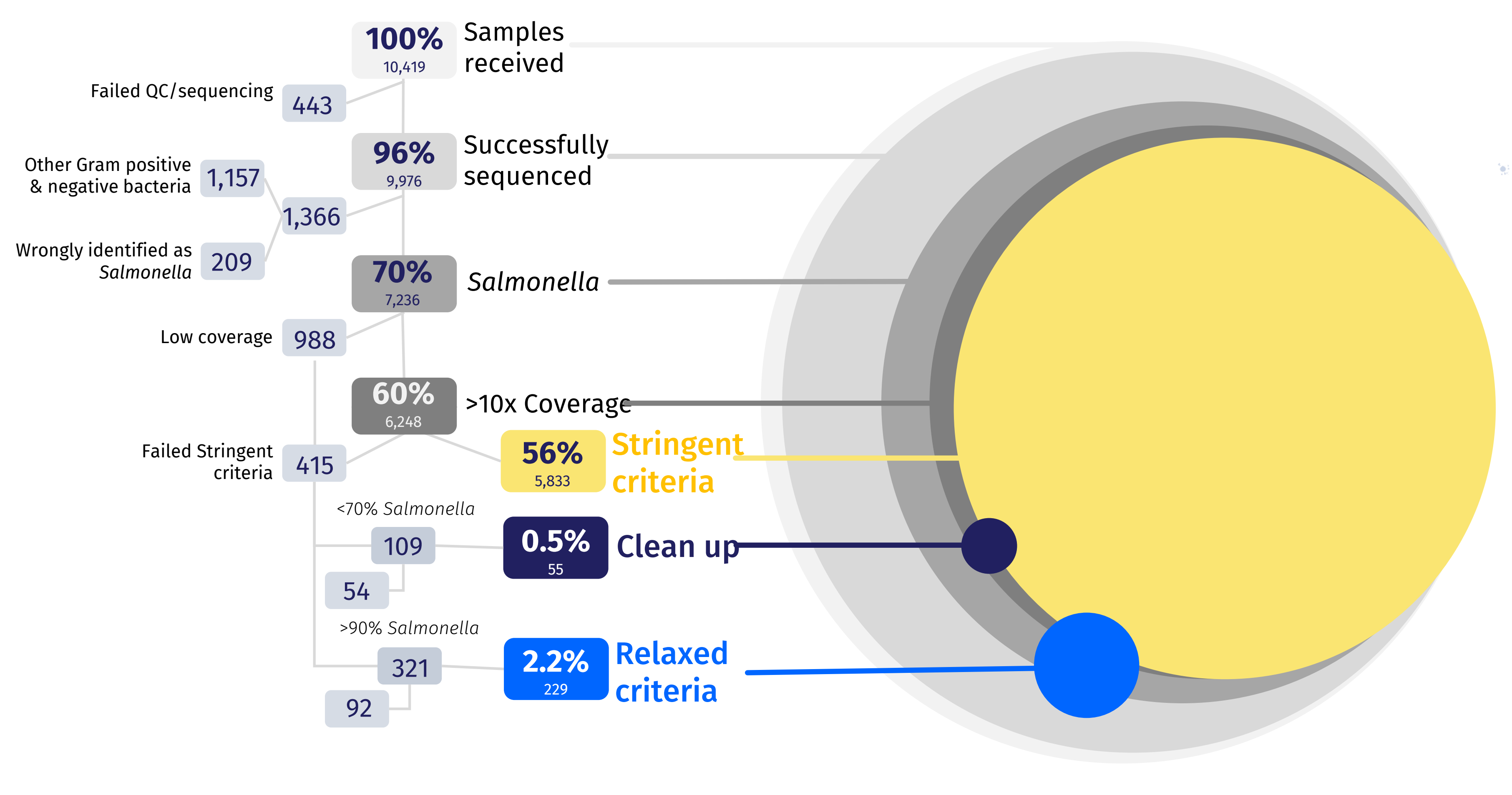
Genome sequence analysis led to the discovery of a pan-susceptible ST313 lineage (L3) in Malawi, closely related to ST313 variants that cause gastrointestinal disease in the UK and Brazil.

In-depth genotypic and phenotypic characterisation identified loss-of-function genetic events that occurred during the stepwise evolution of invasive S. Typhimurium across Africa.



Pulford *et al.* 2021, *Nature Microbiology*

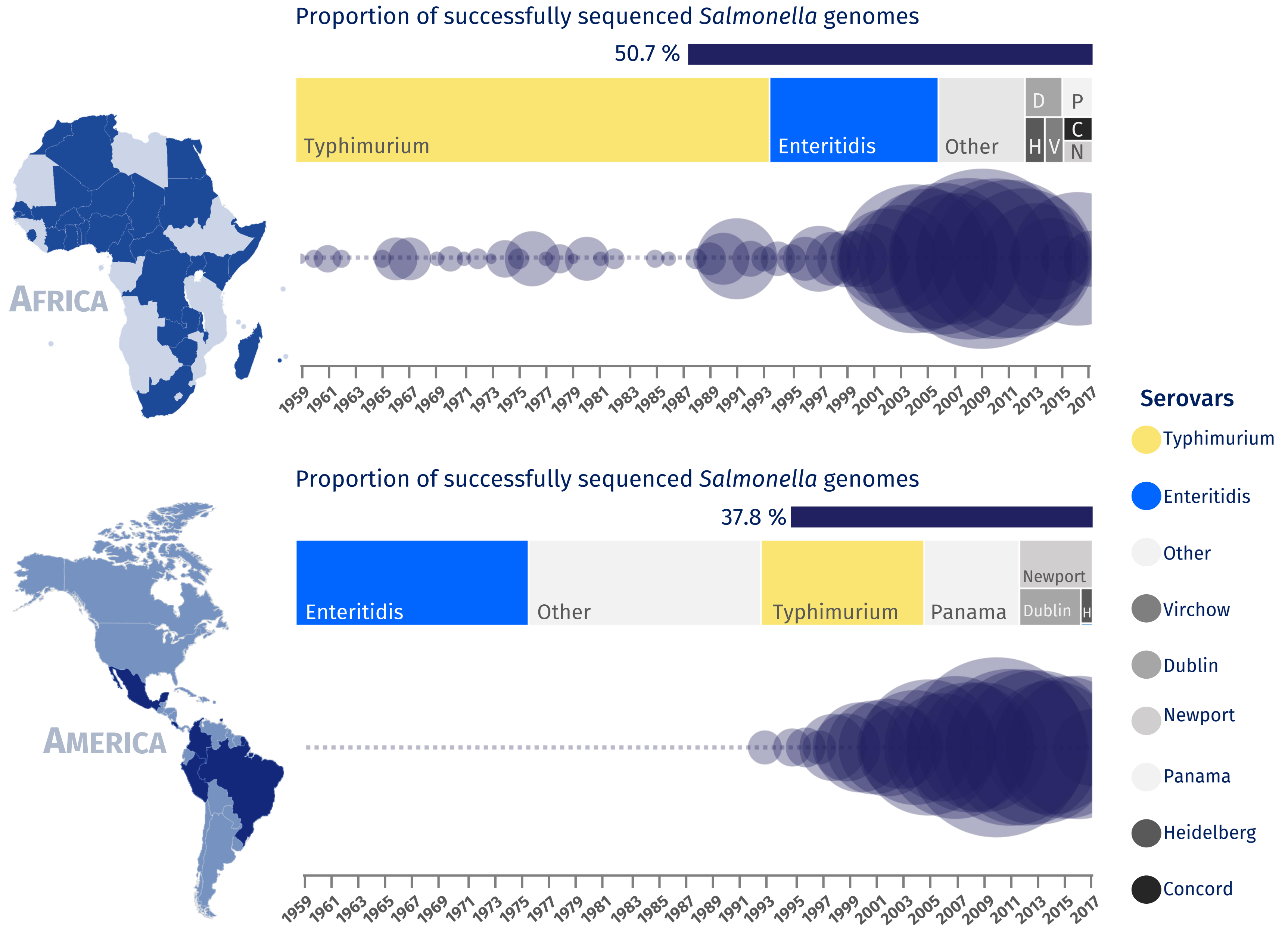
SALMONELLA SEQUENCING QC



Summary

10,419	**7,236**	**6,117**
Bacterial isolates received	Genomes classified as *Salmonella*	**Good quality *Salmonella* genomes**

SALMONELLA GENOMES



***The 10KSGP consortium:**

Blanca M. Perez-Sepulveda[1], Darren Heavens[2], Caisey V. Pulford[1] & María Teresa Acuña[11], Dragan Antic[1], Martin Antonio[5], Kate S. Baker[1], Johan Bernal[8], Hilda Bolaños[11], Marie Chattaway[9], John Cheesbrough[1,‡], Angeziwa Chirambo[4], Karl Costigan[1], Saffiatou Darboe[5], Paula Díaz[10], Pilar Donado[8], Carolina Duarte[10], Gregory Dykes[1], Francisco Duarte[11], Dean Everett[4], Séamus Fanning[12], Nicholas A. Feasey[3,4], Patrick Feglo[13], Adriano M. Ferreira[15], Rachel Floyd[1], Ronnie G. Gavilán[13,26], Melita A. Gordon[1,4], Neil Hall[2], Rodrigo T. Hernandes[15], Gabriela Hernández-Mora[16], Jay C. D. Hinton[1], Daniel Hurley[12], Irene N. Kasumba[17], Benjamin Kumwenda[7], Brenda Kwambana-Adams[24], James Lipscombe[2], Ross Low[2], Salim Mattar[18], Lucy Angeline Montaño[10], Cristiano Gallina Moreira[15], Jaime Moreno[10], Dechamma Mundanda Muthappa[12], Satheesh Nair[9], Chris M. Parry[3], Chikondi Peno[4], Jasnehta Permala-Booth[17], Jelena Petrović[19], Alexander V. Predeus[1], José Luis Puente[20], Getenet Rebrie[21], Martha Redway[1], Will Rowe[1,6], Terue Sadatsune[15], Christian Schudoma[2], Neil Shearer[2], Claudia Silva[20], Anthony M. Smith[22,25], Sharon Tennant[17], Alicia Tran-Dien[23], Chris Watkins[2], Hermione Webster[1], François-Xavier Weill[23], Magdalena Wiesner[10], Catherine Wilson[1,4]

[1]IVES, University of Liverpool, Liverpool, UK, [2]Earlham Institute, Norwich Research Park, Norwich, UK, [3]Liverpool School of Tropical Medicine, Liverpool, UK, [4]Malawi-Liverpool-Wellcome Programme, Blantyre, Malawi, [5]Medical Research Council Unit The Gambia at LSHTM, [6]University of Birmingham, Birmingham, UK, [7]College of Medicine, University of Malawi, Blantyre, Malawi, [8]Corporación Colombiana de investigación Agropecuaria AGROSAVIA, Colombia, [9]Public Health England, UK, [10]Instituto Nacional de Salud (INS), Colombia, [11]Instituto Costarricense de Investigación y Enseñanza en Nutrición y Salud (INCIENSA), Costa Rica, ‡Retired

Dr Blanca Perez-Sepulveda

BlancaPS@liverpool.ac.uk

@BlancaPerSep