How whole genome sequencing can help tackle antibiotic resistance and dissemination of the deadly bacteria?

Case based examples

Claire Chewapreecha

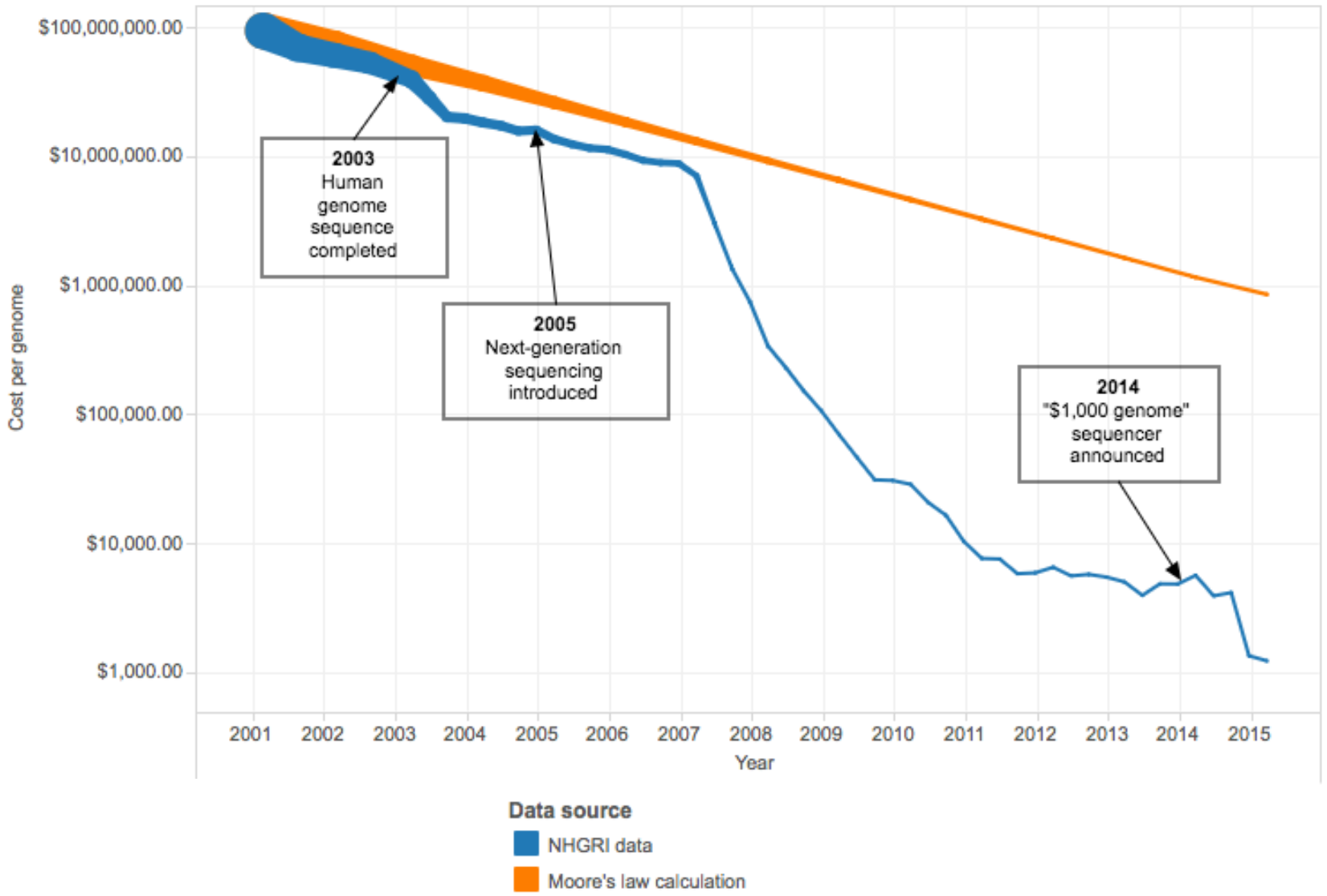
25th July 2017

wellcome trust MORU Tropical Health Network
UNIVERSITY OF OXFORD
wellcome trust sanger institute
KMUTT
wellcome trust Fellow

googlemap

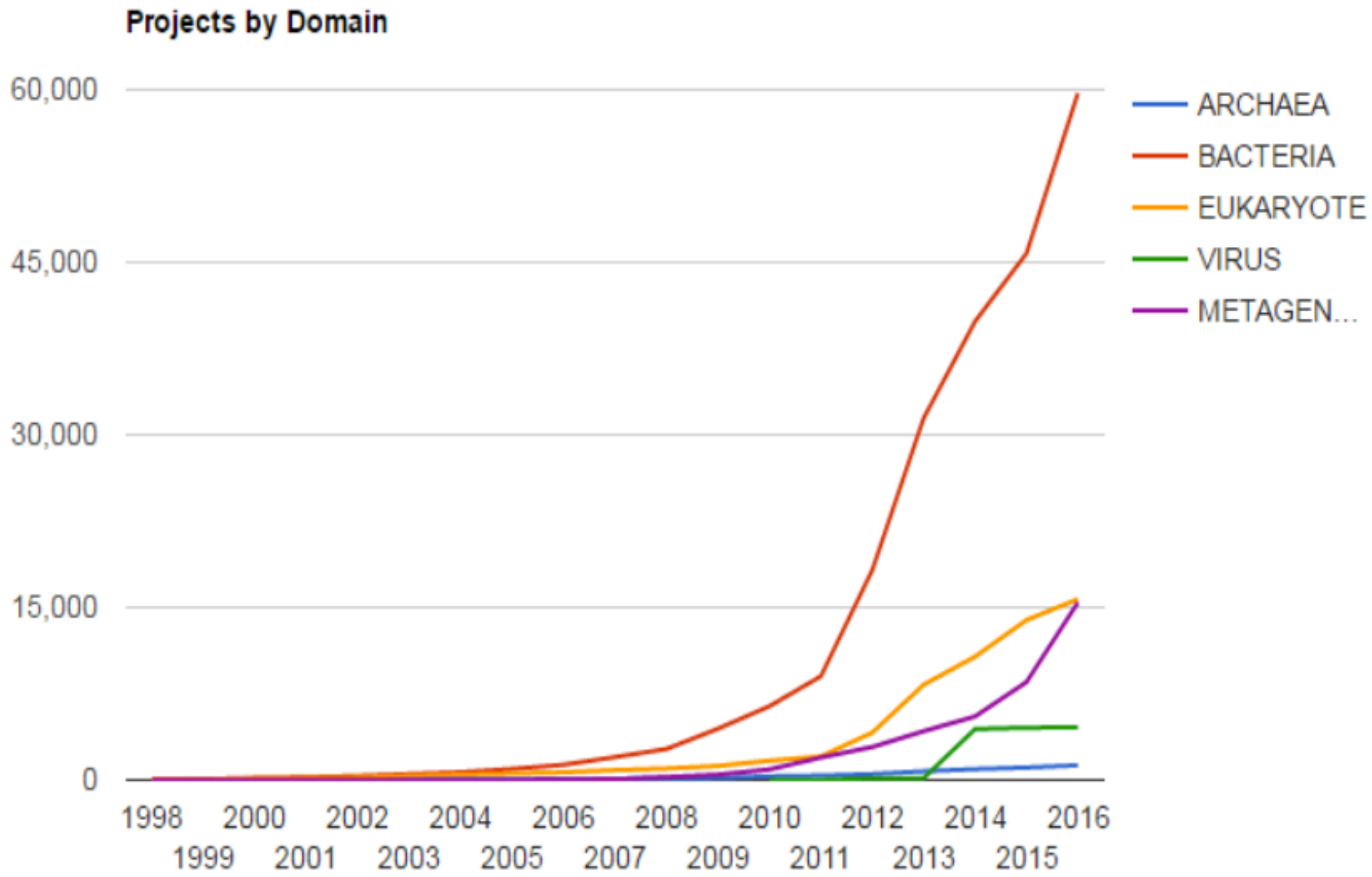
Me, a regular ~~intruder~~ visitor at 6.00 am

Cheaper cost of DNA sequencing makes it possible to sequence more



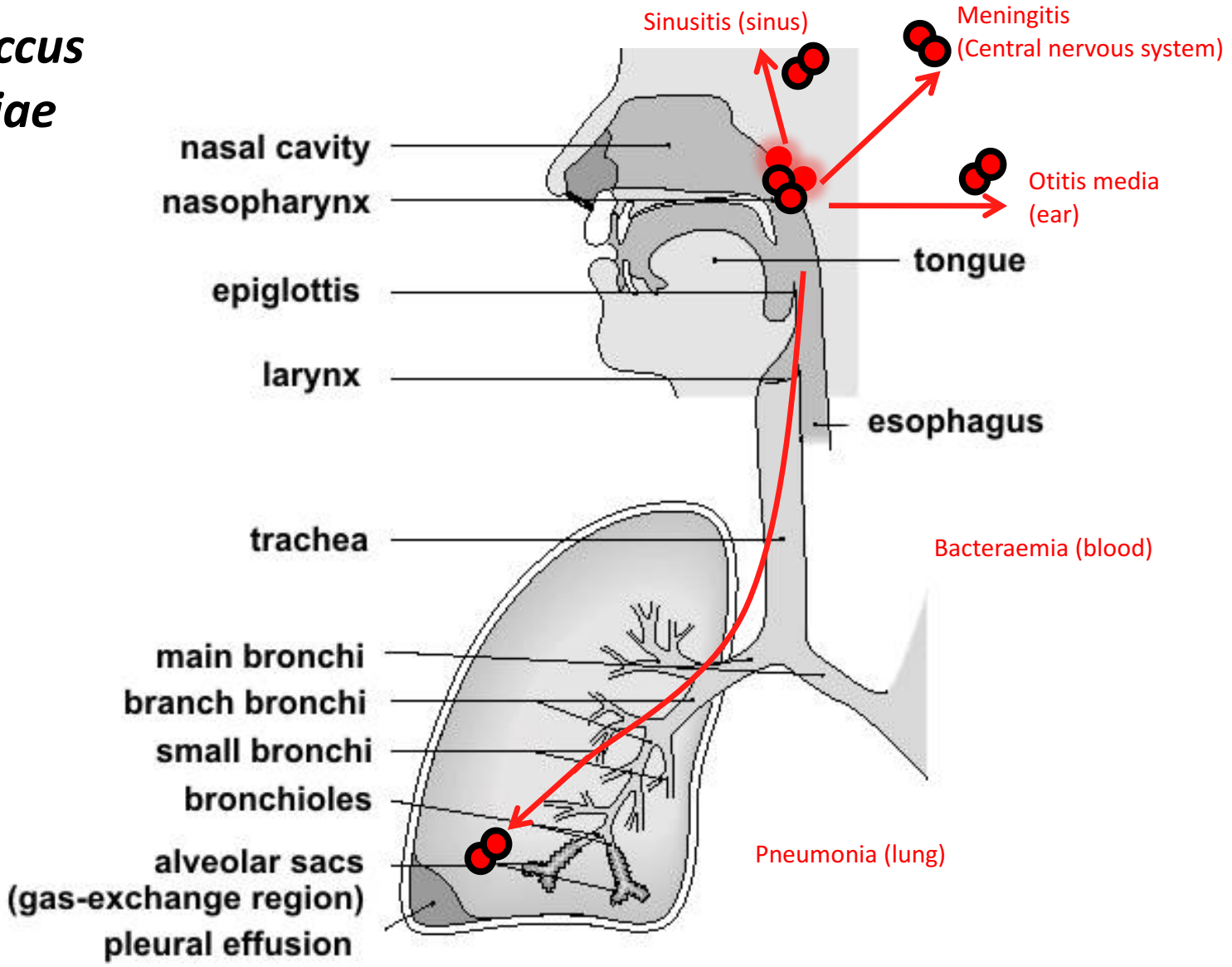
Decline in real costs compared to expected declines based on Moore's Law.
Trend line: Cost per human genome. Line width: Cost per megabase (Mb)
(*Data: NHGRI* https://www.genome.gov/27541954/dna-sequencing-costs-data/)

An exponential increase in projects employing next generation sequencing



Source: NCBI Gold

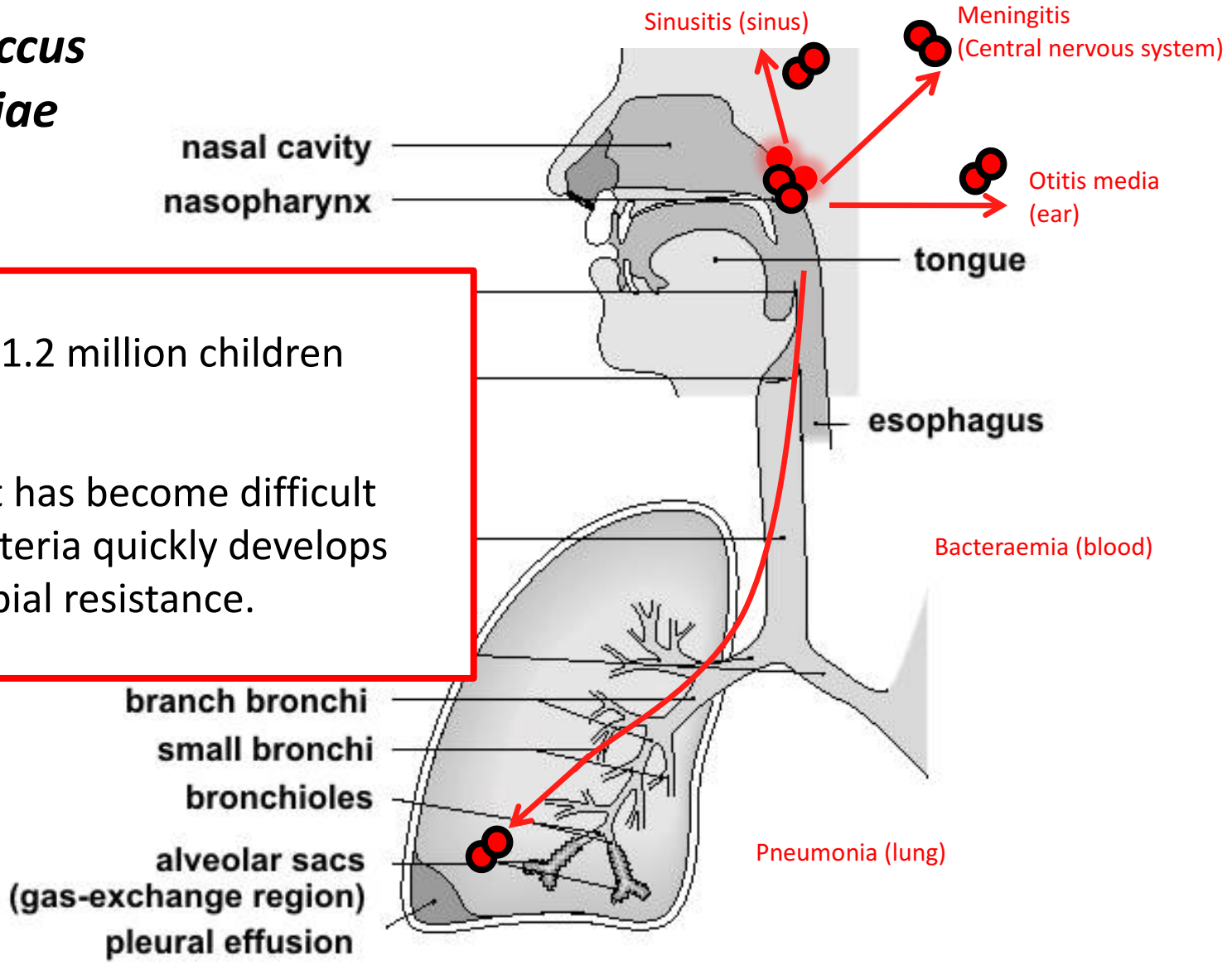
Streptococcus pneumoniae



Streptococcus pneumoniae

A killer of 1.2 million children annually.

Treatment has become difficult as the bacteria quickly develops antimicrobial resistance.



World Health Organization

Feb 2017

WHO priority pathogens list for R&D of new antibiotics

Priority 1: CRITICAL

1. *Acinetobacter baumannii*, carbapenem-resistant
2. *Pseudomonas aeruginosa*, carbapenem-resistant
3. *Enterobacteriaceae*, carbapenem-resistant, ESBL-producing

Priority 2: HIGH

1. *Enterococcus faecium*, vancomycin-resistant
2. *Staphylococcus aureus*, methicillin-resistant, vancomycin-intermediate and resistant
3. *Helicobacter pylori*, clarithromycin-resistant
4. *Campylobacter* spp., fluoroquinolone-resistant
5. *Salmonellae*, fluoroquinolone-resistant
6. *Neisseria gonorrhoeae*, cephalosporin-resistant, fluoroquinolone-resistant

Priority 3: MEDIUM

1. *Streptococcus pneumoniae*, penicillin-non-susceptible
2. *Haemophilus influenzae*, ampicillin-resistant
3. *Shigella* spp., fluoroquinolone-resistant

Streptococcus pneumoniae collection

Paul Turner MD

ClaudiaTurner MD

Myanmar

Thailand

Maela refugee camp: Approx 40,000 Karen refugees, Approx 2.4 km^2

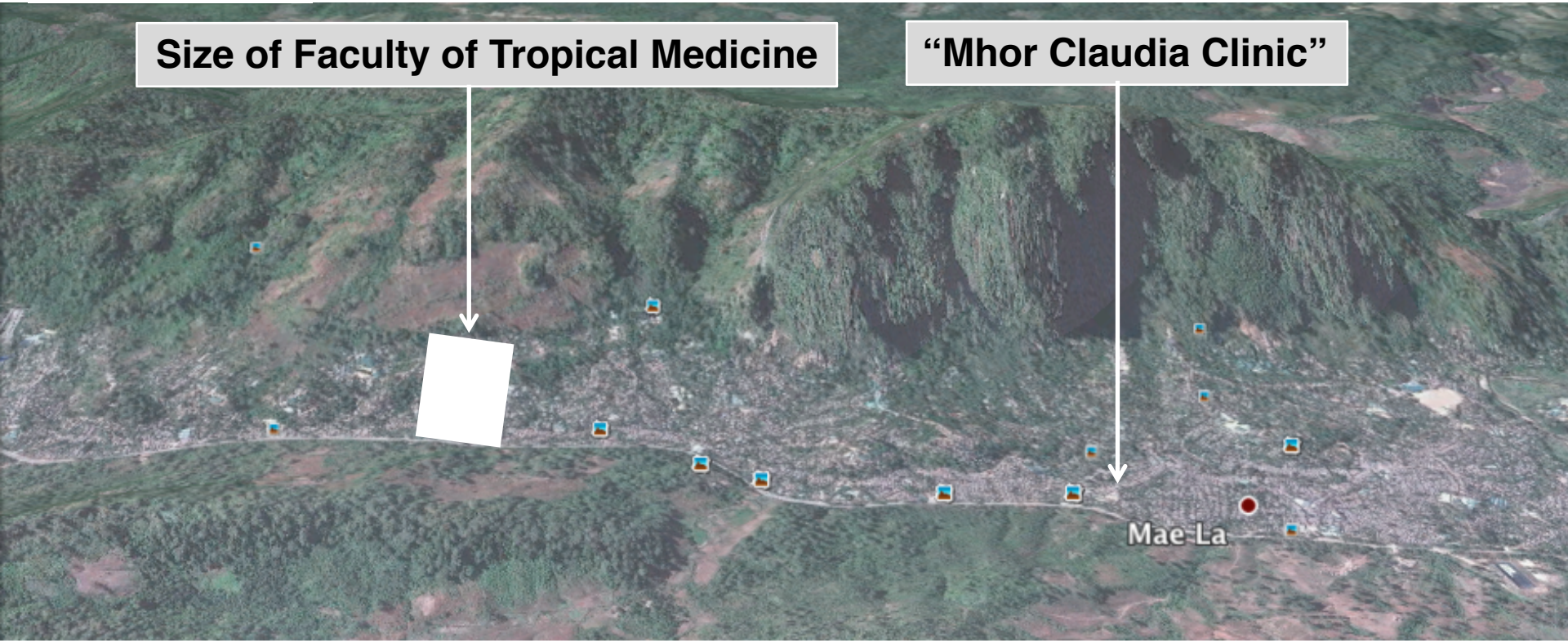
Streptococcus pneumoniae collection

Paul Turner MD

ClaudiaTurner MD

3,085 whole genome sequenced pneumococci collected between 2007-2010.

Myanmar



Thailand

Maela refugee camp: Approx 40,000 Karen refugees, Approx 2.4 km^2

Let's use this 3K collection to identify genetic culprit for penicillin non-susceptibility

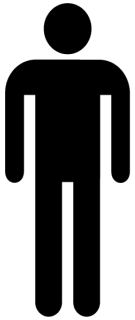
Prof Julian Parkhill

Prof Stephen Bentley

Human Genome wide association study (GWAS)

Disease (case)

Person 1 …ATCTTACCGGGTAAGTGCCTTAGCTAATCGTTAGGAGTCCCA…
Person 2 …ATGTTACCGGGTAAGTGCCTTAGCTAATCGTTAGGAGTCCCA…
Person 3 …ATGTTACCGGGTAAGTGCCTTAGCTAATCGATAGGAGTCCCA…
Person 4 …ATGTTACCGGGTAACTGCCTTAGCTAATCGTTAGGAGTCCCA…
Person 5 …ATGTTACCGGGTAAGTGCCTTAGCTAATCGTTAGGAGTCCCA…



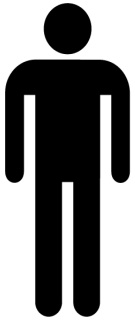
Person 1 …ATGTTACCGGGTAACTGCCTTAGCTAATCGTTAGGAGTCCCA…
Person 2 …ATGTTACCGGGTAACTGCCTTAGATAATCGTTAGGAGTCCCA…
Person 3 …ATGTTACCGGGTAACTGCCTTAGCTAATCGTTAGGAGTCCCA…
Person 4 …ATGTTACCGGGTAACTGCCTTAGCTAATCGTTAGGTGTCCCA…
Person 5 …ATGTTACCGGGTAACTGCCTTAGCTAATCGTTAGGAGTCCCA…

Health (control)

Human Genome wide association study (GWAS)

Disease (case)

Person 1 …ATCTTACCGGGTAAGTGCCTTAGCTAATCGTTAGGAGTCCCA…
Person 2 …ATGTTACCGGGTAAGTGCCTTAGCTAATCGTTAGGAGTCCCA…
Person 3 …ATGTTACCGGGTAAGTGCCTTAGCTAATCGATAGGAGTCCCA…
Person 4 …ATGTTACCGGGTAACTGCCTTAGCTAATCGTTAGGAGTCCCA…
Person 5 …ATGTTACCGGGTAAGTGCCTTAGCTAATCGTTAGGAGTCCCA…

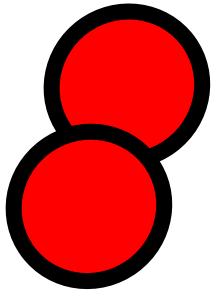


Person 1 …ATGTTACCGGGTAACTGCCTTAGCTAATCGTTAGGAGTCCCA…
Person 2 …ATGTTACCGGGTAACTGCCTTAGATAATCGTTAGGAGTCCCA…
Person 3 …ATGTTACCGGGTAACTGCCTTAGCTAATCGTTAGGAGTCCCA…
Person 4 …ATGTTACCGGGTAACTGCCTTAGCTAATCGTTAGGTGTCCCA…
Person 5 …ATGTTACCGGGTAACTGCCTTAGCTAATCGTTAGGAGTCCCA…

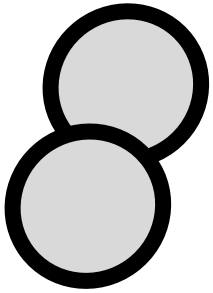
Health (control)

Bacterial Genome wide association study (GWAS)

Penicillin non-susceptible



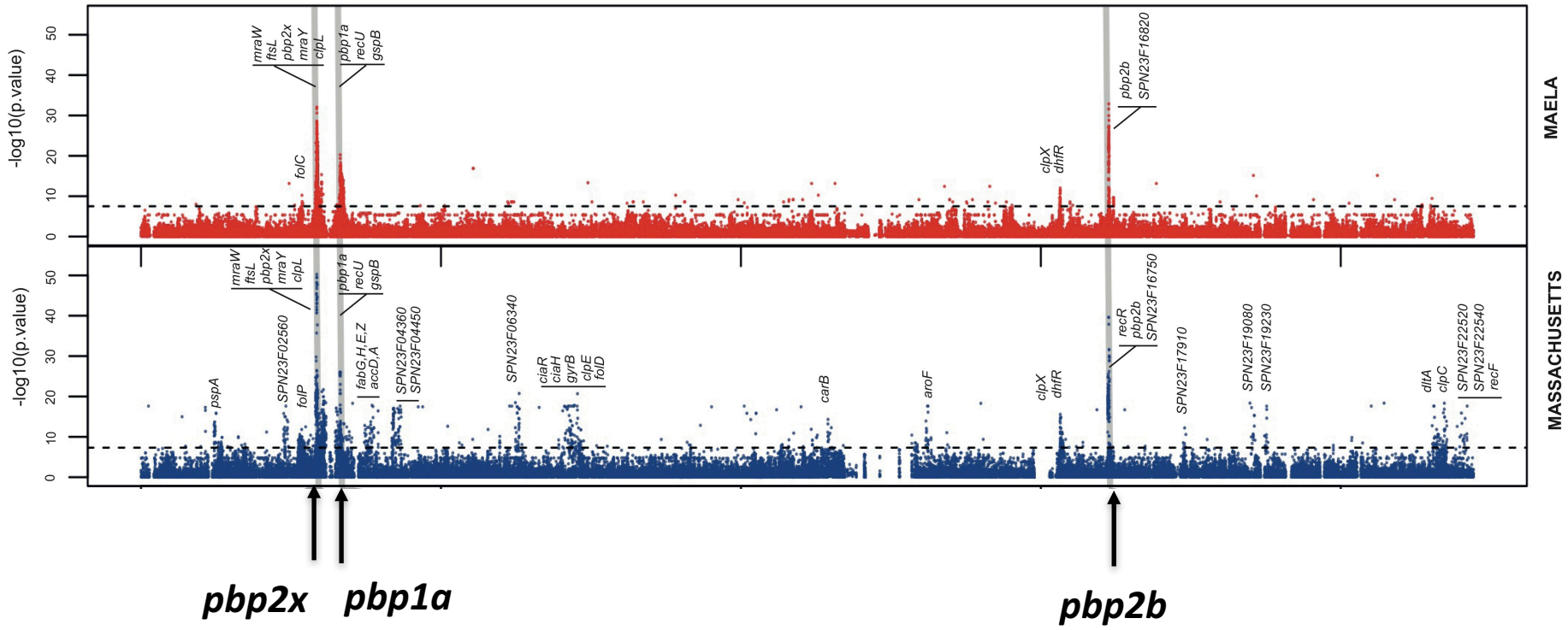
pneumo 1 …ATCTTACCGGGTAAGTGCCTTAGCTAATCGTTAGGAGTCCCA…
pneumo 2 …ATGTTACCGGGTAAGTGCCTTAGCTAATCGTTAGGAGTCCCA…
pneumo 3 …ATGTTACCGGGTAAGTGCCTTAGCTAATCGATAGGAGTCCCA…
pneumo 4 …ATGTTACCGGGTAACTGCCTTAGCTAATCGTTAGGAGTCCCA…
pneumo 5 …ATGTTACCGGGTAAGTGCCTTAGCTAATCGTTAGGAGTCCCA…



pneumo 1 …ATGTTACCGGGTAACTGCCTTAGCTAATCGTTAGGAGTCCCA…
pneumo 2 …ATGTTACCGGGTAACTGCCTTAGATAATCGTTAGGAGTCCCA…
pneumo 3 …ATGTTACCGGGTAACTGCCTTAGCTAATCGTTAGGAGTCCCA…
pneumo 4 …ATGTTACCGGGTAACTGCCTTAGCTAATCGTTAGGTGTCCCA…
pneumo 5 …ATGTTACCGGGTAACTGCCTTAGCTAATCGTTAGGAGTCCCA…

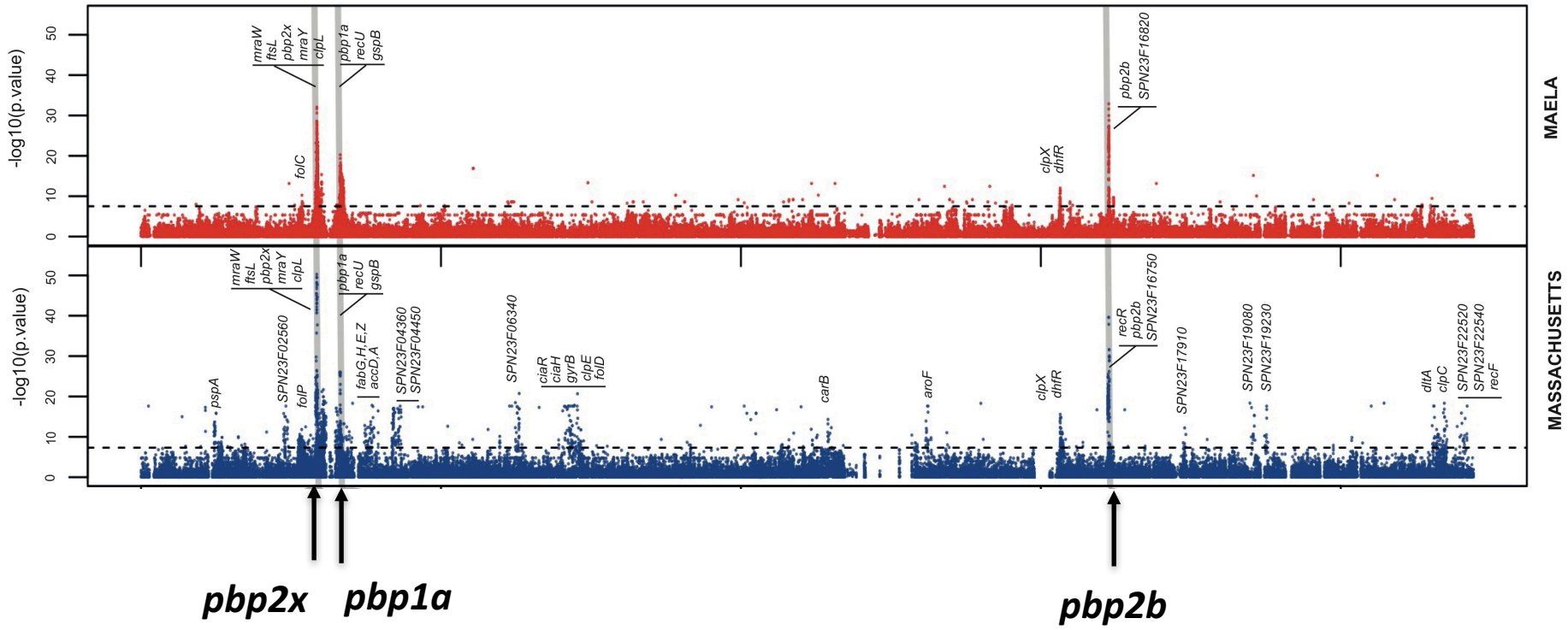
Pennicillin susceptible

Chewapreecha *et al*. PLoS Genetics 2014



It worked!

Chewapreecha *et al*. PLoS Genetics 2014



But are these resistance alleles restricted to one lineage, or they pass on to others ?

KEEP CALM

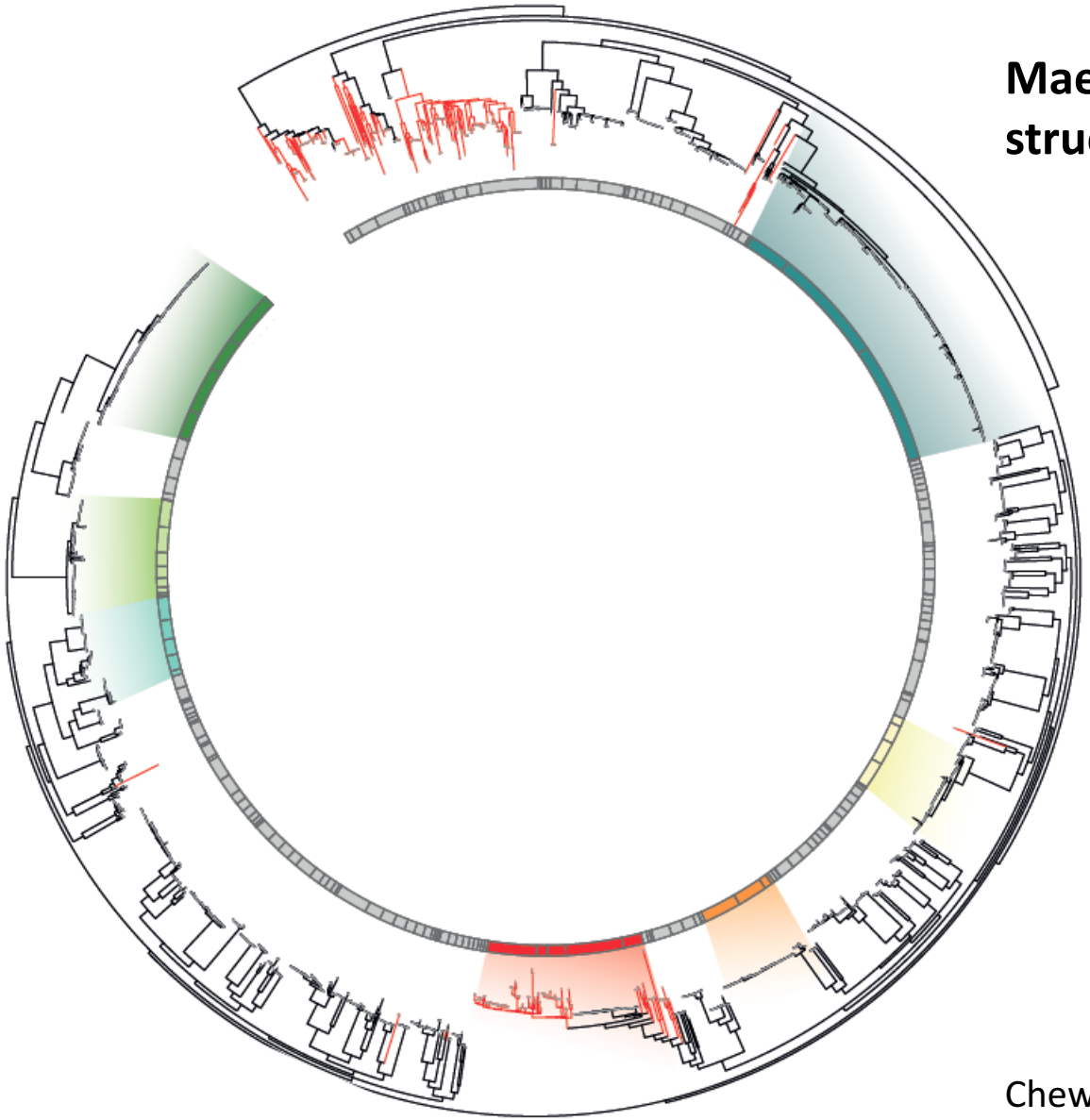
AND

IDENTIFY

RESISTANCE MECHANISM

Prof Julian Parkhill

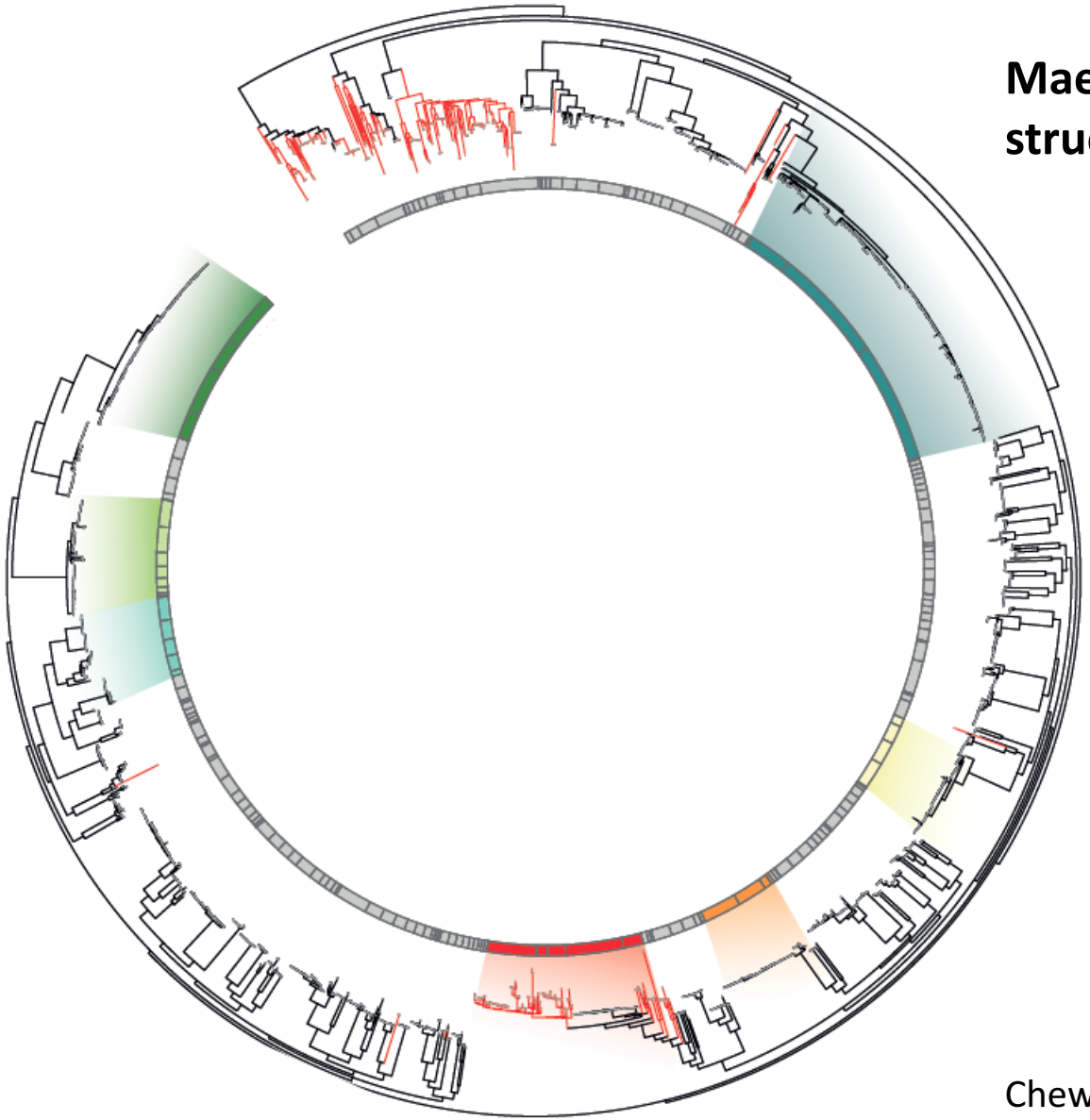
Prof Stephen Bentley



Maela Pneumococcal population structure

- BC1 - 19F
- BC2 - 23F
- BC3 - Non-encapsulated
- BC4 - 6B
- BC5 - 23A/23F
- BC6 - 15B/15C
- BC7 - 14

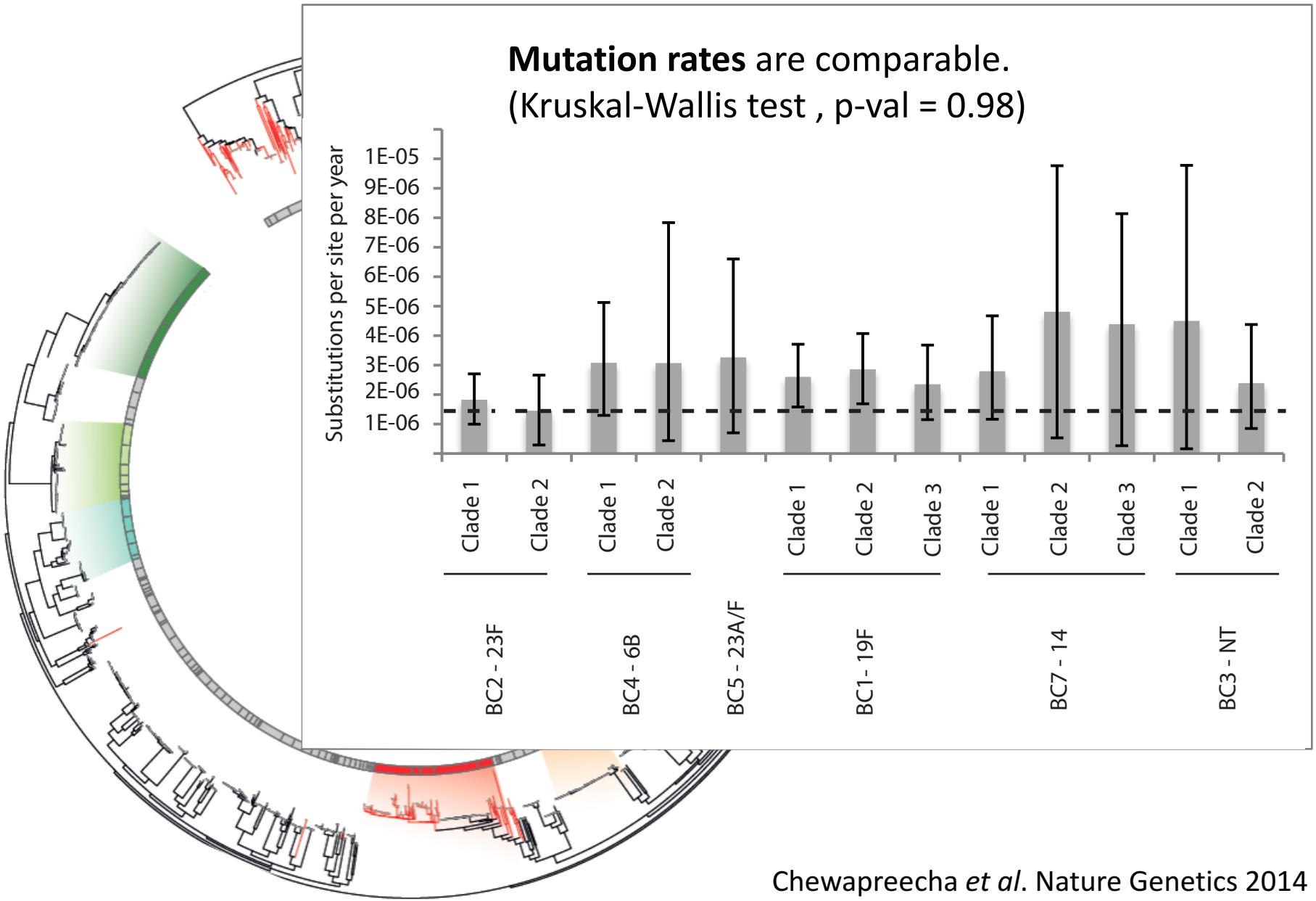
Chewapreecha *et al*. Nature Genetics 2014



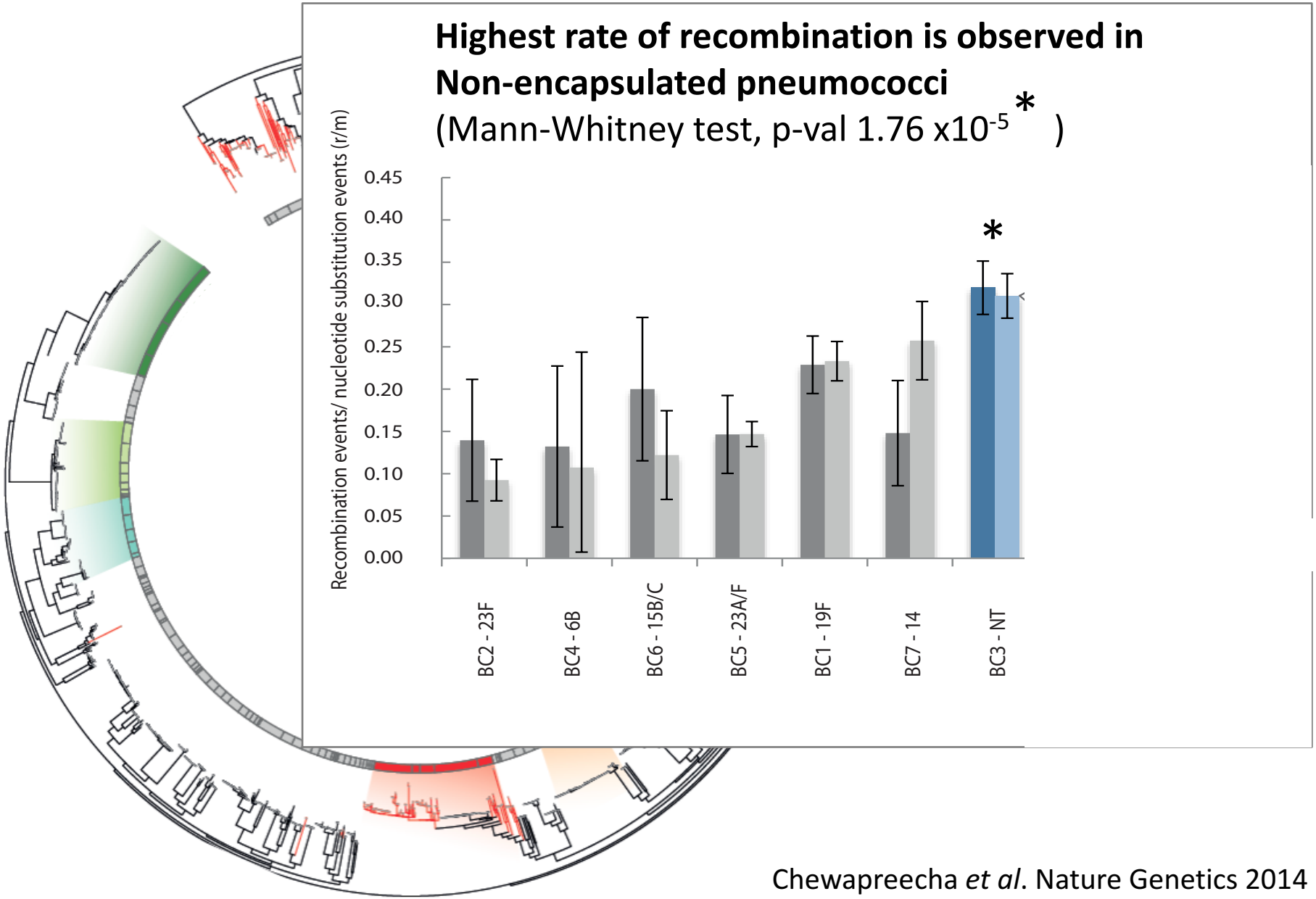
Maela Pneumococcal population structure

- BC1 - 19F
- BC2 - 23F
- BC3 - Non-encapsulated
- BC4 - 6B
- BC5 - 23A/23F
- BC6 - 15B/15C
- BC7 - 14

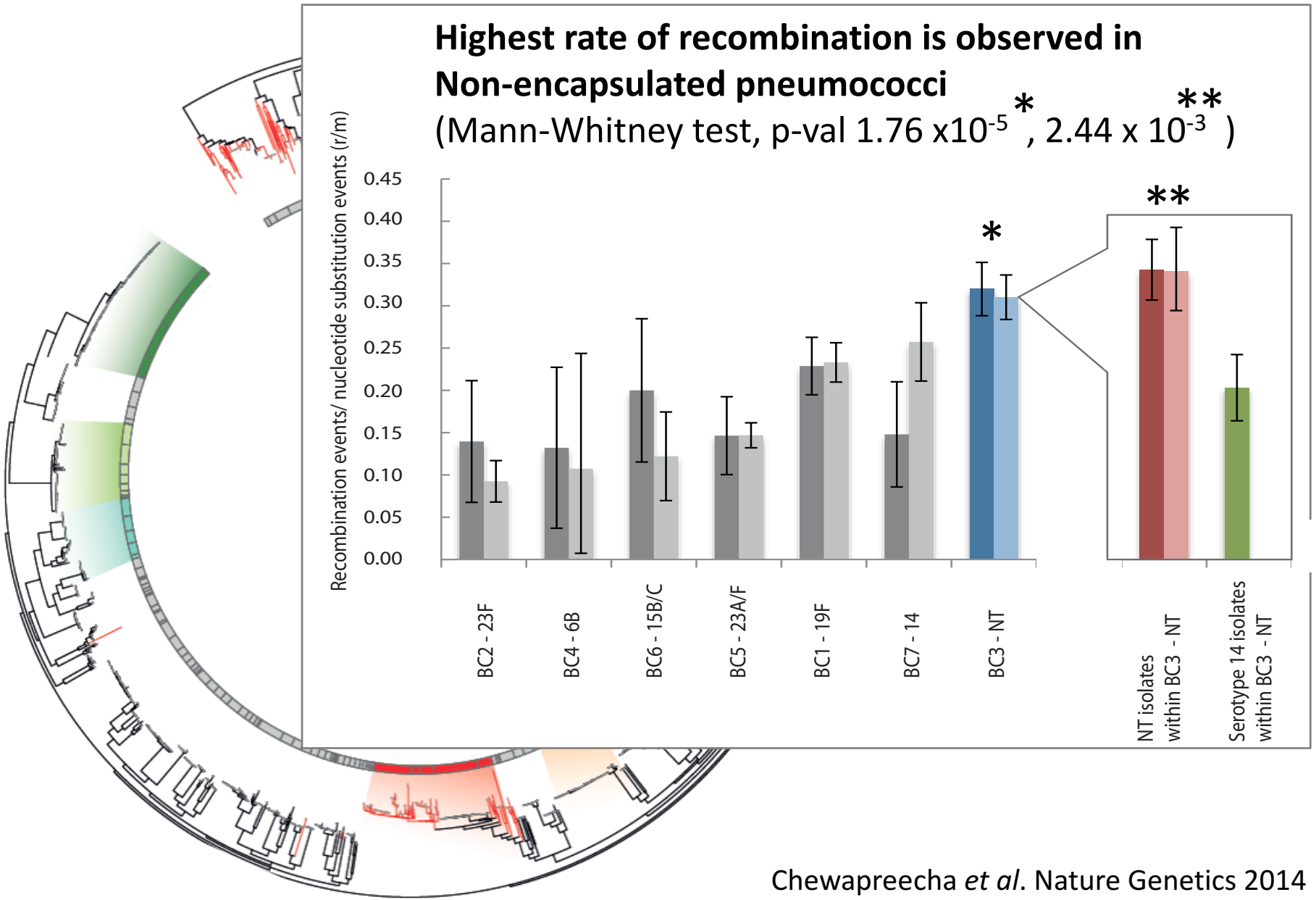
Chewapreecha *et al*. Nature Genetics 2014



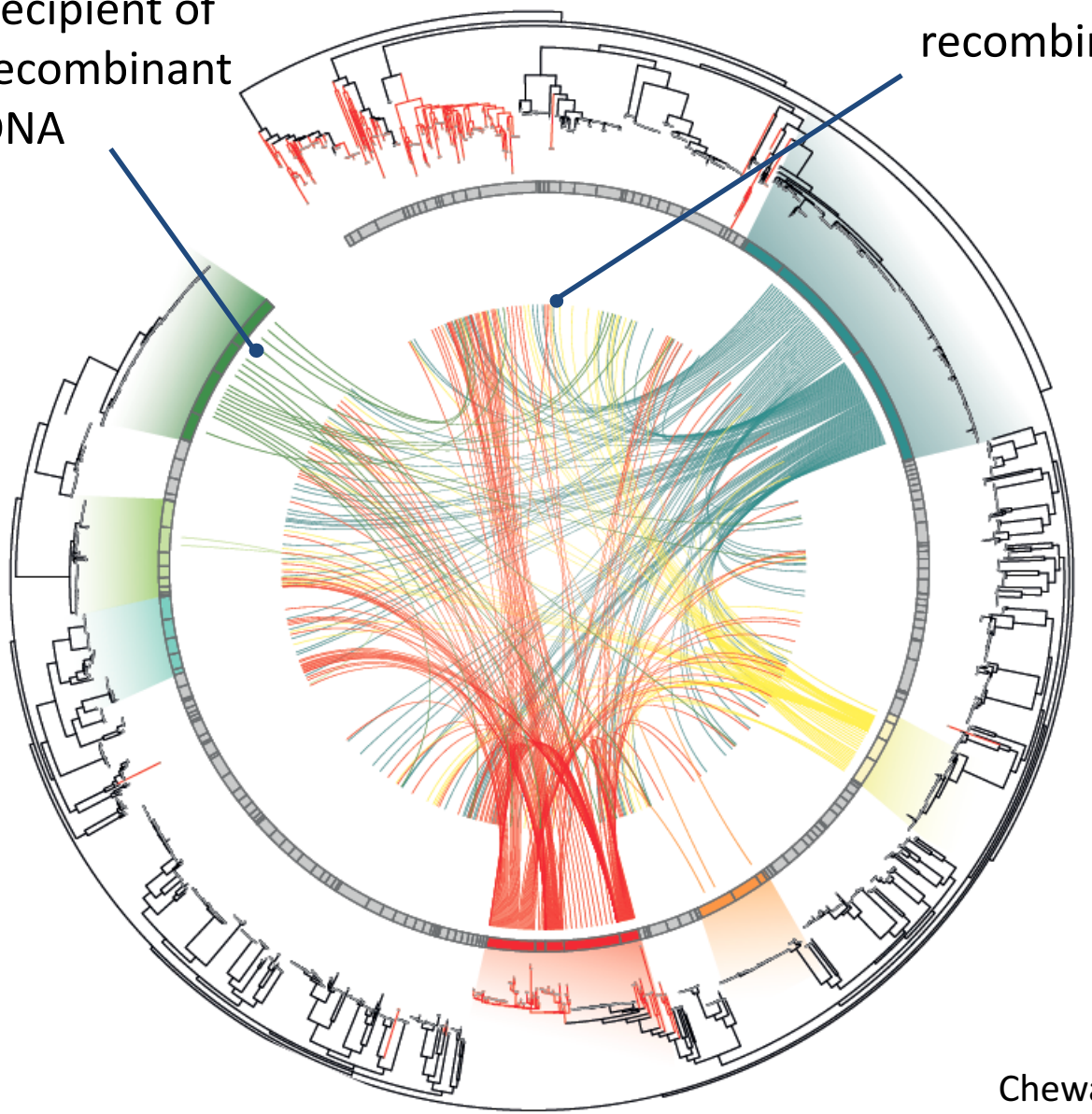
Chewapreecha *et al*. Nature Genetics 2014



Chewapreecha *et al*. Nature Genetics 2014



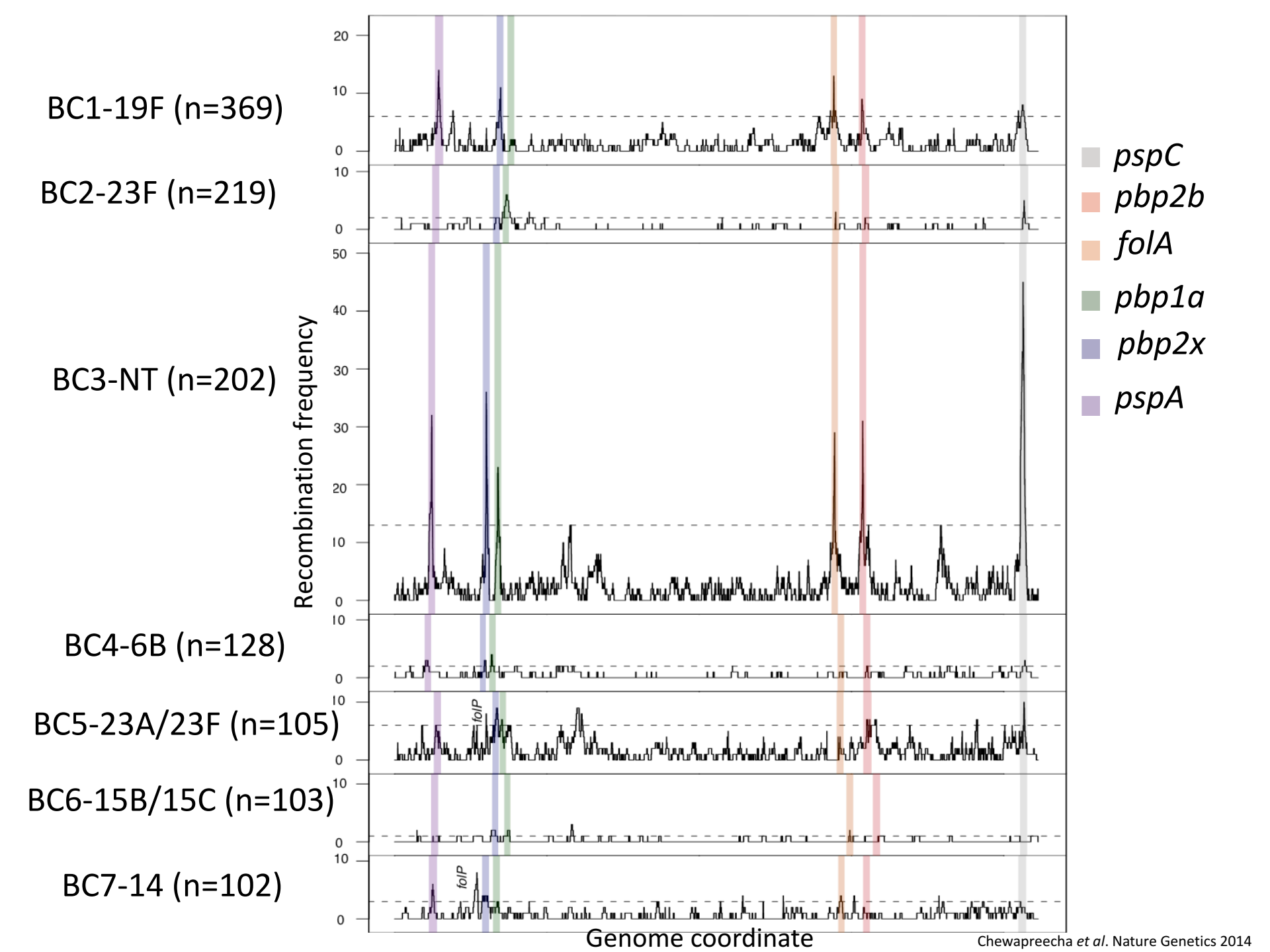
Chewapreecha *et al*. Nature Genetics 2014



- Positive correlation between the probability of becoming donors and cluster population size (Spearman's rank correlation, p-val = 2.69 x 10^{-4})

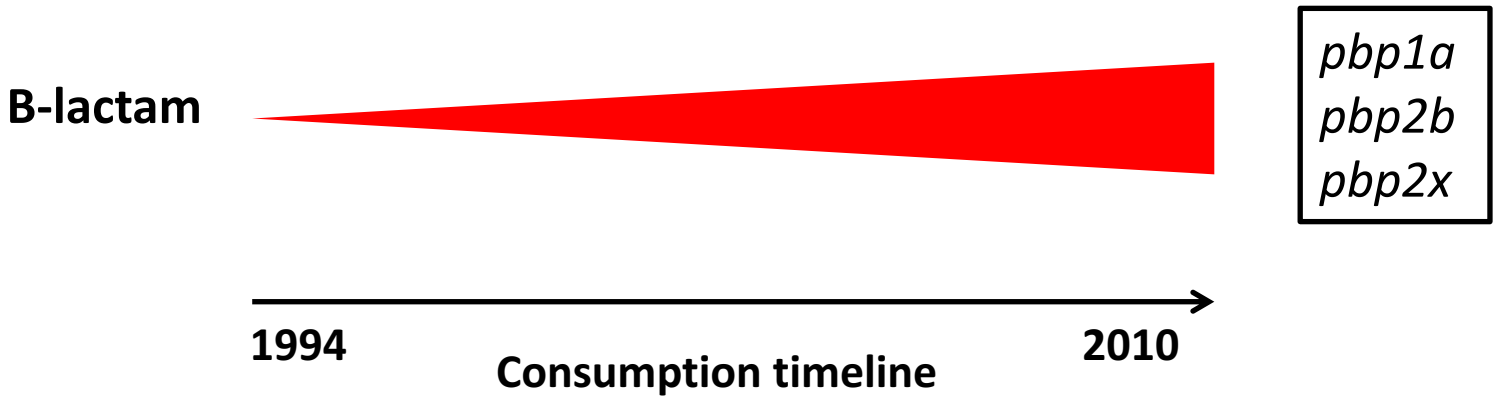
-Positive correlation between the probability of becoming donors and diversity observed within the clusters (Spearman's rank correlation, p-val = 1.45 x 10^{-6})

Chewapreecha *et al*. Nature Genetics 2014

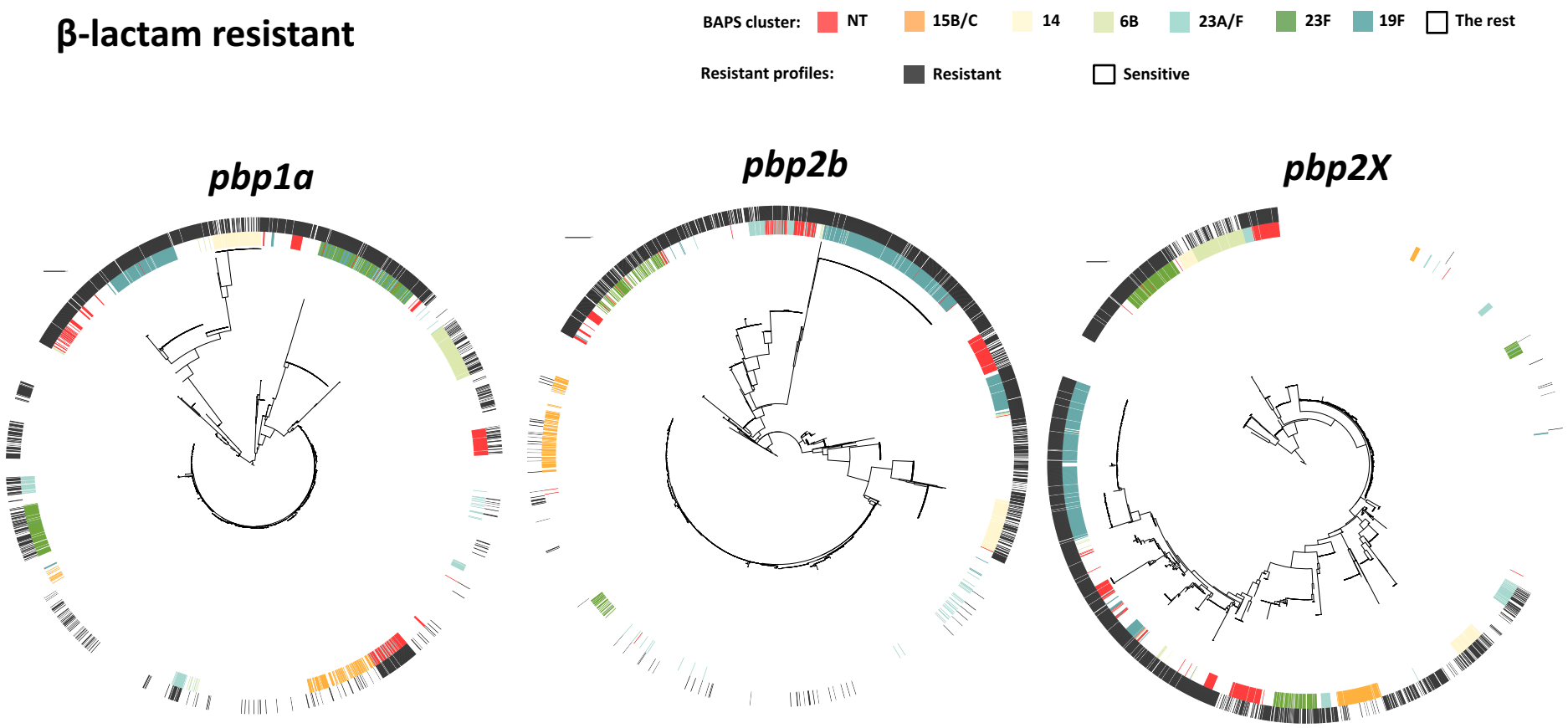


Chewapreecha *et al*. Nature Genetics 2014

Clinical Records



β-lactam resistant



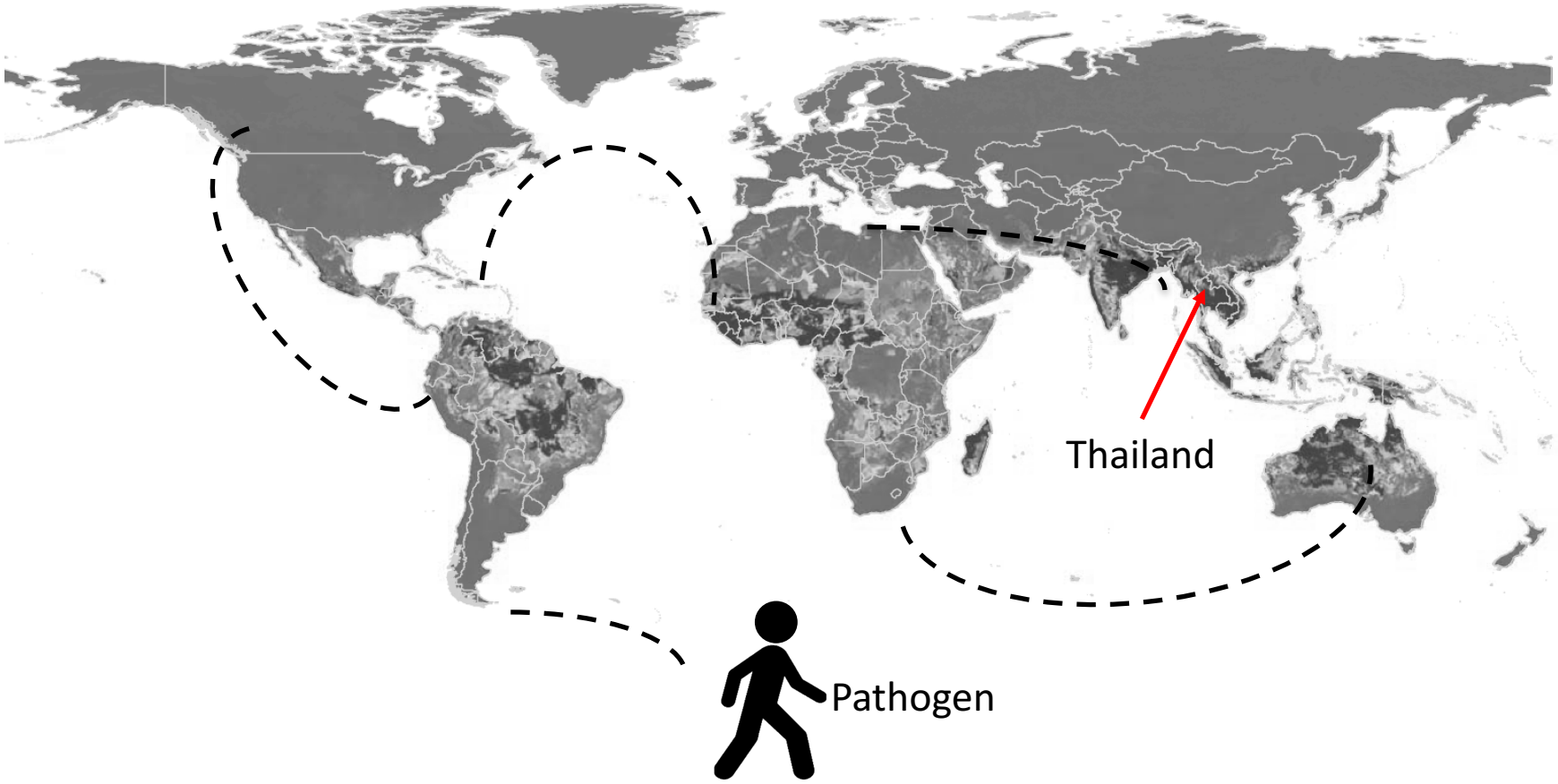
- Isolates undergone recombination at either *pbp1a, pbp2b* and *pbp2x* are phenotypically more resistant to B-lactam compared to those not having undergone recombination (Fisher's exact test: p-val < 2.2e-16)

How whole genome sequencing can help tackle antibiotic resistance and dissemination of the deadly bacteria?

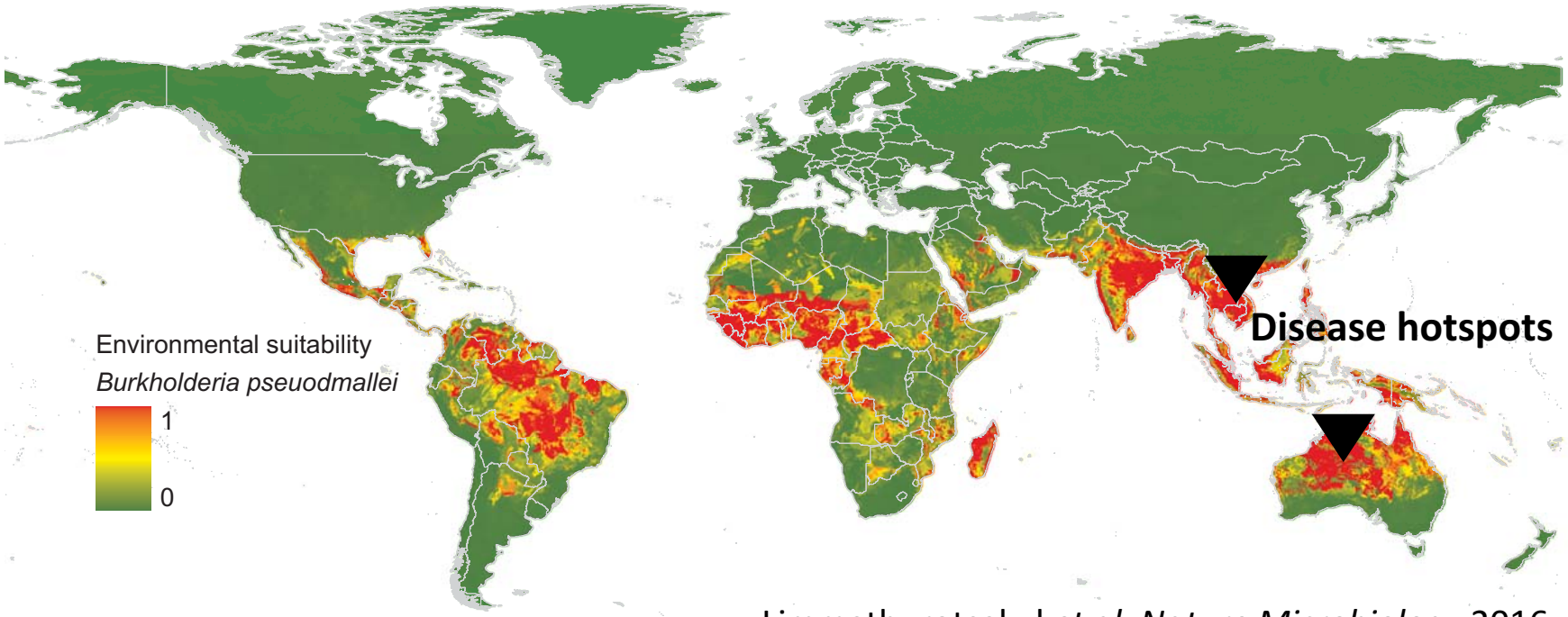
- By providing high resolution for identification of novel resistance alleles.

-By providing high resolution to dissect the mechanism that mediate the resistance alleles; and key players in the bacterial population.

How whole genome sequencing can help tackle antibiotic resistance and dissemination of the deadly bacteria?


Thailand
Pathogen

Burkholderia pseudomallei: A global burden



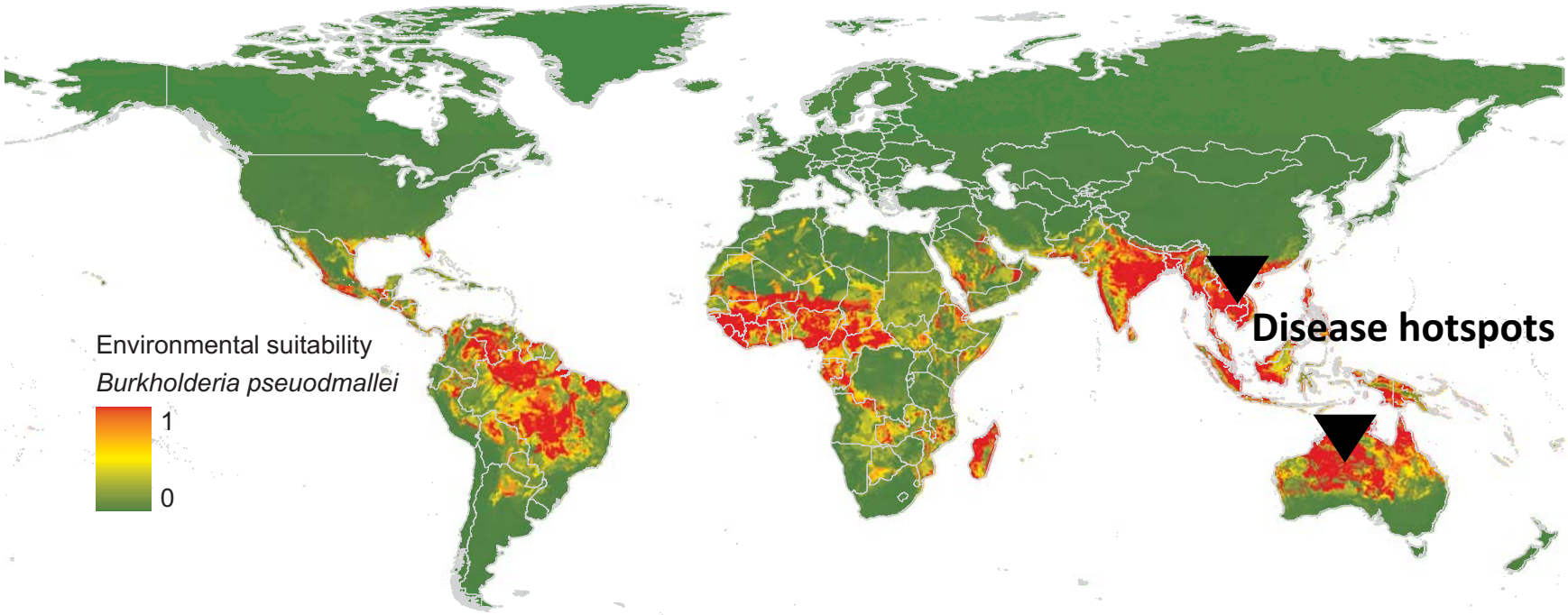
Limmathurotsakul *et al. Nature Microbiology* 2016

-An estimated 165,000 cases of human melioidosis per year world wide
-89,000 were predicted to be fatal.
-*B. pseudomallei* is soil bacterium and thus should be geographically restricted. How did it travel globally?

KEEP CALM AND TRACK TRANSMISSION

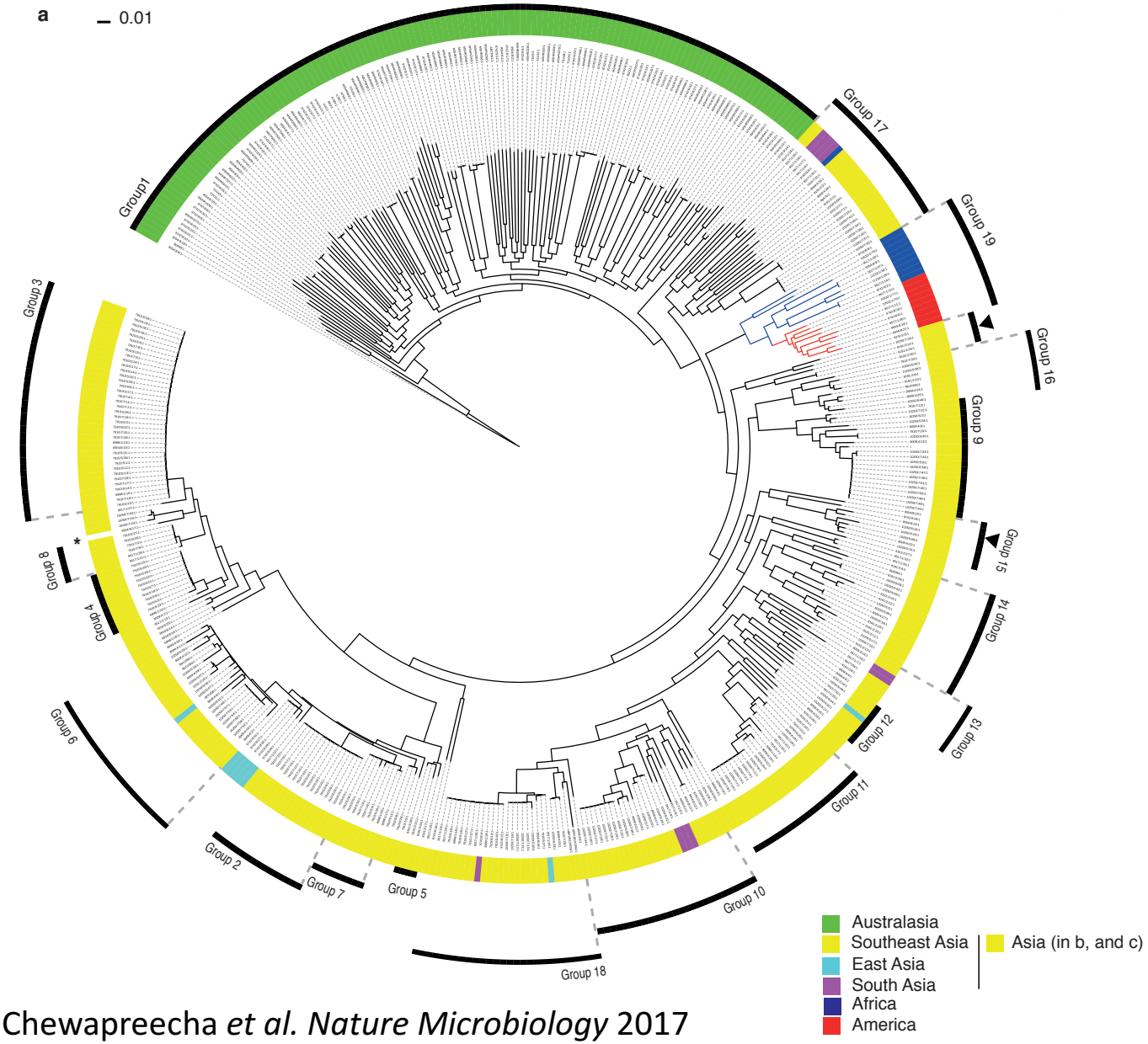
Prof Sharon Peacock

A global burden

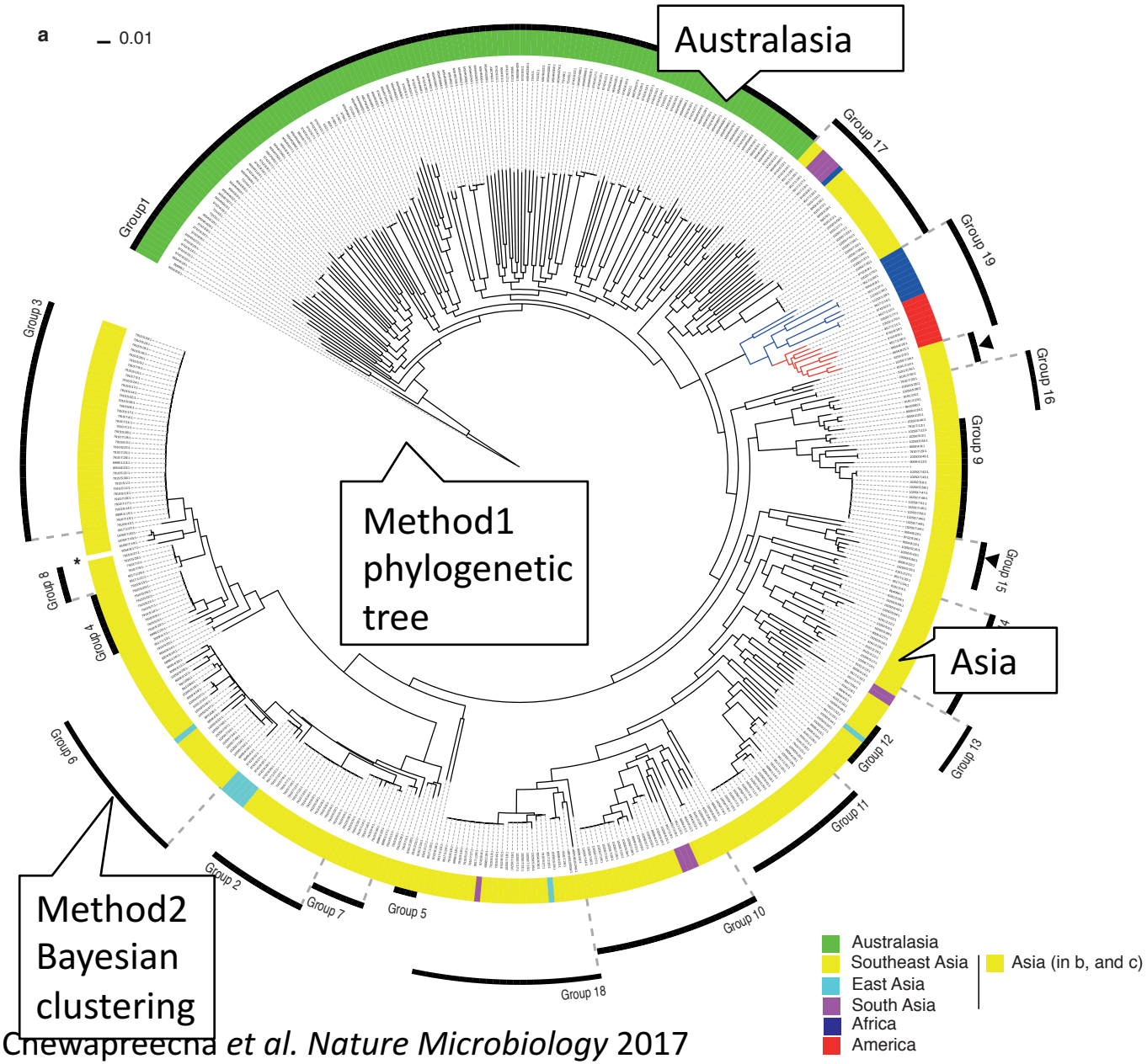


Limmathurotsakul *et al. Nature Microbiology* 2016

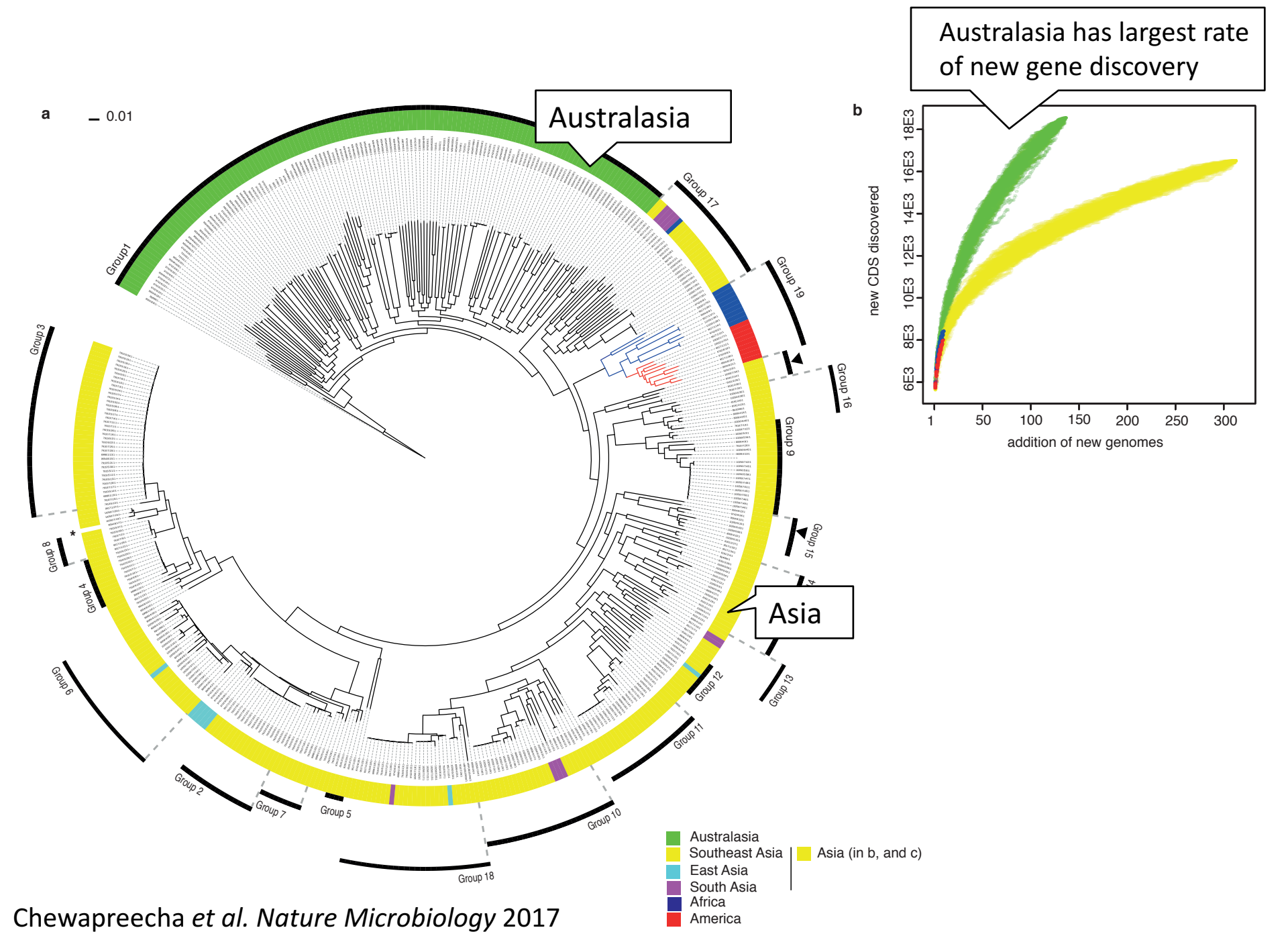
How did *B. pseudomallei* travel globally?
Data:
-whole-genome sequences of 469 isolates
-collected from 30 countries between 1935-2013



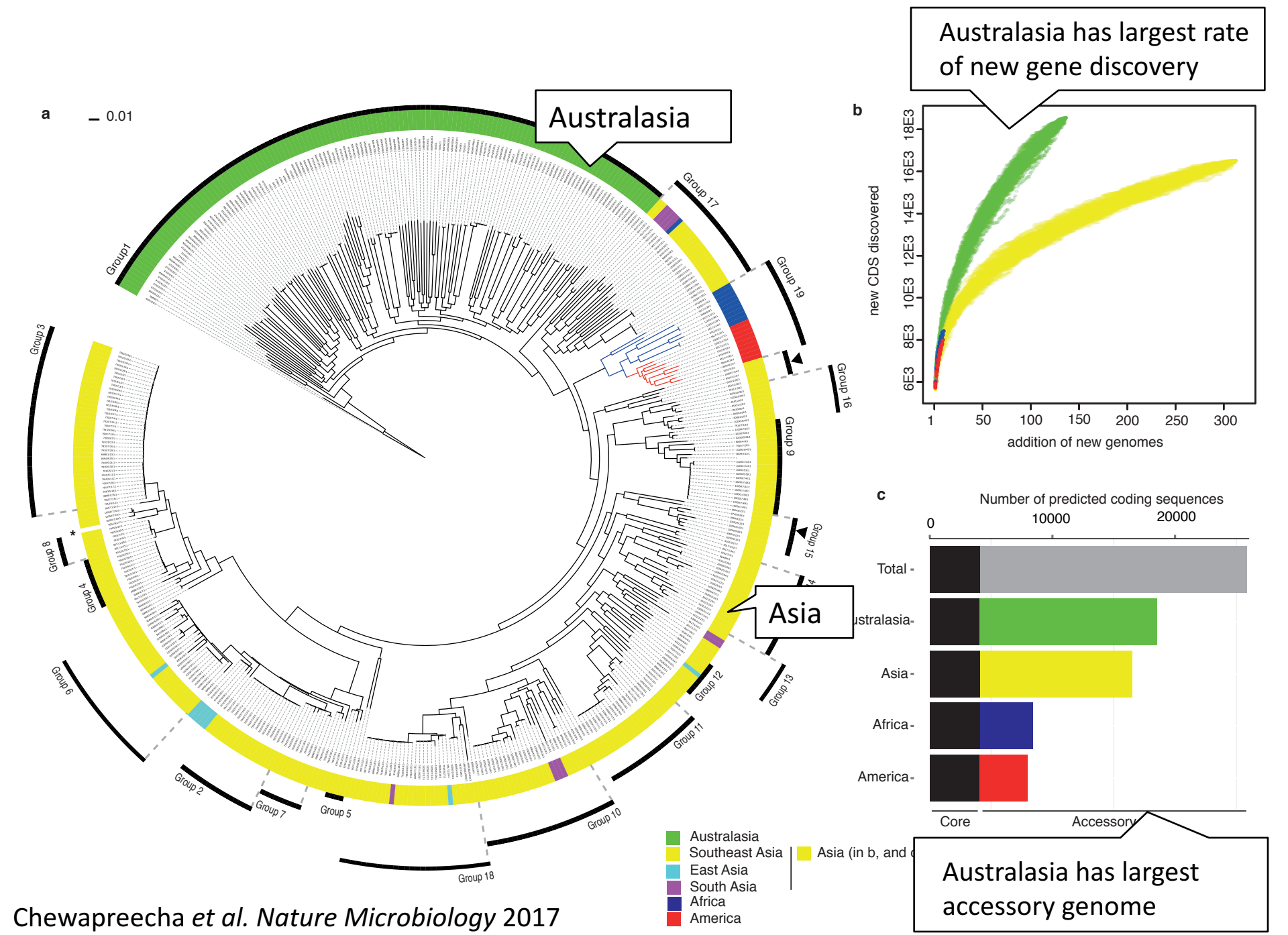
Chewapreecha *et al. Nature Microbiology* 2017



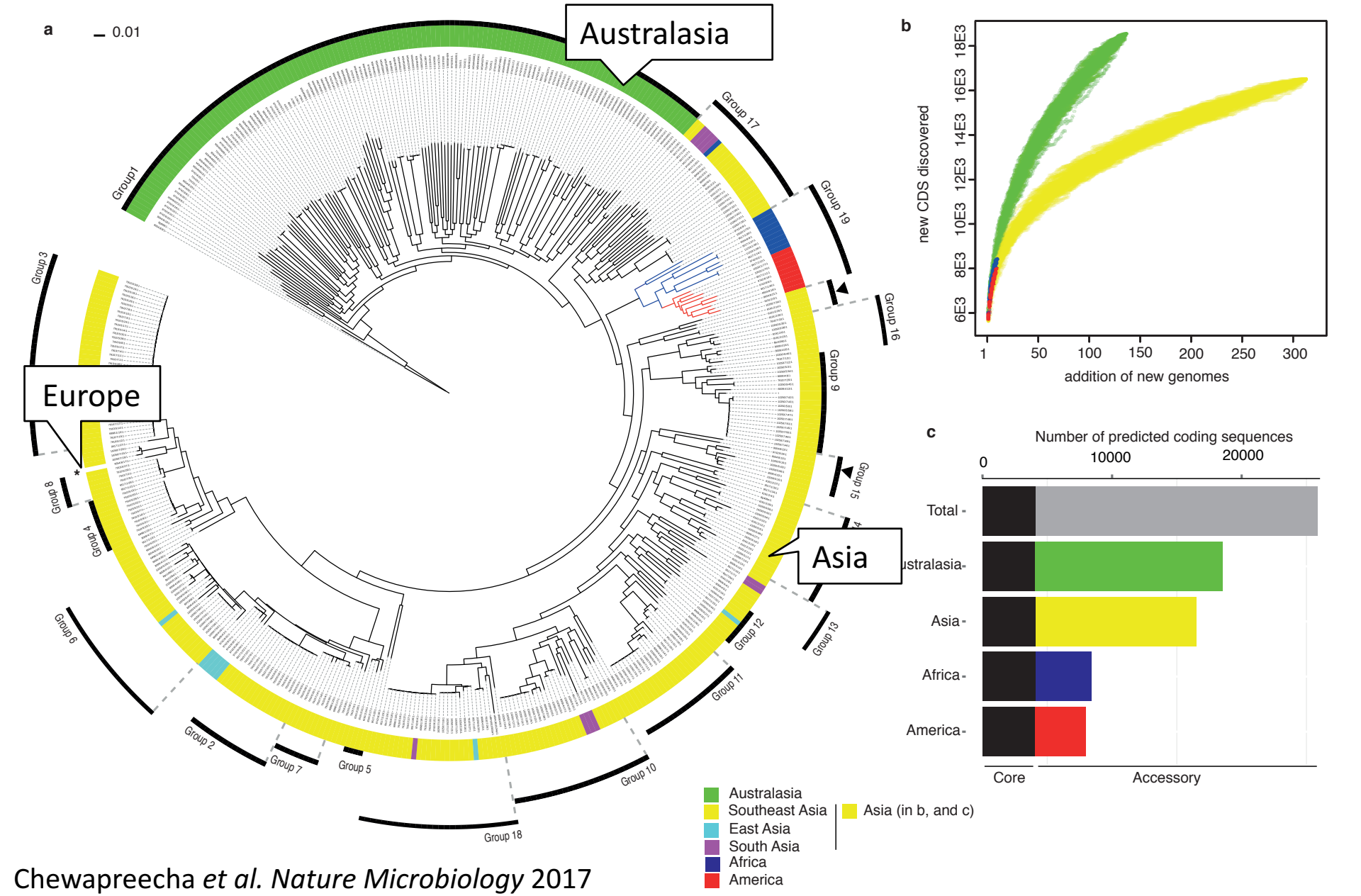
Chewapreecha *et al. Nature Microbiology* 2017



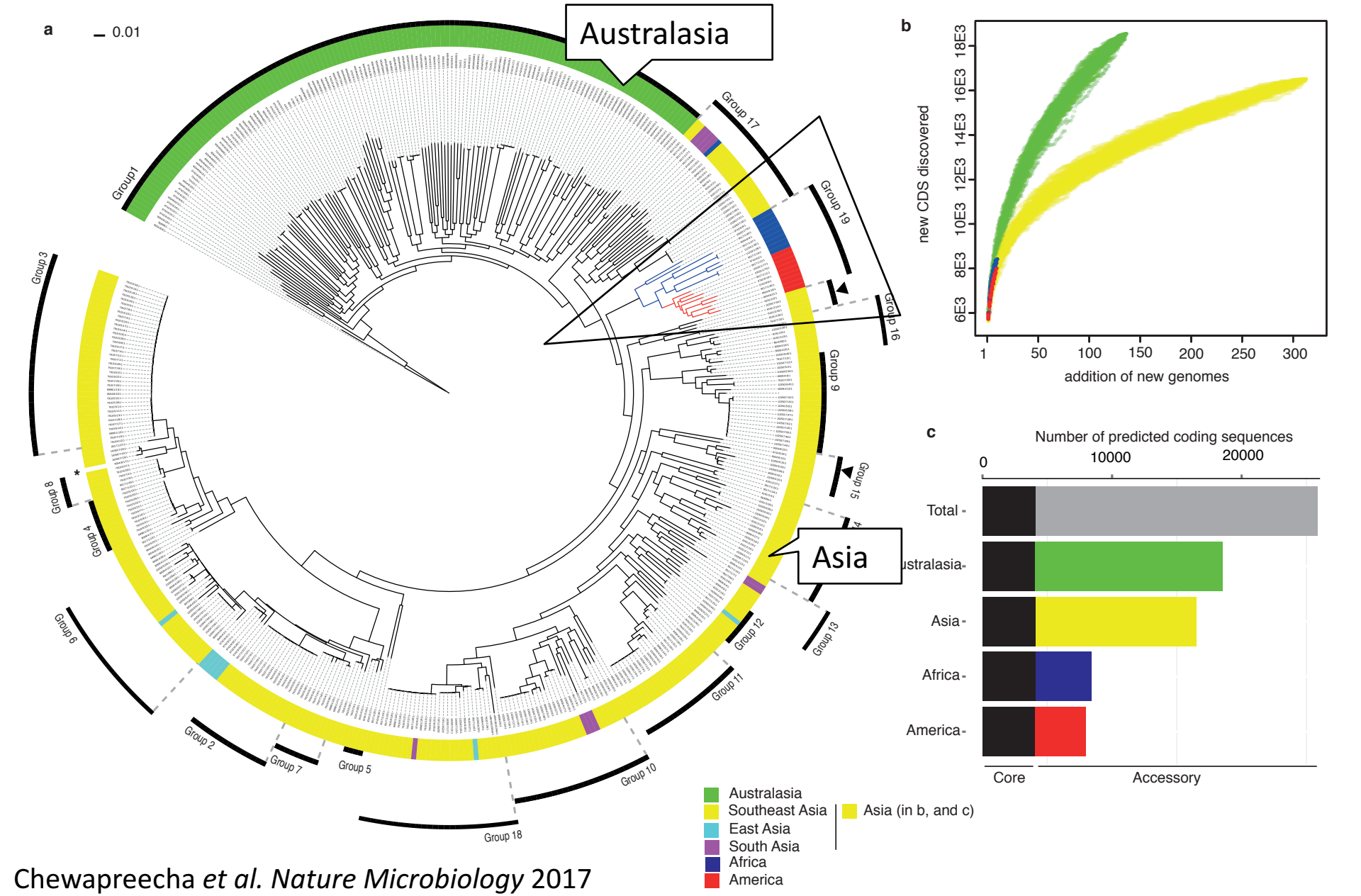
Chewapreecha *et al. Nature Microbiology* 2017



Chewapreecha *et al. Nature Microbiology* 2017

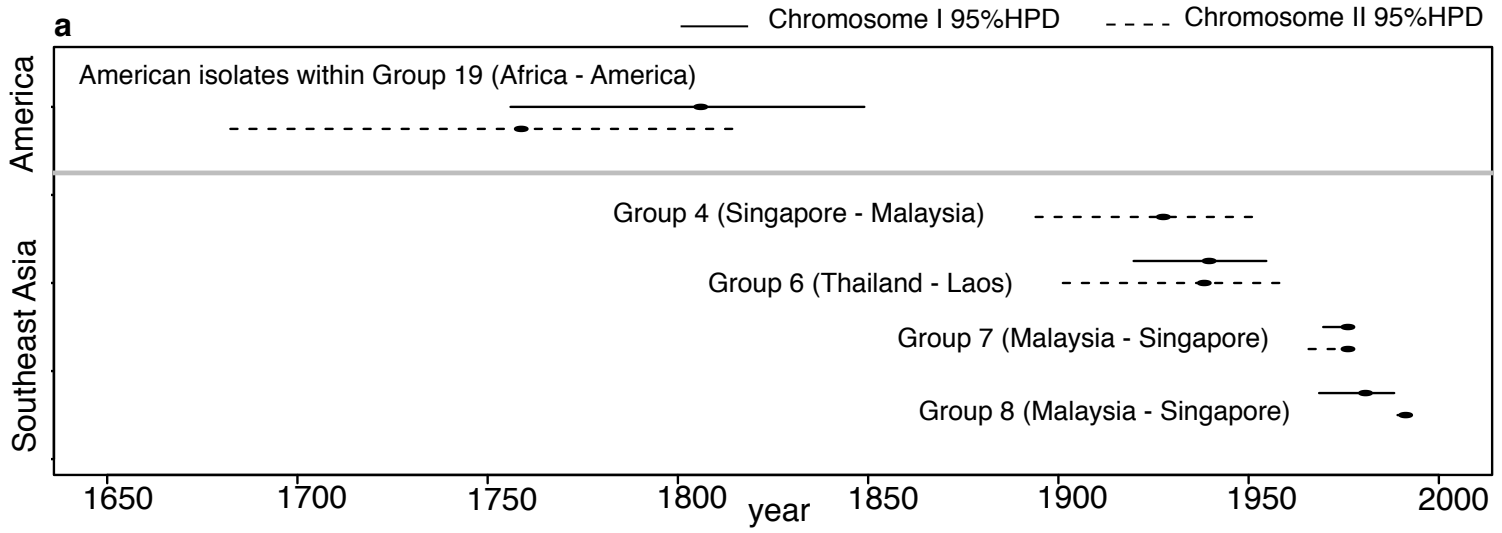


Chewapreecha *et al. Nature Microbiology* 2017

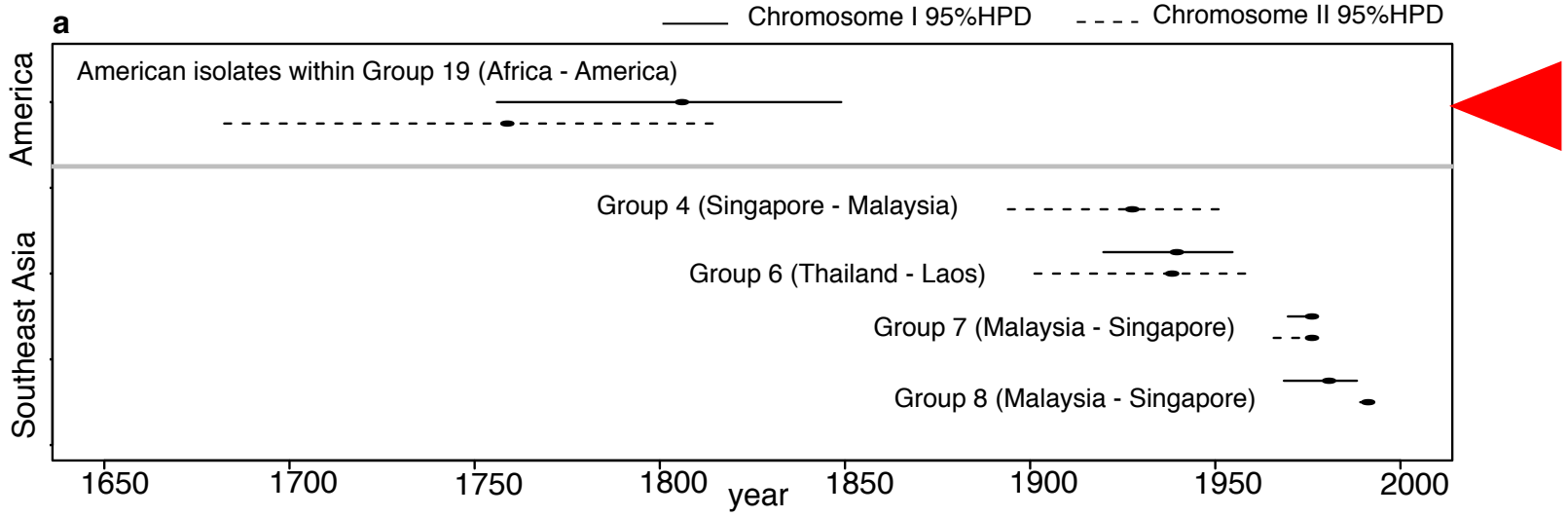


Chewapreecha *et al. Nature Microbiology* 2017

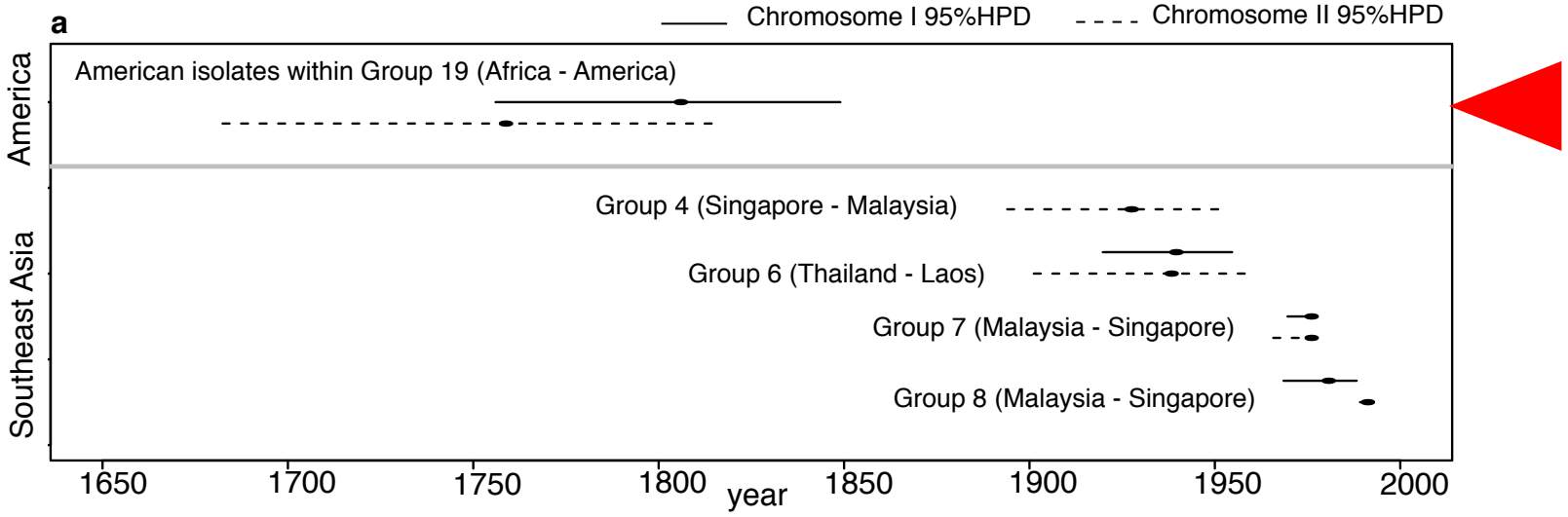
Historical timeline

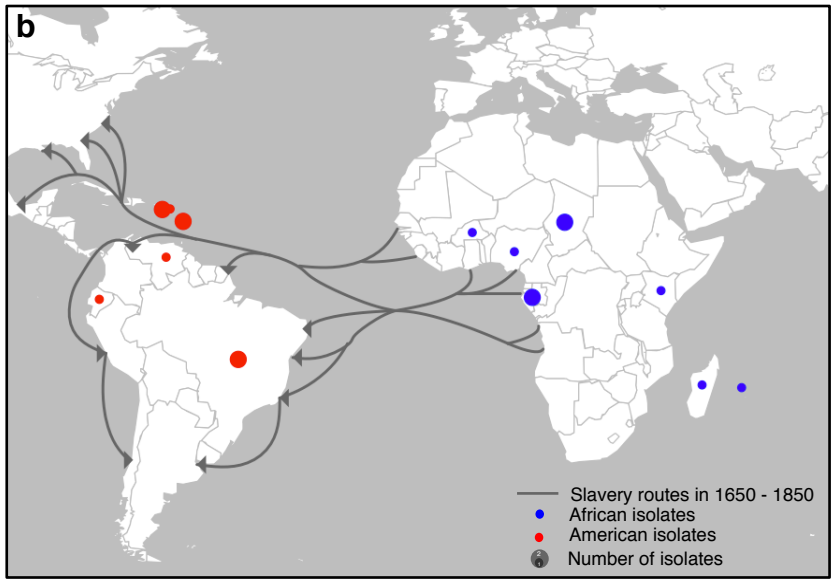


Historical timeline



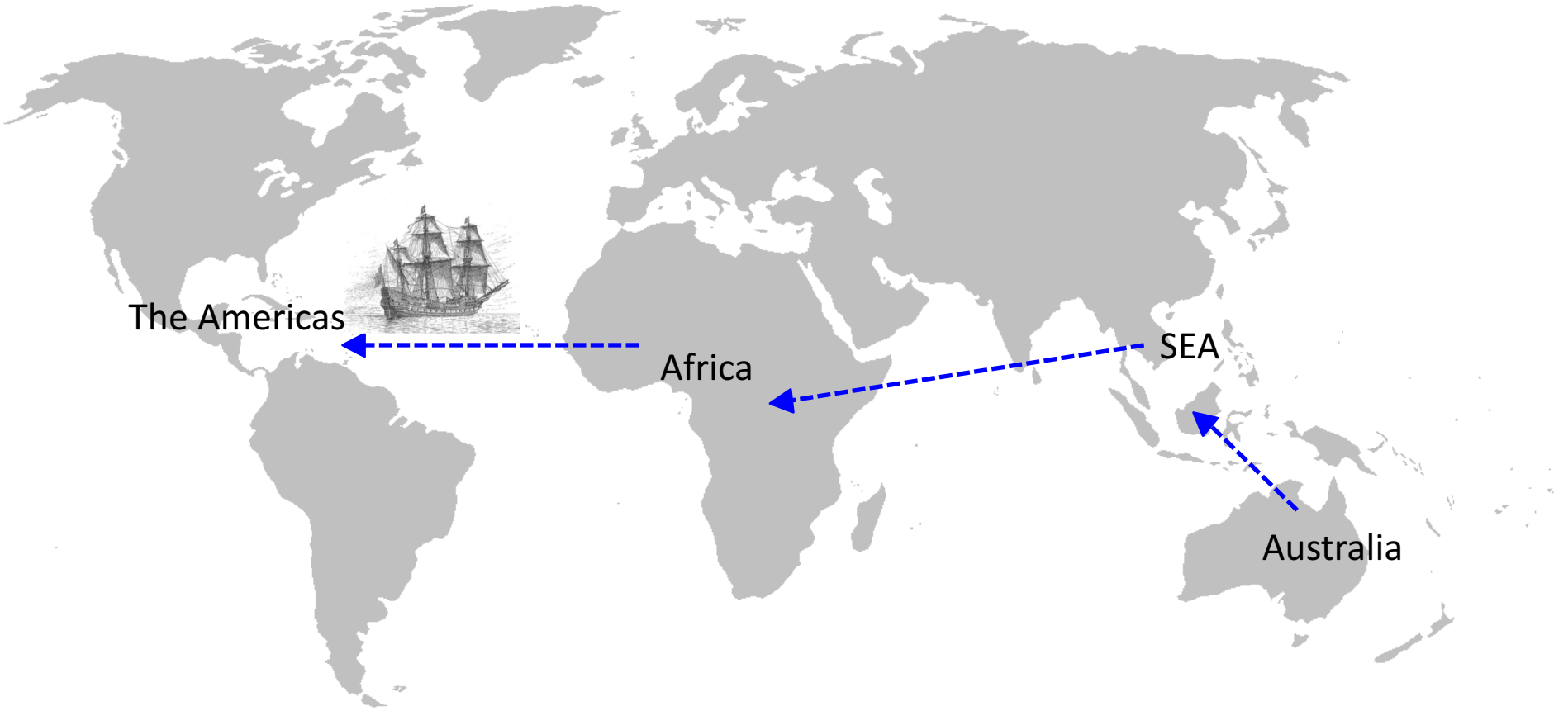
Historical timeline





Slave trade (1650-1850) was estimated to bring 10-15 millions from Africa to the Americas

How whole genome sequencing can help tackle antibiotic resistance and dissemination of the deadly bacteria?



Chewapreecha *et al. Nature Microbiology* 2017

How whole genome sequencing can help tackle antibiotic resistance and dissemination of the deadly bacteria?

-Identify genetic culprits for AMR
-Understand the mechanism of how bacteria develop AMR
-Track how bacteria (and AMR alleles) spread

wellcome trust MORU Tropical Health Network
UNIVERSITY OF OXFORD
wellcome trust sanger institute
KMUTT
wellcome trust Fellow

Acknowledgement

Menzies School of Health Research:
Bart Currie (Australia)

Academic Medical Centre:
Joost Wiersinga (The Netherlands)

Wellcome Trust MORU:
Nick Day (UK/Thailand)
Direk Limmathurotsakul (Thailand)
Vanaporn Wuthiekanun (Thailand)

Wellcome Trust LOMWRU :
David Dance (Laos)
Paul Newton (Laos)

Wellcome Trust COMRU :
Varun Kumar (Cambodia)

KEMRI Wellcome Trust:
Susan Morphet (Kenya)

NCTC: Julie Russell (UK)

CDC: Apichai Tuanyok (USA)

Malawi-Liverpool Wellcome Trust Unit:
Nick Feasey (UK)

NHS Blood and Transplant:
Ty Pitt (UK)

University of Helsinki:
Jukka Corander (Finland)
Zhirong Yang (Finland)
Minna & Niko (Finland)
Elina Numinen (Finland)

Wellcome Trust SMRU
Paul Turner
Claudia Turner
Francois Nosten

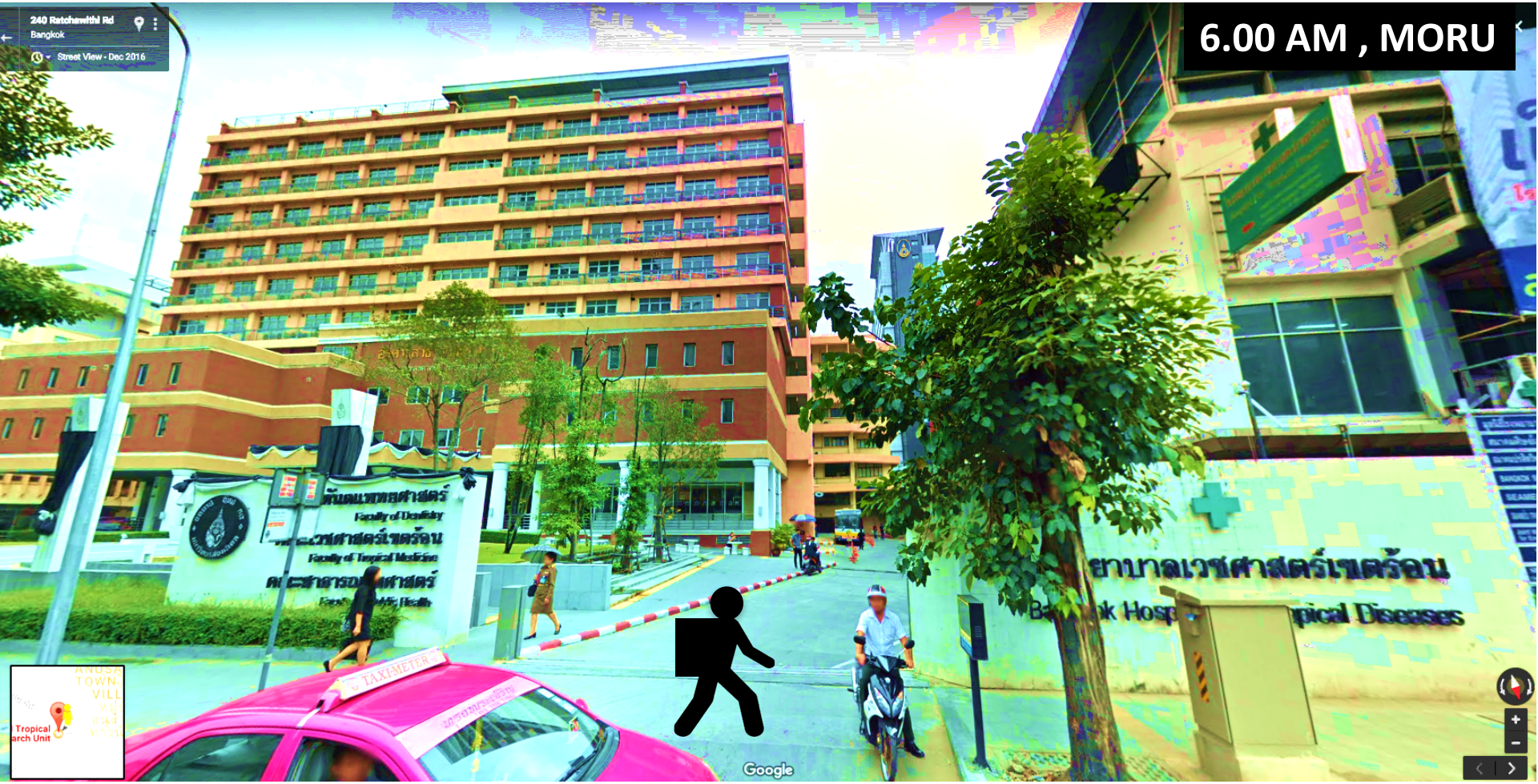
University of St Andrew's:
Matthew Holden (UK)

Wellcome Trust Sanger :
Julian Parkhill (UK)
Simon Harris (UK)
Stephen Bentley (UK)

University of Cambridge:
Sharon Peacock (UK)

THANK YOU SO MUCH!

I am around.



googlemap

Me, a regular ~~intruder~~ visitor at 6.00 am

Sequencing Systems

Sequencing systems for every lab, application, and scale of study.

From the power of the HiSeq X to the speed of MiSeq, Illumina has the sequencer that's just right for you.

MiSeq
Focused power. Speed and simplicity for targeted and small genome sequencing.

NextSeq 500
Flexible power. Speed and simplicity for everyday genomics.

HiSeq 2500
Production power. Power and efficiency for large-scale genomics.

HiSeq X*
Population power. $1,000 human genome and extreme throughput for population-scale sequencing.

Key applications	Small genome, amplicon, and targeted gene panel sequencing.	Everyday genome, exome, transcriptome sequencing, and more.		Production-scale genome, exome, transcriptome sequencing, and more.		Population-scale human whole-genome sequencing.
Run mode	N/A	Mid-Output	High-Output	Rapid Run	High-Output	N/A
Flow cells processed per run	1	1	1	1 or 2	1 or 2	1 or 2
Output range	0.3-15 Gb	20-39 Gb	30-120 Gb	10-300 Gb	50-1000 Gb	1.6-1.8 Tb
Run time	5-55 hours	15-26 hours	12-30 hours	7-60 hours	< 1 day - 6 days	< 3 days
Reads per flow cell†	25 Million‡	130 Million	400 Million	300 Million	2 Billion	3 Billion
Maximum read length	2 × 300 bp	2 × 150 bp	2 × 150 bp	2 × 250 bp	2 × 125 bp	2 × 150 bp

http://systems.illumina.com/systems/sequencing.html

The missing genetic “dark matter”?

More loci of smaller effect, might emerge with ever larger sample sizes?

Apparent heritability from epigenetic effects, rather than mutations?

Rare alleles, not represented in SNP screens, maybe with large effects?

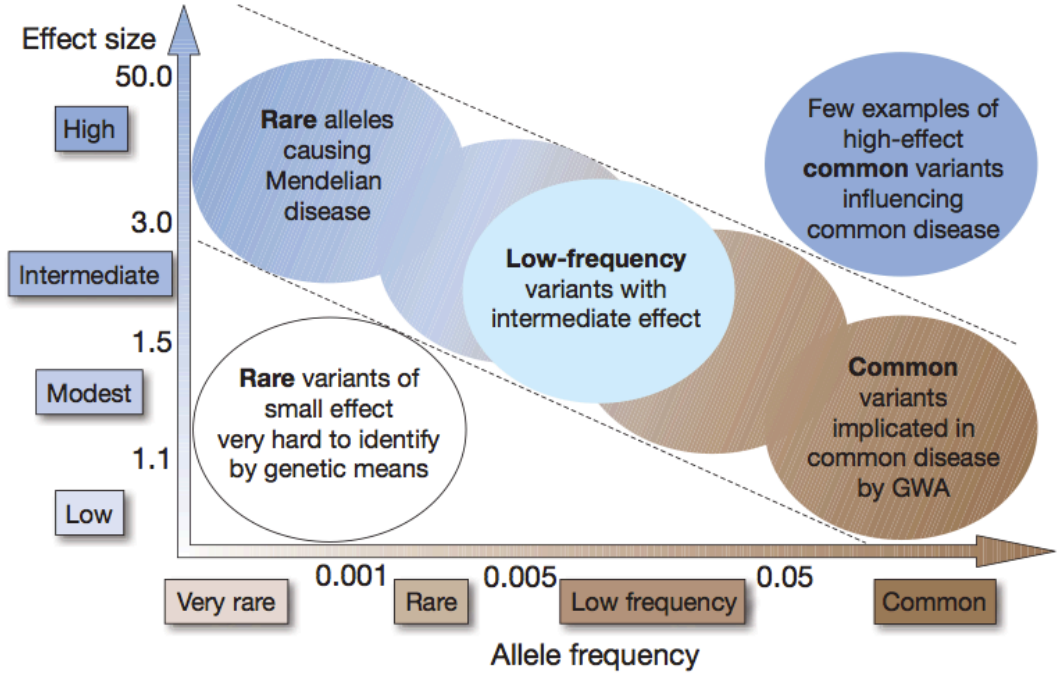
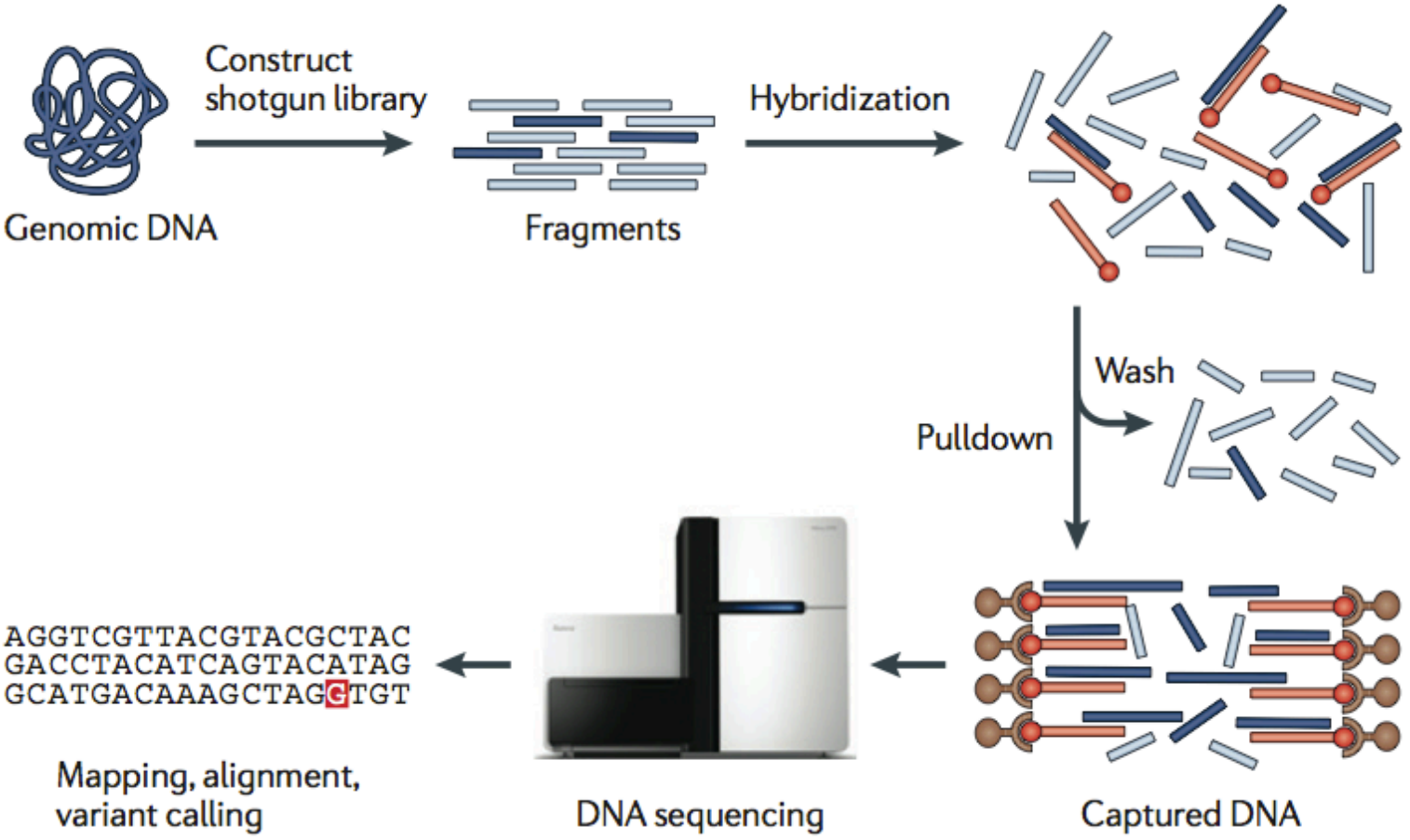


Figure 1 | Feasibility of identifying genetic variants by risk allele frequency and strength of genetic effect (odds ratio). Most emphasis and interest lies in identifying associations with characteristics shown within diagonal dotted lines. Adapted from ref. 42.

Manolio et al (2009) Nature 461:707-13

Detecting (usually rare) mutations of large effect: Exome Sequencing, ~ 2% of genome Faster & cheaper coverage of protein-coding sequence,



Bamshad et al (2011) Nature Reviews Genetics 12:745-755