

My 40 years – from penicillin-binding proteins to molecular epidemiology

Brian G Spratt

Department of Infectious Disease Epidemiology,
Imperial College London,
St. Mary's Hospital Campus.

Mechanism of action of penicillin poorly understood 40 years ago

Tipper and Strominger (1965) had proposed:

Penicillin inhibited the final transpeptidation (cross-linking) step in cell wall (peptidoglycan) synthesis

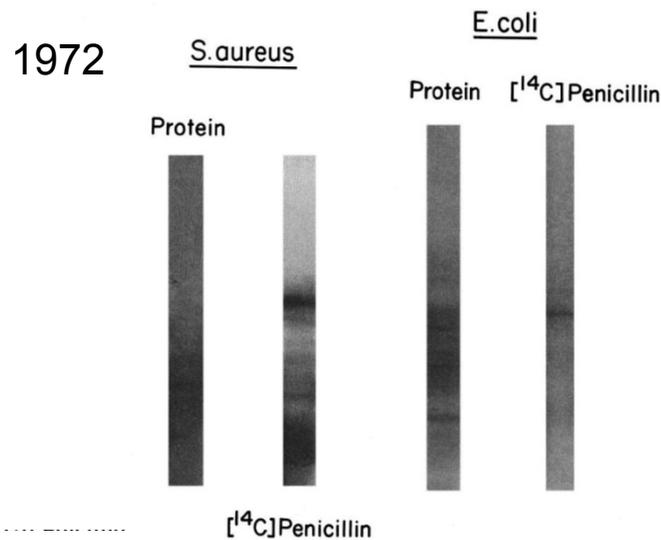
Predicted penicillin would bind covalently (irreversibly) to the cell wall transpeptidase

Tipper DJ, Strominger JL. 1965. Mechanism of action of penicillins: a proposal based on their structural similarity to acyl-D-alanyl-D-alanine. *Proc Natl Acad Sci U S A.* **54**: 1133-1141.

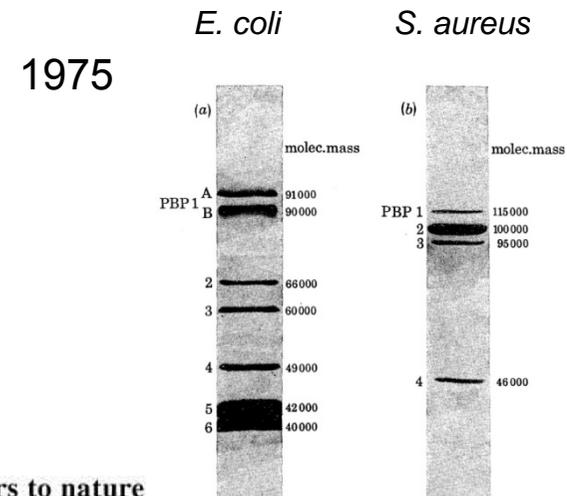
Efforts to identify cell wall transpeptidase by its predicted ability to bind penicillin irreversibly

Seemed to be multiple proteins that bound radioactive penicillin – Jack Strominger 's lab

No convenient way of detecting and studying them



Suginaka et al., 1972. Multiple Penicillin-binding Components in *Bacillus subtilis*, *Bacillus cereus*, *Staphylococcus aureus*, and *Escherichia coli*. *J. Biol. Chem.* **247**: 5279-5288.



letters to nature

Nature **254**, 516 - 517 (10 April 1975); doi:10.1038/254516a0

Penicillin-binding proteins and cell shape in *E. coli*

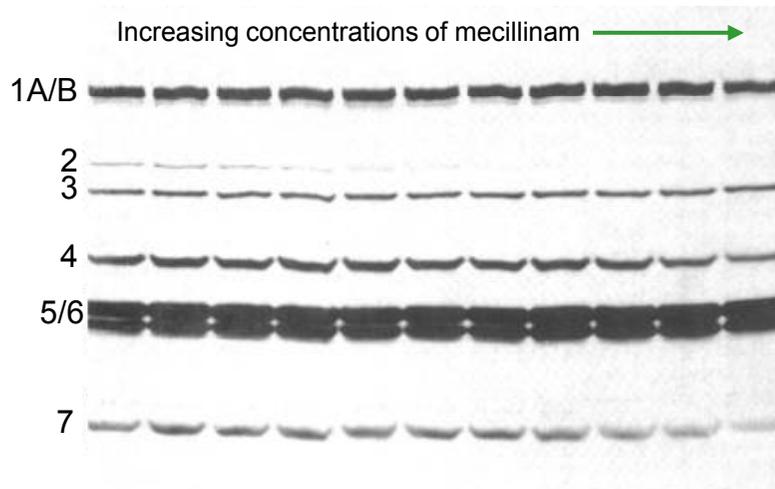
BRIAN G. SPRATT* & ARTHUR B. PARDEE*

2

*Moffett Laboratories, Department of Biochemical Sciences, Princeton University, Princeton, New Jersey 08540

Used new PBP assay to dissect role of different PBPs

1. Binding affinities of β -lactams to individual PBPs



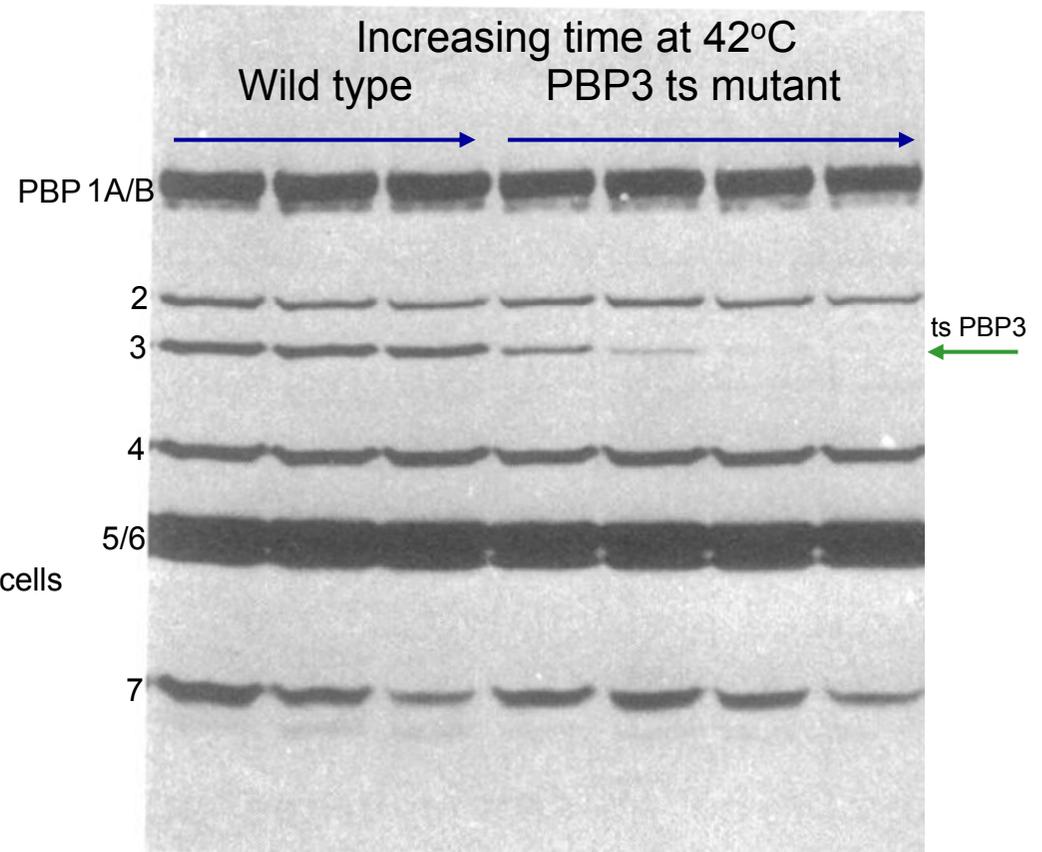
Growth in mecillinam resulted in growth as spherical cells

- and bound only to PBP 2

→ PBP 2 somehow involved in cell shape

Spratt B.G. and Pardee A.B. 1975. Penicillin-binding proteins and cell shape in *E. coli*. *Nature* **254**: 516-517

2. Isolation of mutants producing temperature-sensitive PBPs



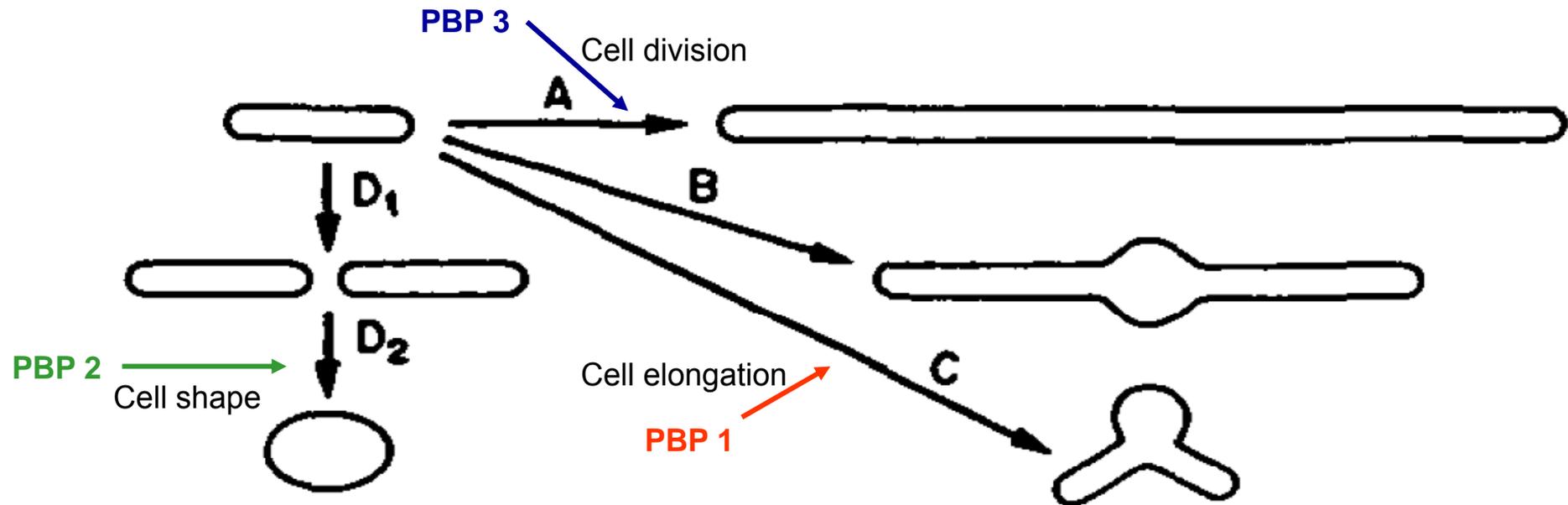
Mutant forms long filamentous cells when PBP3 is inactivated at 42°C

→ PBP 3 involved in cell division

Spratt B. G. 1977. Temperature-sensitive cell division mutants of *Escherichia coli* with thermolabile penicillin-binding proteins. *J Bacteriol.* **131**: 293-305.

Model for how penicillin produces its multiple effects on the growth of *E. coli*

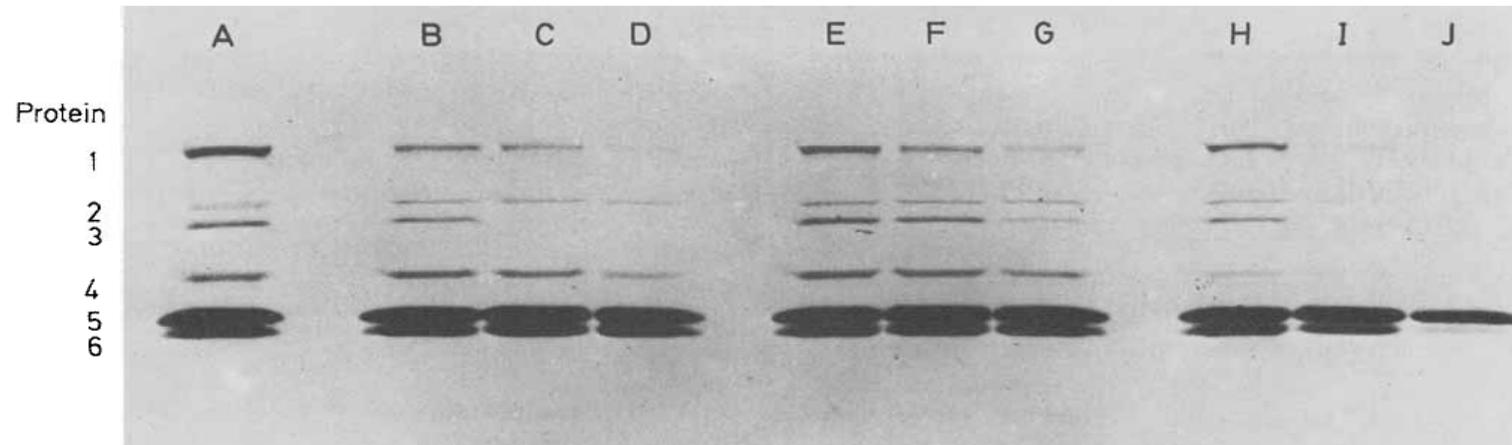
Multiple PBPs (transpeptidases) with different roles in synthesising the cell wall



Spratt, B.G. 1975. Distinct penicillin-binding proteins involved in the division, elongation and shape of *Escherichia coli* K12. Proc. Natl. Acad. Sci. USA **72**: 2999-3003.

Convenient PBP assay allowed:

1. Pharma to examine activity of new β -lactams against individual target enzymes



Spratt, B.G. 1977. Properties of the penicillin-binding proteins of *Escherichia coli* K12. Eur. J. Biochem. **72**: 341-352.

2. Discovery of role of target (PBP) changes in resistance to β -lactams in clinical isolates

Could select β -lactam-resistant mutants in the laboratory that had altered PBPs

Spratt, B.G. 1978. *Escherichia coli* resistance to β -lactam antibiotics through a decrease in the affinity of a target for lethality. Nature **274**:713-715.

Hedge P,J., and Spratt B,G. 1985. Resistance to β -lactam antibiotics by re-modelling the active site of an *E. coli* penicillin-binding protein. Nature. **318**: 478-480.

Lab experiments with *E. coli* - what about clinical isolates?

Involvement of PBPs in resistance to penicillin

1980 – recognition that resistance to penicillin in clinical isolates can be due to alterations of PBPs

1. Acquisition of a novel β -lactam-refractory PBP

Staphylococcus aureus - MRSA

Brown, D.F. and Reynolds, P.E. 1980. Intrinsic resistance to β -lactam antibiotics in *Staphylococcus aureus*. FEBS Lett. **122**: 275-278.

“Therefore the expression of resistance to benzylpenicillin (in a MRSA strain) appears to correlate with the production of large amounts of a protein resistant to the binding of benzylpenicillin, and this protein is present in addition to the normal PBP 3”. (Now called PBP 2a)

2. Reductions in affinity of PBPs for penicillin

Streptococcus pneumoniae

Hakenbeck R, Tarpay M, Tomasz A. 1980. Multiple changes of penicillin-binding proteins in penicillin-resistant clinical isolates of *Streptococcus pneumoniae*. Antimicrob Agents Chemother. **17**: 364-371.

Neisseria gonorrhoeae

Dougherty, T.J., Koller, A.E. and Tomasz, A. 1980. Penicillin-binding proteins of penicillin-susceptible and intrinsically resistant *Neisseria gonorrhoeae*. Antimicrobial Agents and Chemother. **18**, 730-737.

Low affinity PBPs in some penicillin-resistant *Neisseria gonorrhoeae*

Very little binding of penicillin to PBPs 1 or 2 in a penicillin-resistant clinical isolate even at $1\mu\text{g ml}^{-1}$

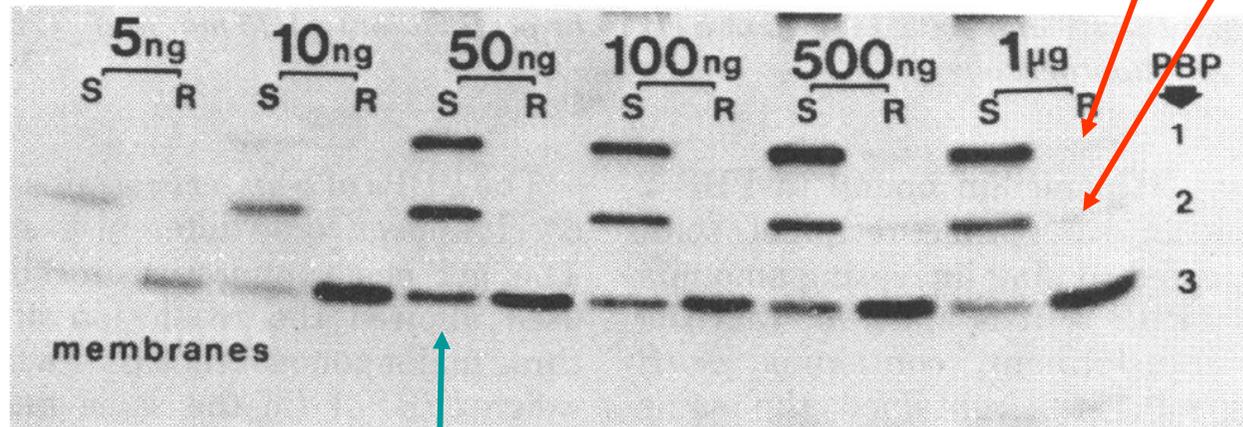
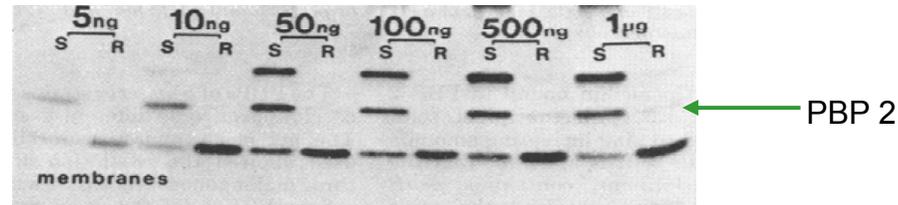


FIG. 5. Binding of [^3H]penicillin to cell membranes. Crude cytoplasmic membranes ($200\mu\text{g}$ of protein) of FA 19 (lanes S) and CDC 77 (lanes R) were incubated for 15 min at 37°C with different [^3H]benzylpenicillin concentrations. Membranes were treated with Sarkosyl and prepared for electrophoresis and fluorography as described in the text.

All three gonococcal PBPs saturated (inhibited) with penicillin (i.e. inhibited) at 50 ng ml^{-1} in a penicillin-susceptible strain

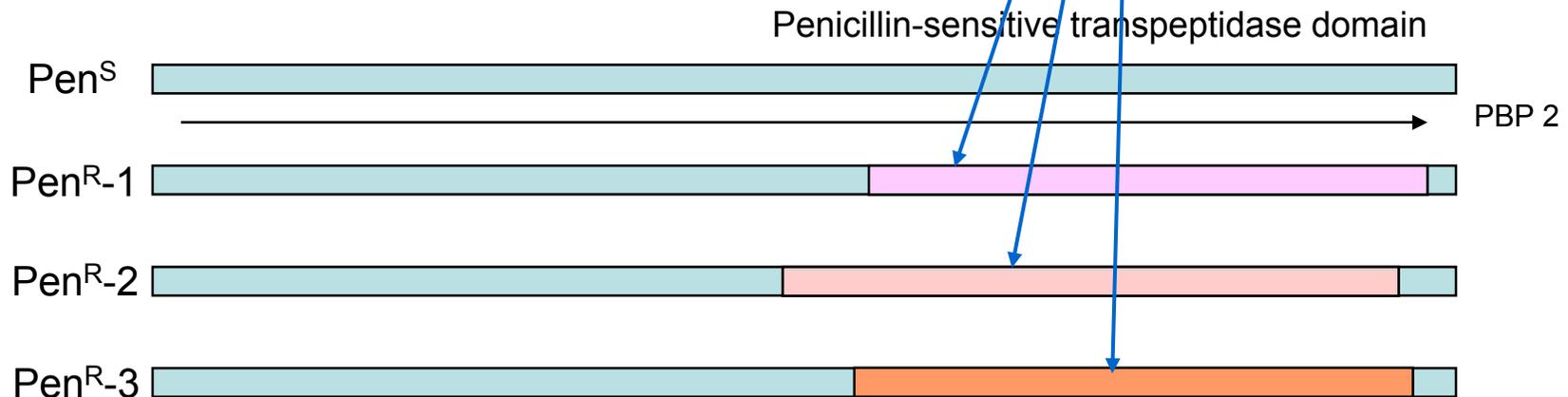
Dougherty, T.J., Koller, A.E. and Tomasz, A. 1980. Penicillin-Binding Proteins of penicillin-susceptible and intrinsically resistant *Neisseria gonorrhoeae*. *Antimicrob. Agents and Chemother.* **18**, 730-737.

Molecular basis of low affinity PBPs – mosaic PBP genes



Divergent sequences from commensal *Neisseria*

Cloned and sequenced PBP 2 gene from several pen^S and pen^R gonococci



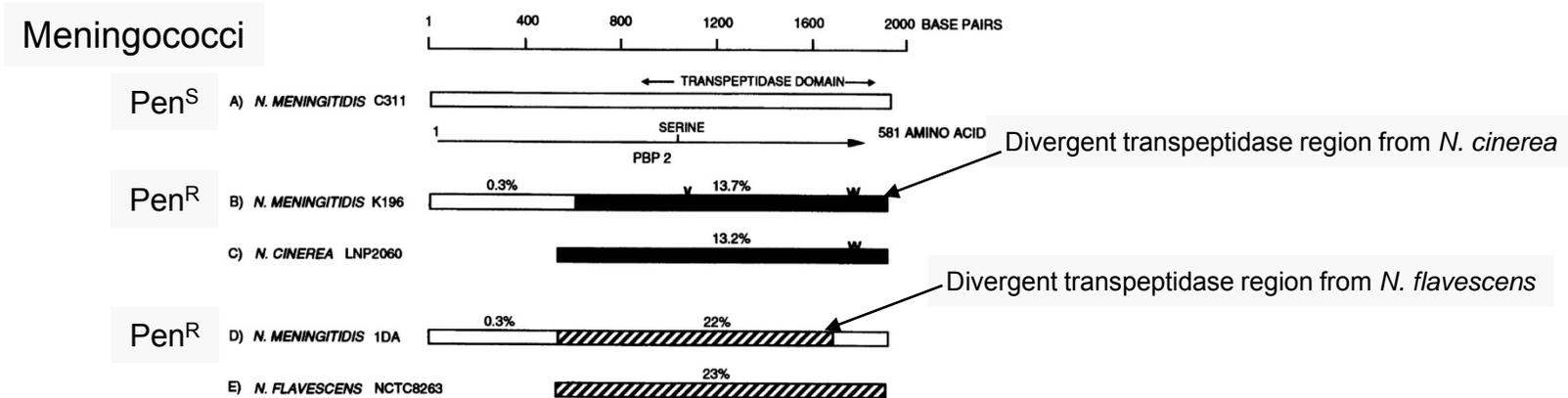
Low affinity of PBP 2 for penicillin not due to accumulation of multiple amino acid changes by mutation

Replacement by homologous recombination with corresponding region from closely related commensal species – hybrid or mosaic PBP genes

Spratt, B.G. 1988. Hybrid penicillin-resistant proteins in penicillin-resistant strains of *Neisseria gonorrhoeae*. Nature **332**: 173-6.

Hybrid or mosaic PBP2s also found in penicillin-resistant meningococci and pneumococci

Meningococci with MICs of up to 1 µg penicillin/ml reported in late 1980's



Spratt, B.G. *et al.*, 1989. Recruitment of a penicillin-binding protein gene from *Neisseria flavescens* during the emergence of penicillin resistance in *Neisseria meningitidis*. *Proc. Natl. Acad. Sci. U S A.* **86**: 8988-8992.

Very similar story in pneumococci

Penicillin resistance due to formation of mosaic PBP genes

Dowson, C.G., Hutchison, A. and Spratt, B.G. 1989. Extensive re-modelling of the transpeptidase domain of penicillin-binding protein 2B of a penicillin-resistant South African isolate of *Streptococcus pneumoniae*. *Mol. Microbiol.* **3**: 95-102.

Dowson CG, Hutchison A, Brannigan JA, George RC, Hansman D, Liñares J, Tomasz A, Smith JM, and Spratt BG. 1989. Horizontal transfer of penicillin-binding protein genes in penicillin-resistant clinical isolates of *Streptococcus pneumoniae*. *Proc. Natl. Acad. Sci. U S A.* **86**: 8842-8846.

~4000-fold increase in MICs for β-lactams in most highly resistant pneumococci

– resistance all due to (multiple) PBP changes

PBP-mediated resistance in gonococci, meningococci and pneumococci

Several questions arose:

Have resistant strains emerged once and spread?

Or have they emerged independently in several countries?

If so, how are resistant strains in UK related to those in other countries?

How best can we track their spread?

Do resistant PBP genes spread or only strains?

More general question of tracking the spread of 'significant' strains of pathogens

= Molecular typing and molecular epidemiology

In 1990's no shortage of molecular typing methods

JOURNAL OF CLINICAL MICROBIOLOGY, July 1996, p. 1870
0095-1137/96/\$04.00+0
Copyright © 1996, American Society for Microbiology

Vol. 34, No. 7

A Surfeit of YATMs?

Mark Achtman

YATM \yat' em, *Brit.* ya' tem\ *n* acronym for *Yet Another Typing Method*

TATBSTM \tate' bee stem, *Brit.* tat' bistem\ *n* acronym for *Tried And True But Stodgy Typing Method*

TBCA \tib' see ay, *Brit.* tib' ka\ *n* acronym for *Totally Boring Clonal Analysis*

I have become increasingly distressed by the number of YATMs described in recent years in the *Journal of Clinical Microbiology*. Although tradition has dictated comment on specific articles, this letter addresses articles too numerous to mention individually and summarizes the limitations and benefits of YATMs, TATBSTMs, and TBCAs.

YATMs are generally based on DNA technology (e.g., ribotyping, random amplified polymorphic DNA analysis, or pulsed-field gel electrophoresis polymorphism) and are designed to rapidly distinguish clonal epidemic outbreaks from hyperendemic disease levels caused by concurrent multiplication of unrelated strains. In many cases in which the results from YATMs were compared with those from TBCAs, the TBCA was as discriminatory as the YATM or the YATM was based on genetically hypervariable properties. Most YATMs have been applied to limited numbers of isolates isolated locally within a short time. Only rarely have globally relevant collections spanning decades been tested, and provisions for the comparison of results between independent laboratories

Many typing methods published – most useless or unsound

Needed a good method for 'Long term' or 'Global epidemiology'

For example:-

Relatedness of MRSA strains from hospitals in different countries

Global spread of drug-resistant pneumococci

Comparing isolates before and after introduction of a new vaccine

'Short term' (outbreak) epidemiology is easy - essentially no difference in genotype is expected for isolates from an outbreak (PFGE etc work fine)

Global epidemiology is more difficult, as isolates of the same strain recovered on different continents or in different decades may have a common ancestor that emerged many years ago

Isolates of the same strain from different countries, or time periods, may have accumulated considerable variation in PFGE profiles

Need to use genetic variation that accumulates relatively slowly

Molecular typing for global epidemiology

Wanted to use 'digital data' to characterise strains

- digital data = precise data

Allows labs to communicate - to have a common language –
e.g. to recognise whether they are talking about the same MRSA strain

- harness power of internet

Interrogate molecular epidemiological databases via the web

DNA sequences are best – unambiguous and understand how sequence variation accumulates

Sequence data can also be used for exploring pathogen population and evolutionary biology

Need to use DNA sequence variation that accumulates slowly

For example, sequence variation in house-keeping genes (e.g. metabolic genes)

Need a multilocus approach

In most bacterial pathogens there is sufficient genetic variation in house-keeping genes to recognise many different sequences (alleles) among strains of the pathogen

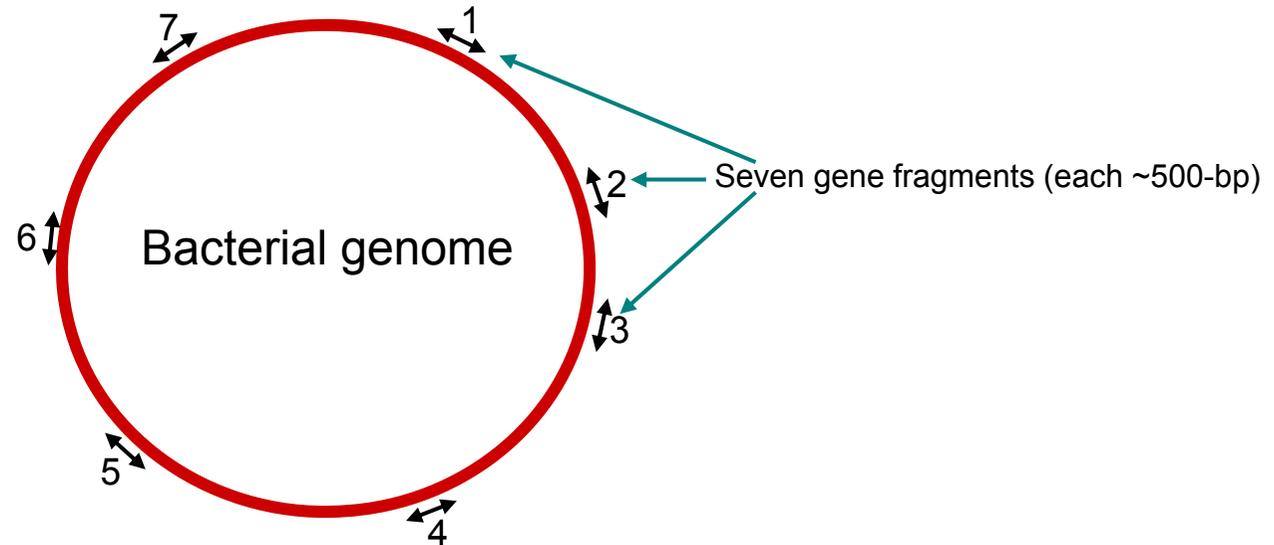
If 40 different alleles (sequences) can be distinguished among strains of a bacterial species by sequencing a single gene

.. with seven genes (loci), each with 40 alleles, can distinguish over 150 billion multilocus genotypes

= Multi-locus sequence typing (MLST)

Maiden, Bygraves, Feil, Morelli, Russell, Urwin, Zhang, Zhou, Zurth, Caugant, Feavers, Achtman and Spratt (1998) PNAS **95**: 3140-3145.

Multi-locus sequence typing

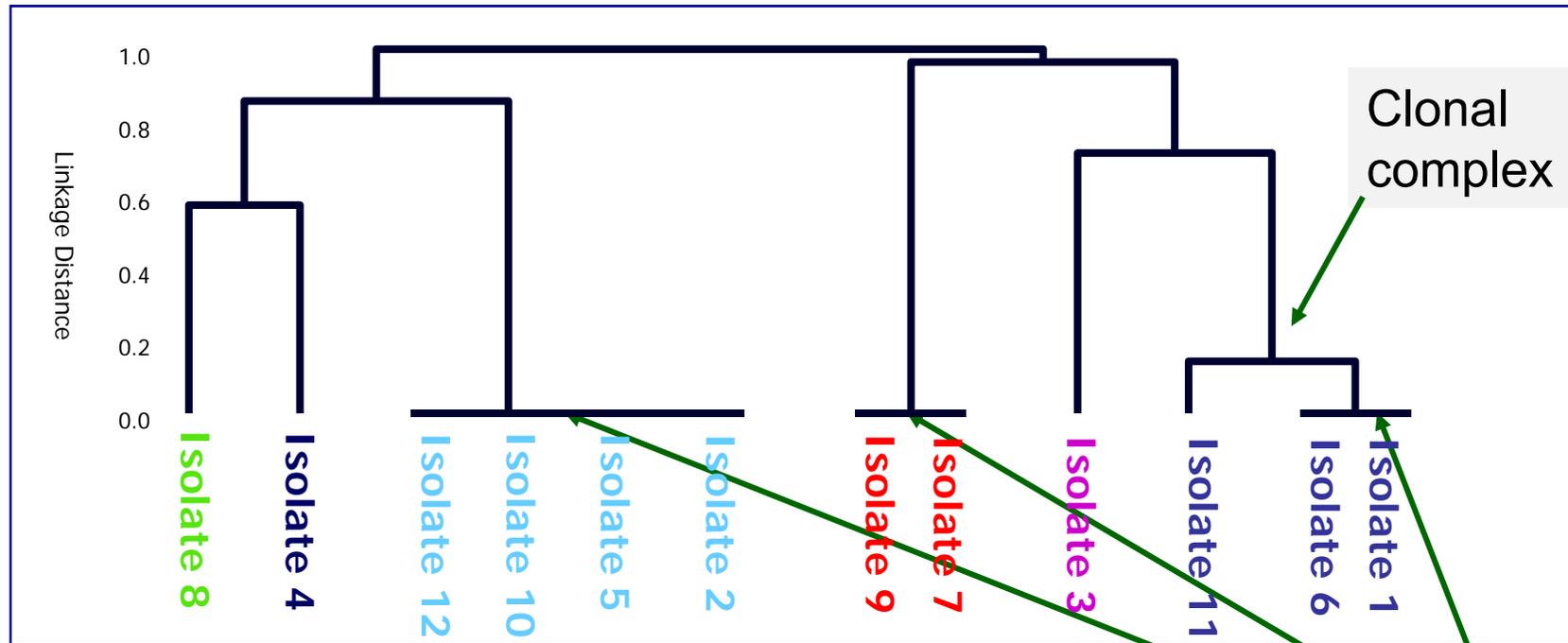


For each isolate, sequence the seven gene fragments (house-keeping genes)

For each gene fragment (locus), each different sequence is given a different allele number

The allele numbers at the seven MLST loci gives the allelic profile of the isolate

Dendrogram based on allelic mismatches between the allelic profiles



Allelic profiles

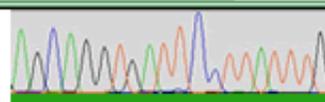
Isolate 1	1	5	3	7	2	8	4
Isolate 2	11	15	30	12	22	11	20
Isolate 3	6	5	9	1	2	13	17
Isolate 4	9	10	1	19	12	18	14
Isolate 5	11	15	30	12	22	11	20
Isolate 6	1	5	3	7	2	8	4
Isolate 7	20	11	29	9	21	13	11
Isolate 8	9	15	13	27	22	18	14
Isolate 9	20	11	29	9	21	13	11
Isolate 10	11	15	30	12	22	11	20
Isolate 11	1	5	3	5	2	8	4
Isolate 12	11	15	30	12	22	11	20

MLST data

Strain/clone

Every different allelic profile is assigned as a sequence type (ST) i.e. as a strain

All data held in online database and can be interrogated through internet



Policy Document

Wednesday 3rd November 2004

DATA ANALYSIS

DATABASES

- B.cereus*
- B.pseudomallei*
- C.albicans*
- C.glabrata*
- C.jejuni*
- E.coli*
- E.faecium*
- H.influenzae*
- H.pylori*
- M.catarrhalis*
- N.meningitidis*
- S.agalactiae*
- S.aureus*
- S.enterica*
- S.epidermidis*
- S.pneumoniae*
- S.pyogenes*
- V.vulnificus*

SUBMISSIONS

NEWS

Streptococcus pneumoniae - Allelic Profile Query ?

Please enter your query below (the figure below each box represents the number of unique alleles in the current database for that locus)-

aroE	gdh_	gki_	recP	spi_	xpt_	ddl_
<input type="text" value="4"/>	<input type="text" value="4"/>	<input type="text" value="2"/>	<input type="text" value="4"/>	<input type="text" value="4"/>	<input type="text" value="1"/>	<input type="text" value="1"/>
67	96	110	70	104	146	161

Please select Query type:

Please choose which dataset to interrogate:

or

Profile Query

Locus Query

DATA ANALYSIS

DATABASES

- B.cereus
- B.pseudomallei
- C.albicans
- C.glabrata
- C.jejuni
- E.coli
- E.faecium
- H.influenzae
- H.pylori
- M.catarrhalis
- N.meningitidis
- S.agalactiae
- S.aureus
- S.enterica
- S.epidermidis
- S.pneumoniae
- S.pyogenes
- V.vulnificus

SUBMISSIONS

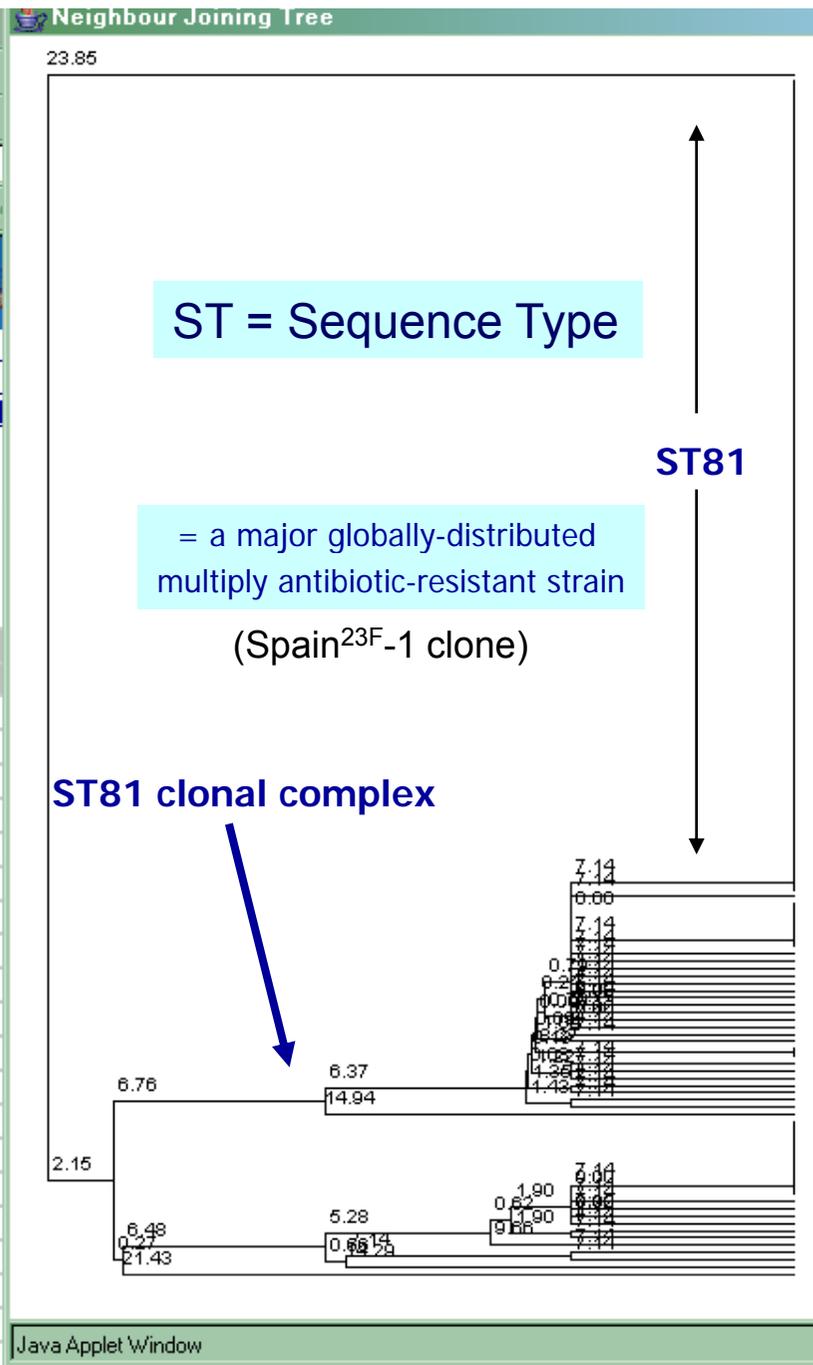
NEWS

LINKS

Streptococcus pneumoniae - Allelic Profiles query resu

Your sequence type is 81

Strain	ST	Serotype	aroe	gdh_	gki_
			4	4	2
SP264-23F	81	23F	4	4	2
UK577-23F	81	23F	4	4	2
GM70-19A	81	19A	4	4	2
GA71-19F	81	19F	4	4	2
M41-23F	81	23F	4	4	2
TW28-19F	81	19F	4	4	2
TW21-19F	81	19F	4	4	2
TW7-23F	81	23F	4	4	2
TW9-19F	81	19F	4	4	2
TW40-23F	81	23F	4	4	2
TW41-23F	81	23F	4	4	2
TW45-23F	81	23F	4	4	2
TW46-23F	81	23F	4	4	2
TW58-19F	81	19F	4	4	2
TW61-23F	81	23F	4	4	2
TW62-23F	81	23F	4	4	2
TW63-19F	81	19F	4	4	2
TW67-23F	81	23F	4	4	2
TW68-23F	81	23F	4	4	2
TW69-23F	81	23F	4	4	2
TW70-23F	81	23F	4	4	2
M329-19F	81	19F	4	4	2



Streptococcus pneumoniae

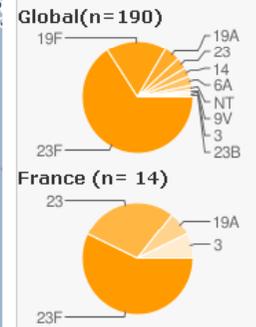


ST Key:

x = number of isolates

Click to return to all isolate map.

Sequence Type 81 Serotype Distribution



Worldwide distribution of ST81 strains

France Sequence Type Summary- [Click to view epidemiological data from isolates from France](#)



- click to view eBURST for Isolates from France only.
- click to view eBURST to compare Isolates from France with the global population
- Select a country to compare Isolates from France to -

[eBURST instructions click here](#)

no. isolates	Sequence Type	aroE	gdh_	gki_	recP	spi_	xpt_	ddl_	other countries (no. isolates elsewhere with same ST)	
45	276	2	19	2	17	6	22	14	4(6)	click to view on map
17	143	7	5	10	18	6	8	1	6(12)	click to view on map
14	81	4	4	2	4	4	1	1	22(178)	click to view on map
9	63	2	5	36	12	17	21	14	7(29)	click to view on map
9	558	18	12	4	44	14	77	97	3(9)	click to view on map
8	156	7	11	10	1	6	8	1	20(72)	click to view on map
8	346	10	41	47	16	6	14	2	3(6)	click to view on map
5	658	7	5	10	18	6	1	103	0	
5	2674	2	19	2	17	6	22	4	0	

MLST very widely used (gold standard) for several major pathogens

Large MLST databases available for many species – mostly pathogenic bacteria but some fungi (e.g. *Candida albicans*)

First MLST scheme - *N. meningitidis* - (Maiden et al., 1998) now very widely used to monitor emergence and spread of new strains and to monitor effects of meningococcal conjugate (MenC) vaccine

Streptococcus pneumoniae

Pneumococcal MLST scheme first used in 1998 to identify major strains associated with invasive disease and for identifying the major antibiotic-resistant strains (Enright and Spratt, 1988)

Enright, M.C., and Spratt, B.G. 1998. A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. *Microbiology* **144**: 3049-3060.

Widely used to monitor effects on new conjugate vaccine – changes in pneumococcal population following vaccination, vaccine escape by serotype switching, differences in attack rates of strains etc

MLST very good at assigning antibiotic-resistant strains to the known resistant strains or as previously unknown resistant strains

Staphylococcus aureus

S. aureus scheme (Enright et al, 2000) widely used to recognise MRSA strains (clones) causing hospital- or community-onset disease in different countries

eMRSA 1,2,3 etc
Berlin strain
Archaic clone
Paediatric clone
New York/Japan clone
Iberian clone
etc

Rationalises chaotic prior nomenclature

MLST data allows evolutionary history of emergence of MRSA strains to be dissected

Enright MC, Day NP, Davies CE, Peacock SJ, Spratt BG. 2000. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. *J. Clin. Microbiol.* **38**: 1008-1015.

Enright, M.C., Robinson, D.A., Randle, G., Feil, E.J., Grundmann, H. and Spratt, B.G. 2002. The evolutionary history of methicillin-resistant *Staphylococcus aureus* (MRSA). *Proc. Natn. Acad. Sci.* **99**: 7687-7692.

Making sense of the molecular epidemiological data

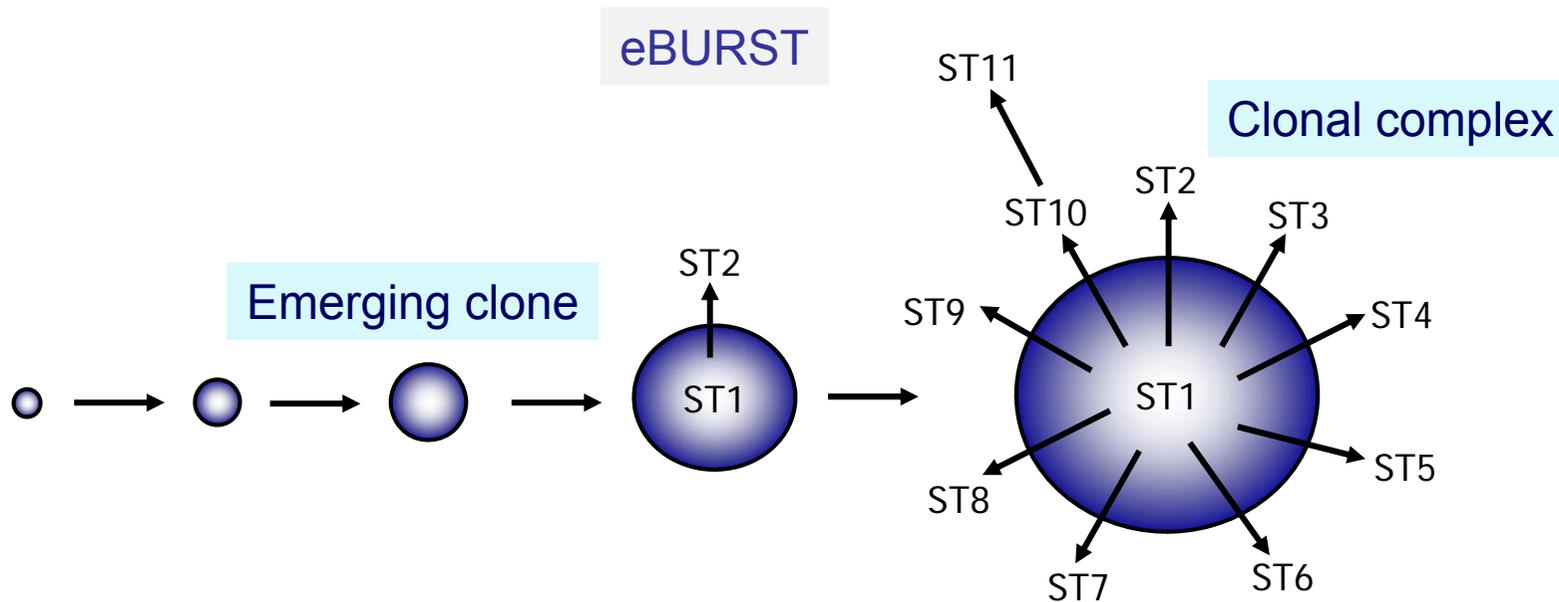
~12,500 isolates in the *S. pneumoniae* database

Dendrogram using only the first 1000 isolates in this database

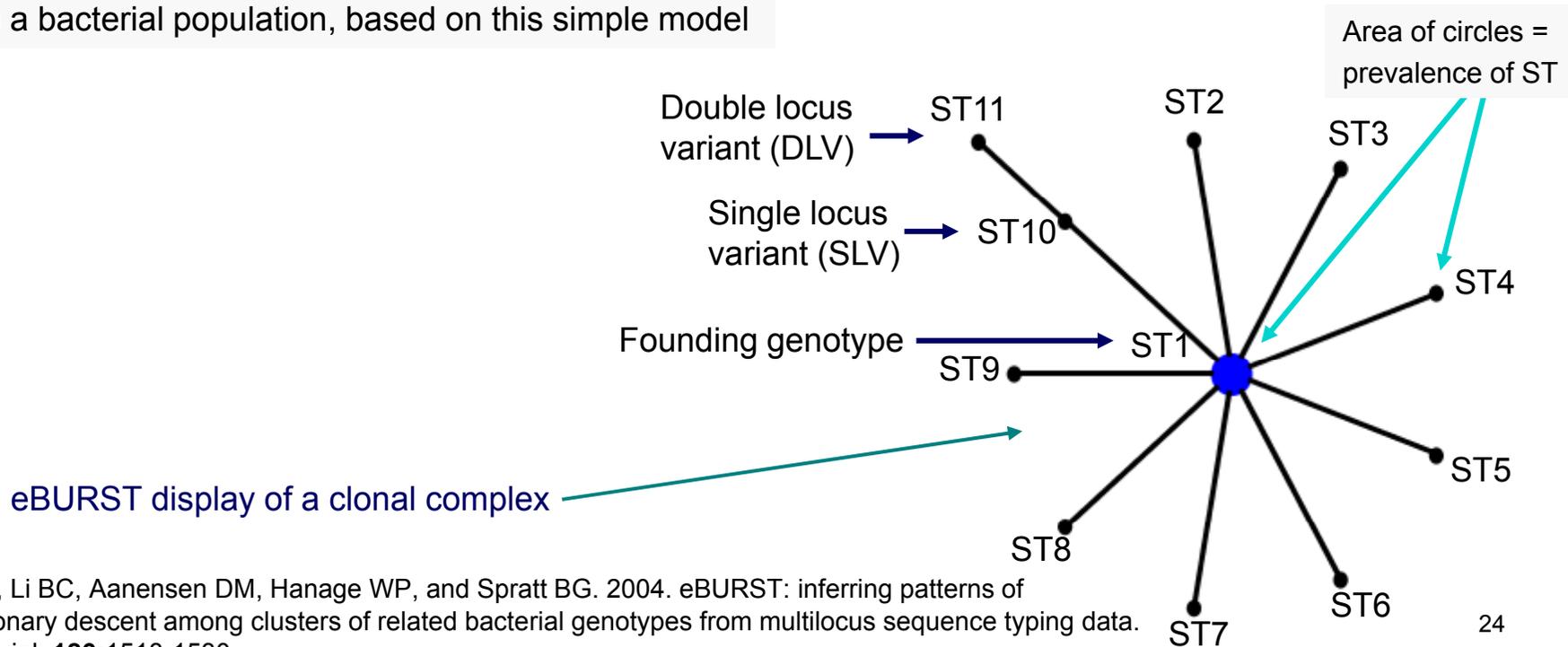
Help!

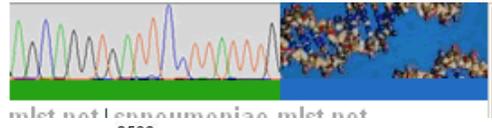
Need tools to make sense of MLST data

The screenshot displays a web browser window with the MLST (Multi Locus Sequence Typing) website. On the left, a 'Neighbour Joining Tree' is visible, showing a complex dendrogram of genetic relationships between isolates. The tree is dense and difficult to read. On the right, the MLST website interface is shown, featuring a search form for *Streptococcus pneumoniae*. The search form includes fields for 'id' and 'serotype', and a 'Submit' button. Below the search form, there is a table of results with columns for 'ST', 'Strain', 'Serotype', and 'Submit for allelic profile query'. The table lists various strains such as NCTC 11906-19F, SAF-17244-19, PJ23/1, SP264-23F, 87-029044-14, SP3026-6B, UKS71-23F, ATCC6323-23, CS111-23F, GM70-19A, GA71-19F, GM169-19F, GM17-6B, KD18-23, KD6-19, PN13-14, PN16-42, and PN18-19.

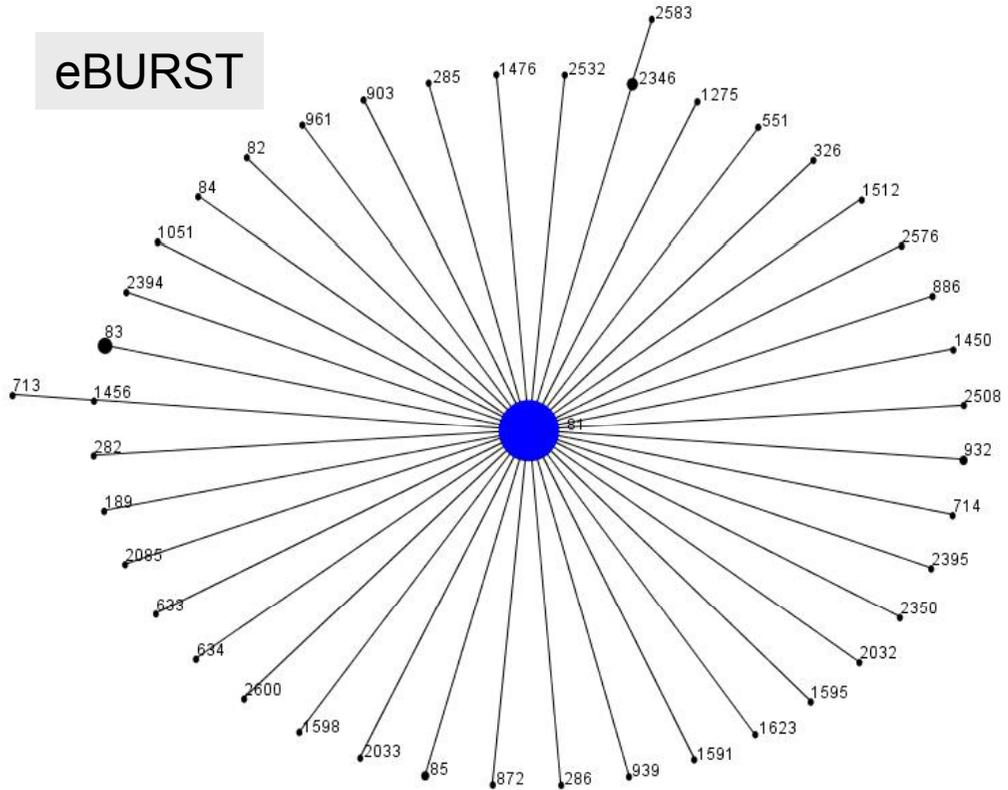


eBURST - identifies and displays the clonal complexes within a bacterial population, based on this simple model





eBURST



- *S. pneumoniae*
- *S. epidermidis*
- *S. pyogenes*
- *S. suis*
- *V. vulnificus*

SUBMISSIONS

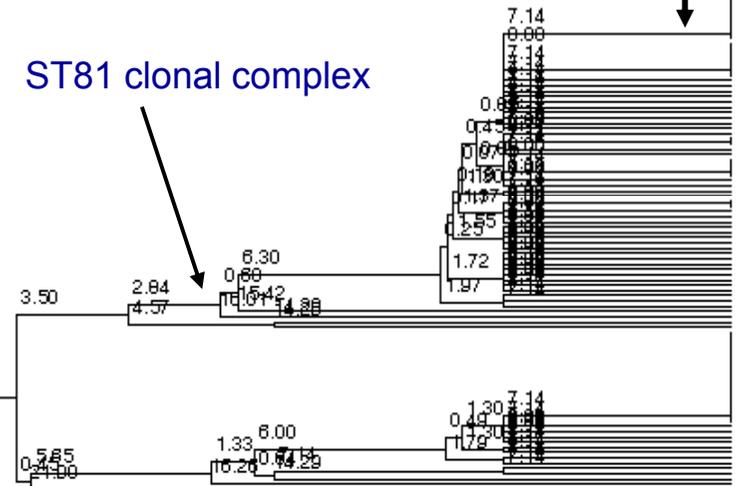
TW21-19F	81	19F	4	4
TW7-23F	81	23F	4	4
TW9-19F	81	19F	4	4
TW39-23F	83	23F	4	4
TW40-23F	81	23F	4	4
TW41-23F	81	23F	4	4
TW42-23F	83	23F	4	4
TW43-23F	83	23F	4	4

24.49

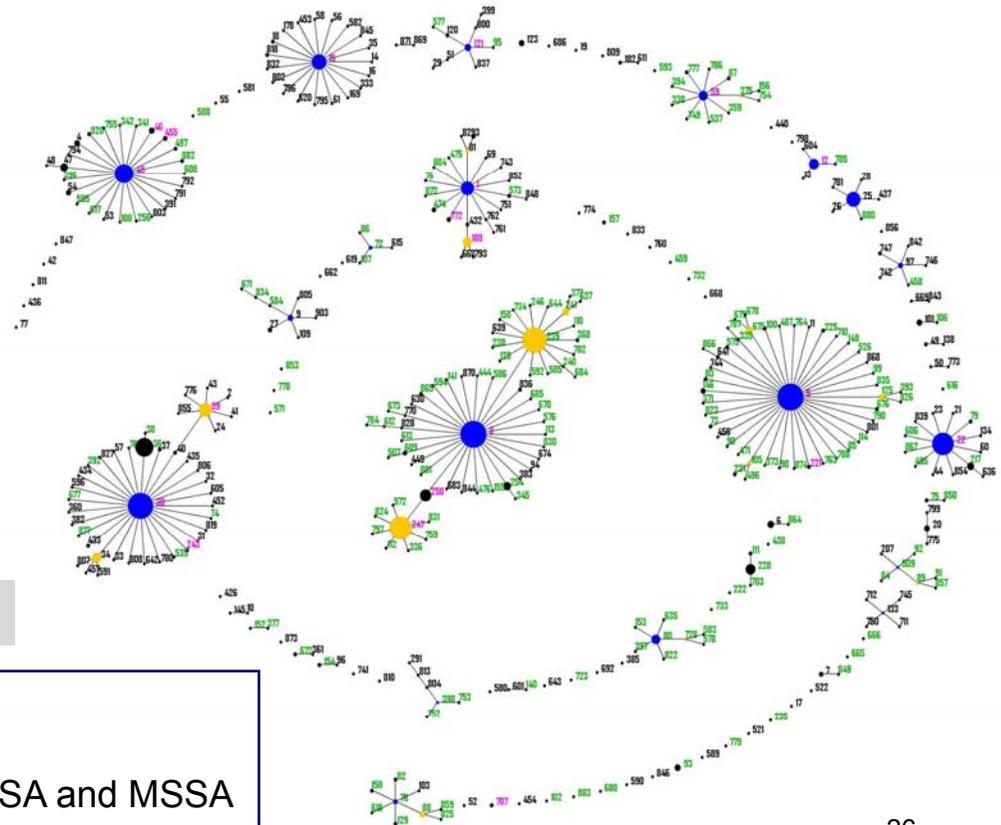
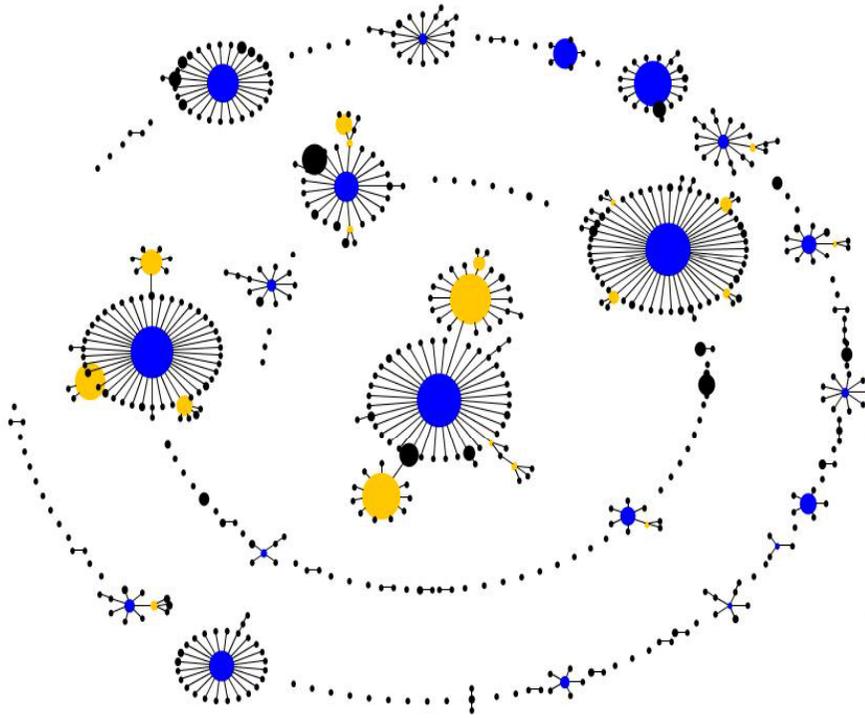
Streptococcus pneumoniae

ST81 = Multiply antibiotic-resistant strain (Spain^{23F}-1)

ST81



Staphylococcus aureus 'population snapshot'



Comparative eBURST

Green ST numbers = MRSA
Black ST numbers = MSSA
Pink ST numbers = Both MRSA and MSSA

Beyond MLST

MLST has been extremely successful (but not for all bacterial pathogens)

However, many questions cannot be answered using MLST, or any other established typing method, as not enough resolution/discrimination

Advances in genome sequencing allow ultimate resolution in molecular typing – comparing the genome sequence of each isolate of a pathogen strain or species

Latest sequencing machines - vast amounts of sequence data (~600,000,000,000 bp per run) – nearly 1 terrabase

Currently, £130/bacterial genome (Sanger Institute FEC costing) – by the end of the year will be ~£40/bacterial genome

Turnaround time is quite long – four weeks minimum

Competitive marketplace - technology will still improve substantially

Prospect of routinely obtaining sequences of isolates of bacterial pathogens in near real time, at a reasonable cost

What could clinical microbiologists/infection control teams do with this revolution?

Applications of 'genomic' molecular epidemiology

For example, the spread of MRSA strains (or strains of other nosocomial pathogens) within a hospital and between local hospitals and nursing homes

Problem: MRSA infection problems in a hospital mostly due to the same strain (e.g. EMRSA-15 in the UK) – almost all isolates in a hospital would be indistinguishable by MLST - and almost all would be the same spa-type

Cannot understand spread of EMRSA-15 in a hospital if all isolates look the same

A method that can discriminate many microvariants of an individual MRSA strain within a hospital would allow such studies to be conducted

Genome sequencing of all EMRSA-15 isolates in the hospital (and in neighbouring hospitals and nursing homes) allows tracking the spread of individual microvariants

Genome sequences for tracking spread of bacterial pathogens

MRSA – single strain – all ST239 by MLST
 – most common MRSA strain in SE Asia

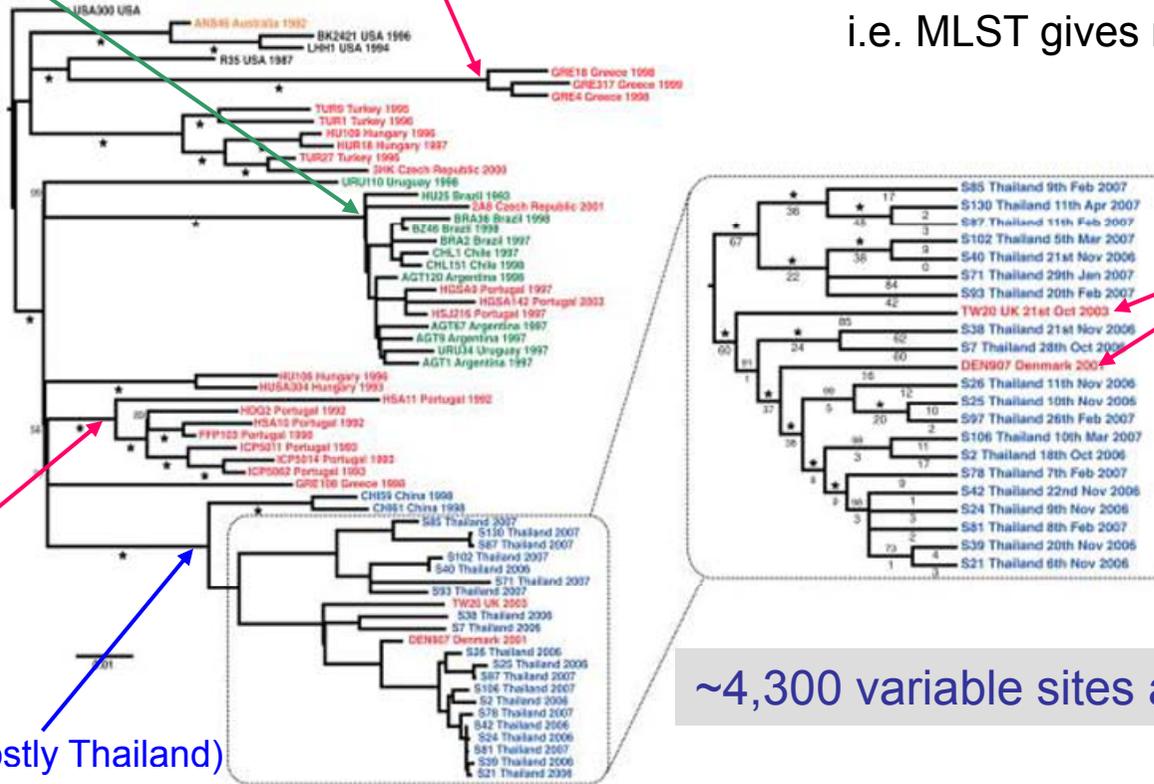
i.e. MLST gives no discrimination

South America (plus Portugal)

Greece

Portugal

Asia (mostly Thailand)



European
 (One was from Thai patient admitted to hospital in Denmark)

Thailand cluster

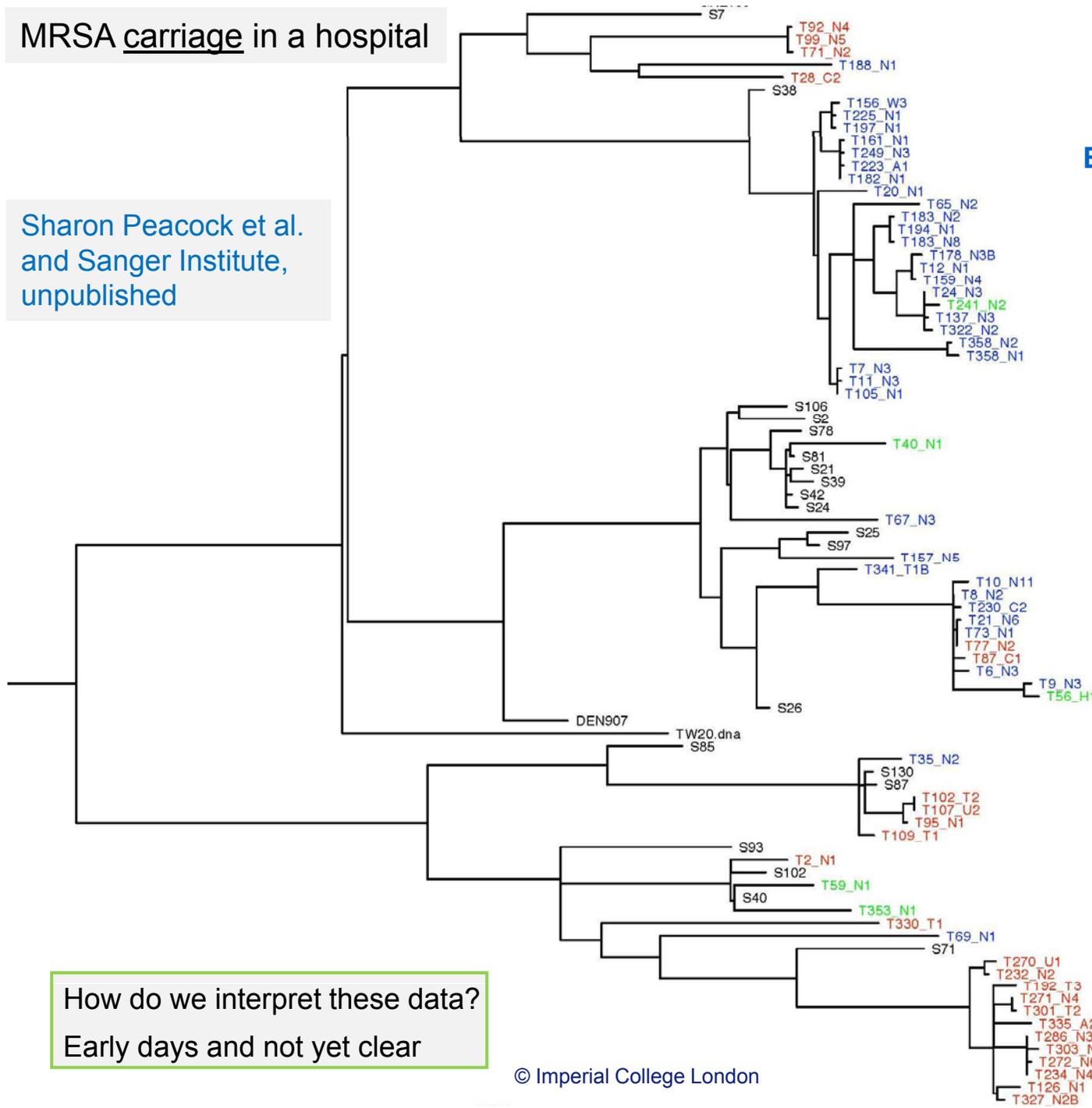
~4,300 variable sites among ST239 isolates

Every isolate of ST239 is different

Harris et al. (2010) Science 327: 469-474

MRSA carriage in a hospital

Sharon Peacock et al.
and Sanger Institute,
unpublished



Blue = Paediatric ICU

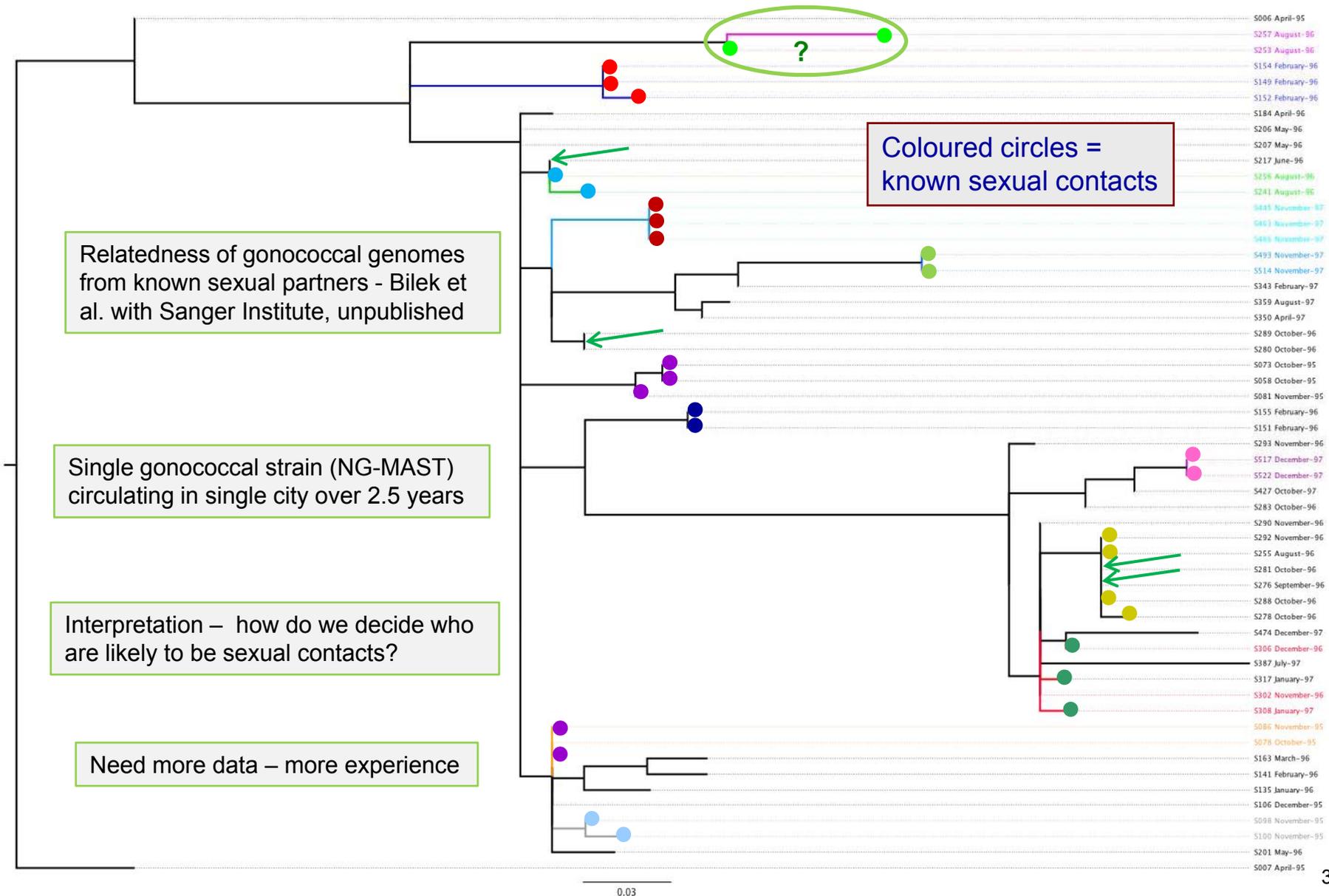
Green = Staff

Red = Adult surgical ICU

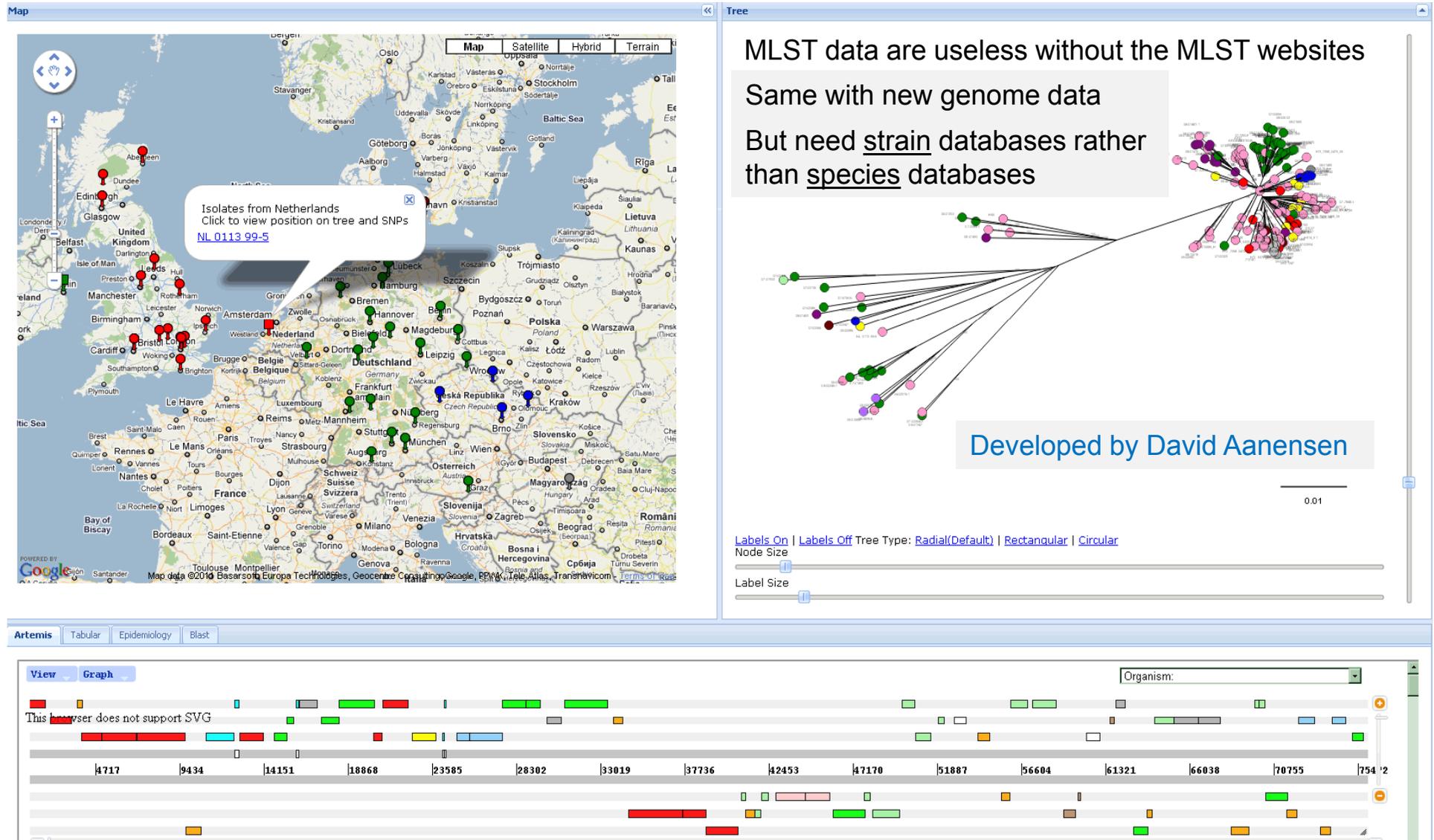
How do we interpret these data?
Early days and not yet clear

Identifying sexual partners and individuals in sexual networks using molecular biology

Neisseria gonorrhoeae



Web application for storing and analysing genomic molecular epidemiology data



Designed to store and analyse the genomes of isolates of the same bacterial strain (e.g. MRSA ST239)

Fully interactive functions – generic – flexible – scalable - public - can keep adding additional functionalities

Acknowledgments



Imperial College – St. Mary's Campus



The Wellcome Trust



WT Sanger Institute