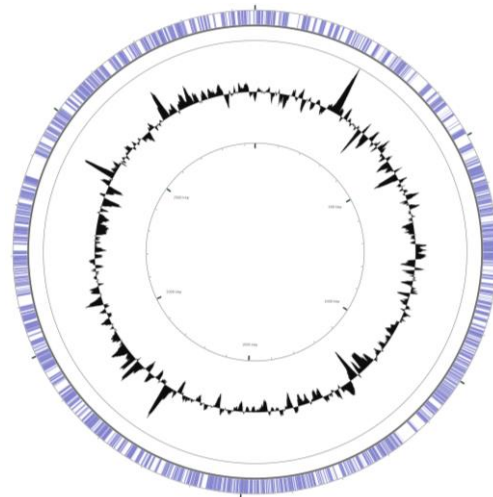# Genomic epidemiology of bacterial pathogens
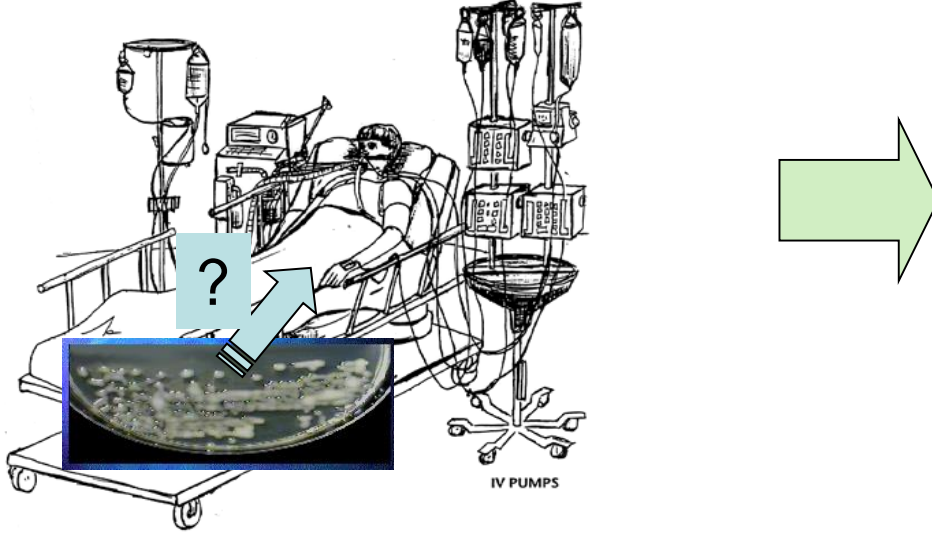
Sylvain BRISSE

Microbial Evolutionary Genomics, Institut Pasteur, Paris

Analysis of strain diversity within species
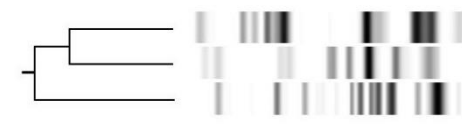
Typing
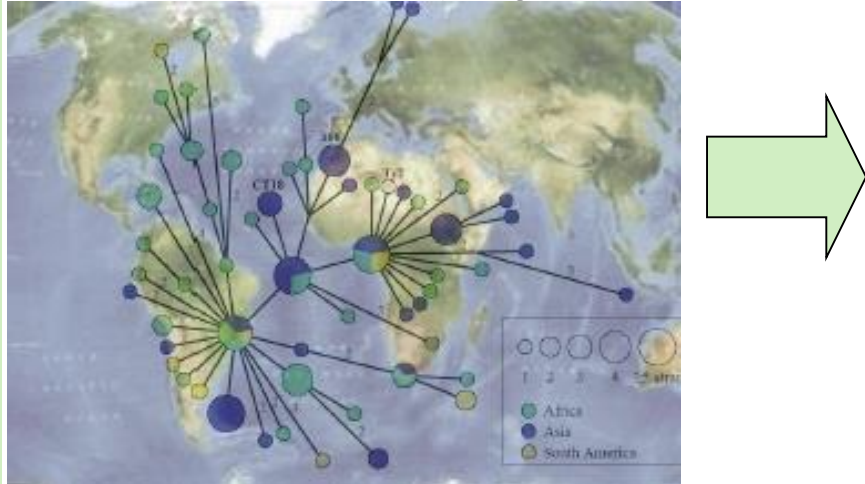
Aim: Local epidemiology

Fingerprinting methods:
- PFGE
- RAPD
- rep-PCR

Population genetics
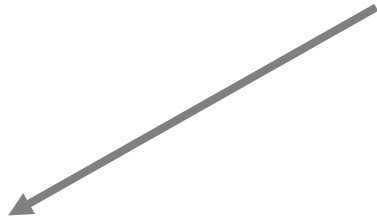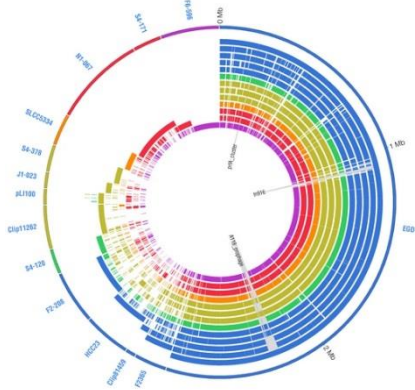
Aim: Diversity, phylogeny, evolution

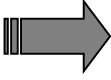Multilocus methods:
- MLST
- MLVA-VNTR
- SNPs

Genome sequencing: 'all in one technology'
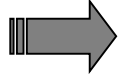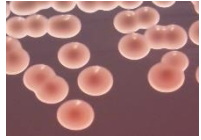
Genotype

Resistance elements
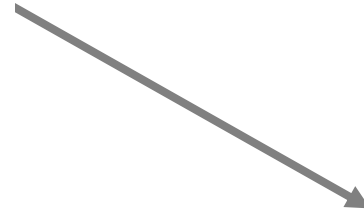
Virulence elements
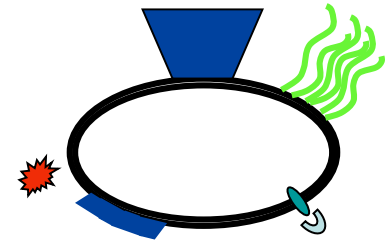
Mapping and SNP calling approaches

- SNP typing to phylogeny

  - Use of reference

  - Recombination removal

  - Quality filtering of SNPs

🙂 Maximal discrimination

☹ Limited forward comparability

☹ Centralized studies

Nucleotide polymorphisms

Centralized model of pathogen genotyping
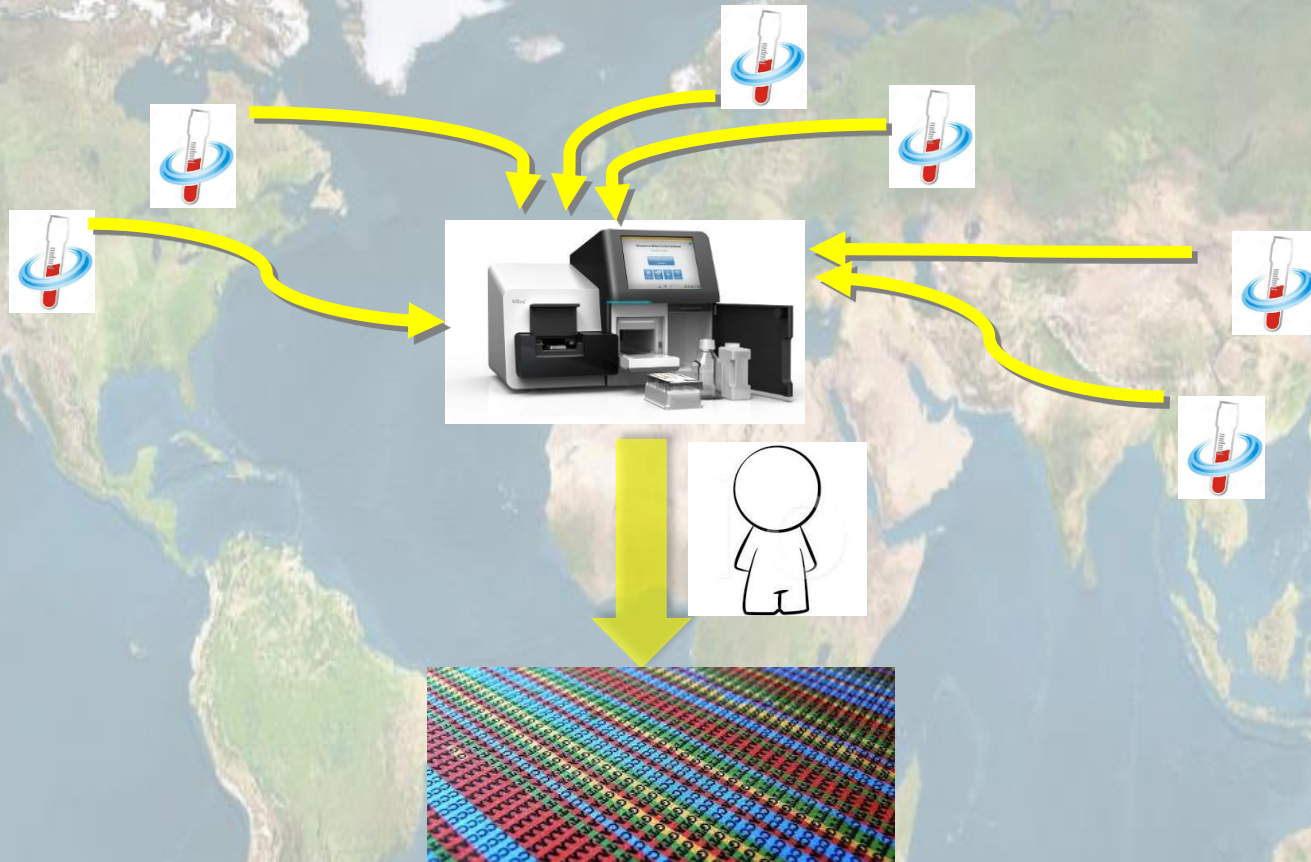
😠 Not efficient
😠 No additivity

Distributed model of pathogen genotyping
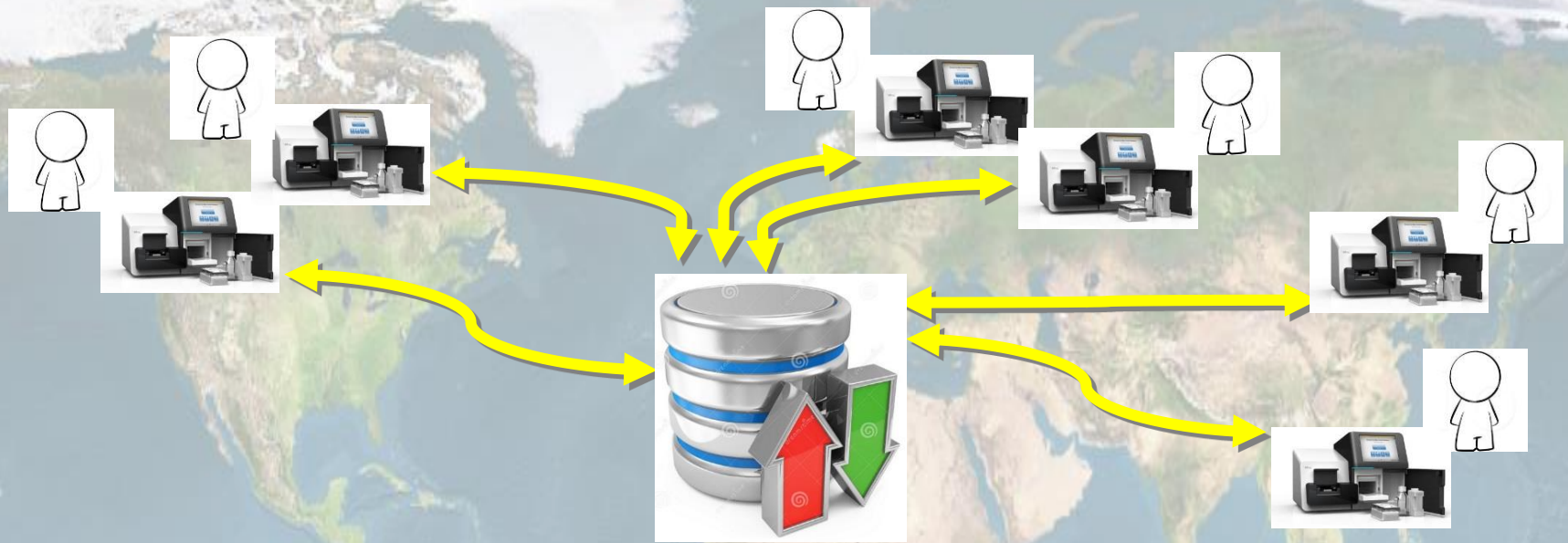
Nomenclature server

☺ Automated
☺ Efficient
☺ Additive
☺ Data confidentiality

MLST
Multilocus
Sequence
Typing

Genomic DNA

PCR amplification of seven housekeeping genes
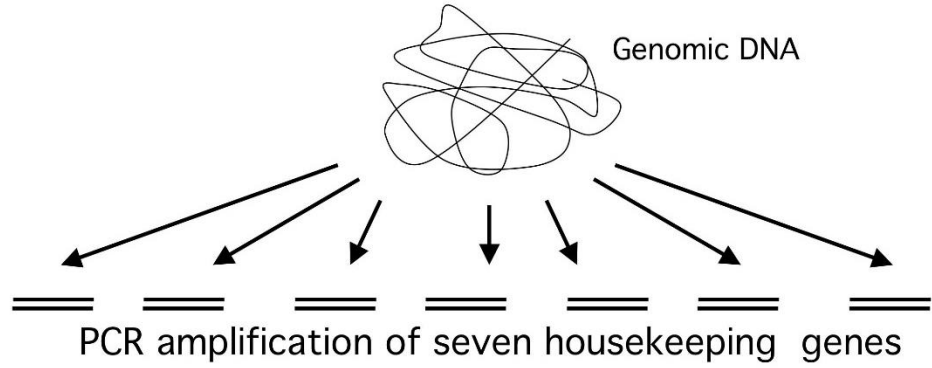
Sequencing on both strands, approx. 500 bp

• Sequence of one gene (*lhkA*, 480 bp)

•Allele

1
1
2
3
3
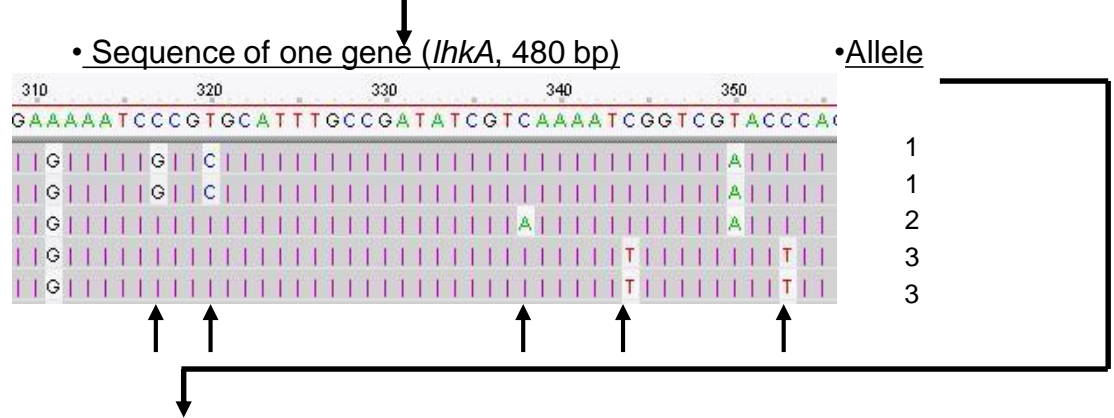
Alleles:

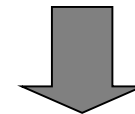| *abcZ*	*bglA*	*cat*	*dapE*	*dat*	*ldh*	*lhkA*
3	1	1	1	3	1	3

Sequence type (ST)	Alleles
1	3 1 1 1 3 1 3
2	1 2 1 7 22 39 7
3	3 1 1 1 3 1 2

Single
Locus
Variants
(SLV)

MLST databases: global epidemiology and population biology

pubmlst.org

PubMLST

www.mlst.net

MLST
Multi Locus Sequence Typing

ST	dinB	icdA	pabB	polB	putP	trpA	trpB	uidA
1	1	1	2	1	1	2	3	1
2	8	2	7	3	7	1	4	2
3	3	8	5	11	8	3	5	3
4	2	4	6	4	1	6	1	1
5	5	3	3	10	5	8	2	5
6	1	7	1	9	2	20	1	6
7	6	6	4	2	6	7	2	4
8	23	9	8	12	9	11	7	13
9	9	20	15	7	4	9	6	9

Warwick databases

- *Escherichia coli*
- *Moraxella catarrhalis*
- *Salmonella enterica*
- *Yersinia pseudotuberculosis*

www.pasteur.fr/mlst

Institut Pasteur MLST Databases
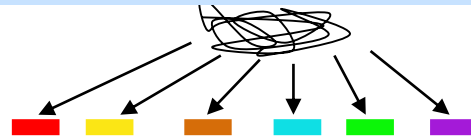
Acinetobacter baumannii | *Bifidobacterium* | *Clostridium difficile* | *Escherichia coli* | *Flavobacterium psychrophilum*

Klebsiella pneumoniae | *Lactobacillus casei* | *Listeria monocytogenes* | *Pantoea agglomerans* | *Plesiomonas shigelloides*

- Commensal
- EPEC
- ETEC
- EHEC
- APEC
- EIEC
- *S. flexneri*
- *S. dysenteriae*
- *S. sonnei*
- *S. boydii*
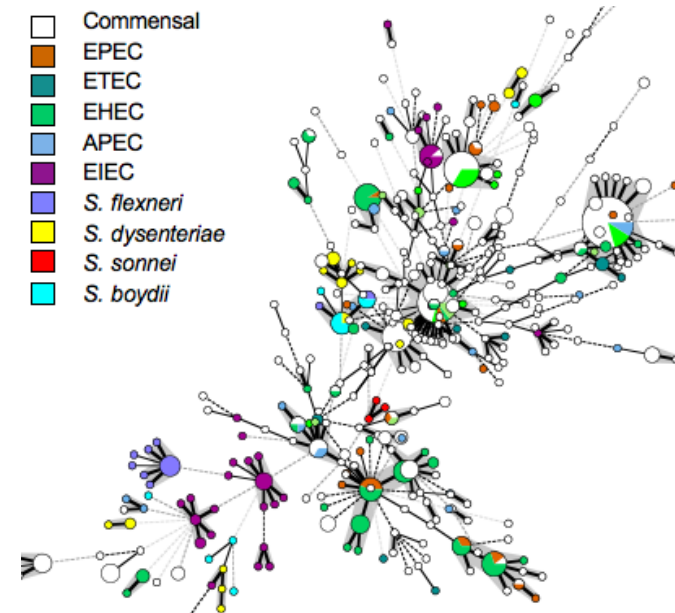
Bacterial nomenclature: 'Biological esperanto'

- Taxonomic ranks

Domain [*Bacteria - Archaea - Eukarya*]
Phylum (Division)
Class
Order
Family
Genus
species
subspecies
sequence type

Molecular dissection of the evolution of carbapenem-resistant multilocus sequence type 258 *Klebsiella pneumoniae*
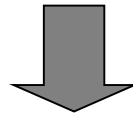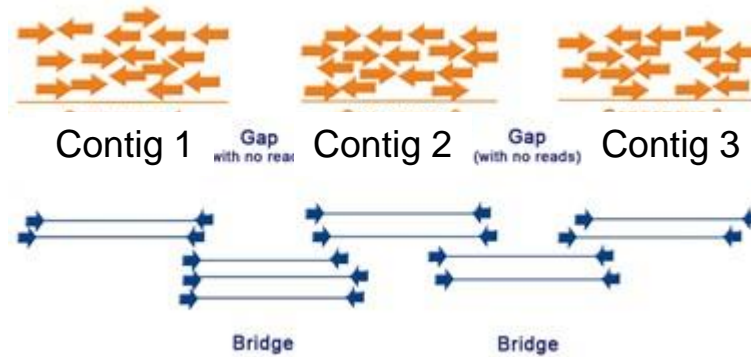
Research article

Open Access

Successful control of a neonatal outbreak caused mainly by ST20 multidrug-resistant SHV-5-producing *Klebsiella pneumoniae*, Greece

Contig 1   Gap (with no reads)   Contig 2   Gap (with no reads)   Contig 3

Bridge   Bridge

Gene 1   …. Gene 2000

	Gene 1									…. Gene 2000
Strain 1	1	15	12	37	3	16	22	11		15
Strain 2	1	5	4	24	4	4	1	1		5
Strain 3	27	28	9	31	1	25	10	15		28

Genomic diversity inside bacterial species
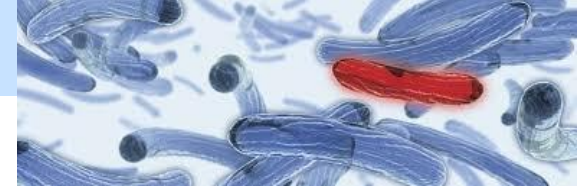
Genomic islands
Phages
Transposons, IS
ICE
Plasmids

...

Welch et al., 2002

O157:H7
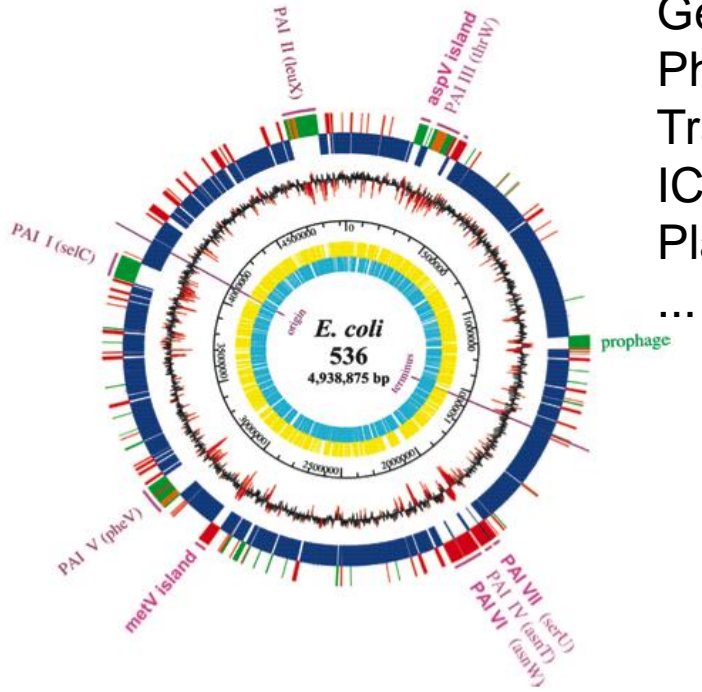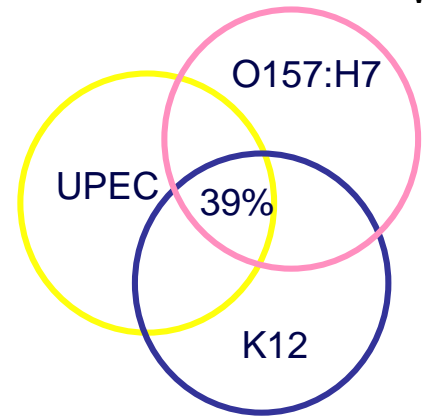
UPEC   39%

K12

Insertion hotspots

Touchon et al., 2009

Escherichia coli

All genes
(pangenome)
~20,000

families
~11,500

IS excluded
~10,000
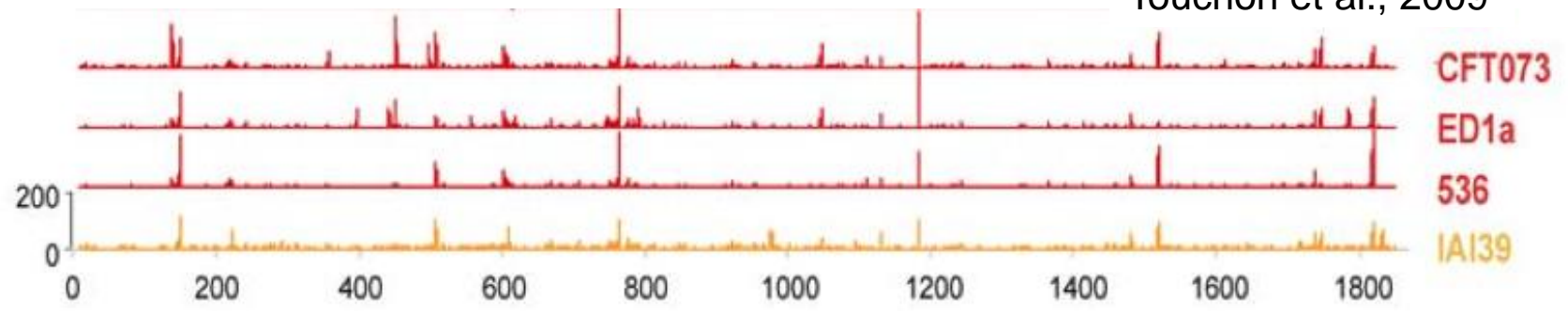
Core genome
~2,000

Touchon et al., 2009

Core genome MLST

Step 1. Define core genes

Step 2. Record variation at core genes; define alleles

Step 3. Record allelic profiles; provide nomenclature

Klebsiella pneumoniae
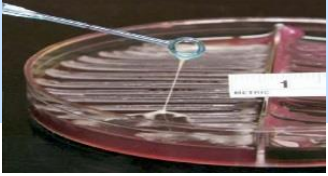
Ubiquitous

* Animal carriage, plants, water, environment

Nosocomial pathogen

• 5 - 8% nosocomial infections Europe / USA

• Urinary, respiratory, bacteremia; outbreaks

• Multidrug resistant (MDR) strains
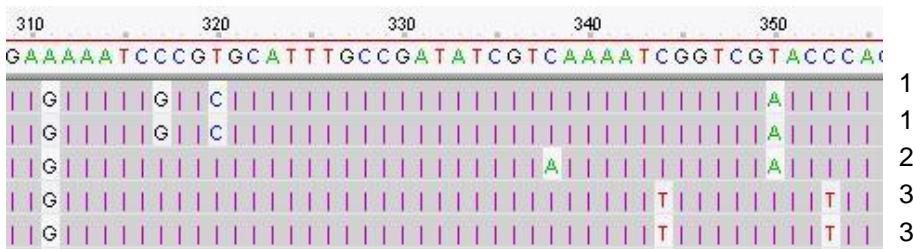

Community pathogen

• Severe pneumonia; pyogenic liver abscess; meningitis

• Hypervirulent *K. pneumoniae* (HVKP)

• Mostly capsular types K1 and K2
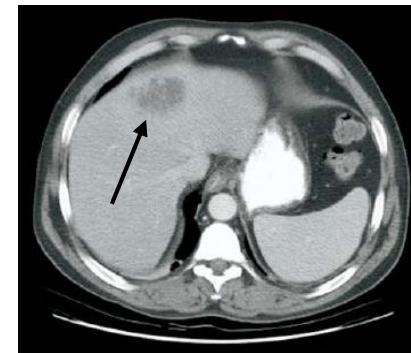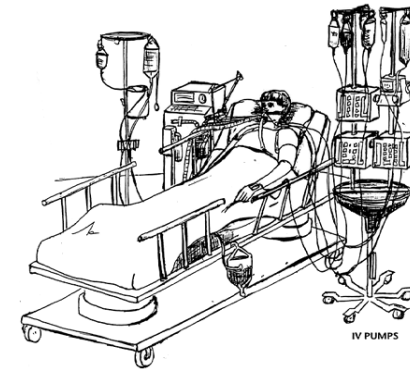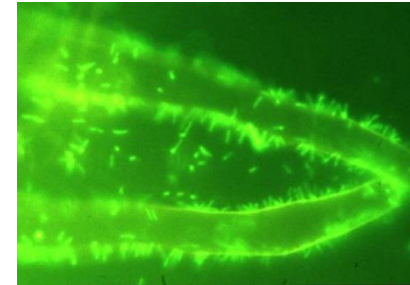
Klebsiella pneumoniae core genome MLST

Hospital strains   Community-onset   Environment   Historical collection
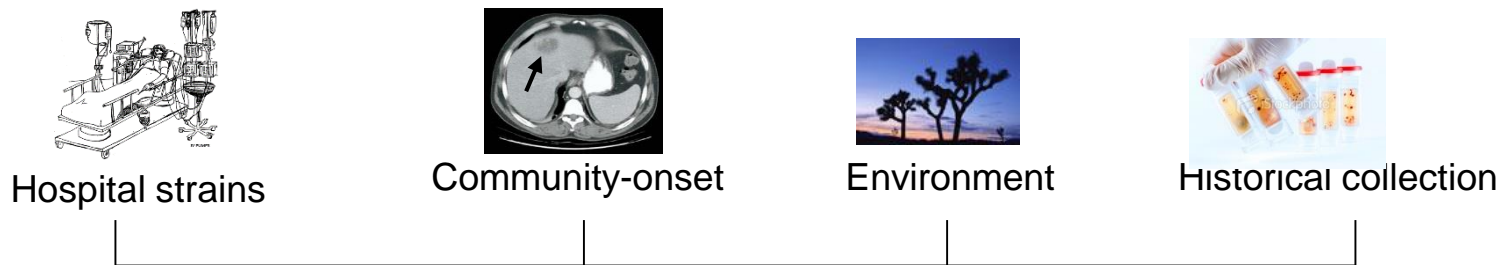
694 core genes (cgMLST)

Core genome MLST
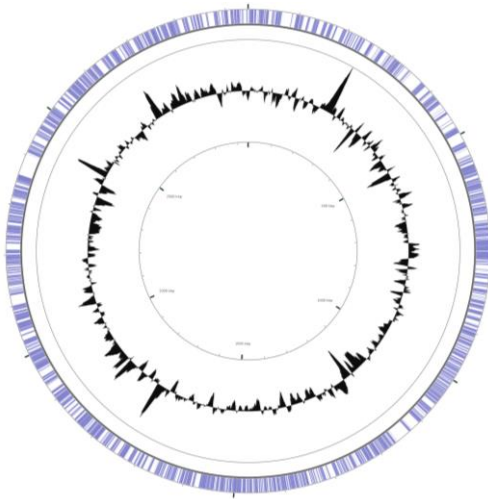
Genomic Definition of Hypervirulent and Multidrug-Resistant *Klebsiella pneumoniae* Clonal Groups

Suzanne Bialek-Davenet,[1] Alexis Criscuolo,[1] Florent Ailloud, Virginie Passet, Louis Jones, Anne-Sophie Delannoy-Vieillard, Benoit Garin, Simon Le Hello, Guillaume Arlet, Marie-Hélène Nicolas-Chanoine, Dominique Decré, and Sylvain Brisse

Bialek, Criscuolo et al., EID, 2014

- Gene selection

Genes in highly conserved, syntenic regions	Ribosomal genes	MLST genes
⬇	⬇	⬇
634 scgMLST	53 rMLST (Jolley et al.)	7 MLST (Diancourt et al.)
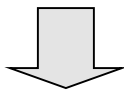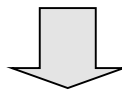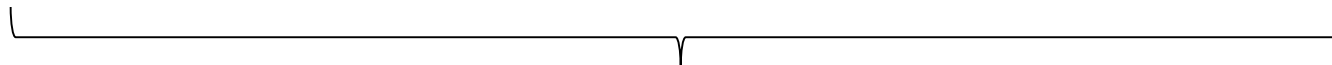
694 core genes (cgMLST)

K. pneumoniae: phylogenetic structure
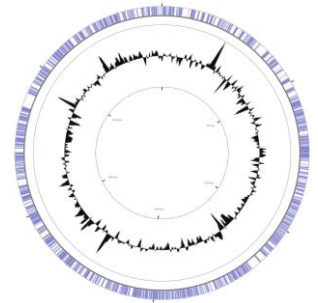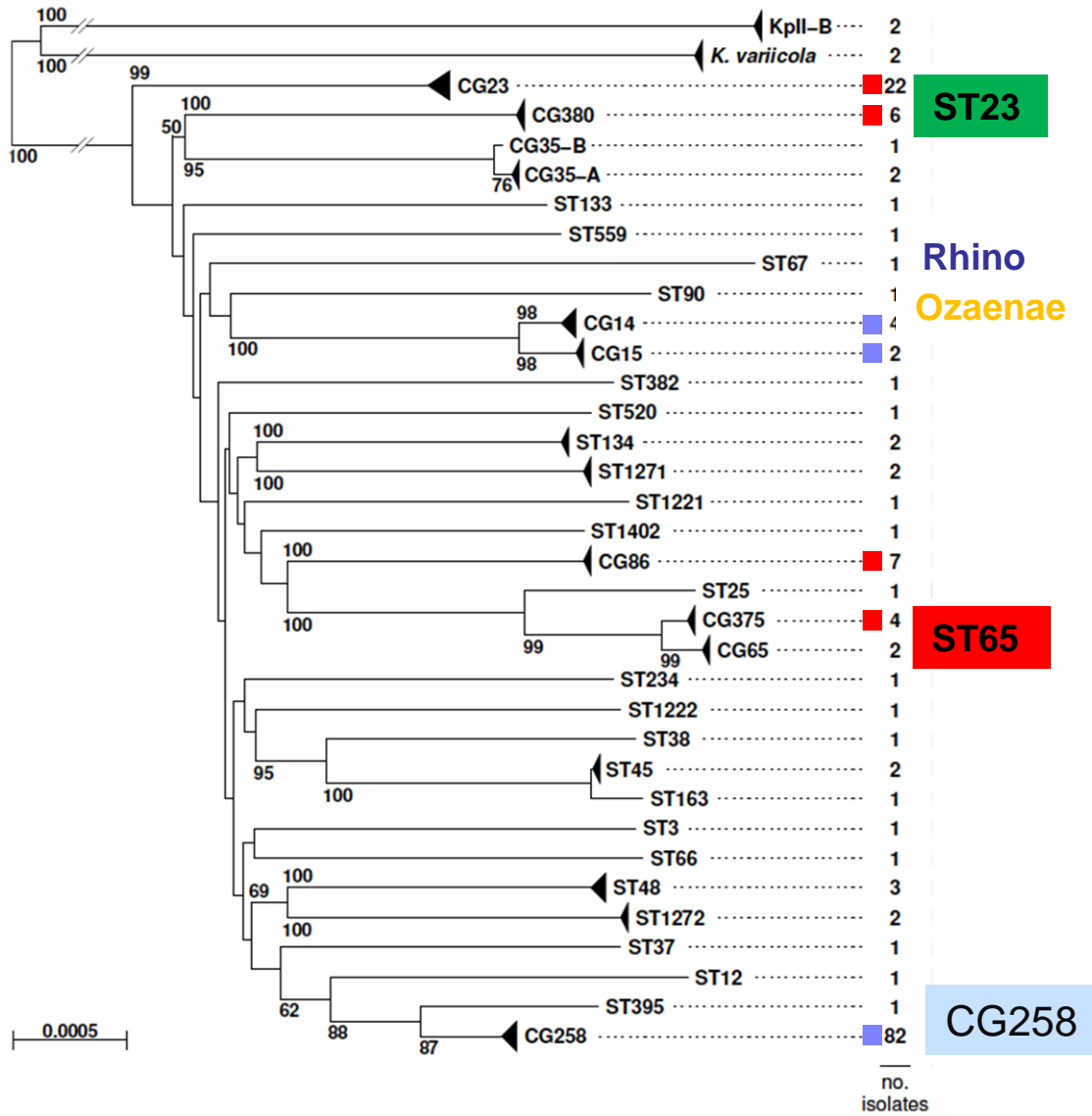
694 core genes
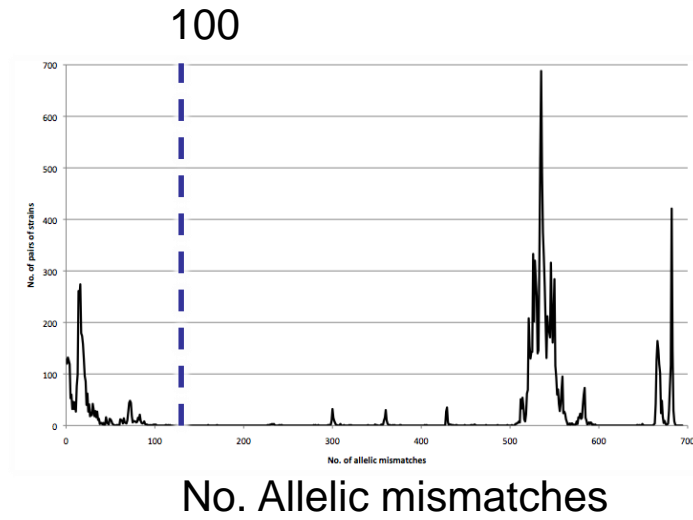
No. Allelic mismatches

Recognition of several hypervirulent *K. pneumoniae* clonal groups

(core-genome MLST, 694 genes)

100

ST11, ST340

HS11286

CG258
(ST258 + ST512
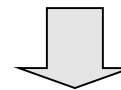+ ST11 + ST340)

CG258

ST395

ST37

ST133

CG35-A
CG35-B

CG380
(ST380 + ST679)

CG380

ST15

ST14

ST90 (*K. pneumoniae* subsp. *ozaenae*)

KpII-B

KpII-A

KpV

KpIII (*K. variicola*)

Kp342
At-22

KpVI

LM21 → ST12

ST559

ST67 (*K. pneumoniae* subsp. *rhinoscleromatis*)

ST1221
ST38
ST163
ST45

MGH78578

ST3
ST382

DSM30104ᵀ

ST48
ST1272
ST234
ST1271
ST520
ST134
ST1402

CG86

CG86

hvKP1,
CG43

CG65

CG375

KCTC 2242

CG375

ST25
ST66

B5055

CG23
(ST23 + ST57)

NTUH-K2044

CG23

K. pneumoniae population structure: classical MLST

Ozaenae

ST23

ST65

ST258

Rhinoscleromatis

http://www.pasteur.fr/mlst
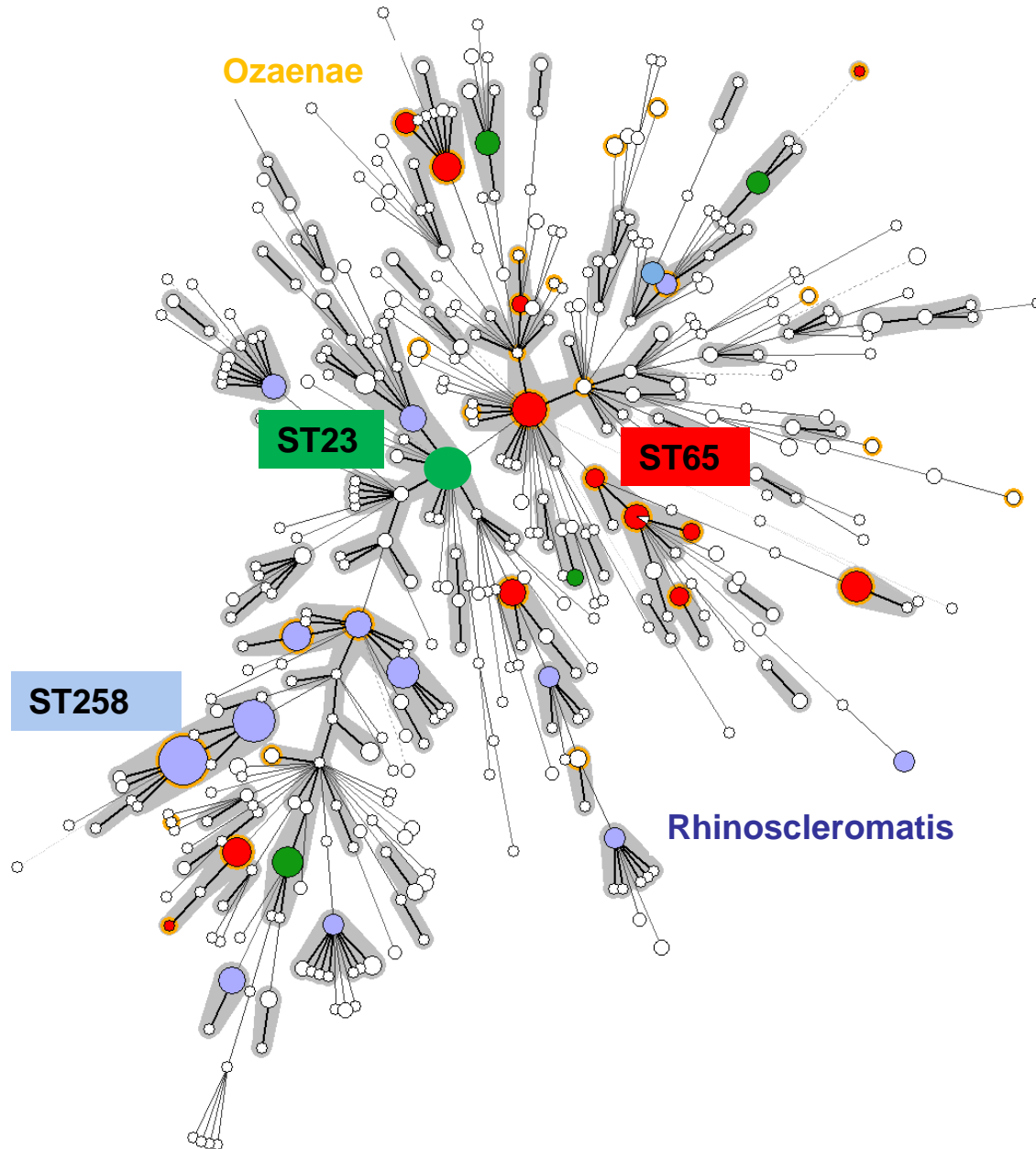
Klebsiella Sequence Typing
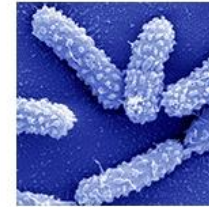
Sequences and profiles database

Isolates database

> 1500 STs
> 2000 isolates
> 250 user labs

- Lack of clear-cut clonal structure

- Merging of majority of sequence types

Global emergence of carbapenem resistance

KPC, NDM, OXA-48, VIM
Few remaining drugs available
Death rate > 50%
'urgent threat' (CDC 2013)
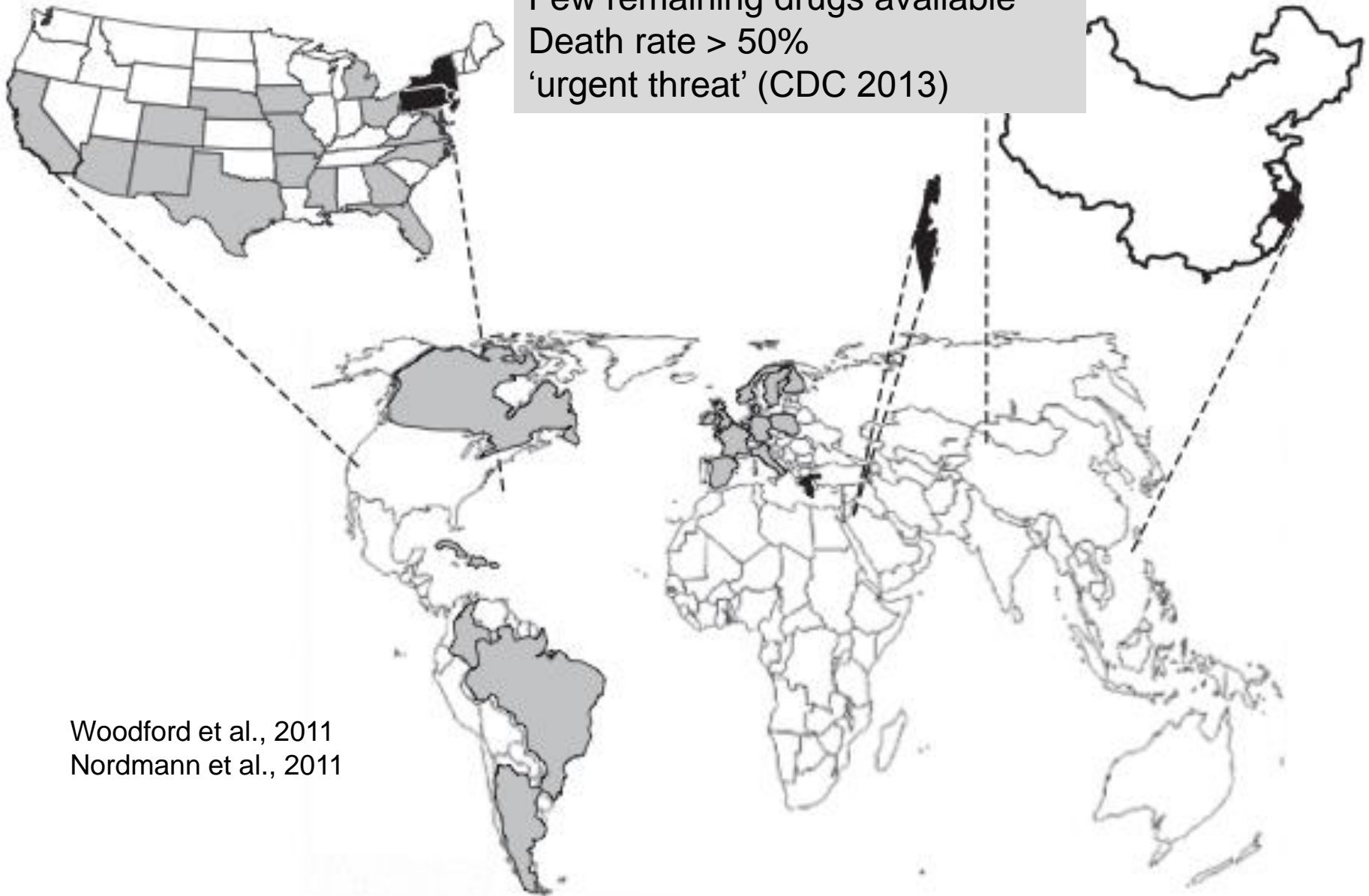
Woodford et al., 2011
Nordmann et al., 2011

Antimicrobial resistance genes

β-lactams

β-lactamases

- Class A (TEM, SHV, CTX-M, KPC)
- Class B (NDM, VIM)
- Class C (AmpC)
- Class D (OXA)

Aminoglycosides

- Inactivating enzymes (AAC, ANT, APH)
- 16S rRNA methylases (ArmA, Rmt)

Quinolones

- Target changes (GyrAB, ParCE)
- Inactivating enzymes(AAC(6')-Ib-cr)
- Target protection (Qnr)
- Efflux pumps (OqxAB)

known variants

BIGSdb-Kp

Klebsiella pneumoniae virulence or colonization factors

✧ Virulence factors:

HYPERMUCOVISCOSITY
(*rmpA, rmpA2*)

FIMBRIAE
type 3 pili (*mrkABCDF*)

METABOLISM
Allantoinase

TOXINS
Microcin E492
Colibactin

SIDEROPHORES /
IRON TRANSPORT
Aerobactin
Yersiniabactin
kvgAS
kfuABC

CAPSULE
77 serotypes : *wzi, wzc*
(serum & phagocytosis resistance)
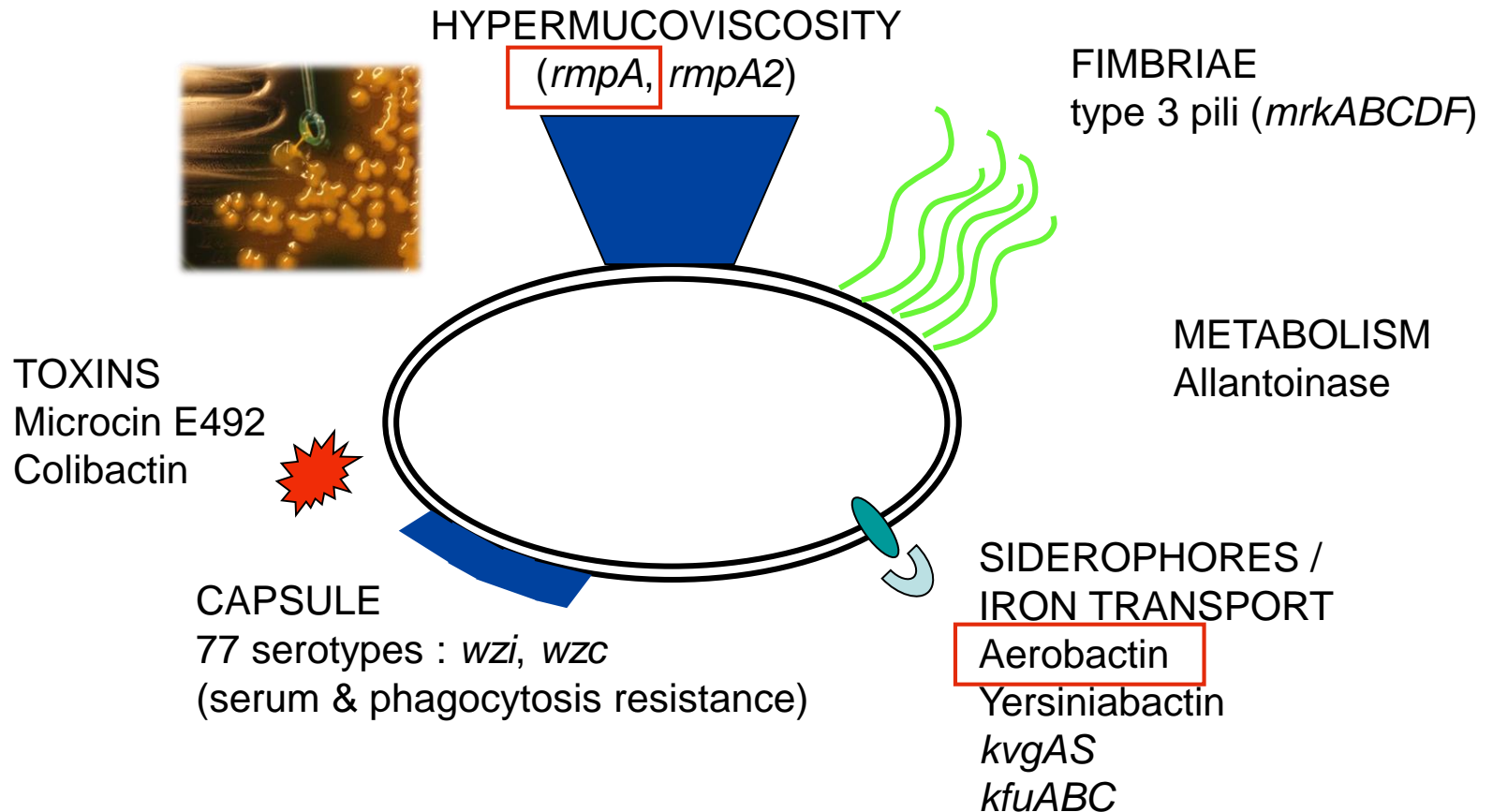
✧ Heavy metal resistance-associated clusters:

pbrABCR, pcoABCDERS, silCERS, and *terWXY*

Gene content breakdown by clonal group of *K. pneumoniae*
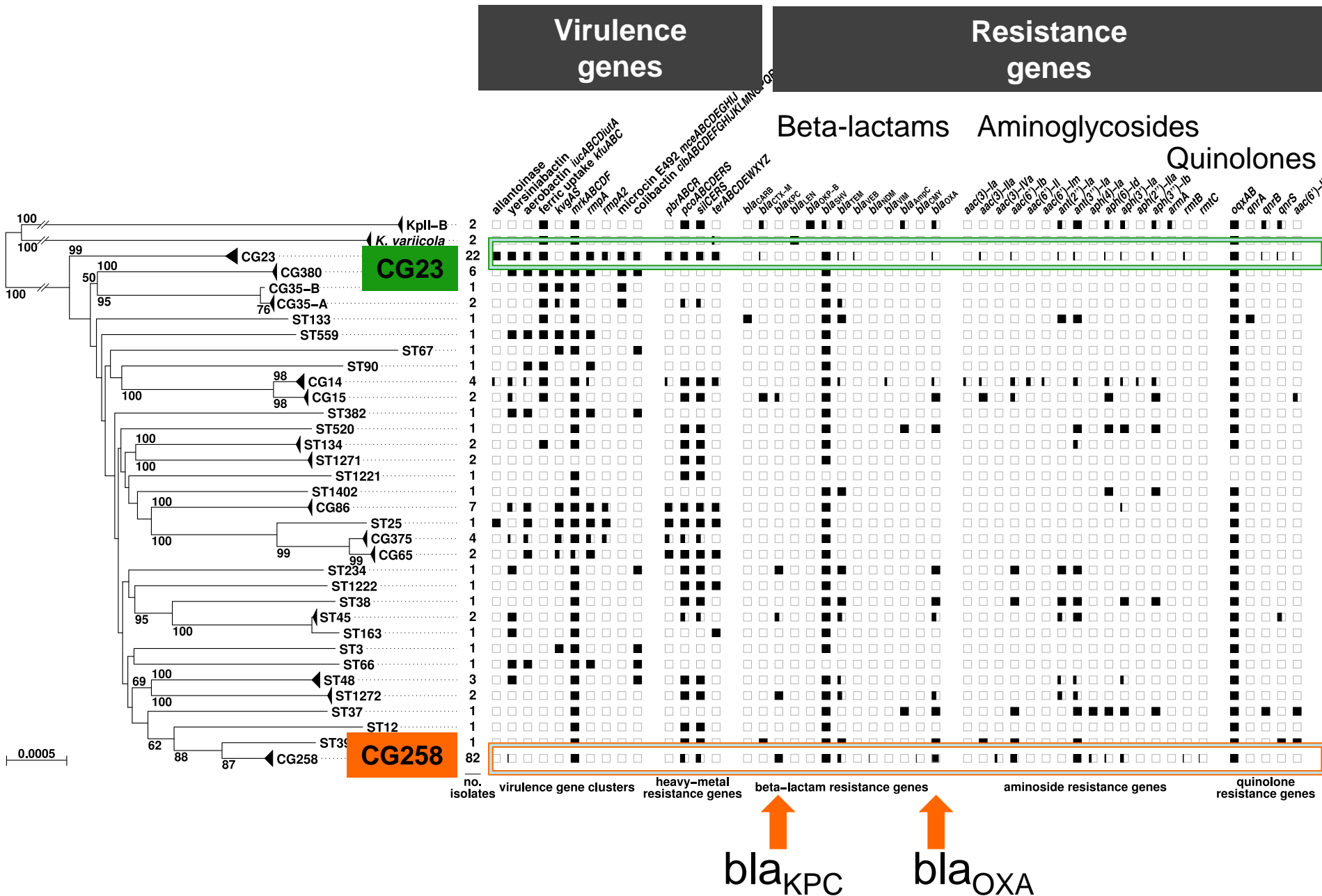
http://bigsdb.web.pasteur.fr

Klebsiella Sequence Typing
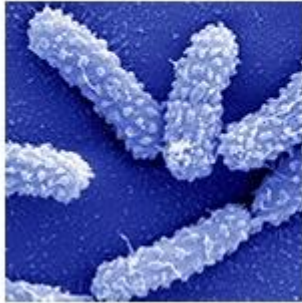
Sequences and profiles database

Isolates database

MLST
cgMLST
Virulence genes
Resistance genes
wzi (capsular typing)

Powered by BIGSdb
(Jolley & Maiden, 2010)

Sequence query - Klebsiella locus/sequence definitions

Please paste in your sequence to query against the database. Query sequences will be checked first for an exact
match is not found. You can query using either DNA or peptide sequences.

Please select locus/scheme
All loci

Order results by
locus

All loci
Allantoinase cluster
Aminoglycoside resistance genes
Beta–lactamase genes
cgMLST694
cps cluster genes
Efflux systems and regulators
Heavy metal resistance genes
hvK2 multiplex PCR
MLST
Quinolone resistance genes
rMLST
scgMLST634
Virulence genes
30S ribosomal protein S10 (rpsJ)
30S ribosomal protein S11 (rpsK)
30S ribosomal protein S12 (rpsL)
30S ribosomal protein S13 (rpsM)
30S ribosomal protein S14 (rpsN)
30S ribosomal protein S15 (rpsO)