Klebsiella virulence typing – part I

Assistant Professor Zoe A. Dyson

London School of Hygiene and Tropical Medicine

zoe.dyson@lshtm.ac.uk

@msmicrobiocode

Today's schedule

Time	Activity
12:10-12:50 (40 mins)	**Lecture: *Klebsiella* virulence typing - part I** • An introduction to *Klebsiella* virulence determinants • The capsule and K-/KL-types • Lipopolysaccharides (LPS) and O-types • An introduction to Kaptive
12:50-13:00 (10 mins)	**Class discussion**
13:00-14:00 (1 hour)	**Lunch**
14:00-15:15 (1 hour 15 mins)	**Kaptive hands on practical**
15:15-15:30 (15 mins)	**Break**
15:30-16:00 (30 mins)	**Kaptive hands on practical (continued)**
16:00-16:30 (30 mins)	**Data sharing workflow mapping (Nicole Dagata)**

Lecture outline: *Klebsiella* virulence typing – part I

1. An introduction to *Klebsiella* virulence determinants
2. The capsule and K-antigen types
3. Lipopolysaccharides (LPS) and O-antigen types
4. An introduction to Kaptive

An introduction to *Klebsiella* virulence determinants

Different infection types driven by *Klebsiella*

Members of the KpSC can cause a variety of different infection types:

- Those causing **healthcare associated infections (HAI)** in immunocompromised individuals are often referred to as '**classical**' strains
- Those causing **community acquired infections (CAI)** in healthy individuals are often referred to as **hypervirulent** strains
- **Pathogenicity/virulence** factors commonly associated with more severe cases can be detected from whole genome sequencing (WGS) data

	Characteristic(s) for strain type	
Parameter	Classical	Hypervirulent
Common types of infection	Pneumonia, UTI, bacteremia	Pyogenic liver abscess; bacteremia; lung, neck, and kidney abscesses; pneumonia; cellulitis; necrotizing fasciitis; myositis, meningitis; endophthalmitis
Susceptible population(s)	Immunosuppressed (diabetics, patients with malignancies)	Diabetics, healthy people
Capsule type(s)	Capsule serotypes K1–K78	Hypercapsule serotype K1 (93%) or K2
Siderophores (% of strains expressing siderophore)	Enterobactin (100), yersiniabactin (17–46), salmochelin (2–4), aerobactin (6)	Enterobactin (100), yersiniabactin (90), salmochelin (>90), aerobactin (93–100)
Geographical concentration	Worldwide	Primarily Taiwan and Southeast Asia
Primarily acquired infection type	Nosocomial	Community acquired
Frequency of reports of antibiotic resistance	Frequent (ESBL and carbapenemase producing)	Infrequent

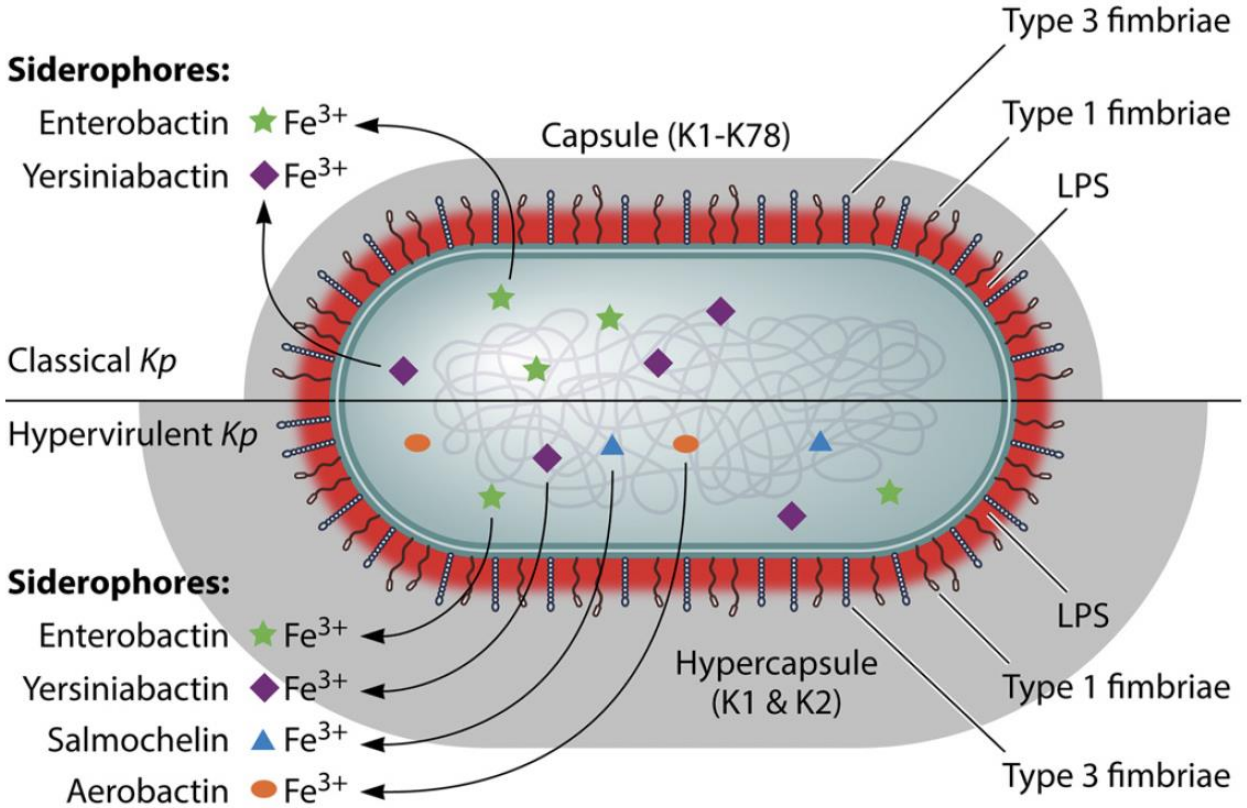
Paczosa and Mecsas 2016, MMBR

Pathogenicity/virulence factors in *Klebsiella*

- Many factors contribute to the ability of *K. pneumoniae* strains to cause disease and evade host defences

- The most well studied of these are:
  - Siderophores
  - Fimbriae/pilli
  - Capsule
  - Lipopolysaccharide
  - Toxins e.g. colibactin

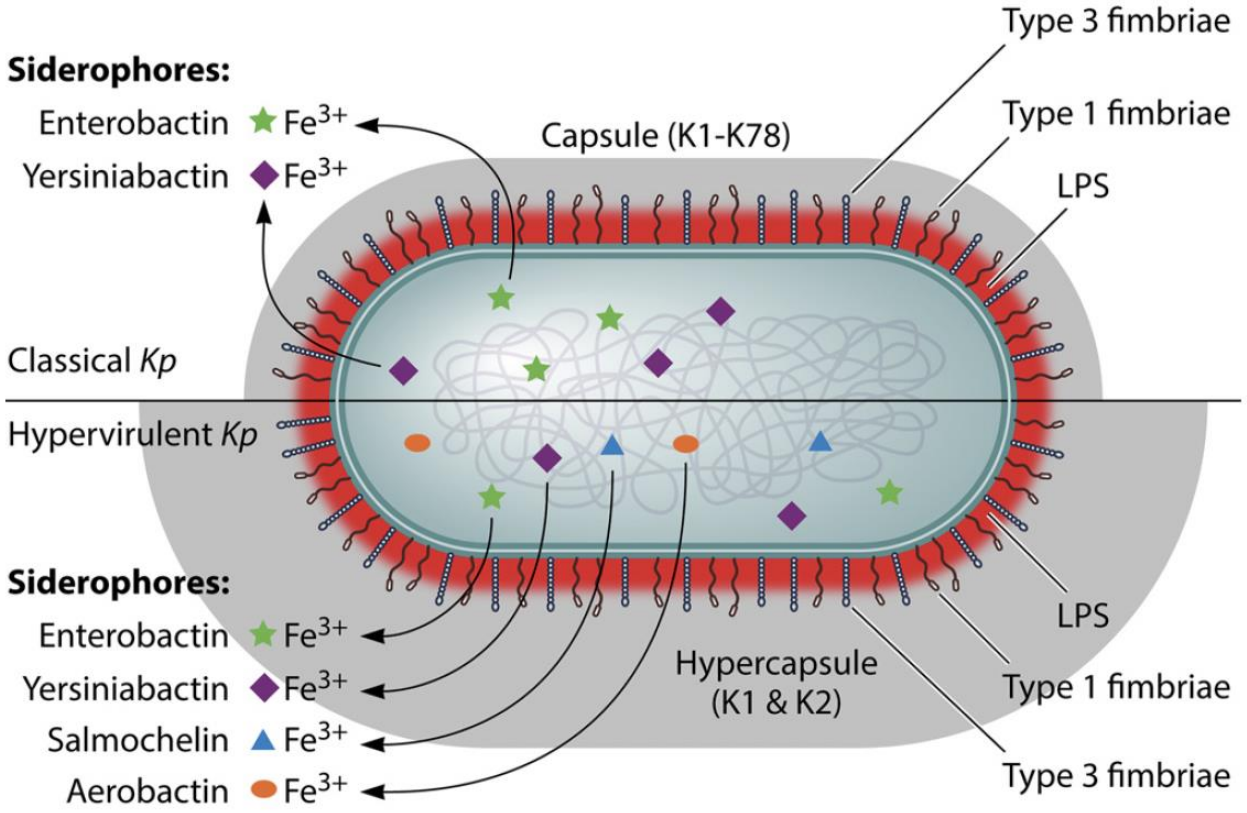
- Other less well studied factors include:
  - Outer membrane proteins (OMPs)
  - Porins
  - Efflux-pumps
  - Iron-transport systems
  - Allantoin metabolism systems
  - Many others…



Paczosa and Mecsas 2016, MMBR

Pathogenicity/virulence factors in *Klebsiella*

- All *K. pneumoniae* encode a subset of four core chromosomally integrated pathogenicity/virulence factors for establishing infections in mammals:
  - *Ent* locus encoding the siderophore enterobactin
  - Types 1and 3 Fimbriae/pilli (*fim* and *mrk* loci)
  - Lipopolysaccharide (O-antigen)
  - Capsular polysaccharide (K-antigen)
- Hypervirulent strains may have:
  - Specific capsule types
  - Other siderophores (e.g. yersiniabactin, aerobactin, salmochelin)
  - The genotoxin colibactin
- This lecture focuses on:
  - K-antigens
  - O-antigens



Paczosa and Mecsas 2016, MMBR; Wyres et al. 2020, Nat Rev Microbiol

The capsule and K-antigen types

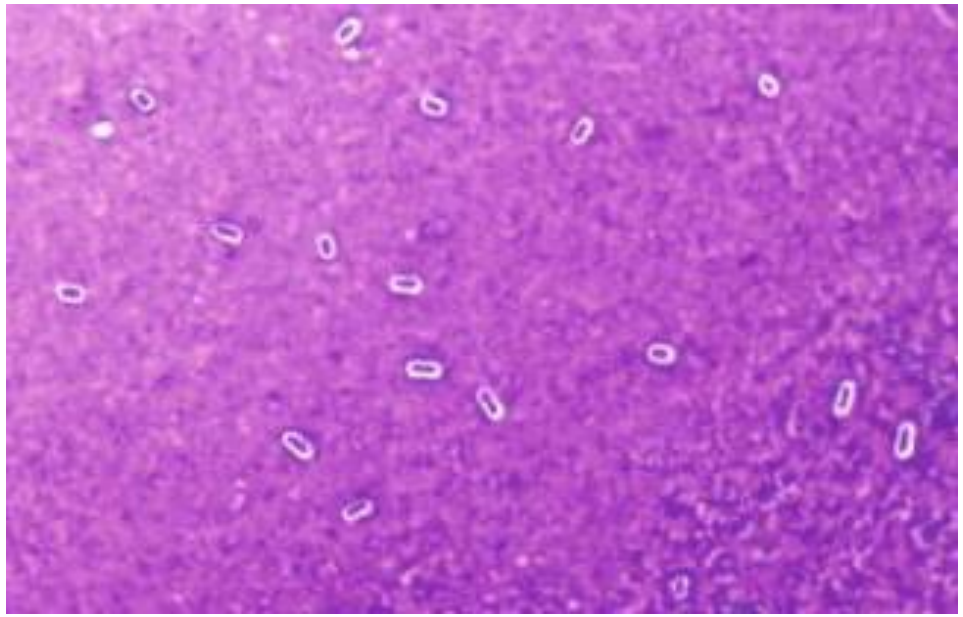
The capsule & K-antigen types

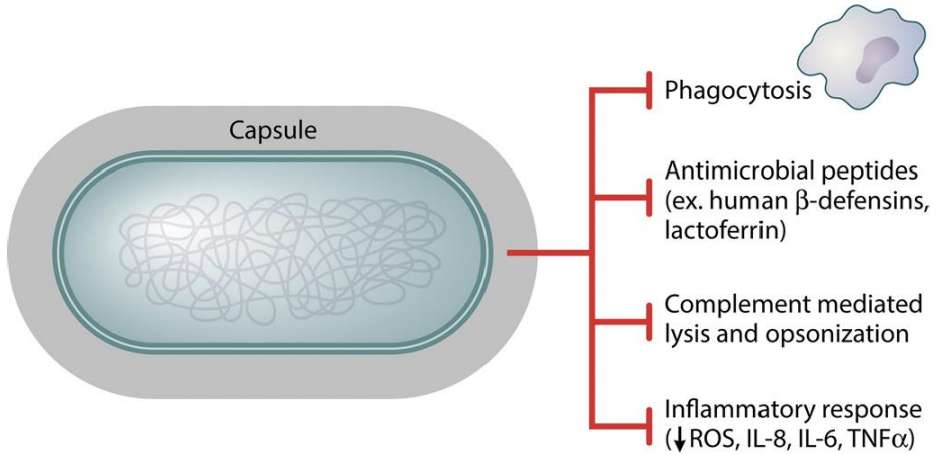
- The **capsule** is a a tight matrix of polysaccharide (repeating sugar units) that is tightly attached to the cell surface

- The capsule is a major virulence/pathogenicity factor that protects *K. pneumoniae*, e.g. from:
  - Phagocytosis
  - Serum killing
  - Desiccation
  - Predation (e.g. phage and protists)

- Laboratory serotyping
  - Developed 1916-1977
  - ~77 serotypes defined
  - Technically difficult
  - Costly & reagent production is complex
  - Some strains (10-70%) untypable
  - Issues with cross reactivity

- Molecular methods, e.g. RFLP, PCR
  - Limited resolution
  - Technically challenging due to genetic structure

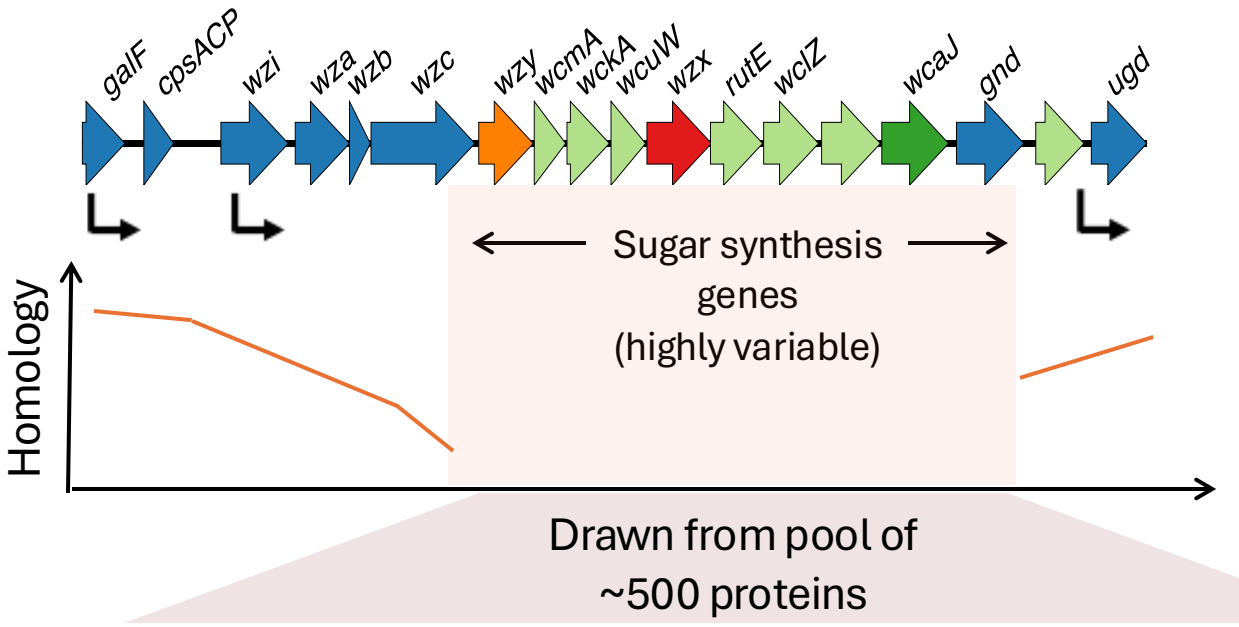




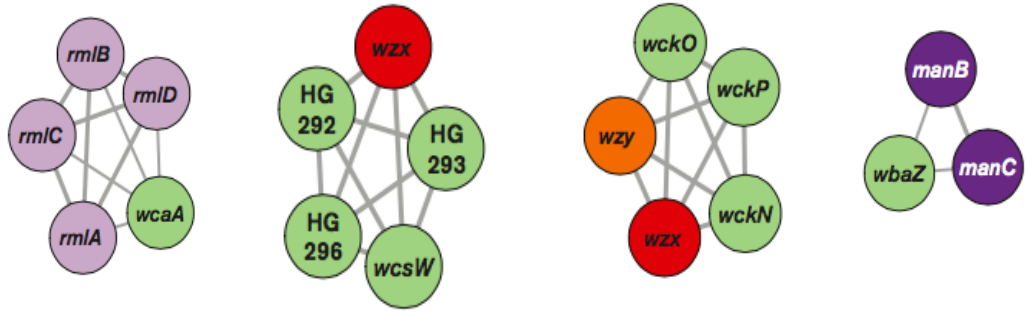
Paczosa and Mecsas 2016, MMBR; *asm.org/image-gallery/capsule-stain*

The capsular polysaccharide biosynthesis locus (*cps*)

- The capsule is encoded by the ***cps*** (capsular polysaccharide biosynthesis) locus
- 10-30kbp in size
- Adjacent to lipopolysaccharides (LPS) O-antigen locus
- Mosaic genetic structure
  - Terminal regions encode conserved genes mostly involved in capsule assembly and translocation
  - Central region is highly variable, encoding polysaccharide biosynthesis genes and other assembly genes (diversifying selection)



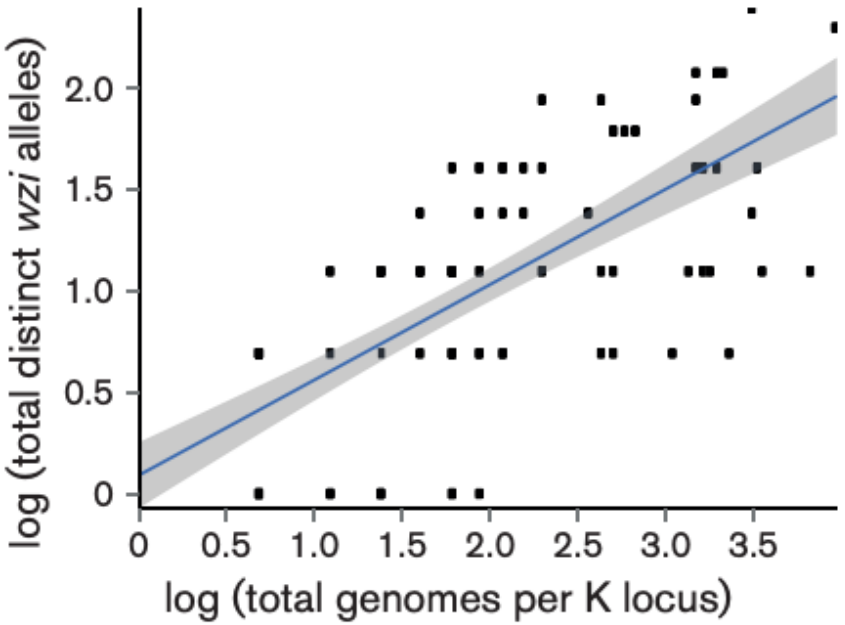
>160 distinct K-loci



Wyres et al. 2016, Microb Genom

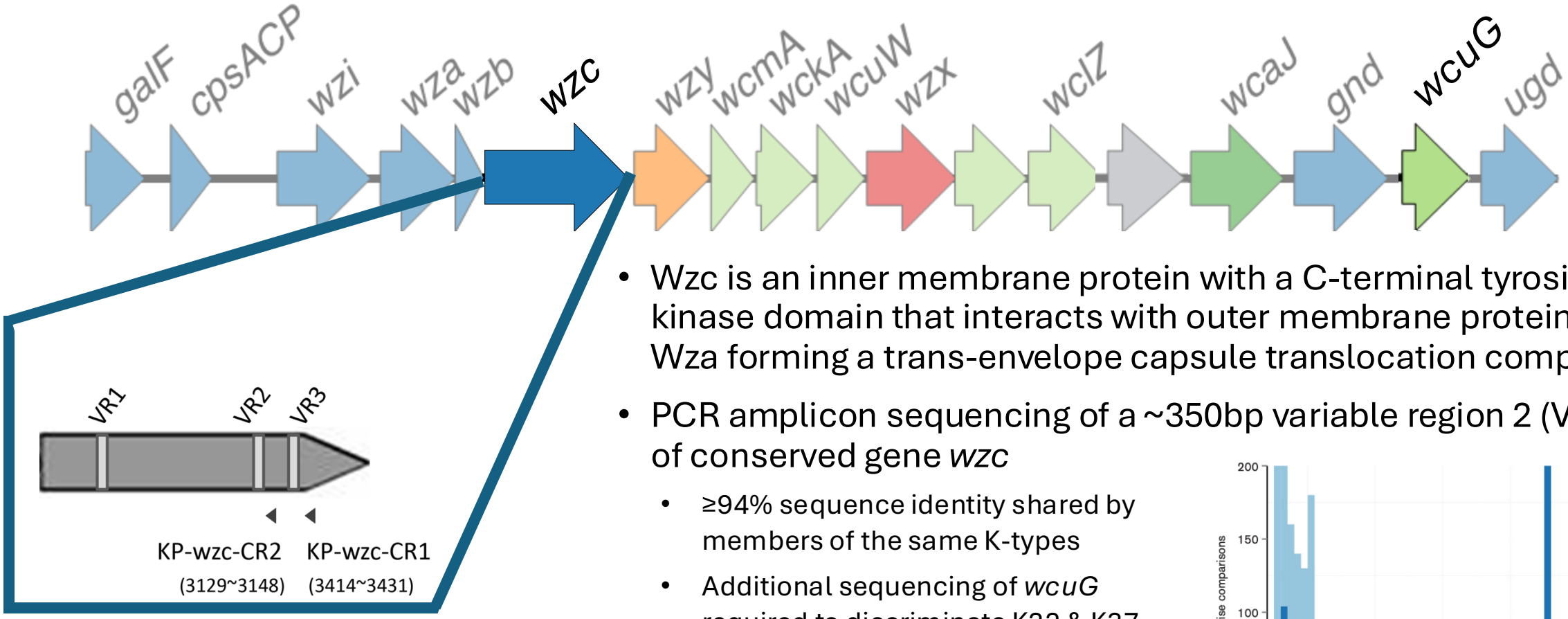
Previous methods of capsule typing: *wzi* typing

- The Wzi outer membrane protein is involved in the attachment of the capsular polysaccharide to the outer membrane
- PCR amplicon sequencing of 447nt 5' region of conserved gene *wzi*, sufficient to distinguish a set of 77 serotypes strains, with 94% accuracy
- Members of the same K-type share near-identical sequences
- Kleborate will use Wzi typing if K-typing with Kaptive is not specifically called
- Not suitable for non-KpSC members

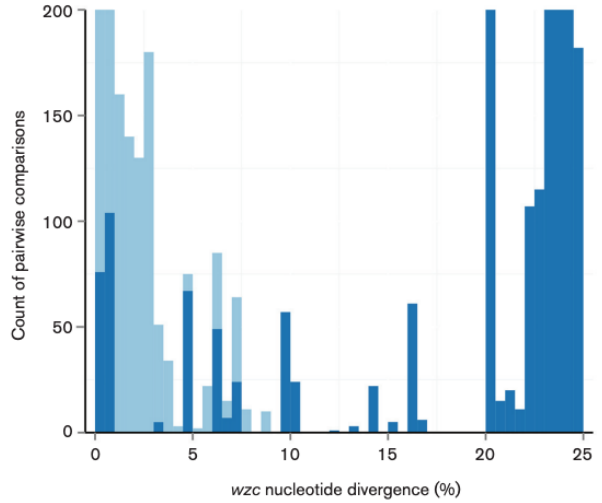


Brisse et al. 2013, J Clin Microbiol; Wyres et al. 2016, Microb Genom

Previous methods of capsule typing: *wzc* typing



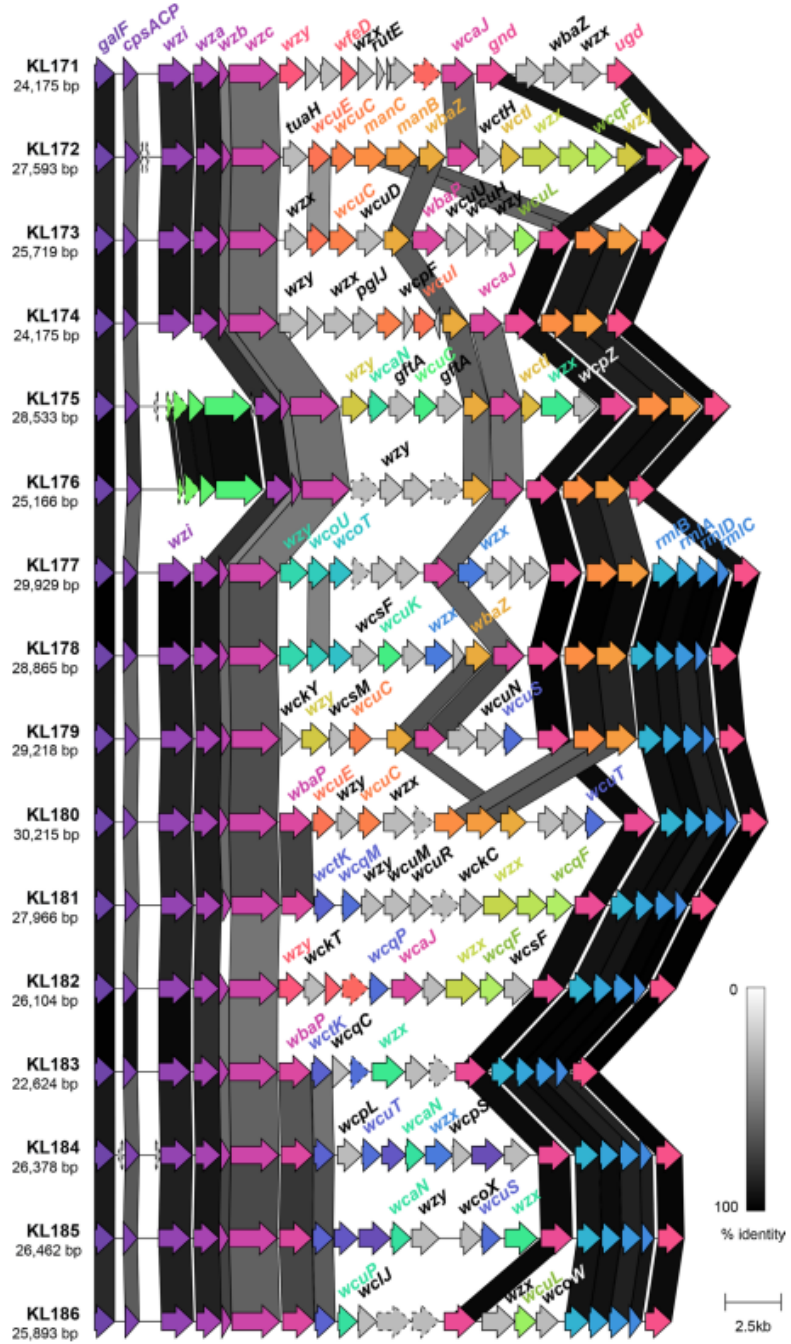
- Wzc is an inner membrane protein with a C-terminal tyrosine kinase domain that interacts with outer membrane protein Wza forming a trans-envelope capsule translocation complex
- PCR amplicon sequencing of a ~350bp variable region 2 (VR2) of conserved gene *wzc*
  - ≥94% sequence identity shared by members of the same K-types
  - Additional sequencing of *wcuG* required to discriminate K22 & K37 (frameshift mutation)
  - Additional analyses required for *wzc* deficient/acapsular K15 & K50 (transposase activity)



Pan et al. 2013, PLOS ONE; Wyres et al. 2016, Microb Genom

Capsule typing in the WGS era

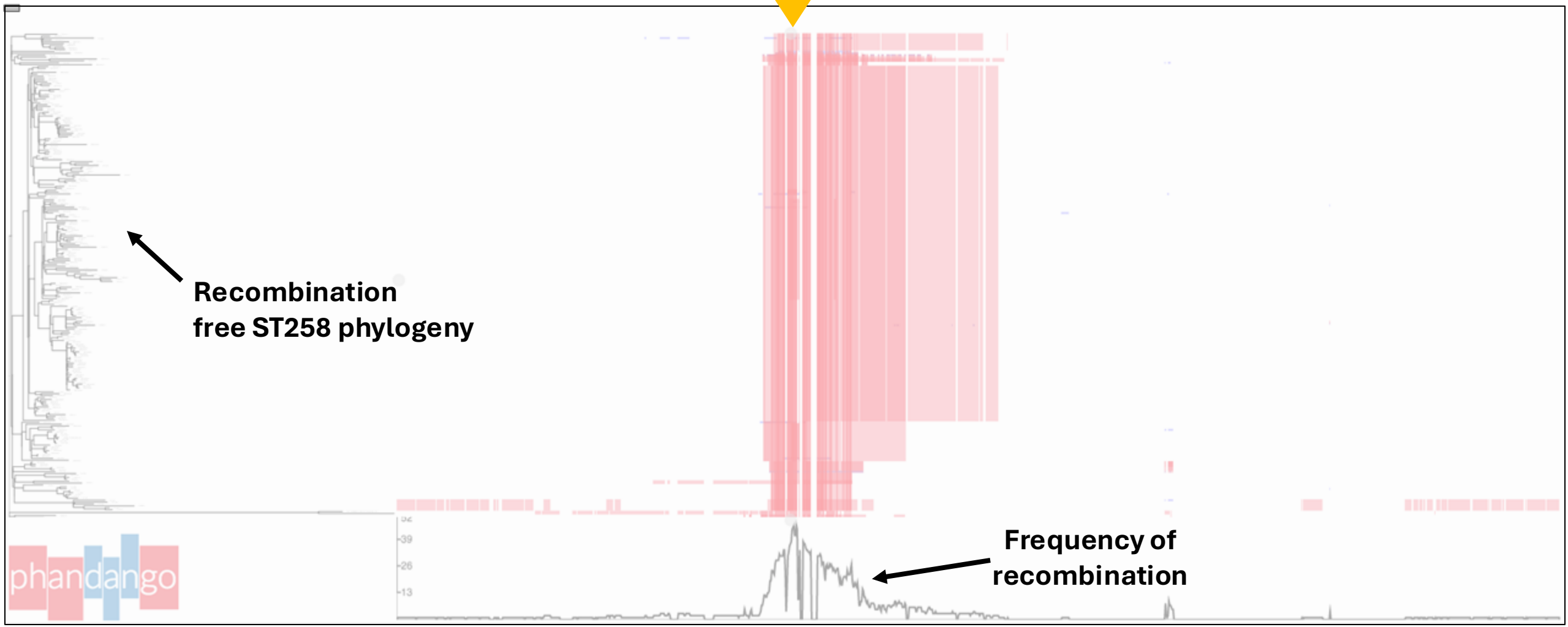
- Higher resolution over previous single-gene based methods as the entire locus can be examined
- More robust than single-gene typing methods, and eliminates cross-reactivity problems
- In addition to ~77 serologically defined (K-types), there are several capsule locus (KL-types) that can be defined genetically
  - >160 capsule types to date
  - KL>100 are serologically defined
  - KL<100 are genetically defined
- Serotype (capsule) switching occurs



Wyres et al. 2016, Microb Genom; Lam et al. 2022, Microb Genom; Gorrie et al. 2022, Nat Commun

The *cps* locus is a recombination hot spot

capsule biosynthesis (*cps*) locus



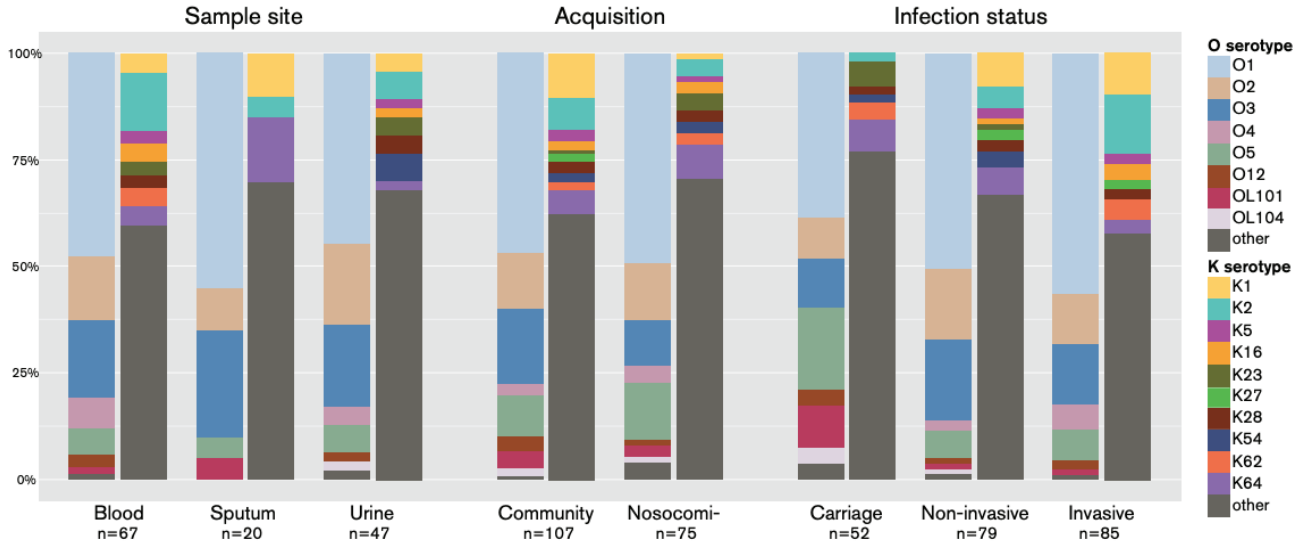
Wyres et al. 2019, PLoS Genetics. ***Phandango data visualization tool:*** https://jameshadfield.github.io/phandango/

Seroepidemiology of K-types

- Capsule has been proposed as a target for infection control strategies, e.g.
  - Vaccines
  - Monoclonal antibodies
  - Phage therapy

- Understanding the prevalence of different capsule types is critical for intervention strategies

- Different capsule types can be associated with infection types



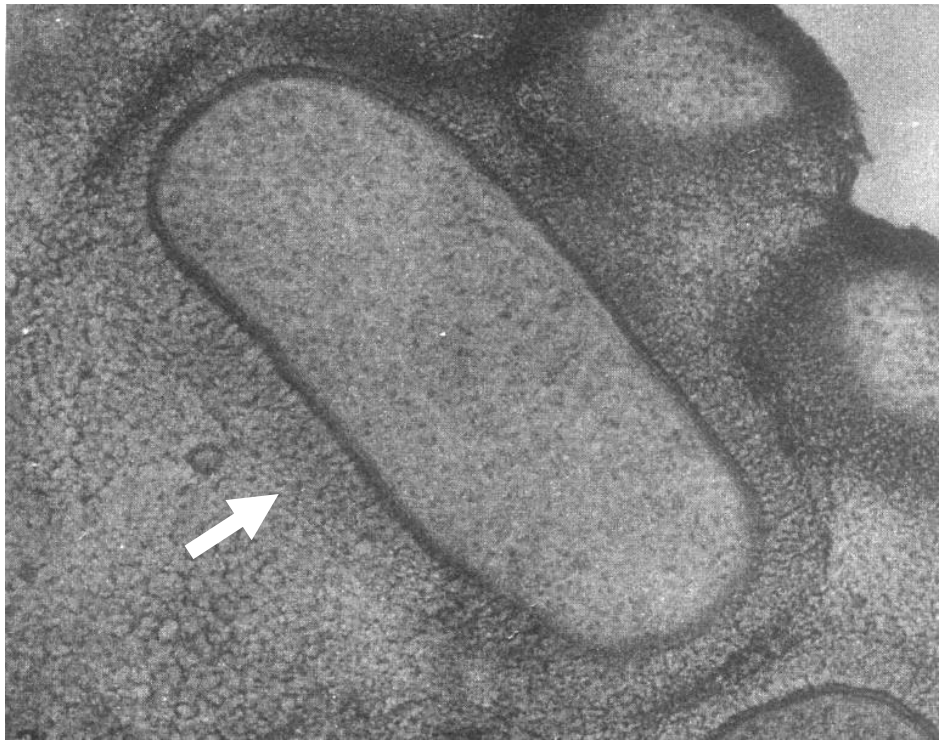
	Sample site						Acquisition				Infectiousness			
	Blood *n*=67		Urine *n*=47		Sputum *n*=20		Community *n*=107		Nosocomial *n*=75		Carriage *n*=52		Infection *n*=79		Invasive *n*=85	
K serotype																
K1	3	4.5 %	2	4.3 %	2	10.0 %	11	10.3 %	1	1.3 %	0	0.0 %	6	7.6 %	8	9.4 %
K2	9	13.4 %	3	6.4 %	1	5.0 %	8	7.5 %	3	4.0 %	1	1.9 %	4	5.1 %	12†	14.1 %
K5	2	3.0 %	1	2.1 %	0	0.0 %	3	2.8 %	1	1.3 %	0	0.0 %	2	2.5 %	2	2.4 %
K16	3	4.5 %	1	2.1 %	0	0.0 %	2	1.9 %	2	2.7 %	0	0.0 %	1	1.3 %	3	3.5 %
K23	0	0.0 %	0	0.0 %	0	0.0 %	1	0.9 %	3	4.0 %	3	5.8 %	1	1.3 %	0	0.0 %
K27	2	3.0 %	2	4.3 %	0	0.0 %	2	1.9 %	0	0.0 %	0	0.0 %	2	2.5 %	2	2.4 %
K28	2	3.0 %	2	4.3 %	0	0.0 %	3	2.8 %	2	2.7 %	1	1.9 %	2	2.5 %	2	2.4 %
K54	0	0.0 %	3	6.4 %	0	0.0 %	2	1.9 %	2	2.7 %	1	1.9 %	3	3.8 %	0	0.0 %
K62	3	4.5 %	0	0.0 %	0	0.0 %	2	1.9 %	2	2.7 %	2	3.8 %	0	0.0 %	4	4.7 %
K64	3	4.5 %	1	2.1 %	3	15.0 %	6	5.6 %	6	8.0 %	4	7.7 %	5	6.3 %	3	3.5 %
Others*	40	59.7 %	32	68.1 %	14	70.0 %	67	62.6 %	53	70.7 %	40	76.9 %	53	67.1 %	49	57.6 %

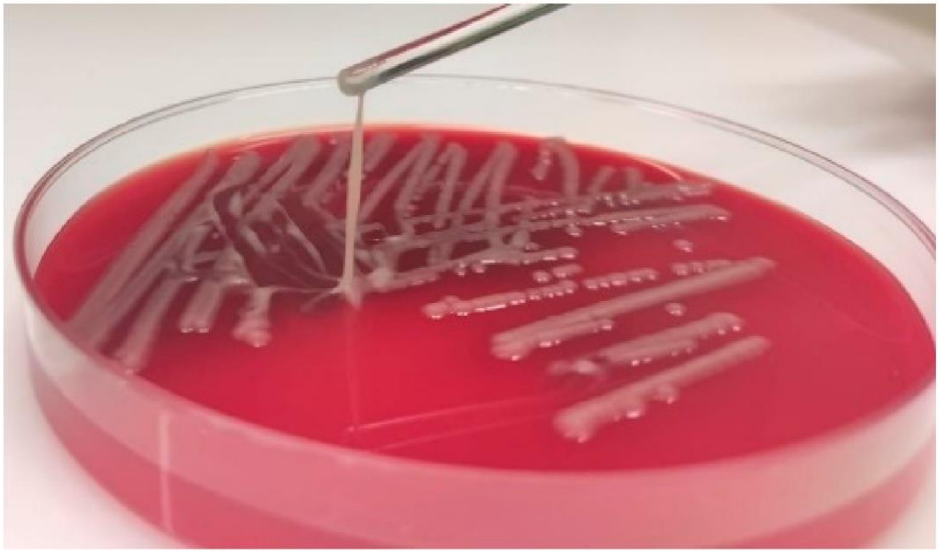
* Unidentified or not listed † Significant correlation (see text)

Follador et al. 2016, Microb Genom

Hypervirulent capsule types

- Some capsule types are associated with more severe infection types, e.g.
  - Pyogenic liver abscess
  - Meningitis
- Also known as a hypercapsule, hypermucoid, or hypermucoviscosity phenotype
  - Colony morphology
  - String test
  - Sedimentation assay
- Common hypervirulent capsule types include K1, K2, and K5
- These capsule types can be commonly found in clonal groups associated with CAI, e.g.
  - Highly serum resistant K1 capsule types are common among hypervirulent CG23





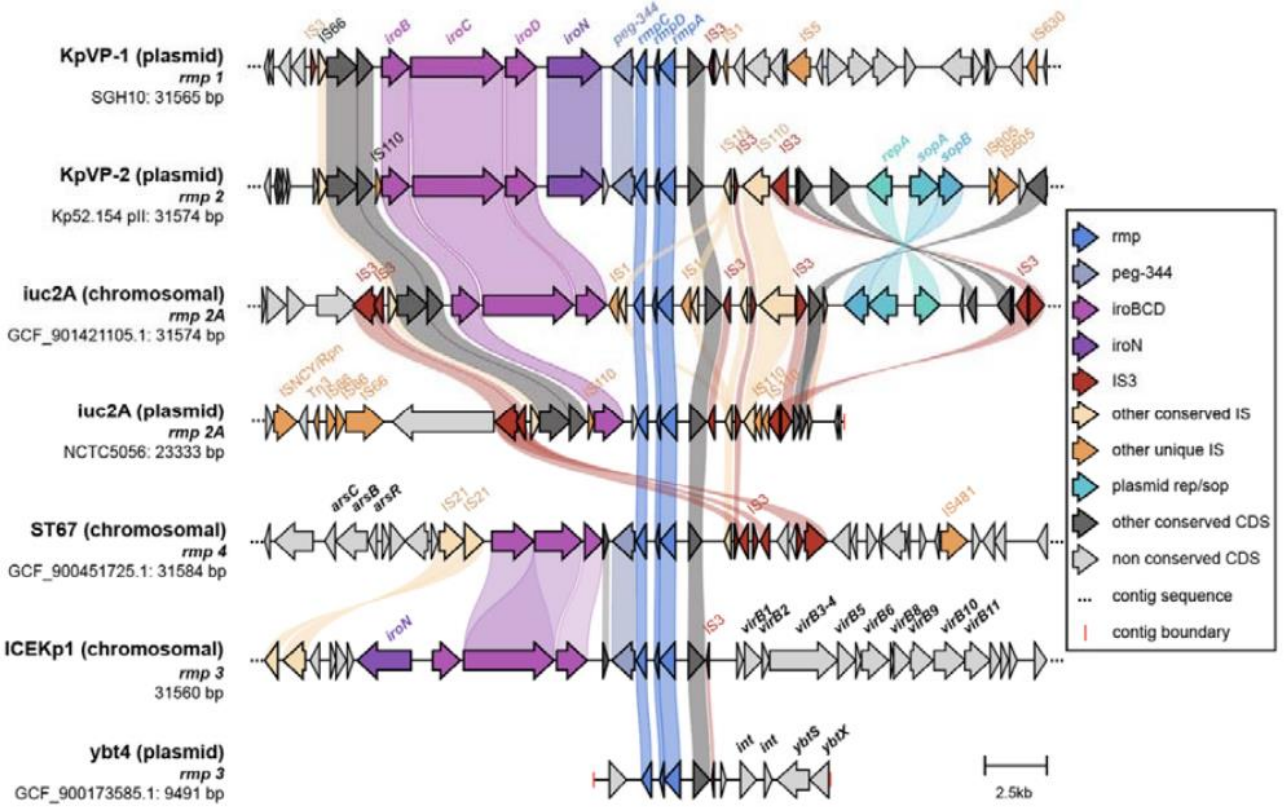
Lam et al. 2023, Nat Commun;
Klaper et al. 2020, Microorganisms; Wacharotayankun et al. 1993

Hypermucoid phenotypes are driven by *rmpADC*

- ***rmp*** genes, located in the *rmpADC* locus, are **r**egulators of **m**ucoid **p**henotype
  - Virulence determinants associated with hypermucoid phenotypes and hypervirulent clonal groups
  - ***rmpA*** regulates *rmpD* & *rmpC* transcription
  - ***rmpC*** is involved in upregulation of capsule expression
  - ***rmpD*** drives hypermucoviscosity

- Present in the chromosome, or on mobile genetic elements
  - Commonly co-localised on mobile elements with siderophores aerobactin & salmochelin

    e.g. Klebsiella virulence plasmid 1; KpVP-1



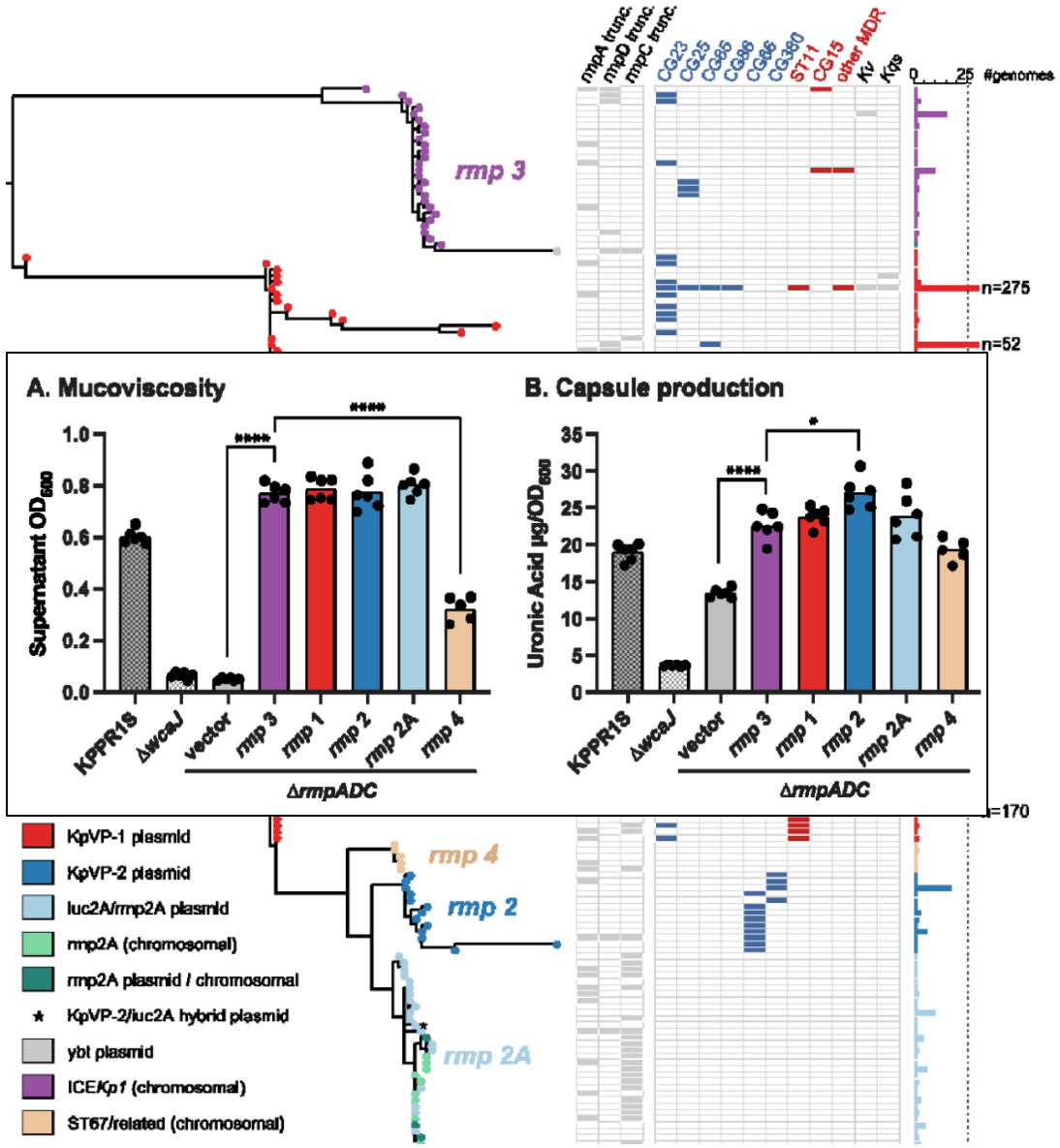
Lam et al. 2024, bioRxiv; Lam et al. 2018 ,Genom Med;
Hsu et al. 2011, Microbiol; Wacharotayankun et al. 1993, Infection Immun

Hypermucoid phenotypes are driven by *rmpADC*

- ***rmp*** genes, located in the *rmpADC* locus, are **r**egulators of **m**ucoid **p**henotype
  - Virulence determinants associated with hypermucoid phenotypes and hypervirulent clonal groups
  - ***rmpA*** regulates *rmpD* & *rmpC* transcription
  - ***rmpC*** is involved in upregulation of capsule expression
  - ***rmpD*** drives hypermucoviscosity

- Present in the chromosome, or on mobile genetic elements
  - Commonly co-localised on mobile elements with siderophores aerobactin & salmochelin
    e.g. Klebsiella virulence plasmid 1; KpVP-1

- Detected and typed by Kleborate
  - Five distinct lineages corresponding to different locations/mobile genetic elements
  - RmST for surveillance and tracking of variants
  - Some phenotypic differences

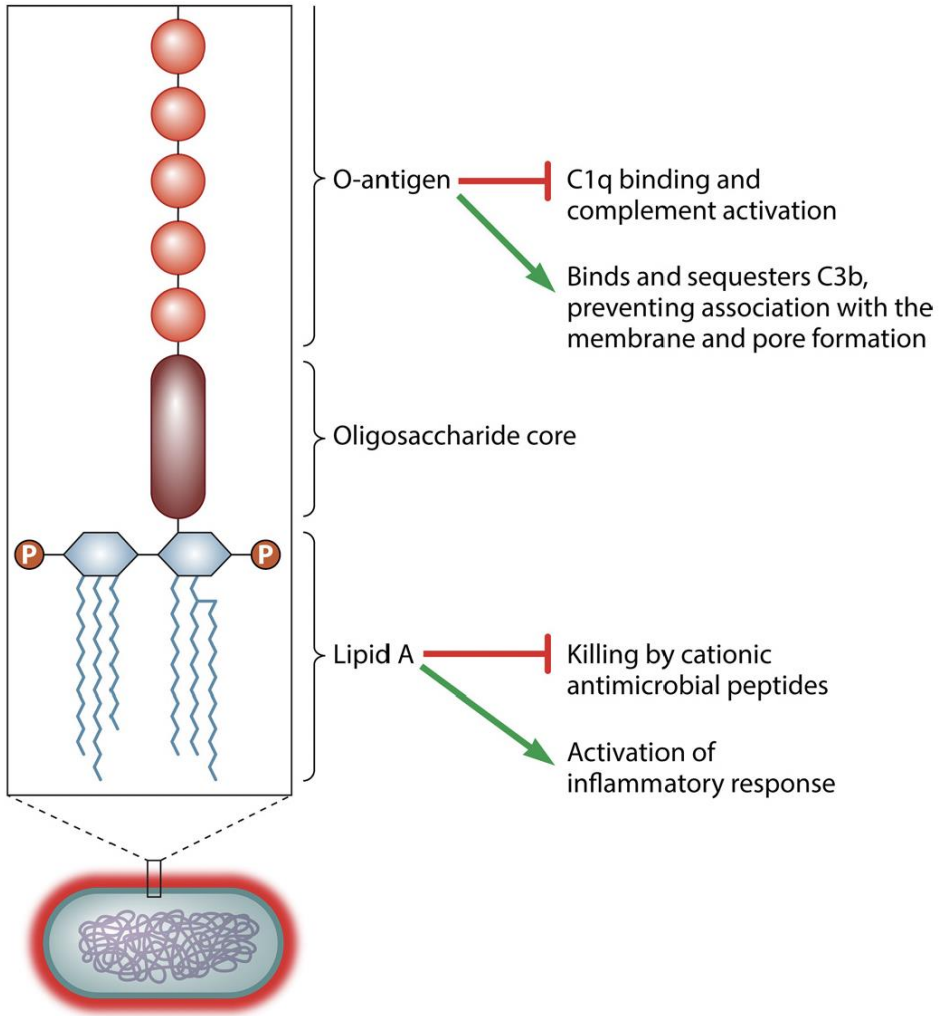


Lam et al. 2024, bioRxiv; Lam et al. 2018 ,Genom Med; Hsu et al. 2011, Microbiol; Wacharotayankun et al. 1993, Infection Immun

Lipopolysaccharide (LPS) and O-antigen types

Lipopolysaccharide (LPS) and O-antigen types

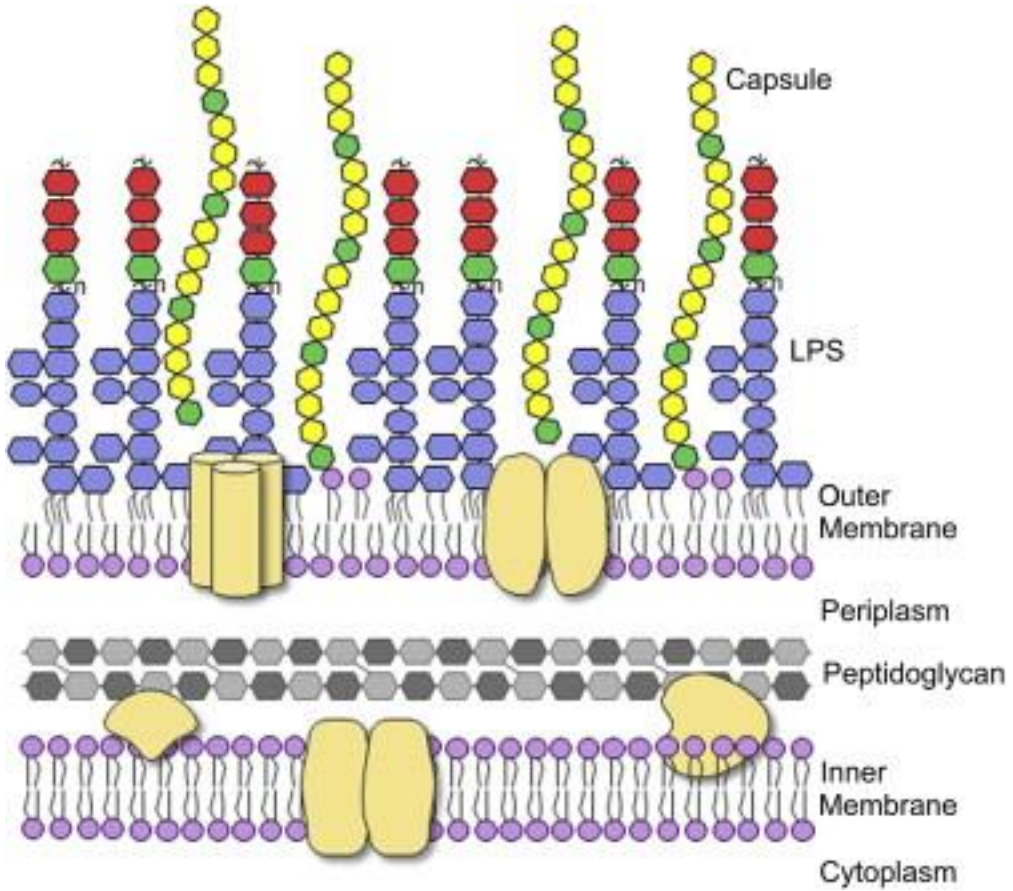
- **Lipopolysaccharide (LPS),** also known as endotoxin, is a major component of the cell membrane of all Gram-negative bacteria. It is comprised of:
  - **O-antigen** (*rfp* locus)
    - 9 serologically defined
    - 4 genetically defined)
  - **Core oligosaccharide** (*waa* genes)
    - 2 types defined
  - **Lipid A** (*lpx* genes)



Paczosa and Mecsas 2016, MMBR

Lipopolysaccharide (LPS) and O-antigen types

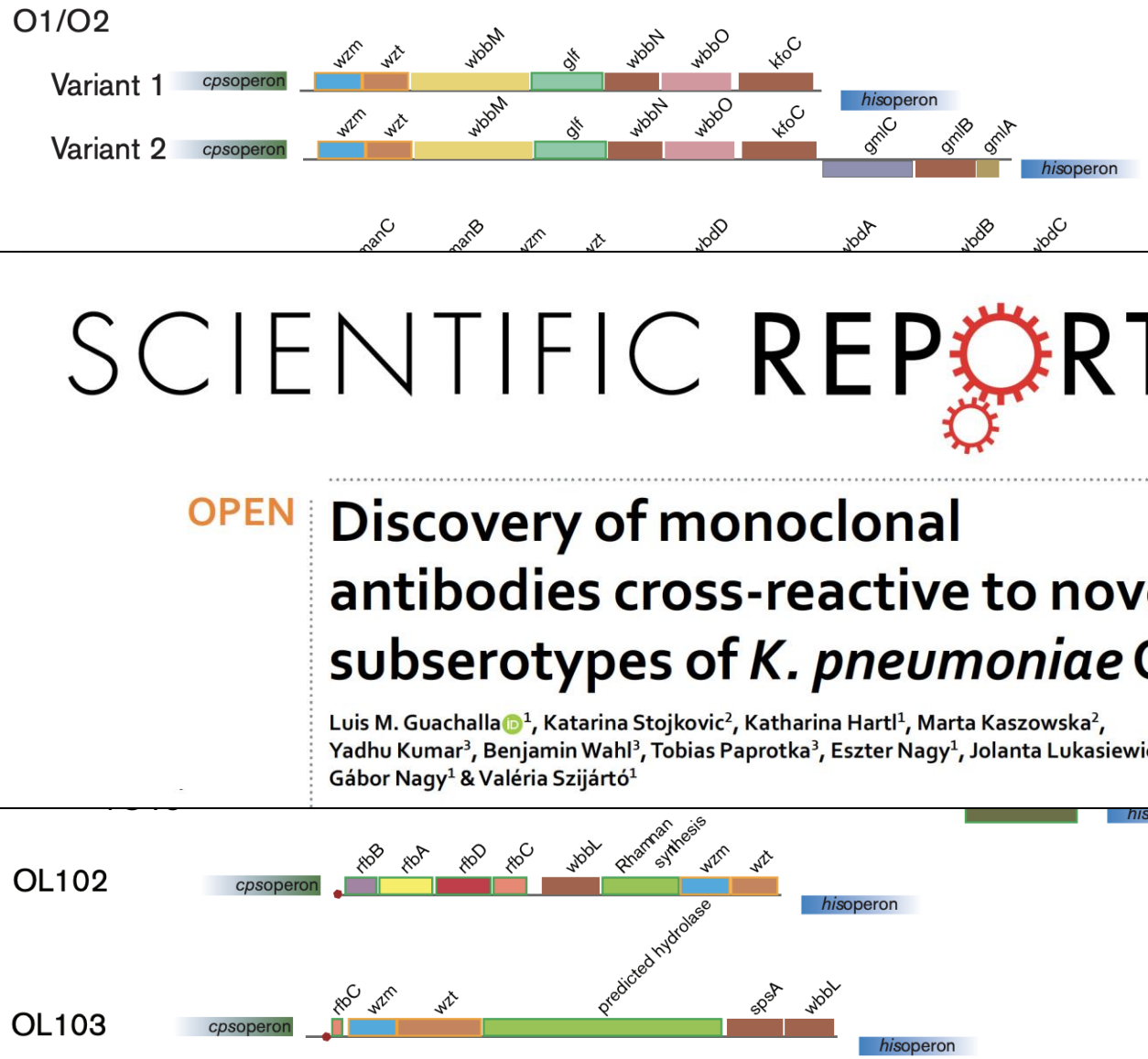
- **Lipopolysaccharide (LPS),** also known as endotoxin, is a major component of the cell membrane of all Gram-negative bacteria. It is comprised of:
  - **O-antigen** (*rfp* locus)
    - 9 serologically defined
    - 4 genetically defined
  - **Core oligosaccharide** (*waa* genes)
    - 2 types defined
  - **Lipid A** (*lpx* genes)
- O-antigen is less diverse than the K-antigen - potential target for control strategies but sometimes inaccessible due to the capsule (e.g. K1, K10, K16)
- Useful epidemiological marker for transmission studies



Willis and Whitfield 2013, Pathotypes and Principles of Pathogenesis

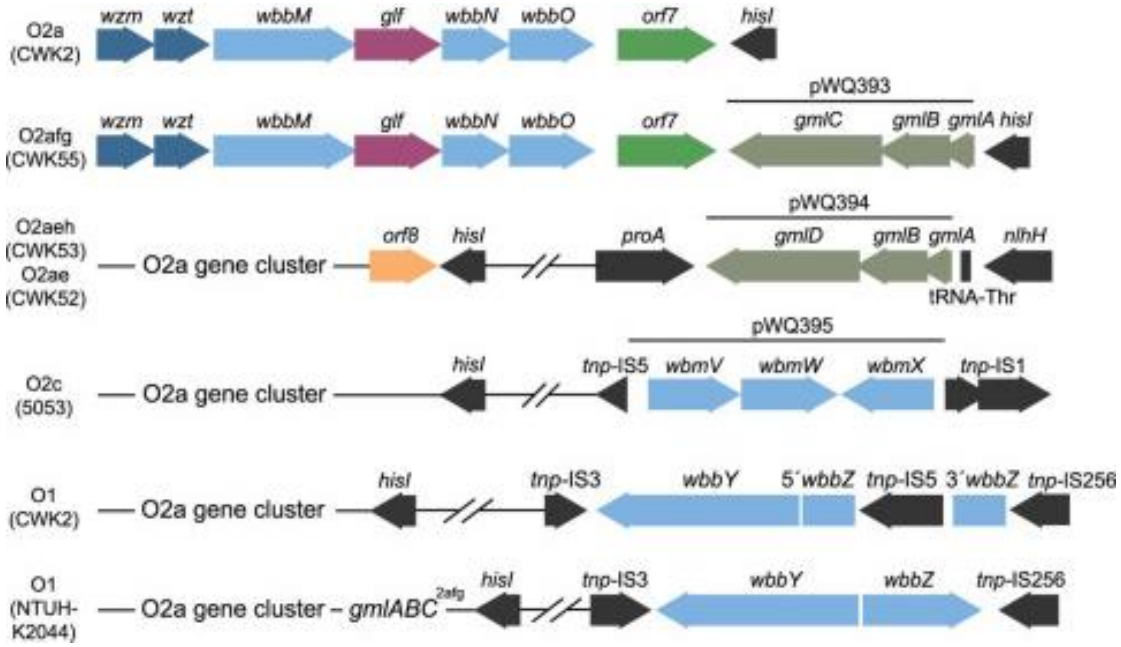
O-antigen and O-locus types

- O-antigen locus is 8-10kbp in size and adjacent to the *cps* locus
- Serologically defined O-antigen types are referred to as O-types
- Genetically defined O-antigen types are defined as O-locus types (prefix OL, numbered from 101 onwards)
- Additional variants of O3 (O3a/O3ab) demonstrate genetic variation in genes encoding mannose polymerization *(wbdD* and *wbdA*) described more recently
- Both O1/O2 O-antigen types can produce either O1 or O2 antigens (including subtypes) based on the presence of genes outside the *rfb* locus



Follader et al. 2016, Microb Genom

O1 and O2 antigen typing



O locus	Extra genes	Kaptive ≥v2.0	
		Locus	Type
O1/O2v1	None	O1/O2v1	O2a
O1/O2v2	None	O1/O2v2	O2afg
O1/O2v3	None	O1/O2v3	O2a
O1/O2v1	*wbbY*	O1/O2v1	O1
O1/O2v2	*wbbY*	O1/O2v2	O1
O1/O2v3	*wbbY*	O1/O2v3	O1
O1/O2v1	*wbmVW*	O1/O2v1	O2ac
O1/O2v2	*wbmVW*	O1/O2v2	O2ac
O1/O2v3	*wbmVW*	O1/O2v3	O2ac
O1/O2v1	*gmlABD*	O1/O2v1	O2aeh
O1/O2v2	*gmlABD*	O1/O2v2	O2aeh
O1/O2v3	*gmlABD*	O1/O2v3	O2aeh
O1/O2v1	*wbbY* AND *wbmVW*	O1/O2v1	O1 (O2ac)§
O1/O2v2	*wbbY* AND *wbmVW*	O1/O2v2	O1 (O2ac)§
O1/O2v3	*wbbY* AND *wbmVW*	O1/O2v3	O1 (O2ac)§

§Predicted antigenic serotype likely O1 but may also be O2ac (there is currently no corresponding type strain with *wbbY* and *wbmVW*).

e.g. For the O1/O2v1 locus variant

O1 = O1/O2v1 locus + *wbbY*

O2a = O1/O2v1 locus + no additional genes

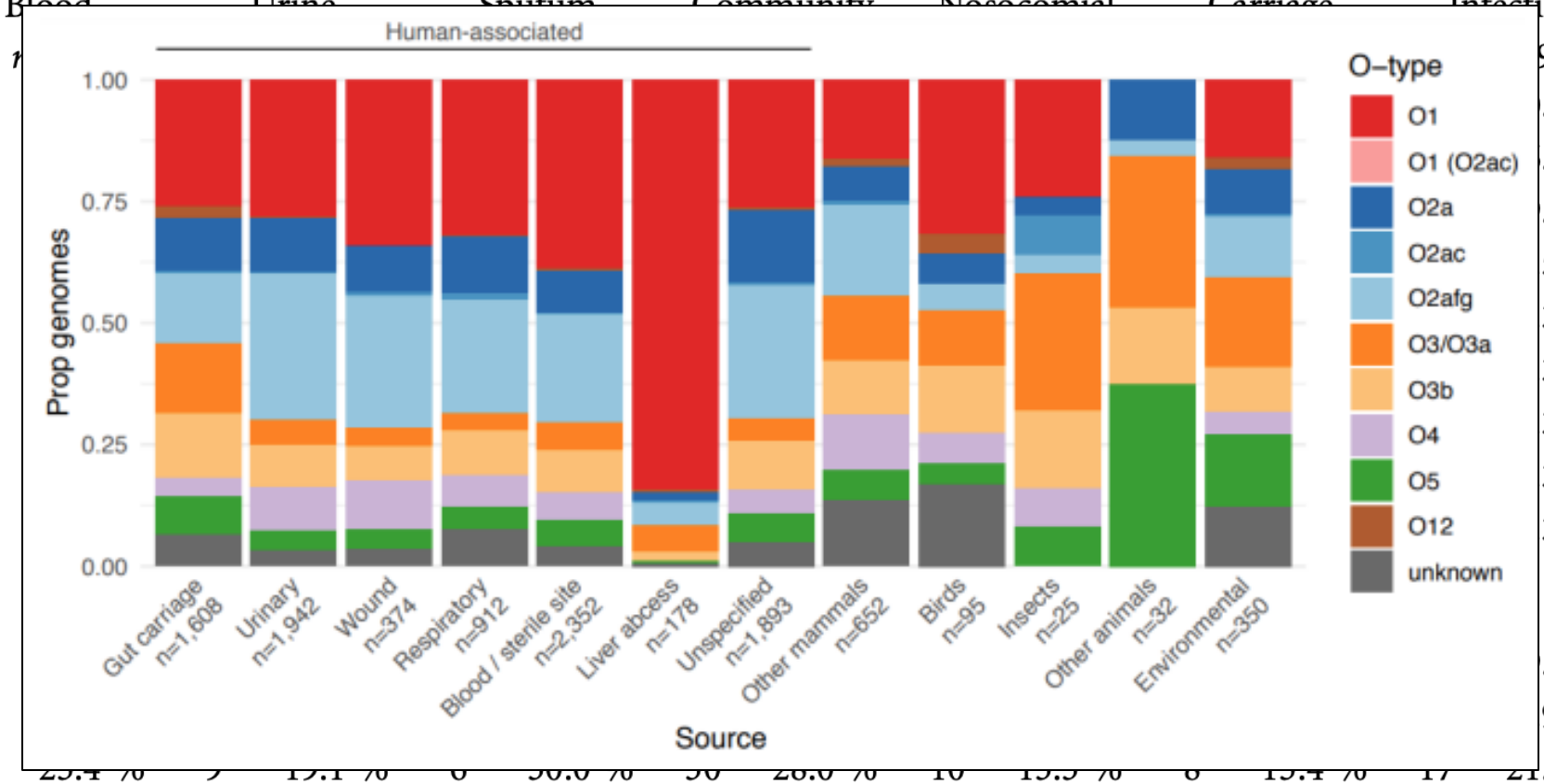
O2ac = O1/O2v1 locus + *wbmVW*

O2ach = O1/O2v1 locus + *gmlABD*

Clarke et al. 2018, JBC; Lam et al. 2022, Microb Genom

Seroepidemiology of O-antigen types

	Sample site				Acquisition			Infectiousness			
O serotype		Blood	Urine	Sputum	Community	Nosocomial	Carriage	Infection		Invasive	
		n						9		*n*=85	
O1	32							.6 %		48	56.5 %
O2	10							.5 %		10	11.8 %
O3	12							.0 %		12	14.1 %
O4	5							5 %		5	5.9 %
O5	4							3 %		6	7.1 %
O12	2							3 %		2	2.4 %
OL101	1							3 %		1	1.2 %
OL104	0							3 %		0	0.0 %
Others*	1							3 %		1	1.2 %
LPS core type											
Type 1	43							.6 %		55	64.7 %
Type 2	7							9 %		11	12.9 %
Others*	17							21.5 %		19	22.4 %



Lam et al. 2022, Microb Genom; Follador et al. 2016, Microb Genom

Typing methods provide useful nomenclature

1. To stratify cases into pathogen subtypes
   - To identify / define those with different genomic / biological traits and assess whether they have distinct epidemiology, so they can be managed in a targeted way
   - May consider phylogenetic relatedness to define groups, or use non-phylogenetic groupings

2. To investigate emergence and spread
   - Of the infectious disease generally, or variants of special clinical interest such as drug resistant or hypervirulent strains
   - Identify sources of infection, track transmission events, investigate outbreaks

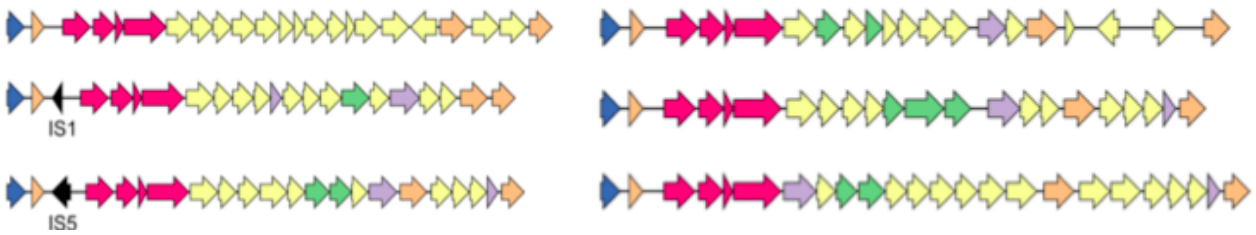
An introduction to Kaptive

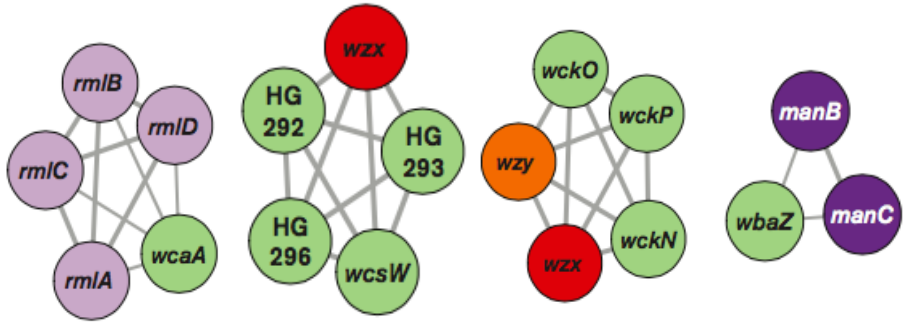
Kaptive: *K. pneumoniae* K- and O-antigen typing

- Kaptive is a bioinformatics software tool for rapidly identifying and typing capsule (K) and outer LPS (O) loci from whole genome sequence data

- Originally designed for use with the KpSC, Kaptive has been updated to include capsule typing for *Acinetobacter baumanni* via species specific databases

How Kaptive works

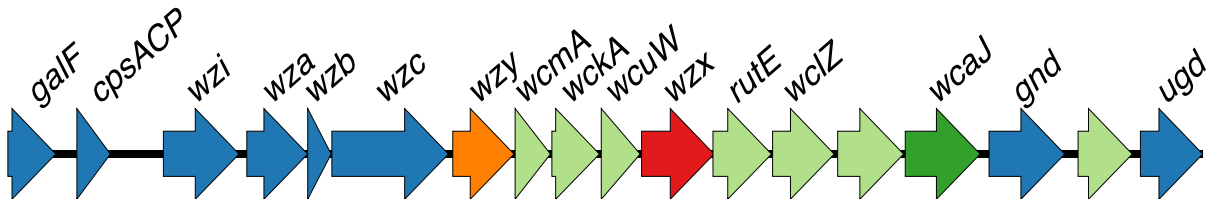




1. Minimap2 alignment for all full length loci

2. Minimap2 alignment for all Amino Acid sequences

- Extract best match
- Report any missing or extra genes



Lam et al. 2022, Microb Genom; Wyres et al. 2016, Microb Genom

How to access Kaptive

github.com/klebgenomics/Kaptive
kaptive-web.erc.monash.edu

Command line tool

- K-types
- O-types

Online tool

- K-types

github.com/klebgenomics/Kleborate

Command line tool

+ AMR
+ Species typing
+ Strain typing
+ Virulence typing

https://pathogen.watch/

Cloud-based (via Kleborate)

+ Phylogenetics
+ Context genomes
+ Species typing
+ Strain typing
+ AMR
+ Virulence typing

Also available via Galaxy Europe via Kleborate!

Kaptive via the command line

- Installation via:
  - conda package manager
  - PyPI/pip

- Three main options:
  - `assembly`
  - `extract`
  - `convert`

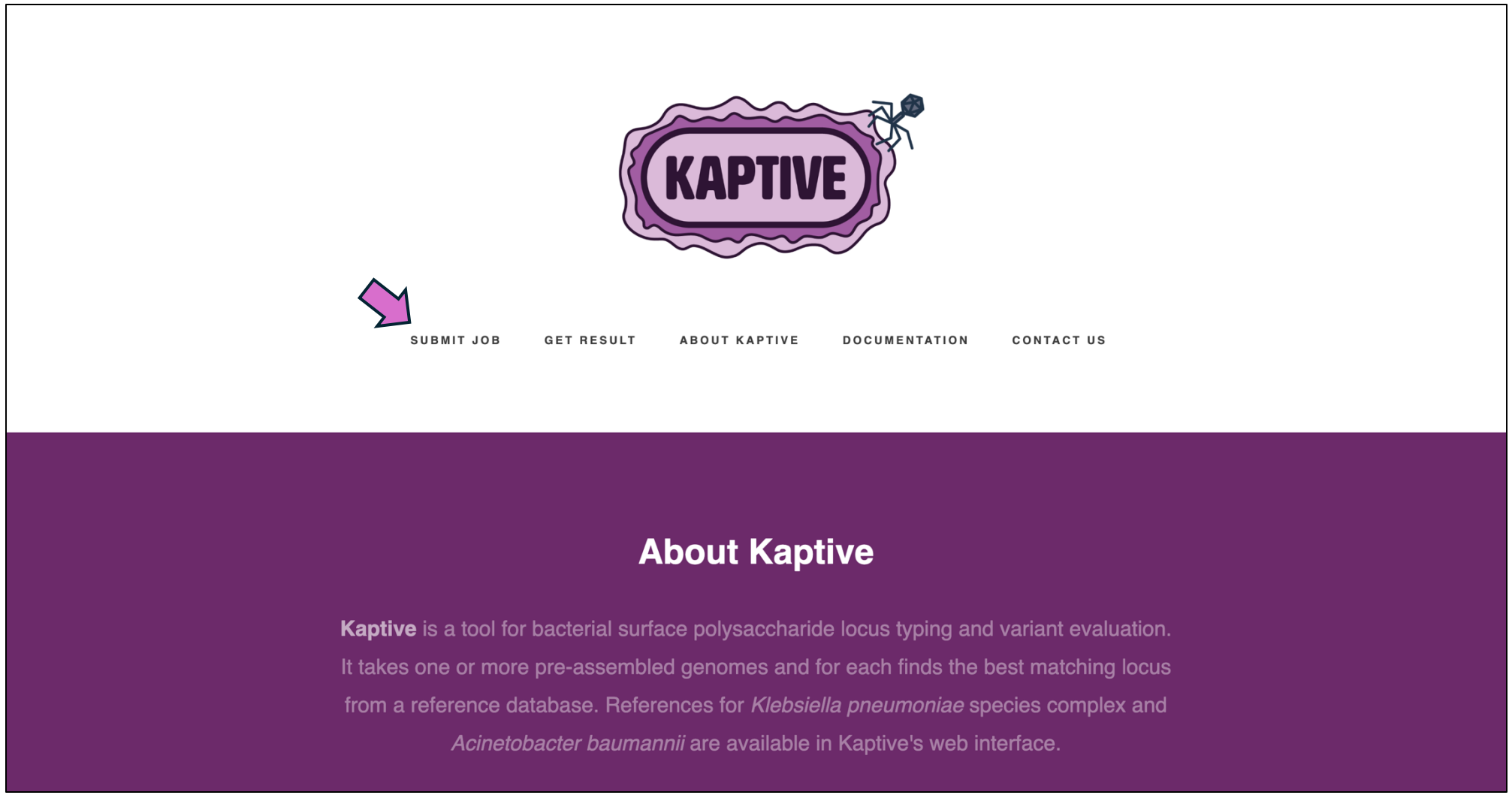
- Example analysis command:

```
kaptive assembly kpsc_k ./assemblies/*.fasta -o kaptive_results.tsv
```

https://github.com/klebgenomics/Kaptive
https://kaptive.readthedocs.io/en/latest/

Lam et al. 2022, Microb Genom; Wyres et al. 2016, Microb Genom

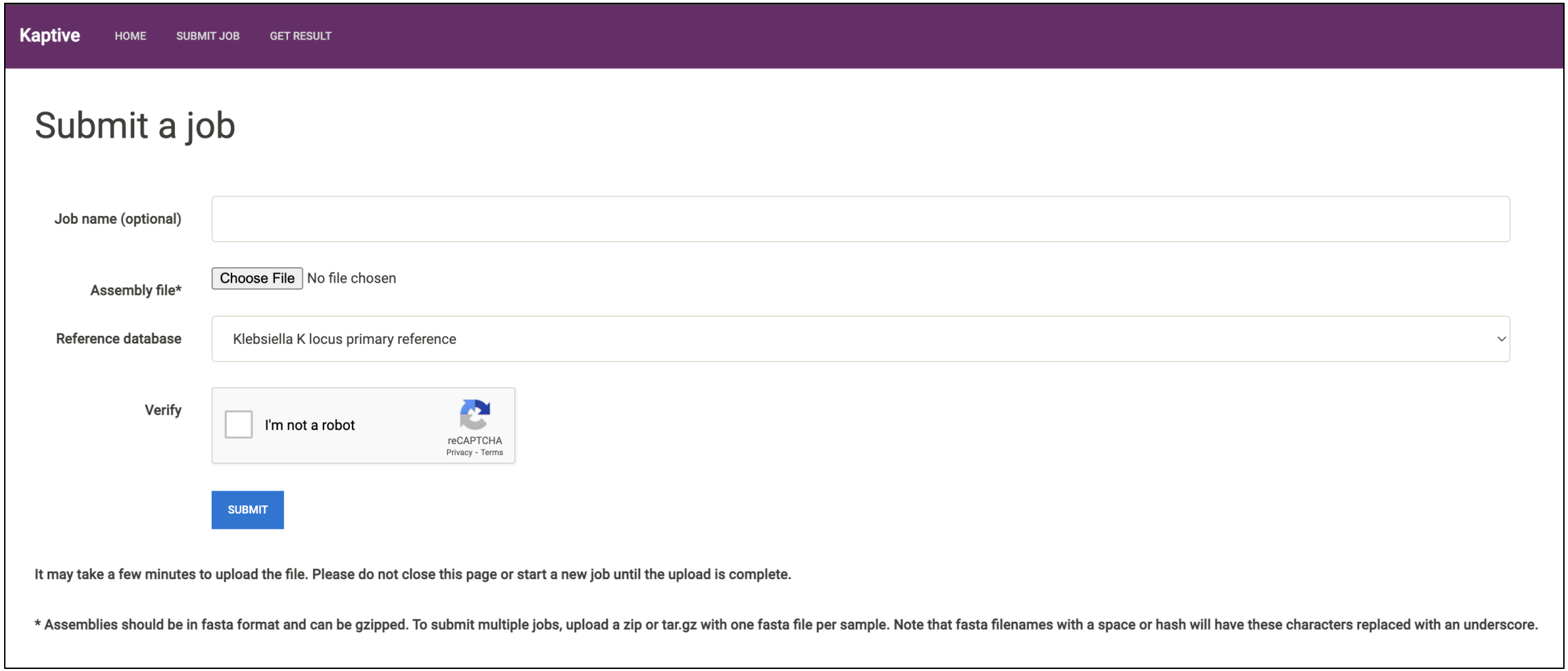
Kaptive via KaptiveWeb



Wick et al. 2018, J Clin Microiol

https://kaptive-web.erc.monash.edu/

Kaptive via KaptiveWeb



Wick et al. 2018, J Clin Microiol

https://kaptive-web.erc.monash.edu/

Kaptive via KaptiveWeb

Wick et al. 2018, J Clin Microiol

https://kaptive-web.erc.monash.edu/

Kaptive quality scores (version 3)

Score	Locus in single contig	Coverage (%)	Identity (%)	Missing genes	Extra genes
Typable	Yes	≥ 50 (nt)	≥ 82.5% (aa)	0	0
Typable	No	≥ 50 (nt)	≥ 82.5% (aa)	≤ 1	≤ 1
Untypable	Does not meet either of the typable criteria above				

Using 'Untypable' calls is NOT recommended!

Potential drivers of low-quality scores include:

- Fragmented K-locus
- Differences in locus length
- Missing genes
- Unexpected genes
- Divergent gene sequence

https://kaptive.readthedocs.io/
https://github.com/klebgenomics/Kaptive

Lam et al. 2022, Microb Genom; Wyres et al. 2016, Microb Genom

Kaptive quality scores (version 2)

Acceptable range = good or above

Score	Locus in single contig?	Coverage (%)	Identity (%)	Missing genes	Extra genes
Perfect	Required	100	100	0	0
Very high	Required	≥ 99	≥ 95	0	0
High	Required	≥ 99	N/A	≤ 3	0
Good	Not required	≥ 95	N/A	≤ 3	≤ 1
Low	Not required	≥ 90	N/A	≤ 3	≤ 2
None	None of the above				

Potential drivers of low-quality scores include:

- Fragmented K-locus
- Differences in locus length
- Missing genes (default threshold: 90% coverage, 80% identity)
- Unexpected genes (defaults threshold: 90% coverage, 80% identity)
- Divergent gene sequences (default threshold: 95% identity)

Lam et al. 2022, Microb Genom; Wyres et al. 2016, Microb Genom

https://github.com/klebgenomics/Kaptive

Any questions or reflections?