Klebsiella pneumoniae genomic epidemiology and antimicrobial resistance

Klebsiella pneumoniae genome analysis with Kleborate

9 September 2025

Intended Learning Objectives

Specific objectives of this session:

1. Overview of key virulence loci in *K. pneumoniae*
2. Learn about the genotyping tool Kleborate
    a. Features of Kleborate
    b. How to interpret the output
    c. Ongoing/future developments

Outline

This session consists of the following elements

1. Overview of the key virulence loci of *K. pneumoniae*
   a. *What are they and what are their roles in (hyper)virulence?*
   b. *How do they mobilize/transmit between strains?*
   c. *Genetic diversity within each loci; development of virulence typing schemes*
2. Overview of Kleborate genotyping/surveillance framework
   a. *Features of the tool*
   b. *How to run the tool via command line or Pathogenwatch*
   c. *Example outputs; how to interpret the output*

Klebsiella pneumoniae virulence

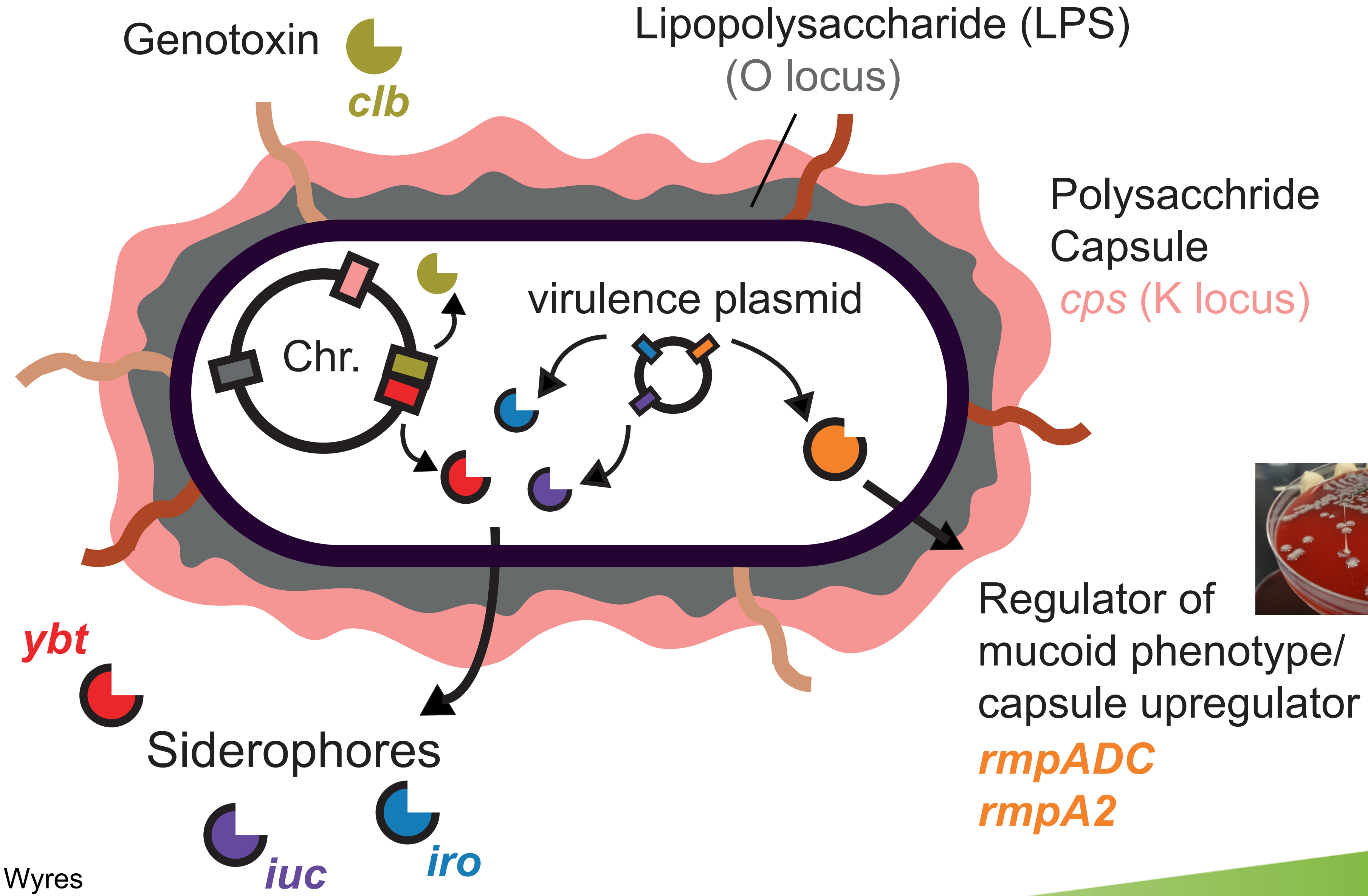
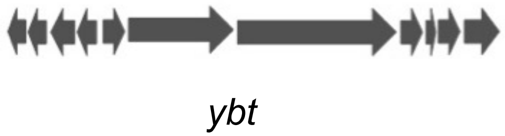


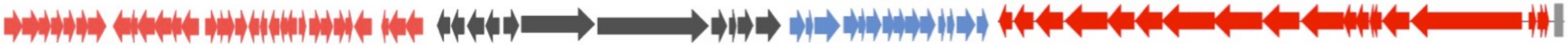
Image credit: A/Prof Kelly Wyres

Yersiniabactin - mobilisation

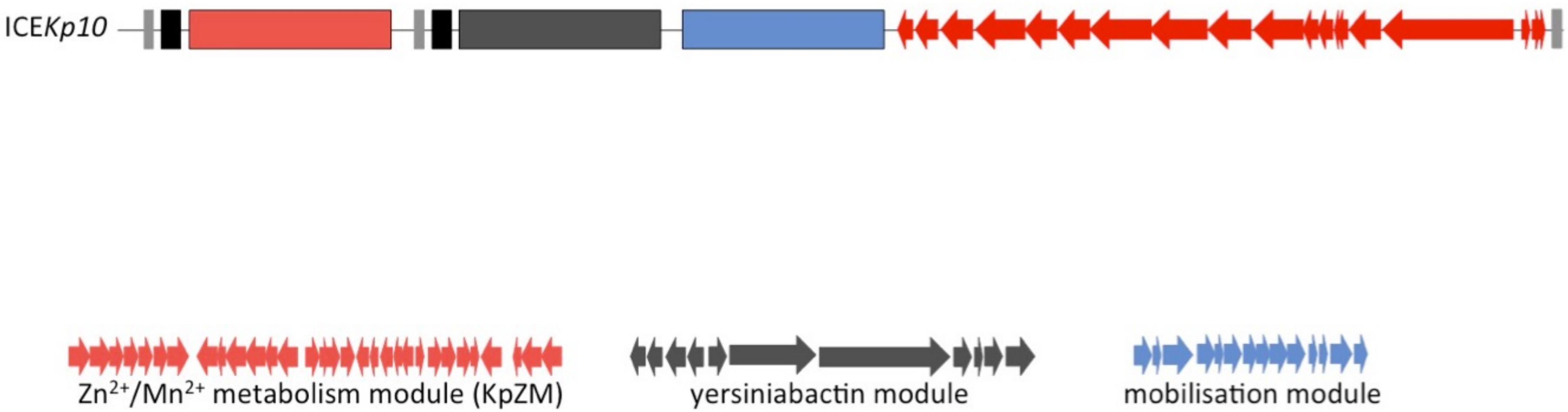


Yersiniabactin - mobilisation

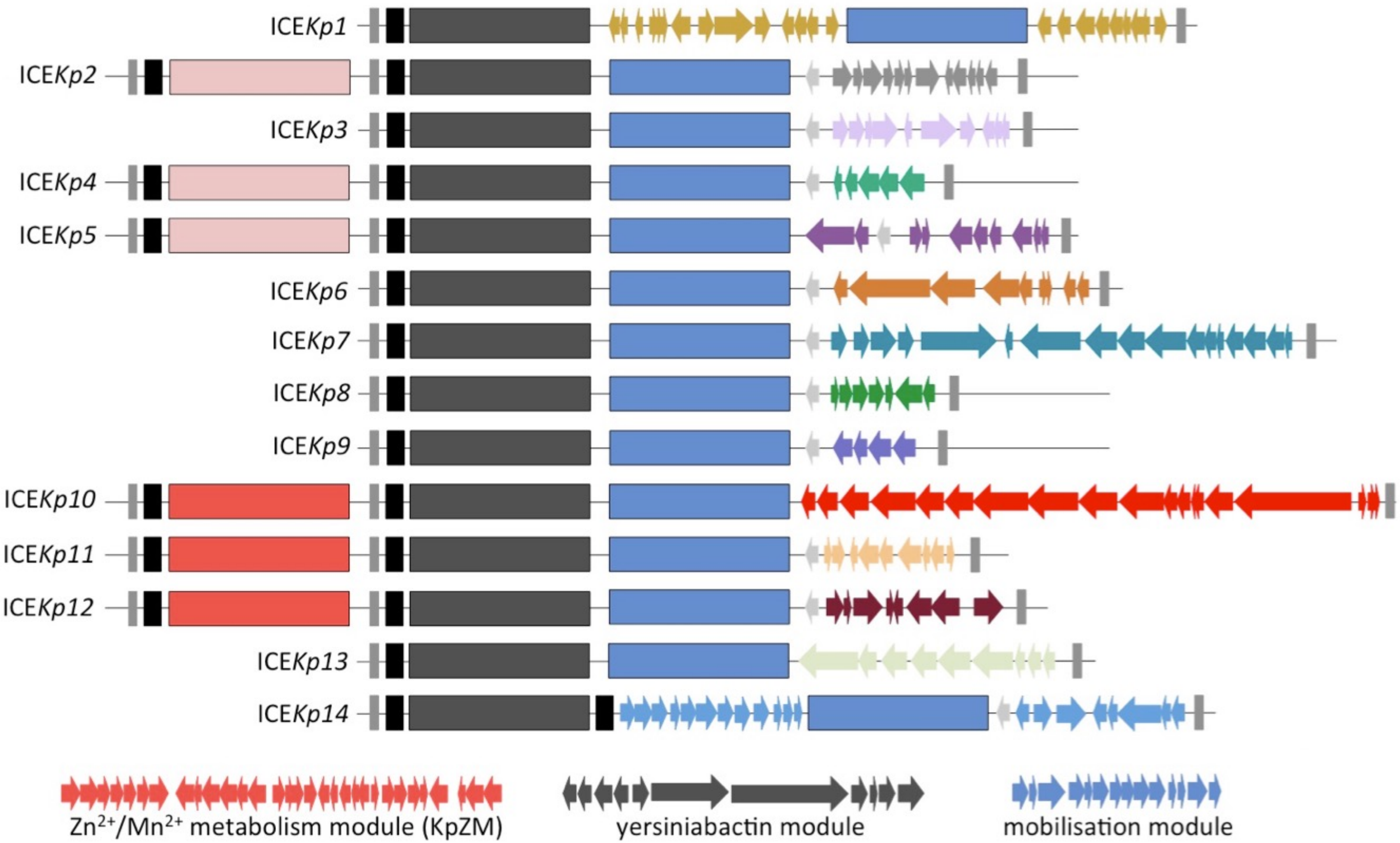
ICE*Kp10:*



Yersiniabactin - mobilisation



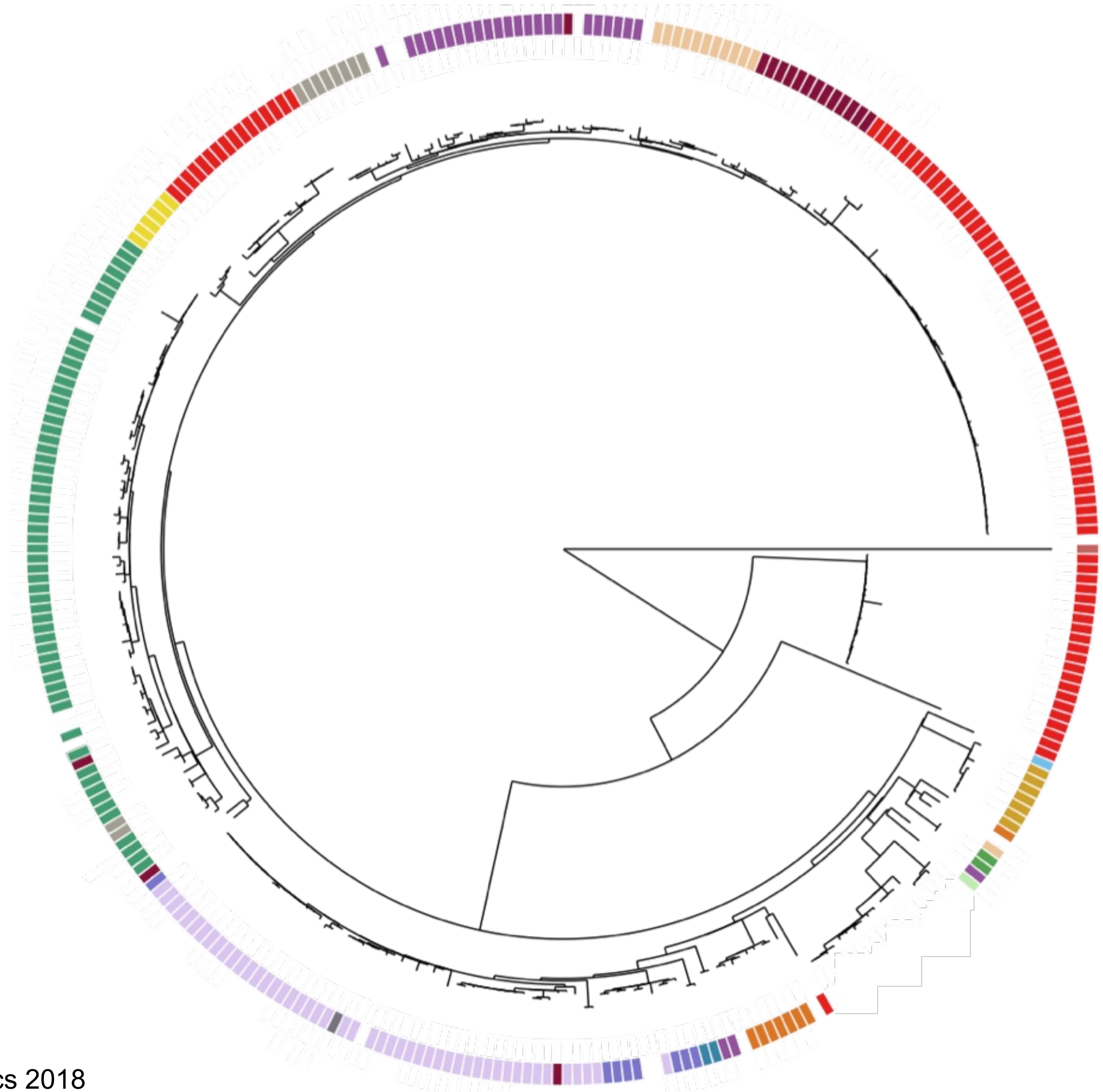
Yersiniabactin - mobilisation; 14 ICE*Kp* variants

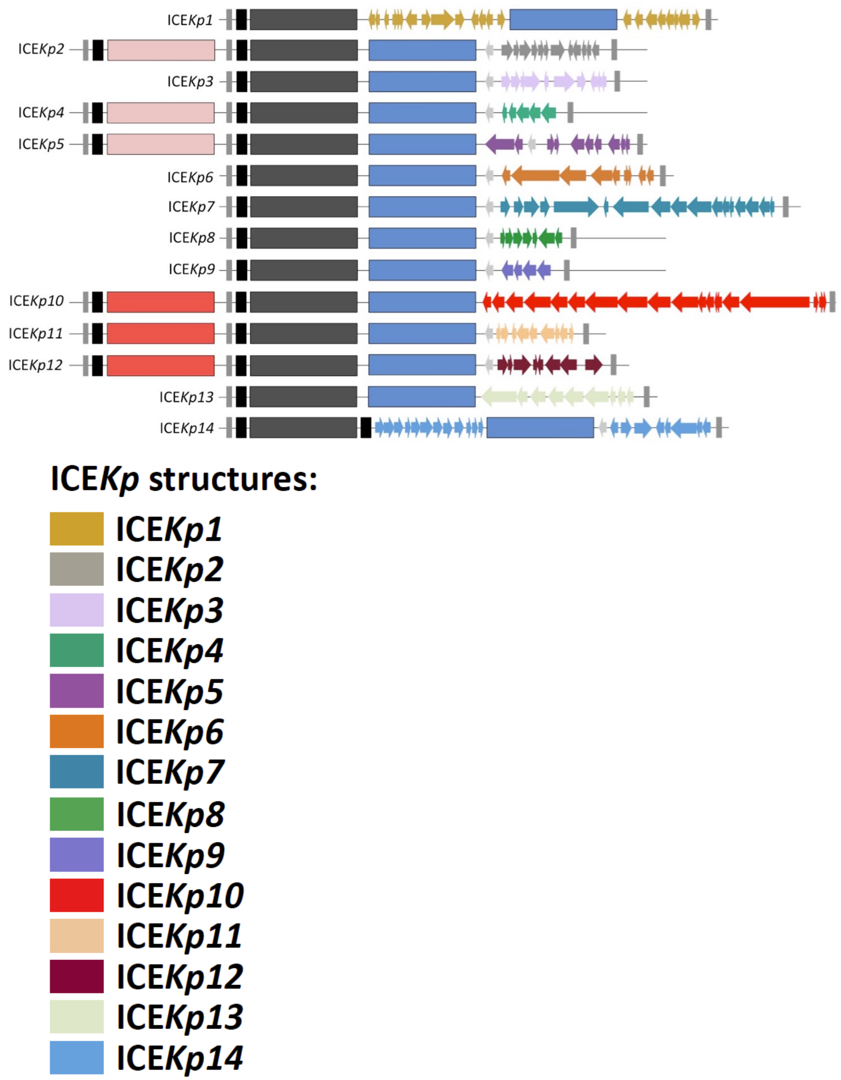


Lam et al, Microbial Genomics 2018

Yersiniabactin - genetic diversity

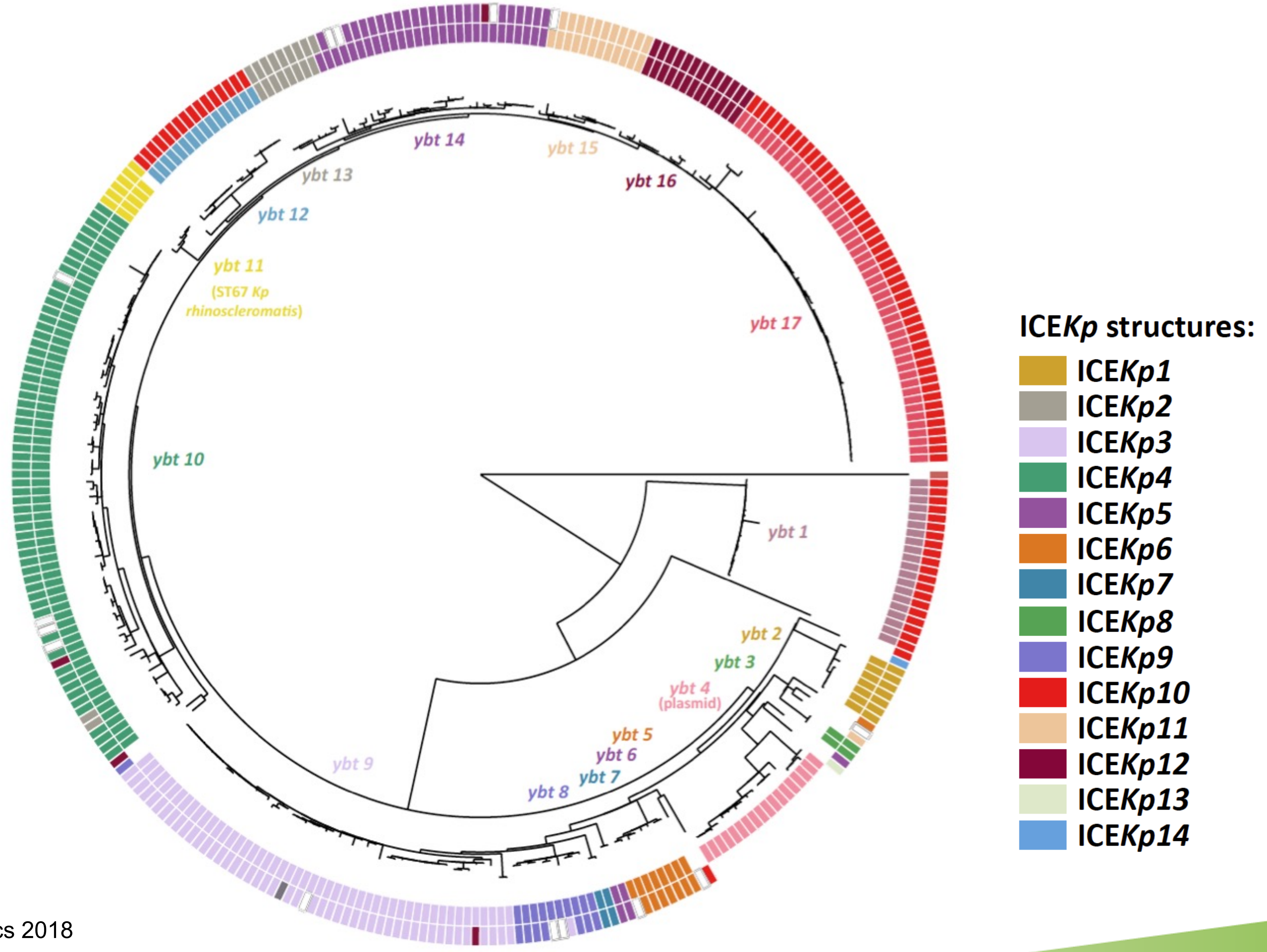
Yersiniabactin - genetic diversity





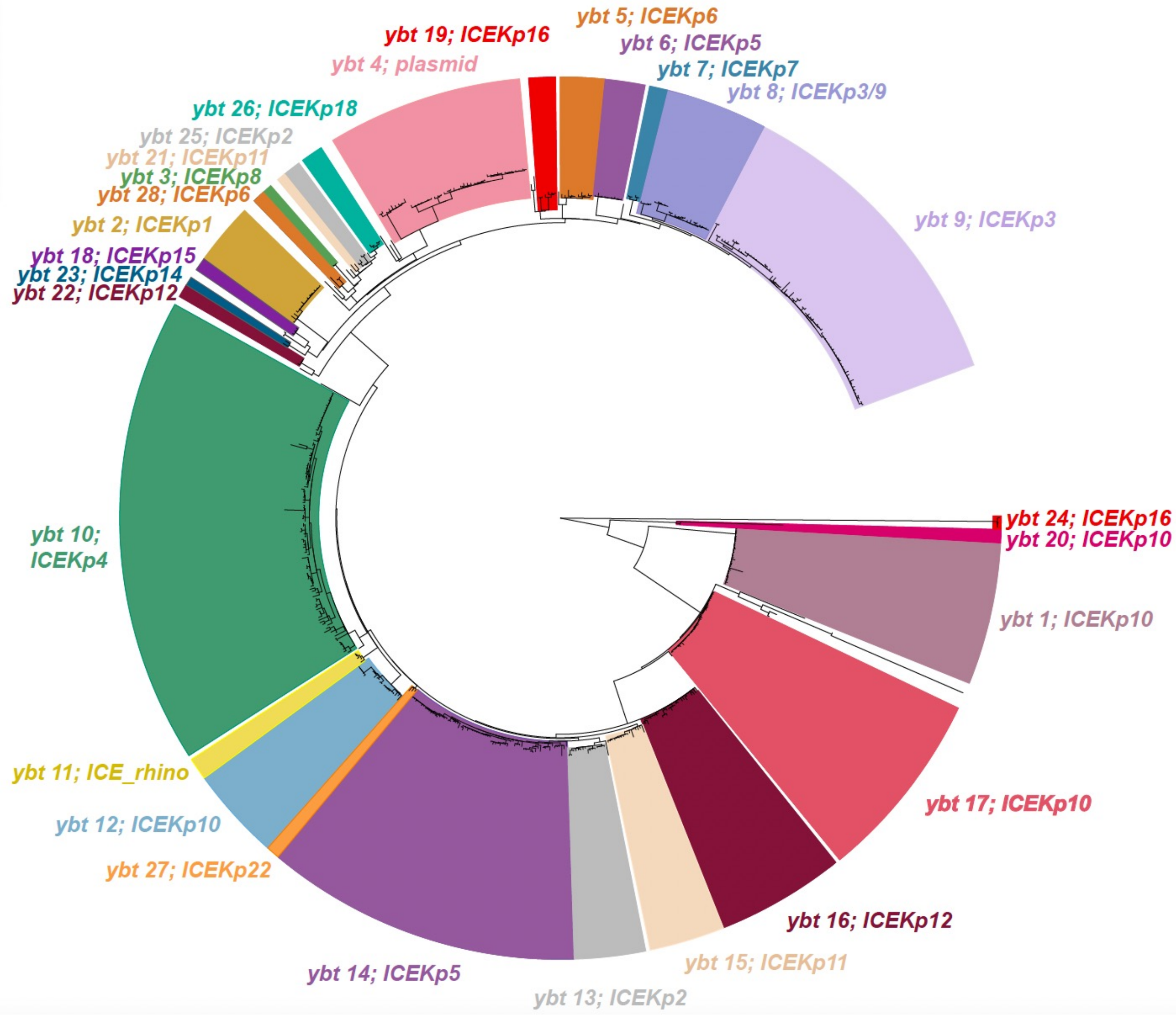
Lam et al, Microbial Genomics 2018

Yersiniabactin - genetic diversity is linked with mobile elements (ICEKp variants and plasmid)



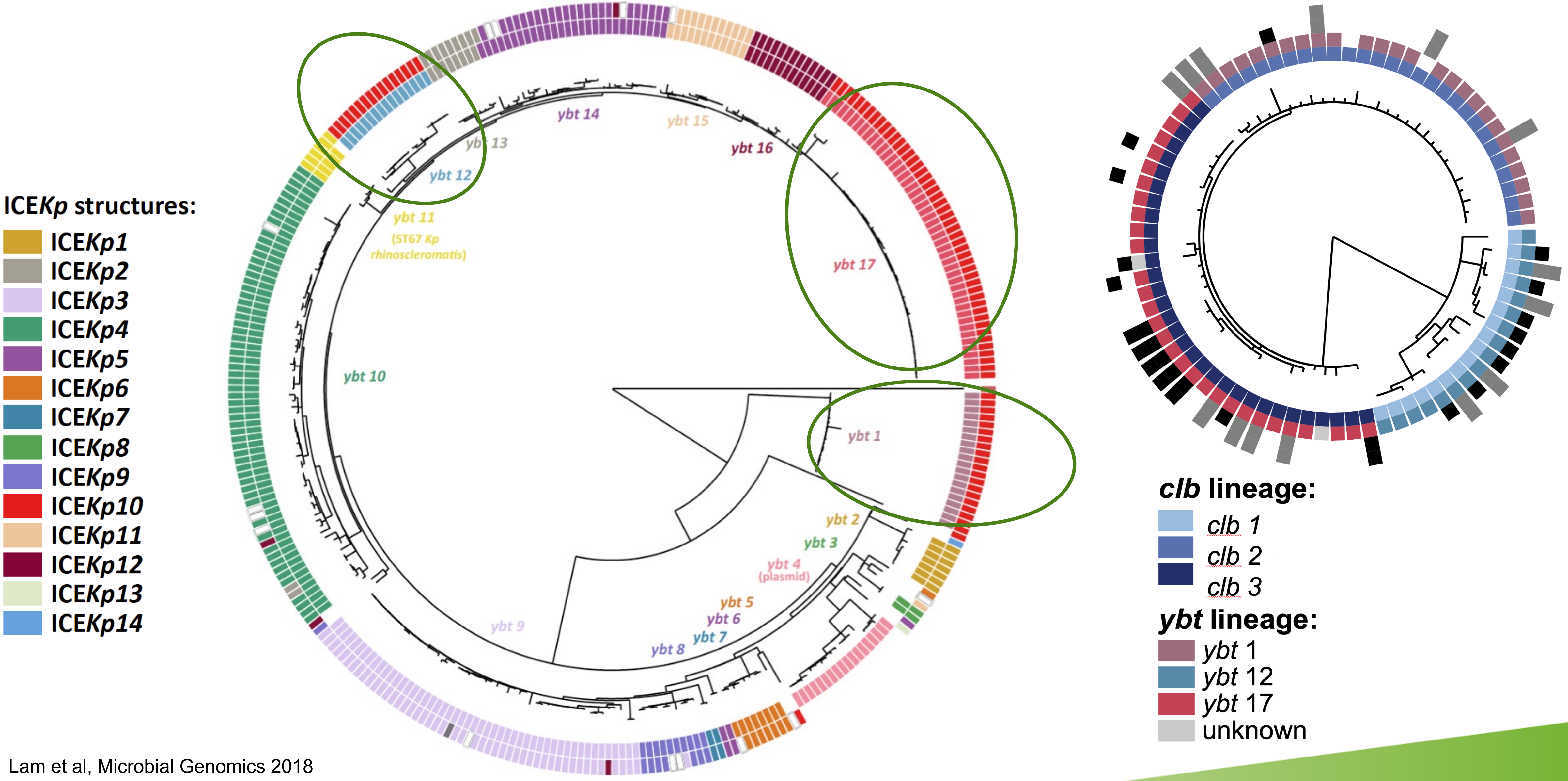
Lam et al, Microbial Genomics 2018

Yersiniabactin - genetic diversity is linked with mobile elements (ICEKp variants and plasmid)



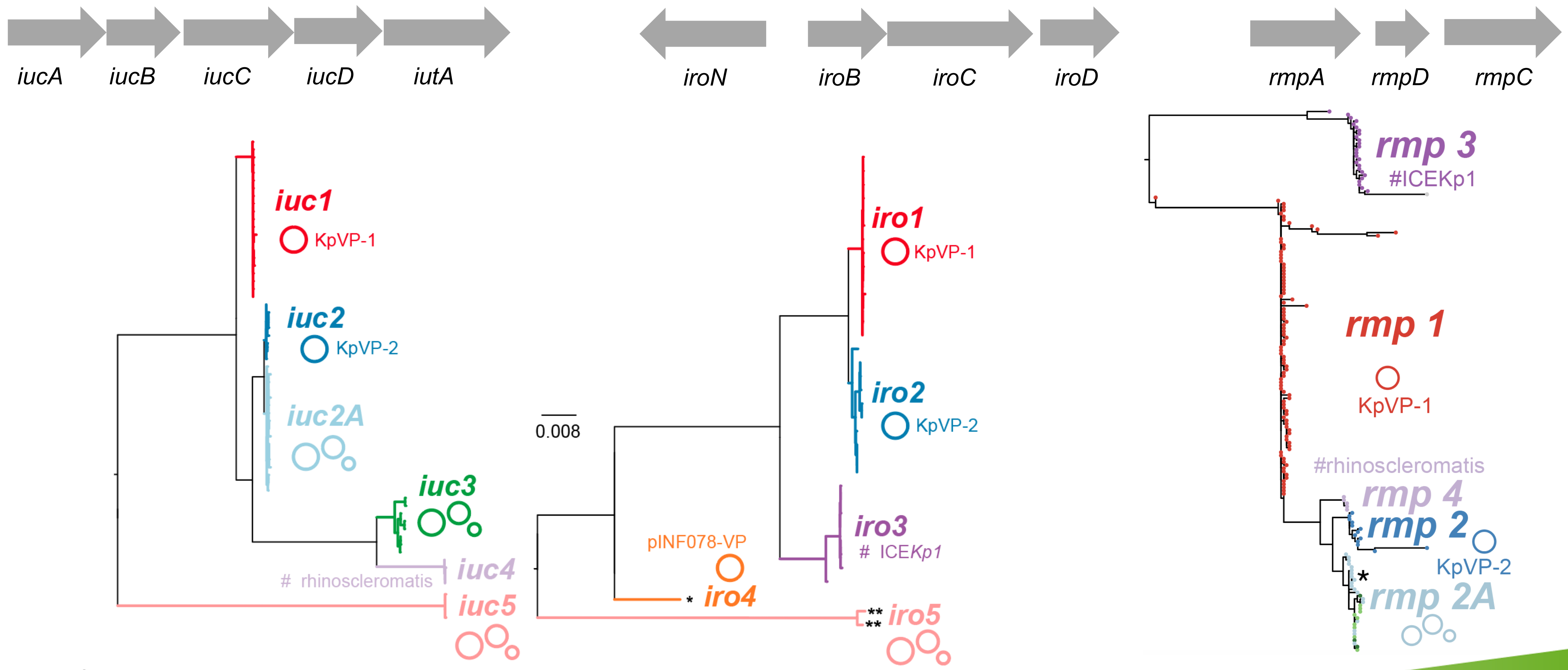
Unpublished

Lineage vs. mobile element associations in other virulence loci - colibactin (*clb*)



Lam et al, Microbial Genomics 2018

Lineage vs. mobile element associations in other virulence loci - aerobactin, salmochelin, *rmp*

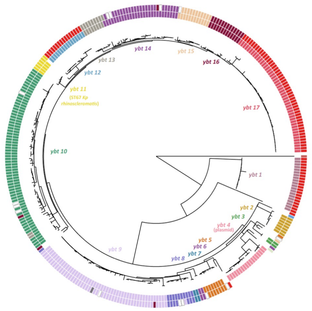


Lam et al, Genome Medicine 2018
Lam et al, Genome Medicine 2025

How can we harness this information for genome surveillance?

Multi-locus sequence typing approach to characterise genetic diversity

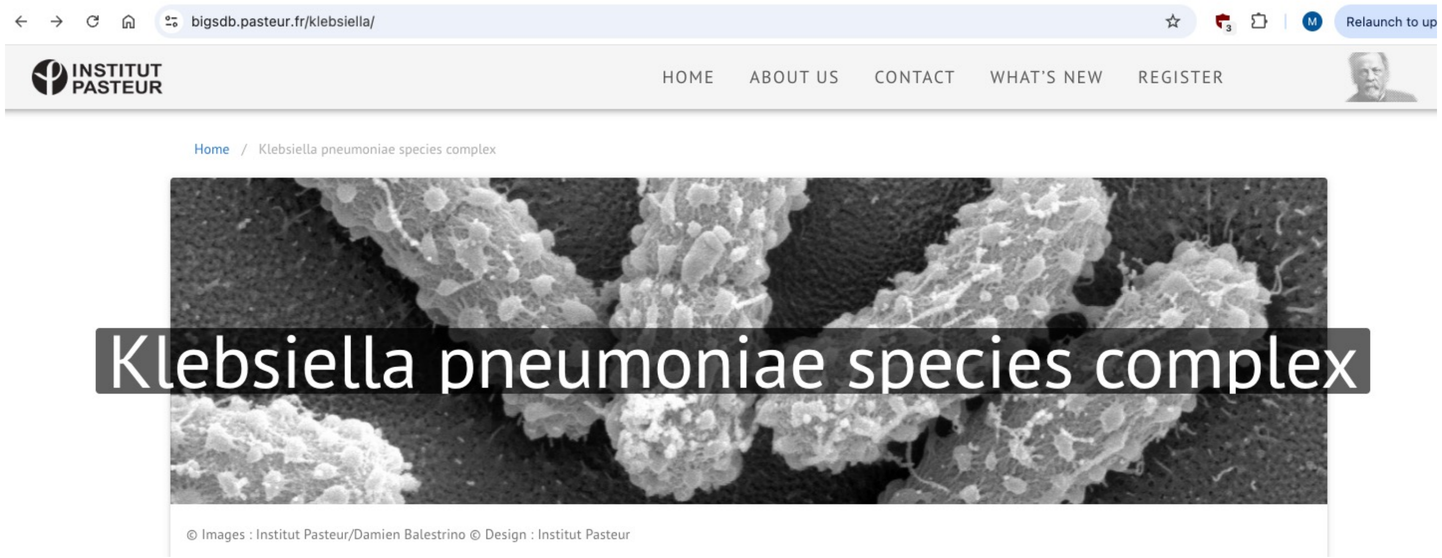
	ybtS	ybtX	ybtQ	ybtP	ybtA	irp2	irp1	ybtU	ybtT	ybtE	fyuA	YbST
Genome 1	1	1	1	1	1	1	1	1	1	1	1	1
Genome 2	1	2	1	1	1	1	1	1	1	1	2	2
Genome 3	1	1	1	2	1	3	1	1	2	8	2	3
Genome 4	1	3	1	1	4	1	1	2	1	1	1	4



YbST = 1, ybt lineage 1, ICEKp10

Virulence typing schemes

AbST – aerobactin
SmST – salmochelin
YbST – yersiniabactin
CbST – colibactin
RmST – *rmp*



Schemes hosted on BIGSdb
Implemented in Kleborate

KLEBORATE

Summary – virulence loci diversity

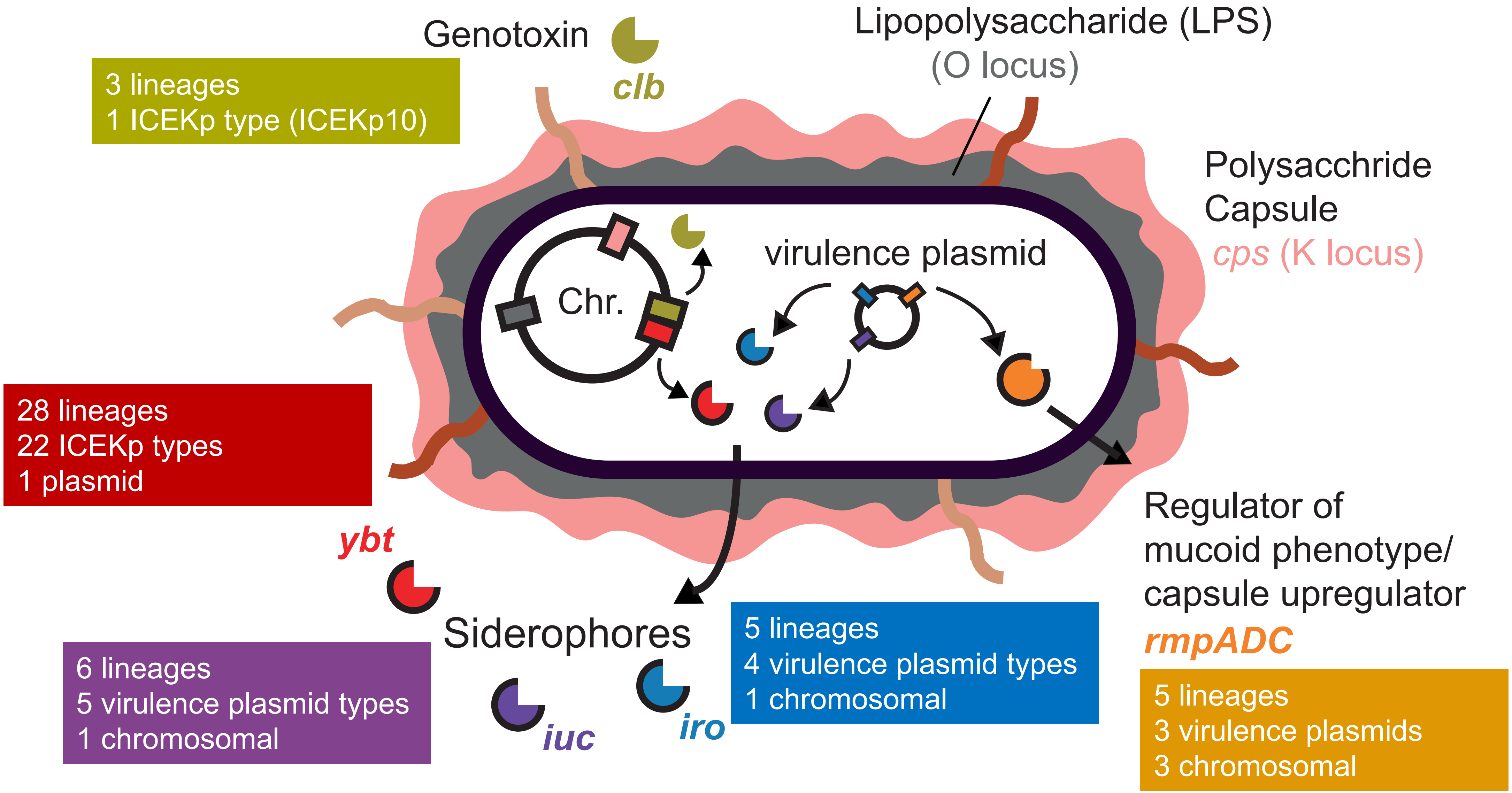


Image credit: A/Prof Kelly Wyres

Kleborate – genotyping and surveillance framework

Input:
Genome (fasta/fna)

KLEBORATE

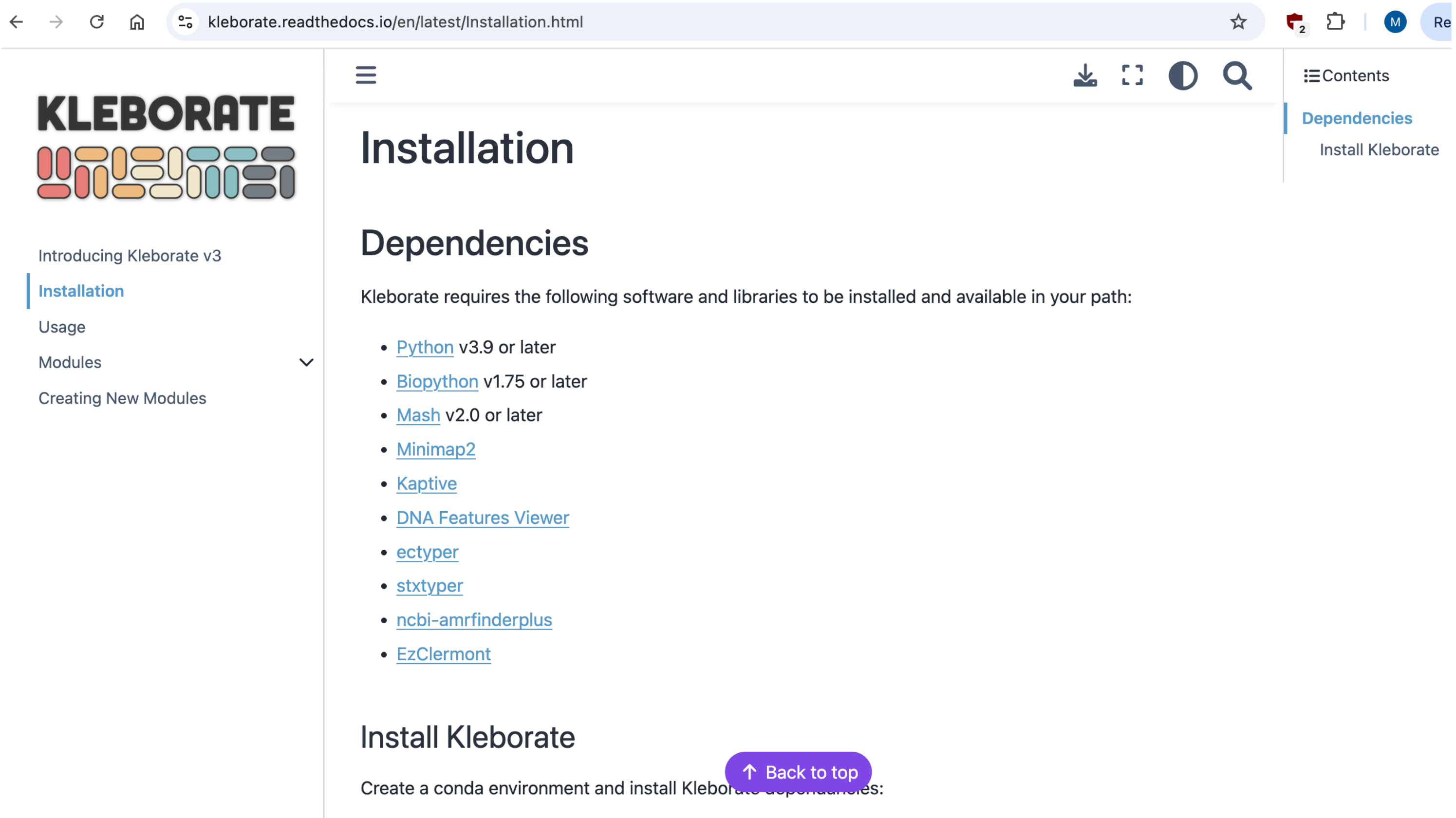
- Assembly QC/stats
- Species prediction
- MLST
- AMR
- Virulence
- Kaptive - In silico serotyping (K/O antigens)

- Initially developed for *Kpn* and related species (KpSC)
- Current version (v3) calls on modules; additional modules for K. oxytoca species complex (KoSC) and *E. coli*

github.com/klebgenomics/Kleborate

Kleborate – run via command line

Kleborate can be installed and used on command line



github.com/klebgenomics/Kleborate

Kleborate – run via command line

Kleborate initially developed for primarily for genotyping K. pneumoniae/KpSC genomes

Latest Kleborate version (v3) has a modular build:

- Genotyping features separated into individual modules
- Depending on species, distinct modules will run
    - KpSC, KoSC and *E. coli*
- Can develop modules specific to other species

github.com/klebgenomics/Kleborate

Kleborate – run via command line

```
kleborate -a *.fasta.gz -o kleborate_results -p kpsc --trim_headers
```

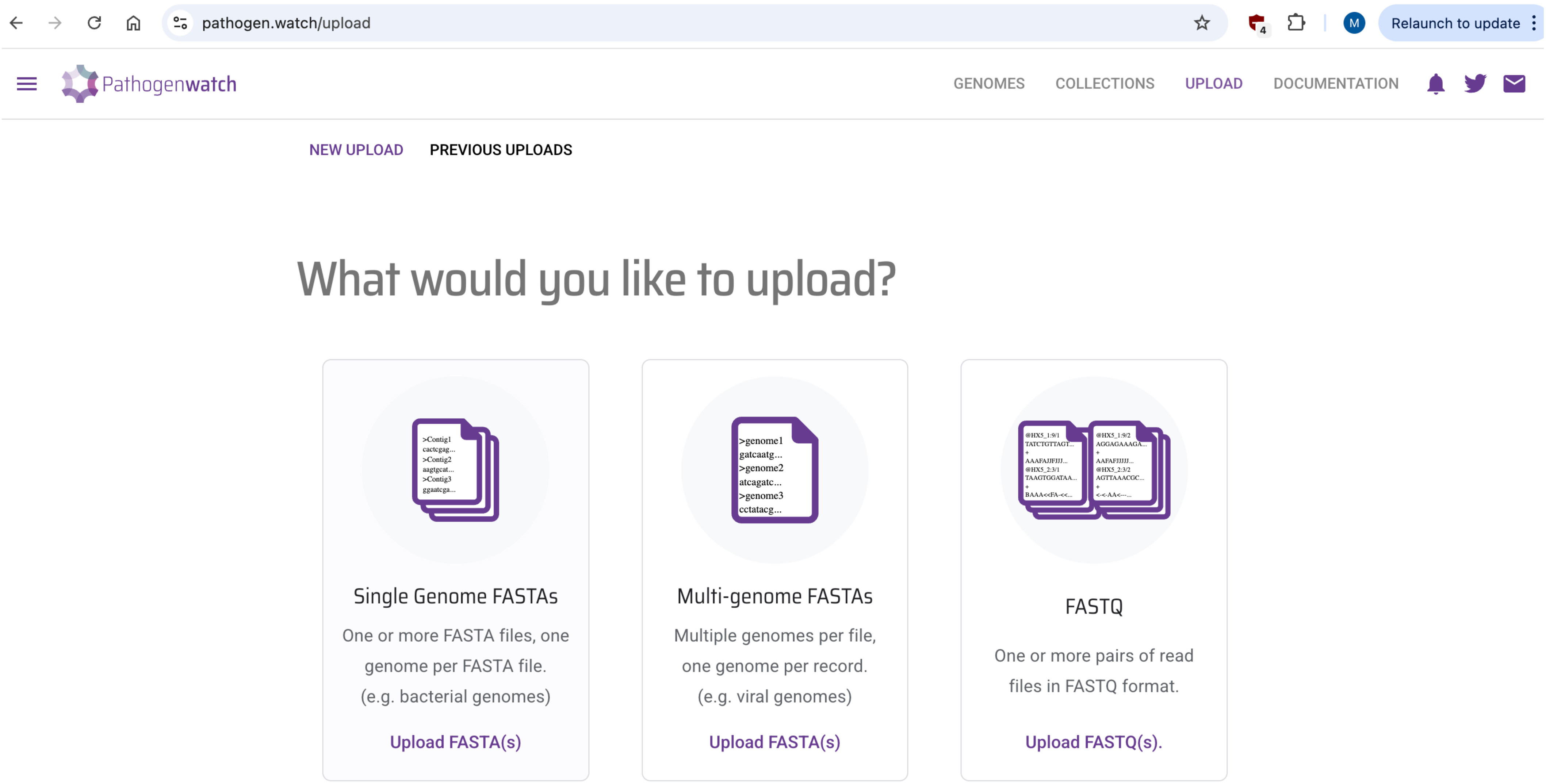
- `-a *.fasta.gz`: Specifies the input files (assemblies) to be analyzed (.fasta or .fasta.gz).
- `-o`: Specifies the directory where the output files will be saved (one output file per species/complex detected).
- `-p`: Specifies the preset modules to run (kpsc, kosc, escherichia).
- `--trim_headers`: Trim module names from column headers in the output.

Run with specified modules only, e.g. AMR typing for *K. pneumoniae* species complex:

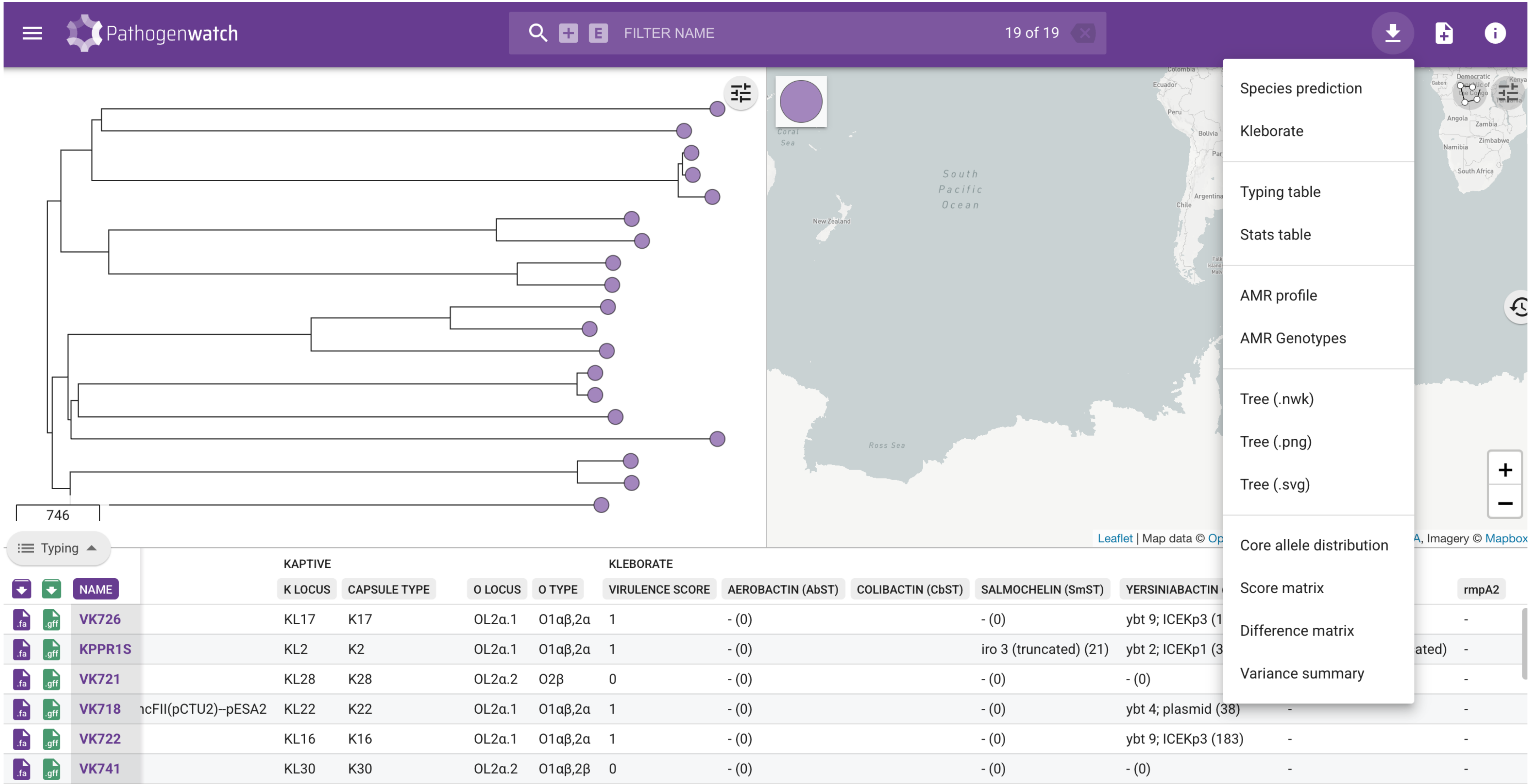
```
kleborate -a *.fasta -o kleborate_results -m klebsiella_pneumo_complex__amr
```

https://kleborate.readthedocs.io/en/latest/Usage.html

Kleborate – run via Pathogenwatch



Kleborate – run via Pathogenwatch



Kleborate – example output

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
Genome Nan	species	species_mat	contig_count	N50	largest_conti	total_size	ambiguous_b	QC_warnings	ST	gapA	infB	mdh	pgi	phoE	rpoB	tonB	YbST	Yersiniabact	ybtS	ybtX
VK741	Klebsiella pneum	strong	133	351082	782217	5528495	no	-	ST29	2	3	2	2	6	4	4	0	-	-	-
VK730	Klebsiella pneum	strong	167	477312	1018715	5596547	no	-	ST45	2	1	1	6	7	1	12	78	ybt 10; ICEKp	3	1
VK735	Klebsiella pneum	strong	4	5356508	5356508	5661247	no	-	ST17	2	1	1	1	4	4	4	232	ybt 15; ICEKp	6	1
VK721	Klebsiella pneum	strong	4	5267448	5267448	5404716	no	-	ST35	2	1	2	1	10	1	19	0	-	-	-
VK733	Klebsiella pneum	strong	3	5315064	5315064	5592223	no	-	ST37	2	9	2	1	13	1	16	34	ybt 4; plasmi	18	2
VK726	Klebsiella pneum	strong	89	370494	735606	5462920	no	-	ST101	2	6	1	5	4	1	6	172	ybt 9; ICEKp3	5	1
VK731	Klebsiella pneum	strong	7	5333974	5333974	5904790	no	-	ST37	2	9	2	1	13	1	16	0	-	-	-
VK743	Klebsiella pneum	strong	71	206413	433333	5187485	no	-	ST147	3	4	6	1	7	4	38	157-2LV	ybt 9; ICEKp3	41	1
VK722	Klebsiella pneum	strong	3	5376827	5376827	5594072	no	-	ST35	2	1	2	1	10	1	19	183	ybt 9; ICEKp3	5	1
VK728	Klebsiella pneum	strong	5	5222404	5222404	5319043	no	-	ST37	2	9	2	1	13	1	16	0	-	-	-
VK779	Klebsiella pneum	strong	41	366051	789681	5401970	yes (195)	ambiguous_b	ST11	3	3	1	1	1	1	4	0	-	-	-
VK719	Klebsiella pneum	strong	3	5270163	5270163	5465612	no	-	ST36	2	1	2	1	7	1	7	214	ybt 9; ICEKp3	5	2
VK723	Klebsiella pneum	strong	96	318387	542148	5557484	no	-	ST36	2	1	2	1	7	1	7	0	-	-	-
VK739	Klebsiella pneum	strong	81	224673	513180	5281750	no	-	ST307	4	1	2	52	1	1	7	0	-	-	-
VK740	Klebsiella pneum	strong	81	345695	755948	5422984	no	-	ST29	2	3	2	2	6	4	4	503	ybt 9; ICEKp3	140	1
VK720	Klebsiella pneum	strong	2	5181281	5181281	5414724	no	-	ST17	2	1	1	1	4	4	4	6	ybt 5; ICEKp6	10	
VK725	Klebsiella pneum	strong	5	5284550	5284550	5465642	no	-	ST45	2	1	1	6	7	1	12	78	ybt 10; ICEKp	3	1
VK718	Klebsiella pneum	strong	84	380351	1052651	5548094	no	-	ST35	2	1	2	1	10	1	19	38	ybt 4; plasmi	18	8
KPPR1S	Klebsiella pneum	strong	44	300807	1053947	5466797	yes (4171)	ambiguous_b	ST493	2	1	70	1	12	1	127	321	ybt 2; ICEKp1	8	5

Output: tab delimited format

Kleborate – output interpretation: virulence score

0	negative for all of yersiniabactin (ybt), colibactin (clb), aerobactin (iuc)
1	yersiniabactin only
2	yersiniabactin and colibactin (or colibactin only)
3	aerobactin (without yersiniabactin or colibactin)
4	aerobactin with yersiniabactin (without colibactin)
5	yersiniabactin, colibactin and aerobactin

- Virulence score corresponds to presence of virulence mobile elements – *ybt* (ICEKp), *clb* (ICEKp10) and *iuc* (virulence plasmid)
- *iro* and *rmp* not included in virulence score

NOT a prediction of hypervirulence

Kleborate – output interpretation: resistance score

0	no ESBL, no carbapenemase (regardless of colistin resistance)
1	ESBL, no carbapenemase (regardless of colistin resistance)
2	Carbapenemase without colistin resistance (regardless of ESBL genes or OmpK mutations)
3	Carbapenemase with colistin resistance (regardless of ESBL genes or OmpK mutations)

Kleborate – output interpretation (virulence)

Species	Klebsiella pneumoniae	Klebsiella pneumoniae	Klebsiella pneumoniae	Klebsiella variicola subsp. variicola
MLST	ST113	ST23	ST2118	ST6768
YbST	65	46	-	183-2LV
Yersiniabactin	ybt 4; plasmid	ybt 1; ICEKp10	-	ybt 9; ICEKp3 (truncated)
CbST	-	29	-	-
Colibactin	-	clb 2	-	-
AbST	-	1	1-3LV	-
Aerobactin	-	iuc 1	iuc unknown	-
iucA	-	1	1*	-
iucB	-	1	1*	-
iucC	-	1	1	-
iucD	-	1	3	-
iutA	-	1	1	-
SmST	-	2	-	-
Salmochelin	-	iro 1	-	-
RmST	-	27	-	-
RmpADC	-	rmp 1; KpVP-1 (truncated)	-	-
rmpA	-	4-47%	-	-
rmpD	-	2	-	-
rmpC	-	2	-	-
virulence score	1	5	3	1
spurious_virulence_hits	iroN_19*0%;iroN_20*0%;iroN_23*0%;iroN_24*0%;iroN_25*0%;iroN_29*0%	-	-	-
rmpA2	-	rmpA2_6*-60%	-	-

Kleborate – output interpretation (AMR)

AGly_acquired	aac(3)-IId^;rmtC	strB.v1*;strA.v1^;aadA5	aadA^	aac(3)IId^;aac(3)IIa.v1^;strB.v1;strA.v1^; aadA2;aac(6')-Ib-cr.v2
Col_acquired	-	-	-	-
Fcyn_acquired	-	-	-	-
Flq_acquired	QnrB76^	-	-	qnrB1.v2^;aac(6')-Ib-cr.v2
Gly_acquired	-	-	-	-
MLS_acquired	Mrx;mphA	Mrx;mphA	-	Mrx;mphA
Phe_acquired	-	-	-	catB3.v2
Rif_acquired	-	-	-	-
Sul_acquired	sul1	sul1;sul2	sul1*	sul1;sul1;sul1;sul2
Tet_acquired	-	tet(A).v1	-	tet(A).v1
Tgc_acquired	-	-	-	-
Tmt_acquired	dfrA14.v2*	dfrA17	dfrA15.v2	dfrA35
Bla_acquired	-	-	-	OXA-1
Bla_inhR_acquired	-	-	-	-
Bla_ESBL_acquired	CTX-M-15	CTX-M-55	-	CTX-M-15;CTX-M-15
Bla_ESBL_inhR_acquired	-	-	-	-
Bla_Carb_acquired	NDM-1	-	-	IMP-4

Kleborate – output interpretation (AMR)

AGly_acquired	aac(3)-IId^;rmtC	strB.v1*;strA.v1^;aadA5	aadA^	aac(3)IId^;aac(3)IIa.v1^;strB.v1;strA.v1^;aadA2;aac(6')-Ib-cr.v2
Col_acquired	-	-	-	-
Fcyn_acquired	-	-	-	-
Flq_acquired	QnrB76^	-	-	qnrB1.v2^;aac(6')-Ib-cr.v2
Gly_acquired	-	-	-	-
MLS_acquired	Mrx;mphA	Mrx;mphA	-	Mrx;mphA
Phe_acquired	-	-	-	catB3.v2
Rif_acquired	-	-	-	-
Sul_acquired	sul1	sul1;sul2	sul1*	sul1;sul1;sul1;sul2
Tet_acquired	-	tet(A).v1	-	tet(A).v1
Tgc_acquired	-	-	-	-
Tmt_acquired	dfrA14.v2*	dfrA17	dfrA15.v2	dfrA35
Bla_acquired	-	-	-	OXA-1
Bla_inhR_acquired	-	-	-	-
Bla_ESBL_acquired	CTX-M-15	CTX-M-55	-	CTX-M-15;CTX-M-15
Bla_ESBL_inhR_acquired	-	-	-	-
Bla_Carb_acquired	NDM-1	-	-	IMP-4
Bla_chr	SHV-11	SHV-1*	SHV-11^	LEN-11
SHV_mutations	SHV:p.Leu35Gln	-	SHV:p.Leu35Gln	SHV:p.Leu35Gln
Omp_mutations	-	-	-	-
Col_mutations	-	-	-	-
Flq_mutations	GyrA:p.Ser83Phe;GyrA:p.Asp87Ala;ParC:p.Ser80Ile	-	GyrA:p.Asp87Asn	GyrA:p.Ser83Ile;ParC:p.Ser80Ile
truncated_resistance_hits	-	-	-	aac(6')-Ib-cr4*-11%
spurious_resistance_hits	DHA-1?-70%	-	-	dfrA14.v1?-0%;dfrA14.v2*?-48%
resistance_score	2	1	0	2
num_resistance_classes	7	7	4	11
num_resistance_genes	9	11	3	23

Kleborate – output interpretation (AMR) ciprofloxacin resistance phenotype prediction

Ongoing Kleborate developments include leveraging large datasets of genomes (>10k) with matched AST data for phenotype predictions as part of KlebNET AMR Geno-Pheno Consortium

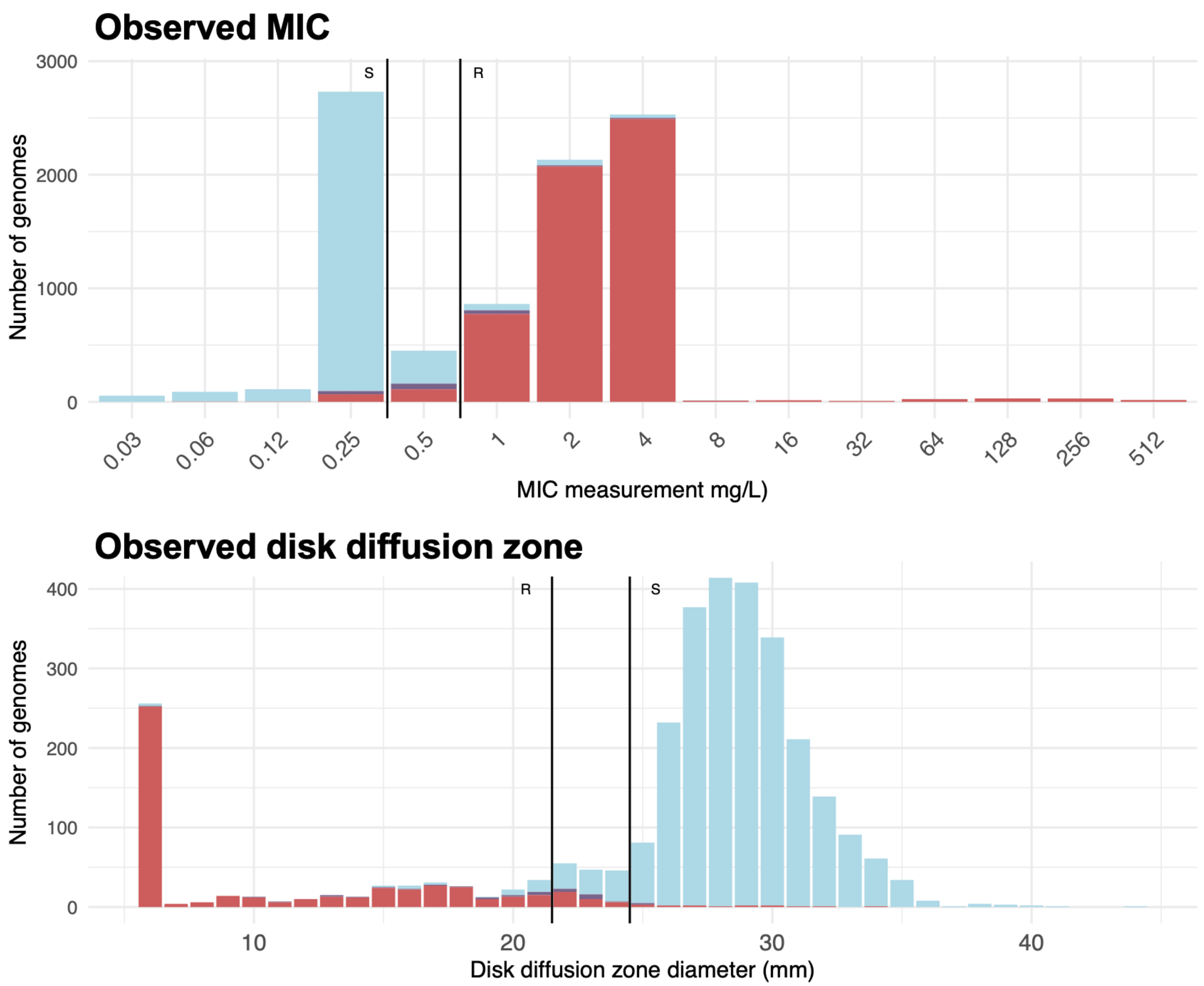
klebnet.org

Ciprofloxacin:

Genomes assigned to one of ten genotypes derived from:

- mutations in QRDR of GyrA and ParC
- acquired/plasmid-mediated quinolone resistance (PMQR) genes
- Presence of aac(6')-Ib-cr

Ciprofloxacin resistance prediction



Predicted phenotype

- S
- I
- R

(available in Kleborate v3.2.2+)

Tsang et al, in preparation (to be preprinted in BioRxiv, September 2025).

Kleborate – output interpretation (AMR) ciprofloxacin resistance phenotype prediction

Each genotype is associated with a resistance phenotype prediction (based on analysis of ~13k genomes; manuscript in prep)

klebnet.org

Genotype profile	Phenotype	Positive predictive value	MIC (mg/L), median [IQR]
0^ QRDR, 0 PMQR, 0 aac(6`)-Ib-cr	wildtype S	90.99% S (N=5168/5680)	0.25 mg/L [0.25-0.25]
0 QRDR, 0 PMQR, 1 aac(6`)-Ib-cr	wildtype S	65.22% S (N=105/161)	0.25 mg/L [0.25-0.5]
0 QRDR, qnrB1, 0 aac(6`)-Ib-cr	nonwildtype I	81.25% I/R (n=130/160)	0.5 mg/L [0.5-1]
1 QRDR, 0 PMQR, 0 aac(6`)-Ib-cr	nonwildtype R	77.67% R (N=80/103)	1 mg/L [1-2]
1 QRDR, 0 PMQR, 1 aac(6`)-Ib-cr	nonwildtype R	86.96% R (N=20/23)	2 mg/L [1-2]
>1 QRDR, 0 PMQR, * aac(6`)-Ib-cr	nonwildtype R	99.22% R (N=2150/2167)	2 mg/L [2-4]
0 QRDR, 1^ PMQR, 0 aac(6`)-Ib-cr	nonwildtype R	77.47% R (N=423/546)	1 mg/L [1-2]
0 QRDR, 1 PMQR, 1 aac(6`)-Ib-cr	nonwildtype R	94.63% R (N=775/819)	2 mg/L [1-2]
0 QRDR, >1 PMQR, * aac(6`)-Ib-cr	nonwildtype R	97.06% R (N=66/68)	2 mg/L [2-4]
>0 QRDR, >0 PMQR, * aac(6`)-Ib-cr	nonwildtype R	99.22% R (N=2421/2440)	4 mg/L [4-4]

Kleborate – output interpretation (AMR) ciprofloxacin resistance phenotype prediction

Flq_acquired	QnrB76^	-	-	qnrB1.v2^;aac(6')-Ib-cr.v2
Flq_mutations	GyrA:p.Ser83Phe;GyrA:p.Asp87Ala;ParC:p.Ser80Ile	-	GyrA:p.Asp87Asn	GyrA:p.Ser83Ile;ParC:p.Ser80Ile
Ciprofloxacin_prediction	nonwildtype R	wildtype S	nonwildtype R	nonwildtype R
Ciprofloxacin_profile_support	99.22% R (N=2424/2443)	90.99% S (N=5168/5680)	77.67% R (N=80/103)	99.22% R (N=2424/2443)
Ciprofloxacin_profile	>0 QRDR, >0 PMQR, * aac(6`)-Ib-cr	0^ QRDR, 0 PMQR, 0 aac(6`)-Ib-cr	1 QRDR, 0 PMQR, 0 aac(6`)-Ib-cr	>0 QRDR, >0 PMQR, * aac(6`)-Ib-cr
Ciprofloxacin_MIC_prediction	4 mg/L [4-4]	0.25 mg/L [0.25-0.25]	1 mg/L [1-2]	4 mg/L [4-4]

Additional notes

- Kleborate output can be used to identify **potential** hypervirulent strains:
Look for rmp AND iuc AND iro

- But what if rmp is truncated?
- Mutations common within poly-G tract due to in-dels
- Ongoing work by Wanford research group suggests these mutations are reversible (phase-variation)

Klebsiella pneumoniae
ST23
46
ybt 1; ICEKp10
29
clb 2
1
iuc 1
1
1
1
1
1
2
iro 1
27
rmp 1; KpVP-1 (truncated)
4-47%
2
2
5
-
rmpA2_6*-60%

Additional notes

While virulence loci/lineages are reported for non-KpSC species (e.g. KoSC, K. aerogenes), the **databases are KpSC specific**

species	species_matc	ST	virulence_scc	resistance_sc	num_resistar	num_resistar	Yersiniabactin	YbST	Colibactin	CbST	Aerobactin	AbST	Salmochelin	SmST
Klebsiella aerogenes	strong	NA	2	0	0	0	ybt 20; ICEKp10	220-1LV	clb 3	64-1LV	-	0	iro unknown	0
Klebsiella aerogenes	strong	NA	2	0	0	0	ybt unknown	0	clb 3	13-2LV	-	0	iro unknown	0
Klebsiella aerogenes	strong	NA	2	0	0	0	ybt unknown	0	clb unknown	0	-	0	iro unknown	0
Klebsiella africana	strong	ST4938	1	0	0	0	ybt unknown	0	-	0	-	0	-	0
Klebsiella africana	strong	ST4938	1	0	0	0	ybt unknown	0	-	0	-	0	-	0

Future plans: develop virulence typing databases for non-KpSC species, non-KpSC modules

In summary

List of learning points in this session:

- Key virulence loci include yersiniabactin and colibactin (largely mobilised by chromosomal ICEKp), aerobactin, salmochelin and *rmp* (largely mobilised by virulence plasmids)
- Genetic diversity linked to mobile elements; information captured in virulence typing schemes
- Kleborate is a genotyping tool for *K. pneumoniae* and related species
    - Interrogation of virulence, AMR, ST, K/O loci, assembly statistics from input genome data
    - Can be used to identify *POTENTIAL* hypervirulent strains
    - Ongoing/future developments include (i) modules for *K. oxytoca* species complex and *E. coli* (other species?), (ii) incorporation of genotype-phenotype predictions

Further reading

Lam et al. A genomic surveillance framework and genotyping tool for *Klebsiella pneumoniae* and its related species complex. Nature Communications 2021 (https://www.nature.com/articles/s41467-021-24448-3)

KlebNET-GSP Kleborate Documentation and Tutorial (kleborate.readthedocs.io/en/latest)

References

- Yersiniabactin and colibactin virulence typing: Lam et al. Microbial Genomics 2018. PMID: 29985125
- Aerobactin and salmochelin virulence typing: Lam et al. Genome Medicine 2018. PMID: 30371343
- *Rmp* virulence typing: Lam et al. Genome Medicine 2025. PMID: 40205597

Acknowledgements

The creation of this training material was commissioned by ECDC to Dr Margaret Lam at Monash University, Australia