

Advanced Molecular Detection Southeast Region Bioinformatics

Outline



Updates



Agenda



Kleborate



Legsta



Questions

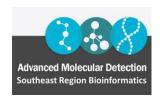
Updates – ABiL Trainings

ABiL in-person courses

- The ABiL in-person (advanced!) bioinformatics course will be held Tuesday Oct. 24th Friday Oct.
 27th at the Georgia Institute of Technology (GT)
- About 2 participants may attend from each state or jurisdiction
- Tanya has sent an email about the courses
- Email us ASAP if you're interested to guarantee a spot!

ABiL online courses

- If additional attendees decide to sign up later, that is not a problem, as the courses will still be available, and they can attend on a rolling basis under the contract
- If you did not receive the invitation for joining the courses, please let us know, we can help to get you enrolled



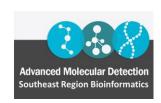
Agenda

October 2 – Bactopia Tools: LisSero and Meningotype

October 16 – Bactopia Tools: Mcroni and MOB_suite

Future Trainings

- ONT & FL's Flisochar pipeline
- StaPH-B Toolkit Programs/Pipelines
- GISAID flagged SARS-CoV-2
- R Training Series
- Dryad pipeline
- ...and more

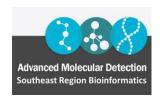


Kleborate

Kleborate is a tool to screen genome assemblies of *Klebsiella pneumoniae* and the *Klebsiella pneumoniae* species complex (KpSC) for:

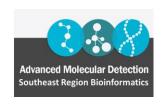
- MLST
- Species (e.g. K. pneumoniae, K. quasipneumoniae, K. variicola, etc.)
- ICEKp associated virulence loci: yersiniabactin (ybt), colibactin (clb), salmochelin (iro), hypermucoidy (rmpA)
- Virulence plasmid associated loci: salmochelin (iro), aerobactin (iuc), hypermucoidy (rmpA, rmpA2)
- Antimicrobial resistance determinants: acquired genes, SNPs, gene truncations and intrinsic β-lactamases
- K (capsule) and O antigen (LPS) serotype prediction, via wzi alleles and Kaptive

klebgenomics/Kleborate (github.com)



Overview

- Klebsiella pneumoniae (Kp) is a commensal bacterium that causes opportunistic infections in hospitals. It has six close relatives (species and subspecies), and together these are known as the K. pneumoniae species complex (KpSC). These related species are often difficult to distinguish from one another in clinical labs using biotyping or MALDI-TOF and consequently can be confused for K. pneumoniae sensu stricto.
- Extracts clinically relevant genotyping information on K. pneumoniae and the KpSC using genome data
- Kleborate a genomic surveillance tool is designed to:
 - Accurately identify species and sequence types
 - Identify the key acquired genetic features for which there is strong evidence of association with either antibiotic resistance or hypervirulence in *K. pneumoniae* sensu stricto
 - Predict K and O types



Installation

Can be installed through conda

conda create –yp /blue/bphl-<state>/<user>/conda_envs/kleborate/ conda activate /blue/bphl-<state>/<user>/conda_envs/kleborate/ conda install –c conda-forge –c bioconda kleborate



Usage

```
thsalikilakshmi@login2;/blue/bphl-florida/thsalikilakshmi/data/jbi/20220825_jax_220629_PLN_WLK_MS/fastqs
                                                                                                               \times
[thsalikilakshmi@login2 fastgs]$ conda activate /blue/bphl-florida/thsalikilakshmi/training/conda envs/kleborate/
(/blue/bphl-florida/thsalikilakshmi/training/conda envs/kleborate) [thsalikilakshmi@login2 fastqs]$ kleborate --hel
usage: kleborate -a ASSEMBLIES [ASSEMBLIES ...] [-r] [--kaptive k] [--kaptive o] [-k] [--all] [-o OUTFILE]
                 [--kaptive k outfile KAPTIVE K OUTFILE] [--kaptive o outfile KAPTIVE O OUTFILE]
                 [--min identity MIN IDENTITY] [--min coverage MIN COVERAGE]
                 [--min spurious identity MIN SPURIOUS IDENTITY] [--min spurious coverage MIN SPURIOUS COVERAGE]
                 [--min kaptive confidence {None, Low, Good, High, Very high, Perfect}] [--force index] [-h]
                 [--version]
Kleborate: a tool for characterising virulence and resistance in Klebsiella
Required arguments:
  -a ASSEMBLIES [ASSEMBLIES ...], --assemblies ASSEMBLIES [ASSEMBLIES ...]
                                       FASTA file(s) for assemblies
Screening options:
  -r, --resistance
                                       Turn on resistance genes screening (default: no resistance gene screening)
                                       Turn on Kaptive screening of K loci (default: do not run Kaptive for K
  --kaptive k
  --kaptive o
                                       Turn on Kaptive screening of O loci (default: do not run Kaptive for O
                                       loci)
  -k, --kaptive
                                       Equivalent to --kaptive k --kaptive o
  --all
                                       Equivalent to --resistance --kaptive
Output options:
  -o OUTFILE, --outfile OUTFILE
                                       File for detailed output (default: Kleborate results.txt)
  --kaptive k outfile KAPTIVE K OUTFILE
                                       File for full Kaptive K locus output (default: do not save Kaptive K locus
                                       results to separate file)
  -- kaptive o outfile KAPTIVE O OUTFILE
                                       File for full Kaptive O locus output (default: do not save Kaptive O locus
                                       results to separate file)
```



Input

 Kleborate takes Klebsiella genome assemblies in FASTA format (can be gzipped). It will work on either draft or completed assemblies, though completed is better because it reduces the risk of fragmented genes/loci

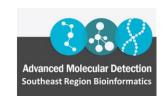
(/blue/bphl-florida/thsalikilakshmi/training/conda_envs/kleborate) [thsalikilakshmi@login2 fastqs]\$ kleborate -o results_res_kaptive.txt --all -a /blue/bphl-

florida/thsalikilakshmi/data/jbi/20220825_jax_220629_PLN_WLK_MS/assemblies/JBI22000619.fasta /blue/bphl-florida/thsalikilakshmi/data/jbi/20220825 jax 220629 PLN WLK MS/assemblies/JBI22000680.fasta



Results

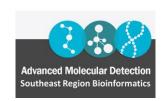
4	А	В	С	D	Е	F	G	Н	-	J	K	L	М	N	0	P	Q	R	S	T	U	V
1 9	strain	species	ST	virulence	resistance	Yersiniaba	YbST	Colibactin	CbST	Aerobacti	AbST	Salmoche	SmST	RmpADC	RmST	rmpA2	wzi	K <u>·</u> locus	K_locus_	O_locus	O_locus_c	AGly_a
2]	BI22000619	Klebsiella pneumoniae	ST34	0	2	-	0	-	0	-	0	-	0	-	0	-	wzi162	unknown	None	01/02v2	Very high	aac(3)-l
3 J	B122000680	Klebsiella pneumoniae	ST258	0	2	-	0	-	0	-	0	-	0	-	0	-	wzi154	unknown	None	01/02v2	Very high	aadA*



Legsta

- In silico Legionella pneumophila Sequence Based Typing (SBT)
- SBT stands for sequence-based typing. The purpose of the *Legionella* pneumophila SBT scheme is to provide a rapid and easily comparable method for the epidemiological typing of clinical and environmental isolates of *Legionella pneumophila* in outbreak investigations

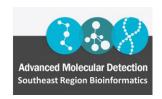
tseemann/legsta: $\ref{thm:legsta}$ In silico Legionella pneumophila Sequence Based Typing (github.com)



Installation

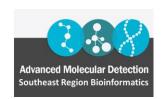
Can be installed through conda

conda create –yp /blue/bphl-<state>/<user>/conda_envs/legsta/ conda activate /blue/bphl-<state>/<user>/conda_envs/legsta/ conda install –c conda-forge –c bioconda legsta



Usage

```
(/blue/bphl-florida/thsalikilakshmi/training/conda envs/legsta) [thsalikilakshmi
@login5 conda envs]$ legsta --help
NAME
  legsta 0.5.1
SYNPOSIS
  Legionella in silico SBT typing of contig sequences
USAGE
  legsta [options] <contigs.fa> ...
OPTIONS
  --help This help.
  --version Print version and exit.
  -- quiet Don't print anything to stderr.
  --debug+ Verbose debug output to stderr (default '0').
  --dbdir=s SBT database folder (default '/blue/bphl-florida/thsalikilakshmi/
training/conda envs/legsta/db').
             Output CSV instead of TSV (default '0').
  --csv
  --noheader Don't print header row (default '0').
HOMEPAGE
  https://github.com/tseemann/legsta
```



Input

Accepts FASTA, Genbank, EMBL, GFF, and many other formats

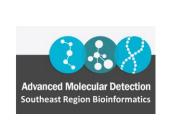
legsta /blue/bphl-

 $florida/thsalikilakshmi/data/legionella/20230719_jax_230531_PLN_WAT_JD/assemblies/JBI23001169.fasta > legsta_out$



Results

- aFILE name of the file
- SBT sequence based typing
- flaA flagellin gene
- pile pilin gene
- asd Aspartate-Beta- Semialdehyde Dehydrogenase gene
- mip macrophage infection enhancer protein
- mompS major outer membrane proteins
- proA major secretory protein, mediates virulence regulation, host tissue degradation and immune evasion
- neuA N-Acylneuraminate Cytidyltransferase Gene



4	Α	В	С	D	E	F	G	Н	
1	aFILE	SBT	flaA	pilE	asd	mip .	mompS	proA	neuA
2	/blue/bpł	1	1	4	3	1	1	1	1



Advanced Molecular Detection
Southeast Region Bioinformatics

Questions?

<u>bphl-sebioinformatics@flhealth.gov</u>

Lakshmi Thsaliki, MS

Bioinformatician

Lakshmi.Thsaliki@flhealth.gov

Molly Mitchell, PhD

Bioinformatician

Molly.Mitchell@flhealth.gov