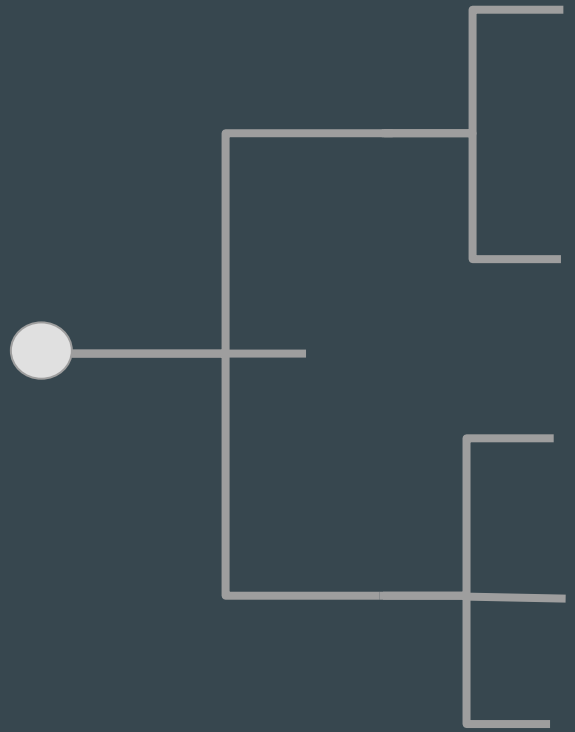
Comparative Genomics Results

Team A Group 4
Members: Anagha, Hina, Lauren, Rohini, Sarth

Agenda

- Overview
- Tools Comparison
- Phylogenetic Tree Results
- Species/Strain identification
- Isolate Relationships
- Virulence and Antibiotic Resistance Profiles
- CDC Recommendations
- Pipeline
- Task Delegation
- References

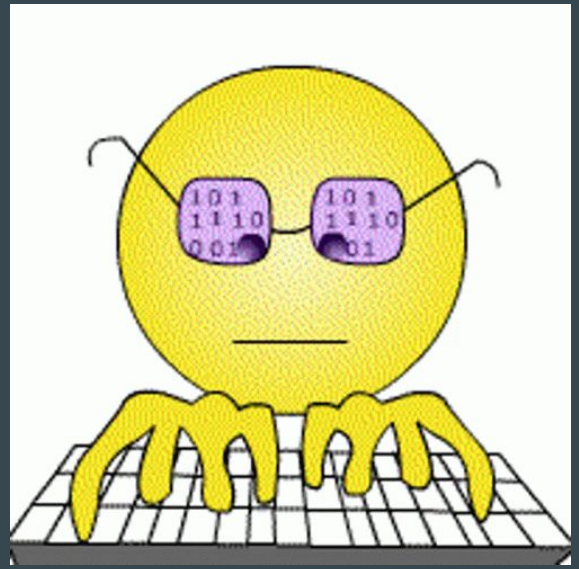


Overview

- Perform comparative studies to infer relatedness among our 28 assigned samples.
- Compare software tools available for genomic comparison and phylogenetics analysis.
- Use the created phylogenetic tree to analyze the samples involved in the outbreak and differentiate them from the outliers.
- Understand the virulence and antibiotic resistance profiles of these pathogens.
- Carry out literature survey on outbreak response, containment strategies and effective treatment options.

Tools evaluated for comparative analysis

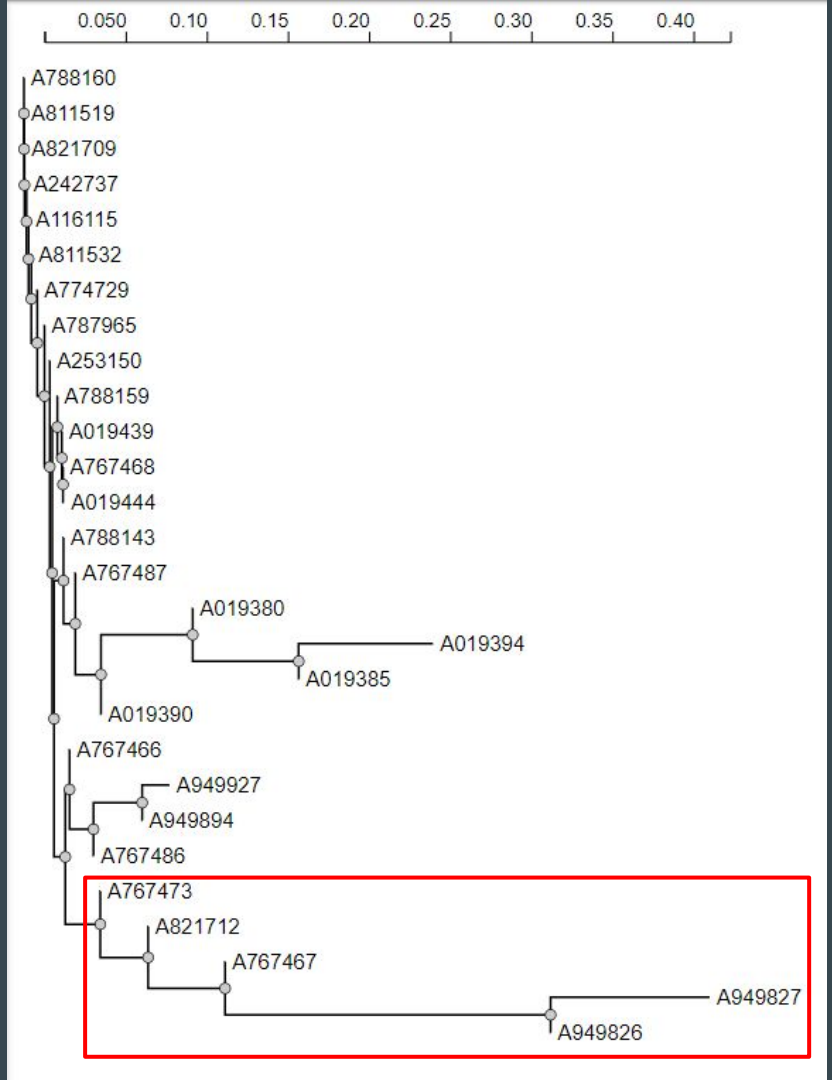
- Lyve-set
- CFSAN
- Snippy
- SNV-Phyl
- Mashtree
- Sourmash
- Skmer
- Co-Phylog
- Parsnp
- OrthoFinder



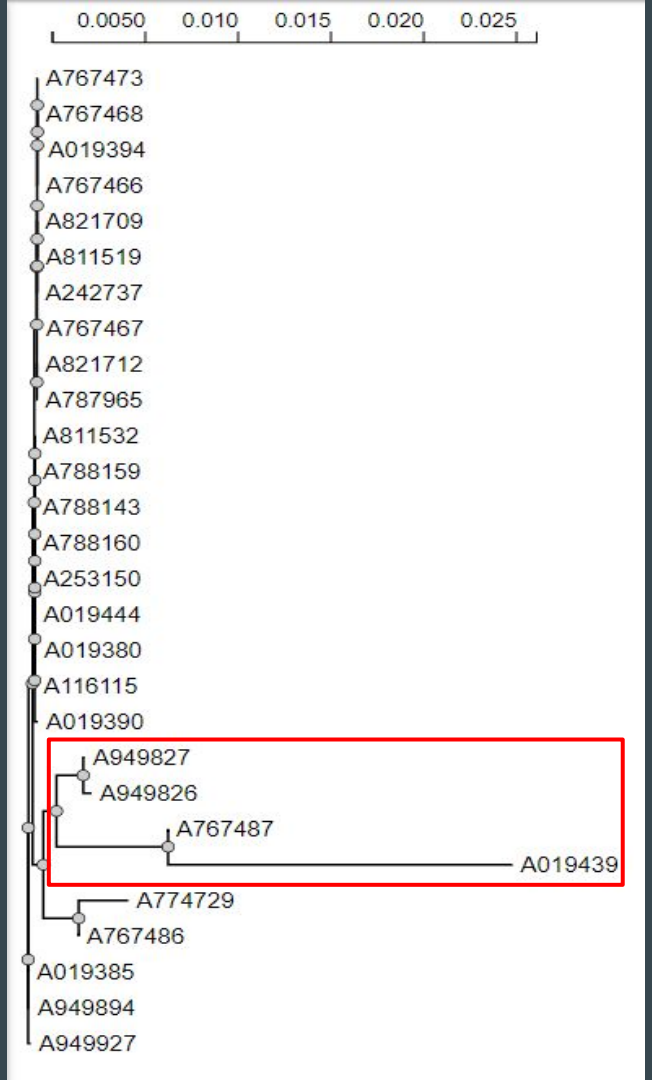
Tool name	Time required	Documentation	Ease of use	Comments
Lyve-SET	-	Comprehensive	Difficult	Too many dependencies, prevented running tool
CFSAN SNP Pipeline	-	Limited	Difficult	Pipeline of scripts/tools
Snippy	~ 1 hour	Comprehensive	Medium	Takes a while to run SNP identification, but tree building is quick
SNV-Phyl 2.0	-	Moderate (for Nextflow version)	Medium	Required too many dependencies
Mashtree	-	Comprehensive	Medium	Not compatible with MacOS
Sourmash	-	Comprehensive	Medium	Not able to visualise
Skmer	~ 20 mins	Comprehensive	Medium	Tree generated was different from the rest of the trees
Parsnp	~ 5 mins	Comprehensive	Easy	Quick to run and the tree was somewhat similar to Snippy.
OrthoFinder	~ 5 mins	Comprehensive	Easy	Very detailed results

Results - Snippy

Results - OrthoFinder



Results - Skmer



Results - Parsnp

What caused the outbreak (species/strain identity)?

Team 2 Results:

16S rRNA extraction and analysis

Team 3 Results:

Cgmlst and bigsdb+

From Team 2 and Team 3 results, we identified that all 28 samples are *Elizabethkingia anophelis*

Lin, J.-N.; Lai, C.-H.; Yang, C.-H.; Huang, Y.-H. Elizabethkingia Infections in Humans: From Genomics to Clinics. Microorganisms 2019, 7, 295. https://doi.org/10.3390/microorganisms7090295

How are the isolates related to each other? How do they differ?

We can a couple of different clusters

There are 2 distinct outliers (**A949826**,**A949827**), these could potentially be

- Other strains of Elizabethkingia like *E. meningoseptica*, *E. miricola*
- Other related species - *Flavobacterium meningosepticum* or *Chryseobacterium meningosepticum*

There could potentially be 3 more outliers which were identified through 4 different tools

What are the Virulence and Antibiotic Resistance Profiles of the outbreak isolates?

ResFinder

Resistance gene	Identity	Alignment length/gene length	Position in reference	Contig or depth	Position in contig	Phenotype	PMID	Accession no.	Notes
blaGOB-3	99.77%	873 /	1...873	contigs_1 OrigDefIn=NODE_1_length_1068317_cov_80.137729	271267...272139	['amoxicillin', 'ampicillin', 'cefoxitin', 'ceftazidime', 'ertapenem', 'imipenem', 'meropenem']	10858348	AF189291	Chromosomal metallo beta-lactamase
blaB-3	100.00%	750 /	1...750	contigs_1 OrigDefIn=NODE_1_length_1068317_cov_80.137729	648013...648762	['amoxicillin', 'ampicillin', 'ticarcillin']	10858348	AF189299	Chromosomal class B carbapenamase
blaCME-1	99.50%	798 /	86...883	contigs_9 OrigDefIn=NODE_9_length_132639_cov_75.980639	82256...83053	['unknown beta-lactam']	10471563	AJ006275	Chromosomal

AMR finder gave us the same results as Abricate except also included the subclass of drugs (CARBAPENEM CEPHALOSPORIN CARBAPENEM) and the method used to scan the assemblies like blastx for finding protein encoding genes.

Abricate

#FILE	SEQUENCE	START	END	STRAND	GENE	COVERAGE	COVERAGE_MAP	GAPS	%COVERAGE	%IDENTITY	DATABASE	ACCESSION	PRODUCT	RESISTANCE
A019439.fna	contigs_12	78901	79773	-	blaGOB-48	1-873/873	==========	0/0	100.00	99.31	ncbi	NG_068012.1	subclass B3 metallo-beta-lactamase GOB-48	CARBAPENEM
A019439.fna	contigs_13	49609	50406	-	blaCME-1	86-883/888	.==========	0/0	89.86	99.50	ncbi	NG_048764.1	class A extended-spectrum beta-lactamase CME-1	BETA-LACTAM
A019439.fna	contigs_3	419933	420682	-	blaB-33	1-750/750	==========	0/0	100.00	98.53	ncbi	NG_067218.1	subclass B1 metallo-beta-lactamase BlaB-33	CARBAPENEM

What are the Virulence and Antibiotic Resistance Profiles of the outbreak isolates?

Resistance gene	Gene Product	Class	Subclass	Antimicrobial Drugs
blaGOB-3	subclass B3 metallo-beta-lactamase GOB-3	BETA-LACTAM	CARBAPENEM	['amoxicillin', 'ampicillin', 'cefoxitin', 'ceftazidime', 'ertapenem', 'imipenem', 'meropenem']
blaB-33	subclass B1 metallo-beta-lactamase BlaB-33	BETA-LACTAM	CARBAPENEM	['amoxicillin', 'ampicillin', 'ticarcillin']
blaCME-1	class A extended-spectrum beta-lactamase CME-1	BETA-LACTAM	CEPHALOSPORIN	['unknown beta-lactam'] ex: cefixime, cefazolin, cephalexin

β-lactam antibiotics (beta-lactam antibiotics) are antibiotics that contain a beta-lactam ring in their chemical structure. This includes penicillin derivatives (penams) and cephalosporins and more.

Carbapenem: the most effective against Gram-positive and Gram-negative bacteria presenting a broad spectrum of antibacterial activity

Cephalosporin: A broad-spectrum antibiotic derived from a specific mold working in a similar way to penicillins. They bind to and block the activity of enzymes responsible for making peptidoglycan, an important component of the bacterial cell wall.

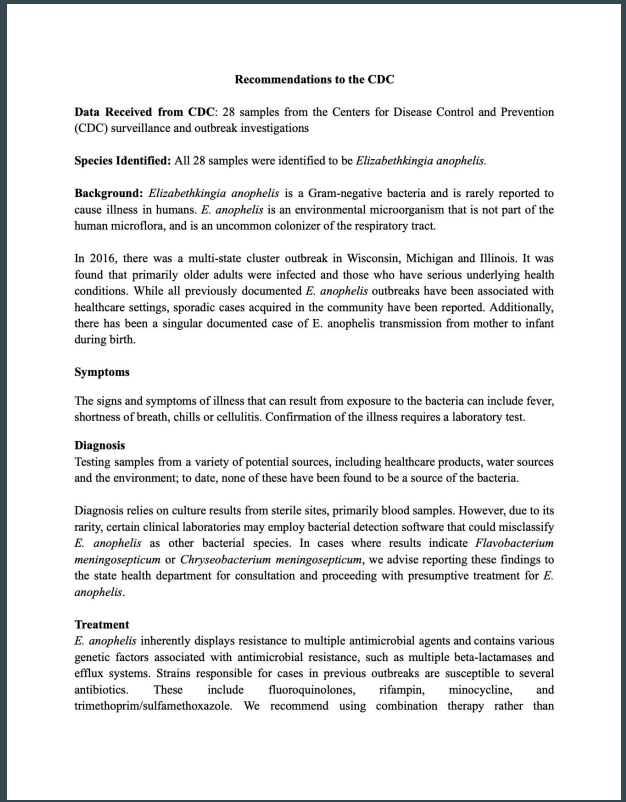
What is the recommended outbreak response and treatment?

Response

- Interviewing patients to gather information about activities and exposures related to healthcare products, food, water, restaurants, and other community settings
- If patients are >65 y.o. and have underlying health conditions, treat as *E. anophelis*

Treatment

- Antibiotics: fluoroquinolones, rifampin, minocycline, and trimethoprim/sulfamethoxazole.
- Combination therapy rather than monotherapy.



Recommendations to the CDC

Data Received from CDC: 28 samples from the Centers for Disease Control and Prevention (CDC) surveillance and outbreak investigations

Species Identified: All 28 samples were identified to be *Elizabethkingia anophelis*.

Background: *Elizabethkingia anophelis* is a Gram-negative bacteria and is rarely reported to cause illness in humans. *E. anophelis* is an environmental microorganism that is not part of the human microflora, and is an uncommon colonizer of the respiratory tract.

In 2016, there was a multi-state cluster outbreak in Wisconsin, Michigan and Illinois. It was found that primarily older adults were infected and those who have serious underlying health conditions. While all previously documented *E. anophelis* outbreaks have been associated with healthcare settings, sporadic cases acquired in the community have been reported. Additionally, there has been a singular documented case of E. anophelis transmission from mother to infant during birth.

Symptoms

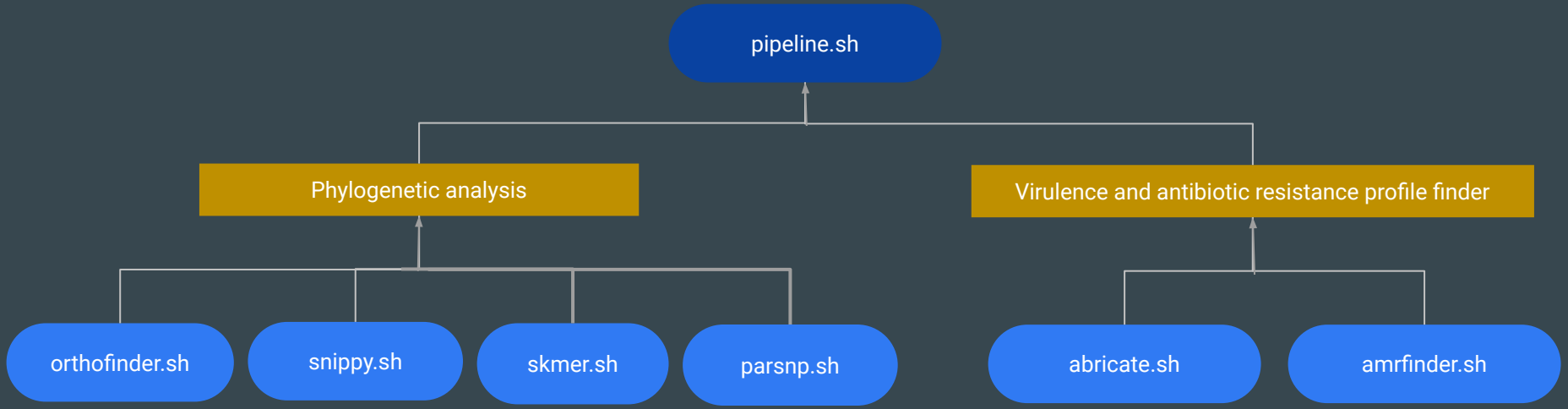
The signs and symptoms of illness that can result from exposure to the bacteria can include fever, shortness of breath, chills or cellulitis. Confirmation of the illness requires a laboratory test.

Diagnosis
Testing samples from a variety of potential sources, including healthcare products, water sources and the environment; to date, none of these have been found to be a source of the bacteria.

Diagnosis relies on culture results from sterile sites, primarily blood samples. However, due to its rarity, certain clinical laboratories may employ bacterial detection software that could misclassify *E. anophelis* as other bacterial species. In cases where results indicate *Flavobacterium meningosepticum* or *Chryseobacterium meningosepticum*, we advise reporting these findings to the state health department for consultation and proceeding with presumptive treatment for *E. anophelis*.

Treatment
E. anophelis inherently displays resistance to multiple antimicrobial agents and contains various genetic factors associated with antimicrobial resistance, such as multiple beta-lactamases and efflux systems. Strains responsible for cases in previous outbreaks are susceptible to several antibiotics. These include fluoroquinolones, rifampin, minocycline, and trimethoprim/sulfamethoxazole. We recommend using combination therapy rather than

Pipeline


pipeline.sh
Phylogenetic analysis
Virulence and antibiotic resistance profile finder
orthofinder.sh
snippy.sh
skmer.sh
parsnp.sh
abricate.sh
amrfinder.sh

```
#!/bin/bash

#run AMRFinderPlus to find antimicrobial resistance and other genes in nucleotide sequences
./amrfinder.sh

#run abricate for mass screening of contigs for antimicrobial resistance or virulence genes
./abricate.sh

#run different tools to generate phylogenetic trees
#run orthofinder
./orthofinder.sh

#run snippy
./snippy.sh

#run parsnp
./parsnp.sh

#run skmer
./skmer.sh
```

References

1. Sarmashghi, S., Bohmann, K., P. Gilbert, M.T. et al. Skmer: assembly-free and alignment-free sample identification using genome skims. Genome Biol 20, 34 (2019). https://doi.org/10.1186/s13059-019-1632-4
2. Yi, H., & Jin, L. (2013). Co-phylog: An assembly-free phylogenomic approach for closely related organisms. Nucleic Acids Research, 41(7), e75. https://doi.org/10.1093/nar/gkt003
3. Hagey J. SNVPhyl Nextflow. 2022-21-4. https://github.com/DHQP/SNVPhyl_Nextflow/
4. Petkau A, Mabon P, Sieffert C, Knox N, Cabral J, Iskander M, Iskander M, Weedmark K, Zaheer R, Katz L, Nadon C, Reimer A, Taboada E, Beiko R, Hsiao W, Brinkman F, Graham M, Van Domselaar G. SNVPhyl: a single nucleotide variant phylogenomics pipeline for microbial genomic epidemiology. 08/06/2017. *M Gen* 3(6): doi:10.1099/mgen.0.000116.
5. Katz LS, Griswold T, Williams-Newkirk AJ, Wagner D, Petkau A, Sieffert C, Van Domselaar G, Deng X, Carleton HA. A Comparative Analysis of the Lyve-SET Phylogenomics Pipeline for Genomic Epidemiology of Foodborne Pathogens. Front Microbiol. 2017 Mar 13;8:375. doi: 10.3389/fmicb.2017.00375. PMID: 28348549; PMCID: PMC5346554. https://github.com/lskatz/lyve-SET/
6. Davis S, Pettengill JB, Luo Y, Payne J, Shpuntoff A, Rand H, Strain E. (2015) CFSAN SNP Pipeline: an automated method for constructing SNP matrices from next-generation sequence data. PeerJ Computer Science 1:e20 https://doi.org/10.7717/peerj-cs.20
7. Emms D.M. & Kelly S. (2019), Genome Biology 20:238, Emms D.M. & Kelly S. (2017), MBE 34(12): 3267-3278, Emms D.M. & Kelly S. (2018), bioRxiv https://doi.org/10.1101/267914

Thank you!