Genomic Tools in Outbreak Investigations - Case Study Cholera Outbreak

Annotation: AMR genes, Virulence factors, and Genetic Elements

August 27, 2025

Intended Learning Objectives

Specific objectives of this session:

1. Predict and interpret AMR and virulence genes from bacterial genomes.
2. Predict genomic islands (GIs) using IslandViewer4 and understand the limits of automated predictions.
3. Detect integrative conjugative elements (ICEs) using ICEfinder and interpret the main result views.
4. Understand that sequence alignment and literature comparison may be needed to identify known GI families in *Vibrio cholerae*.

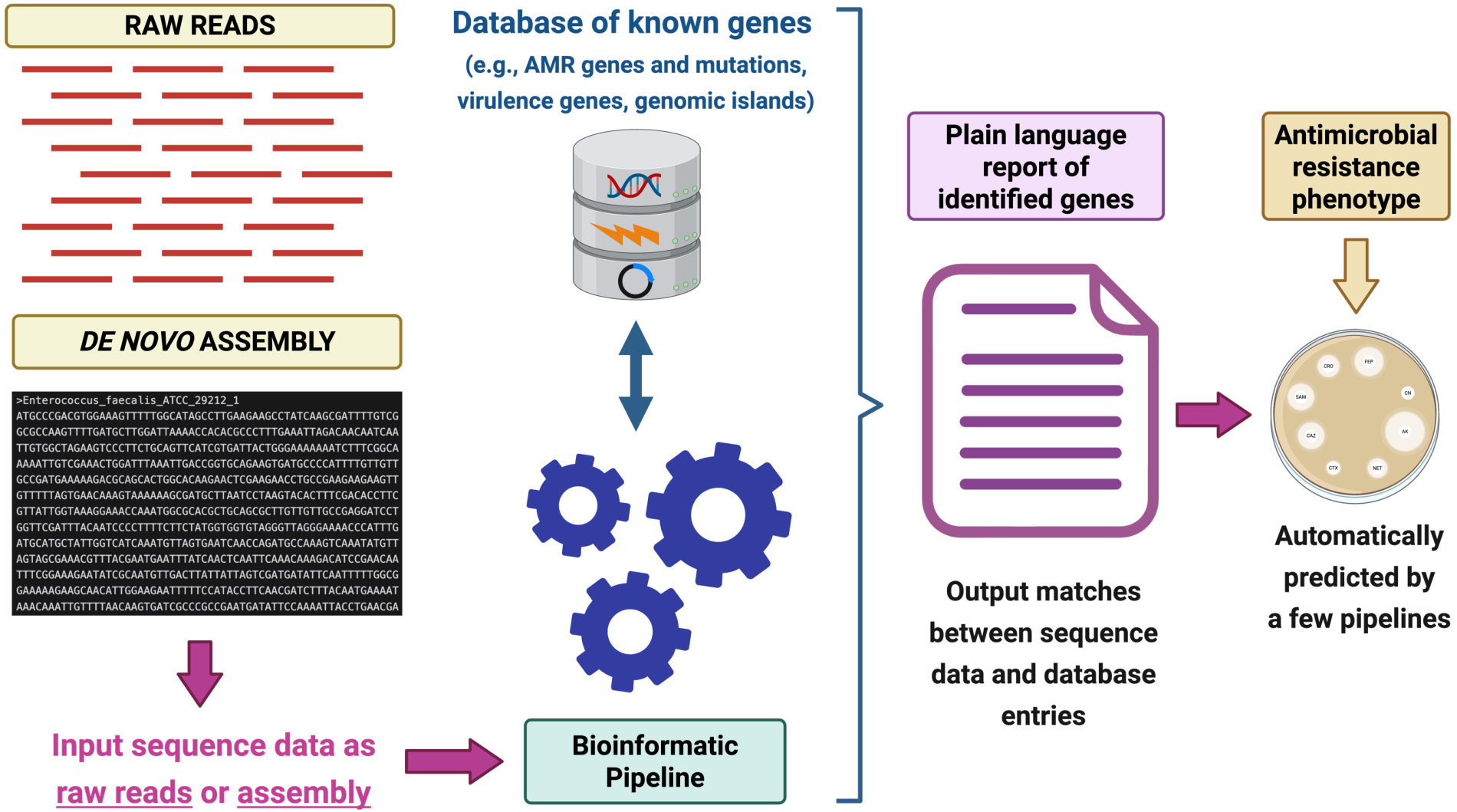
Outline

This session consists of the following elements:

1. Introduction to genome annotation for AMR and virulence genes
2. Predicting ICEs using ICEfinder: input, outputs, and interpretation
3. Predicting genomic islands (GIs) with IslandViewer4: maps, tables, and region details
4. Understanding the limits of automated predictions and the need for literature and sequence alignment to identify known GI families (e.g., in *Vibrio cholerae*)

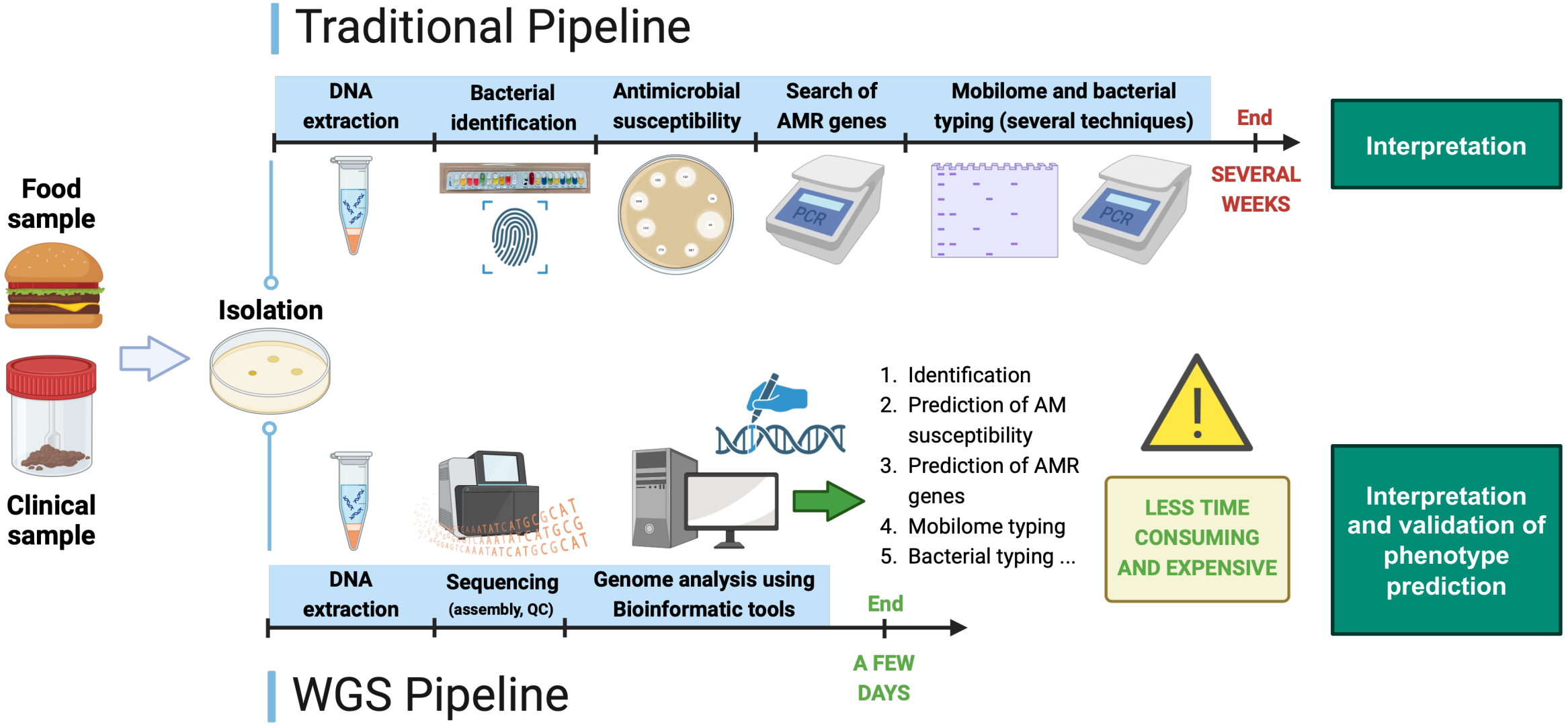
What best describes your experience with bioinformatics or computational tools?

In silico Bacterial Genome Annotation



Hendriksen S. et al. (2019). Front Public Health. 7:242; Image created with BioRender

Why using WGS for Genome Analysis?



Adapted from Ronholm J. et al. (2016). Clin Microbiol Rev. 29(4):837-57; Image created with BioRender

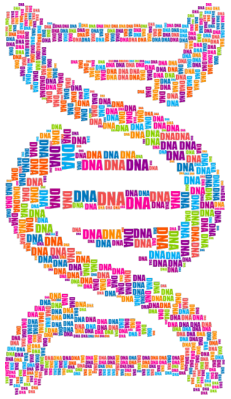
Annotation

Detection of Antimicrobial Resistance (AMR) genes and Virulence factors

AMR Fenotype/Genotype Prediction Tools

Table. Most used (command line-C or server-based-SB) AMR prediction tools for bacterial genomes

Tool name	Target	Input format	Link
ResFinder (C, SB)	**General AMR (genotype, phenotype)**	**FASTA, FASTQ**	**http://genepi.food.dtu.dk/resfinder**
PointFinder (C, SB)	**Selected mutations (genotype)**	**FASTA, FASTQ**	**http://genepi.food.dtu.dk/resfinder**
AMRFinderPlus (C)	**General AMR (genotype)**	**FASTA**	**https://github.com/ncbi/amr/wiki**
RGI (CARD) (C, SB)	General AMR (genotype)	FASTA	https://card.mcmaster.ca/
ABRicate (C)	Multiple databases (AMR, virulence, plasmids)	FASTA	https://github.com/tseemann/abricate
BV-BRC (SB, uses RAST)	General Annotation (genotype, phenotype)	FASTA	https://www.bv-brc.org/
Pathogenwatch (SB)	General AMR (genotype, phenotype – some pathogens)	FASTA, FASTQ	https://pathogen.watch/



Uelze L. et al. (2020). One Health Outlook. 2020; 2:3

Genotype vs Phenotype Concordance

	Pathogen	No. of pathogens	AST method	No. of antimicrobials	Bioinformatic tool	Sequencing data	Concordance	Sensitivity	Specificity	Comment	References
2013	*S. Typhimurium*	49	MIC	17	ResFinder	Assembled, Velvet	99.74%			Disagreement: 7 isolates including 6 *E. coli* resistent to Spec	(7)
	E. coli	48									
	E. faecalis	50		14							
	E. faecium	50									
2013	*E. coli* (ESBL)	74	DD	7	BLASTn, selected panel	Assembled, Velvet		96%	97%	VM rate: 1.2%/M rate: 2.1%	(8)
	K. pneumonia (ESBL)	69									
2014	*S. aureus*	501	DD/MIC (Vitek)	12	BLASTn, selected panel	Assembled, Velvet		97%	99%	VM rate: 0.5%/M rate: 0.7%	(9)
2016	*C. jejuni*	32	MIC	9	BLASTx	Assembled, CLC-bio	99.2%			Lower concordance to	(10)
	C. coli	82								Gen, Azi, Clin, Tel	
2016	*S. enterica*	104	MIC	14	ResFinder/ ARG-ANNOT/ CARD/BLAST	Assembled, CLC-bio	99.0%	99.2%	99.3%	Lower concordance to	(11)
		536						97.6%	98.0%	aminoglycosides/β-lactams	
2017	*E. coli*	31	MIC	4	Custom DB based on ARDB/CARD/β-lactamase allelles			87%	98%	Neg. predictive value: 97%	(12)
	K. pneumonia	24								Pos. Predictive value: 91%	
	P. aeruginosa	22									
	E. cloacae	13									
2017	*S. enterica*	50	MIC	4	ResFinder/ PointFinder	Assembled, SPAdes	98.4%			Disagreement: 2/2 C.jejuni to FQ/ERY	(13)
	E. coli	50		6							
	C. jejuni	50		4						5 *E. coli* to COL (pmrB)	
2018	*E. faecalis*	97	MIC	11	ResFinder/NCBI Pathogen DB/BLAST	Assembled, CLC-bio	96.5%				(14)
	E. faecium	100									
2018	*S. aureus*	501	DD/MIC	12	GeneFinder/ Mykrobe/ Typewriter	FASTQ/assembled, BLAST	98.3%			Disagreements: 0.7% predicted resistant	(15)
		491									
		397	MIC							0.6% predicted susceptible	
2018	*M. tuberculosis*	10,209	MGIT 960	4	Cortex	Assembled	89.5%			97.1%/99.0% predicted R/S	(16)
				4						97.5%/98.8% predicted R/S	
				4						94.6%/93.6% predicted R/S	
				4						91.3%/96.8% predicted R/S	
2019	*H. pylori*	140	MIC (E-test)	5	ARIBA	FASTQ	99%			Phenotype issues to metronidazole	(17)

1) ESBL: Extended Spectrum Beta-Lactamase, 2) MIC: Minimum Inhibitory Concentration, 3) DD: Disk diffusion, 4) VM: Very Major, 5) M: Major, 6) R/S: Resistant/Susceptible, 7) SPEC: Spectinomycin, 8) GEN: Gentamicin, 9) AZI: Azithromycin, 10) CLIN: Clindamycin, 11) TEL: Telithromycin, 12) FQ: Fluoroquinolone, 13) ERY: Erythromycin, 14) COL: colistin.

Several *in silico* AMR prediction tools were used (**ResFinder**, **PointFinder**, ARG-ANNOT, CARD, GeneFinder, Mykrobe, Cortex, ARIBA)

High concordance (>96%) between the presence of known AMR genes or mutations (**genotype**) and the DD and/or MIC data (**phenotype**)

Hendriksen S. et al. (2019). Front Public Health. 7:242; Bortolaia V. et al. (2020). J Antimicrob Chemother. 75(12):3491-3500

Genotype vs Phenotype Concordance

One Day in Denmark: Comparison of Phenotypic and Genotypic Antimicrobial Susceptibility Testing in Bacterial Isolates From Clinical Settings

Ana Rita Rebelo[1], Valeria Bortolaia[1,2], Pimlapas Leekitcharoenphon[1], Dennis Schrøder Hansen[3], Hans Linde Nielsen[4,5], Svend Ellermann-Eriksen[6], Michael Kemp[7], Bent Løwe Røder[8], Niels Frimodt-Møller[9], Turid Snekloth Søndergaard[10], John Eugenio Coia[11], Claus Østergaard[12], Henrik Westh[13,14] and Frank M. Aarestrup[1]*

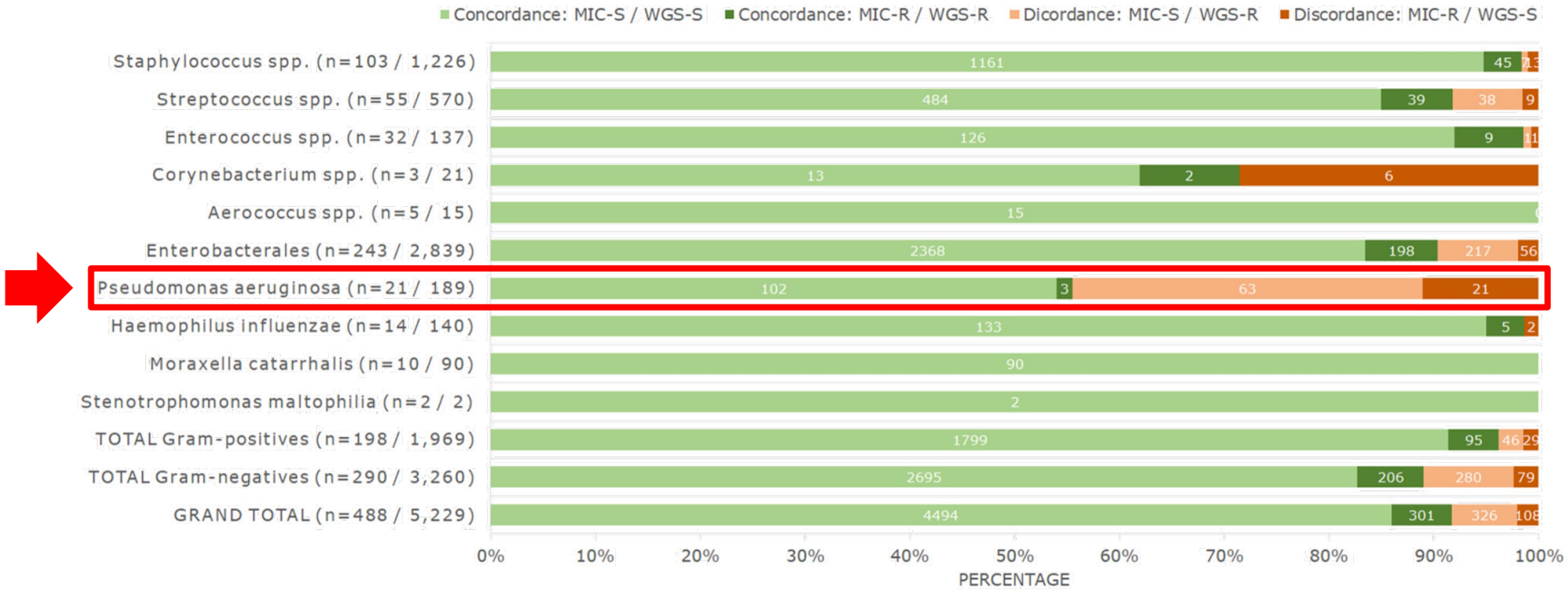
[1] National Food Institute, Technical University of Denmark, Kongens Lyngby, Denmark, [2] Department of Bacteria, Parasites and Fungi, Statens Serum Institut, Copenhagen, Denmark, [3] Department of Clinical Microbiology, Herlev and Gentofte Hospital, Herlev, Denmark, [4] Department of Clinical Microbiology, Aalborg University Hospital, Aalborg, Denmark, [5] Department of Clinical Medicine, Aalborg University, Aalborg, Denmark, [6] Department of Clinical Microbiology, Aarhus University Hospital, Aarhus, Denmark, [7] Department of Clinical Microbiology, Odense University Hospital, Odense, Denmark, [8] Department of Clinical Microbiology, Slagelse Hospital, Slagelse, Denmark, [9] Department of Clinical Microbiology, Rigshospitalet, Copenhagen, Denmark, [10] Department of Clinical Microbiology, Hospital of Southern Jutland, Sønderborg, Denmark, [11] Department of Clinical Microbiology, Hospital of South West Jutland, Esbjerg, Denmark, [12] Department of Clinical Microbiology, Vejle Hospital, Vejle, Denmark, [13] Department of Clinical Microbiology, Hvidovre Hospital, Copenhagen University Hospital – Amager and Hvidovre, Hvidovre, Denmark, [14] Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark*

Abstract

Antimicrobial susceptibility testing (AST) should be fast and accurate, leading to proper interventions and therapeutic success. Clinical microbiology laboratories rely on phenotypic methods, but the continuous improvement and decrease in the cost of whole-genome sequencing (WGS) technologies make them an attractive alternative. Studies evaluating the performance of WGS-based prediction of antimicrobial resistance (AMR) for selected bacterial species have shown promising results. There are, however, significant gaps in the literature evaluating the applicability of WGS as a diagnostics method in real-life clinical settings against the range of bacterial pathogens experienced there. Thus, we compared standard phenotypic AST results with WGS-based predictions of AMR profiles in bacterial isolates without preselection of defined species, to evaluate the applicability of WGS as a diagnostics method in clinical settings. We collected all bacterial isolates processed by all Danish Clinical Microbiology Laboratories in 1 day. We randomly selected 500 isolates without any preselection of species. We performed AST through standard broth microdilution (BMD) for 488 isolates (n = 6,487 phenotypic AST results) and compared results with *in silico* antibiograms obtained through WGS (Illumina NextSeq) followed by bioinformatics analyses using ResFinder 4.0 (n = 5,229 comparisons). A higher proportion of AMR

Genotype vs Phenotype Concordance

Out of the **5.229** isolate-antimicrobial combinations for which MIC and WGS data were compared **8.3% had discordances (not equally distributed)** and **91.7% showed a concordance** between phenotypic and genotypic antimicrobial susceptibility profiles

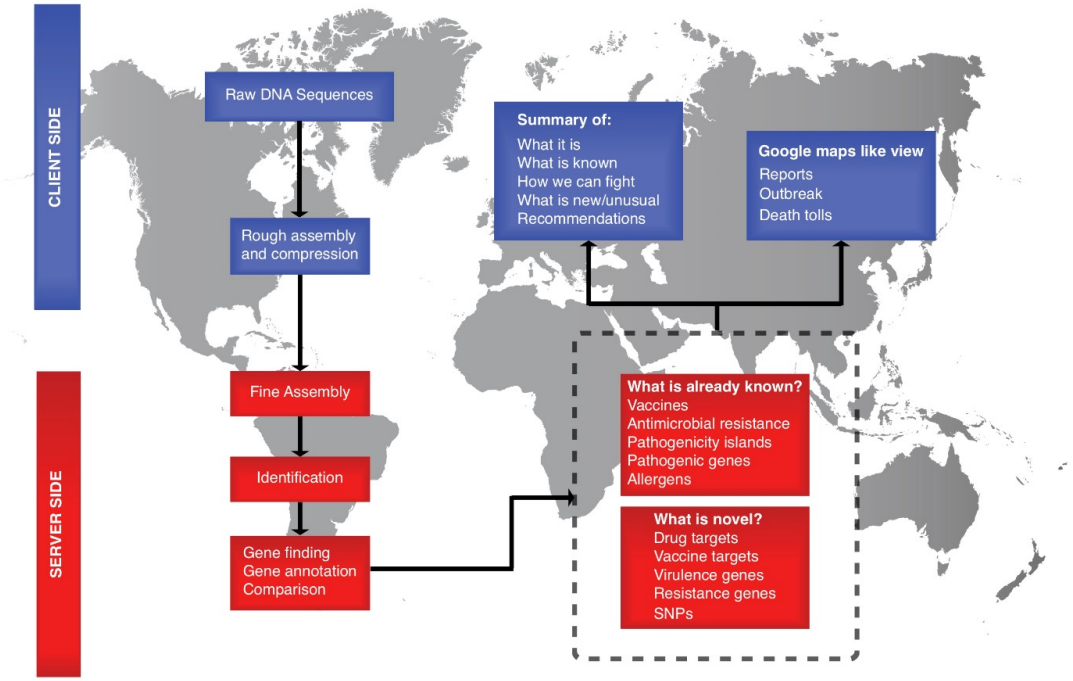


Rebelo AR. et al. (2022). Frontiers in Microbiology. 13

ResFinder

Click **SERVICES** to see all the available tools

1



News

Scaling neighbor joining to one million taxa with dynamic and heuristic neighbor joining
January 2023
Link to article....

SourceFinder: a Machine-Learning-Based Tool for Identification of Chromosomal, Plasmid, and Bacteriophage Sequences from Assemblies
November 2022
Link to article....

PlasmidHostFinder: prediction of plasmid hosts using random forest
April 2022
Link to article....

ResFinder – an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes
January 2022

2

Phenotyping

[ResFinder](#)
Identifcation of acquired antibiotic resistance genes. and chromosomal gene mutations

[ResFinderFG](#)
Identifcation of functional metagenomic antibiotic resistance determinants.

[LRE-finder](#)
Identifcation of genes and mutations leading to linezolid resistance.

[KmerResistance](#)
Identifcation of acquired antibiotic resistance genes using Kmers.

Florensa AF. et al. (2022). Microb Genom. 8(1):000748; https://www.genomicepidemiology.org/

ResFinder: Job Upload

ResFinder

Version

4.5.0

ResFinder identifies acquired genes and/or finds chromosomal mutations mediating antimicrobial resistance in total or partial DNA sequence of bacteria.

ResFinder software: (2024-03-22)
ResFinder database: (2024-03-22)
PointFinder database: (2024-03-08)
DisinFinder database: (2023-05-31)

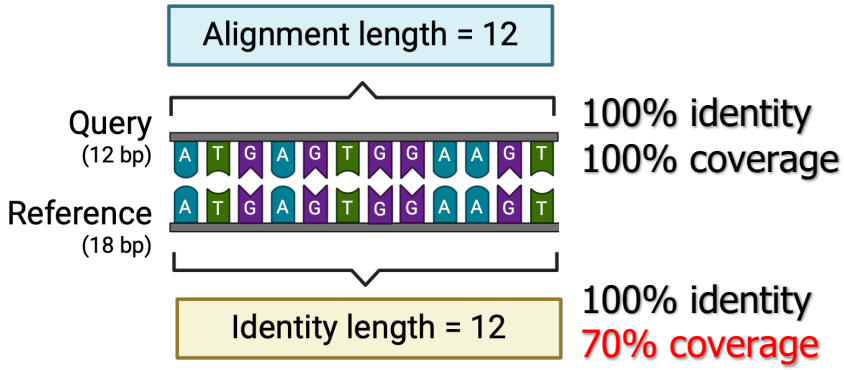
Chromosomal point mutations:

Threshold for %ID: 90%
Minimum length: 60%

Acquired antimicrobial resistance genes:

Threshold for %ID: 90%
Minimum length: 60%

Show unknown mutations
Ignore premature stop codons:
Ignore frameshift indels:

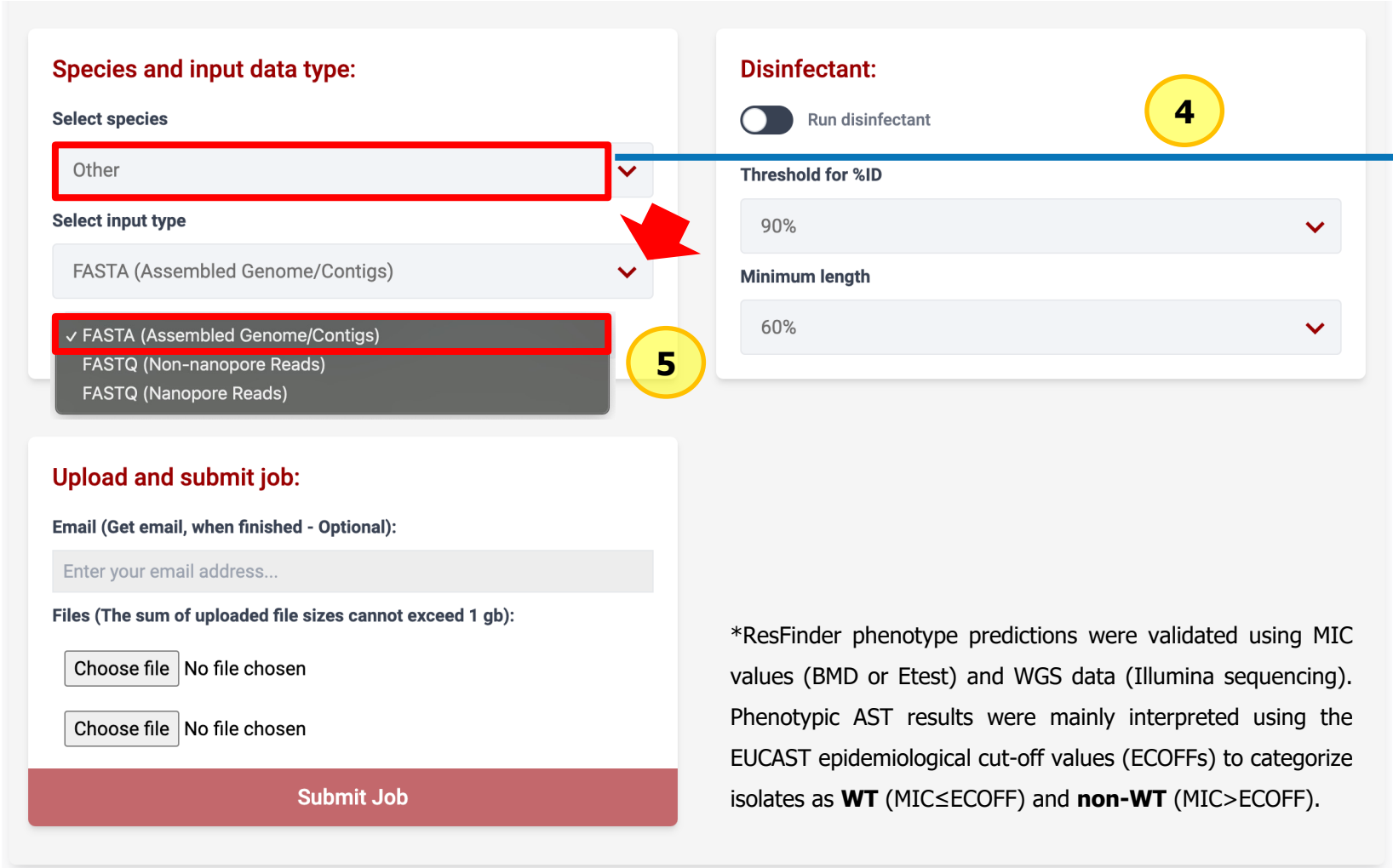


3

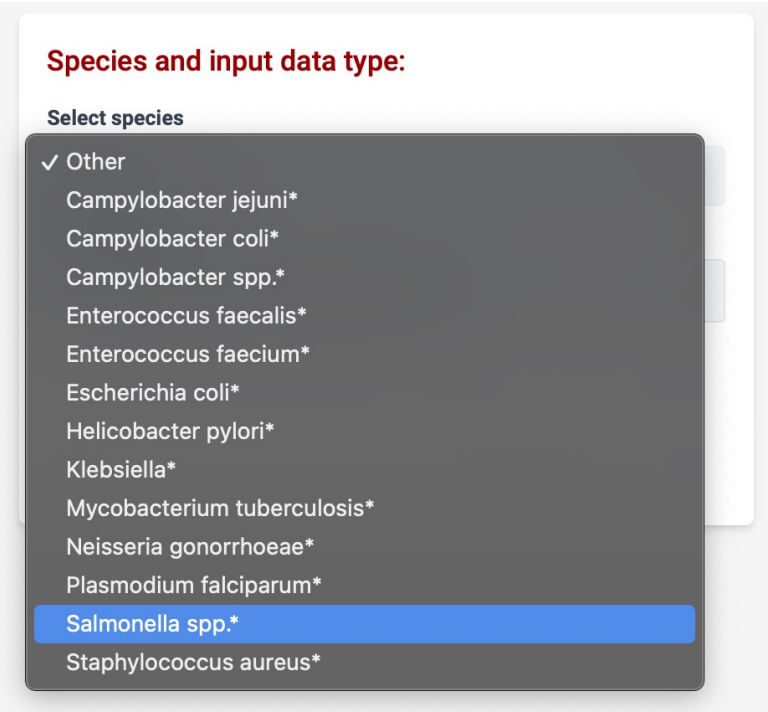
% identity and **coverage** can be adjusted if necessary

Click to view **known** and **unknown** mutations associated with AMR

ResFinder: Job Upload

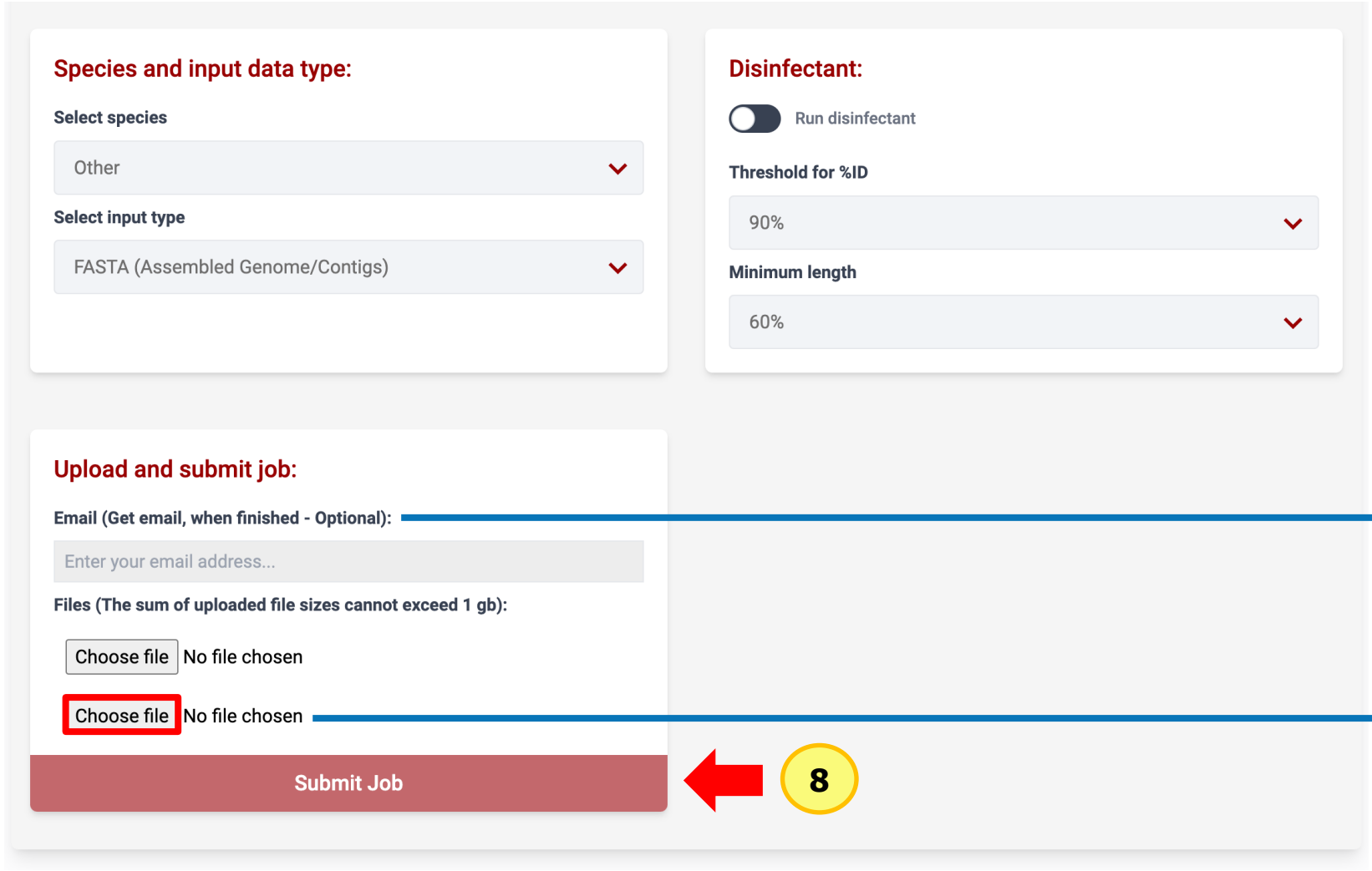


Species-relevant prediction of **clinically relevant phenotypes*** for selected antimicrobial agents



*ResFinder phenotype predictions were validated using MIC values (BMD or Etest) and WGS data (Illumina sequencing). Phenotypic AST results were mainly interpreted using the EUCAST epidemiological cut-off values (ECOFFs) to categorize isolates as **WT** (MIC≤ECOFF) and **non-WT** (MIC>ECOFF).

ResFinder: Job Upload



6

Write your **email** to receive the results when finished

7

Upload the ***.fasta*** or ***.fastq*** **files** of your bacterial genome for analysis

ResFinder: Analysis of Results

Phenotypes

escherichia coli | All

Hide

Antimicrobial	Class	WGS-predicted phenotype	Genetic background
gentamicin	aminoglycoside	No resistance	
tobramycin	aminoglycoside	No resistance	
amikacin	aminoglycoside	No resistance	
ciprofloxacin	quinolone	Resistant	
nalidixic acid	quinolone	Resistant	
ampicillin	beta-lactam	Resistant	blaTEM-1A;;1;;HM749966
ampicillin+clavulanic acid	beta-lactam	No resistance	
cefepime	beta-lactam	No resistance	
cefotaxime	beta-lactam	No resistance	
cefoxitin	beta-lactam	No resistance	
ceftazidime	beta-lactam	No resistance	
ertapenem	beta-lactam	No resistance	
imipenem	beta-lactam	No resistance	
meropenem	beta-lactam	No resistance	
piperacillin+tazobactam	beta-lactam	No resistance	
temocillin	beta-lactam	No resistance	
sulfamethoxazole	folate pathway antagonist	Resistant	sul1;;2;;U12338
trimethoprim	folate pathway antagonist	Resistant	dfrA1;;10;;AF203818
fosfomycin	fosfomycin	No resistance	
azithromycin	macrolide	No resistance	
tetracycline	tetracycline	Resistant	tet(M);;5;;U58985
tigecycline	tetracycline	No resistance	
chloramphenicol	amphenicol	Resistant	cmlA1;;1;;M64556
colistin	polymyxin	Resistant	mcr-1.1;;1;;KP347127

Prediction of species-specific ***in silico*** **antibiograms**

Only **known** AMR mechanisms are detected

When using **("All" option)** interpretation of results must be executed carefully → knowledge on **intrinsic resistance** (often mediated by structural traits) is **essential**

Bortolaia V. et al. (2020). J Antimicrob Chemother. 75(12):3491-3500; http://genepi.food.dtu.dk/resfinder

ResFinder: Analysis of Results

Acquired AMR gene hits

Hide

Resistance gene	Identity	Alignment length/gene length	Position in reference	Contig or depth	Position in contig	Phenotype	PMID	Accession no.	Notes
aadA2b	99.87%	780 / 780	1...780	DTU_2022_1023502_1_SI_S21FP06270_NODE_159_length_4297_cov_16.349400	1726...2505	['streptomycin', 'spectinomycin']	85862590	D43625	
aph(3')-Ia	100.00%	816 / 816	1...816	DTU_2022_1023502_1_SI_S21FP06270_NODE_216_length_1293_cov_9.157804	214...1029	['kanamycin', 'neomycin', 'lividomycin', 'paromomycin', 'ribostamycin']	6270337	V00359	
aadA1	100.00%	722 / 792	1...722	DTU_2022_1023502_1_SI_S21FP06270_NODE_245_length_737_cov_42.177049	16...737	['streptomycin', 'spectinomycin']	23180481	JQ414041	Alternative name ant(3")-Ia
blaTEM-1A	100.00%	861 / 861	1...861	DTU_2022_1023502_1_SI_S21FP06270_NODE_19_length_64335_cov_9.271337	4073...4933	['amoxicillin', 'ampicillin', 'piperacillin', 'ticarcillin', 'cephalothin']	21393132	HM749966	Class A
mcr-1.1	100.00%	1626 / 1626	1...1626	DTU_2022_1023502_1_SI_S21FP06270_NODE_179_length_2787_cov_8.350000	239...1864	['colistin']	26603172	KP347127	
cmlA1	99.92%	1260 / 1260	1...1260	DTU_2022_1023502_1_SI_S21FP06270_NODE_159_length_4297_cov_16.349400	205...1464	['chloramphenicol']	1648560	M64556	MFS transporter
sul1	100.00%	840 / 840	1...840	DTU_2022_1023502_1_SI_S21FP06270_NODE_144_length_6159_cov_10.065816	702...1541	['sulfamethoxazole']	unpublished	U12338	
sul3	100.00%	792 / 792	1...792	DTU_2022_1023502_1_SI_S21FP06270_NODE_155_length_4607_cov_16.789286	1938...2729	['sulfamethoxazole']	12604565	AJ459418	
tet(M)	99.17%	1920 / 1920	1...1920	DTU_2022_1023502_1_SI_S21FP06270_NODE_180_length_2631_cov_8.655351	522...2441	['tetracycline', 'doxycycline', 'minocycline']	9372426	U58985	

Bortolaia V. et al. (2020). J Antimicrob Chemother. 75(12):3491-3500; http://genepi.food.dtu.dk/resfinder

ResFinder: Analysis of Results

Acquired AMR gene hits

Resistance gene	Identity	Alignment length/gene length	Position in reference	Contig or depth
aadA2b	99.87%	780 / 780	1...780	DTU_2022_1023502_1_SI_S21FP06270_NODE_159_length_4297_cov_16.349400
aph(3')-Ia	100.00%	816 / 816	1...816	DTU_2022_1023502_1_SI_S21FP06270_NODE_216_length_1293_cov_9.157804
aadA1	100.00%	722 / 792		DTU_2022_1023502_1_SI_S21FP06270_NODE_245_length_737_cov_42.177049
blaTEM-1A	100.00%	861 / 861	1...861	DTU_2022_1023502_1_SI_S21 ... cov_9.271337
mcr-1.1	100.00%	1626 / 1626	1...1626	DTU_2022_1023502_1_SI_S21FP06270_NODE_179_length_2...
cmlA1	99.92%	1260 / 1260	1...1260	DTU_2022_1023502_1_SI_S21FP06270_NODE_159_length_4297_cov_16.349400
sul1	100.00%	840 / 840	1...840	DTU_2022_1023502_1_SI_S21FP06270_NODE_144_length_6159_cov_10.065816
sul3	100.00%	792 / 792	1...792	DTU_2022_1023502_1_SI_S21FP06270_NODE_155_length_4607_cov_16.789286
tet(M)	99.17%	1920 / 1920	1...1920	DTU_2022_1023502_1_SI_S21FP06270_NODE_180_length_2631_cov_8.655351

The number of aligned bases in our genome/total length of the reference gene in the database

Example: 780/780 means our sequence aligned for 780 bases out of the full reference gene

The dark green color

Indicates a **perfect match** for a give gene. **Identity** is **100%** and the sequence in the genome **covers the entire length** of the resistance gene in the database

ResFinder: Analysis of Results

Acquired AMR gene hits

Resistance gene	Identity	Alignment length/gene length	Position in reference	Contig or depth
aadA2b	99.87%	780 / 780	1...780	DTU_2022_1023502_1_SI_S21FP06270_NODE_159_length_4297_cov_16.349400
aph(3')-Ia	100.00%	816 / 816	1...816	DTU_2022_1023502_1_SI_S21FP...
aadA1	100.00%	722 / 792	1...722	DTU_2022_1023502_1_SI_S21FP06270_NODE_245_length_737_cov_42.177049
blaTEM-1A	100.00%	861 / 861	1...861	DTU_2022_1023502_1_SI_S21FP06270_NODE_19_length_64335_cov_9.271337
mcr-1.1	100.00%	1626 / 1626	1...1626	DTU_2022_1023502_1_SI_S21FP06270_NODE_179_length_2787_cov_8.350000
cmlA1	99.92%	1260 / 1260	1...1260	DTU_2022_1023502_1_SI_S21FP06270_NODE_159_length_4297_cov_16.349400
sul1	100.00%	840 / 840		
sul3	100.00%	792 / 792	1...792	DTU_2022_1023502_1_SI_S21FP06270...
tet(M)	99.17%	1920 / 1920	1...1920	DTU_2022_1023502_1_SI_S21FP06270_NODE_180_length_2631_cov_8.655351

The grey color

Indicates a **warning** due to a non-perfect match, % ID is 100% → alignment length **is different** than resistance gene length

The light green color

Indicates a **warning** due to a non-perfect match, % ID < 100% → alignment length **is equal** to the resistance gene length

Bortolaia V. et al. (2020). J Antimicrob Chemother. 75(12):3491-3500; http://genepi.food.dtu.dk/resfinder

ResFinder: Analysis of Results

Resistance gene	Identity	Alignment length/gene length	Position in reference	Contig or depth
aadA2b	99.87%	780 /	1...780	DTU_2022_1023502_1_SI_S21FP06270_NODE_159_length_4297_cov_16.349400
aph(3')-Ia	100.00%	816 /	1...816	DTU_2022_1023502_1_SI_S21FP06270_NODE_216_length_1293_cov_9.157804
aadA1	100.00%	722 /	1...722	DTU_2022_1023502_1_SI_S21FP06270_NODE_245_length_737_cov_42.177049
blaTEM-1A	100.00%	861 /	1...861	DTU_2022_1023502_1_SI_S21FP06270_NODE_19_length_64335_cov_9.271337
mcr-1.1	100.00%	1626 /	1...1626	DTU_2022_1023502_1_SI_S21FP06270_NODE_179_length_2787_cov_8.350000
cmlA1	99.92%	1260 /	1...1260	DTU_2022_1023502_1_SI_S21FP06270_NODE_159_length_4297_cov_16.349400
sul1	100.00%	840 /	1...840	DTU_2022_1023502_1_SI_S21FP06270_NODE_144_length_6159_cov_10.065816
sul3	100.00%	792 /	1...792	DTU_2022_1023502_1_SI_S21FP06270_NODE_155_length_4607_cov_16.789286
tet(M)	99.17%	1920 /	1...1920	DTU_2022_1023502_1_SI_S21FP06270_NODE_180_length_2631_cov_8.655351

Gene Alignments

Show gene alignments

Click "**Show Gene Alignments**" to see the alignment between the query and the reference gene

aph(3')-Ia

Reference sequence: ATGTTCATGCCGCCTGTTTTTCCTGCTCATTGGCACGTTTCGCAACCTGT
Query sequence: ATGTTCATGCCGCCTGTTTTTCCTGCTCATTGGCACGTTTCGCAACCTGT

Reference sequence: TCTCATTGCGGACACCTTTTCCAGCCTCGTTTGGAAAGTTTCATTGCCAG
Query sequence: TCTCATTGCGGACACCTTTTCCAGCCTCGTTTGGAAAGTTTCATTGCCAG

Reference sequence: ACGGGACTCCTGCAATCGTCAAGGGATTGAAACCTATAGAAGACATTGCT
Query sequence: ACGGGACTCCTGCAATCGTCAAGGGATTGAAACCTATAGAAGACATTGCT

Bortolaia V. et al. (2020). J Antimicrob Chemother. 75(12):3491-3500; http://genepi.food.dtu.dk/resfinder

ResFinder: Analysis of Results

Chromosomal mutations mediating AMR

Hide

Gene / Region	Identity	Alignment Length / Gene Length	Position in reference	Config or Depth	Position in config	PMID	Accession no.	Notes
gyrA	99.51%	2628 / 2628	1..2628	DTU_2022_1023502_1_SI_S21FP06270_NODE_47_length_37275_cov_8.114380	23787..26414	No PMID	CP073768.1	No notes

Mutation	Nucleotide change	Amino acid change	Phenotype	PMID	Notes
p.S83L	tcg>ttg	s -> l	ciprofloxacin,nalidixic acid	8891148,2168148,12654733	undefined
p.D87N	gac>aac	d -> n	ciprofloxacin,nalidixic acid	12654733,22878251,1850972	undefined

Gene / Region	Identity	Alignment Length / Gene Length	Position in reference	Config or Depth	Position in config	PMID	Accession no.	Notes
parC	99.87%	2259 / 2259	1..2259	DTU_2022_1023502_1_SI_S21FP06270_NODE_51_length_35503_cov_8.962517	6191..8449	No PMID	CP084529.1	No notes

Mutation	Nucleotide change	Amino acid change	Phenotype	PMID	Notes
p.S80I	agc>atc	s -> i	ciprofloxacin,nalidixic acid	8851598,21856834-20638608,8524852,25631675	undefined

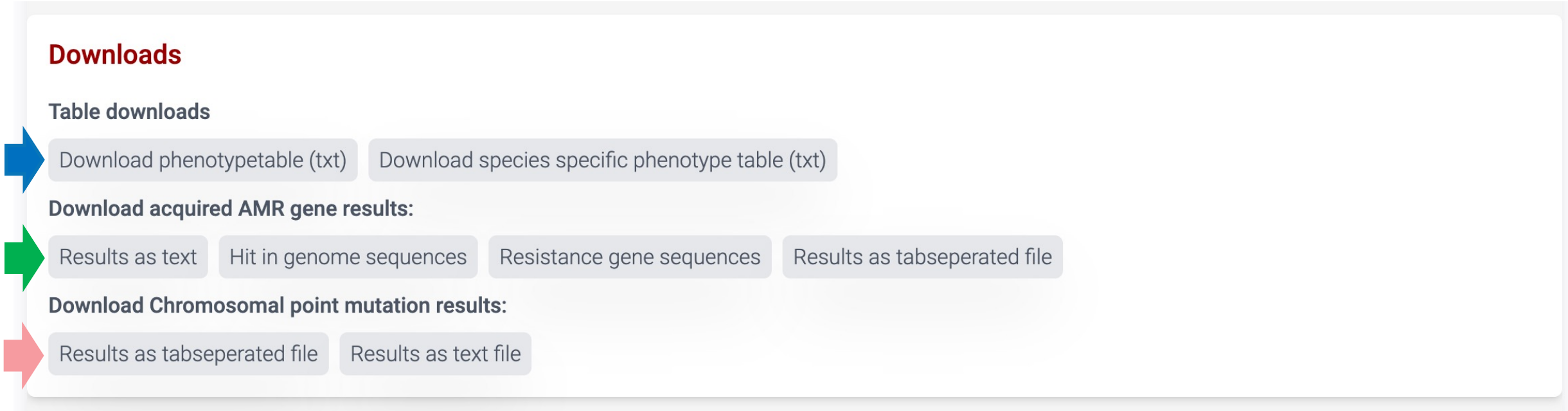
Gene / Region	Identity	Alignment Length / Gene Length	Position in reference	Config or Depth	Position in config	PMID	Accession no.	Notes
gyrB	99.13%	2415 / 2415	1..2415	DTU_2022_1023502_1_SI_S21FP06270_NODE_29_length_49656_cov_8.154071	33866..36280	No PMID	CP047010.1	No notes

Mutations known to confer resistance to antibiotics will appear bellow the corresponding gene (e.g., *gyrA* or *parC* mutations)

No known mutations detected

Bortolaia V. et al. (2020). J Antimicrob Chemother. 75(12):3491-3500; http://genepi.food.dtu.dk/resfinder

ResFinder: Analysis of Results



Download **all** or **species-specific** phenotypic and genotypic data in a *.txt* file format

Download of acquired **AMR genes** (*.txt* file format) and **gene sequences** (*.fasta* file format)

Download of chromosomal **point mutations** in a *.txt* file format

AMRFinderPlus

National Center for Biotechnology Information

Software (command line) and the **database** allow the identification of **acquired antimicrobial resistance** genes in bacterial protein and/or assembled nucleotide sequences as well as **known resistance-associated point mutations** for several taxa

```
# Install AMRFinderPlus and download the last database
$ mamba install ncbi-amrfinderplus
$ amrfinder -U

# See a list of all available organisms
$ amrfinder --list_organisms

# Run AMRFinderPlus in the the assembled genome
$ amrfinder -n file_dna.fasta -O Vibrio_cholerae -o AMRFinderPlus_results

# Run AMRFinderPlus in the the assembled genome with the “plus” option (virulence and metal+biocide tolerance genes)
$ amrfinder -n file_dna.fasta -O Vibrio_cholerae --plus -o AMRFinderPlus_plus_results
```

https://github.com/ncbi/amr/wiki

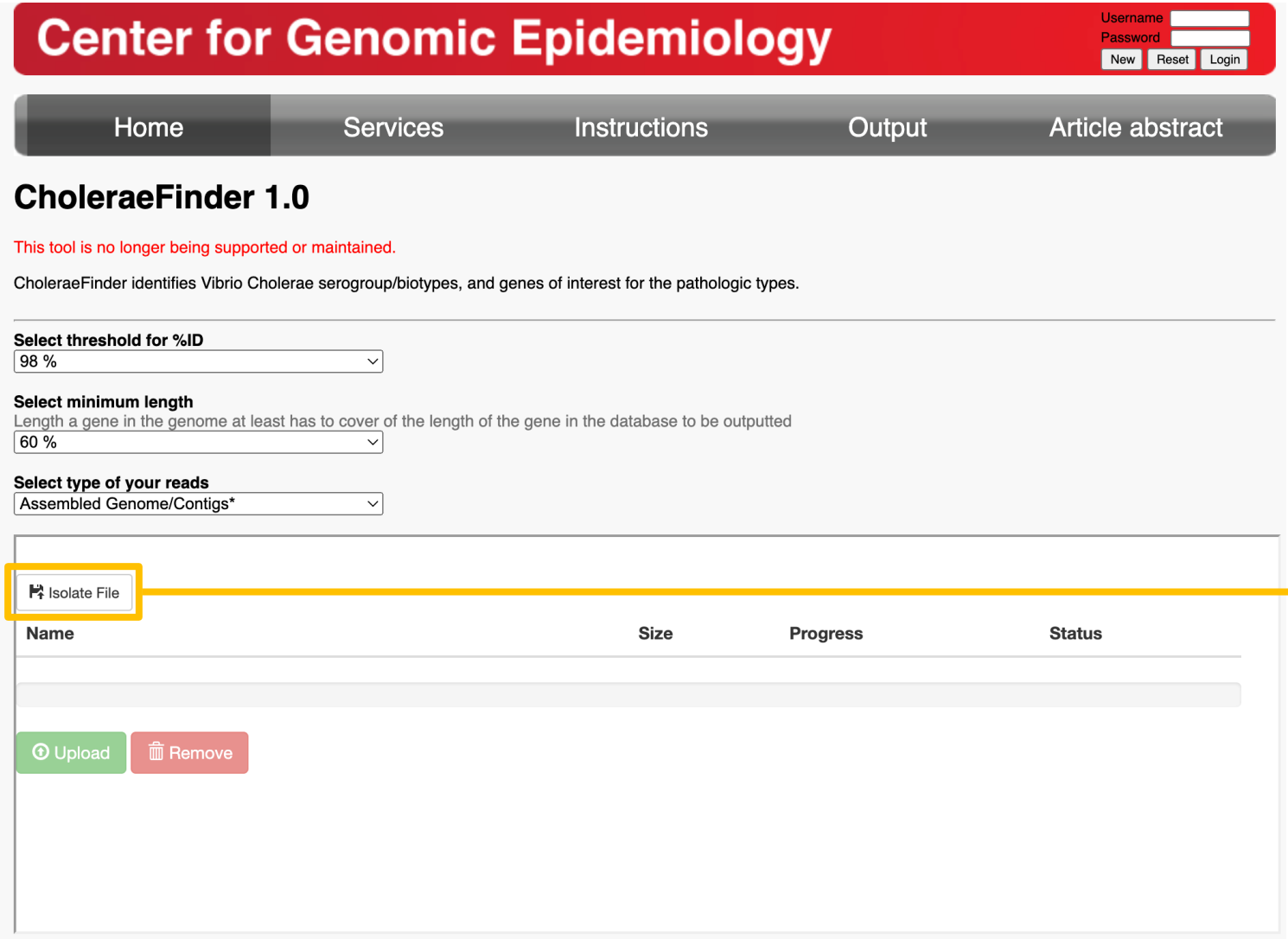
AMRFinderPlus: Output

National Center for Biotechnology Information

Contig id	Start	Stop	Strand	Element symbol	Element name	Scope	Type	Subtype	Class	Subclass	Method	Target length	Reference sequence length	% Coverage of reference	% Identity to reference	Alignment length	Closest reference accession	Closest reference name
NZ_LT906614.1	1566982	1567353	-	ctxB	cholera enterotoxin binding subunit CtxB	plus	VIRULENCE	VIRULENCE	NA	NA	EXACTX	124	124	**100**	**100**	124	AAF94613.1	cholera enterotoxin binding subunit CtxB
NZ_LT906614.1	1567353	1568126	-	ctxA	cholera enterotoxin catalytic subunit CtxA	plus	VIRULENCE	VIRULENCE	NA	NA	EXACTX	258	258	**100**	**100**	258	AAF94614.1	cholera enterotoxin catalytic subunit CtxA
NZ_LT906614.1	1674334	1675455	+	varG	VarG family subclass B1-like metallo-beta-lactamase	core	AMR	AMR	BETA-LACTAM	CARBAPENEM	EXACTX	374	374	**100**	**100**	374	WP_000778180.1	VarG family subclass B1-like metallo-beta-lactamase
NZ_LT906614.1	1687457	1688275	-	almG	glycine--lipid A transferase AlmG	core	AMR	AMR	COLISTIN	COLISTIN	EXACTX	273	273	**100**	**100**	273	WP_001881365.1	glycine--lipid A transferase AlmG
NZ_LT906614.1	1688280	1688513	-	almF	lipid A modification system glycine carrier protein AlmF	core	AMR	AMR	COLISTIN	COLISTIN	EXACTX	78	78	**100**	**100**	78	WP_000806242.1	lipid A modification system glycine carrier protein AlmF
NZ_LT906614.1	1688500	1690167	-	almE	lipid A modification system glycine--protein ligase AlmE	core	AMR	AMR	COLISTIN	COLISTIN	BLASTX	556	556	**100**	**98.38**	556	WP_252944138.1	lipid A modification system glycine--protein ligase AlmE
NZ_LT906615.1	315307	315933	+	catB9	type B-5 chloramphenicol O-acetyltransferase CatB9	core	AMR	AMR	PHENICOL	CHLORAMPHENICOL	EXACTX	209	209	**100**	**100**	209	WP_001009012.1	type B-5 chloramphenicol O-acetyltransferase CatB9

https://github.com/ncbi/amr/wiki

CholeraeFinder



1

Upload your results in **FASTQ** or **FASTA** (assembled genomes) format. The results will take a few minutes to appear

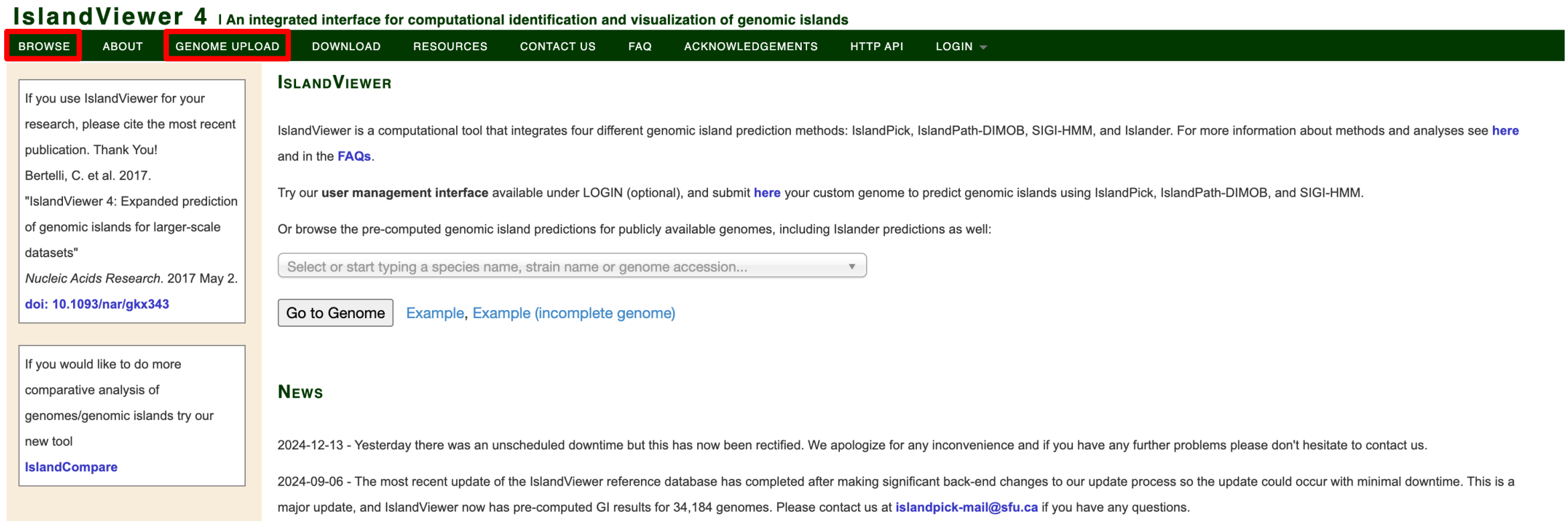
https://cge.food.dtu.dk/services/CholeraeFinder/

Annotation

Identification of Genetic Elements

IslandViewer 4

Predicts **genomic islands** (foreign DNA regions mostly acquired by horizontal transfer) using **four methods** (IslandPick-comparative genomics, IslandPath-DIMOB-mobility genes, SIGI-HMM-codon usage bias, Islander-integration near tRNA genes)

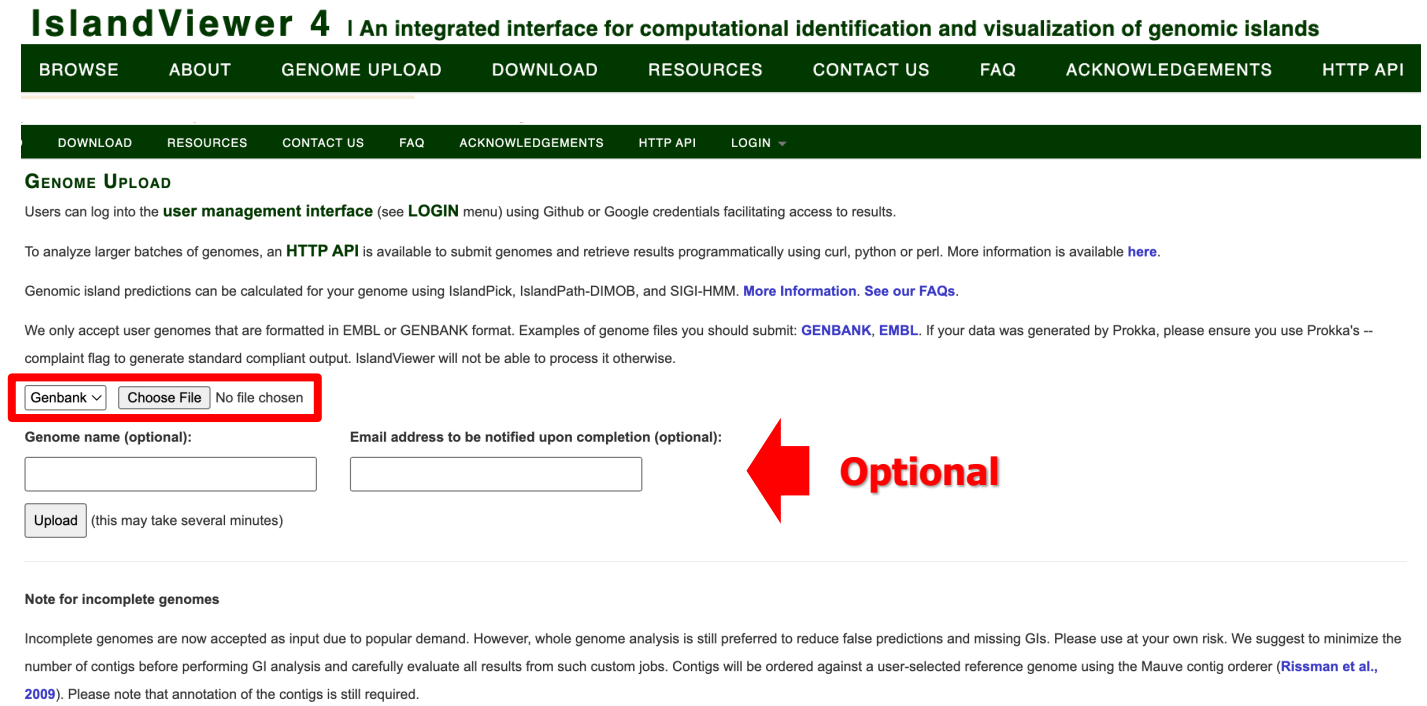


https://www.pathogenomics.sfu.ca/islandviewer

IslandViewer 4: Job Upload

1. **File format:** Use GenBank (.gbk/.gbff) or EMBL file
2. **Choose file:** Click "Choose File" and select your genome file.
3. (Optional) Add a genome name or email to be notified when the run finishes
4. Click "**Upload**"

Note: Whole genomes work best — fragmented assemblies may give false positives or miss GIs

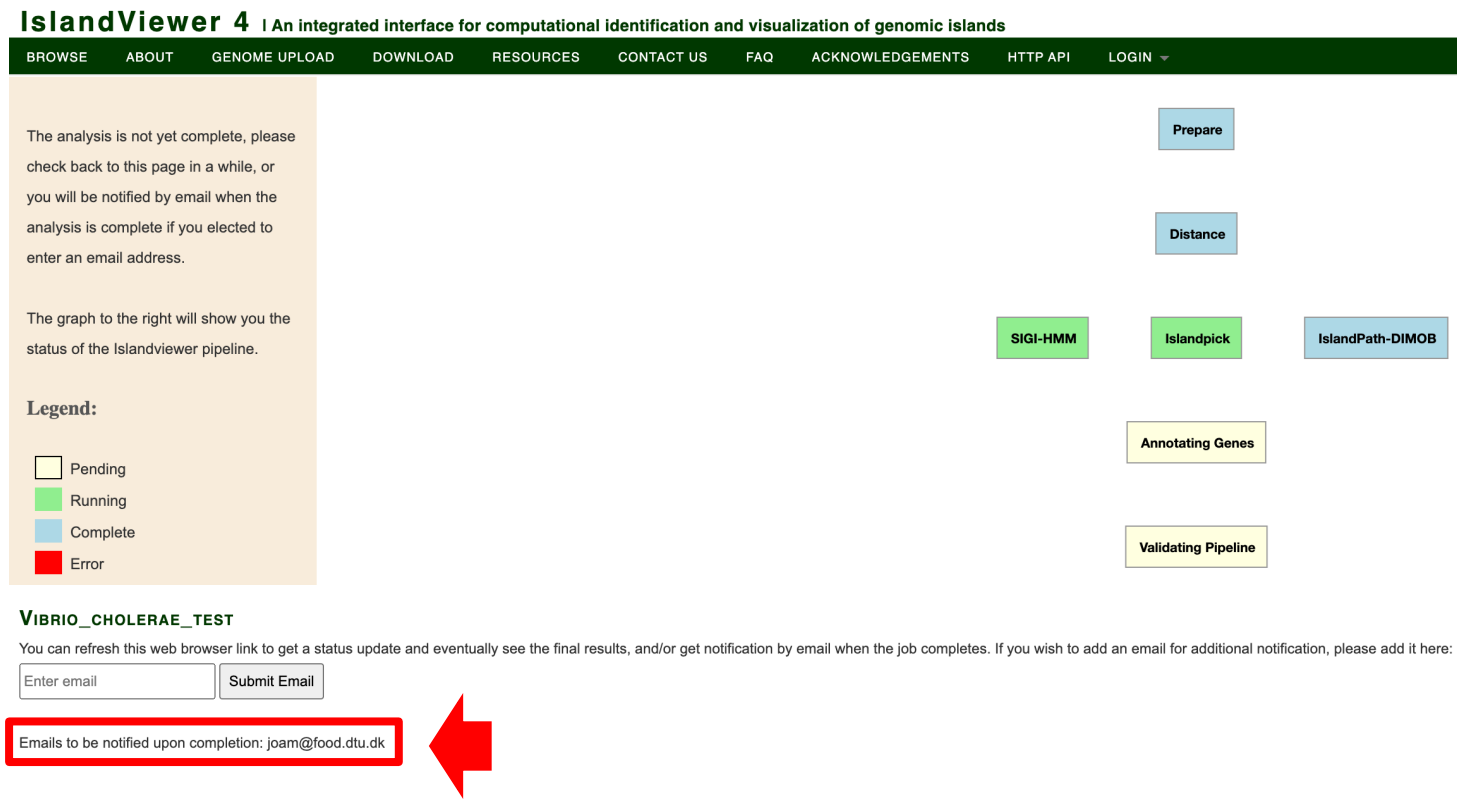


https://www.pathogenomics.sfu.ca/islandviewer

IslandViewer 4: Job Upload

1. Shows the pipeline steps (Prepare → Distance → SIGI-HMM → IslandPick → IslandPath-DIMOB → Annotation → Validation)
2. You can refresh the page or add an email to get notified when the run is complete (it make take some minutes)

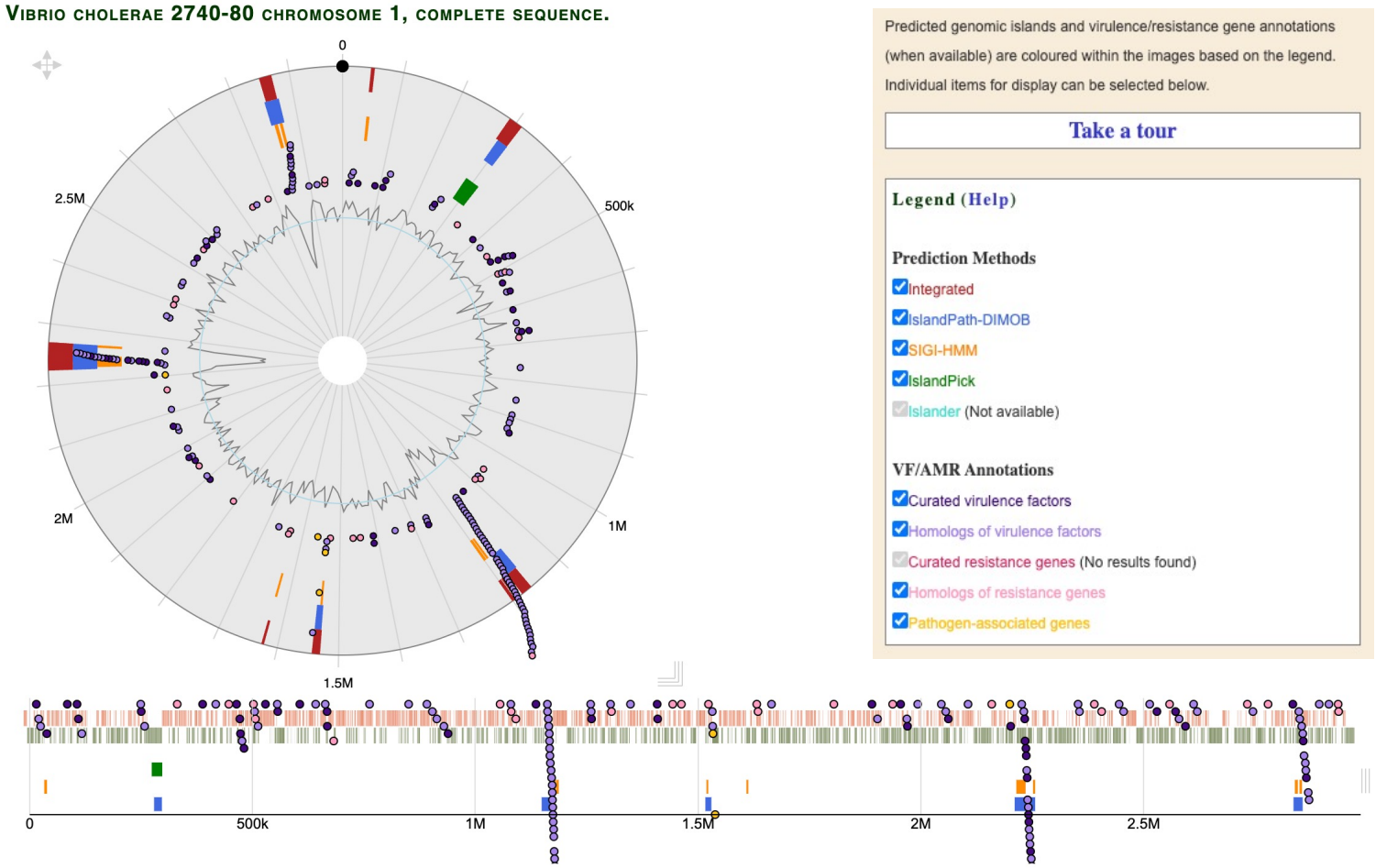
Meaning: This page only tells you the analysis progress. Final genomic island predictions appear once all steps are complete



IslandViewer 4: Analysis of Results

1. **Circular map** (top): Chromosome with predicted genomic islands (colored blocks).
2. **Linear map** (bottom): Same information, but easier to zoom
3. **Colors** = methods used:

- **Red** = integrated predictions (agreement between methods)
- **Blue** = IslandPath-DIMOB
- **Orange** = SIGI-HMM
- **Green** = IslandPick
- **Purple/red dots:** annotated virulence or resistance genes (if present)



https://www.pathogenomics.sfu.ca/islandviewer

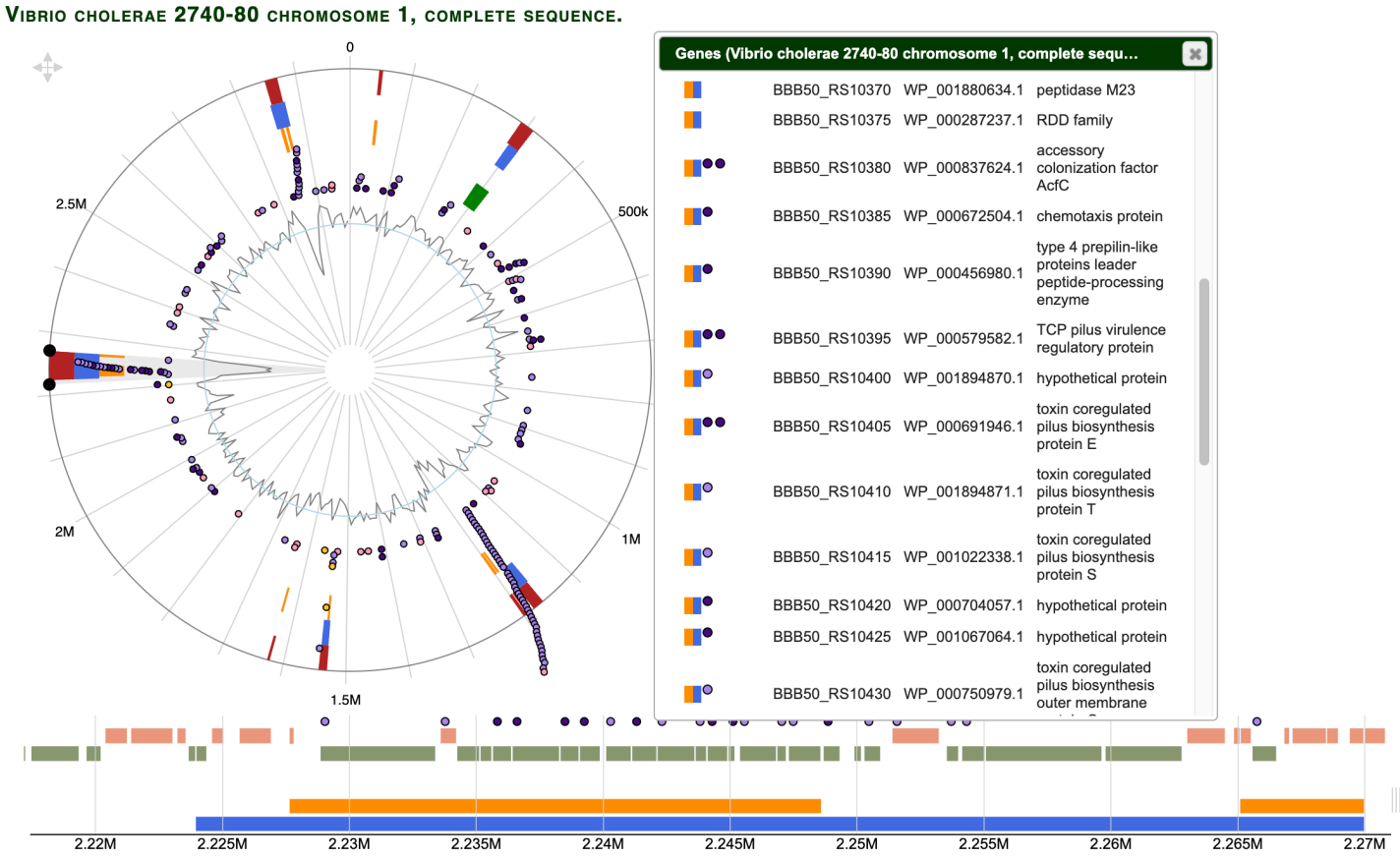
IslandViewer 4: Analysis of Results

1. By clicking in the **region** (≈2.23–2.25 Mb), you can see the genes inside that island
2. The **list shows**:

- Method flags used to identify the GIs
- Gene ID/locus tag
- Protein name/function

3. No antimicrobial resistance genes were identified here

Meaning: This island likely contributes to pathogenicity rather than AMR



IslandViewer 4: Analysis of Results

1. The table lists all **putative genomic islands (GIs)** detected
2. Each **row** = one predicted GI with:

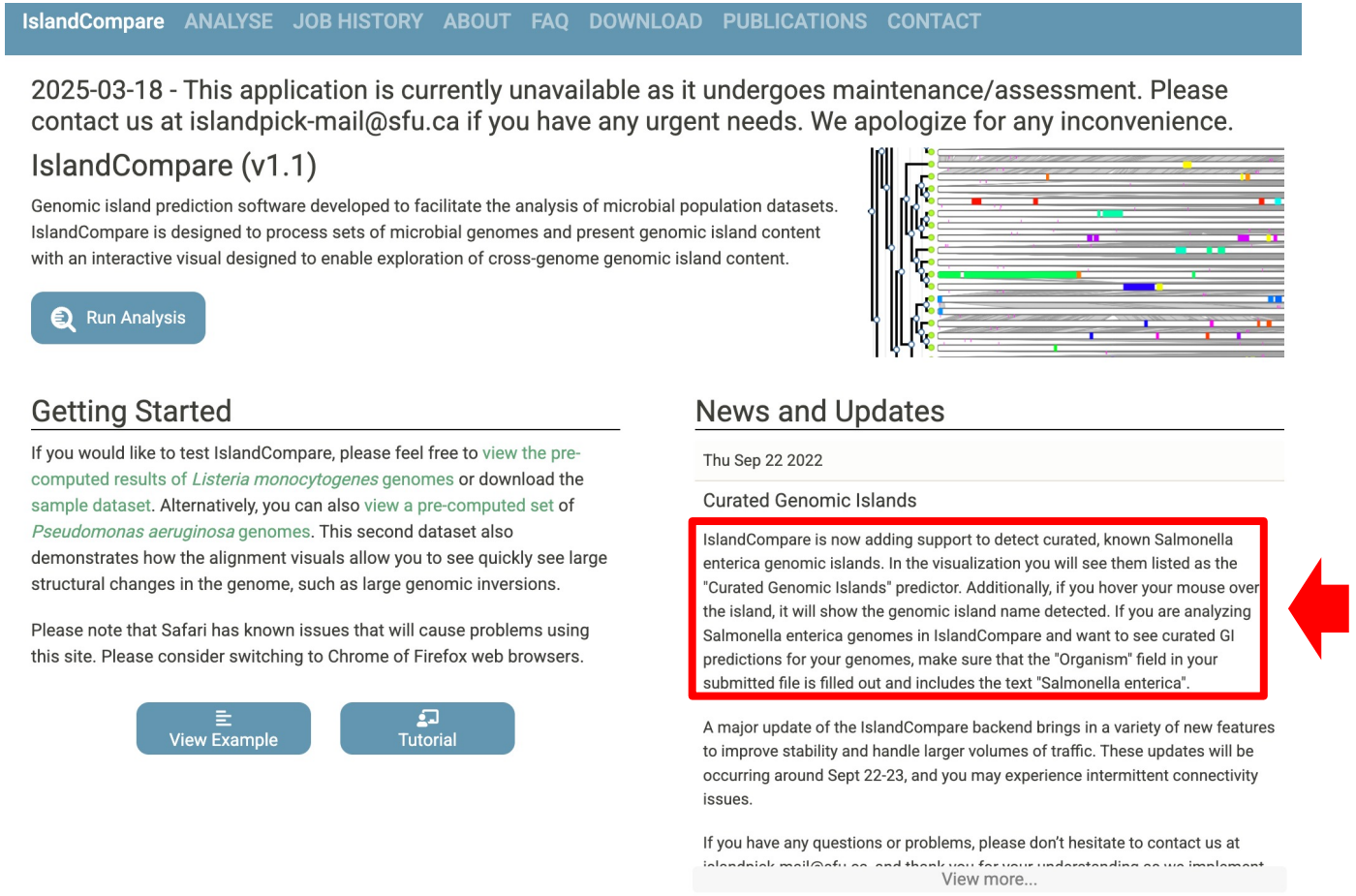
- Start/End coordinates and size
- Which **method(s)** flagged it (e.g. IslandPick, SIGI-HMM, etc.)
- Click "**View**" to see region on the plot
- **External annotations:** genes of interest inside the GI (e.g. virulence genes, pathogen-associated genes)
- **Download links** for sequence, genes, and proteins

Start	End	Size	GI Prediction Method	View	External Annotations	Download (all data)
46,238	52,477	6,239	Predicted by at least one method			Sequence, Genes, Proteins
287,443	311,170	23,727	Predicted by at least one method			Sequence, Genes, Proteins
292,783	310,359	17,576	Predicted by at least one method			Sequence, Genes, Proteins
1,162,634	1,181,481	18,847	Predicted by at least one method		Virulence gene	Sequence, Genes, Proteins
1,190,190	1,194,472	4,282	Predicted by at least one method		Virulence gene	Sequence, Genes, Proteins
1,195,470	1,201,132	5,662	Predicted by at least one method		Virulence gene	Sequence, Genes, Proteins
1,530,047	1,543,566	13,519	Predicted by at least one method		Virulence gene,Pathogen Associated gene	Sequence, Genes, Proteins
1,532,443	1,536,681	4,238	Predicted by at least one method		Virulence gene,Pathogen Associated gene	Sequence, Genes, Proteins
1,622,122	1,626,499	4,377	Predicted by at least one method			Sequence, Genes, Proteins
2,224,200	2,270,213	46,013	Predicted by at least one method		Virulence gene	Sequence, Genes, Proteins
2,227,893	2,248,826	20,933	Predicted by at least one method		Virulence gene	Sequence, Genes, Proteins
2,265,338	2,270,213	4,875	Predicted by at least one method		Virulence gene	Sequence, Genes, Proteins
2,850,054	2,870,627	20,573	Predicted by at least one method		Virulence gene	Sequence, Genes, Proteins
2,853,051	2,859,864	6,813	Predicted by at least one method		Virulence gene	Sequence, Genes, Proteins
2,863,444	2,869,227	5,783	Predicted by at least one method		Virulence gene	Sequence, Genes, Proteins
287,443	311,170	23,727	IslandPick			Sequence, Genes, Proteins
46,238	52,477	6,239	SIGI-HMM			Sequence, Genes, Proteins

IslandViewer 4: Analysis of Results

1. IslandViewer predicts putative genomic islands (GIs) but **does not assign family names**
2. In *Vibrio cholerae*, naming major families (e.g., SXT/R391 ICEs, VPI-1/2) requires **alignment** to **known**, **curated GI** reference sequences (plus identification of hallmark genes)

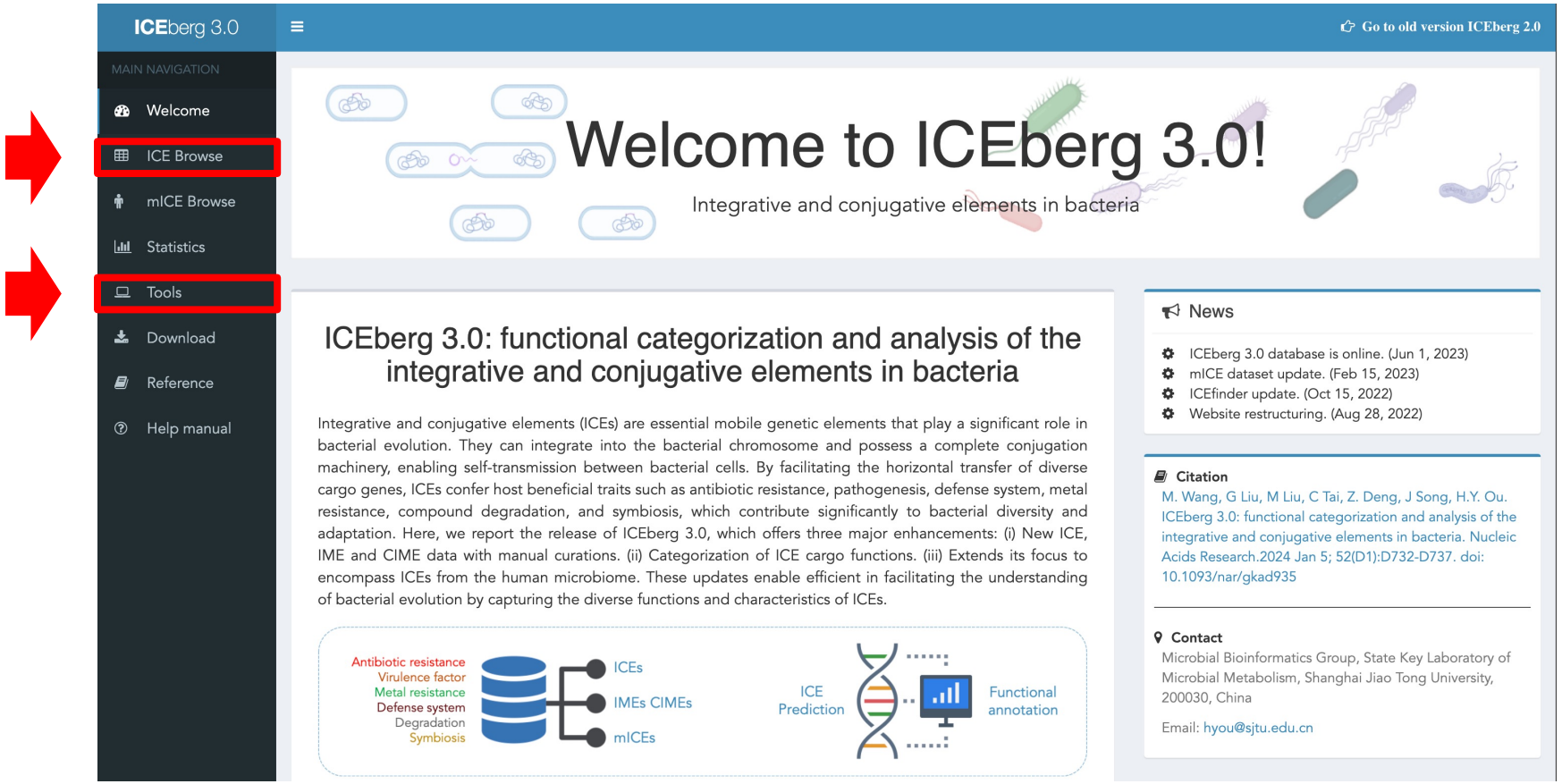
🚨 IslandCompare (within IslandViewer) currently provides curated GI references mainly for *Salmonella*, not *Vibrio*



https://www.pathogenomics.sfu.ca/islandviewer; https://github.com/brinkmanlab/IslandCompare

ICEberg 3.0

Provide a comprehensive, curated database and analysis platform for **Integrative and Conjugative Elements (ICEs)** and related mobile genetic elements, allowing users to explore, identify, and compare ICEs across bacterial genomes



https://tool2-mml.sjtu.edu.cn/ICEberg3/

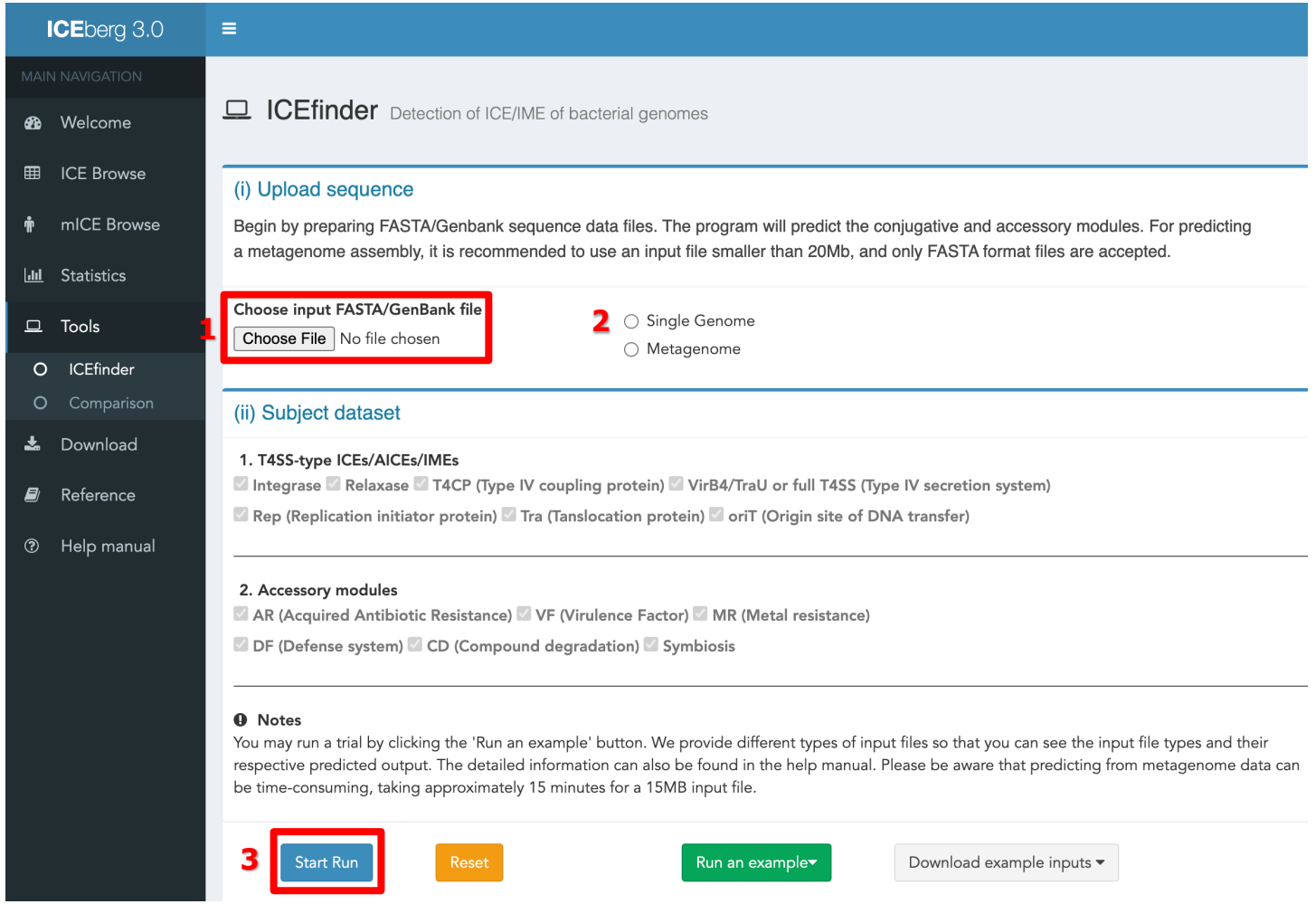
ICEberg 3.0: Job Upload

Click on **Tools** → **ICEfinder** in the top menu

https://tool2-mml.sjtu.edu.cn/ICEberg3/

ICEberg 3.0: Job Upload

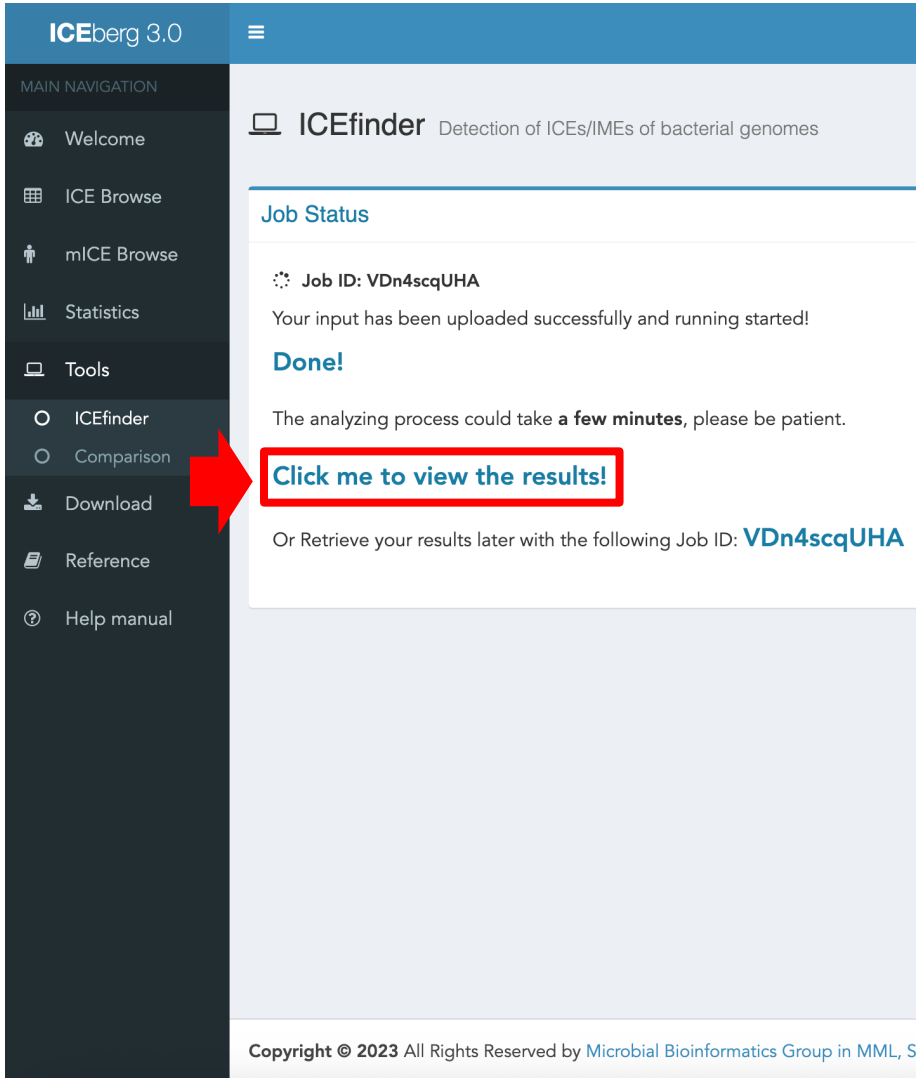
1. In ICEfinder, click "**Choose File**" and select the file containing your genome
2. Accepted formats include:
- **GenBank** format (.gb, .gbk, .gbff) → recommended, since it contains both sequence and feature annotations
- **FASTA** format (.fna, .fa) → accepted, but with fewer annotations (results may be less informative)
3. Click "**Single Genome**" (single cell finished genomes)
4. Press "**Start Run**"



https://tool2-mml.sjtu.edu.cn/ICEberg3/

ICEberg 3.0: Analysis of Results

- After uploading a genome file (FASTA or GenBank), ICEfinder starts an analysis job
- Each run is assigned a unique Job ID (**e.g. VDn4scqUHA**)
- The analysis takes a few minutes, depending on genome size
- When complete, ICEfinder shows "**Done!**" and provides a link
- Click to **view the results!**
- Results can also be retrieved later using the Job ID



https://tool2-mml.sjtu.edu.cn/ICEberg3/

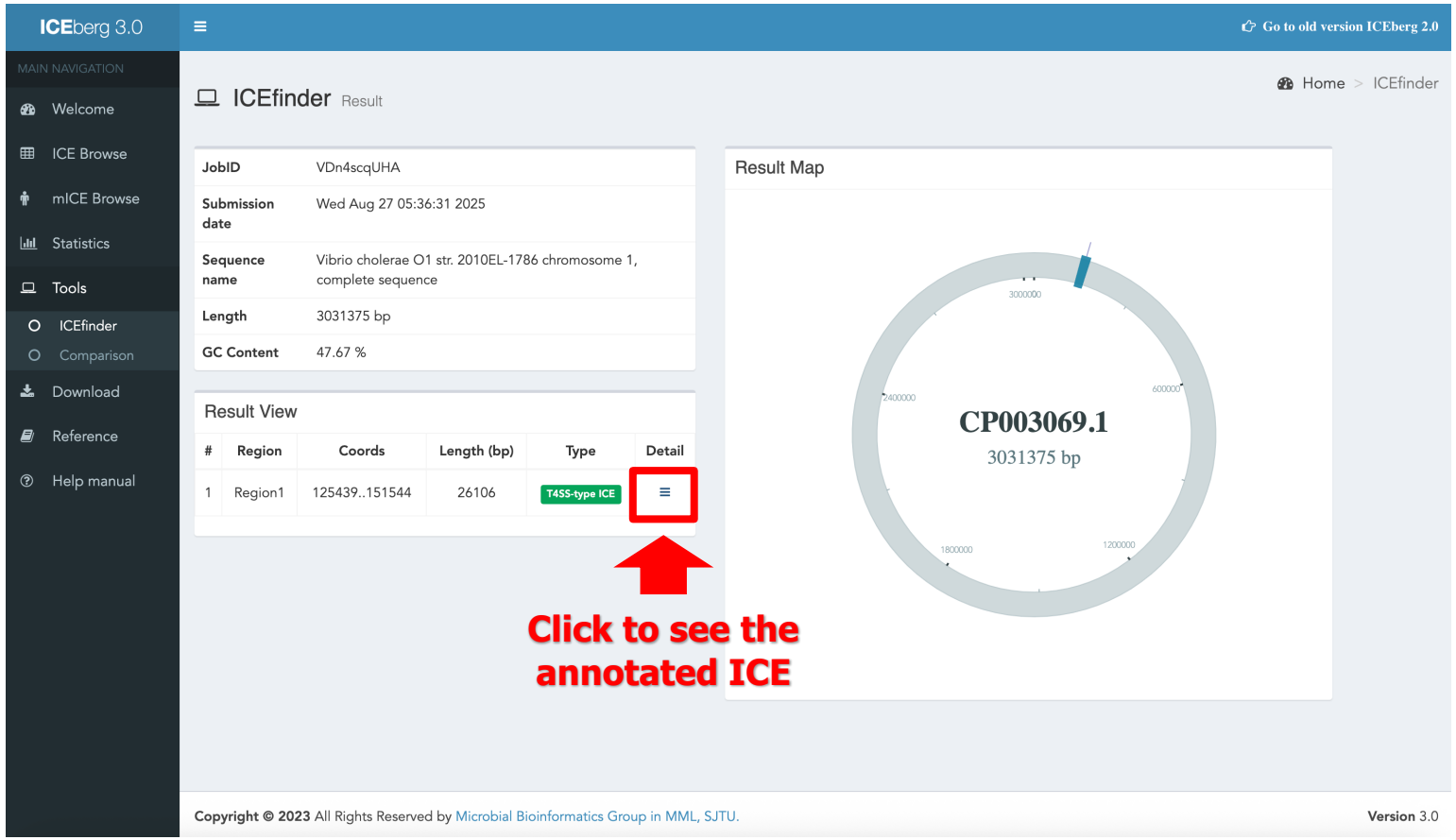
ICEberg 3.0: Analysis of Results

Circle (right): genome map; blue bar = predicted ICE location

Table "Result View" (left):

- Region 1 → coordinates (125–151 kb), length ~26 kb
- Type: T4SS-type ICE = core ICE signatures detected: integrase, relaxase, T4CP, T4SS

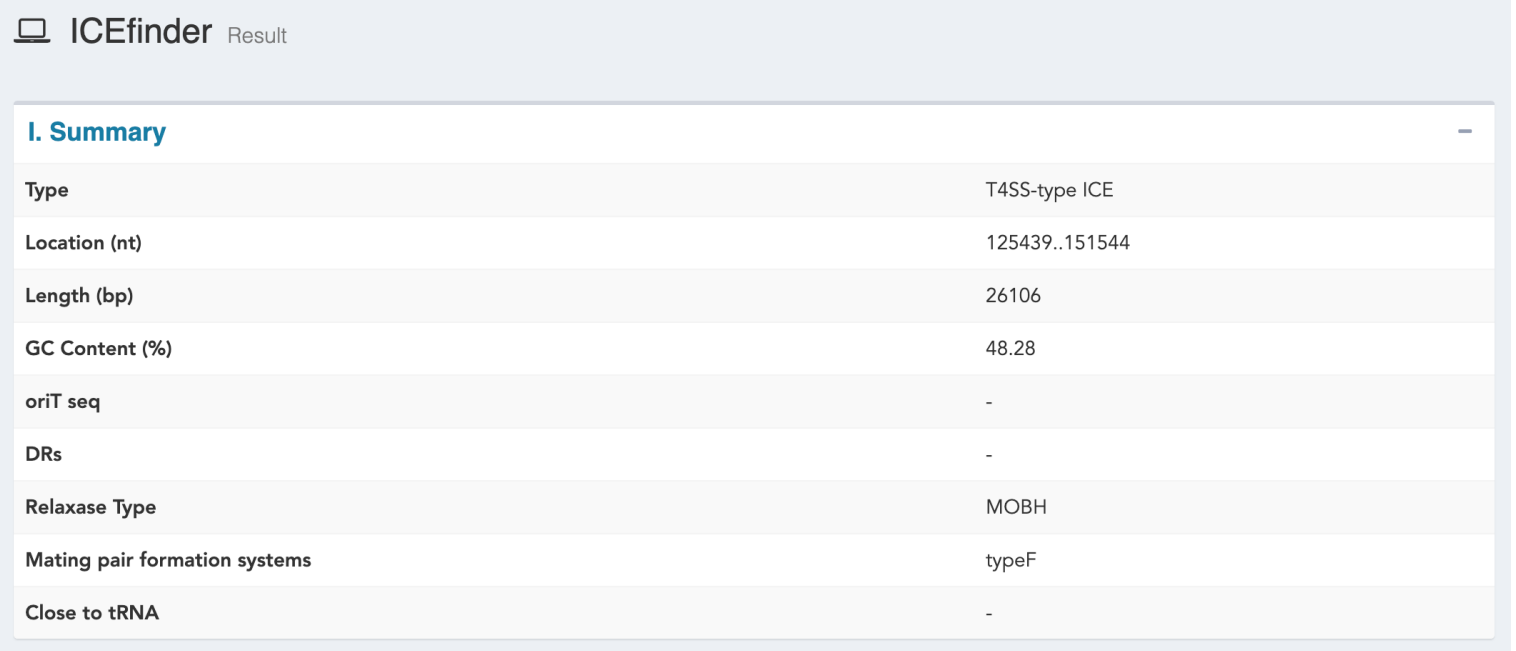
<u>Meaning</u>: Genome contains one ICE-like region



https://tool2-mml.sjtu.edu.cn/ICEberg3/

ICEberg 3.0: Analysis of Results

- **Type:** T4SS-type ICE → detected conjugation machinery (Type IV Secretion System)
- **Location:** coordinates 125–151 kb in the chromosome
- **Length:** ~26 kb region
- **GC Content:** 48.3 %, close to host genome
- **Relaxase type:** MOBH → enzyme that initiates DNA transfer
- **Mating pair formation system:** typeF → conjugative pilus system



ICEfinder Result

I. Summary

Type	T4SS-type ICE
Location (nt)	125439..151544
Length (bp)	26106
GC Content (%)	48.28
oriT seq	-
DRs	-
Relaxase Type	MOBH
Mating pair formation systems	typeF
Close to tRNA	-

Meaning: ICEfinder predicts one ICE-like region, identified as a T4SS-type element with MOBH relaxase and typeF conjugation system

https://tool2-mml.sjtu.edu.cn/ICEberg3/

ICEberg 3.0: Analysis of Results

- **Top graph:** GC% variation across the predicted ICE (differences may indicate foreign DNA)
- **Colored blocks** (bottom): functional genes/modules detected. E.g.,:
  - ➢ **Green** = defense system
  - ➢ **Pink** = T4SS genes
  - ➢ **Brown** = relaxase

Meaning: The region has the typical modular structure of an ICE: conjugation genes + accessory/defense genes

https://tool2-mml.sjtu.edu.cn/ICEberg3/

ICEberg 3.0: Analysis of Results

- **Each row** = one gene in the predicted ICE region
- **Columns:**
  - ➢ **Coordinates/size** → position on the genome
  - ➢ **Product** → predicted protein function
  - ➢ **Feature** → role in the ICE (defense, relaxase, T4CP, T4SS, etc.)

Meaning: Region has conjugation machinery + defense genes → typical ICE structure

III. Gene List

Gene	Coordinates [+/-], size (bp)	Product	Feature
Vch1786_I0098	115524..119265 [+], 3742	conserved BREX,pglX1	Flank; Defense
Vch1786_I0099	119277..121581 [+], 2305	PglZ domain protein; BREX,pglZA	Flank; Defense
Vch1786_I0100	121619..123671 [+], 2053	ATP-dependent Lon protease; BREX,brxL	Flank; Defense
Vch1786_I0101	123697..124468 [+], 772	conserved hypothetical protein	Flank
Vch1786_I0102	124502..125318 [+], 817	conserved AbiJ,AbiJ	Flank; Defense
Vch1786_I0103	125439..127590 [+], 2152	conjugal transfer pilus assembly protein TraI, MOBH	Relaxase
Vch1786_I0104	127638..129459 [+], 1822	conjugal transfer pilus assembly protein TraD, t4cp2	T4CP
Vch1786_I0105	129468..130029 [+], 562	plasmid-related protein	
Vch1786_I0106	130015..130651 [+], 637	plasmid-related protein, G_tfc7	T4SS
Vch1786_I0107	130677..131265 [-], 589	conserved hypothetical protein	
Vch1786_I0108	131278..131518 [-], 241	plasmid-related protein	
Vch1786_I0109	131553..131835 [+], 283	conjugal transfer pilus assembly protein TraL, F_traL	T4SS
Vch1786_I0110	131831..132458 [+], 628	conjugal transfer pilus assembly protein TraE, F_traE	T4SS

https://tool2-mml.sjtu.edu.cn/ICEberg3/

ICEberg 3.0: Analysis of Results

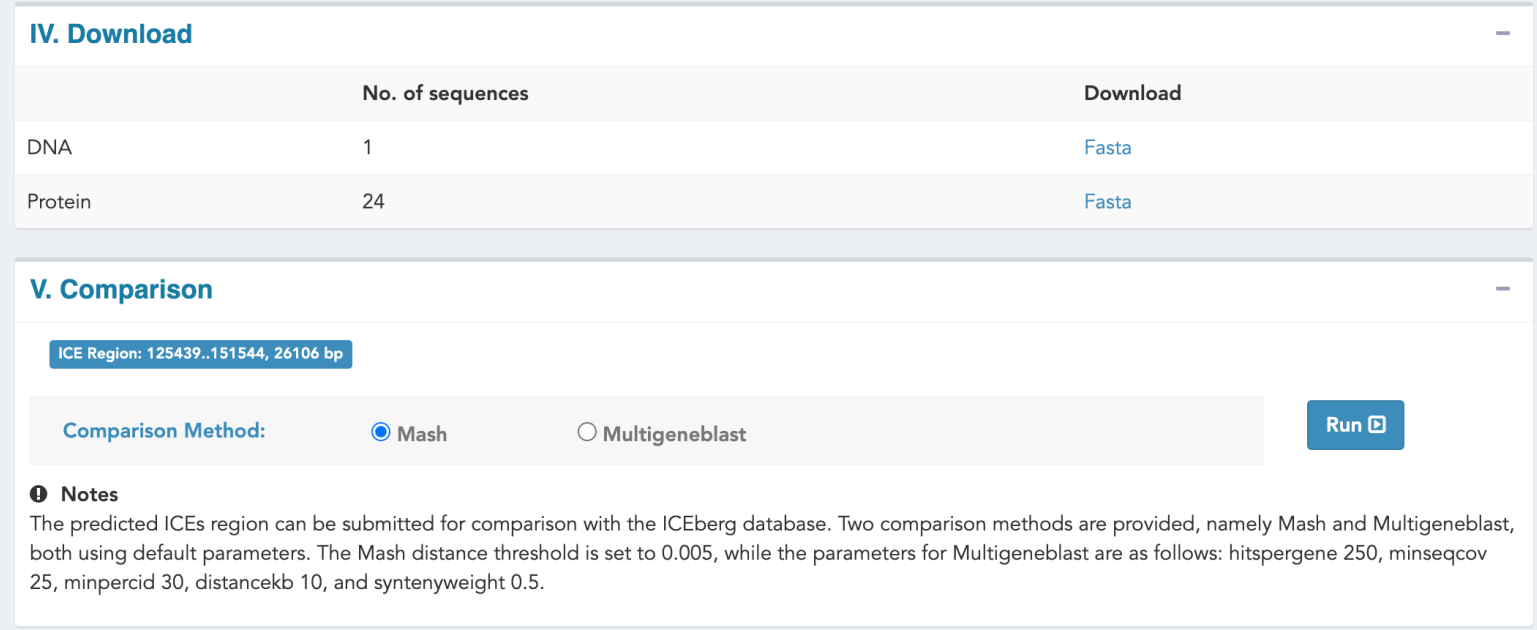
Download section:

- DNA (region sequence) or Protein (all predicted proteins) can be saved as FASTA files

Comparison section:

- Predicted ICE region can be compared with ICEberg's curated database

Two methods available: Mash (fast, k-mer) or Multigeneblast (gene content/synteny)



Key point:

ICEberg detects ICE-like regions but does not always identify the ICE family.

👉 Extra "detective work" is needed — a good next step is to run **Multigeneblast** for similarity to known ICEs present in ICEberg database ("ICE Browse")

https://tool2-mml.sjtu.edu.cn/ICEberg3/

ICEberg 3.0: Analysis of Results

- **Rows** = known ICEs; **colored arrows** = conserved genes
- ➢ **Total score =** weighted similarity of the ICE clusters (coverage + gene order + similarity)
- ➢ **Cumulative BLAST bit score** = sum of all gene-to-gene matches (raw sequence similarity)
- Predicted ICE matches several *V. cholerae* ICEs

Key point: ICEberg suggests related ICEs but may miss full length

👉 Confirm family/boundaries by aligning the full ICE (e.g. with progressiveMauve)

https://tool2-mml.sjtu.edu.cn/ICEberg3/

In summary

List of learning points in this session:

1. Genome annotation tools help detect key antimicrobial resistance and virulence traits.
2. ICEfinder detects ICE-like elements, but family identification may require sequence alignment.
3. IslandViewer predicts where genomic islands are, but it does not name families.
4. In *Vibrio*, identification of known GIs (e.g., VPI-1, VSP-2, SXT) requires comparison with literature-curated references and alignment tools (e.g., BLAST, progressiveMauve).

Acknowledgements

The creation of this training material was commissioned by ECDC to Technical University of Denmark with the direct involvement of Joana Mourão (Postdoc at DTU, PharmD, PhD – joam@food.dtu.dk)