

Automatic annotation platforms

MicroScope

<https://mage.genoscope.cns.fr/microscope/home/index.php>

- Automatic syntactic and functional annotation pipelines
- Metabolic analysis
- Visualization

The screenshot shows the MicroScope web application interface. At the top, there is a navigation bar with the MicroScope logo on the left, a user login status "Logged as Pauline FRANCOIS" with a "Logout" button, and a menu with the following items: MaGe, Genomic Tools, Comparative Genomics, Metabolism, Experimental Data, Search/Export, User Panel, and About. Below the navigation bar, the main content area features the MicroScope logo and the text "MicroScope Microbial Genome Annotation & Analysis Platform". To the right of the logo is a QR code and a text box stating: "The Quality Management System of the LABGeM team has been certified according to the ISO 9001:2015 and NF X50-900:2016 standards (Lloyd's Register Quality Assurance France S.A.S.). The certification applies to LABGeM activities of research, developments, services and MicroScope trainings." Below this is a search bar with the placeholder text "Find a genome among 7778". At the bottom, there is a list of five frequently asked questions (FAQ) items, each with an icon and a question mark:

- ? What is the MicroScope platform?
- 🧪 What type of projects can be handled in MicroScope?
- 📄 How to integrate your own data into MicroScope?
- 📝 Sign up for an account?
- 📖 How to cite us?

IMG

<https://img.jgi.doe.gov/>

Microbial and microbiome datasets (and metadata) sequenced at DOE Joint Genome Institute

Compare with public datasets.

Microbial/microbiome/virus/plasmids (bêta) browse by :

- geolocation
- ecosystem
- sequence length
- ...

JGI **IMG/M** **Integrated Microbial Genomes & Microbiomes**

JGI HOME CONTACT US LOGIN / SIGN-ON

IMG/M Integrated Microbial Genomes & Microbiomes

Bridging the Gap from Sequence to Biology

T G A C

Learn About IMG Explore Find & Analyze Save & Manage

Our Mission
IMG/M system aims to support the annotation, analysis, and distribution of microbial genome and microbiome datasets sequenced at DOE Joint Genome Institute. It also captures the diversity of public datasets to enable comprehensive comparative analysis and to meet the needs of our diverse community of users.

IMG Virus v4

BV-BRC

<https://www.bv-brc.org/>

Automated analysis for bacteria and virus

Genomics, phylogenetics, proteins, metagenomics and transcriptomics

The screenshot displays the BV-BRC website interface. At the top, a dark blue navigation bar contains the BV-BRC logo (version 3.36.16.3) and menu items: ORGANISMS, SEARCHES, TOOLS & SERVICES, WORKSPACES, HELP, and ABOUT. Below the navigation bar, the main heading reads "BACTERIAL AND VIRAL BIOINFORMATICS RESOURCE CENTER". A welcome message follows, describing the center's purpose and providing a link to a "Quick Start Guide". The interface features a search bar with a dropdown menu set to "All Data Types" and a search button. Below the search bar, a "BROWSE" section contains four buttons: BACTERIA, ARCHAEA, VIRUSES, and EUKARYOTIC HOSTS. The "ANALYZE DATA IN BV-BRC" section includes a sub-heading "Upload and analyze your data in the private workspace." and a grid of options categorized into SEARCH, ANALYZE, MANAGE DATA, and BATCH ACCESS. At the bottom, three buttons are visible: QUICK START, REFERENCE GUIDES, and TUTORIALS.

BACTERIAL AND VIRAL BIOINFORMATICS RESOURCE CENTER

Welcome to the Bacterial and Viral Bioinformatics Resource Center (BV-BRC), an information system designed to support research on bacterial and viral infectious diseases. The BV-BRC combines the data and tools from the Legacy BRC resources: PATRIC, the bacterial BRC, and IRD and ViPR, the viral BRCs. If you are transitioning from PATRIC or IRD/ViPR, please refer to the **Quick Start Guide** to learn how to get started with BV-BRC. Please **contact us** if you have any questions or issues.

SEARCH All Data Types *Find a gene, genome, microarray, etc* Q All terms

BROWSE BACTERIA ARCHAEA VIRUSES EUKARYOTIC HOSTS

ANALYZE DATA IN BV-BRC

Upload and analyze your data in the private workspace.

SEARCH	ANALYZE	MANAGE DATA	BATCH ACCESS
Taxa	Assembly	Access Private Workspace	Command-Line Interface
Genomes	Annotation	Save Search Results	Data API
Proteins	BLAST	Upload Data	FTP
Specialty Genes	MSA	Access Analysis Jobs	
All Searches...	All Tools & Services...	Share and Publish	

QUICK START REFERENCE GUIDES TUTORIALS

BV-BRC

Genomics

Genome Assembly
Genome Annotation
Comprehensive Genome Analysis (B)
BLAST
Primer Design
Similar Genome Finder
Genome Alignment
Variation Analysis
Tn-Seq Analysis

Phylogenomics

Bacterial Genome Tree
Viral Genome Tree
Gene/Protein Tree

Protein Tools

MSA and SNP Analysis
Meta-CATS
Proteome Comparison
Protein Family Sorter
Comparative Systems

Metagenomics

Taxonomic Classification
Metagenomic Binning
Metagenomic Read Mapping

Transcriptomics

RNA-Seq Analysis
Expression Import

Utilities

Fastq Utilities
ID Mapper

Viral Tools

SARS-CoV-2 Genome Analysis
SARS-CoV-2 Wastewater Analysis
Influenza Sequence Submission
Influenza HA Subtype Conversion
Subspecies Classification

Outbreak Tracker

Influenza H5N1 2024
SARS-CoV-2

CGE

<http://www.genomicepidemiology.org/services/>

Access to bioinformatics resources :

- Phenotyping
- Typing
- Phylogeny
- Metagenomics
- PCR-tools
- Other typer

Phenotyping

[ResFinder](#)

Identification of acquired antibiotic resistance genes.

[ResFinderFG](#)

Identification of functional metagenomic antibiotic resistance determinants.

[LRE-finder](#)

Identification of genes and mutations leading to linezolid resistance.

[KmerResistance](#)

Identification of acquired antibiotic resistance genes using Kmers.

[PathogenFinder](#)

Prediction of a bacteria's pathogenicity towards human hosts.

[VirulenceFinder](#)

Identification of acquired virulence genes.

[Restriction-ModificationFinder](#)

Determination of Restriction-Modification sites (based on REBASE.)

[SPIFinder](#)

SPIFinder identifies Salmonella Pathogenicity Islands.

[ToxFinder](#)

ToxFinder identifies genes involved in mycotoxin synthesis.

Phylogeny

[MINTyper](#)

Identification of SNPs with automatic filtering, masking and site validation together with inferred phylogeny based on both long and short sequencing data.

[CSIPhylogeny](#)

CSI Phylogeny calls SNPs, filters the SNPs, does site validation and infers a phylogeny based on the concatenated alignment of the high quality* SNPs.

[NDtree](#)

NDtree constructs phylogenetic trees from Single-End or Pair-End FASTQ files.

[Evergreen](#)

Evergreen generates a forest of constantly updated phylogenetic trees with publicly available whole-genome sequencing data from foodborne, bacterial isolates that were deposited in the short sequencing read archives (NCBI SRA/ENA).

[TreeViewer](#)

Phylogeny Tree Viewer.

Metagenomics

[CCMetagen](#)

CCMetagen: Comprehensive and accurate identification of eukaryotes

Specialized annotations

How it works?

± Blast on specialized and curated databases

How it works?

± Blast on specialized and curated databases

Results in tables with homology/identity and overlap/coverage informations

(Homology/identity)

Sequence similarity / percentage of identical nucleotides (or amino acid)

A	T	C	G	T	T	A	T	G	C	Reference
A	T	C	C	T	A	A	T	G	C	Our sequence

	Coverage/Overlap	Identity/homology
	Sequence length / Reference length	Identical bases / Reference length
As base	10/10	8/10
As percent	100%	80%

(Overlap/coverage)

Percentage of reference bases covered by our sequence

A	T	C	G	T	T	A	T	G	C	Reference
A	T	C	G	T	T	A	T			Our sequence

	Coverage/Overlap	Identity/homology
	Sequence length / Reference length	Identical bases / Reference length
As base	8/10	8/8
As percent	80%	100%

Useful databases

Many databases with their advantages and specificities

NCBI

Resfinder

VFDB

Argannot

Ecoli_vf

Card

Megares

EcoH

Plasmidfinder

Pointfinder

PubMLST

Galaxy tools - StarAMR

Querying databases : ResFinder, PlasmidFinder and PointFinder (if exists)
+ ST determination with PubMLST

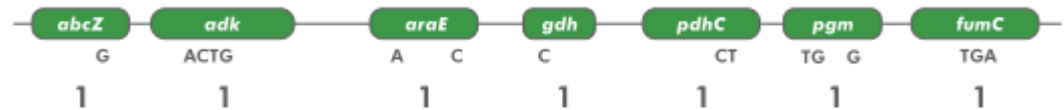
(MultiLocus Sequencing Typing – MLST)

Search allele of 7 housekeeping gene

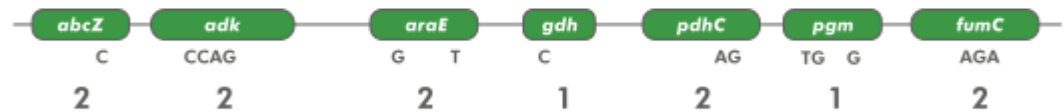
Allelic version combinaison determine the Sequence Type (ST)



Strain A



Strain B



Strain C










<https://www.applied-maths.com/applications/mlst>

(MultiLocus Sequencing Typing – MLST)

Each species have at least 1 scheme. Example Salmonella spp :

MLST

Locus	Download	Type	Alleles	Length (setting)	Min length	Max length	Full name/product	Last updated
aroC		DNA	1622	Variable: No limits set	480	524		2024-05-28
dnaN		DNA	1393	Variable: No limits set	454	537		2024-05-29
hemD		DNA	1332	Variable: No limits set	423	504		2024-05-27
hisD		DNA	1991	Variable: No limits set	482	515		2024-05-31
purE		DNA	1567	Variable: No limits set	315	429		2024-05-29
sucA		DNA	1450	Variable: No limits set	441	510		2024-05-31
thrA		DNA	1683	Variable: No limits set	493	579		2024-05-30

<https://pubmlst.org/>

Galaxy tools - StarAMR

Querying databases : ResFinder, PlasmidFinder and PointFinder (if exists)
+ ST determination with PubMLST

	A	B	C	D	E	F	G	H
	Isolate ID	Quality Module	Genotype	Predicted Phenotype	CGE Predicted Phenotype	Plasmid	Scheme	Sequence Type
1	GCF_000011885.1_ASM1188v1_genomic.fna	Passed	parC (T57S)	None	Nalidixic acid, Ciprofloxacin	None	senterica_achtman_2	85
2	GCF_000018625.1_ASM1862v1_genomic.fna	Passed	None	Susceptible	-	None	senterica_achtman_2	2131
3	GCF_000026565.1_ASM2656v1_genomic.fna	Passed	parC (T57S)	None	Nalidixic acid, Ciprofloxacin	None	senterica_achtman_2	85
4	GCF_000484015.1_ASM48401v1_genomic.fna	Passed	None	Susceptible	-	None	senterica_achtman_2	13
5	GCF_000756465.1_ASM75646v1_genomic.fna	Passed	None	Susceptible	-	None	senterica_achtman_2	2402
6	GCF_000983595.1_PA083_genomic.fna	Passed	parC (T57S)	None	Nalidixic acid, Ciprofloxacin	None	senterica_achtman_2	85
7	GCF_001448925.1_ASM144892v1_genomic.fna	Passed	blaTEM-1B	ampicillin	Amoxicillin, Ampicillin, Cephalothin, Piperacillin, Ticarcillin	None	senterica_achtman_2	2402
8	GCF_001951465.1_ASM195146v1_genomic.fna	Passed	aac(3)-VIa, ant(3'')-Ia, qacE, sul1	gentamicin, spectinomycin, unknown[qacE_1_X68232], sulfisoxazole	Gentamicin, Streptomycin, Benzylkonium Chloride, Ethidium Bromide, Chlorhexidine, Cetylpyridinium Chloride, Sulfamethoxazole	Incl1-I(Alpha)	senterica_achtman_2	869
9								
10								
37								
20								

Summary Detailed_Summary ResFinder PointFinder PlasmidFinder MLST_Summary Settings +

Prêt

Galaxy tools - Abricate

Can querying many different databases one by one (resistance, virulence, plasmids)

DB to use in option (default : resfinder)

Good supplement to StarAMR

Galaxy tools - Abricate

Can querying many different databases one by one (resistance, virulence, plasmids)

#FILE	SEQUENCE	START	END	STRAND	GENE	COVERAGE	COVERAGE_I	GAPS	%COVERAGE	%IDENTITY	DATABASE	ACCESSION	PRODUCT	RESISTANCE
GCF_0019514	NZ_AWRP01	71319	72286	+	ant(3'')-Ia_1	1-972/972	=====	/== 2/4	99.59	99.38	resfinder	X02340	ant(3'')-Ia	Streptomycin
GCF_0019514	NZ_AWRP01	72450	73352	+	aac(3)-VIa_2	1-903/903	=====	0/0	100.00	100.00	resfinder	NC_009838	aac(3)-VIa	Gentamicin
GCF_0019514	NZ_AWRP01	79325	80191	+	sul1_5	1-867/867	=====	0/0	100.00	99.89	resfinder	EU780013	sul1	Sulfamethoxazole

#FILE	SEQUENCE	START	END	STRAND
GCF_0019514	NZ_AWRP01	71319	72286	+
GCF_0019514	NZ_AWRP01	72450	73352	+
GCF_0019514	NZ_AWRP01	79325	80191	+

Galaxy tools - Abricate

Can querying many different databases one by one (resistance, virulence, plasmids)

#FILE	SEQUENCE	START	END	STRAND	GENE	COVERAGE	COVERAGE_I	GAPS	%COVERAGE	%IDENTITY	DATABASE	ACCESSION	PRODUCT	RESISTANCE
GCF_0019514	NZ_AWRP01	71319	72286	+	ant(3'')-Ia_1	1-972/972	=====	/== 2/4	99.59	99.38	resfinder	X02340	ant(3'')-Ia	Streptomycin
GCF_0019514	NZ_AWRP01	72450	73352	+	aac(3)-VIa_2	1-903/903	=====	0/0	100.00	100.00	resfinder	NC_009838	aac(3)-VIa	Gentamicin
GCF_0019514	NZ_AWRP01	79325	80191	+	sul1_5	1-867/867	=====	0/0	100.00	99.89	resfinder	EU780013	sul1	Sulfamethoxazole

GENE	COVERAGE	COVERAGE_I	GAPS	%COVERAGE	%IDENTITY	DATABASE	ACCESSION	PRODUCT	RESISTANCE
ant(3'')-Ia_1	1-972/972	=====	/== 2/4	99.59	99.38	resfinder	X02340	ant(3'')-Ia	Streptomycin
aac(3)-VIa_2	1-903/903	=====	0/0	100.00	100.00	resfinder	NC_009838	aac(3)-VIa	Gentamicin
sul1_5	1-867/867	=====	0/0	100.00	99.89	resfinder	EU780013	sul1	Sulfamethoxazole

Galaxy tools - Abricate

Can summaries multiple output : Each strain with its identified genes with their %identity

Files and number of genes found

#FILE	NUM_FOUND	aac(3)-VIa_2	aac(6)-Iaa_1	ant(3)-Ia_1	blaTEM-1B_1	fosA7_1	sul1_5
ABRicate_on_data_2_report_file	3	100.00	.	99.59	.	.	100.00
ABRicate_on_data_3_report_file	1	.	100.00
ABRicate_on_data_4_report_file	0
ABRicate_on_data_5_report_file	1	.	100.00
ABRicate_on_data_6_report_file	2	.	100.00	.	.	100.00	.
ABRicate_on_data_7_report_file	0
ABRicate_on_data_8_report_file	1	.	100.00
ABRicate_on_data_9_report_file	1	.	.	.	100.00	.	.

Galaxy tools - ABRicate

Can summaries multiple output : Each strain with its identified genes with their %identity

#FILE	NUM_FOUND	Homology per gene					
		aac(3)-VIa_2	aac(6)-Iaa_1	ant(3)-Ia_1	blaTEM-1B_1	fosA7_1	sul1_5
ABRicate_on_data_2_report_file	3	100.00	.	99.59	.	.	100.00
ABRicate_on_data_3_report_file	1	.	100.00
ABRicate_on_data_4_report_file	0
ABRicate_on_data_5_report_file	1	.	100.00
ABRicate_on_data_6_report_file	2	.	100.00	.	.	100.00	.
ABRicate_on_data_7_report_file	0
ABRicate_on_data_8_report_file	1	.	100.00
ABRicate_on_data_9_report_file	1	.	.	.	100.00	.	.

Let's try this

- Import « [data salmonella aa2024](#) » history
- Run StarAMR with Pointfinder on Salmonella
- Run Abricate on VFDB
- Summarize results from Abricate