

ARKbase HelpFile

<https://datascience.imtech.res.in/anshu/arkbase/>

ARKbase is an integrated, curated, value-added knowledgebase developed with the objective of providing a dedicated resource for deep annotation of bacterial priority pathogens (BPP) identified by WHO (<https://www.who.int/publications/i/item/9789240093461>).

ARKbase 1.0
Antimicrobial Resistance Knowledgebase

Home ▾ Pan AMR ▾ Pan Virulence ▾ ARG Structure ▾ PPI HPI Drug Targets ▾ Small Molecule ▾ DTI ▾ Genome Browser ▾ Search ▾

ARKbase: Antimicrobial Resistance Knowledgebase

ARKbase is an integrated, curated, value-added knowledge-base for AMR, with focus on WHO Bacterial Priority Pathogens. ARKbase is a dedicated AMR resource with a potential to provide novel insights towards expanding the drug-target space.

ARKbase is the largest dedicated resource for WHO bacterial priority pathogens, integrating high-quality genomes, proteomes, AMR genes, and comprehensive datasets including biosynthetic gene clusters, drug targets, antibiotic profiles, gene expression data and machine learning models. This centralized platform assists in understanding resistance mechanisms and supports targeted drug discovery for AMR.

Operons • Curated Genome AST • Expression Data • ML Models • Data Summary • Co-Target • Connect • FAQs & Help

Modules in ARKbase

ARKbase offers three core modules—

- Database module
- Insight module
- Comparative Analysis module


Database Module: This module is divided into the reference genome and the curated genome AST.

(i) Reference genomes

1. *Acinetobacter baumannii*
2. *Klebsiella pneumoniae*
3. *Escherichia coli*
4. *Shigella flexneri*
5. *Shigella sonnei*
6. *Enterococcus faecium*
7. *Pseudomonas aeruginosa*
8. *Salmonella enterica*
9. *Neisseria gonorrhoeae*
10. *Staphylococcus aureus*
11. *Streptococcus pyogenes*
12. *Streptococcus pneumoniae*
13. *Haemophilus influenzae*
14. *Streptococcus agalactiae*


Critical Priority Pathogens			
CRITICAL <i>Acinetobacter baumannii</i> Carbapenem-resistant	CRITICAL <i>Klebsiella pneumoniae</i> Carbapenem & 3rd-Gen Cephalosporin-resistant	CRITICAL <i>Escherichia coli</i> Carbapenem & 3rd-Gen Cephalosporin-resistant	
High Priority Pathogens			
HIGH <i>Shigella flexneri</i> Fluoroquinolone-resistant	HIGH <i>Shigella sonnei</i> Fluoroquinolone-resistant	HIGH <i>Enterococcus faecium</i> Vancomycin-resistant	HIGH <i>Pseudomonas aeruginosa</i> Carbapenem-resistant
HIGH <i>Salmonella enterica</i> Fluoroquinolone-resistant	HIGH <i>Neisseria gonorrhoeae</i> Cephalosporin & Fluoroquinolone-resistant	HIGH <i>Staphylococcus aureus</i> Methicillin-resistant	
Medium Priority Pathogens			
MEDIUM <i>Streptococcus pyogenes</i> Macrolide-resistant	MEDIUM <i>Streptococcus pneumoniae</i> Macrolide-resistant	MEDIUM <i>Haemophilus influenzae</i> Ampicillin-resistant	MEDIUM <i>Streptococcus agalactiae</i> Penicillin-resistant

The reference genome page provides pathogen overview, genome overview, genomic representation, COG classification, pangenome distribution, drug resistance profile, ARGs and mechanisms, and protein-protein interactions (PPI).



ARKbase^{1.0}

Antimicrobial Resistance Knowledgebase



[Home](#) ▾ [Pan AMR](#) ▾ [Pan Virulence](#) ▾ [ARG Structure](#) ▾ [PPI](#) [HPI](#) [Drug Targets](#) ▾ [Small Molecule](#) ▾ [DTI](#) ▾ [Genome Browser](#) ▾ [Search](#) ▾

Acinetobacter baumannii

Overview

Operon

ARGs

ARGs Structure

Virulence Factor

Protein-Protein Interaction

Drug Targets

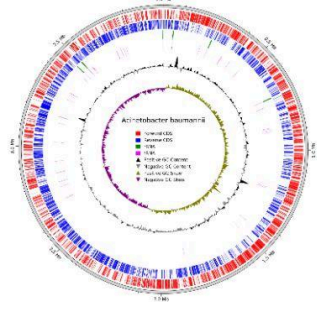
Pathogen Overview

Organism Name: *Acinetobacter baumannii* strain ATCC 19606
 Gram Stain: Gram negative
 Taxon ID: 470
 Country of Isolation: USA
 Year of Collection: 1948

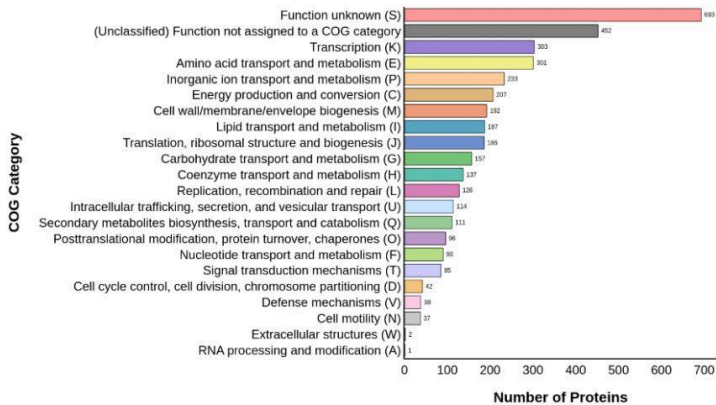
Genome Summary

Assembly Accession: [GCF_009035845.1](#)
 Genome Size: 4 Mb
 GC Content: 57.5%
 Total Genes: 3,833
 Protein Coding Genes: 3,683
 tRNA: 74
 rRNA: 18

Genomic Representation

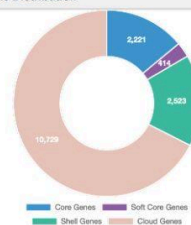


COG Classification

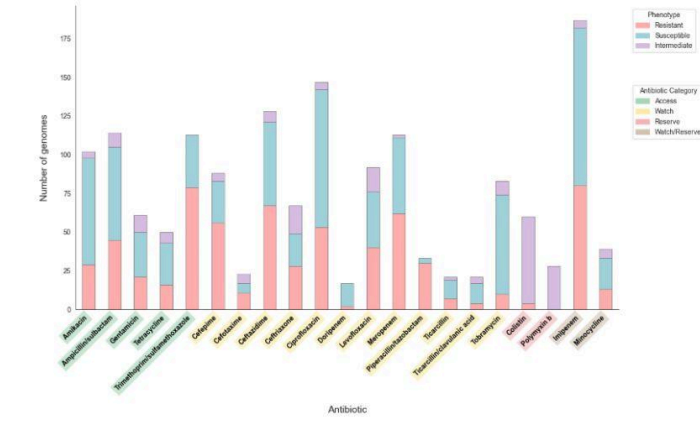


COG Category	Number of Proteins
Function unknown (S)	683
(Unclassified) Function not assigned to a COG category	452
Transcription (K)	303
Amino acid transport and metabolism (E)	301
Inorganic ion transport and metabolism (P)	293
Energy production and conversion (C)	207
Cell wall/membrane/envelope biogenesis (M)	182
Lipid transport and metabolism (I)	187
Translation, ribosomal structure and biogenesis (J)	248
Carbohydrate transport and metabolism (G)	157
Coenzyme transport and metabolism (H)	137
Replication, recombination and repair (L)	126
Intracellular trafficking, secretion, and vesicular transport (U)	114
Secondary metabolites biosynthesis, transport and catabolism (Q)	111
Posttranslational modification, protein turnover, chaperones (O)	96
Nucleotide transport and metabolism (F)	90
Signal transduction mechanisms (T)	85
Cell cycle control, cell division, chromosome partitioning (D)	42
Defense mechanisms (V)	39
Cell motility (N)	27
Extracellular structures (W)	3
RNA processing and modification (A)	1

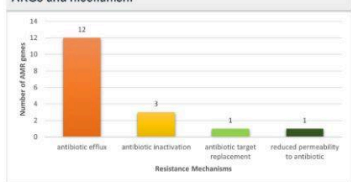
Pan Genome Distribution



Drug Resistance Profile



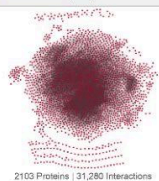
ARGs and mechanism



Operons

[Curated Genome AST](#) [Expression Data](#) [ML Models](#) [Data Summary](#) [Co-Target](#) [Connect](#) [FAQs & Help](#)

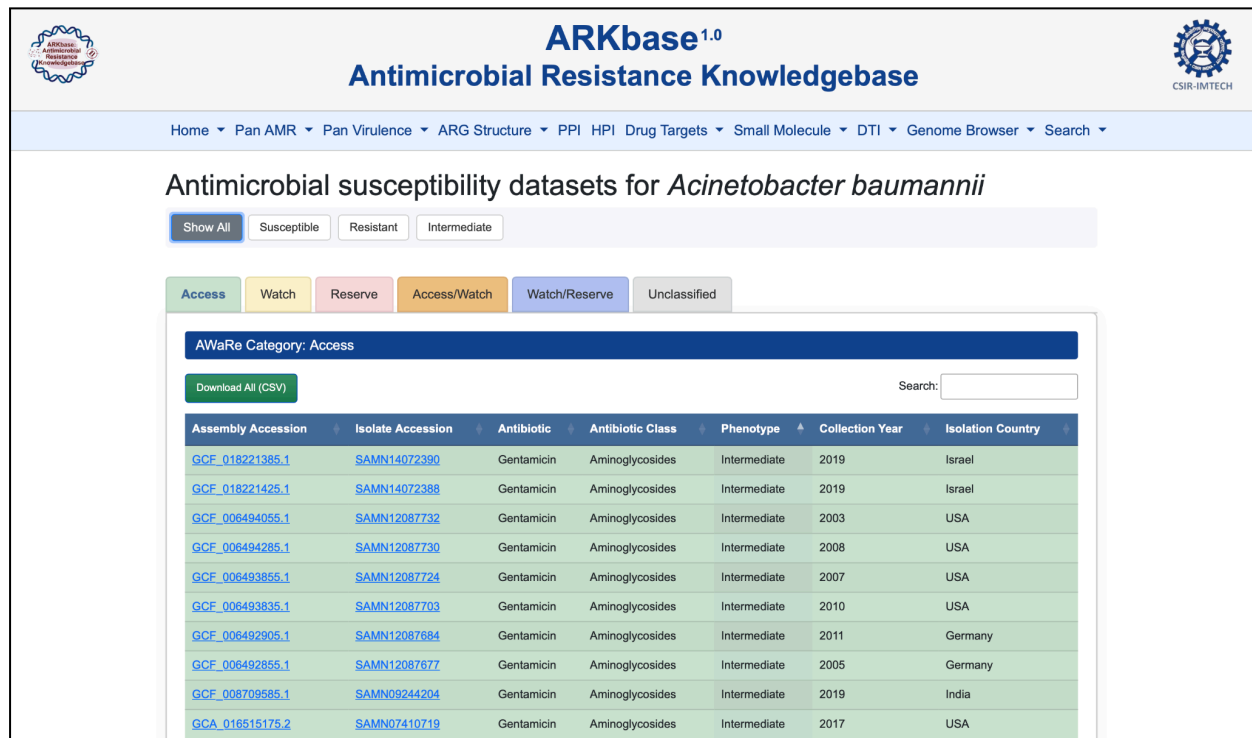
Protein Protein Interactions (PPI)



© 2025 ARKbase^{1.0} | CSIR-IMTECH
 All data are freely accessible to all users, including commercial users.
 This website doesn't use any cookies.

(ii) Curated Genome with Antimicrobial susceptibility (AST)

Genome datasets are obtained from publicly available resources and then quality checked primarily on two criteria - The MICs reported for these strains are quality checked against CLSI breakpoints and the corresponding genome datasets were checked as per EUCAST criteria of good quality genomes. More details can be read at <https://pubmed.ncbi.nlm.nih.gov/36817105/>



The screenshot displays the ARKbase 1.0 Antimicrobial Resistance Knowledgebase interface. The main heading is "Antimicrobial susceptibility datasets for *Acinetobacter baumannii*". Below the heading, there are filter buttons for "Show All", "Susceptible", "Resistant", and "Intermediate". Further down, there are buttons for "Access", "Watch", "Reserve", "Access/Watch", "Watch/Reserve", and "Unclassified". The "Access" button is selected, and the "AWaRe Category: Access" is displayed. A "Download All (CSV)" button and a search box are also present. The main data is presented in a table with the following columns: Assembly Accession, Isolate Accession, Antibiotic, Antibiotic Class, Phenotype, Collection Year, and Isolation Country. The table contains 10 rows of data, all for Gentamicin, with various isolate accessions and collection years from 2003 to 2019, and isolation countries including Israel, USA, Germany, India, and USA.

Assembly Accession	Isolate Accession	Antibiotic	Antibiotic Class	Phenotype	Collection Year	Isolation Country
GCF_018221385.1	SAMN14072390	Gentamicin	Aminoglycosides	Intermediate	2019	Israel
GCF_018221425.1	SAMN14072388	Gentamicin	Aminoglycosides	Intermediate	2019	Israel
GCF_006494055.1	SAMN12087732	Gentamicin	Aminoglycosides	Intermediate	2003	USA
GCF_006494285.1	SAMN12087730	Gentamicin	Aminoglycosides	Intermediate	2008	USA
GCF_006493855.1	SAMN12087724	Gentamicin	Aminoglycosides	Intermediate	2007	USA
GCF_006493835.1	SAMN12087703	Gentamicin	Aminoglycosides	Intermediate	2010	USA
GCF_006492905.1	SAMN12087684	Gentamicin	Aminoglycosides	Intermediate	2011	Germany
GCF_006492855.1	SAMN12087677	Gentamicin	Aminoglycosides	Intermediate	2005	Germany
GCF_008709585.1	SAMN09244204	Gentamicin	Aminoglycosides	Intermediate	2019	India
GCA_016515175.2	SAMN07410719	Gentamicin	Aminoglycosides	Intermediate	2017	USA

Insight Modules

Insight module is a composite module and is further divided into 14 modules as follows:

1. Pan AMR
2. Pan Virulence
3. ARG structure
4. Protein-Protein Interactions
5. Host-Pathogen Interactions
6. Drug Targets
7. Small Molecules
8. Drug-Target Interactions
9. Operons
10. Expression data
11. Machine Learning Models
12. Co-Target
13. Genome Browser
14. Search

ARKbase^{1.0}
Antimicrobial Resistance Knowledgebase

Home ▾ Pan AMR ▾ Pan Virulence ▾ ARG Structure ▾ PPI HPI Drug Targets ▾ Small Molecule ▾ DTI ▾ Genome Browser ▾ Search ▾

About ARKbase
Database Module ▾
Insight Module ▾
Comparative Analysis Module
FAQs & Help

WHO Bacterial Priority Pathogens
Reference Data from NCBI dist

Genome Curation & Deep Annotation of WHO Bacterial Priority Pathogens

Medium group

Genome and Susceptibility data
High-Quality Genome Filtering Criteria

- Clinical Isolates**
Human-origin samples collected from publicly available databases.
- CLSI Classification**
Classification based on Clinical and Laboratory Standards Institute.
- EUGAST Quality Check**
Quality control using European Committee standards.
- Assembly Quality**
Fewer contigs indicate a higher quality assembly.
- Completeness Metrics**
Assessing the completeness and quality of the assembly.

Comprehensive Genome Annotation

- Genomic Annotation**
Identifies gene features like CDS, tRNA, rRNA, and regulatory elements.
- Structure-Based Annotation**
Predicts protein function through 3D structure analysis.
- Operon Detection**
Provides insights into gene regulation and transcriptional units.
- Virulence Factor Detection**
Identifies factors contributing to pathogen virulence.
- Pangenome Analysis**
Reveals core and accessory gene content for comparative genomics.
- BGC Prediction**
Identifies potential for secondary metabolite production.
- AMR Gene Identification**
Detects genes related to antimicrobial resistance.
- Interactome Analysis**
Identifies high central interacting pairs (co-Targets).

ARKbase is an integrated, curated, value-added knowledge-base for AMR, with focus on WHO Bacterial Priority Pathogens. ARKbase is a dedicated AMR resource with a potential to provide novel insights towards expanding the drug-target space.

ARKbase is the largest dedicated resource for WHO bacterial priority pathogens, integrating high-quality genomes, proteomes, AMR genes, and comprehensive datasets including biosynthetic gene clusters, drug targets, antibiotic profiles, gene expression data and machine learning models. This centralized platform assists in understanding resistance mechanisms and supports targeted drug discovery for AMR.

Pan AMR:

This module contains pathogen-specific ARGs, divided into two pages: reference and curated genomes.

The screenshot shows the ARKbase 1.0 interface. The 'Pan AMR' dropdown menu is open, listing three pathogens: *Acinetobacter baumannii*, *Klebsiella pneumoniae*, and *Escherichia coli*. Below the menu, there are buttons for 'Reference Genome' and 'Curated Genome'. The main dashboard features a central 'ARKbase: Antimicrobial Resistance Knowledgebase' logo surrounded by various modules: 'Restriction maps', 'Pan-genome', 'Reference', 'Database Module', 'AMR', 'Virulence', 'Drug Target', 'Analysis & Visualization', 'Comparative Analysis Module', 'Machine learning Models', 'DTIs', 'PPIs', 'Operon', 'Essentiality', 'HPIs', 'Small Molecules', 'Expression datasets', and 'AMR'. At the bottom, there are four status bars: 'AST as per CLSI', 'Genome quality by EUCAST', 'AWaRe classification', and 'FAIR'.

For the reference page, there are five sections: outputs from four different tools, RGI, AMRFinderPlus, MEGARes, and ResFinder, along with a combined tool output. The user can download the data.

ARGs for Reference [Download](#)

All RGI AMRFinderPlus MEGARes ResFinder

Showing 1 - 25 of 28 results.

Pathogen	Assembly_Accession	Protein_ID	Description	Contributing_Tools	Contributing_Subject_IDs
Acinetobacter baumannii	GCF_009035845.1	WP_000010647.1	multidrug efflux RND transporter periplasmic adaptor subunit AdeF [Acinetobacter baumannii]	Megares,RGI	3000778,MEG_3910
Acinetobacter baumannii	GCF_009035845.1	WP_000046678.1	MULTISPECIES: multidrug efflux RND transporter permease subunit AdeJ [Acinetobacter]	Megares,RGI	3000781,MEG_692
Acinetobacter baumannii	GCF_009035845.1	WP_000073361.1	MULTISPECIES: multidrug efflux MATE transporter AbeM [Acinetobacter]	Megares,RGI	3000753,MEG_373
Acinetobacter baumannii	GCF_009035845.1	WP_000121132.1	multidrug efflux SMR transporter AbeS [Acinetobacter baumannii]	Megares,RGI	3000768,MEG_378
Acinetobacter baumannii	GCF_009035845.1	WP_000155956.1	Class A betalactamases	Megares	MEG_1246
Acinetobacter baumannii	GCF_009035845.1	WP_000214022.1	MULTISPECIES: fosfomycin efflux MFS transporter AbaF [Acinetobacter]	Megares,RGI	3004573,MEG_369
Acinetobacter baumannii	GCF_009035845.1	WP_000345073.1	multidrug efflux MFS transporter AbaQ [Acinetobacter baumannii]	Megares,RGI	3004574,MEG_370
Acinetobacter baumannii	GCF_009035845.1	WP_000373596.1	Drug and biocide RND efflux pumps	Megares	MEG_705

For the curated page, there are five sections: outputs from four different tools—RGI, AMRFinderPlus, MEGARes, and ResFinder—along with a combined tool output, the same as the reference page. In addition, the user can filter the data using antibiotics and phenotypes.

ARGs for *Acinetobacter baumannii*

Filter by Antibiotic:

Filter by Phenotype:

All Antibiotics

All Phenotypes

Apply Filters

Download

All
RGI
AMRFinderPlus
MEGARes
ResFinder

Showing 1 - 25 of 30764 results.

#	Pathogen	Assembly Accession	Protein ID	Description	Contributing Tools	Contributing Subject IDs
1	Acinetobacter baumannii	GCA_008709595.1	KAA8938070.1	multidrug efflux RND transporter periplasmic adaptor subunit AdeF [Acinetobacter baumannii]	Megares,RGI	3000778,MEG_3910
2	Acinetobacter baumannii	GCA_008709595.1	KAA8927827.1	chloramphenicol efflux MFS transporter CmlA5 [Acinetobacter baumannii]	AMRFinderPlus,Megares,RGI,ResFinder	3002695,JQ639792,MEG_1766,WP_012300772.1
3	Acinetobacter baumannii	GCA_008709595.1	KAA8927835.1	ABC-F type ribosomal protection protein Msr(E) [Acinetobacter baumannii]	AMRFinderPlus,Megares,RGI,ResFinder	3003109,FR751518,MEG_4075,WP_000052512.1
4	Acinetobacter baumannii	GCA_008709595.1	KAA8927836.1	Mph(E) family macrolide 2'-phosphotransferase [Acinetobacter baumannii]	AMRFinderPlus,Megares,RGI,ResFinder	3003741,DQ839391,MEG_4044,WP_000155092.1
5	Acinetobacter baumannii	GCA_008709595.1	KAA8927840.1	ArmA family 16S rRNA (guanine(1405)-N(7))-methyltransferase [Acinetobacter baumannii]	AMRFinderPlus,Megares,RGI,ResFinder	3000858,AY220558,MEG_1120,WP_000359986.1
6	Acinetobacter baumannii	GCA_008709595.1	KAA8928255.1	multidrug efflux RND transporter permease subunit AdeB [Acinetobacter baumannii]	Megares,RGI	3000775,MEG_678

Pan Virulence:


The PanVirulence module contains pathogen-specific virulence, divided into two pages: reference and curated genomes.

This tab provides the user a consolidated table of virulence factors for each pathogen chosen from the drop-down menu. The user can explore the virulence genes with the corresponding protein id, gene name, virulence factor category, operon id, VF ID and the organism name where the particular gene has matched. The protein_ids are directly linked to the NCBI database. The page also allows the user to search specific fields and to download the tables.

The module includes an interactive filter to search across fields by protein ID, protein name, VF ID and so on. The download tab allows the user to download the complete virulence data for that particular genome.


Pathogen	Genome Accession	Protein ID	Resource Subject Identifier	Subject gene symbol	Subject Description	VF ID	VF Category	Subject Organism	Percent Identity	Query Coverage
Acinetobacter baumannii	GCF_009035845.1	WP_000007335.1	VFG013617(gb WP_200858241)	hemH	ferrochelataase	VF0758	Nutritional/Metabolic factor	Haemophilus somnus 129PT	44.410	95.00
Acinetobacter baumannii	GCF_009035845.1	WP_000007422.1	VFG013607(gb WP_041604533)	hemA	glutamyl-tRNA reductase	VF0758	Nutritional/Metabolic factor	Haemophilus somnus 129PT	40.909	94.00
Acinetobacter baumannii	GCF_009035845.1	WP_000010368.1	VFG002901(gb WP_002209736)	YPA_RS10790	histidine ABC transporter permease HisQ	VF0392	Immune modulation	Yersinia pestis Antiqua	62.389	99.00
Acinetobacter baumannii	GCF_009035845.1	WP_000010647.1	VFG037709(gb WP_000010642)	adeF	membrane-fusion protein	VF0504	Biofilm	Acinetobacter baumannii BJAB0715	99.507	100.00
Acinetobacter baumannii	GCF_009035845.1	WP_000011808.1	VFG037928(gb YP_001083151)	A1S_0054	WbbJ protein	VF0465	Immune modulation	Acinetobacter baumannii ATCC 17978	97.917	100.00
Acinetobacter baumannii	GCF_009035845.1	WP_000013436.1	VFG037807(gb WP_000013434)	lpxB	lipid-A-disaccharide synthase	VF0466	Immune modulation	Acinetobacter baumannii ACICU	99.744	100.00

For the curated page, the module includes an interactive filter to search across fields by protein ID, protein name, VF ID and so on. The download tab allows the user to download the complete virulence data for that particular genome.



ARKbase^{1.0}

Antimicrobial Resistance Knowledgebase



[Home](#) ▾ [Pan AMR](#) ▾ [Pan Virulence](#) ▾ [ARG Structure](#) ▾ [PPI HPI Drug Targets](#) ▾ [Small Molecule](#) ▾ [DTI](#) ▾ [Genome Browser](#) ▾ [Search](#) ▾

Virulence Data for *Acinetobacter baumannii*

Filter by VF Category:

All Categories ▾

[Apply Filter](#)

[Download](#)

Showing 1 - 25 of 205481 results.

Pathogen	Genome_Accession	Protein_ID	Resource_Subject_Identifier	Subject_gene_symbol	Subject_Description	VF_ID	VF_Category	Subject_Organism	Percent_Identity	Query_Coverage
Acinetobacter baumannii	GCA_008709595.1	KAA8927756.1	VFG039536(gb NP_820549)	CBU_1566	Coxiella Dot/Icm type IVB secretion system translocated effector	VF0696	Effector delivery system	Coxiella burnetii RSA 493	60.816	98.00
Acinetobacter baumannii	GCA_008709595.1	KAA8927772.1	VFG048381(gb WP_000820820)	mrkB	fimbrial chaperone protein mrkB precursor	VF0587	Biofilm	Klebsiella oxytoca E718	46.154	89.00
Acinetobacter baumannii	GCA_008709595.1	KAA8927773.1	VFG042535(gb AAA92618)	f17d-A	F17 fimbrial major subunit protein	VF0213	Adherence	Escherichia coli str. 111KH86	48.171	89.00
Acinetobacter baumannii	GCA_008709595.1	KAA8927776.1	VFG047506(gb WP_012281076)	bioB	biotin synthase	VF0552	Nutritional/Metabolic factor	Francisella philomiragia subsp. philomiragia ATCC 25017	50.980	92.00
Acinetobacter baumannii	GCA_008709595.1	KAA8927811.1	VFG013514(gb WP_011272587)	mrsAglmM	phosphoglucosamine mutase	VF0755	Immune modulation	Haemophilus influenzae 86-028NP	51.899	80.00
Acinetobacter baumannii	GCA_008709595.1	KAA8927817.1	VFG042734(gb NP_249343)	vfr	cAMP-regulatory protein	VF0082	Adherence	Pseudomonas aeruginosa PAO1	54.545	83.00

Antimicrobial Resistance Gene (ARG) Structure:

This module gives pathogen-wise structure of the antimicrobial resistance (AMR) gene of the WHO Bacterial Priority Pathogen List.

The user can search for any data field through the search interface. The user can view the consolidated table with AMR gene ID, description, AMR gene family, resistance mechanism, operon and their structure, antibiotic, and GO ID of that particular pathogen. The user can download the whole data in CSV/Excel format for analysis.

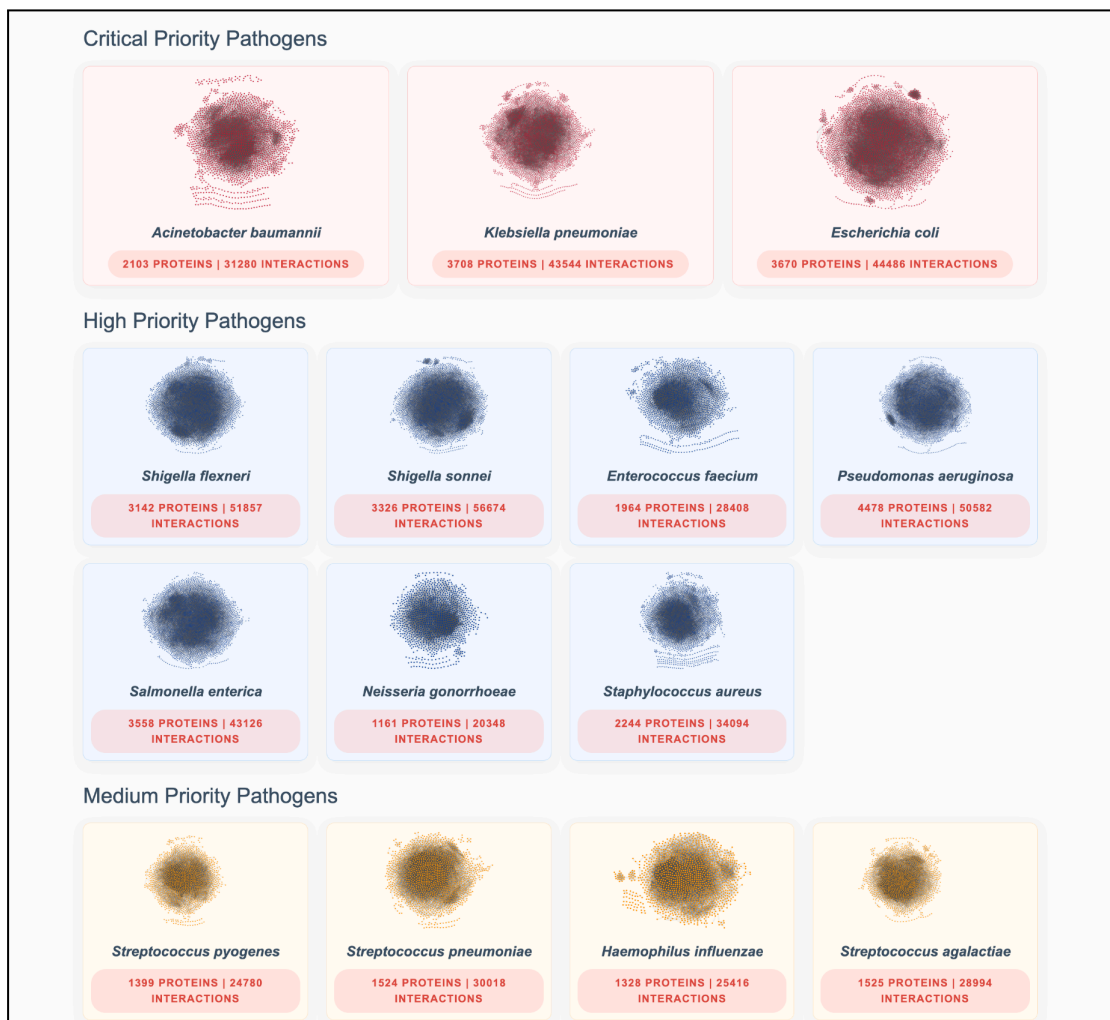
S.No.	Protein ID	Description	AMR Gene Family	Structure	Resistance Mechanism	Antibiotic	GO ID
1	WP_000446678.1	MULTISPECIES: multidrug efflux RND transporter permease subunit AdeJ [Acinetobacter]	resistance-nodulation-cell division (RND) antibiotic efflux pump	7M4P, 7M4Q, 7RY3	antibiotic efflux	tetracycline; rifampin; imipenem; trimethoprim; chloramphenicol; ticarcillin	GO:0009636 , GO:0055085 , GO:0042908 , GO:0016020 , GO:0022857 , GO:0055085 , GO:0015562 , GO:0016020 , GO:0042908 , GO:0005886 , GO:0042910 , GO:0005886
2	WP_00073361.1	MULTISPECIES: multidrug efflux MATE transporter AbeM [Acinetobacter]	multidrug and toxic compound extrusion (MATE) transporter	7php_A	antibiotic efflux	ciprofloxacin; acriflavine; norfloxacin; ofloxacin; triclosan	GO:0042908 , GO:0015297 , GO:0016020 , GO:0042910 , GO:0055085 , GO:0016020 , GO:0042910 , GO:0046677 , GO:0046677 , GO:0015297 , GO:0006811 , GO:0006814 , GO:0005886
3	WP_000214022.1	MULTISPECIES: fosfomycin efflux MFS transporter AbaF [Acinetobacter]	major facilitator superfamily (MFS) antibiotic efflux pump	4gby_A	antibiotic efflux	fosfomycin	GO:0016020 , GO:0055085 , GO:0022857 , GO:0016020 , GO:0022857 , GO:0055085 , GO:0016020 , GO:0022857 , GO:0055085 , GO:0022857 , GO:0015293 , GO:0005886 , GO:0015753 , GO:0005886 , GO:0015519 , GO:0015753 , GO:0005886
4	WP_000459547.1	efflux system response regulator transcription factor AdeR [Acinetobacter baumannii]	resistance-nodulation-cell division (RND) antibiotic efflux pump	7cci_A, 5x5i_E	antibiotic efflux	tigecycline; tetracycline	GO:0000160 , GO:0000976 , GO:0006355 , GO:0000160 , GO:0000976 , GO:0006355
5	WP_000867106.1	glycosyltransferase family 4 protein [Acinetobacter baumannii]	Intrinsic peptide antibiotic resistant Lps	7mi0_A	reduced permeability to antibiotic	colistin A; colistin B; defensin	GO:0016757 , GO:0016757 , GO:0016757 , GO:0016740

Protein-protein interactions (PPIs):

This module provides a protein-protein interactions landscape for BPPs.

The screenshot shows the ARKbase 1.0 website interface. The navigation bar includes 'Home', 'Pan AMR', 'Pan Virulence', 'ARG Structure', 'PPI', 'Drug Targets', 'Small Molecule', 'DTI', 'Genome Browser', and 'Search'. The 'PPI' menu item is highlighted with a red box. The main content area is titled 'Genome Curation & Deep Annotation of WHO Bacterial Priority Pathogens' and features several analysis tools: Genomic Annotation, Structure-Based Annotation, Virulence Factor Detection, Pangenome Analysis, BGC Prediction, and AMR Gene Identification. The page also displays WHO Bacterial Priority Pathogens categorized into Critical, High, and Medium groups.

This module consists of 14 PPIs of WHO BBP classified into three categories - Critical, High, and Medium Priority Pathogens, along with the number of proteins and interactions. The user can select any pathogen to explore centrality and co-targets.



This module is organised into two main sections: Centrality Metrics and Co-Target Interactions.

Interactome Centrality & Co-Target Discovery
Browse, filter, and analyze protein interactivity data.

Centrality Metrics | Co-Target Identification

Centrality Metrics
This table displays key centrality metrics for individual proteins. Select a pathogen to begin.

Pathogen: Acinetobacter baumannii | Assembly Accession: All | Protein ID: All | Protein Description: Type to search... | Protein Group (BW): All | Protein Group (Closeness): All | Protein Group (Degree): All | COG: All | COG Description: Type to search...

Showing 1 - 25 of 2,103 records | Per page: 25 | Download

S.No.	Pathogen	Assembly Accession	Protein ID	Protein Description	Protein Group (BW)	Protein Group (Closeness)	Protein Group (Degree)	COG	COG Description
1	<i>Acinetobacter baumannii</i>	GCF_009035845.1	WP_000672542.1	acetyl/propionyl/methylcrotonyl-CoA carboxylase subunit alpha	Q1	Q1	Q1	I	Lipid transport and metabolism
2	<i>Acinetobacter baumannii</i>	GCF_009035845.1	WP_001125860.1	MULTISPECIES: endolytic transglycosylase MITG	Q1	Q1	Q2	S	Function unknown
3	<i>Acinetobacter baumannii</i>	GCF_009035845.1	WP_000986451.1	MULTISPECIES: aspartate-tRNA ligase	Q1	Q1	Q1	J	Translation, ribosomal structure and biogenesis
4	<i>Acinetobacter baumannii</i>	GCF_009035845.1	WP_001983614.1	aspartate aminotransferase family protein	Q1	Q1	Q1	E	Amino acid transport and metabolism

1. Centrality Metrics

This section provides centrality metrics of each protein for fourteen pathogens. For each network, three centrality metrics are calculated: Betweenness centrality, Closeness centrality, and Degree centrality.

Each centrality metric is classified into four groups:

Q1 – Top 25% of nodes (highest centrality values)

Q2 – 25–50% range

Q3 – 50–75% range

Q4 – 75–100% range (lowest centrality values)

This classification helps in identifying highly influential nodes (Q1) versus peripheral nodes (Q4) for each pathogen's interactome.

The module includes nine interactive filters for targeted data exploration:

- Pathogen
- Assembly Accession
- Protein ID
- Protein Description
- Protein Groups (Betweenness)
- Protein Groups (Closeness)
- Protein Groups (Degree)
- COG (Clusters of Orthologous Groups)
- COG Description

The user can also download the data.

Interactome Centrality & Co-Target Discovery
Browse, filter, and analyze protein interactivity data.

Centrality Metrics **Co-Target Identification**

Co-Target Identification
This table shows interacting protein pairs. Select a pathogen, then click "View Details" to explore potential drug targets.

Pathogen: Select Pathogen First | All | Protein ID 1: All | Protein ID 2: All | Protein Description 1: Type to search... | Protein Description 2: Type to search... | Protein 1 Group (BW): All | Protein 2 Group (BW): All | COG 1: All | COG 2: All | COG Description 1: Type to search... | COG Description 2: Type to search... | Clear

Showing 1 - 25 of 513,607 records | Per page: 25 | Download

S.No.	Pathogen	Protein ID 1	Protein ID 2	Protein Description 1	Protein Description 2	Protein 1 Group (BW)	Protein 2 Group (BW)	COG 1	COG 2	COG Description 1	COG Description 2	More Information
1	Enterococcus faecium	WP_002295492.1	WP_002289912.1	MULTISPECIES: pantetheine-phosphate adenylyltransferase	MULTISPECIES: Holiday junction branch migration DNA helicase RuvB	Q1	Q1	H	L	Coenzyme transport and metabolism	Replication, recombination and repair	View Details
2	Enterococcus faecium	WP_002295135.1	WP_002289751.1	MULTISPECIES: phosphoglycerate kinase	MULTISPECIES: uridine kinase	Q1	Q1	F	F	Nucleotide transport and metabolism	Nucleotide transport and metabolism	View Details
3	Enterococcus faecium	WP_002286627.1	WP_002288809.1	MULTISPECIES: cell division protein SepF	MULTISPECIES: ParB/RepB/SpoU family partition protein	Q1	Q1	D	K	Cell cycle control, cell division, chromosome partitioning	Transcription	View Details
4	Enterococcus faecium	WP_002286806.1	WP_002286627.1	MULTISPECIES: ParB/RepB/SpoU family partition protein	MULTISPECIES: cell division protein SepF	Q1	Q1	K	D	Transcription	Cell cycle control, cell division, chromosome partitioning	View Details
5	Enterococcus faecium	WP_002287001.1	WP_002287361.1	MULTISPECIES: type I glutamate-ammonia ligase	MULTISPECIES: asparaginase	Q1	Q1	E	EJ	Amino acid transport and metabolism	Translation, ribosomal structure and biogenesis; Amino acid transport and metabolism	View Details
6	Enterococcus faecium	WP_002287361.1	WP_002287001.1	MULTISPECIES: asparaginase	MULTISPECIES: type I glutamate-ammonia ligase	Q1	Q1	EJ	E	Translation, ribosomal structure and biogenesis; Amino acid transport and metabolism	Amino acid transport and metabolism	View Details

2. Co-Target Discovery

This section highlights the interacting pairs for each pathogen. The module includes nine interactive filters for targeted data exploration: Pathogen, Protein ID1, Protein ID2, Protein Description 1, Protein Description 2, Protein Group 1(BW), Protein Group 2(BW), COG 1, COG 2, COG Description 1, and COG Description 2. Co-Targets are the co-interactors with drug target features.

Interactome Centrality & Co-Target Discovery
Browse, filter, and analyze protein interactivity data.

Centrality Metrics **Co-Target Identification**

Co-Target Identification
This table shows interacting protein pairs. Select a pathogen, then click "View Details" to explore potential drug targets.

Pathogen: Select Pathogen First | All | Protein ID 1: All | Protein ID 2: All | Protein Description 1: Type to search... | Protein Description 2: Type to search... | Protein 1 Group (BW): All | Protein 2 Group (BW): All | COG 1: All | COG 2: All | COG Description 1: Type to search... | COG Description 2: Type to search... | Clear

Showing 1 - 25 of 513,607 records | Per page: 25 | Download

S.No.	Pathogen	Protein ID 1	Protein ID 2	Protein Description 1	Protein Description 2	Protein 1 Group (BW)	Protein 2 Group (BW)	COG 1	COG 2	COG Description 1	COG Description 2	More Information
1	Enterococcus faecium	WP_002295492.1	WP_002289912.1	MULTISPECIES: pantetheine-phosphate adenylyltransferase	MULTISPECIES: Holiday junction branch migration DNA helicase RuvB	Q1	Q1	H	L	Coenzyme transport and metabolism	Replication, recombination and repair	View Details
2	Enterococcus faecium	WP_002295135.1	WP_002289751.1	MULTISPECIES: phosphoglycerate kinase	MULTISPECIES: uridine kinase	Q1	Q1	F	F	Nucleotide transport and metabolism	Nucleotide transport and metabolism	View Details
3	Enterococcus faecium	WP_002286627.1	WP_002288809.1	MULTISPECIES: cell division protein SepF	MULTISPECIES: ParB/RepB/SpoU family partition protein	Q1	Q1	D	K	Cell cycle control, cell division, chromosome partitioning	Transcription	View Details
4	Enterococcus faecium	WP_002286806.1	WP_002286627.1	MULTISPECIES: ParB/RepB/SpoU family partition protein	MULTISPECIES: cell division protein SepF	Q1	Q1	K	D	Transcription	Cell cycle control, cell division, chromosome partitioning	View Details

View details tab lists drug target features of each co-interactor as shown below.

Details for Protein 1: [WP_002295492.1](#)

Pathogen Name	<i>Enterococcus faecium</i>
Prot Id	WP_002295492.1
Prot Desc	MULTISPECIES: pantetheine-phosphate adenylyltransferase
Non Paralog	Yes
Virulence	Yes
Essential	Yes
Ttd Novel	Yes
Drugbank Novel	Yes
Human Non Homologous	Yes
Anti Target	No
Top 10 Central	Yes
Core	Yes
Not Amr	Yes

Host-Pathogen Interactions:

It provides an interactive data explorer for host-pathogen interaction. It includes an interaction score and experimental information, as well as the top 10 Gene Ontology (GO) terms. The user can search by UniProt identifiers for pathogen and host proteins. This module is divided into four sections:

1. Graphical abstract for the overall summary
2. Top 10 Pathogen Gene Ontology Terms
3. Host-Pathogen Interaction Details
4. Interaction Network

Home ▾ Pan AMR ▾ Pan Virulence ▾ ARG Structure ▾ PPI HPI Drug Targets ▾ Small Molecule ▾ DTI ▾ Genome Browser ▾ Search ▾

Host-Pathogen Interaction

Host-Pathogen Interactions of WHO Bacterial Priority Pathogens

Distribution of HPis from Different Sources

- MINT
- phisto
- HPIDB
- morCVD

Distribution of HPis across Priority Pathogens

- Escherichia coli
- Klebsiella pneumoniae
- Shigella flexneri
- Pseudomonas aeruginosa
- Salmonella enterica
- Neisseria gonorrhoeae
- Staphylococcus aureus
- Haemophilus influenzae
- Streptococcus pyogenes
- Streptococcus agalactiae

Top 10 Pathogen Gene Ontology Terms

GO Term	Frequency
disturbed perturbation of host microtubule cytoskeleton	85
disturbed perturbation of host cell endomembrane system	80
protein ubiquitination	55
protein secretion by the type III secretion system	45
killing of cells of another organism	35
protein autophosphorylation	15
mediated perturbation of host defenses by symbiont	10
inhibition of blood coagulation in another organism	10
inhibition of host NF-kappaB cascade	10
proteolysis	10

Top 10 Pathogen Gene Ontology Terms

Biological Process
Molecular Function
Cellular Component

Interactive Data Explorer

Host-Pathogen Interaction Details

[Download](#)

Filter by Pathogen Organism:

Escherichia coli
 Haemophilus influenzae
 Klebsiella pneumoniae
 Neisseria gonorrhoeae
 Pseudomonas aeruginosa
 Salmonella typhimurium
 Shigella flexneri
 Staphylococcus aureus
 Streptococcus agalactiae
 Streptococcus pyogenes

Search across key columns...

S.No.	Host Protein (ID)	Host Protein (Name)	Pathogen Target (ID)	Pathogen Protein (Name)	Databases	Methods
1	A2A2Y4	FERM domain-containing protein 3 (Band 4,1-like protein 4O) (Ovary type protein 4,1) (4,1O)	Q8XAJ5	T3SS secreted effector NieA/EspI	HPIDB, phisto	two hybrid
2	O00322	Uroplakin-1a (UP1a) (Tetraspanin-21) (Tspan-21) (Uroplakin Ia) (UPIa) (UPKa)	P08191	Type 1 fimbriae D-mannose specific adhesin (Protein FimH)	HPIDB, phisto	solid phase assay, experimental interaction detection

Showing page 1 of 4 (179 total records) Previous 1 Next

Interaction Network

Filter Network by COG Category:

[] Not Classified
 [A] RNA processing and modification
 [B] Chromatin structure and dynamics
 [C] Energy production and conversion

1. Graphical abstract for the overall summary:

This section shows the host-pathogen interaction data sources, pathogen distribution, interaction network, and top functional annotations.

2. Top 10 Pathogen Gene Ontology Terms:

This section displays the bar graph for the top 10 Gene ontology terms. User can select the three gene ontology terms - Biological Process , Molecular Function and Cellular Component.

3. Host-Pathogen Interaction Details:

It provides the pathogen protein's ID and name, the corresponding host protein's ID and name, the source databases, the experimental methods, etc. The user can filter the host-pathogen interaction data by selecting one or multiple pathogen organism names. The result can be displayed in the form of network. The user can download the data.

Interactive Data Explorer

Host-Pathogen Interaction Details

[Download](#)

Filter by Pathogen Organism:

Escherichia coli Haemophilus influenzae Klebsiella pneumoniae Neisseria gonorrhoeae
 Pseudomonas aeruginosa Salmonella typhimurium Shigella flexneri Staphylococcus aureus
 Streptococcus agalactiae Streptococcus pyogenes

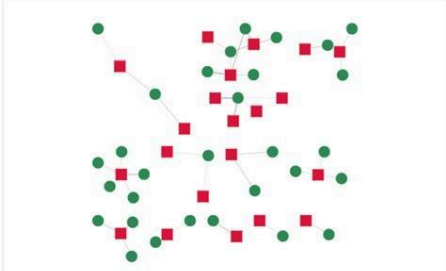
Search across key columns...

S.No.	Host Protein (ID)	Host Protein (Name)	Pathogen Target (ID)	Pathogen Protein (Name)	Databases	Methods	
1	A2A2Y4	FERM domain-containing protein 3 (Band 4.1-like protein 4O) (Ovary type protein 4.1) (4.1O)	Q8XAJ5	T3SS secreted effector NieA/Espl	HPIDB, phisto	two hybrid	2
2	O00322	Uroplakin-1a (UP1a) (Tetraspanin-21) (Tspan-21)	P08191	Type 1 fimbriae D-mannose specific adhesin	HPIDB, phisto	solid phase assay, experimental interaction detection	2

Interaction Network

Filter Network by COG Category:

[-] Not Classified
 [A] RNA processing and modification
 [B] Chromatin structure and dynamics
 [C] Energy production and conversion

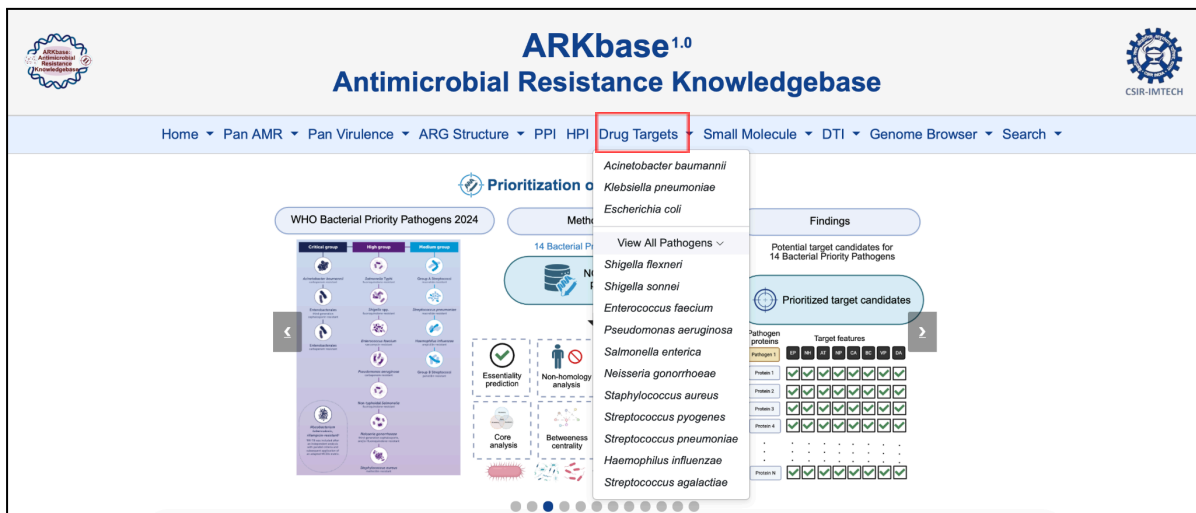


4. Interaction Network:

The user can filter interaction networks by selecting the host proteins according to their COG (Clusters of Orthologous Genes) functional categories.

Drug Targets:

This module provides pathogen-wise potential predicted drug targets for the 14 WHO pathogens; any pathogen can be selected from the drop-down menu of the home page. This module allows users to search by pathogen protein ID and their name (description). The user has the option to filter the targets based upon the provided properties like essentiality, virulence, betweenness centrality metrics, human non-homolog, non-paralogous, and k-mer based non-homolog (anti-target), along with their novelty based upon the Therapeutic Target Database (TTD) and DrugBank.



This module allows users to search by pathogen protein ID and their name (description). The user has the option to filter the targets based upon the provided properties like essentiality, virulence, centrality metrics, human non-homology, non-paralogous, and anti-target, along with their novelty based upon the Therapeutic Target Database (TTD) and DrugBank. The download tab allows the user to download the drug target datasets.

Explore potential drug targets for *Acinetobacter baumannii*

Search by Protein ID or Description:

Items per page: 20 [Evolutionary Analysis](#) [Download](#)

S.No	Protein ID	Protein Description	Essential	Human Non-homolog	Kmer	Non-Paralog	Virulence	Top 10 Central	TTD Novel	DrugBank Novel	Core	Not ARG
1	WP_000064875.1	MULTISPECIES: acyl-ACP-UDP-N-acetylglucosamine O-acyltransferase	●	●	●	●	●	●	●	●	●	●
2	WP_000101096.1	MULTISPECIES: response regulator	●	●	●	●	●	●	●	●	●	●
3	WP_000149523.1	MULTISPECIES: 3-hydroxyacyl-ACP dehydratase FabZ	●	●	●	●	●	●	●	●	●	●
4	WP_000218886.1	histidinol-phosphate transaminase	●	●	●	●	●	●	●	●	●	●
5	WP_000240701.1	UDP-3-O-acyl-N-acetylglucosamine deacetylase	●	●	●	●	●	●	●	●	●	●
6	WP_000273193.1	phosphate regulon sensor histidine kinase PhoR	●	●	●	●	●	●	●	●	●	●
7	WP_000893489.1	MULTISPECIES: phosphatidate cytidyltransferase	●	●	●	●	●	●	●	●	●	●
8	WP_001072477.1	MULTISPECIES: HlyD family secretion protein	●	●	●	●	●	●	●	●	●	●
9	WP_000007335.1	MULTISPECIES: ferrocyclase	●	●	●	●	●	●	●	●	●	●
10	WP_000016600.1	MULTISPECIES: dethiobiotin synthase	●	●	●	●	●	●	●	●	●	●
11	WP_000025981.1	isochromatase family protein	●	●	●	●	●	●	●	●	●	●

In this module, the user can also browse the evolutionary analysis of the prioritized drug target by selecting 'Evolutionary analysis'.

Explore potential drug targets for *Acinetobacter baumannii*

Search by Protein ID or Description...

Filter by: Essential: All | Human NH: All | Anti-Target: All | Non-Paralog: All | Virulence: All | Top 10 Central: All | TTD Novel: All | DrugBank Novel: All | Core: All | Not AMR: All

Items per page: 20 | [Evolutionary Analysis](#) | [Download](#)

S.No	Protein ID	Protein Description	Essential	Human Non-homolog	Anti-Target	Non-Paralog	Virulence	Top 10 Central	TTD Novel	DrugBank Novel	Core	Not AMR
4	WP_000218888.1	histidinol-phosphate transaminase	●	●	●	●	●	●	●	●	●	●
5	WP_000240701.1	UDP-3-O-acetyl-N-acetylglucosamine deacetylase	●	●	●	●	●	●	●	●	●	●
6	WP_000273193.1	phosphate regulon sensor histidine kinase PhoR	●	●	●	●	●	●	●	●	●	●
7	WP_000893489.1	MULTISPECIES: phosphatidate cytidyltransferase	●	●	●	●	●	●	●	●	●	●
8	WP_001072477.1	MULTISPECIES: HyD family secretion protein	●	●	●	●	●	●	●	●	●	●
9	WP_000007335.1	MULTISPECIES: ferrocatalase	●	●	●	●	●	●	●	●	●	●
10	WP_000016606.1	MULTISPECIES: dehydrobiotin synthase	●	●	●	●	●	●	●	●	●	●
11	WP_000026861.1	isochromatase family protein	●	●	●	●	●	●	●	●	●	●
12	WP_000020961.1	MULTISPECIES: tcrR family transcriptional regulator PcaJ	●	●	●	●	●	●	●	●	●	●
13	WP_000051217.1	MULTISPECIES: ATP-binding protein	●	●	●	●	●	●	●	●	●	●
14	WP_000055887.1	malonyl-ACP O-methyltransferase BioC	●	●	●	●	●	●	●	●	●	●
15	WP_000060753.1	MULTISPECIES: two-component system response regulator OmpR	●	●	●	●	●	●	●	●	●	●
16	WP_000078866.1	MULTISPECIES: LpxLpP family Kds(2)-lipid IV(A) lauroylpalmtoeyl acyltransferase	●	●	●	●	●	●	●	●	●	●
17	WP_000080538.1	MULTISPECIES: 3-deoxy-8-phosphoacetyl phosphate synthase	●	●	●	●	●	●	●	●	●	●
18	WP_000089273.1	MULTISPECIES: nitrogen regulation protein NRII	●	●	●	●	●	●	●	●	●	●
19	WP_000097867.1	AraC family transcriptional regulator	●	●	●	●	●	●	●	●	●	●
20	WP_000107128.1	MULTISPECIES: response regulator transcription factor	●	●	●	●	●	●	●	●	●	●

Showing 1 to 20 of 3670 entries

The user can search the data by search interface. In addition, by selecting the view domain, it provides an evolutionary constraint for each site per domain.

ARKbase 1.0
Antimicrobial Resistance Knowledgebase

Home | Pan AMR | Pan Virulence | ARG Structure | PPI HPI Drug Targets | Small Molecule | DTI | Genome Browser | Search

Evolutionary Analysis for Prioritized Drug Target *Acinetobacter baumannii*

Search by Protein ID or Gene Name...

Items per page: 20 | [Download](#)

S.No	Protein ID	Gene Name	FEL	FUBAR	SLAC	MEME	Busted P-Value	Domain
1	WP_000101096.1	response regulator	2,3,4,5,6,7,8,9,10,11,12,13,14... Show More	1,2,3,4,6,7,8,9,10,11,12,13,14... Show More	1,2,3,4,5,6,7,8,9,10,11,12,13,... Show More	12,100,86,104	0.15	View Domain
2	WP_000893489.1	cdsA	139,21,27,66,70,73,76,84,85,91... Show More	6,7,21,27,46,49,58,66,70,73,76... Show More	70,76,91,93,99,102,103,106,110... Show More	182,139,112	0.011	Hide Domain

WP_000893489.1

■ Positive
● Purifying
— Cytidyltransferase family

M L E R T V T A L V L V A V V L G C M F A T Q S H Y P M L V L M T V A A G V A G Y E W Y K L M P R F
V G A V V K P K A W G Y G L L V A F V S G V A L F H D I A L L L W S A S I L T W L V S V Y W V S
F P E F D C W Y N A T L V I G L I L I C A S V T A I F V W Q S S P W W L Y L F L L W G A D S
G A Y F V G R K F G K R K L A P L T V S P L K S V G A L Y G G L T T I I V M L V Q Y L N L L T W
V Q Q L L F L I L S L I T V F G S V L G D L F E S M I K R R A G I K D S G R V L P G H G G V L D R I
D S L L A A A P I F A T G M Y I L K L I G V D L

Small Molecule:

This section is divided into 3 submodules:

1. Biosynthetic Gene Clusters
2. Known Antibiotics
3. Structure Search

The screenshot displays the ARKbase 1.0 Antimicrobial Resistance Knowledgebase interface. The main navigation bar includes Home, Pan AMR, Pan Virulence, ARG Structure, PPI HPI Drug Targets, Small Molecule, DTI, Genome Browser, and Search. The 'Small Molecule' dropdown menu is open, showing 'Biosynthetic Gene Clusters', 'Known Antibiotics', and 'Structure Search'. The central dashboard features a DNA double helix icon and various analysis modules: Deep Annotation, Virulence, Insertion elements, Restriction maps, AST Profiling, Pangenome, Reference, and Insight Module. On the left, there are filters for AMR, Essentiality, HPIs, Small Molecules, and Expression datasets. On the right, there are modules for Co-Target, Evolutionary Analysis, and Genome Browser. At the bottom, there are four main analysis categories: AST as per CLSI, Genome quality by EUCAST, AWaRe classification, and FAIR.

1. Biosynthetic Gene Clusters:

The module gives the user comprehensive data for BGCs—Strain Name, BGC region, BGC genes, Gene Count, Known Metabolites, SMILES, MIBiG Accession, BGC Category, BiG-SCAPE GCF, and antiSMASH Result. BGCs analysis identifies the genomic regions of each pathogen that produce metabolites conferring innate resistance. The user can filter data by pathogen or search any data field. There is also a download option, allowing users to download the entire table and the Bigscape clustering data.

2. Known Antibiotics:

This page features a comprehensive repository of known antibiotics for 14 pathogens with MIC breakpoints, retrieved from the CLSI, with AWaRe annotation. The user can access the antibiotics associated with each pathogen from the left section. It features a search option that enables exploring all 14 pathogens. Additionally, the user can download the data in CSV format. The user can select any annotation, such as Reserve, to see the antibiotics listed under this category.

The screenshot displays the ARKbase 1.0 Antimicrobial Resistance Knowledgebase interface. The top navigation bar includes links for Home, Pan AMR, Pan Virulence, ARG Structure, PPI, HPI, Drug Targets, Small Molecule, DTI, Genome Browser, and Search. The main content area is divided into two columns. The left column, titled 'Known Antibiotics', features a search bar and a 'Download as CSV' button. Below this, a list of pathogens is shown with their antibiotic status: *Acinetobacter baumannii* (Access: 6, Watch: 15, Reserve: 4), *Escherichia coli* (Access: 17, Watch: 44, Reserve: 13), *Klebsiella pneumoniae* (Access: 16, Watch: 45, Reserve: 15), *Shigella flexneri* (Access: 14, Watch: 42, Reserve: 11), and *Shigella sonnei* (Access: 14, Watch: 42, Reserve: 11). The right column displays detailed information for *Acinetobacter baumannii*, including its CLSI breakpoints for Imipenem and Minocycline, and their chemical structures. The Imipenem section shows breakpoints of 2.0 µg/mL for Susceptible (sS), 4 µg/mL for Intermediate (I/SDD), and 8.0 µg/mL for Resistant (rR). The Minocycline section shows breakpoints of 1 µg/mL for Susceptible (sS), 2 µg/mL for Intermediate (I/SDD), and 4 µg/mL for Resistant (rR). Both sections include a 'SMILES Structure Data' link.

3. Structure Search:

This page assists users in identifying matches for their queried SMILES, MOL, or SDF against the structures of secondary metabolites predicted from BGC analysis and known antibiotics from the CLSI database. The user can perform a structure search using either the BGC library or the Known Antibiotic library. The user can also download the output result table as well as the whole library for BGCs and known antibiotics.

The search uses Morgan fingerprints (1024 bits, radius 2) and the Tanimoto threshold of ≥ 0.8

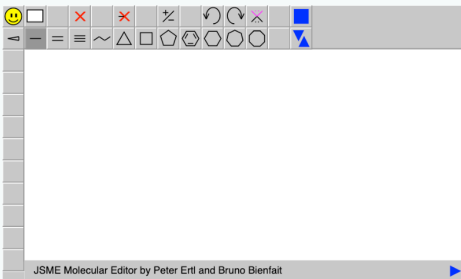
[Download Database](#)

Structure Search Form

Please choose only one option at a time:

1. Draw using JSME editor
2. Paste SMILES
3. Upload SDF/MOL/MOL2 file

Option 1. JSME (Draw Structures & Search)



JSME Molecular Editor by Peter Ertl and Bruno Bienfait

[Clear Editor](#)

Option 2. Paste Structure in SMILES Format

[Load Example SMILES](#)

Option 3. Upload File (MOL/SDF/MOL2)

Browse... No file selected.

[Download Example SDF File](#)

The result page shows all the possible hits along with their structure.

Structure Search Results

Search Query

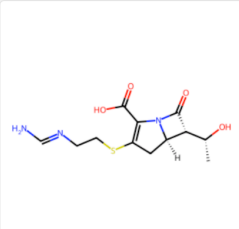
Database: Known Antibiotics

SMILES:

C[C@@H]([C@@H]1[C@@H]2CC(=C(N2C1=O)C(=O)O)SCCN=CN)O

9 hits found


S.No.	Drug / Metabolite Name	Associated Pathogen	Tanimoto Similarity	View Structure
1	Imipenem	<i>Acinetobacter baumannii</i>	1.000	View



Drug-Target Interaction:


This pathogen-wise module provides an interactive, searchable data table for visualising a selected set of targets with their corresponding drugs. This module is divided into four sections:

1. Graphical abstract for the overall summary
2. Top 10 Gene Ontology Terms
3. Explore Drug Target Interactions
4. Target-Drug Interaction Network.



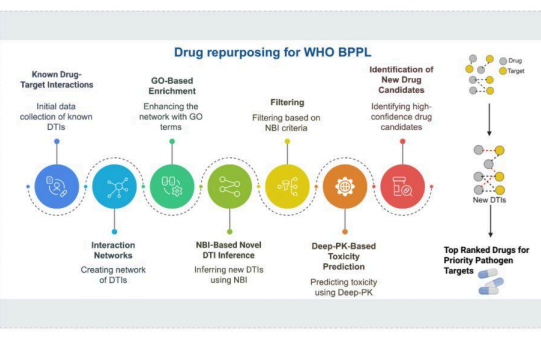
ARKbase^{1.0}

Antimicrobial Resistance Knowledgebase



Home ▾ Pan AMR ▾ Pan Virulence ▾ ARG Structure ▾ PPI HPI Drug Targets ▾ Small Molecule ▾ DTI ▾ Genome Browser ▾ Search ▾

Drug-Target Interactions for *Escherichia coli*



Drug repurposing for WHO BPPL

The workflow consists of several steps: 1. Initial data collection of known DTIs. 2. Interaction Networks (Creating network of DTIs). 3. GO-Based Enrichment (Enhancing the network with GO terms). 4. NBI-Based Novel DTI Inference (Inferring new DTIs using NBI). 5. Filtering (Filtering based on NBI criteria). 6. Deep-PK-Based Toxicity Prediction (Predicting toxicity using Deep-PK). 7. Identification of New Drug Candidates (Identifying high-confidence drug candidates). 8. Top Ranked Drugs for Priority Pathogen Targets.

Top 10 GO terms

GO Molecular Term	Frequency
protein homodimerization activity	35000
identical protein binding	20000
ATP binding	15000
zinc ion binding	12000
magnesium ion binding	10000
DNA binding	8000
manganese ion binding	6000
metal ion binding	5000
thiamine pyrophosphate binding	4000
DNA topoisomerase type II (double strand cut, ATP-hydrolyzing) activity	3000

Interactive Explorer

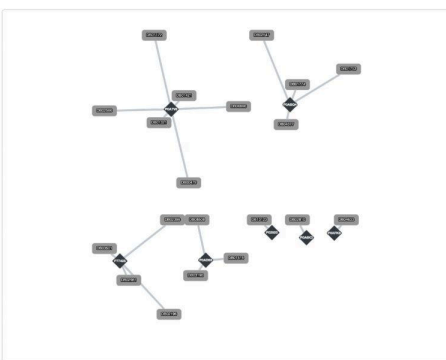
Explore Drug Target Interactions

Search by Target, Drug, Protein, etc. (on current page)

Target ID	DrugBank ID	Drug Name	Score	Protein Names	Drug Degree	Target Degree	Reference Protein ID
P0A6K3	DB02810	N-(2-Acetamido)iminodiacetic Acid	1.00	Peptide deformylase (PDF) (EC 3.5.1.88) (Polypeptide deformylase)	2	1515	NP_312179.1 NA
P0A7V8	DB02586	8-Hydroxy-4-(1-Hydroxyethyl)quinoline-2-Carboxylic Acid	1.00	Small ribosomal subunit protein uS4 (30S ribosomal protein S4)	2	233	NP_312188.1 NA
P0A7V8	DB06696	Arbekacin	1.00	Small ribosomal subunit protein uS4 (30S ribosomal protein S4)	7	233	NP_312188.1 NA

Prev Page 1 of 2922 (58423 records) Next 20 items per page ▾

Target-Drug Interaction Network

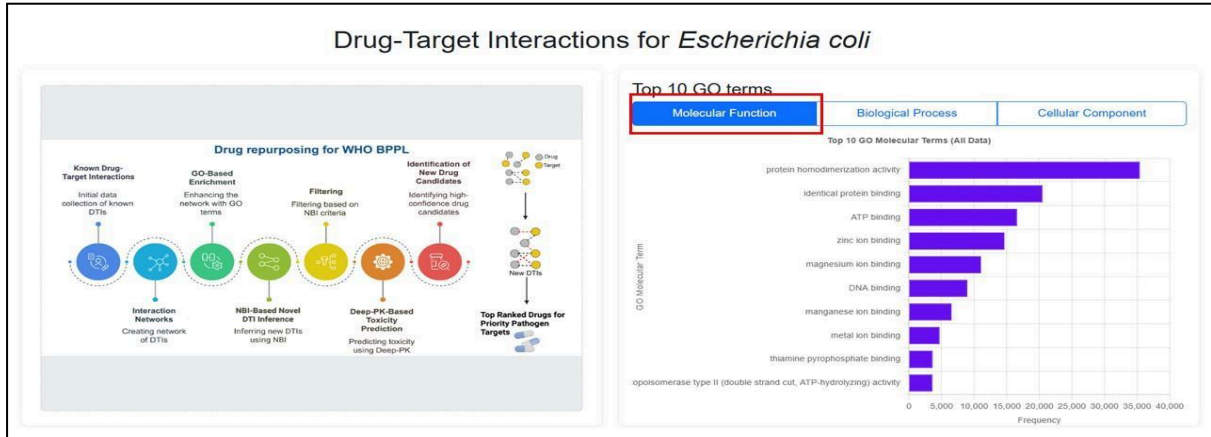


1. Graphical abstract for the overall summary:

This figure depicts the overall methodology and workflow of drug repurposing for the WHO Bacterial Priority Pathogen List (BPPL).

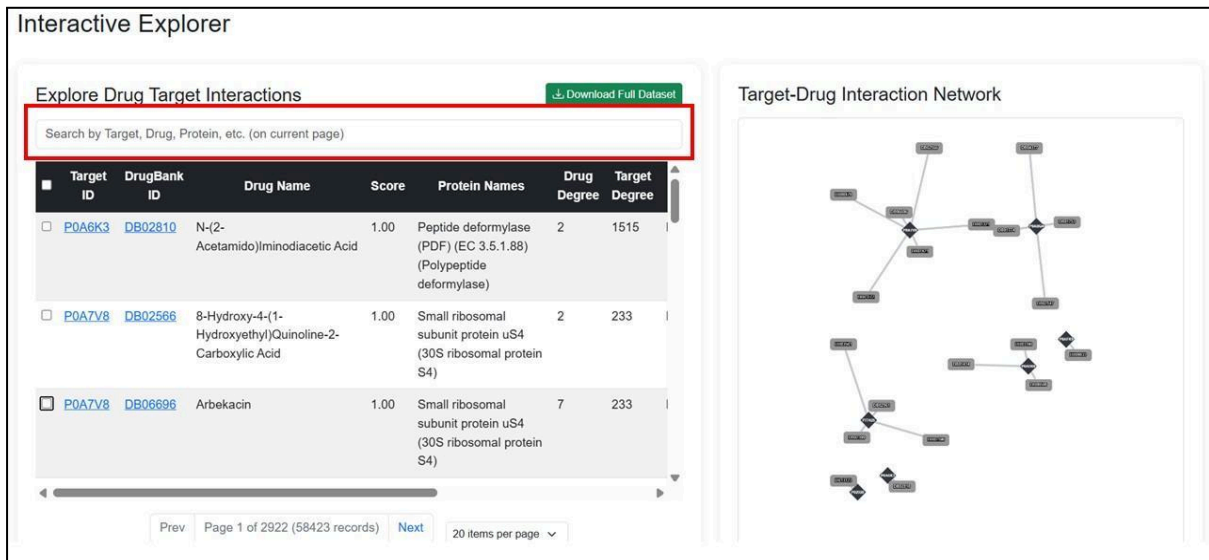
2. Top 10 Gene Ontology Terms:

This section displays the bar graph for the top 10 Gene Ontology terms. The user can select the three gene ontology terms - Molecular Function, Biological Process and Cellular Component.



3. Explore Drug Target Interactions:

The user can select a pathogen from the drop-down menu in the Drug-Target Interaction module. The search interface is provided for exploring any data field. The results will display detailed information about each interaction, including the Target ID, DrugBank ID, drug name, score, protein names, drug and target degree, relevant pathways, and toxicity safety. Additionally, users can view the data in a network format.



The user can select one or multiple target IDs to explore drug-target interactions and the network of selected pathogens. The download tab allows the user to download the complete interaction datasets.

4. Target-Drug Interaction Network:

The interaction network visualizes connections between drug candidates (diamond nodes) and their bacterial protein targets (rectangular nodes), highlighting potential therapeutic relationships for WHO-priority pathogens.

Interactive Explorer

Explore Drug Target Interactions

[Download Full Dataset](#)

Search by Target, Drug, Protein, etc. (on current page)

<input checked="" type="checkbox"/>	P0A7V8	DB02566	8-Hydroxy-4-(1-Hydroxyethyl)Quinoline-2-Carboxylic Acid	1.00	Small ribosomal subunit protein uS4 (30S ribosomal protein S4)	2	233
<input checked="" type="checkbox"/>	P0A7V8	DB06696	Arbekacin	1.00	Small ribosomal subunit protein uS4 (30S ribosomal protein S4)	7	233
<input checked="" type="checkbox"/>	P0A7V8	DB01172	Kanamycin	1.00	Small ribosomal subunit protein uS4 (30S ribosomal protein S4)	7	233
<input checked="" type="checkbox"/>	P0A7V8	DB01321	Josamycin	1.00	Small ribosomal subunit protein uS4 (30S ribosomal protein S4)	2	233

Prev Page 1 of 2922 (58423 records) Next 20 items per page

Target-Drug Interaction Network

The diagram illustrates a network of drug-target interactions. A central diamond-shaped node labeled 'P0A7V8' is connected to three rectangular nodes. One node is labeled 'DB06696', another is 'DB01321', and a third is partially visible on the right. The connections are represented by grey lines.

Genome Browser:

The Genome Browser module provides an interactive platform for visual exploration of the genomic features of 14 bacterial pathogens for the WHO Priority Pathogen List, 2024, included in this database. This tool enables users to visualise, navigate, and interpret genomic regions with ease, offering a detailed view of both reference sequences and associated annotation tracks.

ARKbase 1.0
Antimicrobial Resistance Knowledgebase

Home ▾ Pan AMR ▾ Pan Virulence ▾ ARG Structure ▾ PPI HPI Drug Targets ▾ Small Molecule ▾ DTI ▾ **Genome Browse** ▾ Search ▾

Interactome-Based Co-Target Discovery in WHO-Bacterial Priority Pathogens

Construction of PPINs

- Reference Proteome (Protein-sequence)
- Proteome Annotation using STRING
- Filtering High-Quality Interactions (combined score ≥ 700)

Calculating the Centrality Metrics

NetworkX

Protein (blue circle) Central Protein (red circle) Protein-protein interaction (grey line)

Co-Target Identification

Betweenness Centrality

- Filtered the top 10% of protein betweenness centrality to identify connectors.
- Identification of interacting pairs **Co-Targets**.
- Co-targets represent promising combinatorial targeting, supporting target prioritization.

Further Screening of Co-Targets (Core Genes, AMR Gene Exclusion, Essentiality Assessment, to Human Proteins) for Drug Target Prioritization

14 WHO-Bacterial Priority Pathogen Protein-Protein Interaction Networks

Proteins among Pathogens PPINs among Pathogens

View All Pathogens ▾

- Acinetobacter baumannii*
- Klebsiella pneumoniae*
- Escherichia coli*
- Shigella flexneri*
- Shigella sonnei*
- Enterococcus faecium*
- Pseudomonas aeruginosa*
- Salmonella enterica*
- Neisseria gonorrhoeae*
- Staphylococcus aureus*
- Streptococcus pyogenes*
- Streptococcus pneumoniae*
- Haemophilus influenzae*
- Streptococcus agalactiae*

The module covers genome data for 14 bacterial pathogens, curated from high-quality reference sources. 30+ annotation tracks are available for each pathogen for visualisation, representing a variety of data types such as gene annotation, operons, insertion elements, Antibiotic Resistance Genes (AMR), restriction sites, protein features, Centrality Metrics for protein-protein interactions (Degree, Closeness and Betweenness) and so on. The user can interactively navigate the browser for the chosen pathogen by zooming in/out to view genomic features at nucleotide or broader chromosomal scales, scrolling across genomic regions for comparative exploration and clicking on features to access detailed metadata.

The screenshot displays the ARKbase 1.0 interface. At the top, the logo and name 'ARKbase 1.0 Antimicrobial Resistance Knowledgebase' are visible, along with the CSIR-IMTECH logo. Below the header is a navigation menu with options like 'Home', 'Pan AMR', 'Pan Virulence', 'ARG Structure', 'PPI HPI Drug Targets', 'Small Molecule', 'DTI', 'Genome Browser', and 'Search'. The main content area shows a genomic browser for *Shigella flexneri* with a zoomed-in view of a specific region (NC_004337.2). The browser displays multiple tracks: Reference sequence, Gene Annotations, Essential (DEG), IS Elements, Operons, ARGs (RGI), ARGs (MEGARes), ARGs (AMRFinderPlus), Protein Features, and various Centrality metrics (Betweenness, Degree, Closeness). The right sidebar provides a 'Filter tracks' panel where users can select or deselect these tracks.

How to Use

1. Select an organism from the drop-down list or organism overview page.
2. Choose tracks of interest from the track selector to customize your view.
3. Use the zoom and pan controls or the search bar to locate genes, loci, or specific genomic coordinates.
4. Click any feature to open a pop-up with details, including functional annotation. Users can also hover over the features for their corresponding key information.
5. Download GFF files for the selected organism from the Download section for further analysis in genome analysis pipelines.

Data download of PanGenome annotation for each organism is available in GFF format for offline analysis. Download links are provided directly from the Genome Browser interface and annotation files are organized by pathogen for convenient bulk or individual retrieval.

Search:

The search page is divided into 3 sections: Comparative Analysis Module, Search by Sequence and Advanced Search.

ARKbase^{1.0}
Antimicrobial Resistance Knowledgebase

Home ▾ Pan AMR ▾ Pan Virulence ▾ ARG Structure ▾ PPI HPI Drug Targets ▾ Small Molecule ▾ DTI ▾ Genome Browser ▾ Search ▾

Prioritization of potential targets

WHO Bacterial Priority Pathogens 2024

Methodology

14 Bacterial Priority Pathogens

NCBI RefSeq proteomes

Findings

Potential target candidates for 14 Bacterial Priority Pathogens

Prioritized target candidates

Pathogen proteins

Target features

Essentially prediction

Non-homology analysis

Anti-target analysis

Non-paralog analysis

Core analysis

Betweenness centrality

Virulence prediction

Druggability analysis


Evolutionary Inference

ARKbase is an integrated, curated, value-added knowledge-base for AMR, with focus on WHO Bacterial Priority Pathogens. ARKbase is a dedicated AMR resource with a potential to provide novel insights towards expanding the drug-target space.

ARKbase is the largest dedicated resource for WHO bacterial priority pathogens, integrating high-quality genomes, proteomes, AMR genes, and comprehensive datasets including biosynthetic gene clusters, drug targets, antibiotic profiles, gene expression data and machine learning models. This centralized platform assists in understanding resistance mechanisms and supports targeted drug discovery for AMR.


1. Comparative Analysis Module

This module allows users to compare data related to virulence, antimicrobial resistance (AMR), and drug target proteins. Users can select a query category, choosing from AMR proteins, virulence factor (VF) proteins, or drug target proteins of any pathogen, and compare them against that same category for any other pathogen. For example, users can compare the AMR proteins of *Acinetobacter baumannii* with the AMR proteins of *Escherichia coli*. By selecting Run Analysis, the user can access the results in a new page. The result can be downloaded in tsv format.



ARKbase^{1.0}

Antimicrobial Resistance Knowledgebase



[Home](#) ▾ [Pan AMR](#) ▾ [Pan Virulence](#) ▾ [ARG Structure](#) ▾ [PPI](#) [HPI](#) [Drug Targets](#) ▾ [Small Molecule](#) ▾ [DTI](#) ▾ [Genome Browser](#) ▾ [Search](#) ▾

Comparative Analysis Module

Compare Virulence/AMR/Drug Targets

Step 1: Select Query Category

AMR Proteins
 Virulence Factor (VF) Proteins
 Drug Target Proteins

Step 2: Select Query Proteins

Source Pathogen:

Query Proteins:

- WP_001112624.1 - aminoglycoside nucleotidyltransferase ANT(3^{IIa}) [Acinetobacter baumannii]
- WP_00459547.1 - efflux system response regulator transcription factor AdeR [Acinetobacter baumannii]
- WP_001211227.1 - extended-spectrum class C beta-lactamase ADC-158 [Acinetobacter baumannii]
- WP_000867106.1 - glycosyltransferase family 4 protein [Acinetobacter baumannii]


Step 3: Select Target Organism(s) (Database)

Search against the genome(s) of:

- Acinetobacter baumannii
- Escherichia coli
- Enterococcus faecium
- Haemophilus influenzae
- Klebsiella pneumoniae
- Neisseria gonorrhoeae
- Pseudomonas aeruginosa


This analysis uses standard BLASTP default search criteria. ⓘ

Run Analysis



ARKbase^{1.0}

Antimicrobial Resistance Knowledgebase



[Home](#) ▾ [Pan AMR](#) ▾ [Pan Virulence](#) ▾ [ARG Structure](#) ▾ [PPI](#) [HPI](#) [Drug Targets](#) ▾ [Small Molecule](#) ▾ [DTI](#) ▾ [Genome Browser](#) ▾ [Search](#) ▾

BLAST Analysis Results

BLAST Hits (Tabular Format)

Showing results for selected query proteins against *Escherichia coli* from the AMR database.

⚠ Results are sorted by Bit Score (highest first) to show the most significant hits at the top.

Download Results (TSV)
Download Subject Sequences (FASTA)

Query ID (prot_id)	Subject ID (prot_id)	% Identity	Length (AA)	Mismatch	Gap Open	Q. start	Q. end	S. start	S. end	E-value	Bit Score
WP_001112624.1	reflNP_311276.1	30.952	42	29	0	54	95	740	781	3.5	20.0
WP_001112624.1	reflNP_311172.1	36.667	30	16	1	100	126	499	528	4.6	19.6
WP_001112624.1	reflNP_311276.1	30.769	39	26	1	43	81	254	291	7.8	19.2
WP_001112624.1	reflNP_310914.1	37.500	24	15	0	6	29	204	227	7.1	18.9
WP_001112624.1	reflNP_308458.1	36.000	25	16	0	153	177	330	354	9.2	18.9

There is a download section with files containing the complete all-vs-all BLASTP analysis for each protein category, comparing every protein from one pathogen against all proteins from every other pathogen within the same category. The analysis was performed using the following parameters: E-value 1e-4, Percent Identity 35%, and Query Coverage 50%.

Download Precomputed BLAST Results

These files contain all-versus-all BLASTp results for AMR, VF, and DT proteins. Each protein from one pathogen was compared against all proteins from other pathogens within the same category. The analysis was performed using the following parameters: E-value $1e-4$, Percent Identity 35%, and Query Coverage 50%.

AMR Proteins

Download AMR Results

Virulence Factors (VF)

Download VF Results

Drug Targets

Download DT Results

There is also an Interactive Genome Visualisation section featuring comparative genome maps (Circos plots). Users can interactively explore and select AMR genes, virulence genes, or all genes (AMR and Virulence together) for detailed viewing.

Interactive Genome Visualization

Select two pathogens to compare their interactive genome maps (Circos Plots).

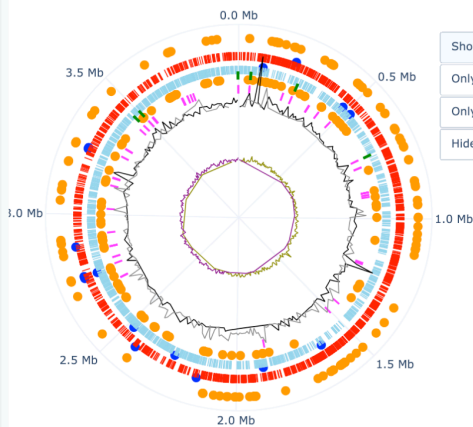
Select Pathogen 1:

A baumannii

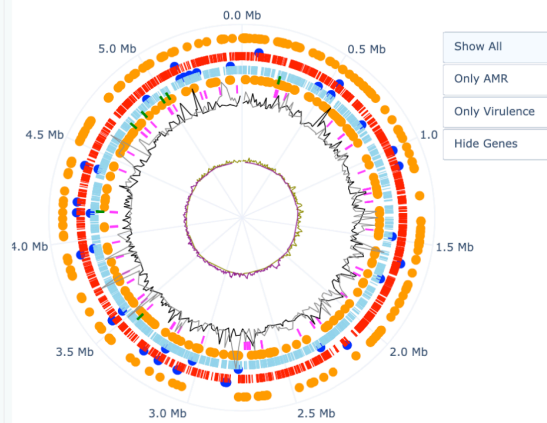
Select Pathogen 2:

E coli

Acinetobacter baumannii – Interactive Genome Map (AMR & Virulence)



Escherichia coli – Interactive Genome Map (AMR & Virulence)



3. Advanced Search

This module enables users to query based on 14 pathogens, protein features (e.g., essential, virulence, top 10 central, etc), COG categories, and ARGs. For example, users can identify proteins that are essential, virulence factors, or ARGs across all 14 pathogens.

ARKbase Advanced Search | Query Builder

Acinetobacter baumannii
 Enterococcus faecium
 Klebsiella pneumoniae
 Pseudomonas aeruginosa
 Escherichia coli
 Haemophilus influenzae
 Neisseria gonorrhoeae
 Streptococcus agalactiae

Essential Gene
 Core Gene
 Non-Paralog
 Top 10 Central
 Human Non-Homolog
 Anti-Target
 Virulence Factor
 TTD Novel Target

[-] Unknown Category
 [B] Chromatin structure and dynamics
 [D] Cell cycle control, cell division, chromosome partitioning
 [A] RNA processing and modification
 [C] Energy production and conversion
 [E] Amino acid transport and metabolism

All Tools
 AMRFinderPlus
 ResFinder
 RGI
 MEGARes

You searched for: Pathogen Name EQUALS Escherichia coli AND Protein Feature is Essential Gene AND Pathogen COG EQUALS K AND ARGs EQUALS All Tools

SR.NO.	PATHOGEN NAME	PROTEIN ID	PROTEIN DESCRIPTION	COG CATEGORY	COG DESCRIPTION	ESSENTIAL GENE	ASSEMBLY ACCESSION	CONTRIBUTING TOOLS
1	Escherichia coli	NP_308107.3	DNA-binding transcriptional activator LeuO	K	Transcription	Yes	GCF_000008865.2	RGI
2	Escherichia coli	NP_308544.1	transcriptional repressor	K	Transcription	Yes	GCF_000008865.2	RGI
3	Escherichia coli	NP_308749.1	two-component regulatory system response regulator KdpE	K	Transcription	Yes	GCF_000008865.2	Megares,RGI
4	Escherichia coli	NP_309766.1	global DNA-binding transcriptional dual regulator H-NS	K	Transcription	Yes	GCF_000008865.2	Megares,RGI

Operons:

This module allows the user to view the total number of operons present in the pathogen. The user can choose a pathogen from the drop-down menu to explore predicted operons. Features like the total length of the operon, locus tags of the gene belonging to the operon, operon_id, gene name with gene coordinates and strand info, protein ID, and the product have been provided. The protein_id is directly linked to the NCBI protein database.

ARKbase 1.0
Antimicrobial Resistance Knowledgebase

Home ▾ Pan AMR ▾ Pan Virulence ▾ ARG Structure ▾ PPI HPI Drug Targets ▾ Small Molecule ▾ DTI ▾ Genome Browser ▾ Search ▾

Genome Curation & Deep Annotation of WHO Bacterial Priority Pathogens

WHO Bacterial Priority Pathogens
Reference DBs from NCBI databases

- Critical group
- High group
- Medium group

Genome and Susceptibility data
High-Quality Genome Filtering Criteria

- Clinical Isolates: Human-origin samples collected from locally antibiotic resistance
- GLI Classification: Classification based on GenBank V. Standards matrix
- EBI/CAZ Quality Check: Quality control using standard genome standards
- Assembly Quality: Filter design indicates higher quality assembly
- Completeness Metrics: Assessing the completeness and quality of the assembly

Comprehensive Genome Annotation

- Genomic Annotation: Identifies gene features like CDS, rRNA, tRNA, and regulatory elements
- Structure-Based Annotation: Predicts protein functions through 3D structure analysis
- Operon Detection: Provides insights into gene regulation and transcriptional units
- Virulence Factor Detection: Identifies factors contributing to pathogen virulence
- Pangenome Analysis: Reveals core and accessory gene content for comparative genomics
- BGC Prediction: Identifies potential secondary metabolite production
- AMR Gene Identification: Detects genes related to antimicrobial resistance
- Interactome Analysis: Identifies high central interacting pairs (co-targets)

ARKbase is an integrated, curated, value-added knowledge-base for AMR, with focus on WHO Bacterial Priority Pathogens. ARKbase is a dedicated AMR resource with a potential to provide novel insights towards expanding the drug-target space.

ARKbase is the largest dedicated resource for WHO bacterial priority pathogens, integrating high-quality genomes, proteomes, AMR genes, and comprehensive datasets including biosynthetic gene clusters, drug targets, antibiotic profiles, gene expression data and machine learning models. This centralized platform assists in understanding resistance mechanisms and supports targeted drug discovery for AMR.

Operons - Curated Genome AST - Expression Data - ML Models - Data Summary - Co-Target - Connect - FAQs & Help

The module includes an interactive filter to search by Operon ID, Locus Tag, Gene name and so on. The download tab allows the user to download the complete operon summary for that particular genome.

Explore Predicted Operon for *Acinetobacter baumannii*

Search by Operon ID, Locus Tag, Gene Name, etc... Items per page: 20

#	Operon ID	Locus Tag	Operon Length	Gene Name	Protein ID	Start	Stop	Strand	Product
1	GCF_009035845.1_op2392	FQU82_RS00005	1682	repM	WP_001208779.1	144	1094	+	replication initiation protein RepM
2	GCF_009035845.1_op2392	FQU82_RS00010	1682	-	WP_001096618.1	1087	1662	+	plasmid replication DNA-binding protein
3	GCF_009035845.1_op2392	FQU82_RS00015	1682	-	WP_001125246.1	1682	1825	+	hypothetical protein
4	GCF_009035845.1_op2394	FQU82_RS00020	586	nadS	WP_000369782.1	3099	3371	-	NadS family protein
5	GCF_009035845.1_op2394	FQU82_RS00025	586	-	WP_000897309.1	3364	3684	-	type II toxin-antitoxin system ReiE/ParE family toxin
6	GCF_009035845.1_op2395	FQU82_RS00030	198	-	WP_000476222.1	3871	4068	-	hypothetical protein
7	GCF_009035845.1_op2396	FQU82_RS00035	228	-	WP_000921865.1	4135	4362	-	hypothetical protein
8	GCF_009035845.1_op2397	FQU82_RS00040	216	-	WP_000071892.1	4460	4675	-	cold-shock protein
9	GCF_009035845.1_op2398	FQU82_RS00045	513	-	WP_000084854.1	5002	5514	-	hypothetical protein
10	GCF_009035845.1_op2399	FQU82_RS00050	360	-	WP_000360920.1	5609	5968	-	hypothetical protein
11	GCF_009035845.1_op2400	FQU82_RS00055	512	-	WP_001985218.1	6077	6217	-	hypothetical protein

Expression Data:

Transcriptomics encompasses the analysis of RNA sequences, their expression levels, and how they are regulated. This module is a compilation of all the datasets to date, involving pathogens exposed to their respective priority-resistant antibiotics. General information about the dataset, including project title, accession, release date, and sample size, has been provided.

ARKbase 1.0
Antimicrobial Resistance Knowledgebase

Home ▾ Pan AMR ▾ Pan Virulence ▾ ARG Structure ▾ PPI HPI Drug Targets ▾ Small Molecule ▾ DTI ▾ Genome Browser ▾ Search ▾

Genome Curation & Deep Annotation of WHO Bacterial Priority Pathogens

WHO Bacterial Priority Pathogens
Reference Data from NCBI databases

- Critical group**
- High group**
- Medium group**

Genome and Susceptibility data
High-Quality Genome Filtering Criteria

- Clinical Isolates**: Human-origin isolates collected from publicly available databases.
- CLS Classification**: Classification based on Clinical and Laboratory Standards Institute.
- EICAST Quality Check**: Quality control using European Committee on Antimicrobial Susceptibility Testing.
- Assembly Quality**: Fewer contigs indicate a higher quality assembly.
- Completeness Metrics**: Assessing the completeness and quality of the assembly.

Comprehensive Genome Annotation

- Genomic Annotation**: Identifies gene features like CDS, tRNA, rRNA, and regulatory elements.
- Structure-Based Annotation**: Predicts protein function through 3D structure analysis.
- Operon Detection**: Provides insights into gene regulation and transcriptional units.
- Virulence Factor Detection**: Identifies factors contributing to pathogen virulence.
- Pangenome Analysis**: Reveals core and accessory gene content for comparative genomics.
- BGC Prediction**: Identifies potential for secondary metabolite production.
- AMR Gene Identification**: Detects genes related to antimicrobial resistance.
- Interactions Analysis**: Identifies high central interacting pairs (co-targets).

ARKbase is an integrated, curated, value-added knowledge-base for AMR, with focus on WHO Bacterial Priority Pathogens. ARKbase is a dedicated AMR resource with a potential to provide novel insights towards expanding the drug-target space.

ARKbase is the largest dedicated resource for WHO bacterial priority pathogens, integrating high-quality genomes, proteomes, AMR genes, and comprehensive datasets including biosynthetic gene clusters, drug targets, antibiotic profiles, gene expression data and machine learning models. This centralized platform assists in understanding resistance mechanisms and supports targeted drug discovery for AMR.

Operons - Curated Genome AST - **Expression Data** - ML Models - Data Summary - Co-Target - Connect - FAQs & Help

The module includes two interactive filters for targeted data exploration:

- Priority antibiotic name
- Pathogen name

The user can also search pathogen and/or priority antibiotic combinations of their interest using the Browse tab.



Transcriptomics Antibiotics Exposure

Antibiotic: Organism:

Show 10 entries

Browse:

#	Organism	Project Title	Project Database	Project Accession	Antibiotic Name	Release Date	Samples
1	Acinetobacter baumannii	The transcriptomic response of <i>Acinetobacter baumannii</i> to colistin and doripenem alone and in combination in an in vitro pharmacokinetics/pharmacodynamics model	GEO	PRJNA265162	Doripenem	Nov-30-2014	22
2	Acinetobacter baumannii	Transcriptional response to carbapenem treatment in <i>Acinetobacter baumannii</i>	GEO	PRJNA787205	Meropenem	May-20-2022	18
3	Escherichia coli	Prediction of antibiotic resistance by large-scale phenotypic and genotypic data	GEO	PRJNA255254	Cefoperazone, Cefixime	Jul-14-2014	42
4	Escherichia coli	Global gene expression profiling and antibiotic susceptibility after repeated exposure to the carbon monoxide-releasing molecule-2 (CORM-2) in multidrug-resistant ESBL-producing uropathogenic <i>Escherichia coli</i>	GEO	PRJNA345430	Cefotaxime	Oct-05-2016	18
5	Escherichia coli	Tellurite-mediated cefotaxime (CTX) potentiation	GEO	PRJNA171495	Cefotaxime	Jul-27-2012	9
6	Escherichia coli	Expression of <i>Escherichia coli</i> treated with cefsulodin and mecillinam	GEO	PRJNA108353	Cefsulodin	Jul-15-2008	61
7	Escherichia coli	Expression of <i>Escherichia coli</i> treated with cefsulodin and mecillinam, alone at the minimum inhibitory concentration	GEO	PRJNA109011	Cefsulodin	Jul-15-2008	18
8	Escherichia coli	Expression of <i>Escherichia coli</i> treated with cefsulodin and mecillinam, alone and in combination	GEO	PRJNA109009	Cefsulodin	Jul-15-2008	43
9	Escherichia coli	Comparative transcriptomic response analysis of <i>Escherichia coli</i> K-12 MG1655 to nine antibiotics	GEO	PRJNA910221	Ceftazidime, Imipenem	Dec-08-2022	33

Machine Learning (ML) Models:

This module highlights machine learning models developed for MIC prediction targeting pathogens from the WHO Bacterial Priority Pathogen List.

ARKbase 1.0
Antimicrobial Resistance Knowledgebase

Home ▾ Pan AMR ▾ Pan Virulence ▾ ARG Structure ▾ PPI HPI Drug Targets ▾ Small Molecule ▾ DTI ▾ Genome Browser ▾ Search ▾

Genome Curation & Deep Annotation of WHO Bacterial Priority Pathogens

WHO Bacterial Priority Pathogens
Reference Data from NCBI datasets

- Critical group**
- High group**
- Medium group**

Genome and Susceptibility data
High-Quality Genome Filtering Criteria

- EUCAST Quality Check**: Quality control using European Committee standards
- Assembly Quality**: Fewer contigs indicate a higher quality assembly
- Completeness Metrics**: Assessing the completeness and quality of the assembly

Comprehensive Genome Annotation

- Clinical Isolates**: Human-derived samples collected from publicly available databases
- CLS Classification**: Classification based on Clinical and Laboratory Standards Institute
- Genomic Annotation**: Identifies gene features like CDS, rRNA, tRNA, and regulatory elements
- Structure-Based Annotation**: Predicts protein function through 3D structure analysis
- Operon Detection**: Provides insights into gene regulation and transcriptional units
- Virulence Factor Detection**: Identifies factors contributing to pathogen virulence
- Pangenome Analysis**: Reveals core and accessory gene content for comparative genomics
- BGC Prediction**: Identifies potential for secondary metabolite production
- AMR Gene Identification**: Detects genes related to antimicrobial resistance
- Interactome Analysis**: Identifies high central interacting pairs (top-Targets)

ARKbase is an integrated, curated, value-added knowledge-base for AMR, with focus on WHO Bacterial Priority Pathogens. ARKbase is a dedicated AMR resource with a potential to provide novel insights towards expanding the drug-target space.

ARKbase is the largest dedicated resource for WHO bacterial priority pathogens, integrating high-quality genomes, proteomes, AMR genes, and comprehensive datasets including biosynthetic gene clusters, drug targets, antibiotic profiles, gene expression data and machine learning models. This centralized platform assists in understanding resistance mechanisms and supports targeted drug discovery for AMR.

Operons - Curated Genome AST - Expression Data **ML Models** Data Summary Co-Target Connect FAQs & Help

The user can get an idea of what the different ML models reported for different antibiotics divided into AWaRe classes. Users can access the curated models with GitHub repositories or without GitHub repositories.

When user select **With GitHub repositories**, models are categorized into

- Access
- Watch
- Reserve
- Access/Watch
- Watch/Reserve
- Unclassified/Not recommended

Curated Machine Learning Models for MIC Prediction

With GitHub Repositories Without GitHub Repositories

Tetracycline

<p>Pathogen: <i>Escherichia coli</i></p> <p>Algorithm: GDBT,LR,RF, SVC</p> <p>Score: Average AUC: For EC: close to 1 (GDBT)</p> <p>View Github Page</p>	<p>Pathogen: <i>Mycobacterium tuberculosis</i></p> <p>Algorithm: GDBT,LR,RF, SVC</p> <p>Score: Average AUC: For EC: close to 1 (GDBT)</p> <p>View Github Page</p>	<p>Pathogen: <i>Neisseria gonorheae</i></p> <p>Algorithm: GDBT,LR,RF, SVC</p> <p>Score: Average AUC: For EC: close to 1 (GDBT)</p> <p>View Github Page</p>	<p>Pathogen: <i>Klebsiella pneumoniae</i></p> <p>Algorithm: LGB</p> <p>Score: Average F1 score: close 1</p> <p>View Github Page</p>
---	---	--	---

Users can view **Without GitHub repositories**, which are divided into

- Access
- Watch
- Watch/Reserve
- Unclassified/Not Recommended

Curated Machine Learning Models for MIC Prediction

With GitHub Repositories **Without GitHub Repositories**

Access Watch Watch/Reserve Unclassified/Not Recommended

Gentamicin

Pathogen	Algorithm	Score	Action
<i>Escherichia coli</i>	SVM, NB, AB, RF	AUC: 0.64 - 0.71 (SVM); 0.93 - 0.97 (SVM with GA)	View Article
<i>Pseudomonas aeruginosa</i>	RF and BioWeka	Score: Mean accuracy: ≥98% (BioWeka); ≥96% (Random Forest)	View Article
<i>Pseudomonas aeruginosa</i>	XGB, AB, RF, LR, SVM	Score: Accuracy: 0.678 - 0.837	View Article
<i>Acinetobacter baumannii</i>	DNN, LR	Score: Accuracy: 0.9122 - 0.9864	View Article

Connect:

ARKbase^{1.0}
Antimicrobial Resistance Knowledgebase

Home ▾ Pan AMR ▾ Pan Virulence ▾ ARG Structure ▾ PPI HPI Drug Targets ▾ Small Molecule ▾ DTI ▾ Genome Browser ▾ Search ▾

Prioritization of potential targets

WHO Bacterial Priority Pathogens 2024

Methodology

14 Bacterial Priority Pathogens
NCBI RefSeq proteomes

Findings

Potential target candidates for 14 Bacterial Priority Pathogens

Prioritized target candidates

Pathogen, Proteins, Target features

Essentiality prediction, Non-homology analysis, Anti-target analysis, Non-perturb analysis, Core analysis, Betweenness centrality, Virulence prediction, Druggability analysis, Evolutionary Inference

ARKbase is an integrated, curated, value-added knowledge-base for AMR, with focus on WHO Bacterial Priority Pathogens. ARKbase is a dedicated AMR resource with a potential to provide novel insights towards expanding the drug-target space.

ARKbase is the largest dedicated resource for WHO bacterial priority pathogens, integrating high-quality genomes, proteomes, AMR genes, and comprehensive datasets including biosynthetic gene clusters, drug targets, antibiotic profiles, gene expression data and machine learning models. This centralized platform assists in understanding resistance mechanisms and supports targeted drug discovery for AMR.

Operons - Curated Genome AST - Expression Data - ML Models - Data Summary - Co-Target - **Connect** - FAQs & Help

ARKbase^{1.0}
Antimicrobial Resistance Knowledgebase

Home ▾ Pan AMR ▾ Pan Virulence ▾ ARG Structure ▾ PPI HPI Drug Targets ▾ Small Molecule ▾ DTI ▾ Genome Browser ▾ Search ▾

Community Contributions

We invite researchers, clinicians, and domain experts to contribute to ARKbase. Your contributions are vital in helping the global community accelerate discoveries.

Submit Your Contribution

Your Name*

Email*

Genotype

Pathogen Name*

Genome Accession Number*

Genome Accession URL Link*

Publication DOI (Optional)

Antibiotic Susceptibility Testing (AST) Data

Antibiotic Name*

Minimum Inhibitory Concentration (µg/mL)

Disk Diffusion (mm)

Short Description or Notes (Optional)

Our Team

Shivan Gambhir, Project Associate I, CSIR-IMTech

Sheela Pandey, Senior Research Fellow, CSIR-IMTech

Harsh Bajetha, Project Associate I, CSIR-IMTech

Jatleen Kaur, Project Candidate, CSIR-IMTech

Ankita Das, Project Associate I, CSIR-IMTech

Dr. Phaniavithiyani G., Project Scientist I, CSIR-IMTech

Rupali Aggarwal, Dissertation Student, CSIR-IMTech

Upasana Malty, Project Associate II, CSIR-IMTech

Suhani Dange, Project Associate I, CSIR-IMTech

Dr. Vipan Singh, Independent Researcher, CSIR-IMTech

Mayur Zarkar, Project Associate I, CSIR-IMTech

Raghav Gankhedar, Senior Project Associate, CSIR-IMTech

Dr. Bhupender Singh, Project Scientist I, CSIR-IMTech

Shivani Seth, Junior Research Fellow, CSIR-IMTech

Dr. Anshu Bhardwaj, Project PI, CSIR-IMTech

The Connect page is divided into two sections: the left section features community contributions, while the right section displays team members. The community contribution section invites researchers, clinicians, and domain experts to contribute to ARKbase, an open-access Antimicrobial Resistance Knowledgebase. The

submission allows contributions exclusively for genotype and phenotype data related to antimicrobial resistance in WHO Bacterial Priority Pathogens. For genotype data, users can select a specific pathogen from the BPPL list and provide a valid Genome Accession Number (e.g., from NCBI), and include the corresponding accession URL. For phenotype data, users must provide Antimicrobial Susceptibility Testing (AST) information for one or more antibiotics, including either Minimum Inhibitory Concentration (MIC) or Disk Diffusion (DD) values.

The unique features of ARKbase

ARKbase is a unique resource linking curated datasets with multi-omics, systems biology, AI/ML methods for predicting AMR determinants, and Foldseek/AlphaFold-based approaches towards understanding the molecular basis of resistance, identification of novel drug resistance determinants, and prioritizing potential therapeutic targets. It offers several unique features, like implementation of CLSI/EUCAST standards for reporting AST and genome datasets, identification of co-targets, providing a structured database for ARGs, offering a curated list of AI/ML models as per AWaRe classification, operon mapping of AMR and virulence genes, systems-level data for host-pathogen interactions, and PPIs for WHO priority pathogens.

ARKbase is an open database

Thank you