

# How To: Download all the reference sequences for a set of proteins (e.g., all *aac(6')-Ib* sequences)

NCBI Pathogen Detection

<https://www.ncbi.nlm.nih.gov/pathogens>



U.S. National Library of Medicine  
National Center for Biotechnology Information

# How do I download all the reference sequences for a set of proteins? (e.g., all *aac(6')-Ib* sequences)

- Use [MicroBIGG-E](#)
- Use filters or search to select gene family “*aac(6')-Ib\**”
- Download table containing protein accessions
- Use Batch Entrez to download proteins

# Pathogen Detection BETA



To assist the National Database of Antibiotic Resistant Organisms (NDARO), NCBI Pathogen Detection identifies the antimicrobial resistance, stress response, and virulence genes found in bacterial genomic sequences. This enables scientists to track the spread of resistance genes and to understand the relationships between antimicrobial resistance and virulence.

NCBI Pathogen Detection integrates bacterial pathogen genomic sequences originating in food, environmental sources, and patients. It quickly clusters and identifies related sequences to uncover potential food contamination sources, helping public health scientists investigate foodborne disease outbreaks.



There has been a change to the *Isolation type / epi\_type* attribute that affects *min-same/min-diff* computation. Now where an isolate has no information to support the setting *environmental/other* the Browser will no longer default to that value but instead present as *NULL*. This means that the *min-same/min-diff* values for this isolate will present as *n/a*, and other *min-same/min-diff* values for isolates clustered with this isolate may change. Please see the [Help](#) text for more details.

## Learn More

[About](#)

[FAQ](#)

[Browser Factsheet](#)

[Antimicrobial Resistance Factsheet](#)

[Antimicrobial Resistance](#)

[Contributors](#)

[Help](#)

## Data Resources

[Isolates Browser](#)

[Microbial Browser](#)

[Genomic Elements \(MIR, IS, CRISPR, etc.\)](#)

[Reference Gene Catalog](#)

NEW [Reference HMM Catalog](#)

1. Click Reference Gene Catalog

Search

2. Click Filters bar to show filters

db version: 2021-08-11.1 [Changelog](#)

Filters **4**

Available filters

Filter
<input type="checkbox"/> allele
<input checked="" type="checkbox"/> gene family
<input type="checkbox"/> product name
<input type="checkbox"/> scope
<input type="checkbox"/> subclass
<input type="checkbox"/> curated refseq start
<input type="checkbox"/> whitelisted taxa
<input type="checkbox"/> blacklisted taxa
<input type="checkbox"/> genbank strand orientation
<input type="checkbox"/> genbank cds start
<input type="checkbox"/> genbank cds stop
<input type="checkbox"/> pubmed reference
<input type="checkbox"/> synonyms

3. Select gene family to show the gene family filter

4 gene family

Search

gene family	Count
<input checked="" type="checkbox"/> aac(6')-Ib	
<input checked="" type="checkbox"/> aac(6')-Ib-cr	10
<input checked="" type="checkbox"/> aac(6')-Ib11	2
<input type="checkbox"/> aac(3)-Ib/aac(6')-Ib"	1
<input type="checkbox"/> aac(6')-30/aac(6')-Ib'	1
<input checked="" type="checkbox"/> aac(6')-Ib'	1
<input type="checkbox"/> ant(3'')-Ij/aac(6')-Ib	1
<input type="checkbox"/> catB8/aac(6')-Ib'	1

4. Enter "aac(6')-Ib" and select the gene families of interest

Total unique values: 8

Search

gene\_family:aac(6')-Ib\*

5. Alternatively, use the search term "gene\_family" to find the aac(6')-Ib family members

db version: 2021-08-11.1 [Changelog](#)

Filters

Page 1 of 4 | Records per Page 20 | Choose columns | Download

#	allele	gene family	product name	scope	type	subtype	class
1		<a href="#">aac(6')-Ib</a>	AAC(6')-Ib family aminoglycoside 6'-N-acetyltransferase	core	AMR	AMR	AMIN
2		<a href="#">aac(6')-Ib</a>	AAC(6')-Ib family aminoglycoside 6'-N-acetyltransferase	core	AMR	AMR	AMIN
3		<a href="#">aac(6')-Ib</a>	AAC(6')-Ib family aminoglycoside 6'-N-acetyltransferase	core	AMR	AMR	AMIN
4		<a href="#">aac(6')-Ib</a>	AAC(6')-Ib family aminoglycoside 6'-N-acetyltransferase	core	AMR	AMR	AMIN
5		<a href="#">aac(6')-Ib</a>	AAC(6')-Ib family aminoglycoside 6'-N-acetyltransferase	core	AMR	AMR	AMIN
6		<a href="#">aac(6')-Ib</a>	AAC(6')-Ib family aminoglycoside 6'-N-acetyltransferase	core	AMR	AMR	AMIN
7		<a href="#">aac(6')-Ib</a>	AAC(6')-Ib family aminoglycoside 6'-N-acetyltransferase	core	AMR	AMR	AMIN
8		<a href="#">aac(6')-Ib</a>	AAC(6')-Ib family aminoglycoside 6'-N-acetyltransferase	core	AMR	AMR	AMIN
9		<a href="#">aac(6')-Ib</a>	AAC(6')-Ib family aminoglycoside 6'-N-acetyltransferase	core	AMR	AMR	AMIN
10		<a href="#">aac(6')-Ib</a>	AAC(6')-Ib family aminoglycoside 6'-N-acetyltransferase	core	AMR	AMR	AMIN
11		<a href="#">aac(6')-Ib</a>	AAC(6')-Ib family aminoglycoside 6'-N-acetyltransferase	core	AMR	AMR	AMIN
12		<a href="#">aac(6')-Ib</a>	AAC(6')-Ib family aminoglycoside 6'-N-acetyltransferase	core	AMR	AMR	AMIN
13		<a href="#">aac(6')-Ib</a>	AAC(6')-Ib family aminoglycoside 6'-N-acetyltransferase	core	AMR	AMR	AMIN
14		<a href="#">aac(6')-Ib</a>	AAC(6')-Ib family aminoglycoside 6'-N-acetyltransferase	core	AMR	AMR	AMIN
15		<a href="#">aac(6')-Ib</a>	AAC(6')-Ib family aminoglycoside 6'-N-acetyltransferase	core	AMR	AMR	AMIN
16		<a href="#">aac(6')-Ib</a>	AAC(6')-Ib family aminoglycoside 6'-N-acetyltransferase	core	AMR	AMR	AMIN
17		<a href="#">aac(6')-Ib</a>	AAC(6')-Ib family aminoglycoside 6'-N-acetyltransferase	core	AMR	AMR	AMIN
18		<a href="#">aac(6')-Ib</a>	AAC(6')-Ib family aminoglycoside 6'-N-acetyltransferase	core	AMR	AMR	AMIN
19		<a href="#">aac(6')-Ib</a>	AAC(6')-Ib family aminoglycoside 6'-N-acetyltransferase	core	AMR	AMR	AMIN

Search

db version: 2021-08-11.1 [Changelog](#)

Filters **4**

Available filters ↻

Filter
<input type="checkbox"/> allele
<input checked="" type="checkbox"/> gene family
<input type="checkbox"/> product name
<input type="checkbox"/> scope
<input type="checkbox"/> type
<input type="checkbox"/> subtype
<input type="checkbox"/> class
<input type="checkbox"/> subclass
<input type="checkbox"/> curated refseq start
<input type="checkbox"/>
<input type="checkbox"/>
<input type="checkbox"/>
<input type="checkbox"/>
<input type="checkbox"/>
<input type="checkbox"/> pubmed reference
<input type="checkbox"/> synonyms

Search  ↻ ✕

gene family	Count
<input checked="" type="checkbox"/> aac(6')-Ib	55
<input checked="" type="checkbox"/> aac(6')-Ib-cr	10
<input checked="" type="checkbox"/> aac(6')-Ib11	2
<input type="checkbox"/> aac(3)-Ib/aac(6')-Ib"	1
<input type="checkbox"/> aac(6')-30/aac(6')-Ib'	1
<input checked="" type="checkbox"/> aac(6')-Ib'	1
<input type="checkbox"/> ant(3'')-Ij/aac(6')-Ib	1
<input type="checkbox"/> catB8/aac(6')-Ib'	1

Total unique values: 8

6. Select Download to download table

⏪
⏴
Page 1 of 4
⏵
⏩
↻
Records per Page 20
☰ Choose columns
Download

#	allele	gene family	product name	scope	type	subtype	class
1		<a href="#">aac(6')-Ib</a>	AAC(6')-Ib family aminoglycoside 6'-N-acetyltransferase	core	AMR	AMR	AMIN
2		<a href="#">aac(6')-Ib</a>	AAC(6')-Ib family aminoglycoside 6'-N-acetyltransferase	core	AMR	AMR	AMIN

Available filters

- allele
- gene family
- product name
- scope
- type
- subtype
- class
- subclass
- curated refseq start
- whitelisted taxa
- blacklisted taxa
- genbank strand orientation
- ge
- ge
- pu
- sy

Search aac(6')-Ib

gene family	Count
<input checked="" type="checkbox"/> aac(6')-Ib	55
<input checked="" type="checkbox"/> aac(6')-Ib-cr	10
<input checked="" type="checkbox"/> aac(6')-Ib11	2
<input type="checkbox"/> aac(3)-Ib/aac(6')-Ib"	1
<input type="checkbox"/> aac(6')-30/aac(6')-Ib'	1
<input checked="" type="checkbox"/> aac(6')-Ib'	1
<input type="checkbox"/> ant(3'')-Ij/aac(6')-Ib	1
<input type="checkbox"/> catB8/aac(6')-Ib'	1

Total unique values: 8

7. Click Download to download table

Page 1 of 4 | Records per Page 20 | Choose columns | Download

#	allele	gene family	product name	type	subtype	class
1		aac(6')-Ib	AAC(6')-Ib family aminoglycos	AMR	AMR	AMIN
2		aac(6')-Ib	AAC(6')-Ib family aminoglycos	AMR	AMR	AMIN
3		aac(6')-Ib	AAC(6')-Ib family aminoglycos	AMR	AMR	AMIN
4		aac(6')-Ib	AAC(6')-Ib family aminoglycos	AMR	AMR	AMIN
5		aac(6')-Ib	AAC(6')-Ib family aminoglycos	AMR	AMR	AMIN
6		aac(6')-Ib	AAC(6')-Ib family aminoglycoside 6'-N-acetyltransferase	AMR	AMR	AMIN

Download

Type: Tab-delimited (.tsv)

Name: refgenes.tsv

Download Cancel

gene family	product name	scope	type	subtype	class	subclass	refseq protein	refseq
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/G	WP_071846254.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/G	WP_065187201.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/K	WP_063840283.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/K	WP_063978482.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/K	WP_063840319.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/K	WP_071846332.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/K	WP_101516660.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/K	WP_085844237.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/K	WP_085843736.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/K	WP_063612062.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/K	WP_071846208.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/K	WP_071846209.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/K	WP_071846373.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/K	WP_071846241.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/K	WP_064765190.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/K	WP_071593230.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/K	WP_063840280.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/K	WP_069067473.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/K	WP_063840281.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/K	WP_159287722.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/K	WP_113613509.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/K	WP_123085936.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/K	WP_079452738.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/K	WP_132629969.1	NG_01
aac(6)-Ib	AAC(6)-Ib family core		AMR	AMR	AMINOGLYC	AMIKACIN/K	WP_117046752.1	NG_01



```


WP_071846254.1
WP_065187201.1
WP_063840283.1
WP_063978482.1
WP_063840319.1
WP_071846332.1
WP_101516660.1
WP_085844237.1
WP_085843736.1
WP_063612062.1
WP_071846208.1
WP_071846209.1
WP_071846373.1
WP_071846241.1

```

8. Place the protein accessions into a single column text file, with one accession per row. Then go to Batch Entrez



All Databases PubMed Nucleotide Protein  
 Database Protein File: Choose File aac6primelbaccessions.txt Retrieve


**COVID-19 Information**  
[Public health information \(CDC\)](#) | [Research information \(NIH\)](#) | [SARS-CoV-2 data \(NCBI\)](#) | [Prevention a](#)

9. Select the Protein database.

10. Upload the text file containing the accession.

Batch Entrez downloads the

11. Select Retrieve.

### Instructions

1. Start with a local file containing a list of accession numbers or identifiers
2. Select the database corresponding to the type of accession numbers or identifiers in your input file
3. Use the **Browse** or **Choose File...** button to select the input file
4. Press the **Retrieve** button to see a list of document summaries
5. Select a format in which to display the data for viewing, and/or saving
6. Select 'Send to file' to save the file.

Received lines: 68

Rejected lines: 0

Removed duplicates: 0

Passed to Entrez: 68

[Retrieve records for 68 UID\(s\).](#)



12. Select Retrieve.



## COVID-19 Information

[Public health information \(CDC\)](#) | [Research information \(NIH\)](#) | [SARS-CoV-2 data \(NCBI\)](#) | [Prevention and treatment information \(HHS\)](#) | [Español](#)

### Species

Bacteria (67)

[Customize ...](#)

### Source databases

RefSeq (68)

[Customize ...](#)

### Sequence length

[Custom range...](#)

### Molecular weight

[Custom range...](#)

### Release date

[Custom range...](#)

### Revision date

[Custom range...](#)

[Clear all](#)

[Show additional filters](#)

Summary ▾ 20 per page ▾ Sort by Default order ▾

Send to: ▾

Filters: [Manage Filters](#)

Items: 1 to 20 of 68

15. Select the "Send to:"  
pulldown menu.

- [MULTISPECIES: aminoglycoside N-acetyltransferase \[Proteobacteria\]](#)
1. 184 aa protein  
Accession: WP\_003159191.1 GI: 489251133  
[BioProject](#) [Nucleotide](#) [Taxonomy](#)  
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

- [MULTISPECIES: AAC\(6'\)-Ib family aminoglycoside 6'-N-acetyltransferase \[Proteobacteria\]](#)
2. 184 aa protein  
Accession: WP\_012695458.1 GI: 502019073  
[BioProject](#) [Nucleotide](#) [Taxonomy](#)  
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

- [MULTISPECIES: AAC\(6'\)-Ib family aminoglycoside 6'-N-acetyltransferase \[Enterobacteriaceae\]](#)
3. 184 aa protein  
Accession: WP\_013250882.1 GI: 503015906  
[BioProject](#) [Nucleotide](#) [Taxonomy](#)

1 of 4

[Next >](#)

[Last >>](#)

## Results by taxon

### Top Organisms

[Pseudomonas aeruginosa](#) (2)  
[Klebsiella pneumoniae](#) (2)  
[Enterobacteriaceae](#) (2)  
[g-proteobacteria](#) (2)  
[Escherichia coli](#) (1)  
[All other taxa](#) (31)  
[More...](#)

## Find related data

Database:

[Find items](#)

## Recent activity

[DNA-binding tra](#)



## COVID-19 Information

[Public health information \(CDC\)](#) | [Research information \(NIH\)](#) | [SARS-CoV-2 data \(NCBI\)](#) | [Prevention and treatment information \(HHS\)](#) | [Español](#)

### Species

Bacteria (67)

Customize ...

### Source databases

RefSeq (68)

Customize ...

### Sequence length

Custom range...

### Molecular weight

Custom range...

### Release date

Custom range...

### Revision date

Custom range...

[Clear all](#)

[Show additional filters](#)

Summary ▾ 20 per page ▾

Items: 1 to 20 of 68

[MULTISPECIES: aminoglycoside N-acetyltransferase AAC\(6\)-Ib4 \[Proteobacteria\]](#)

1. 184 aa protein

Accession: WP\_003159191.1 GI: 489251133

[BioProject](#) [Nucleotide](#) [Taxonomy](#)

[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

[MULTISPECIES: AAC\(6\)-Ib family aminoglycoside 6'-N-acetyltransferase \[Enterobacteriaceae\]](#)

2. 184 aa protein

Accession: WP\_012695458.1

[BioProject](#) [Nucleotide](#) [Taxonomy](#)

[GenPept](#) [Identical Proteins](#)

[MULTISPECIES: AAC\(6\)-Ib family aminoglycoside 6'-N-acetyltransferase \[Enterobacteriaceae\]](#)

3. 184 aa protein

Accession: WP\_013250882.1 GI: 503015906

[BioProject](#) [Nucleotide](#) [Taxonomy](#)

14. Select FASTA.

15. Select Create File.

### Choose Destination

File

Clipboard

Collections

Download 68 items.

Format

FASTA ▾

Sort by

Default order ▾

Show GI

Create File

Find items

Recent activity

DNA-binding tra

# More information

- For full help documentation of the Reference Gene Catalog see:  
[https://www.ncbi.nlm.nih.gov/pathogens/pathogens\\_help/#reference-gene-catalog](https://www.ncbi.nlm.nih.gov/pathogens/pathogens_help/#reference-gene-catalog)
- For details about filters see:  
[https://www.ncbi.nlm.nih.gov/pathogens/pathogens\\_help/#refgene-filters](https://www.ncbi.nlm.nih.gov/pathogens/pathogens_help/#refgene-filters)
- For details about the table downloads see:  
[https://www.ncbi.nlm.nih.gov/pathogens/pathogens\\_help/#refgene-access-download](https://www.ncbi.nlm.nih.gov/pathogens/pathogens_help/#refgene-access-download)

Questions and further help: email [pd-help@ncbi.nlm.nih.gov](mailto:pd-help@ncbi.nlm.nih.gov)