



A joint venture between The University of Melbourne and The Royal Melbourne Hospital

Genomic Databases

Ammar Aziz

Title

- Words

Objectives

- Introduction to genomic databases: NCBI Virus, PathoPlexus and GISAID
- Comparison between the databases
- Use PathoPlexus to search and download MPXV data (meta and sequencing)
- Use NCBI Virus to search for other (non-MPXV) viruses

What is a genomic database?

- A collection of nucleic acid sequences for any organism (human, virus, bacterial, etc)
- Database also contain meta (non-sequencing) data that is very important.
- Used every day, genomic databases power:
 - NCBI BLAST
 - Nextstrain/Nextclade
 - Primer design
 - Surveillance reporting
 - And many more

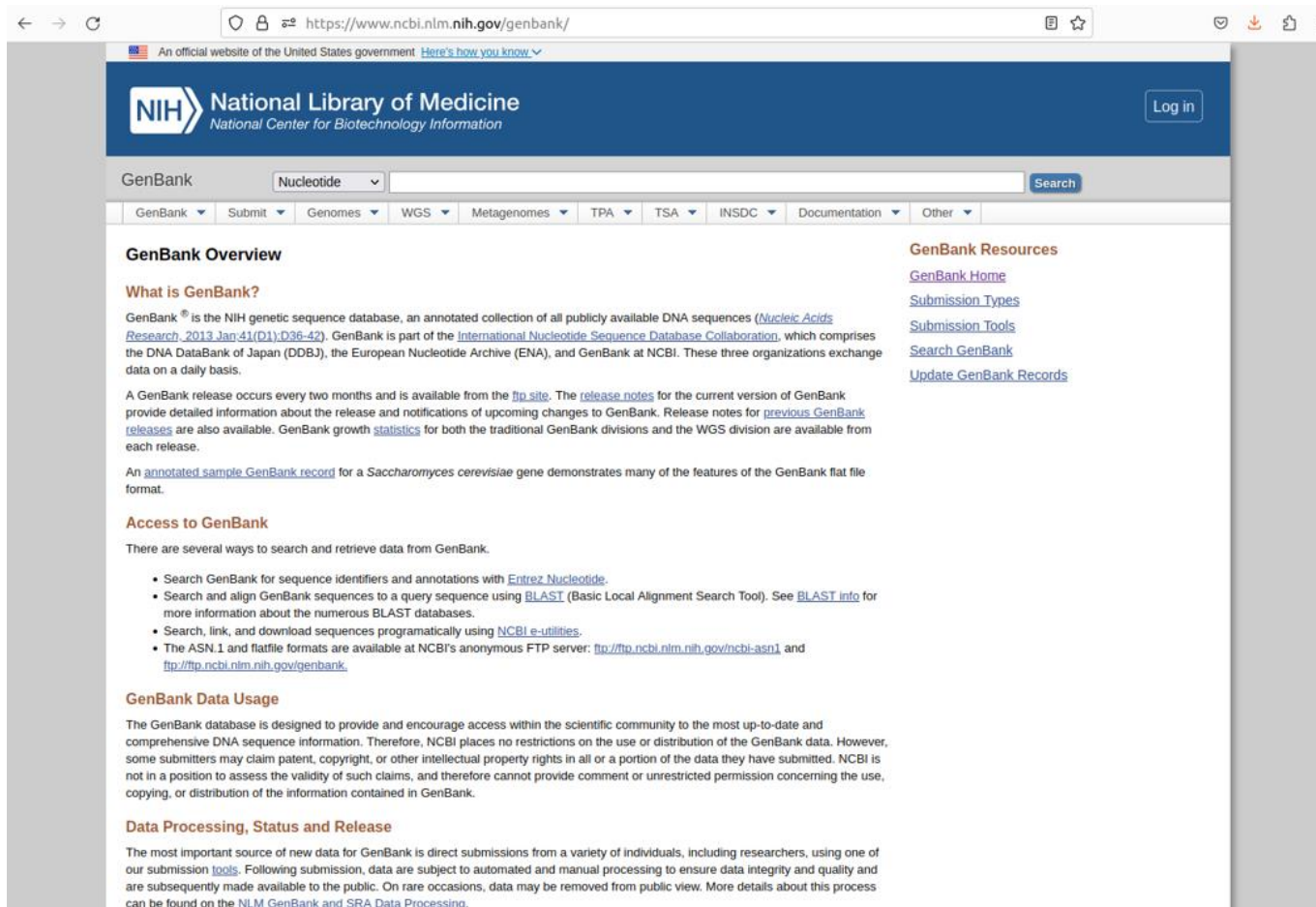
Databases for Viruses

General database:

- The INSDC (International Nucleotide Sequence Database Collaboration) - same data different interface
 - NCBI
 - ENA
 - DDBJ
- Virus specific:
 - NCBI Virus
 - GISAID
 - PathoPlexus

Note: there are many different databases, we will only focus on the most relevant to MPXV

INSDC - NCBI



The screenshot shows the NCBI GenBank homepage. At the top, there's a navigation bar with the NIH logo and 'National Library of Medicine' text. Below this is a search bar with 'Nucleotide' selected. A horizontal menu contains links for GenBank, Submit, Genomes, WGS, Metagenomes, TPA, TSA, INSDC, Documentation, and Other. The main content area is titled 'GenBank Overview' and includes a 'What is GenBank?' section with a paragraph about the database's history and purpose. To the right, there's a 'GenBank Resources' section with links to GenBank Home, Submission Types, Submission Tools, Search GenBank, and Update GenBank Records. Below the overview, there's an 'Access to GenBank' section with a list of search methods. Further down, 'GenBank Data Usage' and 'Data Processing, Status and Release' sections provide information on data access and submission processes.

GenBank Overview

What is GenBank?

GenBank[®] is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences ([Nucleic Acids Research, 2013 Jan 41\(D1\):D36-42](#)). GenBank is part of the [International Nucleotide Sequence Database Collaboration](#), which comprises the DNA DataBank of Japan (DDBJ), the European Nucleotide Archive (ENA), and GenBank at NCBI. These three organizations exchange data on a daily basis.

A GenBank release occurs every two months and is available from the [ftp site](#). The [release notes](#) for the current version of GenBank provide detailed information about the release and notifications of upcoming changes to GenBank. Release notes for [previous GenBank releases](#) are also available. GenBank growth [statistics](#) for both the traditional GenBank divisions and the WGS division are available from each release.

An [annotated sample GenBank record](#) for a *Saccharomyces cerevisiae* gene demonstrates many of the features of the GenBank flat file format.

Access to GenBank

There are several ways to search and retrieve data from GenBank.

- Search GenBank for sequence identifiers and annotations with [Entrez Nucleotide](#).
- Search and align GenBank sequences to a query sequence using [BLAST](#) (Basic Local Alignment Search Tool). See [BLAST info](#) for more information about the numerous BLAST databases.
- Search, link, and download sequences programmatically using [NCBI e-utilities](#).
- The ASN.1 and flatfile formats are available at NCBI's anonymous FTP server: <ftp://ftp.ncbi.nlm.nih.gov/ncbi-asn1> and <ftp://ftp.ncbi.nlm.nih.gov/genbank>.

GenBank Data Usage

The GenBank database is designed to provide and encourage access within the scientific community to the most up-to-date and comprehensive DNA sequence information. Therefore, NCBI places no restrictions on the use or distribution of the GenBank data. However, some submitters may claim patent, copyright, or other intellectual property rights in all or a portion of the data they have submitted. NCBI is not in a position to assess the validity of such claims, and therefore cannot provide comment or unrestricted permission concerning the use, copying, or distribution of the information contained in GenBank.

Data Processing, Status and Release

The most important source of new data for GenBank is direct submissions from a variety of individuals, including researchers, using one of our submission [tools](#). Following submission, data are subject to automated and manual processing to ensure data integrity and quality and are subsequently made available to the public. On rare occasions, data may be removed from public view. More details about this process can be found on the [NCBI GenBank and SRA Data Processing](#).

NCBI has many databases:

- Genbank/Nucleotide – Consensus sequences
- RefSeq – curated reference sequences
- SRA – raw read sequencing data
- Taxonomy – Taxnomic data
- PubMed – Publications

NCBI also has lots of bioinformatics tools such as BLAST.

- Allows for uploading a variety of data.
- Disadvantages include sometimes complex data uploading process.
- But simplified for key organisms:
 - SC2
 - Influenza
 - MPXV
 - others

GISAID – Global Initiative on Sharing Avian Influenza Data



- Established to provide access to genomic data of influenza viruses.
- Expanded to include the SC2, RSV, MPXV, and other viruses.
- Simplifies uploading of sequencing and metadata
- Provides protection for data creators by forcing users to credit in public work.
- Not open access but open registration

- Pathoplexus is a new, open-source database for sharing of human viral pathogen genomic data.
 - Uses NCBI Genbank data
- Simplifies uploading viral data.
 - Deposits sequences back into NCBI
- Provides strong protection for data generators
 - users can choose to share their data openly, or with time-limited protections to help ensure proper attribution and credit.
- Free to download data without an account
- Need an account to upload
- Supports different viruses: MPXV, RSV, Ebola, HMPV

NCBI Virus

- Uses the sequencing data of NCBI Genbank/Nucleotide
- User friendly interface
- Allows for bulk downloads of viral data
- Not for uploading – only for access

NCBI Virus

Community portal for viral sequence data from RefSeq, GenBank and other NCBI repositories.
To find, retrieve and analyze data, please select an option below.



Search by virus

Use virus name or taxid to find viral nucleotide and protein sequences.



NCBI BLAST™ search

Find viral nucleotide and protein sequences using the BLAST™ sequence similarity tool.

Search GenBank virus/viroid sequences and explore metadata in Results Table with advanced filtering and data visualizations options

Begin typing a virus name, viral taxonomy group, or taxid to select from a list of suggestions.

Popular searches

All viruses

Human viruses

H5N1 Influenza A virus

Released in past month

SARS-CoV-2

Which database to use and when?

- For sharing MPXV sequencing data – your choice:
 - PathoPlexus
 - GISAID
- For bulk download data:
 - NCBI Virus or PathoPlexus are easiest to access
 - GISAID downloading can be complicated due to slow speeds and limitation on download (generally 5000 sequences)

Demonstration of each database.


Demo: Using GISAID for MPXV

Website: www.gisaid.org


Registered Users | EpiFlu™ | EpiCoV™ | EpiRSV™ | **EpiPox™** | EpiArbo™ | My Profile

Mpox | Search | Upload


Surveillance of MPox viruses




Emerging Variants



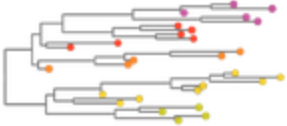
EpiCharts



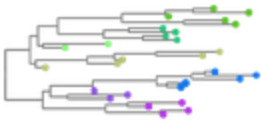
PoxSurver



Search & Browse



Phylogeny
Clade I




Phylogeny
Clade II

Demo: Using GISAID for MPXV - Searching


Registered Users | EpiFlu™ | EpiCoV™ | EpiRSV™ | **EpiPox™** | EpiArbo™ | My Profile

Mpox | Search | Upload


Surveillance of MPox viruses




Emerging Variants



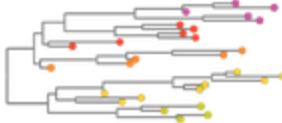
EpiCharts



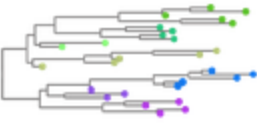
PoxSurver



Search & Browse



Phylogeny
Clade I



Phylogeny
Clade II

Demo: Using GISAID for MPXV - Searching

Mpox | Search | Upload

Search MPox

Accession ID Virus name

Location Host

Collection to Submission to

Clade / Lineage Substitutions

Publications 10.1590/... or select publication →

☐ Cpl. CDS ☐ low cov. excl ☐ high cov. ☐ compl coll. date

<input type="checkbox"/>	VIRUS NAME	PASSAGE	ACCESSION ID	COLLECTION	SUBMISSION		LENGTH	LOCATION	ORIGINATING LAB
<input type="checkbox"/>	hMpxV/USA/IL-RIPHL-004-0223/2025	Original	EPI_ISL_19891609	2025-04-14	2025-06-02		196,151	North America / U.S.	Northwestern Medical Center
<input type="checkbox"/>	hMpxV/USA/IL-RIPHL-050-0222/2025	Original	EPI_ISL_19891608	2025-04-15	2025-06-02		196,288	North America / U.S.	Quest Diagnostics
<input type="checkbox"/>	hMpxV/USA/IL-RIPHL-050-0221/2025	Original	EPI_ISL_19891607	2025-04-17	2025-06-02		195,611	North America / U.S.	Quest Diagnostics
<input type="checkbox"/>	hMpxV/USA/IL-RIPHL-050-0220/2024	Original	EPI_ISL_19891606	2025-04-17	2025-06-02		195,589	North America / U.S.	Quest Diagnostics
<input type="checkbox"/>	hMpxV/USA/IL-RIPHL-050-0219/2024	Original	EPI_ISL_19891605	2025-03-28	2025-06-02		195,591	North America / U.S.	Quest Diagnostics
<input type="checkbox"/>	hMpxV/Togo/BMV-09/2025	Original	EPI_ISL_19888233	2025-05-16	2025-05-30		197,044	Africa / Togo / Gr	Institut National d'H
<input type="checkbox"/>	hMpxV/Togo/BMV-17/2025	Original	EPI_ISL_19888238	2025-05-20	2025-05-30		197,044	Africa / Togo / Gr	Institut National d'H
<input type="checkbox"/>	hMpxV/Togo/BMV-24/2025	Original	EPI_ISL_19888241	2025-05-22	2025-05-30		197,044	Africa / Togo / Gr	Institut National d'H
<input type="checkbox"/>	hMpxV/Togo/BMV-19/2025	Original	EPI_ISL_19888240	2025-05-20	2025-05-30		197,028	Africa / Togo / Ag	Institut National d'H
<input type="checkbox"/>	hMpxV/Pakistan/DUHS-01/2025	Original	EPI_ISL_19883106	2025-03-22	2025-05-27		197,143	Asia / Pakistan /	Provincial Public H
<input type="checkbox"/>	hMpxV/USA/WA-UW-52066/2025	Original	EPI_ISL_19890345	2025-04-22	2025-05-23		197,184	North America / U.S.	Laboratory Medicir

Demo: GISAID – Download data

Note: There is a limit on the number of sequences you can download.

Message

Select 5,000 submissions at maximum!

OK

Accession ID

Virus name

☐ Cpl. CDS ? ☐ low cov. excl ? ☐ high cov. ?

Location

Host

☐ compl coll. date ?


Collection to

Submission to

Clade / Lineage

Substitutions

Publications



Reset

Fulltext ▲

<input type="checkbox"/>	VIRUS NAME	PASSAGE	ACCESSION ID	COLLECTION	SUBMISSION	?	LENGTH	LOCATION	ORIGINATING LAB
<input checked="" type="checkbox"/>	hMpxV/Australia/NSW-030/2025	Original	EPI_ISL_19876908	2025-05	2025-05-18	?	190,977	Oceania / Austra	Douglass Hanly M
<input checked="" type="checkbox"/>	hMpxV/Australia/NSW-002/2024	Original	EPI_ISL_19876907	2024-05	2025-05-18	?	192,484	Oceania / Austra	NSWHP ICPMR
<input checked="" type="checkbox"/>	hMpxV/Australia/NSW-145/2024	Original	EPI_ISL_19876906	2024-09	2025-05-18	?	192,504	Oceania / Austra	Lavery Pathology
<input checked="" type="checkbox"/>	hMpxV/Australia/NSW-139/2024	Original	EPI_ISL_19876905	2024-09	2025-05-18	?	188,132	Oceania / Austra	Lavery Pathology
<input checked="" type="checkbox"/>	hMpxV/Australia/NSW-138/2024	Original	EPI_ISL_19876904	2024-09	2025-05-18	?	197,193	Oceania / Austra	NSWHP ICPMR
<input checked="" type="checkbox"/>	hMpxV/Australia/NSW-133/2024	Original	EPI_ISL_19876903	2024-09	2025-05-18	?	197,116	Oceania / Austra	Australian Clinical
<input checked="" type="checkbox"/>	hMpxV/Australia/NSW-132/2024	Original	EPI_ISL_19876902	2024-09	2025-05-18	?	189,426	Oceania / Austra	NSWHP ICPMR
<input type="checkbox"/>	hMpxV/Australia/NSW-130/2024	Original	EPI_ISL_19876901	2024-09	2025-05-18	?	186,751	Oceania / Austra	NSWHP ICPMR
<input type="checkbox"/>	hMpxV/Australia/NSW-127/2024	Original	EPI_ISL_19876900	2024-09	2025-05-18	?	185,472	Oceania / Austra	NSWHP ICPMR
<input type="checkbox"/>	hMpxV/Australia/NSW-126/2024	Original	EPI_ISL_19876899	2024-09	2025-05-18	?	197,200	Oceania / Austra	NSWHP ICPMR

Demo: GISAID – Download data

Total: 480 viruses

<< < 1 2 3 4 5 > >>

EPI_SET

Select

Analysis

Download

Download

Format

☒ Sequences (FASTA)

☐ Patient status metadata

☐ Sequencing technology metadata

☐ Dates and Location

☒ Replace spaces with underscores in FASTA header

For Acknowledgement and Data Availability use the [EPI_SET data generator](#)

Back

Download

- **Most important:**
 - Sequences (FASTA) option for consensus sequences (.fasta)
 - Dates and Locations (.csv)

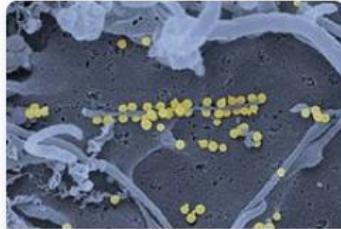
Click Download to begin downloading.
It can take a long time to download.

Demo: Using PathoPlexus

Website:
<https://pathoplexus.org>

Welcome to Pathoplexus!

Pathoplexus is a new, open-source database dedicated to the efficient sharing of human viral pathogen genomic data, fostering global collaboration and public health response.



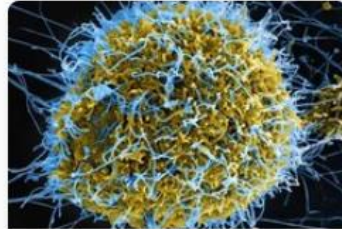
Crimean-Congo Hemorrhagic Fever Virus

2,765 sequences
+3 in last 30 days



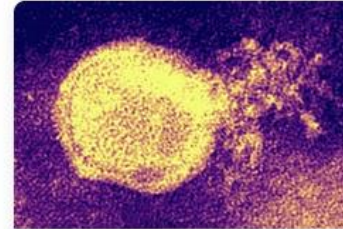
Ebola Sudan

166 sequences
+0 in last 30 days



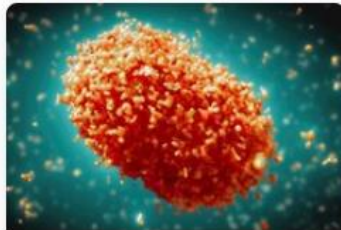
Ebola Zaire

3,698 sequences
+0 in last 30 days



HMPV

12,215 sequences
+12,215 in last 30 days



Mpox Virus

10,520 sequences
+202 in last 30 days



RSV-A

32,537 sequences
+32,537 in last 30 days



RSV-B

23,880 sequences
+23,880 in last 30 days



West Nile Virus

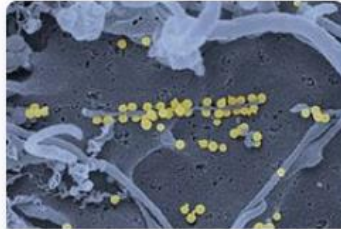
7,901 sequences
+71 in last 30 days

Demo: Using PathoPlexus

Website:
<https://pathoplexus.org>

Welcome to Pathoplexus!

Pathoplexus is a new, open-source database dedicated to the efficient sharing of human viral pathogen genomic data, fostering global collaboration and public health response.



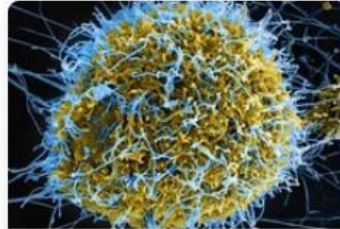
Crimean-Congo Hemorrhagic Fever Virus

2,765 sequences
+3 in last 30 days



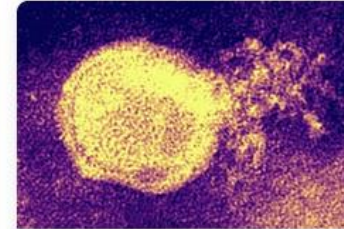
Ebola Sudan

166 sequences
+0 in last 30 days



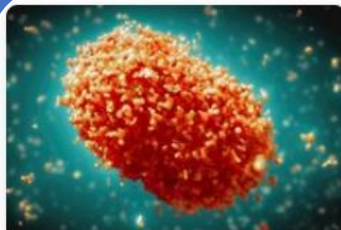
Ebola Zaire

3,698 sequences
+0 in last 30 days



HMPV

12,215 sequences
+12,215 in last 30 days



Mpox Virus

10,520 sequences
+202 in last 30 days



RSV-A

32,537 sequences
+32,537 in last 30 days



RSV-B

23,880 sequences
+23,880 in last 30 days



West Nile Virus

7,901 sequences
+71 in last 30 days

Demo: Using PathoPlexus

Website:

<https://pathoplexus.org>

Search

[Add Search Fields](#) [Reset](#) [Help](#)

Accession

Mutations

Data use terms

Clade

Collection subdivision level 1

Collection country

Length

From

To

Lineage

Collection date

strict ☒

From dd/mm/yyyy

To dd/mm/yyyy

Search returned 10,520 sequences

[Customize columns](#)

[Download all entries](#)

	ACCESSION VERSION	CLADE	LINEAGE	COLLECTION DATE ▼	COLLECTION COUNTRY	COLLECTION SUBDIVISION LEVEL 1	AUTHORS	AUTHOR AFFILIATIONS	LENGTH	EARLIEST RELEASE DATE
<input type="checkbox"/>	PP_002XLCT.1	I1b	A.2.2	2025-05-20	Sierra Leone	Kenema	Sandi, John D.;...	National Public...	197,136	2025-05-27
<input type="checkbox"/>	PP_002Y2Z0.1	I1b	A.2.2	2025-05-07	Sierra Leone	Western Area Ur...	Campbell, Allan...	Central Public ...	197,180	2025-05-31
<input type="checkbox"/>	PP_002Y2SE.1	I1b	A.2.2	2025-05-07	Sierra Leone	Western Area Ur...	Campbell, Allan...	Central Public ...	197,184	2025-05-31
<input type="checkbox"/>	PP_002Y2V8.1	I1b	A.2.2	2025-05-07	Sierra Leone	Western Area Ur...	Campbell, Allan...	Central Public ...	197,178	2025-05-31
<input type="checkbox"/>	PP_002Y2TC.1	I1b	A.2.2	2025-05-07	Sierra Leone	Western Area Ur...	Campbell, Allan...	Central Public ...	197,189	2025-05-31
<input type="checkbox"/>	PP_002Y2MQ.1	I1b	A.2.2	2025-05-07	Sierra Leone	Western Area Ur...	Campbell, Allan...	Central Public ...	197,174	2025-05-31
<input type="checkbox"/>	PP_002Y30Y.1	I1b	A.2.2	2025-05-07	Sierra Leone	Western Area Ur...	Campbell, Allan...	Central Public ...	197,189	2025-05-31
<input type="checkbox"/>	PP_002Y2RG.1	I1b	A.2.2	2025-05-07	Sierra Leone	Western Area Ur...	Campbell, Allan...	Central Public ...	197,180	2025-05-31
<input type="checkbox"/>	PP_002Y2PL.1	I1b	A.2.2	2025-05-07	Sierra Leone	Western Area Ur...	Campbell, Allan...	Central Public ...	197,180	2025-05-31
<input type="checkbox"/>	PP_002Y2NN.1	I1b	A.2.2	2025-05-07	Sierra Leone	Western Area Ur...	Campbell, Allan...	Central Public ...	197,180	2025-05-31
<input type="checkbox"/>	PP_002Y2W6.1	I1b	A.2.2	2025-05-07	Sierra Leone	Western Area Ur...	Campbell, Allan...	Central Public ...	197,181	2025-05-31
<input type="checkbox"/>	PP_002Y2X4.1	I1b	A.2.2	2025-05-07	Sierra Leone	Western Area Ur...	Campbell, Allan...	Central Public ...	197,183	2025-05-31
<input type="checkbox"/>	PP_002Y2UA.1	I1b	A.2.2	2025-05-07	Sierra Leone	Western Area Ur...	Campbell, Allan...	Central Public ...	197,193	2025-05-31
<input type="checkbox"/>	PP_002Y2QJ.1	I1b	A.2.2	2025-05-07	Sierra Leone	Western Area Ur...	Campbell, Allan...	Central Public ...	197,180	2025-05-31
<input type="checkbox"/>	PP_002Y2Y2.1	I1b	A.2.2	2025-05-07	Sierra Leone	Western Area Ur...	Campbell, Allan...	Central Public ...	197,179	2025-05-31
<input type="checkbox"/>	PP_002XLM8.1	I1b	A.2.2	2025-05-05	Sierra Leone	Kailahun	Sandi, John D.;...	National Public...	197,163	2025-05-27
<input type="checkbox"/>	PP_002XLJE.1	I1b	A.2.2	2025-05-05	Sierra Leone	Kenema	Sandi, John D.;...	National Public...	197,177	2025-05-27

Demo: Using PathoPlexus

Website:
<https://pathoplexus.org>

Search

Add Search Fields Reset Help

Clade

Collection subdivision level 1

Collection country

Length

From

To

Lineage

Collection date

strict ☒

From

To

Clade: **I1b** Collection country: **Australia**

Search returned 204 sequences

	ACCESSION VERSION	CLADE	LINEAGE	COLLECTION DATE ▼	COLLECTION COUNTRY	COLLECTION SUBDIVISION LEVEL 1	AUTHORS	AUTHOR AFFILIATIONS	LENGTH	EARLIEST RELEASE DATE
<input type="checkbox"/>	PP_000ZTFC.1	I1b	E.3	2024-07-31	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,066	2024-10-03
<input type="checkbox"/>	PP_000ZTEE.1	I1b	E.3	2024-07-31	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,055	2024-10-03
<input type="checkbox"/>	PP_000ZWAK.1	I1b	F.2	2024-07-26	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,196	2024-10-03
<input type="checkbox"/>	PP_000ZW9M.1	I1b	F.2	2024-07-25	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,196	2024-10-03
<input type="checkbox"/>	PP_000ZW7R.1	I1b	F.2	2024-07-25	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,196	2024-10-03
<input type="checkbox"/>	PP_000ZW8P.1	I1b	F.2	2024-07-25	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,196	2024-10-03
<input type="checkbox"/>	PP_000ZTDG.1	I1b	E.3	2024-07-24	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,065	2024-10-03
<input type="checkbox"/>	PP_000ZW5V.1	I1b	F.2	2024-07-24	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,195	2024-10-03
<input type="checkbox"/>	PP_000ZW6T.1	I1b	F.2	2024-07-23	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,196	2024-10-03
<input type="checkbox"/>	PP_000ZW3Z.1	I1b	F.2	2024-07-22	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,196	2024-10-03
<input type="checkbox"/>	PP_000ZW4X.1	I1b	F.2	2024-07-22	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,196	2024-10-03
<input type="checkbox"/>	PP_000ZW21.1	I1b	F.2	2024-07-18	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,194	2024-10-03
<input type="checkbox"/>	PP_000ZW13.1	I1b	F.2	2024-07-17	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,187	2024-10-03
<input type="checkbox"/>	PP_000ZW05.1	I1b	F.2	2024-07-17	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,195	2024-10-03
<input type="checkbox"/>	PP_000ZVZ7.1	I1b	F.2	2024-07-16	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,196	2024-10-03
<input type="checkbox"/>	PP_000ZVY9.1	I1b	F.2	2024-07-16	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,176	2024-10-03
<input type="checkbox"/>	PP_000ZVXB.1	I1b	F.2	2024-07-15	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,195	2024-10-03
<input type="checkbox"/>	PP_000ZVWD.1	I1b	F.2	2024-07-15	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,195	2024-10-03

[Customize columns](#)

[Download all entries](#)

Demo: Using PathoPlexus

Website:
<https://pathoplexus.org>

Search

 Add Search Fields  Reset  Help

Accession

Mutations

Data use terms

Clade

I1b

Collection subdivision level 1

Collection country

Australia

Length

From

To

Lineage

Collection date

strict

From

dd/mm/yyyy

To

dd/mm/yyyy

Clade: I1b x Collection country: Australia x

Search returned 204 sequences

	ACCESSION VERSION	CLADE	LINEAGE	COLLECTION DATE ▼	COLLECTION COUNTRY	COLLECTION SUBDIVISION LEVEL 1	AUTHORS	AUTHOR AFFILIATIONS	LENGTH	EARLIEST RELEASE DATE
<input type="checkbox"/>	PP_000ZTFC.1	I1b	E.3	2024-07-31	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,066	2024-10-03
<input type="checkbox"/>	PP_000ZTEE.1	I1b	E.3	2024-07-31	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,055	2024-10-03
<input type="checkbox"/>	PP_000ZWAK.1	I1b	F.2	2024-07-26	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,196	2024-10-03
<input type="checkbox"/>	PP_000ZW9M.1	I1b	F.2	2024-07-25	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,196	2024-10-03
<input type="checkbox"/>	PP_000ZW7R.1	I1b	F.2	2024-07-25	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,196	2024-10-03
<input type="checkbox"/>	PP_000ZW8P.1	I1b	F.2	2024-07-25	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,196	2024-10-03
<input type="checkbox"/>	PP_000ZTDG.1	I1b	E.3	2024-07-24	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,065	2024-10-03
<input type="checkbox"/>	PP_000ZW5V.1	I1b	F.2	2024-07-24	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,195	2024-10-03
<input type="checkbox"/>	PP_000ZW6T.1	I1b	F.2	2024-07-23	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,196	2024-10-03
<input type="checkbox"/>	PP_000ZW3Z.1	I1b	F.2	2024-07-22	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,196	2024-10-03
<input type="checkbox"/>	PP_000ZW4X.1	I1b	F.2	2024-07-22	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,196	2024-10-03
<input type="checkbox"/>	PP_000ZW21.1	I1b	F.2	2024-07-18	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,194	2024-10-03
<input type="checkbox"/>	PP_000ZW13.1	I1b	F.2	2024-07-17	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,187	2024-10-03
<input type="checkbox"/>	PP_000ZW05.1	I1b	F.2	2024-07-17	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,195	2024-10-03
<input type="checkbox"/>	PP_000ZVZ7.1	I1b	F.2	2024-07-16	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,196	2024-10-03
<input type="checkbox"/>	PP_000ZVY9.1	I1b	F.2	2024-07-16	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,176	2024-10-03
<input type="checkbox"/>	PP_000ZVXB.1	I1b	F.2	2024-07-15	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,195	2024-10-03
<input type="checkbox"/>	PP_000ZVWD.1	I1b	F.2	2024-07-15	Australia	Victoria	Edenborough, K....	Melbourne Healt...	197,195	2024-10-03

[Customize columns](#)

[Download all entries](#)

Demo: Using PathoPlexus

Download

Active filters

Clade: **IIb**

Collection country: **Australia**

Include restricted data?

- ☒ No, only download open data
- ☐ Yes, include restricted data
([What does it mean?](#))

Data Type

- ☒ Metadata Choose fields (24)
- ☐ Raw nucleotide sequences
FASTA header style
- ☒ Accession
- ☐ Display name
- ☐ Aligned nucleotide sequences
- ☐ Aligned amino acid sequences

OPG001

Compression

- ☒ None
- ☐ Zstandard
- ☐ Gzip

☐ I agree to the [data use terms](#).

Download

- No limit on download

Demo: Using PathoPlexus

Download

Active filters

Clade: **I1b**

Collection country: **Australia**

Include restricted data?

- ☒ No, only download open data
- ☐ Yes, include restricted data
([What does it mean?](#))

Data Type

- ☒ Metadata [Choose fields \(24\)](#)
- ☐ Raw nucleotide sequences
FASTA header style
 - ☒ Accession
 - ☐ Display name
- ☐ Aligned nucleotide sequences
- ☐ Aligned amino acid sequences

OPG001

Compression

- ☒ None
- ☐ Zstandard
- ☐ Gzip

☐ I agree to the [data use terms](#).

Download

- Select restriction as appropriate.
- Restricted data means you can not publish without permission.

Demo: Using PathoPlexus

PP_002XLCT.1



Display Name: Sierra_Leone/PP_002XLCT.1/2025-05-20

Sandi, John D.; Parker, Edyth; Harding, Doris Harding; Kallom, Tiangay P.M.S; Ope-ewe, Oludayo Oluwaseyi; Kamara, Mohamed S.; Sijuwola, Ayotunde Elijah; Saibu, Femi; Specht, Ivan; Fofanah, Ibrahim U.; Eromon, Philomena Ehiaghe; Ozonoff, Al; Soumare, Harouna; Squire, James; Wilkason, Colby; Park, Danny; Levy, Josh; Baudi, Ian; Marrs, Libby; Cronan, Paul; Shin, Christopher; Varilly, Patrick; Nosamiefan, Dolo; Paye, ..., Donald S. [Show more](#)

Submission details

Submission ID	G-29009-1
Submitting group	KGHSL_VHF_LAB
Date submitted	2025-05-27 19:51:18 UTC
Date released	2025-05-27 19:52:23 UTC

Data use terms

Data use terms	OPEN (history)
Data use terms URL	https://pathoplexus.org/about/terms-of-use/open-data

Authors

Author affiliations	National Public Health Agency (NPHA) Sierra Leone, Sierra Leone Ministry of Health, Kenema Government Hospital Sierra Leone, Institute of Genomics and Global Health Redeemer's University Ede Osun State Nigeria, The Broad Institute of MIT and Harvard Cambridge MA USA, The Scripps Research Institute La Jolla CA 92037 USA, School of Community Health Sciences Njala University Sierra Leone, Fathom Information Design, Boston, MA, USA
---------------------	---

Clade & Lineage

Clade	IIb
Lineage	A.2.2
Outbreak	hMPXV-1

Alignment and QC metrics ⓘ

Completeness	99.99%
Frame shifts	OPG174:290-347(nt:148355-148528)
Length	197136
Total ambiguous nucs	3
Total deleted nucs	84
Total frame shifts	1
Total inserted nucs	19
Total SNPs	81
Total stop codons	0
Total unknown nucs	2

Sample details

Earliest release date	2025-05-27
Collection subdivision level 1	Kenema
Collection country	Sierra Leone
Collection date	2025-05-20
Collection date (lower bound)	2025-05-20
Collection date (upper bound)	2025-05-20

Provides additional per sample information and annotations

Demo: Using PathoPlexus

PP_002XLCT.1

Display Name: Sierra_Leone/PP_002XLCT.1/2025-05-20

Sandi, John D.; Parker, Edyth; Harding, Doris Harding; Kallom, Tiangay P.M.S; Ope-ewe, Oludayo Oluwaseyi; Kamara, Mohamed S.; Sijuwola, Ayotunde Elijah; Saibu, Femi; Specht, Ivan; Fofanah, Ibrahim U.; Eromon, Philomena Ehiaghe; Ozonoff, Al; Soumare, Harouna; Squire, James; Wilkason, Colby; Park, Danny; Levy, Josh; Baudi, Ian; Marrs, Libby; Cronan, Paul; Shin, Christopher; Varilly, Patrick; Nosamiefan, Dolo; Paye, ..., Donald S. [Show more](#)

Submission details

Submission ID

Submitting group

Date submitted

Date released

G-29009-1

[KGHSL_VHF_LAB](#)

2025-05-27 19:51:18 UTC

2025-05-27 19:52:23 UTC

Data use terms

Data use terms

Data use terms URL

OPEN ([history](#))

<https://pathoplexus.org/about/terms-of-use/open-data>

Authors

Author affiliations

National Public Health Agency (NPHA) Sierra Leone, Sierra Leone Ministry of Health, Kenema Government Hospital Sierra Leone, Institute of Genomics and Global Health Redeemer's University Ede Osun State Nigeria, The Broad Institute of MIT and Harvard Cambridge MA USA, The Scripps Research Institute La Jolla CA 92037 USA, School of Community Health Sciences Njala University Sierra Leone, Fathom Information Design, Boston, MA, USA

Clade & Lineage

Clade

Lineage

Outbreak

lIb

A.2.2

hMPXV-1

Alignment and QC metrics ⓘ

Completeness

Frame shifts

Length

Total ambiguous nucs

Total deleted nucs

Total frame shifts

Total inserted nucs

Total SNPs

Total stop codons

Total unknown nucs

99.99%

OPG174:290-347(nt:148355-148528)

197136

3

84

1

19

81

0

2

Sample details

Earliest release date

Collection subdivision level

Collection country

Collection date

Collection date (lower bound)

Collection date (upper bound)

2025-05-27

1

Sierra Leone

2025-05-20

2025-05-20

2025-05-20

Provides additional per sample information and annotations

PP_002XLCT.1

Clade & Lineage

Clade

Lineage

Outbreak

lIb

A.2.2

hMPXV-1

Host

Host name - scientific

Host taxon ID

Homo sapiens

[9606](#)

Demo: NCBI Virus

- Website: <https://www.ncbi.nlm.nih.gov/labs/virus/vssi/>

Search GenBank virus/viroid sequences and explore metadata in Results Table with advanced filtering and data visualizations options

Monkeypox virus, taxid:10244

Potyvirus plum~~p~~oxi, taxid:3240600

All viruses

Human viruses

H5N* Influenza A virus

Released in past month

SARS-CoV-2

Note: NCBI Virus and PathoPlexus can potentially have the exact same data. PathoPlexus will deposit sequences in NCBI on the submitter's behalf after a specified period of time.

Demo: NCBI Virus

Advanced Filters for GenBank Sequences

Visual Filters for GenBank Sequences

Outbreak Sequence Statistics

Refine Results

Virus/Taxonomy

Monkeypox virus, taxid:10244

Accession

Sequence Length

Ambiguous Characters

GenBank/RefSeq

Assembly Completeness

Nucleotide Completeness

MPXV lineage

Isolate/Strain Name

Genotype

Has Proteins

Provirus/Integrated

Geographic Region

Host

Applied Filters: Virus/Taxonomy (1)

Selected Results: 0 / 10,355

Align

Build Phylogenetic Tree

Download All Results

Nucleotide (10,355)

Protein (1,531,335)

NCBI Virus Assembly (2)

	Accession	Organism Name	GenBank/RefSeq	Assembly	Submitters	Organization	Release Date	MPXV LI	
<input type="checkbox"/>	NC_063383.1	Monkeypox virus	RefSeq	GCF_014621545.1	Mauldin,M.R....	National Center for Biotec...	2022-05-30	A	MPXV-M5312
<input type="checkbox"/>	NC_003310.1	Monkeypox virus	RefSeq	GCF_000857045.1	Senkevich,T....	National Center for Biotec...	2001-12-12		
<input type="checkbox"/>	PV697337.1	Monkeypox virus	GenBank		Roychoudhu...	University of Washington,...	2025-05-28	F.4	WA-UW-5206
<input type="checkbox"/>	PV697338.1	Monkeypox virus	GenBank		Roychoudhu...	University of Washington,...	2025-05-28	C.1	WA-UW-8933
<input type="checkbox"/>	LC876638.1	Monkeypox virus	GenBank			Shinjuku, Tokyo 162-8640...	2025-05-23	C.1	hMpoxV/Jap
<input type="checkbox"/>	LC876639.1	Monkeypox virus	GenBank			Shinjuku, Tokyo 162-8640...	2025-05-23	C.1	hMpoxV/Jap
<input type="checkbox"/>	LC876640.1	Monkeypox virus	GenBank			Shinjuku, Tokyo 162-8640...	2025-05-23	C.1	hMpoxV/Jap
<input type="checkbox"/>	LC876641.1	Monkeypox virus	GenBank			Shinjuku, Tokyo 162-8640...	2025-05-23	C.1	hMpoxV/Jap
<input type="checkbox"/>	LC876642.1	Monkeypox virus	GenBank			Shinjuku, Tokyo 162-8640...	2025-05-23	C.1	hMpoxV/Ja
<input type="checkbox"/>	LC876643.1	Monkeypox virus	GenBank			Shinjuku, Tokyo 162-8640...	2025-05-23	C.1	hMpoxV/Ja
<input type="checkbox"/>	LC876644.1	Monkeypox virus	GenBank			Shinjuku, Tokyo 162-8640...	2025-05-23	C.1	hMpoxV/Jap

Select Columns

Feedback