BV-BRC Test Report

S1. Global Search

Item to test	Genomes Search	
URL	https://www.bv-brc.org/, top right on any page	
Prerequisites	None	
References	https://www.bv-brc.org/docs/quick_references/global_search.html	
Tester(s)	Christian Zmasek, Anna Niewiadomska, Rebecca Wattam, Ron Kenyon	
Test date	04-Jan-2022 and 04-Feb-2022 (original-passed), 21-Apr-2022 (follow-up-passed)	
Test result	Passed (minor issues reported)	

Overview

- Test the Global Search using representative search criteria for bacterial and viral data.
- Test using example Keywords.
- Test using combinations Data Type and Keywords.
- Inspect search results to verify that they match search criteria.
- For viral data, test using representative criteria and verify that returned genome counts are comparable with IRD/ViPR.

Test Results

- Test results were verified by examination of returned data via inspection, filters, and/or comparison with the same search in the legacy PATRIC and IRD/ViPR resources (if present in those resources).
- All searches returned expected data, with results matching search criteria.
- Search results were comparable between BV-BRC and IRD/ViPR.

Q ② All terms All Data Types Mycobacterium tuberculosis H37Rv

Results:

Mycobacterium AND tuberculosis AND H37Rv

Taxa: 1 Domains and Motifs: 0 Surveillance: 0 Genomes: 52 Epitopes: 136 Serology: 0 Strains: 0 Protein Structures: 714 Experiments: 36 Proteins: 108592 Pathways: 36942 Antibiotics: 1 Specialty Genes: 23524569 Subsystems: 22008 Genomic Sequences: 1155

Top Matches

Taxa (1)

Mycobacterium tuberculosis H37Rv

Genomes (52)

Mycobacterium tuberculosis H37Rv (Broad)

Genome ID: 83332.111 | 1 Contigs SEQUENCED: 7/11/12 (Broad Institute)

Genomic DNA, wild types, parents and mutants provided by John Aquadro at the Broad Institute.

Mycobacterium tuberculosis H37Rv [PRJNA218189]

Genome ID: 83332.115 | 50 Contigs SEQUENCED: 4/28/14 (Broad Institute) HOST: Human, Homo sapiens

TB-ARC - Development

Mycobacterium tuberculosis H37Rv

Genome ID: 83332.12 | 1 Contigs SEQUENCED: 6/19/98 (Wellcome Trust Sanger Institute)

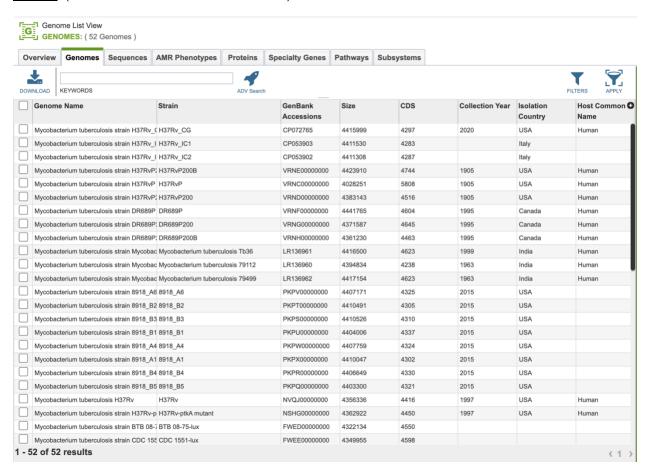
Mycobacterium tuberculosis strain H37RV. This strain has been derived from the original human-lung H37 isolate in 1934, and has been used extensively worldwide in biomedical research. Unlike some clinical isolates, it retains full virulence in animal models of tuberculosis and is susceptible to drugs and receptive to genetic manipulation.

Proteins (108592)

Data Type: Genomes

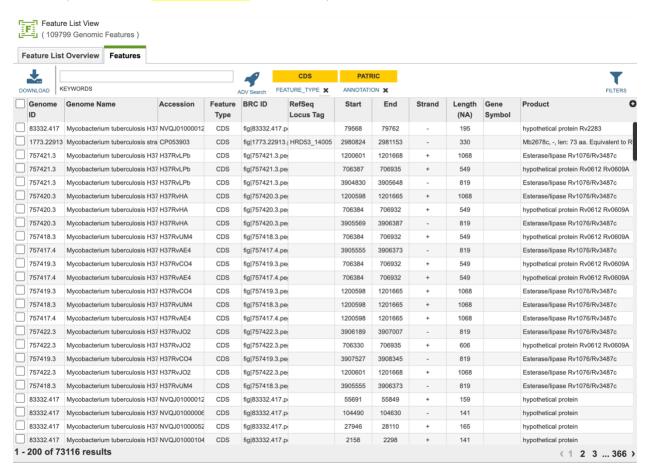


Results: (Genome count matches count above)



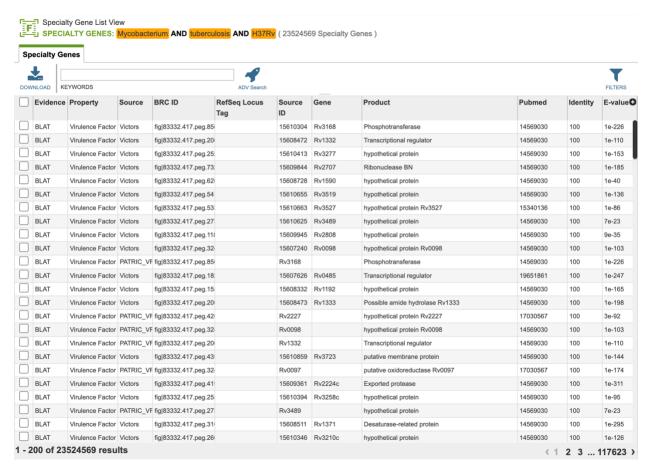
Data Type: Proteins (features)

Results: (Protein count **does not match** count above)



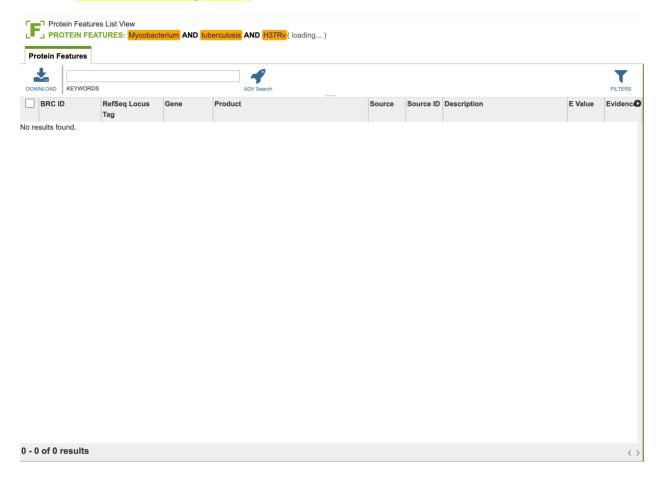
Data Type: Specialty Genes

Results: (Specialty Genes count matches count above)

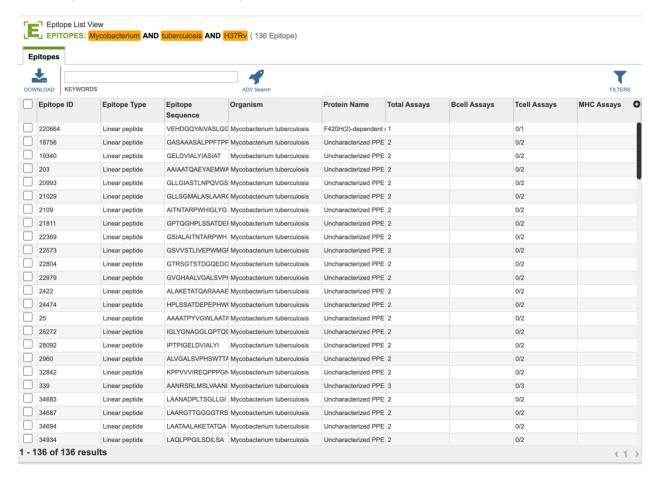


Data Type: Domains and Motifs

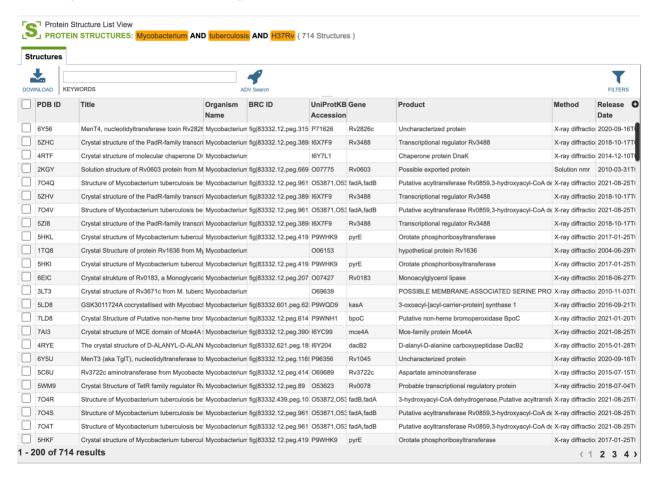
Results: (Does not return any results)



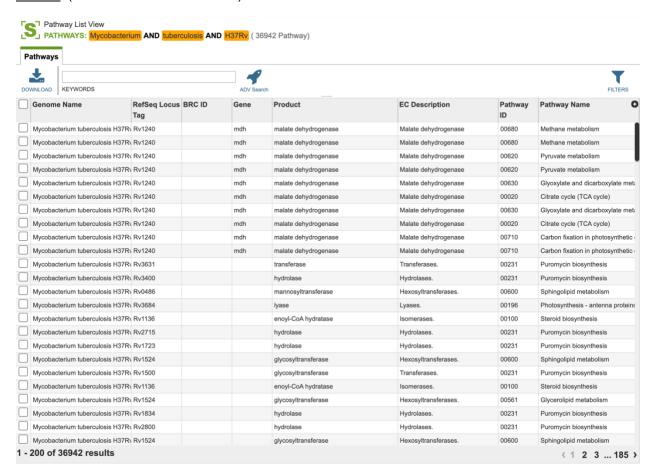
Data Type: Epitope



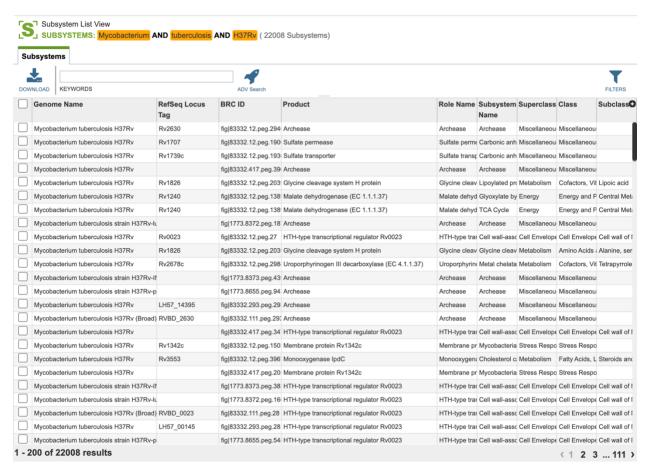
Data Type: Protein Structures



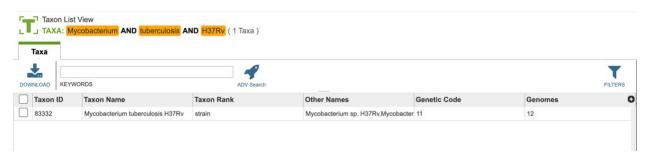
Data Type: Pathways



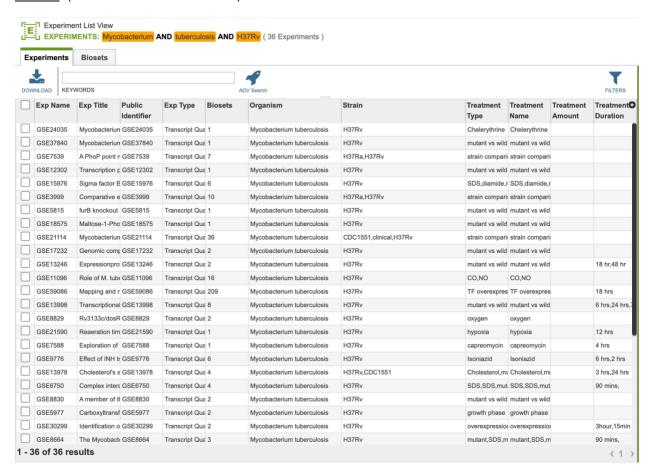
Data Type: Subsystems



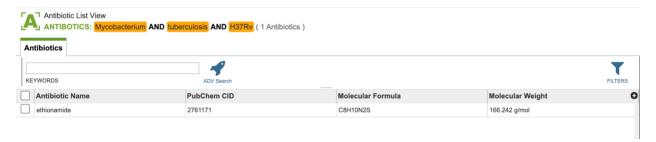
<u>Keyword</u>: Mycobacterium tuberculosis H37Rv <u>Data Type</u>: Taxa



Data Type: Experiments



Data Type: Antibiotics



Keyword: Influenza

Results: Counts are comparable, with difference attributable to different data ingestion cycles

BV-BRC: 169,407 strains

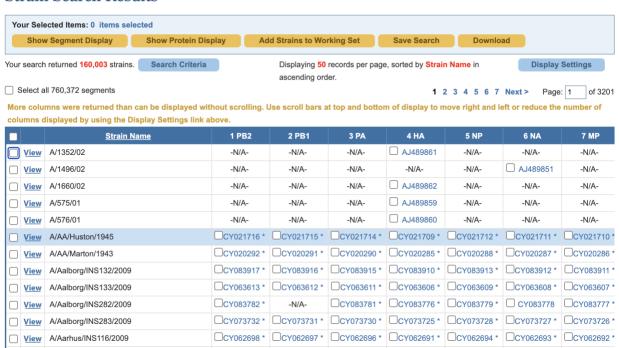
Taxa: 129900 Domains and Motifs: 391 Surveillance: 1665853

Genomes: 939401 Epitopes: 9209 Serology: 47936 Strains: 169407 Protein Structures: 1249 Experiments: 25

Proteins: 5057826 Pathways: 0 Antibiotics: 0

Specialty Genes: 1132 Subsystems: 0 Genomic Sequences: 986547

IRD/ViPR: 160,003 strains Strain Search Results



Keyword: Dengue

Results: Counts are comparable, with difference attributable to different data ingestion cycles

BV-BRC: 40,609 genomes

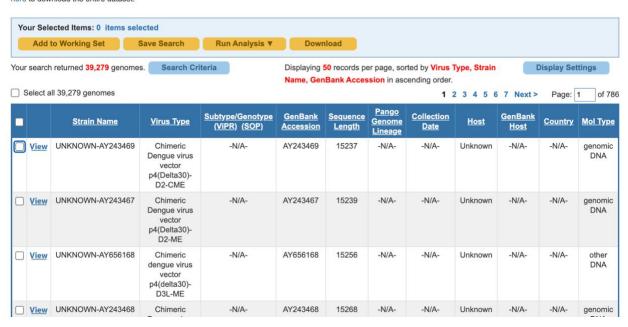
Taxa: 22 Domains and Motifs: 0 Surveillance: 0
Genomes: 40609 Epitopes: 9991 Serology: 0
Strains: 0 Protein Structures: 210 Experiments: 0

Proteins: 408549 Pathways: 0 Antibiotics: 0

Specialty Genes: 43 Subsystems: 0 Genomic Sequences: 39574

IRD/ViPR: 39,279 genomes Genome Search Result

ViPR has clinical data linked to genomic sequences for ~2300 Dengue virus isolates, including patient demographics, measures of disease severity, immune response, etc. Click here to download the entire dataset.



Keyword: Ebolavirus

Results: Counts are comparable, with difference attributable to different data ingestion cycles

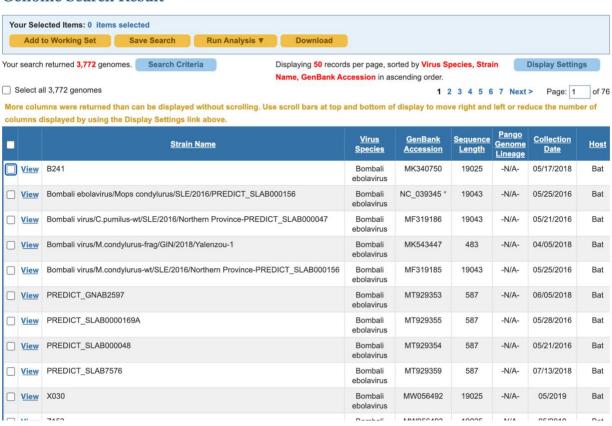
BV-BRC: 3,969 genomes

ebolavirus

Taxa: 10 Domains and Motifs: 0 Surveillance: 0
Genomes: 3969 Epitopes: 2506 Serology: 0
Strains: 0 Protein Structures: 86 Experiments: 0
Proteins: 180300 Pathways: 0 Antibiotics: 0

Specialty Genes: 11 Subsystems: 0 Genomic Sequences: 3678

IRD/ViPR: 3,772 genomes Genome Search Result



Keyword: Enterovirus

Results: Counts are comparable, with difference attributable to different data ingestion cycles

BV-BRC: 111,454 genomes

enterovirus

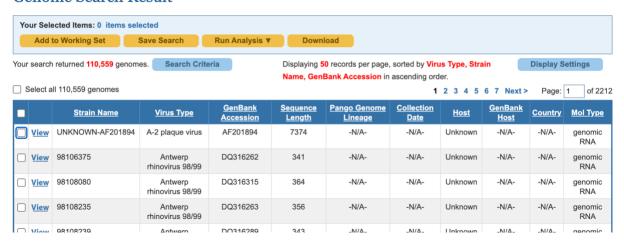
Taxa: 330 Domains and Motifs: 0 Surveillance: 0
Genomes: 111454 Epitopes: 1868 Serology: 0
Strains: 0 Protein Structures: 148 Experiments: 0
Proteins: 125487 Pathways: 0 Antibiotics: 0

Specialty Genes: 1 Subsystems: 0 Genomic Sequences: 30527

IRD/ViPR: 110,559 genomes

<u>ViPR Home</u> ▶ <u>Enterovirus Home</u> ▶ <u>Genome Search</u> ▶ Results

Genome Search Result



Keyword: Lassa

Results: Counts are comparable, with difference attributable to different data ingestion cycles

BV-BRC: 2,384 genomes

lassa

Taxa: 3 Domains and Motifs: 0 Surveillance: 0
Genomes: 2384 Epitopes: 1310 Serology: 0
Strains: 1138 Protein Structures: 45 Experiments: 0

Proteins: 12470 Pathways: 0 Antibiotics: 0

Specialty Genes: 0 Subsystems: 0 Genomic Sequences: 2378

IRD/ViPR: 2,285 genomes

Genome Search Result

	Your Selected Items: 0 items selected Add to Working Set Save Search Run Analysis ▼ Download												
Vour search returned 2,285 genomes. Search Criteria Displaying 50 records per page, sorted by Virus Species, Strain Name, GenBank Accession in ascending order. Select all 2,285 genomes 1 2 3 4 5 6 7 Next > Page: 1 of 4													
-		Strain Name	<u>Virus Species</u>	GenBank Accession	Sequence Length	<u>Pango</u> <u>Genome</u> <u>Lineage</u>	Collection Date	<u>Host</u>	GenBank Host	Country	Mol Type	<u>Segment</u>	Protein Name
	<u>View</u>	1007	Lassa mammarenavirus	KP339095	756	-N/A-	05/29/2004	Mouse	Mastomys natalensis	Guinea	viral cRNA	S	nucleoproteir
	View	1008	Lassa mammarenavirus	KP339096	756	-N/A-	05/29/2004	Mouse	Mastomys natalensis	Guinea	viral cRNA	S	nucleoproteir
	<u>View</u>	1010	Lassa mammarenavirus	KP339097	756	-N/A-	05/29/2004	Mouse	Mastomys natalensis	Guinea	viral cRNA	S	nucleoproteir
	<u>View</u>	1011	Lassa mammarenavirus	KP339098	756	-N/A-	05/29/2004	Mouse	Mastomys natalensis	Guinea	viral cRNA	S	nucleoprotei
	<u>View</u>	1033	Lassa mammarenavirus	KP339099	756	-N/A-	05/30/2004	Mouse	Mastomys natalensis	Guinea	viral cRNA	S	nucleoprotei
	View	1034	Lassa mammarenavirus	KP339100	756	-N/A-	05/30/2004	Mouse	Mastomys natalensis	Guinea	viral cRNA	S	nucleoprotei
	<u>View</u>	1066	Lassa mammarenavirus	KP339101	756	-N/A-	05/30/2004	Mouse	Mastomys natalensis	Guinea	viral cRNA	S	nucleoprotei
	View	1074	Lassa mammarenavirus	KP339102	756	-N/A-	05/31/2004	Mouse	Mastomys natalensis	Guinea	viral cRNA	S	nucleoprotei
	<u>View</u>	1198	Lassa mammarenavirus	KP339103	756	-N/A-	10/04/2004	Mouse	Mastomys natalensis	Guinea	viral cRNA	S	nucleoprotei

Keyword: Zika

Results: Counts are comparable, with difference attributable to different data ingestion cycles

BV-BRC: 2,222 genomes

zika

Taxa: 2 Domains and Motifs: 0 Surveillance: 0
Genomes: 2222 Epitopes: 2895 Serology: 0
Strains: 0 Protein Structures: 150 Experiments: 0
Proteins: 37880 Pathways: 0 Antibiotics: 0

Specialty Genes: 0 Subsystems: 0 Genomic Sequences: 2225

IRD/ViPR: 2,214 genomes Genome Search Result

