

---

# BV-BRC

Bacterial and Viral (BV) -  
Bioinformatics Resource Center (BRC)

## Monthly Usage Metrics Report

**Performance Period:** February 1, 2021 – February 28, 2021

*Issued to:*

National Institute of Allergy and Infectious Diseases  
National Institute of Health

**Contract No.:** 75N93019C00076

**Contract Title:** Bioinformatics Resource Centers for Infectious Diseases

**Submission Date:**

March 10, 2021

## BV-BRC Usage Metrics Report

This monthly usage metrics report provides a summary of the BV-BRC usage for the current reporting period in accordance with the Joint-BRC Common Usage Metrics Plan developed by the BRCs and subsequently approved by NIAID.

As per the plan, each BRC will aggregate and report usage metrics for their constituent parts, *i.e.*, PATRIC and IRD/ViPR for BV-BRC. These metrics will serve as a basis for collecting quantitative measures of usage of the BRC resources to identify trends, areas that are performing well, and areas for improvement. Usage metrics will be reported to NIAID individually by each BRC on a monthly basis, and in combination on the BRC Gateway website once this is publicly available. In addition, annual summaries will be included in the Annual Progress Reports.

*It is important to note that usage metrics across the two BRCs are highly dependent on the relative sizes of the respective research communities, the associated quantities and types of available public data, and how each of the resources delivers the data and tools to the user. Thus, cross-BRC comparisons of individual metrics are not necessarily indicative of relative usage or performance.*

**Common** usage metrics covering both BRCs (note that this list is subject to modification, based on feasibility of collection, changes in availability technologies, BRC website development, suggestions from NIAID program and other stakeholders, *etc.*):

### Website Usage Metrics

Website usage is a key measure for evaluating use of the resource by the research communities. The number of website sessions unique users in a given period provide insights into trends, such as increased traffic resulting from outreach activities and prominent research topics and endeavors. Both the BRCs will use **AWStats** to monitor and track website usage by and report the number of unique visitors, visits, page views, pages/visit and visits/visitors for a given reporting period, aggregated across all constituent BRC websites, as summarized in the table below. In addition, we will also provide links to the live website usage statistics pages generated by AWStats from respective BRC websites, which will provide more detailed usage statistics by day of the week/month, country, browser / operating system, and more.

- **Total visits**
  - *Definition* - Number of visits made by all visitors. Think "session" here, say a unique IP accesses a page, and then requests three other pages within an hour. All of the "pages" are included in the visit, therefore you should expect multiple pages per visit and multiple visits per unique visitor (assuming that some of the unique IPs are logged with more than an hour between requests)
  - *Measurement mechanism* - AWStats.
  - *Measure* - Total number of visits per month.
- **Total unique visitors**
  - *Definition* - A unique visitor is a person or computer (host) that has made at least 1 hit on 1 page of your web site during the current period shown by the report. If this user makes several visits during this period, it is counted only once. Visitors are tracked by IP address, so if multiple users are accessing your site from the same IP (such as a home or office network), they will be counted as a single unique visitor
  - *Measurement mechanism* - AWStats.
  - *Measure* - Total number of unique visitors per month.
- **Total page views**

- *Definition* - The number of "pages" viewed by visitors. Pages are usually HTML, PHP or ASP files, not images or other files requested as a result of loading a "Page" (like js,css... files).
- *Measurement mechanism* - AWStats.
- *Measure* - Total pageviews per month.
- **Average pages per visit**
  - *Definition* - The average number of pages viewed during a visit. Repeated views of a single page are counted.
  - *Measurement mechanism* - AWStats.
  - *Measure* - Average number of pages per visit per month.
- **Average visits per visitor**
  - *Definition* - The average number of visits per visitor.
  - *Measurement mechanism* - AWStats.
  - *Measure* - Average number of visits per visitor per month.
- **Average visit duration**
  - *Definition* - The average time a visitor spent on the site for each visit, measured in seconds.
  - *Measurement mechanism* - AWStats.
  - *Measure* - Average visit duration per month.
- **Total bandwidth**
  - *Definition* - Total number of bytes for pages, images and files downloaded by web browsing. This number includes traffic for web only (or mail only, or ftp only depending on value of LogType). This number does not include technical header data size used inside the HTTP or HTTPS protocol or by protocols at a lower level (TCP, IP...). Note that this number is often lower than the bandwidth usually reported by internet providers as it is counted at a lower level and includes all IP and UDP traffic.
  - *Measurement mechanism* - AWStats.
  - *Measure* - Total bandwidth per month.

**Table 1. BRC Website Usage Metrics**

<b>Metric</b>	<b>Result</b>
Total visits	116,091
Total unique visitors	32,177
Total pageviews	4,153,341
Avg. pages / visit	35.77
Avg. visits / visitor	3.6
Avg. visit duration (seconds)	922
Bandwidth (GB)	246.98

### Website Usage by Taxa

BRCs support a variety of organism taxa containing human pathogens and their vectors, along with related genomic and other omics data types. These taxa vary widely in the number of species and genomes they contain, availability of omics data, as well as the size of the research communities studying them. Measuring the BRC website usage by taxa allows us to understand how BRC resources are used by various organism communities. We will report the number of pageviews by taxa, which will be measured by querying the website usage statistics in Google Analytics by taxa name.

**Table 2. BV-BRC Website Usage by Taxa**

<b>Taxa</b>	<b>Domain</b>	<b>Species</b>	<b>Genomes</b>	<b>Page Views</b>
Acinetobacter	Bacteria	672	9,818	1,820
Bacillus	Bacteria	794	6,288	5,123
Bartonella	Bacteria	75	177	1,255
Borrelia	Bacteria	17	687	24
Brucella	Bacteria	64	1,025	3,063
Burkholderia	Bacteria	284	4,055	480
Campylobacter	Bacteria	262	5,868	1,052
Chlamydia	Bacteria	19	545	542
Clostridium	Bacteria	401	2,287	1,130
Coxiella	Bacteria	10	106	363
Ehrlichia	Bacteria	7	40	639
Escherichia	Bacteria	153	29,707	5,869
Francisella	Bacteria	32	866	242
Helicobacter	Bacteria	81	2,196	836
Listeria	Bacteria	57	5,027	700
Mycobacterium	Bacteria	320	30,204	2,155
Pseudomonas	Bacteria	1,573	12,520	1,918
Rickettsia	Bacteria	55	180	992
Salmonella	Bacteria	206	26,393	2,132
Shigella	Bacteria	45	3,970	979
Staphylococcus	Bacteria	340	19,816	1,691
Streptococcus	Bacteria	408	34,203	2,188
Vibrio	Bacteria	372	5,007	1,142
Yersinia	Bacteria	31	1,325	477
Bunyavirales	Virus	3,345	16,648	974
Caliciviridae	Virus	3,530	58,400	281

Coronaviridae	Virus	5,252	112,855	8,475
Filoviridae	Virus	391	4,239	614
Flaviviridae	Virus	59,042	360,860	7,482
Hepeviridae	Virus	2,019	19,635	189
Herpesviridae	Virus	30,557	62,702	2,963
Influenza	Virus	0	758,366	33,438
Paramyxoviridae	Virus	19,855	81,171	1028
Picornaviridae	Virus	13,793	140,435	1,379
Pneumoviridae	Virus	0	41,843	266
Poxviridae	Virus	3,505	10,979	760
Reoviridae	Virus	13,791	127,082	1,946
Rhabdoviridae	Virus	5,457	36,340	320
Togaviridae	Virus	1,945	14,266	834

### Website Usage by Data Types

BRCs support genomic and a variety of other omics data types, providing an integrated view of these multi-omics data and related analysis tools. Tracking the website usage by primary data types allows us to understand how these data types are used. We will report the number of website pageviews by primary data types, which will be measured by querying the website usage statistics in Google Analytics by data type.

**Table 3. BRC Website Usage by Data Type (BV-BRC)**

Data Type	BRC Domain	Page Views
Taxonomy	PATRIC	27,910
Genome	PATRIC	66,747
Genome sequence	PATRIC	1,538
Feature (Genes/Proteins)	PATRIC	21,296
Specialty gene	PATRIC	5,990
Protein families	PATRIC	3,656
Pathway	PATRIC	7,656
Subsystems	PATRIC	3,077
Transcriptomics	PATRIC	1,103
Interactions	PATRIC	712
Phylogeny	PATRIC	1,815

Antibiotics	PATRIC	0
Genome	IRD/ViPR	10,662
Gene/Protein	IRD/ViPR	8,961
Strain	IRD/ViPR	7,842
Immune epitopes	IRD/ViPR	2,539
Ortholog groups	IRD/ViPR	125
Antiviral drugs	IRD/ViPR	295
Host factors	IRD/ViPR	391
Protein structures	IRD/ViPR	614
Protein domains and motifs	IRD/ViPR	717
Plasmids	IRD/ViPR	74
SFVT	IRD/ViPR	129
Surveillance	IRD/ViPR	412
Serology	IRD/ViPR	19
Phenotypes	IRD/ViPR	32
PCR Primers	IRD/ViPR	388

### Service/Tool Usage

Both BRC analysis services and tools allow users to analyze data pulled from the respective BRC databases and their own private data, compare to other datasets, and save the results in their private workspaces. Since the types of tools vary across the BRCs, we will report aggregated usage of all tools in each BRC, and also a breakdown by service/tool. We will also report the total amount of storage used for user data.

- **Total number of analysis tasks submitted and completed successfully by users**
  - *Definition* - The total number of analysis tasks submitted and completed successfully by users for a given month. An analysis task usually involves users providing input data/search terms and/or parameters to initiate a search or analysis task, which may perform one or more searches, data transformations, or data analysis steps, generate results that provide additional insights into the data and present it back to the user in structured view and/or file formats via web interface and/or user workspace.
  - *Measurement mechanism* - Analysis tasks are recorded via website and server logs, which are used to tally the number.
  - *Measure* - Analysis tasks submitted and completed successfully per month.
- **Analysis tasks submitted and successfully completed by service/tool**
  - *Definition* - A breakdown of total number of analysis tasks (see metric above), summarized by service/tool during the specified date range.
  - *Measurement mechanism* - Analysis tasks submitted by users and successfully completed are captured via website and server logs, which are used to tally the number.
  - *Measure* - Jobs per month, tallied by service/tool.

**Table 4. BRC Tools/Services Usage Metrics**

<b>Tool/Service</b>	<b>BRC Domain</b>	<b>Submitted</b>	<b>Completed</b>
Codon Tree	PATRIC	532	521
Comprehensive Genome Analysis	PATRIC	1,541	1,385
Differential Expression	PATRIC	2	1
FastqUtils	PATRIC	369	348
Genome Alignment	PATRIC	178	168
Genome Annotation	PATRIC	172,683	160,024
Genome Annotation GenBank	PATRIC	38,130	37,788
Genome Assembly	PATRIC	9,615	6,291
Genome Comparison	PATRIC	177	169
Metagenome Binning	PATRIC	42	41
Metagenomic Read Mapping	PATRIC	413	407
Phylogenetic Tree	PATRIC	18	13
RNASeq Analysis	PATRIC	296	291
Taxonomic Classification	PATRIC	470	469
TnSeq Analysis	PATRIC	5	1
Variation Analysis	PATRIC	227	196
Alignment Viewer	IRD/ViPR	55	55
Antiviral-Resistance-Risk	IRD/ViPR	15	15
BLAST	IRD/ViPR	418	413
Enrichment	IRD/ViPR	5	5
Genotype-Recombination	IRD/ViPR	31	8
H1-Clade Classifier	IRD/ViPR	80	79
H1N1-classifier	IRD/ViPR	15	15
H5N1-classifier	IRD/ViPR	62	60
Ha Numbering	IRD/ViPR	120	120
MGC	IRD/ViPR	70	65
MSA	IRD/ViPR	606	587

Mutation-analysis	IRD/VI PR	41	41
Primer3	IRD/VI PR	29	29
RaxML hpc2_tgb	IRD/VI PR	0	0
Read-seq	IRD/VI PR	72	72
Rva Genotyper	IRD/VI PR	450	450
Short-seqsearch	IRD/VI PR	134	125
SNP-analysis	IRD/VI PR	341	340
Surveillance-data-mapping	IRD/VI PR	7	7
Tbl-formatter	IRD/VI PR	5	1
Tree	IRD/VI PR	927	911
VIGOR Annotator	IRD/VI PR	142	142
SARS-2 Genome Assembly and Annotation	BV-BRC	37	33

## Publications and Citations

Publications and citations provide insights into how the BRC is moving science and technology forward and how the resources are serving their respective research communities. Lists of BRC-generated publications (including publications supported by the BRC program in collaboration with various partners) are updated when new manuscripts are accepted and published. Citations to BRC resources are measured using Google Scholar and augmented using PubMed and custom queries as needed to identify citations to the resource that do not cite the official reference publication(s).

### ● Citations to BRC publications

- *Definition* - Citations to the BRC as measured by citations to key BRC publications, which describe the overall BRC resources, new data and/or analysis tools, or novel use cases supported by them.
- *Measurement mechanism* - Set up a common Google Scholar profile covering key BRC resource publications (grouped by BRC) and show aggregated citations for each group. The use of Google Scholar profile makes it easier to view the list of publications used to track citations, update the list with new publications, and provide citation counts for individual publications as well as aggregated counts for each resource. Below is the link to the common BRC Google Scholar Profile.
  - <https://scholar.google.com/citations?user=kXLGwkYAAAAJ>
- *Measure* - Cumulative number of citations.

### ● Citations to BRC resources

- *Definition* - Citations to the BRC resource as measured Google Scholar searches using predetermined set of keywords based on name and/or acronym of each of the BRC resources, and additional keywords to filter out any false positive or negative results to the extent possible. This is complementary to the citations to the BRC publications described above and necessary because, often, users cite BRC resources by mentioning the resource name or URL in the manuscript text, instead of citing relevant publications.
- *Measurement mechanism* - Define set of keywords based on name and/or acronym of each of the BRC resources and additional keywords to filter out any false positive or negative results to the extent possible. Using these keywords as search terms, create Google Scholar

URLs for each of the BRC resources, which will be checked every month to report a cumulative number of citations for each resource. Because of the limitations of the logical and advanced query operations supported by Google Scholar search interface, we are dividing BV-BRC query into three distinct sub queries as shown below.

- VEuPathDB (merged DB, including legacy VectorBase, FungiDB & parasite resources):  
<https://scholar.google.com/scholar?q=OrthoMCL+OR+PlasmoDB+OR+ToxoDB+OR+CryptoDB+OR+TrichDB+OR+GiardiaDB+OR+TriTrypDB+OR+AmoebaDB+OR+MicrosporidiaDB+OR+%22FungiDB%22+OR+PiroplasmaDB+OR+%22vectorbase%22+OR+veupathdb+OR+ApiDB+OR+EuPathDB+-encrypt+-cryptography+-hymenoptera>
- BV-BRC:
  - PATRIC BRC:  
[https://scholar.google.com/scholar?hl=en&as\\_sdt=0%2C39&q=%28PATRIC+AND+Wattam%29+OR+%E2%80%9Cpatricbrc%22+OR+%22pathosystems+resource+integration+center%22](https://scholar.google.com/scholar?hl=en&as_sdt=0%2C39&q=%28PATRIC+AND+Wattam%29+OR+%E2%80%9Cpatricbrc%22+OR+%22pathosystems+resource+integration+center%22)
  - RAST/RASTtk:  
[https://scholar.google.com/scholar?hl=en&as\\_sdt=0%2C39&q=%28RAST+AND+overbeek%29+OR+%22rast.nmpdr.org%22](https://scholar.google.com/scholar?hl=en&as_sdt=0%2C39&q=%28RAST+AND+overbeek%29+OR+%22rast.nmpdr.org%22)
  - IRD/ViPR:  
[https://scholar.google.com/scholar?hl=en&as\\_sdt=0%2C39&q=%22viprbrc%22+OR+%22virus+pathogen+resource%22+OR+%E2%80%9Cfludb%22+OR+%22influenza+research+database%22](https://scholar.google.com/scholar?hl=en&as_sdt=0%2C39&q=%22viprbrc%22+OR+%22virus+pathogen+resource%22+OR+%E2%80%9Cfludb%22+OR+%22influenza+research+database%22)

- *Measure* - Cumulative number of citations, cumulative.

	Number of Citations (YTD)	Number of Citations (Cumulative)
Citations to BRC publications	496	13,097
Citations to BRC resources	523	14,030

## User Activities

Outreach activities provide additional channels to engage users. User requests for help typically come in through the help desk functionality available from both BRC websites and are tracked using ticketing software tools. Webinar and workshop participants are counted at the time of registration and participation at the event. Counts of access to recorded webinars may be used to augment the total. Followers on social media (Twitter, Facebook, YouTube) are counted using the built-in mechanisms those platforms provide.

- **Total registered users**
  - *Definition* - Total cumulative number of users who have registered with the BRC via the website registration mechanism, from inception to the specified date.
  - *Measurement mechanism* - The registration process creates an entry in the registered user database for each BRC. Total number of registered users is queried from the database at the specified date.
  - *Measure* - Total number of registered users (cumulative).
- **Total storage used for user data**
  - *Definition* - Total amount of disk storage in use to host user data at the specified date. This metric provides an additional indication of resource usage that may not be reflected by website traffic or analysis jobs.
  - *Measurement mechanism* - Inspection of disk usage via query or automated script.

- *Measure* - Total terabytes (TB) currently in use.
- **User requests for help**
  - *Definition* - Total number of user-initiated contacts to the BRC to request help or information during the specified date range. In addition to summarizing total user requests, we will also summarize them by the following categories: Requests for help, Bug reports, and New features / enhancements.
  - *Measurement mechanism* - Manual tally of the auto-generated helpdesk tickets triggered by user requests. Tallies may be augmented with manual counts of interactions where the user bypassed the helpdesk system, e.g. via direct email or messaging to BRC team members.
  - *Measure* - Requests per month.
- **Webinar/workshop events and participants**
  - *Definition* - Total number of outreach events (i.e. BRC webinars, workshops, and online courses) held per month and total number of participants who attended those events.
  - *Measurement mechanism* - Manual tally of participants in attendance at the time of the webinar or workshop, summed over all of the events held per month.
  - *Measure* - Cumulative number of participants per month
- **Followers on social media**
  - *Definition* - Total number of followers, by social media outlet, at the specified date. Current active BRC social media outlets are Twitter, Facebook, and YouTube.
  - *Measurement mechanism* - Inspection of the number of followers reported by the media outlet at the specified date.
  - *Measure* - Total number of followers, by media outlet.

	<b>PATRIC</b>	<b>IRD/ViPR</b>	<b>Total</b>
Total registered users	17,134	11,170	28,304
Total storage used for user data	147 TB	461 GB	148 TB
Total user requests for help	75	28	103
Webinar/workshop events	0	1	1
Total webinar/workshop participants	0	76	76
Total MOOC participants (cumulative)	1,743	0	1,743
Twitter followers	430	266	696
Facebook followers	237	1,542	1,779
YouTube followers	180	133	313
YouTube views	1000	180	1180