Takara Bio USA

Cogent™ NGS Immune Viewer v1.0 User Manual

Cat. Nos. 634466, 634467, 634478, 634479, 634776, 634777, & 634778
software v1.0
(100622)

Table of Contents

Table of Figures

Table of Tables

Introduction

A. Welcome to the Cogent NGS Immune Viewer

Cogent NGS Immune Viewer (referred to as Immune Viewer in this manual) is a post-analysis tool designed to visualize and generate tabulated data output and publication-ready plots for download from the CSV file output created by Cogent NGS Immune Profiler v1.5 (CogentIP).

This document describes the required data, environment, and how to operate.

B. What's New

- **Cogent NGS Immune Viewer v1.0**
  - Visualization of data output from CogentIP v1.5
  - Download tabulated analysis results in CSV format and plot images as PDFs

NOTE: Refer to the software compatibility table on the bioinformatics portal page at takarabio.com for more details on supported kits for CogentIP and Immune Viewer.

II. Before You Begin

A. Supported Operating Systems

- Mac OS X: El Capitan (Version 10.11 and up)
- Window 8 and up

B. Supported Web Browsers

- Google Chrome
- Microsoft Edge
- Mozilla Firefox

C. Hardware Requirements

- Memory: 2 GB RAM
- Free disk space: at least 2 GB available hard drive space

D. Additional Software Dependencies

- A program that can zip and unzip files (e.g., 7-zip, WinRAR, Stuffit, etc.)
- A PDF file reader (e.g., Adobe Acrobat Reader)

E. Required Input Files

Immune Viewer requires comma-separated value (CSV) files generated by CogentIP v1.5 as input.

- **(Recommended)** `*_result.csv` files generated by Immune Profiler from the `report/<sampleID>` and `airr_report/<sampleID>` folders, -or-
- `airr_report/<output name>_<sampleID>_mig_[cdr3|fl]_report.csv`

The `report/` and `airr_report/` subfolders can be found under the directory name specified when configuring an Immune Viewer run ("Output name" in the GUI, the value of the `-o` option on the CLI).

IMPORTANT: Maximum input file size is 1 GB.

III. Software Overview

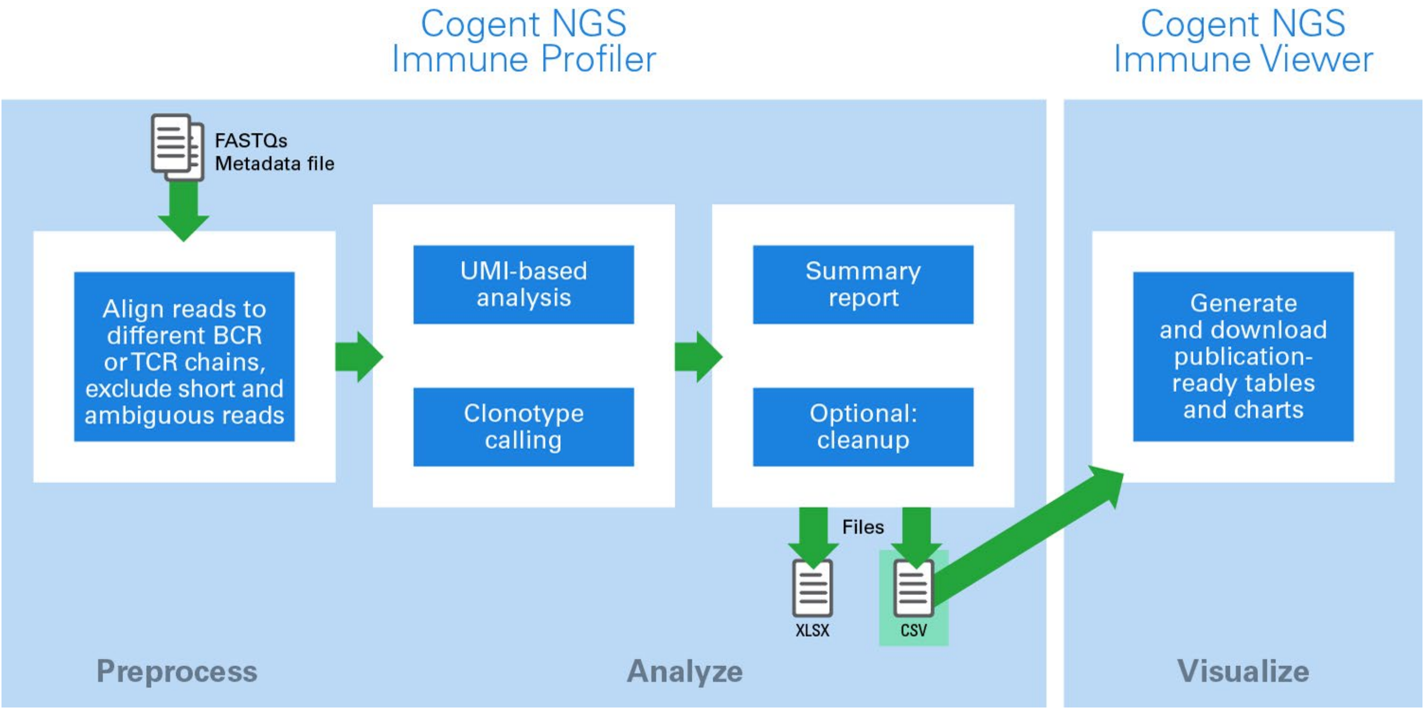


Figure 1. Cogent NGS Immune Profiler and Cogent NGS Immune Viewer workflow overview.

Figure 1 illustrates at a high-level the workflow of taking sequencing data from Illumina® sequencing platforms from libraries prepared using Takara Bio's human repertoire immune profiling kits through CogentIP and then into Immune Viewer for reporting.

The steps specific to Immune Viewer are the following:

1. Import of CSV file (Section II.E, above) outputs by CogentIP into the web-based Immune Viewer.
2. Processing of the import file and generation of tabulated data output and plot images for display on the web page.
3. Availability of download functions to save the table data and plot images to your local computer.

IV. Using Immune Viewer

Access the Immune Viewer user interface through the Takara Bio website at takarabio.com/ngs-immune-viewer. No login is required.

A. User Interface Overview

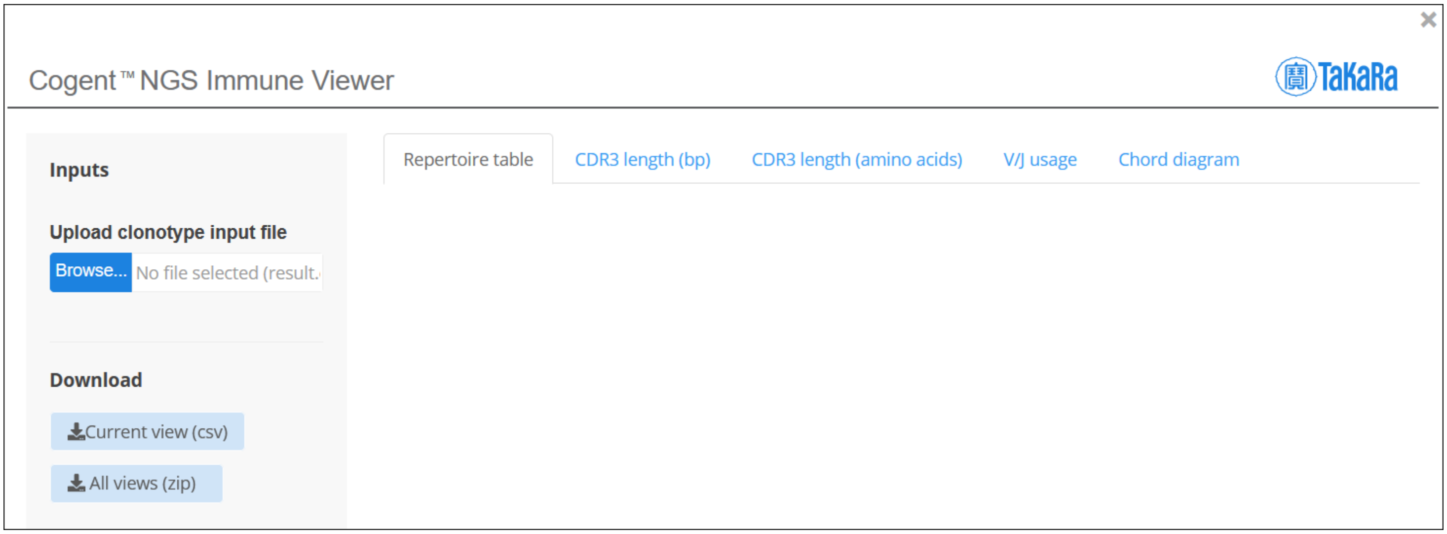


Figure 2. Cogent NGS Immune Viewer start screen. The left side of the frame lists “Inputs” and “Download” functions. The main page displays tabs for the available report types.

- Left-side menu
- Main page
    - Repertoire table
    - CDR3 length (bp) plot
    - CDR3 length (amino acids) plot
    - V/J usage distribution plots
    - Chord diagram

B. Upload a Clonotype File

1. Click on the [Browse] button. A window will pop up prompting you to select a CSV file to upload.
2. Navigate to the CogentIP output folder of interest then into `reports/` or `airr_reports/`.
3. Select the file name of interest and click [Open] to begin the upload.

   While the file is uploading to the tool, you will see a progress bar under the [Browse] button and file name input field. When it finishes, it will change to an "Upload complete" message (Figure 3).

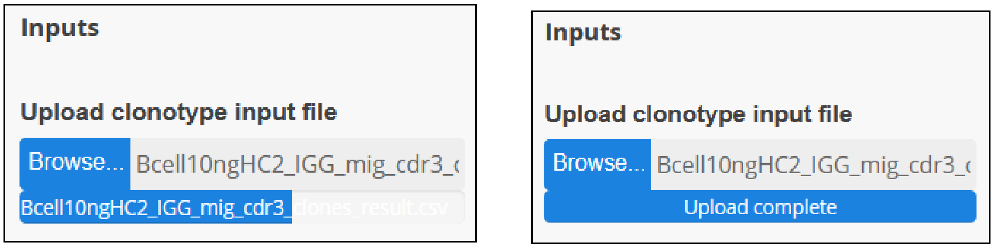
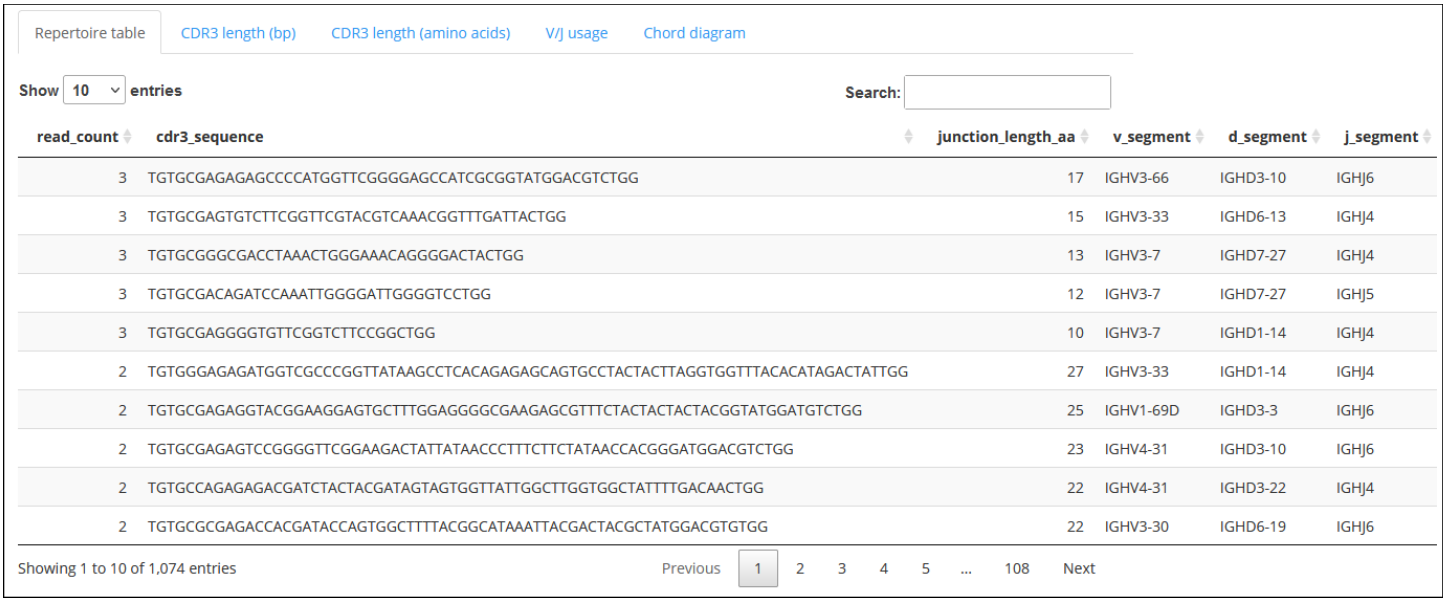


Figure 3. Input UI during and after completion of file upload.

After the upload completes, the main frame of the browser will display the Repertoire Table

C. Repertoire Table



read_count	cdr3_sequence	junction_length_aa	v_segment	d_segment	j_segment
3	TGTGCGAGAGAGCCCCATGGTTCGGGGAGCCATCGCGGTATGGACGTCTGG	17	IGHV3-66	IGHD3-10	IGHJ6
3	TGTGCGAGTGTCTTCGGTTCGTACGTCAAACGGTTTGATTACTGG	15	IGHV3-33	IGHD6-13	IGHJ4
3	TGTGCGGGCGACCTAAACTGGGAAACAGGGGACTACTGG	13	IGHV3-7	IGHD7-27	IGHJ4
3	TGTGCGACAGATCCAAATTGGGGATTGGGGTCCTGG	12	IGHV3-7	IGHD7-27	IGHJ5
3	TGTGCGAGGGGTGTTCGGTCTTCCGGCTGG	10	IGHV3-7	IGHD1-14	IGHJ4
2	TGTGGGAGAGATGGTCGCCCGGTTATAAGCCTCACAGAGAGCAGTGCCTACTACTTAGGTGGTTTACACATAGACTATTGG	27	IGHV3-33	IGHD1-14	IGHJ4
2	TGTGCGAGAGGTACGGAAGGAGTGCTTTGGAGGGGCGAAGAGCGTTTCTACTACTACTACGGTATGGATGTCTGG	25	IGHV1-69D	IGHD3-3	IGHJ6
2	TGTGCGAGAGTCCGGGGTTCGGAAGACTATTATAACCCTTTCTTCTATAACCACGGGATGGACGTCTGG	23	IGHV4-31	IGHD3-10	IGHJ6
2	TGTGCCAGAGAGACGATCTACTACGATAGTAGTGGTTATTGGCTTGGTGGCTATTTTGACAACTGG	22	IGHV4-31	IGHD3-22	IGHJ4
2	TGTGCGCGAGACCACGATACCAGTGGCTTTTACGGCATAAATTACGACTACGCTATGGACGTGTGG	22	IGHV3-30	IGHD6-19	IGHJ6

Showing 1 to 10 of 1,074 entries

Figure 4. Example repertoire table display. The column headers shown here are specific to the import of an AIRR-compliant report file.

The data on the *Repertoire table* tab provides tabulated summary statistics of the six major data categories from the input file. The table lists annotated immune receptor sequences with the major data output categories. The table is searchable and can be filtered for the various data outputs.

The six major data categories represented as columns in the table are:

- Read count
- CDR3 sequence
- Amino acid length (junction_length_aa)
- V segment
- D segment
- J segment

NOTE: If you're unable to see all the columns in the report, scroll vertically down the browser page to find a second, horizontal scroll bar and scroll to the right.

There are a number of page elements to assist with navigating the report.

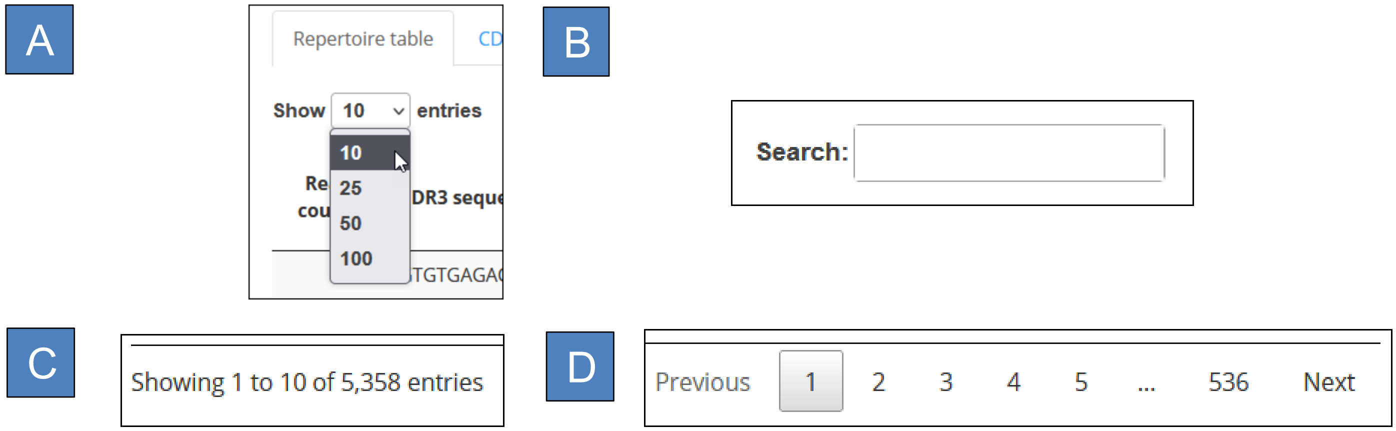


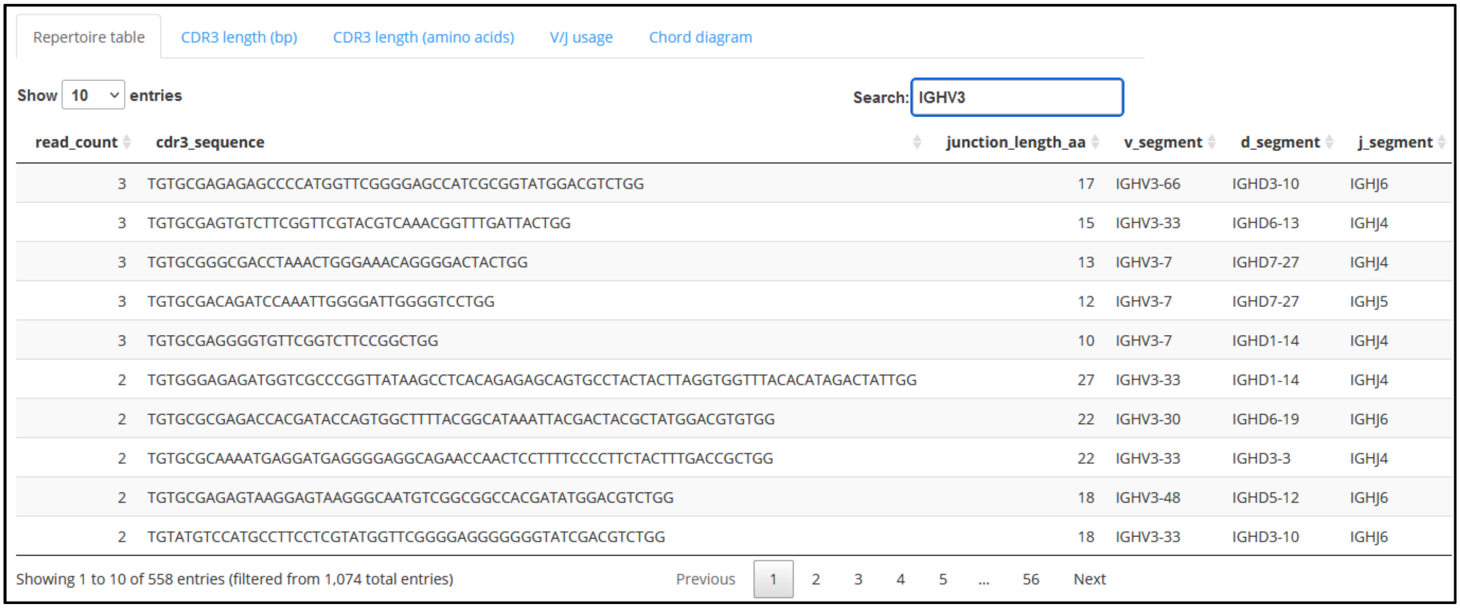
Figure 5. Navigation elements on the *Repertoire table* view. (Panel A) Entries per page. **(Panel B)** Search input box. **(Panel C)** Number of entries displaying of the total number. **(Panel D)** Navigation by page number.

- Show *N* entries**—**Adjusts the number of data rows displayed per page
- "Search"**—**Allows the use of text strings to filter the rows displayed in the table (Figure 6)

NOTES:

- The search function is case sensitive.
- Filtering on the Repertoire table will also affect plots on the CDR3 length and V/J usage distribution plots and the chord diagram.

- Showing *X* to *Y* of *Z* entries—Indicates not only where the records fall in the list but also the total number of records in the input file
- Previous / Next—Allows for navigation between pages



read_count	cdr3_sequence	junction_length_aa	v_segment	d_segment	j_segment
3	TGTGCGAGAGAGCCCCATGGTTCGGGGAGCCATCGCGGTATGGACGTCTGG	17	IGHV3-66	IGHD3-10	IGHJ6
3	TGTGCGAGTGTCTTCGGTTCGTACGTCAAACGGTTTGATTACTGG	15	IGHV3-33	IGHD6-13	IGHJ4
3	TGTGCGGGCGACCTAAACTGGGAAACAGGGGACTACTGG	13	IGHV3-7	IGHD7-27	IGHJ4
3	TGTGCGACAGATCCAAATTGGGGATTGGGGTCCTGG	12	IGHV3-7	IGHD7-27	IGHJ5
3	TGTGCGAGGGGTGTTCGGTCTTCCGGCTGG	10	IGHV3-7	IGHD1-14	IGHJ4
2	TGTGGGAGAGATGGTCGCCCGGTTATAAGCCTCACAGAGAGCAGTGCCTACTACTTAGGTGGTTTACACATAGACTATTGG	27	IGHV3-33	IGHD1-14	IGHJ4
2	TGTGCGCGAGACCACGATACCAGTGGCTTTTACGGCATAAATTACGACTACGCTATGGACGTGTGG	22	IGHV3-30	IGHD6-19	IGHJ6
2	TGTGCGCAAAATGAGGATGAGGGGAGGCAGAACCAACTCCTTTTCCCCTTCTACTTTGACCGCTGG	22	IGHV3-33	IGHD3-3	IGHJ4
2	TGTGCGAGAGTAAGGAGTAAGGGCAATGTCGGCGGCCACGATATGGACGTCTGG	18	IGHV3-48	IGHD5-12	IGHJ6
2	TGTATGTCCATGCCTTCCTCGTATGGTTCGGGGAGGGGGGGTATCGACGTCTGG	18	IGHV3-33	IGHD3-10	IGHJ6

Showing 1 to 10 of 558 entries (filtered from 1,074 total entries)

Figure 6. Example of "Search" field filtering on the Repertoire table. Input was an AIRR-compliant report file from analysis performed by CogentIP on BCR sequencing data.

D. CDR3 Length (bp) Plot

The data on the *CDR3 length (bp)* tab (along with amino acids, Section IV.E) is an important determinant of immune cell repertoire diversity and is indicative of repertoire variations between individuals and over time.

As mentioned in Section IV.C, the plot is affected by filtering done on the *Repertoire table* tab by the "Search" field. Figure 7 shows the difference between the plots based on an unfiltered (all results) view and results filtered to match the string 'IGHV3', which is displayed in the chart title.

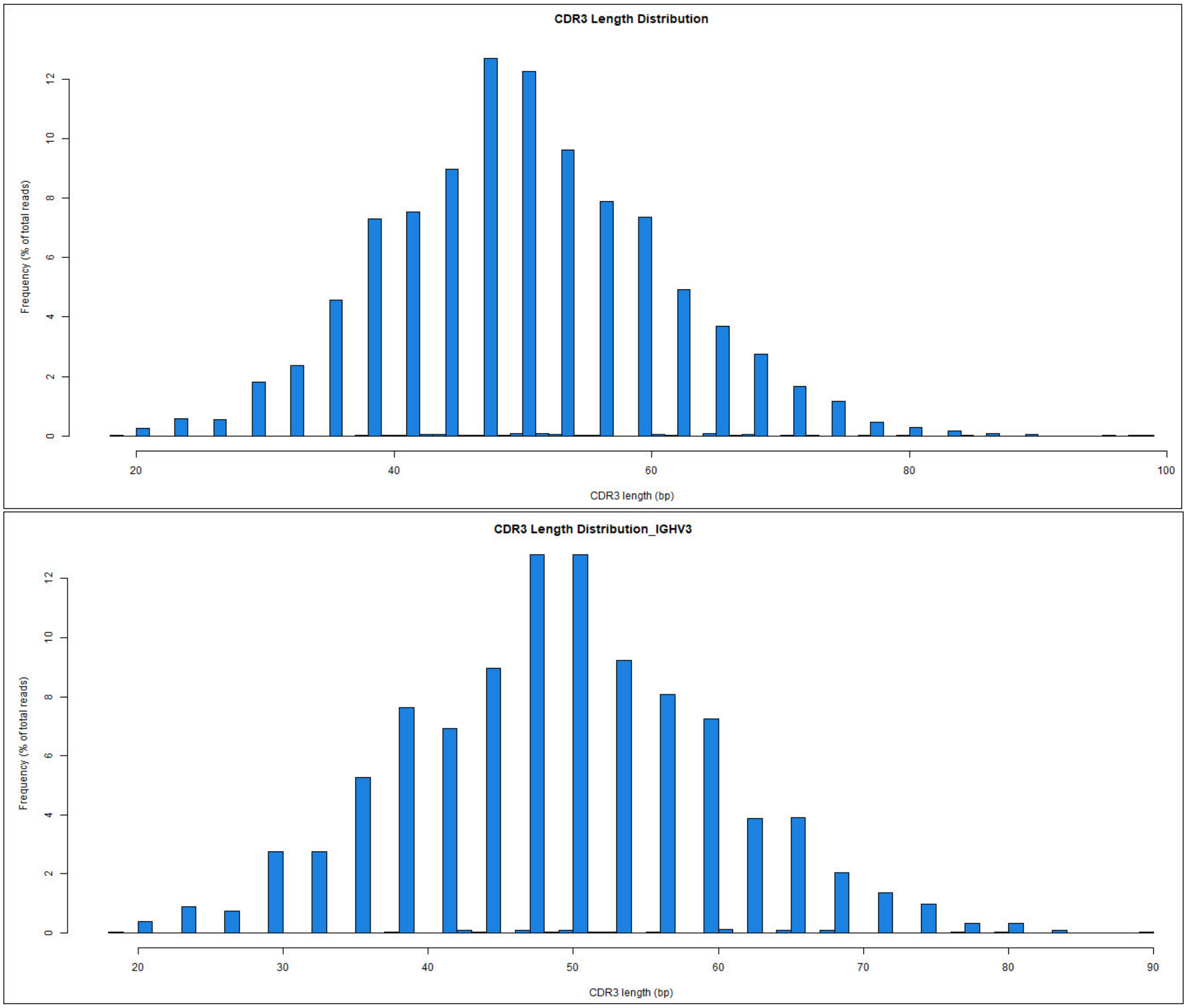


Figure 7. Example CDR3 Length (bp) plots. Input file was from analysis performed by CogentIP on BCR sequencing data. **(Top)** Unfiltered. **(Bottom)** Filtered results.

E. CDR3 Length (Amino Acids) Plot

The data on the *CDR3 length (amino acids)* tab (along with bp, Section IV.D) is an important determinant of immune cell repertoire diversity and is indicative of repertoire variations between individuals and over time.

As mentioned in Section IV.C, the plot is affected by filtering done on the *Repertoire table* tab by the "Search" field. Figure 8 (next page) shows the difference between the plots based on an unfiltered (all results) view and results filtered to match the string 'IGHV3', which is displayed in the chart title.

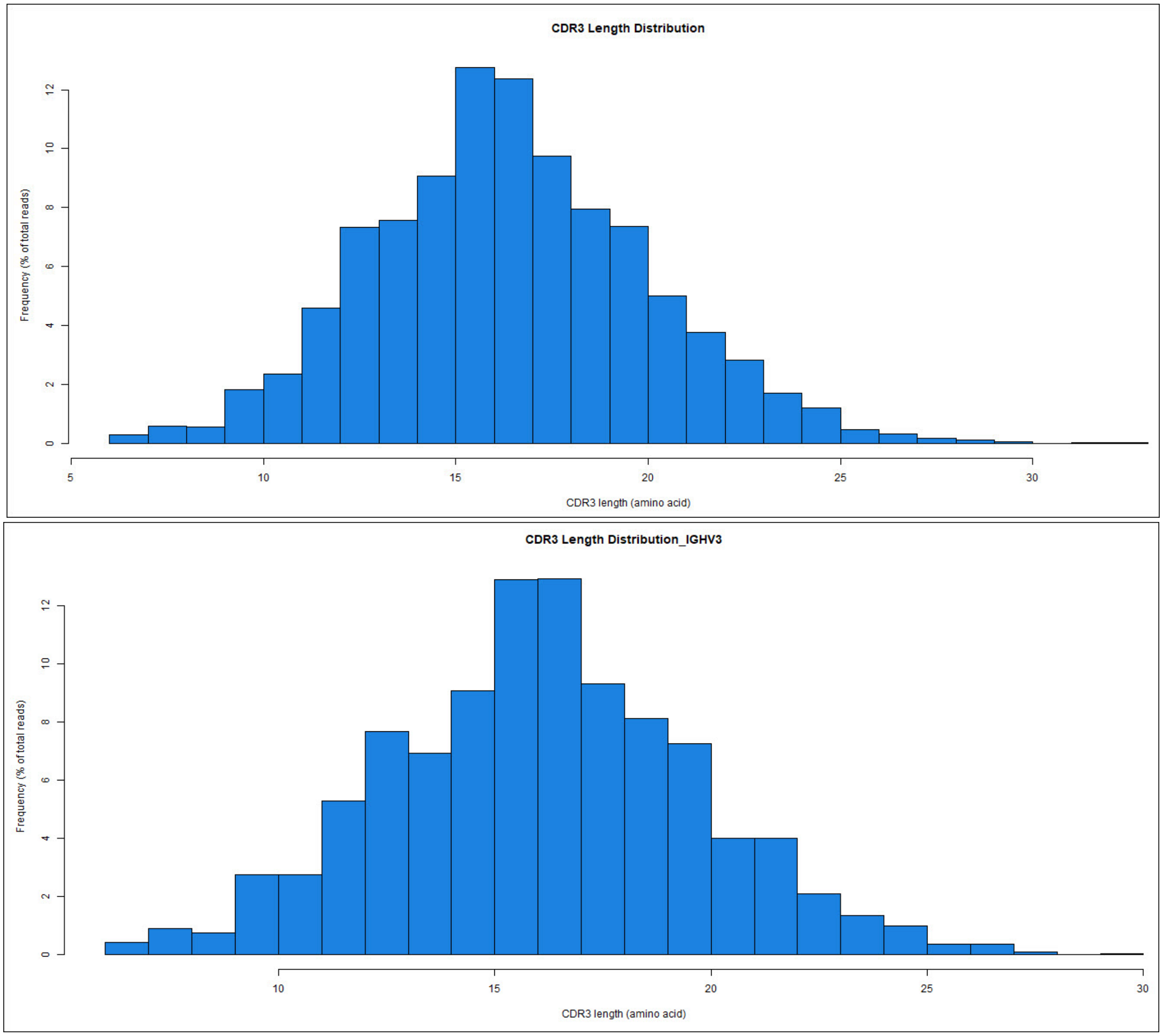


Figure 8. Example CDR3 Length (amino acids) plots. Input file was from analysis performed by CogentIP on BCR sequencing data. **(Top)** Unfiltered. **(Bottom)** Filtered results.

F. V/J Distribution Plots

The data on the *V/J usage* tab consists of two graphs: one for V distribution and one for J distribution.

The data in the V and J distribution plots provide distribution of V/J gene usage across clonotypes. This information allows evaluation of the distribution and co-occurrence of different V/J germline gene segments.

As mentioned in Section IV.C, the plot is affected by filtering done on the *Repertoire table* tab by the "Search" field. The figures below show the difference between the plots based on an unfiltered (all results) view and results filtered to match the string 'IGHV3', which is displayed in the chart titles.

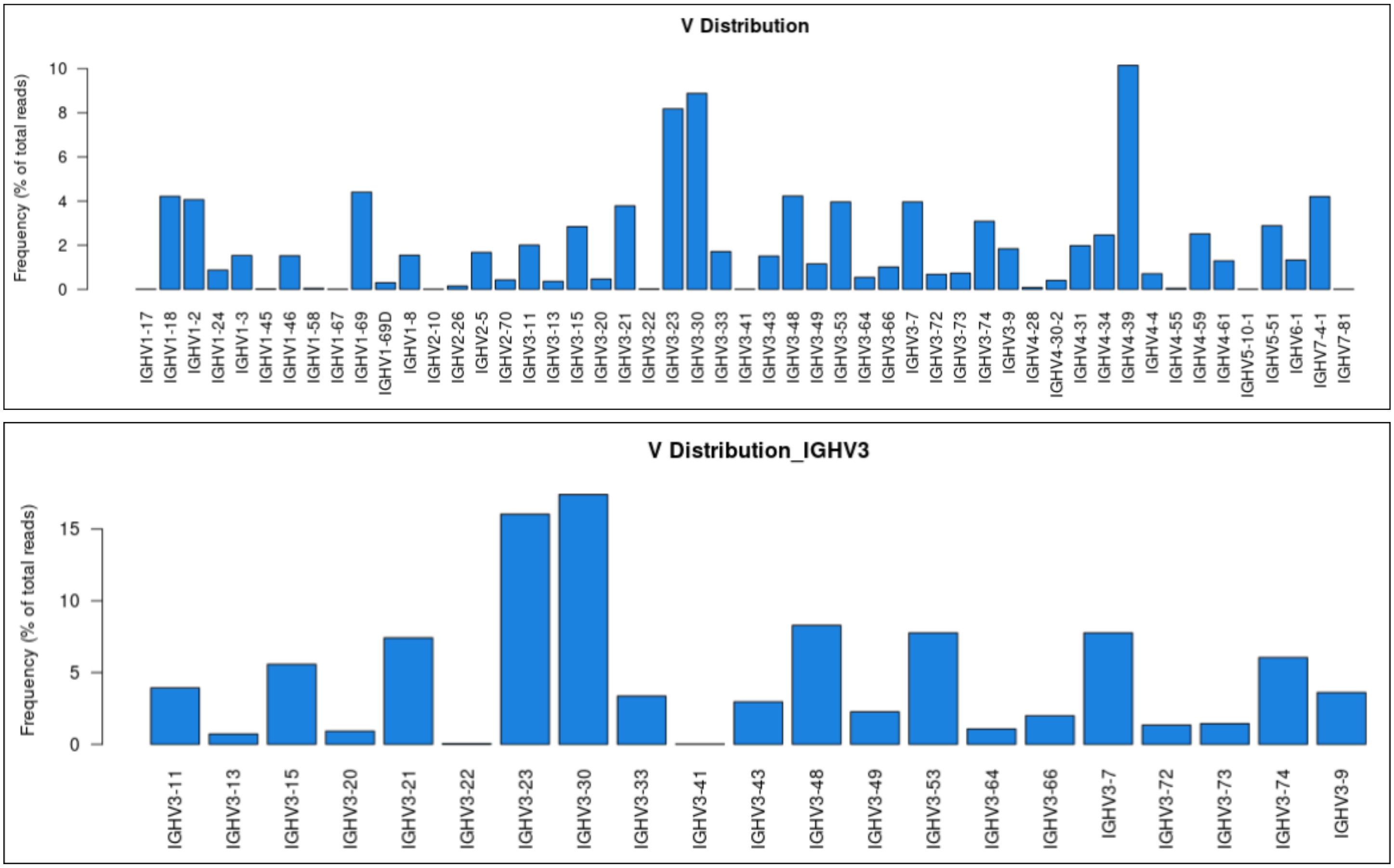


Figure 9. Example V Distribution plots. Input file was from analysis performed by CogentIP on BCR sequencing data. **(Top)** Unfiltered. **(Bottom)** Filtered results.

(continued on next page)

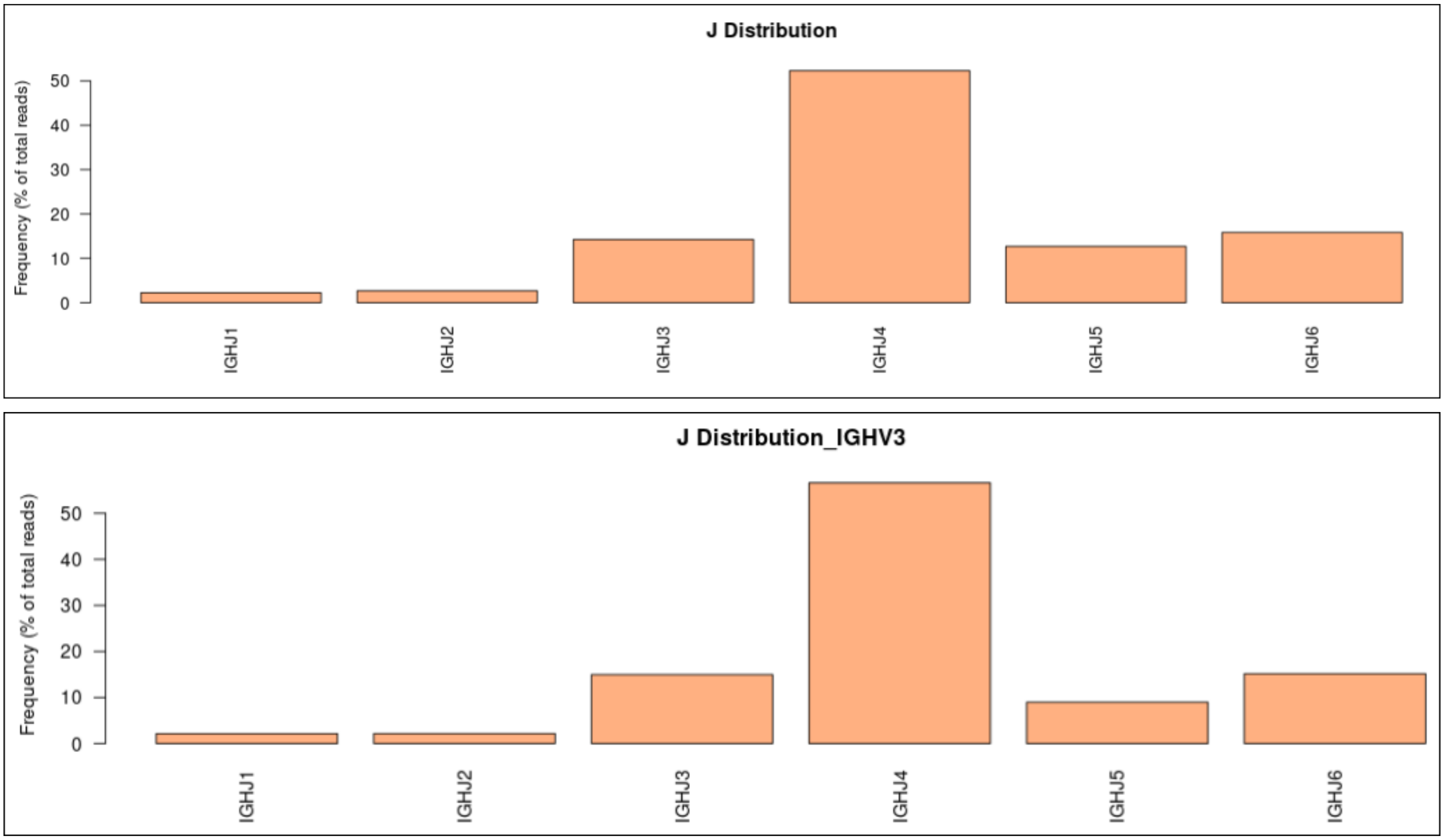


Figure 10. Example J Distribution plots. Input file was from analysis performed by CogentIP on BCR sequencing data. **(Top)** Unfiltered. **(Bottom)** Filtered results.

(continued on next page)

G. Chord Diagram

The data on the *Chord diagram* tab provides a visual representation of clonotype distribution and shows pairing of V and J segments. Chord widths represent the pairing frequency.

As mentioned in Section IV.C, the plot is affected by filtering done on the *Repertoire table* tab by the "Search" field. Figure 11 shows the difference between the plots based on an unfiltered (all results) view and results filtered to match the string 'IGHV3'.

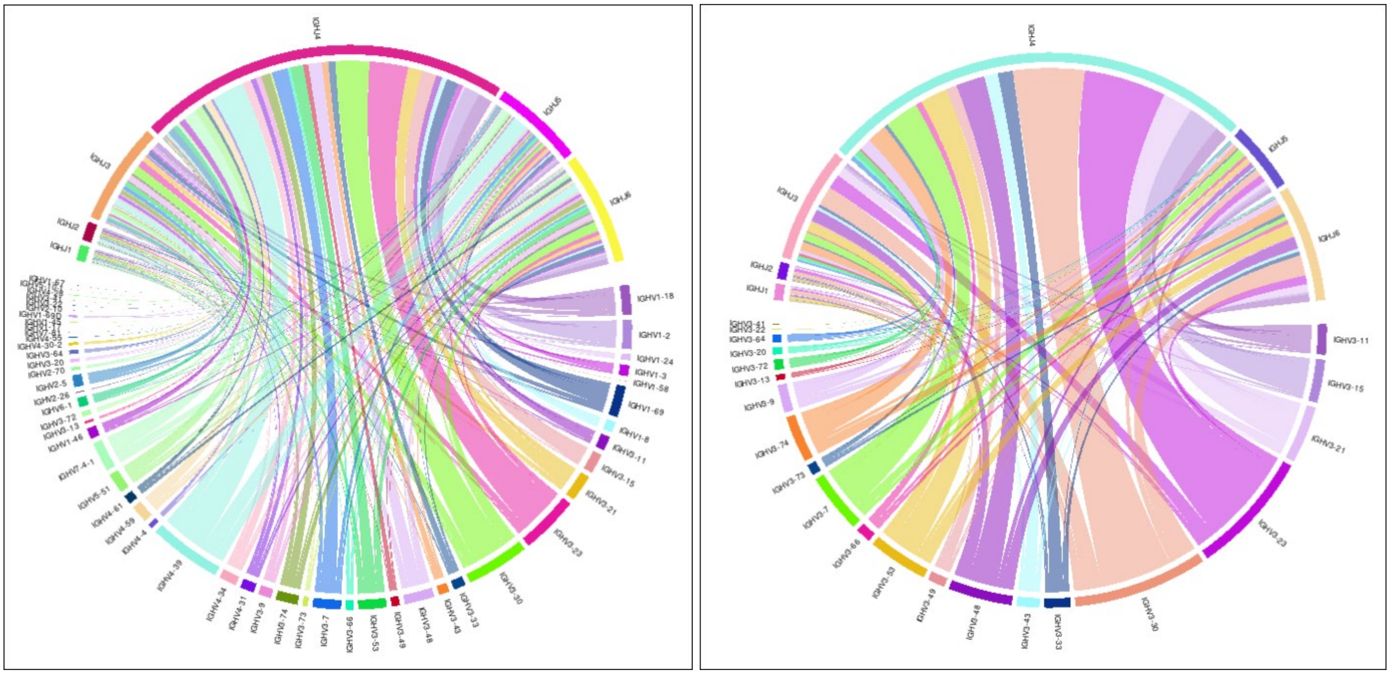


Figure 11. Example chord diagrams. Input file was from analysis performed by CogentIP on BCR sequencing data. **(Left)** Unfiltered. **(Right)** Filtered results.

H. Saving the Reports

After generating the reports, either filtered or unfiltered, there are two options with which to save them to your local server via Download.

1. Download

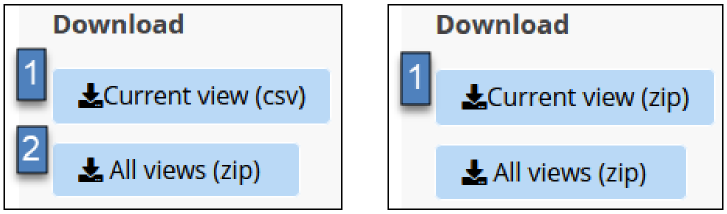


Figure 12. Download options. **(Left)** Options present for the Repertoire table. **(Right)** Options displayed on the plot tabs.

Click on the button of your choice to save the output in the local folder of your choice.

The two options (Current and All views) are described below.

- [Current view (csv|zip)]—Save only the table or plot on the tab you are currently viewing.
  - The Repertoire table is downloaded as a CSV file
  - The plots are downloaded as ZIP files containing both a CSV of the underlying data and a PDF file of the plot

- [All views (zip)]**—**Bundles together files for all reports as a single ZIP file. The files included depend on whether the data is unfiltered or filtered.
    - Unfiltered—Downloads two files each (CSV and PDF) for the four chart tabs (eight total)
    - Filtered—Downloads the eight files (CSV and PDF) for filtered chart data, plus a CSV of the filtered Repertoire table (nine total)

2. Output File Naming

The output files are named with the following structures:

- The first part of the output file names is a shared base name (prefix), constructed with the following rule:

  `<input file name>_YYYY-MM-DD`

  where:

  – `<input file name>` is the name of the file used for input into Immune Viewer, minus the `.csv` file extension.

  Example: If the file name used as input is `Bcell10ng_mig_fl_mapping_stats.csv`, the value of `<input file name>` would be: `Bcell10ng_mig_fl_mapping_stats`

  – `YYYY` is the four-digit year the report was generated
  – `MM` is the two-digit month the report was generated
  – `DD` is the two-digit day the report was generated

Example

If the `Bcell10ng_mig_fl_mapping_stats.csv` output from CogentIP was used as input to Immune Viewer on February 28, 2023, the file name prefix for the downloaded files would be:

`Bcell10ng_mig_fl_mapping_stats_2023-02-28`

The second part of the output file names depends on (a) the report or plot type and (b) whether or not the reports have been generated from unfiltered or filtered data.

- The table below represents the possible file name suffixes from reports generated from unfiltered data and the output file associated with it, plus the suffix of repertoire table generated from filtered data.

Table 1. Output file name base suffixes based on report type.

Report type	File type	File name suffix
CDR3 length (bp) Plot	CSV PDF	`_CDR3_length_bp.csv` `_CDR3_length_bp.pdf`
CDR3 length (amino acids) Plot	CSV PDF	`_CDR3_length_aa.csv` `_CDR3_length_aa.pdf`
V/J distribution Plots	CSV PDF	`_VJ_usage.csv` `_VJ_usage.pdf`
Chord Diagram	CSV PDF	`_Chord_diagram.csv` `_Chord_diagram.pdf`
Repertoire table (filtered results)	CSV	`_Filtered_table_`**`<filter-parameter>`**`.csv`

Example

The chord diagram reports for the input file in the previous example would have the full names:

`Bcell10ng_mig_fl_mapping_stats_2023-02-28_Chord_diagram.csv`

`Bcell10ng_mig_fl_mapping_stats_2023-02-28_Chord_diagram.pdf`

- Downloaded files for reports based on filtered data will have an additional `_<filter-parameter>` suffix appended to the full file name, just before the file extension. The `<filter-parameter>` text will match the text string used in the "Search" field to apply the filtering.

Example

After applying the filter 'IGHV3' to the repertoire table from the initial example, the resulting filtered chord diagram reports would have the full names:

`Bcell10ng_mig_fl_mapping_stats_2023-02-28_Chord_diagram_IGHV3.csv`

`Bcell10ng_mig_fl_mapping_stats_2023-02-28_Chord_diagram_IGHV3.pdf`

V. Technical Support

If you encounter an error message or unexpected behavior in the Immune Viewer, if possible, take a screenshot of the problem, then contact Technical Support.

Please be ready to provide the following information, if possible:

- A brief description of the issue
- What steps you may have taken just prior to seeing the problem, to help us with recreating it
- Screenshot(s) of the error message or behavior

Contact Us	
Customer Service/Ordering	**Technical Support**
tel: 800.662.2566 (toll-free)	tel: 800.662.2566 (toll-free)
fax: 800.424.1350 (toll-free)	fax: 800.424.1350 (toll-free)
web: takarabio.com/service	web: takarabio.com/support
e-mail: ordersUS@takarabio.com	e-mail: technical_support@takarabio.com

---

Notice to Purchaser