Variant Navigator Description

Variant Navigator (VN) is a software package for visual assessment of the quality of structural variant (SV) (>= 50 bp) and copy number variant (CNV) calls by GROM and other variant callers of the deletion, duplication, insertion, and inversion types.

VN is written in Java and will run on Windows, Linux, and Mac if Java 7 or better is installed. Click [here](https://java.com/en/download/help/version_manual.xml) for instructions on determining if Java is installed. Java can be downloaded from this [website](http://www.java.com/en/download/).

Each project will have the following directory structure:

One or more data folders containing text files, an “etc” folder, a file with a “.bat” extension, and a “.jar” file.

A screenshot of a computer

Description automatically generated

Windows:

Double-click on the “.bat” file in the project folder. If a warning message such as that shown at the end of this document appears follow the instructions below.

VN can also be run by running the following commands:

cd <project folder>

java -jar vn020821.jar <data\_folder>

In this case <data\_folder> would be “example”.

Mac:

Open a Terminal. To find this, open up Launchpad and search for Terminal.

Type in "cd [name of file location]/[project folder]". For example,

"cd Desktop/608\_chr1\_022321".

Click enter and in a new line, type "java -jar [name of jar file] [data folder]". For example,

"java -jar vn020821.jar 608\_chr1\_022321".

Now, Variant Navigator should open!

VN consists of:

1. A data table which displays information about SV calls (below). This table displays data from a Variant Call Format (VCF) file. Descriptions of some column contents can be found in tool tips for column headers. More information can be found in the GROM paper <https://academic.oup.com/gigascience/article/6/10/gix091/4160384> and this URL describing VCF files <https://www.internationalgenome.org/wiki/Analysis/Variant%20Call%20Format/vcf-variant-call-format-version-40/>.

Graphical user interface, application, table, Excel

Description automatically generated

1. Reads windows which displays reads for selected samples (shown below). Components of this window are from top to bottom – read depth histogram, variant calls, gene features, reference sequence, and reads. Light-colored rectangles join read pairs if input data is from paired-end reads; smaller light-colored rectangles indicate paired-end reads where the mate is offscreen at initial magnification. The color of rectangle indicates SV type. Clicking the KEY button will open a window that describes components in this window. Fig. 1 and Fig. 2 show many of the features in the VN display.

A picture containing text, screenshot, indoor

Description automatically generated

Fig. 1.

1. Features (Exons, CDS, etc. not shown) and Variant calls.

2. Overlapping reads – light green area indicates mismatches in overlapping portion.

3. Single nucleotide variant indicated as colored thin vertical bar.

4. Overlapping read pairs outlined with a green rectangle indicating full overlap.

5. Partially overlapping read pair, overlapping portion outlined with a green rectangle.

6. Non-primary alignment – dark gray with arrow indicating direction of read.

7. Histogram showing read depth.

8. Alignment details of selected read.

9. Red outline indicates read is selected.

10. Read with mate offscreen.

11. Soft-clipped region at high magnification.

12. Normal read pair. Two primary alignments pointing in concordant directions.

A picture containing timeline

Description automatically generated

Fig. 2.

1. Indel deletion.
2. Primary alignment pointing in discordant direction.
3. Soft-clipping at low magnification.

The data table contains rows that are color coded by entries in the “SV Type” column. “SV Type” entries are DEL – deletion, DUP – duplication, INS – insertion, and INV – inversion. The table can be sorted by clicking the column header. Variants are shown in one window or two windows depending on the size of the variant. The maximum variant size shown in one window is 6000 bp. The minimum size of SVs that can be viewed in two windows is 300 bp. The “Show Variant” button at the bottom will open a “Reads” window for the selected row. The “Show Variant in Two Windows” will open two windows, one centered on the “From” column entry, the other on the To column entry.

Reads windows have zoom and scroll navigation and text field showing information about components under the mouse cursor and any object that is selected. All menus can be closed using the “Esc” key.

Reads window buttons and controls perform the following functions.

1. DISPLAY button – displays the names of all components in the window. If a stripe is unchecked and Apply is clicked, a component can be hidden.
2. ABOUT button – when a feature such as an Exon or CDS is selected and this button is clicked, information from databases is accessible. Currently the NCBI page for the selected feature is opened. If no feature is selected the ABOUT button has no function.
3. KEY button – opens a window which explains the meaning of symbols.
4. PAIR button – opens window with mate of read if data is available.
5. ZOOM slider – allows user to change the magnification of the display.
6. SELECTED OBJECT text field – displays name of selected object or “None” if nothing is selected.
7. MOUSE OVER text field – displays information about object where mouse is located or x coordinate if mouse is in an empty area.
8. Edit menu (top of window) Find item or Ctrl F will open a Find dialog to find features by name in the display.

Variant Navigator Tutorial

Examples of deletions, duplications, and insertions will be described. Items will be referred to by “SV Type” column and “Size” column entries with the exception of INS entries where the “From” column entry will be used since all INS have a size of 0.

1. DEL 1266 – Example of a homozygous deletion. This entry is about 14 rows down in the data table. Click on this row and the row color will become light blue and both buttons will become enabled.

Graphical user interface, application, table, Excel

Description automatically generated

Click the “Show Variant” button. A Reads window will open.

Graphical user interface, calendar

Description automatically generated

Read depth is the number of reads that align to a base in the reference sequence. Notice the read depth of 0 in the region of the DEL symbol immediately below the histogram. Pink bars joining read pairs indicate abnormally large insert sizes (arrow).

1. DUP 176 – Click on the entry 7 rows from the top and click the “Show Variant” button.

A screenshot of a computer

Description automatically generated

Notice the high read depth in the region of the DUP symbol. A large number of overlapping reads are often present in the region of a duplication.

1. INS 12625703 - Click on the entry about 17 rows from the top and click the “Show Variant” button.

Chart

Description automatically generated

The small blue rectangle indicates the position of the insertion call. There is a small amplification in the histogram in the region of the insertion call indicating a possible duplication. Soft-clipping and overlapping read pairs are often seen in regions with insertion calls. Some read pairs may be connected by blue bars in regions with insertions, but blue bars are not found in this sample.

SVs that affect gene features such as exons and coding regions (CDS) are most interesting since the effect of the SVs can sometimes be predicted if the gene has a known function. The VN data table has a filtering function that can remove all SVs that do not overlap gene features. Click the Edit menu at the top of the data table window and click the Filter menu. A dialog will appear:

Text

Description automatically generated with low confidence

Select the “Only Overlap” and “Only Nearby” check boxes and then click the OK button. The data table will reload and only contain SVs overlapping gene features or nearby to gene features. It is a good idea to look at the SVs that are nearby features since variant calls can often be inaccurate and visual examination may find that the SV actually overlaps the feature.

Graphical user interface

Description automatically generated

Click on the deletion DEL 4287 about 22 rows from the top. This deletion clearly overlaps the gene feature (yellow rectangle) indicated by the arrow above.

Here is an example of a gene feature that is present in the data table as nearby an SV but likely overlaps the SV. If you select DUP 9950 the only button available to click will be the “Show Variant in Two Windows” button since the size of this variant is > 6000. Click this button. Two windows will be displayed.

Graphical user interface

Description automatically generated

The blue arrow indicates where the DUP call ends but the actual end of the duplication is likely near the red arrow. The gene feature in the table listed as nearby overlaps this region and should be annotated.

Windows warning message:

![A screenshot of a computer

Description automatically generated]()

If a message like this pops up when double-clicking the bat file, open the bat file in a text editor such as Notepad++ or Sublime text editor and save the file. This will turn off this message.