



African Genome Variation Database

Tutorial

Searching

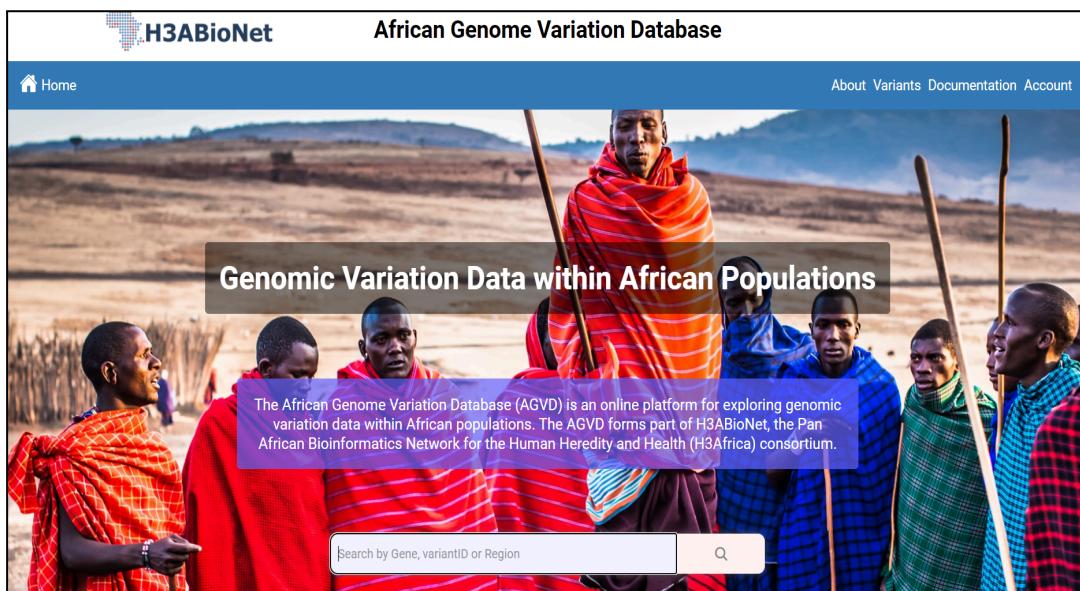


Figure 1: AGVD main search page

1. On the homepage, users can explore a world map view presenting the number of samples per country along with their corresponding projects. This feature utilizes data to visually represent the sample counts and their associated projects, all displayed based on their geographical locations.



Figure 2: Interactive World Map Displaying Sample Counts and Project Associations

2. **Search by Data category:**

- 4.1 **Search by variant:** The search bar enables users to search for variants using various formats, including Reference Sequence (RefSeq) identifiers (e.g., rs116600158), Ensembl gene or transcript identifiers (e.g., ENSG00000100342, ENST00000397278), or genomic locations or ranges (e.g., 19:7177679:C:T)
- 4.2 **Search by Gene name:** The search bar enables the users to search from the list of Genes names, e.g WASH2P
- 4.3 **Search by specific genomic region:** The search bar enable users to search for region in the format chr:start-end, e.g. 1:1000-40000

A screenshot of the AGVD homepage. At the top, there's a navigation bar with the H3ABioNet logo, the text "African Genome Variation Database", and links for "Home", "About", "Variants", "Documentation", and "Account". Below the navigation is a banner with a photograph of several Maasai men in traditional red and blue shukas. A dark overlay on the banner contains the text "Genomic Variation Data within African Populations". Underneath the banner, a purple box provides a brief description of the database. In the bottom right corner of the banner area, there's a search interface. A search bar contains the text "brca", and below it, three gene suggestions are listed: "BRCA1 (ENSG00000012048)", "BRCA1P1 (ENSG00000267595)", and "BRCA2 (ENSG00000139618)".

The African Genome Variation Database (AGVD) is an online platform for exploring genomic variation data within African populations. The AGVD forms part of H3ABioNet, the Pan African Bioinformatics Network for the Human Heredity and Health (H3Africa) consortium.

Figure 3: AGVD search by Variant. The search bar also includes a “type-hint” feature with the ability to auto-complete a search

Results

The following section discusses the results pages in relation to results obtained from WASH2P, 1:1000-40000, and 2-113601195-G-A searches.

1. **Gene:** When searching by gene, a list of relevant genes will appear below the search box. Simply click on your desired gene and press enter to access the gene results page



AGVD_v1.0.0 ▾ WASH2P

File upload

Upload a filter file of either rsIDs, PositionalIDs or gene names

Select File

Results

Total Variants Matching Query: 27

Export Results

Filters

	PosID	rsID	Gene(s)	AGVD MAF	GnomAD MAF
1	2-113580154-A-G ⓘ	rs1224155658	FAM138B,MIR1302-3,WASH2P	0.0000	-
2	2-113580202-G-T ⓘ	rs7355738	FAM138B,MIR1302-3,WASH2P	0.1089	0.2632
3	2-113594327-A-G ⓘ	rs560343909	WASH2P;DDX11L2	0.0000	-
4	2-113595958-C-T ⓘ	rs560542800	WASH2P;DDX11L2	0.0000	0.0001
5	2-113600668-A-G ⓘ	rs1207536552	WASH2P;DDX11L2	0.0000	0.0001
6	2-113601195-G-A ⓘ	rs11897134	WASH2P;DDX11L2	0.1046	0.2639
7	2-113601229-G-A ⓘ	rs1200055798	WASH2P;DDX11L2	0.0003	0.0000

Fig 4: *WASH2P* gene results page

The gene results page displays a list of variants matching the selected gene, along with the AGVD MAF and GenomAD MAF for each variant.

❖ Filters

On the results page, users are presented with a list of filter parameters. Users can access these filters by interacting with the accordion menu in the filter panel. The available filter menus in the pilot release include “Cluster,” “Allele Frequency,” “VariantIDs,” , “Region,” and “Genes.” Each filter menu, when expanded, provides explanatory text to assist users in understanding the acceptable types of parameters, the correct format for inputting them, and examples of valid search terms for each type of filter parameter.



Cluster

Show Frequencies and Plots for the Selected Clusters Only

Allele Frequency

Minor Allele Frequency CutOff Threshold

Less Than or Equal To Threshold
 Greater Than or Equal To Threshold

MAF Threshold = 0.00

ID

Variants IDs to search for in the format chrom:start:ref:alt, e.g.
2:113601195:G:A

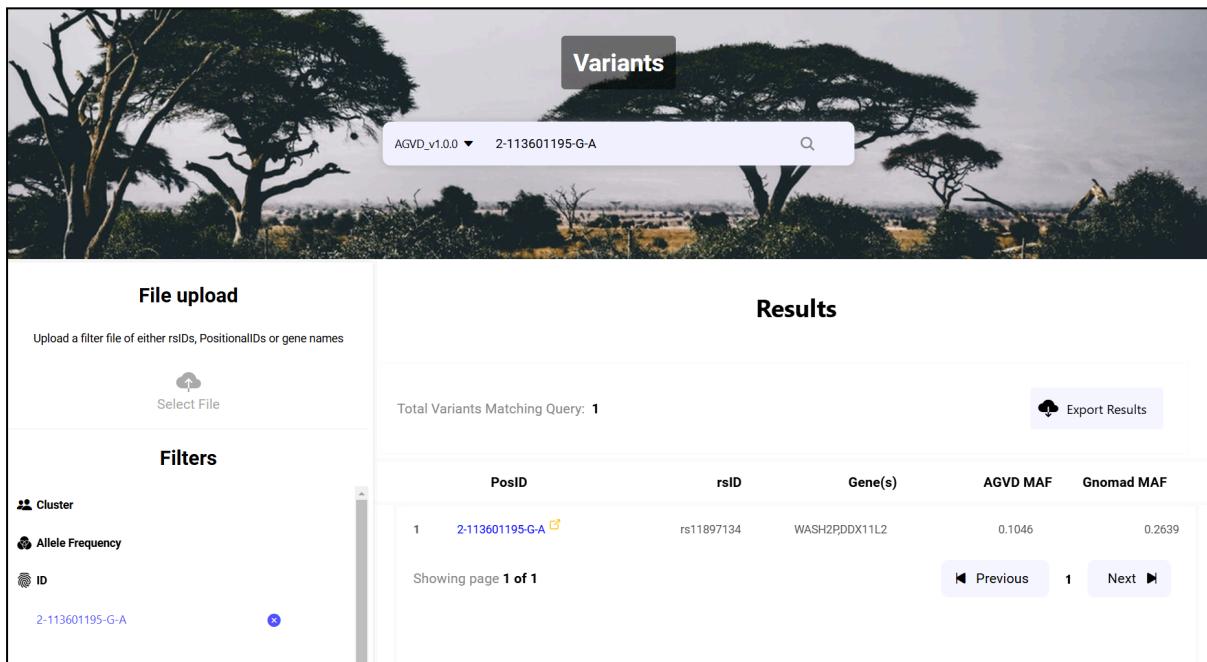
Region

A region to search for in the format chr:start-end, e.g. 1:1000-40000

Genes

List of Genes names, e.g WASH2P

2. **Variant ID:** Type your variantID on the search box. Then, click on the search button to access the variant results page



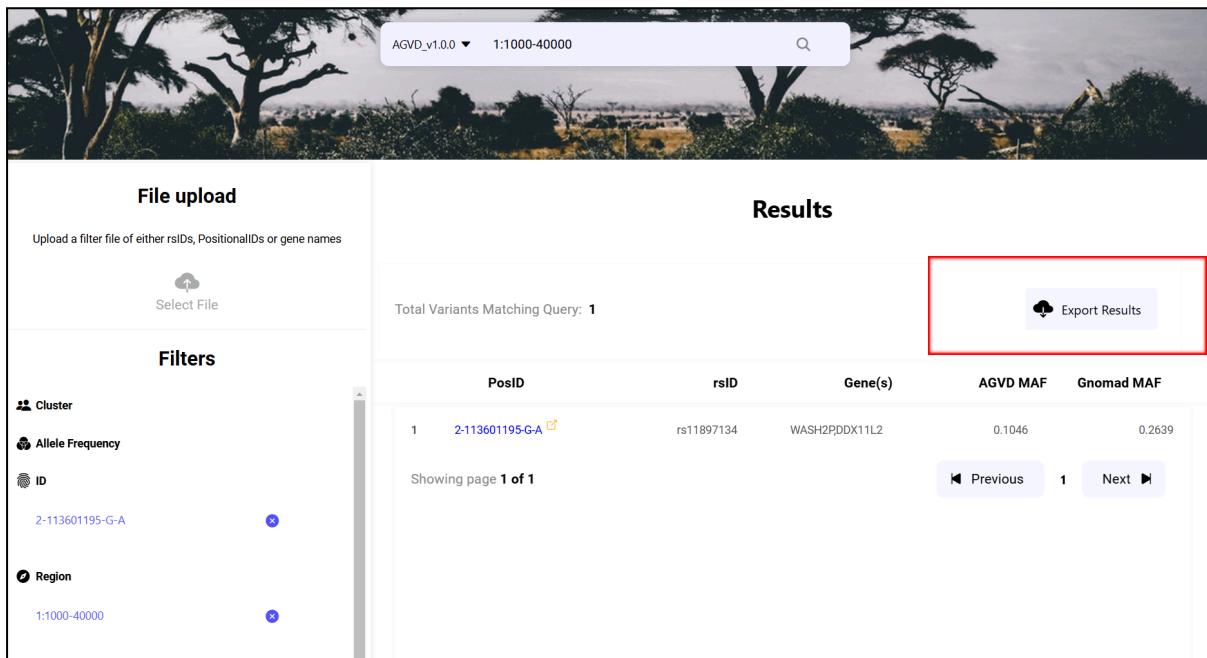
The screenshot shows the H3ABioNet Variant results page. At the top, there is a search bar with the query "2-113601195-G-A". Below the search bar, the results table has the following columns: PosID, rsID, Gene(s), AGVD MAF, and Gnomad MAF. One result is listed: PosID 1, rsID rs11897134, Gene(s) WASH2PDDX11L2, AGVD MAF 0.1046, and Gnomad MAF 0.2639. The "Export Results" button is highlighted with a red box.

PosID	rsID	Gene(s)	AGVD MAF	Gnomad MAF
1	2-113601195-G-A	WASH2PDDX11L2	0.1046	0.2639

Fig: 2-113601195-G-A variant results page

The variant results page displays the variant and the related gene, along with the AGVD MAF and GenomAD MAF for this variant.

- 3. Region:** Type the region you are searching for on the search box. Then, click on the search button to access the region results page



The screenshot shows the H3ABioNet Region results page. At the top, there is a search bar with the query "1:1000-40000". Below the search bar, the results table has the same columns as the variant results page. One result is listed: PosID 1, rsID rs11897134, Gene(s) WASH2PDDX11L2, AGVD MAF 0.1046, and Gnomad MAF 0.2639. The "Export Results" button is highlighted with a red box.

PosID	rsID	Gene(s)	AGVD MAF	Gnomad MAF
1	2-113601195-G-A	WASH2PDDX11L2	0.1046	0.2639

Fig: 2-113601195-G-A region results page



The region results page displays a list of variants and their related genes matching the selected region, along with the AGVD MAF and GenomAD MAF for each variant.

Filter file

- If you have a list of genes, rsids or positions (and alleles), this can be used as a filter byt using the File Upload feature...

File Upload

Upload a valid variant file or gene file, select appropriate filters and apply them below

rsIDs.tsv



Select Again

- You can export your research results as a tsv file

Results

Total Variants Matching Query: 1

Export Results