# Just-DNA-Seq

*Release 0.1*

**VovixSW**

**Nov 09, 2022**

# CONTENTS

**1 Contents** **3**
  1.1 Getting Started . . . 3
  1.2 Filtering Variants . . . 9
  1.3 Working With Annotated Variants . . . 14
  1.4 Viewing Reports . . . 16

**Just-DNA-Seq** is a project to facilitate working with human genomes at all levels - from clinical cases to personal curiosity, education and longevity. We envision the future when genomics is becoming more *available*, *understandable* and *useful* for everybody, especially those interested in life extension and improving human condition.

For this purpose we use OakVar framework to integrate *annotators* which are tools or databases accumulating what we actually know about the genome - genes, their influence on health or drug response, polygenic risk scores and so on.

We created the **Oakvar-Longevity** module for OakVar. It annotates user genome with longevity associated gene variants using the longevitymap annotator, and provides the longevity PRS, variants, drugs and major risks (oncorisk) reports with the longevity-combined reporter.

Check out the *Getting Started* section for further information, including the assemblies section to understand whether you have the appropriate genome assembly.

**NOTE:** This project is under active development.

CHAPTER

ONE

# CONTENTS

## 1.1 Getting Started

**NOTE:** *Both OpenCravat and OakVar can be used to annotate a human genome. At the beginning of the project we used OpenCravat as a framework. However, as OakVar is based on OpenCravat and contains more advanced features customized specially for personal longevity genomics, we decided to base further development of the project on OakVar.*

### 1.1.1 Installing OakVar

Since our module is based on OakVar, you have to install OakVar first to run our module, if it is not already installed. OakVar docs: https://oakvar.readthedocs.io/en/latest/

**Pre-requirements for Oakvar:**

- installed conda/mamba environment management systems, or you can use their lighter versions: miniconda/micromamba
- installed python and pip

You can find documentation for mamba here: https://mamba.readthedocs.io/en/latest/index.html

And for conda here: https://docs.conda.io/en/latest/

*The installation of OakVar and the further work should proceed after activation of an environment created by Conda/Mamba or Miniconda/Micromamba.*

### 1.1.2 Installing Annotators

For the Longevity module to work, you need to install the following annotators:

1. clinvar
2. dbsnp
3. gnomad
4. ncbigene
5. omim
6. pubmed
7. longevitymap

You can install them by using terminal or Oakvar GUI.

**Installation using terminal:**

Use the following command:

```
ov module install module_name
```

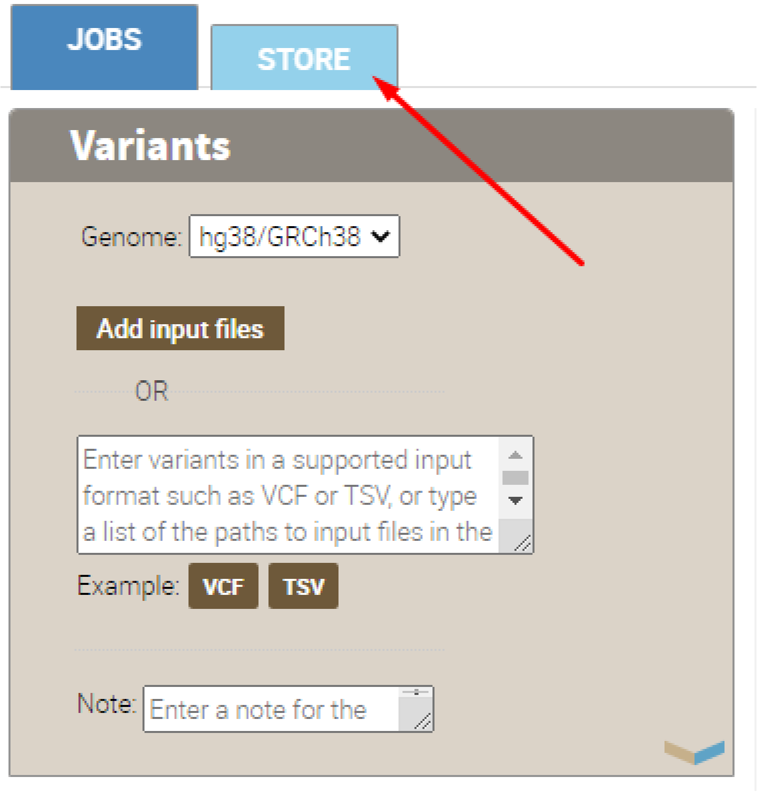
**Installation using GUI:**

To activate Oakvar GUI, use the following command:

```
ov gui
```

After the execution GUI will be opened in your browser.

Go to "Store":

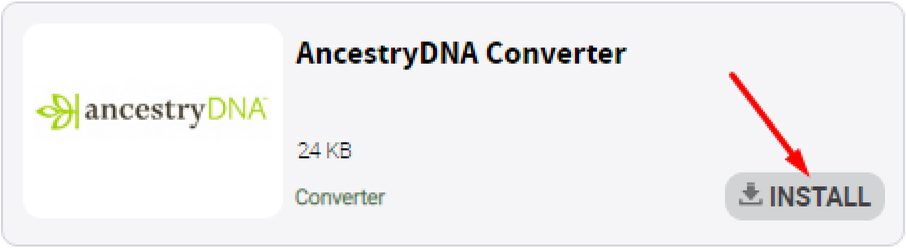

JOBS
STORE
Variants
Add input files
OR
Example:
VCF
TSV

Genome: hg38/GRCh38

Enter variants in a supported input format such as VCF or TSV, or type a list of the paths to input files in the

Note: Enter a note for the

Find annotators and install them:


ancestryDNA
AncestryDNA Converter
24 KB
Converter
INSTALL

### 1.1.3 Installing the Reporter

**Installation using terminal:**

Use the following command in terminal:

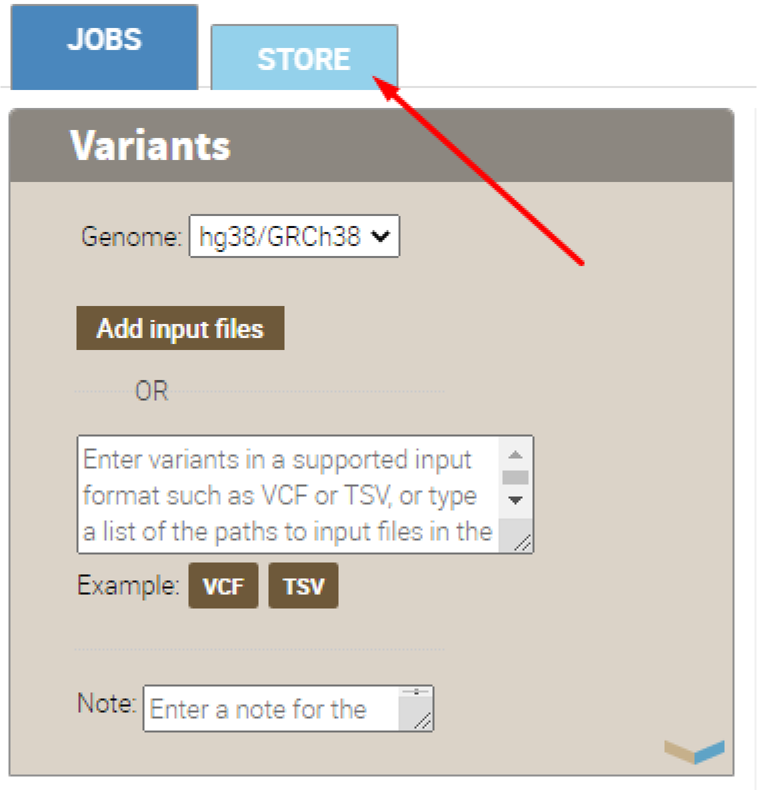
```
ov module install longevity-combinedreporter
```

**Installation using GUI:**

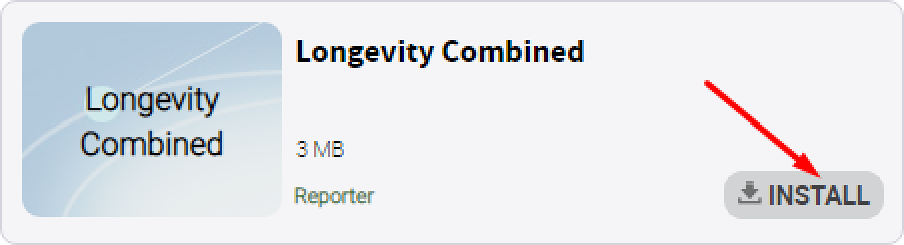
To activate Oakvar GUI, use the following command:

```
ov gui
```

Go to "Store":

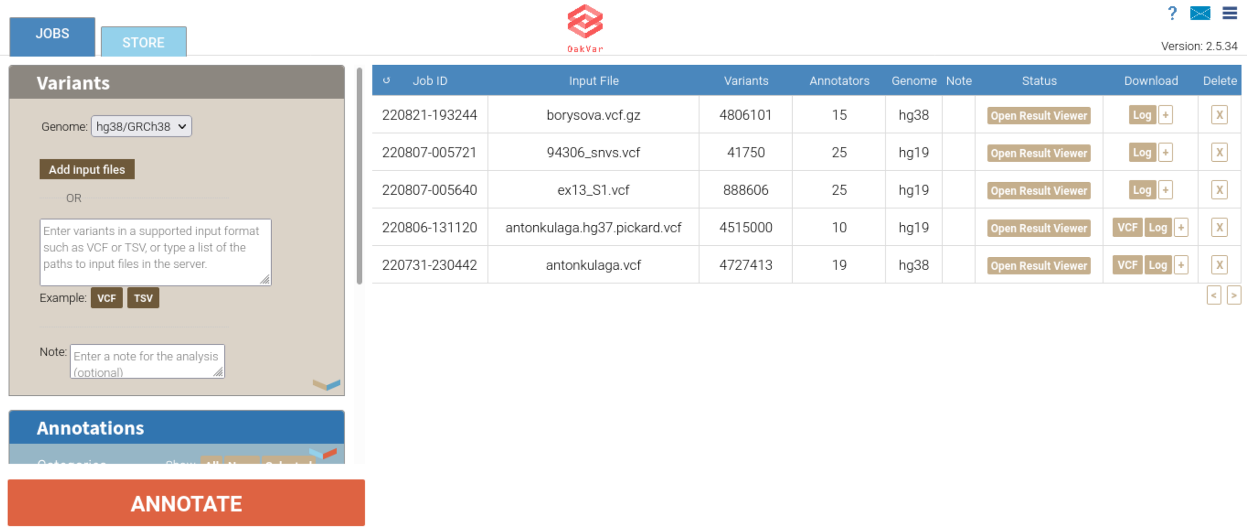

JOBS
STORE
Variants
Genome:
hg38/GRCh38
Add input files
OR
Enter variants in a supported input format such as VCF or TSV, or type a list of the paths to input files in the
Example:
VCF
TSV
Note:
Enter a note for the

Find the reporter called "longevity-combinedreporter" and install it:


Longevity Combined
Longevity Combined
3 MB
Reporter
INSTALL

### 1.1.4 Loading Genome Files

1. Open OakVar in your browser. You will see the index page:


JOBS
STORE
OakVar
Version: 2.5.34
Variants
Add input files
OR
Enter variants in a supported input format such as VCF or TSV, or type a list of the paths to input files in the server.
Example: VCF TSV
Note: Enter a note for the analysis (optional)
Annotations
ANNOTATE

Genome: hg38/GRCh38

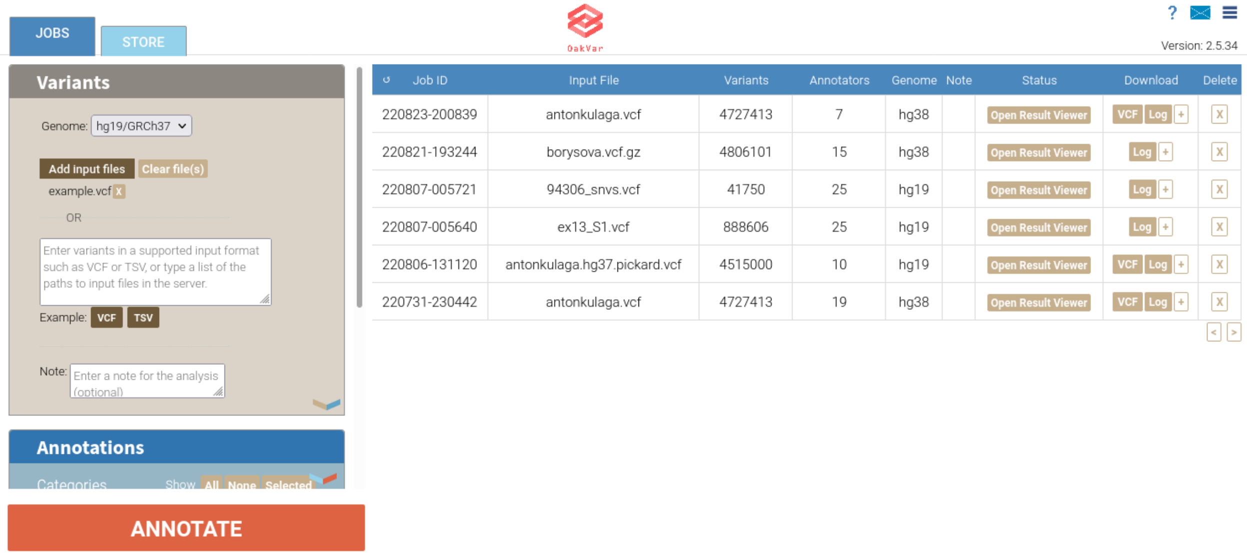
| Job ID | Input File | Variants | Annotators | Genome | Note | Status | Download | Delete |
|---|---|---|---|---|---|---|---|---|
| 220821-193244 | borysova.vcf.gz | 4806101 | 15 | hg38 | | Open Result Viewer | Log + | X |
| 220807-005721 | 94306_snvs.vcf | 41750 | 25 | hg19 | | Open Result Viewer | Log + | X |
| 220807-005640 | ex13_S1.vcf | 888606 | 25 | hg19 | | Open Result Viewer | Log + | X |
| 220806-131120 | antonkulaga.hg37.pickard.vcf | 4515000 | 10 | hg19 | | Open Result Viewer | VCF Log + | X |
| 220731-230442 | antonkulaga.vcf | 4727413 | 19 | hg38 | | Open Result Viewer | VCF Log + | X |

2. In the **Variants** section you should choose the right assembly version of the **Genome:** hg38/GRCh38, hg19/GRCh37, or hg18/GRCh36.

For example we'll take a small VCF file of the hg19/GRCh37 version named example.vcf.

3. Click **Add input files**. A file upload dialog will open, allowing to browse and select the vcf file (or multiple files at once).

After loading, the file(s) will be shown next to the **Add input files** button along with another button **Clear file(s)** and a small **X** button next to each file name. If you click that **X**, the appropriate file will be deleted. If you click **Clear file(s)**, all the files you loaded will be deleted.


JOBS
STORE
OakVar
Version: 2.5.34
Variants
Add input files
Clear file(s)
OR
Enter variants in a supported input format such as VCF or TSV, or type a list of the paths to input files in the server.
Example: VCF TSV
Note: Enter a note for the analysis (optional)
Annotations
Categories
Show All None Selected
ANNOTATE

Genome: hg19/GRCh37

example.vcf X

| Job ID | Input File | Variants | Annotators | Genome | Note | Status | Download | Delete |
|---|---|---|---|---|---|---|---|---|
| 220823-200839 | antonkulaga.vcf | 4727413 | 7 | hg38 | | Open Result Viewer | VCF Log + | X |
| 220821-193244 | borysova.vcf.gz | 4806101 | 15 | hg38 | | Open Result Viewer | Log + | X |
| 220807-005721 | 94306_snvs.vcf | 41750 | 25 | hg19 | | Open Result Viewer | Log + | X |
| 220807-005640 | ex13_S1.vcf | 888606 | 25 | hg19 | | Open Result Viewer | Log + | X |
| 220806-131120 | antonkulaga.hg37.pickard.vcf | 4515000 | 10 | hg19 | | Open Result Viewer | VCF Log + | X |
| 220731-230442 | antonkulaga.vcf | 4727413 | 19 | hg38 | | Open Result Viewer | VCF Log + | X |

### 1.1.5 Installing and Selecting Necessary Annotators

Scroll the left area down to the **Annotations** section.

Here you can see the categories of annotators available for selection (above) and checkboxes for particular annotators.

Annotators are software modules which can be developed, added and installed as needed. If any necessary annotator is not yet installed, you can install it on the **STORE** tab in the upper left corner.

All annotators can be divided into 2 groups:

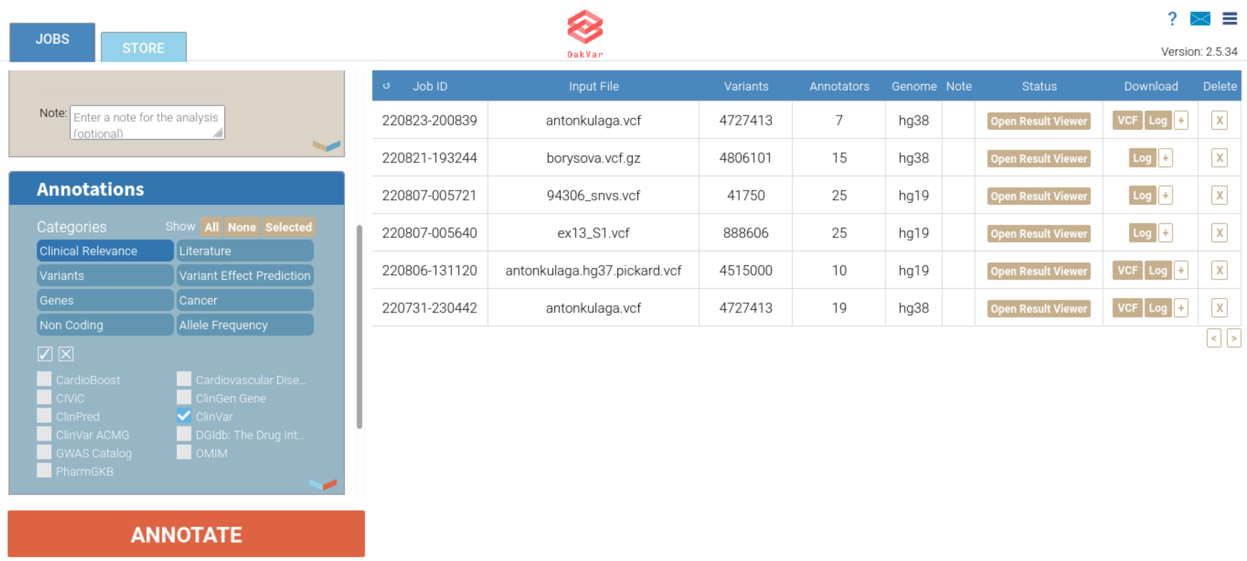
1) Tools that predict pathogenicity (**bold**)
2) Tools that provide information like databases

Here are their internal (coded) module names:

- **cadd_exome** (1.6.1) - CADD is a tool for scoring the deleteriousness of single nucleotide variants as well as insertion/deletions variants in the human genome
- gnomad_gene (2.2.1) - gene level population statistics from gnomAD
- pubmed (1.1.5) - articles related to a particular gene
- **clingen** (1.0.1) - NIH-funded resource that defines the clinical relevance of genes and variants
- **clinpred** (1.0.0) - prediction tool to identify disease-relevant nonsynonymous single nucleotide variants
- clinvar (2021.10.01) - ClinVar is an archive of reports of the relationships among human variations and phenotypes, as well as interpretations of clinically relevant variants (Uncertain significance, Likely pathogenic, Pathogenic etc.)
- mitomap (1.1.0) - a human mitochondrial genome database
- ncbigene (2019.08.02) - gene descriptions from NCBI (National Center for Biotechnology Information) Gene database
- omim (1.0.0) - catalog of human genes and genetic disorders and traits
- **prec** (3.6.0) - provides a database identifying rare and likely deleterious loss-of-function (LoF) alleles
- **provean** (1.0.0) - tool which predicts whether an amino acid substitution or indel has an impact on the biological function of a protein
- **revel** (2020.12.02) - ensemble method for predicting the pathogenicity of missense variants based on a combination of scores from 13 individual tools
- **sift** (1.2.0) - predicts whether an amino acid substitution affects protein function based on sequence homology and the physical properties of amino acids
- GnomADD - aggregating and harmonizing both exome and genome sequencing data from a wide variety of large-scale sequencing projects
- PharmGKB - an NIH-funded resource that provides information about how human genetic variation affects response to medications
- dbSNP - the Single Nucleotide Polymorphism Database is a free public archive for genetic variation within and across different species developed and hosted by the National Center for Biotechnology Information (NCBI) in collaboration with the National Human Genome Research Institute (NHGRI)

Once an annotator is installed, you can select in on the **JOBS** tab in the upper left corner.

For example, let's select the **ClinVar** annotator from the **Clinical Relevance** category:


JOBS
STORE
OakVar
Version: 2.5.34
Note:
Enter a note for the analysis (optional)
Annotations
Categories
Show All None Selected
Clinical Relevance
Literature
Variants
Variant Effect Prediction
Genes
Cancer
Non Coding
Allele Frequency
ANNOTATE

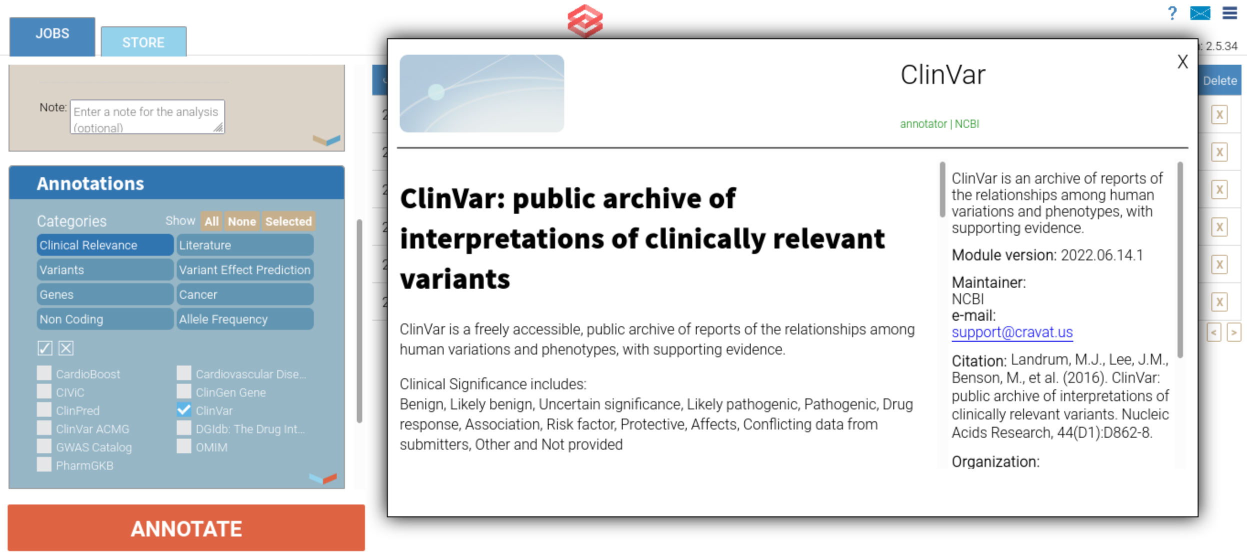
☐ CardioBoost
☐ CIViC
☐ ClinPred
☐ ClinVar ACMG
☐ GWAS Catalog
☐ PharmGKB
☐ Cardiovascular Dise...
☐ ClinGen Gene
☑ ClinVar
☐ DGIdb: The Drug Int...
☐ OMIM

| Job ID | Input File | Variants | Annotators | Genome | Note | Status | Download | Delete |
|---|---|---|---|---|---|---|---|---|
| 220823-200839 | antonkulaga.vcf | 4727413 | 7 | hg38 | | Open Result Viewer | VCF Log + | X |
| 220821-193244 | borysova.vcf.gz | 4806101 | 15 | hg38 | | Open Result Viewer | Log + | X |
| 220807-005721 | 94306_snvs.vcf | 41750 | 25 | hg19 | | Open Result Viewer | Log + | X |
| 220807-005640 | ex13_S1.vcf | 888606 | 25 | hg19 | | Open Result Viewer | Log + | X |
| 220806-131120 | antonkulaga.hg37.pickard.vcf | 4515000 | 10 | hg19 | | Open Result Viewer | VCF Log + | X |
| 220731-230442 | antonkulaga.vcf | 4727413 | 19 | hg38 | | Open Result Viewer | VCF Log + | X |

**Note:** An annotator may belong to multiple categories at once.

The checkbox and **X** buttons between the categories and the annotators sections allow to select all of the displayed annotator chechboxes or to clear all of them.

If you right-click any annotator, a pop-up window with its description will open in the right area:


JOBS
STORE
Annotations
ANNOTATE
X
ClinVar
annotator | NCBI

**ClinVar: public archive of interpretations of clinically relevant variants**

ClinVar is a freely accessible, public archive of reports of the relationships among human variations and phenotypes, with supporting evidence.

Clinical Significance includes:
Benign, Likely benign, Uncertain significance, Likely pathogenic, Pathogenic, Drug response, Association, Risk factor, Protective, Affects, Conflicting data from submitters, Other and Not provided

ClinVar is an archive of reports of the relationships among human variations and phenotypes, with supporting evidence.

Module version: 2022.06.14.1

Maintainer:
NCBI
e-mail:
support@cravat.us

Citation: Landrum, M.J., Lee, J.M., Benson, M., et al. (2016). ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Research, 44(D1):D862-8.

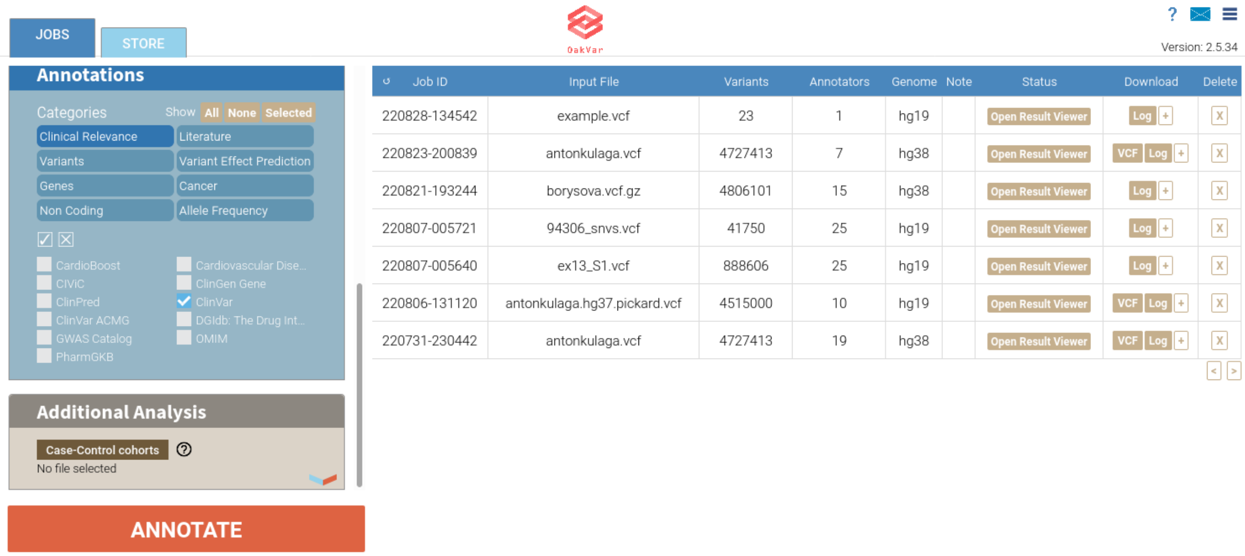
Organization:

For our purposes we will need the following annotators: **ClinVar** (clinvar), **dbSNP** (dbsnp), **gnomAD3** (gnomad), **LongevityMap** (longevitymap), **NCBI Gene** (ncbigene), **OMIM** (omim), and **PubMed** (pubmed). If any of them are missing, install them on the **STORE** tab, then go back to **JOBS**, in the **Annotations** section select **All** categories, and then select each of the annotator checkboxes.

### 1.1.6 Annotating Your Genome

When you select all the annotators you need, click the large **ANNOTATE** button below in the left area.

Annotating a large genome file may take some time. While loading, it will appear in the right area on the top of the list, displaying different stages of the processing in the **Status** column, and when finished, the **Open Results Viewer** button will appear in that column of the particular genome row:


JOBS
STORE
OakVar
Version: 2.5.34
Annotations
Categories
Show All None Selected
Clinical Relevance
Literature
Variants
Variant Effect Prediction
Genes
Cancer
Non Coding
Allele Frequency
Additional Analysis
Case-Control cohorts
No file selected
ANNOTATE

☐ CardioBoost
☐ CIViC
☐ ClinPred
☐ ClinVar ACMG
☐ GWAS Catalog
☐ PharmGKB
☐ Cardiovascular Dise...
☐ ClinGen Gene
☑ ClinVar
☐ DGIdb: The Drug Int...
☐ OMIM

| Job ID | Input File | Variants | Annotators | Genome | Note | Status | Download | Delete |
|---|---|---|---|---|---|---|---|---|
| 220828-134542 | example.vcf | 23 | 1 | hg19 | | Open Result Viewer | Log + | X |
| 220823-200839 | antonkulaga.vcf | 4727413 | 7 | hg38 | | Open Result Viewer | VCF Log + | X |
| 220821-193244 | borysova.vcf.gz | 4806101 | 15 | hg38 | | Open Result Viewer | Log + | X |
| 220807-005721 | 94306_snvs.vcf | 41750 | 25 | hg19 | | Open Result Viewer | Log + | X |
| 220807-005640 | ex13_S1.vcf | 888606 | 25 | hg19 | | Open Result Viewer | Log + | X |
| 220806-131120 | antonkulaga.hg37.pickard.vcf | 4515000 | 10 | hg19 | | Open Result Viewer | VCF Log + | X |
| 220731-230442 | antonkulaga.vcf | 4727413 | 19 | hg38 | | Open Result Viewer | VCF Log + | X |

### 1.1.7 Opening Your Annotated Genome

Now click the **Open Results Viewer** button, and the annotated genome will open in a new browser tab/window.

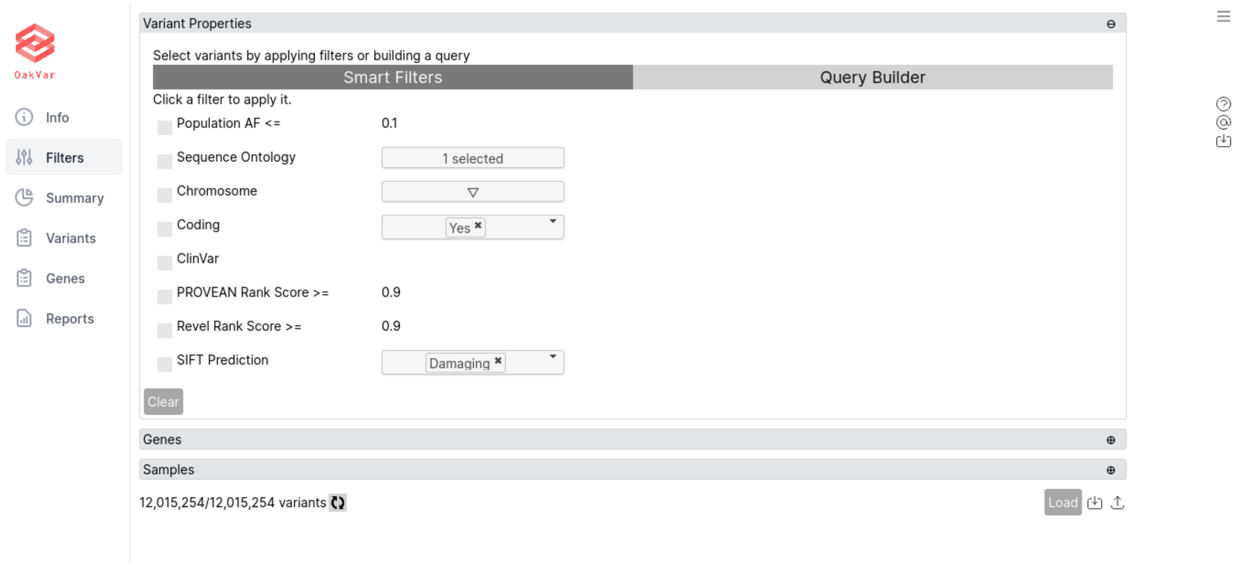
## 1.2 Filtering Variants

Filters in OakVar allow to select those variants which are relevant. As the number of variants in a genome usually is very large, you need to filter them first. OakVar cannot load more than 100,000 variants at once.

### 1.2.1 The Filter page

Select the **Filter** page in the Result Viewer. There are sections where you can filter the variants:

- *Variant Properties*, with **Smart Filters** and **Query Builder** tabs
- *Genes*, where you can type in any particular gene names
- *Samples*, which is used mainly for oncological purposes and is not used in Just-DNA-Seq.



Select variants by applying filters or building a query

Click a filter to apply it.

☐ Population AF <= 0.1
☐ Sequence Ontology 1 selected
☐ Chromosome
☐ Coding Yes
☐ ClinVar
☐ PROVEAN Rank Score >= 0.9
☐ Revel Rank Score >= 0.9
☐ SIFT Prediction Damaging

12,015,254/12,015,254 variants

### 1.2.2 Using Smart Filters

Here are various useful filters:

**Population AF «==*** allows to set the maximum allele frequency in population.

**Sequence Ontology** allows to choose one or more sequence ontologies.

**Chromosome** allows to choose one or more chromosomes *or their specific versions*. E.g. chr1, chr10 and so on.

**Coding** allows to include only coding or noncoding variants.

**ClinGen** allows to include only variants with data from ClinGen.

**ClinVar** allows to include only variants with data from ClinVar.

**dbSNP Common ID** allows to include only variants with dbSNP common IDs.

**PROVEAN Rank Score >=** includes variants with PROVEAN rank score not less than 0.9.

**Revel Rank Score >=** includes variants with Revel rank score not less than 0.9.

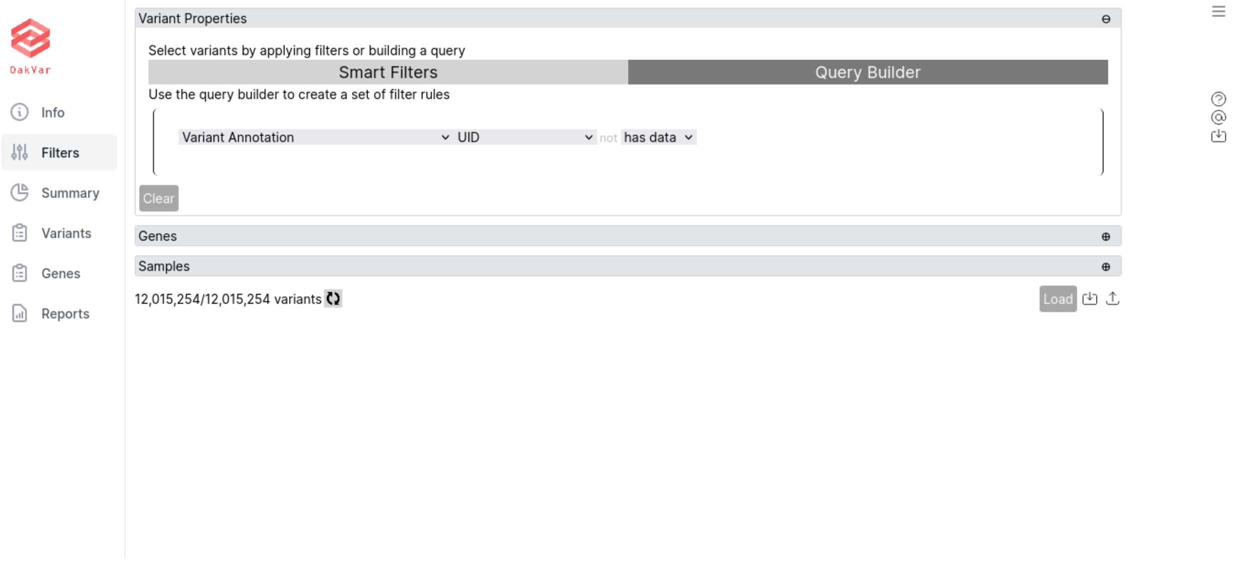
**SIFT Prediction** can be set as Damaging or Tolerated.

### 1.2.3 Using Query Builder

Here you can create a set of filter rules.

By default, an opening (left) parenthesis appears with **+** and **(** buttons in the lower left corner, and a greyed out **NOT** switch appears if you hover the mouse in the upper left corner, which allows to make the following rule negative by clicking on it. Clicking **NOT** once again deactivates it.

Click **+** to add a rule. A line of boxes will appear:


OakVar
Info
Filters
Summary
Variants
Genes
Reports
Variant Properties
Select variants by applying filters or building a query
Smart Filters
Query Builder
Use the query builder to create a set of filter rules
Variant Annotation
UID
not
has data
Clear
Genes
Samples
12,015,254/12,015,254 variants
Load

The first drop-down box is the source to which the rule will apply. For example: Variant Annotation, ClinVar, PharmGKB etc. The second drop-down box allow to select an item in the source to apply the rule. E.g. UID, Chrom, Position, Gene etc. The following "not" switch, greyed out (inactive) by default, allows to select if the following condition should apply or should not apply. For the latter, click the "not" word, and it will become black (active). To remove "not" from the condition, just click it again, and it will be greyed out. The next drop-down allows to select the condition from one of the following:

**has data** - if the item being searched contains any data

**equals** - opens one more box where you can enter what the item should be equal to

**is empty** - if the item being searched is empty

**in range** - opens two boxes where you enter the boundaries of the range where the item should be

**<=** - if the item is less or equal to the value in the following box

**>=** - if the item is greater or equal to the value in the following box

At the end of the line, a small "x" allows to delete the whole rule by clicking on it.

If you click **+** once again, another rule is added, and between them the **and** operator is displayed by default, meaning that to satisfy the filter, both rules should apply. You can change it to **or** by clicking on it, so that to satisfy the filter, one of rules being true may be enough. Clicking **or** once again turns it to **and** again.

You can add as many rules as you wish, and the operators **and / or** between them will follow the general priority logic of boolean operations, i.e. **and** has the priority over **or**, as in any program code.

To change the priority and build more complex logical rules, you can click **(** making a separate set of rules (in parentheses), which have higher priority, as in mathematical operations. Note than the **and / or** operator which appears before the parentheses depends of the previous operator selected, i.e. if it was **or**, the next one will also be **or**, and vice versa. You can always change the operators by clicking on them.

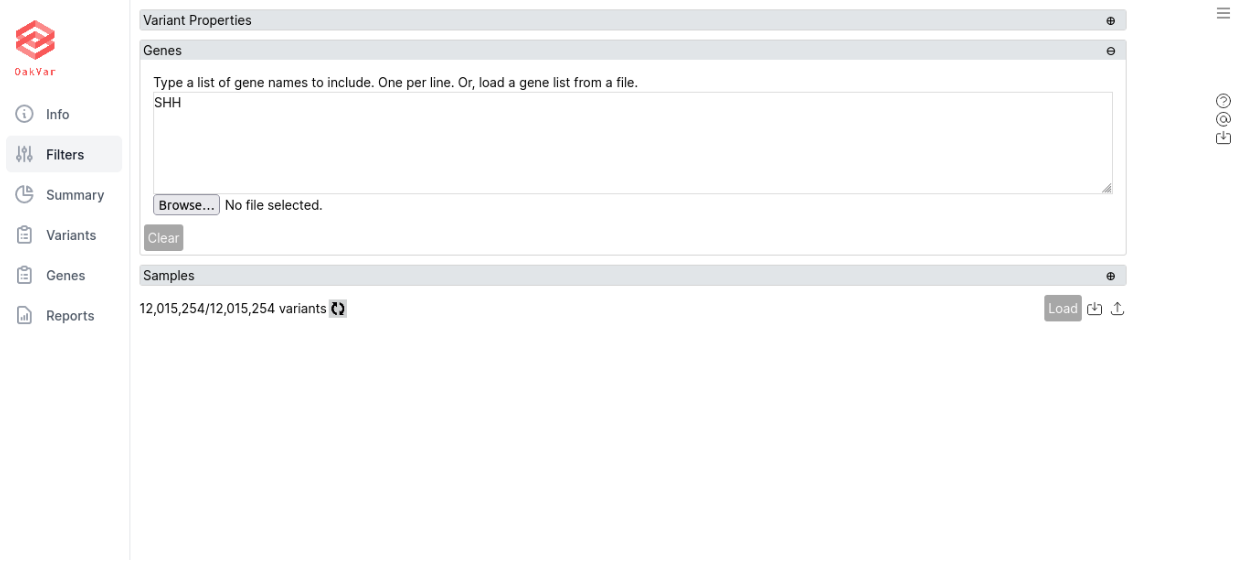
Within the parentheses, you can create any number of rules, and there are separate **+** and **(** buttons to add new rules and nested parentheses inside the parentheses. Also in the upper left corner a separate **NOT** switch appears if you hover the mouse over it.

You can also move any rule to another rule. To do this, drag an anchor **||** which appears from the left side of the rule if you hover the mouse there, and drop it on any rounded **+** anchors which appear between rules and/or parentheses (not on the **+** button that adds rules).

**BUG NOTE: If dropping a rule just before or after itself, it redirects browser to an error page. In this case all the previously made filter settings may not be saved. Please avoid dropping a rule before or after itself until this bug is fixed in Oakvar.**

### 1.2.4 Filtering by Genes

Switch to the **Genes** section and enter any particular gene names, one per line. Also you can load them from a file by clicking **Browse…**



Type a list of gene names to include. One per line. Or, load a gene list from a file.

SHH

No file selected.

12,015,254/12,015,254 variants

### 1.2.5 Clearing Filters

Under any section you can click the **Clear** button to remove any filter settings from that section.

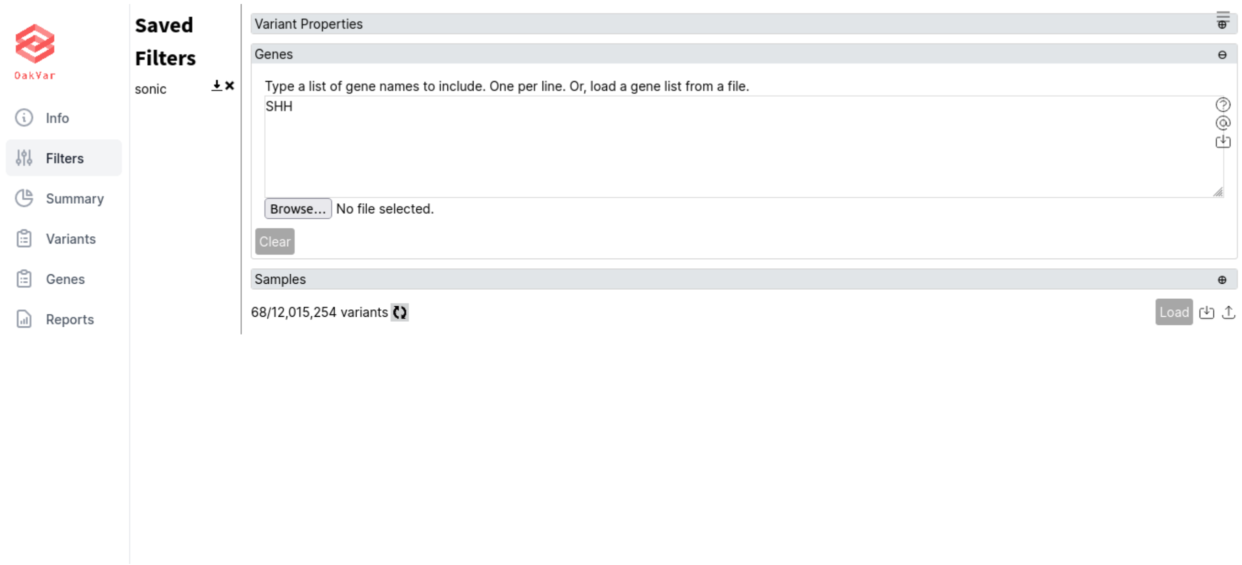
### 1.2.6 Saving and Importing Filters

You can save the filter (the whole set of rules) in OakVar for further loading, as well as exporting to a file, or import it from a file.

To save the filter, click the middle button ("inward arrow") in the lower right corner of the page, and enter the filter name.

**NOTE:** Filters are saved internally in OakVar, i.e. on the server if using a remote installation. To have a filter saved into a local file, export it after saving.

The saved filter appears in the left part of the page in the **Saved Filters** list:



Type a list of gene names to include. One per line. Or, load a gene list from a file.

SHH

No file selected.

68/12,015,254 variants

To load a saved filter, just click its name. To export a saved filter into a file, click the icon with a down arrow next to its name. To delete a saved filter, click the **X** icon in its line.

To import a filter from a file, click the "up arrow" (rightmost) button in the lower right corner of the page, and browse for a file to import (e.g. pathogenic.json). Clicking **Open** in the browse window loads the filter. **NOTE:** the filter is not saved automatically, you need to save it using the "Save filter" (inward arrow) button if you want to keep it on the server for further working.

### 1.2.7 Loading Filtered Variants

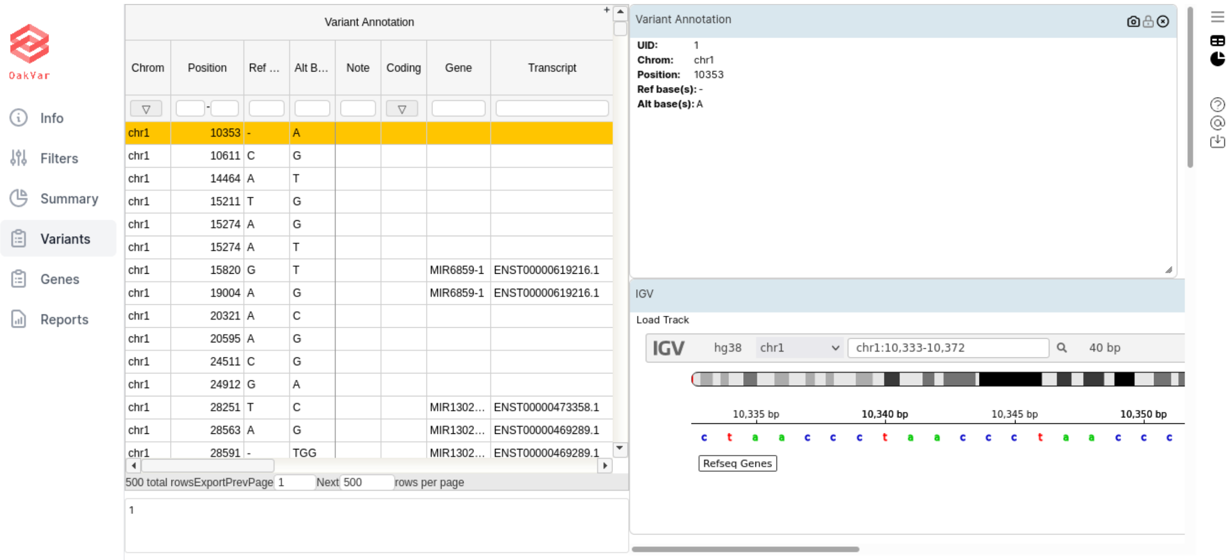
When building a filter, you can click the refresh button next to the number of variants (e.g. **68/12,015,254 variants**) in the lower left corner of the page to check how many results the filter provides. If the number is small enough, when the filter is ready, click **Load** in the lower right corner of the page. After loading the filter, the number of variants in the lower left corner (the first number before the slash, while the second one is the total number of variants and doesn't change) may be updated.

When the filtered variants are loaded, you can proceed to the **Variant** tab to analyse them (see the next section).

## 1.3 Working With Annotated Variants

After applying the necessary filters, select the **Variants** page.

By default a combined view is displayed, with both table and widgets:



| Variant Annotation | | | | | | | |
|---|---|---|---|---|---|---|---|
| Chrom | Position | Ref ... | Alt B... | Note | Coding | Gene | Transcript |
| chr1 | 10353 | - | A | | | | |
| chr1 | 10611 | C | G | | | | |
| chr1 | 14464 | A | T | | | | |
| chr1 | 15211 | T | G | | | | |
| chr1 | 15274 | A | G | | | | |
| chr1 | 15274 | A | T | | | | |
| chr1 | 15820 | G | T | | | MIR6859-1 | ENST00000619216.1 |
| chr1 | 19004 | A | G | | | MIR6859-1 | ENST00000619216.1 |
| chr1 | 20321 | A | C | | | | |
| chr1 | 20595 | A | G | | | | |
| chr1 | 24511 | C | G | | | | |
| chr1 | 24912 | G | A | | | | |
| chr1 | 28251 | T | C | | | MIR1302... | ENST00000473358.1 |
| chr1 | 28563 | A | G | | | MIR1302... | ENST00000469289.1 |
| chr1 | 28591 | - | TGG | | | MIR1302... | ENST00000469289.1 |

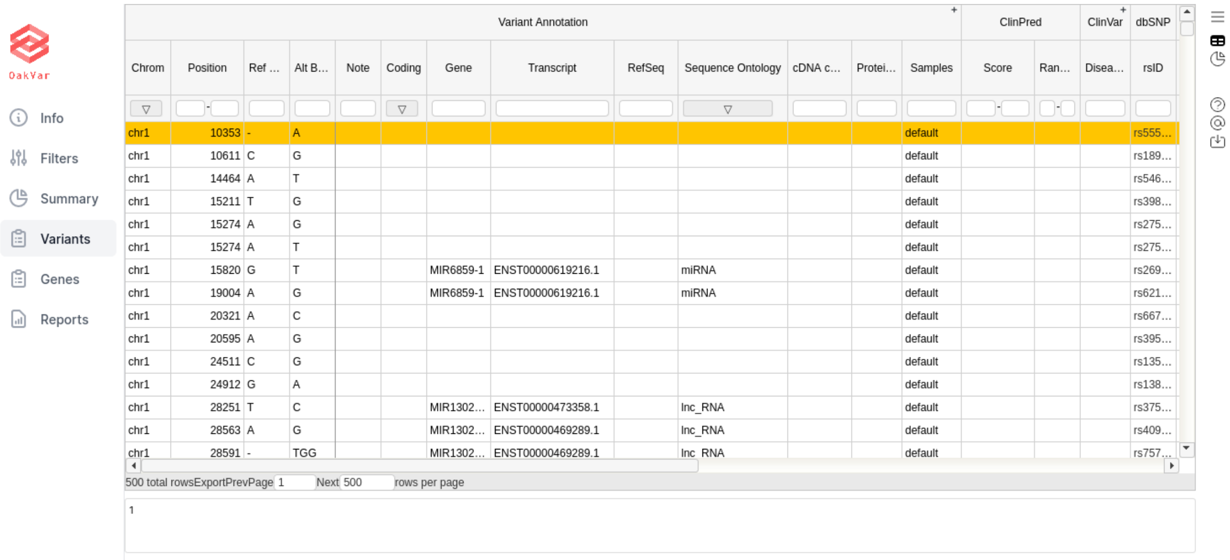
500 total rowsExportPrevPage 1 Next 500 rows per page

1

Variant Annotation

**UID:** 1
**Chrom:** chr1
**Position:** 10353
**Ref base(s):** -
**Alt base(s):** A

By clicking the icons in the upper right corner, you can toggle on/off the table view (window-like icon) and the widgets view (piechart-like icon). For our purposes first of all we need the table view:



| Variant Annotation | | | | | | | | | | | | | ClinPred | | ClinVar | dbSNP |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Chrom | Position | Ref ... | Alt B... | Note | Coding | Gene | Transcript | RefSeq | Sequence Ontology | cDNA c... | Protei... | Samples | Score | Ran... | Disea... | rsID |
| chr1 | 10353 | - | A | | | | | | | | | default | | | | rs555... |
| chr1 | 10611 | C | G | | | | | | | | | default | | | | rs189... |
| chr1 | 14464 | A | T | | | | | | | | | default | | | | rs546... |
| chr1 | 15211 | T | G | | | | | | | | | default | | | | rs398... |
| chr1 | 15274 | A | G | | | | | | | | | default | | | | rs275... |
| chr1 | 15274 | A | T | | | | | | | | | default | | | | rs275... |
| chr1 | 15820 | G | T | | | MIR6859-1 | ENST00000619216.1 | | miRNA | | | default | | | | rs269... |
| chr1 | 19004 | A | G | | | MIR6859-1 | ENST00000619216.1 | | miRNA | | | default | | | | rs621... |
| chr1 | 20321 | A | C | | | | | | | | | default | | | | rs667... |
| chr1 | 20595 | A | G | | | | | | | | | default | | | | rs395... |
| chr1 | 24511 | C | G | | | | | | | | | default | | | | rs135... |
| chr1 | 24912 | G | A | | | | | | | | | default | | | | rs138... |
| chr1 | 28251 | T | C | | | MIR1302... | ENST00000473358.1 | | lnc_RNA | | | default | | | | rs375... |
| chr1 | 28563 | A | G | | | MIR1302... | ENST00000469289.1 | | lnc_RNA | | | default | | | | rs409... |
| chr1 | 28591 | - | TGG | | | MIR1302... | ENST00000469289.1 | | lnc_RNA | | | default | | | | rs757... |

500 total rowsExportPrevPage 1 Next 500 rows per page

1

The table contains columns and column sets with general information about the filtered variants, as well as those connected to certain annotators. Some logically grouped column sets (by a particular annotator) can be extended or collapsed by clicking the **+/-** sign in the upper right corner of the column set (the topmost row). If you filtered by particular annotators, especially using "has data" condition, for other annotators it may show nothing for that particular variants, and they can be collapsed for convenience.

Each row of the table represents a variant that you can research.

The most important column groups for us are listed below, along with columns:

### 1.3.1 Variant Annotation

UID - the variant number in this (filtered) sequence

Chrom - chromosome where the variant is located. Chromosome names are ‘chr1’ to ‘chr22’, ‘chrX’, ‘chrY’ and ‘chrM’.

Position - chromosomal position of the variant. The first position in each chromosome is position 1.

Ref Base - reference allele at this chromosomal position (one of A, C, G, T, and N).

Alt Base - alternative allele; called based on reads mapping to this chromosomal position.

Note - note for the variant, if available.

Coding - whether this gene variant is coding.

Gene - the gene this variant belongs to.

Transcript - GENCODE transcript.

RefSeq - the reference sequence.

Sequence Ontology - could be: missense variant, start lost, stop gained, or stop lost.

cDNA change - change of coding DNA.

Protein Change - change of protein being synthesized.

All Mappings - expression showing all the mappings.

Sample Count - the number of samples which contain the variant.

Samples - samples which contain the variant.

Tags - variant tags from the input file.

### 1.3.2 ClinVar

Clinical Significance - the level of clinical significance of the variant.

Disease Ref Nums - disease reference numbers.

Disease Names - names of diseases associated with the variant.

Review Status - the level of review supporting clinical significance.

ClinVar ID - ID in the ClinVar database.

Significance Detail - additional detail on clinical significance used when it is conflicting.

### 1.3.3 dbSNP

rsID - the database identifier (“rs” number) of this variant in dbSNP.

This column is empty if the observed variant is not described in dbSNP. Such variants can be extremely rare variants or technical artifacts.

### 1.3.4 LongevityMap Annotator

LongevityMap ID - ID(s) of the variant in LongevityMap.

Significance - could be: significant, non-significant, or conflicted.

Source Population - population the data were obtained on, e. g. Danish or American (Caucasian).

dbSNP id - ID of the variant in dbSNP.

Associated Genes - genes associated with the variant.

PubMed ID - ID of the variant in PubMed.

Info - additional information.

Description - detailed description of the research.

allele - allele associated with the variant.

state - state of the variant.

zygosity - zygosity of the variant vs OV.

weight - weight of the variant.

priority - priority of the variant.

### 1.3.5 VCF Info

Phred - Phred quality score.

VCF Filter - if the VCF filter is passed (PASS).

Zygosity - most likely zygosity of the variant in this chromosomal position, computed from the observed variant frequency (column 8). Can be “FP/HET” (<15%), “HET” (15-75%), “HET/HOM” (75-85%), or “HOM” (>85%).

Alternate reads - the number of reads showing the alternative allele.

Total reads - the total number of reads.

Variant AF - the variant allele frequency.

Haplotype block ID - ID of the haplotype block.

Haplotype strand ID - ID of the haplotype strand.

## 1.4 Viewing Reports

Just-DNA-Seq makes a set of reports on a genome which contain data about longevity-related gene variants, known cancer risks and drug responses.

### 1.4.1 Opening Reports

1. Go to the **Reports** page.
2. In the **Longevity Combined** section, click **Download**.
3. When asked by your web browser what to do with the report file, open it in the browser. The report will be opened on a new tab.
4. Switch to the report tab in your browser.

### 1.4.2 Part 1: Longevity Significant Variations

This report contains gene variants which have significant influence on longevity. It has the following columns:

**+** - clicking this green button opens detailed information on each entry (row), and the button becomes red with a **-** sign. Clicking this **-** closes the details. Clicking **+** in the header opens the details for all rows and behaves in the same way (clicking **-** in the header closes all detail sections).

**RSID** - the reference sequence ID of the variant.

**Population** - population(s) on which the research was conducted, e.g. Greek, Ashkenazi Jewish etc., or multiple (for more details, open **+**).

**Gene** - gene the variant belongs to.

**Your Genotype** - which variants your genome contains. Note that in case of homozygosity two letters should be the same, and for heterozygosity they differ.

**Ref allele** - the reference allele.

**Alt allele** - the alternative allele.

**Zygosity** - hom (homozygosity) or het (heterozygosity).

**Weight** - weight of this variant (the degree of significance).

### 1.4.3 Part 2: Cancer Report

This report shows variants with known cancer risks and contains the following columns:

**+** - acting the same way in all reports (see above).

**#** - the number of an entry (row).

**Chrom** - chromosome the variant belongs to, e.g. **chr1**.

**Position** - position of the variant on the chromosome (number).

**Gene** - gene the variant belongs to, like in the previous report.

**RSID** - the reference sequence ID of the variant.

**cDNA Change** - change in the coding DNA by the variant.

**Zygosity** - hom or het (see above).

**Allele Frequency** - frequency of the allele.

**Phenotype Name** - description of condition(s) associated with the variant.

**Significance** - description of significance of this variant.

### 1.4.4 Part 3: Drugs Report

This report contains known issues of response to certain drugs associated with gene variants. It has the following columns:

**#** - the number of an entry (row).

**Variant/Haplotypes** - by rsID.

**Drug(s)** - name(s) of drug(s) response to which is affected by the variants.

**Phenotype Category**: Efficacy, Dosage, or Other.

**Significance** - yes or no.

**Sentence** - description of the case.

**Allele Of Frequency In Cases** - allele of the variant (one or more letters A, T, C or G) in cases involved.

**Allele Of Frequency In Controls** - allele of the variant in controls.

**Ratio Stat Type** - the ratio statistics type.

**Effect** - information on the effect.

Note: In some browsers the last one or two columns may be found beyond the visible area at 100% zoom level; in such cases try zooming out to 90%, 80% and so on, until everything is visible.