GETTING STARTED WITH SEAVE

This will get you started using Seave, a comprehensive variant filtration platform for clinical genomics, developed at the KCCG.

For a more detailed introduction, you can watch a demonstration video here.

Every case is different, so experiment with different selections and see how they change your results.

If you have any questions or feedback, please contact us at kccgeducation@garvan.org.au.

1. Login

Navigate to Seave at https://seave.bio/

Click "Log In" on the top right of the page and enter your username and password (login details for the case studies are in the downloadable instructions). There are also publically available datasets that don't require a login.

2. Go to the data

Click 'Take me to the data' in the middle of the page.

3. Select a database

Scroll to the bottom of the page to see the databases that you have access to.

Click on the relevant database.

4. Select a family

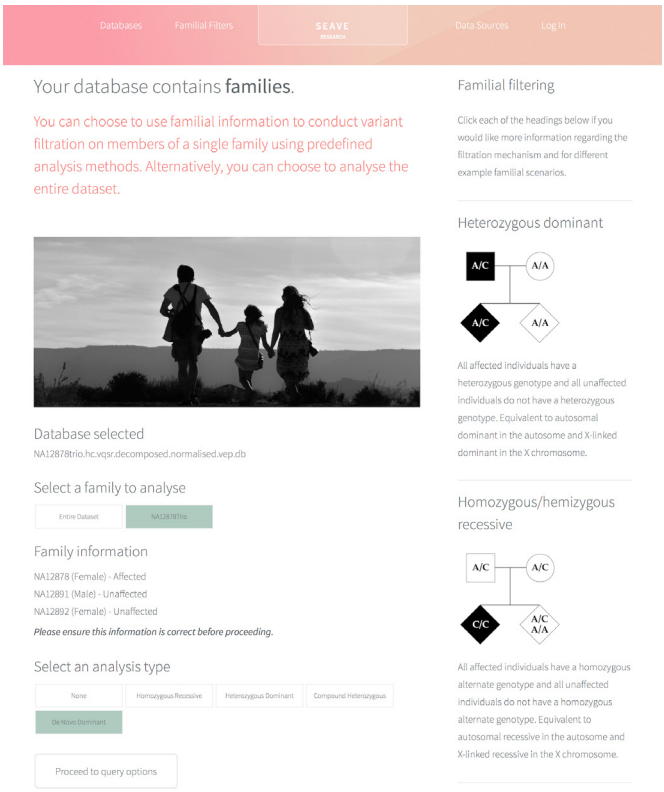
Databases can include multiple individuals, grouped into multiple families.

Select the relevant family and review their affected or unaffected status.

5. Select an analysis type

Select a relevant analysis type.

Multiple analysis types might fit the family's pedigree, meaning you may need to filter the data multiple times.



To start with, try homozygous recessive, compound heterozygous and de novo dominant. If multiple generations are affected, replace de novo dominant with heterozygous dominant.

You can read about inheritance patterns on the right of the page, or by clicking Familial Filters at the top.

Click 'Proceed to query options'.

6. Set up your query

These steps will determine which variants are filtered out, and which are kept for you to manually review.

Remember, you can experiment with these settings.

Genomic locations (optional)

Inclusion: You can enter a location to only return variants in that region, e.g. 'chr1:1-100000' for chromosome 1, bases 1 to 100000,or 'MT' for the mitochondrial genome.

Exclusion: This will exclude all variants in this region, which is a useful way of reducing incidental findings.

Connect with KCCG:

Visit us http://www.garvan.org.au/kccg
Follow our newsfeed on Twitter @kinghorngenomes
Email us at kccgenquiries@garvan.org.au

KINGHORN CENTRE FOR CLINICAL GENOMICS

GARVAN INSTITUTE OF MEDICAL RESEARCH

Gene lists (optional)

Gene lists are compiled by researchers, clinicians and commercial providers.

Select multiple gene lists by holding command (mac) or control (pc).

You can also enter your own custom gene list.

Impact

Leave 'Impact' at high & medium.

Some variants will have little to no impact, e.g if they code for the same amino acid. Others are likely to be harmful, e.g if they disrupt the transcription of a gene.

Seave follows the Ensembl variant impact prediction, which you can learn more about here.

CADD score

Set 'CADD score' at 10.

The CADD (combined annotation dependent depletion) score predicts a variant's functional significance. Read more about CADD scores here.

Prevalence

Set 'Prevalence' at 1%.

If a variant is found frequently in control databases, it is less less likely to be damaging.

Quality

Leave 'Minimum sequencing depth' at 0.

Set 'Minimum variant quality' at 200.

Leave 'Exclude failed variants' selected.

These filters allow you to exclude variants that were not well sequenced. Note: when looking at a family, these filters may exclude variants that were poorly sequenced in only one sample.

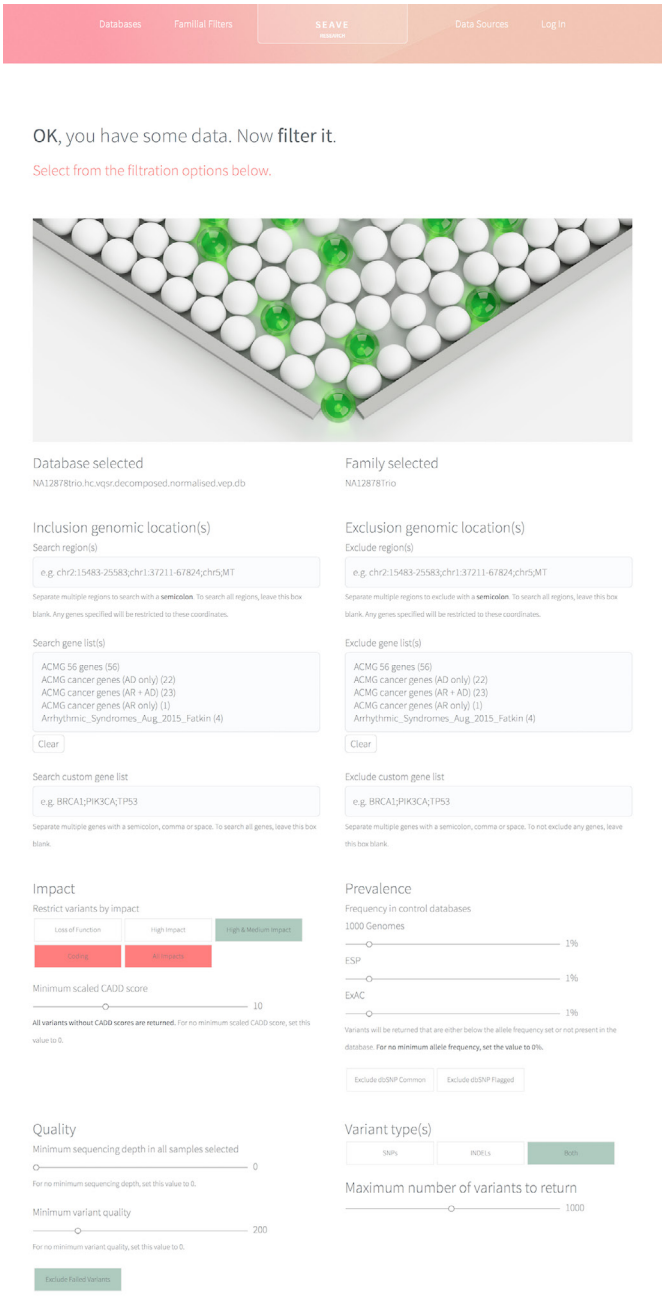
Variant type(s)

Leave 'Variant type(s)' at both to include single nucleotide polymorphisms (SNPs) and insertions and deletions (INDELs).

Maximum variants to return

Leave 'Maximum variants to return' at 200.

This gives you a manageable list of variants to review. The maximum may need to be higher if you chose no analysis type in step 5.



7. Get some results

On this screen you will see a table with the filtered variants, ready for you to investigate more closely There are many approaches to variant analysis, but to get started you could try the following steps.

Sort by the 'Impact Summary' column to see which variants 'light up' with red, meaning multiple tools predict they are damaging.

Scroll down to the 'Functional prediction and conservation scores' menu and **select** CADD Scores (look for scores over 10), SIFT (look for deleterious), and PolyPhen (look for possibly and probably damaging)

Scroll down to the 'Disease phenotypes' menu to see if any databases have disorders associated with the filtered variants. **Select** OMIM to start with, clicking on the OMIM number to get more information.

Scroll down to 'Allele frequencies' to see a variant's frequency in unaffected populations. **Select** EXAC and 1000 Genomes to start with.

The results table

Click on the headings to sort the table by any column.

Variant: location of the variant and details of the change

Quality: quality of the data showing the variant.

Gene: gene the variant is within or affects.

Type: type of variant (insertion, deletion, SNP).

Impact: impact of the variant.

MGRB: presence of the variant in the Medical Genome Reference Bank control cohort.

Impact summary: A summary of the evidence in key databases, hover over a square for more details.
Red = evidence for pathogenicity
Grey = no evidence/entry
White = evidence for non-pathogenicity.

Searching

Search by any of the terms in the table, including gene, genomic location and syndrome.

Additional columns

Scroll down to see further options and click on any of these to bring up additional columns in the table.

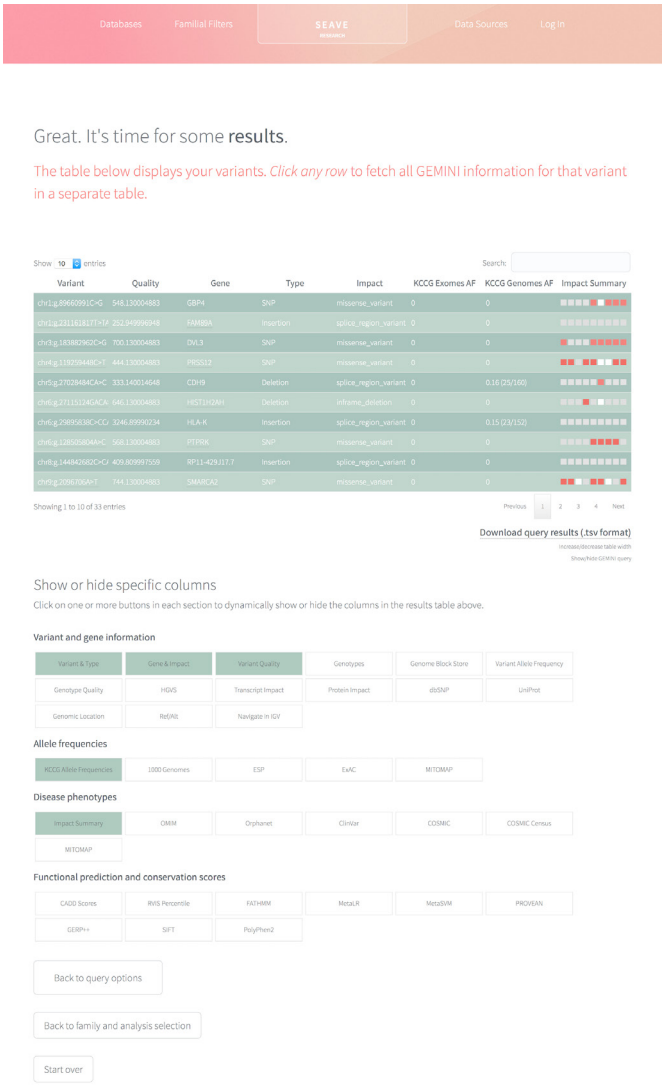
For example, selecting OMIM will bring up four columns showing the variant's number, title, status and associated disorders in that database.

Download the data

Click 'Download query results' to download all of the data for closer investigation. This TSV file can be opened in Excel.

8. Edit your search

Click 'Back to query options' to edit your selections if you can't find what you are looking for, or would like to see how different selections change your results.



Connect with KCCG:

Visit us http://www.garvan.org.au/kccg
Follow our newsfeed on Twitter @kinghorngenomes
Email us at kccgenquiries@garvan.org.au

KINGHORN CENTRE FOR CLINICAL GENOMICS

GARVAN INSTITUTE OF MEDICAL RESEARCH