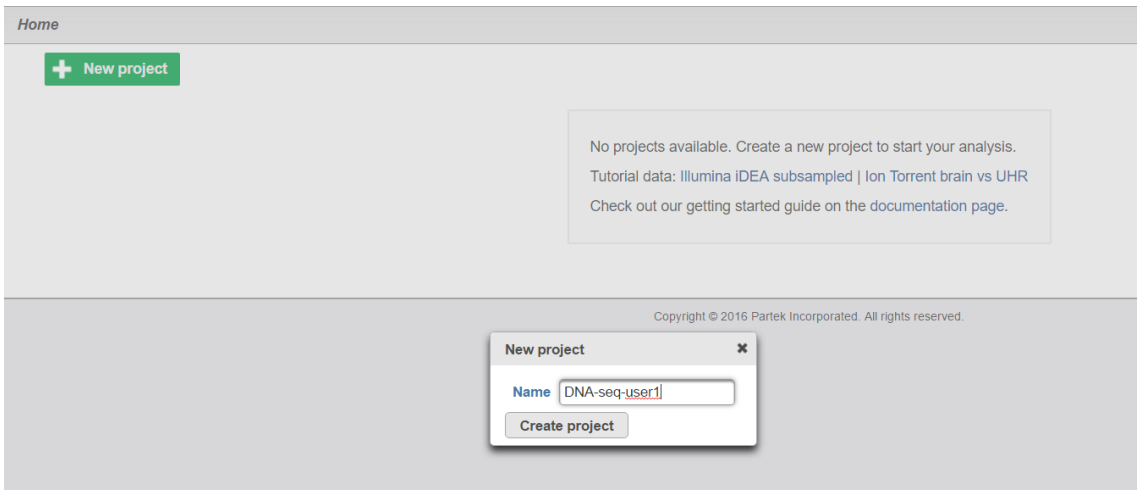
DNA-Seq Analysis in Partek® Flow®

HANDS-ON TRAINING

Partek®
turning data into discovery®

Login and Project Setup

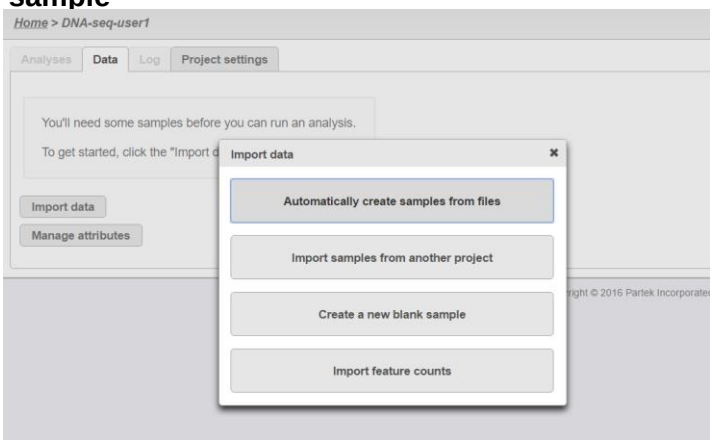
- Log in using the username and password given to you
- This will open to the Partek Flow homepage
- Click **New project** and enter project name: DNA-seq-[username]
- This will create a new project (push **Create project**)

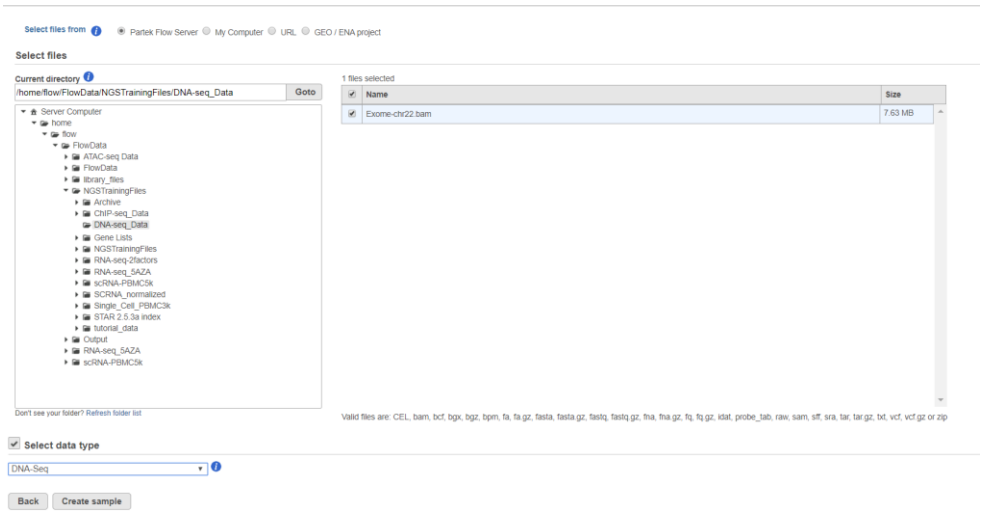


Notes:_____

Data Upload

- Creating a new project automatically opens up the **Data** tab
- To upload your data, click **Import Data > Automatically create samples from files**
- Browse to /home/flow/FlowData/DNA-seq (or equivalent)
- Select the available **.bam** file. Check mark **Select data type** and use **DNA-Seq**. Click **Create sample**

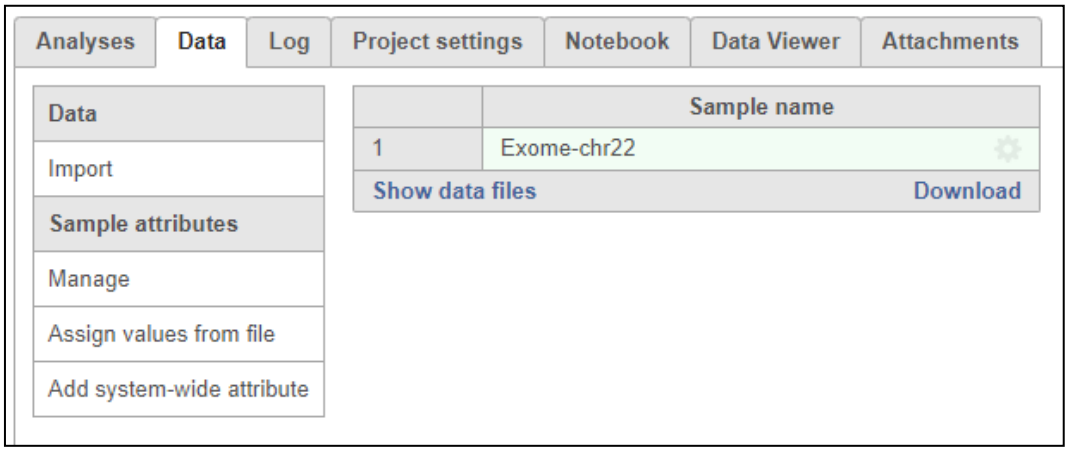


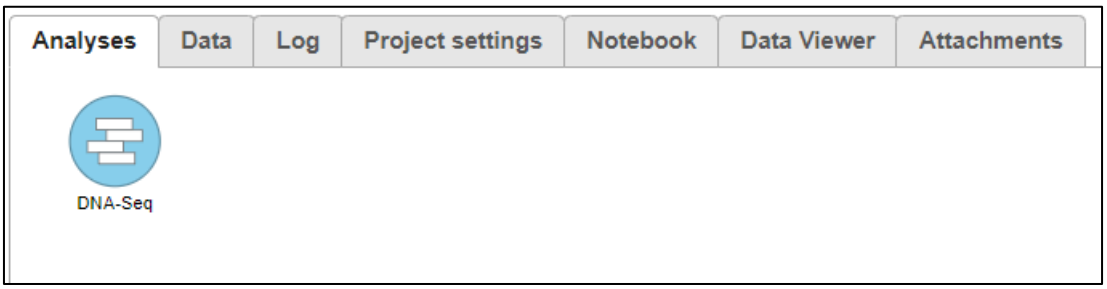


Notes:

Analyses Tab Overview

- Once Partek Flow processes the data, the *Aligned reads* node will appear on the canvas under the Analyses tab

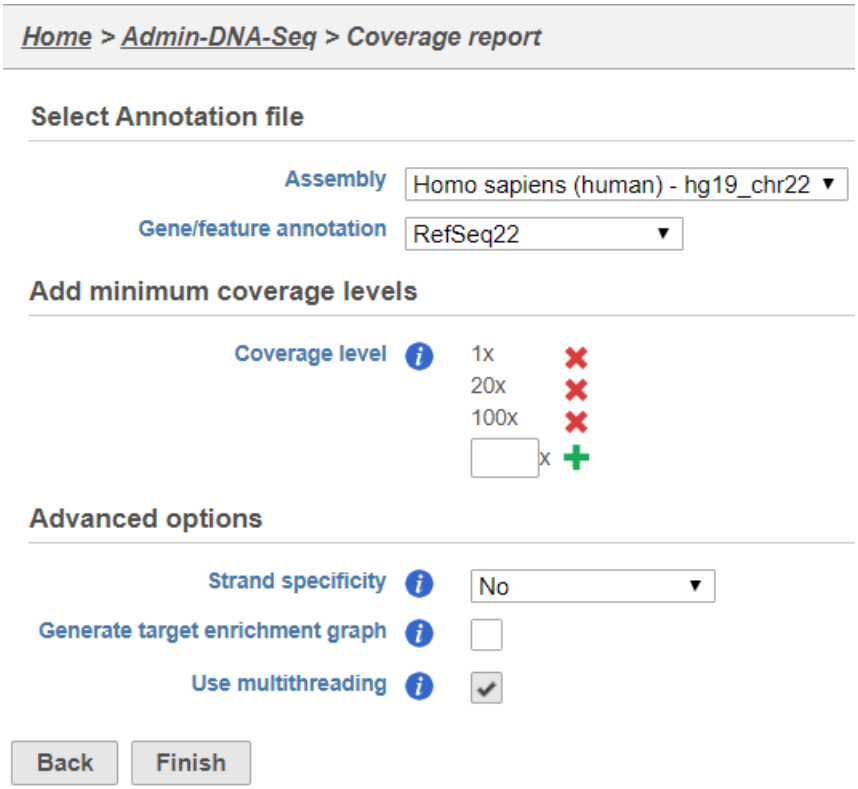




Notes:_____

Coverage Report Setup

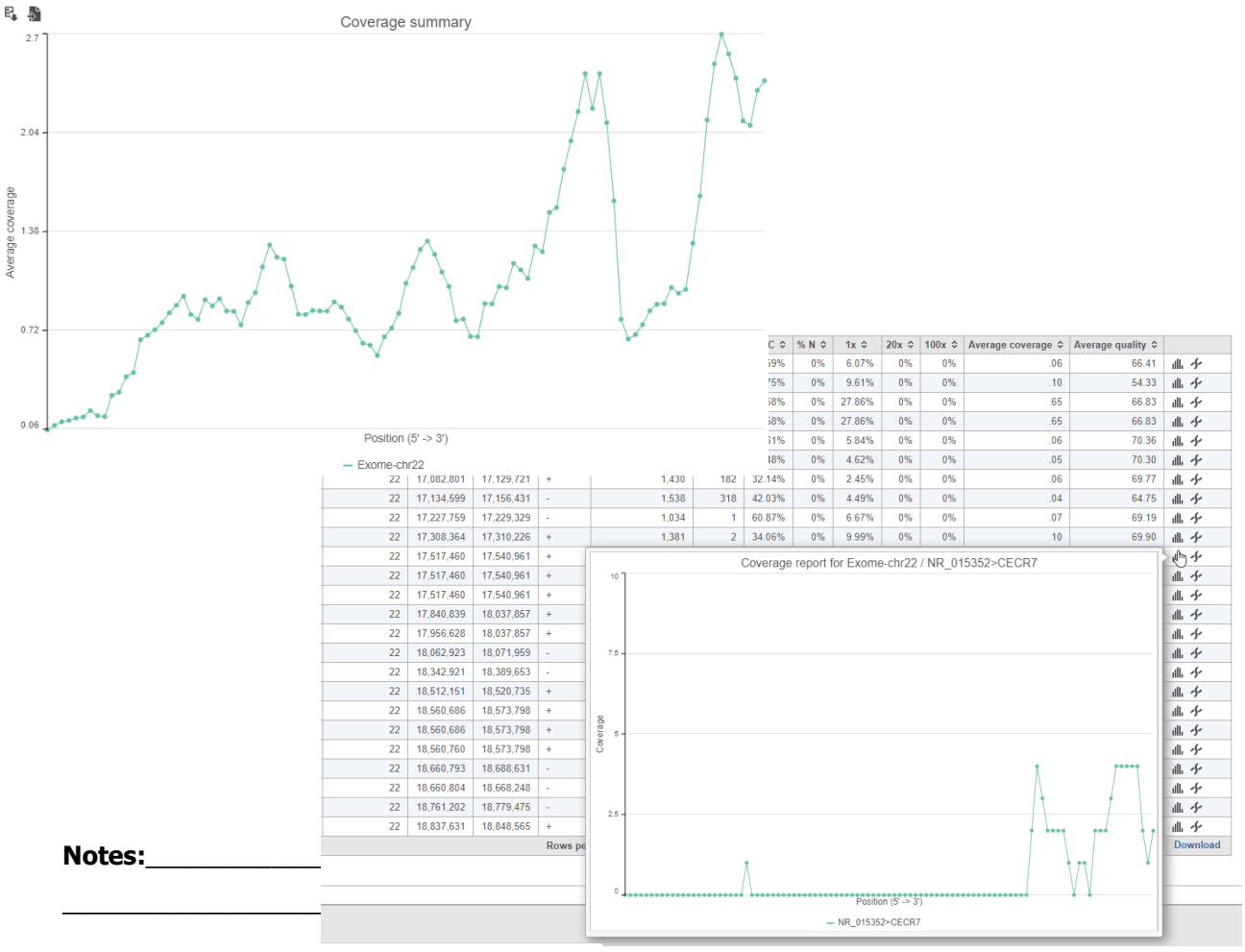
- *Coverage report* provides overview of sequencing coverage across specified regions (e.g. exome or targeted panel)
- Select the **Aligned reads** node and then **Coverage report**
- Set the *Assembly* to **Homo sapiens (human) – hg19_chr22** (or similar, depending on the training setup) and the *Gene/feature annotation* to **refseq** (or similar). The report will, consequently, contain information on coverage on RefSeq genes on the chromosome 22
- Check mark **Generate target enrichment graph** and push **Finish** to start



Notes:_____

Coverage Report Results

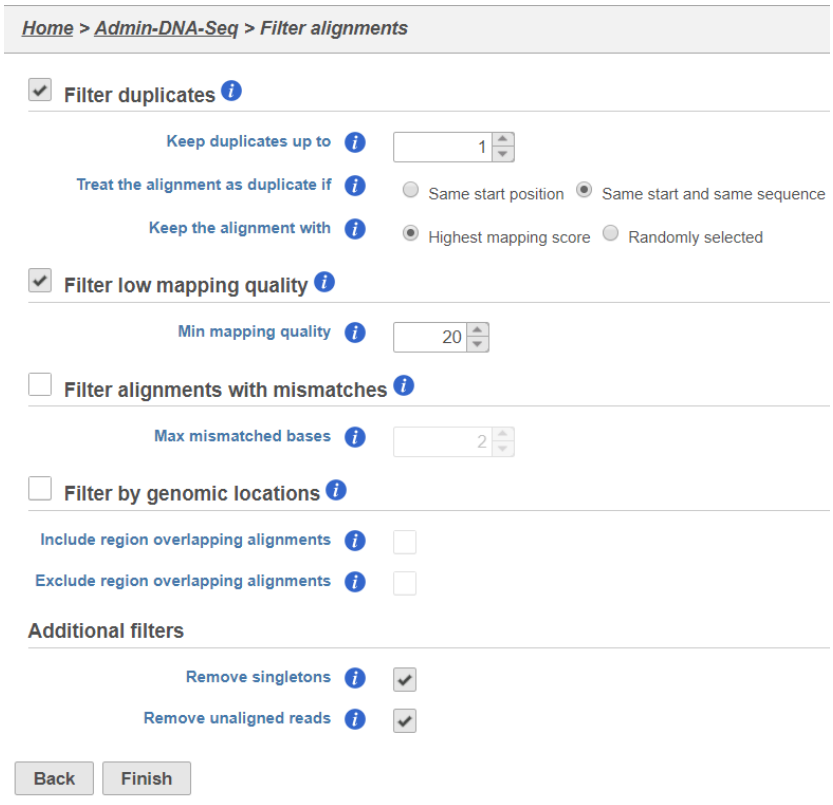
- When the task completes open the **Coverage report** task node
- The first page contains an overview of the coverage in the study samples
- To get details per targeted regions select the sample name in the table on the top
- For target enrichment plot, click / mouse over the on the **histogram** icon ( )



Notes:_____

Deduplication

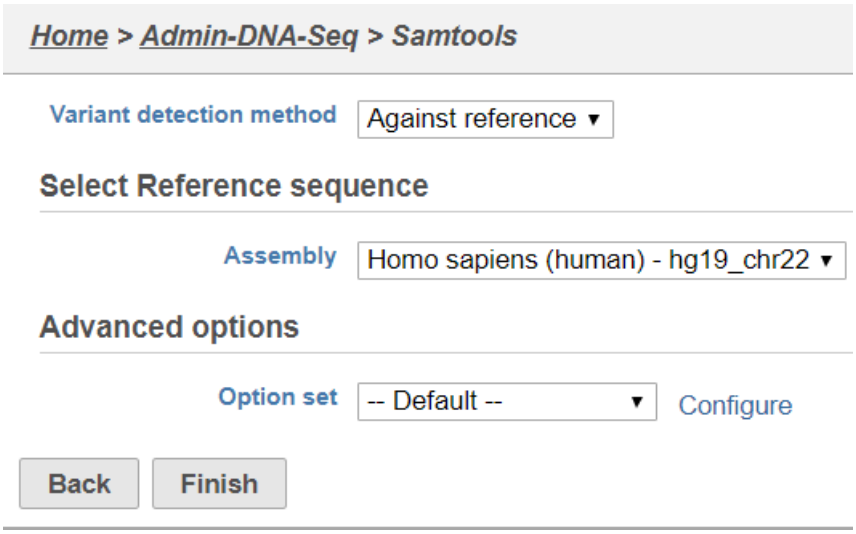
- Select the **Aligned reads** node and then **Filter alignments** (in the *Post-alignment tools*)
- To get rid of the PCR duplicates check the **Filter duplicates** option
- Choose duplicate as **Same start and same sequence** option
- Select **Filter low mapping quality**
- Push **Finish**
- The filter will create a *Filtered reads* data node



Notes:

Detect Variants Against Reference

- Select the **Filtered reads** data node > **Variant Callers** > **Samtools**
- Choose *Assembly* to **Homo sapiens (human) – hg19_chr22** as reference, click **Finish**
- The result is a *Variants* data node



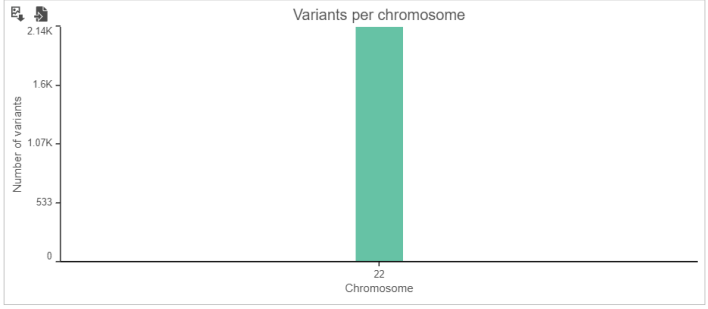
Notes:_____

Inspecting Variants

- Select the **Variants** data node > **Variant analysis** > **View variants**
- Use default settings and click **Finish**
- To inspect the result open the **View variants** task node
- Ti/Tv ratio: ratio of transitions vs. transversions in SNPs
  - Transition: mutation within the same type of nucleotide (C<->T; A<->G)
  - Transversion: mutation from pyrimidine to purine or vice versa

Sample name	Average variant quality	Variants	SNVs	Indels	Ti/Tv ratio	Het/Hom ratio	Estimated SNV Precision	Lower limit	Upper limit	Warnings
Exome-chr22	23.74	2,132	1,997	135	1.35773 (1150/847)	0.25 (428/1704)	39.70%	37.55%	41.85%	Precision estimation is unreliable.

Rows per page 25 (1 of 1) Download

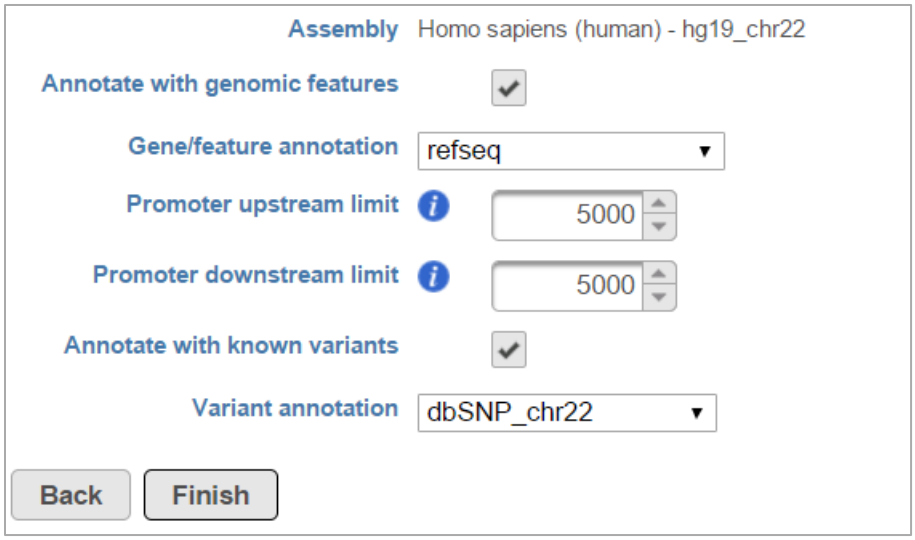


Task details

Notes:

Annotating Variants

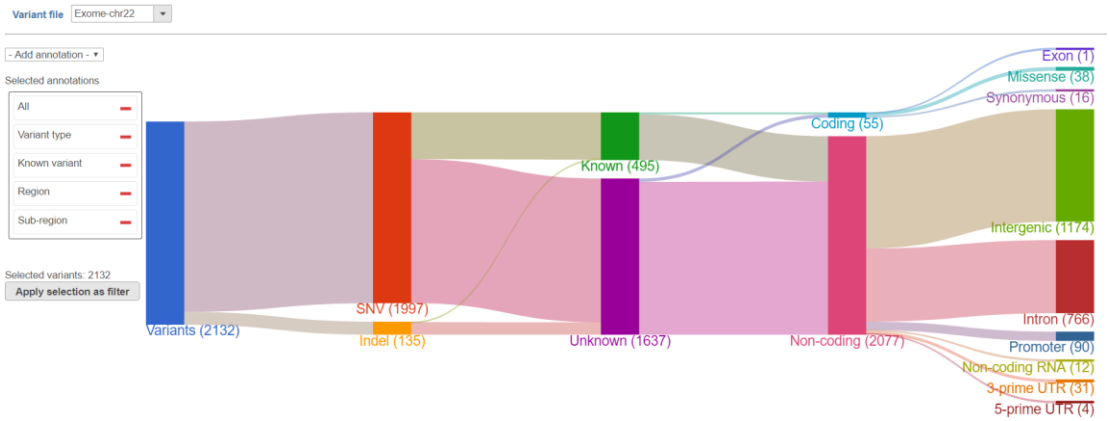
- Select the **Variants** data node > **Variant analysis** > **Annotate variants**
- Select **Annotate with genomic features** and point to **refseq**
- Select **Annotate with known variants** and point to **dbSNP**
- Click **Finish** to start. The output is *Annotated variants* node



Notes:

Variants Report

- Select the **Annotated variants** data node, choose **Task report** on the menu
- Sankey plot presenting different classes of variant based on the annotation
  - Each vertical block is one type of annotation
  - Annotation can be added or removed
  - Click on any colored section to apply filter
  - Click on the any blank to remove the filter
- The record in the bottom table will reflect the filter applied



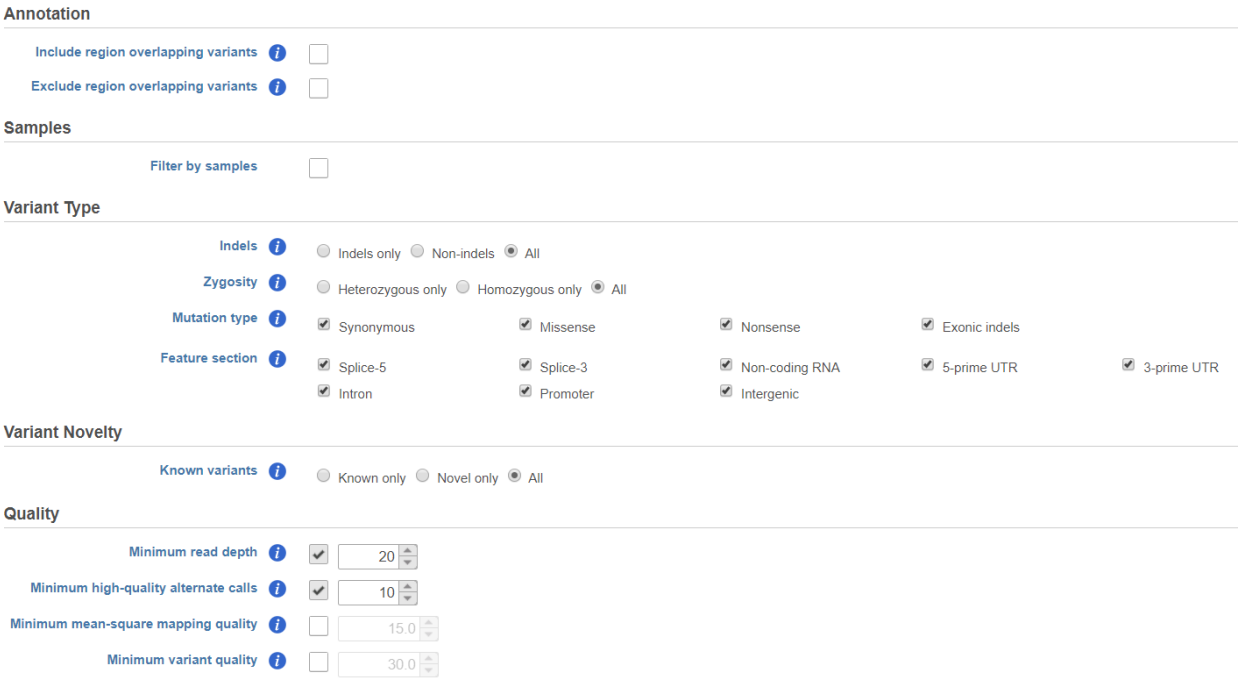
Optional columns

	View	Chr	Position	ID	Ref	Alts	VarQual	Type	Gene symbol	Transcript	Strand	Gene section	Nt change	AA change	Genotype
		Search...	Search...	Search...	Search...	Search...	Search...	Search...	Search...	Search...	Search...	Search...	Search...	Search...	Search...
1		22	16053317	.	G	C	14.57	.	.	.	.	.	.	.	CC
2		22	16058812	.	G	C	14.57	.	.	.	.	.	.	.	CC
3		22	16058891	.	C	T	6.51	.	.	.	.	.	.	.	TT
4		22	16062817	.	G	A	14.57	.	.	.	.	.	.	.	AA
5		22	16062906	.	A	T	9.00	.	.	.	.	.	.	.	TT
6		22	16063044	.	G	T	5.76	.	.	.	.	.	.	.	TT
7		22	16063077	.	C	G	5.76	.	.	.	.	.	.	.	GG
8		22	16064159	.	C	A	5.76	.	.	.	.	.	.	.	AA
9		22	16064166	.	T	A	24.43	.	.	.	.	.	.	.	AA
10		22	16069573	.	T	C	5.76	.	.	.	.	.	.	.	CC
11		22	16069589	.	C	T	5.76	.	.	.	.	.	.	.	TT
12		22	16077491	.	T	C	5.76	.	.	.	.	.	.	.	CC

Notes:

Filtering Based on Annotation

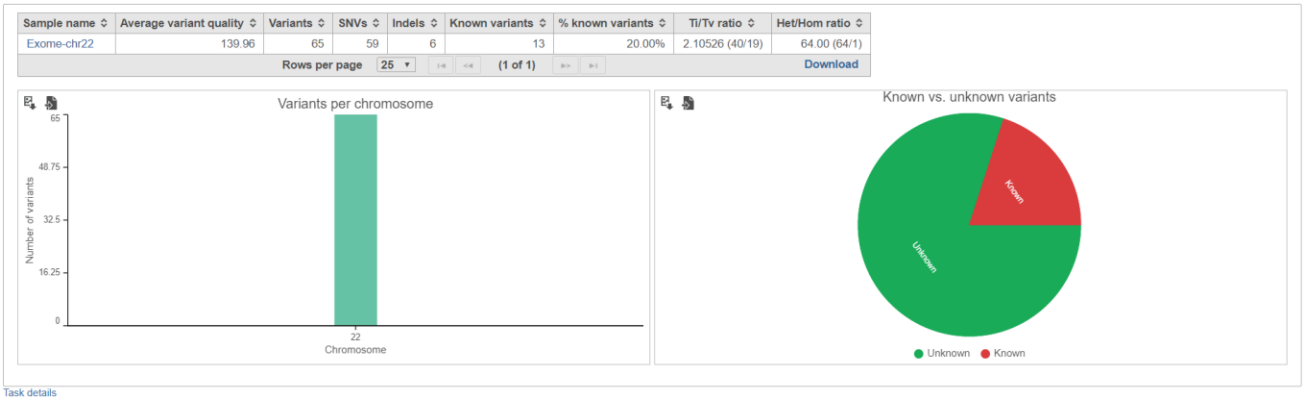
- Select the **Annotated variants** data node > **Variant analysis** > **Filter variants**
- Set *Minimum read depth* to **20**. The filter shows variants covered at least 20 ×.
- Set *Minimum alternate calls* to **10**. The filter shows variants with at least 10 supporting reads.
- Push **Finish**. The output is a *Variants* data node



Notes:_____

Inspecting Variants

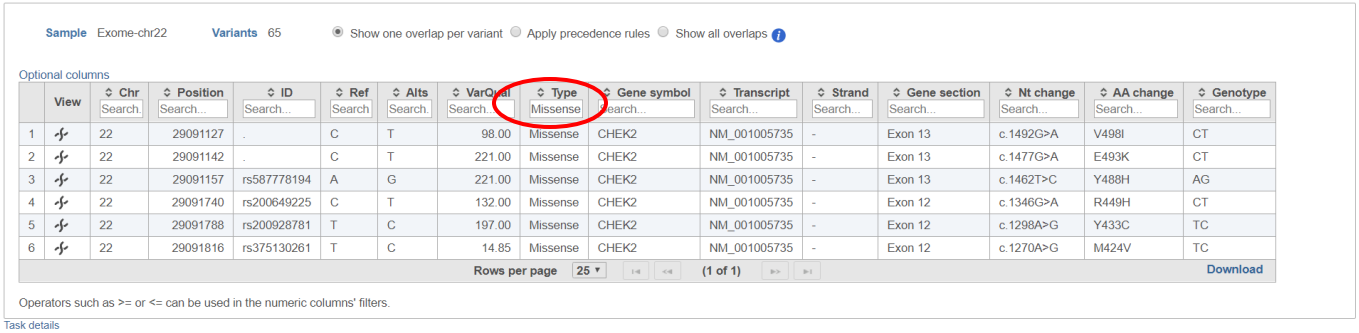
- Select the **Variants** data node > **Variant analysis** > **View variants**
- Use default settings and click **Finish**
- To inspect the result open the **View variants** task node
- To get detailed output click on the **sample name**



Notes:_____

Filtering the Results Table

- To select missense variants only, select the **Missense** colum in the *Variant occurences by type* chart
- When the table gets refreshed, all the entries in the *Type* column are set to *Missense*



Sample Exome-chr22 **Variants** 65 Show one overlap per variant | Apply precedence rules | Show all overlaps

Optional columns

	View	Chr	Position	ID	Ref	Alts	VarQual	Type	Gene symbol	Transcript	Strand	Gene section	Nt change	AA change	Genotype
		Search...	Search...	Search...	Search...	Search...	Search...	Missense	Search...	Search...	Search...	Search...	Search...	Search...	Search...
1		22	29091127	.	C	T	98.00	Missense	CHEK2	NM_001005735	-	Exon 13	c.1492G>A	V498I	CT
2		22	29091142	.	C	T	221.00	Missense	CHEK2	NM_001005735	-	Exon 13	c.1477G>A	E493K	CT
3		22	29091157	rs587778194	A	G	221.00	Missense	CHEK2	NM_001005735	-	Exon 13	c.1462T>C	Y488H	AG
4		22	29091740	rs200649225	C	T	132.00	Missense	CHEK2	NM_001005735	-	Exon 12	c.1346G>A	R449H	CT
5		22	29091788	rs200928781	T	C	197.00	Missense	CHEK2	NM_001005735	-	Exon 12	c.1298A>G	Y433C	TC
6		22	29091816	rs375130261	T	C	14.85	Missense	CHEK2	NM_001005735	-	Exon 12	c.1270A>G	M424V	TC

Rows per page 25 (1 of 1) Download

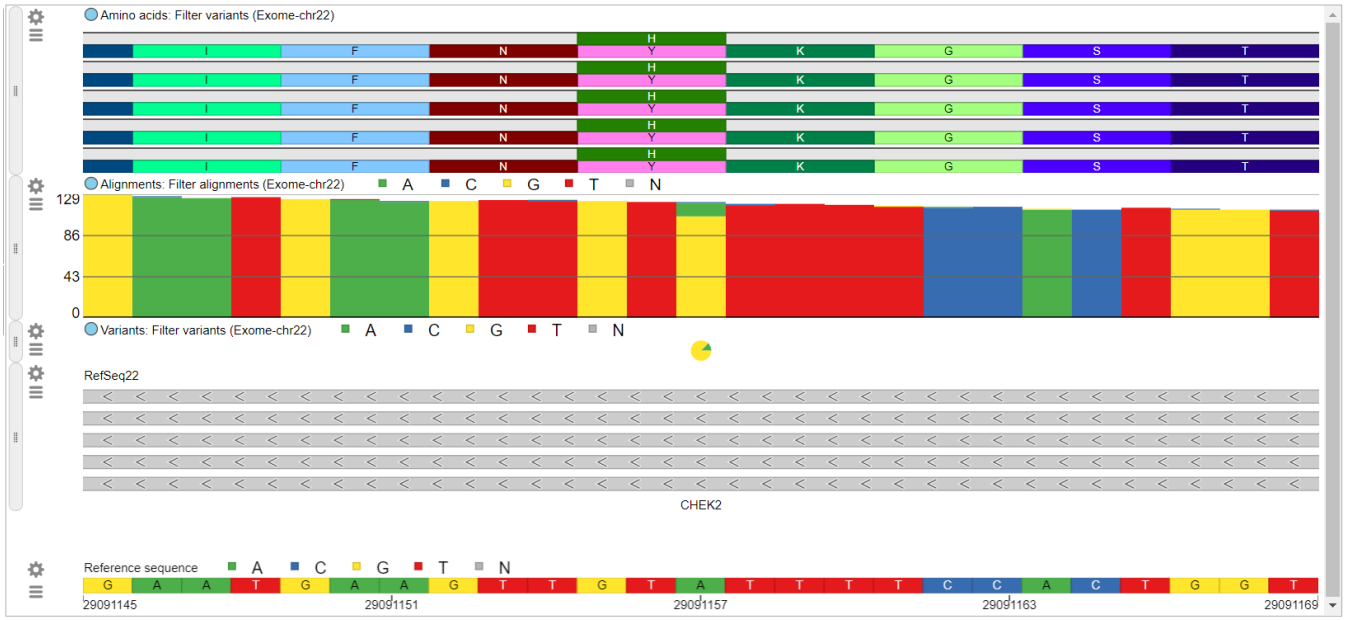
Operators such as >= or <= can be used in the numeric columns' filters.

Task details

Notes:

Exploring the Browser View

- Click on to browse in chromosome view
- Click on Select tracks to add or remove data tracks or annotation tracks
- On each track, click on button to configure the track
- Click on button to hide or pin track
- Drag and drop to change the order of the tracks



Notes:_____

Further Training

Self-learning

- Check out http://www.partek.com/resources-partek-flow for resources
- Recorded webinars available on Partek Incorporated's YouTube page

Regional Technical Support

- Email: support@partek.com
- Phone: +1-314-878-2329

Notes: