Annotate Variants

The *Annotate variants* task in Partek Flow provides a means to add information with regards to genomic features, such as transcript models, and existing variant databases to the variants contained in the projects. This information can be useful for filtering, interpreting, and prioritizing variants for downstream investigation. The *Annotate variants* task can be invoked from any *Variants* or *Annotated variants* data node, and the task will be added to and supplement any existing annotation in the underlying vcf files. Annotation information will also be visible in the View variants *Variant report* and the Summarize Cohort Mutations *Cohort mutation summary report*.

Annotate variants dialog

The task dialog for **Annotate variants** contains three sections: *Assembly*, *Annotate with genomic features*, and *Annotate with known variants* (Figure 1). If variant detection was performed in Partek Flow, the *Assembly* will be displayed as text in the section, and you do not have the option to change the reference. In the event that variant detection was performed outside of Partek Flow, you will need to select the appropriate Assembly utilized for variant detection in the drop-down list. Assemblies previously added to library files (see Library File Management) will be available for selection or *New assembly…* can be utilized to import the reference sequence from within the task.

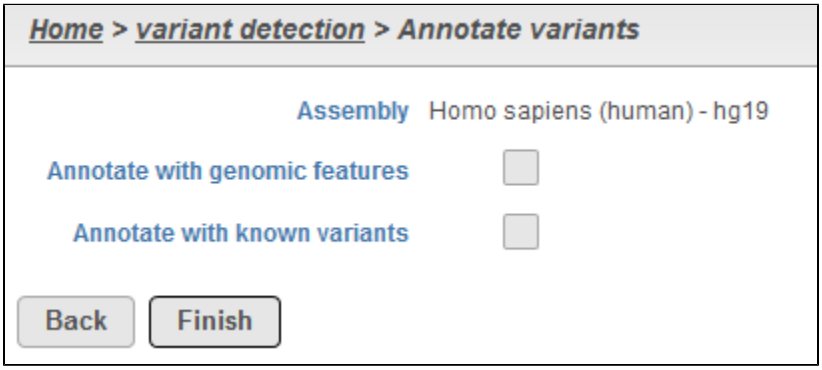

Figure 1. Components of the annotate variants dialog

Selecting *Annotate with genomic feature*s provides the means to add gene/feature information to the variants (Figure 2). This typically takes the form of overlaying a transcript model (such as RefSeq). Annotation models previously added to library files (see Library File Management) will be available for selection or *Add annotation model* in the drop-down list can be utilized to import an annotation model to library files within the task. *Promoter upstream limit* and *Promoter downstream limit* provides a means to set the number of bases flanking the transcription start site, and this region will considered the promoter of a feature.

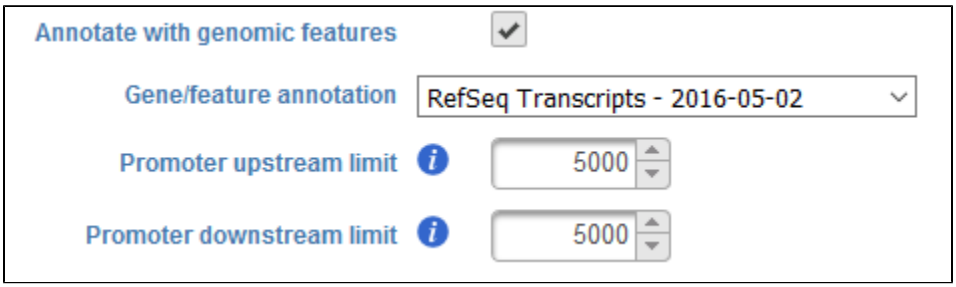

Figure 2. Configuration of Annotate with genomic features

Selecting *Annotate with known variants* will provide the ability to specify a *Variant annotation* database (Figure 3). Known variant databases in vcf format, such as dbSNP[1] and 1000 Genomes[2] for human variants, can be used in the task. Additional databases not provided for automated download in Partek® Flow®, such as the Catalogue of Somatic Mutations in Cancer (COSMIC)[3], can be obtained and employed by the user. Variant databases previously added to library files (see Library File Management) will be available for selection or *Add variant database* in the menu can be utilized to import the variant database to library files from within the task.

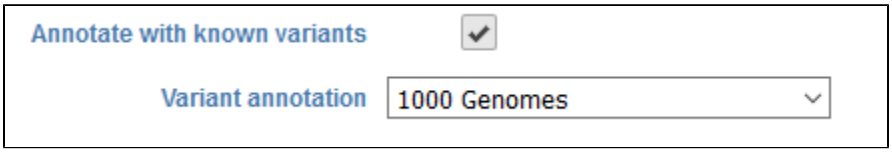

Figure 3. Configuration of Annotate with known variants

References

1. Sherry ST. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Research*. 2001;29(1):308-311. doi:10.1093/nar/29.1.308
2. Auton A, Abecasis GR, Altshuler DM, et al. A global reference for human genetic variation. *Nature*. 2015;526(7571):68-74. doi:10.1038/nature15393.
3. Forbes SA, Bhamra G, Bamford S, et al. The Catalogue of Somatic Mutations in Cancer (COSMIC). In: Haines JL, Korf BR, Morton CC, Seidman CE, Seidman JG, Smith DR, eds. *Current Protocols in Human Genetics*. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2008. http://doi.wiley.com/10.1002/0471142905.hg1011s57.

Additional Assistance

If you need additional assistance, please visit our support page to submit a help ticket or find phone numbers for regional support.