illumina®

BaseSpace VariantStudio™ v2.2 Software User Guide

FOR RESEARCH USE ONLY

ILLUMINA PROPRIETARY
Part # 15047059 Rev. A
June 2014

Read Before Using this Product

This Product, and its use and disposition, is subject to the following terms and conditions. If Purchaser does not agree to these terms and conditions then Purchaser is not authorized by Illumina to use this Product and Purchaser must not use this Product.

1 **Definitions**. "**Application Specific IP**" means Illumina owned or controlled intellectual property rights that pertain to this Product (and use thereof) only with regard to specific field(s) or specific application(s). Application Specific IP excludes all Illumina owned or controlled intellectual property that cover aspects or features of this Product (or use thereof) that are common to this Product in all possible applications and all possible fields of use (the "**Core IP**"). Application Specific IP and Core IP are separate, non-overlapping, subsets of all Illumina owned or controlled intellectual property. By way of non-limiting example, Illumina intellectual property rights for specific diagnostic methods, for specific forensic methods, or for specific nucleic acid biomarkers, sequences, or combinations of biomarkers or sequences are examples of Application Specific IP. "**Consumable(s)**" means Illumina branded reagents and consumable items that are intended by Illumina for use with, and are to be consumed through the use of, Hardware. "**Documentation**" means Illumina's user manual for this Product, including without limitation, package inserts, and any other documentation that accompany this Product or that are referenced by the Product or in the packaging for the Product in effect on the date of shipment from Illumina. Documentation includes this document. "**Hardware**" means Illumina branded instruments, accessories or peripherals. "**Illumina**" means Illumina, Inc. or an Illumina affiliate, as applicable. "**Product**" means the product that this document accompanies (e.g., Hardware, Consumables, or Software). "**Purchaser**" is the person or entity that rightfully and legally acquires this Product from Illumina or an Illumina authorized dealer. "**Software**" means Illumina branded software (e.g., Hardware operating software, data analysis software). All Software is licensed and not sold and may be subject to additional terms found in the Software's end user license agreement. "**Specifications**" means Illumina's written specifications for this Product in effect on the date that the Product ships from Illumina.

2 **Research Use Only Rights**. Subject to these terms and conditions and unless otherwise agreed upon in writing by an officer of Illumina, Purchaser is granted only a non-exclusive, non-transferable, personal, non-sublicensable right under Illumina's Core IP, in existence on the date that this Product ships from Illumina, solely to use this Product in Purchaser's facility for Purchaser's internal research purposes (which includes research services provided to third parties) and solely in accordance with this Product's Documentation, **but specifically excluding any use that** (a) would require rights or a license from Illumina to Application Specific IP, (b) is a re-use of a previously used Consumable, (c) is the disassembling, reverse-engineering, reverse-compiling, or reverse-assembling of this Product, (d) is the separation, extraction, or isolation of components of this Product or other unauthorized analysis of this Product, (e) gains access to or determines the methods of operation of this Product, (f) is the use of non-Illumina reagent/consumables with Illumina's Hardware (does not apply if the Specifications or Documentation state otherwise), or (g) is the transfer to a third-party of, or sub-licensing of, Software or any third-party software. All Software, whether provided separately, installed on, or embedded in a Product, is licensed to Purchaser and not sold. Except as expressly stated in this Section, no right or license under any of Illumina's intellectual property rights is or are granted expressly, by implication, or by estoppel.

Purchaser is solely responsible for determining whether Purchaser has all intellectual property rights that are necessary for Purchaser's intended uses of this Product, including without limitation, any rights from third parties or rights to Application Specific IP. Illumina makes no guarantee or warranty that purchaser's specific intended uses will not infringe the intellectual property rights of a third party or Application Specific IP.

3 **Regulatory**. This Product has not been approved, cleared, or licensed by the United States Food and Drug Administration or any other regulatory entity whether foreign or domestic for any specific intended use, whether research, commercial, diagnostic, or otherwise. This Product is labeled For Research Use Only. Purchaser must ensure it has any regulatory approvals that are necessary for Purchaser's intended uses of this Product.

4 **Unauthorized Uses**. Purchaser agrees: (a) to use each Consumable only one time, and (b) to use only Illumina consumables/reagents with Illumina Hardware. The limitations in (a)-(b) do not apply if the Documentation or Specifications for this Product state otherwise. Purchaser agrees not to, nor authorize any third party to, engage in any of the following activities: (i) disassemble, reverse-engineer, reverse-compile, or reverse-assemble the Product, (ii) separate, extract, or isolate components of this Product or subject this Product or components thereof to any analysis not expressly authorized in this Product's Documentation, (iii) gain access to or attempt to determine the methods of operation of this Product, or (iv) transfer to a third-party, or grant a sublicense, to any Software or any third-party software. Purchaser further agrees that the contents of and methods of operation of this Product are proprietary to Illumina and this Product contains or embodies trade secrets of Illumina. The conditions and restrictions found in these terms and conditions are bargained for conditions of sale and therefore control the sale of and use of this Product by Purchaser.

5 **Limited Liability. TO THE EXTENT PERMITTED BY LAW, IN NO EVENT SHALL ILLUMINA OR ITS SUPPLIERS BE LIABLE TO PURCHASER OR ANY THIRD PARTY FOR COSTS OF PROCUREMENT OF SUBSTITUTE PRODUCTS OR SERVICES, LOST PROFITS, DATA OR BUSINESS, OR FOR ANY INDIRECT, SPECIAL, INCIDENTAL, EXEMPLARY, CONSEQUENTIAL, OR PUNITIVE DAMAGES OF ANY KIND ARISING OUT OF OR IN CONNECTION WITH, WITHOUT LIMITATION, THE SALE OF THIS PRODUCT, ITS USE, ILLUMINA'S PERFORMANCE HEREUNDER OR ANY OF THESE TERMS AND CONDITIONS, HOWEVER ARISING OR CAUSED AND ON ANY THEORY OF LIABILITY (WHETHER IN CONTRACT, TORT (INCLUDING NEGLIGENCE), STRICT LIABILITY OR OTHERWISE).**

6 **ILLUMINA'S TOTAL AND CUMULATIVE LIABILITY TO PURCHASER OR ANY THIRD PARTY ARISING OUT OF OR IN CONNECTION WITH THESE TERMS AND CONDITIONS, INCLUDING WITHOUT LIMITATION, THIS PRODUCT (INCLUDING USE THEREOF) AND ILLUMINA'S PERFORMANCE HEREUNDER, WHETHER IN CONTRACT, TORT (INCLUDING NEGLIGENCE), STRICT LIABILITY OR OTHERWISE, SHALL IN NO EVENT EXCEED THE AMOUNT PAID TO ILLUMINA FOR THIS PRODUCT.**

7 **Limitations on Illumina Provided Warranties. TO THE EXTENT PERMITTED BY LAW AND SUBJECT TO THE EXPRESS PRODUCT WARRANTY MADE HEREIN ILLUMINA MAKES NO (AND EXPRESSLY DISCLAIMS ALL) WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THIS PRODUCT, INCLUDING WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, NONINFRINGEMENT, OR ARISING FROM COURSE OF PERFORMANCE, DEALING, USAGE OR TRADE. WITHOUT LIMITING THE GENERALITY OF THE FOREGOING, ILLUMINA MAKES NO CLAIM, REPRESENTATION, OR WARRANTY OF ANY KIND AS TO THE UTILITY OF THIS PRODUCT FOR PURCHASER'S INTENDED USES.**

8 **Product Warranty**. All warranties are personal to the Purchaser and may not be transferred or assigned to a third-party, including an affiliate of Purchaser. All warranties are facility specific and do not transfer if the Product is moved to another facility of Purchaser, unless Illumina conducts such move.

a **Warranty for Consumables**. Illumina warrants that Consumables, other than custom Consumables, will conform to their Specifications until the later of (i) 3 months from the date of shipment from Illumina, and (ii) any expiration date or the end of the shelf-life pre-printed on such Consumable by Illumina, but in no event later than 12 months from the date of shipment. With respect to custom Consumables (i.e., Consumables made to specifications or designs made by Purchaser or provided to Illumina by, or on behalf of, Purchaser), Illumina only warrants that the custom Consumables will be made and tested in accordance with Illumina's standard manufacturing and quality control processes. Illumina makes no warranty that custom Consumables will work as intended by Purchaser or for Purchaser's intended uses.

b **Warranty for Hardware**. Illumina warrants that Hardware, other than Upgraded Components, will conform to its Specifications for a period of 12 months after its shipment date from Illumina unless the Hardware includes Illumina provided installation in which case the warranty period begins on the date of installation or 30 days after the date it was delivered, whichever occurs first ("Base Hardware Warranty"). "Upgraded Components" means Illumina provided components, modifications, or enhancements to Hardware that was previously acquired by Purchaser. Illumina warrants that Upgraded Components will conform to their Specifications for a period of 90 days from the date the Upgraded Components are installed. Upgraded Components do not extend the warranty for the Hardware unless the upgrade was conducted by Illumina at Illumina's facilities in which case the upgraded Hardware shipped to Purchaser comes with a Base Hardware Warranty.

c **Exclusions from Warranty Coverage**. The foregoing warranties do not apply to the extent a non-conformance is due to (i) abuse, misuse, neglect, negligence, accident, improper storage, or use contrary to the Documentation or Specifications, (ii) improper handling, installation, maintenance, or repair (other than if performed by Illumina's personnel), (iii) unauthorized alterations, (iv) Force Majeure events, or (v) use with a third party's good not provided

by Illumina (unless the Product's Documentation or Specifications expressly state such third party's good is for use with the Product).

d **Procedure for Warranty Coverage**. In order to be eligible for repair or replacement under this warranty Purchaser must (i) promptly contact Illumina's support department to report the non-conformance, (ii) cooperate with Illumina in confirming or diagnosing the non-conformance, and (iii) return this Product, transportation charges prepaid to Illumina following Illumina's instructions or, if agreed by Illumina and Purchaser, grant Illumina's authorized repair personnel access to this Product in order to confirm the non-conformance and make repairs.

e **Sole Remedy under Warranty**. Illumina will, at its option, repair or replace non-conforming Product that it confirms is covered by this warranty. Repaired or replaced Consumables come with a 30-day warranty. Hardware may be repaired or replaced with functionally equivalent, reconditioned, or new Hardware or components (if only a component of Hardware is non-conforming). If the Hardware is replaced in its entirety, the warranty period for the replacement is 90 days from the date of shipment or the remaining period on the original Hardware warranty, whichever is shorter. If only a component is being repaired or replaced, the warranty period for such component is 90 days from the date of shipment or the remaining period on the original Hardware warranty, whichever ends later. The preceding states Purchaser's sole remedy and Illumina's sole obligations under the warranty provided hereunder.

f **Third-Party Goods and Warranty**. Illumina has no warranty obligations with respect to any goods originating from a third party and supplied to Purchaser hereunder. Third-party goods are those that are labeled or branded with a third-party's name. The warranty for third-party goods, if any, is provided by the original manufacturer. Upon written request Illumina will attempt to pass through any such warranty to Purchaser.

9 **Indemnification**.

a **Infringement Indemnification by Illumina**. Subject to these terms and conditions, including without limitation, the Exclusions to Illumina's Indemnification Obligations (Section 9(b) below), the Conditions to Indemnification Obligations (Section 9(d) below), Illumina shall (i) defend, indemnify and hold harmless Purchaser against any third-party claim or action alleging that this Product when used for research use purposes, in accordance with these terms and conditions, and in accordance with this Product's Documentation and Specifications infringes the valid and enforceable intellectual property rights of a third party, and (ii) pay all settlements entered into, and all final judgments and costs (including reasonable attorneys' fees) awarded against Purchaser in connection with such infringement claim. If this Product or any part thereof, becomes, or in Illumina's opinion may become, the subject of an infringement claim, Illumina shall have the right, at its option, to (A) procure for Purchaser the right to continue using this Product, (B) modify or replace this Product with a substantially equivalent non-infringing substitute, or (C) require the return of this Product and terminate the rights, license, and any other permissions provided to Purchaser with respect this Product and refund to Purchaser the depreciated value (as shown in Purchaser's official records) of the returned Product at the time of such return; provided that, no refund will be given for used-up or expired Consumables. This Section states the entire liability of Illumina for any infringement of third party intellectual property rights.

b **Exclusions to Illumina Indemnification Obligations**. Illumina has no obligation to defend, indemnify or hold harmless Purchaser for any Illumina Infringement Claim to the extent such infringement arises from: (i) the use of this Product in any manner or for any purpose outside the scope of research use purposes, (ii) the use of this Product in any manner not in accordance with its Specifications, its Documentation, the rights expressly granted to Purchaser hereunder, or any breach by Purchaser of these terms and conditions, (iii) the use of this Product in combination with any other products, materials, or services not supplied by Illumina, (iv) the use of this Product to perform any assay or other process not supplied by Illumina, or (v) Illumina's compliance with specifications or instructions for this Product furnished by, or on behalf of, Purchaser (each of (i) – (v), is referred to as an "Excluded Claim").

c **Indemnification by Purchaser**. Purchaser shall defend, indemnify and hold harmless Illumina, its affiliates, their non-affiliate collaborators and development partners that contributed to the development of this Product, and their respective officers, directors, representatives and employees against any claims, liabilities, damages, fines, penalties, causes of action, and losses of any and every kind, including without limitation, personal injury or death claims, and infringement of a third party's intellectual property rights, resulting from, relating to, or arising out of (i) Purchaser's breach of any of these terms and conditions, (ii) Purchaser's use of this Product outside of the scope of research use purposes, (iii) any use of this Product not in accordance with this Product's Specifications or Documentation, or (iv) any Excluded Claim.

d **Conditions to Indemnification Obligations**. The parties' indemnification obligations are conditioned upon the party seeking indemnification (i) promptly notifying the other party in writing of such claim or action, (ii) giving the other party exclusive control and authority over the defense and settlement of such claim or action, (iii) not admitting infringement of any intellectual property right without prior written consent of the other party, (iv) not entering into any settlement or compromise of any such claim or action without the other party's prior written consent, and (v) providing reasonable assistance to the other party in the defense of the claim or action; provided that, the party reimburses the indemnified party for its reasonable out-of-pocket expenses incurred in providing such assistance.

e **Third-Party Goods and Indemnification**. Illumina has no indemnification obligations with respect to any goods originating from a third party and supplied to Purchaser. Third-party goods are those that are labeled or branded with a third-party's name. Purchaser's indemnification rights, if any, with respect to third party goods shall be pursuant to the original manufacturer's or licensor's indemnity. Upon written request Illumina will attempt to pass through such indemnity, if any, to Purchaser.

Revision History

Part #	Revision	Date	Description of Change
15047059	A	June 2014	Initial release.

Table of Contents

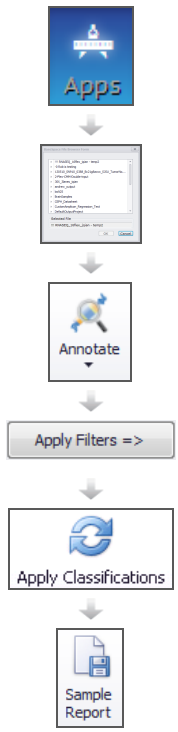
Getting Started

Chapter 1

Introduction

The BaseSpace VariantStudio™ software is an executable that gets installed on your desktop computer when you launch the BaseSpace VariantStudio App. BaseSpace VariantStudio imports SNPs and indels from your projects in BaseSpace or BaseSpace Onsite. After import, BaseSpace VariantStudio provides commands to annotate variants, filter results using filtering options, and export data to a report.

Figure 1 BaseSpace VariantStudio Workflow



Launch the BaseSpace VariantStudio App and associate to a project.

Import SNPs and indels from BaseSpace or BaseSpace Onsite. Import one sample or multiple samples to a single project.

Annotate variants for the current sample. An internet connection is required to annotate variants.

Filter data based on any combination of filtering options. Save filter combinations for use in other projects.

Classify variants according to their biological impact.

Generate a sample report using a customized report template. Export data and filtering history to text files. Export data to a histogram or pie chart.

NOTE
When you have already installed the BaseSpace VariantStudio App, it is best to launch BaseSpace VariantStudio from the desktop instead of going through BaseSpace or BaseSpace Onsite.

System Requirements

Installing the BaseSpace VariantStudio software requires the following system specifications:

- 64-bit Windows OS (Windows 7, or later)
- 2 GB RAM minimum; 4 GB RAM recommended
- 25 MB hard drive space for installation
- Internet connection required

NOTE
An internet connection is required for annotating variants. After variants have been annotated and saved in a project, an internet connection is no longer required.

Installation

To install BaseSpace VariantStudio, launch the BaseSpace VariantStudio App from your account in BaseSpace or BaseSpace Onsite. For instructions, see the *BaseSpace User Guide* or *BaseSpace Onsite User Guide*.

NOTE
Installation of VariantStudio v2.2, or later, does not overwrite the version previously installed on your computer, allowing side-by-side installation of different software versions.

VCF Input Requirements

BaseSpace VariantStudio imports SNPs and indels reported in VCF v4.0, or later, file formats. If analysis software other than Illumina analysis software is used to generate data, the VCF file might not contain the required columns.

VCF Column	Required Value
CHROM	The chromosome number. Values are #, c#, or chr#, where # is the chromosome number, as in 1–22, or name, as in X or Y, or M for mitochondrial.
POS	The position of the variant. Values are numeric with the first base having position 1 (1-based).
ID	The ID is the rs number for the SNP as recorded in dbSNP.txt. A value must be present. If a dbSNP entry does not exist, a missing value marker '.' is an acceptable value. Although the ID column and valid values are required, the values are not imported. The software applies dbSNP annotations with the Annotate command.
REF	The reference allele.
ALT	The alternate allele.
QUAL	The quality score assigned by the variant caller. A value of '.' is acceptable, and is reported as a 0.
INFO	Recognized fields are VF (alt variant freq), DP (read depth), AD (allelic depth), SOMATIC, and '.' (none). • VF—Represented in the Alt Variant Freq column in the Variants table. • DP—Represented in the Read Depth column in the Variants table. • DPI—Represented in the Read Depth column in the Variants table for insertion and deletion events called by the Illumina Isaac Alignment and Variant Calling workflow. • AD—Represented in the Alt Read Depth and Allelic Depth columns in the Variants table. • SOMATIC—Represented in the Genotype column in the Variants table. This value applies only to somatic variants.

VCF Column	Required Value
FORMAT	A list of fields that define values in the Sample column. Possible values are VF (alt variant frequency), DP (read depth), AD (allelic depth), GT (genotype), and '.' (none). • VF—Represented in the Alt Variant Freq column in the Variants table. • DP—Represented in the Read Depth column in the Variants table. • DPI—Represented in the Read Depth column in the Variants table for insertion and deletion events called by the Illumina Isaac Alignment and Variant Calling workflow. • AD—Represented in the Alt Read Depth and Allelic Depth columns in the Variants table. **Genotype Values:** (Not present in somatic VCF files.) • Acceptable GT values are 0/0, 0/1, and 1/1. Non-numeric GT values, or './.' as in a no-call, are not imported. • Hemizygous alt GT values, '1', are accepted. Hemizygous reference calls, '0', are not imported. • If FORMAT and Sample are not empty, then a GT value is required. • IF FORMAT and Sample are empty, the software assumes that GT is heterozygous, 0/1.

Genome VCF Files

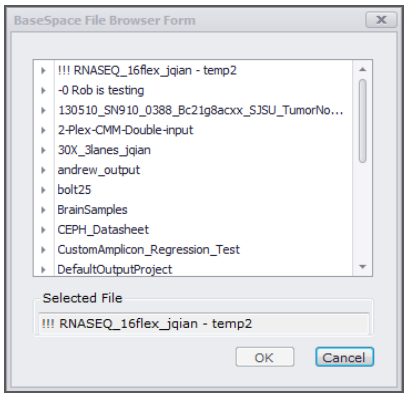
Importing genome VCF (gVCF) files is supported as of BaseSpace VariantStudio v2.1 for targeted enrichment data.

Using gVCF is not recommended for whole genomes without pre-processing with gVCF tools. Alternatively, you can load only exonic regions, or only regions from a gene list or BED file without the need for pre-processing. For more information, see *VCF Import Options* on page 10.

BaseSpace VariantStudio Software Interface

When the BaseSpace VariantStudio software launches, the interface opens with BaseSpace File Browser. This file browser allows you to import a BaseSpace VCF file to VariantStudio.

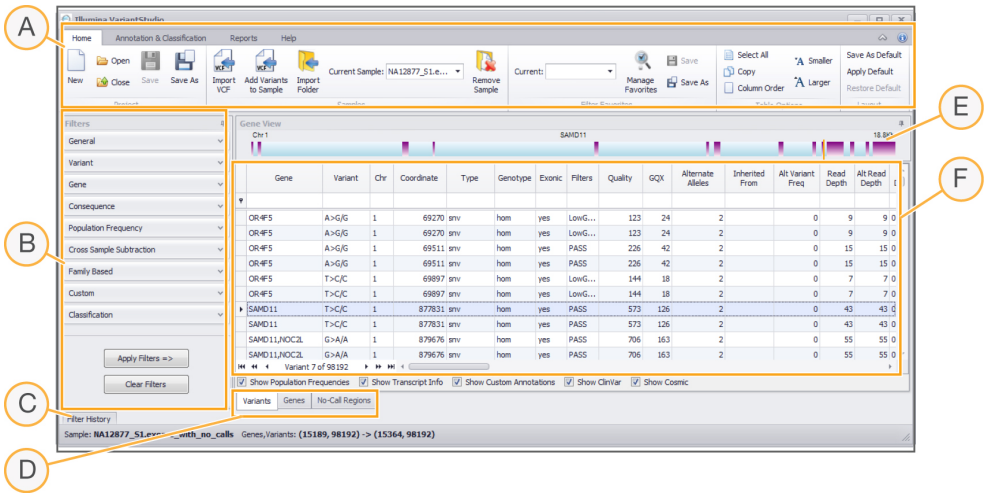
Figure 2 BaseSpace File Browser



Interface Commands

The BaseSpace VariantStudio interface is an interactive view of genes and variants in a selected sample. Use the interface commands to import VCF files, sort data, apply filters, and export data to a report.

Figure 3 BaseSpace VariantStudio Interface

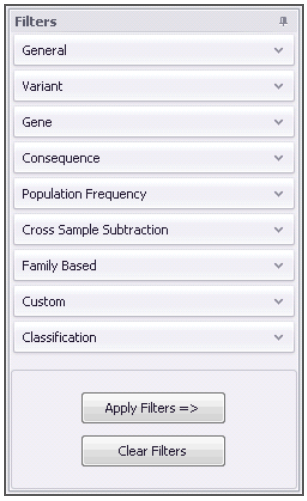


- **A** **Menu and commands**—Contains commands for managing the project, annotating variants, and reporting results. Commands are organized in four tabs: Home, Annotation and Classification, Reports, and Help.
- **B** **Filters pane**—Provides options for filtering data using any combination of filters.
- **C** **Filter history**—Opens the history panel that shows all filters applied to the project.
- **D** **Table tabs**—Navigation between the Variants table, Genes table, and No-Call Regions table.
- **E** **Gene view**—Shows a graphical representation of the selected gene.
- **F** **Table views**—View of data shown in the Variants table, Genes table, and No-Call Regions table. Use the table tabs to toggle between table views.

Filters Pane

The Filters pane provides various filtering options to narrow results to your area of interest. Combine any number of filtering options from the filter categories and click **Apply Filters**. Filters are applied to the current sample only, not to all samples that are imported into the project. To clear filters, click **Clear Filters**. For more information, see *Apply Filters* on page 40.

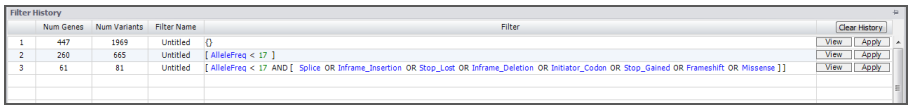
Figure 4 Filters Pane



Filter History

The Filter History pane shows filters that have been applied to the samples in this project. Filters can be a single filter, a combination of filters, or a saved filter from the favorite filter list.

Figure 5 Filter History Pane



Column Heading	Description
Num Genes	The number of genes showing with the filters applied.
Num Variants	The number of variants showing with the filters applied.
Filter Name	The name of the filter applied. The filter name appears only if the filter was saved as a favorite. Otherwise, the filter name is Untitled.
Filter	The description of the filter applied, which can describe one filter or a combination of filters.

The Filter History pane includes three buttons: Clear History, View, and Apply:

- **Clear History**—Clears entries in the filter history pane.
- **View**—Shows a block diagram illustration of the filter.
- **Apply**—Applies the filter to the variants table.

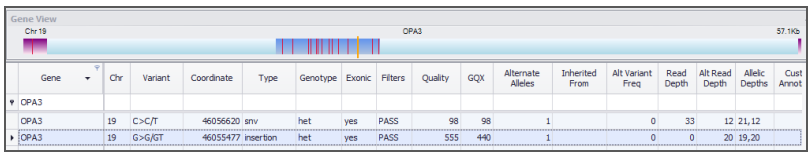
You can export information from the Filter History pane in a comma-separated values (CSV) file format. For more information, see *Generating Reports* on page 57.

Gene View

The Gene View shows a graphical representation of the gene with the following indicators:

- Exons are indicated in dark blue.
- Variants are indicated with a red line.
- The selected variant is indicated with an orange line.
- The selected transcript is indicated in purple.
- No-call regions are indicated in gray.

Figure 6 Gene View



The Gene View is interactive. Using your mouse, hover over the view to see the coordinate. Click and drag your mouse to slide the view from end to end. Use the scrolling feature on your mouse to zoom in and zoom out.

Menus and Commands

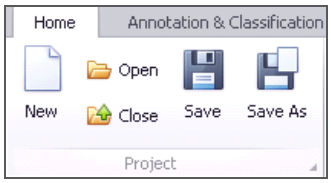
BaseSpace VariantStudio commands are arranged in the following four tabs:

- **Home tab**—Contains commands for saving projects, importing data, managing favorite filters, and changing layout options. For more information, see the following sections:
  - *Import Variant Call Files* on page 10
  - *Modify Table Views* on page 18
  - *Create Favorite Filters* on page 55
- **Annotation and Classification**—Contains commands to annotate variants and apply classifications. For more information, see *Annotate Variants* on page 26.
- **Reports tab**—Contains commands for exporting results to reports. For more information, see *Introduction* on page 58.
- **Help tab**—Contains information about the software version and a link to online help. An internet connection is required to access the help files.

Create or Open a Project

The Project menu includes commands to create, open, save, and name projects.

Figure 7 Project Menu Commands

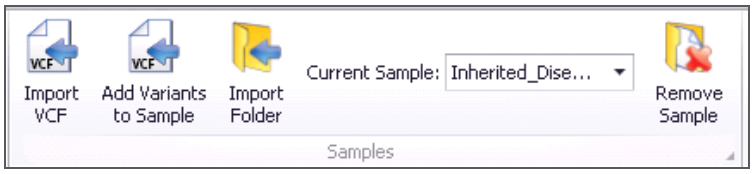


Command	Description
New	Creates a project. Starting a new project closes the current project. If you have not yet saved changes to the current project, a reminder to save your changes appears.
Open	Opens a project. Opening another project closes the current project. If you have not yet saved changes to the current project, a reminder to save your changes appears.
Close	Closes the current project. This command does not close the software application. If you have not yet saved changes to the current project, a reminder to save your changes appears.
Save	Saves changes made to an open project. If your project is new, you are prompted to name the project.
Save As	**Save As...**: Provides the option to save the current project with a different name. **Save As Reduced Project**: Provides the option to save the current project with a different name, at a reduced size. You cannot recover variants that were filtered out when saving at a reduced size.

Import Variant Call Files

From the Home tab, use commands on the Samples menu to import variant call files in VCF file format and manage samples in the project.

Figure 8 Samples Menu Commands

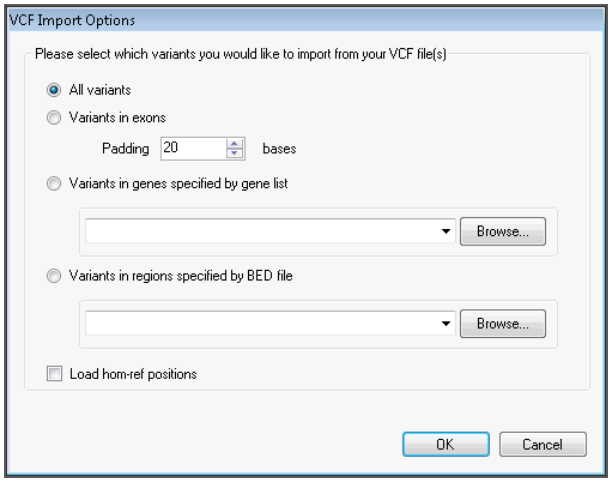


Command	Description
Import VCF	Opens a window to browse to a file location and import one selected VCF file.
Add Variants to Sample	Opens a window to browse to a file location and import SNPs and indels from another VCF file. This command imports data from the selected VCF file and adds it to the current sample. **Important**: There is no change to the sample name to denote that variant calls have been merged.
Import Folder	Opens a window to browse to a folder location and import all VCF files in the selected folder.
Current Sample	Shows the active sample name. The Current Sample drop-down list shows all samples in the project. To change to another sample in the project, select a sample name from the drop-down list.
Remove Sample	Removes the current sample from the project. A confirmation dialog box opens before the sample is removed from the project.

VCF Import Options

With any command to import variant calls, the VCF Import Options dialog box opens. From this dialog box, specify which variants to import using one of four options.

Figure 9 VCF Import Options

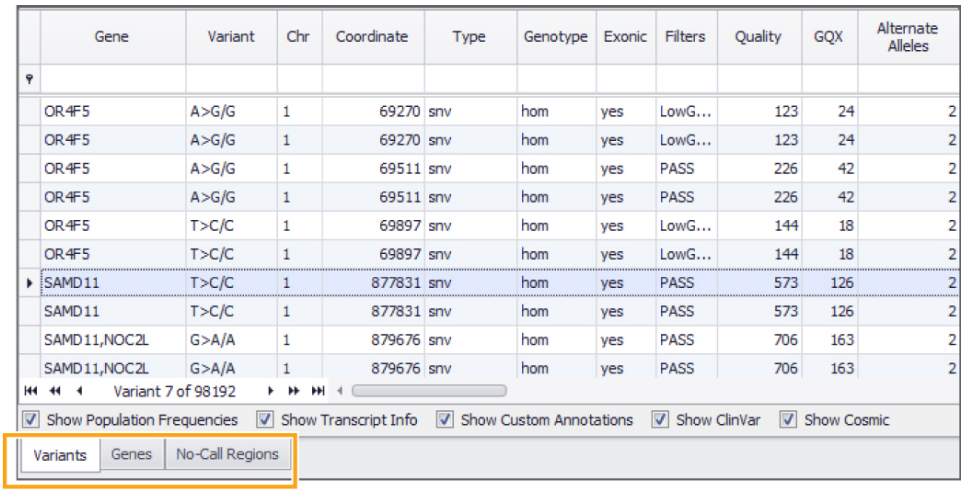


Command	Description
All variants	Select the radio button to import all variants in the selected VCF files.
Variants in exons	Select the radio button to import only variants found in exonic regions. With this option, set the number of bases, or padding, to include on both sides of the exon.
Variants specified in a gene list	Select the radio button and click **Browse** to navigate to the location of a gene list file. A gene list file must be a text file with a *.txt file extension that lists one gene per line.
Variants in regions specified in a BED file	Select the radio button and click **Browse** to navigate to the location of the BED file.
Load hom-ref positions	This setting applies to gVCF files. Select the checkbox to import all homozygous reference positions, 0/0. Clear the checkbox to omit homozygous reference positions from the import.

Data in BaseSpace VariantStudio Tables

Imported and annotated information for the visible sample is arranged in three tables on the BaseSpace VariantStudio interface: Variants table, Genes table, and No-Call Regions table. Use the tabs below the table area to navigate between tables.

Figure 10 Navigation Tabs for Variants Table, Genes Table, and No-Call Regions Table



Variants Table

The Variants table lists the genes that overlap variants identified in the selected sample. Each row of the table contains the gene and reported variant. Genes that include multiple variants are listed multiple times in the table, one time for each variant.

The following information is provided in the Variants table. If a column described in the following is not visible in your instance of BaseSpace VariantStudio, click **Column Order** from the Table Options menu to view hidden columns.

Column Heading	Description
Gene	The name of the gene.
Variant	Lists the reference allele and the diploid genotype call for the sample as Reference > AlleleA/AlleleB. AlleleA and AlleleB are explicitly defined from the REF, ALT, and GT fields of the VCF file. For example, at a heterozygous position noted as GT=0/1 is represented as REF > REF/ALT, and a homozygous non-reference position noted as GT=1/1 is represented as REF > ALT/ALT.
Chr	The chromosome number in which the gene occurs.
Coordinate	The genomic location of the variant (1-based).
Classification	The classification assigned to the variant. This field is populated for variants that match criteria specified in the classification database.
Type	The type of variant, which is either a single nucleotide variant (SNV), insertion, deletion, or ref (for reference call).

Column Heading	Description
Genotype	The genotype, which is either heterozygous (het), homozygous (hom) or somatic (som).
Exonic	A variant found within a coding region, ±20 bp on both sides of the coding region.
Filters	The status of the variant call quality as annotated in the VCF file. PASS indicates that all filters were passed; otherwise the variant call filter is listed. The filter listed and threshold for passing filter depends on the method used to generate the VCF file.
Quality	The numeric value of variant call quality as written to the QUAL column of the VCF file. Determination of variant quality depends on the variant caller.
GQX	The conservative measure of genotype quality derived from the minimum of the GQ and QUAL values listed in the VCF file. This field is not populated for somatic VCF files. For more information, see *Somatic VCF Fields Reported in the Variants Table* on page 22.
Inherited From	The inherited source of the variant. Possible values are father, mother, both, indeterminate, or ambiguous. If the variant is heterozygous in the father, mother, and child, a variant is listed as ambiguous. If the variant is homozygous in the child and heterozygous in both parents, a variant is listed as both. If the inheritance of the variant cannot be determined from the other VCFs, the variant is listed as indeterminate. Entries in this column are meaningful only when the family-based filter is applied.
Alt Variant Freq	The frequency of the Alt Allele.
Read Depth	The total number of reads passing quality filters at this position.
Alt Read Depth	The number of reads called at this position.
Allelic Depth	The number of reads called for the Ref Allele and the Alt Allele.
Custom Annotation	Annotations according to values provided in the Annotation column of an optional custom annotation file.
Custom Gene Annotation	Annotations according to values provided in the Annotation column of an optional custom gene annotation file.
Custom Gene Annotation 2	Annotations according to values provided in the Gene Annotation 2 column of an optional custom annotation file.
Num Transcripts	The number of transcripts reported in the annotation, which includes overlapping transcripts and upstream and downstream transcripts within 5 kb of the variant.
Transcript	The name of the transcript, usually a database identifier from RefSeq or Ensembl.

Column Heading	Description
Consequence	Consequence of the variant, described in Sequence Ontology standardized vocabulary.
cDNA Position	Position of the variant in cDNA.
CDS Position	Position of the variant in the coding region.
Protein Position	Position of the amino acid in the protein
Amino Acids	Amino acid or amino acid change. If the variant is synonymous, then there is no change and one amino acid is listed.
Codons	Specific codon noted with and without the variant, highlighted in uppercase.
HGNC	The gene name, expressed as official HGNC nomenclature.
Transcript HGNC	The transcript name, expressed as official HGNC nomenclature.
Canonical	Indicates whether the transcript is the canonical transcript.
SIFT	SIFT score.
PolyPhen	PolyPhen score.
ENSP	Protein ID (Ensembl ID).
HGVSc	Human Genome Variation Society (HGVS) notation in the cDNA.
HGVSp	Human Genome Variation Society (HGVS) notation in the protein.
dbSNP ID	The rsID entry in dbSNP.
Ancestral Allele	The inferred allele ancestral to the chimpanzee/human lineage. For more information, see www.1000genomes.org/faq/where-does-ancestral-allele-information-your-variants-come.
Allele Freq	The allele frequency from all populations of 1000 genomes data; April 2012 phase 1 call set (v3 update).
Allele Freq Global Minor	Global minor allele frequency (GMAF); technically, the frequency of the second most frequent allele. For more information, see www.ncbi.nlm.nih.gov/projects/SNP/docs/rs_attributes.html#gmaf.
Global Minor Allele	The specific allele with the reported GMAF.
Allele Freq Amr	The allele frequency from 1000 Genomes (Ad Mixed American population).
Allele Freq Asn	The allele frequency from 1000 Genomes (East Asian population).
Allele Freq Af	The allele frequency from 1000 Genomes (African population).
Allele Freq Eur	The allele frequency from 1000 Genomes (European population).
Allele Freq Evs	The allele frequency from the NHLBI exome sequencing project. Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA (evs.gs.washington.edu/EVS) [November 2012 accessed]

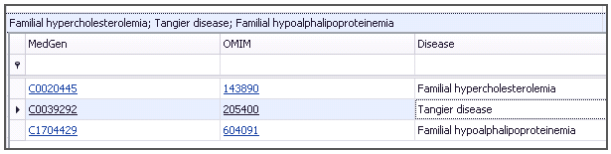
Column Heading	Description
EVS Coverage	The average depth of coverage for SNVs that were called at this position from the Exome Variant Server (EVS).
EVS Samples	The number of samples that were called at this position from the Exome Variant Server (EVS).
Conserved Sequence	Denotes if the variant is an identical or similar sequence that occurs between species and maintained between species throughout evolution.
COSMIC ID	The numeric identifier for the variant in Catalogue of Somatic Mutations in Cancer (COSMIC) database, if the genomic position of the variant overlaps a variant listed in COSMIC. The COSMIC ID links to the COSMIC page associated with the identifier.
COSMIC Wildtype	The allele in unaffected individuals as reported in the COSMIC database.
COSMIC Allele	The allele as reported in the COSMIC database.
COSMIC Gene	The gene name as reported in the COSMIC database.
COSMIC Primary Site	The primary tissue type associated with the allele as reported in the COSMIC database.
COSMIC Histology	The tissue type associated with the allele as reported in the COSMIC database.
ClinVar Accession	The alpha-numeric ID assigned to the allele in the ClinVar database and link to the associated page of the ClinVar database.
ClinVar Ref	The Reference Allele as reported in the ClinVar database.
ClinVar Alleles	The name of the allele as reported in the ClinVar database.
ClinVar Allele Type	The type of allele, either single nucleotide variant (SNV), insertion, deletion, or duplication as reported in the ClinVar database.
ClinVar Significance	The clinical significance or classification assigned to the allele as reported in the ClinVar database.
Regulatory Feature	A link to regulatory information in Ensembl for that genomic region.
Alternate Alleles	The number of nucleotide bases called for Allele A and Allele B that differ from the RefAllele.
Google Scholar	Link to the Google Scholar search page for the selected variant, cDNA, and amino acid. The page opens with an auto-populated search field. Click search to continue. The search is transcript-dependent. If a rs number is present, the number is included in the search.

Column Heading	Description
PubMed	Link to PubMed search page for the selected variant, cDNA, and amino acid. The page opens with an auto-populated search field. Click search to continue. The search is transcript-dependent. If a rs number is present, the number is included in the search.
UCSC Browser	Link to UCSC Browser search page for the selected chromosome and position.
ClinVar RS	The numeric rsID assigned to the allele in the ClinVar database.
ClinVar Disease Name	The disease associated with the allele as reported in the ClinVar database.
ClinVar MedGen	The alpha-numeric identifier of the disease as reported by MedGen and link to the associated page of the MedGen database.
ClinVar OMIM	The numeric identifier for the disease as reported by Online Mendelian Inheritance in Man (OMIM) and link to the associated page of the OMIM database.
ClinVar Orphanet	The numeric identifier of the disease as reported by Orphanet and link to the associated page in the Orphanet database.
ClinVar Gene Reviews	The alpha-numeric identifier of the disease as reported by Gene Reviews and link to the associated page in the Gene Review database.
ClinVar SnoMedCt ID	The numeric identifier of the disease and associated clinical terms as reported by SnoMedCt.
Exon	The exon number in which the variant is present.
Intron	The intron number in which the variant is present.
Distance	Distance between the variant and the nearest end of the gene. • For upstream variants, this value is the distance to the beginning of the first exon. • For downstream variants, this value is the distance to the end of the last exon.

Genes Table

The Genes table lists the genes that contain variants identified in the selected sample. Each row of the table contains the gene and number of variants reported, along with the following information reported for each gene.

Column Heading	Description
Name	The name of the gene.
Gene ID	The Entrez Gene ID for the gene and link to the associated entry in the NCBI database.
Chr	The chromosome number in which the gene occurs.

Column Heading	Description
Start	The start coordinate of the gene (1-based).
Stop	The end coordinate of the gene.
Length	The length of the gene.
Num Alleles	The number of alleles reported.
Num Variants	The number of variants reported.
Paternal Variants	The number of variants inherited from the father.
Maternal Variants	The number of variants inherited from the mother.
Ambiguous Variants	The number of ambiguous variants.
Custom Gene Annotation	Gene annotations according to values provided in the Annotation column of an optional custom gene annotation file.
Custom Gene Annotation 2	Annotations according to values provided in the Gene Annotation 2 column of an optional custom annotation file.
PubMed	Link to PubMed search page for the selected gene. The page opens with an auto-populated search field. Click search to continue.
GeneReviews	Link to GeneReviews website. Clicking the entry provides a drop-down list with links to GeneReviews.
Disease	Diseases associated with the gene. Clicking the disease name provides a drop-down list with links to MedGen and OMIM.  Familial hypercholesterolemia; Tangier disease; Familial hypoalphalipoproteinemia MedGen / OMIM / Disease C0020445 / 143890 / Familial hypercholesterolemia C0039292 / 205400 / Tangier disease C1704429 / 604091 / Familial hypoalphalipoproteinemia
Description	The protein name associated with gene function.

No-Call Regions Table

The No-Call Regions table shows regions where calls could not be confidently made due to a low read depth or failing a quality filter. The No-Call Regions table is populated directly from the non-variant regions reported in the genome VCF (gVCF) file. Non-variant regions are reported as 0/0 in the gVCF file. For more information, see sites.google.com/site/gvcftools/home/about-gvcf/gvcf-conventions

Each row of the table contains the gene and information reported for the gene. The following information is provided in the No-Call Regions table.

Column Heading	Description
Gene	The name of the gene located within the no-call region.
Chr	The chromosome number in which the no-call region occurs.
Start	The start chromosomal coordinate of the no-call region.

Column Heading	Description
Stop	The end chromosomal coordinate of the no-call region.
Length	The length of the no-call region.
Depth	The read depth of the no-call region.
Quality	The numeric value of variant call quality as written to the QUAL column of the VCF file. Determination of variant quality depends on the variant caller.
Filter	The filter associated with the variant call quality as annotated in the VCF file.

Modify Table Views

To modify how data appear in the tables, click the column headings. Options include sorting in descending or ascending order, showing only selected data based on listed values, or adjusting column order.

NOTE
Modifying how data appear in the Variants table only affects how information is arranged in the table. Modifying views does not change the underlying data.

Sort Data in Ascending or Descending Order

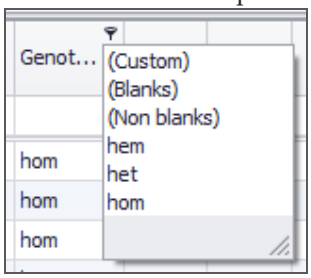
To change the order in which data appear in the Variants table, click a column heading. Data are sorted in either descending or ascending order of values listed in that column. Click again to reverse the order.

- When the table is sorted in ascending order, the up arrow icon appears in the column heading.
- When the table is sorted in descending order, the down arrow icon appears in the column heading.

Show Only Selected Data

To show only selected data based on information in the Variants table, use the show/hide icon in the column heading.

1 Click the show/hide icon in the column heading. A drop-down list opens that contains all values present in that column.



2 Select a value from the drop-down list. The Variants table shows only data that contains your selection.

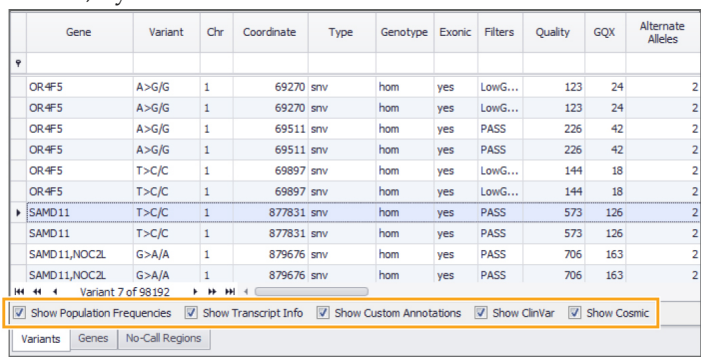
To restore the default view of the Variants table, use one of the following methods:

- Click the show/hide icon in the column heading used to modify the table and select **All** from the drop-down list.
- Click the icon at the bottom of the Variants table.

A history of previous selections appears at the top of the column heading drop-down list for quick access to frequently used selections. To remove a selection from history, click the delete icon.

Show or Hide Selected Columns

Use the checkboxes below the table tabs to show or hide specific columns in the Variants table. Select the checkbox to show data, and clear the checkbox to hide data. All options are set to show, by default.

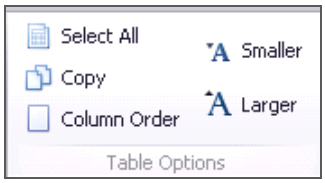


Show/hide options include the following sections of the Variants table:

- **Show Population Frequencies**—Shows and hides Allele Freq, Allele Freq Global Minor, global Minor Allele, Allele Freq Amr, Allele Freq Asn, Allele Freq Af, Allele Freq Eur, and Allele Freq Evs.
- **Show Transcript Info**—Shows and hides Num Transcripts, Transcript, Consequence, cDNA Position, CDS Position, Amino Acids, Codons, Exon, Intron, Transcript HGNC, Distance, Canonical, Sift, PolyPhen, ENSP, HGVSc, and HGVSp.
- **Show Custom Annotations**—Shows and hides Custom Annotation, Custom Annotation 2, Custom Annotation 3, Custom Annotation 4, and Custom Gene Annotation.
- **Show ClinVar**—Shows and hides ClinVar RS, ClinVar Ref, ClinVar Alleles, ClinVar Significance, ClinVar Disease Name, ClinVar Accession, ClinVar MedGen, ClinVar OMIM, ClinVar Orphanet, ClinVar Gene Reviews, and ClinVar SnoMedCt ID.
- **Show COSMIC**—Shows and hides COSMIC ID, COSMIC Wildtype, COSMIC Allele, COSMIC Gene, COSMIC Primary Site, and COSMIC Histology.

Set Table Options

Figure 11 Table Options Menu



The Table Options menu includes the following commands:

- **Select All**, which selects all rows in a table.

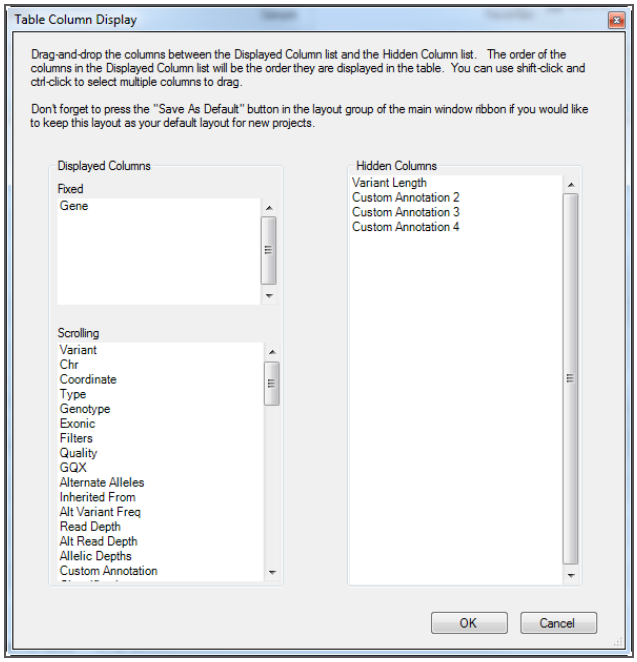
- **Copy**, which copies selected data to the clipboard.
- **Smaller** and **Larger**, which changes text size in a table.
- **Column Order**, which includes commands to change table layout.

From the Table Options menu, click **Column Order** to open the Table Column Display window. From this window, drag and drop column headings to specify table layout:

1. To show or hide columns, drag and drop column headings from the Displayed Columns list to the Hidden Columns list.
2. To prevent selected columns from scrolling horizontally, drag and drop column headings from the Scrolling list to the Fixed list.
3. Click **OK** when you are finished.
4. To save this layout for use in other BaseSpace VariantStudio projects, click **Save As Default** from the Layout menu.

In the following example, the Gene column is set to Fixed, and variant length and optional custom annotation columns are hidden.

Figure 12 Table Column Display Window



VCF Fields Reported in the Variants Table

Several columns of the Variants table are populated from columns or fields in the VCF file, as described in the following table.

Variants Table Column Heading	VCF File Column or Field Description
Allelic Depth	Based on values listed for AD in INFO or FORMAT/[Sample Name].
Alt Read Depth	Based on the second value listed for AD in INFO or FORMAT/[Sample Name].
Alt Variant Freq	Based on values listed for VF in INFO or FORMAT/[Sample Name].
Chr	Based on values in the CHROM column.
Coordinate	Based on values in the POS column.
Exonic	Based on values in the CHROM and POS columns, and calculated from a list of exonic regions.
Filters	Based on values in the FILTER column.
Gene	Based on values in the CHROM and POS columns, and calculated using a list of gene coordinates.
Genotype	Based on values listed for GT in FORMAT/[Sample Name].
GQX	Based on values listed for GQX in FORMAT/[Sample Name].
Quality	Based on values in the QUAL column.
Read Depth	Based on values listed for DP in FORMAT/[Sample Name], DPI in FORMAT/[Sample Name] for insertion and deletion events called by the Illumina Isaac Alignment and Variant Calling workflow.
Type	Based on the number of bases in the REF and ALT columns.
Variant	Based on values in the REF and ALT columns. • At a heterozygous position, the value is REF > REF/ALT. • At a homozygous position, the value is REF > ALT/ALT.

NOTE

Some fields reported in the Variants table differ for somatic VCF files. For more information, see *Somatic VCF Fields Reported in the Variants Table* on page 22.

Somatic VCF Fields Reported in the Variants Table

Information reported in BaseSpace VariantStudio for VCF files generated by the Illumina cancer analysis pipeline differs from what is reported for other VCF files.

For these files, there is no genotype (GT) or genotype score (GQX). Instead, allelic depths are listed.

Each VCF includes two samples, one of which is a reference and the other is the cancer sample. All reported values are specific to the cancer sample.

The following table lists the VCF fields that are unique to somatic VCF files.

Variants Table Column Heading	VCF File Column or Field Description
Allelic Depth	Based on values in the FORMAT column. Allelic Depth is calculated differently for SNVs and indels: • **For SNVs**—Based on four values listed as AU:CU:GU:TU in the FORMAT column. These values are listed as two numbers each, separated by a comma, and represent each possible allele in the cancer sample. The Allelic depth column is populated with the full set of numbers, 0,0:0,0:10,10:3,4. • **For indels**—Two values listed as TAR:TIR in the FORMAT column represent the Ref Allele and Alt Allele, respectively. Only the first number in each value is used. In the example 0,0:12,12, the Ref Allele is 0 and the Alt Allele is 12. Allelic Depth is listed as 0,12.
Alt Read Depth	Based on values in the FORMAT column. Alt Read Depth is calculated differently for SNVs and indels: • **For SNVs**—Based on the first value from the appropriate Allelic Depth (AU:CU:GU:TU). In the example 0,0:0,0:10,10:3,4, the values are 10,10 for GU and 3,4 for TU. If the Ref Allele is G and the Alt Allele is T, the Alt Read Depth is 3. • **For indels**—Based on the first value from the appropriate Allelic Depth (TAR:TIR). In the example 0,0:12,12, the Ref Allele is 0 and the Alt Allele is 12. Alt Read Depth is listed as 12.
Alt Variant Freq	For somatic VCF files, allele frequency is calculated from values in the VCF file before data are reported in the Variants table. • **For SNVs**—Using only the first values for AU:CU:GU:TU, allele frequency is calculated as (alt allelic depth/(alt allelic depth + ref allelic depth))*100. In the example 0,0:0,0:10,10:3,4, Alt Variant Freq is 23.08% by calculating (3/(3+10))*100. • **For indels**—Using only the first values for TAR and TIR, allele frequency is calculated as (TIR/(TIR+TAR))*100. In the example 0,0:12,12, Alt Variant Freq is 100% by calculating (12/(12+0))*100.
Genotype	Based on values listed in the INFO column. If SOMATIC is listed in the INFO column, the genotype is listed as somatic (som) in the Variants table.

Variants Table Column Heading	VCF File Column or Field Description
Quality	Quality is based on different values for SNVs and indels: • **For SNVs**—Quality is based on the QSS_NT field in the INFO column. This score represents the probability that the SNV exists and is somatic. • **For indels**—Quality is based on the QSI_NT field in the INFO column. This score represents the probability that the indel exists and is somatic.
Read Depth	For SNVs and indels, Read Depth is extracted from values listed for DP in the FORMAT column of the cancer sample.

Chapter 2

Applying Annotations and Classifications

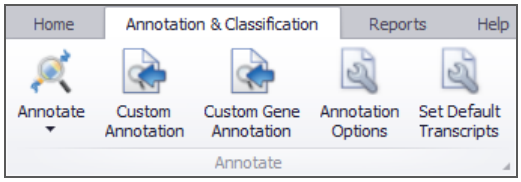
Annotate Variants

From the Annotations and Classification tab, use commands on the Annotate menu to annotate variants in the current sample with options to annotate all variants or only those variants specified. All coordinates used in BaseSpace VariantStudio are genomic coordinates on the positive strand.

Always annotate variants before applying filters.

NOTE
An internet connection is required to annotate variants. After annotating, an internet connection is not necessary.

Figure 13 Annotate Menu



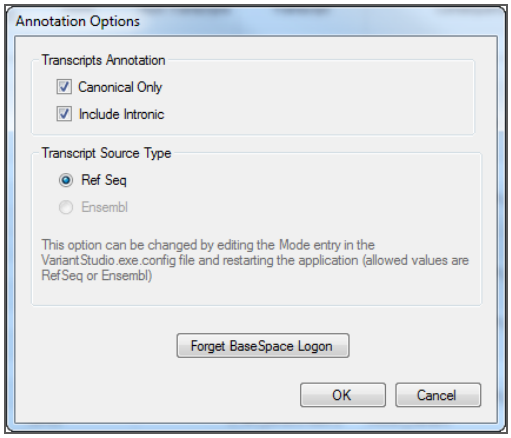
Command	Description
Annotate	Annotates variants in the project using the following options: • **All Variants of Current Sample**—Annotates all variants in the current sample. The current sample is listed in the Current Sample field of the Samples menu. • **Exonic Variants of Current Sample**—Annotates variants found within an exon plus 20 bp on either side of the exonic region to include the annotation of splice site variants. • **Selected Variants of Current Sample**—Annotates only the variants that you have selected or filtered. • **All Samples**—Annotates all variants within each sample imported into the project. This process can take time to complete depending on the number of samples in the project.
Custom Annotation	Opens a window to browse to the location of the custom annotations file for variant-level annotation. For more information, see *Input File for Custom Variant Annotations* on page 29.
Custom Gene Annotation	Opens a window to browse to the location of the custom annotations file for gene-level annotation. For more information, see *Input File for Custom Gene Annotations* on page 29.
Set Default Transcripts	Opens a window that lists the default transcript for each gene and options for changing to other than the default transcript.

NOTE
Although mitochondrial variants can be imported, the annotation database does not provide annotations for these variants.

Annotation Options

From the Annotate menu, click **Annotation Options**. The Annotation Options form opens with options to annotate only certain variants.

Figure 14 Variant Option Form

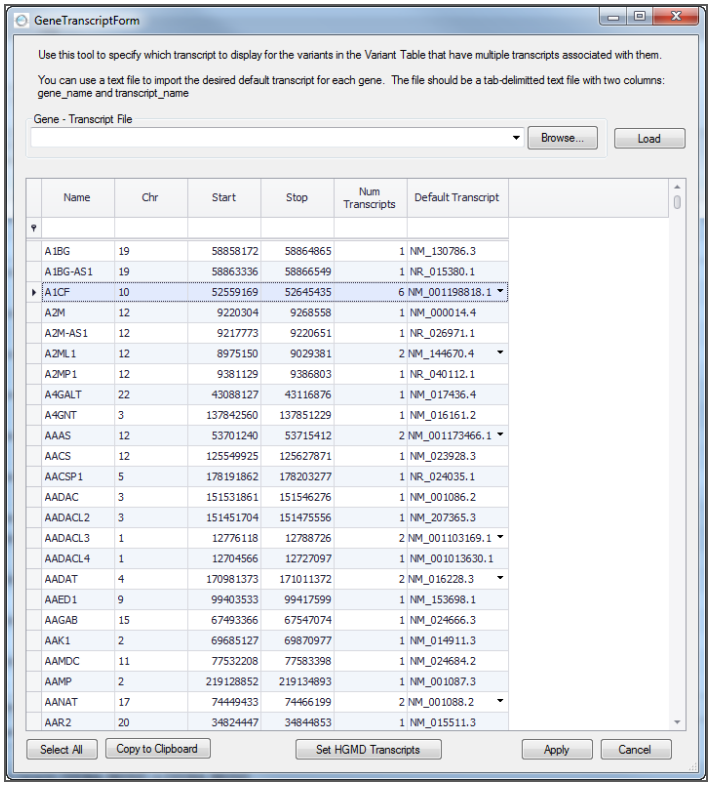


Option	Description
Transcript Annotation	Provides options to annotate only variants in the canonical transcript, which is the longest translated transcript in the gene, and variants in intronic regions.
Transcript Source Type	Annotates variants identified in a specific annotation source, which is RefSeq, by default. The default can be changed to Ensembl by editing the mode entry in the BaseSpace VariantStudio configuration file (VariantStudio.exe.config), as follows: <add key="Mode" value="Ensembl"/> Close and reopen BaseSpace VariantStudio to enable the change.
Forget BaseSpace Logon	Clears BaseSpace login information, such as ID and password.

Set Default Transcripts

1. Click **Set Default Transcripts**. A window opens that lists the default transcript for each gene. By default, BaseSpace VariantStudio lists the canonical transcript, which is the longest translated transcript in the gene.
2. For genes with multiple transcripts, use the drop-down list to set the default to another transcript.

Figure 15 Set Default Transcripts



Alternatively, click **Browse** to navigate to a tab-delimited text file containing your preferred default transcripts and click **Load**.

Input File for Default Transcripts

The input file for default transcripts requires two columns: Gene_Name and Transcript_Name, as shown in the following example.

Gene_Name	Transcript_Name
ACTN3	NM_003793.3
ADH1B	NM_000668.4
AKAP10	NM_007202.3

Create Custom Annotations

Applying custom annotations requires a tab-delimited input file with a *.txt extension. One input file is required for variant-level annotations and one for gene-level annotations.

NOTE
BaseSpace VariantStudio assumes that all annotations are expressed in the genomic coordinates on the positive strand, including any custom annotations that are imported into BaseSpace VariantStudio.

Input File for Custom Variant Annotations

The input file for custom variant annotations requires five columns: Chr, Position, Ref, Variant, and Annotation, as shown in the following example.

Chr	Position	Ref	Variant	Annotation
1	11046855	G	T	Good
1	11046909	A	T	Bad
1	14096821	T	C	Confirmed

- **Chr**—The chromosome for the variant (1–22, X, Y, or M).
- **Position**—The genomic coordinate of the variant on the chromosome (1-based).
- **Ref**—The reference base, or bases for an insert or deletion, at the specified position.
- **Variant**—The base, or bases for an insert or deletion, at the specified position.
- **Annotation**—The value assigned to a variant with matching values for chr, position, ref, and variant.
- (Optional) Three additional annotation columns are recognized input for custom variant annotations. Use the headings **Annotation2**, **Annotation3**, and **Annotation4**.

NOTE
The Ref field and Variant field must be expressed in VCF format, where indels contain the preceding base in common between the reference and variant allele. For more information, see www.1000genomes.org/wiki/Analysis/Variant Call Format/vcf-variant-call-format-version-41. On this site, go to step 3, **Data Lines, Fixed Fields**, and then step 4 **Ref**.

Input File for Custom Gene Annotations

The input file for gene annotations requires two columns: Gene and Annotation, as shown in the following example.

Gene	Annotation
AGRN	Myasthenia, limb-girdle, familial
CCDC39	Ciliary dyskinesia, primary, 14
DHTKD1	2-aminoadipic 2-oxoadipic aciduria

- **Gene**—The gene symbol.
- **Annotation**—The value assigned to the specified gene.

Apply Custom Annotations

1. Create a custom annotations file using a text editor, such as Notepad, and save it with a *.txt file extension.
2. From the Annotate menu, click **Custom Annotations** to apply annotations to variants or **Custom Gene Annotations** to apply annotations to genes.

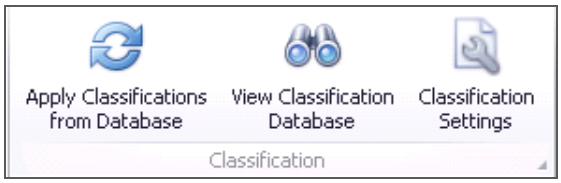
3 Browse to the custom annotations file and click **OK**. This step links the custom annotations file to the project.

4 Use the custom filters in the Filters pane to filter data based on custom annotations. For more information, see *Custom Filters* on page 47.

Apply Variant Classifications

Introduced in BaseSpace VariantStudio v2.1, you can apply classifications to variants according to their biological impact. Classifications are stored in the classification database.

Figure 16 Classification Menu



Command	Description
Apply Classifications from Database	Use this command to apply classifications to any variants in the current sample that are listed in the classification database. For more information, see *Apply Classifications from Database* on page 32.
View Classification Database	Use this command to open the classification database, edit entries in the database, or import classifications from an external file. For more information, see *View Classifications Database* on page 35.
Classification Settings	BaseSpace VariantStudio provides five classifications: Benign, Presumed Benign, Presumed Pathogenic, Pathogenic, and Unknown Significance. Use this command to add or remove classification categories. For more information, see *Add or Remove Classification Categories* on page 35.

Variant classifications can be changed at any time in the classification database or changed locally in the current project without changing database entries. For more information, see *Edit Variant Classifications* on page 33.

By default, the classification database is saved locally for use with any BaseSpace VariantStudio project that is opened locally. For more information, see *Classification Database Location* on page 36.

A backup of the classification database is created with the first change of each day. For more information, see *Classification Database Backup* on page 36.

For each classified variant, two text fields are available for recording comments about the variant, the Notes field and the Report Fragment field:

- **Notes**—Information in Notes field is stored in the classification database only.
- **Report Fragment**—Information in the Report Fragment field is stored in the classification database and exported as a column in the sample report. For more information, see *Sample Report Overview* on page 59.

There are three ways to apply classifications to variants in a project:

- From the menu, apply classifications to variants in the current sample that are listed in the classification database.
- From the Variants table, apply a classification to a selected variant in the Variants table and save the classification to the database.
- From the Variants table, apply a classification to multiple selected variants in the Variants table and save the classification to the database. The same classification must apply to all selected variants.

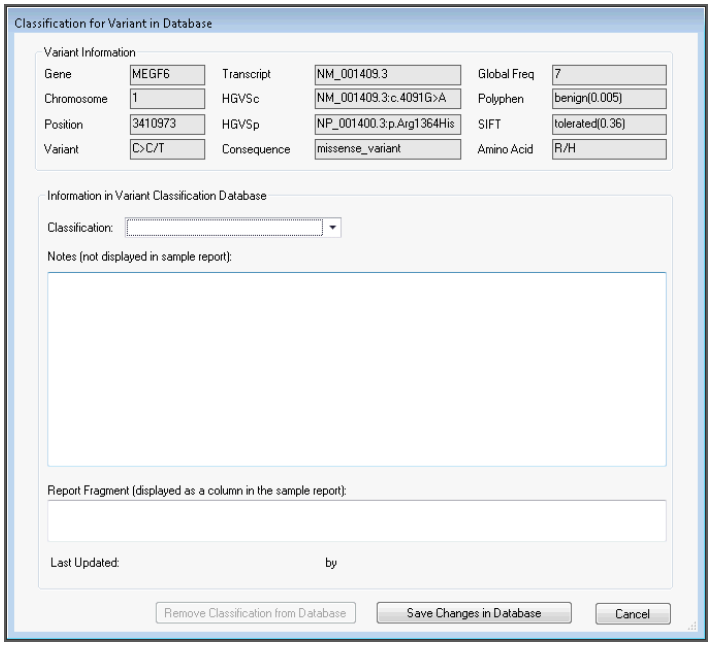
Apply Classifications from Database

Click **Apply Classifications from Database**. Any variants in the current project that have matching criteria in the classification database are annotated with the classification as specified in the database.

Apply Classifications in the Variants Table

1 Click the ... icon in the Classification column for the variant you want to classify. The Classification for Variant in Database window opens, which shows information for the variant and provides a drop-down list of available classification categories.

Figure 17 Classification for Variant in Database



2 Select a classification category from the Classification drop-down list, and enter any applicable comments in the Notes field and Report Fragment field.

3 Click **Save Changes to Database**. The classification can later be removed from the classification database.

Apply Classifications to Multiple Variants

1 Use shift-click or ctrl-click to select more than one row in the Variants table.

2 Right-click in the Classifications column over a selected row, and then select **Classify Selected Variants**.

3 From the Classify Selected Variants window, use the drop-down list to assign a classification. Enter any applicable comments in the Notes field and Report Fragment field.

4 Click **OK**. The classification assignments are saved to the database automatically.

Edit Variant Classifications

There are two ways to edit classifications for variants with assigned classifications:

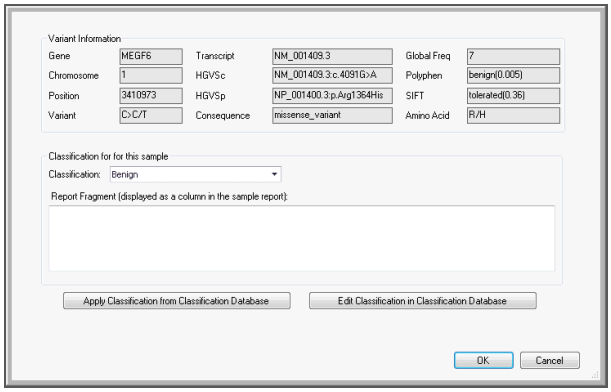
- Edit variant classifications in the database.
- Edit classifications locally in the current project without changing database entries.

Edit Classifications in the Database

1 Click the ... icon in the Classification column for the variant you want to change. A window opens that shows information about the variant, the current classification, and any comments in the Report Fragment field.

2 Click **Edit Classification in Classification Database**. The Classification for Variant in Database window opens.

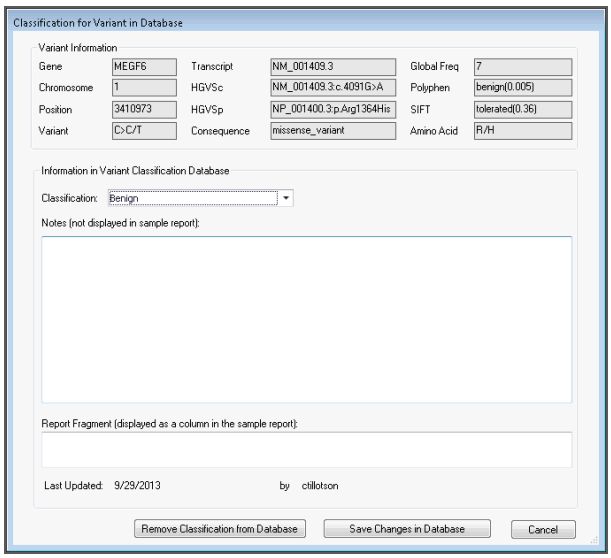
Figure 18 Edit Classifications



3 Do one of the following:

- To change the classification in the database, select a new classification from the Classification drop-down list and click **Save Changes in Database**.
- To remove the classification from the database, click **Remove Classification from Database**.

Figure 19 Remove Classification from Database

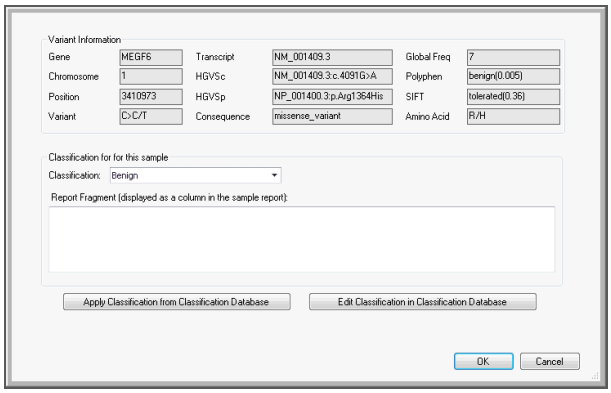


Edit Variant Classifications Locally

To change variant classifications locally, the variant must already have a classification assigned in the classification database.

1 Click the ... icon in the Classification column for the variant that you want to edit locally. A window opens that shows information about the variant, the current classification, and any comments in the Report Fragment field.

Figure 20 Reapply Classifications from Database



2 Select a different classification category from the Classification drop-down list.

3 Click **OK**. The classification is applied to the variant in the current project only. The variant classification recorded in the database can be reapplied to the variant later.

4 To revert the classification to what is assigned in the database, click the ... icon in the Classification column.

5 Click **Apply Classification from Classification Database**. The classification recorded in the database appears in the Classification field. Click **OK**.

Manage Classifications

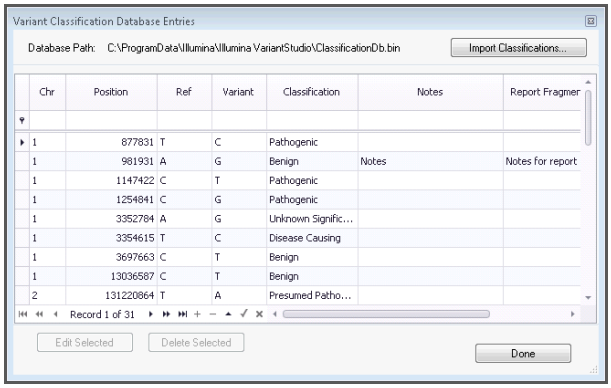
From the Annotations and Classification tab, use commands on the Classifications menu to view the classification database and manage classification settings.

View Classifications Database

Click **View Classification Database** to view the entries in the classification database. From this window, you can edit an entry, delete an entry, or import classifications from an internal file.

- **Edit an entry**—Select a row, or use shift-click or ctrl-click to select multiple rows. Click **Edit Selected**. Reassign a classification or add comments.
- **Delete an entry**—Select a row, or use shift-click or ctrl-click to select multiple rows. Click **Delete Selected**. The entry is permanently deleted from the database.
- **Import classifications**—Click **Import Classifications** and browse to the location of your external classifications file. For more information, see *Import Classifications* on page 37.

Figure 21 View Classification Database

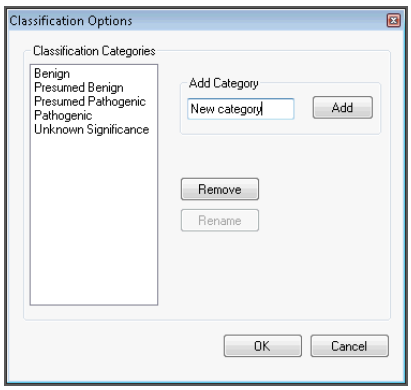


Add or Remove Classification Categories

Click **Classification Settings** to add, remove, or rename classification categories.

- **Add**—In the Add Category field, enter a new category name. Click **Add**.
- **Remove**—Select a category from the list. Click **Remove**.
- **Rename**—Select a category from the list. Click **Rename** and enter a new name.

Figure 22 Classifications Options Dialog Box



Classification Database Location

By default, the classification database is saved locally in C:\ProgramData\Illumina\Illumina VariantStudio\ClassificationDb.bin.

When the database is stored locally, classifications are available for the current project and any future projects opened on that computer.

If the classification database is stored on a network location, classifications are available to projects opened in any installation of BaseSpace VariantStudio with access to that network location.

1. To change the default setting, open the BaseSpace VariantStudio configuration file in C:\Program Files\Illumina\Illumina VariantStudio\VariantStudio.exe.config.
2. In the value field of the **ClassificationDatabaseFilePath** key, enter the preferred network path.
3. Save and close the configuration file.
4. Close and reopen BaseSpace VariantStudio to enable the change.

Classification Database Backup

A backup of the classification database is created the first time the database is changed on any given day.

The backup is named DDMMYYYY.bin and is stored in the folder DatabaseBackups, which is located in the same folder as the classification database, C:\ProgramData\Illumina\Illumina VariantStudio\DatabaseBackups.

Import Classifications

To import classifications to the classification database from an external file, create an input file in a tab-delimited text format (TSV) using a *.tsv file extension.

The input file requires five columns: Chr, Position, Ref, Variant, and Classification, as shown in the following example. Optionally, include a Notes column and a Fragment column.

Chr	Position	Ref	Variant	Classification	Notes
1	11046855	G	T	Classification 1	Note 1
1	11046868	C	G	Classification 2	Note 2
1	11046909	A	T	Classification 3	Note 3

- **Chr**—The chromosome for the variant (1–22, X, Y, or M).
- **Position**—The genomic coordinate of the variant on the chromosome.
- **Ref**—The reference base, or bases, for an insert or deletion at the specified position.
- **Variant**—The base, or bases, for an insert or deletion at the specified position.
- **Classification**—The value assigned to a variant with matching values for chr, position, ref, and variant. The classification name must match one of the classifications listed in your database.
- **Notes**—Note about the entry. Information in this field is not included in the sample report.
- **Fragment**—Notes about the entry that are intended for the sample report.

NOTE
Make sure that you add any new classification names to the database using the Classifications Settings command.

Best Practices for Importing Classifications

NOTE
All coordinates used in BaseSpace VariantStudio are genomic coordinates on the positive strand.

Before importing previously classified variants into the BaseSpace VariantStudio classification database, convert classifications to genomic coordinates. This step is especially important for variants that were classified based on HGVSc notations and transcripts.

Importing classifications before converting to genomic coordinates can result in some variants not being annotated with imported classifications when you use the command Apply Classifications from Database. Because BaseSpace VariantStudio assumes that all annotations are expressed in genomic coordinates on the positive strand, the classification database requires an exact match for variants to be annotated with stored classifications.

Applying Filters

Apply Filters

The Filters pane provides options for applying any combination of filters to the data in your project. Filters are grouped in nine expandable sections: General, Variant, Gene, Consequence, Population Frequency, Cross Sample Subtraction, Family Based, Custom, and Classification.

Figure 23 Filters Pane

1 Click the down arrow icon to expand a filter section.

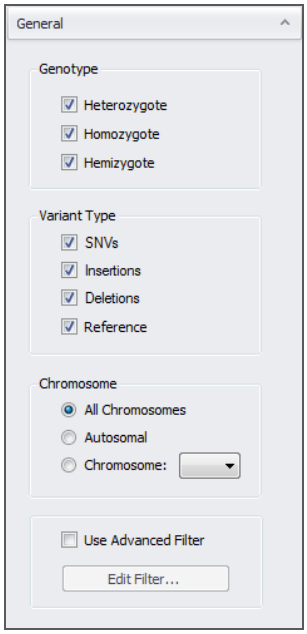
2 From the available options, select filter settings. Use any combination of settings from any number of filters.

3 Click **Apply Filters**. Filters are applied to the current sample only, not to all samples that are imported into the project.

4 Click **Clear Filters** to remove applied filters.

NOTE
You can create a filter using any combination of the filter options in the Filters pane, and then save the combination as a single filter. Saved filters can later be applied to other samples. For more information, see *Create Favorite Filters* on page 55.

General Filters



Use the General filters to filter data by genotype, variant type, and chromosome.

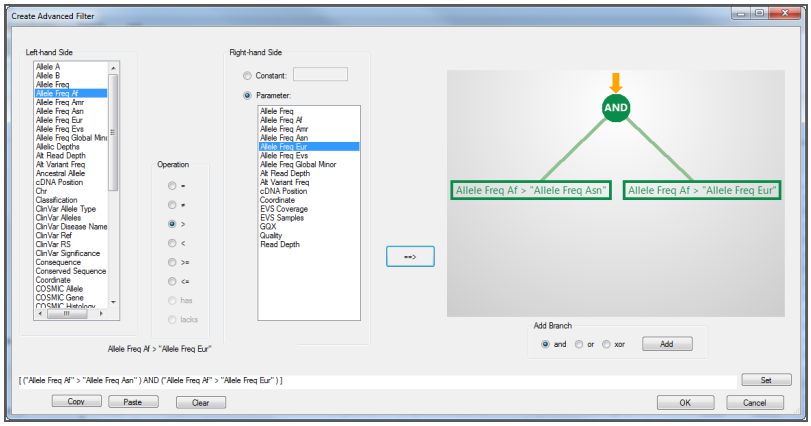
Filter Name	Setting Description
Genotype	Filters data to show any combination of heterozygote, homozygote, or hemizygote. All options are selected by default.
Variant Type	Filters data to show any combination of SNVs, Insertions, Deletions, or Reference calls.
Chromosome	Filters data to show all chromosomes (default), autosomal chromosomes, or a specific chromosome number.
Advanced	Filters data based on selections that you make in the Advanced Filter window.

Advanced Filter Options

Use the Advanced filter options to create a multi-branched Boolean expression for filtering data in the Variants table. As you build the advanced filter, a diagram appears to illustrate the filter and branches in the expression.

1 Select the checkbox labeled **Use Advanced Filter** and then click **Edit Filter**. The Create Advanced Filter window opens.

Figure 24 Create Advanced Filter Window



2 Select a column heading from the list on the left-hand side.

3 Select an operation from the Operation list.

4 Select either **Constant** or **Parameter**.

- To filter on a constant, enter a constant associated with the selection from the left-hand column.
- To filter on a parameter, select a column heading from the list on the right-hand side.

5 Click the generate filter ==> button. A diagram of the filters appears.

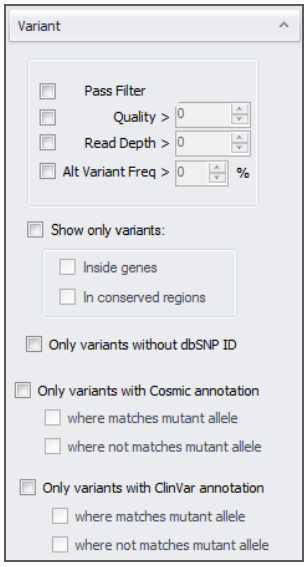
6 To add another branch to the advanced filter, select the radio button for either **and**, **or**, **xor** (exclusive). Then, click **Add**. A new branch is added to the diagram.

7 Continue selecting options and operators until you have completed the filter.

8 When the advanced filter is complete, click **OK**.

9 From the Filters pane, click **Apply Filters**.

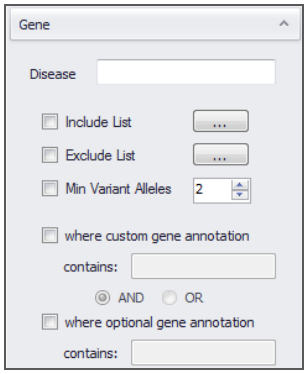
Variant Filters



Use the Variant filters to filter by variant call attributes, variant positions, and variants with specific annotation.

Filter Name	Setting Description
Variant Call	Filters data based on a specified value for variant call quality: pass filter, quality score, read depth, or percentage of variant frequency for the minor allele. Select the checkbox, and then use the up/down arrows to specify a minimum threshold.
Show only variants	Filters data based on variant position. Options include inside genes and in conserved regions.
Only variants with	Filters data based on the source of annotation. Options include variants without dbSNP ID, with COSMIC annotation, and with ClinVar annotation. COSMIC and ClinVar annotation enable two more choices: • If ***where matches mutant allele*** is selected and variant has multiple records, only pass variant if at least one matches. • If ***where not matches mutant allele*** is selected and variant has multiple records, only pass variant if none matches.

Gene Filters



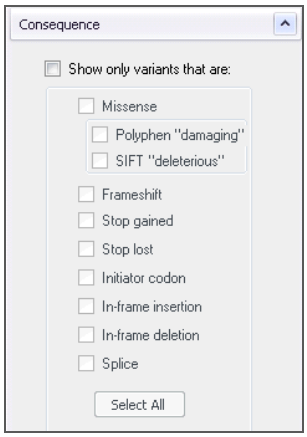
Use the Gene filters to filter data by disease, or include or exclude specific genes.

Filter Name	Setting Description
Disease	Filters data to show genes associated with the specified disease. Enter the disease name. This field is not case-sensitive.
Include List	Filters data to include specified genes. To include genes, click the ... button to open the gene list field next to the Include List options, and enter the gene name. This field is not case-sensitive.
Exclude List	Filters data to exclude specified genes. To exclude genes, click the ... button to open the gene list field next to the Exclude List options, and enter the gene name. This field is not case-sensitive.

Filter Name	Setting Description
Min Variant Alleles	Filters data to show only variants that overlap genes with the specified number of variant alleles. A homozygous variant counts as two variant alleles, while a heterozygous variant counts as one variant allele.
Custom and Optional Gene Annotation	Filters data to show only genes with as specified custom annotation. You can use Boolean logic (AND and OR) to between the Custom and Optional Gene Annotation filters. The Optional Gene Annotation filter is based on the optional second annotation column that you can import.

NOTE
If you click **Clear Filters**, the gene list is also cleared. To save a gene list, create a favorite filter. For more information, see *Create Favorite Filters* on page 55.

Consequence Filters



Use the Consequence filters to filter data by variants that alter the coding potential of the transcript.

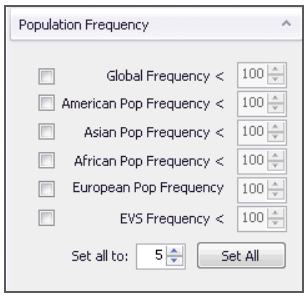
1 Select the checkbox **Show only variants that are**.

2 Select the checkbox for each individual consequence setting or click **Select All**.

Filter Name	Setting Definition
Missense	A single base pair substitution that results in the translation of a different amino acid at that position. Note: PolyPhen and SIFT report only SNVs.
PolyPhen, damaging	A prediction of a damaging effect of an amino acid substitution on the function of a human protein based on PolyPhen.
SIFT, deleterious	A prediction of a deleterious effect of an amino acid substitution on the function of a human protein based on SIFT.

Filter Name	Setting Definition
Frameshift	An insertion or deletion involving a number of base pairs that are not a multiple of three, which disrupts the triple reading frame.
Stop gained	The gain of a stop codon in the coding sequence.
Stop loss	The loss of a stop codon in the coding sequence.
Initiator codon	A codon that acts as a start signal for the synthesis of a protein.
In-frame insertion	An insertion that does not alter the reading frame as a result of the insertion.
In-frame deletion	A deletion that does not alter the reading frame as a result of the deletion.
Splice	An insertion, deletion, or substitution that occurs in a splice region of the gene. A splice is not in a coding region.

Population Frequency Filters



Use the Population Frequency filters to filter data based on the allele frequency in population studies.

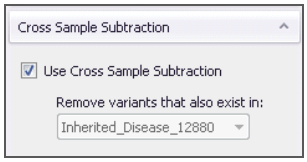
Options include global frequency, American, Asian, African, European, and EVS.

- American, Asian, African, and European are allele frequency from 1000 Genomes.
- EVS is allele frequency from the NHLBI exome sequencing project.

1 Select the checkbox and then use the up/down arrows to specify a value expressed as percentage.

2 To set the same value to all populations, use the up/down arrows in the **Set all to** field. Click **Set All**.

Cross Sample Subtraction Filter



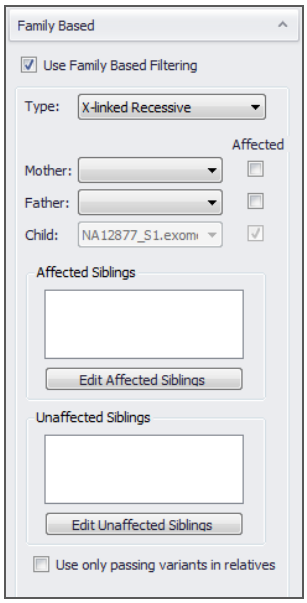
If multiple samples are present in the project, use the cross sample filter to exclude variants that are also present in another sample.

1 Select the checkbox **Use Cross Sample Subtraction**.

2 From the drop-down list, select a sample in the project. Only one sample can be selected as the cross sample filter.

This filter is helpful when filtering variants present in tumor-normal samples.

Family Based Filter



Use the Family Based filter to filter for variants that are consistent with user-specified inheritance mode and provided variant data for available family members. The Family Based Filter requires input of at least one parent or sibling. This filter is useful in identifying candidate disease causing variants.

1 Select the checkbox **Use Family Based Filtering**.

2 Using the Type drop-down list, select a type from the following choices:

- **X-linked Recessive**—Variant-level filtering of heterozygous variants in affected females that are not present in the father and hemizygous in affected males.
- **Autosomal Recessive transmission**—Gene-level filtering of different heterozygous variants in the same gene in relatives, or variant-level filtering of the same heterozygous variants in both parents.
- ***De novo* mutation**—Filters variants not present in the relatives. This filter can also be applied using the cross-sample subtraction filter.
- **Autosomal Dominant transmission**—Variant-level filtering of heterozygous variants that are present in the affected relatives, and not present in the unaffected relatives. This filter requires that you indicate the affected relatives.

3 With the child sample set as current, use the drop-down lists to select at least one parent or sibling. All samples to be used in the family-based filtering must be present in the current project.

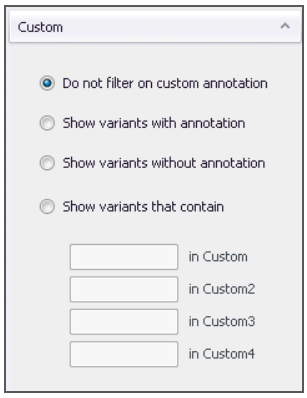
For more information, see *Family-Based Filtering Workflows* on page 49.

Best Practices When Using the Family Based Filter

- When using gVCF files for family-based filtering, variants that were not called in the parents are included if the variants are otherwise consistent with the selected inheritance mode.

- If you have a gVCF file, use a gVCF viewer such as the Integrative Genomics Browser (IGV) to examine the no coverage regions. Check for the presence of a disease gene of interest in samples from the child and other family members. For more information, see www.broadinstitute.org/igv/.

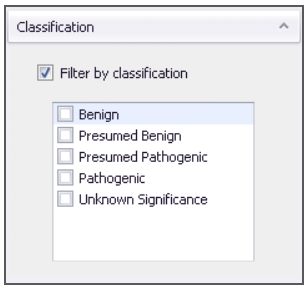
Custom Filters



Custom filters enable filtering based on input provided in the custom annotations input file. For more information, see *Create Custom Annotations* on page 29.

Filter Name	Setting Description
Do not filter on custom annotation	Turns off custom annotations. This setting is on by default.
Show variants with annotation	Filters data to show variants that match criteria provided in the custom annotations input file with an assigned annotation value in the annotations column.
Show variants without annotation	Filters data to show variants that match criteria provided in the custom annotations input file without an assigned annotation value in the annotations column.
Show variants that contain	Filters data to show variants that match criteria provided in the custom annotation input file. Options include annotations from any of the four possible annotation columns.

Classification Filter



Use the classification filter to filter by classifications assigned in the classification database. Any customized classifications appear in the classification filters list.

1 Select the **Filter by classification** checkbox.

2 Select the checkbox next to any number of available classifications.

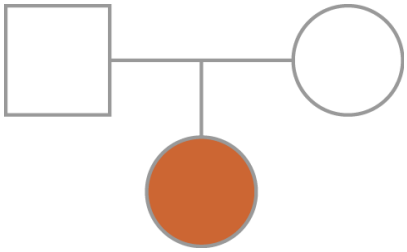
Family-Based Filtering Workflows

The mode of inheritance, which is the inheritance pattern of a genetic trait or disorder as passed down through generations, is typically one of the following:

- Autosomal recessive
- Autosomal dominant
- X-linked recessive
- *De novo* mutation

Disease-causing variants co-exist with the disorder according to the mode of inheritance. Family-based filtering requires at least two samples, the affected person, also known as the proband, and at least one parent or sibling.

Figure 25 Example: Father, Mother, and Proband



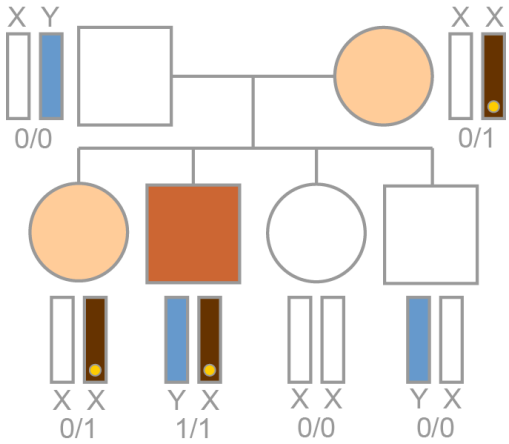
- Affected

Because the proband contains thousands of variants that appear deleterious, filtering is necessary to remove variants that are not disease-causing and identify disease-causing variants.

X-Linked Recessive Transmission Workflow

- A variant is on the X chromosome
- The variant is heterozygous (0/1) in the mother
- The variant is not present in the father
- The variant is homozygous (1/1) in the affected child

Figure 26 X-Linked Recessive Transmission Logic



- Unaffected reference
- Unaffected carrier
- Affected
- Mutation

X-Linked Recessive Transmission Workflow

Proband

↓ Subtract

All variants not on X

↓ Subtract

All variants that are not heterozygous (0/1) in mother

↓ Subtract

All variants that are homozygous (1/1) in father

↓ Subtract

All variants that are not homozygous (1/1) in affected siblings

↓ Subtract

All variants that are homozygous (1/1) in unaffected siblings

↓

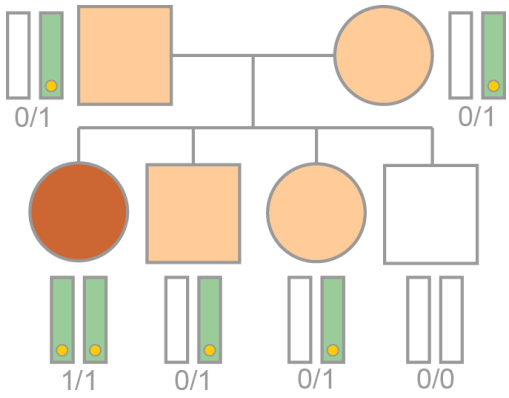
Filtering results: Deleterious variants

Autosomal Recessive Transmission Workflow

There are two possibilities for recessive transmission.

1 A single gene contains a variant that is:
- Heterozygous (0/1) in the mother
- Heterozygous (0/1) in the father
- Homozygous (1/1) in the affected children

Figure 27 Autosomal Recessive Transmission Logic #1



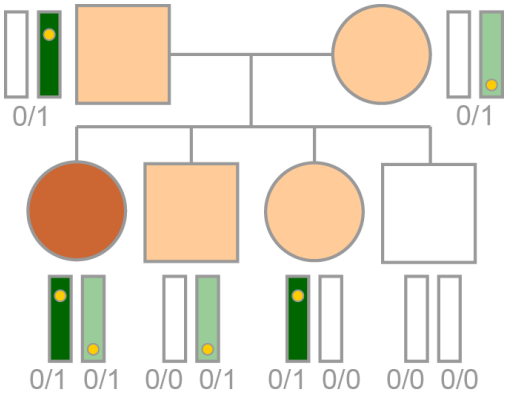
- Unaffected reference
- Unaffected carrier
- Affected
- Mutation

Autosomal Recessive Transmission Workflow #1

Proband
↓ Subtract
All variants that are homozygous (1/1) in the father, mother, and unaffected siblings
↓ Include
Homozygous (1/1) variants in the child that are heterozygous (0/1) in the mother and father
↓ Subtract
All variants that are not homozygous (1/1) in affected siblings
↓ Subtract
All variants that are homozygous (1/1) in unaffected siblings
↓
Filtering results: Deleterious variants

2 A single gene contains:
- One variant that is heterozygous (0/1) in the mother
- The same gene contains a different variant that is heterozygous (0/1) in the father
- Both variants are present in the affected child (0/1 and 0/1)

Figure 28 Autosomal Recessive Transmission Logic #2



Autosomal Recessive Transmission Workflow #2

Proband

↓ Subtract

All variants that are homozygous (1/1) in the father, mother, and unaffected siblings

↓ Include

Compound heterozygous (0/1) variants, if at least two variants are in the same gene, at least one variant is heterozygous in the father, and other variants are heterozygous in the mother.

↓ Subtract

All variants that are not compound heterozygous (0/1) in affected siblings

↓ Subtract

All variants that are compound heterozygous (0/1) in unaffected siblings

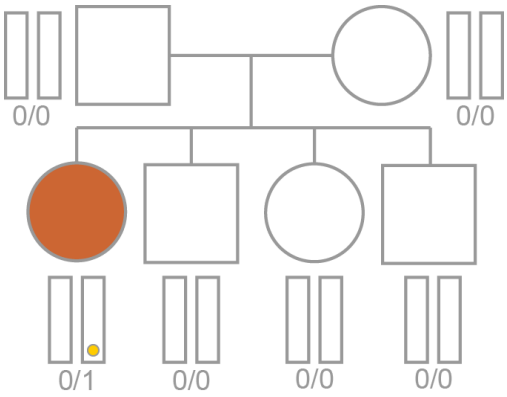
↓

Filtering results: Deleterious variants

De Novo Mutation Workflow

- A variant is present (0/1 or 1/1) in the proband
- The variant is not present (0/0) in either parent or siblings
- Only one child in the family is affected

Figure 29 *De Novo* Mutation Logic



- Unaffected reference
- Unaffected carrier
- Affected
- Mutation

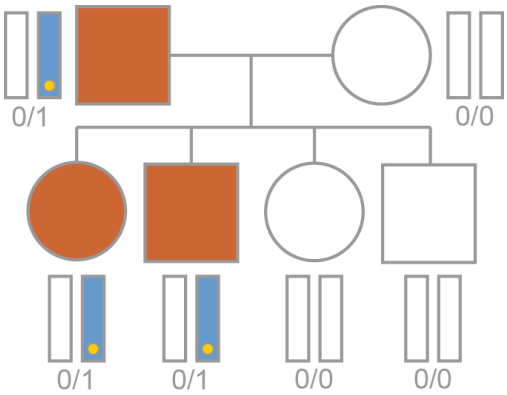
De Novo Mutation Workflow

Proband
↓ Subtract
All variants that are heterozygous (0/1) or homozygous (1/1) in the mother, father, and unaffected siblings
↓
Filtering results: Deleterious variants

Autosomal Dominant Transmission Workflow

- A variant is heterozygous (0/1) in the affected parent
- The variant is not present (0/0) in the unaffected parent
- The variant is heterozygous (0/1) in the affected children

Figure 30 Autosomal Dominant Transmission Logic



- Unaffected reference
- Unaffected carrier
- Affected
- Mutation

Autosomal Dominant Transmission Workflow

Proband

↓ Subtract

All variants that are heterozygous (0/1) in unaffected parent or siblings

↓ Subtract

All variants that are not heterozygous (0/1) in affected parent or siblings

↓

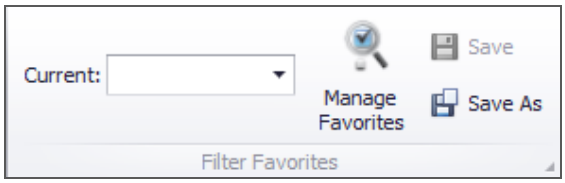
Filtering results: Deleterious variants

Create Favorite Filters

To save any combination of filtering options for use with a different sample or for later use in another project, save the filtering options as a favorite filter.

The Filter Favorites menu includes commands to save, apply, modify, and manage saved filters.

Figure 31 Filter Favorites Menu Commands



Command	Description
Current	Shows the current filter that is applied and a list of available saved filters. Select a saved favorite filter from the drop-down list to apply it to the current sample. If you change to another sample, a favorite filter applied to the previous sample is not applied automatically to the next sample.
Manage Filters	Opens tools for renaming, duplicating, or deleting saved filters.
Save	Saves changes to the currently applied filter.
Save As	Opens a dialog box for naming a favorite filter.

Save a Favorite Filter

1. With any combination of filters specified in the Filters pane, select **Apply Filter**.
2. Click **Save As**  in the Filter Favorites menu.
3. Enter a name for the new filter. Click **OK**. When a saved filter is applied, the saved filter name appears in the Current field.

Apply a Favorite Filter

1. To apply a saved filter, expand the Current field drop-down list.
2. Select a filter name from the list. The filter is applied automatically.
3. To change to another saved filter, expand the drop-down list in the Current field, and select a different filter name.
4. Alternatively, click the blank entry at the top of the saved filters list to remove the currently applied filter. The variants table is restored to an unfiltered view.

NOTE
The favorite filter is not automatically applied when you move to another sample in your project. To apply a favorite filter, reselect the favorite filter name from the Current drop-down list.

Modify a Favorite Filter

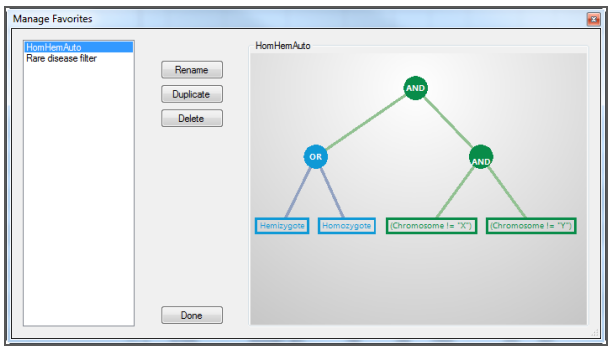
1 Select additional filtering options from the Filters pane, and click **Apply Filter**. An asterisk appears next to the saved filter name, which indicates that changes have been applied while the saved filter was selected.

2 To modify the saved filter with the applied filtering options, click **Save** in the Filter Favorites menu. The selected saved filter is modified to include the additional filtering options.

Manage Favorite Filters

To rename, duplicate, or delete favorite filters, use the Manage Favorites feature.

Figure 32 Managing Favorite Filters



1 Click **Manage Favorites**. Names of saved filters appear on the left panel and a block diagram of the selected filter appears on the right panel. To adjust the view of the block diagram, click anywhere on the right panel and use the scrolling feature on your mouse to zoom in or zoom out.

2 Click **Done** to apply changes.

Generating Reports

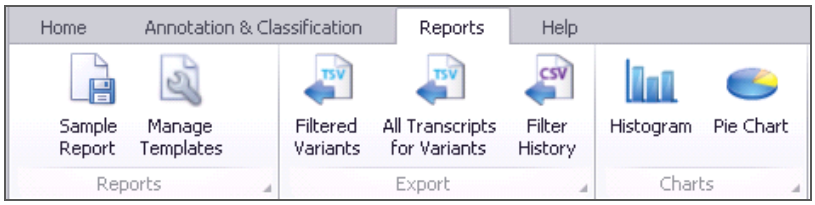
Chapter 4

Introduction

BaseSpace VariantStudio provides tools to export results from a project to an external report.

Use the commands on the Reports tab to create a sample report, and to export data to text files and graphical representations.

Figure 33 Reports Tab



Sample Report Overview

The sample report consists of five sections plus a footer and is generated as a PDF file or RTF file, depending on your preference. A sample report includes the following sections:

- **Lab information**—Typically, this section is defined in the template and appears as a header in the sample report.
- **Sample information**—This section contains details about the sample and appears as a two-column table in the sample report. The first column contains the field name and the second column contains the value. There are two ways to populate this section of the report:
  - Specify field names using the Manage Templates feature and then manually enter the value in the text fields.
  - Import sample information from an external text file. The text file must have two tab-delimited or comma-separated columns, one for the field name and one for the field value, and use a *.txt, *.csv, or *.tsv extension.
- **Test summary**—This section is reserved for a description of the test performed. Set up preferred content and formatting in the template. Add information specific to the report when you create the report.
- **Results**—This section lists variants in the open project that have a classification assigned. This section is blank if the project does not contain assigned classifications. Results are formatted in a four-column table with headings of Gene, Variant, Classification, and Details. Information in the Details column comes from the Report Fragment field in the classification database.

Figure 34 Example of a Results Section

Gene	Variant	Classification	Details
C3	NM_000064.2:c.941C>T	Benign	
GHRHR	NM_000823.3:c.169G>A	Benign	
GHRHR	NM_000823.3:c.363G>T	Benign	
ABCC9	NM_020297.2:c.1165-6delT	Presumed Benign	
C3	NM_000064.2:c.304C>G	Presumed Benign	
APOE	NM_000041.2:c.388T>C	Pathogenic	
CASP10	NM_032977.3:c.1228G>A	Pathogenic	
CLCN1	NM_000083.2:c.2680C>T	Pathogenic	
NCF1	NM_000265.4:c.73_74delGT	Pathogenic	
POLG	NM_002693.2:c.1399G>A	Pathogenic	
SLC4A1	NM_000342.3:c.166A>G	Pathogenic	
BCHE	NM_000055.2:c.1699G>A	Unknown Significance	
MEFV	NM_000243.2:c.1772T>C	Unknown Significance	

- **Methodology**—This section is reserved for a description of the methodology specific to the report. Set up preferred content and formatting in the template. Add information specific to the report when you create the report.
- **References**—This section is reserved for references applicable to the contents of the report.
- **Page footer**—Typically, this section is defined in the template. For example, the footer can contain the facility address and contact information, or it can be blank.

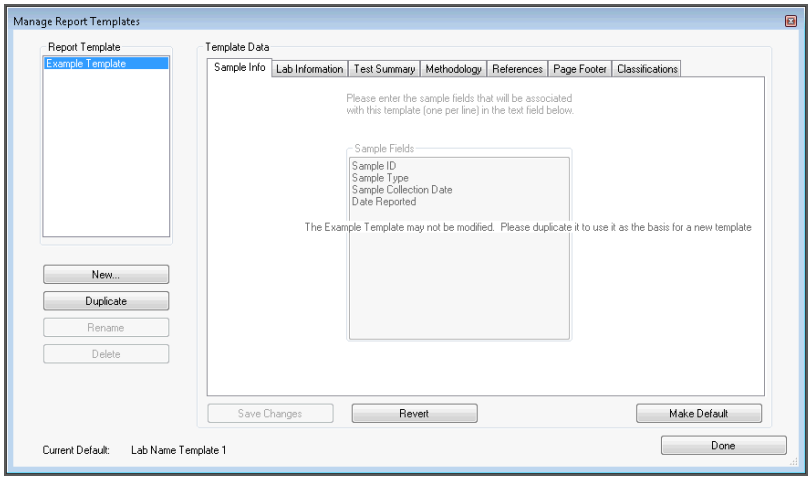
Create a Sample Report Template

Use the Manage Templates feature to create a customized template for sample reports. After templates are created, use Manage Templates to duplicate, edit, rename, or delete templates in the template library.

BaseSpace VariantStudio includes an example report template to help in creating a template. The example report template cannot be edited. Instead, create a copy of the example report template. From this copy, rename the template and customize each section of the template using the template tabs.

1 Click **Manage Templates**. The Manage Report Templates window opens.

Figure 35 Manage Report Templates Window



2 Do one of the following:

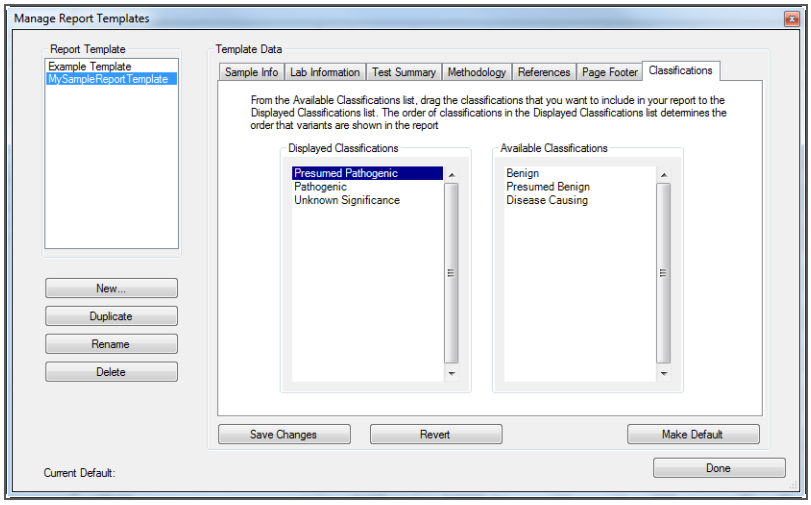
- Highlight Example Template in the Report Template field and click **Duplicate** and enter a template name. Click **OK**.
- To create a template without using the example template, click **New** and enter a template name. Click **OK**.
- To edit an existing template, click to highlight the template name in the Report Template list. This template is now the active template and ready for editing.

3 For each of the following tabs, enter the information to be included in reports using this template. Use the formatting tools to customize the layout. Information included in the following sections of the template are editable when creating the sample report.

- **Sample Info tab**—Specify the sample fields to include in the report. Each field name generates a row in the sample information table.
- **Lab Information tab**—Enter the lab name and location, or other preferred information for the report header.
- **Test Summary tab**—Enter preferred introductory content to begin this section. Otherwise, leave this section blank in the template.
- **Methodology tab**—Similar to Test Summary, enter preferred introductory content to begin this section. Otherwise, leave this section blank in the template.
- **References tab**—Similar to Test Summary, enter preferred introductory content to begin this section. Otherwise, leave this section blank in the template.
- **Page Footer tab**—Enter preferred content for the template footer, such as contact information. This information appears at the bottom of each page in the report.

4 Click the **Classifications** tab. Drag and drop classification names from the Available Classification list to the Displayed Classification list. The selected classifications are included in any reports using this template and they appear in the order listed.

Figure 36 Classifications for Reporting



5 Click **Save Changes**.

6 (Optional) With the template name highlighted in the Report Template list, click **Make Default**. The current default is listed in the lower-left corner of the Manage Report Templates window.

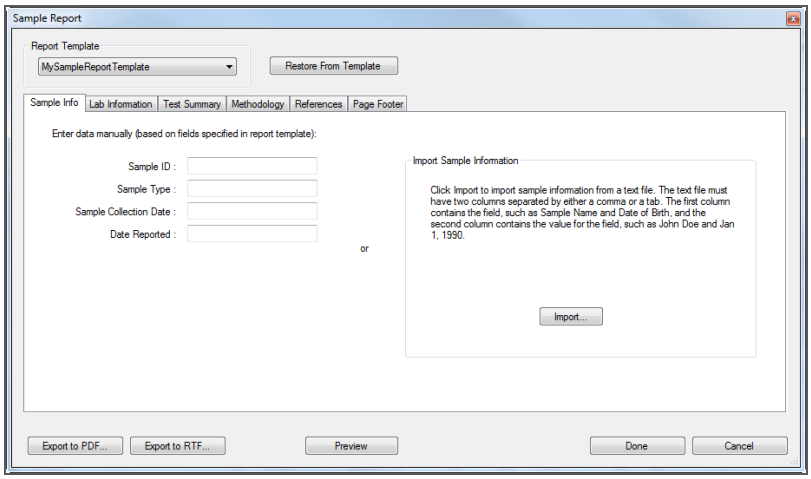
7 Click **Done**. The Manage Report Templates window closes.

Create a Sample Report

Before proceeding, consider creating a template using the Manage Templates feature. For more information, see *Create a Sample Report Template* on page 60.

1 From the Reports menu, click **Sample Report**. The Sample Report window opens.

Figure 37 Sample Report, Sample Info Tab



2 From the Report Template drop-down list, select an appropriate template for the report.

3 On the Sample Info tab, enter information in the fields provided or click **Import** to browse to the location of the text file containing the information.

4 Enter information for the remaining tabs that are not already populated in the selected template.

5 Click **Preview** to preview the report before generating it.

6 Click **Export to PDF** or **Export to RTF** to generate the report.
To save the report contents without generating the report, click **Done**.

Export Text Files and Charts

In addition to sample reports, BaseSpace VariantStudio provides tools for exporting to text files and graphical representations of data.

Export Data Files

Exporting filtered variants and all transcripts for variants generates a tab-separated values file. Exporting filter history generates a comma-separated values (CSV) file. These text file formats are not application-specific and can be opened in any text editor. The ANT file is a binary file that contains

Command	Description
Filtered Variants (TSV)	Exports filtered variants from the current sample. For variants that overlap multiple genes, only the transcripts that appear on the interface are exported.
All Transcripts for Variants (TSV)	Exports all transcripts for filtered variants in the sample.
Filter History (CSV)	Exports a report of all filters applied to the project.

Export Charts

From the Charts menu, select a preferred format to export results in a histogram or a pie chart.

Command	Description
Histogram	Generates a histogram of filtered results from the Variants table. • Use the Plot drop-down list on the generated histogram (Figure 38) to represent variant quality values or indel variant length. • Use Data Source options to show all variants or only filtered variants.
Pie Chart	Generates a pie chart of filtered results from the Variants table. • Use the Plot drop-down list on the generated pie chart (Figure 39) to represent percentages of variant call filters, variant type, genotype, coding regions, or filtered variants. • Use Data Source options to show all variants or only filtered variants.

From the generated chart, click **Copy to Clipboard** to transfer the image from the BaseSpace VariantStudio software to an application that supports images.

Figure 38 Histogram

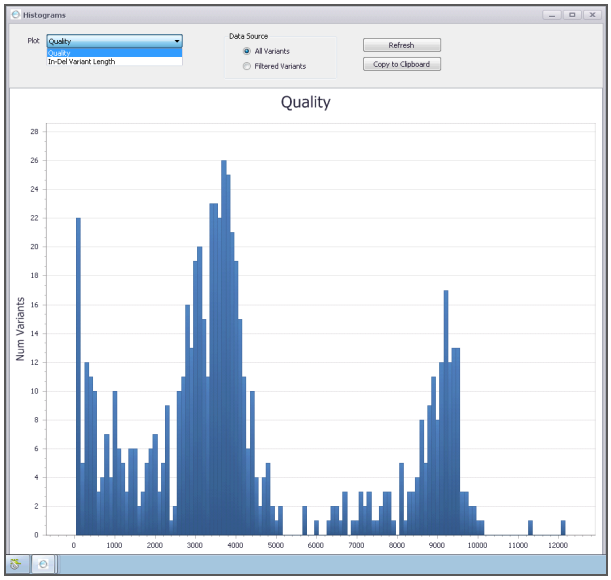
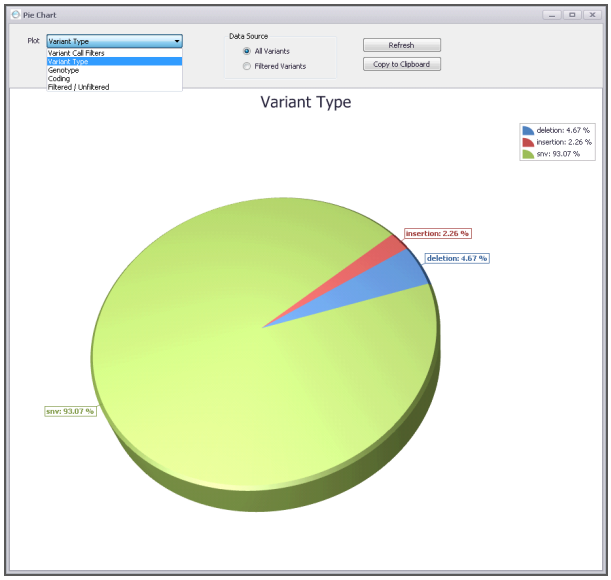


Figure 39 Pie Chart



Annotation Sources

Annotation Sources

Annotation sources are static in the BaseSpace VariantStudio software. Any changes to source databases, such as ClinVar, do not automatically update the annotation sources connected to BaseSpace VariantStudio.

The BaseSpace VariantStudio software includes the following annotation sources:

- Variant Effect Predictor (VEP)
- 1000 Genomes Project
- Catalogue of Somatic Mutations in Cancer (COSMIC)
- ClinVar
- National Center for Biotechnology (NCBI)
- National Heart, Lung, and Blood Institute (NHLBI) Exome Variant Server
- UCSC

VEP v2.8

- Source:
  - Uses data from the Ensembl infrastructure (release 72) ftp.ncbi.nih.gov/snp/organisms/human_9606/VCF/00-All.vcf.gz
  - Ensembl infrastructure pulls information from RefSeq (release 56)
- Values:
  - Positional (specific to the position, not necessarily matching the allele):
    - GMAF—Global minor allele frequency www.ncbi.nlm.nih.gov/variation/tools/reporter/docs/faq#gmaf
    - GMAF allele
    - HGNC of overlapping transcripts (for Ensembl only) For RefSeq, see *NCBI* on page 67.
  - Transcript specific:
    - Feature
    - Feature Type
    - Consequence
    - cDNA Position
    - CDS Position
    - Protein Position
    - Amino Acids
    - Codons
    - Exon
    - Intron
    - HGNC—Direct from Ensembl for ENST For RefSeq, same as NCBI.
    - Distance
    - Canonical
    - Sift
    - PolyPhen
    - ENSP
    - Domains
    - CCDS—Reported for Ensemble annotations only
    - HGVSc
    - HGVSp
  - Positional—Specific to the position, not necessarily matching the allele:

- Feature ID
- Feature Type
- Consequence
- Motif Name
- Motif Position
- High Influence Position
- Motif Score Change
- Cell Type

1000 Genomes (April 2012 v3)

- Source:
  - ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/20110521/ALL.wgs.phase1_release_v3.20101123.snps_indels_sv.sites.vcf.gz
- Values:
  - Ancestral Allele [AA]
  - Global allele Frequency [AF]
  - Population allele frequencies: (www.1000genomes.org/category/frequently-asked-questions/population)
    - African [AFR_AF]
    - Ad Mixed American [AMR_AF]
    - East Asian [ASN]
    - European [EUR]

COSMIC (v65)

- Source:
  - ngs.sanger.ac.uk/production/cosmic/*_noLimit.vcf.gz

  Additional annotation for COSMIC entries were obtained from ftp.sanger.ac.uk/pub/CGP/cosmic/data_export/CosmicCompleteExport_*.tsv.gz
  From the CosmicCompleteExport data file, the following four fields were added:
  - primary_site
  - site_subtype
  - primary_histology
  - histology_subtype

ClinVar

- Version September 5, 2013
- www.ncbi.nlm.nih.gov/clinvar/

NCBI

dbSNP (v137)

- Source:
  - ftp.ncbi.nih.gov/snp/organisms/human_9606/VCF/00-All.vcf.gz
- Values:
  - rsID

RefSeq Transcript ID and **HGNC**

- Sources:
  - ftp.ncbi.nlm.nih.gov/gene/DATA/gene_info.gz
  - ftp.ncbi.nlm.nih.gov/gene/DATA/gene2refseq.gz
- Values:
  - Transcript ID and gene name are mapped as follows: Transcript ID > Gene ID > HGNC, using the NCBI unique Gene ID to link the two databases.

NHLBI Exome Variant Server

- Version ESP6500SI-V2 (updated June 7, 2013)
- Source: evs.gs.washington.edu/EVS/
  - ESP6500SI-V2-SSA137.dbSNP138.snps_indels.vcf.tar.gz
  - ESP6500SI-V2.coverage.all_sites.txt.tar.gz
- Values:
  - For SNVs and indels—Allele frequency computed from the TAC field, which reports alternate alleles observed and reference alleles observed
  - From the all_sites file (identifies captured positions, even if a variant is not present)
    - TotalSamplesCovered
    - AvgSampleReadDepth

UCSC (hg19)

Placental mammalian phastCons elements (downloaded from UCSC table browser)

Index

Technical Assistance

For technical assistance, contact Illumina Technical Support.

Table 1 Illumina General Contact Information

Illumina Website	www.illumina.com
Email	techsupport@illumina.com

Table 2 Illumina Customer Support Telephone Numbers

Region	Contact Number	Region	Contact Number
North America	1.800.809.4566	Italy	800.874909
Austria	0800.296575	Netherlands	0800.0223859
Belgium	0800.81102	Norway	800.16836
Denmark	80882346	Spain	900.812168
Finland	0800.918363	Sweden	020790181
France	0800.911850	Switzerland	0800.563118
Germany	0800.180.8994	United Kingdom	0800.917.0041
Ireland	1.800.812949	Other countries	+44.1799.534000

Safety Data Sheets

Safety data sheets (SDSs) are available on the Illumina website at support.illumina.com/sds.ilmn.

Product Documentation

Product documentation in PDF is available for download from the Illumina website. Go to support.illumina.com, select a product, then click **Documentation & Literature**.

Illumina
Headquartered in San Diego, California, U.S.A.
+1.800.809.ILMN (4566)
+1.858.202.4566 (outside North America)
techsupport@illumina.com
www.illumina.com