

Filter by SIFT Synonymous Classification

Author: Gabe Rudy, Golden Helix, Inc.

Overview

This filter inactivates mapped markers that are either predicted as synonymous or are predicted as nonsynonymous, depending on the inactivation option selected.

Note: This script has been deprecated by **Filter on NS Functional Predictions**, shipped in SVS version 7.6.7. SIFT filtering can now be applied simultaneously with other filtering tools such as PolyPhen2 and MutationTaster.

Recommended Directory Location

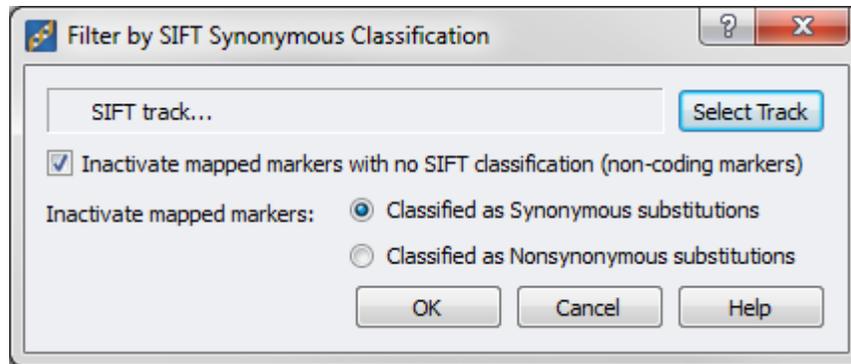
Save the script to the following directory:

*..\Application Data\Golden Helix SVS\UserScripts\Spreadsheet\Select\Filter by Annotation

Note: The **Application Data** folder is a hidden folder on Windows operating systems and its location varies between XP and Vista. The easiest way to locate this directory on your computer is to open SVS and go to **Tools > Open Folder > User Scripts Folder**. If saved to the proper folder, this script will be accessible from the spreadsheet **Select > Filter by Annotation** menu.

Using the Script

1. From a spreadsheet with a marker map applied to the columns, go to **Select > Filter by Annotation > Filter by Synonymous Classification**.
2. Select the SIFT track (required to be a local track).
 - a. If the SIFT track has not yet been downloaded to your local machine, download it from the network using the annotation track manager.
3. Specify whether mapped markers with no SIFT score (non-coding markers) should be inactivated or not.
 - a. If no SIFT score is available at the loci of a given marker, then it is not in a coding region (exon). These markers can be inactivated as well or left as active.
4. Indicate which markers should be inactivated, either:
 - a. Those classified as synonymous (removes markers predicted to have no effect on protein sequence).
 - b. Those classified as nonsynonymous (removes markers predicted to have some effect on protein sequence).



In the original spreadsheet, the columns that meet the inactivation criteria are inactivated. The other columns are left active.

A marker mapped filtering results spreadsheet is also created as a child of the original spreadsheet. This spreadsheet contains the following columns:

- **Drop?** Indicates whether the marker is dropped based on the specified filtering criteria.
- **Coding?** Indicates whether the marker is in a coding region or not.
- **Synonymous?** Will be one if the marker is synonymous and zero if the marker is non-synonymous. This will be missing for non-coding markers.
- **Ref/Alt Alleles** The reference and alternate nucleotide bases used for the calculations of the SIFT and median scores.

Unsort		B 1	B 2	B 3	C 4
Map	Probe	Drop?	Coding?	Synonymous?	Ref/Alt Alleles
728	11:103413564	0	1	0	G/A
729	11:103413722	0	1	0	T/C
730	11:110661855	0	1	0	G/C
731	11:110661976	0	1	0	G/A
732	11:111117993	0	1	0	T/C
733	11:111123947	0	1	0	C/A
734	11:111131378	0	1	0	C/A
735	11:111136822	0	1	0	G/T
736	11:111140829	1	1	1	A/G
737	11:111140830	0	1	0	T/A
738	11:111141275	0	1	0	G/A
739	11:111461341	0	1	0	C/T
740	11:113819833	0	1	0	A/C
741	11:113897904	0	1	0	C/A
742	11:113898271	0	1	0	G/A

