Find de Novo Candidate Variants

Author: Autumn Laughbaum, Golden Helix, Inc.

Overview

This tool uses pedigree information to identify candidate functional polymorphism, defined as the offspring in a trio having a genotype classified as a Mendelian error. By default, only heterozygous errors are considered candidates. Optionally, homozygous non-reference errors can be considered and require a reference allele field to be present in the marker map. Another option allows the user to restrict computation to affected offspring.
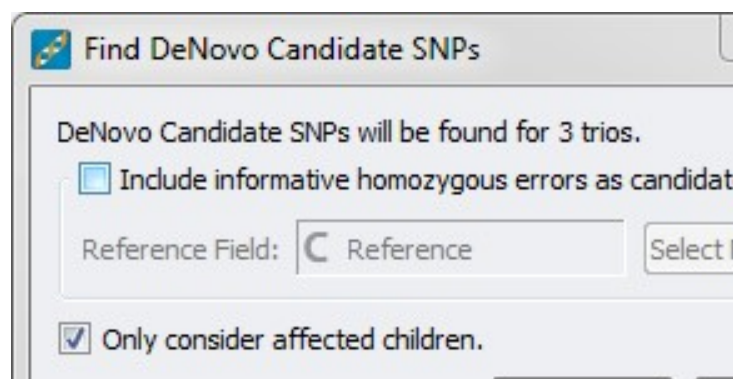
Recommended Directory Location

Save the script to the following directory:
*..**\Application Data\Golden Helix SVS\UserScripts\Spreadsheet\DNA_Seq**

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 select **Tools > Open Folder > UserScripts Folder**. If saved to the proper folder, this script will be accessible from the spreadsheet's **DNA-Seq** menu.

Using the Script

1. Open a spreadsheet containing Pedigree columns and several (mapped) genotypic columns and choose **DNA-Seq > Find de Novo Candidate Variants**.
2. The spreadsheet is then scanned for validity and the number of trios is displayed in the prompt.

3. Choose whether or not to include informative homozygous candidates (requires reference allele field) and whether or not to only consider affected children.
4. Click **OK**.

A new child node is created for each trio found in the original spreadsheet, assuming the trio has candidate variants. A message is added to the log if a trio does not have any candidates. In each new spreadsheet, every genotypic column represents a Mendelian Error (heterozygous and optionally non-reference homozygous) found in the trio's genotypic data. An additional report spreadsheet is created containing all variants that were found to be candidates in any trio. The variants are listed in the row labels and there is a binary column for each trio (with the child's label) signifying if that variant was considered a candidate for the trio.