Golden Helix

varSEQ®

Introduction to VSWarehouse Tutorial

Release 1.4.3

Golden Helix, Inc.

February 21, 2017

CONTENTS

1 Overview 3

2 Connecting to the VSWarehouse 5

3 Adding a Project as an Annotation 9

4 Using a VSWarehouse Assessment Catalogs 11

5 Using a VSWarehouse Report 17

6 Uploading Projects to VSWarehouse 25

7 Exploring VSWarehouse 29

8 Conclusion 35

Updated: February 21, 2017

Level: Advanced

Version: 1.4.3 or higher

Product: VSWarehouse

This tutorial covers the basic VSWarehouse workflow. This tutorial focuses on connecting to a VSWarehouse instance from VarSeq, adding an existing VSWarehouse project as an annotation source, and using reports and assessment catalogs hosted on VSWarehouse.

Important

To perform the workflow described you will need access to a demo VSWarehouse server, in addition to an active VarSeq license. You can go to Discover VarSeq to request an evaluation license, and a demo VSWarehouse server.

Requirements

To complete this tutorial you can use the project created in the Cancer Panel tutorial or you can download a copy of the completed project from the following link.

Download

VSWarehouse_Intro_Tutorial.zip

Files included in the above ZIP file:

- **Cancer Panel Demo** - Starter project matching the project created in the Cancer Panel Tutorial

Note: VarSeq version 1.4.3 was used to create this tutorial. While every attempt will be made to keep this content relevant, it is possible that certain features or icons may change with newer releases.

CHAPTER ONE

OVERVIEW

This tutorial focuses on several of the basic use cases for VarSeq and its integration with a VSWarehouse server. The VSWarehouse server allows for the simple centralized management of the variants and samples previously imported in VarSeq. Using VSWarehouse as part of an existing VarSeq workflow allows users to add previously seen variants to the current project. This can include variants from previous projects, reports or assessment catalogs. These resource are all centrally located allowing for easy management and collaboration.

The typical VarSeq workflow with VSWarehouse can be summarized with the following .

- The variants are imported in VarSeq where they are filtered down to the ones which are relevant.
- Annotation sources from VSWarehouse are added to autofill previous variants or assessments in order to speed up the final interpretation step.
- Once the initial variants have been reduced down to an actionable result set, they can be included in a report that is rendered and stored on the VSWarehouse.
- Additionally variant specific notes can be added to a VSWarehouse assessment catalog.
- Finally once the reporting is finished the project is uploaded to the VSWarehouse where its variants are added to a cohort of samples that underwent similar analysis.

The VarSeq project that is included in the ZIP file that accompanies this tutorial can be used to walk through this tutorial. Alternatively if you have previously completed the Cancer Panel VarSeq tutorial you can use that project to complete this tutorial.

CHAPTER TWO

CONNECTING TO THE VSWAREHOUSE

The first step to using VSWarehouse with VarSeq is setting VarSeq to talk to your VSWarehouse instance. This is completed through the VSWarehouse Manage Dialog in VarSeq.

1. The VSWarehouse manage dialog can be opened by going to **Tools > Manage VSWarehouse...** (as noted in *Figure 2-1*).

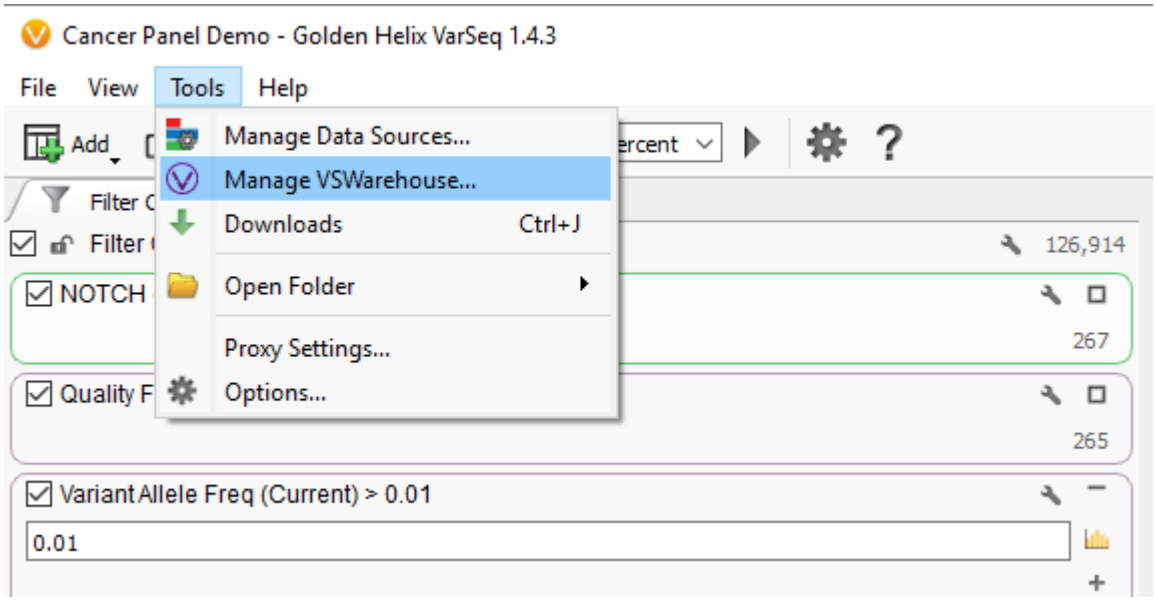


Figure 2-1: The location of the Manage VSWarehouse Dialog

2. The first time that you open this dialog it will provide a splash page with VSWarehouse information, and a URL bar *Figure 2-2*.

   Enter the URL of the warehouse demo server that you have been provide with. And click the **Connect** button. If the connection is successful the splash screen will be replaced with a view with four tabs *Figure 2-3*.

3. The **Projects**, **Reports** and **Catalogs** correspond to the three types of data that you can store on the warehouse. The **Annotations** tab allows you to add those three types of data as annotation sources. Cancer Gene Panel should be one of the projects listed in the **Projects** tab. This is an active project that represents all of the cancer data that we have seen so far for this gene panel.

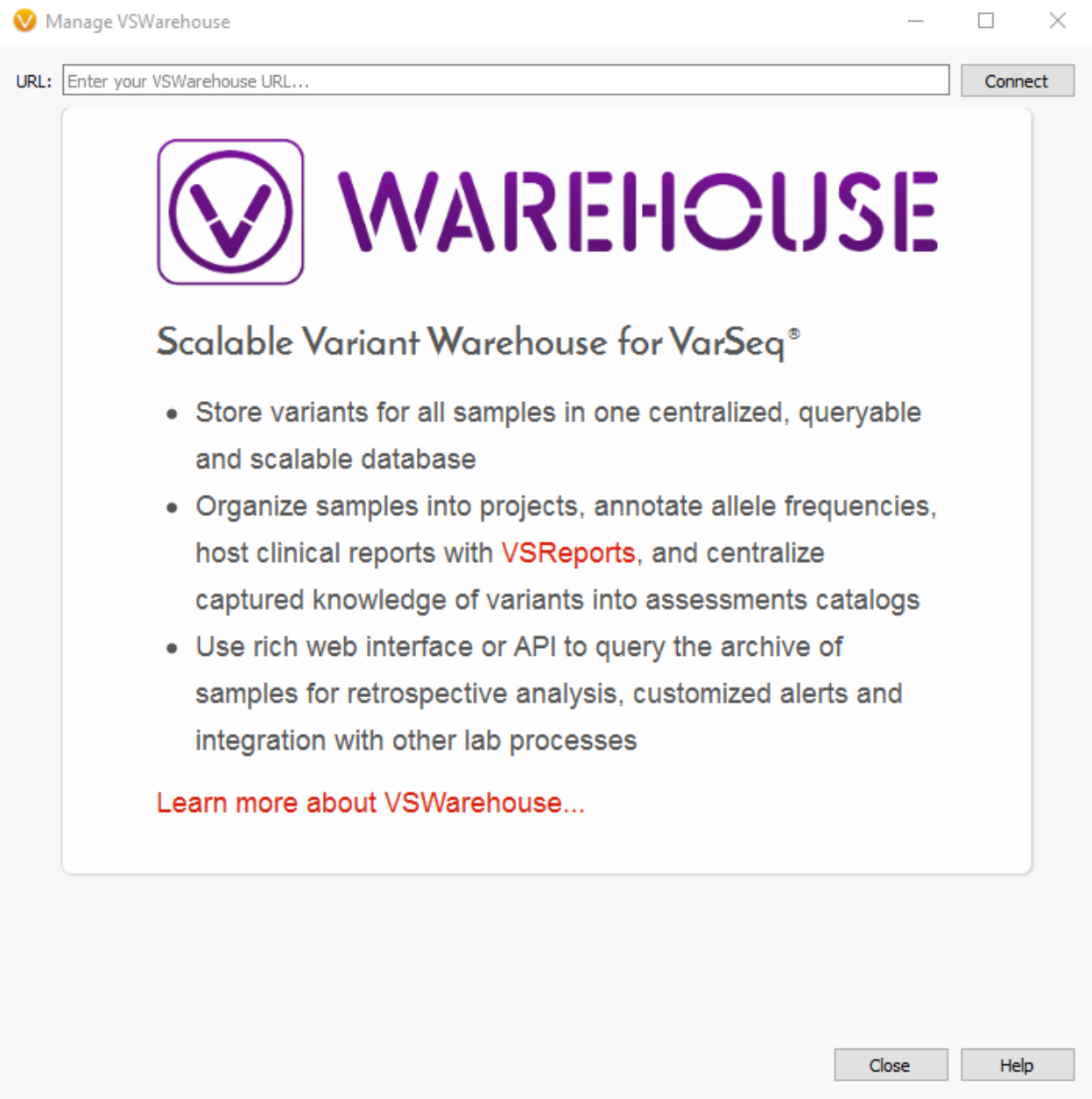


Figure 2-2: The VSWarehouse connection dialog the first time it is opened.

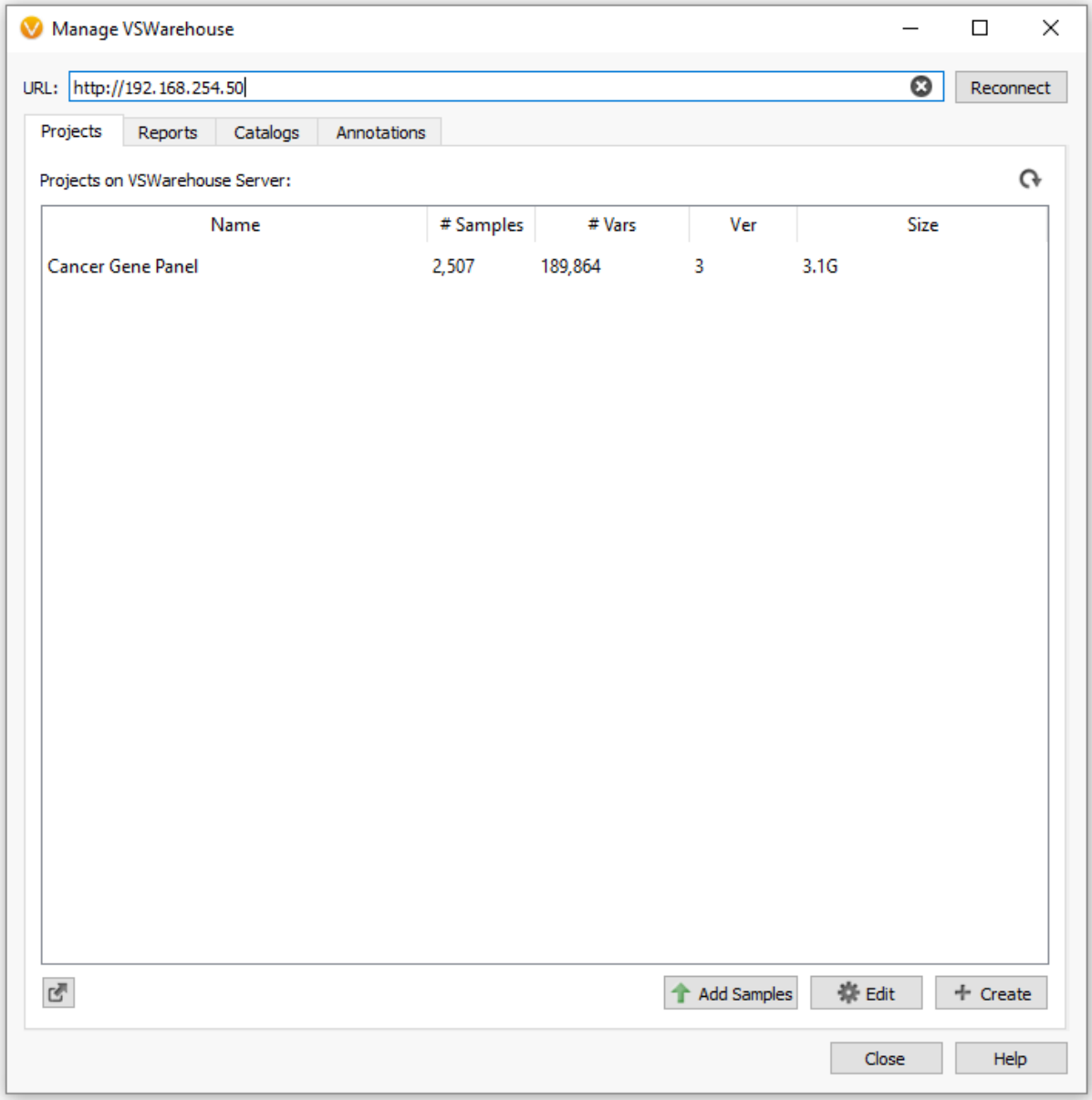


Figure 2-3: The VSWarehouse Manage Dialog after a connection has been established.

CHAPTER THREE

ADDING A PROJECT AS AN ANNOTATION

Any of the Projects, Reports, or Assessment Catalogs that are hosted on the VSWarehouse can be added as annotations to your VarSeq project. Adding a VSWarehouse project as an annotation, allows you to check which of the variants that your current project has in common with the previous projects. The same goes for assessment catalogs and reports which when stored on warehouse keep a record of every variant that has been previously reported along with any interpretation entered by the user.

1. From the VSWarehouse management dialog, switch to the **Annotations** tab. Check the *Cancer Gene Panel* source and add it as an annotation. This allows us to see if any of the variants that were in previous sample sets share variants with the samples in our current project.

   The *Cancer Gene Panel* will then appear in the table like any other annotation source, *Figure 3-2*. As you can see there are two variants that are in the current filtered set that have been seen previously. Both of the allele frequencies of the two matched variants are above 0.5 indicating that they are both common variants.

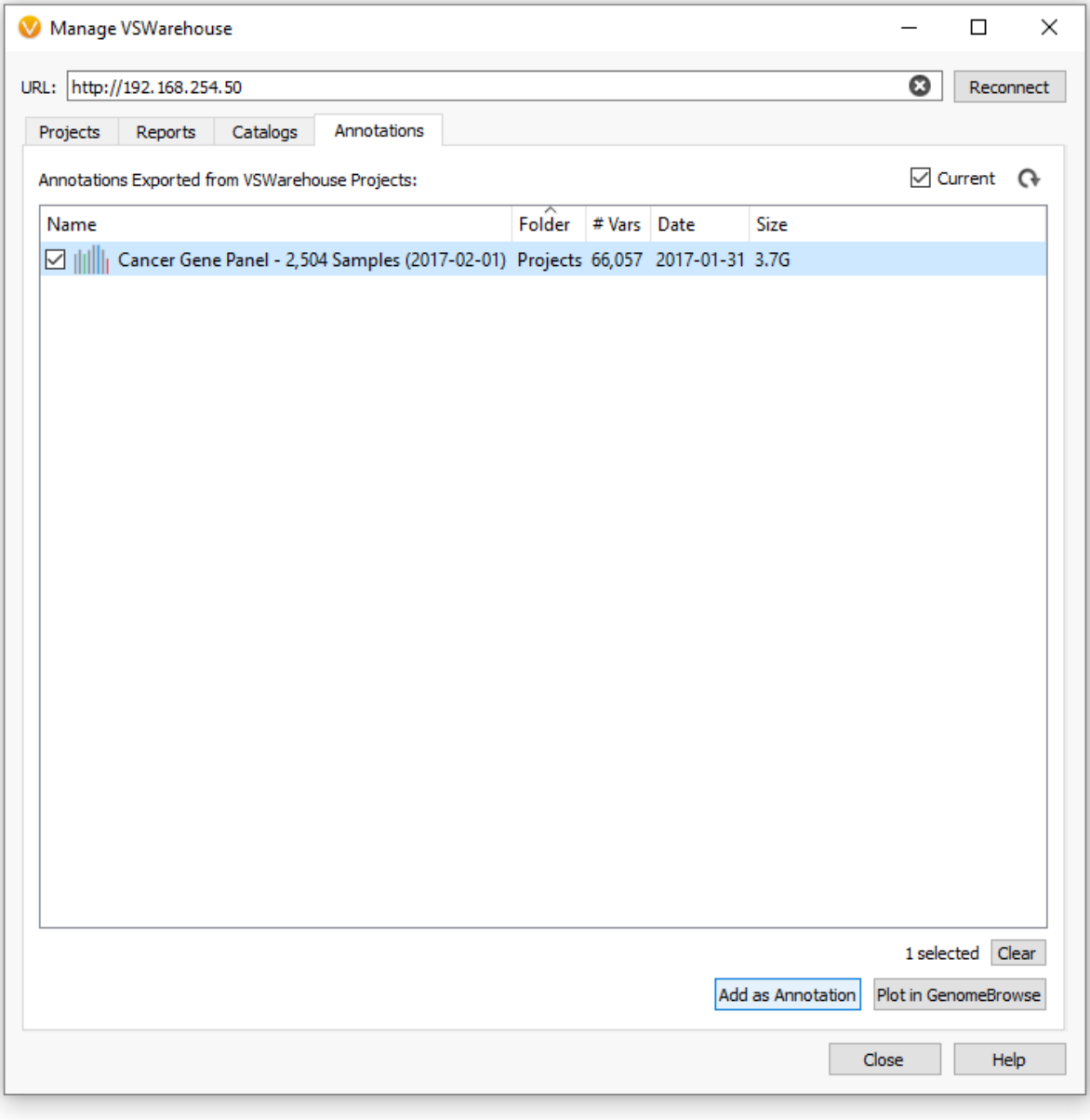


Figure 3-1: The annotations tab shows the VSWarehouse sources that are available to be added as an annotation.

Cancer Gene Panel - 2,504 Samples (2017-02-01)						
Ref/Alt	Project ID	Allele Counts	Allele Frequencies	# Alleles	# Het	# HomoVar
G/A	27933	3328	0.664537	5008	880	1224
G/A	28223	2685	0.536142	5008	1019	833
?	?	?	?	?	?	?

Figure 3-2: The VSWarehouse annotation source in a VarSeq table.

CHAPTER FOUR

USING A VSWAREHOUSE ASSESSMENT CATALOGS

Assessment catalogs are modifiable annotation sources. By hosting an assessment on a VSWarehouse instance it makes it easy to share the changes to each variant in the catalog between multiple users. You can create assessment catalogs with custom fields which provides the flexibility to use them many different ways. The more common use cases center around using them as sets of white-listed or black-listed variants to speed up your variant interpretation. For this tutorial we will be using the *Clinical Significance* catalog that is already on the VSWarehouse.

1. The catalog that we will be using in this tutorial can be found in the **Catalogs** tab of the Manage VSWarehouse tab.

   Click **Open** to create an assessment catalog view in VarSeq.

2. Next switch to the 50 Percent sample by selecting it from the current sample selector. This will update the filters allowing an additional variant into the final results.

   Select the third variant, and notice that the assessment catalog view updates to the record at that position.

3. Fill out the assessment for this variant with a *Pathogenic Classification* and a *Reported Classification* as *Confirmed somatic variant* (see *Figure 4-3*). Click the **Save** button to update the VSWarehouse assessment catalog.

   This assessment will then be placed on VSWarehouse were anyone can view or add the catalog as an annotation.

   Additionally you will see that your change has been logged in the recent assessments section for that variant. This allows for tracking of users and their changes.

4. To view the assessment catalog data on the VSWarehouse. Open the internet browser of your choice and navigate to the VSWarehouse address. After login in you should see the *Clinical Significance* Catalog in the Assessment catalog section of the home page.

   Select the Assessment catalog and find the record that you changed in VarSeq. Click on the record that was just changed in VarSeq, the record will allow you to change the variant assessment. Change the *Reported Classification* to *Unconfirmed somatic variant* and click save.

5. Now, switching back to VarSeq refresh the catalog to load the changes made from the browser.

   The reported classification has now changed, notice the additional edit in the Recent Assessments Log.

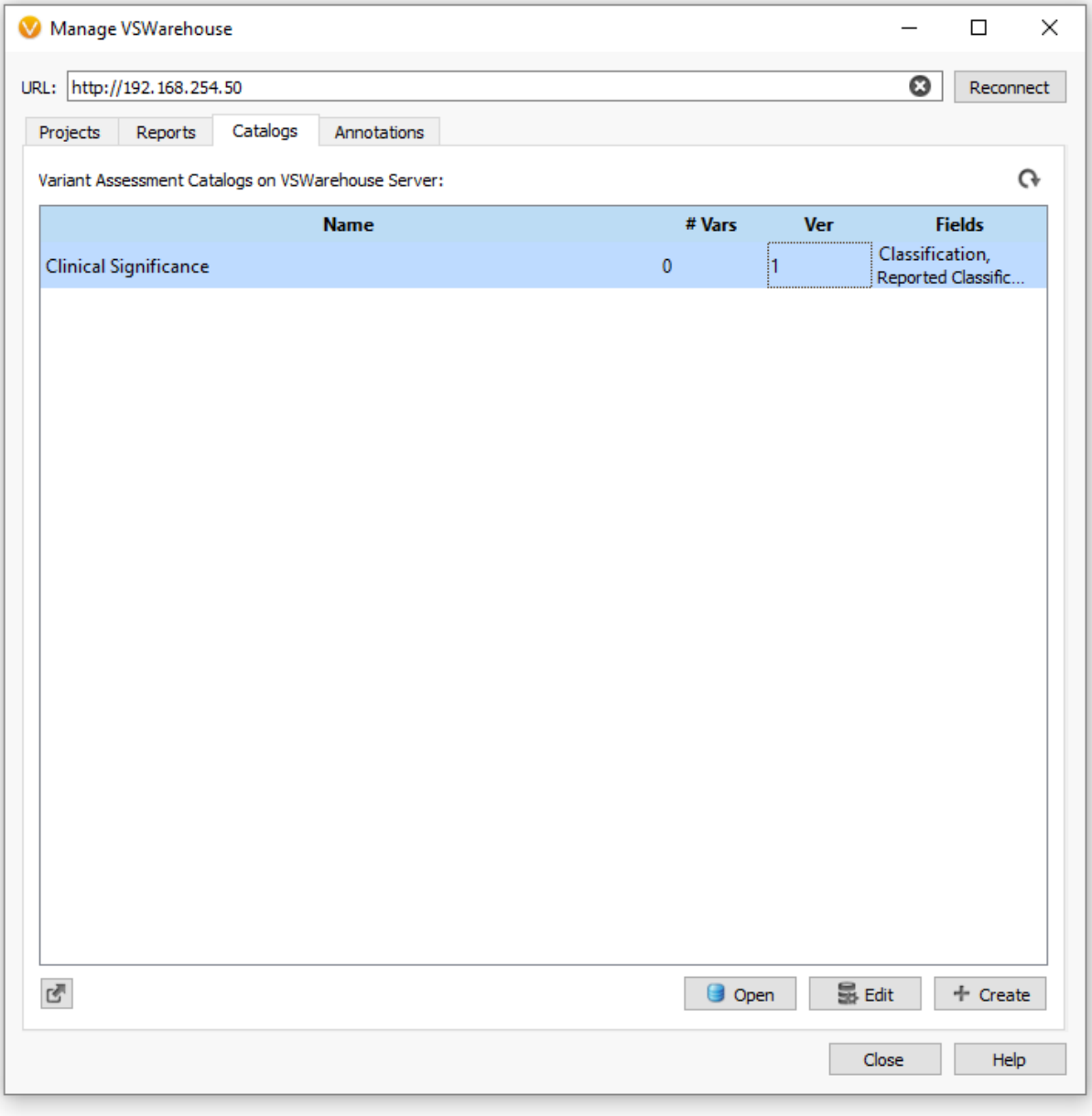


Figure 4-1: Open the assessment catalog in VarSeq.

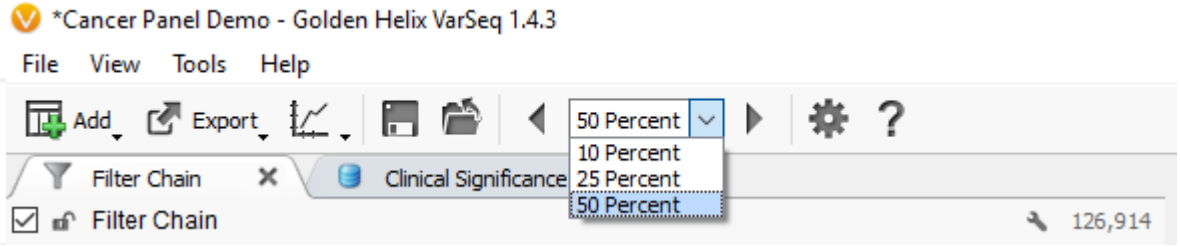


Figure 4-2: Selecting the current sample

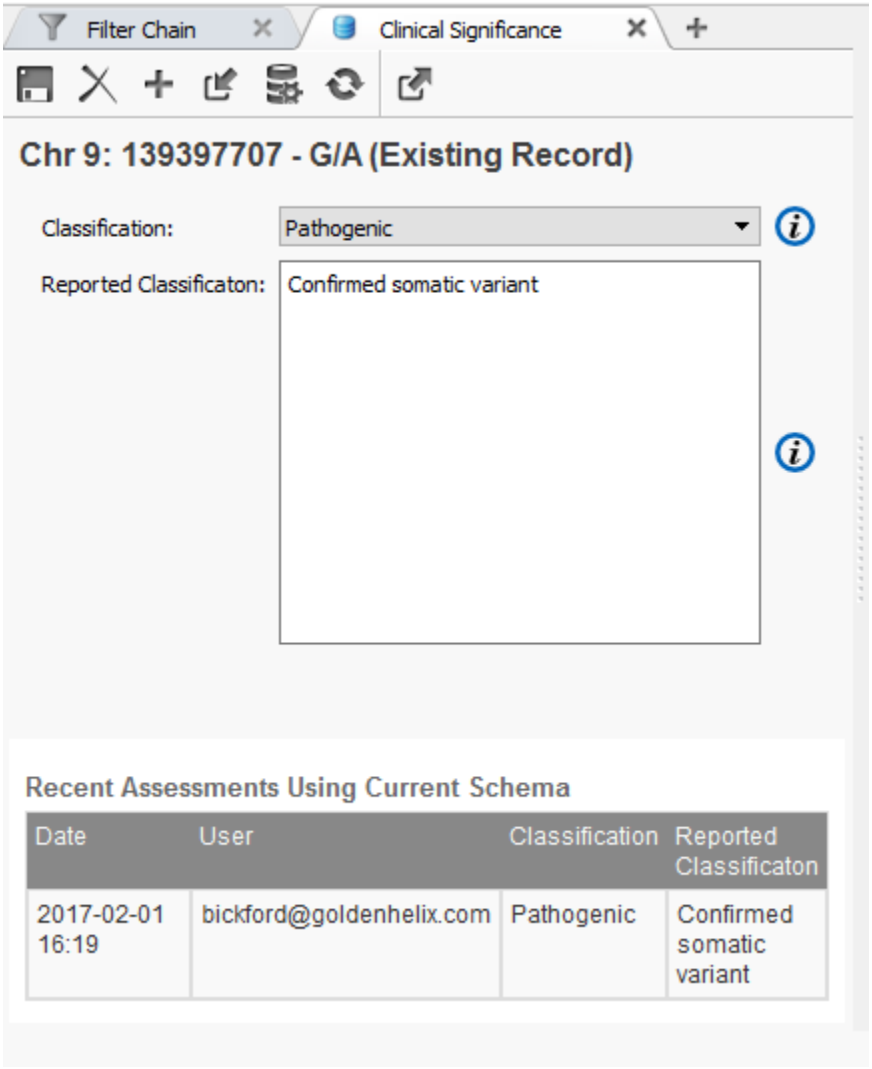


Figure 4-3: Selecting the current sample

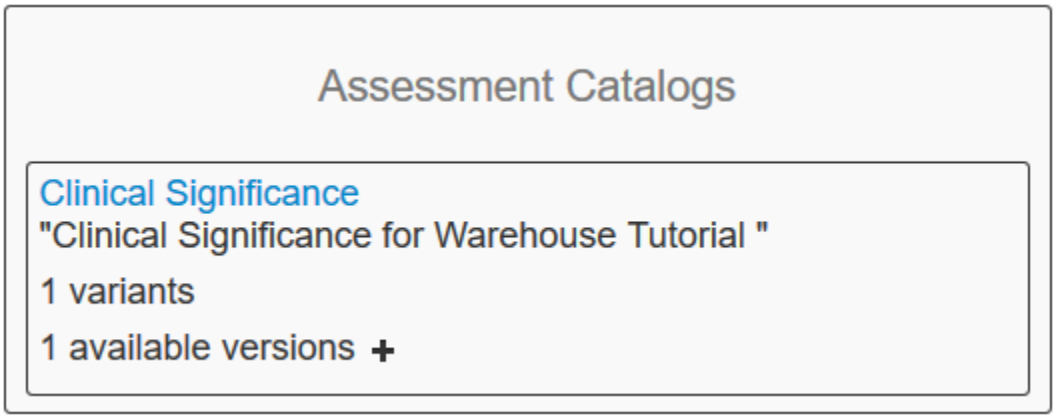


Figure 4-4: Catalog on the VSWarehouse home page

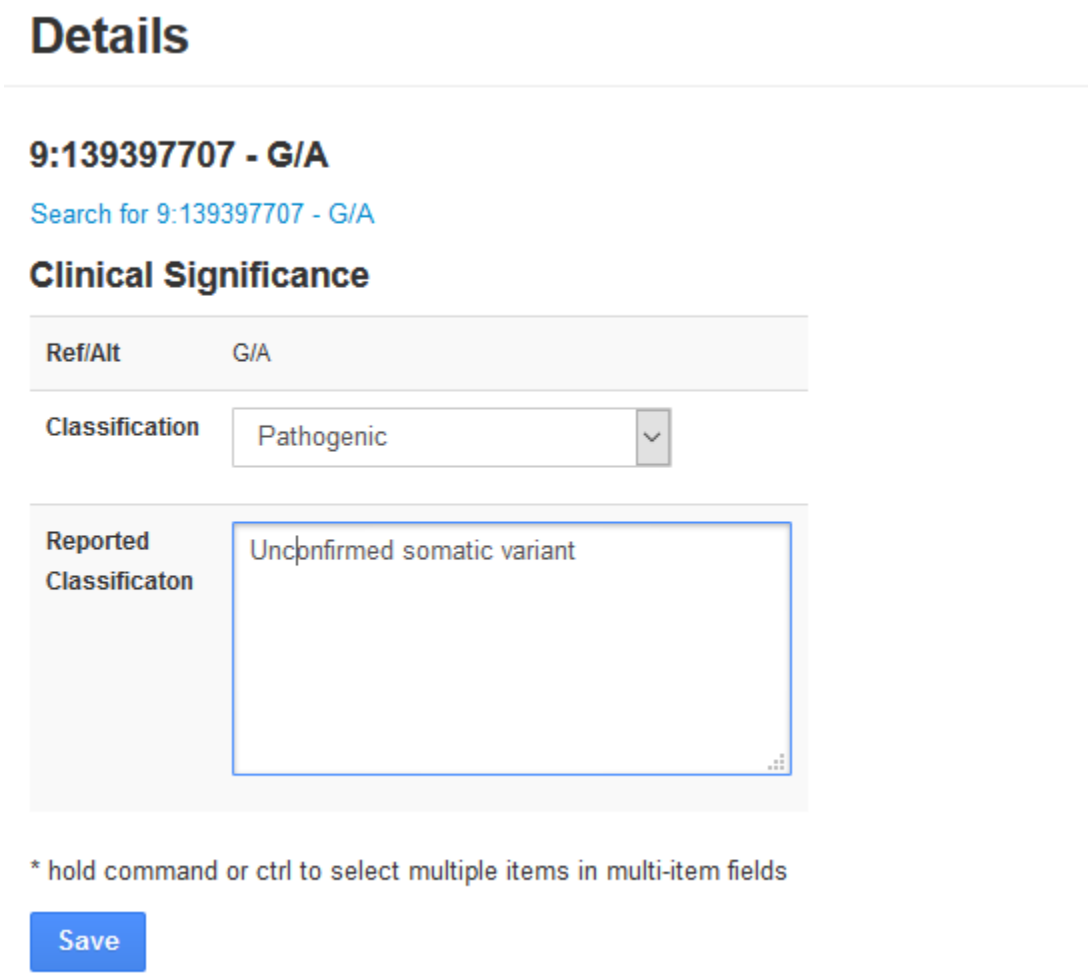


Figure 4-5: Editing the catalog the VSWarehouse.

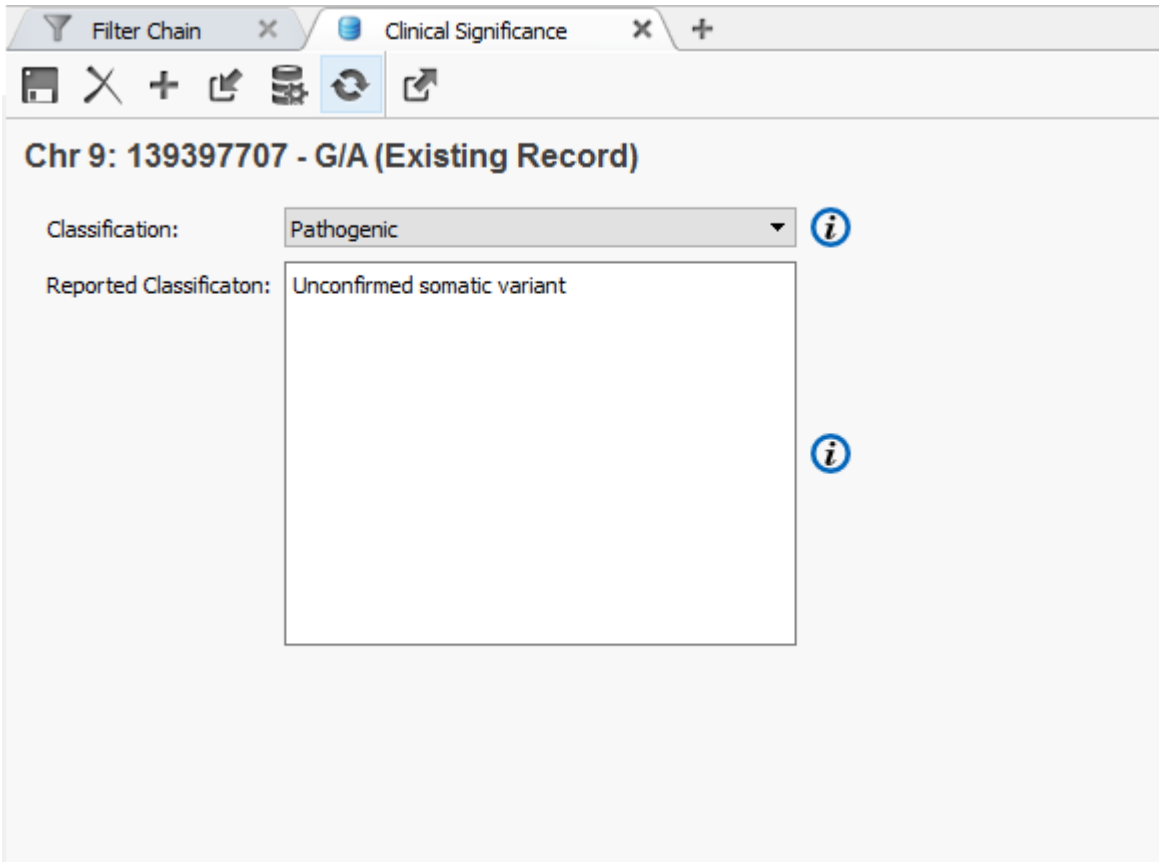


Figure 4-6: Refreshing the catalog in VarSeq

CHAPTER FIVE

USING A VSWAREHOUSE REPORT

Warehouse can be used to store sample report data. This allows you to view previous reports, as well as query across the variants that have been included in the report findings. Now that we have finished filtering the variants we will create a report with the filtered variants and upload it to VSWarehouse.

1. First, we will open a VSWarehouse hosted report template. Open the manage VSWarehouse dialog and navigate to the **Reports** tab *Figure5-1*

   Select the *Cancer Panel Report* and click **Open**. This will open that report in a new VSWarehouse tab in your VarSeq project.

2. The next step will be to add sources required by the report that are missing from the project. These sources are used to autofill the data in the report. Click the red triangle in the upper right hand corner of the report view to view the current errors. The menu should display that two required sources are missing, and four optional sources are missing. Click each of the messages and add the missing annotations to your project.

   Adding these premium annotations will allow information from these annotations to be included in the report.

3. While the annotations are running you can start to fill out the report template with some basic sample information.

   The values that are placed in the form are used to render the report once it has been filled out. Additionally these are saved on the warehouse where they are aggregated across all of the reports that have been created using this template.

4. After the patient and physician information has been filled out the next step is to select the variants that will be included in the report using *Variant Sets*. To add a variant set click the square icon in the variant table view.

   This will open the Create Variant Set dialog. Name the record set **Primary Findings** change the Initials to **PF** and click **OK**. Add an additional variant set named **Incidental Findings** with the initials **IN** colored green, following the same steps.

   This will create two new flags fields. Add the top two variants to the Incidental Findings flags by clicking the colored square for each record under the IN field. Add the last variant to the Primary Findings category.

   Finally, select the variants sets that correspond to the Primary and Incidental section in the reports view. This will take the information from the selected variants and fill it in the report.

5. Next check the signoff checkbox. Then click the render report button at the top of the report view. This will open a web view prompting you to log into VSWarehouse. After you log in the rendered html report will load from the VSWarehouse.

   The report has now been saved on the warehouse. All of the samples and the variants in their findings can be viewed and queried from the warehouse. Additionally the variant set in the report can be used as an annotation source.

6. To add the variants from the findings section as an annotation source open the Manage VSWarehouse Dialog. Navigate to the annotations tab and find the **Cancer Panel Report - Primary Findings** source.

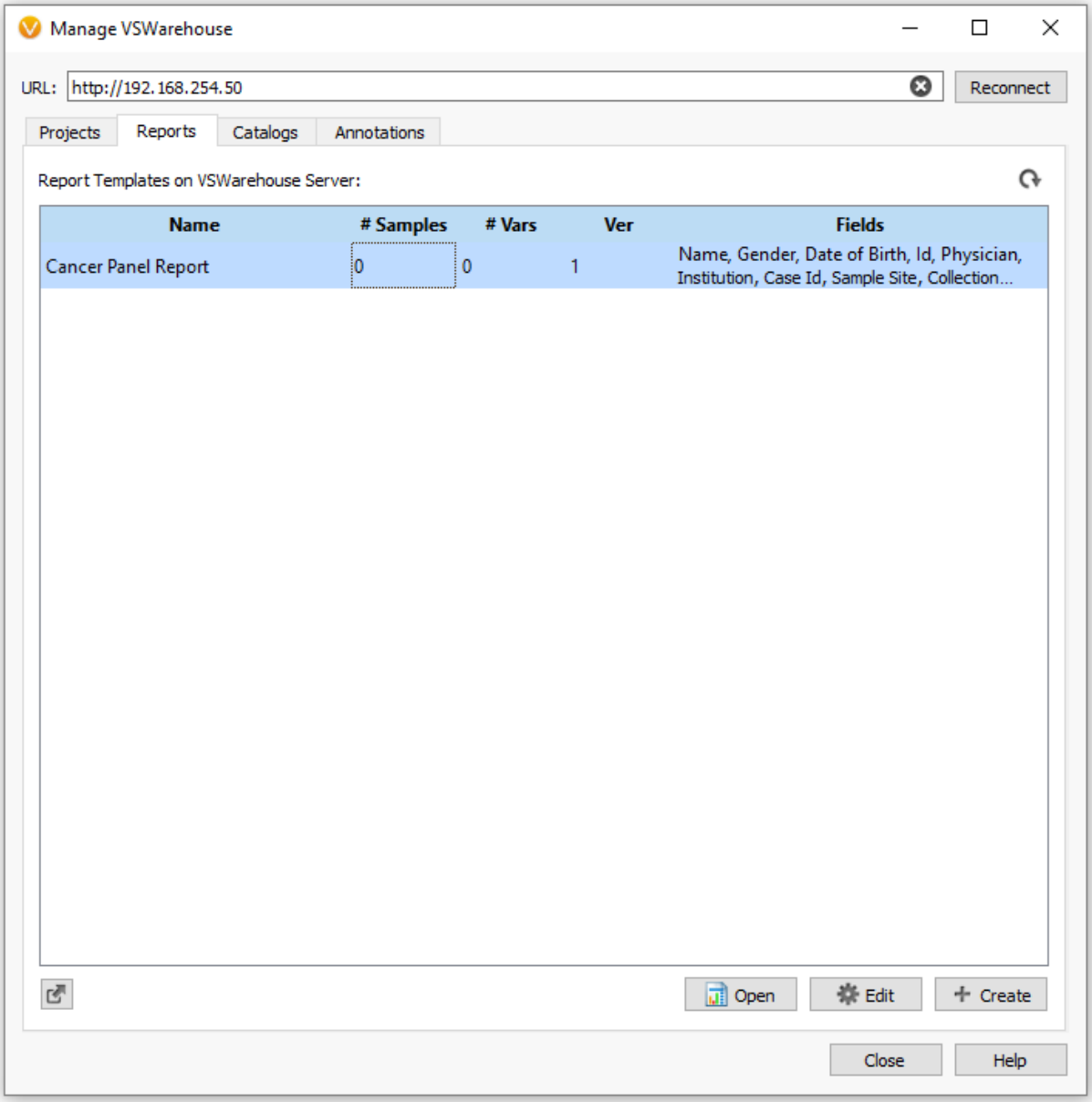


Figure 5-1: Open the reports tab in the VSWarehouse management dialog.

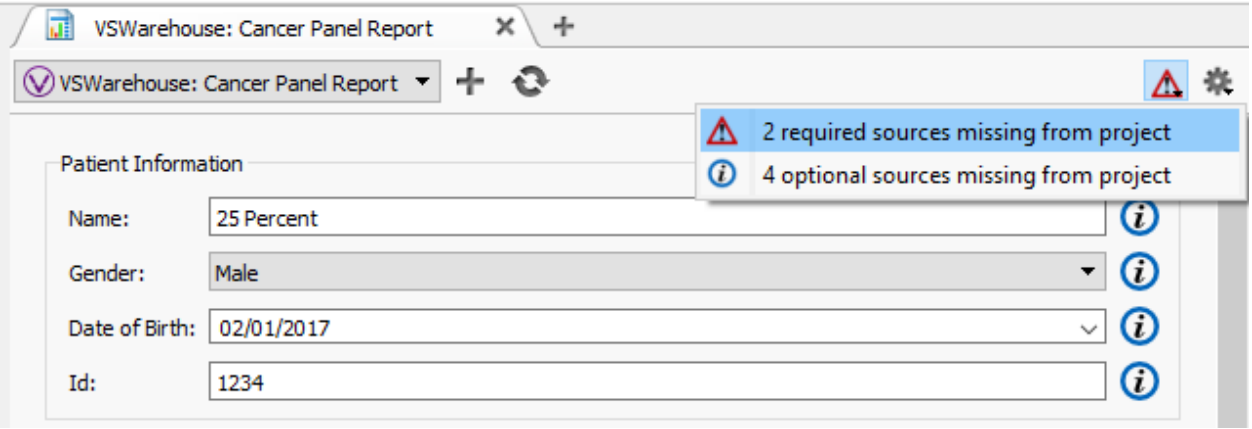


Figure 5-2: The report is missing required OncoMD annotations, and optional OMIM sources.

Adding this source will add the primary findings from the source that was just uploaded to the warehouse as an annotation source. Scrolling all the way to the right of the table you will see the variant included in the Primary Findings appear as an annotated variant.

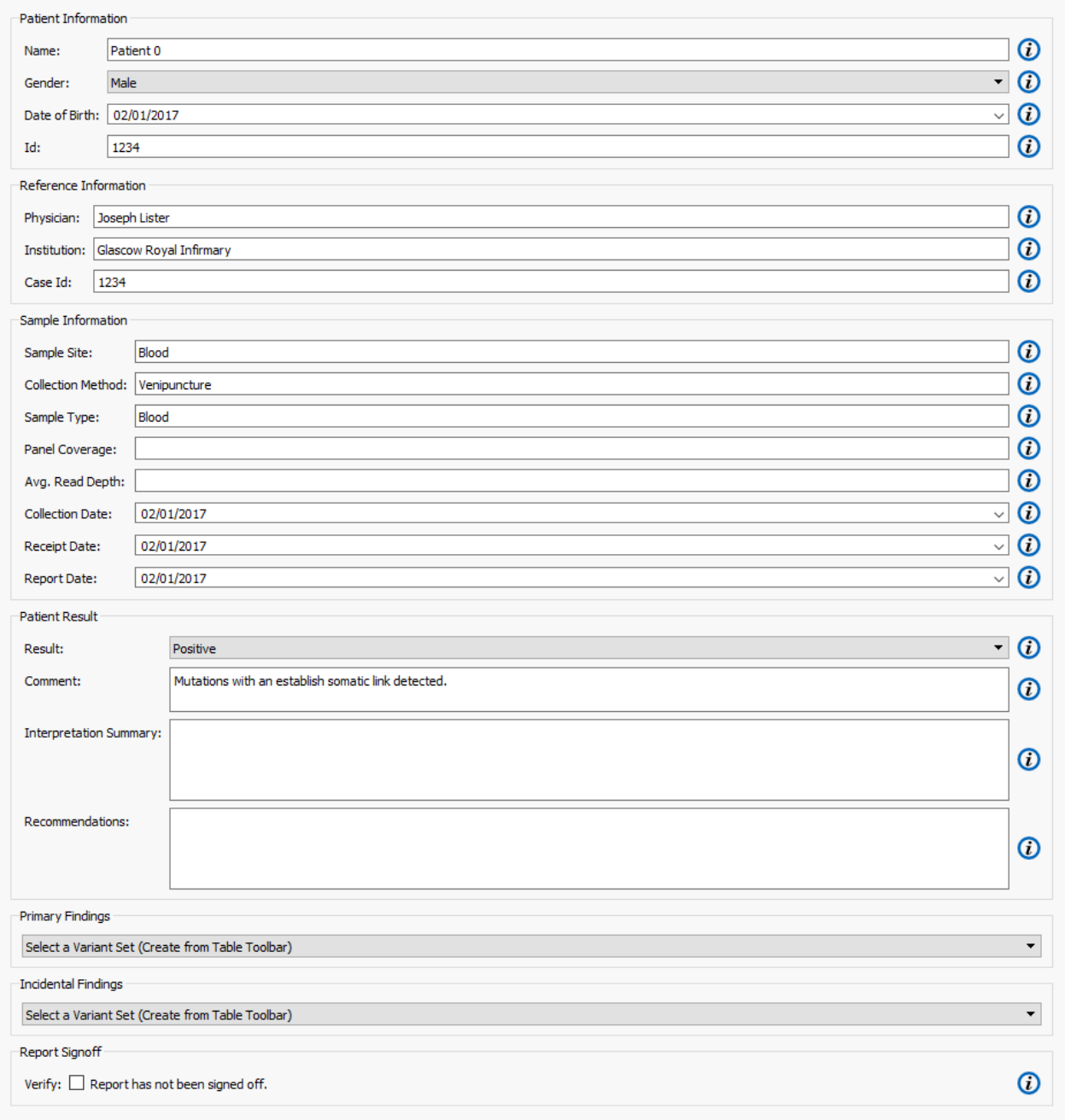


Figure 5-3: Report with the patient and physician information.

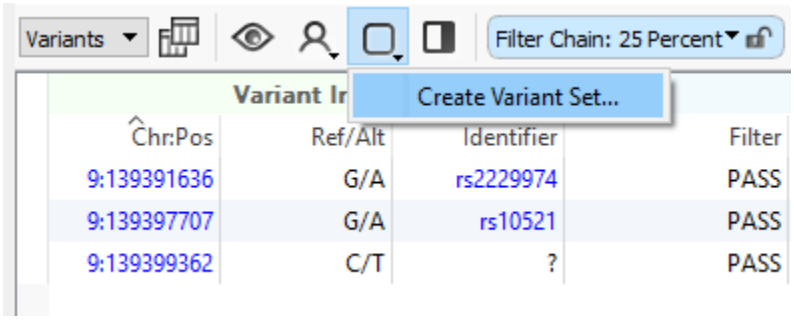


Figure 5-4: Adding a variant set to the variant table

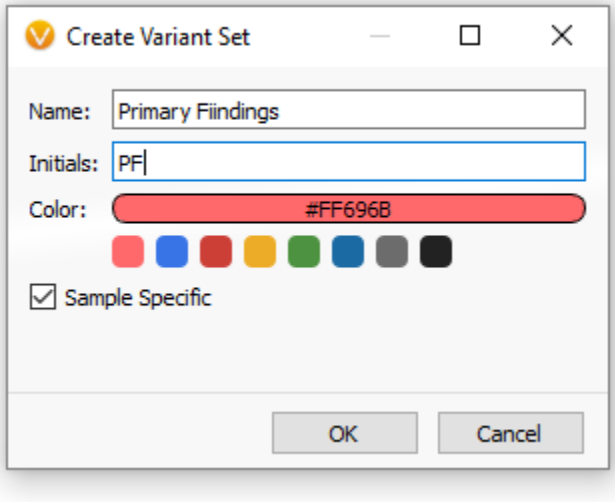


Figure 5-5: Naming the new variant set

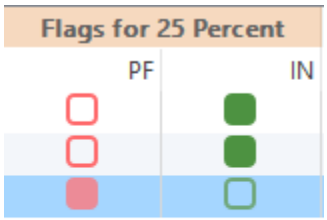


Figure 5-6: Flagging the filtered Variants

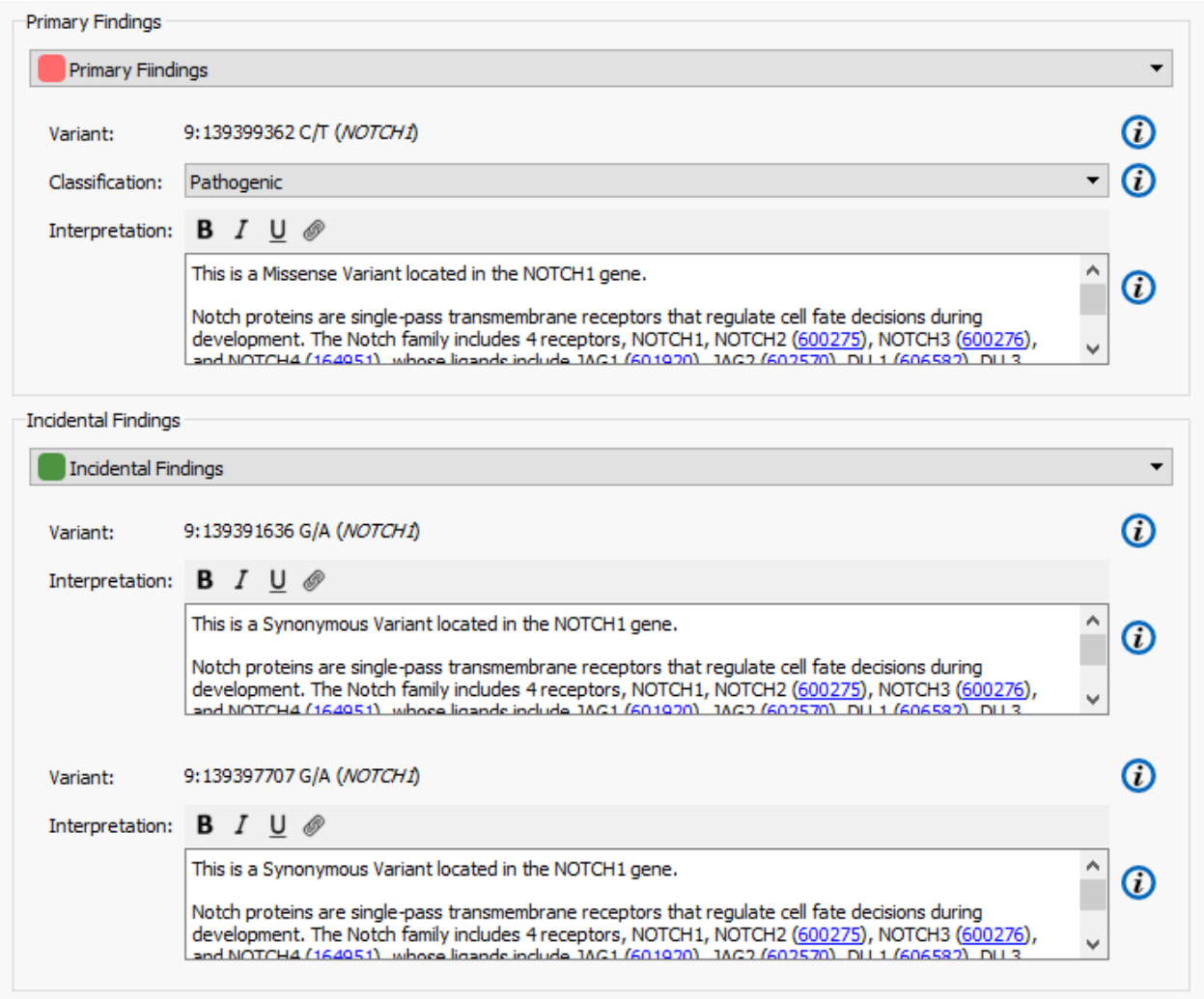


Figure 5-7: Selecting the report record sets

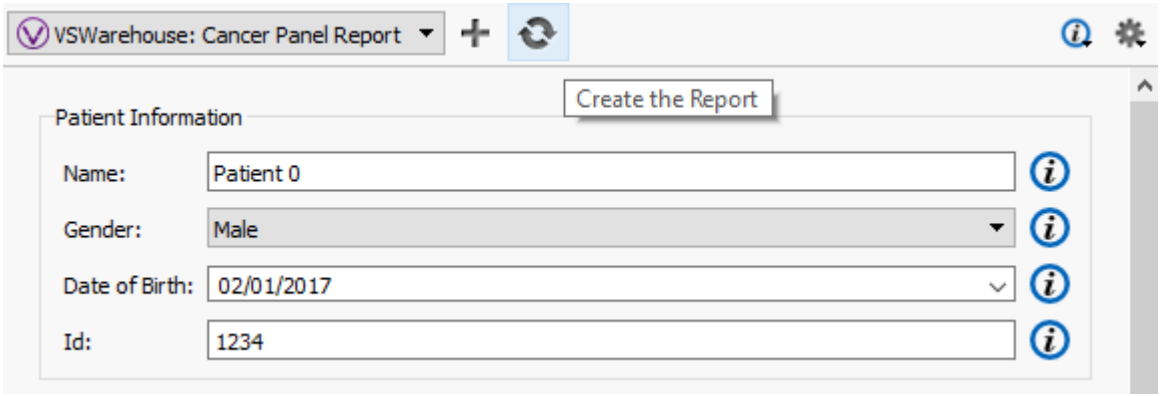
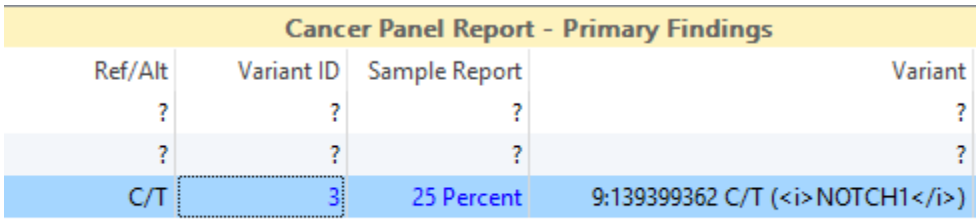


Figure 5-8: Click the render button to load the rendered report from VSWarehouse.



Cancer Panel Report - Primary Findings

Ref/Alt	Variant ID	Sample Report	Variant
?	?	?	?
?	?	?	?
C/T	3	25 Percent	9:139399362 C/T (NOTCH1)

Figure 5-9: Adding the Primary Findings from the report

CHAPTER SIX

UPLOADING PROJECTS TO VSWAREHOUSE

1. After you have finished the analysis and the final report in VarSeq the final step is to upload the project to VSWarehouse. To upload the current project open the Manage VSWarehouse dialog through the Tools menu and select the **Projects** tab.

   Next, select the *Cancer Gene Panel* project and click the add samples button. This will open the Add Samples to Warehouse wizard.

2. The first step of the Add Samples to Warehouse wizard is editing the sample fields. This allows you to change the sample fields in the same manner that they are added when the first importing data.

   The fields can be left with their existing values. Click **Next** to progress to the next page of the wizard.

3. The last page of the wizard allows you to change the *Upload Source* for the samples. This is a tag that is applied as an additional field to all of the samples. Additionally, you can change when to run the import on warehouse.

   Change this value to *Run Immediately* to import these samples to the VSWarehouse project as soon as they are uploaded. Click **Upload** to begin the process.

4. When the upload is complete you will be presented with a link which will allow you to follow the progress of the import on the VSWarehouse.

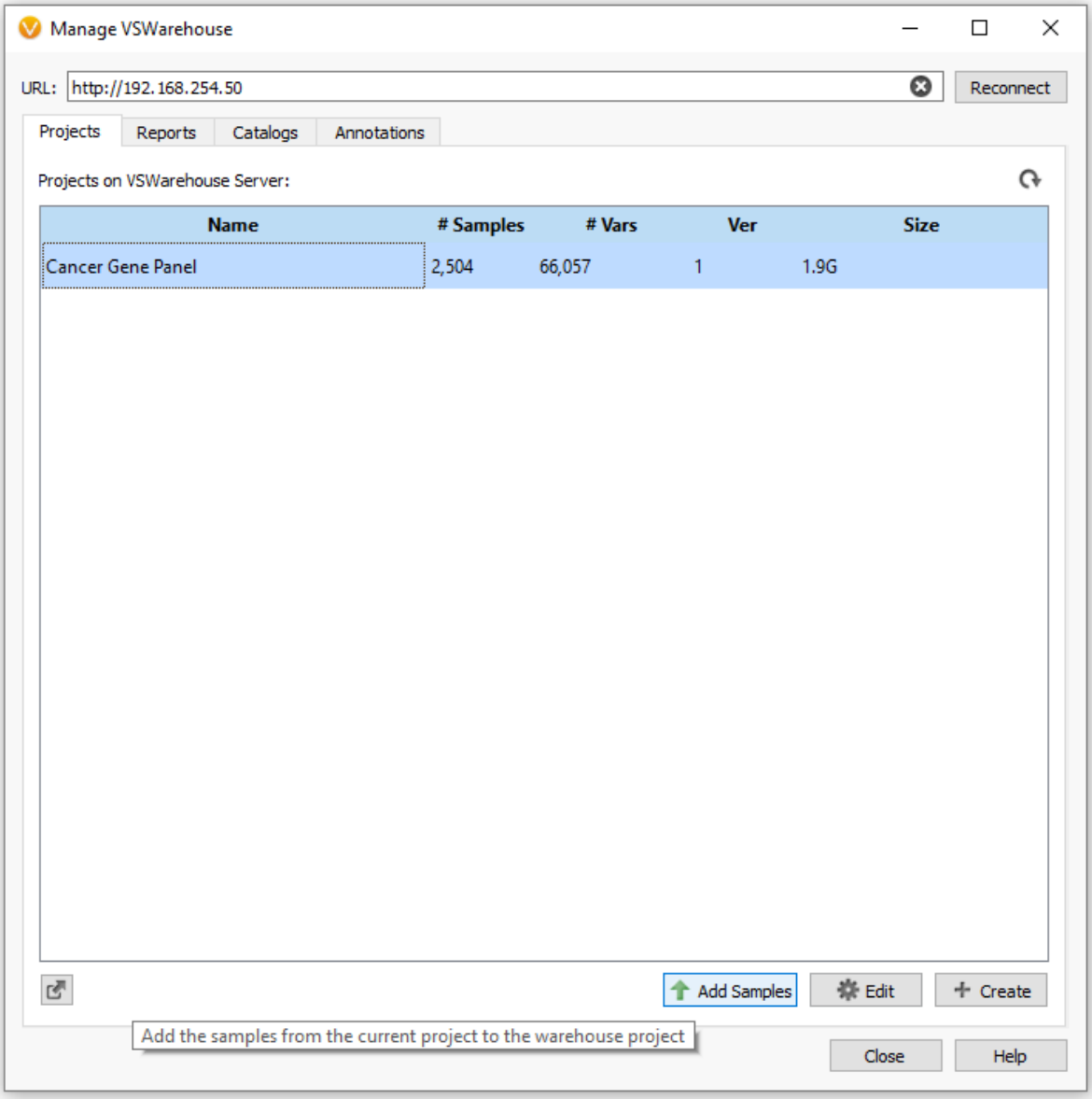


Figure 6-1: Add samples to the VSWarehouse

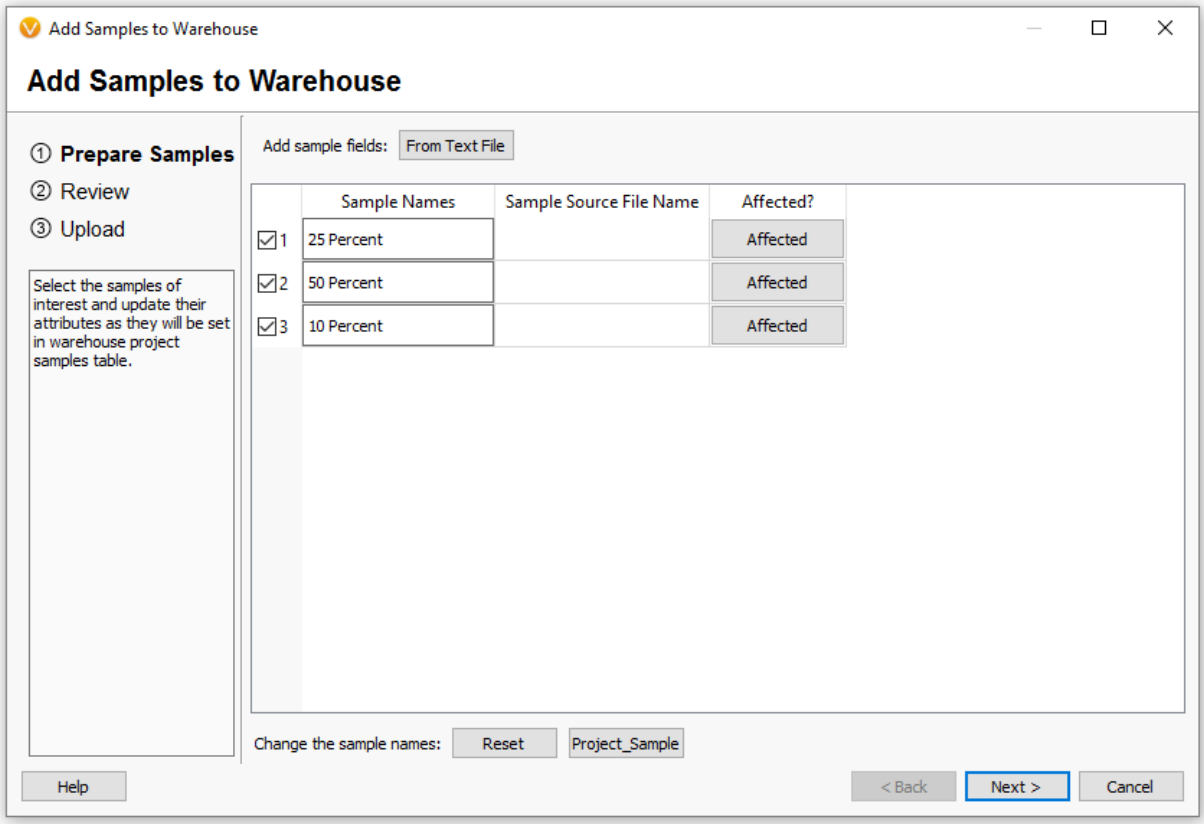


Figure 6-2: Editing the samples attributes before upload.

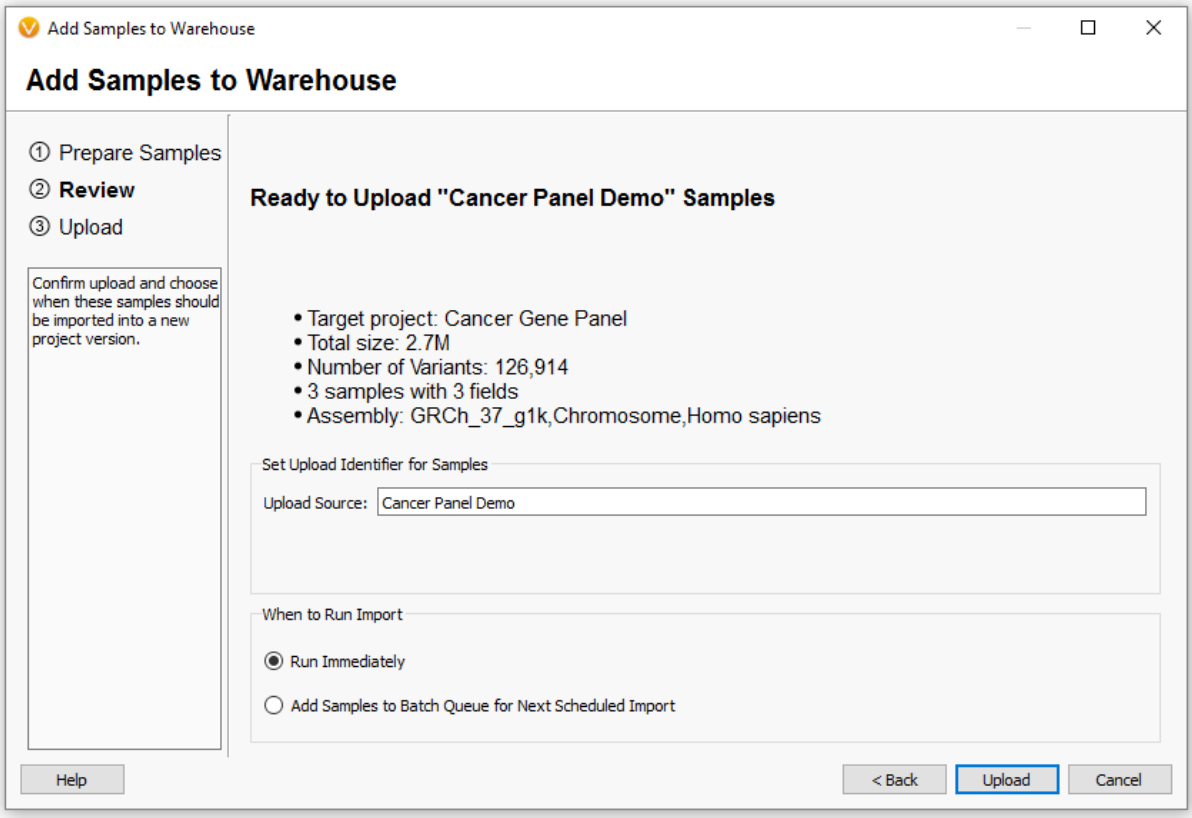


Figure 6-3: Editing the samples attributes before upload.

CHAPTER SEVEN

EXPLORING VSWAREHOUSE

From the VSWarehouse web interface you can view the samples, variants, assessment catalogs, and reports that have been uploaded to the VSWarehouse. Each of these categories can be queried and exported from using the same **Query** and **Result** interfaces.

1. Open the web browser of your choice and log into the VSWarehouse. This will open the VSWarehouse homepage, which is divided into four sections. One for each category of data stored on the warehouse.

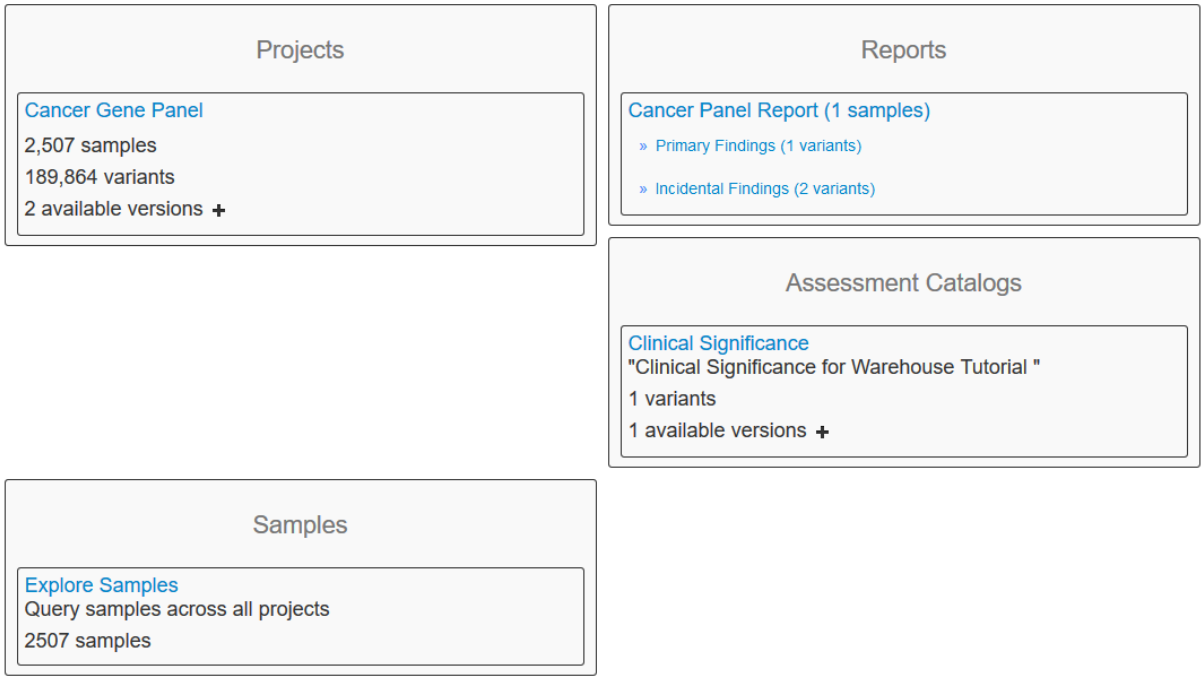


Figure 7-1: VSWarehouse home screen

   In the project section, select the *Cancer Gene Panel*. This will open the VSWarehouse project that corresponds to the VarSeq project was previously uploaded.

2. To find other rare variants in the BRAF gene opened the query interface by selecting **Query** at the top of the page. This will open the query interface where you can construct filters based on the fields in the data.

   Next add a filter in the *Allele Counts* section for the Allele Frequency with an upper bound of **0.01**, and uncheck include missings.

   This will filter to rare variants that occur in less than 1 percent of the samples that been added to this project. Next switch to the *RefSeq Genes 105v2, NCBI* section to add a filter for the *Gene Names* equal to **BRAF**. Finally

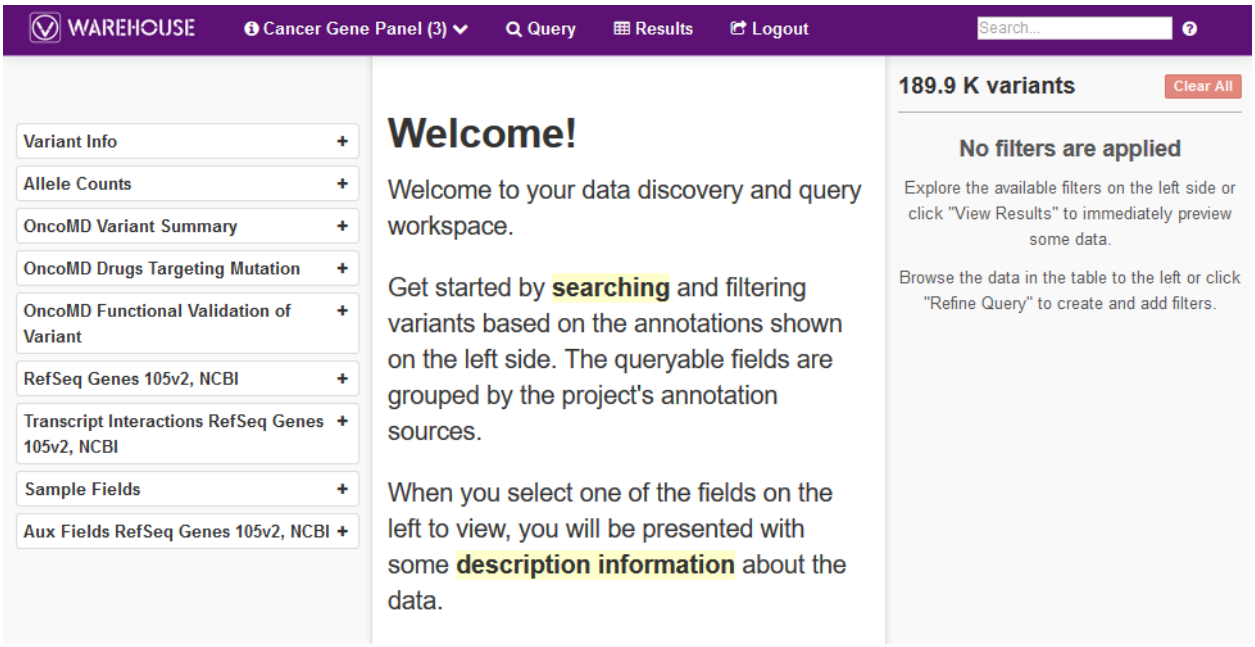


Figure 7-2: VSWarehouse query screen

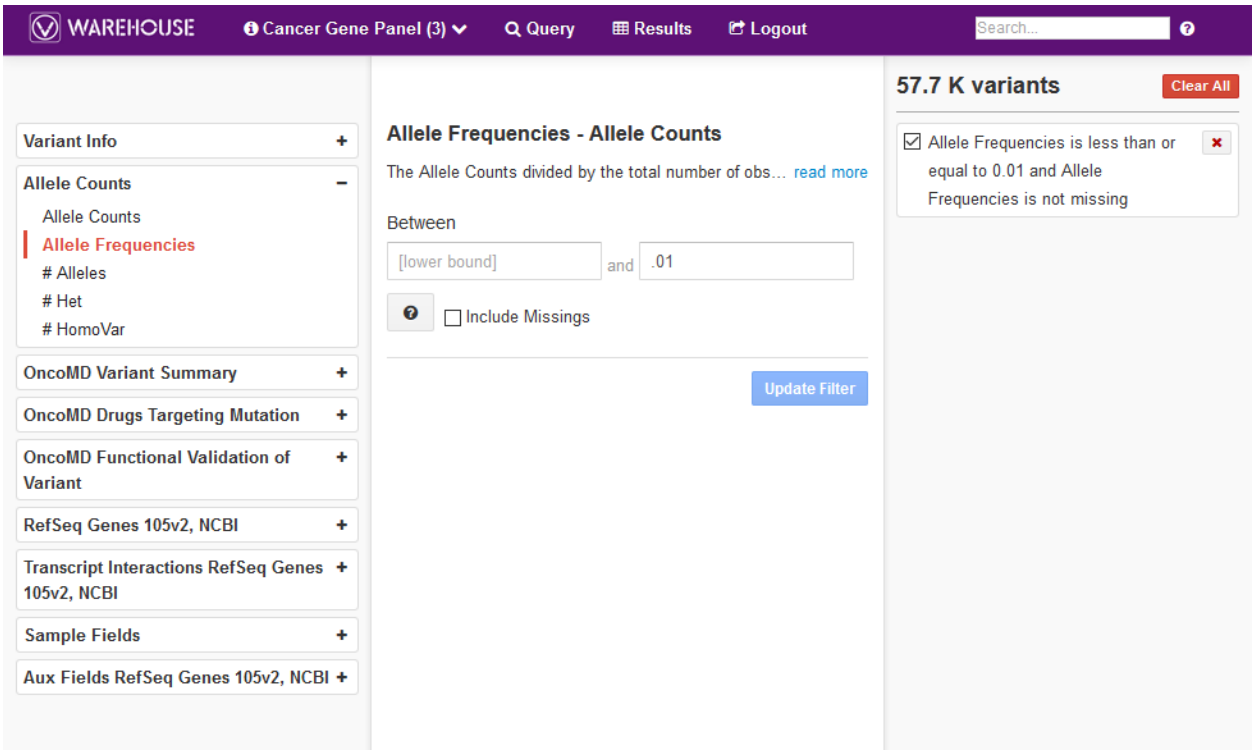


Figure 7-3: VSWarehouse query screen

add a filter for the *RefSeq Genes 105v2, NCBI* field *Effect (Combined)* equal to **Missense**.

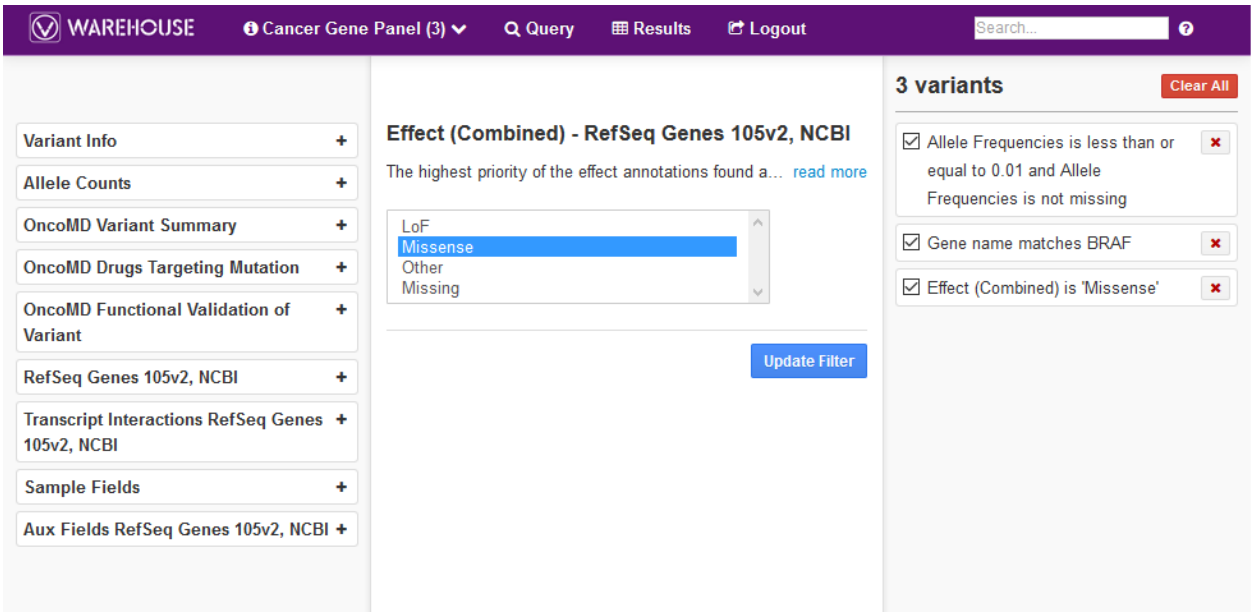


Figure 7-4: Add the Effect (combined) filter equal to Missense.

This should reduce the number of variants to 3. To view the variants click **Results** at the top of the page.

3. The results page will display the three records that pass the filters. Change the visible fields by selecting **Change Columns**. This will display the published columns that are available on the left and the current visible columns on the right.

   Add the **Allele Frequencies** field from the *Allele Counts* category, as well as the **Drug**, **Generic Name** and **Response Rate** fields from the *OncoMD Drugs Targeting Mutation* category.

4. Next select **Export** at the top of the screen to save the results. Select **Excel** as the file type and under *Select Samples* click *Unselect All*. This will remove the sample fields from the output.

   Clicking **Export** will start a job to perform the export. Click **View Task** to see the status of the job, and to download the results after it has completed. You can then open the file in Excel shown in the VSWarehouse results.

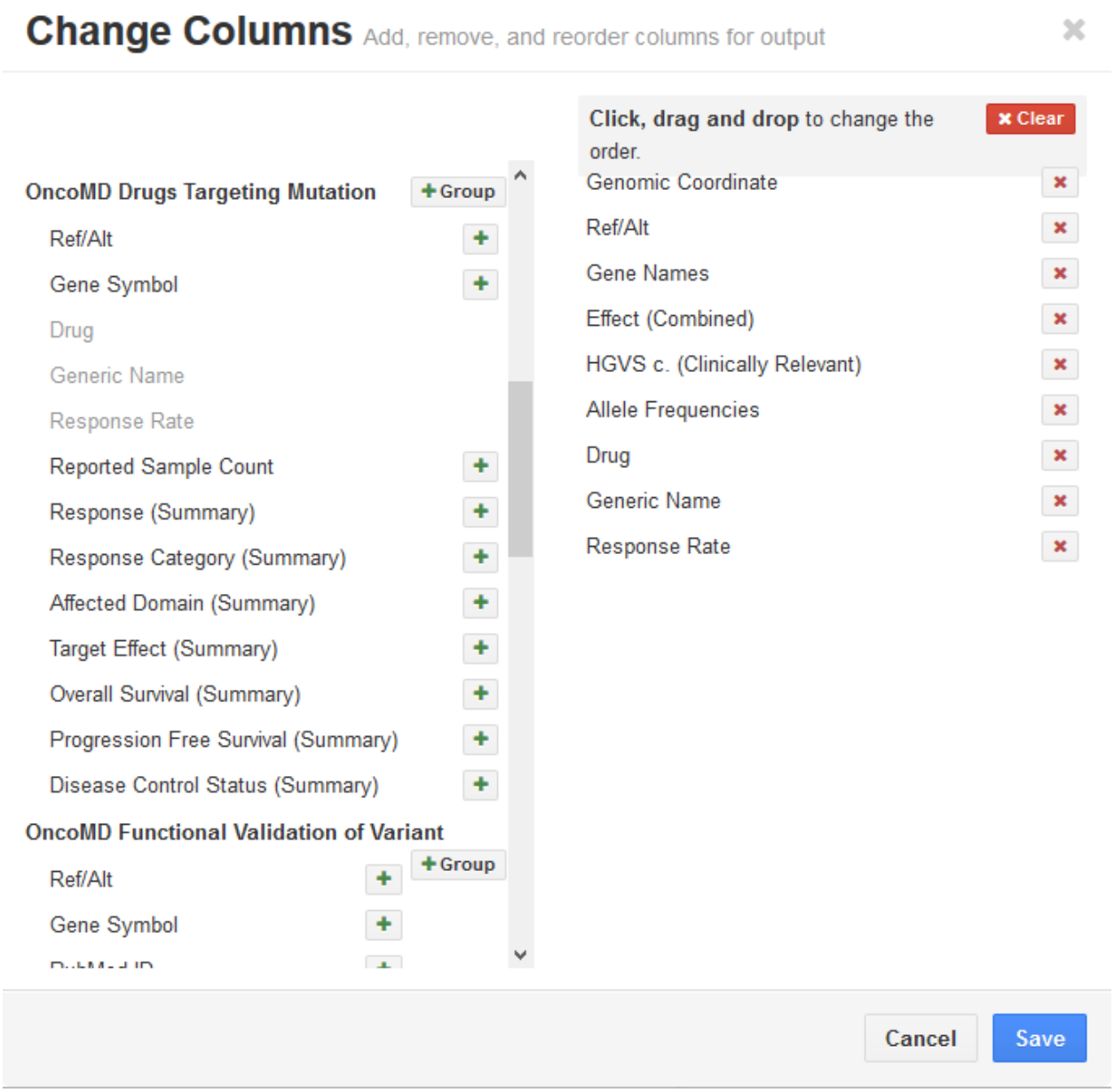


Figure 7-4: Changing the columns visible in the results

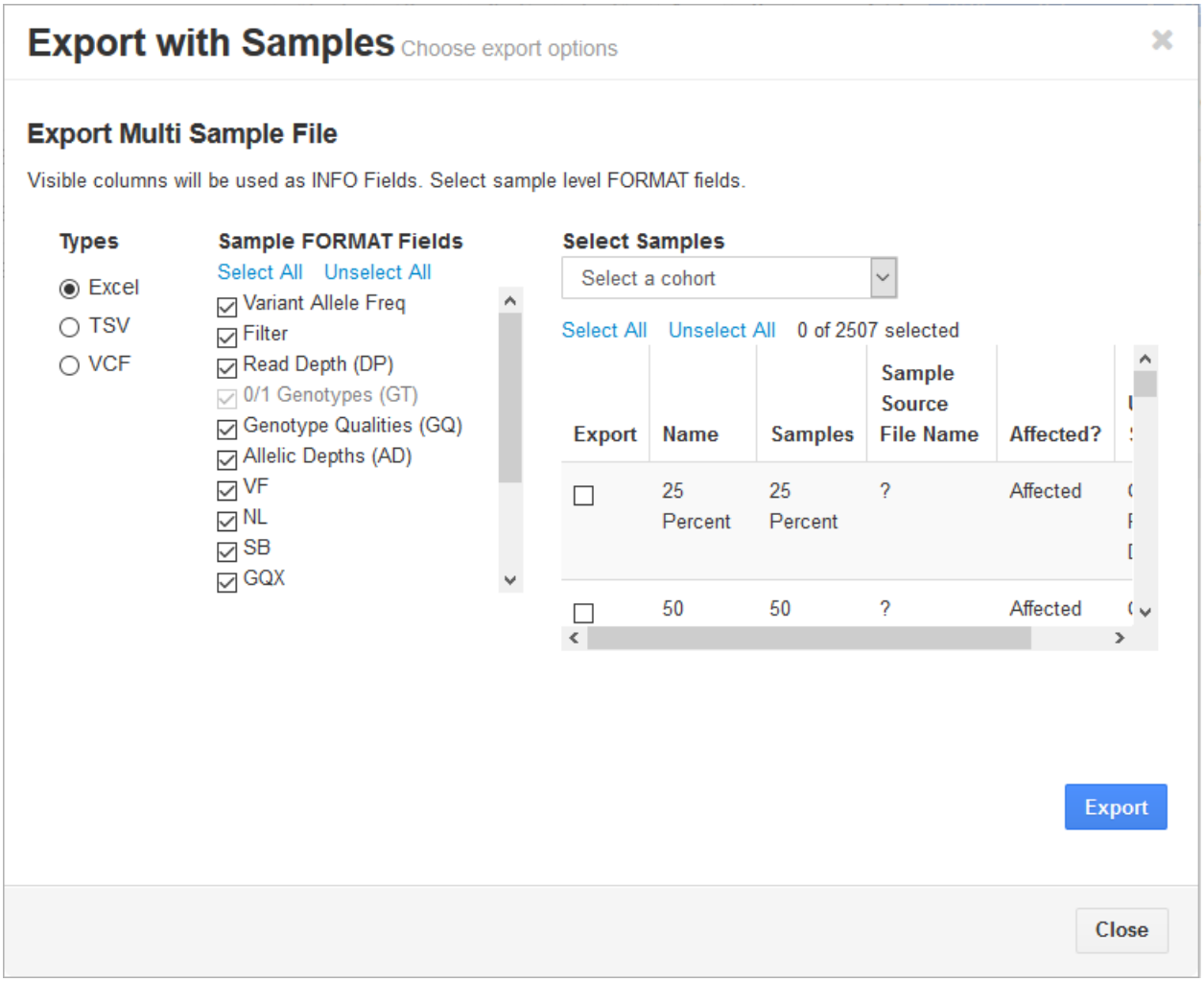


Figure 7-4: Export the filtered results

CHAPTER EIGHT

CONCLUSION

This tutorial was designed to give a taste of all the features and capabilities of VSWarehouse and how it can be integrated in a VarSeq workflow.

Additional features and capabilities are being added all the time, so if you do not see a feature you need for your workflows please do not hesitate to let us know!