Adding annotations to a gene list

There are many useful visualizations, annotations, and biological interpretation tools that can operate on a gene list. In order for these features work with an imported list, an annotation file must be associated with the gene list. Additionally, many operations that work with a list of significant genes (like GO- or Pathway-Enrichment) require comparison against a background of “non-significant” genes. The quickest way to accomplish both is to use the background of “all genes” for that organism provided by an annotation source like RefSeq, Ensembl, etc. in .pannot (Partek annotation), .gff, .gtf, .bed, tab- or comma-delimited format. If the file is not already in a tab-separated or comma delimited format, you may import, modify, and save the file in the proper file format.

Associating a spreadsheet with an annotation file

- Select **File** from the main toolbar
- Select **Genomic Database** under **Import** (Figure 1)

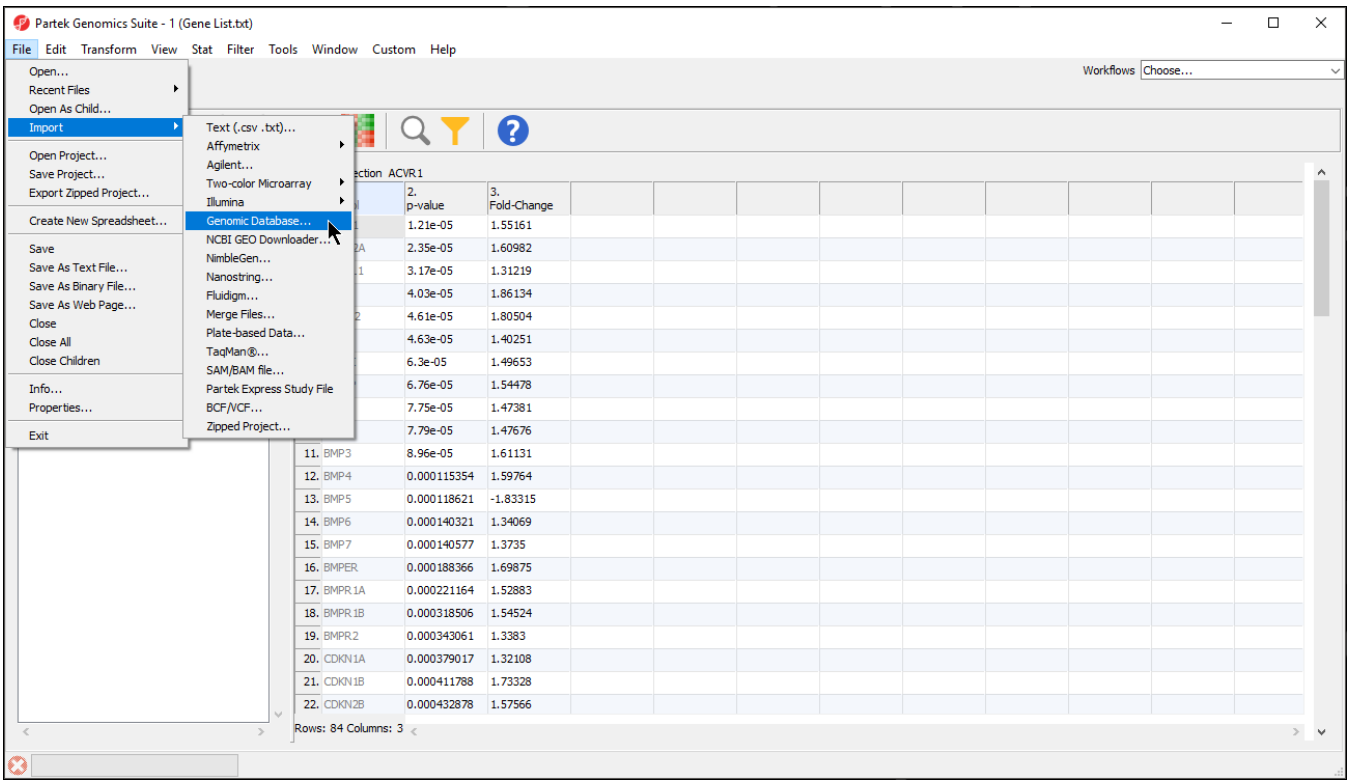
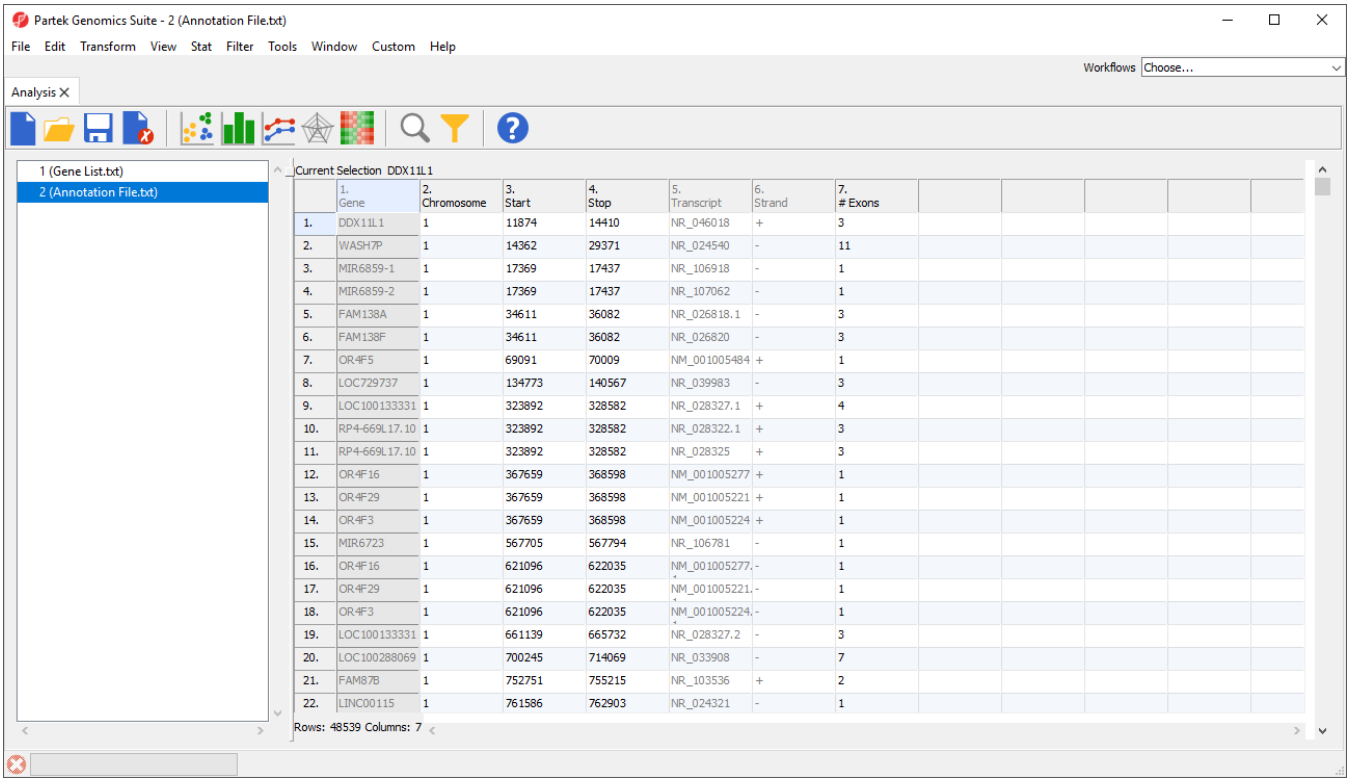


Figure 8. Importing an annotation file

- Select the annotation file; in this example, we select a .pannot file downloaded from Partek distributed library file repository – *hg19_refseq_14_01_03_v2.pannot*
- Delete or rearrange the columns as necessary; we have placed the column with identifiers (should be unique ID) that correspond to our gene list first
- Select **File** then **Save As Text File...** to save the annotation file; we have named it *Annotation File* (Figure 2)



	1. Gene	2. Chromosome	3. Start	4. Stop	5. Transcript	6. Strand	7. # Exons
1.	DDX11L1	1	11874	14410	NR_046018	+	3
2.	WASH7P	1	14362	29371	NR_024540	-	11
3.	MIR6859-1	1	17369	17437	NR_106918	-	1
4.	MIR6859-2	1	17369	17437	NR_107062	-	1
5.	FAM138A	1	34611	36082	NR_026818.1	-	3
6.	FAM138F	1	34611	36082	NR_026820	-	3
7.	OR4F5	1	69091	70009	NM_001005484	+	1
8.	LOC729737	1	134773	140567	NR_039983	-	3
9.	LOC100133331	1	323892	328582	NR_028327.1	+	4
10.	RP4-669L17.10	1	323892	328582	NR_028322.1	+	3
11.	RP4-669L17.10	1	323892	328582	NR_028325	+	3
12.	OR4F16	1	367659	368598	NM_001005277	+	1
13.	OR4F29	1	367659	368598	NM_001005221	+	1
14.	OR4F3	1	367659	368598	NM_001005224	+	1
15.	MIR6723	1	567705	567794	NR_106781	-	1
16.	OR4F16	1	621096	622035	NM_001005277.	-	1
17.	OR4F29	1	621096	622035	NM_001005221.	-	1
18.	OR4F3	1	621096	622035	NM_001005224.	-	1
19.	LOC100133331	1	661139	665732	NR_028327.2	-	3
20.	LOC100288069	1	700245	714069	NR_033908	-	7
21.	FAM87B	1	752751	755215	NR_103536	+	2
22.	LINC00115	1	761586	762903	NR_024321	-	1

Figure 9. Modified annotation file

- Select ( ) to close the annotation file

Now we can add the annotation file to our imported gene list.

- Right click *1 (gene_list.txt)* in the spreadsheet tree
- Select **Properties** from the pop-up menu

This brings up the *Configure Genomic Properties* dialog (Figure 3).

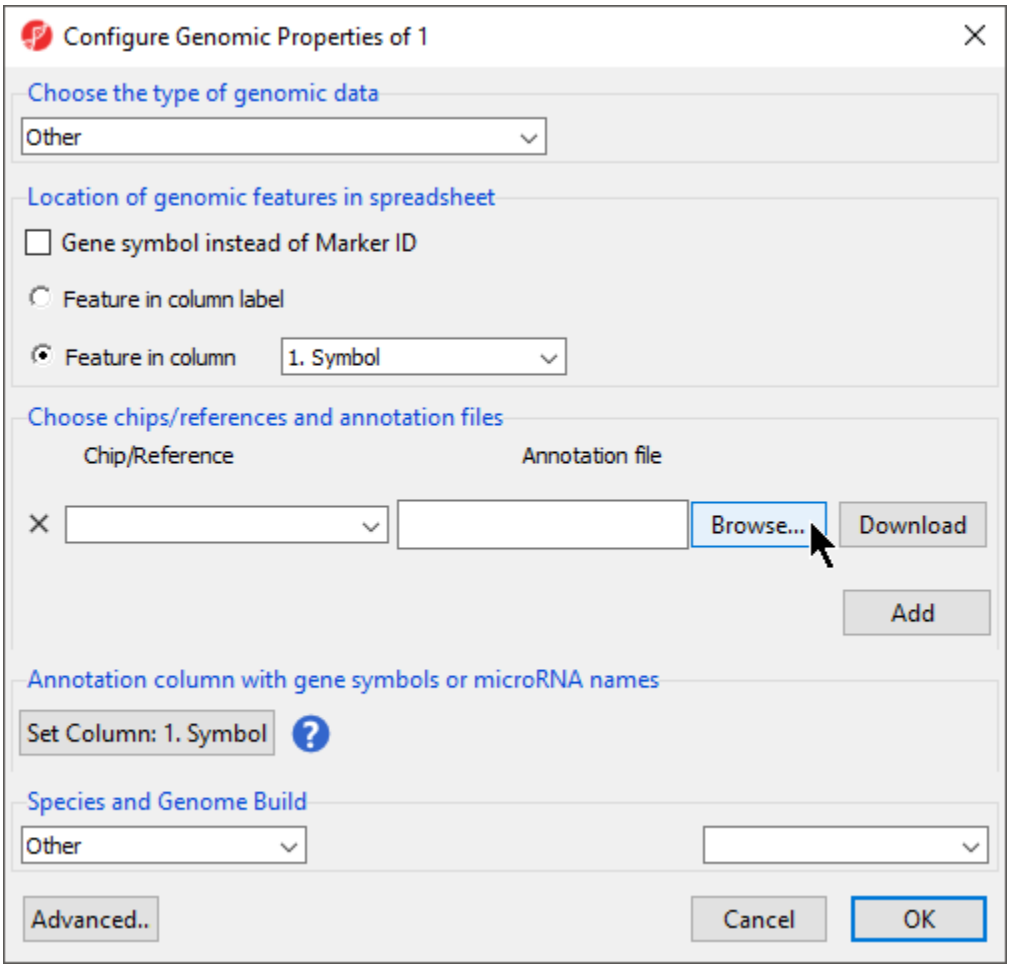


Figure 10. Selecting an annotation file using the Configure Genomic Properties dialog

- Select **Browse** under *Annotation File*
- Choose the annotation file; we have chosen *Annotation File.txt*

If this is the first time you have used an annotation, the *Configure Annotation* dialog will launch. This is used to choose the columns with the chromosome number and position information for each feature. Our example annotation file has chromosome, start, and stop in separate columns.

- Select the proper column configuration options (Figure 4)

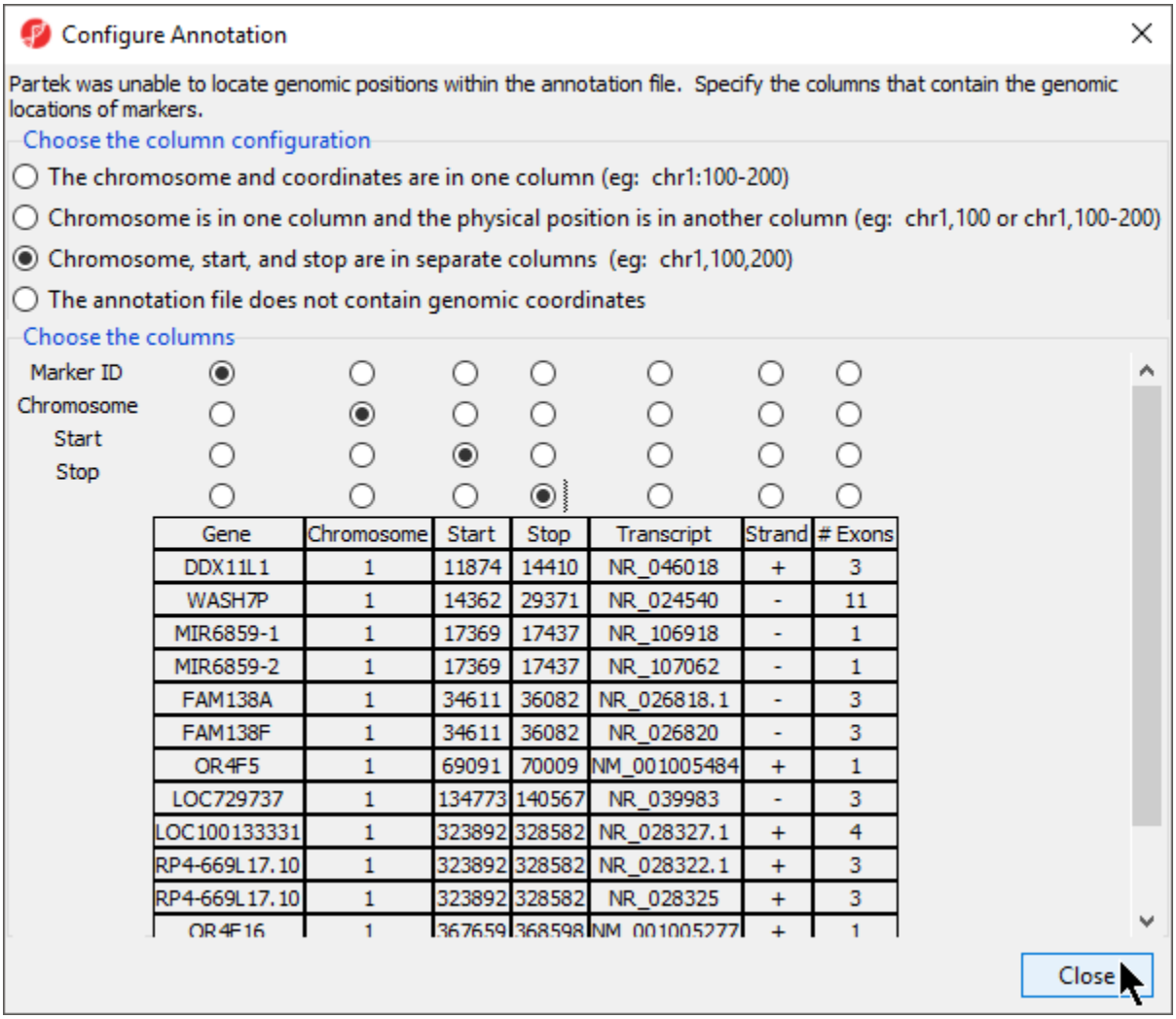


Figure 11. Assigning columns for chromosome and genomic positions in the annotation file

- Select **Close** to return to the *Configure Genomic Properties* dialog
- Select **Set Column:** to open the *Choose column with gene symbols or microRNA names* dialog (Figure 5)

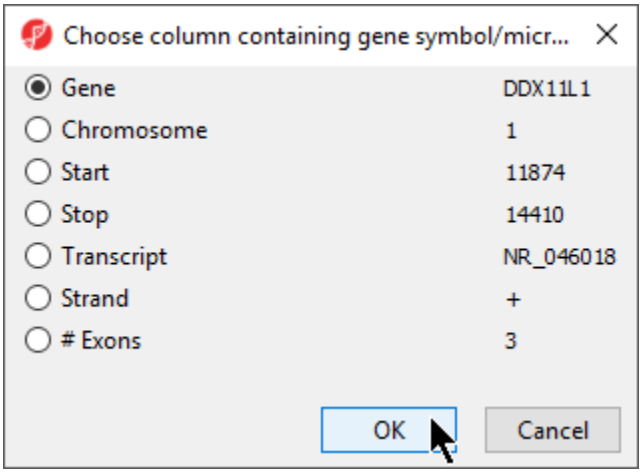


Figure 12. Choosing the column in the annotation file with gene symbols or microRNA names

- Select the appropriate column; here the default choice of **1. Symbol** is appropriate
- Select **OK** to return to the *Configure Genomic Properties* dialog
- Select the appropriate species and genome build options; we have selected **Homo sapiens** and **hg19** (Figure 6)

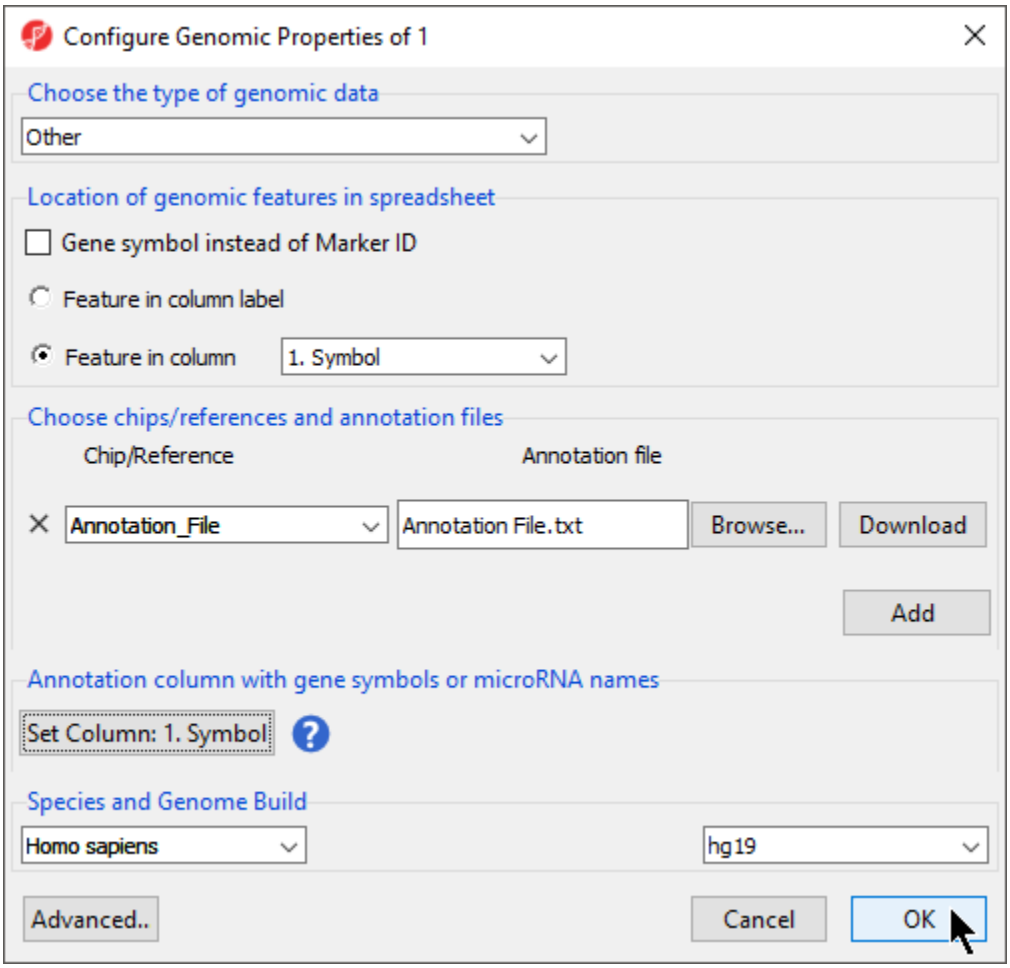


Figure 13. The gene list is now fully configured with an annotation file and reference genome selected

- Select **OK**
- Select ( ) to save the spreadsheet

The annotation file has been associated with the spreadsheet and additional tasks can now be performed on the data, e.g. since the annotation has genomic location, you can draw chromosome view on this data.

Adding annotations to a spreadsheet

Inserting annotations from an annotation file

If an annotation file has been associated with a spreadsheet, annotations from the file can be added as columns in the spreadsheet when each identifier is on a row.

- Right click on a column header
- Select **Insert Annotation**
- Select columns to add from *Column* Configuration; we have selected **Chromosome**, **Start**, and **Stop** (Figure 7)
- Select **OK**

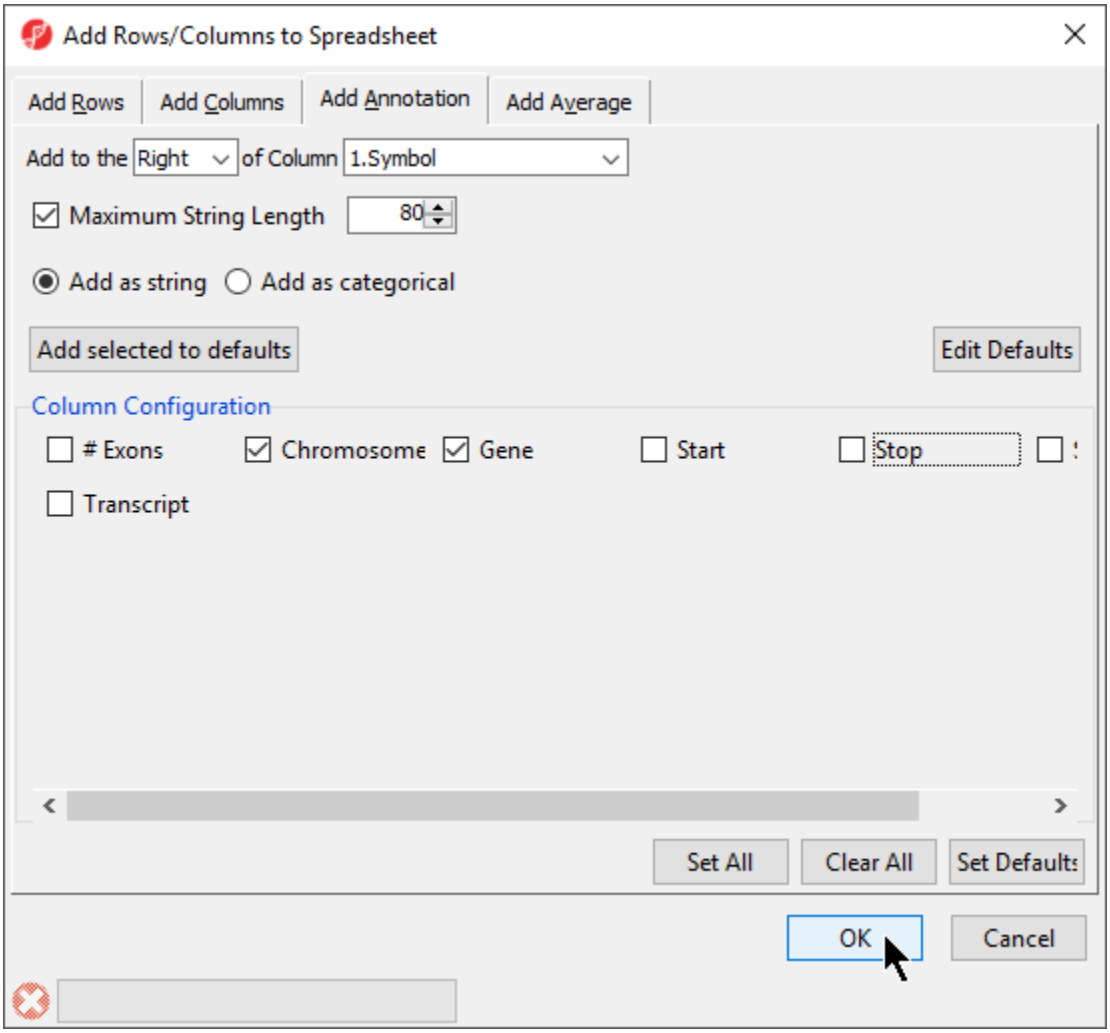


Figure 14. Adding an annotation column from the annotation file

« Importing a text file list Tasks available for a gene list »

Additional Assistance

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

Your Rating:

Results: 35 rates