Adding sample attributes

- Modifying sample attributesAdding sample attributes
- Choosing Sample ID column

Now that the data has been imported, we need to make a few changes to the data annotation before analysis.

Modifying sample attributes

Notice that the Sample ID names in column 1 are gray (Figure 1). This indicates that Sample ID is a text factor. Text factors cannot be used as a variable in downstream analysis so we need to change Sample ID to a categorical factor.

🤣 Partek Genomics Suite - 1 (RNA-seq)	– 🗆 X
<u>Eile Edit Transform View Stat Filter Tools Window Custom H</u> elp	Workflows RNA-Seg ~
Analysis X	RNA-Seq
■ → ↓ ½ ↓	└─ Import
	Import and Manage Samples
1 (RNA-seq)	Add Sample Attributes
1. 2. Sample ID Number of	> Analyze Known Genes
Alignments	> Allele-Specific Analysis
1. Heart 11199653 2. Brain 10984232	> Visualization
3. Musde 14486541	> Biological Interpretation
4. Liver 11583902	
× ·	
< > Rows: 4 Columns: 2 <	> V Related: Analyze a Partek Flow project

Figure 7. Viewing the imported data in a spreadsheet

- Right-click on the column header to invoke the pop-up menu
- Select Properties (Figure 2)

🤣 Partek Genomics Suite - 1 (RNA File Edit Transform View Sta	Filter T <u>o</u> ols <u>W</u> indow <u>C</u> ustom <u>H</u> elp	>
		Workflows RNA-Seq
Analysis X		RNA-Seq
🗅 🧀 🗖 🗙 🛛 💒	 2 ⊕ 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	✓ Import
		Import and Manage Samples
1 (RNA-seq)	^ _Current Selection Heart	Add Sample Attributes
	1. 2. San Copy	> Analyze Known Genes
		> Allele-Specific Analysis
		> Visualization
	2. Brai Sort Ascending 3. Mus Sort Descending	> Biological Interpretation
	4. Live Fill Column Split Column Split Column Find / Replace / Select Change Case Change Case Filter Include Filter Include Filter Exclude Insert Delete Delete Filt Columns Create List from Column Labels Create List With Occurrence Counts Create List Properties Properties Selection Selection Selection <th></th>	
< >	Rows: 4 Columns: 2 <	> V Related: Analyze a Partek Flow project

Figure 8. Changing column properties

• Configure the Properties of Column 1 in Spreadsheet 1 dialog as shown (Figure 3) with Type set to categorical and Attribute set to factor

🥬 Prop	erties of Column 1 in Spreadsheet 1 X
Column La	abel:
Sample II	0
Type:	categorical V String Size: 15
Attribute	factor V Random Effect
	OK Cancel

Figure 9. Changing column 1 properties

Select OK

The samples names in column 1 are now black, indicating that they have been changed to a categorical variable. Next, we will add attributes for grouping the data.

Adding sample attributes

• From the RNA-seq workflow panel, select Add Sample Attributes to bring up the Add Sample Attributes dialog (Figure 4)

🤣 Add Sample Attributes	×
Specify Type Select a type of sample attribute to add to the spreadsheet.	
O Add Attributes from an Existing Column	
Add a Categorical Attribute	
○ Add a Numeric Attribute	
OK Cancel	

Figure 10. Add Sample Attributes dialog

- Select Add a Categorical Attribute
- Select OK to bring up the Create categorical attribute dialog

Creating a categorical sample attribute allows us to group samples. This is useful for designating samples as replicates, as members of an experimental group, or as sharing a phenotype of interest. In this tutorial, we have four different samples from different tissues and different donors, but to illustrate the available statistical analysis options, we need to divide the samples into two groups: *muscle* (Heart and Muscle) and *not muscle* (Brain and Liver).

- Set Attribute name: as Tissue
- Rename Group 1 to muscle and Group 2 to not muscle
- Select and drag the samples from the Unassigned panel to the correct group panel (Figure 5)

🤣 Create categorical attribute						×
Specify the name of the new attribute to be created. Attribute name: Tissue						
Attribute groups Select and drag the samples from the "Unassigned" list o	n the left to	the appropr	riate group on the right	Use the shift and control	l keys to select multi	nle
samples.	in the left to	ле арргорг	hate group on the right.	ose the shint and control	i keys to select mara	JIC
Unassigned	Gro	oup Name:	muscle	(2 samples)	Remove Grou	up
		Heart Muscle		-		
						-
	Gro	oup Name:	not muscle	(2 samples)	Remove Grou	up
		Brain Liver				<u> </u>
	-					~
					New	Group
					ок	Cancel
						Galicei

Copyright © 2018 by Partek Incorporated. All Rights Reserved. Reproduction of this material without express written consent from Partek Incorporated is strictly prohibited.

Figure 11. Creating a categorical attribute

- Select OK
- Select No from the Add another attribute? dialog
- Select Yes from the Save spreadsheet 1 dialog

The attribute will now appear as a new column in the RNA-seq spreadsheet with the heading Tissue and the groups muscle and not muscle.

Choosing Sample ID column

The next available step in the *Import* panel of the *RNA-seq* workflow is **Choose Sample ID Column**. Verifying the correct column is designated the Sample ID becomes particularly important when data from multiple experiments is being combined.

- Select Choose Sample ID Column from the Import panel of the RNA-Seq workflow
- Select OK (Figure 6)

🤣 Choose Sample ID Column			
The sample ID column is required for integrated analysis (using the filename is not recommended). The specified sample IDs must match the sample IDs from the spreadsheet with which you want to integrate. Sample IDs are case sensitive. If you don't have a column that you want to use as Sample ID, please use "Add sample attributes" from the workflow to add a Sample ID column.			
sample attributes from the worknow to	add a Sample 15 column.		
Sample ID Column	1. Sample ID	\sim	
	Sample ID Preview	^	
	Heart		
	Brain		
	Muscle		
	Liver		
		ς.	
	< >		
	OK Cancel		

Figure 12. Choosing the correct column as Sample ID

« Importing aligned reads RNA-Seq mRNA quantification »

Additional Assistance

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

