

## TO PROCESS AFFYMETRIX CEL FILES USING ASCAT PROCESSING:

To process Affymetrix data from CEL files using the Allele Specific Copy Number Analysis of Tumors (ASCAT) algorithm (Van Loo et. al. Proc Natl Acad Sci USA. 2010 Sep 28;107(39):16910-5), you must load the data using a sample descriptor (the one-click method of data loading does not work for ASCAT data types). With the sample descriptor method, you will specify the input file location and assign factors to your samples by specifying them in the sample descriptor. When data is loaded factors will already be associated and will appear in the **Data Set** tab.

### DATA LOADING USING A SAMPLE DESCRIPTOR

1. You need to create a sample descriptor file (a tab delimited text file) to load data into Nexus. In the **Templates** folder of the Nexus installation directory, you will find several arrays types that Nexus is able to process. Open the folder titled **Affymetrix** and select the **AffyCEL-ASCAT-Template.txt** template file for your sample descriptor. It may be easiest to edit this file by opening it in a spreadsheet application such as Excel. If you have technical replicates, please refer to the document called *Processing Replicate Data* to see how to specify this in the sample descriptor.
2. In the first column, **Sample Name**, specify a unique name for each sample that you are loading (e.g. sample1, sample2, etc.). Note that you cannot use the following special characters `\/:*?"<>|` within the sample name. In the second column, **File**, specify the location of your input data file. In the **Matched File** column, specify the location of the matched normal file for your sample. You can use either the full path (e.g. C:\Nexus Projects\Brain Tumor \sample1.CEL) or a relative path (relative to where you will save this sample descriptor file). If the descriptor is in the same folder as the sample data files then you can specify just the name (e.g. sample1.CEL) without any path qualifiers. You can add an unlimited number of additional columns to specify Factors (clinical annotations) for each sample. The column header is **Factor:** followed by a name for the Factor. Save this file with a new name and make sure to save it as a tab-delimited text file.

Example of a sample descriptor for Affymetrix CEL ASCAT data type:

Data Type:	Affymetrix CEL ASCAT	
Sample Name	File	Matched File
S1	293040T.CEL	293040N.CEL
S2	293042T.CEL	293042N.CEL

For 500K arrays, you need to provide Nsp and Sty files for both the tumor and normal samples. Below is an example of a sample descriptor for such data:

Data Type:	Affymetrix CEL ASCAT			
Sample Name	File1	Matched File1	File2	Matched File2
S1	293040T_nsp.CEL	293040N_nsp.CEL	293040T_sty.CEL	293040N_sty.CEL
S2	293042T_nsp.CEL	293042N_nsp.CEL	293042T_sty.CEL	293042N_sty.CEL

3. Run Nexus and click the **Create a New Project** or **Open Existing Project** button. If creating a new project, in the resulting dialog, specify a name for your project and choose a genome and then the appropriate genome build based on the mapping information for your probes. If adding additional samples to an existing project (of the same genome build), select your project folder in the resulting dialog.

4. On the next screen, click the **Load** button, select **Load Descriptor** from the dropdown and choose the sample descriptor file you saved in step 2 above. You will see all your samples loaded into Nexus.
5. Click **View** to process the data and see the results.