To Process Data for Matched Paired Analysis:

Matched paired analysis is typically performed on a pair of samples (tumor vs. normal) from the same individual. The tumor is compared to the normal to identify those changes that are present in the tumor sample only. To load matched paired data into Nexus, a sample descriptor file must be used. The Load Data method cannot be used. 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 (a tab delimited text file). When data is loaded, factors will already be associated with the samples and will appear in the **Data Set** tab.

The Sample Descriptor file basically states what type of data is being loaded, the sample name, where to find the files on the hard drive and potentially and clinical factors. The Sample Descriptor for paired analysis is basically the same as that for standard analysis but requires an additional "Matched File" column which specifies the location of the matched normal file.

1. You need to create a sample descriptor file (a tab delimited text file) to load data into Nexus. In the T**emplates** folder of the Nexus installation directory, you will find several arrays types that Nexus is able to process. Select the appropriate template for your data type. For Affymetrix and Illumina data, please see below for special instructions. It may be easiest to edit this file by opening it in a spreadsheet application such as Excel.

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 and in the **Matched File** column provide the file for the matched normal 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. Those are the only required columns. 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.

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 Descriptor** button 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.

NOTES FOR AFFYMETRIX ARRAYS
Open the folder titled **Affymetrix** and select one of the files with "paired matching" in the name.

For 6.0 CEL files use **AffyCEL-Template-for-paired-matching.txt**

For 500K Nsp/Sty CEL files, use **AffyCEL-Template-for-paired-matching-with-Sty-Nsp.txt**

For CNCHP files, use **AffyCNCHP-Template-for-paired-matching.txt**

For 500K Nsp/Sty CNCHP files, use **AffyCNCHP-Template-for-paired-matching-with-Sty-Nsp.txt**

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					
Sample Name	File1	Matched File1	File2	Matched File2	Factor:Age	Factor:Gender
Sample 1	tumor1_nsp.CEL	normal1_nsp.CEL	tumor1_sty.CEL	tumor1_sty.CEL	58	Male
Sample 2	tumor2_nsp.CEL	normal2_nsp.CEL	tumor2_sty.CEL	tumor2_sty.CEL	47	Male
Sample 3	tumor3_nsp.CEL	normal3_nsp.CEL	tumor3_sty.CEL	tumor3_sty.CEL	55	Female

NOTES FOR ILLUMINA ARRAYS

Open the folder titled **Illumina** and select one of the files with "paired matching" in the name. The template selected depends on whether Sample Name or Sample ID was the column chosen to create the Final Report from BeadStudio or GenomeStudio.

If Sample ID was selected, use **Illumina-SampleID-paired_matching-Template.txt**

If Sample Name was selected use **Illumina-SampleName-paired_matching-Template.txt**

Example sample descriptor using Sample ID:

Data Type:	Illumina			
Sample Name	File	File Sample ID	Matched File	Matched File Sample ID
S1	Tumor.txt	S1_Tumor	Blood.txt	S1_Blood
S2	Tumor.txt	S2_Tumor	Blood.txt	S2_Blood
S3	Tumor.txt	S3_Tumor	Blood.txt	S3_Blood
S4	Tumor.txt	S4_Tumor	Blood.txt	S4_Blood
S5	Tumor.txt	S5_Tumor	Blood.txt	S5_Blood
S6	Tumor.txt	S6_Tumor	Blood.txt	S6_Blood