Partek® Methylation User Guide

Introduction

This user guide will explain the different types of workflow that can be used to analyze methylation datasets.

Under the Partek® Methylation workflow there are three different sub-workflows (Figure 1) aimed at different types of data, but also different objectives. They are:

1. Illumina BeadArray Methylation
2. Next-Gen Sequencing Methylation
3. ChIP-Chip Methylation

Figure 1: Methylation sub-workflows

We can categorize the methylation data into two categories: Next Generation Sequencing (NGS) methylation datasets, and methylation array datasets.

Next Generation Sequencing (NGS) Methylation Data

For NGS methylation data, the sub-workflow to use is "Next-Gen Sequencing Methylation". The workflow is shown in Figure 2, and is mainly used to analyze NGS methylation data from the MeDIP-seq (Methylated DNA Immunoprecipitation) assay, i.e. an assay to enrich methylated DNA sequences.

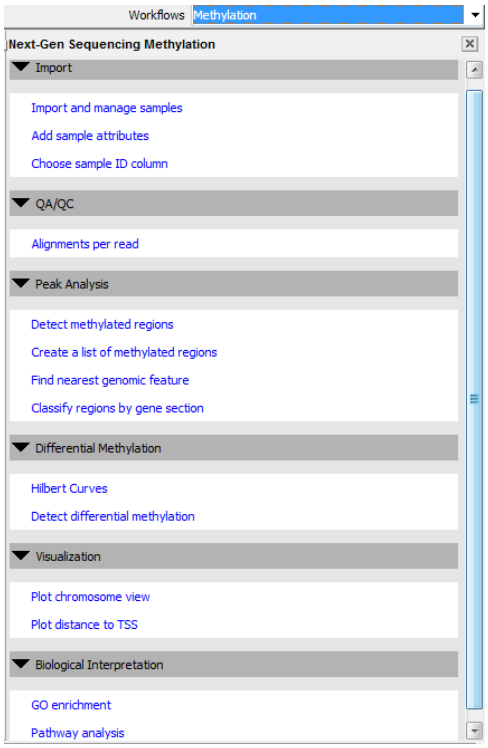


Figure 2: Next-Gen Sequencing Methylation Workflow

MeDIP-seq is a "pull-down" assay, therefore the "Next-Gen Sequencing Methylation" workflow is very much similar to peak detection in the ChIP-Seq workflow. In the "Peak Analysis" section you will find the function "Detect methylated regions" which is the same as peak detection; for further information, please refer to the ChIP-seq data analysis tutorial.

In addition, in the "Next-Gen Sequencing Methylation" workflow there is also an option for ***Detecting differential methylation***. In order to detect regions of differential methylation, the genome is divided into non-overlapping windows of a specified window length. A scaled fold-change and a binomial p-value are then calculated for each window; differential methylation will be reported as significant if it meets the p-value cut-off. The dialog box for detecting differential methylated regions is shown in Figure 3.

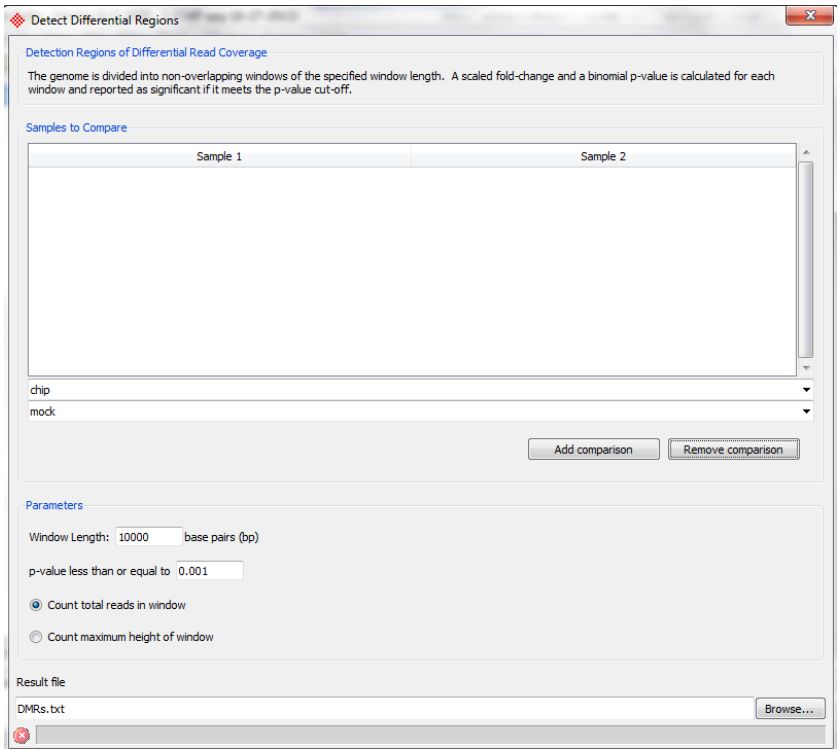


Figure 3: Dialog box for detecting differentially methylated region in "Next-Gen Sequencing Methylation" workflow

What shall I do if I have RRBS (Reduced Representation Bisulfite Sequencing) data?

If you have RRBS data and you would like to detect differentially methylated regions, you will need the "percentage of methylation" value for each of your samples. This percentage methylation value can be a single-site specific percentage methylation value, or a CpG island percentage methylation value. For Illumina data, you can get the percentage methylation value using Illumina CASAVA pipeline.

The percentage methylation value is in a text format and therefore you can import into Partek® as a text file (File > Import > Text (.csv .txt) for data analysis. Subsequently, use either the "Gene Expression" workflow or the "Illumina BeadArray Methylation" sub-workflow (under the Methylation workflow) to detect differentially methylated sites/regions.

Methylation Array Data

Illumina BeadArray Methylation

If you are analyzing an Illumina Infinium Methylation array you can simply follow the Partek® "Illumina BeadArray Methylation" workflow (Figure 4). This workflow is very similar to the "Gene Expression" workflow.

Note: For the import of Illumina Methylation Data you are advised to load a Partek® project following Illumina GenomeStudio export using the Partek Plugin. This Partek® project file consists of two spreadsheets: i) a beta value spreadsheet and ii) an intensity spreadsheet. The beta value spreadsheet contains the percentage methylation values, and this will be the spreadsheet used in the downstream analysis for detecting differentially methylated sites.

Figure 4: Illumina BeadArray Methylation workflow

Note: In order to load a project following Illumina GenomeStudio export, refer to: http://www.partek.com/Tutorials/microarray/User_Guides/GenomeStudioMethylationPlugin.pdf

I would like to summarize the single site percentage methylation value into a region percentage methylation value. How can I do that?

Assuming that you would like to summarize the single site percentage methylation value into a **CpG island** percentage methylation value, you will need the file specifying the genomic coordinates of the CpG island (typically a .bed file).

For example, if you are interested in looking at hg18 CpG islands, you can download the bed file from the UCSC site:
http://hgdownload.cse.ucsc.edu/goldenPath/hg18/database/cpgIslandExt.txt.gz

Unzip the file and import into Partek (refer to the "Import Region File" user guide).

After importing the region file you can now insert the average "methylation value" for the CpG island by right-clicking on the CpG island column header and then choosing "Insert Average" from the contextual menu (Figure 5).

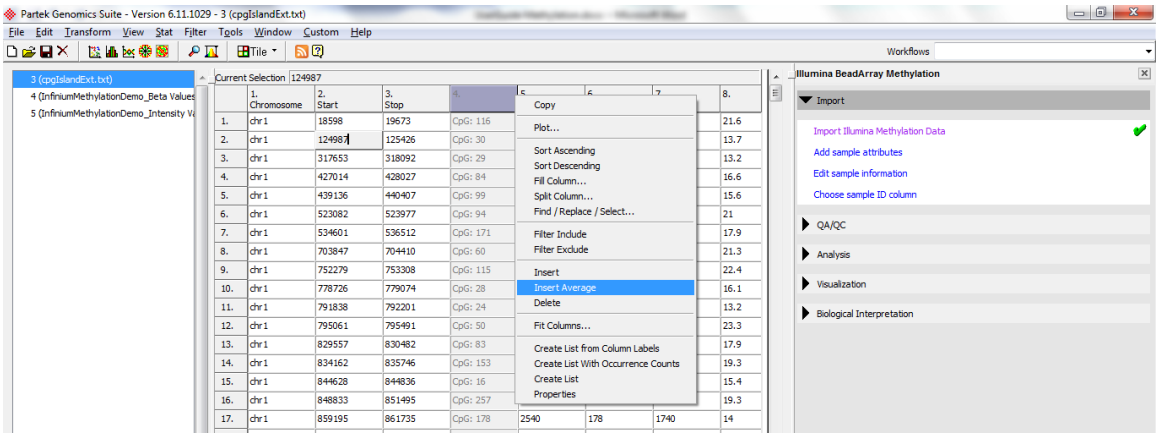


Figure 5: Insert Average

In the dialogue box choose to add the column to the right of column 4. Get average from spreadsheet "beta value" and choose "mean value for all samples separately" (Figure 6).

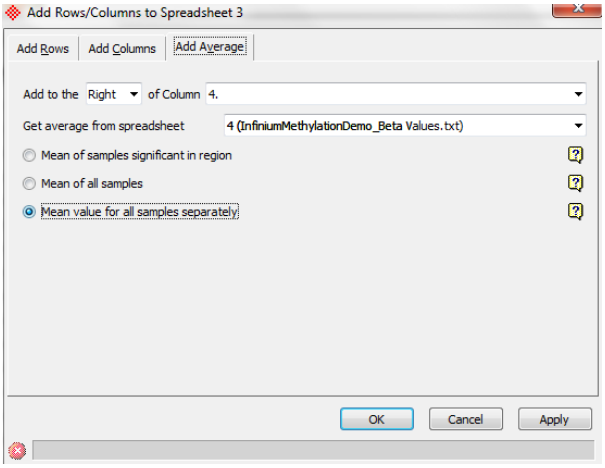
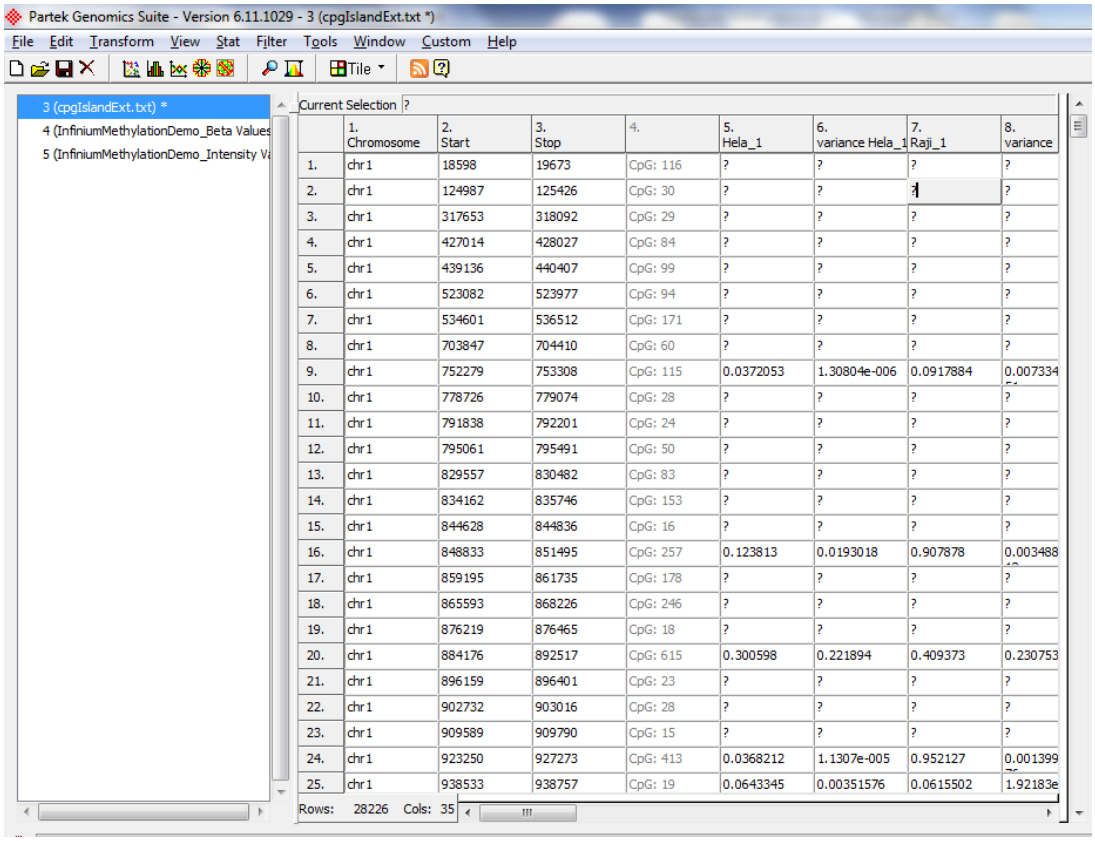


Figure 6: Insert the mean value for all samples separately

The result will appear as below (Figure 7):



	1. Chromosome	2. Start	3. Stop	4.	5. Hela_1	6. variance Hela_1	7. Raji_1	8. variance
1.	chr1	18598	19673	CpG: 116	?	?	?	?
2.	chr1	124987	125426	CpG: 30	?	?	?	?
3.	chr1	317653	318092	CpG: 29	?	?	?	?
4.	chr1	427014	428027	CpG: 84	?	?	?	?
5.	chr1	439136	440407	CpG: 99	?	?	?	?
6.	chr1	523082	523977	CpG: 94	?	?	?	?
7.	chr1	534601	536512	CpG: 171	?	?	?	?
8.	chr1	703847	704410	CpG: 60	?	?	?	?
9.	chr1	752279	753308	CpG: 115	0.0372053	1.30804e-006	0.0917884	0.007334
10.	chr1	778726	779074	CpG: 28	?	?	?	?
11.	chr1	791838	792201	CpG: 24	?	?	?	?
12.	chr1	795061	795491	CpG: 50	?	?	?	?
13.	chr1	829557	830482	CpG: 83	?	?	?	?
14.	chr1	834162	835746	CpG: 153	?	?	?	?
15.	chr1	844628	844836	CpG: 16	?	?	?	?
16.	chr1	848833	851495	CpG: 257	0.123813	0.0193018	0.907878	0.003488
17.	chr1	859195	861735	CpG: 178	?	?	?	?
18.	chr1	865593	868226	CpG: 246	?	?	?	?
19.	chr1	876219	876465	CpG: 18	?	?	?	?
20.	chr1	884176	892517	CpG: 615	0.300598	0.221894	0.409373	0.230753
21.	chr1	896159	896401	CpG: 23	?	?	?	?
22.	chr1	902732	903016	CpG: 28	?	?	?	?
23.	chr1	909589	909790	CpG: 15	?	?	?	?
24.	chr1	923250	927273	CpG: 413	0.0368212	1.1307e-005	0.952127	0.001399
25.	chr1	938533	938757	CpG: 19	0.0643345	0.00351576	0.0615502	1.92183e

Figure 7: Inserted average methylation value for each CpG island in each sample

Please note that "?" appears for those CpG islands that do not contain any value in the beta-value spreadsheet. As there is more than 1 probe in each CpG island in the original spreadsheet, you will see not only the mean percentage methylation value, but also the variance of the percentage methylation value is added.

Note: To provide a more informative name to Column 4, right-click on the column header and rename as "CpG name".

ChIP-Chip Methylation

This workflow is used to analyse methylation arrays based on affinity pull down assays, i.e. MeDIP (Methylated DNA Immunoprecipitation). Therefore, instead of detecting differentially methylated regions, the workflow looks for enriched methylated region.

The "ChIP-Chip Methylation" workflow is the same as the "Tiling" workflow, therefore please refer to the "Tiling" workflow tutorial for further information.

Figure 8: ChIP-Chip Methylation Workflow

End of User Guide

This is the end of the user guide. If you need additional assistance, you may call our technical support staff at +1-314-878-2329 or email *support@partek.com*.