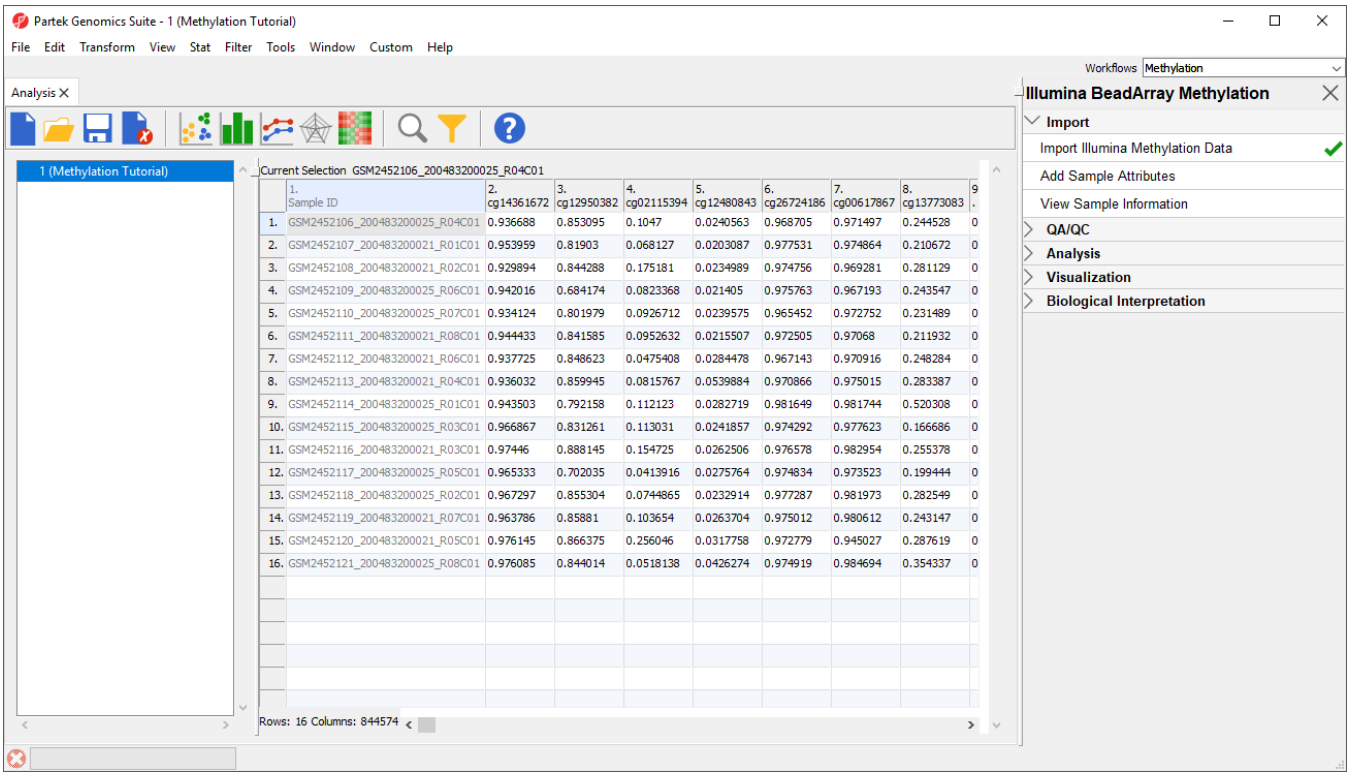
# Annotate samples

Each row of the spreadsheet (Figure 1) corresponds to a single sample. The first column is the names of the .idat files and the remaining columns are the array probes. The table values are -values, which correspond to the percentage methylation at each site. A -value is calculated as the ratio of methylated probe intensity over the overall intensity at each site (the overall intensity is the sum of methylated and unmethylated probe intensities).




| | 1. Sample ID | 2. cg14361672 | 3. cg12950382 | 4. cg02115394 | 5. cg12480843 | 6. cg26724186 | 7. cg00617867 | 8. cg13773083 | 9. |
|---|---|---|---|---|---|---|---|---|---|
| 1. | GSM2452106_200483200025_R04C01 | 0.936688 | 0.853095 | 0.1047 | 0.0240563 | 0.968705 | 0.971497 | 0.244528 | 0 |
| 2. | GSM2452107_200483200021_R01C01 | 0.953959 | 0.81903 | 0.068127 | 0.0203087 | 0.977531 | 0.974864 | 0.210672 | 0 |
| 3. | GSM2452108_200483200021_R02C01 | 0.929894 | 0.844288 | 0.175181 | 0.0234989 | 0.974756 | 0.969281 | 0.281129 | 0 |
| 4. | GSM2452109_200483200025_R06C01 | 0.942016 | 0.684174 | 0.0823368 | 0.021405 | 0.975763 | 0.967193 | 0.243547 | 0 |
| 5. | GSM2452110_200483200025_R07C01 | 0.934124 | 0.801979 | 0.0926712 | 0.0239575 | 0.965452 | 0.972752 | 0.231489 | 0 |
| 6. | GSM2452111_200483200021_R08C01 | 0.944433 | 0.841585 | 0.0952632 | 0.0215507 | 0.972505 | 0.97068 | 0.211932 | 0 |
| 7. | GSM2452112_200483200021_R06C01 | 0.937725 | 0.848623 | 0.0475408 | 0.0284478 | 0.967143 | 0.970916 | 0.248284 | 0 |
| 8. | GSM2452113_200483200021_R04C01 | 0.936032 | 0.859945 | 0.0815767 | 0.0539884 | 0.970866 | 0.975015 | 0.283387 | 0 |
| 9. | GSM2452114_200483200025_R01C01 | 0.943503 | 0.792158 | 0.112123 | 0.0282719 | 0.981649 | 0.981744 | 0.520308 | 0 |
| 10. | GSM2452115_200483200025_R03C01 | 0.966867 | 0.831261 | 0.113031 | 0.0241857 | 0.974292 | 0.977623 | 0.166686 | 0 |
| 11. | GSM2452116_200483200021_R03C01 | 0.97446 | 0.888145 | 0.154725 | 0.0262506 | 0.976578 | 0.982954 | 0.255378 | 0 |
| 12. | GSM2452117_200483200025_R05C01 | 0.965333 | 0.702035 | 0.0413916 | 0.0275764 | 0.974834 | 0.973523 | 0.199444 | 0 |
| 13. | GSM2452118_200483200025_R02C01 | 0.967297 | 0.855304 | 0.0744865 | 0.0232914 | 0.977287 | 0.981973 | 0.282549 | 0 |
| 14. | GSM2452119_200483200021_R07C01 | 0.963786 | 0.85881 | 0.103654 | 0.0263704 | 0.975012 | 0.980612 | 0.243147 | 0 |
| 15. | GSM2452120_200483200021_R05C01 | 0.976145 | 0.866375 | 0.256046 | 0.0317758 | 0.972779 | 0.945027 | 0.287619 | 0 |
| 16. | GSM2452121_200483200025_R08C01 | 0.976085 | 0.844014 | 0.0518138 | 0.0426274 | 0.974919 | 0.984694 | 0.354337 | 0 |

*Figure 1. Spreadsheet after .idat file import: samples on rows (Sample IDs are based on file names), probes on columns, cell values are functionally normalized beta values (default settings)*

Before we can perform any analysis, the study samples need to be organized into their experimental groups.

- Select **Add Sample Attributes** from the *Import* section of the *Illumina BeadArray Methylation* workflow
- Select **Add a Categorical Attribute** from the *Add Sample Attributes* dialog (Figure 2)

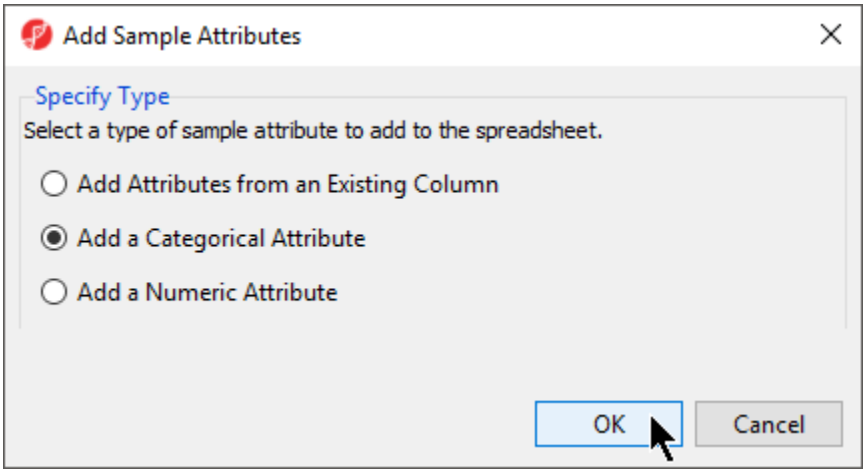



*Figure 2. Adding sample attributes. Adding Attributes from an Existing Column can be used to split file names into sections, based on delimiters (e.g. _, -, space etc.). Adding a Numeric or Categorical Attribute enables the user to manually specify sample attributes*

- Select **OK**

The *Create categorical attribute* dialog allows us to create groups for a categorical attribute. By default, two groups are created, but additional groups can be added.

- Set *Attribute name:* to *Cell Type*
- Rename the groups *B cells* and *LCLs*
- Drag and drop the samples from the *Unassigned* list to their groups as listed in the table below

| Sample ID | Cell Type |
|---|---|
| GSM2452106_200483200025_R04C01 | B cells |
| GSM2452107_200483200021_R01C01 | B cells |
| GSM2452108_200483200021_R02C01 | B cells |
| GSM2452109_200483200025_R06C01 | B cells |
| GSM2452110_200483200025_R07C01 | B cells |
| GSM2452111_200483200021_R08C01 | B cells |
| GSM2452112_200483200021_R06C01 | B cells |
| GSM2452113_200483200021_R04C01 | B cells |
| GSM2452114_200483200025_R01C01 | LCLs |
| GSM2452115_200483200025_R03C01 | LCLs |
| GSM2452116_200483200021_R03C01 | LCLs |
| GSM2452117_200483200025_R05C01 | LCLs |
| GSM2452118_200483200025_R02C01 | LCLs |
| GSM2452119_200483200021_R07C01 | LCLs |
| GSM2452120_200483200021_R05C01 | LCLs |
| GSM2452121_200483200025_R08C01 | LCLs |

There should now be two groups with eight samples in each group (Figure 3).

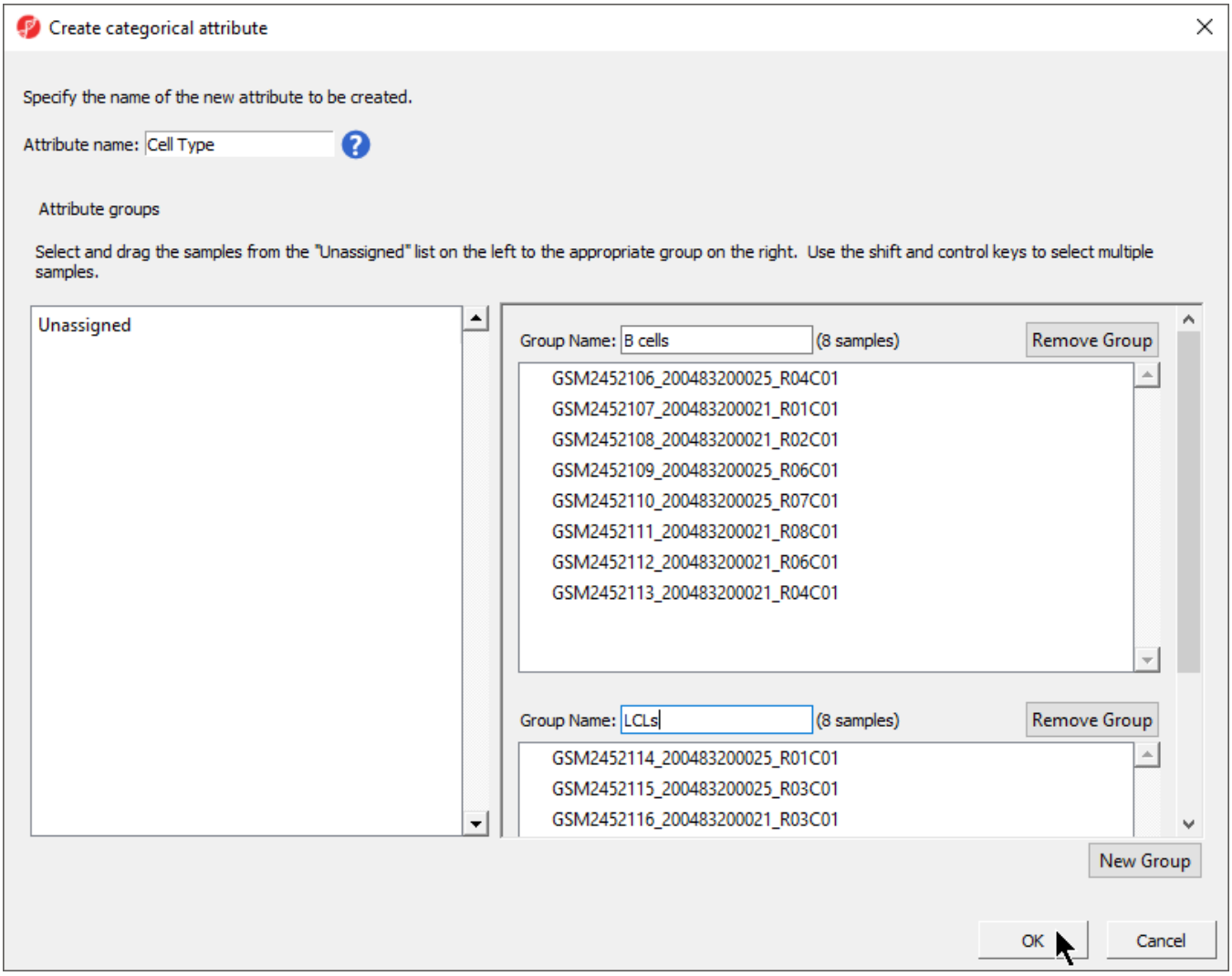



*Figure 3. Adding Cell Type attribute, assigning samples.*

*Figure 3. Adding Cell Type attribute as a categorical group*

- Select **OK**
- Select **Yes** from the *Add another categorical attribute* dialog
- Set *Attribute name:* to *Gender*
- Rename the groups *Male* and *Female*
- Drag and drop the samples from the *Unassigned* list to their groups as listed in the table below

| Sample ID | Gender |
|---|---|
| GSM2452106_200483200025_R04C01 | Female |
| GSM2452107_200483200021_R01C01 | Female |
| GSM2452108_200483200021_R02C01 | Male |
| GSM2452109_200483200025_R06C01 | Female |
| GSM2452110_200483200025_R07C01 | Female |
| GSM2452111_200483200021_R08C01 | Female |
| GSM2452112_200483200021_R06C01 | Female |
| GSM2452113_200483200021_R04C01 | Male |
| GSM2452114_200483200025_R01C01 | Female |
| GSM2452115_200483200025_R03C01 | Female |
| GSM2452116_200483200021_R03C01 | Male |
| GSM2452117_200483200025_R05C01 | Female |
| GSM2452118_200483200025_R02C01 | Female |
| GSM2452119_200483200021_R07C01 | Female |
| GSM2452120_200483200021_R05C01 | Female |
| GSM2452121_200483200025_R08C01 | Male |

There should now be two groups with four samples in *Male* and twelve samples in *Female* (Figure 4).

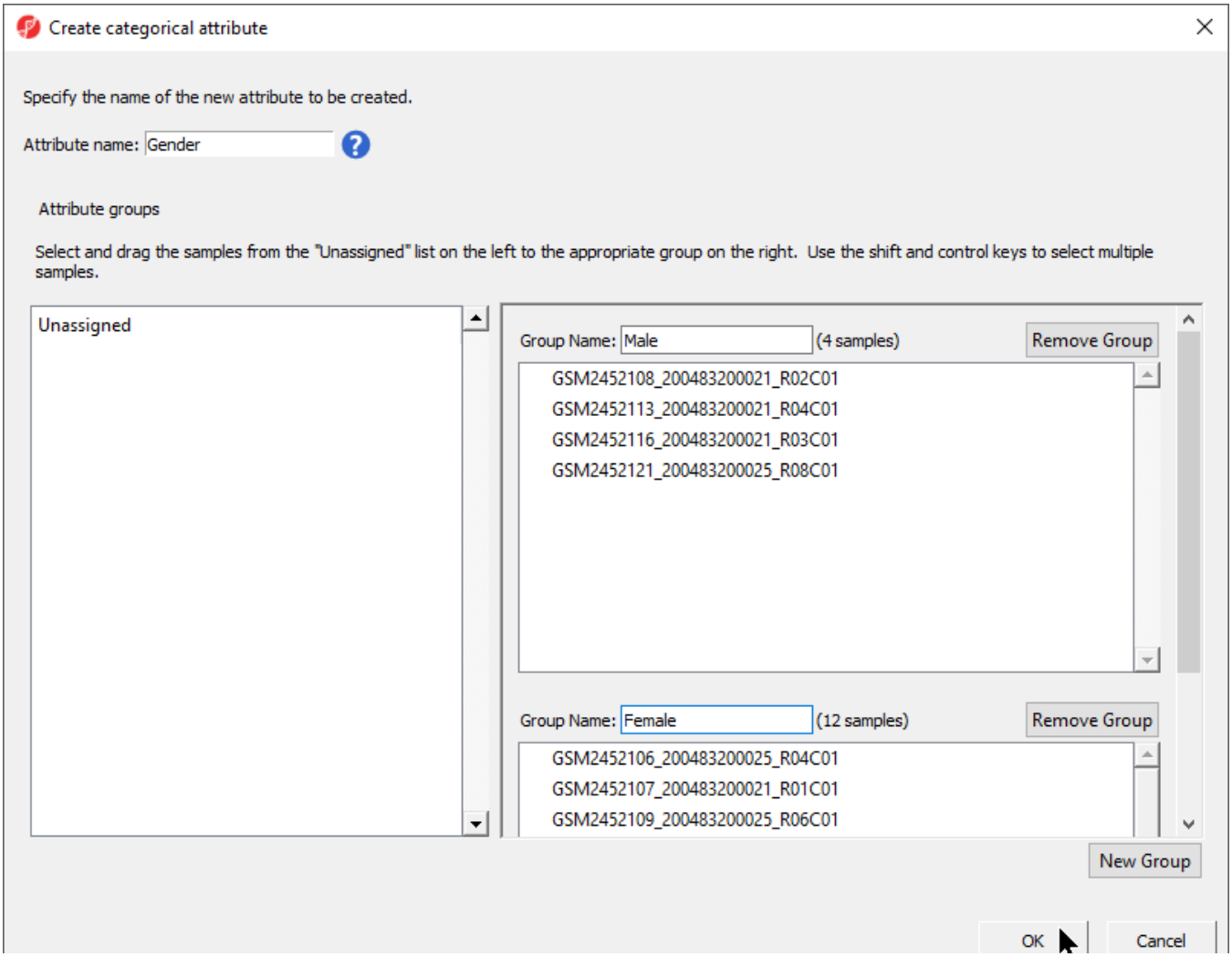
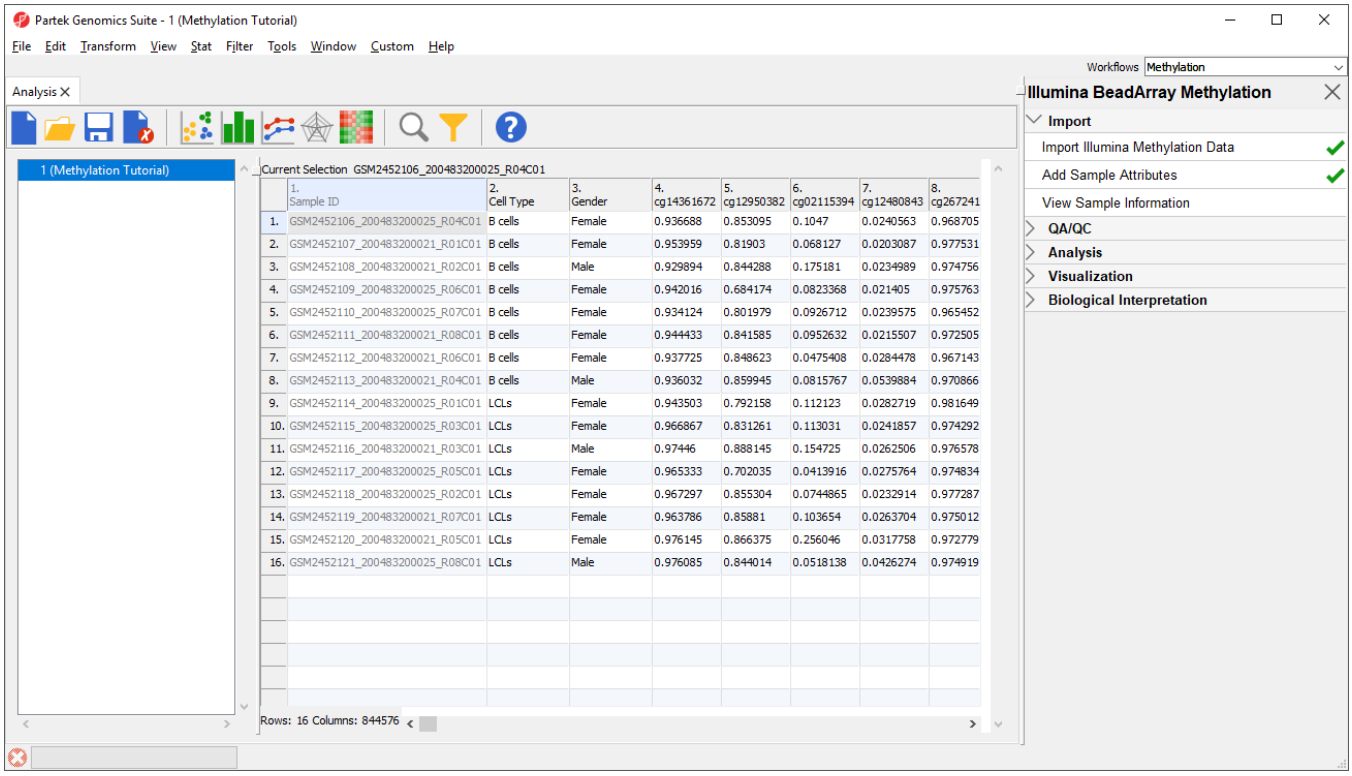


*Figure 4. Adding Gender attribute as a categorical group*

- Select **OK**
- Select **No** from the *Add another categorical attribute* dialog
- Select **Yes** to save the spreadsheet

Two new columns have been added to spreadsheet *1 (Methylation)* with the cell type and gender of each sample (Figure 5).




| | 1. Sample ID | 2. Cell Type | 3. Gender | 4. cg14361672 | 5. cg12950382 | 6. cg02115394 | 7. cg12480843 | 8. cg267241 |
|---|---|---|---|---|---|---|---|---|
| 1. | GSM2452106_200483200025_R04C01 | B cells | Female | 0.936688 | 0.853095 | 0.1047 | 0.0240563 | 0.968705 |
| 2. | GSM2452107_200483200021_R01C01 | B cells | Female | 0.953959 | 0.81903 | 0.068127 | 0.0203087 | 0.977531 |
| 3. | GSM2452108_200483200021_R02C01 | B cells | Male | 0.929894 | 0.844288 | 0.175181 | 0.0234989 | 0.974756 |
| 4. | GSM2452109_200483200025_R06C01 | B cells | Female | 0.942016 | 0.684174 | 0.0823368 | 0.021405 | 0.975763 |
| 5. | GSM2452110_200483200025_R07C01 | B cells | Female | 0.934124 | 0.801979 | 0.0926712 | 0.0239575 | 0.965452 |
| 6. | GSM2452111_200483200021_R08C01 | B cells | Female | 0.944433 | 0.841585 | 0.0952632 | 0.0215507 | 0.972505 |
| 7. | GSM2452112_200483200021_R06C01 | B cells | Female | 0.937725 | 0.848623 | 0.0475408 | 0.0284478 | 0.967143 |
| 8. | GSM2452113_200483200021_R04C01 | B cells | Male | 0.936032 | 0.859945 | 0.0815767 | 0.0539884 | 0.970866 |
| 9. | GSM2452114_200483200025_R01C01 | LCLs | Female | 0.943503 | 0.792158 | 0.112123 | 0.0282719 | 0.981649 |
| 10. | GSM2452115_200483200025_R03C01 | LCLs | Female | 0.966867 | 0.831261 | 0.113031 | 0.0241857 | 0.974292 |
| 11. | GSM2452116_200483200021_R03C01 | LCLs | Male | 0.97446 | 0.888145 | 0.154725 | 0.0262506 | 0.976578 |
| 12. | GSM2452117_200483200025_R05C01 | LCLs | Female | 0.965333 | 0.702035 | 0.0413916 | 0.0275764 | 0.974834 |
| 13. | GSM2452118_200483200025_R02C01 | LCLs | Female | 0.967297 | 0.855304 | 0.0744865 | 0.0232914 | 0.977287 |
| 14. | GSM2452119_200483200021_R07C01 | LCLs | Female | 0.963786 | 0.85881 | 0.103654 | 0.0263704 | 0.975012 |
| 15. | GSM2452120_200483200021_R05C01 | LCLs | Female | 0.976145 | 0.866375 | 0.256046 | 0.0317758 | 0.972779 |
| 16. | GSM2452121_200483200025_R08C01 | LCLs | Male | 0.976085 | 0.844014 | 0.0518138 | 0.0426274 | 0.974919 |

*Figure 5. Annotated beta values spreadsheet*

« Import and normalize methylation data  Perform data quality analysis and quality control »

## 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