Differential methylation analysis (of Processed Data) with limma

Nivretta Thatra

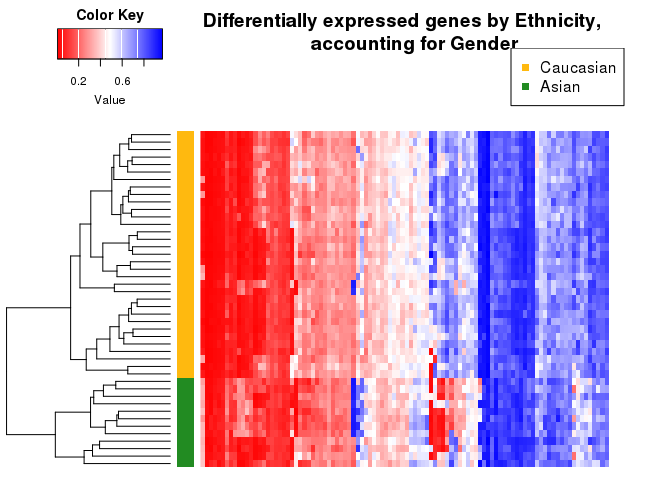
Here we used a linear model to identify differentially methylated probes with limma, as shown in this older 540 [seminar](http://www.ugrad.stat.ubc.ca/~stat540/seminars/seminar08_methylation.html).

We first fit a linear model with ethnicity as the only covariate to obtain top differentially methylated CpG sites (testing the null hypothesis that samples in the two ethnic groups are drawn from populations with the same mean CpG site methylation). Using a cutoff of FDR = 0.01, we identified 106 CpG sites that are differentially methylated between Caucasian and Asian genetic ancestry.

However, DNA methylation is known to associate with gender, so the interaction effect ethnicity and gender was accounted for in another linear model. Using a cutoff of FDR = 0.01, we identified just 13 CpG sites that are differentially methylated between Caucasian and Asian genetic ancestry, when the interaction effect of ethnicity and gender was accounted for. Here are 6 of those sites:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | logFC | AveExpr | t | P.Value | adj.P.Val | B |
| cg16329197 | 0.5368513 | 0.4996451 | 9.546477 | 0 | 0.0000020 | 17.319139 |
| cg25025879 | 0.4343004 | 0.4535298 | 9.205678 | 0 | 0.0000028 | 16.272412 |
| cg05393297 | 0.4229273 | 0.6325893 | 8.211144 | 0 | 0.0000427 | 13.142162 |
| cg14581129 | 0.2265901 | 0.5049442 | 6.992750 | 0 | 0.0016940 | 9.185289 |
| cg26513180 | -0.0294624 | 0.0339633 | -6.732052 | 0 | 0.0025085 | 8.327817 |
| cg19041462 | 0.1018915 | 0.8764931 | 6.689685 | 0 | 0.0025085 | 8.188292 |

## Plot top 100 Limma hits when accounting for Ethnicity and gender interaction effect

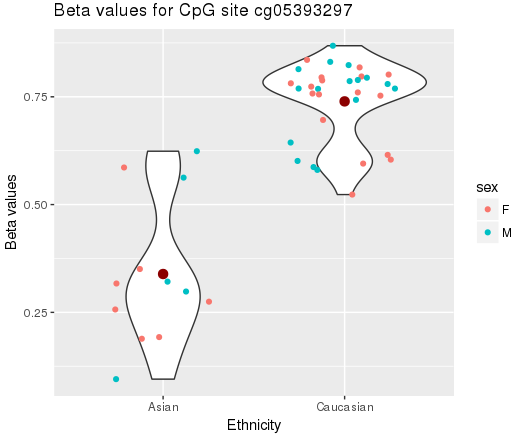
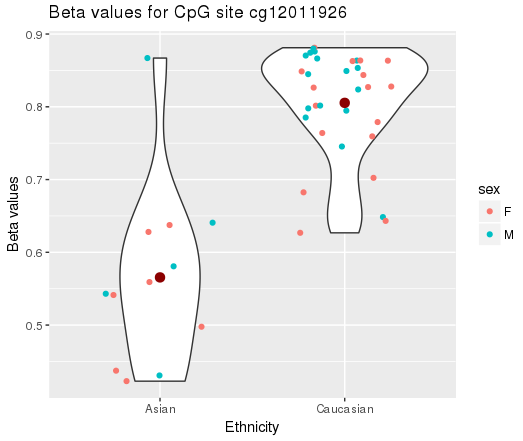


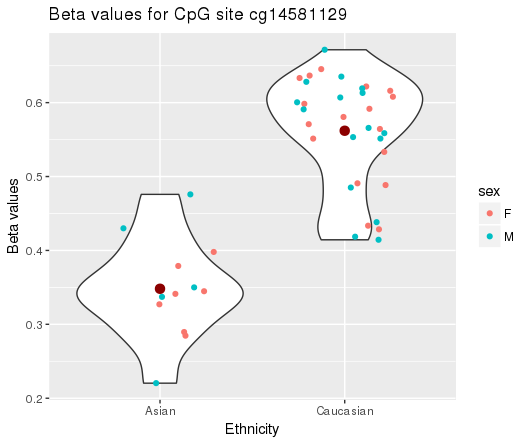
## Check overlap with GLMnet predictions

There is an overlap of 5 CpG sites between the those detected by GLMnet and the ones detected in linear regression analysis.

## SiteID  
## 1 cg05393297  
## 2 cg12011926  
## 3 cg14581129  
## 4 cg16329197  
## 5 cg25025879

**Plots of 3 of these CpG sites:**

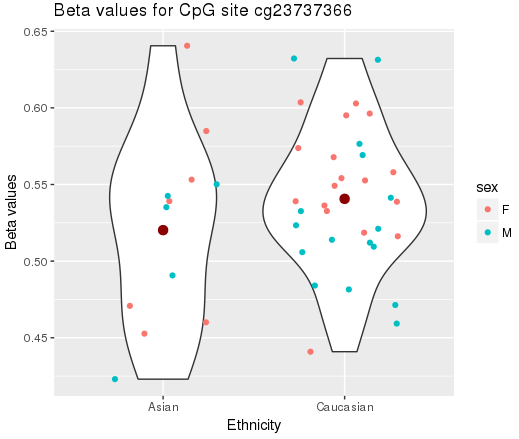


Plot of random CpG site:

RANDOM CPG SITES for comparison:

# 

# 



# Other: table and plot for methylated sites by Ethnicity only

We fit a linear model with ethnicity as the only covariate to obtain top differentially methylated CpG sites. We are testing the null hypothesis that samples in the two ethnic groups are drawn from populations with the same mean CpG site methylation.

Using a cutoff of FDR = 0.01, we identified 106 CpG sites that are differentially methylated between Caucasian and Asian genetic ancestry. Here are the top 6 of these sites:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | logFC | AveExpr | t | P.Value | adj.P.Val | B |
| cg16329197 | 0.4916578 | 0.4996451 | 11.656642 | 0 | 0.0e+00 | 23.94872 |
| cg25025879 | 0.3917075 | 0.4535298 | 10.947130 | 0 | 0.0e+00 | 21.86189 |
| cg05393297 | 0.4004352 | 0.6325893 | 10.444695 | 0 | 0.0e+00 | 20.34470 |
| cg16808927 | -0.2639234 | 0.1481974 | -9.190328 | 0 | 1.0e-06 | 16.41929 |
| cg06903451 | 0.1702557 | 0.5767780 | 9.038860 | 0 | 1.2e-06 | 15.93273 |
| g14581129 | 0.2138548 | 0.5049442 | 8.955025 | 0 | 1.3e-06 | 15.66234 |

