

**10-17-2018**  
**log transformation and FDR 0.05 CpGs**  
**Guannan Shen**  
**ECCHO**

 Zhang, Weiming  
Yesterday, 12:26 PM

t2d\_mom.txt  
185 bytes

top\_CpG\_from\_DMR\_adj...  
23 KB

dmr\_male\_formatted00...  
13 KB

3 attachments (36 KB) Download all Save all to OneDrive - The University of Colorado Denver

Hi,

I just found that I did not exclude the subjects with wrong sex in this DMR analysis. Sorry about that. Attached is the updated lists. We also removed children whose mom has T2D from previous analysis. Attached is the list of T2D moms.

Weiming

---

 Zhang, Weiming  
Yesterday, 10:47 AM  
Shen, Guannan; Starling, Anne; Kechris, Katerina; Yang, Ivana; Borengasser, Sarah J. 5

top\_CpG\_from\_DMR\_adj...  
23 KB

dmr\_male\_formatted00...  
13 KB

wrongsex.txt  
224 bytes

3 attachments (36 KB) Download all Save all to OneDrive - The University of Colorado Denver

Dataset updated

Hi,

Here are the lists for dmr with cutoff 0.05 and their top CpGs. I also attached the list of subjects with wrong sex. They were removed from analyses.

Thanks.

Weiming

## Data Pre-processing

Male 305 subjects

Female 278 subjects

```
## [1] 583 320
## [1] 305 320
## [1] 278 320
## [1] "cg16884940"
## [1] "cg16884940 are not in the TOP300 M values CpG list"
## [1] 583 140
## [1] "now the CpGs are 120 and whole sample size is 583"
## [1] 305 140
## [1] 278 140
```

```
cpg_reg <- function(outcome, data, name, Topn, Gender, ncp) {

  ## outcome lm
  outcome_lm = lapply(21:(ncp + 20), function(i) {
    lm = lm(outcome ~ data[, i] + maternal_age + race_4 +
      Bcell + CD4T + CD8T + Gran + Mono + NK + nRBC, data = data)
    coef = round(summary(lm)$coefficients[2, ], 4)
    return(coef)
  })
  outcome_lm = data.frame(matrix(unlist(outcome_lm), ncol = 4,
    byrow = TRUE, dimnames = list(c(colnames(data)[21:(ncp +
      20)]), c("Estimate", "Std.Error", "t.statistic",
      "p.value"))))

  # adjusted p-value
  outcome_lm = outcome_lm %>% mutate(FDR = p.adjust(p.value,
    "BH", ncp), names = colnames(data)[21:(ncp + 20)]) %>%
```

```

select(names, everything())
# sort by p.value
outcome_lm = outcome_lm[order(outcome_lm$p.value), ]

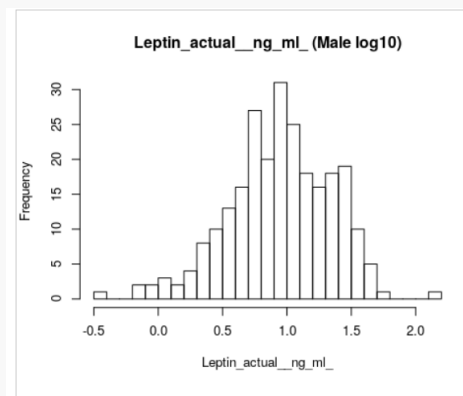
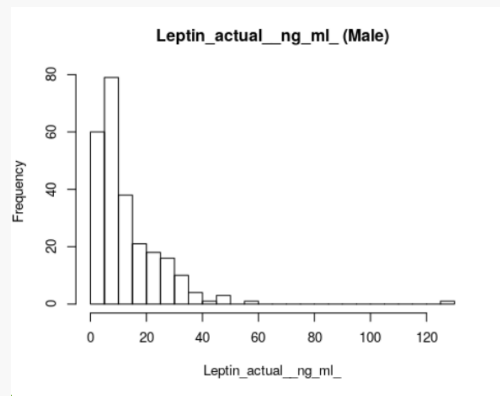
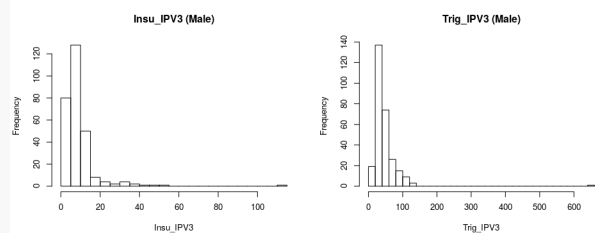
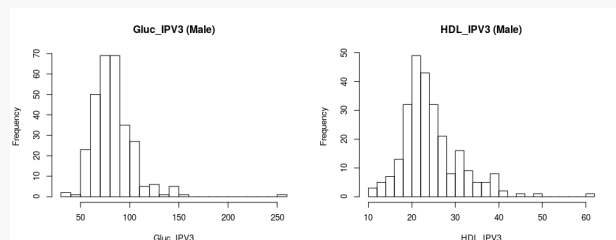
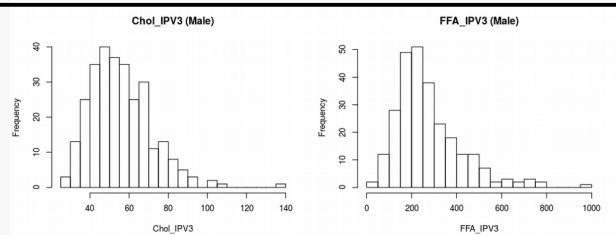
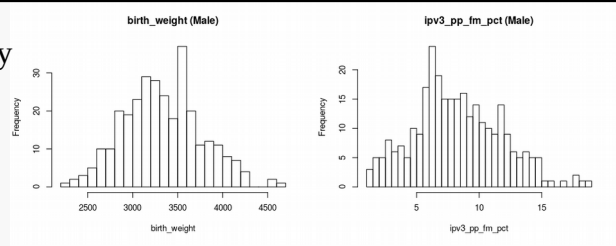
## sample size
size = length(outcome) - sum(is.na(outcome))

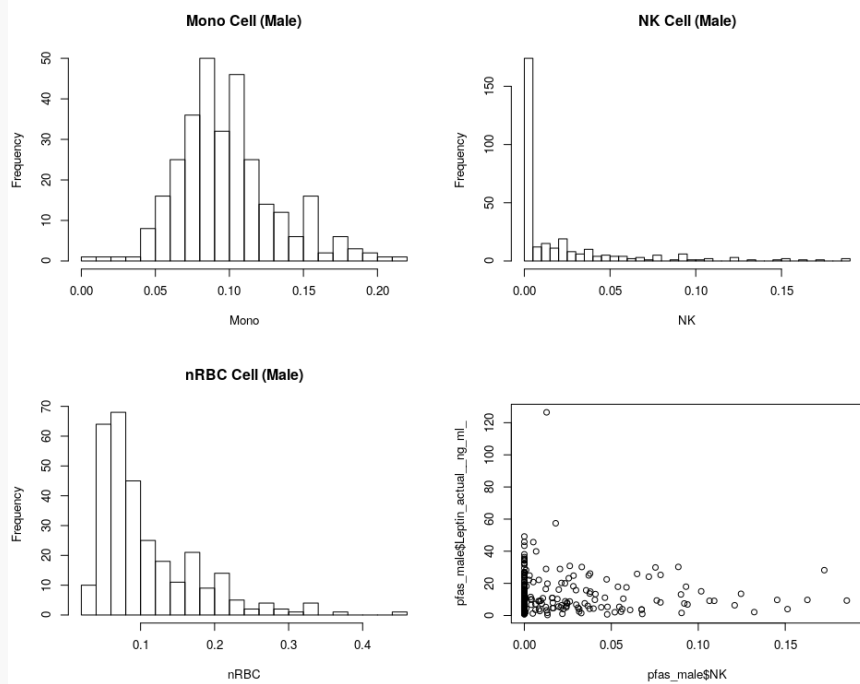
## summary table
kable(head(outcome_lm, Topn), caption = paste("Top10 CpGs from ",
ncpg, " for ", name, " of ", Gender, " by p.value", " (Sample Size
= ",
size, ") ", sep = "", collapse = ""))
}

```

## Test normality of Outcomes

May  
not





## Results Tables

- 1 - 9 Male TOP300 CpGs
- 10 - 18 Male TOP300 CpGs log10
- 19 - 27 Female TOP300 CpGs log10
- 28 - 36 Male (TOP300 and FDR 0.05) 120 CpGs log 10
- 37 - 45 Female (TOP300 and FDR 0.05) 120 CpGs log 10