583 top CpG

Guannan Shen

November 1, 2018

Contents

```
## set up workspace
library(knitr)
library(tidyverse)
options(stringsAsFactors = F)
options(dplyr.width = Inf)
getwd()
```

```
## [1] "/home/guanshim/Documents/gitlab/ECCHO_github/Reports"
```

```
"%nin%" <- Negate("%in%")
```

```
## top CpG for male and female male 103 and female 38
cpg_male <- read.delim("~/Documents/gitlab/ECCHO_github/DataRaw/583_top_CpG_from_DMR_adjusted_M005.txt"
dim(cpg_male)
```

```
## [1] 103  15
```

```
cpg_female <- read.delim("~/Documents/gitlab/ECCHO_github/DataRaw/583_top_CpG_from_DMR_adjusted_F005.tx
dim(cpg_female)
```

```
## [1] 38 15
```

```
## the list of top cpg
male103 <- cpg_male$CpG
female38 <- cpg_female$CpG

### import M values for top CpGs of female and male
m103 <- read_csv("~/Documents/gitlab/ECCHO_github/DataRaw/HS_450K_CB_Mval_normbatch_StarlingSubset_MALE_
dim(m103)
```

```
## [1] 600 104
```

```
f38 <- read_csv("~/Documents/gitlab/ECCHO_github/DataRaw/HS_450K_CB_Mval_normbatch_StarlingSubset_FEMALE
dim(f38)
```

```
## [1] 600  39
```

```
## double check the CpG list
sum(male103 %in% colnames(m103[, -1])) == length(male103)
```

```
## [1] TRUE
```

```
sum(female38 %in% colnames(f38[, -1])) == length(female38)
```

```
## [1] TRUE
```

```
## get the 583 subjects list
pid583 <- read.csv("~/Documents/gitlab/ECCHO_github/DataRaw/pid583.csv")
pid583 <- as.vector(unlist(pid583))
## clinical data and cord blood cell types
```

```
pfas <- read_csv("~/Documents/gitlab/ECCHO_github/DataRaw/pfas_methyl_di.csv")
celltype <- read_csv("~/Documents/gitlab/ECCHO_github/DataProcessed/healthy_start_cordblood_cellcounts_

## inner joint, in the clinical data, 2 levels Race.
pfas_cell <- merge(pfas, celltype, by = "pid") %>% rename(Race = race_4) %>%
    select(pid, infant_sex, maternal_age, Race, everything()) %>%
    mutate(Race = fct_collapse(as.factor(Race), `Non-Hispanic white` = "2",
        `All others` = c("1", "3", "4")))
## clinical, cell types and 103 CpGs for male
pfas_malecpg <- merge(pfas_cell, m103, by = "pid") %>% filter(pid %in%
    pid583) %>% filter(infant_sex == 2)
dim(pfas_malecpg)
```

```
## [1] 305 123
```

```
## clinical, cell types and 38 CpGs for female
pfas_femalecpg <- merge(pfas_cell, f38, by = "pid") %>% filter(pid %in%
    pid583) %>% filter(infant_sex == 1)
dim(pfas_femalecpg)
```

```
## [1] 278  58
```

```
## variables besides CpGs not cpg part
out_cov <- colnames(pfas_cell)
## the outcomes vector
Outcomes <- colnames(pfas_cell)[5:13]

## save dataset
write.csv(pfas_malecpg, "~/Documents/gitlab/ECCHO_github/DataProcessed/11_05_pfas_malecpg.csv")
write.csv(pfas_femalecpg, "~/Documents/gitlab/ECCHO_github/DataProcessed/11_05_pfas_femalecpg.csv")
```

```
################################### No maternal_age ########### equal length of outcomes and
################################### covariates
cpg_FunRegSim <- function(outcome, CpGdata, Xs, Outcome_name,
    Topn, Gender) {
    ## number of CpG to test, also the number of multiple test
    n_cpg = ncol(CpGdata)
    ## outcome lm
    outcome_lm = lapply(1:n_cpg, function(i) {
        lm = lm(outcome ~ CpGdata[, i] + Race + Bcell + CD4T +
            CD8T + Gran + Mono + NK + nRBC, data = Xs)
        coef = summary(lm)$coefficients[2, ]
        return(coef)
    })
    outcome_lm = data.frame(matrix(unlist(outcome_lm), ncol = 4,
        byrow = TRUE, dimnames = list(c(colnames(CpGdata)), c("Estimate",
            "Std.Error", "t.statistic", "p.value"))))

    # adjusted p-value
    outcome_lm = outcome_lm %>% mutate(FDR = p.adjust(p.value,
        "BH", n_cpg), names = colnames(CpGdata)) %>% mutate(Estimate = round(Estimate,
        4), Std.Error = round(Std.Error, 4), t.statistic = round(t.statistic,
        4)) %>% 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("Top CpGs from ",
        n_cpg, " for Outcome: ", Outcome_name, " of ", Gender,
        " by p.value", " (Sample Size = ", size, ") ", sep = "",
        collapse = ""))
}


cpg_RegSimRaw <- function(outcome, CpGdata, Xs, Outcome_name,
    Topn, Gender) {
    ## number of CpG to test, also the number of multiple test
    n_cpg = ncol(CpGdata)
    ## outcome lm
    outcome_lm = lapply(1:n_cpg, function(i) {
        lm = lm(outcome ~ CpGdata[, i] + Race + Bcell + CD4T +
            CD8T + Gran + Mono + NK + nRBC, data = Xs)
        coef = summary(lm)$coefficients[2, ]
        return(coef)
    })
    outcome_lm = data.frame(matrix(unlist(outcome_lm), ncol = 4,
        byrow = TRUE, dimnames = list(c(colnames(CpGdata)), c("Estimate",
            "Std.Error", "t.statistic", "p.value"))))

    # adjusted p-value
    outcome_lm = outcome_lm %>% mutate(FDR = p.adjust(p.value,
        "BH", n_cpg), names = colnames(CpGdata)) %>% mutate(Estimate = round(Estimate,
        4), Std.Error = round(Std.Error, 4), t.statistic = round(t.statistic,
        4)) %>% 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))

    ## raw_data
    head(outcome_lm, max(c(sum(outcome_lm$p.value <= 0.05))))
}
```

```
## equal length of outcomes and covariates
cpg_FunReg <- function(outcome, CpGdata, Xs, Outcome_name, Topn,
    Gender) {
    ## number of CpG to test, also the number of multiple test
    n_cpg = ncol(CpGdata)
    ## outcome lm
    outcome_lm = lapply(1:n_cpg, function(i) {
        lm = lm(outcome ~ CpGdata[, i] + maternal_age + Race +
            Bcell + CD4T + CD8T + Gran + Mono + NK + nRBC, data = Xs)
        coef = summary(lm)$coefficients[2, ]
        return(coef)
    })
```

```
    outcome_lm = data.frame(matrix(unlist(outcome_lm), ncol = 4,
        byrow = TRUE, dimnames = list(c(colnames(CpGdata)), c("Estimate",
            "Std.Error", "t.statistic", "p.value"))))

    # adjusted p-value
    outcome_lm = outcome_lm %>% mutate(FDR = p.adjust(p.value,
        "BH", n_cpg), names = colnames(CpGdata)) %>% mutate(Estimate = round(Estimate,
        4), Std.Error = round(Std.Error, 4), t.statistic = round(t.statistic,
        4)) %>% 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("Top CpGs from ",
        n_cpg, " for Outcome: ", Outcome_name, " of ", Gender,
        " by p.value", " (Sample Size = ", size, ") ", sep = "",
        collapse = ""))

}
```

```
## # the regression summary table for log10 outcomes
lapply(Outcomes, function(x) {
    cpg_FunRegSim(log10(pfas_malecpg[, x]), pfas_malecpg[, 21:123],
        pfas_malecpg[, out_cov], x, 10, "Male log10")
})
```

[[1]]

Table 1: Top CpGs from 103 for Outcome: birth_weight of Male log10 by p.value (Sample Size = 305)

	names	Estimate	Std.Error	t.statistic	p.value	FDR
9	cg25195288	0.0864	0.0231	3.7332	0.0002268	0.0233628
54	cg25137968	0.0460	0.0157	2.9334	0.0036159	0.1862181
86	cg15302379	0.0511	0.0185	2.7636	0.0060765	0.2086279
43	cg04029532	0.0545	0.0231	2.3555	0.0191545	0.4679804
60	cg25138412	-0.0195	0.0086	-2.2525	0.0250254	0.4679804
85	cg00798281	-0.0273	0.0123	-2.2188	0.0272610	0.4679804
22	cg15355952	0.0297	0.0157	1.8910	0.0596040	0.7418986
68	cg06407657	0.0247	0.0134	1.8420	0.0664827	0.7418986
31	cg16312900	0.0273	0.0150	1.8242	0.0691389	0.7418986
79	cg22946159	0.0444	0.0251	1.7678	0.0781254	0.7418986

[[2]]

Table 2: Top CpGs from 103 for Outcome: ipv3_pp_fm_pct of Male log10 by p.value (Sample Size = 292)

	names	Estimate	Std.Error	t.statistic	p.value	FDR
56	cg15066197	-0.1860	0.0693	-2.6835	0.0077168	0.7123866

	names	Estimate	Std.Error	t.statistic	p.value	FDR
98	cg20078119	0.1862	0.0752	2.4771	0.0138328	0.7123866
37	cg08743751	0.1291	0.0566	2.2819	0.0232409	0.7695147
95	cg04804814	0.1808	0.0860	2.1022	0.0364197	0.7695147
54	cg25137968	0.1272	0.0635	2.0028	0.0461561	0.7695147
73	cg01969701	0.0839	0.0425	1.9728	0.0494981	0.7695147
69	cg24440898	0.1429	0.0733	1.9489	0.0522971	0.7695147
80	cg20732198	0.0926	0.0513	1.8073	0.0717803	0.8546206
29	cg19529074	-0.1253	0.0716	-1.7497	0.0812548	0.8546206
6	cg22199628	-0.2070	0.1288	-1.6076	0.1090447	0.8546206

[[3]]

Table 3: Top CpGs from 103 for Outcome: Chol_IPV3 of Male log10 by p.value (Sample Size = 287)

	names	Estimate	Std.Error	t.statistic	p.value	FDR
70	cg12857407	0.0900	0.0323	2.7911	0.0056184	0.5786901
84	cg08162803	0.0980	0.0394	2.4888	0.0134053	0.6903724
72	cg04061372	0.0226	0.0107	2.1078	0.0359507	0.8529190
93	cg17132124	0.0538	0.0270	1.9945	0.0470804	0.8529190
34	cg23467747	0.0671	0.0373	1.8002	0.0729246	0.8529190
10	cg13699963	0.0844	0.0476	1.7719	0.0775137	0.8529190
22	cg15355952	-0.0552	0.0334	-1.6517	0.0997311	0.8529190
79	cg22946159	-0.0873	0.0537	-1.6258	0.1051334	0.8529190
65	cg09825146	-0.0345	0.0213	-1.6243	0.1054518	0.8529190
91	cg23964682	-0.0353	0.0222	-1.5886	0.1132986	0.8529190

[[4]]

Table 4: Top CpGs from 103 for Outcome: FFA_IPV3 of Male log10 by p.value (Sample Size = 265)

	names	Estimate	Std.Error	t.statistic	p.value	FDR
80	cg20732198	-0.1141	0.0528	-2.1631	0.0314664	0.7869043
82	cg16375541	0.2510	0.1169	2.1467	0.0327577	0.7869043
59	cg00910503	-0.2589	0.1239	-2.0892	0.0376804	0.7869043
29	cg19529074	-0.1419	0.0708	-2.0045	0.0460760	0.7869043
47	cg15404665	0.0695	0.0353	1.9669	0.0502757	0.7869043
40	cg13771313	0.0655	0.0374	1.7510	0.0811413	0.7869043
86	cg15302379	0.1321	0.0776	1.7021	0.0899549	0.7869043
16	cg23208443	-0.0189	0.0112	-1.6912	0.0920267	0.7869043
78	cg05227616	-0.1114	0.0664	-1.6772	0.0947374	0.7869043
63	cg03015672	0.0979	0.0585	1.6730	0.0955528	0.7869043

[[5]]

Table 5: Top CpGs from 103 for Outcome: Gluc_IPV3 of Male log10 by p.value (Sample Size = 295)

	names	Estimate	Std.Error	t.statistic	p.value	FDR
67	cg11196848	-0.0683	0.0265	-2.5801	0.0103794	0.6030879
13	cg01816336	-0.0994	0.0392	-2.5371	0.0117104	0.6030879
18	cg01060409	0.1271	0.0569	2.2347	0.0262098	0.8171162
80	cg20732198	0.0483	0.0233	2.0722	0.0391488	0.8171162
36	cg23478547	0.0353	0.0175	2.0170	0.0446368	0.8171162
22	cg15355952	0.0573	0.0288	1.9895	0.0475990	0.8171162
51	cg26781129	0.0446	0.0242	1.8398	0.0668388	0.8342697
29	cg19529074	-0.0555	0.0307	-1.8108	0.0712312	0.8342697
60	cg25138412	-0.0286	0.0163	-1.7564	0.0800873	0.8342697
43	cg04029532	0.0743	0.0424	1.7505	0.0811040	0.8342697

[[6]]

Table 6: Top CpGs from 103 for Outcome: HDL_IPV3 of Male log10 by p.value (Sample Size = 261)

	names	Estimate	Std.Error	t.statistic	p.value	FDR
22	cg15355952	-0.1030	0.0330	-3.1196	0.0020221	0.2082809
72	cg04061372	0.0293	0.0107	2.7521	0.0063533	0.3271960
84	cg08162803	0.0987	0.0383	2.5751	0.0105933	0.3350125
79	cg22946159	-0.1356	0.0543	-2.4960	0.0132042	0.3350125
2	cg19549232	0.0872	0.0362	2.4093	0.0167049	0.3350125
85	cg00798281	-0.0608	0.0259	-2.3506	0.0195153	0.3350125
92	cg03989507	0.0706	0.0309	2.2830	0.0232659	0.3359445
62	cg06230206	-0.0644	0.0288	-2.2381	0.0260928	0.3359445
33	cg13095300	0.0434	0.0215	2.0231	0.0441208	0.4239298
42	cg25206725	0.0627	0.0320	1.9559	0.0515888	0.4239298

[[7]]

Table 7: Top CpGs from 103 for Outcome: Insu_IPV3 of Male log10 by p.value (Sample Size = 282)

	names	Estimate	Std.Error	t.statistic	p.value	FDR
60	cg25138412	-0.1088	0.0438	-2.4840	0.0135948	0.9725865
96	cg09630142	0.1215	0.0576	2.1106	0.0357229	0.9725865
81	cg20276750	-0.1990	0.1000	-1.9892	0.0476860	0.9725865
67	cg11196848	-0.1404	0.0723	-1.9410	0.0532904	0.9725865
91	cg23964682	-0.0924	0.0515	-1.7945	0.0738488	0.9725865
46	cg01541565	0.1250	0.0720	1.7358	0.0837332	0.9725865
41	cg17578309	-0.1830	0.1191	-1.5372	0.1254022	0.9725865
30	cg05906144	0.0862	0.0566	1.5228	0.1289713	0.9725865
62	cg06230206	0.1007	0.0666	1.5135	0.1313116	0.9725865
102	cg01607625	-0.1572	0.1046	-1.5032	0.1339569	0.9725865

[[8]]

Table 8: Top CpGs from 103 for Outcome: Trig_IPV3 of Male log10 by p.value (Sample Size = 284)

	names	Estimate	Std.Error	t.statistic	p.value	FDR
96	cg09630142	-0.1331	0.0443	-3.0042	0.0029087	0.2995944
39	cg14349977	-0.0837	0.0326	-2.5662	0.0108126	0.3851148
36	cg23478547	0.0868	0.0363	2.3884	0.0175993	0.3851148
43	cg04029532	-0.2016	0.0887	-2.2725	0.0238303	0.3851148
32	cg18373158	0.1147	0.0508	2.2568	0.0248096	0.3851148
59	cg00910503	-0.2626	0.1169	-2.2468	0.0254485	0.3851148
82	cg16375541	0.2390	0.1069	2.2358	0.0261728	0.3851148
9	cg25195288	-0.1831	0.0902	-2.0291	0.0434134	0.4770835
33	cg13095300	-0.0816	0.0402	-2.0268	0.0436489	0.4770835
58	cg19711268	-0.0852	0.0426	-2.0016	0.0463188	0.4770835

[[9]]

Table 9: Top CpGs from 103 for Outcome: Leptin_actual_*ng_ml* of Male log10 by p.value (Sample Size = 252)

	names	Estimate	Std.Error	t.statistic	p.value	FDR
74	cg24280832	0.4215	0.1139	3.7018	0.0002651	0.0273025
86	cg15302379	0.4836	0.1436	3.3683	0.0008797	0.0453057
22	cg15355952	0.3193	0.1246	2.5635	0.0109691	0.3766047
4	cg07818713	-0.1664	0.0724	-2.2976	0.0224368	0.5777483
81	cg20276750	-0.2979	0.1544	-1.9292	0.0548807	0.7675843
14	cg07694864	0.1764	0.0963	1.8320	0.0681765	0.7675843
91	cg23964682	-0.1481	0.0810	-1.8288	0.0686586	0.7675843
90	cg11618577	-0.1608	0.0896	-1.7948	0.0739329	0.7675843
79	cg22946159	-0.3351	0.1923	-1.7427	0.0826599	0.7675843
102	cg01607625	0.2751	0.1606	1.7129	0.0880208	0.7675843

```
lapply(Outcomes, function(x) {
    cpg_RegSimRaw(log10(pfas_malecpg[, x]), pfas_malecpg[, 21:123],
        pfas_malecpg[, out_cov], x, 10, "Male log10")
})
```

```
## [[1]]
##         names Estimate Std.Error t.statistic      p.value        FDR
## 9  cg25195288   0.0864    0.0231      3.7332 0.0002268232 0.02336279
## 54 cg25137968   0.0460    0.0157      2.9334 0.0036158860 0.18621813
## 86 cg15302379   0.0511    0.0185      2.7636 0.0060765419 0.20862794
## 43 cg04029532   0.0545    0.0231      2.3555 0.0191545488 0.46798044
## 60 cg25138412  -0.0195    0.0086     -2.2525 0.0250254154 0.46798044
## 85 cg00798281  -0.0273    0.0123     -2.2188 0.0272609966 0.46798044
##
## [[2]]
##         names Estimate Std.Error t.statistic     p.value       FDR
## 56 cg15066197  -0.1860    0.0693     -2.6835 0.007716796 0.7123866
## 98 cg20078119   0.1862    0.0752      2.4771 0.013832750 0.7123866
## 37 cg08743751   0.1291    0.0566      2.2819 0.023240947 0.7695147
## 95 cg04804814   0.1808    0.0860      2.1022 0.036419712 0.7695147
```

```
## 54 cg25137968   0.1272    0.0635      2.0028 0.046156093 0.7695147
## 73 cg01969701   0.0839    0.0425      1.9728 0.049498064 0.7695147
##
## [[3]]
##         names Estimate Std.Error t.statistic     p.value       FDR
## 70 cg12857407   0.0900    0.0323      2.7911 0.005618351 0.5786901
## 84 cg08162803   0.0980    0.0394      2.4888 0.013405290 0.6903724
## 72 cg04061372   0.0226    0.0107      2.1078 0.035950684 0.8529190
## 93 cg17132124   0.0538    0.0270      1.9945 0.047080449 0.8529190
##
## [[4]]
##         names Estimate Std.Error t.statistic    p.value       FDR
## 80 cg20732198  -0.1141    0.0528     -2.1631 0.03146636 0.7869043
## 82 cg16375541   0.2510    0.1169      2.1467 0.03275769 0.7869043
## 59 cg00910503  -0.2589    0.1239     -2.0892 0.03768036 0.7869043
## 29 cg19529074  -0.1419    0.0708     -2.0045 0.04607601 0.7869043
##
## [[5]]
##         names Estimate Std.Error t.statistic    p.value       FDR
## 67 cg11196848  -0.0683    0.0265     -2.5801 0.01037936 0.6030879
## 13 cg01816336  -0.0994    0.0392     -2.5371 0.01171045 0.6030879
## 18 cg01060409   0.1271    0.0569      2.2347 0.02620976 0.8171162
## 80 cg20732198   0.0483    0.0233      2.0722 0.03914883 0.8171162
## 36 cg23478547   0.0353    0.0175      2.0170 0.04463684 0.8171162
## 22 cg15355952   0.0573    0.0288      1.9895 0.04759900 0.8171162
##
## [[6]]
##         names Estimate Std.Error t.statistic     p.value       FDR
## 22 cg15355952  -0.1030    0.0330     -3.1196 0.002022145 0.2082809
## 72 cg04061372   0.0293    0.0107      2.7521 0.006353319 0.3271960
## 84 cg08162803   0.0987    0.0383      2.5751 0.010593298 0.3350125
## 79 cg22946159  -0.1356    0.0543     -2.4960 0.013204212 0.3350125
## 2  cg19549232   0.0872    0.0362      2.4093 0.016704862 0.3350125
## 85 cg00798281  -0.0608    0.0259     -2.3506 0.019515293 0.3350125
## 92 cg03989507   0.0706    0.0309      2.2830 0.023265933 0.3359445
## 62 cg06230206  -0.0644    0.0288     -2.2381 0.026092773 0.3359445
## 33 cg13095300   0.0434    0.0215      2.0231 0.044120830 0.4239298
##
## [[7]]
##         names Estimate Std.Error t.statistic    p.value       FDR
## 60 cg25138412  -0.1088    0.0438     -2.4840 0.01359482 0.9725865
## 96 cg09630142   0.1215    0.0576      2.1106 0.03572290 0.9725865
## 81 cg20276750  -0.1990    0.1000     -1.9892 0.04768602 0.9725865
##
## [[8]]
##         names Estimate Std.Error t.statistic     p.value       FDR
## 96 cg09630142  -0.1331    0.0443     -3.0042 0.002908683 0.2995944
## 39 cg14349977  -0.0837    0.0326     -2.5662 0.010812622 0.3851148
## 36 cg23478547   0.0868    0.0363      2.3884 0.017599320 0.3851148
## 43 cg04029532  -0.2016    0.0887     -2.2725 0.023830313 0.3851148
## 32 cg18373158   0.1147    0.0508      2.2568 0.024809578 0.3851148
## 59 cg00910503  -0.2626    0.1169     -2.2468 0.025448491 0.3851148
## 82 cg16375541   0.2390    0.1069      2.2358 0.026172848 0.3851148
## 9  cg25195288  -0.1831    0.0902     -2.0291 0.043413396 0.4770835
```

```
## 33 cg13095300  -0.0816    0.0402     -2.0268 0.043648950 0.4770835
## 58 cg19711268  -0.0852    0.0426     -2.0016 0.046318791 0.4770835
##
## [[9]]
##         names Estimate Std.Error t.statistic      p.value        FDR
## 74 cg24280832   0.4215    0.1139      3.7018 0.0002650726 0.02730247
## 86 cg15302379   0.4836    0.1436      3.3683 0.0008797220 0.04530568
## 22 cg15355952   0.3193    0.1246      2.5635 0.0109690694 0.37660471
## 4  cg07818713  -0.1664    0.0724     -2.2976 0.0224368266 0.57774828
```

```
## # the regression summary table for log10 outcomes
lapply(Outcomes, function(x) {
    cpg_FunRegSim(log10(pfas_femalecpg[, x]), pfas_femalecpg[,
        21:58], pfas_femalecpg[, out_cov], x, 10, "Female log10")
})
```

[[1]]

Table 10: Top CpGs from 38 for Outcome: birth_weight of Female log10 by p.value (Sample Size = 278)

	names	Estimate	Std.Error	t.statistic	p.value	FDR
23	cg19425295	0.0794	0.0201	3.9407	0.0001037	0.0039417
15	cg18802138	0.0877	0.0418	2.0961	0.0370138	0.4370985
36	cg05194362	0.0473	0.0232	2.0373	0.0426009	0.4370985
6	cg06561666	-0.0275	0.0137	-2.0046	0.0460104	0.4370985
35	cg06755596	0.0095	0.0059	1.6224	0.1058848	0.6416820
1	cg19729649	0.0274	0.0170	1.6082	0.1089782	0.6416820
32	cg07670516	-0.0159	0.0101	-1.5674	0.1182046	0.6416820
20	cg04999036	-0.0199	0.0142	-1.4016	0.1621974	0.7107154
26	cg05767411	-0.0188	0.0136	-1.3813	0.1683273	0.7107154
24	cg11108474	0.0157	0.0119	1.3202	0.1878967	0.7140074

[[2]]

Table 11: Top CpGs from 38 for Outcome: ipv3_pp_fm_pct of Female log10 by p.value (Sample Size = 271)

	names	Estimate	Std.Error	t.statistic	p.value	FDR
32	cg07670516	-0.0970	0.0382	-2.5384	0.0117187	0.4453091
29	cg00399059	-0.0748	0.0375	-1.9954	0.0470413	0.7582772
23	cg19425295	0.1427	0.0791	1.8039	0.0723974	0.7582772
37	cg08897388	0.1748	0.1065	1.6415	0.1018880	0.7582772
2	cg14220537	-0.1085	0.0675	-1.6071	0.1092307	0.7582772
11	cg18583565	0.0765	0.0490	1.5610	0.1197280	0.7582772
35	cg06755596	0.0294	0.0227	1.2931	0.1971090	0.7811947
10	cg13732083	0.0214	0.0173	1.2363	0.2174465	0.7811947
16	cg05076221	-0.0488	0.0400	-1.2197	0.2236980	0.7811947
21	cg05688348	0.0921	0.0828	1.1115	0.2673683	0.7811947

[[3]]

Table 12: Top CpGs from 38 for Outcome: Chol_IPV3 of Female log10 by p.value (Sample Size = 257)

	names	Estimate	Std.Error	t.statistic	p.value	FDR
16	cg05076221	0.0624	0.0239	2.6134	0.0095138	0.3615227
18	cg15927196	-0.0899	0.0429	-2.0943	0.0372556	0.5777871
38	cg03967627	-0.0524	0.0267	-1.9612	0.0509787	0.5777871
37	cg08897388	-0.1203	0.0639	-1.8834	0.0608197	0.5777871
4	cg15889061	0.0525	0.0297	1.7671	0.0784391	0.5961370
26	cg05767411	0.0515	0.0310	1.6623	0.0977269	0.6189369
33	cg01319701	0.0924	0.0606	1.5257	0.1283554	0.6533471
23	cg19425295	-0.0688	0.0471	-1.4618	0.1450567	0.6533471
13	cg15506703	0.0278	0.0195	1.4241	0.1556820	0.6533471
17	cg27370696	0.0407	0.0297	1.3700	0.1719334	0.6533471

[[4]]

Table 13: Top CpGs from 38 for Outcome: FFA_IPV3 of Female log10 by p.value (Sample Size = 237)

	names	Estimate	Std.Error	t.statistic	p.value	FDR
22	cg21503202	0.1545	0.0836	1.8471	0.0660283	0.9144910
28	cg10742801	-0.1452	0.0824	-1.7611	0.0795608	0.9144910
2	cg14220537	0.1252	0.0736	1.7024	0.0900510	0.9144910
36	cg05194362	0.1352	0.0978	1.3818	0.1683927	0.9144910
15	cg18802138	0.2394	0.1780	1.3451	0.1799366	0.9144910
5	cg27238852	-0.0700	0.0521	-1.3442	0.1802132	0.9144910
7	cg19838087	0.0351	0.0264	1.3271	0.1858232	0.9144910
17	cg27370696	0.0738	0.0565	1.3070	0.1925244	0.9144910
24	cg11108474	-0.0588	0.0505	-1.1653	0.2451034	0.9924946
11	cg18583565	0.0549	0.0523	1.0509	0.2944058	0.9924946

[[5]]

Table 14: Top CpGs from 38 for Outcome: Gluc_IPV3 of Female log10 by p.value (Sample Size = 263)

	names	Estimate	Std.Error	t.statistic	p.value	FDR
31	cg22968966	0.0297	0.0150	1.9746	0.0494006	0.9798371
6	cg06561666	-0.0435	0.0228	-1.9064	0.0577266	0.9798371
14	cg04859241	-0.0691	0.0485	-1.4251	0.1553675	0.9798371
2	cg14220537	0.0370	0.0291	1.2733	0.2040771	0.9798371
38	cg03967627	-0.0246	0.0195	-1.2613	0.2083526	0.9798371
36	cg05194362	0.0444	0.0393	1.1305	0.2593316	0.9798371
12	cg06988897	-0.0126	0.0115	-1.0947	0.2746983	0.9798371
3	cg19602187	-0.0404	0.0375	-1.0761	0.2828946	0.9798371
32	cg07670516	-0.0182	0.0171	-1.0652	0.2878051	0.9798371
26	cg05767411	-0.0241	0.0226	-1.0633	0.2886517	0.9798371

[[6]]

Table 15: Top CpGs from 38 for Outcome: HDL_IPV3 of Female log10 by p.value (Sample Size = 242)

	names	Estimate	Std.Error	t.statistic	p.value	FDR
38	cg03967627	-0.0468	0.0258	-1.8146	0.0708796	0.8834627
16	cg05076221	0.0348	0.0236	1.4762	0.1412549	0.8834627
29	cg00399059	-0.0328	0.0228	-1.4415	0.1507989	0.8834627
23	cg19425295	-0.0646	0.0457	-1.4121	0.1592524	0.8834627
24	cg11108474	-0.0315	0.0261	-1.2080	0.2282624	0.8834627
28	cg10742801	-0.0476	0.0431	-1.1044	0.2705675	0.8834627
11	cg18583565	-0.0315	0.0291	-1.0848	0.2791573	0.8834627
18	cg15927196	-0.0457	0.0425	-1.0758	0.2831436	0.8834627
15	cg18802138	0.0935	0.0944	0.9905	0.3229734	0.8834627
33	cg01319701	0.0555	0.0585	0.9484	0.3439169	0.8834627

[[7]]

Table 16: Top CpGs from 38 for Outcome: Insu_IPV3 of Female log10 by p.value (Sample Size = 255)

	names	Estimate	Std.Error	t.statistic	p.value	FDR
30	cg23188684	0.1602	0.0525	3.0486	0.0025515	0.0969585
35	cg06755596	0.0641	0.0270	2.3720	0.0184672	0.2351718
13	cg15506703	0.0887	0.0374	2.3699	0.0185662	0.2351718
37	cg08897388	0.2585	0.1251	2.0657	0.0399126	0.3791696
8	cg01434025	0.2098	0.1117	1.8783	0.0615269	0.4676042
10	cg13732083	-0.0361	0.0202	-1.7901	0.0746752	0.4729427
34	cg13912721	-0.1826	0.1072	-1.7028	0.0898667	0.4878477
29	cg00399059	-0.0676	0.0434	-1.5583	0.1204510	0.5162588
14	cg04859241	-0.2032	0.1310	-1.5507	0.1222718	0.5162588
2	cg14220537	-0.0856	0.0789	-1.0857	0.2786918	0.8524610

[[8]]

Table 17: Top CpGs from 38 for Outcome: Trig_IPV3 of Female log10 by p.value (Sample Size = 252)

	names	Estimate	Std.Error	t.statistic	p.value	FDR
20	cg04999036	0.1725	0.0521	3.3096	0.0010767	0.0409127
37	cg08897388	-0.2697	0.1026	-2.6293	0.0091039	0.1729746
13	cg15506703	0.0723	0.0315	2.2952	0.0225790	0.2860006
29	cg00399059	0.0813	0.0374	2.1752	0.0305822	0.2905312
7	cg19838087	0.0429	0.0237	1.8054	0.0722565	0.4434715
2	cg14220537	0.1155	0.0650	1.7775	0.0767420	0.4434715
12	cg06988897	0.0448	0.0256	1.7482	0.0816921	0.4434715
19	cg08861456	0.0447	0.0287	1.5591	0.1202768	0.5645078
17	cg27370696	0.0722	0.0480	1.5047	0.1336992	0.5645078
1	cg19729649	-0.0923	0.0641	-1.4409	0.1509043	0.5734364

[[9]]

Table 18: Top CpGs from 38 for Outcome: Leptin_actual_*ng_ml* of Female log10 by p.value (Sample Size = 226)

	names	Estimate	Std.Error	t.statistic	p.value	FDR
1	cg19729649	0.3250	0.1196	2.7175	0.0071104	0.2304729
37	cg08897388	0.4868	0.1924	2.5297	0.0121302	0.2304729
9	cg25236028	-0.1517	0.0759	-1.9983	0.0469429	0.5263734
24	cg11108474	0.1468	0.0817	1.7980	0.0735800	0.5263734
14	cg04859241	0.3526	0.2008	1.7559	0.0805166	0.5263734
17	cg27370696	-0.1516	0.0941	-1.6123	0.1083613	0.5263734
5	cg27238852	-0.1385	0.0868	-1.5959	0.1119637	0.5263734
38	cg03967627	-0.1247	0.0818	-1.5237	0.1290450	0.5263734
36	cg05194362	0.2467	0.1635	1.5085	0.1328911	0.5263734
23	cg19425295	0.2143	0.1442	1.4868	0.1385193	0.5263734

```
lapply(Outcomes, function(x) {
    cpg_RegSimRaw(log10(pfas_femalecpg[, x]), pfas_femalecpg[,
        21:58], pfas_femalecpg[, out_cov], x, 10, "Female log10")
})
```

```
## [[1]]
##         names Estimate Std.Error t.statistic      p.value         FDR
## 23 cg19425295   0.0794    0.0201      3.9407 0.0001037282 0.003941671
## 15 cg18802138   0.0877    0.0418      2.0961 0.0370137582 0.437098474
## 36 cg05194362   0.0473    0.0232      2.0373 0.0426009345 0.437098474
## 6  cg06561666  -0.0275    0.0137     -2.0046 0.0460103657 0.437098474
##
## [[2]]
##         names Estimate Std.Error t.statistic    p.value       FDR
## 32 cg07670516  -0.0970    0.0382     -2.5384 0.01171866 0.4453091
## 29 cg00399059  -0.0748    0.0375     -1.9954 0.04704127 0.7582772
##
## [[3]]
##         names Estimate Std.Error t.statistic     p.value       FDR
## 16 cg05076221   0.0624    0.0239      2.6134 0.009513755 0.3615227
## 18 cg15927196  -0.0899    0.0429     -2.0943 0.037255560 0.5777871
##
## [[4]]
## [1] names       Estimate    Std.Error   t.statistic p.value     FDR
## <0 rows> (or 0-length row.names)
##
## [[5]]
##         names Estimate Std.Error t.statistic    p.value       FDR
## 31 cg22968966   0.0297     0.015      1.9746 0.04940061 0.9798371
##
## [[6]]
## [1] names       Estimate    Std.Error   t.statistic p.value     FDR
## <0 rows> (or 0-length row.names)
##
## [[7]]
##         names Estimate Std.Error t.statistic     p.value        FDR
## 30 cg23188684   0.1602    0.0525      3.0486 0.002551539 0.09695848
## 35 cg06755596   0.0641    0.0270      2.3720 0.018467222 0.23517179
```

```
## 13 cg15506703   0.0887    0.0374      2.3699 0.018566194 0.23517179
## 37 cg08897388   0.2585    0.1251      2.0657 0.039912585 0.37916955
##
## [[8]]
##         names Estimate Std.Error t.statistic     p.value        FDR
## 20 cg04999036   0.1725    0.0521      3.3096 0.001076651 0.04091274
## 37 cg08897388  -0.2697    0.1026     -2.6293 0.009103927 0.17297461
## 13 cg15506703   0.0723    0.0315      2.2952 0.022578992 0.28600056
## 29 cg00399059   0.0813    0.0374      2.1752 0.030582232 0.29053121
##
## [[9]]
##         names Estimate Std.Error t.statistic     p.value       FDR
## 1  cg19729649   0.3250    0.1196      2.7175 0.007110431 0.2304729
## 37 cg08897388   0.4868    0.1924      2.5297 0.012130152 0.2304729
## 9  cg25236028  -0.1517    0.0759     -1.9983 0.046942891 0.5263734
```