

The results below are generated from an R script.

```
# Load results for 15 traits for 266K SNPs
setwd("/home/lima/Sodbo/OmicABEL_paper/final_cut/tables/")
load("3_methods_15_traits_266K_SNPs.RData")

# List of objects
ls()

## [1] "EMMAX_15traits" "GWFGSL_15traits" "OMICABEL_15traits"

# Head of table containing EMMAX results
head(EMMAX_15traits)

##           Chisq      P-value    beta_SNP    se_SNP
## rs3737728_pcp_47 4.36218563 0.03674513 -0.12006189 0.05748481
## rs6687776_pcp_47 0.35498937 0.55130227  0.04012389 0.06734346
## rs12726255_pcp_47 0.01751924 0.89469933  0.00902336 0.06817272
## rs4970362_pcp_47 0.18559329 0.66660987  0.02290982 0.05317905
## rs9660710_pcp_47 0.01372266 0.90674612 -0.01143067 0.09757811
## rs4970420_pcp_47 6.19236104 0.01283029 -0.15935771 0.06403906

# Create vector of differences between GWFGSL and OmicABEL estimation of
# standart error of SNPs. Using formula (a-b) / b
gwfgslVSomica = abs(GWFGSL_15traits[,"se_SNP"]-OMICABEL_15traits[,"se_SNP"])/
  OMICABEL_15traits[,"se_SNP"]

# Brief summary of differences
summary(gwfgslVSomica)

##      Min.   1st Qu.   Median     Mean   3rd Qu.    Max.
## 0.0000000 0.0006751 0.0010600 0.0009720 0.0012400 0.1932000

# Create vector of differences between EMMAX and OmicABEL estimation of
# standart error of SNPs. Using formula (a-b) / b
emmaxVSomica = abs(EMMAX_15traits[,"se_SNP"]-OMICABEL_15traits[,"se_SNP"])/
  OMICABEL_15traits[,"se_SNP"]

# Brief summary of differences
summary(emmaxVSomica)

##      Min.   1st Qu.   Median     Mean   3rd Qu.    Max.
## 0.0000000 0.0005470 0.0007900 0.0007447 0.0009633 0.1819000

# Select SNPs for which differences between standart errors are more than 2%
TwoProcDiffSNPs = cbind(OMICABEL_15traits[which(gwfgslVSomica>0.02),2:4],
  diff = gwfgslVSomica[which(gwfgslVSomica>0.02)])

# Print head of TwoProcDiffSNPs
head(TwoProcDiffSNPs)

##           P_value    beta_SNP    se_SNP    diff
## rs509360_pcp_25 3.307857e-16  0.4219920 0.05170498 0.04236777
## rs102275_pcp_25 4.336341e-58 -0.8621287 0.05365793 0.18074562
## rs412334_pcp_25 2.314809e-12   0.5587760 0.07966498 0.03076672
## rs174546_pcp_25 1.555490e-61  -0.8925267 0.05392344 0.19322518
## rs174556_pcp_25 2.242274e-53  -0.8894322 0.05783179 0.16397953
## rs968567_pcp_25 2.368272e-14  -0.5830343 0.07642476 0.03674679
```

```
# Sort DwaProcRaz by differences between standart errors. You can see that
# differences between se_SNP estimated by OmicABEL and GWFGLS are growing up
# with increasing of Chi-square and beta_SNP.
TwoProcDiffSNPs[order((TwoProcDiffSNPs[, "diff"]),),]
```

##	P_value	beta_SNP	se_SNP	diff
## rs968567_pc_38	6.798219e-09	-0.4391189	0.07576456	0.02048700
## rs12889954_spm_1	5.872073e-09	0.4219019	0.07248725	0.02067360
## rs12221836_pcp_25	5.404342e-09	-0.5351502	0.09172611	0.02077955
## rs1466448_spm_8	4.557403e-09	-0.3712051	0.06331772	0.02099699
## rs968567_spm_18_AND_pcp_25	4.341861e-09	0.4517291	0.07694754	0.02105884
## rs1466448_spm_3_AND_spm_8	4.000476e-09	0.3743031	0.06361180	0.02116336
## rs174611_pc_38	2.288916e-09	-0.3294776	0.05513485	0.02187657
## rs12889954_spm_1_AND_spm_18	4.439378e-10	0.4538860	0.07276446	0.02397772
## rs509360_spm_18_AND_pcp_25	3.545796e-10	-0.3268071	0.05209896	0.02426633
## rs174611_spm_18_AND_pcp_25	3.312160e-10	0.3518054	0.05598951	0.02435395
## rs509360_pc_38	3.283283e-10	0.3222979	0.05128231	0.02436518
## rs174450_spm_18_AND_pcp_25	2.404998e-10	-0.3314577	0.05233852	0.02476523
## rs174449_spm_18_AND_pcp_25	9.392995e-11	0.3414013	0.05271457	0.02597537
## rs2851682_pc_38	2.327351e-12	-0.5561987	0.07930610	0.03075979
## rs412334_pcp_25	2.314809e-12	0.5587760	0.07966498	0.03076672
## rs2851682_spm_18_AND_pcp_25	3.700730e-13	0.5853170	0.08055472	0.03315285
## rs4902242_spm_1	3.078635e-13	0.5337301	0.07320488	0.03339286
## rs174449_pcp_25	2.819387e-13	-0.3824866	0.05237569	0.03350758
## rs174450_pcp_25	1.652445e-13	0.3834552	0.05199907	0.03420491
## rs4902242_spm_1	1.105094e-13	0.5443100	0.07328115	0.03473061
## rs968567_pcp_25	2.368272e-14	-0.5830343	0.07642476	0.03674679
## rs4902242_spm_1_AND_spm_18	1.270965e-14	0.5666248	0.07350449	0.03756316
## rs174570_pc_38	1.121222e-14	-0.5342898	0.06916647	0.03772775
## rs174570_spm_18_AND_pcp_25	3.634845e-15	0.5525909	0.07024238	0.03920794
## rs2072114_pc_38	1.378325e-15	-0.5314611	0.06653759	0.04048462
## rs174611_pcp_25	6.787587e-16	-0.4491140	0.05562280	0.04141869
## rs509360_pcp_25	3.307857e-16	0.4219920	0.05170498	0.04236777
## rs2072114_spm_18_AND_pcp_25	2.992127e-19	0.6060835	0.06757537	0.05168423
## rs2851682_pcp_25	1.438392e-21	-0.7629675	0.07998172	0.05885507
## rs174570_pcp_25	1.745761e-23	-0.6967911	0.06977304	0.06482993
## rs174556_pc_38	4.900701e-29	-0.6410568	0.05732068	0.08238955
## rs2072114_pcp_25	5.699808e-30	-0.7633443	0.06711902	0.08538118
## rs102275_pc_38	2.517949e-33	-0.6397595	0.05318707	0.09620988
## rs1535_pc_38	2.448529e-34	-0.6592433	0.05395056	0.09950359
## rs174556_spm_18_AND_pcp_25	7.930555e-35	0.7165727	0.05820723	0.10110172
## rs174546_pc_38	7.587002e-35	-0.6582190	0.05345157	0.10116473
## rs102275_spm_18_AND_pcp_25	3.525456e-37	0.6881317	0.05401181	0.10882063
## rs174546_spm_18_AND_pcp_25	2.064136e-38	0.7034822	0.05428126	0.11289388
## rs1535_spm_18_AND_pcp_25	1.035820e-38	0.7128190	0.05477865	0.11388608
## rs174556_pcp_25	2.242274e-53	-0.8894322	0.05783179	0.16397953
## rs102275_pcp_25	4.336341e-58	-0.8621287	0.05365793	0.18074562
## rs1535_pcp_25	1.153723e-60	-0.8944834	0.05443985	0.19005464
## rs174546_pcp_25	1.555490e-61	-0.8925267	0.05392344	0.19322518

```
# Let's check this hypothethis by testing linear regression of
# differences against Chisq and compare results with testing against
# se_SNP and beta_SNP
```

```

# Linear regression of differences against Chisq
# We can see that there is strong effect of Chisq on se_differences.
# As Chisq is squared (beta_SNP / se_SNP) we can conclude that se_differences
# are growing up with increasing of ratio between beta_SNP and se_SNP.
summary(lm(gwfglsVSomica ~ GWFGLS_15traits[, "Chisq"]))

##
## Call:
## lm(formula = gwfglsVSomica ~ GWFGLS_15traits[, "Chisq"])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.001110 -0.000352  0.000147  0.000385  0.134344
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      8.339e-04  3.335e-07  2500.4  <2e-16 ***
## GWFGLS_15traits[, "Chisq"] 1.379e-04  1.846e-07   747.2  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0005554 on 4003168 degrees of freedom
## Multiple R-squared:  0.1224, Adjusted R-squared:  0.1224
## F-statistic: 5.582e+05 on 1 and 4003168 DF,  p-value: < 2.2e-16

# This is actually what we can see on the plot of minus log10 of p.value
# estimated by OmicABEL against estimated by GWFGLS
# Let's plot it

# Temporary function to plot y=x line
#my_line <- function(x,y,...){
#  points(x,y,...)
#  abline(a = 0,b = 1,...)
#}

# Plotting minus log10 of pvalue OmicABEL vs GWFGLS
#pairs(cbind(OmicABEL = (-log10(OMICABEL_15traits$P_value)),
# GWFGLS = (-log10(GWFGLS_15traits[, "P-value"]))), lower.panel = my_line,
# upper.panel=my_line, cex.labels=4.0, cex=1.5, cex.main=2.0, cex.axis=2.5,
# )

# What about effect of
# Linear regression of differences against beta_SNP
summary(lm(gwfglsVSomica ~ GWFGLS_15traits[, "beta_SNP"]))

##
## Call:
## lm(formula = gwfglsVSomica ~ GWFGLS_15traits[, "beta_SNP"])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.000978 -0.000297  0.000088  0.000268  0.192217
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      9.720e-04  2.963e-07 3280.351  <2e-16 ***

```

```

## GWFGLS_15traits[, "beta_SNP"] -4.003e-05  4.435e-06  -9.027  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0005929 on 4003168 degrees of freedom
## Multiple R-squared:  2.036e-05, Adjusted R-squared:  2.011e-05
## F-statistic: 81.49 on 1 and 4003168 DF,  p-value: < 2.2e-16

# Linear regression of differences against se_SNP
summary(lm(gwfglsVSomica ~ GWFGLS_15traits[, "se_SNP"]))

##
## Call:
## lm(formula = gwfglsVSomica ~ GWFGLS_15traits[, "se_SNP"])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.000973 -0.000297  0.000088  0.000268  0.192252
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      9.764e-04  1.301e-06  750.681 < 2e-16 ***
## GWFGLS_15traits[, "se_SNP"] -6.674e-05  1.945e-05  -3.431 0.000601 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0005929 on 4003168 degrees of freedom
## Multiple R-squared:  2.941e-06, Adjusted R-squared:  2.691e-06
## F-statistic: 11.77 on 1 and 4003168 DF,  p-value: 0.000601

```

The R session information (including the OS info, R version and all packages used):

```

sessionInfo()

## R version 3.2.2 (2015-08-14)
## Platform: x86_64-pc-linux-gnu (64-bit)
## Running under: CentOS Linux 7 (Core)
##
## locale:
##  [1] LC_CTYPE=en_US.UTF-8      LC_NUMERIC=C              LC_TIME=en_US.UTF-8
##  [4] LC_COLLATE=en_US.UTF-8   LC_MONETARY=en_US.UTF-8  LC_MESSAGES=en_US.UTF-8
##  [7] LC_PAPER=en_US.UTF-8    LC_NAME=C                 LC_ADDRESS=C
## [10] LC_TELEPHONE=C          LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C
##
## attached base packages:
## [1] stats      graphics  grDevices  utils      datasets  methods    base
##
## other attached packages:
## [1] knitr_1.11
##
## loaded via a namespace (and not attached):
## [1] magrittr_1.5  formatR_1.2.1 tools_3.2.2  stringi_0.5-5 highr_0.5.1  stringr_1.0.0
## [7] evaluate_0.8

Sys.time()

```

[1] "2015-10-07 04:24:06 NOVT"