

HDAC inhibitor screen analysis

F. Heigwer

06/02/2021

Overview

This vignette is used to analyze a CRISPR screen performed by treating SW480 cells using 8 nM panobinostat in 0.1 % DMSO and 0.1 % DMSO as a control sample. It was screened in duplicates per treatment. Raw counts per sgRNA are the basis of the analysis.

All plots were used as given for the main figure panels. The compiled vignette pdf serves as an overview.

Data input

```
treat1<-read_delim(file = "../raw_data_counts/counts_panobinostat_screen/P1_S7_L001_R1_001.txt",delim =
treat2<-read_delim(file = "../raw_data_counts/counts_panobinostat_screen/P2_S8_L001_R1_001.txt",delim =

dms01<-read_delim(file = "../raw_data_counts/counts_panobinostat_screen/DMS01_S1_L001_R1_001.txt",delim
dms02<-read_delim(file = "../raw_data_counts/counts_panobinostat_screen/DMS02_S2_L001_R1_001.txt",delim

mydata<-treat1 %>%
  left_join(treat2,by="sgRNA") %>%
  left_join(dms01,by="sgRNA") %>%
  left_join(dms02,by="sgRNA") %>%
  extract(sgRNA,c("sgRNA","gene_id"),"((\\w+)_\\.*)"

names(mydata)<-c("sgRNA","gene_id","treat_1","treat_2","dms0_1","dms0_2")

mydata %<>%
  gather(treatment,count,-sgRNA,-gene_id) %>%
  extract(treatment,c("treatment","replicate"),"((\\w+)_((\\d)"))

anno<-read_delim("../meta_data/gene_names.txt",delim="\t")
```

Lets normalize the data

```
mydata %<>%
  mutate(count=if_else(count>10,count,NA_real_)) %>%      #filter low read counts
  group_by(treatment,replicate) %>% #group by treatment and replicate
  mutate(normval=asinh(count/median(count,na.rm = T)))
```

Lets do some data QC

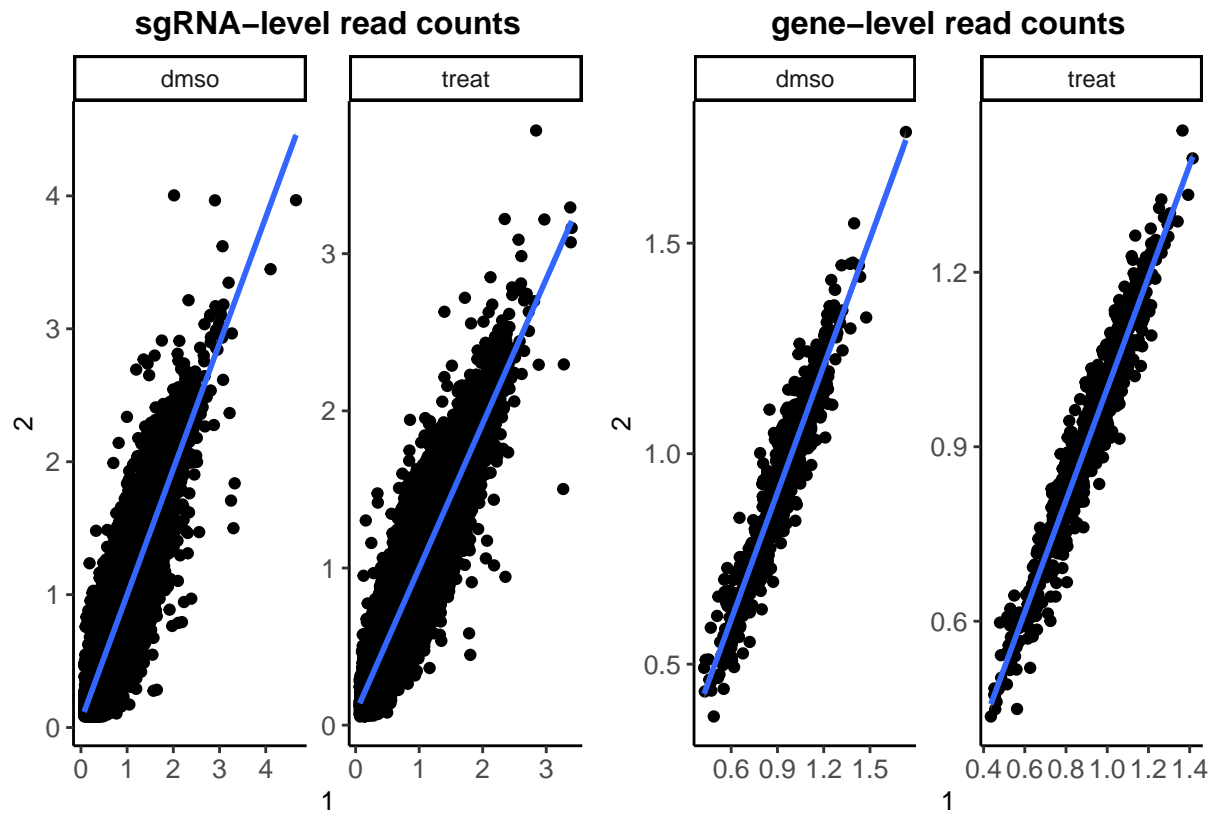
sgRNA-level replicate correlation

```
## $dmso
##   rowname   X1 X2
## 1         1
## 2         2 .88
##
## $treat
##   rowname   X1 X2
## 1         1
## 2         2 .91
```

gene-level replicate correlation

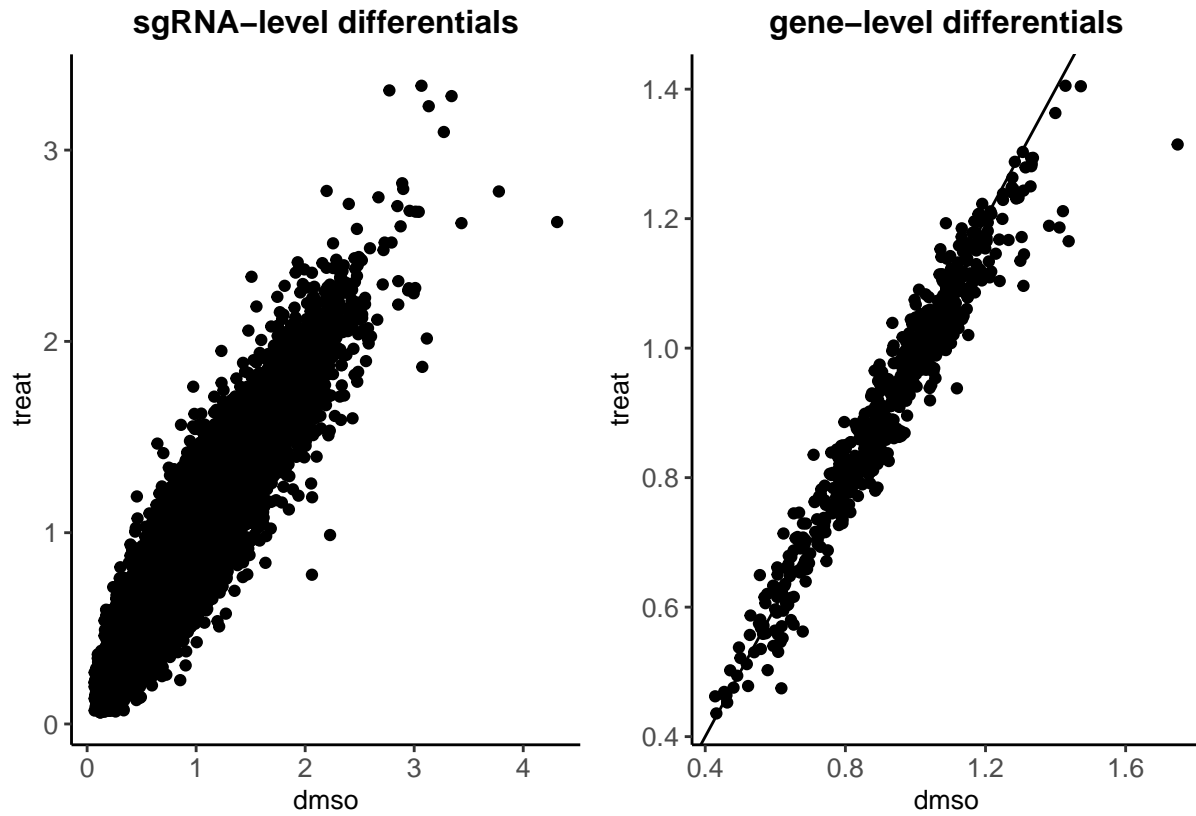
```
## $dmso
##   rowname   X1 X2
## 1         1
## 2         2 .94
##
## $treat
##   rowname   X1 X2
## 1         1
## 2         2 .97
```

graphics



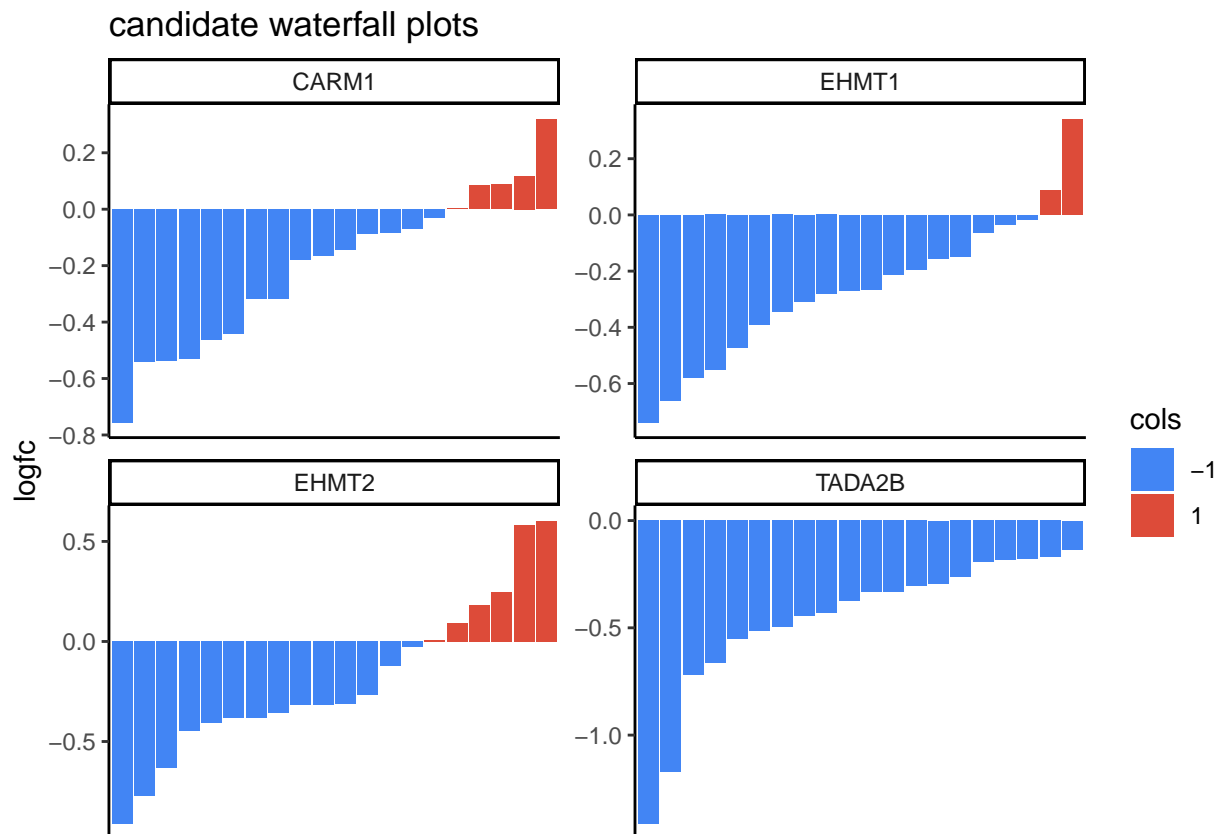
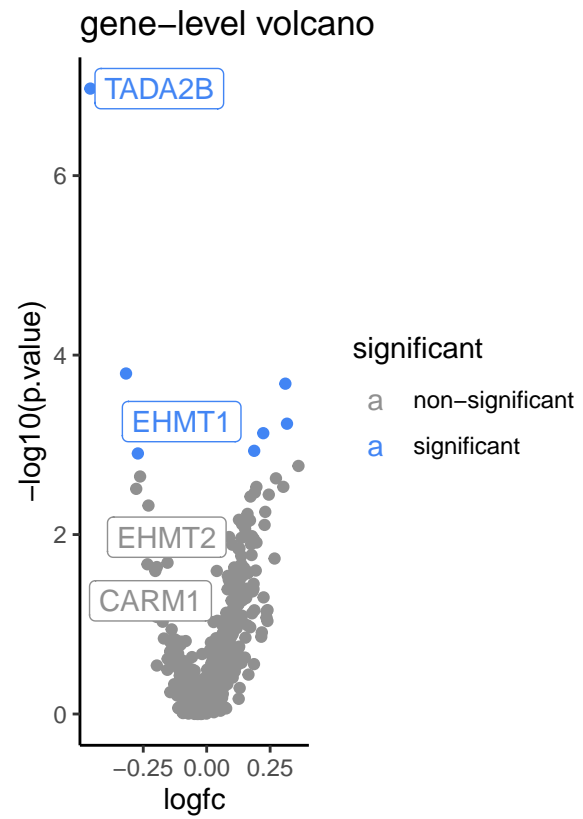
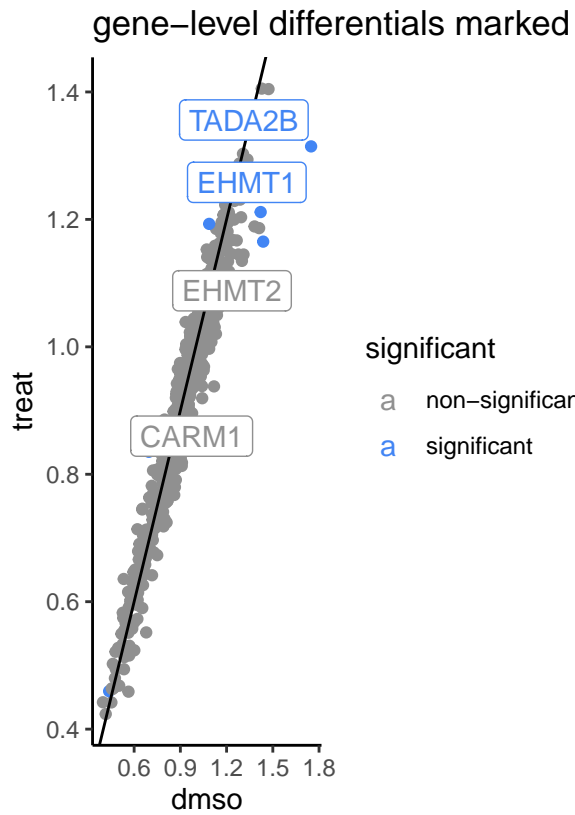
QC looks very nice and good so lets calls some foldchanges and hits thereof

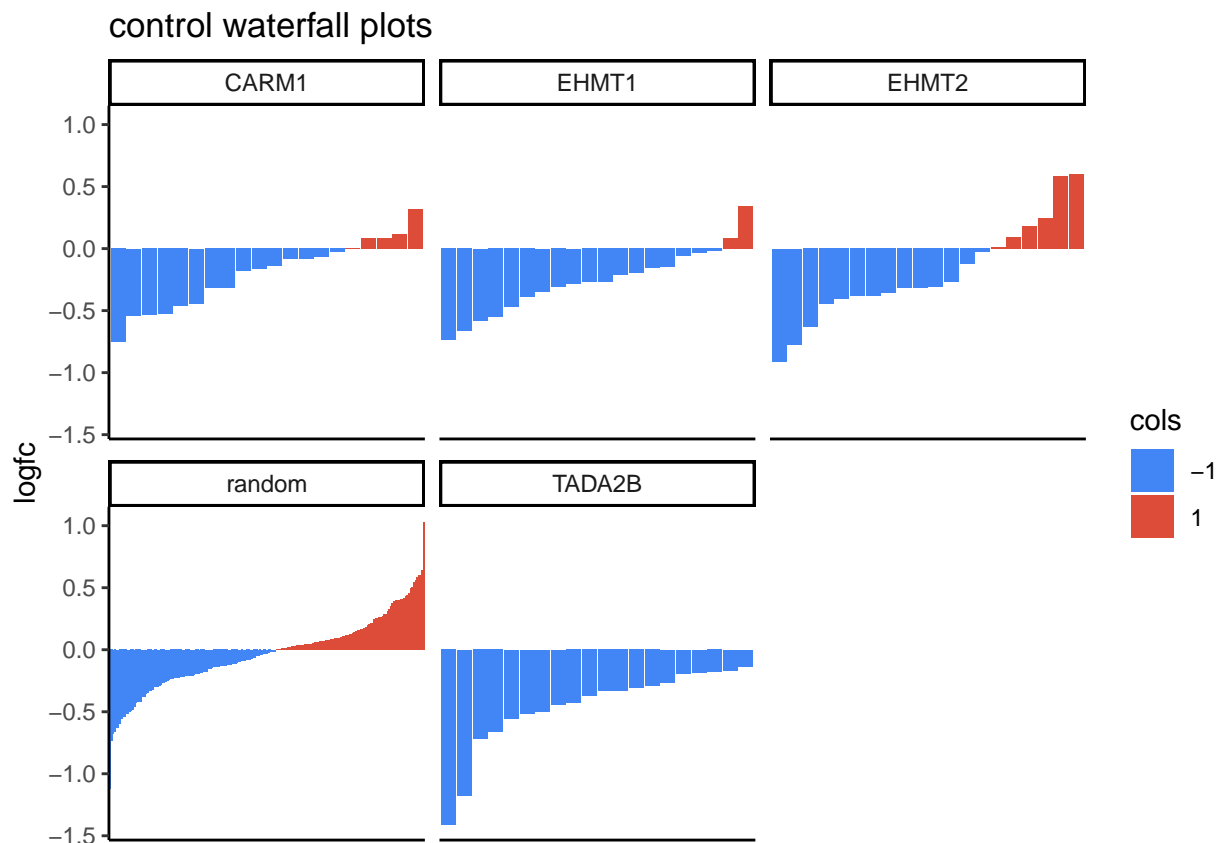
Hit-calling



Statistical analysis

Here we use pairwise wilcoxon rank-sum test between to fold changes of the individual sgRNAs for each gene and the individual negative control sgRNA to determine if the foldchange of the gene represented by independent sgRNAs is significantly different from the non-targeting controls.





Session info

```
## R version 4.0.0 (2020-04-24)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 19041)
##
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=German_Germany.1252 LC_CTYPE=German_Germany.1252
## [3] LC_MONETARY=German_Germany.1252 LC_NUMERIC=C
## [5] LC_TIME=German_Germany.1252
##
## attached base packages:
## [1] stats      graphics  grDevices  utils      datasets  methods   base
##
## other attached packages:
## [1] patchwork_1.0.0 broom_0.5.6      ggrepel_0.8.2   ggrastr_0.2.3
## [5] corr_0.4.2      FitAR_1.94      bestglm_0.37.3  ltsa_1.4.6
## [9] leaps_3.1       lattice_0.20-41 readxl_1.3.1    forcats_0.5.0
## [13] stringr_1.4.0  dplyr_0.8.5     purrr_0.3.4     readr_1.3.1
## [17] tidyr_1.0.3     tibble_3.0.1    ggplot2_3.3.0   tidyverse_1.3.0
##
## loaded via a namespace (and not attached):
```

```

## [1] httr_1.4.2      jsonlite_1.7.2  splines_4.0.0   foreach_1.5.1
## [5] modelr_0.1.7    assertthat_0.2.1 vipor_0.4.5     cellranger_1.1.0
## [9] yaml_2.2.1      pillar_1.4.4    backports_1.1.6 glue_1.4.1
## [13] digest_0.6.25   rvest_0.3.5     colorspace_1.4-1 htmltools_0.4.0
## [17] Matrix_1.2-18   pkgconfig_2.0.3 haven_2.2.0     scales_1.1.1
## [21] grpreg_3.3.0    mgcv_1.8-31     generics_0.1.0 farver_2.0.3
## [25] ellipsis_0.3.1  withr_2.4.1     cli_2.0.2       survival_3.1-12
## [29] magrittr_1.5    crayon_1.3.4    evaluate_0.14   fs_1.4.1
## [33] fansi_0.4.1     nlme_3.1-147    xml2_1.3.2      beeswarm_0.3.1
## [37] registry_0.5-1  tools_4.0.0     hms_0.5.3       lifecycle_1.0.0
## [41] munsell_0.5.0   reprex_0.3.0    glmnet_4.1-1    pls_2.7-3
## [45] compiler_4.0.0  rlang_0.4.10    grid_4.0.0      iterators_1.0.13
## [49] rstudioapi_0.11 labeling_0.3      rmarkdown_2.1   gtable_0.3.0
## [53] codetools_0.2-16 DBI_1.1.0        TSP_1.1-10      R6_2.4.1
## [57] seriation_1.2-9 lubridate_1.7.8 knitr_1.28       shape_1.4.5
## [61] stringi_1.4.6   ggbeeswarm_0.6.0 Rcpp_1.0.4.6     vctrs_0.3.0
## [65] dbplyr_1.4.3    tidysselect_1.1.0 xfun_0.13

```