test_rankings

John Stansfield

November 30, 2017

set up

```
knitr::opts_chunk$set(fig.height = 4)
```

```
library(readr)
library(data.table)
library(HiCcompare)
```

```
## Loading required package: dplyr
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:data.table':
##
##     between, first, last
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(dplyr)
```

Test Rankings

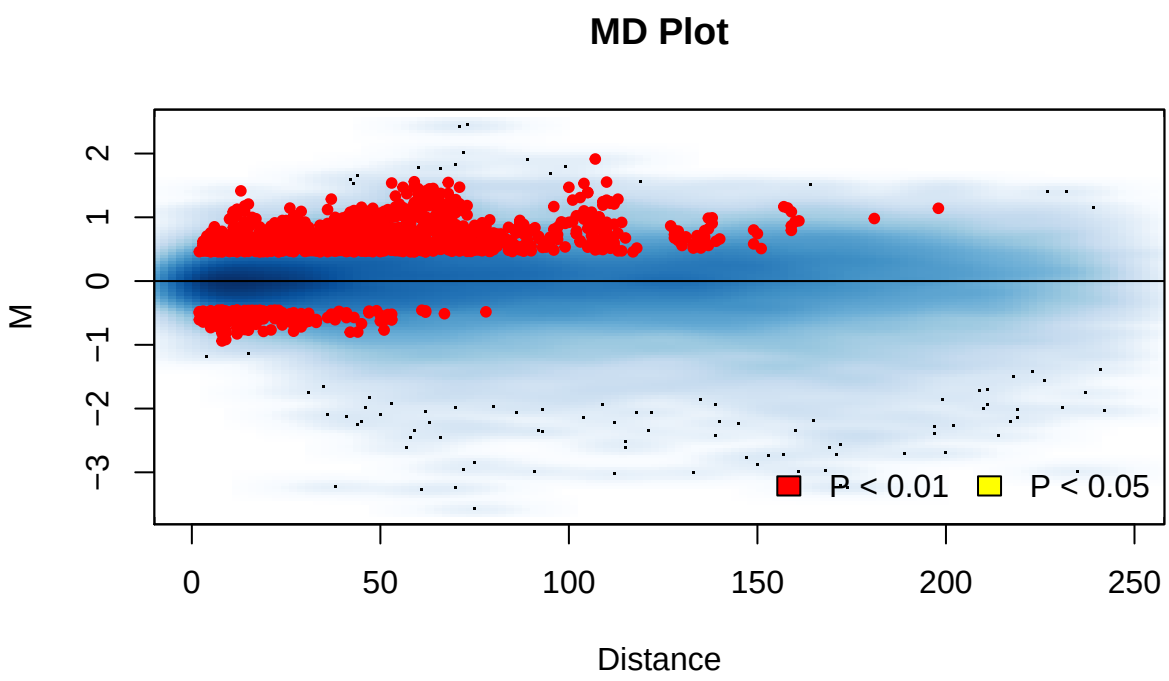
MD plot of p-values using old methods

```
# MD plot of p-values
hic.table <- amyg_dplfc1[[1]]
# MD.plot2(hic.table$adj.M, hic.table$D, hic.table$p.value)
```

Top 5% max ranks where max is taken from M, raw difference, and Avg expression

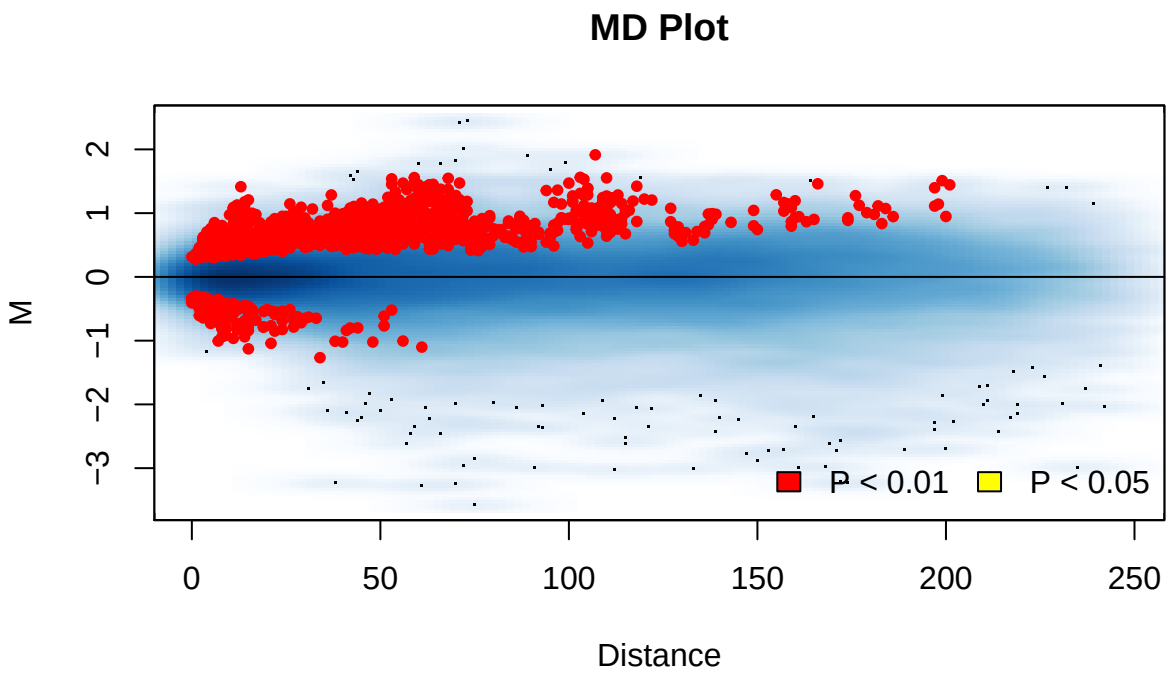
```
# set up
alpha <- 0.05
idx <- 1:(nrow(hic.table) * alpha)


# get top 5% of max ranks where max is taken from M, raw difference, and Avg expression
```

```
hic.table <- hic.table[order(rnkMax),]
topRanks <- rep(1, nrow(hic.table)) # make indicator for top ranks
topRanks[idx] <- 0 # set top ranking rows to 0 indicator for plotting on MD plot
MD.plot2(hic.table$adj.M, hic.table$D, p.val = topRanks)
```



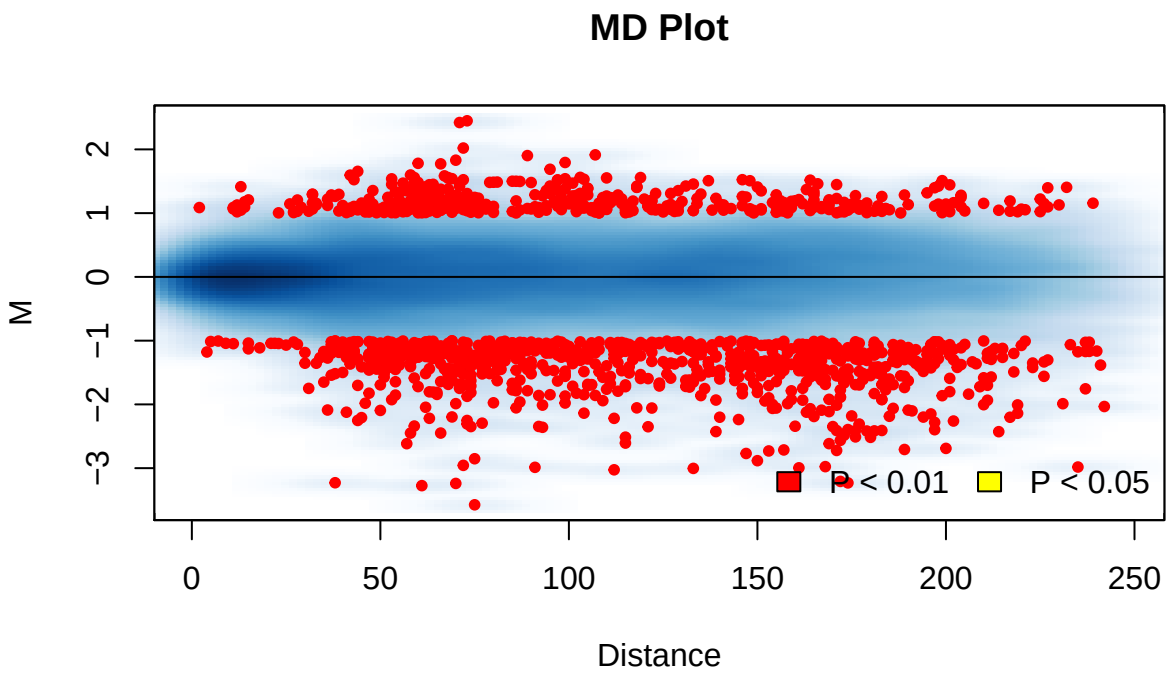
top 5% mean ranks where mean is taken from M, raw difference, and Avg expression

```
# get top 5% of mean ranks
hic.table <- hic.table[order(rnkMean),] # order by mean rank
topRanks <- rep(1, nrow(hic.table)) # make indicator for top ranks
topRanks[idx] <- 0 # set top ranking rows to 0 indicator for plotting on MD plot
MD.plot2(hic.table$adj.M, hic.table$D, p.val = topRanks)
```



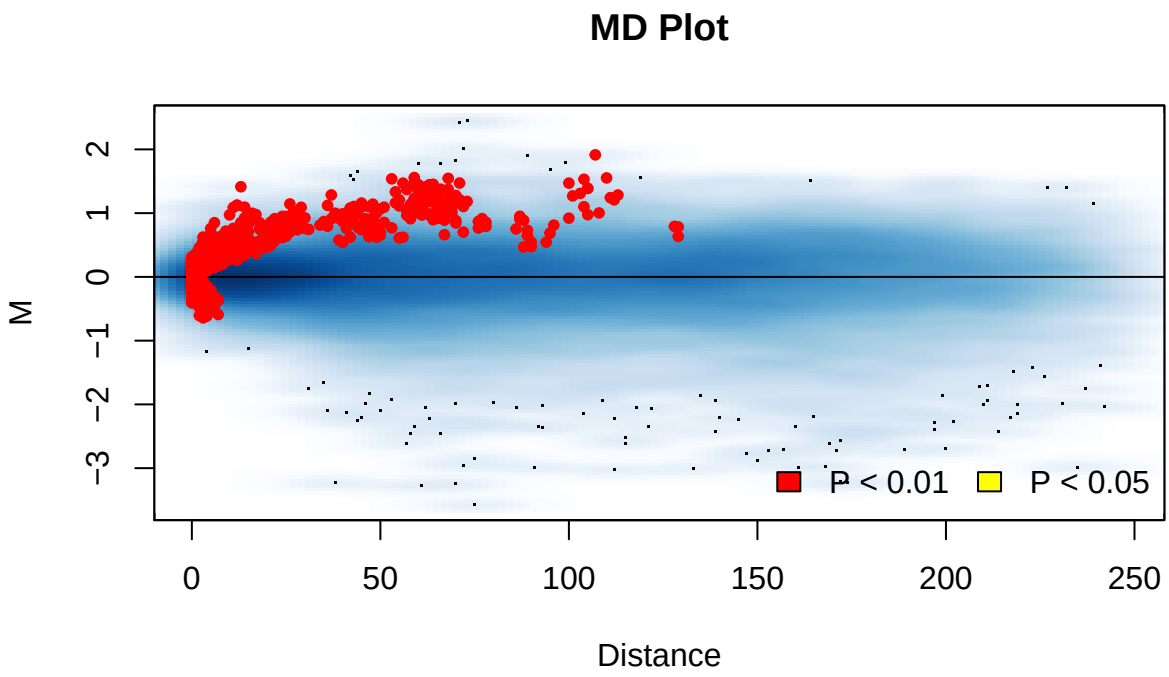
top 5% rnkM and rnkDiff

```
# by top 5% of rnkM and rnkDiff
hic.table <- hic.table[order(rnkM, rnkDiff),] # order by mean rank
topRanks <- rep(1, nrow(hic.table)) # make indicator for top ranks
topRanks[idx] <- 0 # set top ranking rows to 0 indicator for plotting on MD plot
MD.plot2(hic.table$adj.M, hic.table$D, p.val = topRanks)
```



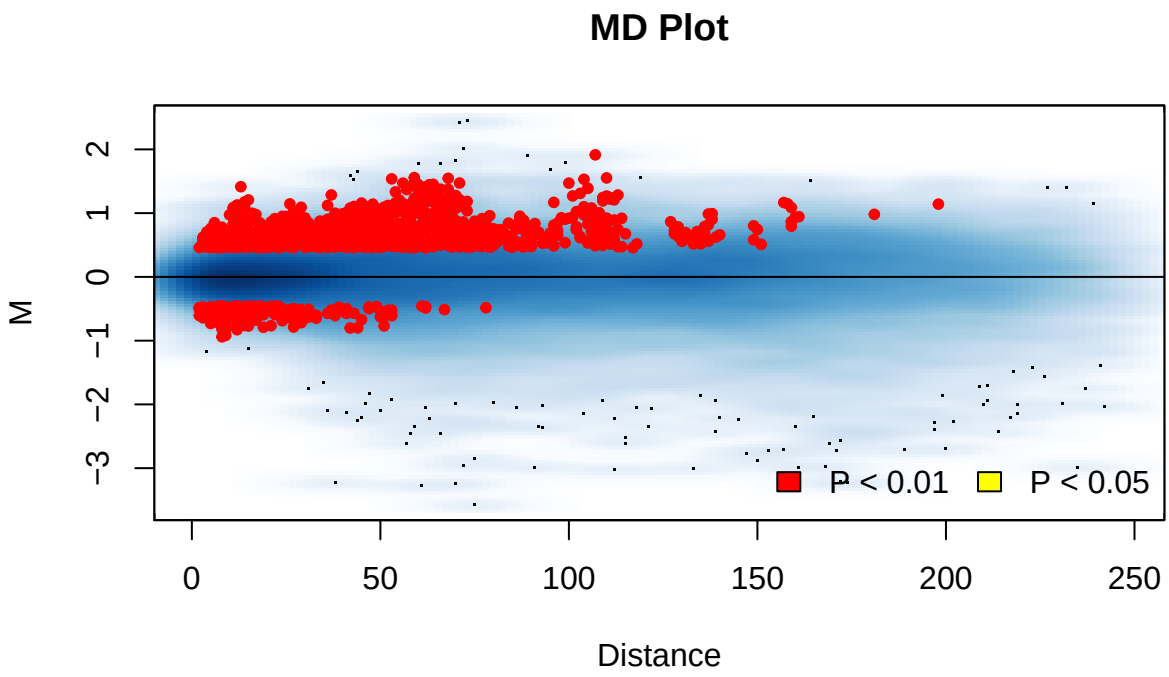
top 5% of rnkDiff

```
# by top 5% of rnkDiff
hic.table <- hic.table[order(rnkDiff),] # order by mean rank
topRanks <- rep(1, nrow(hic.table)) # make indicator for top ranks
topRanks[idx] <- 0 # set top ranking rows to 0 indicator for plotting on MD plot
MD.plot2(hic.table$adj.M, hic.table$D, p.val = topRanks)
```



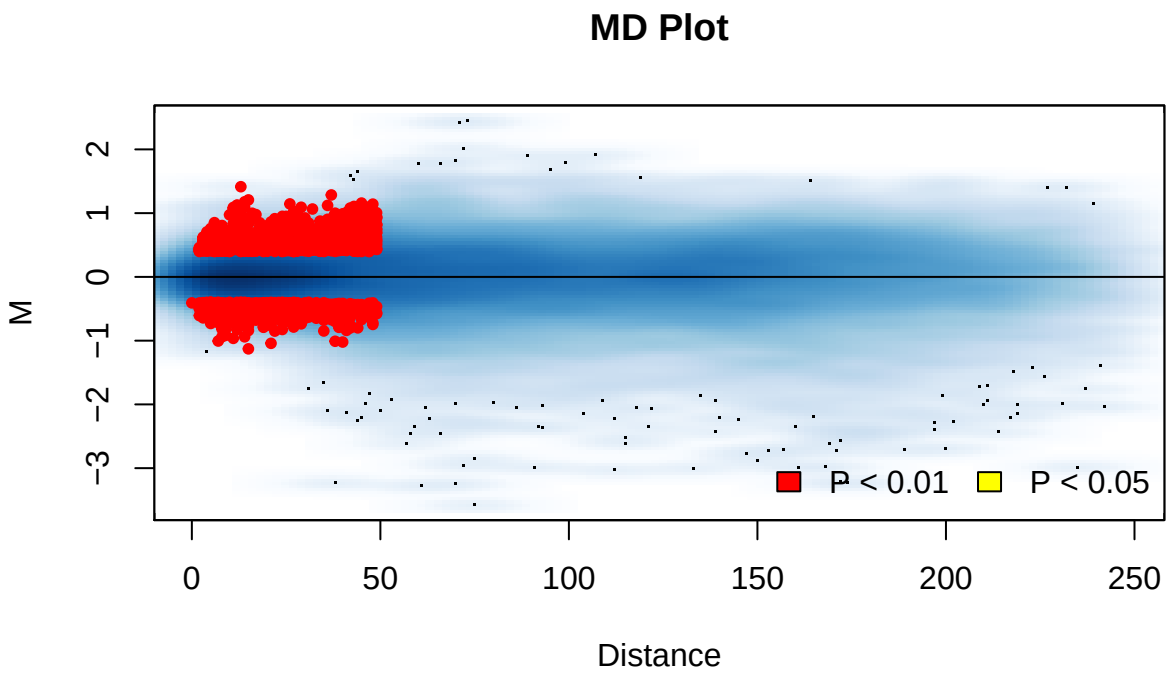
top 5% of max(rnkM, rnkA)

```
# by top 5% of rnkDiff
max_rank <- hic.table %>% dplyr::select(rnkM, rnkA) %>% as.matrix() %>% apply(., 1, max)
hic.table[, rnkMax := max_rank]
hic.table <- hic.table[order(rnkMax),] # order by mean rank
topRanks <- rep(1, nrow(hic.table)) # make indicator for top ranks
topRanks[idx] <- 0 # set top ranking rows to 0 indicator for plotting on MD plot
MD.plot2(hic.table$adj.M, hic.table$D, p.val = topRanks)
```



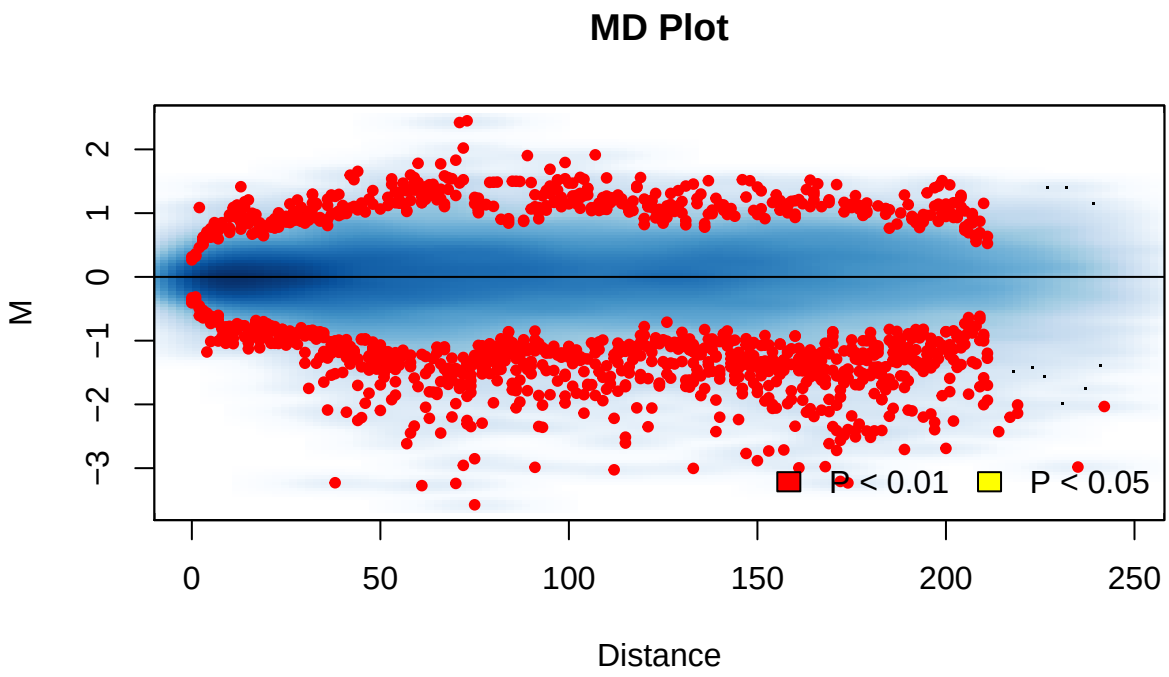
top 5% of max(rnkM, rnkD, rnkA, rnkDiff)

```
# by top 5% of rnkDiff
max_rank <- hic.table %>% dplyr::select(rnkM, rnkA, rnkD, rnkDiff) %>% as.matrix() %>% apply(., 1, max)
hic.table[, rnkMax := max_rank]
hic.table <- hic.table[order(rnkMax),] # order by mean rank
topRanks <- rep(1, nrow(hic.table)) # make indicator for top ranks
topRanks[idx] <- 0 # set top ranking rows to 0 indicator for plotting on MD plot
MD.plot2(hic.table$adj.M, hic.table$D, p.val = topRanks)
```



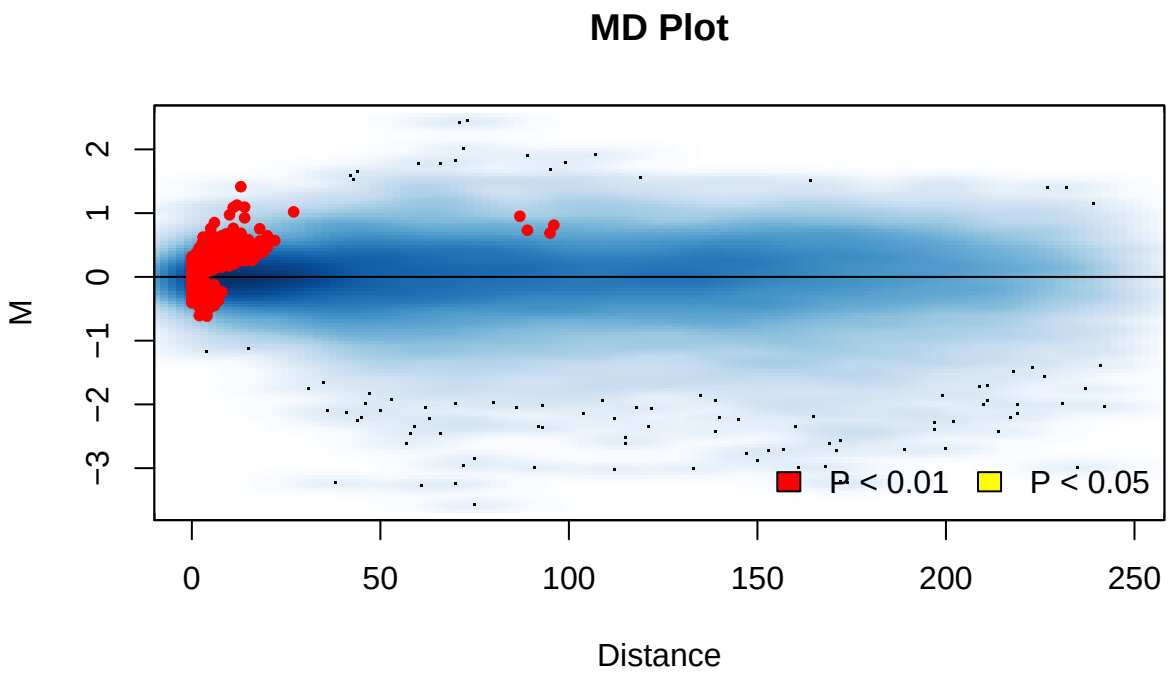
top 5% of rnkM_D

```
hic.table <- hic.table[order(rnkM_D),] # order by mean rank
topRanks <- rep(1, nrow(hic.table)) # make indicator for top ranks
topRanks[idx] <- 0 # set top ranking rows to 0 indicator for plotting on MD plot
MD.plot2(hic.table$adj.M, hic.table$D, p.val = topRanks)
```



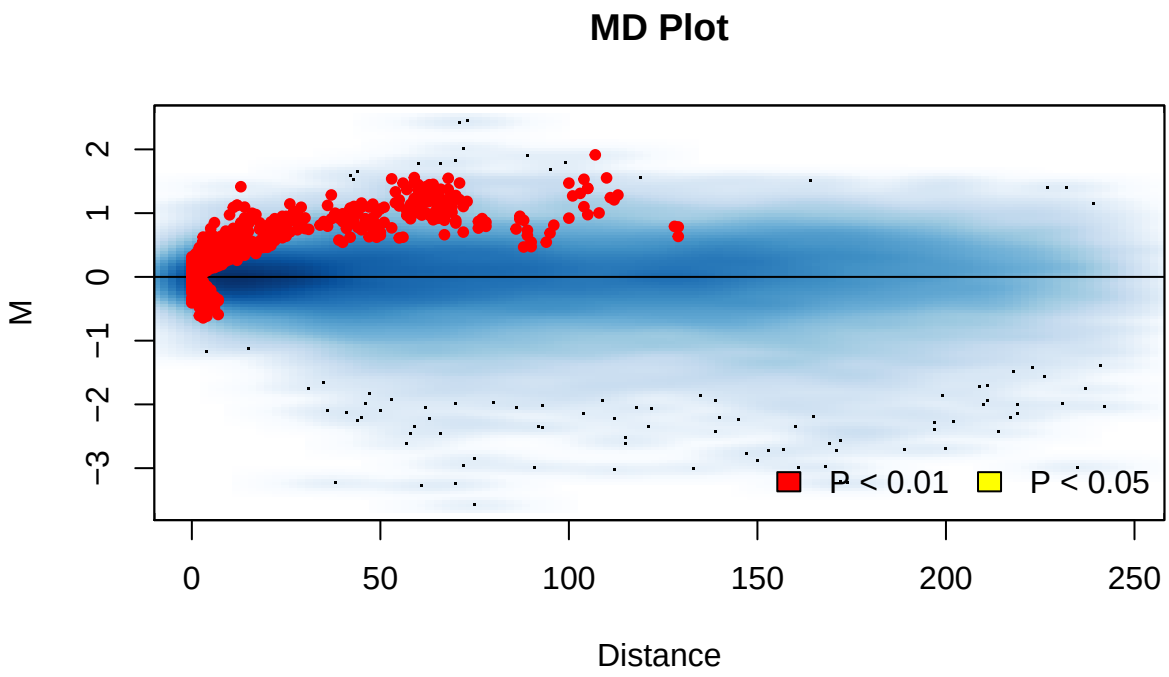
top 5% of max(rankM_D, A, rnkDiff)

```
# by top 5% of rnkDiff
max_rank <- hic.table %>% dplyr::select(rnkM_D, rnkA, rnkDiff) %>% as.matrix() %>% apply(., 1, max)
hic.table[, rnkMax := max_rank]
hic.table <- hic.table[order(rnkMax),] # order by mean rank
topRanks <- rep(1, nrow(hic.table)) # make indicator for top ranks
topRanks[idx] <- 0 # set top ranking rows to 0 indicator for plotting on MD plot
MD.plot2(hic.table$adj.M, hic.table$D, p.val = topRanks)
```



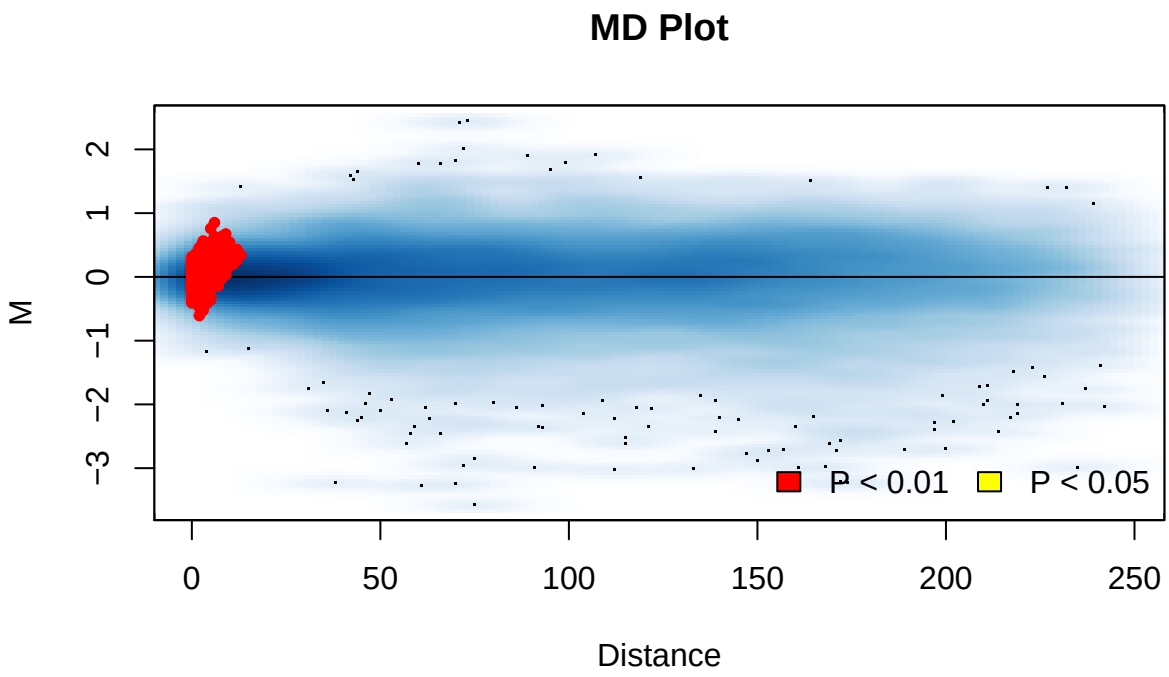
top 5% of max(rankM_D, rnkDiff)

```
# by top 5% of rnkDiff
max_rank <- hic.table %>% dplyr::select(rnkM_D, rnkDiff) %>% as.matrix() %>% apply(., 1, max)
hic.table[, rnkMax := max_rank]
hic.table <- hic.table[order(rnkMax),] # order by mean rank
topRanks <- rep(1, nrow(hic.table)) # make indicator for top ranks
topRanks[idx] <- 0 # set top ranking rows to 0 indicator for plotting on MD plot
MD.plot2(hic.table$adj.M, hic.table$D, p.val = topRanks)
```



top 5% of max(rankM_D, rnkA)

```
# by top 5% of rnkDiff
max_rank <- hic.table %>% dplyr::select(rnkM_D, rnkA) %>% as.matrix() %>% apply(., 1, max)
hic.table[, rnkMax := max_rank]
hic.table <- hic.table[order(rnkMax),] # order by mean rank
topRanks <- rep(1, nrow(hic.table)) # make indicator for top ranks
topRanks[idx] <- 0 # set top ranking rows to 0 indicator for plotting on MD plot
MD.plot2(hic.table$adj.M, hic.table$D, p.val = topRanks)
```

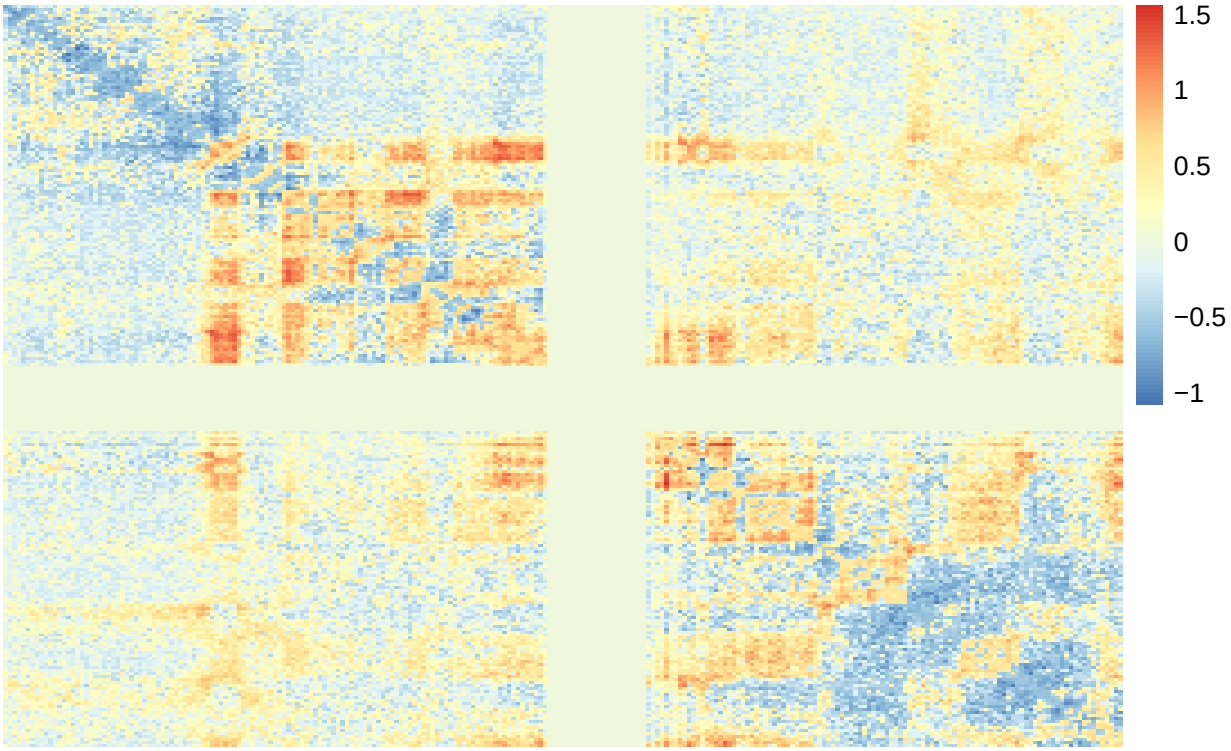


Visualization

Mean rank

```
hic.table = amyg_dplfc1[[1]]
visualize_differences(hic.table, which_rank = 'rnkMean', only_toprank = FALSE)
```

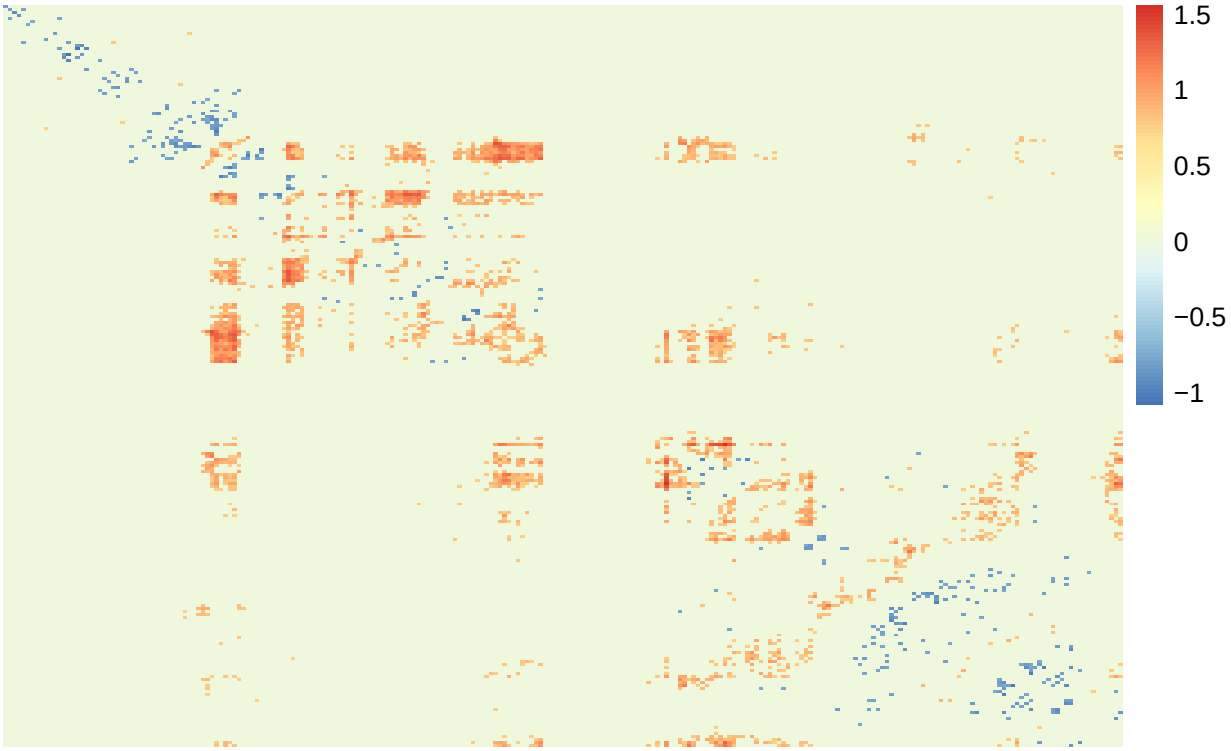
```
## Matrix dimensions: 249x249
## Matrix dimensions: 249x249
```



Top 5% mean ranks

```
visualize_differences(hic.table, which_rank = 'rnkMean', only_toprank = TRUE, proportion = 0.05)
```

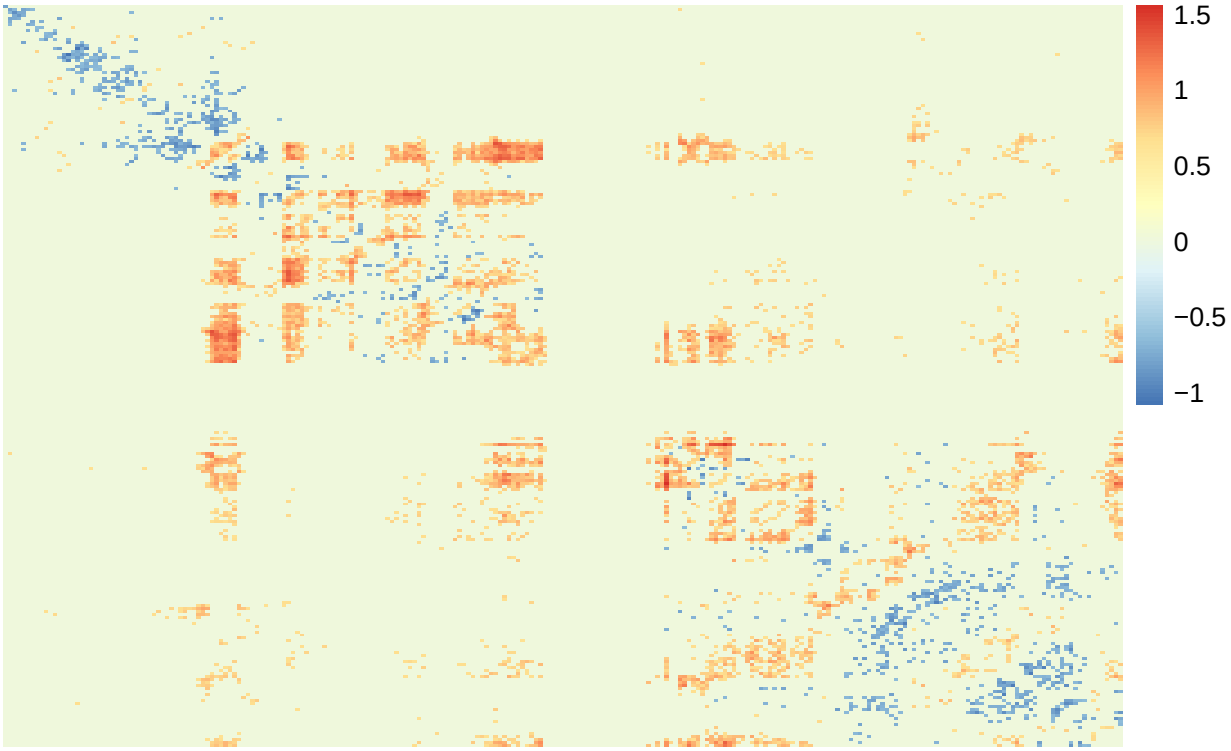
```
## Matrix dimensions: 249x249
## Matrix dimensions: 249x249
## Matrix dimensions: 249x249
```



Top 10% mean ranks

```
visualize_differences(hic.table, which_rank = 'rnkMean', only_toprank = TRUE, proportion = 0.1)
```

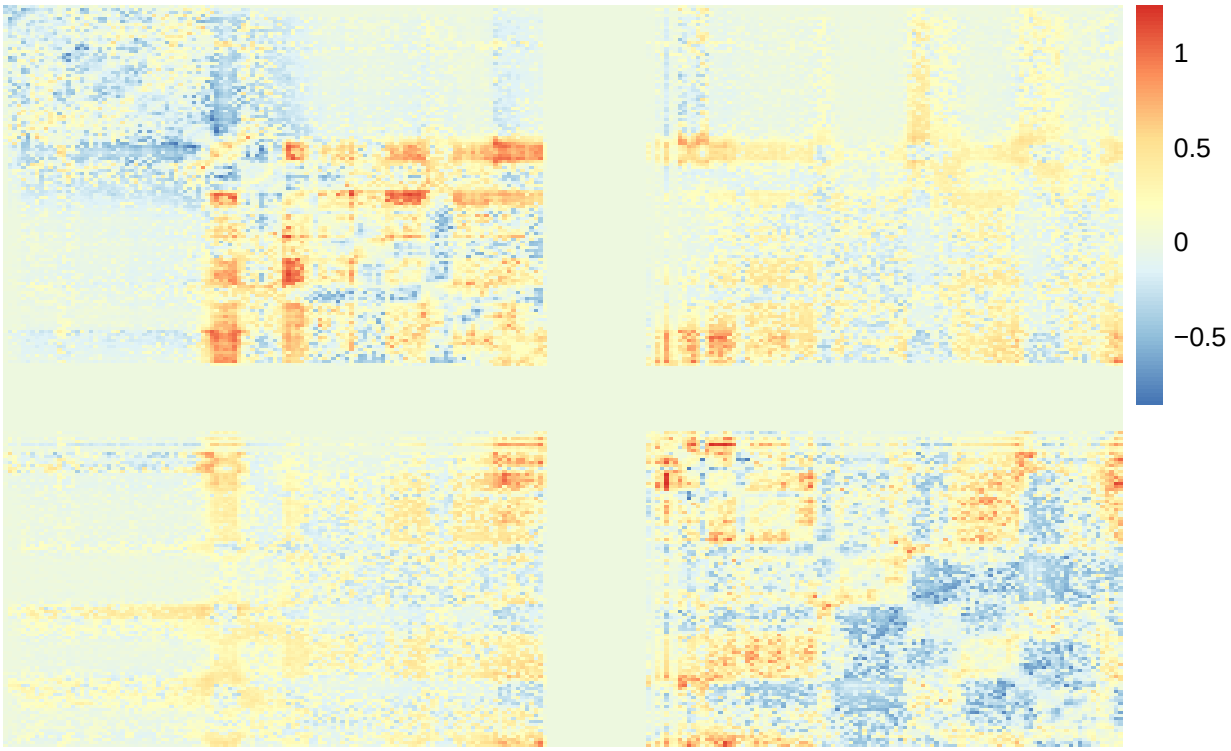
```
## Matrix dimensions: 249x249
## Matrix dimensions: 249x249
## Matrix dimensions: 249x249
```



Max rank

```
hic.table = amyg_dplfc1[[1]]
visualize_differences(hic.table, which_rank = 'rnkMax', only_toprank = FALSE)
```

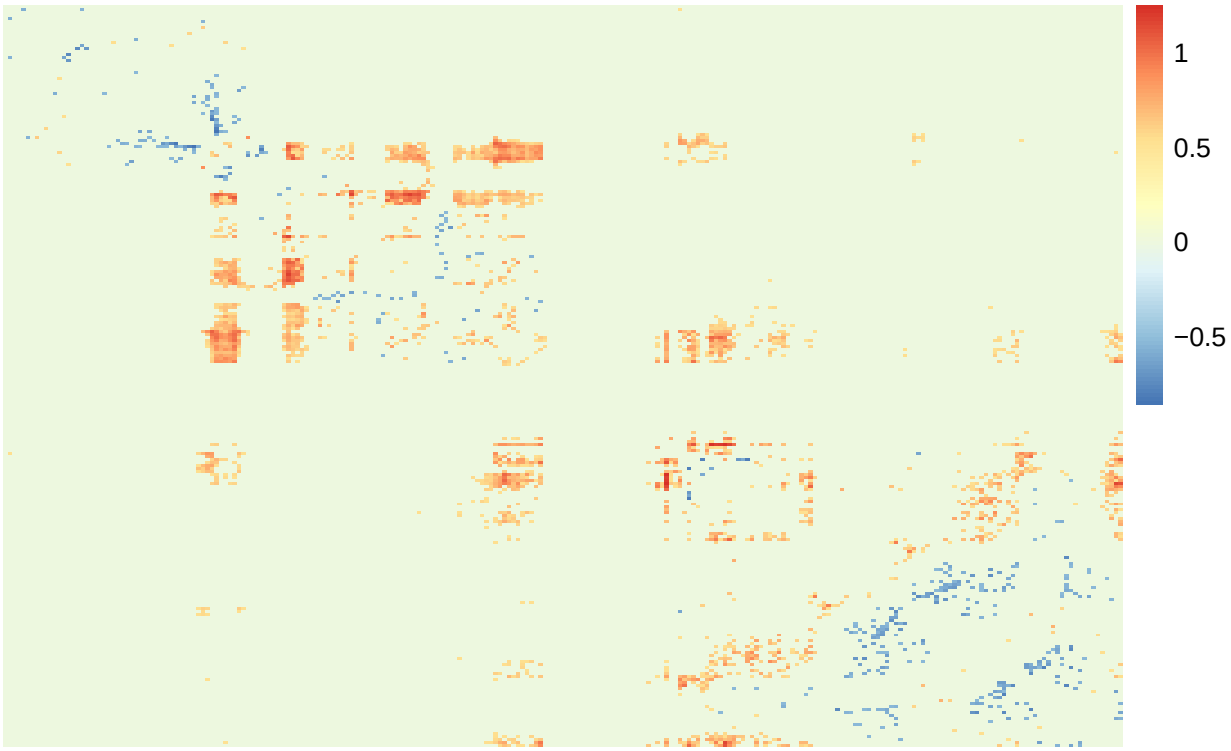
```
## Matrix dimensions: 249x249
## Matrix dimensions: 249x249
```



Top 5% mean ranks

```
visualize_differences(hic.table, which_rank = 'rnkMax', only_toprank = TRUE, proportion = 0.05)
```

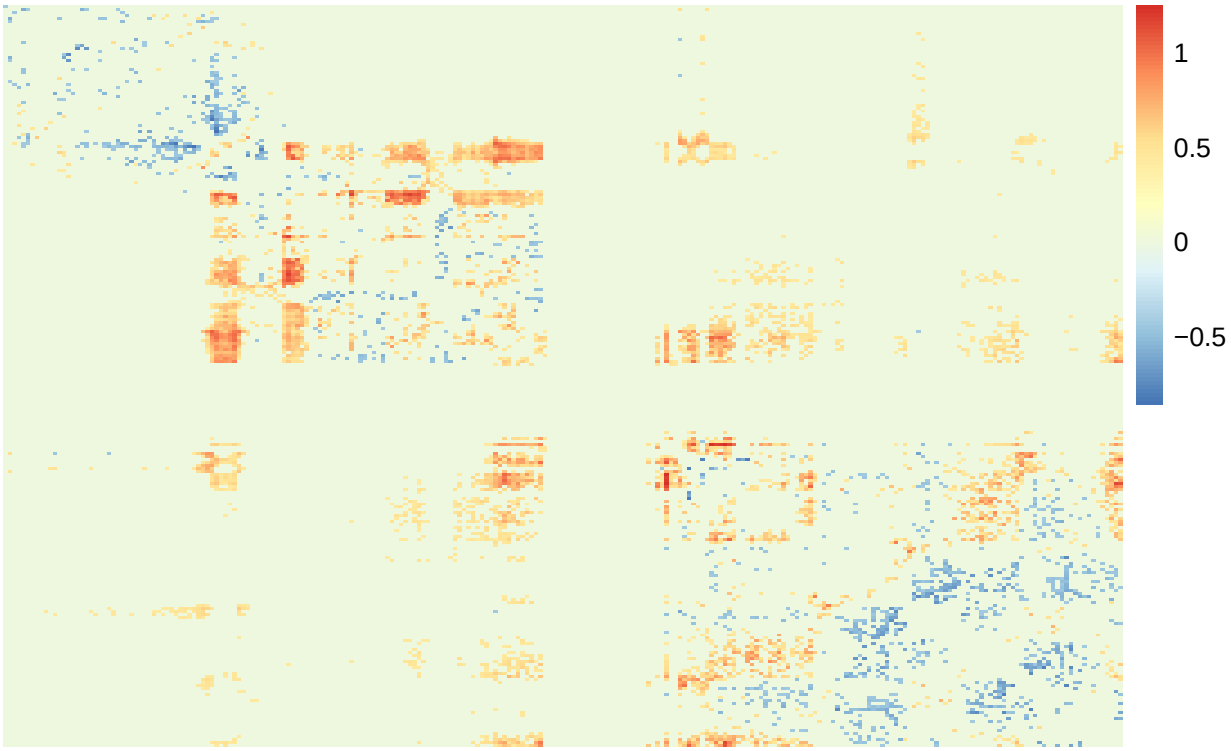
```
## Matrix dimensions: 249x249
## Matrix dimensions: 249x249
## Matrix dimensions: 249x249
```



Top 10% mean ranks

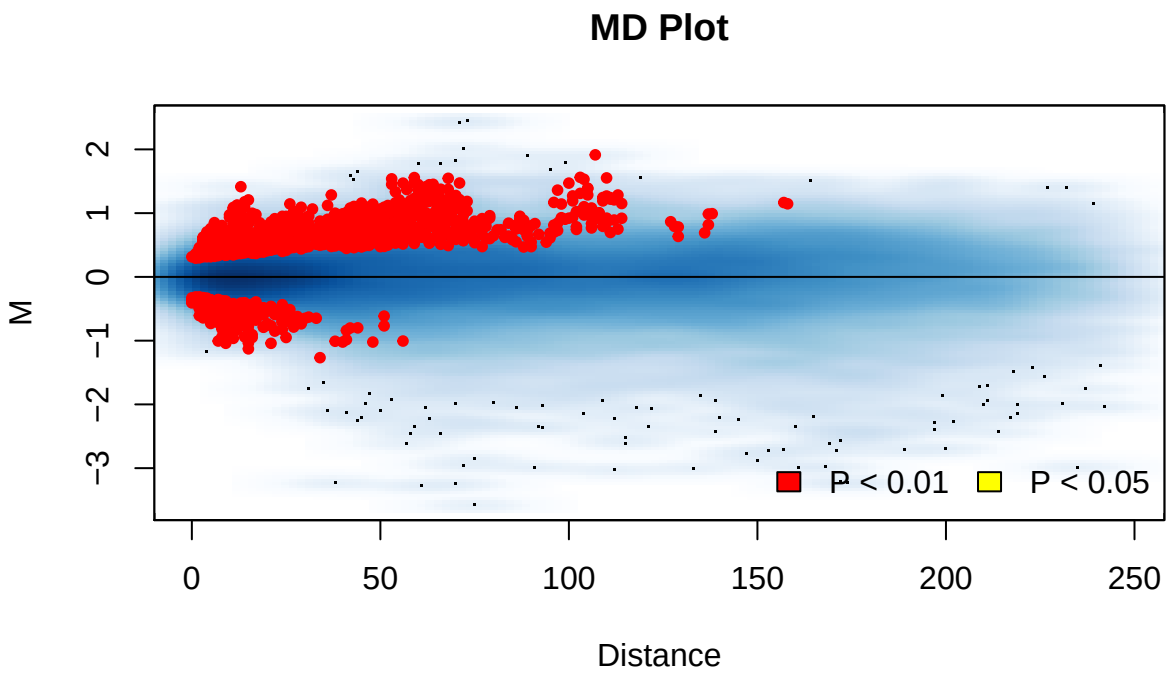
```
visualize_differences(hic.table, which_rank = 'rnkMax', only_toprank = TRUE, proportion = 0.1)
```

```
## Matrix dimensions: 249x249
## Matrix dimensions: 249x249
## Matrix dimensions: 249x249
```



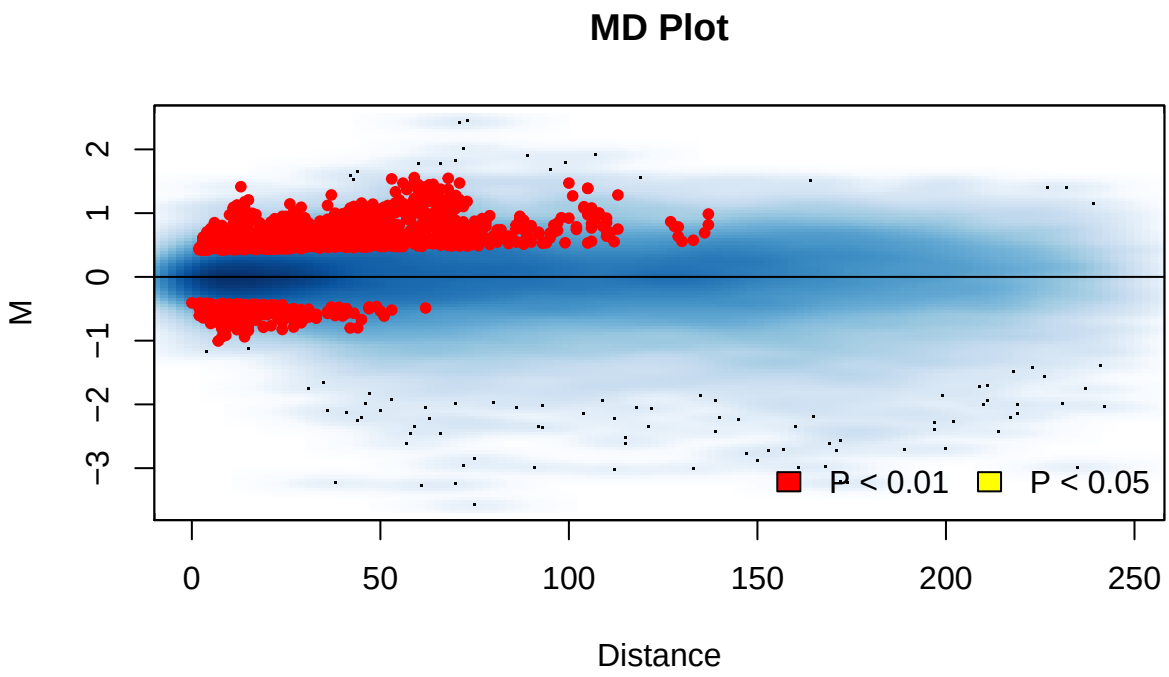
Test using mean A and M with distance weighting

```
hic.table <- amyg_dplfc1[[1]]

mean_rank <- hic.table %>% dplyr::select(rnkM, rnkA) %>% as.matrix() %>% apply(., 1, mean)
hic.table[, rnkMean := mean_rank]
hic.table <- hic.table[order(rnkMean),] # order by mean rank
# create weight for distance
dist_weight <- 1+((hic.table$D + 1)/max(hic.table$D + 1))
hic.table[, rnkMean := dist_weight * rnkMean]
hic.table <- hic.table[order(rnkMean),]

topRanks <- rep(1, nrow(hic.table)) # make indicator for top ranks
topRanks[idx] <- 0 # set top ranking rows to 0 indicator for plotting on MD plot
MD.plot2(hic.table$adj.M, hic.table$D, p.val = topRanks)
```



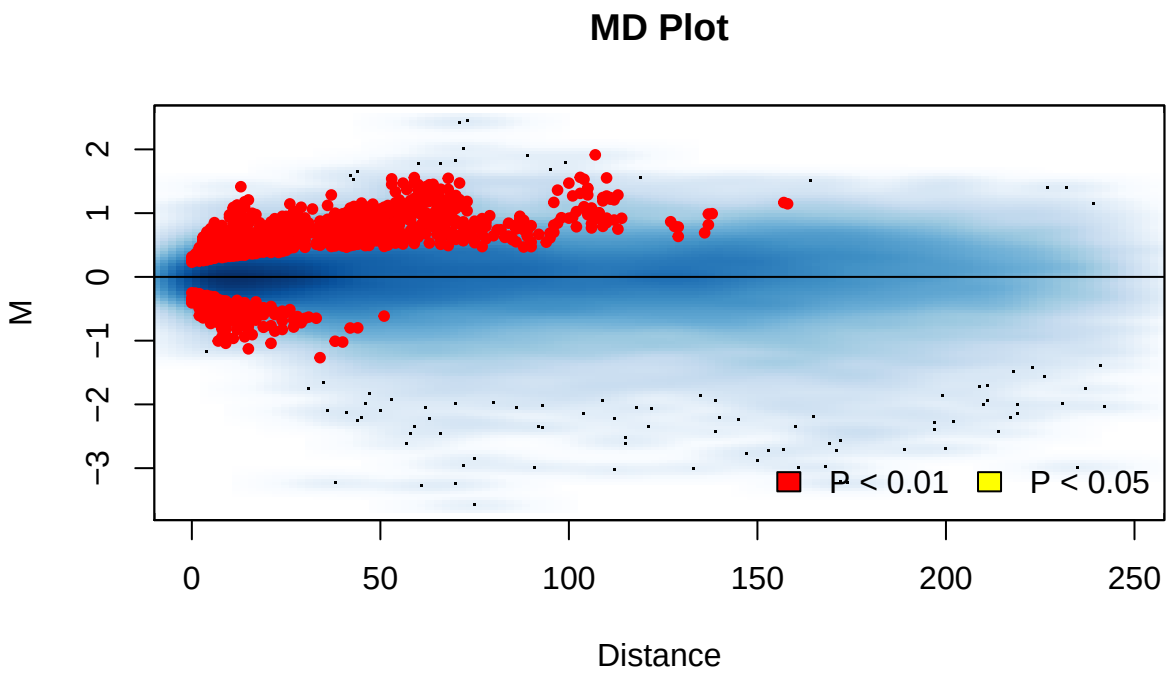
Test using max A and M with distance weighting

```
mean_rank <- hic.table %>% dplyr::select(rnkM, rnkA) %>% as.matrix() %>% apply(., 1, max)
hic.table[, rnkMean := mean_rank]
hic.table <- hic.table[order(rnkMean),] # order by mean rank
# create weight for distance
dist_weight <- 1+((hic.table$D + 1)/max(hic.table$D + 1))
hic.table[, rnkMean := dist_weight * rnkMean]
hic.table <- hic.table[order(rnkMean),]

topRanks <- rep(1, nrow(hic.table)) # make indicator for top ranks
topRanks[idx] <- 0 # set top ranking rows to 0 indicator for plotting on MD plot
MD.plot2(hic.table$adj.M, hic.table$D, p.val = topRanks)
```



Test using mean A, M, raw difference with distance weighting

```
mean_rank <- hic.table %>% dplyr::select(rnkM, rnkA, rnkDiff) %>% as.matrix() %>% apply(., 1, mean)
hic.table[, rnkMean := mean_rank]
hic.table <- hic.table[order(rnkMean),] # order by mean rank
# create weight for distance
dist_weight <- 1+((hic.table$D + 1)/max(hic.table$D + 1))
hic.table[, rnkMean := dist_weight * rnkMean]
hic.table <- hic.table[order(rnkMean),]

topRanks <- rep(1, nrow(hic.table)) # make indicator for top ranks
topRanks[idx] <- 0 # set top ranking rows to 0 indicator for plotting on MD plot
MD.plot2(hic.table$adj.M, hic.table$D, p.val = topRanks)
```


MD Plot
M
2
1
0
−1
−2
−3
0
50
100
150
200
250
Distance
P < 0.01
P < 0.05