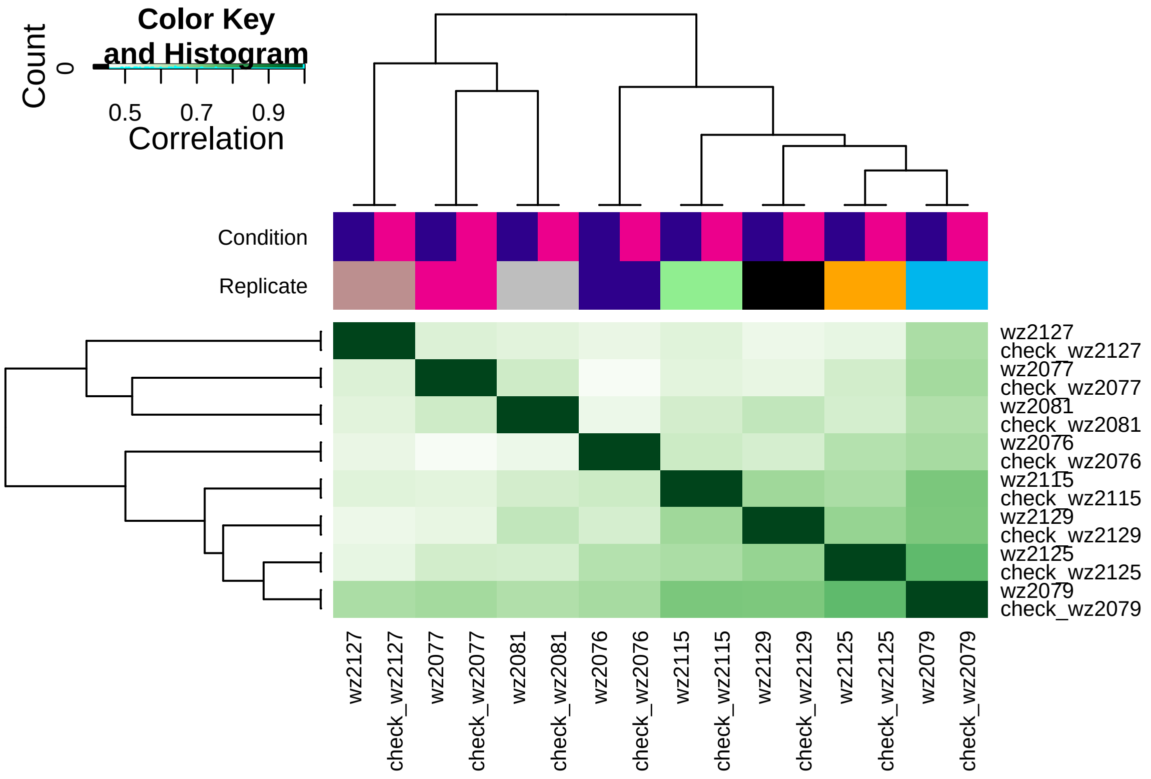
Checking for diffs between old analyzed data (Yanyun primary prostate tumor, AR ChIP-seq) and rhpc_fix_snakemake_ChIPseq of the same input bams

```
setwd("/DATA/t.severson/rhpc_conda_check/")
library(DiffBind)
```

Load data

Check correlation plot

```
plot(mdb)
```



make files

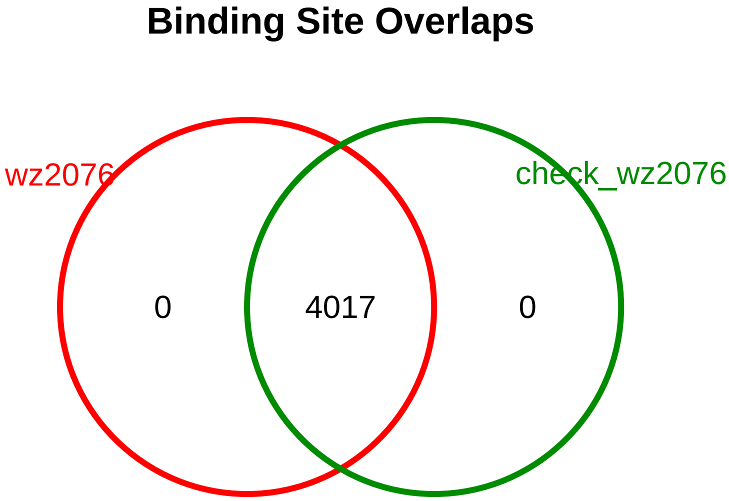
```
wz1 <- file[c(1,9),]
write.csv(wz1, file="/DATA/t.severson/rhpc_conda_check/wz2076_samplesheet.csv")
wz2 <- file[c(2,10),]
write.csv(wz2, file="/DATA/t.severson/rhpc_conda_check/wz2077_samplesheet.csv")
wz3 <- file[c(3,11),]
```

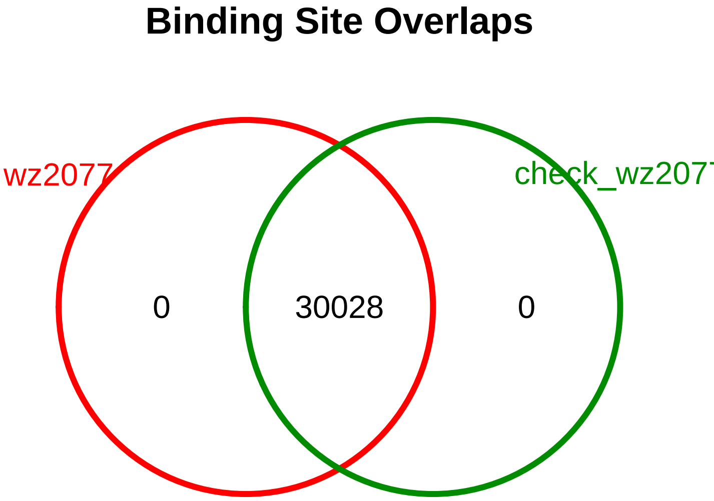
```
write.csv(wz3, file="/DATA/t.severson/rhpc_conda_check/wz2079_samplesheet.csv")
wz4 <- file[c(4,12),]
write.csv(wz4, file="/DATA/t.severson/rhpc_conda_check/wz2081_samplesheet.csv")
wz5 <- file[c(5,13),]
write.csv(wz5, file="/DATA/t.severson/rhpc_conda_check/wz2115_samplesheet.csv")
wz6 <- file[c(6,14),]
write.csv(wz6, file="/DATA/t.severson/rhpc_conda_check/wz2125_samplesheet.csv")
wz7 <- file[c(7,15),]
write.csv(wz7, file="/DATA/t.severson/rhpc_conda_check/wz2127_samplesheet.csv")
wz8 <- file[c(8,16),]
write.csv(wz8, file="/DATA/t.severson/rhpc_conda_check/wz2129_samplesheet.csv")
```

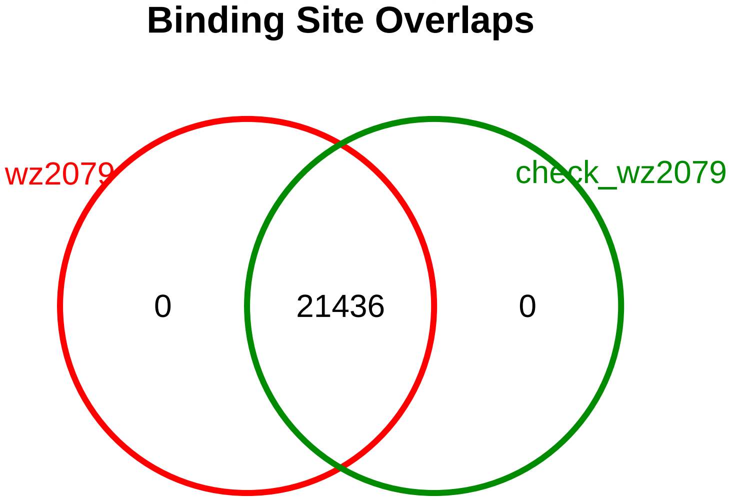
Analyze the old versus new pipeline peaks

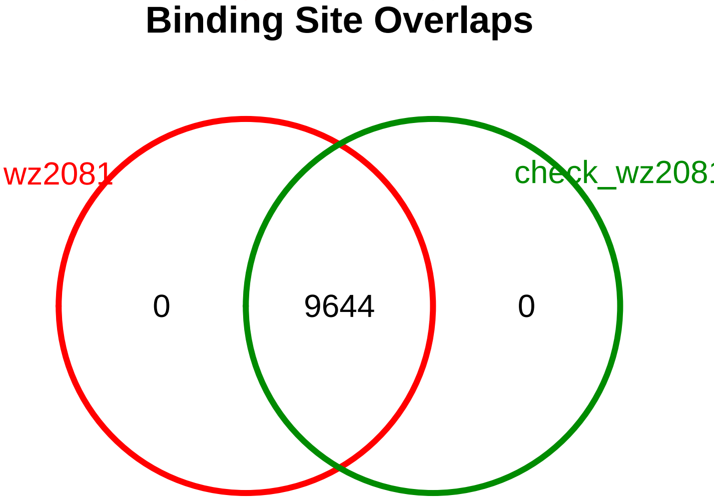
```
files <- list.files(path="/DATA/t.severson/rhpc_conda_check/", pattern="*samplesheet.csv", full.names=T


db <- lapply(files, function(df){dba(sampleSheet = df)})
#double check input files, if they are right, print them to screen
# for (i in seq_along(db)) {
#   bams   <- db[[i]]$samples$bamReads
#   peaks  <- db[[i]]$samples$Peaks
#   inputs <- db[[i]]$samples$bamControl
#
#   print("sample bams")
#   print(bams)
#   print("sample peak files")
#   print(peaks)
#   print("sample input files")
#   print(inputs)
#
# }
```

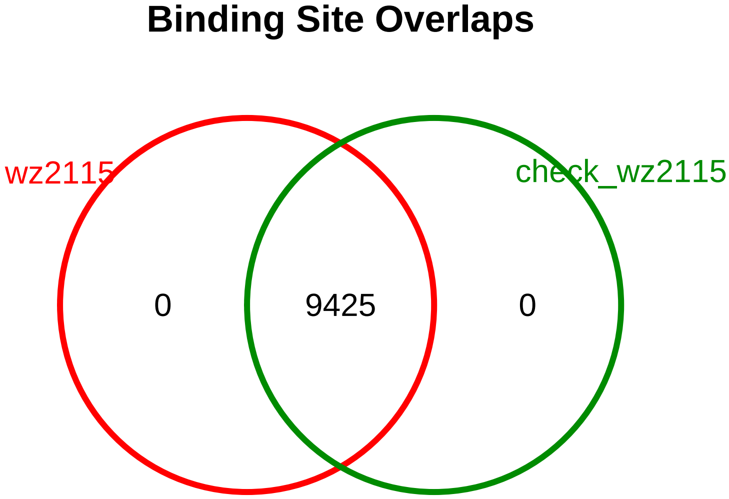
```
outpath <- "/DATA/t.severson/rhpc_conda_check/diffbind_check_out/"
for (i in seq_along(db)){
  # get output files ready (use name from input files)
  x <- unlist(strsplit(files[i],"/"))#split name
  y <- unlist(strsplit(x[6],"_"))
  sample <- y[1]
  pdffile <- paste(outpath,sample,"_venn.pdf", sep = '')
  outfile1 <- paste(outpath,sample,"_sample1_only.txt", sep='')
  outfile2 <- paste(outpath,sample,"_sample2_only.txt", sep='')
  outfile4 <- paste(outpath,sample,"_input_inAll_only.txt", sep='')

  # run dba and output venn file for each sample sheet
  pdf(file=pdffile, width=10, height=10)
  dba.plotVenn(db[[i]], db[[i]]$masks$AR)
  dev.off()

    #output data files
        tf <-dba.plotVenn(db[[i]], db[[i]]$masks$AR, DataType = DBA_DATA_FRAME)
        write.table(tf$onlyA, row.names=F, sep="\t", col.names = FALSE, file=outfile1)
        write.table(tf$onlyB, row.names=F, sep="\t", col.names = FALSE, file=outfile2)
        write.table(tf$inAll, row.names=F, sep="\t", col.names = FALSE, file=outfile4)
}
```

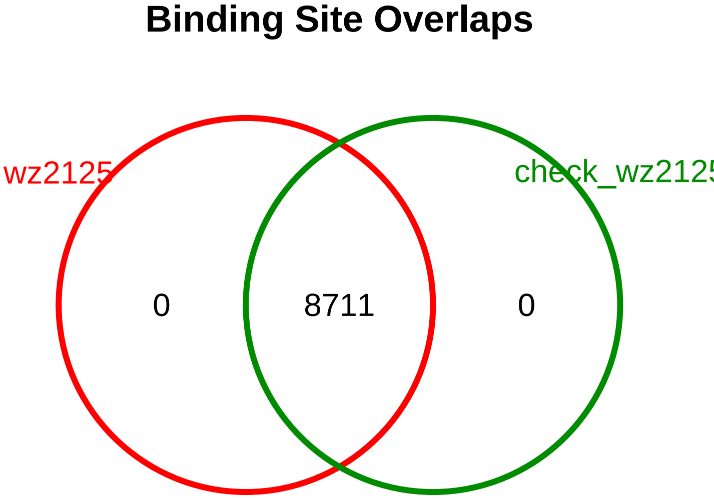


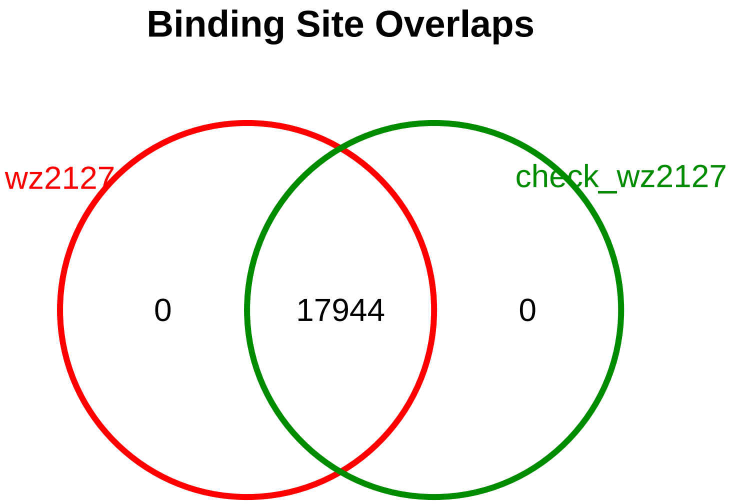


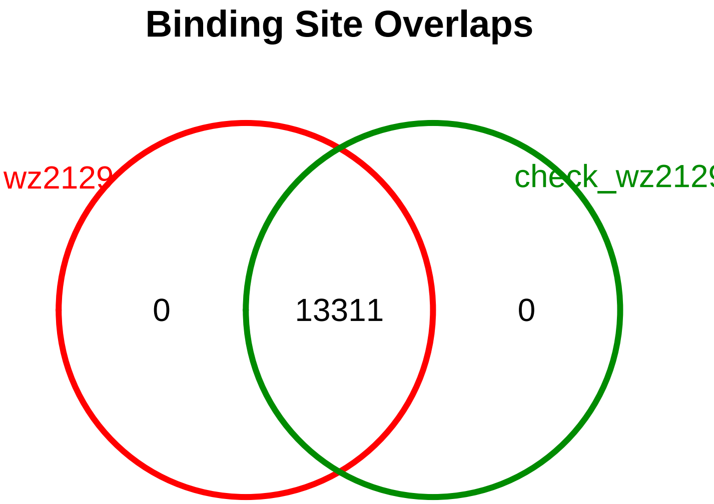












```
print(sessionInfo())
```

```
## R version 3.4.4 (2018-03-15)
## Platform: x86_64-pc-linux-gnu (64-bit)
## Running under: Ubuntu 16.04.6 LTS
##
## Matrix products: default
## BLAS: /usr/lib/openblas-base/libblas.so.3
## LAPACK: /usr/lib/libopenblasp-r0.2.18.so
##
## locale:
##  [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
##  [3] LC_TIME=en_US.UTF-8        LC_COLLATE=en_US.UTF-8
##  [5] LC_MONETARY=en_US.UTF-8    LC_MESSAGES=en_US.UTF-8
##  [7] LC_PAPER=en_US.UTF-8       LC_NAME=C
##  [9] LC_ADDRESS=C               LC_TELEPHONE=C
## [11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C
##
## attached base packages:
## [1] parallel  stats4    stats     graphics  grDevices utils     datasets
## [8] methods   base
##
## other attached packages:
##  [1] DiffBind_2.4.8             SummarizedExperiment_1.6.5
##  [3] DelayedArray_0.2.7         matrixStats_0.54.0
##  [5] Biobase_2.36.2             GenomicRanges_1.28.6
##  [7] GenomeInfoDb_1.12.3        IRanges_2.10.5
##  [9] S4Vectors_0.14.7           BiocGenerics_0.22.1
##
## loaded via a namespace (and not attached):
##  [1] Category_2.42.1          bitops_1.0-6
##  [3] bit64_0.9-7              RColorBrewer_1.1-2
##  [5] progress_1.2.0           httr_1.3.1
##  [7] rprojroot_1.3-2          tools_3.4.4
##  [9] backports_1.1.2          R6_2.3.0
## [11] KernSmooth_2.23-15       DBI_1.0.0
## [13] lazyeval_0.2.1           colorspace_1.3-2
## [15] tidyselect_0.2.5         prettyunits_1.0.2
## [17] bit_1.1-14               compiler_3.4.4
## [19] sendmailR_1.2-1          graph_1.54.0
## [21] rtracklayer_1.36.6       checkmate_1.8.5
## [23] caTools_1.17.1.1         scales_1.0.0
## [25] BatchJobs_1.7            genefilter_1.58.1
## [27] RBGL_1.52.0              stringr_1.3.1
## [29] digest_0.6.18            Rsamtools_1.28.0
## [31] rmarkdown_1.10           AnnotationForge_1.18.2
## [33] XVector_0.16.0           base64enc_0.1-3
## [35] pkgconfig_2.0.2          htmltools_0.3.6
## [37] limma_3.32.10            rlang_0.3.0.1
## [39] RSQLite_2.1.1            BBmisc_1.11
## [41] bindr_0.1.1              GOstats_2.42.0
## [43] hwriter_1.3.2            BiocParallel_1.10.1
## [45] gtools_3.8.1             dplyr_0.7.8
## [47] RCurl_1.95-4.11          magrittr_1.5
## [49] GO.db_3.4.1              GenomeInfoDbData_0.99.0
## [51] Matrix_1.2-14            Rcpp_1.0.0
## [53] munsell_0.5.0            stringi_1.2.4
## [55] yaml_2.2.0               edgeR_3.18.1
## [57] zlibbioc_1.22.0          gplots_3.0.1
## [59] plyr_1.8.4               grid_3.4.4
## [61] blob_1.1.1               ggrepel_0.8.0
## [63] gdata_2.18.0             crayon_1.3.4
## [65] lattice_0.20-38          splines_3.4.4
## [67] Biostrings_2.44.2        GenomicFeatures_1.28.5
## [69] annotate_1.54.0          hms_0.4.2
## [71] locfit_1.5-9.1           knitr_1.20
## [73] pillar_1.3.0             rjson_0.2.20
## [75] systemPipeR_1.10.2       biomaRt_2.37.9
## [77] XML_3.98-1.16            glue_1.3.0
## [79] evaluate_0.12            ShortRead_1.34.2
## [81] latticeExtra_0.6-28      data.table_1.11.8
## [83] gtable_0.2.0             purrr_0.2.5
## [85] amap_0.8-16              assertthat_0.2.0
## [87] ggplot2_3.1.0            xtable_1.8-3
## [89] survival_2.44-1.1        pheatmap_1.0.10
## [91] tibble_1.4.2             GenomicAlignments_1.12.2
## [93] AnnotationDbi_1.38.2     memoise_1.1.0
## [95] bindrcpp_0.2.2           brew_1.0-6
## [97] GSEABase_1.38.2
```