main

July 10, 2021

1 Visualize results, local splicing

1.1 Load Libraries

[1]:
```
library(tidyverse)
library(ggplot2)
library(DT)
library(leafcutter)
library(reshape2)
library(gridExtra)
library(intervals) # needed for pretty strand arrow placement
library(foreach)
library(grid)
library(gtable)
library(ggrepel)
```

```
  Attaching packages                                   tidyverse
1.3.1

  ggplot2 3.3.5      purrr   0.3.4
  tibble  3.1.2      dplyr   1.0.7
  tidyr   1.1.3      stringr 1.4.0
  readr   1.4.0      forcats 0.5.1

  Conflicts
tidyverse_conflicts()
  dplyr::filter() masks stats::filter()
  dplyr::lag()    masks stats::lag()

Loading required package: Rcpp


Attaching package: ‘reshape2’


The following object is masked from ‘package:tidyr’:

    smiths
```

```
Attaching package: 'gridExtra'

The following object is masked from 'package:dplyr':

    combine

Attaching package: 'intervals'

The following object is masked from 'package:purrr':

    reduce

The following object is masked from 'package:tidyr':

    expand

Attaching package: 'foreach'

The following objects are masked from 'package:purrr':

    accumulate, when
```

1.2 Summary of results

[2]:
```
lname = load('../../_m/leafviz.RData')
lname
```

1. 'introns' 2. 'clusters' 3. 'counts' 4. 'meta' 5. 'exons_table' 6. 'pca' 7. 'intron_summary' 8. 'cluster_summary' 9. 'introns_to_plot' 10. 'cluster_ids' 11. 'sample_table' 12. 'annotation_code' 13. 'code'

[3]:
```
levels(meta$group) <- c("Female", "Male")
sample_table
```

A data.frame: 2 × 2

group	count
<chr>	<int>
F	114
M	245

[4]:
```
cluster_summary
```

A data.frame: 3 × 2

Results	n
<chr>	<int>
Number of differentially spliced clusters at FDR = 0.05	352
Fully annotated	160
Contain unannotated junctions	192

[5]:
```
intron_summary
```

A data.frame: 5 × 2

Results	n
<chr>	<int>
Number of fully annotated junctions	1249
Number of junctions with cryptic 5' splice site	119
Number of junctions with cryptic 3' splice site	145
Number of junctions with two cryptic splice sites	38
Number of novel junctions that connect two annotated splice sites	118

[6]:
```
clusters['gene'] <- gsub("</i>", "", gsub("<i>", "", clusters$gene))
head(clusters)
```

A data.frame: 6 × 6

	clusterID	N	coord	gene	annotation	FDR
	<chr>	<dbl>	<chr>	<chr>	<chr>	<dbl
352	clu_755_-	3	chrX:53217966-53220839	KDM5C	annotated	2.27e
350	clu_744_-	10	chrX:53176622-53193437	KDM5C	cryptic	3.64e
282	clu_109096_+	9	chr6:72292046-72390598	RIMS1	cryptic	1.89e
291	clu_143938_-	22	chr6:130863718-130885096	EPB41L2	cryptic	1.89e
340	clu_61504_+	3	chrX:47199106-47199480	UBA1	cryptic	2.45e
227	clu_21090_-	11	chr3:197066754-197081051	DLG1	cryptic	2.32e

[7]:
```
write.table(clusters, file="cluster_ds_results_annotated.txt",
            sep="\t", quote=FALSE, row.names=FALSE)
```

1.3 Generate plots

1.3.1 Define functions

[8]:
```
filter_intron_table <- function(introns, clu){
    d < - introns %>% filter(clusterID == clu) %>%
        select(chr, start, end, verdict, deltapsi) %>%
        arrange(desc(abs(deltapsi))) %>%
        rename("ΔPSI" = deltapsi)
    row.names(d) <- letters[1:nrow(d)] # letters is just a:z
    return(d)
```

```
}


getGeneLength <- function(gene_name, exons_table){
    exons      <- exons_table[ exons_table$gene_name == gene_name, ]
    geneStart  <- min(exons$start)
    geneEnd    <- max(exons$end)
    geneLength <- geneEnd - geneStart
    if( geneLength >1e6){
        pixels <- 5000
    } else if ( geneLength > 5e5 & geneLength < 1e6){
        pixels <- 3000
    } else if ( geneLength > 1.5e5 & geneLength <= 5e5){
        pixels <- 2000
    } else {
        stopifnot(geneLength <= 1.5e5)
        pixels <- "auto"
    }
    return(pixels)
}


select_data <- function(sel, clusters, exons_table){
    gene <- clusters[ sel, ]$gene
    width <- getGeneLength(gene, exons_table)
    clusterID <- clusters[ sel, ]$clusterID
    coord <- clusters[ sel, ]$coord
    return(list(gene = gene, width = width, cluster = clusterID, coord = coord))
}
```

1.3.2 Plot top 6 clusters

[9]:
```
plot_cluster <- function(num, clusters, dir='./'){
    mydata = select_data(num, clusters, exons_table)
    if(mydata$gene == '.'){
        mydata$gene = gsub("-", "_", gsub(":", "_", mydata$coord))
    }
    gene_name = mydata$gene
    plotTitle <- paste0(gene_name, '_', mydata$cluster, '_top_',num,'.pdf')
    pdf(file=paste0(dir, plotTitle), width = 10, height = 5)
    print(make_cluster_plot(mydata$cluster,
                            main_title = plotTitle,
                            meta = meta,
                            cluster_ids = cluster_ids,
                            exons_table = exons_table,
                            counts = counts,
                            introns = introns))
```

```
    dev.off()

    if (is.numeric(mydata$width)) {
        new_width = mydata$width / 100
    } else {
        new_width = mydata$width
    }

    pdf(file=paste0(dir, gene_name, '_allClusters_top_',num,'.pdf'),
        width=new_width, height=6)
    try(print(make_gene_plot(mydata$gene,
                         counts = counts,
                         introns = introns,
                         exons_table = exons_table,
                         cluster_list = clusters,
                         clusterID = mydata$cluster,
                         introns_to_plot = introns_to_plot, debug=F)))
    dev.off()
}
```

1.4 Plot splicing

[10]:
```
dir.create("top10")
for(num in 1:10){
    ii = plot_cluster(num, clusters, "top10/")
}
```

```
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."
```

```
TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]
```

```
Warning message:
"'mode(width)' differs between new and previous
         ==> NOT changing 'width'"
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."
```

```
TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]
```

```
Warning message:
"'mode(width)' differs between new and previous
```

```
         ==> NOT changing ‘width’"
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."

TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]

Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."

TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]

Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."

TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]

Warning message:
"‘mode(width)’ differs between new and previous
         ==> NOT changing ‘width’"
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."

TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]

Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."

TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]

Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."
```

```
TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]
```

```
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."
```

```
TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]
```

```
Warning message:
"'mode(width)' differs between new and previous
         ==> NOT changing 'width'"
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."
```

```
TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]
```

[11]:
```
dir.create("x_chromosome")
x_clu = clusters %>% filter(str_detect(coord, "chrX"), FDR < 0.05)
x_clu
```

A data.frame: 22 × 6

clusterID <chr>	N <dbl>	coord <chr>	gene <chr>	annotation <chr>	FDR <dbl>
clu_755_-	3	chrX:53217966-53220839	KDM5C	annotated	2.27e-43
clu_744_-	10	chrX:53176622-53193437	KDM5C	cryptic	3.64e-40
clu_61504_+	3	chrX:47199106-47199480	UBA1	cryptic	2.45e-16
clu_62902_+	10	chrX:153765580-153769162	PLXNB3	cryptic	1.06e-09
clu_1810_-	3	chrX:152989331-152991916	PNMA5	cryptic	4.00e-09
clu_752_-	3	chrX:53210576-53211497	KDM5C	cryptic	8.61e-09
clu_62904_+	7	chrX:153771081-153771864	PLXNB3	cryptic	1.38e-07
clu_62916_+	5	chrX:153777688-153778396	PLXNB3	cryptic	1.14e-06
clu_62431_+	3	chrX:115636978-115640408	PLS3	annotated	7.28e-06
clu_261_-	11	chrX:15825984-15845379	AP1S2	cryptic	1.37e-04
clu_206_-	5	chrX:13785808-13938507	GPM6B	annotated	1.53e-04
clu_1841_-	9	chrX:154032557-154097604	MECP2	cryptic	9.14e-04
clu_1482_-	3	chrX:120428626-120441730	LAMP2	annotated	1.94e-03
clu_60959_+	9	chrX:2861722-2910874	GYG2	cryptic	8.68e-03
clu_62906_+	5	chrX:153772287-153773230	PLXNB3	cryptic	8.68e-03
clu_62231_+	4	chrX:103310981-103330984	TCEAL7	cryptic	1.14e-02
clu_480_-	3	chrX:41610025-41626604	CASK	annotated	1.23e-02
clu_12601_?	4	chrX:120428626-120441730	LAMP2	cryptic	1.70e-02
clu_62948_+	5	chrX:154399941-154400702	RPL10	cryptic	2.73e-02
clu_61084_+	6	chrX:13735220-13736656	OFD1	cryptic	3.09e-02
clu_62377_+	3	chrX:108584525-108586615	COL4A5	cryptic	4.60e-02
clu_61903_+	2	chrX:71167800-71171654	NLGN3	cryptic	4.93e-02

[12]:
```
for(num in 1:dim(x_clu)[1]){
    plot_cluster(num, x_clu, "x_chromosome/")
}
```

```
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."

TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]

Warning message:
"‘mode(width)’ differs between new and previous
         ==> NOT changing ‘width’"
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."

TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]
```

```
Warning message:
"‘mode(width)’ differs between new and previous
         ==> NOT changing ‘width’"
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."
```

```
TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]
```

```
Warning message:
"‘mode(width)’ differs between new and previous
         ==> NOT changing ‘width’"
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."
```

```
TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]
```

```
Warning message:
"‘mode(width)’ differs between new and previous
         ==> NOT changing ‘width’"
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."
Warning message:
"Removed 1 row(s) containing missing values (geom_path)."
Warning message:
"Removed 1 row(s) containing missing values (geom_path)."
```

```
TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]
```

```
Warning message:
"‘mode(width)’ differs between new and previous
         ==> NOT changing ‘width’"
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."
```

```
TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]
```

```
Warning message:
"‘mode(width)’ differs between new and previous
	 ==> NOT changing ‘width’"
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."

TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]

Warning message:
"‘mode(width)’ differs between new and previous
	 ==> NOT changing ‘width’"
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."

TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]

Warning message:
"‘mode(width)’ differs between new and previous
	 ==> NOT changing ‘width’"
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."

TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]

Warning message:
"‘mode(width)’ differs between new and previous
	 ==> NOT changing ‘width’"
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."

TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]

Warning message:
"‘mode(width)’ differs between new and previous
	 ==> NOT changing ‘width’"
Warning message:
```

```
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."

TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]

Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."

TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]

Warning message:
"‘mode(width)’ differs between new and previous
         ==> NOT changing ‘width’"
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."

TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]

Warning message:
"‘mode(width)’ differs between new and previous
         ==> NOT changing ‘width’"
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."

TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]

Warning message:
"‘mode(width)’ differs between new and previous
         ==> NOT changing ‘width’"
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."

TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]
```

```
Warning message:
"‘mode(width)’ differs between new and previous
         ==> NOT changing ‘width’"
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."
TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]
Warning message:
"‘mode(width)’ differs between new and previous
         ==> NOT changing ‘width’"
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."
TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."
TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]
Warning message:
"‘mode(width)’ differs between new and previous
         ==> NOT changing ‘width’"
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."
TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]
Warning message:
"‘mode(width)’ differs between new and previous
         ==> NOT changing ‘width’"
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."
```

```
TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]
```

```
Warning message:
"‘mode(width)’ differs between new and previous
         ==> NOT changing ‘width’"
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."
```

```
TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]
```

```
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."
Warning message:
"Removed 1 row(s) containing missing values (geom_path)."
Warning message:
"Removed 1 row(s) containing missing values (geom_path)."
```

```
TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]
```

```
Warning message:
"‘mode(width)’ differs between new and previous
         ==> NOT changing ‘width’"
```

```
Error : Aesthetics must be either length 1 or the same as the data (1): x, xend,
group and colour
```

[13]:
```
dir.create("ank3")
ank3 = clusters %>% filter(gene == 'ANK3')
ank3
```

A data.frame: 2 × 6

clusterID <chr>	N <dbl>	coord <chr>	gene <chr>	annotation <chr>	FDR <dbl>
clu_22299_-	3	chr10:60059430-60063111	ANK3	annotated	0.00159
clu_22304_-	3	chr10:60139087-60166824	ANK3	annotated	0.00895

[14]:
```
for(num in 1:dim(ank3)[1]){
    plot_cluster(num, ank3, "ank3/")
}
```

```
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
```

```
instead."
```

```
TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]
```

```
Warning message:
"`guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")`
instead."
```

```
TableGrob (2 x 1) "arrange": 2 grobs
  z     cells    name           grob
1 1 (1-1,1-1) arrange gtable[layout]
2 2 (2-2,1-1) arrange gtable[layout]
```

[ ]: