

main

August 20, 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
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]: sample_table
```

	group <chr>	count <int>
A data.frame: 2 × 2	AA	47
	EA	43

```
[4]: cluster_summary
```

	Results <chr>	n <int>
A data.frame: 3 × 2	Number of differentially spliced clusters at FDR = 0.05	655
	Fully annotated	208
	Contain unannotated junctions	447

```
[5]: intron_summary
```

	Results <chr>	n <int>
A data.frame: 5 × 2	Number of fully annotated junctions	1850
	Number of junctions with cryptic 5' splice site	430
	Number of junctions with cryptic 3' splice site	479
	Number of junctions with two cryptic splice sites	296
	Number of novel junctions that connect two annotated splice sites	336

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

	clusterID <chr>	N <dbl>	coord <chr>	gene <chr>	annotation <chr>	F <chr>
A data.frame: 6 × 6	clu_71861_+	6	chr3:129488397-129499902	IFT122	cryptic	9.
	clu_97155_-	12	chr12:124911899-124913725	UBC	cryptic	2.
	clu_56723_+	19	chr22:45323427-45332425	FAM118A	cryptic	1.
	clu_97156_-	24	chr12:124911952-124913549	UBC	cryptic	7.
	clu_87410_+	10	chr6:26440648-26443957	BTN3A3	cryptic	8.
	clu_112823_+	18	chr2:39518710-39665202	AC007246.3	cryptic	1.

```
[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 Plotting functions

```

[9]: plot_cluster <- function(num, clusters, dir='./'){
  mydata = select_data(num, clusters, exons_table)
  while(mydata$gene == '.'){
    num = num+1
    mydata = select_data(num, clusters, exons_table)
  }
  gene_name = mydata$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)
print(make_gene_plot(mydata$gene,
                     counts = counts,
                     introns = introns,
                     exons_table = exons_table,
                     cluster_list = clusters,
                     clusterID = mydata$clusterID,
                     introns_to_plot = introns_to_plot, debug=F))

dev.off()
}

```

1.3.3 Plot splicing

```

[10]: dir.create("top10")
      for(num in 1:10){
        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:
"ggrepel: 22 unlabeled data points (too many overlaps). Consider increasing
max.overlaps"
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:
"ggrepel: 22 unlabeled data points (too many overlaps). Consider increasing
max.overlaps"
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:
"ggrepel: 178 unlabeled data points (too many overlaps). Consider increasing
max.overlaps"

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’"
```

```
[ ]:
```