ABCB1_distalTFs_scan

charvi khanna

2025-11-03

```r
# load the required libraries/packages
library(TFBSTools)
library(Biostrings)
```

```
## Loading required package: BiocGenerics


## Loading required package: generics


##
## Attaching package: 'generics'


## The following objects are masked from 'package:base':
##
##      as.difftime, as.factor, as.ordered, intersect, is.element, setdiff,
##      setequal, union


##
## Attaching package: 'BiocGenerics'


## The following objects are masked from 'package:stats':
##
##      IQR, mad, sd, var, xtabs


## The following objects are masked from 'package:base':
##
##      anyDuplicated, aperm, append, as.data.frame, basename, cbind,
##      colnames, dirname, do.call, duplicated, eval, evalq, Filter, Find,
##      get, grep, grepl, is.unsorted, lapply, Map, mapply, match, mget,
##      order, paste, pmax, pmax.int, pmin, pmin.int, Position, rank,
##      rbind, Reduce, rownames, sapply, saveRDS, table, tapply, unique,
##      unsplit, which.max, which.min


## Loading required package: S4Vectors


## Loading required package: stats4


##
## Attaching package: 'S4Vectors'
```

```
## The following object is masked from 'package:utils':
##
##     findMatches


## The following objects are masked from 'package:base':
##
##     expand.grid, I, unname


## Loading required package: IRanges


##
## Attaching package: 'IRanges'


## The following object is masked from 'package:grDevices':
##
##     windows


## Loading required package: XVector


## Loading required package: GenomeInfoDb


##
## Attaching package: 'Biostrings'


## The following object is masked from 'package:base':
##
##     strsplit
```

```r
library(JASPAR2022)
```

```
## Loading required package: BiocFileCache


## Loading required package: dbplyr
```

```r
library(dplyr)
```

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


## The following objects are masked from 'package:dbplyr':
##
##     ident, sql


## The following objects are masked from 'package:Biostrings':
##
##     collapse, intersect, setdiff, setequal, union


## The following object is masked from 'package:GenomeInfoDb':
##
##     intersect
```

```
## The following object is masked from 'package:XVector':
##
##      slice


## The following objects are masked from 'package:IRanges':
##
##      collapse, desc, intersect, setdiff, slice, union


## The following objects are masked from 'package:S4Vectors':
##
##      first, intersect, rename, setdiff, setequal, union


## The following objects are masked from 'package:BiocGenerics':
##
##      combine, intersect, setdiff, setequal, union


## The following object is masked from 'package:generics':
##
##      explain


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


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

```r
library(writexl)
library(knitr)
library(kableExtra)
```

```
## Warning: package 'kableExtra' was built under R version 4.5.2


##
## Attaching package: 'kableExtra'


## The following object is masked from 'package:dplyr':
##
##      group_rows
```

```r
# Using biostrings package store DNA sequence to be scanned across for TFs
ABCB1_distal <- DNAString("AAATTAGCTTTATGGAAATTCATTTGTTTACTTTGCAATTATATCAGTATTTAATTTATAATGAAAAAGTTGTGTT

# Get human motifs
opts <- list(species = 9606)
motifs <- getMatrixSet(JASPAR2022, opts)

# Convert motifs to PWM using log2 probability ratios
pwm <- toPWM(motifs, type = "log2probratio")
```

```r
# Applying searchseq command from TFBSTools to scan stored DNA sequence
hitsdist <- searchSeq(pwm, ABCB1_distal, seqname = "ABCB1_distal", min.score = "95%", strand = "*")
hits_filtered_distal <- hitsdist[sapply(hitsdist, length) > 0]
hits_df_distal <- do.call(rbind, lapply(hits_filtered_distal,
                                        as.data.frame))
hits_df_distal$TF <- rep(names(hits_filtered_distal),
                         sapply(hits_filtered_distal, length))
# View result in table format using kableExtra
kable(head(hits_df_distal, 42), booktabs = TRUE, longtable = TRUE) %>%
kable_styling(latex_options = c("striped", "scale_down"))
```

```
## Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be
## resized.
```

	seqnames	source	feature	start	end	absScore	relScore	strand	ID	TF
MA0030.1	ABCB1_distal	TFBS	TFBS	23	36	15.962982	0.9508261	-	MA0030.1	MA0030
MA0031.1	ABCB1_distal	TFBS	TFBS	24	31	13.567858	0.9958595	-	MA0031.1	MA0031
MA0084.1	ABCB1_distal	TFBS	TFBS	23	31	11.350859	0.9696310	-	MA0084.1	MA0084
MA0091.1	ABCB1_distal	TFBS	TFBS	958	969	13.931702	0.9509646	+	MA0091.1	MA0091
MA0130.1	ABCB1_distal	TFBS	TFBS	925	930	9.675124	0.9940480	+	MA0130.1	MA0130
MA0259.1.1	ABCB1_distal	TFBS	TFBS	763	770	10.298448	0.9721358	+	MA0259.1	MA0259
MA0259.1.2	ABCB1_distal	TFBS	TFBS	1012	1019	11.790654	1.0000000	-	MA0259.1	MA0259
MA0468.1.1	ABCB1_distal	TFBS	TFBS	1385	1395	10.486880	0.9545984	-	MA0468.1	MA0468
MA0468.1.2	ABCB1_distal	TFBS	TFBS	1409	1419	10.486880	0.9545984	-	MA0468.1	MA0468
MA0468.1.3	ABCB1_distal	TFBS	TFBS	1433	1443	10.486880	0.9545984	-	MA0468.1	MA0468
MA0468.1.4	ABCB1_distal	TFBS	TFBS	1457	1467	10.486880	0.9545984	-	MA0468.1	MA0468
MA0468.1.5	ABCB1_distal	TFBS	TFBS	1481	1491	10.486880	0.9545984	-	MA0468.1	MA0468
MA0144.2	ABCB1_distal	TFBS	TFBS	796	806	11.913090	0.9659668	-	MA0144.2	MA0144
MA0593.1	ABCB1_distal	TFBS	TFBS	23	33	16.559326	1.0000000	-	MA0593.1	MA0593
MA0595.1	ABCB1_distal	TFBS	TFBS	515	524	13.417160	0.9644551	-	MA0595.1	MA0595
MA0599.1	ABCB1_distal	TFBS	TFBS	896	905	13.639327	0.9867113	+	MA0599.1	MA0599
MA0613.1	ABCB1_distal	TFBS	TFBS	24	31	14.280506	1.0000000	-	MA0613.1	MA0613
MA0618.1.1	ABCB1_distal	TFBS	TFBS	1127	1134	12.068281	0.9855810	+	MA0618.1	MA0618
MA0618.1.2	ABCB1_distal	TFBS	TFBS	1384	1391	9.402401	0.9518206	-	MA0618.1	MA0618
MA0618.1.3	ABCB1_distal	TFBS	TFBS	1408	1415	9.402401	0.9518206	-	MA0618.1	MA0618
MA0618.1.4	ABCB1_distal	TFBS	TFBS	1432	1439	9.402401	0.9518206	-	MA0618.1	MA0618
MA0618.1.5	ABCB1_distal	TFBS	TFBS	1456	1463	9.402401	0.9518206	-	MA0618.1	MA0618
MA0618.1.6	ABCB1_distal	TFBS	TFBS	1480	1487	9.402401	0.9518206	-	MA0618.1	MA0618
MA0630.1.1	ABCB1_distal	TFBS	TFBS	1127	1134	9.968525	0.9710462	+	MA0630.1	MA0630
MA0630.1.2	ABCB1_distal	TFBS	TFBS	1384	1391	9.636288	0.9616395	+	MA0630.1	MA0630
MA0630.1.3	ABCB1_distal	TFBS	TFBS	1408	1415	9.636288	0.9616395	+	MA0630.1	MA0630
MA0630.1.4	ABCB1_distal	TFBS	TFBS	1432	1439	9.636288	0.9616395	+	MA0630.1	MA0630
MA0630.1.5	ABCB1_distal	TFBS	TFBS	1456	1463	9.636288	0.9616395	+	MA0630.1	MA0630
MA0630.1.6	ABCB1_distal	TFBS	TFBS	1480	1487	9.636288	0.9616395	+	MA0630.1	MA0630
MA0630.1.7	ABCB1_distal	TFBS	TFBS	201	208	9.519959	0.9583458	-	MA0630.1	MA0630
MA0630.1.8	ABCB1_distal	TFBS	TFBS	1384	1391	9.454322	0.9564874	-	MA0630.1	MA0630
MA0630.1.9	ABCB1_distal	TFBS	TFBS	1408	1415	9.454322	0.9564874	-	MA0630.1	MA0630
MA0630.1.10	ABCB1_distal	TFBS	TFBS	1432	1439	9.454322	0.9564874	-	MA0630.1	MA0630
MA0630.1.11	ABCB1_distal	TFBS	TFBS	1456	1463	9.454322	0.9564874	-	MA0630.1	MA0630

MA0630.1.12	ABCB1_distal	TFBS	TFBS	1480	1487	9.454322	0.9564874	-	MA0630.1	MA0630
MA0635.1.1	ABCB1_distal	TFBS	TFBS	450	459	10.026987	0.9691135	-	MA0635.1	MA0635
MA0635.1.2	ABCB1_distal	TFBS	TFBS	857	866	8.699951	0.9532006	-	MA0635.1	MA0635
MA0635.1.3	ABCB1_distal	TFBS	TFBS	911	920	9.395098	0.9615364	-	MA0635.1	MA0635
MA0027.2.1	ABCB1_distal	TFBS	TFBS	35	42	9.158663	0.9606159	+	MA0027.2	MA0027
MA0027.2.2	ABCB1_distal	TFBS	TFBS	201	208	9.792497	0.9693421	+	MA0027.2	MA0027
MA0027.2.3	ABCB1_distal	TFBS	TFBS	1127	1134	10.560649	0.9799176	+	MA0027.2	MA0027
MA0027.2.4	ABCB1_distal	TFBS	TFBS	1369	1376	8.551844	0.9522616	+	MA0027.2	MA0027

```r
View(hits_filtered_distal)
```