Ribo-seq quality check by riboWaltz

Toshihiro Arae

General directory setting

```
wd <- here::here()
shared <- fs::path(fs::path_dir(wd), "shared")
```

Loading packages

```
library(magrittr)
library(ggplot2)
```

Load common R scripts

```
#source(fs::path(wd, "script_r", "MISC.R"))
#source(fs::path(here::here(), "script_r", "MISC_PALETTE.R"))
```

Load custom scripts

```
# Create annotation data table
source(fs::path(wd, "script_r", "create_annotation_mod.R"))
```

```
# Original code are from riboWaltz@v1.1.0 and slightly modified.
# Instead of specifying the package name in the `txdb` argument,
# we changed it so that we can pass our own txdb object directly to the `txdbanno` argument.
create_annotation_mod <- function(txdbanno = NULL) {
  exon <-
    suppressWarnings(GenomicFeatures::exonsBy(txdbanno, by = "tx", use.names = T))
  utr5 <-
    suppressWarnings(GenomicFeatures::fiveUTRsByTranscript(txdbanno, use.names = T))
  cds <-
    suppressWarnings(GenomicFeatures::cdsBy(txdbanno, by = "tx", use.names = T))
  utr3 <-
    suppressWarnings(GenomicFeatures::threeUTRsByTranscript(txdbanno, use.names = T))
  exon <- data.table::as.data.table(exon[unique(names(exon))])
  utr5 <- data.table::as.data.table(utr5[unique(names(utr5))])
  cds <- data.table::as.data.table(cds[unique(names(cds))])
  utr3 <- data.table::as.data.table(utr3[unique(names(utr3))])
  anno_df <- exon[, list(l_tr = sum(width)), by = list(transcript = group_name)]
  l_utr5 <- utr5[, list(l_utr5 = sum(width)), by = list(transcript = group_name)]
  l_cds <- cds[, list(l_cds = sum(width)), by = list(transcript = group_name)]
  l_utr3 <- utr3[, list(l_utr3 = sum(width)), by = list(transcript = group_name)]
  merge_allx <- function(x, y) merge(x, y, all.x = TRUE)
  anno_df <- BiocGenerics::Reduce(merge_allx, list(anno_df, l_utr5, l_cds, l_utr3))
  anno_df[is.na(anno_df)] <- 0
  return(anno_df)
}
```

Import genomic sequences and annotation data

```

```

```
bsg_tair <- BSgenome::getBSgenome("BSgenome.Athaliana.TAIR.TAIR9")
##
## 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, append, as.data.frame, basename, cbind, colnames,
##     dirname, do.call, duplicated, eval, evalq, Filter, Find, get, grep,
##     grepl, intersect, is.unsorted, lapply, Map, mapply, match, mget,
##     order, paste, pmax, pmax.int, pmin, pmin.int, Position, rank,
##     rbind, Reduce, rownames, sapply, setdiff, sort, table, tapply,
##     union, unique, unsplit, which.max, which.min
##
## Attaching package: 'S4Vectors'
## The following objects are masked from 'package:base':
##
##     expand.grid, I, unname
##
## Attaching package: 'GenomicRanges'
## The following object is masked from 'package:magrittr':
##
##     subtract
##
## Attaching package: 'Biostrings'
## The following object is masked from 'package:base':
##
##     strsplit
path_gff <- fs::path(wd, "misc", "gff_gtf",
                     "Araport11_GFF3_genes_transposons.201606_mod.gff")
txdb_araport <-
  GenomicFeatures::makeTxDbFromGFF(
    file = path_gff,
    format = "gff3",
    dataSource = "Araport11",
    organism = "Arabidopsis thaliana",
    circ_seqs = c("ChrC", "ChrM"),
    chrominfo = GenomeInfoDb::seqinfo(bsg_tair)
  )
## Import genomic features from the file as a GRanges object ...
## OK
## Prepare the 'metadata' data frame ... OK
## Make the TxDb object ...
## Warning in .get_cds_IDX(mcols0$type, mcols0$phase): The "phase" metadata column contains non-NA
values for features of type
##   exon. This information was ignored.
## OK

annot_dt <- create_annotation_mod(txdb = txdb_araport)
```

Run the riboWaltz pipeline

```
dir_output <- fs::path("analysis", "out_ribowaltz")
path_out <- function(...) fs::path(wd, dir_output, ...)
fs::dir_create(path_out())

sample_labels <-
  c(
    "zt0_1_ribo", "zt0_2_ribo",
    "zt3_1_ribo", "zt3_2_ribo",
```

```
    "zt6_1_ribo", "zt6_2_ribo",
    "zt12_1_ribo", "zt12_2_ribo",
    "zt18_1_ribo", "zt18_2_ribo",
    "zt21_1_ribo", "zt21_2_ribo"
  ) %>%
  paste0(".tr.sort")
```

```
for(bam in sample_labels){
  names(bam) <- bam
  label <- stringr::str_remove(bam, ".tr.sort")

  reads_list <-
    riboWaltz::bamtolist(
      bamfolder = fs::path(wd, "data_preproc", "bam_transcriptome"),
      annotation = annot_dt,
      name_samples = bam
    )

  # read ends heatmaps
  path_out_ <- function(...) path_out("rends_heat", ...)
  fs::dir_create(path_out_())

  gp <- riboWaltz::rends_heat(reads_list, annot_dt, sample = bam)
  ggplot2::ggsave(plot = gp$plot, filename = path_out_(paste0(label, ".png")),
                  width = 11, height = 6)
  readr::write_csv(x = tibble::as_tibble(gp$dt), path_out_(paste0(label, ".csv")))

  # p-site prediction
  path_out_ <- function(...) path_out("psite_predict", ...)
  fs::dir_create(path_out_())

  psite_offset_start <-
    riboWaltz::psite(reads_list, start = TRUE, plot_dir = path_out_(), plot = TRUE)
  readr::write_csv(
    x = tibble::as_tibble(psite_offset_start),
    file = path_out_(paste0(label, ".csv"))
  )

  reads_psite_list <- riboWaltz::psite_info(reads_list, psite_offset_start)
  rm(reads_list, gp)
  gc();gc()

  # p-site region
  path_out_ <- function(...) path_out("region_psite", ...)
  fs::dir_create(path_out_())

  gp <- riboWaltz::region_psite(reads_psite_list, annot_dt, sample = bam)
  ggsave(plot = gp$plot, filename = path_out_(paste0(label, ".png")),
         width = 8, height = 6)
  readr::write_csv(x = tibble::as_tibble(gp$dt), path_out_(paste0(label, ".csv")))

  # frame p-site length
  path_out_ <- function(...) path_out("frame_psite_len", ...)
  fs::dir_create(path_out_())

  gp <-
    riboWaltz::frame_psite_length(
      data = reads_psite_list,
      sample = bam,
      region = "all", cl = 90
    )
```

```
  ggsave(plot = gp$plot, filename = path_out_(paste0(label, ".png")),
         width = 10, height = 5)
  readr::write_csv(x = tibble::as_tibble(gp$dt), path_out_(paste0(label, ".csv")))

  # frame p-site
  path_out_ <- function(...) path_out("frame_psite", ...)
  fs::dir_create(path_out_())

  gp <- riboWaltz::frame_psite(data = reads_psite_list, sample = bam,
                               region = "all")
  gp$plot <- gp$plot + scale_y_continuous(limits = c(0, 50))
  ggsave(plot = gp$plot, filename = path_out_(paste0(label, ".png")),
         width = 10, height = 4)
  readr::write_csv(x = tibble::as_tibble(gp$dt), path_out_(paste0(label, ".csv")))

  # metaprofile p-site
  path_out_ <- function(...) path_out("metaplof_psite", ...)
  fs::dir_create(path_out_())

  gp <-
    riboWaltz::metaprofile_psite(
      data = reads_psite_list,
      annotation = annot_dt,
      sample = bam, plot_title = label,
      utr5l = 20, cdsl = 40, utr3l = 20
    )
  ggsave(plot = gp$plot, filename = path_out_(paste0(label, ".png")),
         width = 11, height = 6)
  readr::write_csv(x = tibble::as_tibble(gp$dt), path_out_(paste0(label, "_all.csv")))

  for(nt in 32:35) {
    gp <-
      riboWaltz::metaprofile_psite(
        data = reads_psite_list,
        annotation = annot_dt,
        length_range = nt,
        sample = bam, plot_title = label,
        utr5l = 20, cdsl = 40, utr3l = 20
      )
    ggsave(plot = gp$plot, filename = path_out_(paste0(label, "_", nt, ".png")),
           width = 11, height = 6)
    readr::write_csv(x = tibble::as_tibble(gp$dt), path_out_(paste0(label, "_", nt, ".csv")))
  }

  rm(reads_psite_list)
  gc();gc()
}
Reading zt0_1_ribo.tr.sort.bam
Input reads: 58.905 M
44.000 (< 0.001 %) reads removed: exceeding indel_threshold.
Great! All reads' reference transcript IDs were found in the annotation table. No reads removed.
1.112 M  (1.887 %) reads removed: mapping on negative strand.
Output reads: 57.793 M
Done! zt0_1_ribo.tr.sort.bam has been loaded as zt0_1_ribo.tr.sort
Warning: Removed 6666 rows containing missing values (`geom_tile()`).
processing zt0_1_ribo.tr.sort
best offset: 13 nts from the 5' end

plotting                          <<-- 8% -->>

plotting                         <<---- 16% ---->>
```

```
plotting                       <<----- 20% ----->>

plotting                      <<------- 28% ------->>

plotting                     <<--------- 32% --------->>

plotting                   <<---------- 40% ---------->>

plotting                  <<----------- 44% ----------->>

plotting                <<------------- 52% ------------->>

plotting              <<--------------- 60% --------------->>

plotting            <<----------------- 64% ----------------->>

plotting          <<------------------- 72% ------------------->>

plotting         <<--------------------- 76% --------------------->>

plotting       <<----------------------- 84% ----------------------->>

plotting      <<------------------------ 88% ------------------------>>

plotting    <<-------------------------- 96% -------------------------->>

plotting   <<--------------------------- 100% --------------------------->>

plotting   <<--------------------------- 100% --------------------------->>
processing zt0_1_ribo.tr.sort
1. adding p-site position
2. adding transcript region
Reading zt0_2_ribo.tr.sort.bam
Input reads: 43.505 M
32.000 (< 0.001 %) reads removed: exceeding indel_threshold.
Great! All reads' reference transcript IDs were found in the annotation table. No reads removed.
1.026 M  (2.357 %) reads removed: mapping on negative strand.
Output reads: 42.479 M
Done! zt0_2_ribo.tr.sort.bam has been loaded as zt0_2_ribo.tr.sort
Warning: Removed 6565 rows containing missing values (`geom_tile()`).
processing zt0_2_ribo.tr.sort
best offset: 14 nts from the 5' end

plotting                           <<-- 8% -->>

plotting                        <<---- 16% ---->>

plotting                       <<----- 20% ----->>

plotting                      <<------- 28% ------->>

plotting                     <<--------- 32% --------->>

plotting                   <<---------- 40% ---------->>

plotting                  <<----------- 44% ----------->>

plotting                <<------------- 52% ------------->>

plotting              <<--------------- 60% --------------->>

plotting            <<----------------- 64% ----------------->>
```

```
plotting          <<------------------- 72% ------------------->>

plotting         <<-------------------- 76% -------------------->>

plotting       <<---------------------- 84% ---------------------->>

plotting      <<------------------------ 88% ------------------------>>

plotting    <<-------------------------- 96% -------------------------->>

plotting   <<--------------------------- 100% --------------------------->>

plotting   <<--------------------------- 100% --------------------------->>
processing zt0_2_ribo.tr.sort
1. adding p-site position
2. adding transcript region
Reading zt3_1_ribo.tr.sort.bam
Input reads: 53.506 M
40.000 (< 0.001 %) reads removed: exceeding indel_threshold.
Great! All reads' reference transcript IDs were found in the annotation table. No reads removed.
1.160 M  (2.168 %) reads removed: mapping on negative strand.
Output reads: 52.346 M
Done! zt3_1_ribo.tr.sort.bam has been loaded as zt3_1_ribo.tr.sort
Warning: Removed 7171 rows containing missing values (`geom_tile()`).
processing zt3_1_ribo.tr.sort
best offset: 13 nts from the 5' end

plotting                           <<-- 8% -->>

plotting                          <<--- 12% --->>

plotting                        <<----- 20% ----->>

plotting                       <<------ 24% ------>>

plotting                     <<-------- 32% -------->>

plotting                    <<--------- 36% --------->>

plotting                 <<------------ 44% ------------>>

plotting                <<------------- 48% ------------->>

plotting              <<--------------- 56% --------------->>

plotting             <<---------------- 60% ---------------->>

plotting           <<------------------ 68% ------------------>>

plotting          <<------------------- 72% ------------------->>

plotting        <<---------------------- 80% ---------------------->>

plotting       <<----------------------- 84% ----------------------->>

plotting     <<------------------------- 92% ------------------------->>

plotting    <<-------------------------- 96% -------------------------->>

plotting   <<--------------------------- 100% --------------------------->>
```

```
plotting   <<------------------------- 100% ------------------------->>
processing zt3_1_ribo.tr.sort
1. adding p-site position
2. adding transcript region
Reading zt3_2_ribo.tr.sort.bam
Input reads: 39.231 M
17.000 (< 0.001 %) reads removed: exceeding indel_threshold.
Great! All reads' reference transcript IDs were found in the annotation table. No reads removed.
1.046 M  (2.667 %) reads removed: mapping on negative strand.
Output reads: 38.185 M
Done! zt3_2_ribo.tr.sort.bam has been loaded as zt3_2_ribo.tr.sort
Warning: Removed 6161 rows containing missing values (`geom_tile()`).
processing zt3_2_ribo.tr.sort
best offset: 14 nts from the 5' end

plotting                          <<-- 8% -->>

plotting                        <<---- 16% ---->>

plotting                       <<----- 20% ----->>

plotting                     <<------- 28% ------->>

plotting                    <<-------- 32% -------->>

plotting                 <<---------- 40% ---------->>

plotting                <<----------- 44% ----------->>

plotting              <<------------- 52% ------------->>

plotting            <<--------------- 60% --------------->>

plotting           <<---------------- 64% ---------------->>

plotting         <<------------------ 72% ------------------>>

plotting        <<------------------- 76% ------------------->>

plotting      <<---------------------- 84% ---------------------->>

plotting     <<----------------------- 88% ----------------------->>

plotting   <<------------------------- 96% ------------------------->>

plotting   <<------------------------- 100% ------------------------->>

plotting   <<------------------------- 100% ------------------------->>
processing zt3_2_ribo.tr.sort
1. adding p-site position
2. adding transcript region
Reading zt6_1_ribo.tr.sort.bam
Input reads: 37.576 M
78.000 (< 0.001 %) reads removed: exceeding indel_threshold.
Great! All reads' reference transcript IDs were found in the annotation table. No reads removed.
1.024 M  (2.726 %) reads removed: mapping on negative strand.
Output reads: 36.551 M
Done! zt6_1_ribo.tr.sort.bam has been loaded as zt6_1_ribo.tr.sort
Warning: Removed 5252 rows containing missing values (`geom_tile()`).
processing zt6_1_ribo.tr.sort
best offset: 13 nts from the 5' end
```

```
plotting                          <<-- 8% -->>

plotting                         <<---- 16% ---->>

plotting                        <<------ 20% ------>>

plotting                      <<-------- 28% -------->>

plotting                     <<--------- 32% --------->>

plotting                   <<----------- 40% ----------->>

plotting                  <<------------ 44% ------------>>

plotting                <<-------------- 52% -------------->>

plotting              <<---------------- 60% ---------------->>

plotting             <<----------------- 64% ----------------->>

plotting           <<------------------- 72% ------------------->>

plotting          <<-------------------- 76% -------------------->>

plotting        <<---------------------- 84% ---------------------->>

plotting       <<----------------------- 88% ----------------------->>

plotting     <<------------------------- 96% ------------------------->>

plotting    <<-------------------------- 100% -------------------------->>

plotting    <<-------------------------- 100% -------------------------->>
processing zt6_1_ribo.tr.sort
1. adding p-site position
2. adding transcript region
Reading zt6_2_ribo.tr.sort.bam
Input reads: 38.352 M
37.000 (< 0.001 %) reads removed: exceeding indel_threshold.
Great! All reads' reference transcript IDs were found in the annotation table. No reads removed.
0.923 M  (2.407 %) reads removed: mapping on negative strand.
Output reads: 37.429 M
Done! zt6_2_ribo.tr.sort.bam has been loaded as zt6_2_ribo.tr.sort
Warning: Removed 6060 rows containing missing values (`geom_tile()`).
processing zt6_2_ribo.tr.sort
best offset: 14 nts from the 5' end

plotting                          <<-- 8% -->>

plotting                         <<---- 16% ---->>

plotting                        <<------ 20% ------>>

plotting                      <<-------- 28% -------->>

plotting                     <<--------- 32% --------->>

plotting                   <<----------- 40% ----------->>

plotting                  <<------------ 44% ------------>>

plotting                <<-------------- 52% -------------->>
```

```
plotting             <<---------------- 60% ---------------->>

plotting            <<----------------- 64% ----------------->>

plotting          <<------------------- 72% ------------------>>

plotting         <<--------------------- 76% -------------------->>

plotting       <<----------------------- 84% ----------------------->>

plotting      <<------------------------ 88% ------------------------>>

plotting    <<-------------------------- 96% -------------------------->>

plotting   <<--------------------------- 100% --------------------------->>

plotting   <<--------------------------- 100% --------------------------->>
processing zt6_2_ribo.tr.sort
1. adding p-site position
2. adding transcript region
Reading zt12_1_ribo.tr.sort.bam
Input reads: 49.633 M
58.000 (< 0.001 %) reads removed: exceeding indel_threshold.
Great! All reads' reference transcript IDs were found in the annotation table. No reads removed.
0.574 M  (1.156 %) reads removed: mapping on negative strand.
Output reads: 49.059 M
Done! zt12_1_ribo.tr.sort.bam has been loaded as zt12_1_ribo.tr.sort
Warning: Removed 9292 rows containing missing values (`geom_tile()`).
processing zt12_1_ribo.tr.sort
best offset: 13 nts from the 5' end

plotting                          <<-- 8% -->>

plotting                        <<---- 16% ---->>

plotting                      <<------ 24% ------>>

plotting                    <<-------- 32% -------->>

plotting                  <<---------- 40% ---------->>

plotting                <<------------ 48% ------------>>

plotting              <<-------------- 56% -------------->>

plotting            <<---------------- 64% ---------------->>

plotting          <<------------------ 72% ------------------>>

plotting        <<--------------------- 80% --------------------->>

plotting      <<----------------------- 88% ----------------------->>

plotting    <<-------------------------- 96% -------------------------->>

plotting   <<--------------------------- 100% --------------------------->>

plotting   <<--------------------------- 100% --------------------------->>
processing zt12_1_ribo.tr.sort
1. adding p-site position
2. adding transcript region
```

```
Reading zt12_2_ribo.tr.sort.bam
Input reads: 39.851 M
78.000 (< 0.001 %) reads removed: exceeding indel_threshold.
Great! All reads' reference transcript IDs were found in the annotation table. No reads removed.
1.266 M  (3.177 %) reads removed: mapping on negative strand.
Output reads: 38.585 M
Done! zt12_2_ribo.tr.sort.bam has been loaded as zt12_2_ribo.tr.sort
Warning: Removed 4444 rows containing missing values (`geom_tile()`).
processing zt12_2_ribo.tr.sort
best offset: 13 nts from the 5' end

plotting                          <<-- 8% -->>

plotting                        <<---- 16% ---->>

plotting                       <<------ 20% ------>>

plotting                     <<-------- 28% -------->>

plotting                    <<--------- 32% --------->>

plotting                  <<----------- 40% ----------->>

plotting                 <<------------ 44% ------------>>

plotting              <<-------------- 52% -------------->>

plotting            <<---------------- 60% ---------------->>

plotting           <<----------------- 64% ----------------->>

plotting         <<------------------- 72% ------------------->>

plotting        <<-------------------- 76% -------------------->>

plotting      <<---------------------- 84% ---------------------->>

plotting     <<----------------------- 88% ----------------------->>

plotting   <<------------------------- 96% ------------------------->>

plotting  <<-------------------------- 100% -------------------------->>

plotting  <<-------------------------- 100% -------------------------->>
processing zt12_2_ribo.tr.sort
1. adding p-site position
2. adding transcript region
Reading zt18_1_ribo.tr.sort.bam
Input reads: 47.327 M
40.000 (< 0.001 %) reads removed: exceeding indel_threshold.
Great! All reads' reference transcript IDs were found in the annotation table. No reads removed.
0.688 M  (1.453 %) reads removed: mapping on negative strand.
Output reads: 46.639 M
Done! zt18_1_ribo.tr.sort.bam has been loaded as zt18_1_ribo.tr.sort
Warning: Removed 7070 rows containing missing values (`geom_tile()`).
processing zt18_1_ribo.tr.sort
best offset: 13 nts from the 5' end

plotting                          <<-- 8% -->>

plotting                        <<---- 16% ---->>
```

```
plotting                       <<----- 20% ----->>

plotting                      <<------- 28% ------->>

plotting                     <<-------- 32% -------->>

plotting                   <<---------- 40% ---------->>

plotting                  <<----------- 44% ----------->>

plotting                <<------------- 52% ------------->>

plotting              <<--------------- 60% --------------->>

plotting             <<---------------- 64% ---------------->>

plotting           <<------------------ 72% ------------------>>

plotting          <<------------------- 76% ------------------->>

plotting        <<---------------------- 84% ---------------------->>

plotting       <<----------------------- 88% ----------------------->>

plotting     <<------------------------- 96% ------------------------->>

plotting    <<-------------------------- 100% -------------------------->>

plotting    <<-------------------------- 100% -------------------------->>
processing zt18_1_ribo.tr.sort
1. adding p-site position
2. adding transcript region
Reading zt18_2_ribo.tr.sort.bam
Input reads: 45.696 M
28.000 (< 0.001 %) reads removed: exceeding indel_threshold.
Great! All reads' reference transcript IDs were found in the annotation table. No reads removed.
0.644 M  (1.409 %) reads removed: mapping on negative strand.
Output reads: 45.052 M
Done! zt18_2_ribo.tr.sort.bam has been loaded as zt18_2_ribo.tr.sort
Warning: Removed 7474 rows containing missing values (`geom_tile()`).
processing zt18_2_ribo.tr.sort
best offset: 14 nts from the 5' end

plotting                           <<-- 8% -->>

plotting                        <<---- 16% ---->>

plotting                       <<----- 20% ----->>

plotting                      <<------- 28% ------->>

plotting                    <<--------- 36% --------->>

plotting                   <<---------- 40% ---------->>

plotting                 <<------------ 48% ------------>>

plotting               <<-------------- 56% -------------->>

plotting              <<---------------- 60% ---------------->>

plotting            <<------------------ 68% ------------------>>
```

```
plotting         <<-------------------- 76% -------------------->>

plotting        <<--------------------- 80% --------------------->>

plotting      <<----------------------- 88% ----------------------->>

plotting    <<------------------------- 96% ------------------------->>

plotting   <<-------------------------- 100% -------------------------->>

plotting   <<-------------------------- 100% -------------------------->>
processing zt18_2_ribo.tr.sort
1. adding p-site position
2. adding transcript region
Reading zt21_1_ribo.tr.sort.bam
Input reads: 41.954 M
48.000 (< 0.001 %) reads removed: exceeding indel_threshold.
Great! All reads' reference transcript IDs were found in the annotation table. No reads removed.
0.710 M  (1.691 %) reads removed: mapping on negative strand.
Output reads: 41.245 M
Done! zt21_1_ribo.tr.sort.bam has been loaded as zt21_1_ribo.tr.sort
Warning: Removed 7070 rows containing missing values (`geom_tile()`).
processing zt21_1_ribo.tr.sort
best offset: 13 nts from the 5' end

plotting                          <<-- 8% -->>

plotting                        <<---- 16% ---->>

plotting                       <<----- 20% ----->>

plotting                     <<------- 28% ------->>

plotting                   <<--------- 36% --------->>

plotting                  <<---------- 40% ---------->>

plotting                <<------------ 48% ------------>>

plotting              <<-------------- 56% -------------->>

plotting             <<--------------- 60% --------------->>

plotting           <<----------------- 68% ----------------->>

plotting         <<------------------- 76% ------------------->>

plotting        <<--------------------- 80% --------------------->>

plotting      <<----------------------- 88% ----------------------->>

plotting    <<------------------------- 96% ------------------------->>

plotting   <<-------------------------- 100% -------------------------->>

plotting   <<-------------------------- 100% -------------------------->>
processing zt21_1_ribo.tr.sort
1. adding p-site position
2. adding transcript region
Reading zt21_2_ribo.tr.sort.bam
Input reads: 47.392 M
```

```
29.000 (< 0.001 %) reads removed: exceeding indel_threshold.
Great! All reads' reference transcript IDs were found in the annotation table. No reads removed.
0.747 M  (1.576 %) reads removed: mapping on negative strand.
Output reads: 46.645 M
Done! zt21_2_ribo.tr.sort.bam has been loaded as zt21_2_ribo.tr.sort
Warning: Removed 7272 rows containing missing values (`geom_tile()`).
processing zt21_2_ribo.tr.sort
best offset: 14 nts from the 5' end

plotting                          <<-- 8% -->>

plotting                        <<---- 16% ---->>

plotting                       <<----- 20% ----->>

plotting                     <<------- 28% ------->>

plotting                   <<--------- 36% --------->>

plotting                  <<---------- 40% ---------->>

plotting                <<------------ 48% ------------>>

plotting              <<-------------- 56% -------------->>

plotting             <<--------------- 60% --------------->>

plotting           <<----------------- 68% ----------------->>

plotting         <<------------------- 76% ------------------->>

plotting        <<-------------------- 80% -------------------->>

plotting      <<---------------------- 88% ---------------------->>

plotting    <<------------------------ 96% ------------------------>>

plotting   <<------------------------- 100% ------------------------->>

plotting   <<------------------------- 100% ------------------------->>
processing zt21_2_ribo.tr.sort
1. adding p-site position
2. adding transcript region
```

```
rm(annot_dt, bsg_tair, psite_offset_start, txdb_araport)
gc(); gc()
```

```
           used   (Mb) gc trigger    (Mb) limit (Mb)   max used    (Mb)
Ncells  11772353  628.8   24850738  1327.2         NA   24850738  1327.2
Vcells 799173066 6097.3 1960606682 14958.3      16384 2123736709 16202.9
```

```
           used   (Mb) gc trigger    (Mb) limit (Mb)   max used    (Mb)
Ncells  11772518  628.8   24850738  1327.2         NA   24850738  1327.2
Vcells 799169844 6097.2 1960606682 14958.3      16384 2123736709 16202.9
```

Sessioninfo

```
sessionInfo()
```

```
R version 4.2.1 (2022-06-23)
Platform: aarch64-apple-darwin20 (64-bit)
Running under: macOS Ventura 13.1

Matrix products: default
BLAS:   /Library/Frameworks/R.framework/Versions/4.2-arm64/Resources/lib/libRblas.0.dylib
LAPACK: /Library/Frameworks/R.framework/Versions/4.2-arm64/Resources/lib/libRlapack.dylib

locale:
[1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8

attached base packages:
[1] stats4    stats     graphics  grDevices datasets  utils     methods
[8] base

other attached packages:
 [1] BSgenome.Athaliana.TAIR.TAIR9_1.3.1000
 [2] BSgenome_1.64.0
 [3] rtracklayer_1.56.1
 [4] Biostrings_2.64.1
 [5] XVector_0.36.0
 [6] GenomicRanges_1.48.0
 [7] GenomeInfoDb_1.32.4
 [8] IRanges_2.30.1
 [9] S4Vectors_0.34.0
[10] BiocGenerics_0.42.0
[11] ggplot2_3.4.2
[12] magrittr_2.0.3

loaded via a namespace (and not attached):
 [1] bitops_1.0-7                matrixStats_0.62.0
 [3] fs_1.5.2                    bit64_4.0.5
 [5] filelock_1.0.2              progress_1.2.2
 [7] httr_1.4.5                  rprojroot_2.0.3
 [9] tools_4.2.1                 utf8_1.2.2
[11] R6_2.5.1                    DBI_1.1.3
[13] colorspace_2.0-3            withr_2.5.0
[15] tidyselect_1.2.0            prettyunits_1.1.1
[17] bit_4.0.5                   curl_4.3.3
[19] compiler_4.2.1              textshaping_0.3.6
[21] cli_3.6.0                   Biobase_2.56.0
[23] xml2_1.3.3                  DelayedArray_0.22.0
[25] labeling_0.4.2              scales_1.2.1
[27] readr_2.1.4                 rappdirs_0.3.3
[29] systemfonts_1.0.4           stringr_1.5.0
[31] digest_0.6.31               Rsamtools_2.12.0
[33] rmarkdown_2.24              pkgconfig_2.0.3
[35] htmltools_0.5.3             riboWaltz_1.2.0
[37] MatrixGenerics_1.8.1        dbplyr_2.3.2
[39] fastmap_1.1.0               rlang_1.1.0
[41] rstudioapi_0.14             RSQLite_2.2.18
[43] BiocIO_1.6.0                generics_0.1.3
[45] farver_2.1.1                jsonlite_1.8.4
[47] vroom_1.6.0                 BiocParallel_1.30.4
[49] dplyr_1.1.1                 RCurl_1.98-1.9
[51] GenomeInfoDbData_1.2.8      Matrix_1.6-4
[53] Rcpp_1.0.11                 munsell_0.5.0
[55] fansi_1.0.3                 lifecycle_1.0.3
[57] stringi_1.7.12              yaml_2.3.6
[59] SummarizedExperiment_1.26.1 zlibbioc_1.42.0
[61] BiocFileCache_2.4.0         grid_4.2.1
[63] blob_1.2.3                  parallel_4.2.1
```

```
[65] crayon_1.5.2                lattice_0.20-45
[67] GenomicFeatures_1.48.4      hms_1.1.3
[69] KEGGREST_1.36.3             knitr_1.42
[71] pillar_1.9.0                rjson_0.2.21
[73] codetools_0.2-18            biomaRt_2.52.0
[75] XML_3.99-0.11               glue_1.6.2
[77] evaluate_0.20               data.table_1.14.4
[79] renv_1.0.3                  BiocManager_1.30.18
[81] tzdb_0.3.0                  png_0.1-7
[83] vctrs_0.6.1                 gtable_0.3.1
[85] cachem_1.0.6                xfun_0.40
[87] restfulr_0.0.15             ragg_1.2.5
[89] tibble_3.2.1                GenomicAlignments_1.32.1
[91] AnnotationDbi_1.58.0        memoise_2.0.1
[93] here_1.0.1
```