

Figure1B_VideoReversals

MOD

2023-09-26

```
library(rhdf5)
library(tidyverse)

create_state_column <- function(df) {
  df %>%
    mutate(
      state = case_when(
        cumsum(!is.na(realSpeed) & realSpeed > 0) >= 1 ~ "post-reversal",
        TRUE ~ "pre-reversal"),
      censored = case_when(
        cumsum(!is.na(realSpeed) & realSpeed < 0 & state == "post-reversal") >= 1 ~ "censored",
        TRUE ~ "kept"
      )
    )
}

files <- fs::dir_ls(path='datasets/tracking', glob = "*cest-1.2*|*N2*")

mergedData <- map_df(files, read_csv, .id = "filename")
```

```
## Rows: 1246 Columns: 11
## -- Column specification -----
## Delimiter: ","
## chr (3): genotype, condition, flag
## dbl (8): worm_index, timestamp, realSpeed, coord_x_midbody, coord_y_midbody,...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
## Rows: 1155 Columns: 11
## -- Column specification -----
## Delimiter: ","
## chr (3): genotype, condition, flag
## dbl (8): worm_index, timestamp, realSpeed, coord_x_midbody, coord_y_midbody,...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
## Rows: 1063 Columns: 11
## -- Column specification -----
## Delimiter: ","
## chr (3): genotype, condition, flag
## dbl (8): worm_index, timestamp, realSpeed, coord_x_midbody, coord_y_midbody,...
##
```

```

## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
## Rows: 1120 Columns: 11
## -- Column specification -----
## Delimiter: ","
## chr (3): genotype, condition, flag
## dbl (8): worm_index, timestamp, realSpeed, coord_x_midbody, coord_y_midbody,...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
## Rows: 1141 Columns: 11
## -- Column specification -----
## Delimiter: ","
## chr (3): genotype, condition, flag
## dbl (8): worm_index, timestamp, realSpeed, coord_x_midbody, coord_y_midbody,...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
## Rows: 1056 Columns: 11
## -- Column specification -----
## Delimiter: ","
## chr (3): genotype, condition, flag
## dbl (8): worm_index, timestamp, realSpeed, coord_x_midbody, coord_y_midbody,...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
## Rows: 1180 Columns: 11
## -- Column specification -----
## Delimiter: ","
## chr (3): genotype, condition, flag
## dbl (8): worm_index, timestamp, realSpeed, coord_x_midbody, coord_y_midbody,...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
## Rows: 1059 Columns: 11
## -- Column specification -----
## Delimiter: ","
## chr (3): genotype, condition, flag
## dbl (8): worm_index, timestamp, realSpeed, coord_x_midbody, coord_y_midbody,...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.

```

#calculate a running loess fit over the data, vary the span

```

mergedData <- mergedData |>
  filter(!(genotype == "cest-1.2" & worm_index == 561 & filename == "datasets/tracking/2022-07-12_cest-1.2"))
  #make a unique wormID
  mutate(wormID = as.factor(as.numeric(as.factor(interaction(genotype, worm_index, filename))))) |>
  group_by(genotype, worm_index, filename) |>
  create_state_column() |>
  filter(censored == "kept") |>
  nest() |>
  mutate(loess_fit = purrr::map(data,
                                loess,

```

```

        formula = realSpeed ~ realTime,
        span = .5),
    fitted = purrr::map(loess_fit, ~predict(.x, newdata = seq(1,15, by = 0.1))))

# get the x intercepts:
Xints <- mergedData |>
  unnest(data) |>
  filter(state != lag(state) | state != lead(state)) |>
group_by(worm_index, genotype, filename) |>
  nest() |>
  dplyr::mutate(
    # get slope and y-intercept to calculate x intercept
    m = purrr::map(data, lm,
      formula = realSpeed ~ realTime),
    xint = purrr::map(m, function(x) {
      -x$coefficients[1]/x$coefficients[2]
    }) |>
  select(worm_index, genotype, xint) |>
  unnest(cols = c(xint))

```

Adding missing grouping variables: 'filename'

```

plot <- mergedData |>
  unnest(data) |>
ggplot(aes(x = realTime, y = realSpeed)) +
  geom_hline(yintercept = 0, linetype = 2, color = "grey") +
  geom_line(aes(color = genotype, group = wormID), size = 0.5, alpha = 0.5) +
  geom_point(data = Xints, aes(x = xint, y = 0, color = genotype), alpha = 0.5) +
  geom_smooth(aes(color = genotype, group = genotype, span = 0.1)) +
  # scale_color_manual(values = c( "darkgoldenrod2", "black")) +
  scale_color_manual(values = c( "royalblue2", "black")) +
  coord_cartesian(xlim = c(0,7.5), ylim = c(-.4,.4))

```

Warning: Using 'size' aesthetic for lines was deprecated in ggplot2 3.4.0.
 ## i Please use 'linewidth' instead.
 ## This warning is displayed once every 8 hours.
 ## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
 ## generated.

Warning in geom_smooth(aes(color = genotype, group = genotype, span = 0.1)):
 ## Ignoring unknown aesthetics: span

```

plot2 <- ggplot(filter(Xints, xint < 7.5), aes(x = xint, y = genotype, fill = genotype)) +
  ggdist::stat_halfeye(
    # adjust bandwidth
    adjust = 1,
    # move to the right
    justification = 0,
    # remove the slub interval
    #.width = 0,
    #point_colour = NA
    alpha = 0.5) +

```

```
# scale_fill_manual(values = c( "darkgoldenrod2", "black")) +
scale_fill_manual(values = c( "royalblue2", "black")) +
theme_void()
```

```
library(patchwork)
```

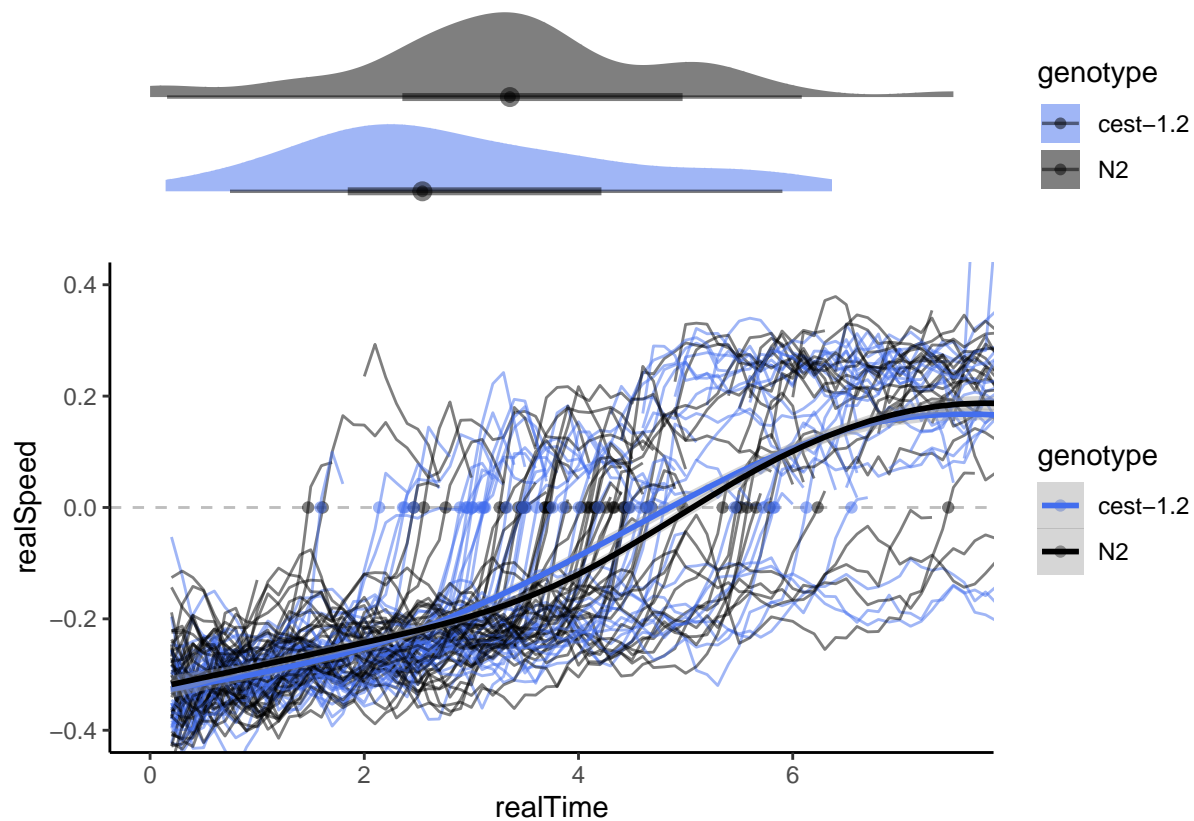
```
plot2 / plot + plot_layout(heights = c(1,2))
```

```
## 'geom_smooth()' using method = 'gam' and formula = 'y ~ s(x, bs = "cs")'
```

```
## Warning: Removed 1467 rows containing non-finite values ('stat_smooth()').
```

```
## Warning: Removed 430 rows containing missing values ('geom_line()').
```

```
## Warning: Removed 5 rows containing missing values ('geom_point()').
```



```
# first we want to get initial x/y positions of each animal post-reversal:
# start with midbody, may want to switch to head
#
```

```
Pos_start <- unnest(mergedData, data) |>
```

```
group_by(wormID) |>
```

```
summarize(X0 = first(coord_x_midbody), Y0 = first(coord_y_midbody))
```

```
mergedData <- unnest(mergedData, data) |>
```

```
full_join(Pos_start, by = join_by(wormID)) |>
```

```

# now use the X0, Y0 to calculate total (net) distance traveled at each time point
mutate(distance = sqrt((X0 - coord_x_midbody)^2 + (Y0 - coord_y_midbody)^2))

#correct distance for pixelSize in 2022-07-12 data
mergedData <- mergedData |>
  mutate(
    distance = case_when(
      stringr::str_detect(filename, "2022-07-12") ~ distance * (9.7/11.1) / 10,
      TRUE ~ distance
    )
  )

# get max distance during reversal
maxD <- mergedData |>
  filter(state == "pre-reversal") |>
  group_by(wormID, genotype) |>
  summarise(maxD = max(distance), realTime = max(realTime))

```

'summarise()' has grouped output by 'wormID'. You can override using the
'.groups' argument.

```

# now plot distance over time:

plot3 <- mergedData |>
  filter(state == "pre-reversal") |>
  ggplot() +
  geom_line(aes(x = realTime,
                y = distance,
                group = wormID,
                color = genotype), alpha = 0.5) +
  geom_point(data = maxD,
             aes(x = realTime, y = maxD, color = genotype), alpha = 0.5) +
  coord_cartesian(xlim = c(0,7.5)) +
  scale_color_manual(values = c( "royalblue2", "black")) +
  guides(color = 'none')

plot4 <- ggplot(filter(maxD, realTime < 7.5),
                aes(y = maxD, x = genotype, fill = genotype)) +
  ggdist::stat_halfeye(
    # adjust bandwidth
    adjust = 1,
    # move to the right
    justification = 0,
    # remove the slub interval
    #.width = 0,
    #point_colour = NA
  ) +
  alpha = 0.5) +
  # scale_fill_manual(values = c( "darkgoldenrod2", "black")) +
  scale_fill_manual(values = c( "royalblue2", "black")) +
  theme_void()

```

```
plot3 + plot4 + plot_layout(widths = c(5,2))
```

```
## Warning: Removed 207 rows containing missing values ('geom_line()').
```

```
## Warning: Removed 8 rows containing missing values ('geom_point()').
```

```
## Warning: Removed 5 rows containing missing values ('stat_slabinterval()').
```

