

TractMLESummary

This script is used to summarize IGC tract length model with 14 Yeast data sets.

```
rm(list=ls()) # clean up workspace
setwd("/Users/xji3/GitFolders/YeastIGCTract/Code/")

finished.pairs <- readLines('./Finished_Pairs.txt')

JS.HKY.nonclock.summary <- NULL
PSJS.HKY.nonclock.summary <- NULL
for(pair in finished.pairs){
  summary.file.name <- paste("JS_HKY", pair, "One_rate_nonclock_summary.txt", sep = "_")
  summary_file <- paste("./summary/", summary.file.name, sep = "")
  all <- readLines(summary_file, n = -1)
  row.names <- strsplit(all[length(all)], ' ')[[1]][-1]
  summary_mat <- as.matrix(read.table(summary_file,
                                     row.names = row.names,
                                     col.names = pair))
  JS.HKY.nonclock.summary <- cbind(JS.HKY.nonclock.summary, summary_mat)

  summary.file.name <- paste("PSJS_HKY", pair, "One_rate_nonclock_summary.txt", sep = "_")
  summary_file <- paste("./summary/", summary.file.name, sep = "")
  all <- readLines(summary_file, n = -1)
  row.names <- strsplit(all[length(all)], ' ')[[1]][-1]
  summary_mat <- as.matrix(read.table(summary_file,
                                     row.names = row.names,
                                     col.names = pair))
  PSJS.HKY.nonclock.summary <- cbind(PSJS.HKY.nonclock.summary, summary_mat)
}
```

Now show the estimated tract length and IGC initiation rates.

```
# Estimated Tract length (unit: nucleotide)
PSJS.HKY.nonclock.summary["tract_length", ]

## YER131W_YGL189C YLR406C_YDL075W YNL301C_YOL120C YBL087C_YER117W
##      6.5343489      2.6954536      20.2063946      12.3521620
## YER074W_YIL069C YLR333C_YGR027C YJL177W_YKL180W YML026C_YDR450W
##      43.7449667      18.6378714      16.8112335      0.6730206
## YDR418W_YEL054C YMR142C_YDL082W
##      2.2464840      28.0580220

# initial guess tract length
JS.HKY.nonclock.summary["Tau",] / exp(-0.6)

## YER131W_YGL189C YLR406C_YDL075W YNL301C_YOL120C YBL087C_YER117W
##      14.06610      14.58950      20.00219      25.39543
## YER074W_YIL069C YLR333C_YGR027C YJL177W_YKL180W YML026C_YDR450W
##      43.49963      20.21153      18.43512      28.46189
```

```
## YDR418W_YEL054C YMR142C_YDL082W
##          14.71203          28.59658
```

```
# Initiation rate
```

```
PSJS.HKY.nonclock.summary["init_rate", ]
```

```
## YER131W_YGL189C YLR406C_YDL075W YNL301C_YOL120C YBL087C_YER117W
##          1.1815526          2.9728313          0.5430014          1.1278237
## YER074W_YIL069C YLR333C_YGR027C YJL177W_YKL180W YML026C_YDR450W
##          0.5450514          0.5878825          0.5989426          23.2115751
## YDR418W_YEL054C YMR142C_YDL082W
##          3.5948859          0.5692985
```

```
# Now compare the product of init_rate and tract length with IGC+HKY model estimated Tau value
```

```
illustract.mat <- rbind(PSJS.HKY.nonclock.summary['init_rate', ] * PSJS.HKY.nonclock.summary["tract_length", ],
row.names(illustract.mat) <- c('product', 'estimated Tau')
illustract.mat
```

```
##          YER131W_YGL189C YLR406C_YDL075W YNL301C_YOL120C
## product          7.720677          8.013129          10.97210
## estimated Tau    7.719639          8.006885          10.97744
##          YBL087C_YER117W YER074W_YIL069C YLR333C_YGR027C
## product          13.93106          23.84326          10.95688
## estimated Tau    13.93731          23.87310          11.09232
##          YJL177W_YKL180W YML026C_YDR450W YDR418W_YEL054C
## product          10.06896          15.62187          8.075854
## estimated Tau    10.11741          15.62022          8.074132
##          YMR142C_YDL082W
## product          15.97339
## estimated Tau    15.69413
```

```
# Now export data
```

```
save.image(file = "./TractSummary.RData")
```