

# EDN\_ECP\_summary

Xiang Ji

January 12, 2016

1, Read in tables

```
rm(list=ls()) # clean up workspace
path <- "/Users/xji3/GitFolders/EDN_ECP/Summary/MG94"

summary.list <- c("_clock_summary",
                  "_nonclock_summary",
                  "_Force_clock_summary",
                  "_Force_nonclock_summary"
                  )
pair = c("EDN", "ECP")
for (target.summary in summary.list){
  summary_file <- paste(path, "_EDN_ECP", target.summary, '.txt', sep = '')
  all <- readLines(summary_file, n = -1)
  row.names <- strsplit(all[length(all)], ' ')[[1]][-1]
  col.name <- paste("MG94", target.summary, sep = "")
  summary_mat <- as.matrix(read.table(summary_file,
                                      row.names = row.names,
                                      col.names = col.name))
  assign(paste("MG94", target.summary, sep = ""), summary_mat)
}
ECP.EDN.MG94 <- cbind(MG94_nonclock_summary, MG94_clock_summary,
                      MG94_Force_nonclock_summary, MG94_Force_clock_summary)
ECP.EDN.MG94
```

##	MG94_nonclock_summary	MG94_clock_summary
## length	1.570000e+02	1.570000e+02
## ll	-1.700519e+03	-1.703650e+03
## pi_a	2.910093e-01	2.916945e-01
## pi_c	2.433912e-01	2.426929e-01
## pi_g	2.069226e-01	2.068703e-01
## pi_t	2.586769e-01	2.587424e-01
## kappa	2.062436e+00	2.089065e+00
## omega	8.270292e-01	8.389169e-01
## tau	6.312271e-01	6.207583e-01
## (N0,N1)	1.990860e-01	1.960748e-01
## (N0,Tamarin)	3.252961e-01	3.268887e-01
## (N1,N2)	3.193865e-02	5.181832e-02
## (N1,Macaque)	1.564817e-01	1.250430e-01
## (N2,N3)	3.455331e-02	5.565978e-02
## (N2,Orangutan)	8.979048e-02	7.322467e-02
## (N3,Chimpanzee)	1.427697e-02	1.756489e-02
## (N3,Gorilla)	1.596838e-02	1.756489e-02
## (N0,N1,tau)	5.202655e-01	5.194297e-01
## (N0,Tamarin,tau)	0.000000e+00	0.000000e+00
## (N1,N2,tau)	4.162337e-01	4.911286e-01
## (N1,Macaque,tau)	3.747596e-01	3.373610e-01
## (N2,N3,tau)	4.485431e-01	4.106192e-01

## (N2,Orangutan,tau)	1.071400e+00	1.231183e+00
## (N3,Chimpanzee,tau)	4.868393e-02	6.296886e-02
## (N3,Gorilla,tau)	4.722111e-01	4.331299e-01
## (N0,N1,1->2)	2.592440e+00	2.515012e+00
## (N0,Tamarin,1->2)	0.000000e+00	0.000000e+00
## (N1,N2,1->2)	3.635547e-01	7.182819e-01
## (N1,Macaque,1->2)	1.512961e+00	1.002534e+00
## (N2,N3,1->2)	5.827584e-01	8.879946e-01
## (N2,Orangutan,1->2)	6.399548e+00	6.069097e+00
## (N3,Chimpanzee,1->2)	6.505905e-02	1.065007e-01
## (N3,Gorilla,1->2)	5.130258e-02	6.176839e-02
## (N0,N1,2->1)	2.592440e+00	2.515012e+00
## (N0,Tamarin,2->1)	0.000000e+00	0.000000e+00
## (N1,N2,2->1)	9.049477e-01	1.701750e+00
## (N1,Macaque,2->1)	5.101368e+00	3.726697e+00
## (N2,N3,2->1)	1.046395e+00	1.519445e+00
## (N2,Orangutan,2->1)	3.413858e+00	3.137466e+00
## (N3,Chimpanzee,2->1)	1.286481e-02	1.756355e-02
## (N3,Gorilla,2->1)	7.941026e-01	7.916351e-01
## (N0,N1,mut)	6.295535e+01	6.198371e+01
## (N0,Tamarin,mut)	5.115805e+01	5.141753e+01
## (N1,N2,mut)	1.045748e+01	1.399356e+01
## (N1,Macaque,mut)	5.231267e+01	4.851329e+01
## (N2,N3,mut)	1.122474e+01	1.245078e+01
## (N2,Orangutan,mut)	2.350807e+01	2.198678e+01
## (N3,Chimpanzee,mut)	5.280903e+00	5.203542e+00
## (N3,Gorilla,mut)	5.147585e+00	5.162176e+00
##	MG94_Force_nonclock_summary	MG94_Force_clock_summary
## length	1.570000e+02	1.570000e+02
## ll	-1.714099e+03	-1.716567e+03
## pi_a	2.927431e-01	2.927869e-01
## pi_c	2.425981e-01	2.421026e-01
## pi_g	2.076225e-01	2.078204e-01
## pi_t	2.570363e-01	2.572901e-01
## kappa	2.100482e+00	2.102322e+00
## omega	9.044276e-01	9.065773e-01
## tau	0.000000e+00	0.000000e+00
## (N0,N1)	1.440343e-01	1.400382e-01
## (N0,Tamarin)	3.556100e-01	3.579323e-01
## (N1,N2)	4.519787e-02	6.042200e-02
## (N1,Macaque)	1.777791e-01	1.480029e-01
## (N2,N3)	4.510421e-02	6.707861e-02
## (N2,Orangutan)	9.981750e-02	8.758086e-02
## (N3,Chimpanzee)	1.699996e-02	2.050224e-02
## (N3,Gorilla)	1.880088e-02	2.050224e-02
## (N0,N1,tau)	0.000000e+00	0.000000e+00
## (N0,Tamarin,tau)	0.000000e+00	0.000000e+00
## (N1,N2,tau)	0.000000e+00	0.000000e+00
## (N1,Macaque,tau)	0.000000e+00	0.000000e+00
## (N2,N3,tau)	0.000000e+00	0.000000e+00
## (N2,Orangutan,tau)	0.000000e+00	0.000000e+00
## (N3,Chimpanzee,tau)	0.000000e+00	0.000000e+00
## (N3,Gorilla,tau)	0.000000e+00	0.000000e+00
## (N0,N1,1->2)	0.000000e+00	0.000000e+00

```

## (N0,Tamarin,1->2) 0.000000e+00 0.000000e+00
## (N1,N2,1->2) 0.000000e+00 0.000000e+00
## (N1,Macaque,1->2) 0.000000e+00 0.000000e+00
## (N2,N3,1->2) 0.000000e+00 0.000000e+00
## (N2,Orangutan,1->2) 0.000000e+00 0.000000e+00
## (N3,Chimpanzee,1->2) 0.000000e+00 0.000000e+00
## (N3,Gorilla,1->2) 0.000000e+00 0.000000e+00
## (N0,N1,2->1) 0.000000e+00 0.000000e+00
## (N0,Tamarin,2->1) 0.000000e+00 0.000000e+00
## (N1,N2,2->1) 0.000000e+00 0.000000e+00
## (N1,Macaque,2->1) 0.000000e+00 0.000000e+00
## (N2,N3,2->1) 0.000000e+00 0.000000e+00
## (N2,Orangutan,2->1) 0.000000e+00 0.000000e+00
## (N3,Chimpanzee,2->1) 0.000000e+00 0.000000e+00
## (N3,Gorilla,2->1) 0.000000e+00 0.000000e+00
## (N0,N1,mut) 4.544445e+01 4.418651e+01
## (N0,Tamarin,mut) 5.598401e+01 5.635691e+01
## (N1,N2,mut) 1.427259e+01 1.584501e+01
## (N1,Macaque,mut) 5.608860e+01 5.462142e+01
## (N2,N3,mut) 1.422474e+01 1.573075e+01
## (N2,Orangutan,mut) 3.153865e+01 3.006394e+01
## (N3,Chimpanzee,mut) 5.352806e+00 5.299775e+00
## (N3,Gorilla,mut) 5.923543e+00 5.952135e+00

```

2, Now show branch specific % changes due to IGC

```
(ECP.EDN.MG94[26:33, ] + ECP.EDN.MG94[34:41, ])/(ECP.EDN.MG94[42:49, ] + ECP.EDN.MG94[26:33, ] + ECP.EDN.MG94[34:41, ])
```

```

## MG94_nonclock_summary MG94_clock_summary
## (N0,N1,1->2) 0.07609132 0.07505960
## (N0,Tamarin,1->2) 0.00000000 0.00000000
## (N1,N2,1->2) 0.10817880 0.14744073
## (N1,Macaque,1->2) 0.11224614 0.08882434
## (N2,N3,1->2) 0.12674399 0.16202747
## (N2,Orangutan,1->2) 0.29450698 0.29514515
## (N3,Chimpanzee,1->2) 0.01454122 0.02328706
## (N3,Gorilla,1->2) 0.14106566 0.14186554
## MG94_Force_nonclock_summary MG94_Force_clock_summary
## (N0,N1,1->2) 0 0
## (N0,Tamarin,1->2) 0 0
## (N1,N2,1->2) 0 0
## (N1,Macaque,1->2) 0 0
## (N2,N3,1->2) 0 0
## (N2,Orangutan,1->2) 0 0
## (N3,Chimpanzee,1->2) 0 0
## (N3,Gorilla,1->2) 0 0

```

3, % changes due to IGC in all branches

```
colSums(ECP.EDN.MG94[26:33, ] + ECP.EDN.MG94[34:41, ])/colSums(ECP.EDN.MG94[42:49, ] + ECP.EDN.MG94[26:33, ] + ECP.EDN.MG94[34:41, ])
```

```

## MG94_nonclock_summary MG94_clock_summary
## 0.1027710 0.1009066
## MG94_Force_nonclock_summary MG94_Force_clock_summary
## 0.0000000 0.0000000

```

12212017 update

HKY+PS-IGC results

```
all <- readLines("./Summary/PSJS_HKY_EDN_ECP_One_rate_Guess_1_rv_SCOK_nonclock_summary.txt", n = -1)
col.names <- "Guess_1"
row.names <- strsplit(all[length(all)], ' ')[[1]][-1]
EDN.ECP.guess.1.result <- as.matrix(read.table("./Summary/PSJS_HKY_EDN_ECP_One_rate_Guess_1_rv_SCOK_nonclock_summary.txt", n = -1))

all <- readLines("./Summary/PSJS_HKY_EDN_ECP_One_rate_Guess_2_rv_SCOK_nonclock_summary.txt", n = -1)
col.names <- "Guess_2"
row.names <- strsplit(all[length(all)], ' ')[[1]][-1]
EDN.ECP.guess.2.result <- as.matrix(read.table("./Summary/PSJS_HKY_EDN_ECP_One_rate_Guess_2_rv_SCOK_nonclock_summary.txt", n = -1))

gradient.file <- "./Summary/PSJS_HKY_EDN_ECP_One_rate_Guess_1_nonclock_gradient.txt"
gradient.list <- read.table(gradient.file)
cat("Verify gradient ~ 0: Gradient = ", colSums(gradient.list), ". Gradient/Objective = ", colSums(gradient.list/EDN.ECP.guess.1.result))

## Verify gradient ~ 0: Gradient = 0.03022336 -0.1039073 . Gradient/Objective = 3.753859e-08 -1.290561
Godambe.matrix.guess.1 <- read.table("./Summary/PSJS_HKY_EDN_ECP_One_rate_Guess_1_nonclock_godambe.txt", n = -1)

gradient.file <- "./Summary/PSJS_HKY_EDN_ECP_One_rate_Guess_2_nonclock_gradient.txt"
gradient.list <- read.table(gradient.file)
cat("Verify gradient ~ 0: Gradient = ", colSums(gradient.list), ". Gradient/Objective = ", colSums(gradient.list/EDN.ECP.guess.2.result))

## Verify gradient ~ 0: Gradient = -0.1178978 0.3038744 . Gradient/Objective = -1.464336e-07 3.774238
Godambe.matrix.guess.2 <- read.table("./Summary/PSJS_HKY_EDN_ECP_One_rate_Guess_2_nonclock_godambe.txt", n = -1)

Results <- cbind(EDN.ECP.guess.1.result, EDN.ECP.guess.2.result)
Results
```

	Guess_1	Guess_2
## ll	8.051278e+05	8.051278e+05
## length	4.710000e+02	4.710000e+02
## Pi_A	2.803335e-01	2.803501e-01
## Pi_C	2.550526e-01	2.550557e-01
## Pi_G	2.088842e-01	2.088806e-01
## Pi_T	2.557297e-01	2.557136e-01
## kappa	2.115032e+00	2.114975e+00
## r2	1.518936e+00	1.519417e+00
## r3	1.557700e+00	1.558265e+00
## init_rate	3.625929e-01	3.525476e-01
## tract_length	6.699765e+00	6.895731e+00
## D1__N1	5.251823e-02	5.252127e-02
## NO__D1	7.670355e-03	7.668180e-03
## NO__Tamarin	6.903320e-02	6.901362e-02
## N1__Macaque	3.820307e-02	3.819756e-02
## N1__N2	6.289331e-03	6.276964e-03
## N2__N3	8.150475e-03	8.145238e-03
## N2__Orangutan	2.213472e-02	2.212855e-02
## N3__Chimpanzee	3.386406e-03	3.381945e-03
## N3__Gorilla	3.717720e-03	3.715324e-03

```
Godambe.inverse <- cbind(c(solve(Godambe.matrix.guess.1)/dim(gradient.list)[1]),
                          c(solve(Godambe.matrix.guess.2)/dim(gradient.list)[1]))
```

```
Godambe.inverse
```

```
##           [,1]           [,2]
## [1,]  1.124734e-04  1.124108e-04
## [2,] -1.477086e-05 -9.732373e-06
## [3,] -1.477086e-05 -9.732373e-06
## [4,]  1.273179e-01  1.199276e-01
```

```
# Guess 2 has higher lnL
which.max(Results["ll",])
```

```
## Guess_2
##      2
```

```
# effective Tau
```

```
Results["init_rate", ] * Results["tract_length", ] * 3 / (1.0 + colSums(Results[c("r2", "r3"), ]))
```

```
## Guess_1 Guess_2
## 1.787715 1.788570
```

```
Results["init_rate", ] * Results["tract_length", ] * 3 / exp(1.96*sqrt(Godambe.inverse[1, ])) / (1.0 + c
```

```
## Guess_1 Guess_2
## 1.750938 1.751786
```

```
Results["init_rate", ] * Results["tract_length", ] * 3 * exp(1.96*sqrt(Godambe.inverse[1, ])) / (1.0 + c
```

```
## Guess_1 Guess_2
## 1.825264 1.826127
```

```
# Tract length
```

```
Results["tract_length", ]
```

```
## Guess_1 Guess_2
## 6.699765 6.895731
```

```
exp(log(Results["tract_length", ]-1.0)-1.96*sqrt(Godambe.inverse[4,]))+1.0
```

```
## Guess_1 Guess_2
## 3.832231 3.990586
```

```
exp(log(Results["tract_length", ]-1.0)+1.96*sqrt(Godambe.inverse[4,]))+1.0
```

```
## Guess_1 Guess_2
## 12.47058 12.62302
```

```
HKY + IS-IGC results
```

```
all <- readLines("./Summary/JS_HKY_EDN_ECP_One_rate_rv_nonclock_summary.txt", n = -1)
```

```
col.names <- c("HKY+IS-IGC")
```

```
row.names <- strsplit(all[length(all)], ' ')[[1]][-1]
```

```
EDN.ECP.HKY.result <- as.matrix(read.table("./Summary/JS_HKY_EDN_ECP_One_rate_rv_nonclock_summary.txt",
```

```
all <- readLines("./Summary/Force_JS_HKY_EDN_ECP_One_rate_rv_nonclock_summary.txt", n = -1)
```

```
col.names <- c("Force_HKY+IS-IGC")
```

```
row.names <- strsplit(all[length(all)], ' ')[[1]][-1]
```

```
Force.EDN.ECP.HKY.result <- as.matrix(read.table("./Summary/Force_JS_HKY_EDN_ECP_One_rate_rv_nonclock_s
```

```
gradient.file <- "./Summary/JS_HKY_EDN_ECP_One_rate_rv_nonclock_gradient.txt"
```

```

gradient.list <- read.table(gradient.file)
cat("Verify gradient ~ 0: Gradient = ", colSums(gradient.list), ". Gradient/Objective = ", colSums(gradient.list))

## Verify gradient ~ 0: Gradient = 0.01009568 -0.0004857768 -7.803258e-05 -0.0002322224 0.0001139664 0.0001139664 0.0001139664 0.0001139664 0.0001139664

Godambe.matrix <- read.table("./Summary/JS_HKY_EDN_ECP_One_rate_rv_nonclock_godambe.txt")
Godambe.JS <- cbind(diag(solve(Godambe.matrix)/dim(gradient.list)[1]))

show.mat <- c(EDN.ECP.HKY.result["l1", ],
              EDN.ECP.HKY.result["l1", ] - Force.EDN.ECP.HKY.result["l1", ],
              EDN.ECP.HKY.result["Tau", ]*3.0/(1. + sum(EDN.ECP.HKY.result[c("r2", "r3"),])),
              EDN.ECP.HKY.result["Tau", ]/exp(1.96*sqrt(Godambe.JS[7, ]))*3.0/(1. + sum(EDN.ECP.HKY.result[c("r2", "r3"),])),
              EDN.ECP.HKY.result["Tau", ]*exp(1.96*sqrt(Godambe.JS[7, ]))*3.0/(1. + sum(EDN.ECP.HKY.result[c("r2", "r3"),])),
              EDN.ECP.HKY.result["r2", ],
              EDN.ECP.HKY.result["r2", ]/exp(1.96*sqrt(Godambe.JS[5, ])),
              EDN.ECP.HKY.result["r2", ]*exp(1.96*sqrt(Godambe.JS[5, ])),
              EDN.ECP.HKY.result["r3", ],
              EDN.ECP.HKY.result["r3", ]/exp(1.96*sqrt(Godambe.JS[6, ])),
              EDN.ECP.HKY.result["r3", ]*exp(1.96*sqrt(Godambe.JS[6, ]))
              )
names(show.mat) <- c("l1", "Diff", "Tau", "min", "max",
                    "r2", "min", "max",
                    "r3", "min", "max")
round(show.mat, digits = 2)

```

```

##      l1      Diff      Tau      min      max      r2      min      max
## -1713.06  12.61    1.79    0.90    3.56    1.52    1.03    2.23
##      r3      min      max
##      1.56    1.06    2.30

```

### 12302017 %changes due to IGC

```

num.IGC <- sum(EDN.ECP.HKY.result[20:28,])
num.Mut <- sum(EDN.ECP.HKY.result[c(29, 32:37), ])
# % changes due to IGC
num.IGC/(num.IGC+num.Mut)

```

```
## [1] 0.1608216
```

Now save workspace.

```
save.image("./EDN_ECP_Summary.RData")
```