# Fit mutant with binomial aging model

*h qin*

*2017, May 2 ~ June*

```
rm(list=ls())
host = "Applejack" #"Ridgeside"
if (host == "AppleJack") {
 setwd("/Users/hqin/github/bmc_netwk_aging_manuscript/R1/1.mutants")
}
if (host == "Ridgeside") {
 setwd("/hong/hqin/github/bmc_netwk_aging_manuscript/R1/1.mutants")
}

library('flexsurv')
```

```
## Loading required package: survival
```

```
library('stringr')
source("../lifespan.r")
```

## parse the strains from files

Change mydir to either YPD or rls05D and run this code twice will generate the two output files in sandbox.

```
mydir = 'rlsYPD'
#mydir = 'rls05D'
files = list.files(path=paste(mydir,"/", sep=''), pattern="csv")

genotypes = c();
media = c();
for( i in 1:length(files)) {
 elements = unlist(str_split(files[i], "_"))
 genotypes = c(genotypes, elements[1])
 media = c(media, elements[length(elements)])
}
genotypes = unique(genotypes)
genotypes
```

```
## [1] "BY4742" "sir2"   "sod2"   "tor1"
```

```
#media
```

Take files from natural isolates

```
report = data.frame(files)
report$samplesize = NA; report$R=NA; report$t0=NA; report$n=NA; report$G=NA;
#report$longfilename=NA;
```

## Explore the fitting outcomes of 'flexsurv'.

```
i=2
 # tb = read.table( paste("../qinlab_rls/",files[i],sep=''), sep="\t")
  tb = read.csv( paste(mydir,"/",files[i] ,sep=''), sep="\t")
  GompFlex = flexsurvreg(formula = Surv(tb[,1]) ~ 1, dist = 'gompertz')
  WeibFlex = flexsurvreg(formula = Surv(tb[,1]) ~ 1, dist = 'weibull')
#str(GompFlex)
GompFlex$res
```

```
##              est       L95%       U95%          se
## shape 0.15084682 0.13763862 0.16405502 0.006739004
## rate  0.01709645 0.01400899 0.02086437 0.001737344
```

```
GompFlex$res.t
```

```
##              est       L95%      U95%          se
## shape  0.1508468  0.1376386  0.164055 0.006739004
## rate  -4.0688842 -4.2680560 -3.869712 0.101620127
```

```
GompFlex$opt$hessian
```

```
##           shape      rate
## shape 90953.233 5250.9845
## rate   5250.984  399.9908
```

```
#str(WeibFlex)
```

## Now, fit all RLS data sets by strains

```
for( i in 1:length(report[,1])){
#for( i in 3:4){
  tb = read.csv( paste(mydir,"/",files[i] ,sep=''), sep="\t")
  #tb = read.table( paste("../qinlab_rls/",my.files[1],sep=''), sep="\t")
  report$samplesize[i] = length(tb[,1])

  GompFlex = flexsurvreg(formula = Surv(tb[,1]) ~ 1, dist = 'gompertz')
  WeibFlex = flexsurvreg(formula = Surv(tb[,1]) ~ 1, dist = 'weibull')

  report$avgLS[i] =  mean(tb[,1])
  report$stdLS[i] =  sd(tb[,1])
  report$CV[i] = report$stdLS[i] / report$avgLS[i]

  report$GompGFlex[i] = GompFlex$res[1,1]
  report$GompRFlex[i] = GompFlex$res[2,1]
  report$GompLogLikFlex[i] = round(GompFlex$loglik, 1)
  report$GompAICFlex[i] = round(GompFlex$AIC)

  report$WeibShapeFlex[i] = WeibFlex$res[1,1]
  report$WeibRateFlex[i] = WeibFlex$res[2,1]
  report$WeibLogLikFlex[i] = round(WeibFlex$loglik, 1)
  report$WeibAICFlex[i] = round(WeibFlex$AIC)
```

```
  #set initial values
  Rhat = report$GompRFlex[i]; # 'i' was missing. a bug costed HQ a whole afternoon.
  Ghat = report$GompGFlex[i];
  nhat = 6;
  t0= (nhat-1)/Ghat;
  fitBinom = optim ( c(Rhat, t0, nhat),  llh.binomialMortality.single.run,
                     lifespan=tb[,1],
                     #method='SANN') #SANN needs control
                     method="L-BFGS-B",
                     lower=c(1E-10, 1, 1), upper=c(1,200,20) );
  report[i, c("R", "t0", "n")] = fitBinom$par[1:3]
  report$G[i] = (report$n[i] - 1)/report$t0[i]
}
```

Show the results

```
#report[ grep("tBY", report$strains), ]
report
```

```
##                                       files samplesize           R
## 1 BY4742_MATalpha_temp30_media_YPD_n19930.csv      19930 0.004634869
## 2         sir2_MATa_temp30_media_YPD_n400.csv        400 0.003872312
## 3     sod2_MATalpha_temp30_media_YPD_n542.csv        542 0.007511820
## 4    tor1_MATalpha_temp30_media_YPD_n2823.csv       2823 0.003328287
##          t0        n          G    avgLS     stdLS        CV  GompGFlex
## 1 50.910004 7.564035 0.12893408 26.22910  9.565108 0.3646754 0.08747631
## 2  3.924184 3.509807 0.63957431 13.12750  4.665368 0.3553889 0.15084682
## 3 74.206648 7.536815 0.08808934 26.78598 12.940442 0.4831051 0.06710602
## 4 56.603468 7.479705 0.11447541 32.12256 11.737375 0.3653935 0.08270110
##     GompRFlex GompLogLikFlex GompAICFlex WeibShapeFlex WeibRateFlex
## 1 0.006510853       -74535.4      149075      2.953187     29.36164
## 2 0.017096454        -1235.4        2475      2.962374     14.67888
## 3 0.008849582        -2130.0        4264      2.110887     30.01028
## 4 0.003925530       -10963.6       21931      3.035092     35.94141
##   WeibLogLikFlex WeibAICFlex
## 1       -73218.0      146440
## 2        -1182.3        2369
## 3        -2165.5        4335
## 4       -10950.5       21905
```

Output

```
write.csv(report, file = paste('sandbox/_report_', mydir, '.csv', sep='' ), row.names = FALSE)
```