

# stats\_pairwise\_corr\_assembly

```
data <- read.csv('Z:/home/ryanh/projects/ripple_heterogeneity/assembly_unit_corrs/results/df.csv')
```

```
data$basepath = factor(data$basepath)
data$name = factor(data$name)
data$label = factor(data$label)
data$membership = factor(data$membership)
data$assembly_id <- factor(as.character(data$assembly_id))

summary(data)
```

```
##           X           Unnamed..0           rho
## Min.      : 13936   Min.      :      2   Min.      :-0.3839111
## 1st Qu.:2719625   1st Qu.: 43805   1st Qu.: 0.0000783
## Median :3285234   Median : 578155   Median : 0.0202598
## Mean    :3351281   Mean    : 981266   Mean    : 0.0273914
## 3rd Qu.:4496409   3rd Qu.:1973326   3rd Qu.: 0.0470042
## Max.    :6028057   Max.    :2538935   Max.    : 1.0000000
##
##           label           assembly_id
## member_deep      :    1166   362      : 28920
## member_deep_sup  :     863   364      : 28920
## member_sup       :     345   373      : 28920
## non_member_deep  :1097583   374      : 28920
## non_member_deep_sup: 835156   382      : 28920
## non_member_sup   : 327324   388      : 28920
##                                     (Other):2088917
##                                     basepath
## Z:\\Data\\GrosmarkAD\\Achilles\\Achilles_10252013:986292
## Z:\\Data\\ORproject\\OR15\\day10                      :253517
## Z:\\Data\\GrosmarkAD\\Achilles\\Achilles_11012013: 79022
## Z:\\Data\\ORproject\\OR22\\day3                      : 44201
## Z:\\Data\\ORproject\\OR22\\day5                      : 43710
## Z:\\Data\\Kenji\\i01_maze05_MS.001_003                : 33554
## (Other)                                                :822141
##           membership           name
## member      :    2374   Achilles:1065314
## non_member:2260063   OR15      : 313295
##                                     OR22      : 119848
##                                     AB3       : 116629
##                                     ec014     : 115635
##                                     Rat11     : 77557
##                                     (Other)   : 454159
```

```
# data$assembly_id
```

# make var for simple effect

```
data$simple_effect = "unknown"
data$simple_effect[data$label == "member_deep" | data$label == "member_sup"] = "member_single_layer"
data$simple_effect[data$label == "member_deep_sup"] = "member_cross_layer"
data$simple_effect = factor(data$simple_effect)

summary(data)
```

```
##           X           Unnamed..0           rho
## Min.      : 13936   Min.      :    2   Min.      :-0.3839111
## 1st Qu.:2719625   1st Qu.: 43805   1st Qu.: 0.0000783
## Median :3285234   Median : 578155   Median : 0.0202598
## Mean    :3351281   Mean    : 981266   Mean    : 0.0273914
## 3rd Qu.:4496409   3rd Qu.:1973326   3rd Qu.: 0.0470042
## Max.    :6028057   Max.    :2538935   Max.    : 1.0000000
##
##           label           assembly_id
## member_deep      :    1166   362      : 28920
## member_deep_sup  :     863   364      : 28920
## member_sup       :     345   373      : 28920
## non_member_deep  :1097583   374      : 28920
## non_member_deep_sup: 835156   382      : 28920
## non_member_sup   : 327324   388      : 28920
##                  (Other):2088917
##
##                               basepath
## Z:\\Data\\GrosmarkAD\\Achilles\\Achilles_10252013:986292
## Z:\\Data\\ORproject\\OR15\\day10                :253517
## Z:\\Data\\GrosmarkAD\\Achilles\\Achilles_11012013: 79022
## Z:\\Data\\ORproject\\OR22\\day3                  : 44201
## Z:\\Data\\ORproject\\OR22\\day5                  : 43710
## Z:\\Data\\Kenji\\i01_maze05_MS.001_003          : 33554
## (Other)                                           :822141
##
## membership           name           simple_effect
## member      :    2374   Achilles:1065314   member_cross_layer :    863
## non_member:2260063   OR15      : 313295   member_single_layer:   1511
##
##                               OR22      : 119848   unknown              :2260063
##
##                               AB3       : 116629
##
##                               ec014    : 115635
##
##                               Rat11    :  77557
##
##                               (Other)  : 454159
```

```
length(unique(data$rho))
```

```
## [1] 187253
```

```
length(data$rho)
```

```
## [1] 2262437
```

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5      v purrr   0.3.4
## v tibble  3.1.6      v dplyr   1.0.7
## v tidyr   1.1.4      v stringr 1.4.0
## v readr   2.1.1      v forcats 0.5.1
```

```
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(lme4)
```

```
## Loading required package: Matrix
```

```
##
## Attaching package: 'Matrix'
```

```
## The following objects are masked from 'package:tidyr':
##
##   expand, pack, unpack
```

```
library(lmerTest)
```

```
##
## Attaching package: 'lmerTest'
```

```
## The following object is masked from 'package:lme4':
##
##   lmer
```

```
## The following object is masked from 'package:stats':
##
##   step
```

```
library(MASS)
```

```
##
## Attaching package: 'MASS'
```

```
## The following object is masked from 'package:dplyr':
##
##      select
```

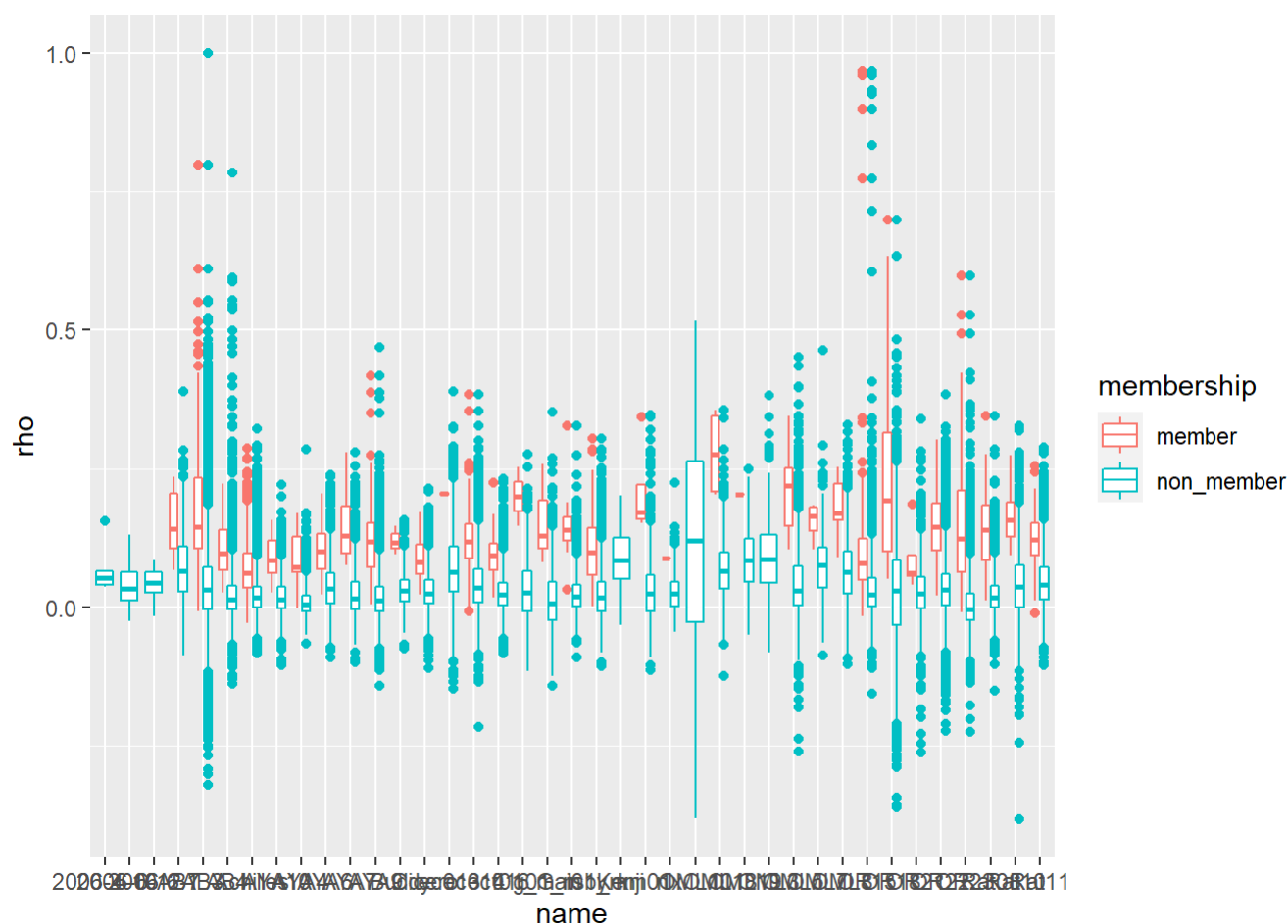
```
library(DHARMA)
```

```
## This is DHARMA 0.4.5. For overview type '?DHARMA'. For recent changes, type news(package = 'D
HARMA')
```

```
library(ggplot2)
```

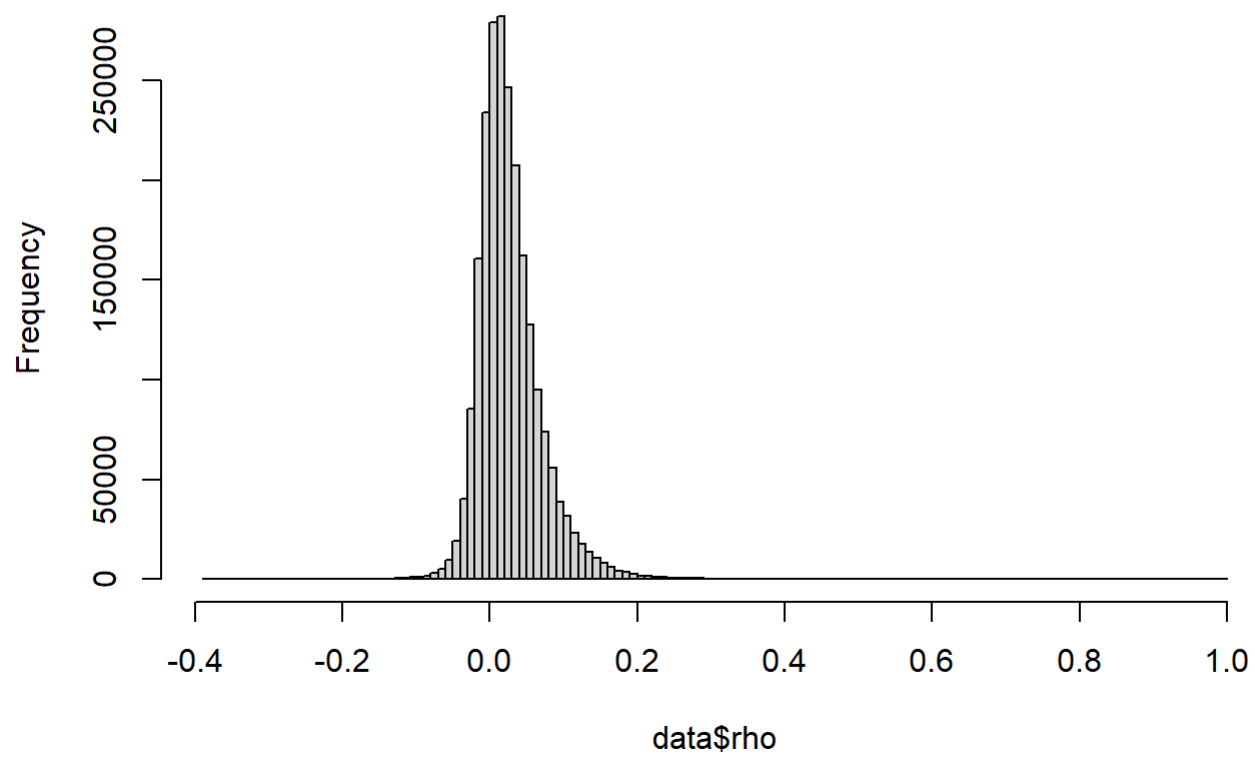
```
library(ggplot2)

p<-ggplot(data, aes(x=name, y=rho, color=membership)) +
  geom_boxplot()
p
```



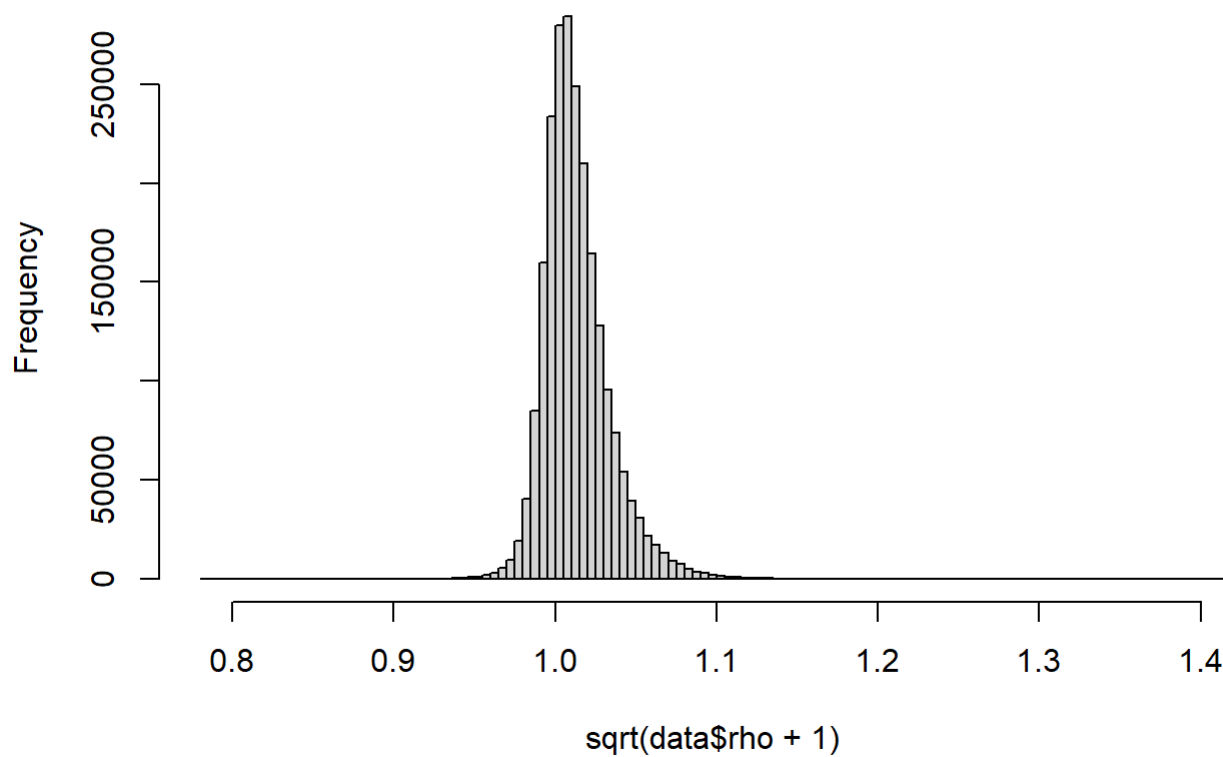
```
hist(data$rho,100)
```

## Histogram of data\$rho



```
hist(sqrt(data$rho + 1),100)
```

## Histogram of $\sqrt{\text{data}\$rho + 1}$

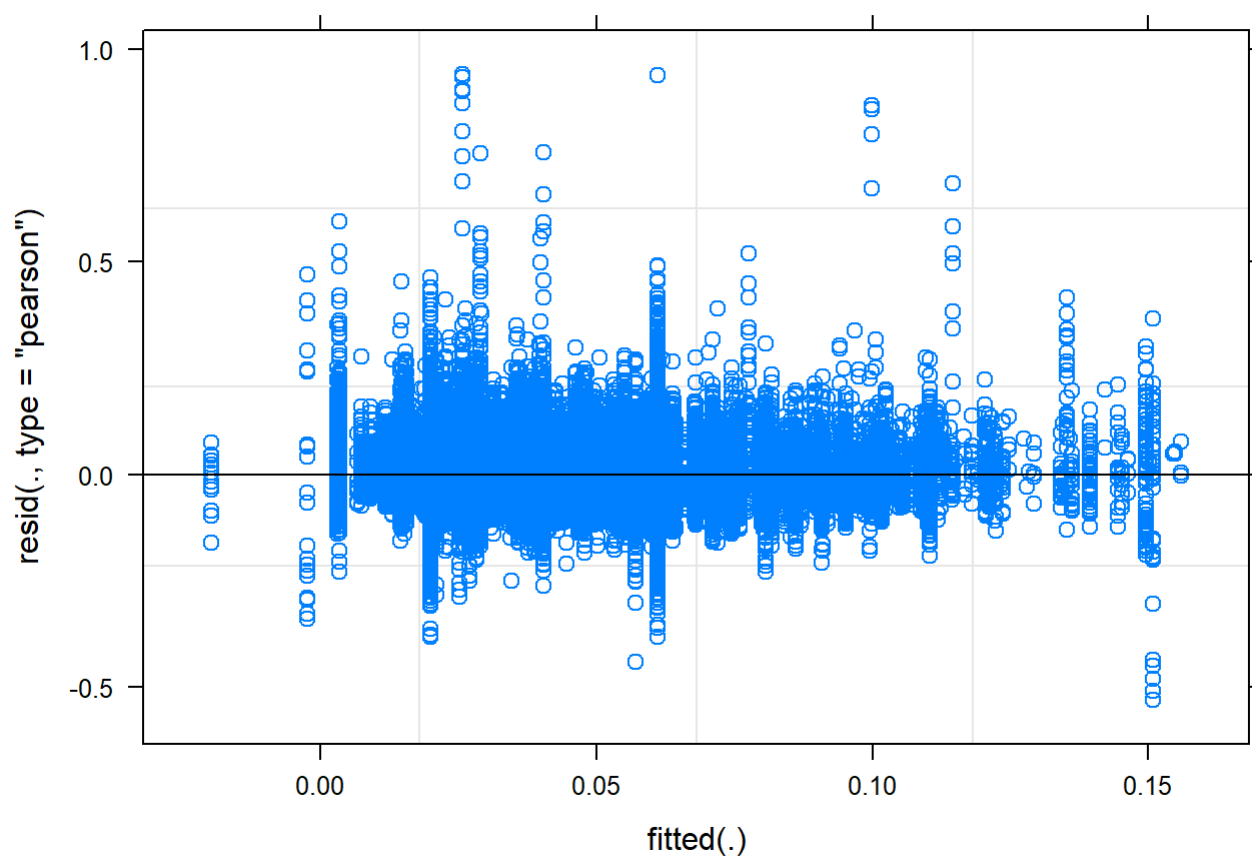


```
data$rho_trans <- log(data$rho+1)
data$rho_trans <- data$rho
```

```
m1 = lmer(rho_trans ~ membership + (1 | name/basepath), data = data, REML=FALSE)
summary(m1)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
## method [lmerModLmerTest]
## Formula: rho_trans ~ membership + (1 | name/basepath)
## Data: data
##
##      AIC      BIC    logLik deviance df.resid
## -7943982 -7943918  3971996 -7943992  2262432
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -12.7271  -0.5942  -0.1238   0.4654  22.5495
##
## Random effects:
## Groups      Name      Variance Std.Dev.
## basepath:name (Intercept) 0.0004585 0.02141
## name          (Intercept) 0.0003733 0.01932
## Residual                0.0017473 0.04180
## Number of obs: 2262437, groups: basepath:name, 209; name, 39
##
## Fixed effects:
##              Estimate Std. Error    df t value Pr(>|t|)
## (Intercept)      1.181e-01  3.772e-03 4.026e+01  31.31  <2e-16 ***
## membershipnon_member -7.419e-02  8.584e-04 2.262e+06 -86.42  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr)
## mmbbrshpnn_m -0.227
```

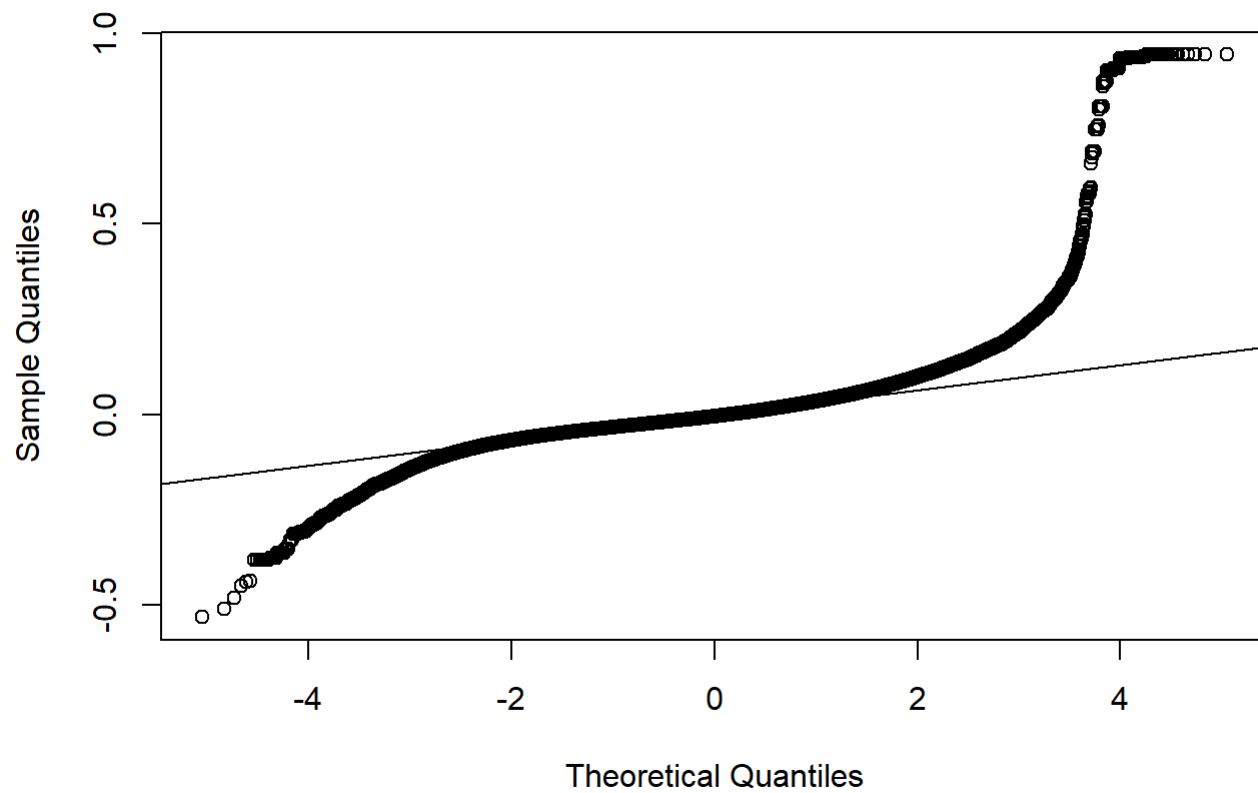
```
plot(m1)
```



```
qqnorm(resid(m1))  
qqline(resid(m1))
```



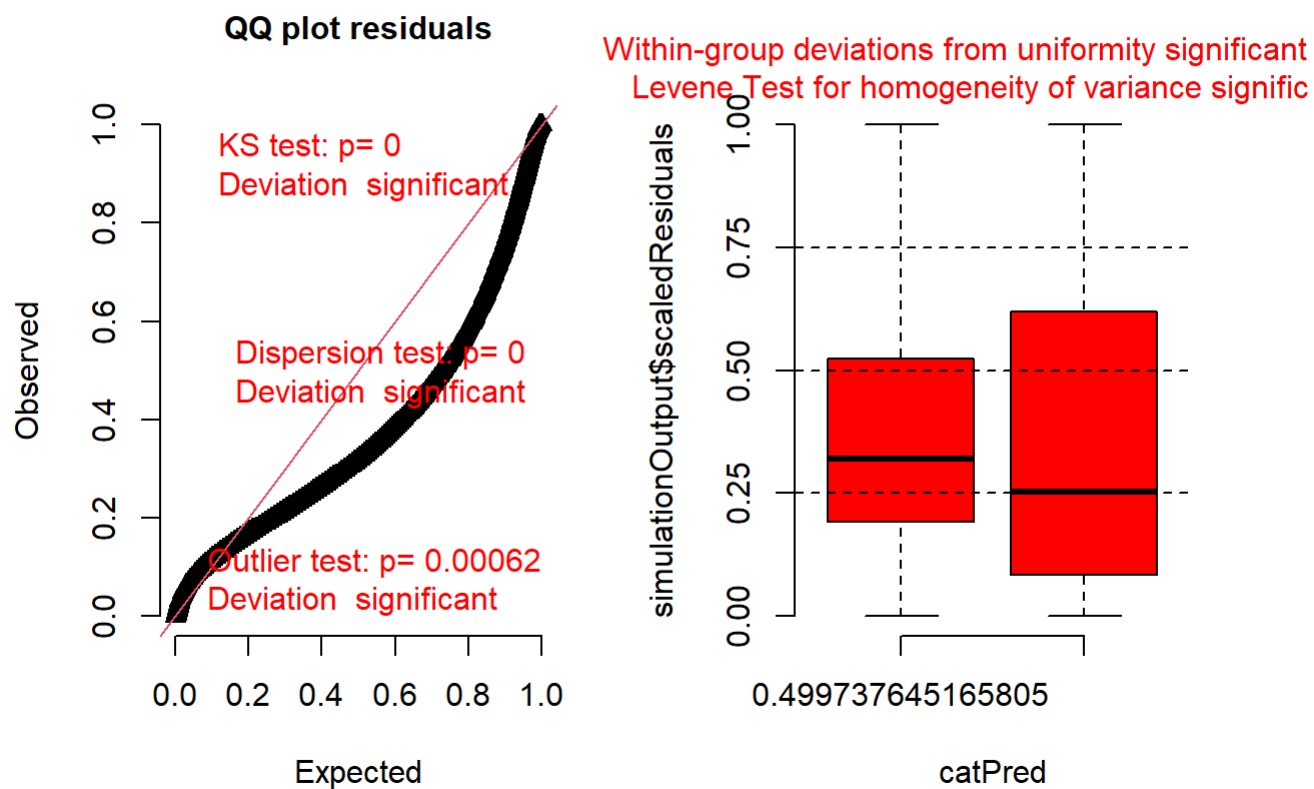
## Normal Q-Q Plot



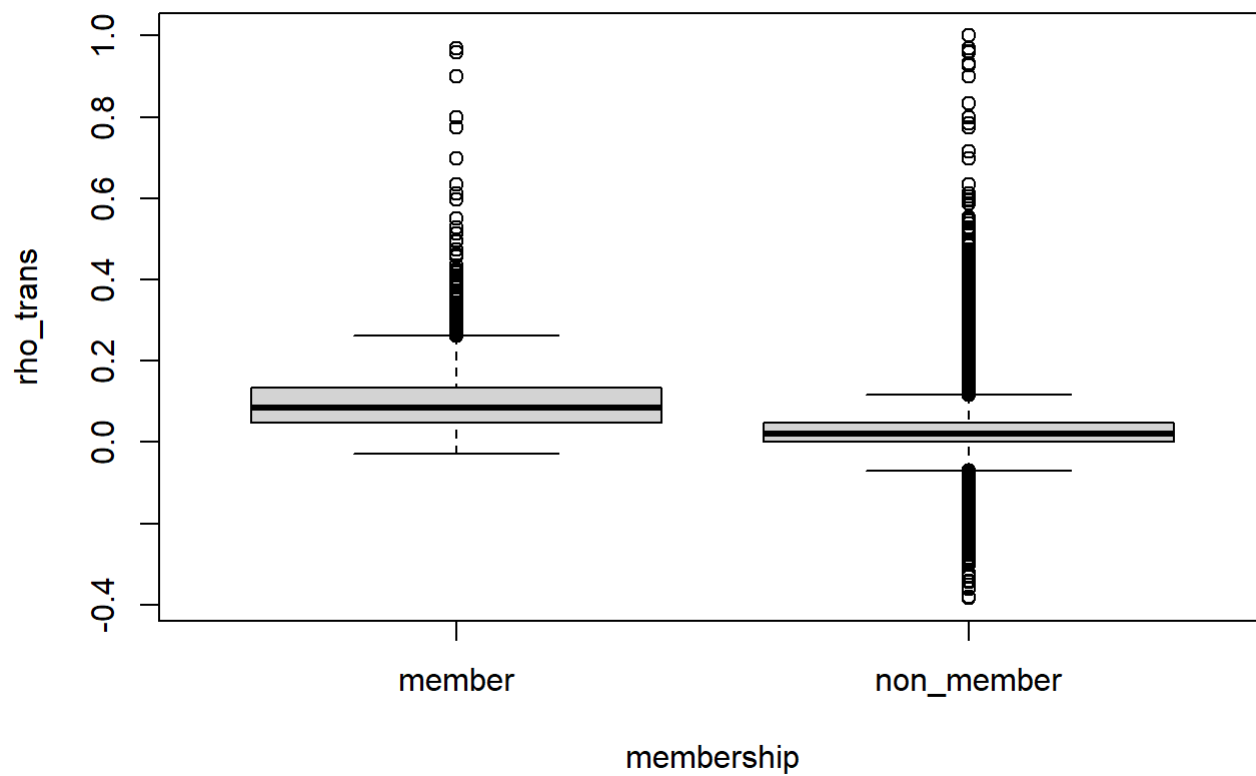
```
simulationOutput <- simulateResiduals(fittedModel = m1, plot = F)
```

```
plot(simulationOutput)
```

## DHARMa residual



```
plot(rho_trans ~ membership, data = data)
```



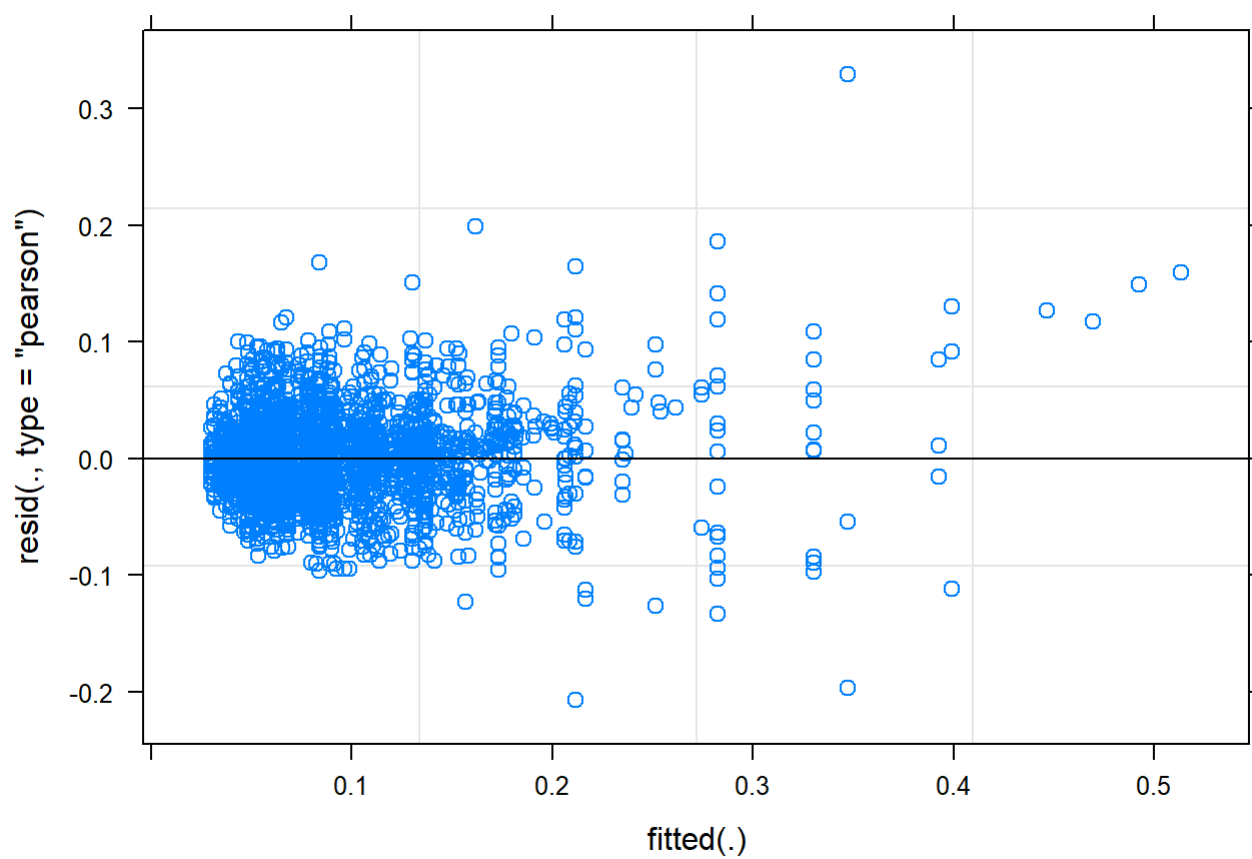
```
data_2 = data[!data$simple_effect == "unknown",]

data_2$rho_trans = log(data_2$rho+1)

m1 = lmer(rho_trans ~ simple_effect + (1 | name/basepath/assembly_id), data = data_2, REML=FALSE)
summary(m1)
```

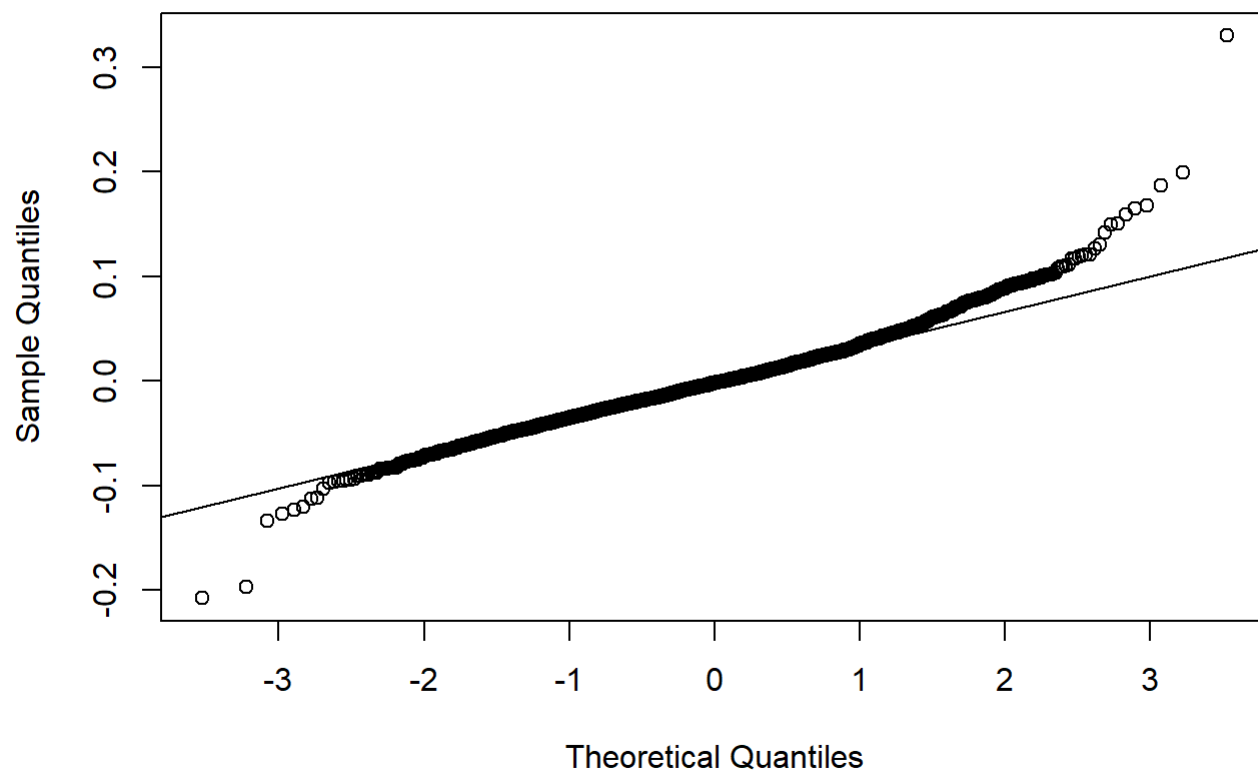
```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
## method [lmerModLmerTest]
## Formula: rho_trans ~ simple_effect + (1 | name/basepath/assembly_id)
## Data: data_2
##
##      AIC      BIC    logLik deviance df.resid
## -7536.6 -7501.9   3774.3  -7548.6     2368
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -4.9421 -0.5771 -0.0408  0.5087  7.8535
##
## Random effects:
## Groups                Name            Variance Std.Dev.
## assembly_id:(basepath:name) (Intercept) 0.0036223 0.06019
## basepath:name              (Intercept) 0.0005893 0.02428
## name                       (Intercept) 0.0004600 0.02145
## Residual                   0.0017631 0.04199
## Number of obs: 2374, groups:
## assembly_id:(basepath:name), 354; basepath:name, 111; name, 33
##
## Fixed effects:
##
##              Estimate Std. Error      df t value
## (Intercept)    1.272e-01  6.500e-03 3.042e+01   19.57
## simple_effectmember_single_layer 4.884e-03  2.061e-03 2.112e+03    2.37
##
##              Pr(>|t|)
## (Intercept)    <2e-16 ***
## simple_effectmember_single_layer  0.0179 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr)
## smpl_ffct__ -0.232
```

```
plot(m1)
```

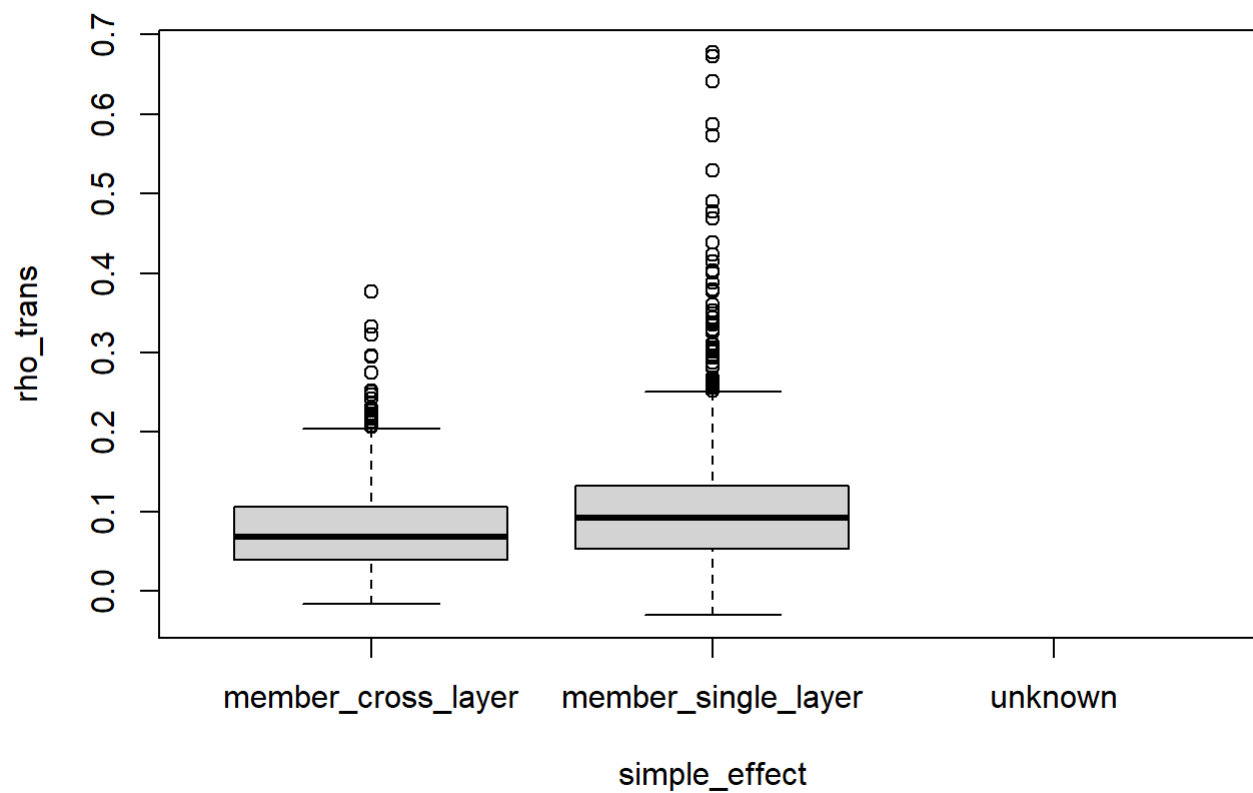


```
qqnorm(resid(m1))  
qqline(resid(m1))
```

## Normal Q-Q Plot



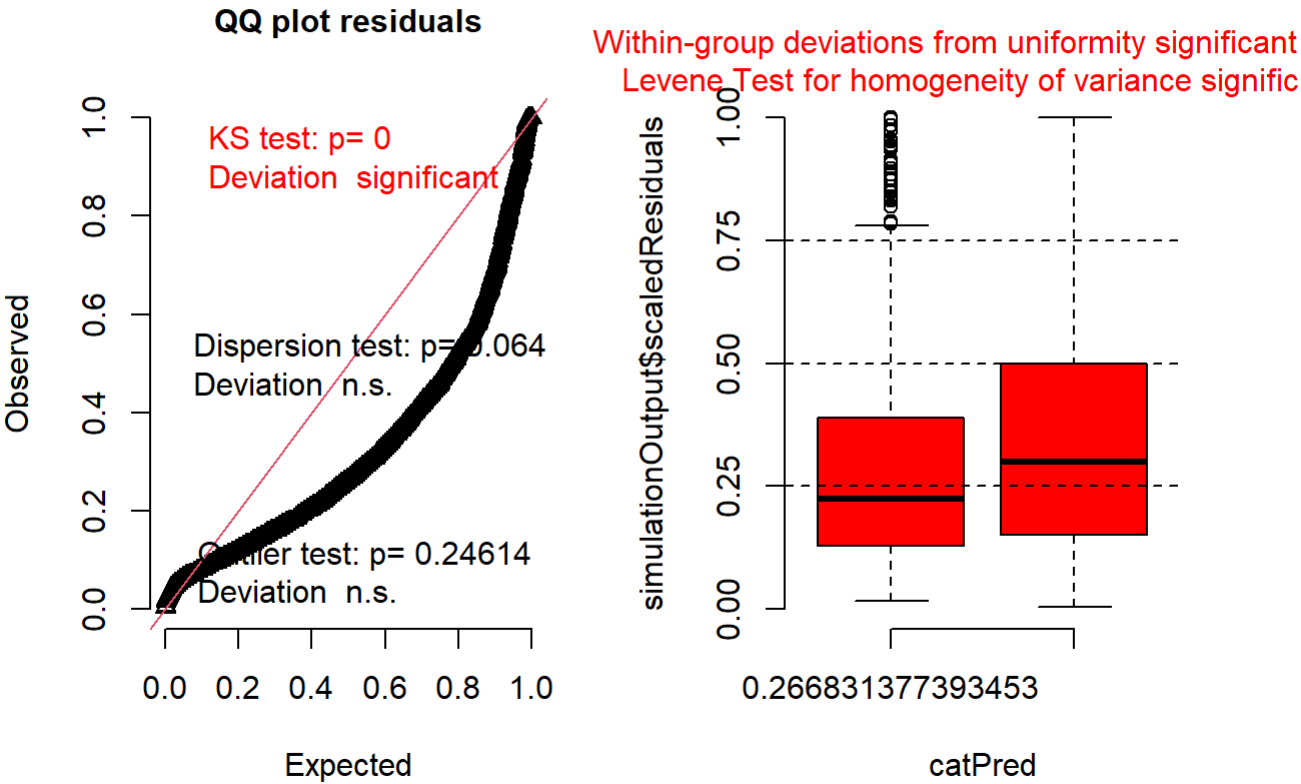
```
plot(rho_trans ~ simple_effect, data = data_2)
```



```
# testDispersion(m1)
simulationOutput <- simulateResiduals(fittedModel = m1, plot = F)
```

```
plot(simulationOutput)
```

DHARMa residual



```
unique(data_2$name)
```

```
## [1] AB1      AB3      AB4      AYA10    AYA4     AYA6     AYA7     AYA9
## [9] Rat08    Rat10    Rat11    Achilles Buddy    Cicero    Gatsby    ec013
## [17] ec014    ec016    f01_m    g01_m    i01_m    km01     nlx       OML18
## [25] OML19    OML5     OML7     OML8     OR15     OR18     OR21     OR22
## [33] OR23
## 39 Levels: 2006-4-10 2006-6-12 2006-6-7 AB1 AB3 AB4 Achilles AYA10 ... Rat11
```

```
unique(data_2$basepath)
```



```
## [1] Z:\\Data\\AYAoId\\AB1\\day1
## [2] Z:\\Data\\AYAoId\\AB3\\AB3_38_41
## [3] Z:\\Data\\AYAoId\\AB3\\AB3_42_46
## [4] Z:\\Data\\AYAoId\\AB3\\AB3_47_49
## [5] Z:\\Data\\AYAoId\\AB3\\AB3_50_51
## [6] Z:\\Data\\AYAoId\\AB3\\AB3_55_57
## [7] Z:\\Data\\AYAoId\\AB3\\AB3_58_59
## [8] Z:\\Data\\AYAoId\\AB3\\AB3_60
## [9] Z:\\Data\\AYAoId\\AB4\\day03
## [10] Z:\\Data\\AYAoId\\AB4\\day07
## [11] Z:\\Data\\AYAoId\\AB4\\day09
## [12] Z:\\Data\\AYAoId\\AB4\\day11
## [13] Z:\\Data\\AYAoId\\AYA10\\day25
## [14] Z:\\Data\\AYAoId\\AYA10\\day27
## [15] Z:\\Data\\AYAoId\\AYA10\\day32
## [16] Z:\\Data\\AYAoId\\AYA10\\day34
## [17] Z:\\Data\\AYAoId\\AYA4\\day150726
## [18] Z:\\Data\\AYAoId\\AYA4\\day150728
## [19] Z:\\Data\\AYAoId\\AYA6\\day17
## [20] Z:\\Data\\AYAoId\\AYA6\\day19
## [21] Z:\\Data\\AYAoId\\AYA6\\day20
## [22] Z:\\Data\\AYAoId\\AYA7\\day19
## [23] Z:\\Data\\AYAoId\\AYA7\\day20
## [24] Z:\\Data\\AYAoId\\AYA7\\day22
## [25] Z:\\Data\\AYAoId\\AYA7\\day24
## [26] Z:\\Data\\AYAoId\\AYA7\\day25
## [27] Z:\\Data\\AYAoId\\AYA7\\day27
## [28] Z:\\Data\\AYAoId\\AYA7\\day30
## [29] Z:\\Data\\AYAoId\\AYA9\\day12
## [30] Z:\\Data\\AYAoId\\AYA9\\day15
## [31] Z:\\Data\\AYAoId\\AYA9\\day16
## [32] Z:\\Data\\AYAoId\\AYA9\\day17
## [33] Z:\\Data\\AYAoId\\AYA9\\day20
## [34] Z:\\Data\\GirardeauG\\Rat08\\Rat08-20130709
## [35] Z:\\Data\\GirardeauG\\Rat08\\Rat08-20130712
## [36] Z:\\Data\\GirardeauG\\Rat08\\Rat08-20130713
## [37] Z:\\Data\\GirardeauG\\Rat08\\Rat08-20130717
## [38] Z:\\Data\\GirardeauG\\Rat08\\Rat08-20130718
## [39] Z:\\Data\\GirardeauG\\Rat10\\Rat10-20140619
## [40] Z:\\Data\\GirardeauG\\Rat10\\Rat10-20140620
## [41] Z:\\Data\\GirardeauG\\Rat10\\Rat10-20140622
## [42] Z:\\Data\\GirardeauG\\Rat10\\Rat10-20140624
## [43] Z:\\Data\\GirardeauG\\Rat10\\Rat10-20140704
## [44] Z:\\Data\\GirardeauG\\Rat10\\Rat10-20140705
## [45] Z:\\Data\\GirardeauG\\Rat11\\Rat11-20150313
## [46] Z:\\Data\\GirardeauG\\Rat11\\Rat11-20150316
## [47] Z:\\Data\\GirardeauG\\Rat11\\Rat11-20150317
## [48] Z:\\Data\\GirardeauG\\Rat11\\Rat11-20150327
## [49] Z:\\Data\\GirardeauG\\Rat11\\Rat11-20150328
## [50] Z:\\Data\\GirardeauG\\Rat11\\Rat11-20150330
## [51] Z:\\Data\\GirardeauG\\Rat11\\Rat11-20150331
## [52] Z:\\Data\\GrosmarkAD\\Achilles\\Achilles_10252013
```

```
## [53] Z:\\Data\\GrosmarkAD\\Achilles\\Achilles_11012013
## [54] Z:\\Data\\GrosmarkAD\\Buddy\\Buddy_06272013
## [55] Z:\\Data\\GrosmarkAD\\Cicero\\Cicero_09012014
## [56] Z:\\Data\\GrosmarkAD\\Cicero\\Cicero_09102014
## [57] Z:\\Data\\GrosmarkAD\\Cicero\\Cicero_09172014
## [58] Z:\\Data\\GrosmarkAD\\Gatsby\\Gatsby_08022013
## [59] Z:\\Data\\GrosmarkAD\\Gatsby\\Gatsby_08282013
## [60] Z:\\Data\\Kenji\\ec013.771_777
## [61] Z:\\Data\\Kenji\\ec014.171_188
## [62] Z:\\Data\\Kenji\\ec014.192_204
## [63] Z:\\Data\\Kenji\\ec014.207_234
## [64] Z:\\Data\\Kenji\\ec014.254_269
## [65] Z:\\Data\\Kenji\\ec014.271_287
## [66] Z:\\Data\\Kenji\\ec014.329_340
## [67] Z:\\Data\\Kenji\\ec014.345_366
## [68] Z:\\Data\\Kenji\\ec014.427_456
## [69] Z:\\Data\\Kenji\\ec014.459_480
## [70] Z:\\Data\\Kenji\\ec016.228_240
## [71] Z:\\Data\\Kenji\\ec016.267_278
## [72] Z:\\Data\\Kenji\\ec016.281_298
## [73] Z:\\Data\\Kenji\\ec016.390_405
## [74] Z:\\Data\\Kenji\\ec016.425_437
## [75] Z:\\Data\\Kenji\\ec016.479_487
## [76] Z:\\Data\\Kenji\\ec016.491_508
## [77] Z:\\Data\\Kenji\\ec016.532_540
## [78] Z:\\Data\\Kenji\\ec016.577_590
## [79] Z:\\Data\\Kenji\\f01_maze09_MS.001_003
## [80] Z:\\Data\\Kenji\\g01_maze05_MS.001_002
## [81] Z:\\Data\\Kenji\\g01_maze11_MS.001_004
## [82] Z:\\Data\\Kenji\\i01_maze05_MS.001_003
## [83] Z:\\Data\\Kenji\\i01_maze15_MS.001_004
## [84] Z:\\Data\\Kenji\\km01.004_011
## [85] Z:\\Data\\Kenji\\km01.012_020
## [86] Z:\\Data\\Kenji\\nlx_070329
## [87] Z:\\Data\\OMLproject\\OML18\\day1
## [88] Z:\\Data\\OMLproject\\OML18\\day4
## [89] Z:\\Data\\OMLproject\\OML18\\day5
## [90] Z:\\Data\\OMLproject\\OML19\\day2
## [91] Z:\\Data\\OMLproject\\OML5\\day12
## [92] Z:\\Data\\OMLproject\\OML7\\day9
## [93] Z:\\Data\\OMLproject\\OML8\\day4
## [94] Z:\\Data\\OMLproject\\OML8\\day6
## [95] Z:\\Data\\OMLproject\\OML8\\day8
## [96] Z:\\Data\\OMLproject\\OML8\\day9
## [97] Z:\\Data\\ORproject\\OR15\\day1
## [98] Z:\\Data\\ORproject\\OR15\\day10
## [99] Z:\\Data\\ORproject\\OR15\\day2
## [100] Z:\\Data\\ORproject\\OR15\\day3
## [101] Z:\\Data\\ORproject\\OR15\\day4
## [102] Z:\\Data\\ORproject\\OR18\\day1
## [103] Z:\\Data\\ORproject\\OR18\\day3
## [104] Z:\\Data\\ORproject\\OR21\\day2
```

```
## [105] Z:\\Data\\ORproject\\OR21\\day4
## [106] Z:\\Data\\ORproject\\OR22\\day1
## [107] Z:\\Data\\ORproject\\OR22\\day3
## [108] Z:\\Data\\ORproject\\OR22\\day4
## [109] Z:\\Data\\ORproject\\OR22\\day5
## [110] Z:\\Data\\ORproject\\OR23\\day1
## [111] Z:\\Data\\ORproject\\OR23\\day5
## 209 Levels: Z:\\Data\\AYAold\\AB1\\day1 ... Z:\\Data\\ORproject\\OR23\\day5
```