

Alpha diversity ITS2 DINO analyses for Pocillopora species across the Indo-Pacific, 1 Nov 2021

```
rm(list = ls())  
library(tidyr)  
library(purrr)  
library(dplyr)
```

```
##  
## Attaching package: 'dplyr'  
  
## The following objects are masked from 'package:stats':  
##  
##   filter, lag  
  
## The following objects are masked from 'package:base':  
##  
##   intersect, setdiff, setequal, union
```

```
library(metagMisc)
```

```
##  
## Attaching package: 'metagMisc'  
  
## The following object is masked from 'package:purrr':  
##  
##   some
```

```
library(kableExtra)
```

```
##  
## Attaching package: 'kableExtra'  
  
## The following object is masked from 'package:dplyr':  
##  
##   group_rows
```

```
library(reshape2)
```

```
##  
## Attaching package: 'reshape2'  
  
## The following object is masked from 'package:tidyr':  
##  
##   smiths
```

```

library(stringr)
library(tidyverse)

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

## v ggplot2 3.3.5      v readr  2.0.2
## v tibble  3.1.6      v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter()      masks stats::filter()
## x kableExtra::group_rows() masks dplyr::group_rows()
## x dplyr::lag()         masks stats::lag()
## x metagMisc::some()    masks purrr::some()

library(phyloseq)
library(magrittr)

##
## Attaching package: 'magrittr'

## The following object is masked from 'package:purrr':
##
##   set_names

## The following object is masked from 'package:tidyr':
##
##   extract

library(metagMisc)
library(randomForest)

## randomForest 4.6-14

## Type rfNews() to see new features/changes/bug fixes.

##
## Attaching package: 'randomForest'

## The following object is masked from 'package:ggplot2':
##
##   margin

## The following object is masked from 'package:dplyr':
##
##   combine

library(knitr)
library(seqinr)

```

```

##
## Attaching package: 'seqinr'

## The following object is masked from 'package:dplyr':
##
##      count

library(phyloseq)
library(ggplot2)
library(gridExtra)

##
## Attaching package: 'gridExtra'

## The following object is masked from 'package:randomForest':
##
##      combine

## The following object is masked from 'package:dplyr':
##
##      combine

library(vegan)

## Loading required package: permute

##
## Attaching package: 'permute'

## The following object is masked from 'package:seqinr':
##
##      getType

## Loading required package: lattice

## This is vegan 2.5-7

library(DESeq2)

## Loading required package: S4Vectors

## Loading required package: stats4

## Loading required package: BiocGenerics

## Loading required package: parallel

##
## Attaching package: 'BiocGenerics'

```

```

## The following objects are masked from 'package:parallel':
##
##   clusterApply, clusterApplyLB, clusterCall, clusterEvalQ,
##   clusterExport, clusterMap, parApply, parCapply, parLapply,
##   parLapplyLB, parRapply, parSapply, parSapplyLB

## The following object is masked from 'package:gridExtra':
##
##   combine

## The following object is masked from 'package:randomForest':
##
##   combine

## The following objects are masked from 'package:dplyr':
##
##   combine, intersect, setdiff, union

## The following objects are masked from 'package:stats':
##
##   IQR, mad, sd, var, xtabs

## The following objects are masked from 'package:base':
##
##   anyDuplicated, append, as.data.frame, basename, cbind, colnames,
##   dirname, do.call, duplicated, eval, evalq, Filter, Find, get, grep,
##   grepl, intersect, is.unsorted, lapply, Map, mapply, match, mget,
##   order, paste, pmax, pmax.int, pmin, pmin.int, Position, rank,
##   rbind, Reduce, rownames, sapply, setdiff, sort, table, tapply,
##   union, unique, unsplit, which.max, which.min

##
## Attaching package: 'S4Vectors'

## The following objects are masked from 'package:dplyr':
##
##   first, rename

## The following object is masked from 'package:tidyr':
##
##   expand

## The following objects are masked from 'package:base':
##
##   expand.grid, I, unname

## Loading required package: IRanges

##
## Attaching package: 'IRanges'

```

```

## The following object is masked from 'package:phyloseq':
##
##     distance

## The following objects are masked from 'package:dplyr':
##
##     collapse, desc, slice

## The following object is masked from 'package:purrr':
##
##     reduce

## Loading required package: GenomicRanges

## Loading required package: GenomeInfoDb

## Loading required package: SummarizedExperiment

## Loading required package: MatrixGenerics

## Loading required package: matrixStats

##
## Attaching package: 'matrixStats'

## The following object is masked from 'package:seqinr':
##
##     count

## The following object is masked from 'package:dplyr':
##
##     count

##
## Attaching package: 'MatrixGenerics'

## The following objects are masked from 'package:matrixStats':
##
##     colAlls, colAnyNAs, colAnys, colAvgPerRowSet, colCollapse,
##     colCounts, colCummaxs, colCummins, colCumprods, colCumsums,
##     colDiffs, colIQRDiffs, colIQRs, colLogSumExps, colMadDiffs,
##     colMads, colMaxs, colMeans2, colMedians, colMins, colOrderStats,
##     colProds, colQuantiles, colRanges, colRanks, colSdDiffs, colSds,
##     colSums2, colTabulates, colVarDiffs, colVars, colWeightedMads,
##     colWeightedMeans, colWeightedMedians, colWeightedSds,
##     colWeightedVars, rowAlls, rowAnyNAs, rowAnys, rowAvgPerColSet,
##     rowCollapse, rowCounts, rowCummaxs, rowCummins, rowCumprods,
##     rowCumsums, rowDiffs, rowIQRDiffs, rowIQRs, rowLogSumExps,
##     rowMadDiffs, rowMads, rowMaxs, rowMeans2, rowMedians, rowMins,
##     rowOrderStats, rowProds, rowQuantiles, rowRanges, rowRanks,
##     rowSdDiffs, rowSds, rowSums2, rowTabulates, rowVarDiffs, rowVars,
##     rowWeightedMads, rowWeightedMeans, rowWeightedMedians,
##     rowWeightedSds, rowWeightedVars

```

```

## Loading required package: Biobase

## Welcome to Bioconductor
##
##     Vignettes contain introductory material; view with
##     'browseVignettes()'. To cite Bioconductor, see
##     'citation("Biobase")', and for packages 'citation("pkgname)".

##
## Attaching package: 'Biobase'

## The following object is masked from 'package:MatrixGenerics':
##
##     rowMedians

## The following objects are masked from 'package:matrixStats':
##
##     anyMissing, rowMedians

## The following object is masked from 'package:phyloseq':
##
##     sampleNames

library(phyloseq); packageVersion("phyloseq")

## [1] '1.36.0'

library(ggplot2); packageVersion("ggplot2")

## [1] '3.3.5'

library(picante)

## Loading required package: ape

##
## Attaching package: 'ape'

## The following objects are masked from 'package:seqinr':
##
##     as.alignment, consensus

## Loading required package: nlme

##
## Attaching package: 'nlme'

## The following object is masked from 'package:IRanges':
##
##     collapse

```

```

## The following object is masked from 'package:seqinr':
##
##     gls

## The following object is masked from 'package:dplyr':
##
##     collapse

library(ggrepel)
library(igraph)

##
## Attaching package: 'igraph'

## The following objects are masked from 'package:ape':
##
##     edges, mst, ring

## The following object is masked from 'package:GenomicRanges':
##
##     union

## The following object is masked from 'package:IRanges':
##
##     union

## The following object is masked from 'package:S4Vectors':
##
##     union

## The following objects are masked from 'package:BiocGenerics':
##
##     normalize, path, union

## The following object is masked from 'package:vegan':
##
##     diversity

## The following object is masked from 'package:permute':
##
##     permute

## The following object is masked from 'package:tibble':
##
##     as_data_frame

## The following objects are masked from 'package:dplyr':
##
##     as_data_frame, groups, union

```

```
## The following objects are masked from 'package:purrr':
##
##   compose, simplify

## The following object is masked from 'package:tidyr':
##
##   crossing

## The following objects are masked from 'package:stats':
##
##   decompose, spectrum

## The following object is masked from 'package:base':
##
##   union
```

```
library(plotly)
```

```
##
## Attaching package: 'plotly'

## The following object is masked from 'package:igraph':
##
##   groups

## The following object is masked from 'package:IRanges':
##
##   slice

## The following object is masked from 'package:S4Vectors':
##
##   rename

## The following object is masked from 'package:ggplot2':
##
##   last_plot

## The following object is masked from 'package:stats':
##
##   filter

## The following object is masked from 'package:graphics':
##
##   layout
```

```
library(ggtree)
```

```
## ggtree v3.0.4 For help: https://yulab-smu.top/treedata-book/
##
## If you use ggtree in published research, please cite the most appropriate paper(s):
##
## 1. Guangchuang Yu. Using ggtree to visualize data on tree-like structures. Current Protocols in Bioinformatics. 2018;
## 2. Guangchuang Yu, Tommy Tsan-Yuk Lam, Huachen Zhu, Yi Guan. Two methods for mapping and visualizing genomic data on phylogenetic trees. Molecular Biology and Evolution. 2019;
## 3. Guangchuang Yu, David Smith, Huachen Zhu, Yi Guan, Tommy Tsan-Yuk Lam. ggtree: an R package for visualizing phylogenetic trees. Bioinformatics. 2017;
```



```

##
## Attaching package: 'ggtree'

## The following object is masked from 'package:nlme':
##
##     collapse

## The following object is masked from 'package:ape':
##
##     rotate

## The following object is masked from 'package:IRanges':
##
##     collapse

## The following object is masked from 'package:S4Vectors':
##
##     expand

## The following object is masked from 'package:randomForest':
##
##     margin

## The following object is masked from 'package:magrittr':
##
##     inset

## The following object is masked from 'package:tidyr':
##
##     expand

library(ellipse)

##
## Attaching package: 'ellipse'

## The following object is masked from 'package:graphics':
##
##     pairs

library(dplyr)
library(indicspecies)

##
## Attaching package: 'indicspecies'

## The following object is masked from 'package:SummarizedExperiment':
##
##     coverage

```

```
## The following object is masked from 'package:GenomicRanges':  
##  
## coverage
```

```
## The following object is masked from 'package:IRanges':  
##  
## coverage
```

```
library(yhat)
```

```
## Registered S3 methods overwritten by 'yacca':  
## method from  
## plot.cca vegan  
## print.cca vegan  
## print.summary.cca vegan  
## summary.cca vegan
```

```
library("dunn.test"); packageVersion("dunn.test")
```

```
## [1] '1.3.5'
```

```
library(metagenomeSeq)
```

```
## Loading required package: limma
```

```
##  
## Attaching package: 'limma'
```

```
## The following object is masked from 'package:DESeq2':  
##  
## plotMA
```

```
## The following object is masked from 'package:BiocGenerics':  
##  
## plotMA
```

```
## The following object is masked from 'package:seqinr':  
##  
## zscore
```

```
## Loading required package: glmnet
```

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

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

```
## The following object is masked from 'package:ggtree':  
##  
## expand
```

```
## The following object is masked from 'package:S4Vectors':  
##  
##     expand
```

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

```
## Loaded glmnet 4.1-2
```

```
## Loading required package: RColorBrewer
```

```
library("multcompView")  
library(stats)  
library(emmeans)  
library(dbstats)
```

```
## Loading required package: cluster
```

```
## Loading required package: pls
```

```
##  
## Attaching package: 'pls'
```

```
## The following object is masked from 'package:ape':  
##  
##     mvr
```

```
## The following object is masked from 'package:vegan':  
##  
##     scores
```

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

```
library(geodist)  
library(radiant.data)
```

```
## Loading required package: lubridate
```

```
##  
## Attaching package: 'lubridate'
```

```
## The following objects are masked from 'package:igraph':  
##  
##     %--%, union
```

```
## The following objects are masked from 'package:GenomicRanges':  
##  
##     intersect, setdiff, union
```

```

## The following object is masked from 'package:GenomeInfoDb':
##
## intersect

## The following objects are masked from 'package:IRanges':
##
## %within%, intersect, setdiff, union

## The following objects are masked from 'package:S4Vectors':
##
## intersect, second, second<-, setdiff, union

## The following objects are masked from 'package:BiocGenerics':
##
## intersect, setdiff, union

## The following objects are masked from 'package:base':
##
## date, intersect, setdiff, union

##
## Attaching package: 'radiant.data'

## The following objects are masked from 'package:lubridate':
##
## month, wday

## The following object is masked from 'package:ggtree':
##
## flip

## The following objects are masked from 'package:plotly':
##
## ggplotly, subplot

## The following object is masked from 'package:igraph':
##
## normalize

## The following object is masked from 'package:picante':
##
## psd

## The following object is masked from 'package:BiocGenerics':
##
## normalize

## The following object is masked from 'package:magrittr':
##
## set_attr

```



```
##This CSVs has all N sequences removed, as not informative
```

```
## taxonomy table into matrix
taxmat<-as.matrix(taxtable[,2:3])
rownames(taxmat)<-taxtable$ASV
```

```
##Combine the taxonomy matrix and the otu_table (otus) into a phyloseq object
```

```
TAX = tax_table(taxmat)
```

```
## Check if any duplicated row names
```

```
duplicated(TAX)
```

```
## asv1001 asv1074 asv1188 asv1235 asv1261 asv1341 asv137 asv148 asv1488 asv149
## FALSE FALSE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## asv1617 asv1660 asv1689 asv1997 asv2005 asv221 asv224 asv228 asv2280 asv2357
## TRUE TRUE FALSE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## asv2378 asv25 asv2553 asv2754 asv2830 asv3083 asv3798 asv4369 asv4370 asv448
## TRUE TRUE TRUE TRUE FALSE TRUE TRUE TRUE TRUE TRUE
## asv449 asv454 asv492 asv547 asv604 asv606 asv618 asv625 asv640 asv693
## TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## asv756 asv765 asv779 asv790 asv793 asv809 asv831 asv838 asv860 asv1
## TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE FALSE
## asv1005 asv101 asv1019 asv102 asv1023 asv1033 asv1037 asv1039 asv1042 asv1046
## FALSE TRUE FALSE TRUE TRUE TRUE FALSE TRUE TRUE TRUE
## asv105 asv1055 asv106 asv1060 asv1067 asv107 asv1078 asv108 asv1085 asv1090
## TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## asv1093 asv1098 asv1108 asv1110 asv112 asv1124 asv114 asv1144 asv1149 asv1153
## TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## asv1156 asv1158 asv1167 asv117 asv1172 asv1175 asv1185 asv119 asv12 asv120
## TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## asv1202 asv1205 asv1209 asv121 asv1216 asv1227 asv1228 asv123 asv1236 asv1239
## TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## asv1247 asv125 asv1252 asv1256 asv1265 asv127 asv128 asv1284 asv1285 asv1293
## TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## asv1295 asv1296 asv13 asv1302 asv1309 asv131 asv1310 asv132 asv133 asv134
## TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## asv135 asv136 asv1360 asv1363 asv1378 asv138 asv1389 asv139 asv1405 asv1408
## TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## asv1426 asv1436 asv144 asv1441 asv1448 asv1457 asv146 asv147 asv1476 asv1482
## TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## asv1497 asv15 asv150 asv1502 asv1505 asv1511 asv1517 asv1519 asv152 asv1523
## TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## asv153 asv154 asv1544 asv1545 asv155 asv1561 asv157 asv159 asv1594 asv1595
## TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## asv1599 asv160 asv1601 asv161 asv1610 asv1616 asv163 asv1633 asv164 asv1653
## TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## asv1668 asv167 asv1678 asv1686 asv1687 asv1690 asv1693 asv1696 asv17 asv1704
## TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## asv171 asv1711 asv1715 asv1719 asv1729 asv173 asv1737 asv175 asv176 asv1762
## TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## asv1768 asv177 asv178 asv1784 asv1786 asv179 asv1792 asv1799 asv18 asv180
```

##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv1809	asv1810	asv182	asv1827	asv1828	asv1843	asv185	asv1854	asv1858	asv1863
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv187	asv1870	asv1875	asv188	asv1886	asv1896	asv1897	asv19	asv1903	asv1904
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv1905	asv1906	asv191	asv1913	asv192	asv1920	asv1921	asv193	asv1938	asv194
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv1952	asv1960	asv1969	asv1970	asv1982	asv1983	asv1990	asv1995	asv2	asv200
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv2013	asv202	asv2023	asv204	asv2043	asv2049	asv2051	asv2052	asv206	asv2071
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv2075	asv2076	asv209	asv2095	asv211	asv2117	asv215	asv2151	asv2157	asv2158
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv218	asv2187	asv22	asv2226	asv223	asv2248	asv2255	asv226	asv2268	asv227
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv229	asv2309	asv232	asv2329	asv233	asv2330	asv2339	asv2340	asv2341	asv235
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv2356	asv236	asv2367	asv2368	asv2377	asv239	asv2399	asv24	asv240	asv2400
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv2402	asv2403	asv241	asv242	asv2427	asv243	asv2441	asv246	asv248	asv2486
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv2488	asv250	asv254	asv2552	asv2564	asv257	asv258	asv26	asv2600	asv2608
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv2609	asv2610	asv2622	asv2629	asv2631	asv2632	asv264	asv2648	asv2649	asv265
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv2662	asv2663	asv2665	asv267	asv2682	asv269	asv2698	asv2699	asv2701	asv271
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv2717	asv2718	asv272	asv273	asv2744	asv2755	asv276	asv280	asv2804	asv2805
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv2806	asv281	asv2815	asv2825	asv2826	asv2827	asv283	asv284	asv288	asv2881
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv29	asv290	asv2915	asv2916	asv2918	asv2938	asv295	asv296	asv2965	asv298
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv2987	asv3	asv30	asv300	asv3003	asv301	asv302	asv3033	asv3049	asv305
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv3068	asv31	asv310	asv3104	asv3106	asv3107	asv312	asv3134	asv3135	asv3138
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv314	asv3160	asv317	asv3180	asv3181	asv3200	asv3201	asv321	asv3217	asv3219
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv322	asv323	asv3245	asv3246	asv3247	asv3249	asv325	asv3251	asv328	asv33
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv3304	asv3306	asv3307	asv332	asv3329	asv3331	asv334	asv335	asv3357	asv3358
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv336	asv337	asv340	asv341	asv3413	asv3414	asv343	asv3442	asv3443	asv3444
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv347	asv3474	asv348	asv35	asv351	asv3512	asv3539	asv355	asv3577	asv3578
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv3580	asv3581	asv3582	asv3583	asv3584	asv3608	asv361	asv362	asv3638	asv3640
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv365	asv367	asv3670	asv368	asv3698	asv37	asv3705	asv371	asv375	asv3753
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv377	asv378	asv3795	asv3796	asv3799	asv38	asv381	asv3828	asv3829	asv383
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv3830	asv3831	asv3832	asv3871	asv388	asv39	asv3925	asv393	asv395	asv399

##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv4	asv4005	asv4006	asv402	asv404	asv4048	asv4049	asv405	asv406	asv4094
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv4095	asv4096	asv4097	asv41	asv410	asv411	asv414	asv4154	asv4157	asv4159
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv42	asv4204	asv4208	asv4209	asv423	asv424	asv4253	asv4254	asv426	asv427
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv428	asv43	asv4307	asv4308	asv4310	asv4311	asv433	asv4367	asv4368	asv4371
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv44	asv440	asv4428	asv4429	asv4430	asv4433	asv4434	asv444	asv445	asv446
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv45	asv450	asv46	asv460	asv462	asv468	asv47	asv471	asv472	asv475
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv476	asv477	asv479	asv48	asv486	asv490	asv491	asv493	asv495	asv498
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv499	asv50	asv500	asv505	asv507	asv51	asv514	asv516	asv519	asv52
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv53	asv531	asv532	asv535	asv537	asv54	asv540	asv543	asv554	asv556
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv56	asv561	asv562	asv563	asv565	asv587	asv60	asv605	asv607	asv609
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv611	asv613	asv63	asv630	asv632	asv633	asv638	asv649	asv653	asv654
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv657	asv660	asv661	asv666	asv67	asv68	asv691	asv699	asv70	asv700
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv703	asv706	asv71	asv714	asv715	asv718	asv73	asv732	asv735	asv742
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv746	asv75	asv752	asv753	asv76	asv764	asv77	asv770	asv772	asv774
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv777	asv78	asv784	asv787	asv788	asv789	asv79	asv792	asv799	asv8
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv80	asv801	asv806	asv807	asv808	asv818	asv82	asv825	asv827	asv83
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv834	asv837	asv85	asv859	asv86	asv863	asv864	asv87	asv873	asv874
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv877	asv88	asv884	asv887	asv89	asv899	asv9	asv90	asv902	asv905
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv911	asv914	asv916	asv92	asv922	asv925	asv929	asv934	asv94	asv940
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv943	asv945	asv946	asv948	asv950	asv956	asv961	asv967	asv968	asv970
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv976	asv977	asv985	asv989	asv992	asv999	asv1031	asv104	asv1049	asv1050
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	TRUE
##	asv109	asv110	asv111	asv1131	asv1171	asv1194	asv1203	asv1204	asv1221	asv1258
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv1286	asv129	asv1319	asv1342	asv1354	asv1359	asv1364	asv1391	asv14	asv142
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv1420	asv145	asv1460	asv1465	asv1466	asv1471	asv1475	asv1498	asv1532	asv1533
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv156	asv1581	asv1589	asv16	asv1600	asv1609	asv166	asv1661	asv1667	asv1738
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv1749	asv1757	asv1763	asv1777	asv1778	asv1785	asv1791	asv1805	asv183	asv1852
##	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	asv1853	asv1864	asv1874	asv1879	asv1931	asv1948	asv1973	asv201	asv2026	asv2032


```
##      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE
## asv2044 asv205 asv2063 asv2105 asv2124 asv2150 asv2169 asv2213 asv2290 asv23
##      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE
## asv231 asv2310 asv2312 asv2328 asv237 asv238 asv2401 asv245 asv2487 asv249
##      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      FALSE      TRUE      TRUE
## asv2523 asv2524 asv2525 asv259 asv2599 asv260 asv2621 asv2630 asv2664 asv27
##      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE
## asv270 asv2700 asv279 asv2828 asv2829 asv286 asv2862 asv293 asv2937 asv294
##      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE
## asv297 asv2985 asv2986 asv3008 asv3084 asv3105 asv3136 asv3137 asv318 asv3248
##      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE
## asv327 asv3279 asv3305 asv3328 asv345 asv357 asv3579 asv359 asv3639 asv3702
##      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE
## asv3797 asv390 asv3926 asv396 asv40 asv4003 asv4004 asv415 asv4155 asv4156
##      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE
## asv417 asv422 asv429 asv4309 asv431 asv437 asv4431 asv4432 asv447 asv453
##      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE
## asv459 asv463 asv467 asv474 asv49 asv501 asv504 asv529 asv533 asv538
##      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE
## asv542 asv55 asv553 asv558 asv559 asv567 asv571 asv574 asv575 asv585
##      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE
## asv598 asv6 asv62 asv639 asv66 asv674 asv675 asv680 asv682 asv7
##      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE
## asv702 asv730 asv738 asv74 asv767 asv780 asv81 asv816 asv820 asv822
##      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE
## asv836 asv84 asv848 asv857 asv862 asv904 asv91 asv920 asv95 asv952
##      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE
## asv955 asv984 asv1059 asv1305 asv2311 asv2342 asv3330 asv385 asv461 asv489
##      TRUE      TRUE      FALSE      TRUE      TRUE      FALSE      TRUE      TRUE      TRUE      TRUE      TRUE
## asv791 asv928
##      TRUE      TRUE
```

```
taxa_are_rows(TAX)
```

```
## NULL
```

```
any(duplicated(rownames(TAX)))
```

```
## [1] FALSE
```

```
which(duplicated(rownames(TAX)))
```

```
## integer(0)
```

```
## If no duplicated row names:
```

```
ps_tax = phyloseq(otu, TAX, sam, tre)
```

```
## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
```

```
## Also defined by 'tidytree'
```

```
## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

## Also defined by 'tidytree'
```

```
ps_tax
```

```
## phyloseq-class experiment-level object
## otu_table() OTU Table: [ 882 taxa and 416 samples ]
## sample_data() Sample Data: [ 416 samples by 21 sample variables ]
## tax_table() Taxonomy Table: [ 882 taxa by 2 taxonomic ranks ]
## phy_tree() Phylogenetic Tree: [ 882 tips and 881 internal nodes ]
```

```
# Compare phyloseq's raw and now SymPortal amended ps's
ps_tax
```

```
## phyloseq-class experiment-level object
## otu_table() OTU Table: [ 882 taxa and 416 samples ]
## sample_data() Sample Data: [ 416 samples by 21 sample variables ]
## tax_table() Taxonomy Table: [ 882 taxa by 2 taxonomic ranks ]
## phy_tree() Phylogenetic Tree: [ 882 tips and 881 internal nodes ]
```

```
ps
```

```
## phyloseq-class experiment-level object
## otu_table() OTU Table: [ 11374 taxa and 416 samples ]
## sample_data() Sample Data: [ 416 samples by 21 sample variables ]
## phy_tree() Phylogenetic Tree: [ 11374 tips and 11372 internal nodes ]
```

```
## Only 8% of taxa kept from ps to ps_tax
```

```
#Remove samples less 1000 reads
```

```
ps1 = prune_samples(sample_sums(ps_tax) > 1000, ps_tax)
ps1
```

```
## phyloseq-class experiment-level object
## otu_table() OTU Table: [ 882 taxa and 343 samples ]
## sample_data() Sample Data: [ 343 samples by 21 sample variables ]
## tax_table() Taxonomy Table: [ 882 taxa by 2 taxonomic ranks ]
## phy_tree() Phylogenetic Tree: [ 882 tips and 881 internal nodes ]
```

```
#Remove taxa not seen more than 1 times in at least 5% of the samples
```

```
ps2 = filter_taxa(ps1, function(x) sum(x > 1) > (0.05*length(x)), TRUE)
```

```
## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
```

```
## Also defined by 'tidytree'
```

```
ps2
```

```
## phyloseq-class experiment-level object
## otu_table() OTU Table: [ 435 taxa and 343 samples ]
## sample_data() Sample Data: [ 343 samples by 21 sample variables ]
## tax_table() Taxonomy Table: [ 435 taxa by 2 taxonomic ranks ]
## phy_tree() Phylogenetic Tree: [ 435 tips and 434 internal nodes ]
```

```
# keep only taxa that were observed at least twice
```

```
ps3 = prune_taxa(taxa_sums(ps2) >= 2, ps2)
```

```
## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
```

```
## Also defined by 'tidytree'
```

```
ps3
```

```
## phyloseq-class experiment-level object
## otu_table() OTU Table: [ 435 taxa and 343 samples ]
## sample_data() Sample Data: [ 343 samples by 21 sample variables ]
## tax_table() Taxonomy Table: [ 435 taxa by 2 taxonomic ranks ]
## phy_tree() Phylogenetic Tree: [ 435 tips and 434 internal nodes ]
```

```
# remove taxonomy samples that were Ns
```

```
ps4 = subset_samples(ps3, Spec != "P_spp")
```

```
## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
```

```
## Also defined by 'tidytree'
```

```

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

## Also defined by 'tidytree'

## Remove Fr Poly as ASVs all = 0
ps5 = subset_samples(ps4, Loc != "FrenPoly")

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

## Also defined by 'tidytree'

## Assign ASV numeric values to replace sequences
##n_seqs <- seq(ntaxa(ps5))
##len_n_seqs <- nchar(max(n_seqs))
##taxa_seqs <- taxa_names(ps5)
##asvs <- paste("ASV", formatC(n_seqs,
##width = len_n_seqs,
##flag = "0"), sep = "_")
##taxa_names(ps5) <- asvs

## remove samples with otu = 0
ps6 <- prune_samples(sample_sums(ps5) >= 1, ps5)
ps6

## phyloseq-class experiment-level object
## otu_table() OTU Table: [ 435 taxa and 338 samples ]
## sample_data() Sample Data: [ 338 samples by 21 sample variables ]
## tax_table() Taxonomy Table: [ 435 taxa by 2 taxonomic ranks ]
## phy_tree() Phylogenetic Tree: [ 435 tips and 434 internal nodes ]

```

```
##dfASV_seq <- data.frame(asv=asvs, seq=taxa_seqs, stringsAsFactors = FALSE)
##write.csv(dfASV_seq, file="dfASV_seq.csv", row.names = FALSE)
##write.fasta(as.list(taxa_seqs), asvs, "asv_seq.fasta", open = "w", nbchar = 60, as.string =TRUE)

##saveRDS(dfASV_seq, file="dfASV_Seq.rds")
```

```
## Remove singletons
ps6_filt <- filter_taxa(ps6, function(x) sum(x > 1) > 1, TRUE)
```

Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

Also defined by 'tidytree'

```
ps6_filt
```

```
## phyloseq-class experiment-level object
## otu_table() OTU Table: [ 435 taxa and 338 samples ]
## sample_data() Sample Data: [ 338 samples by 21 sample variables ]
## tax_table() Taxonomy Table: [ 435 taxa by 2 taxonomic ranks ]
## phy_tree() Phylogenetic Tree: [ 435 tips and 434 internal nodes ]
```

```
## CSS transformation
```

```
ps6_filt_css <- phyloseq_transform_css(ps6_filt, norm = TRUE, log = FALSE)
```

Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

Also defined by 'tidytree'

Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

Also defined by 'tidytree'

Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

Also defined by 'tidytree'

Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

Also defined by 'tidytree'

```
ps_normalized <- list()
normalization <- 'css'
ps_normalized[[normalization]] <- ps6_filt_css
ps_normalized[[normalization]]
```

```
## phyloseq-class experiment-level object
## otu_table() OTU Table: [ 435 taxa and 338 samples ]
## sample_data() Sample Data: [ 338 samples by 21 sample variables ]
## tax_table() Taxonomy Table: [ 435 taxa by 2 taxonomic ranks ]
## phy_tree() Phylogenetic Tree: [ 435 tips and 434 internal nodes ]
```

```
colnames(sample_data(ps6_filt_css))
```

```
## [1] "Loc"      "Yr"      "Spec"    "Exp_cond" "Code"
## [6] "Repro"    "Month"   "Season"  "S_region" "L_region"
## [11] "Exact.date" "Tbl_bin" "T_bleach" "DHW"      "DHW_cat"
## [16] "SST_a"    "Coord_X" "Coord_Y" "Primer"   "Pub"
## [21] "Note"
```

```
saveRDS(ps6_filt_css, file="ps6_filt_css.rds")
```

```
##Binary transformation, did not use
```

```
ps_filt_binary <- ps6_filt_css
otups2 <- otu_table(ps_filt_binary)
```

```
# if >= 1, then transform value to 1
otups2[otups2 >= 1] <- 1
otu_table(ps_filt_binary) <- otups2
```

```
## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
```

```
## Also defined by 'tidytree'
```

```
## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
```

```
## Also defined by 'tidytree'
```

```
## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
```

```
## Also defined by 'tidytree'
```

```
## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
```

```
## Also defined by 'tidytree'
```

```
input_data <- t(as(otu_table(ps_filt_binary), "matrix"))
sampleData <- sample_data(ps_filt_binary)
all_data <- merge(sampleData, input_data, by = 0)
rownames(all_data) <- all_data$Row.names
all_data <- subset(all_data, select = -c(Row.names))
all_nonNA_data <- all_data[,colSums(is.na(all_data))<nrow(all_data)]
all_nonNA_data$yrSeason <- as.factor(all_nonNA_data$Loc)
kable(head(all_nonNA_data[,c(1:8)]))
```

	Loc	Yr	Spec	Exp_cond	Code	Repro	Month	Season
SRR5963022	NewCal	2014	P_dam	30C	NC2	B	Nov	Spring
SRR5963023	NewCal	2014	P_dam	30C	NC1	B	Nov	Spring
SRR5963024	Oman	2014	P_dam	31C	Om2	B	June	Winter
SRR5963025	Oman	2014	P_dam	31C	Om2	B	June	Winter
SRR5963026	Oman	2014	P_dam	31C	Om3	B	June	Winter
SRR5963027	Oman	2014	P_dam	31C	Om2	B	June	Winter

```
tax_info <- as.data.frame(tax_table(ps_filt_binary))
saveRDS(ps_filt_binary, file="ps_filt_binary.rds")
normalization <- 'binary'
```

```
ps_normalized[[normalization]] <- ps_filt_binary
ps_normalized[[normalization]]
```

```
## phyloseq-class experiment-level object
## otu_table() OTU Table: [ 435 taxa and 338 samples ]
## sample_data() Sample Data: [ 338 samples by 21 sample variables ]
## tax_table() Taxonomy Table: [ 435 taxa by 2 taxonomic ranks ]
## phy_tree() Phylogenetic Tree: [ 435 tips and 434 internal nodes ]
```

```
saveRDS(ps_normalized, file="ps_normalized.rds")

tax_info <- as.data.frame(tax_table(ps6_filt_css))
write.csv(tax_info, file="tax_info.csv")
```

```
## Set random seed for reproducibility
```

```
set.seed(8765)
```

```
##Alpha div
```

```
#Remove Panama
```

```
ps7 <- subset_samples(ps6_filt_css, Loc != "Panam")
```

```
## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
```

```
## Also defined by 'tidytree'
```

```
## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
```

```
## Also defined by 'tidytree'
```

```
## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
```

```
## Also defined by 'tidytree'
```

```
## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
```

```
## Also defined by 'tidytree'
```

```
#Remove P. me
```

```
ps8 <- subset_samples(ps7, Spec != "P_me")
```

```
## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
```

```
## Also defined by 'tidytree'
```

```
## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
```

```

## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

## Also defined by 'tidytree'

#Keep only P.dam
ps9 = subset_samples(ps8, Spec == "P_dam")

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

## Also defined by 'tidytree'

ps9

## phyloseq-class experiment-level object
## otu_table() OTU Table: [ 435 taxa and 187 samples ]
## sample_data() Sample Data: [ 187 samples by 21 sample variables ]
## tax_table() Taxonomy Table: [ 435 taxa by 2 taxonomic ranks ]
## phy_tree() Phylogenetic Tree: [ 435 tips and 434 internal nodes ]

colnames(sample_data(ps9))

## [1] "Loc" "Yr" "Spec" "Exp_cond" "Code"
## [6] "Repro" "Month" "Season" "S_region" "L_region"
## [11] "Exact.date" "Tbl_bin" "T_bleach" "DHW" "DHW_cat"
## [16] "SST_a" "Coord_X" "Coord_Y" "Primer" "Pub"
## [21] "Note"

head(sample_data(ps9))

```



```
##          Loc   Yr  Spec Exp_cond Code  Repro Month Season S_region L_region
## SRR5963022 NewCal 2014 P_dam    30C  NC2    B   Nov Spring   NewCal   WPac
## SRR5963023 NewCal 2014 P_dam    30C  NC1    B   Nov Spring   NewCal   WPac
## SRR5963024 Oman 2014 P_dam    31C  Om2    B   June Winter IndianOc IndianOc
## SRR5963025 Oman 2014 P_dam    31C  Om2    B   June Winter IndianOc IndianOc
## SRR5963026 Oman 2014 P_dam    31C  Om3    B   June Winter IndianOc IndianOc
## SRR5963027 Oman 2014 P_dam    31C  Om2    B   June Winter IndianOc IndianOc
##          Exact.date Tbl_bin T_bleach DHW DHW_cat SST_a Coord_X Coord_Y
## SRR5963022      <NA>    Long    15y 0.00      N  27.1  -22.29  166.43
## SRR5963023      <NA>    Long    15y 0.00      N  27.1  -22.30  166.43
## SRR5963024      <NA>    Long    15y 4.24    Mod  30.8   23.52   58.74
## SRR5963025      <NA>    Long    15y 4.24    Mod  30.8   23.52   58.74
## SRR5963026      <NA>    Long    15y 3.79    Mod  30.8   23.62   58.60
## SRR5963027      <NA>    Long    15y 4.24    Mod  30.8   23.52   58.74
##          Primer
## SRR5963022 ITS-DINO https://www.biorxiv.org/content/10.1101/398602v4.full.pdf
## SRR5963023 ITS-DINO https://www.biorxiv.org/content/10.1101/398602v4.full.pdf
## SRR5963024 ITS-DINO https://www.biorxiv.org/content/10.1101/398602v4.full.pdf
## SRR5963025 ITS-DINO https://www.biorxiv.org/content/10.1101/398602v4.full.pdf
## SRR5963026 ITS-DINO https://www.biorxiv.org/content/10.1101/398602v4.full.pdf
## SRR5963027 ITS-DINO https://www.biorxiv.org/content/10.1101/398602v4.full.pdf
##
## SRR5963022 Colonies said to be "Pocillopora damicornis-like"; based on ORF and microsatellites, all
## SRR5963023 Colonies said to be "Pocillopora damicornis-like"; based on ORF and microsatellites, all
## SRR5963024 Colonies said to be "Pocillopora damicornis-like"; based on ORF and microsatellites, all
## SRR5963025 Colonies said to be "Pocillopora damicornis-like"; based on ORF and microsatellites, all
## SRR5963026 Colonies said to be "Pocillopora damicornis-like"; based on ORF and microsatellites, all
## SRR5963027 Colonies said to be "Pocillopora damicornis-like"; based on ORF and microsatellites, all
```

#Move into Vegan

```
set.seed(423542)
asv_css <- t(otu_table(ps9))
asv_css_hell <- decostand((asv_css), "hell") #not sure we are going to do this.
meta = as(sample_data(ps9), "data.frame")
```

```
Richness <- specnumber(asv_css) #calculates richness from css
Shannon <- vegan::diversity(asv_css, index = "shannon")
InvSimpson <- vegan::diversity(asv_css, index = "invsimpson") #inverted simpson
#Beta <- vegdist(asv_css, "bray") #beta dissimilarity
alpha_vegan <- cbind(Richness, Shannon, InvSimpson, meta)
alpha_vegan <- rownames_to_column(alpha_vegan, var = "id") %>% as_tibble()
head(alpha_vegan)
```

```
## # A tibble: 6 x 25
##   id   Richness Shannon InvSimpson Loc   Yr Spec Exp_cond Code  Repro Month
##   <chr>   <int>   <dbl>   <dbl> <chr> <int> <chr> <chr>   <chr> <chr> <chr>
## 1 SRR5~    88    1.81    4.19 NewC~ 2014 P_dam 30C    NC2    B   Nov
## 2 SRR5~    99    1.84    4.41 NewC~ 2014 P_dam 30C    NC1    B   Nov
## 3 SRR5~   173    2.39    6.46 Oman 2014 P_dam 31C    Om2    B   June
## 4 SRR5~   161    2.17    4.66 Oman 2014 P_dam 31C    Om2    B   June
## 5 SRR5~   155    2.32    5.33 Oman 2014 P_dam 31C    Om3    B   June
## 6 SRR5~   188    2.18    4.15 Oman 2014 P_dam 31C    Om2    B   June
## # ... with 14 more variables: Season <chr>, S_region <chr>, L_region <chr>,
```

```
## #   Exact.date <chr>, Tbl_bin <chr>, T_bleach <chr>, DHW <dbl>, DHW_cat <chr>,
## #   SST_a <dbl>, Coord_X <dbl>, Coord_Y <dbl>, Primer <chr>, Pub <chr>,
## #   Note <chr>
```

```
shapiro.test(alpha_vegan$Richness)
```

```
##
## Shapiro-Wilk normality test
##
## data:  alpha_vegan$Richness
## W = 0.89757, p-value = 4.731e-10
```

```
shapiro.test(alpha_vegan$Shannon)
```

```
##
## Shapiro-Wilk normality test
##
## data:  alpha_vegan$Shannon
## W = 0.95422, p-value = 9.696e-06
```

```
shapiro.test(alpha_vegan$InvSimpson)
```

```
##
## Shapiro-Wilk normality test
##
## data:  alpha_vegan$InvSimpson
## W = 0.9429, p-value = 8.839e-07
```

```
##Sean's approach
```

```
#Step 1: run regular LM models
```

```
mod.alpha <- lm(Richness ~ Loc + Tbl_bin + SST_a, data = alpha_vegan)
mod.even <- lm(Shannon ~ Loc + Tbl_bin + SST_a, data = alpha_vegan)
mod.inv <- lm(InvSimpson ~ Loc + Tbl_bin + SST_a, data = alpha_vegan)
```

```
##transformation as slight right-skew
```

```
mod.inv2 <- lm(log(InvSimpson) ~ Loc + Tbl_bin + SST_a, data = alpha_vegan)
mod.inv3 <- lm(sqrt(InvSimpson) ~ Loc + Tbl_bin + SST_a, data = alpha_vegan)
```

```
##Nested models with interaction effects, skew accounted for for InvSimpson
```

```
###Removed DHW as only one location have non-zero DHW experienced
```

```
###Nest Location within region
```

```
###Expect seasons to have a given range of SST_a, so have an interaction effect to reflect such
```

```
mod.alpha.nest <- lm(Richness ~ S_region/Loc + SST_a + Tbl_bin, data = alpha_vegan)
mod.eve.nest <- lm(Shannon ~ S_region/Loc + SST_a + Tbl_bin, data = alpha_vegan)
mod.inv.nest <- lm(log(InvSimpson) ~ S_region/Loc + SST_a + Tbl_bin, data = alpha_vegan)
```

```
##Compare AIC scores as sanity check
```

```
library(stats)
```

```
#Alpha
```

```
extractAIC(mod.alpha)
```

```
## [1] 10.000 1329.362
```

```
extractAIC(mod.alpha.nest)
```

```
## [1] 10.000 1329.362
```

```
#Even
```

```
extractAIC(mod.even)
```

```
## [1] 10.0000 -412.8677
```

```
extractAIC(mod.eve.nest)
```

```
## [1] 10.0000 -412.8677
```

```
#InvSimp
```

```
extractAIC(mod.inv)
```

```
## [1] 10.000 121.139
```

```
##extractAIC(mod.inv2)
```

```
##extractAIC(mod.inv3)
```

```
extractAIC(mod.inv.nest)
```

```
## [1] 10.0000 -429.4705
```

```
##Step 2: standardize residuals
```

```
mod.alpha.stdres = rstandard(mod.alpha)
```

```
mod.even.stdres = rstandard(mod.even)
```

```
mod.inv.stdres = rstandard(mod.inv)
```

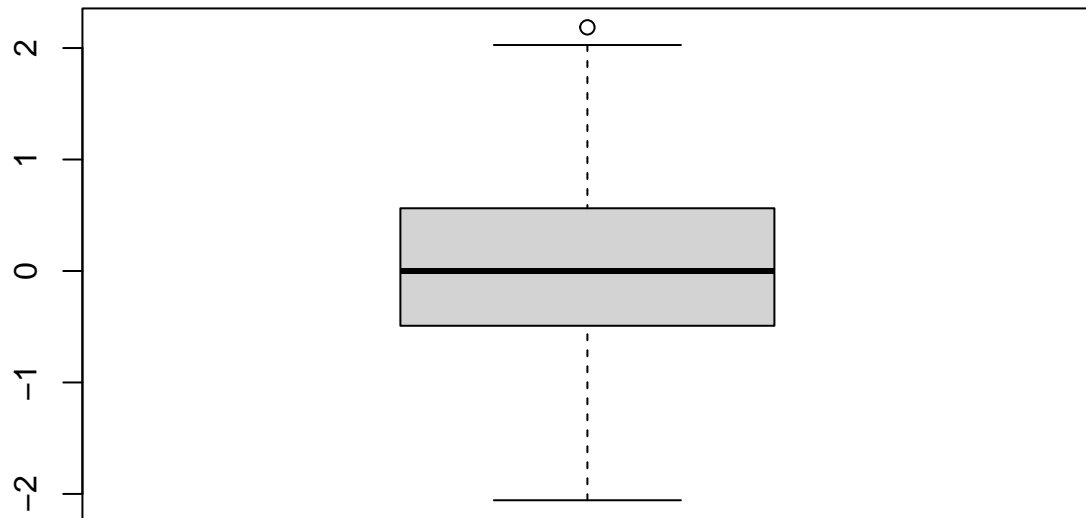
```
mod.inv2.stdres = rstandard(mod.inv2)
```

```
mod.alpha.nest.stdres = rstandard(mod.alpha.nest)
```

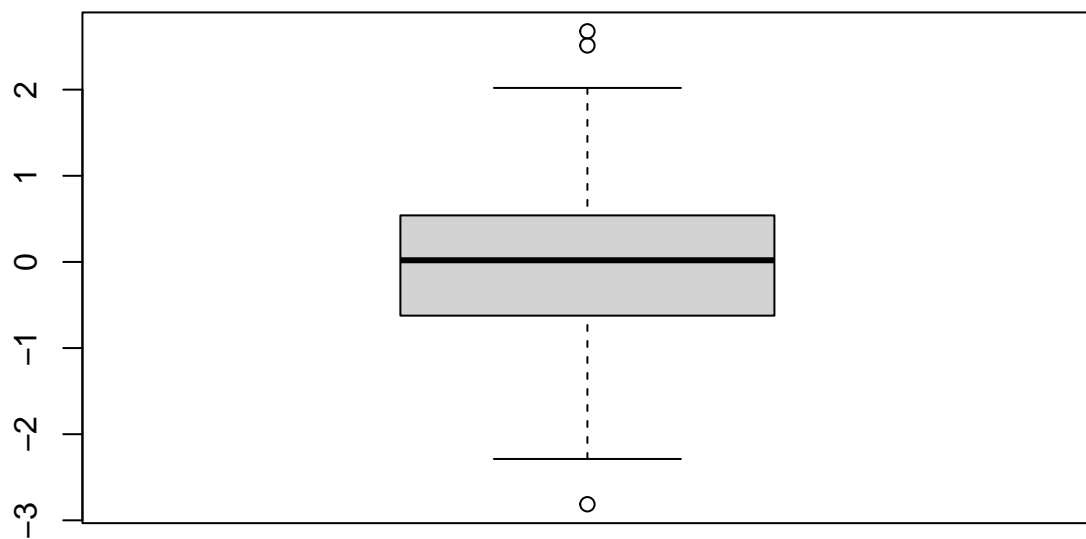
```
mod.eve.nest.stdres = rstandard(mod.eve.nest)
```

```
mod.inv.nest.stdres = rstandard(mod.inv.nest)
```

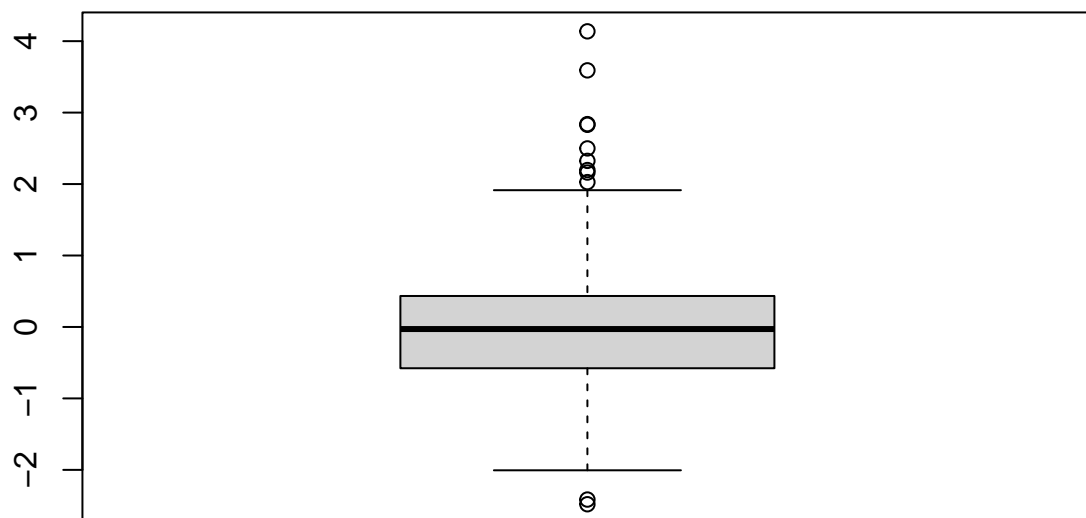
```
boxplot(mod.alpha.stdres)
```



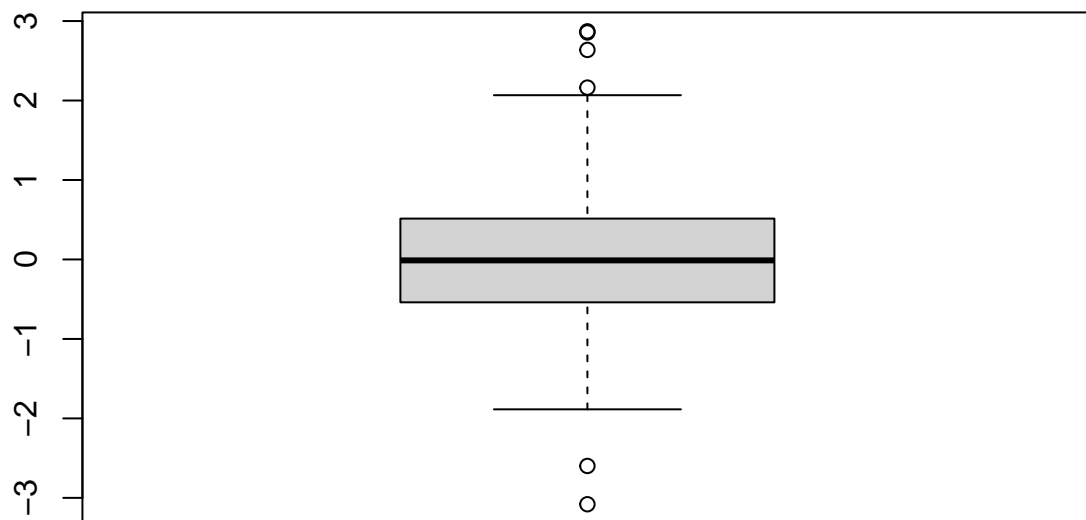
```
boxplot(mod.even.stdres)
```



```
boxplot(mod.inv.stdres)
```

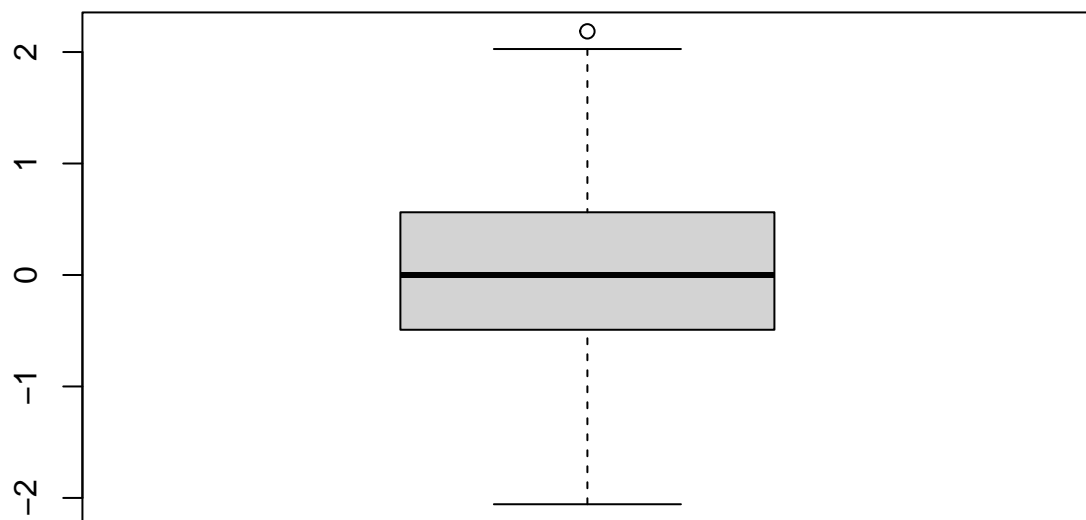


```
boxplot(mod.inv2.stdres)
```

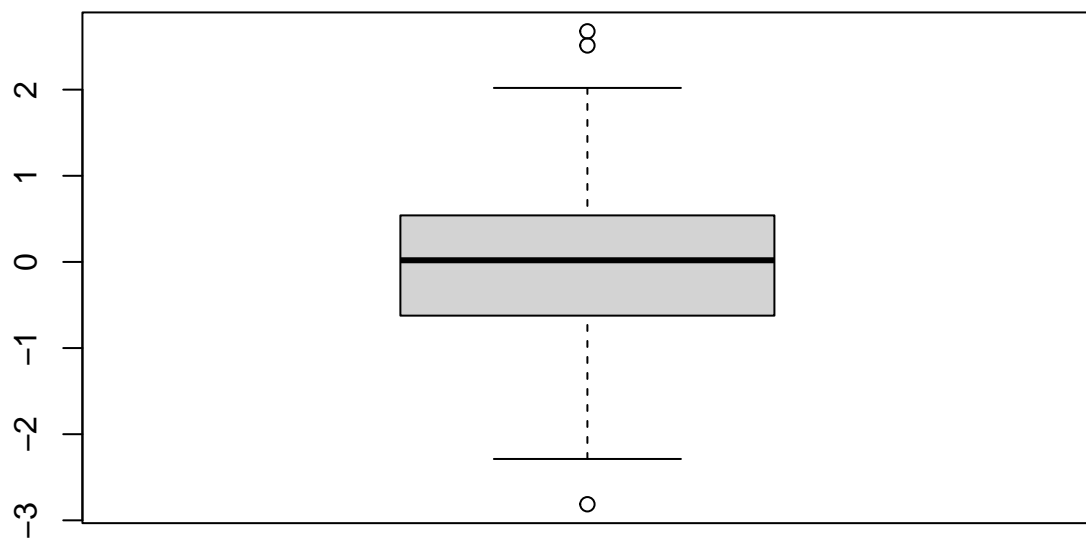


```
##Nested and interaction effect
```

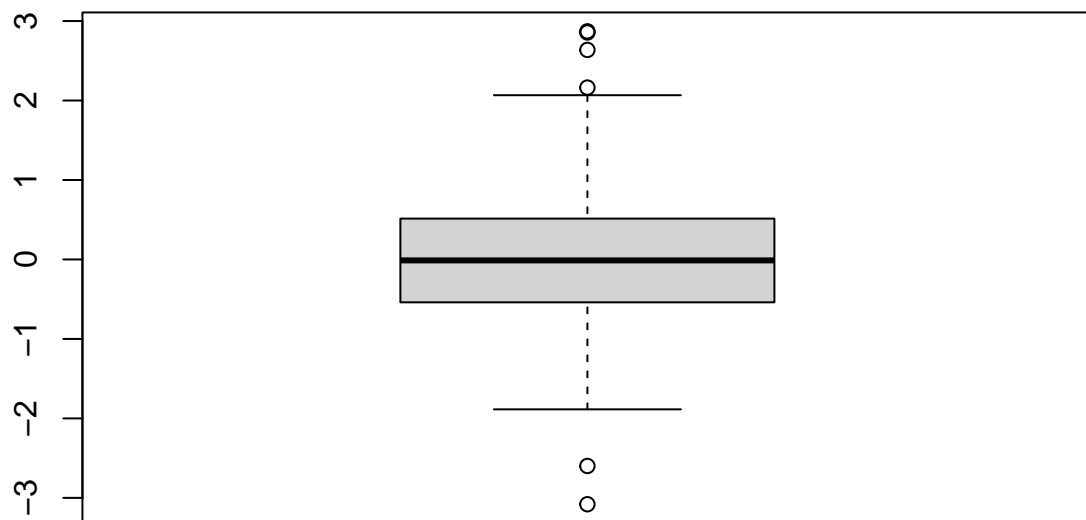
```
boxplot(mod.alpha.nest.stdres)
```



```
boxplot(mod.eve.nest.stdres)
```

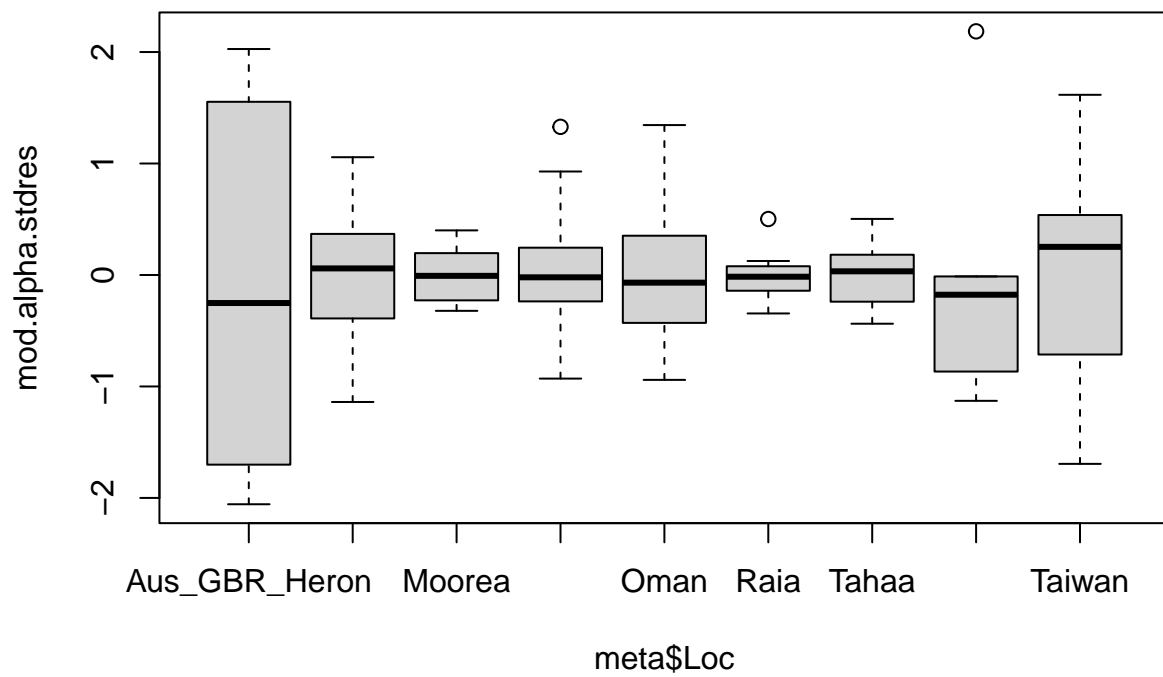



```
boxplot(mod.inv.nest.stdres)
```

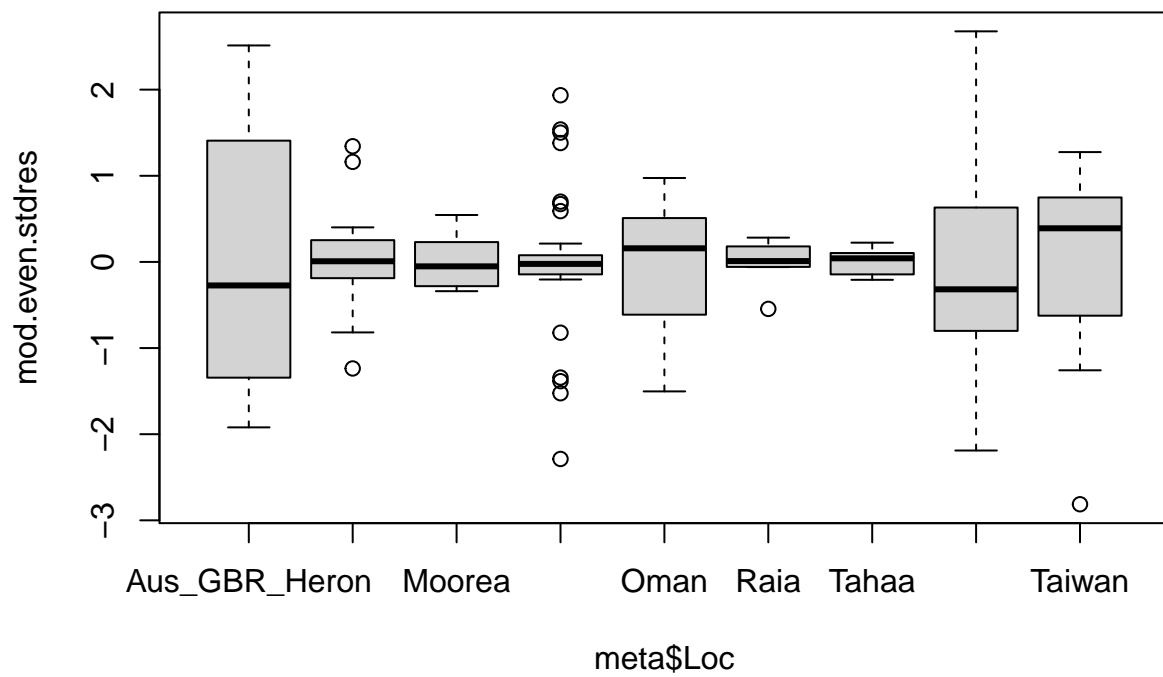


#Step 3: Boxplot standardized residual, by location

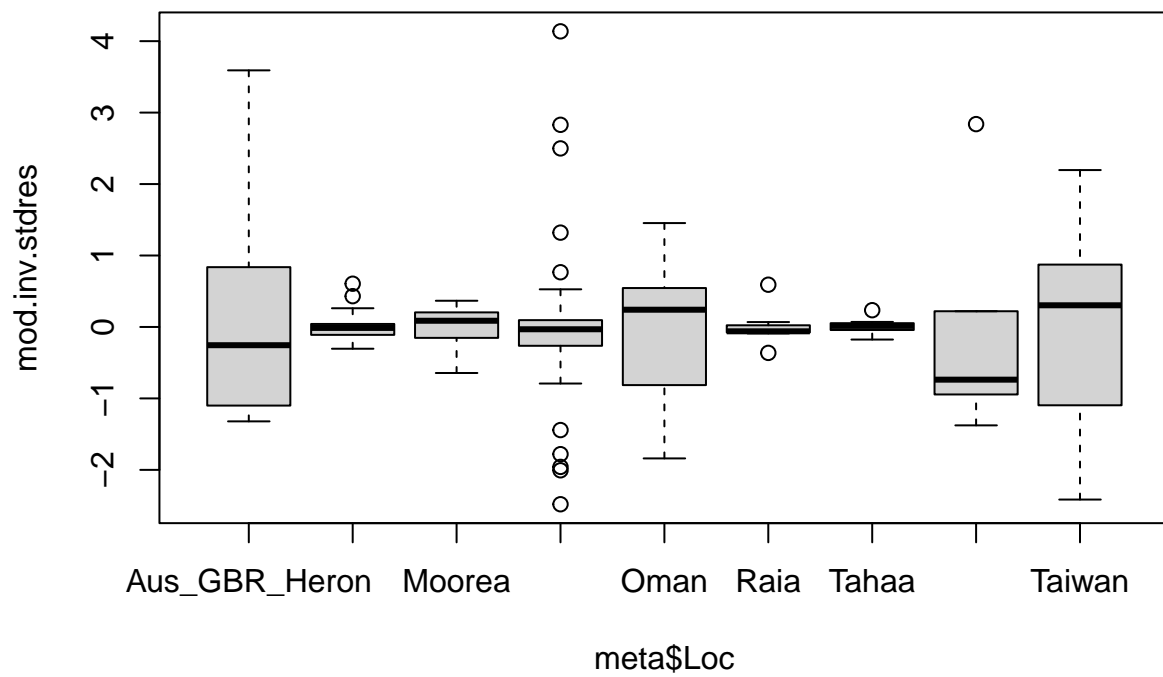
```
boxplot(mod.alpha.stdres ~ meta$Loc)
```



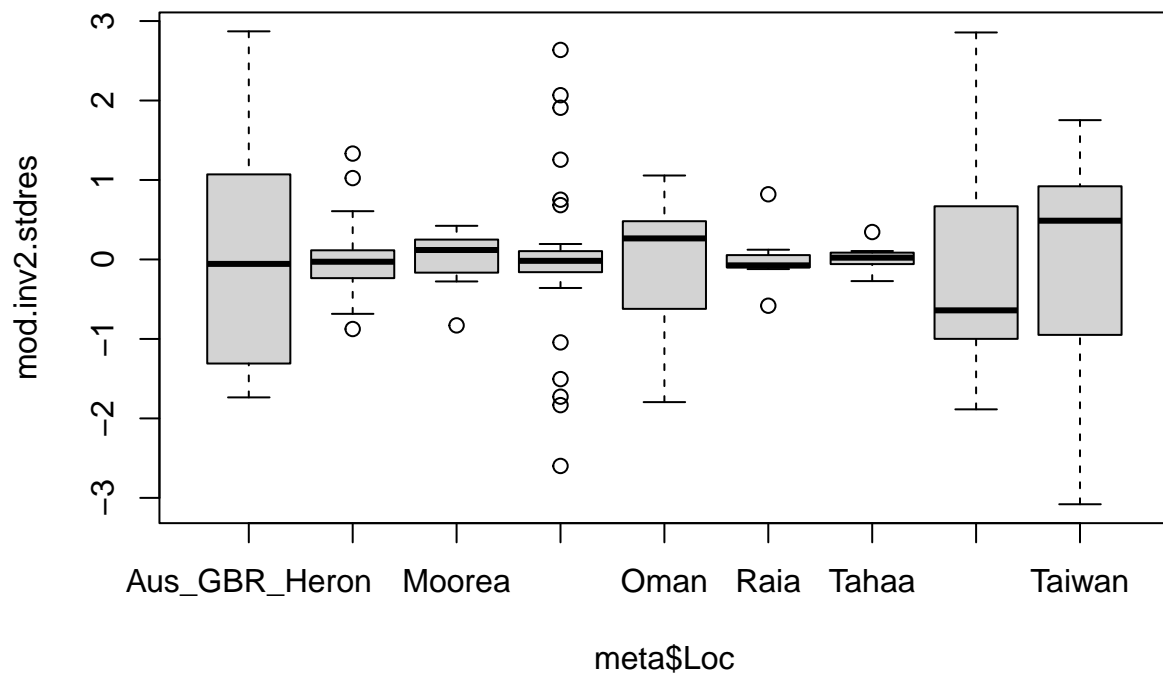
```
boxplot(mod.even.stdres ~ meta$Loc)
```



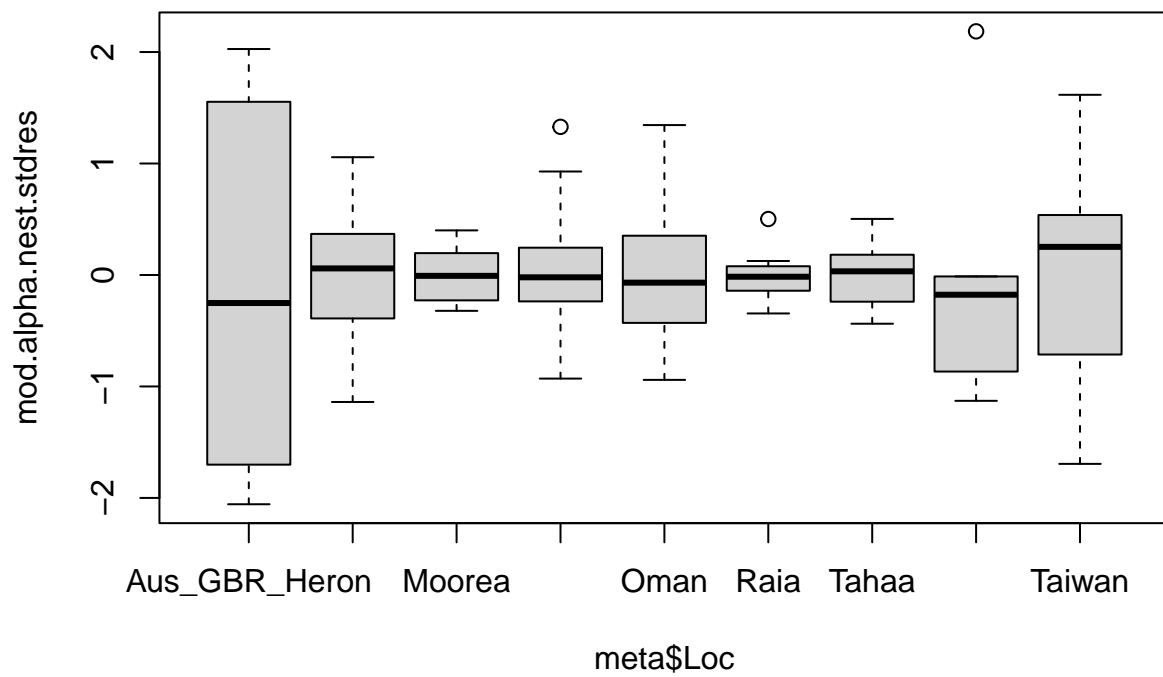
```
boxplot(mod.inv.stdres ~ meta$Loc)
```



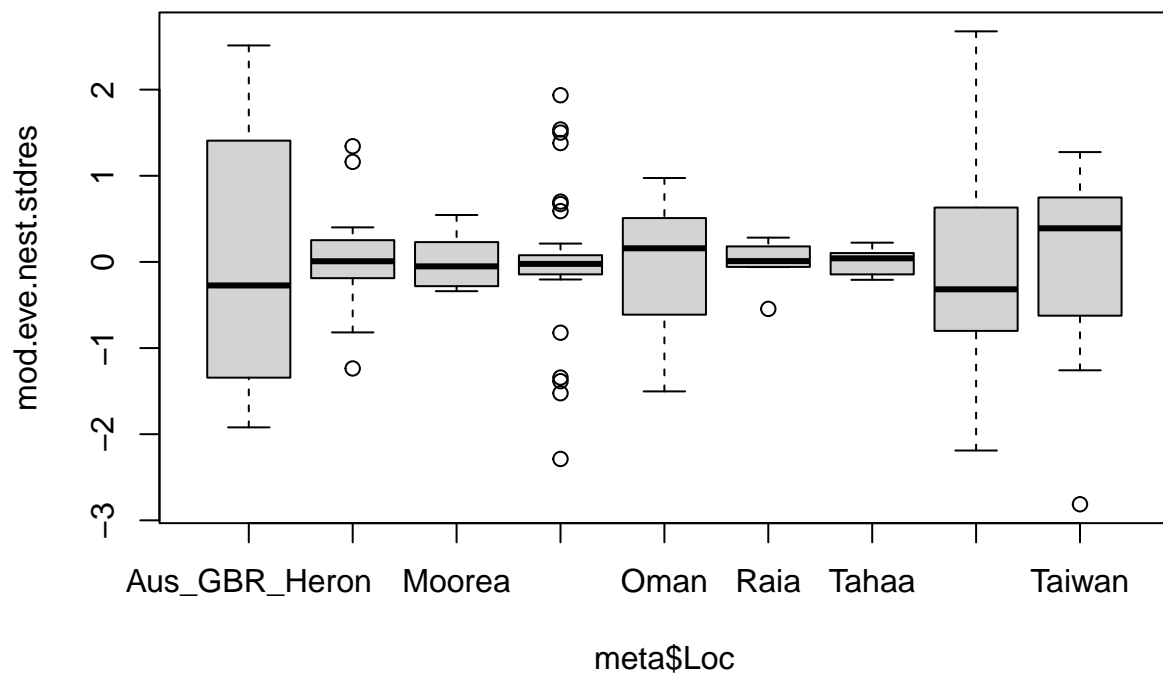
```
boxplot(mod.inv2.stdres ~ meta$Loc)
```



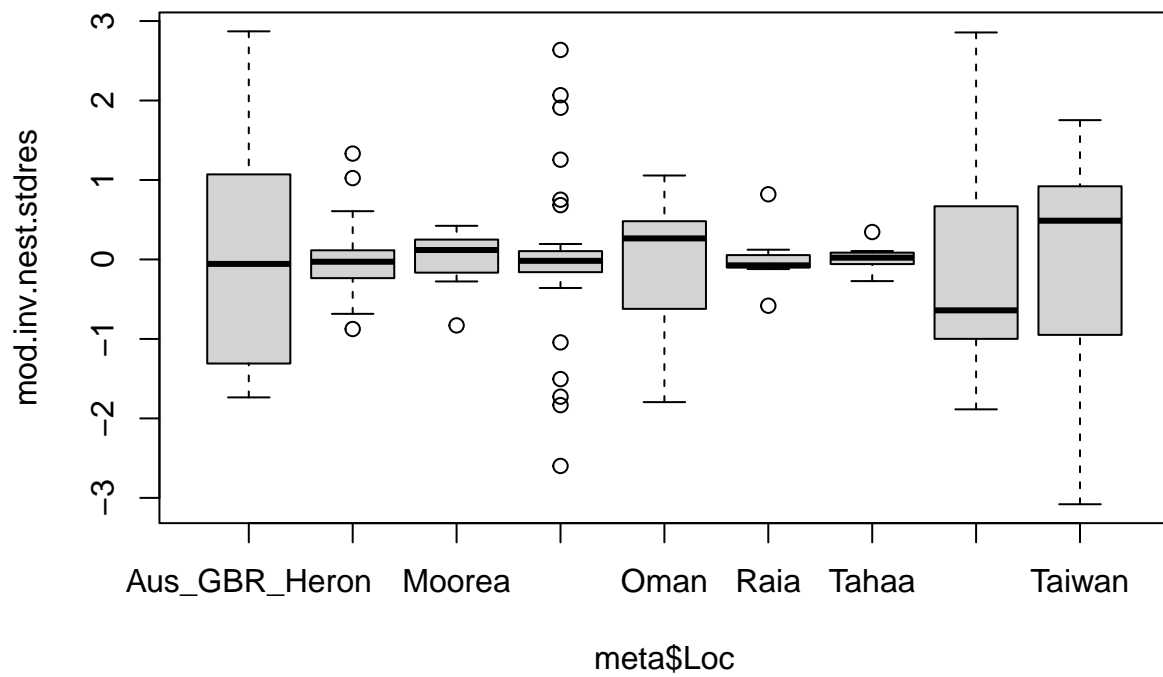
```
boxplot(mod.alpha.nest.stdres ~ meta$Loc)
```



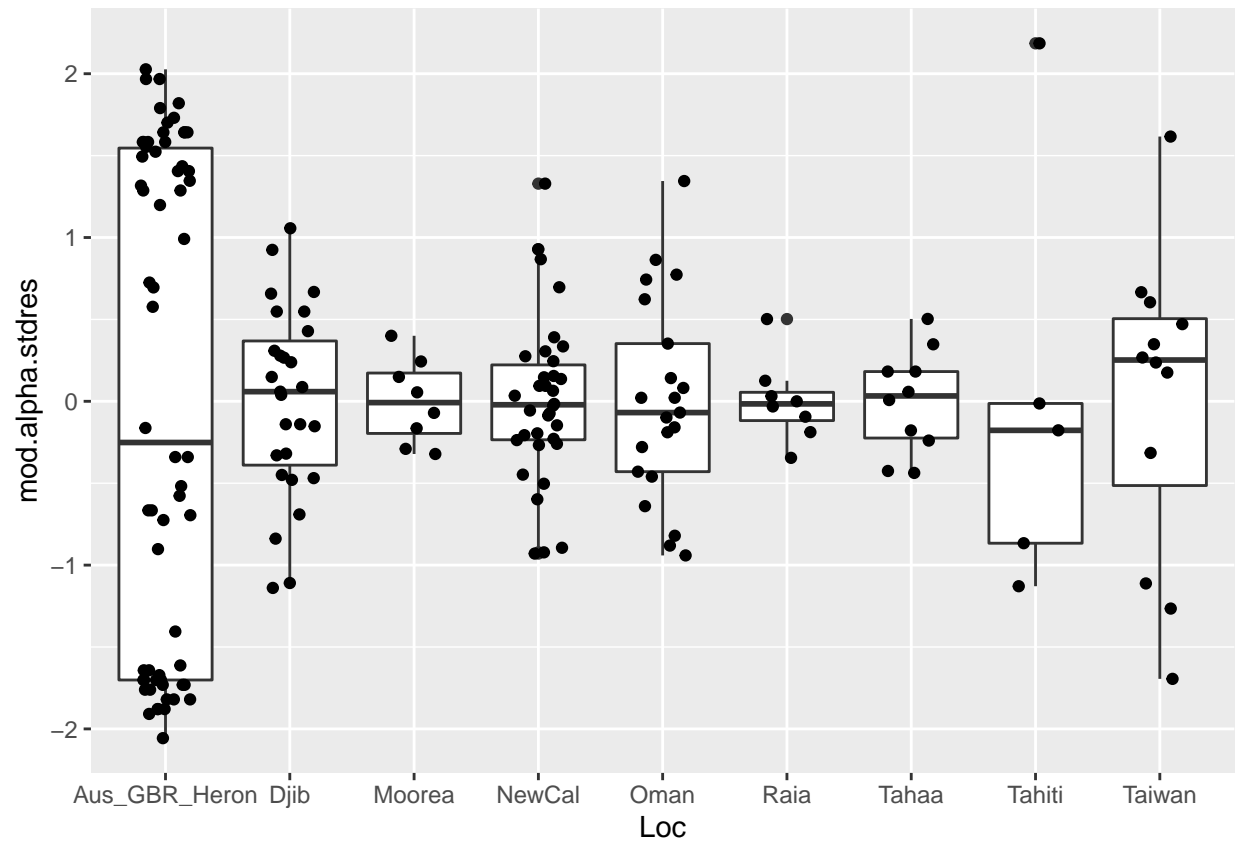
```
boxplot(mod.eve.nest.stdres ~ meta$Loc)
```



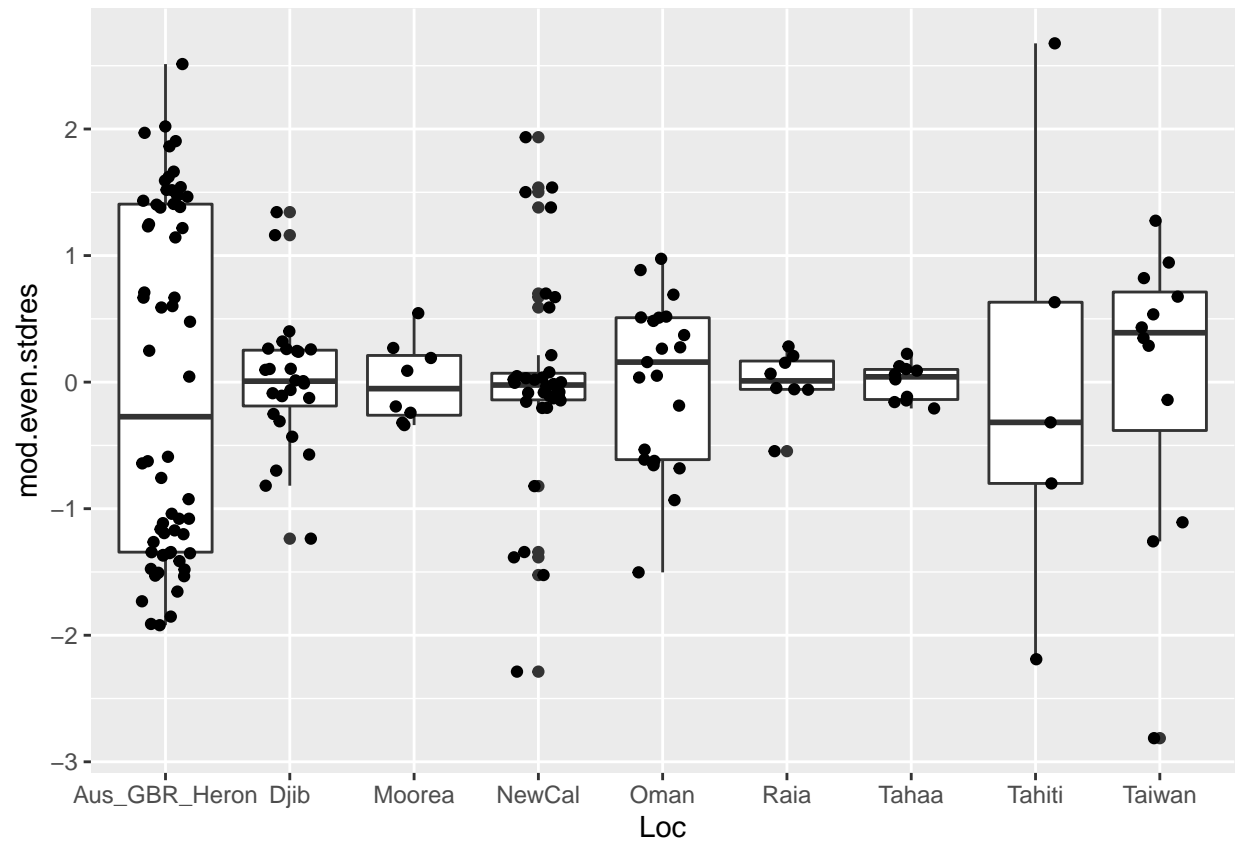
```
boxplot(mod.inv.nest.stdres ~ meta$Loc)
```

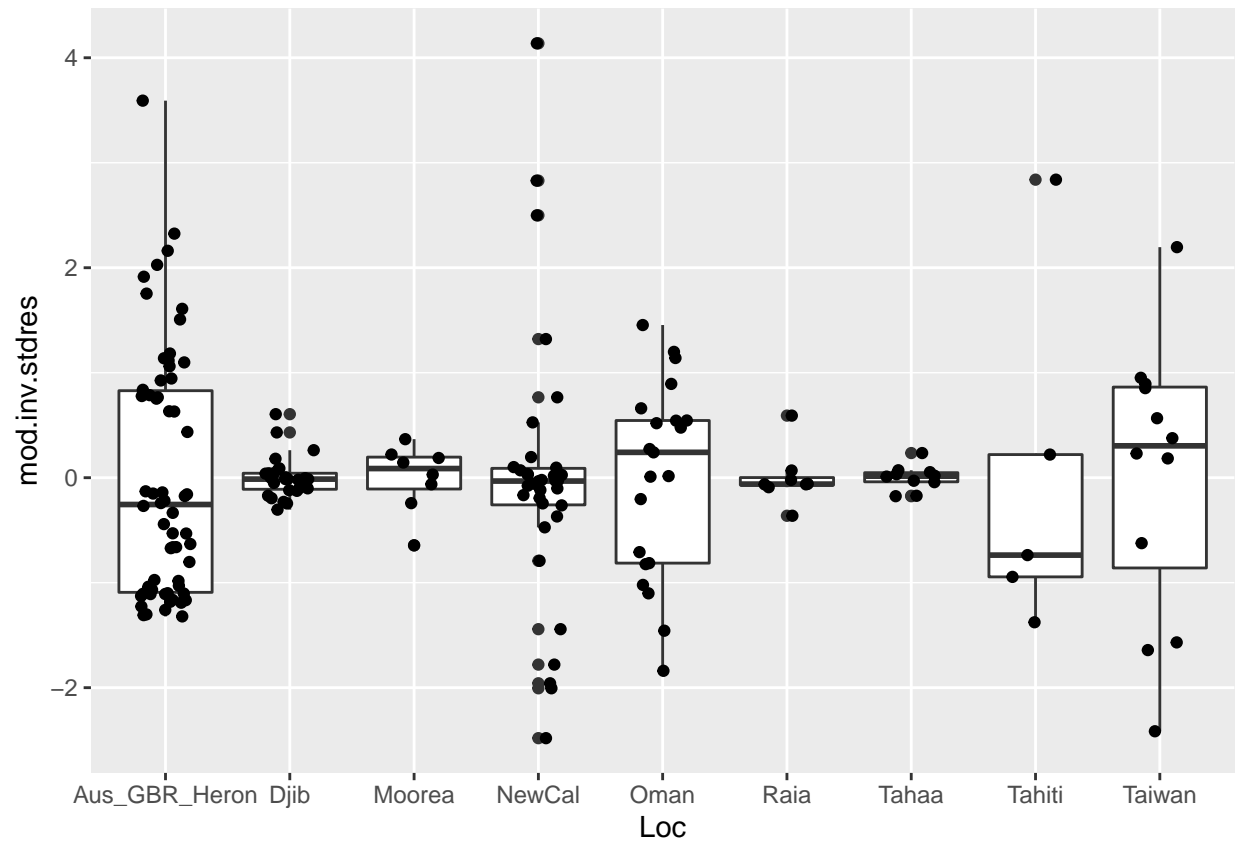
```
ggplot(meta, aes(x = Loc, y = mod.alpha.stdres)) + geom_boxplot() +
  geom_jitter(height = 0, width = .2)
```



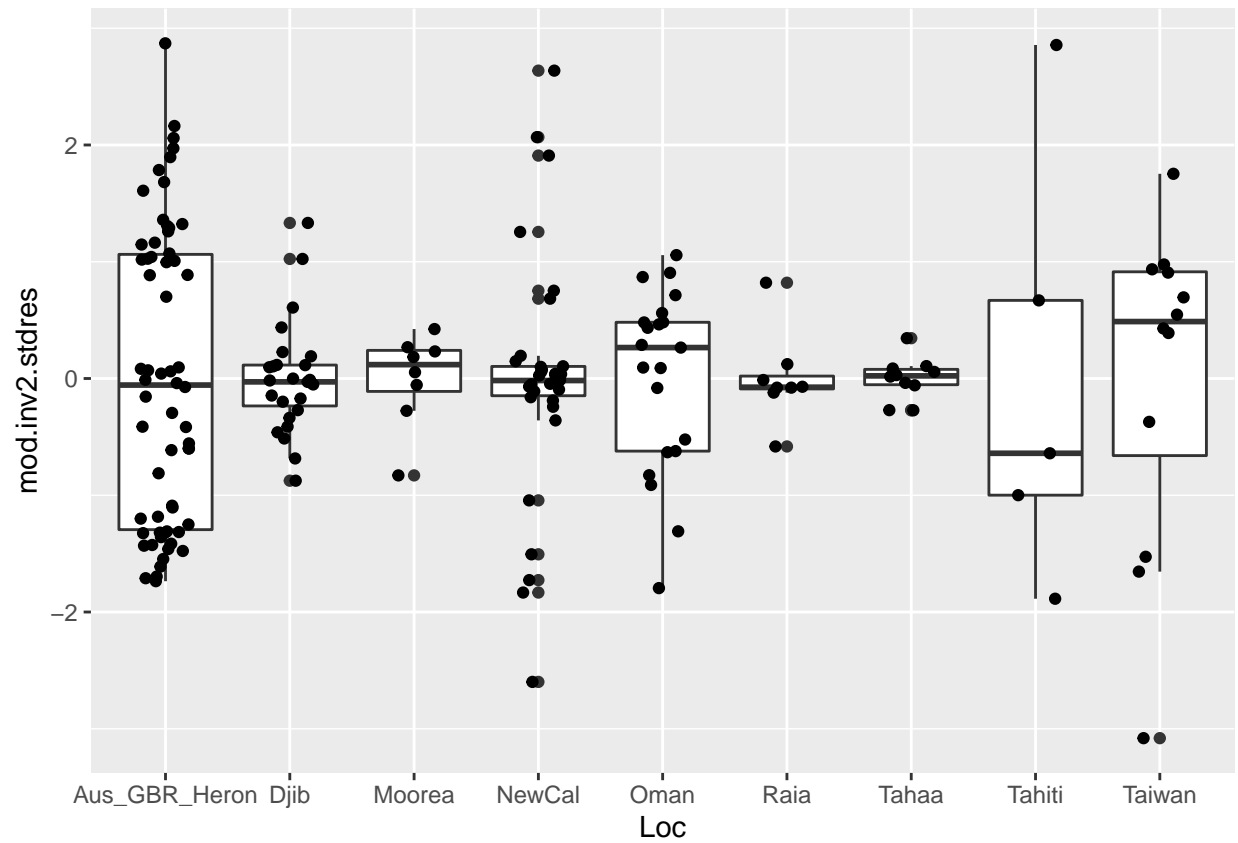
```
ggplot(meta, aes(x = Loc, y = mod.even.stdres)) + geom_boxplot() +
  geom_jitter(height = 0, width = .2)
```



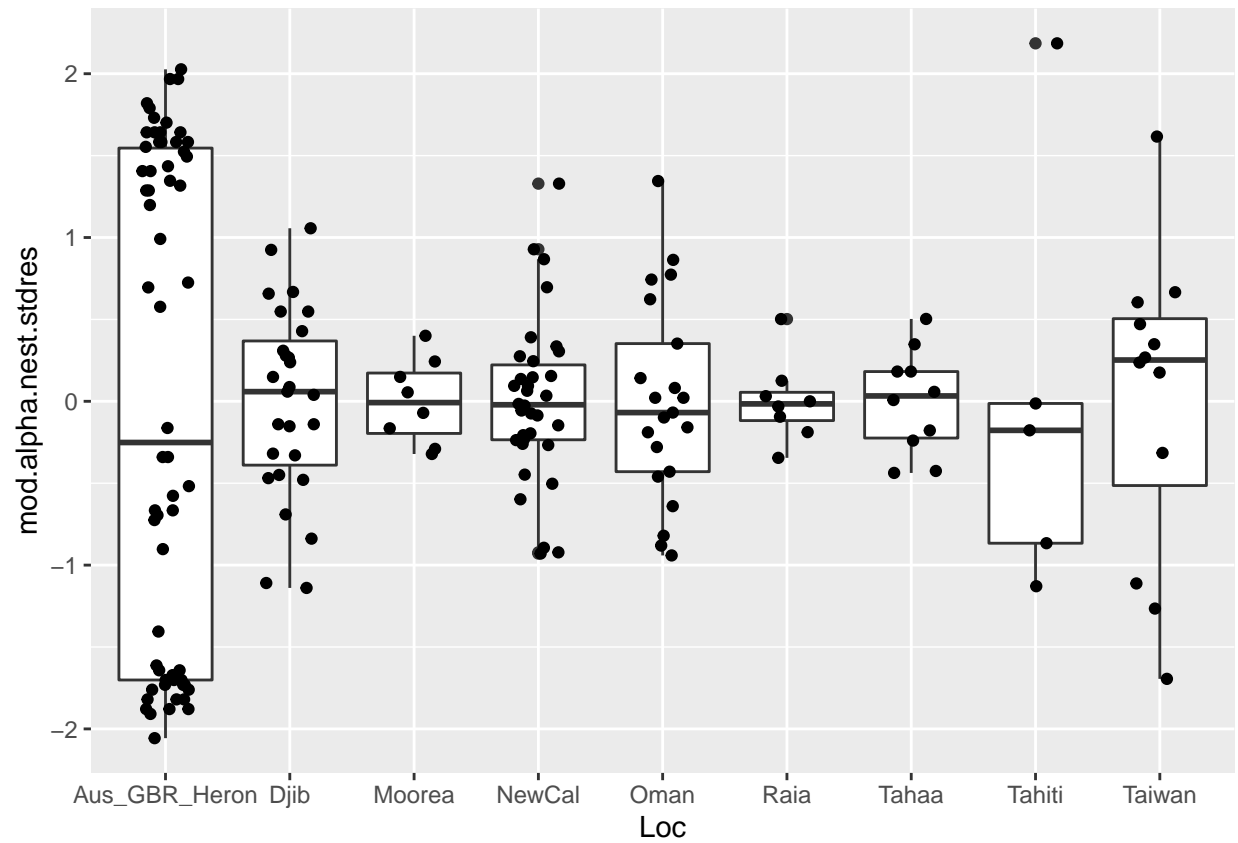
```
ggplot(meta, aes(x = Loc, y = mod.inv.stdres)) + geom_boxplot() +
  geom_jitter(height = 0, width = .2)
```



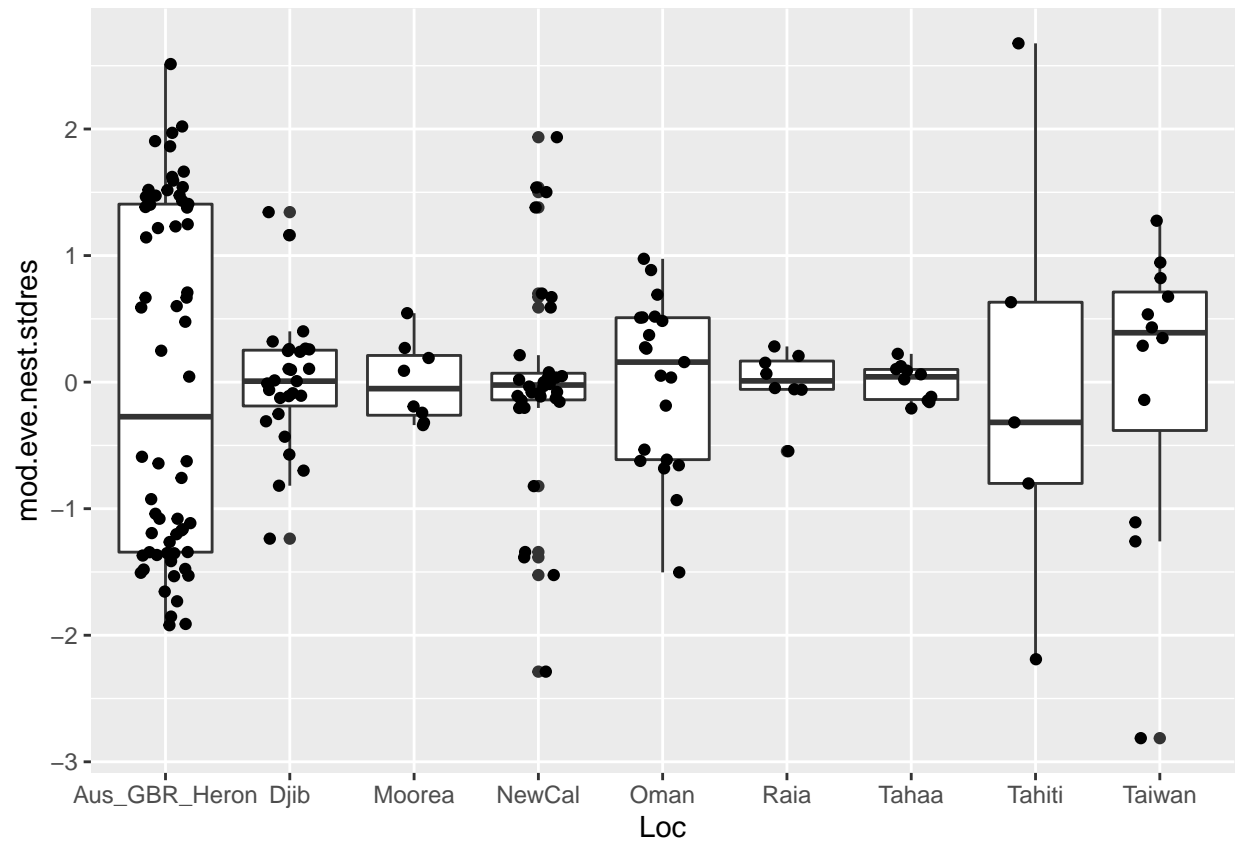
```
ggplot(meta, aes(x = Loc, y = mod.inv2.stdres)) + geom_boxplot() +
  geom_jitter(height = 0, width = .2)
```



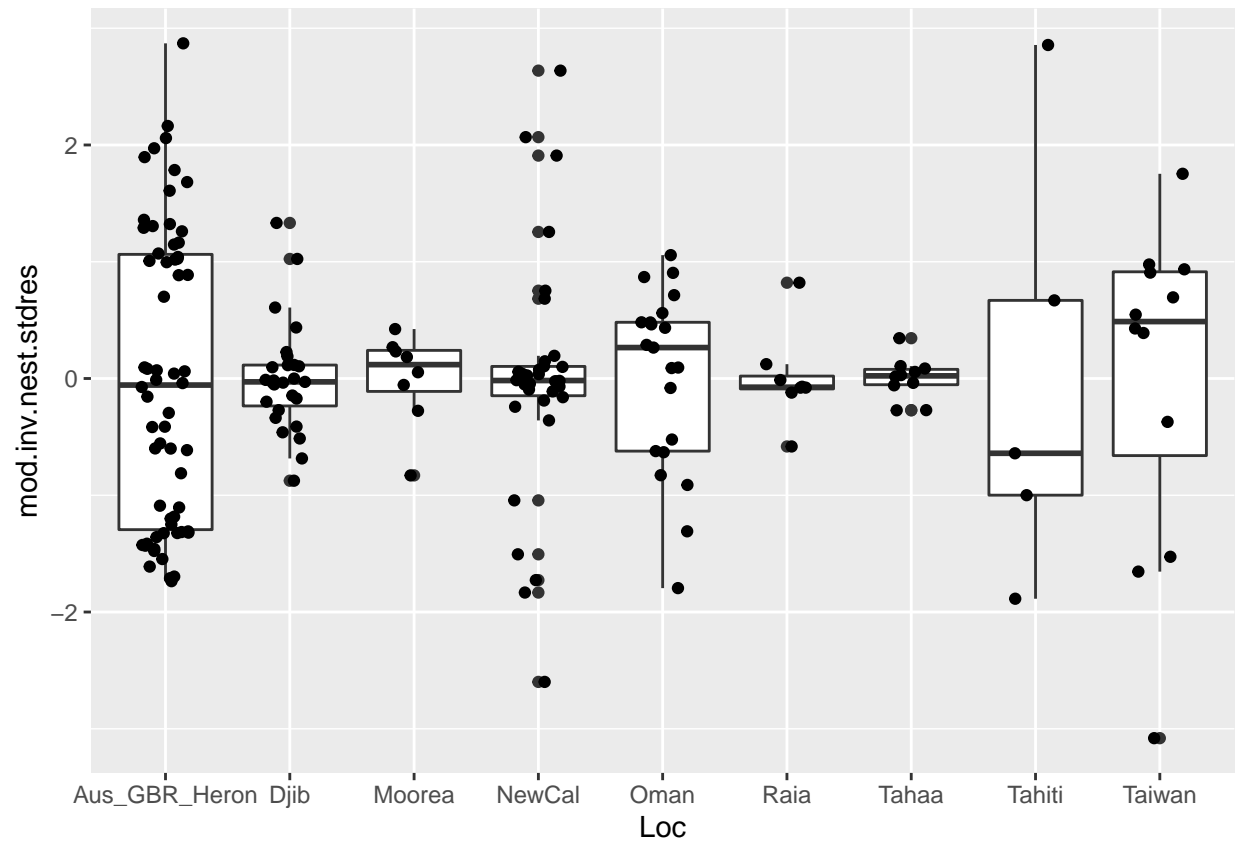
```
ggplot(meta, aes(x = Loc, y = mod.alpha.nest.stdres)) + geom_boxplot() +
  geom_jitter(height = 0, width = .2)
```



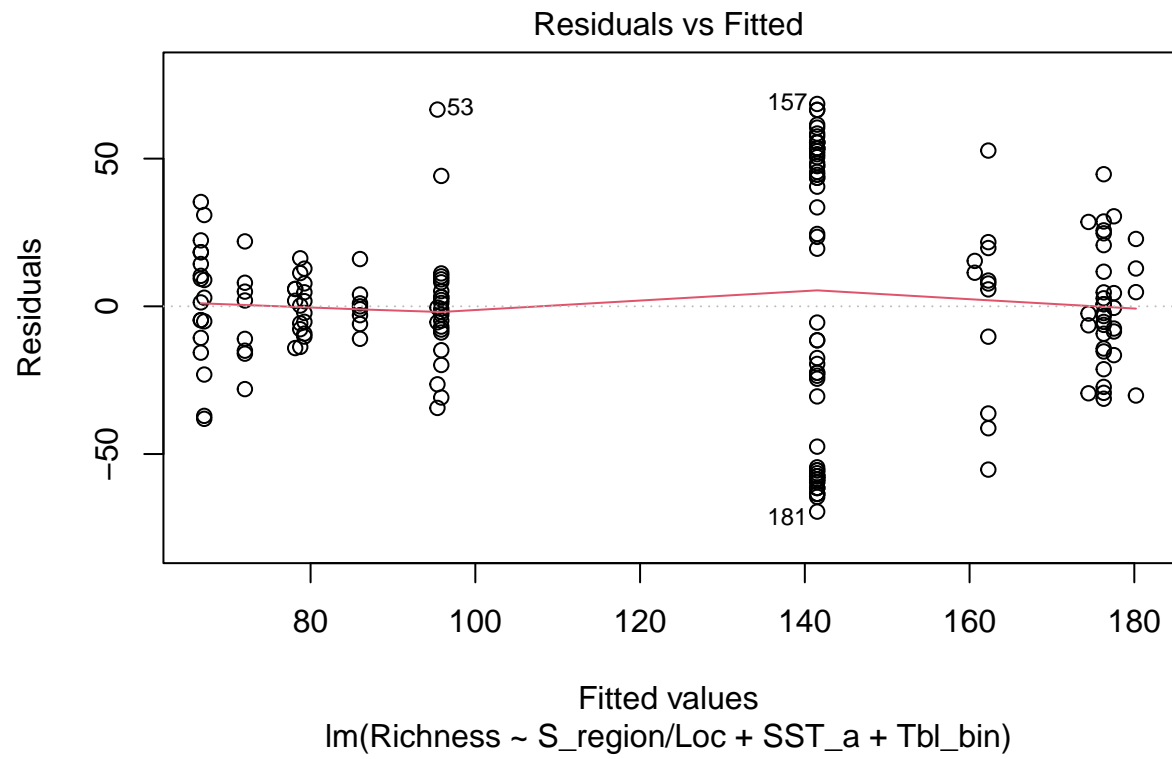
```
ggplot(meta, aes(x = Loc, y = mod.eve.nest.stdres)) + geom_boxplot() +
  geom_jitter(height = 0, width = .2)
```

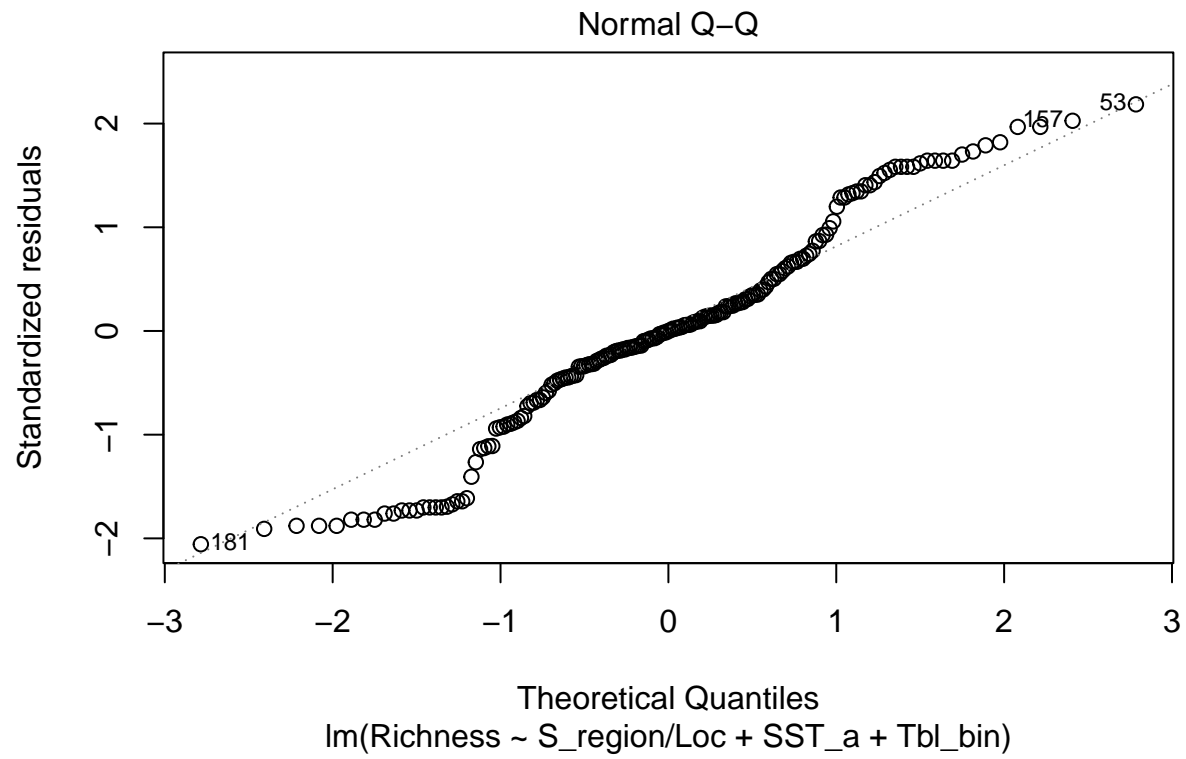


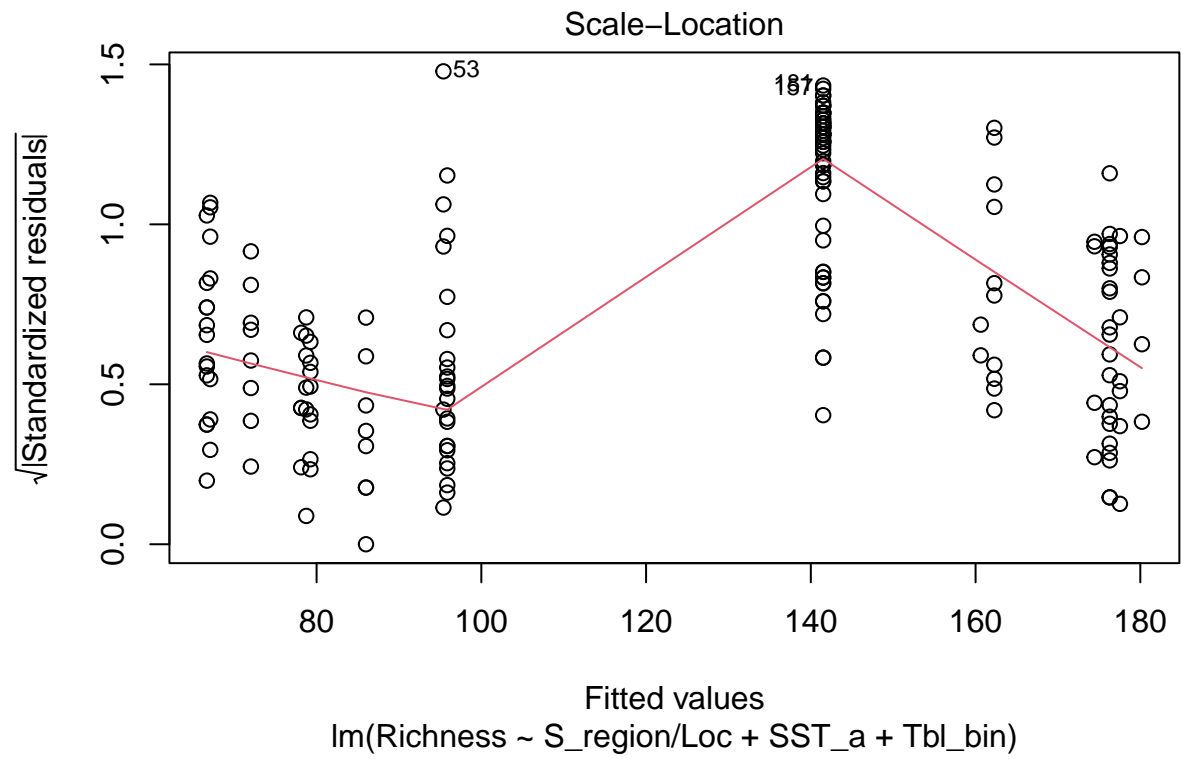
```
ggplot(meta, aes(x = Loc, y = mod.inv.nest.stdres)) + geom_boxplot() +
  geom_jitter(height = 0, width = .2)
```

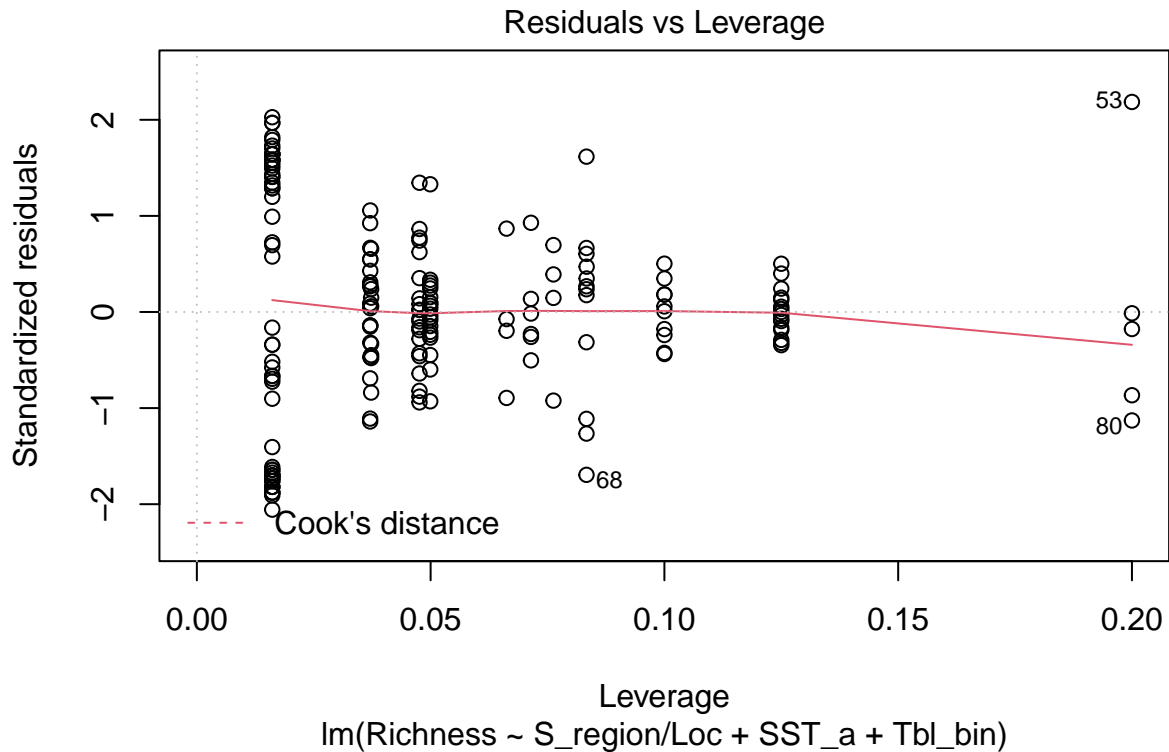


```
##qqplot for ref
plot(mod.alpha.nest)
```







```
summary(mod.alpha, cor=T)
```

```
##
## Call:
## lm(formula = Richness ~ Loc + Tbl_bin + SST_a, data = alpha_vegan)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -69.50 -16.27   0.00  18.91  68.50
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  653.4572    74.5160   8.769 1.46e-15 ***
## LocDjib       17.3273    15.3127   1.132  0.259
## LocMoorea    -18.0271    14.3220  -1.259  0.210
## LocNewCal     -0.1773     7.4730  -0.024  0.981
## LocOman      156.3473    19.6470   7.958 2.02e-13 ***
## LocRaia       -6.9577    14.6143  -0.476  0.635
## LocTahaa     -15.4450    13.5113  -1.143  0.255
## LocTahiti     -5.5795    16.8984  -0.330  0.742
## LocTaiwan     96.8788    15.4477   6.271 2.66e-09 ***
## Tbl_binRecent      NA           NA      NA      NA
## SST_a         -20.5688     2.9888  -6.882 9.84e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 34.07 on 177 degrees of freedom
## Multiple R-squared:  0.5816, Adjusted R-squared:  0.5603
## F-statistic: 27.34 on 9 and 177 DF,  p-value: < 2.2e-16
##
## Correlation of Coefficients:
##      (Intercept) LocDjib LocMoorea LocNewCal LocOman LocRaia LocTahaa
## LocDjib      0.84
## LocMoorea    0.43      0.47
## LocNewCal    0.20      0.36      0.28
## LocOman      0.88      0.83      0.47      0.34
## LocRaia      0.46      0.50      0.31      0.28      0.50
## LocTahaa     0.49      0.53      0.33      0.30      0.53      0.34
## LocTahiti    0.33      0.37      0.23      0.23      0.37      0.24      0.26
## LocTaiwan    0.70      0.70      0.41      0.33      0.71      0.43      0.46
## SST_a       -1.00      -0.86     -0.45     -0.23     -0.90     -0.48     -0.51
##      LocTahiti LocTaiwan
## LocDjib
## LocMoorea
## LocNewCal
## LocOman
## LocRaia
## LocTahaa
## LocTahiti
## LocTaiwan  0.32
## SST_a      -0.35     -0.72
```

```
summary(mod.even, cor=T)
```

```
##
## Call:
## lm(formula = Shannon ~ Loc + Tbl_bin + SST_a, data = alpha_vegan)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.87015 -0.19828  0.00597  0.16510  0.80528
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   4.86890    0.70656   6.891 9.36e-11 ***
## LocDjib       -0.66888    0.14519  -4.607 7.79e-06 ***
## LocMoorea     -0.12643    0.13580  -0.931 0.353108
## LocNewCal     -0.05813    0.07086  -0.820 0.413104
## LocOman        0.89017    0.18629   4.778 3.70e-06 ***
## LocRaia       -0.21315    0.13857  -1.538 0.125794
## LocTahaa      -0.23449    0.12811  -1.830 0.068881 .
## LocTahiti     -0.28815    0.16023  -1.798 0.073829 .
## LocTaiwan      0.49135    0.14647   3.355 0.000973 ***
## Tbl_binRecent      NA          NA      NA      NA
## SST_a         -0.10978    0.02834  -3.874 0.000151 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3231 on 177 degrees of freedom
```

```
## Multiple R-squared:  0.6589, Adjusted R-squared:  0.6415
## F-statistic: 37.99 on 9 and 177 DF,  p-value: < 2.2e-16
##
## Correlation of Coefficients:
##      (Intercept) LocDjib LocMoorea LocNewCal LocOman LocRaia LocTahaa
## LocDjib      0.84
## LocMoorea    0.43      0.47
## LocNewCal    0.20      0.36      0.28
## LocOman      0.88      0.83      0.47      0.34
## LocRaia      0.46      0.50      0.31      0.28      0.50
## LocTahaa     0.49      0.53      0.33      0.30      0.53      0.34
## LocTahiti    0.33      0.37      0.23      0.23      0.37      0.24      0.26
## LocTaiwan    0.70      0.70      0.41      0.33      0.71      0.43      0.46
## SST_a       -1.00     -0.86     -0.45     -0.23     -0.90     -0.48     -0.51
##      LocTahiti LocTaiwan
## LocDjib
## LocMoorea
## LocNewCal
## LocOman
## LocRaia
## LocTahaa
## LocTahiti
## LocTaiwan  0.32
## SST_a     -0.35     -0.72
```

```
summary(mod.inv2, cor=T)
```

```
##
## Call:
## lm(formula = log(InvSimpson) ~ Loc + Tbl_bin + SST_a, data = alpha_vegan)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.91135 -0.16292 -0.00353  0.15351  0.87988
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   1.98073    0.67588   2.931  0.00383 **
## LocDjib       -0.61963    0.13889  -4.461 1.44e-05 ***
## LocMoorea      0.09044    0.12990   0.696  0.48721
## LocNewCal      0.27120    0.06778   4.001 9.26e-05 ***
## LocOman        0.58212    0.17820   3.267  0.00131 **
## LocRaia       -0.16892    0.13256  -1.274  0.20421
## LocTahaa      -0.16177    0.12255  -1.320  0.18852
## LocTahiti     -0.07575    0.15327  -0.494  0.62177
## LocTaiwan      0.36069    0.14011   2.574  0.01086 *
## Tbl_binRecent      NA          NA      NA      NA
## SST_a         -0.02811    0.02711  -1.037  0.30111
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.309 on 177 degrees of freedom
## Multiple R-squared:  0.5722, Adjusted R-squared:  0.5504
## F-statistic: 26.3 on 9 and 177 DF,  p-value: < 2.2e-16
```

```
##
## Correlation of Coefficients:
##      (Intercept) LocDjib LocMoorea LocNewCal LocOman LocRaia LocTahaa
## LocDjib      0.84
## LocMoorea    0.43      0.47
## LocNewCal    0.20      0.36      0.28
## LocOman      0.88      0.83      0.47      0.34
## LocRaia      0.46      0.50      0.31      0.28      0.50
## LocTahaa     0.49      0.53      0.33      0.30      0.53      0.34
## LocTahiti    0.33      0.37      0.23      0.23      0.37      0.24      0.26
## LocTaiwan    0.70      0.70      0.41      0.33      0.71      0.43      0.46
## SST_a       -1.00      -0.86     -0.45     -0.23     -0.90     -0.48     -0.51
##      LocTahiti LocTaiwan
## LocDjib
## LocMoorea
## LocNewCal
## LocOman
## LocRaia
## LocTahaa
## LocTahiti
## LocTaiwan  0.32
## SST_a      -0.35     -0.72
```

```
summary(mod.alpha.nest, cor=T)
```

```
##
## Call:
## lm(formula = Richness ~ S_region/Loc + SST_a + Tbl_bin, data = alpha_vegan)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -69.50 -16.27   0.00  18.91  68.50
##
## Coefficients: (37 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    653.4572    74.5160   8.769 1.46e-15 ***
## S_regionFrPoly  -18.0271    14.3220  -1.259 0.209795
## S_regionIndianOc 156.3473    19.6470   7.958 2.02e-13 ***
## S_regionNewCal   -0.1773     7.4730  -0.024 0.981096
## S_regionTaiwan    84.4312    25.0626   3.369 0.000926 ***
## SST_a          -20.5688     2.9888  -6.882 9.84e-11 ***
## Tbl_binRecent    12.4476    19.4309   0.641 0.522605
## S_regionAus:LocDjib      NA         NA      NA      NA
## S_regionFrPoly:LocDjib    NA         NA      NA      NA
## S_regionIndianOc:LocDjib -139.0199    10.8949 -12.760 < 2e-16 ***
## S_regionNewCal:LocDjib    NA         NA      NA      NA
## S_regionTaiwan:LocDjib    NA         NA      NA      NA
## S_regionAus:LocMoorea     NA         NA      NA      NA
## S_regionFrPoly:LocMoorea  NA         NA      NA      NA
## S_regionIndianOc:LocMoorea NA         NA      NA      NA
## S_regionNewCal:LocMoorea  NA         NA      NA      NA
## S_regionTaiwan:LocMoorea  NA         NA      NA      NA
## S_regionAus:LocNewCal     NA         NA      NA      NA
## S_regionFrPoly:LocNewCal  NA         NA      NA      NA
```

```

## S_regionIndianOc:LocNewCal      NA      NA      NA      NA
## S_regionNewCal:LocNewCal         NA      NA      NA      NA
## S_regionTaiwan:LocNewCal         NA      NA      NA      NA
## S_regionAus:LocOman              NA      NA      NA      NA
## S_regionFrPoly:LocOman           NA      NA      NA      NA
## S_regionIndianOc:LocOman         NA      NA      NA      NA
## S_regionNewCal:LocOman           NA      NA      NA      NA
## S_regionTaiwan:LocOman           NA      NA      NA      NA
## S_regionAus:LocRaia              NA      NA      NA      NA
## S_regionFrPoly:LocRaia          -1.3782   19.4584  -0.071  0.943616
## S_regionIndianOc:LocRaia         NA      NA      NA      NA
## S_regionNewCal:LocRaia           NA      NA      NA      NA
## S_regionTaiwan:LocRaia           NA      NA      NA      NA
## S_regionAus:LocTahaa             NA      NA      NA      NA
## S_regionFrPoly:LocTahaa          -9.8655   18.6894  -0.528  0.598255
## S_regionIndianOc:LocTahaa        NA      NA      NA      NA
## S_regionNewCal:LocTahaa          NA      NA      NA      NA
## S_regionTaiwan:LocTahaa          NA      NA      NA      NA
## S_regionAus:LocTahiti            NA      NA      NA      NA
## S_regionFrPoly:LocTahiti         NA      NA      NA      NA
## S_regionIndianOc:LocTahiti       NA      NA      NA      NA
## S_regionNewCal:LocTahiti         NA      NA      NA      NA
## S_regionTaiwan:LocTahiti         NA      NA      NA      NA
## S_regionAus:LocTaiwan            NA      NA      NA      NA
## S_regionFrPoly:LocTaiwan         NA      NA      NA      NA
## S_regionIndianOc:LocTaiwan       NA      NA      NA      NA
## S_regionNewCal:LocTaiwan         NA      NA      NA      NA
## S_regionTaiwan:LocTaiwan         NA      NA      NA      NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 34.07 on 177 degrees of freedom
## Multiple R-squared:  0.5816, Adjusted R-squared:  0.5603
## F-statistic: 27.34 on 9 and 177 DF,  p-value: < 2.2e-16
##
## Correlation of Coefficients:
##              (Intercept) S_regionFrPoly S_regionIndianOc
## S_regionFrPoly      0.43
## S_regionIndianOc    0.88      0.47
## S_regionNewCal      0.20      0.28      0.34
## S_regionTaiwan      0.45      0.66      0.46
## SST_a               -1.00     -0.45     -0.90
## Tbl_binRecent       -0.03     -0.53     -0.02
## S_regionIndianOc:LocDjib -0.41     -0.19     -0.63
## S_regionFrPoly:LocRaia  0.06      0.03      0.05
## S_regionFrPoly:LocTahaa 0.05      0.02      0.05
##              S_regionNewCal S_regionTaiwan SST_a Tbl_binRecent
## S_regionFrPoly
## S_regionIndianOc
## S_regionNewCal
## S_regionTaiwan      0.21
## SST_a               -0.23     -0.46
## Tbl_binRecent       -0.01     -0.79      0.03
## S_regionIndianOc:LocDjib -0.10     -0.19      0.41  0.01

```



```
## S_regionFrPoly:LocRaia    0.01          0.50          -0.06 -0.62
## S_regionFrPoly:LocTahaa   0.01          0.52          -0.05 -0.64
##                               S_regionIndianOc:LocDjib S_regionFrPoly:LocRaia
## S_regionFrPoly
## S_regionIndianOc
## S_regionNewCal
## S_regionTaiwan
## SST_a
## Tbl_binRecent
## S_regionIndianOc:LocDjib
## S_regionFrPoly:LocRaia   -0.02
## S_regionFrPoly:LocTahaa -0.02          0.64
```

```
summary(mod.eve.nest, cor=T)
```

```
##
## Call:
## lm(formula = Shannon ~ S_region/Loc + SST_a + Tbl_bin, data = alpha_vegan)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.87015 -0.19828  0.00597  0.16510  0.80528
##
## Coefficients: (37 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      4.86890    0.70656   6.891 9.36e-11 ***
## S_regionFrPoly    -0.12643    0.13580  -0.931 0.353108
## S_regionIndianOc   0.89017    0.18629   4.778 3.70e-06 ***
## S_regionNewCal    -0.05813    0.07086  -0.820 0.413104
## S_regionTaiwan     0.65306    0.23764   2.748 0.006615 **
## SST_a            -0.10978    0.02834  -3.874 0.000151 ***
## Tbl_binRecent    -0.16171    0.18424  -0.878 0.381289
## S_regionAus:LocDjib      NA         NA      NA      NA
## S_regionFrPoly:LocDjib    NA         NA      NA      NA
## S_regionIndianOc:LocDjib -1.55905    0.10330 -15.092 < 2e-16 ***
## S_regionNewCal:LocDjib    NA         NA      NA      NA
## S_regionTaiwan:LocDjib    NA         NA      NA      NA
## S_regionAus:LocMoorea     NA         NA      NA      NA
## S_regionFrPoly:LocMoorea   NA         NA      NA      NA
## S_regionIndianOc:LocMoorea NA         NA      NA      NA
## S_regionNewCal:LocMoorea   NA         NA      NA      NA
## S_regionTaiwan:LocMoorea   NA         NA      NA      NA
## S_regionAus:LocNewCal     NA         NA      NA      NA
## S_regionFrPoly:LocNewCal   NA         NA      NA      NA
## S_regionIndianOc:LocNewCal NA         NA      NA      NA
## S_regionNewCal:LocNewCal   NA         NA      NA      NA
## S_regionTaiwan:LocNewCal   NA         NA      NA      NA
## S_regionAus:LocOman       NA         NA      NA      NA
## S_regionFrPoly:LocOman     NA         NA      NA      NA
## S_regionIndianOc:LocOman   NA         NA      NA      NA
## S_regionNewCal:LocOman     NA         NA      NA      NA
## S_regionTaiwan:LocOman     NA         NA      NA      NA
## S_regionAus:LocRaia       NA         NA      NA      NA
## S_regionFrPoly:LocRaia     0.07500    0.18450   0.406 0.684874
```

```

## S_regionIndianOc:LocRaia      NA      NA      NA      NA
## S_regionNewCal:LocRaia        NA      NA      NA      NA
## S_regionTaiwan:LocRaia        NA      NA      NA      NA
## S_regionAus:LocTahaa          NA      NA      NA      NA
## S_regionFrPoly:LocTahaa      0.05366  0.17721  0.303  0.762417
## S_regionIndianOc:LocTahaa     NA      NA      NA      NA
## S_regionNewCal:LocTahaa       NA      NA      NA      NA
## S_regionTaiwan:LocTahaa       NA      NA      NA      NA
## S_regionAus:LocTahiti         NA      NA      NA      NA
## S_regionFrPoly:LocTahiti      NA      NA      NA      NA
## S_regionIndianOc:LocTahiti    NA      NA      NA      NA
## S_regionNewCal:LocTahiti      NA      NA      NA      NA
## S_regionTaiwan:LocTahiti      NA      NA      NA      NA
## S_regionAus:LocTaiwan         NA      NA      NA      NA
## S_regionFrPoly:LocTaiwan      NA      NA      NA      NA
## S_regionIndianOc:LocTaiwan    NA      NA      NA      NA
## S_regionNewCal:LocTaiwan      NA      NA      NA      NA
## S_regionTaiwan:LocTaiwan      NA      NA      NA      NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3231 on 177 degrees of freedom
## Multiple R-squared:  0.6589, Adjusted R-squared:  0.6415
## F-statistic: 37.99 on 9 and 177 DF,  p-value: < 2.2e-16
##
## Correlation of Coefficients:
##              (Intercept) S_regionFrPoly S_regionIndianOc
## S_regionFrPoly      0.43
## S_regionIndianOc    0.88      0.47
## S_regionNewCal      0.20      0.28      0.34
## S_regionTaiwan      0.45      0.66      0.46
## SST_a              -1.00     -0.45     -0.90
## Tbl_binRecent      -0.03     -0.53     -0.02
## S_regionIndianOc:LocDjib -0.41     -0.19     -0.63
## S_regionFrPoly:LocRaia   0.06      0.03      0.05
## S_regionFrPoly:LocTahaa  0.05      0.02      0.05
##              S_regionNewCal S_regionTaiwan SST_a Tbl_binRecent
## S_regionFrPoly
## S_regionIndianOc
## S_regionNewCal
## S_regionTaiwan      0.21
## SST_a              -0.23     -0.46
## Tbl_binRecent      -0.01     -0.79      0.03
## S_regionIndianOc:LocDjib -0.10     -0.19      0.41  0.01
## S_regionFrPoly:LocRaia   0.01      0.50     -0.06 -0.62
## S_regionFrPoly:LocTahaa  0.01      0.52     -0.05 -0.64
##              S_regionIndianOc:LocDjib S_regionFrPoly:LocRaia
## S_regionFrPoly
## S_regionIndianOc
## S_regionNewCal
## S_regionTaiwan
## SST_a
## Tbl_binRecent
## S_regionIndianOc:LocDjib

```

```
## S_regionFrPoly:LocRaia    -0.02
## S_regionFrPoly:LocTahaa   -0.02                0.64
```

```
summary(mod.inv.nest, cor=T)
```

```
##
## Call:
## lm(formula = log(InvSimpson) ~ S_region/Loc + SST_a + Tbl_bin,
##     data = alpha_vegan)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.91135 -0.16292 -0.00353  0.15351  0.87988
##
## Coefficients: (37 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.98073    0.67588   2.931  0.00383 **
## S_regionFrPoly      0.09044    0.12990   0.696  0.48721
## S_regionIndianOc    0.58212    0.17820   3.267  0.00131 **
## S_regionNewCal      0.27120    0.06778   4.001 9.26e-05 ***
## S_regionTaiwan      0.52688    0.22732   2.318  0.02161 *
## SST_a             -0.02811    0.02711  -1.037  0.30111
## Tbl_binRecent     -0.16619    0.17624  -0.943  0.34699
## S_regionAus:LocDjib      NA         NA      NA      NA
## S_regionFrPoly:LocDjib    NA         NA      NA      NA
## S_regionIndianOc:LocDjib -1.20175    0.09882 -12.161 < 2e-16 ***
## S_regionNewCal:LocDjib    NA         NA      NA      NA
## S_regionTaiwan:LocDjib    NA         NA      NA      NA
## S_regionAus:LocMoorea     NA         NA      NA      NA
## S_regionFrPoly:LocMoorea  NA         NA      NA      NA
## S_regionIndianOc:LocMoorea NA         NA      NA      NA
## S_regionNewCal:LocMoorea  NA         NA      NA      NA
## S_regionTaiwan:LocMoorea  NA         NA      NA      NA
## S_regionAus:LocNewCal     NA         NA      NA      NA
## S_regionFrPoly:LocNewCal  NA         NA      NA      NA
## S_regionIndianOc:LocNewCal NA         NA      NA      NA
## S_regionNewCal:LocNewCal  NA         NA      NA      NA
## S_regionTaiwan:LocNewCal  NA         NA      NA      NA
## S_regionAus:LocOman       NA         NA      NA      NA
## S_regionFrPoly:LocOman    NA         NA      NA      NA
## S_regionIndianOc:LocOman  NA         NA      NA      NA
## S_regionNewCal:LocOman    NA         NA      NA      NA
## S_regionTaiwan:LocOman    NA         NA      NA      NA
## S_regionAus:LocRaia       NA         NA      NA      NA
## S_regionFrPoly:LocRaia    -0.09317    0.17649  -0.528  0.59821
## S_regionIndianOc:LocRaia  NA         NA      NA      NA
## S_regionNewCal:LocRaia    NA         NA      NA      NA
## S_regionTaiwan:LocRaia    NA         NA      NA      NA
## S_regionAus:LocTahaa      NA         NA      NA      NA
## S_regionFrPoly:LocTahaa   -0.08603    0.16952  -0.507  0.61246
## S_regionIndianOc:LocTahaa NA         NA      NA      NA
## S_regionNewCal:LocTahaa   NA         NA      NA      NA
## S_regionTaiwan:LocTahaa   NA         NA      NA      NA
## S_regionAus:LocTahiti     NA         NA      NA      NA
```

```

## S_regionFrPoly:LocTahiti      NA      NA      NA      NA
## S_regionIndianOc:LocTahiti    NA      NA      NA      NA
## S_regionNewCal:LocTahiti      NA      NA      NA      NA
## S_regionTaiwan:LocTahiti      NA      NA      NA      NA
## S_regionAus:LocTaiwan         NA      NA      NA      NA
## S_regionFrPoly:LocTaiwan      NA      NA      NA      NA
## S_regionIndianOc:LocTaiwan    NA      NA      NA      NA
## S_regionNewCal:LocTaiwan      NA      NA      NA      NA
## S_regionTaiwan:LocTaiwan      NA      NA      NA      NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.309 on 177 degrees of freedom
## Multiple R-squared:  0.5722, Adjusted R-squared:  0.5504
## F-statistic: 26.3 on 9 and 177 DF, p-value: < 2.2e-16
##
## Correlation of Coefficients:
##              (Intercept) S_regionFrPoly S_regionIndianOc
## S_regionFrPoly      0.43
## S_regionIndianOc    0.88      0.47
## S_regionNewCal      0.20      0.28      0.34
## S_regionTaiwan      0.45      0.66      0.46
## SST_a              -1.00     -0.45     -0.90
## Tbl_binRecent      -0.03     -0.53     -0.02
## S_regionIndianOc:LocDjib -0.41     -0.19     -0.63
## S_regionFrPoly:LocRaia   0.06      0.03      0.05
## S_regionFrPoly:LocTahaa  0.05      0.02      0.05
##              S_regionNewCal S_regionTaiwan SST_a Tbl_binRecent
## S_regionFrPoly
## S_regionIndianOc
## S_regionNewCal
## S_regionTaiwan      0.21
## SST_a              -0.23      -0.46
## Tbl_binRecent      -0.01      -0.79      0.03
## S_regionIndianOc:LocDjib -0.10      -0.19      0.41  0.01
## S_regionFrPoly:LocRaia   0.01      0.50     -0.06 -0.62
## S_regionFrPoly:LocTahaa  0.01      0.52     -0.05 -0.64
##              S_regionIndianOc:LocDjib S_regionFrPoly:LocRaia
## S_regionFrPoly
## S_regionIndianOc
## S_regionNewCal
## S_regionTaiwan
## SST_a
## Tbl_binRecent
## S_regionIndianOc:LocDjib
## S_regionFrPoly:LocRaia  -0.02
## S_regionFrPoly:LocTahaa -0.02      0.64

```

#Step 3.1: Check if missing cells for some interactions

```
coef(lm(Richness ~ S_region/Loc + SST_a + Tbl_bin, data = alpha_vegan))
```

```

##              (Intercept)      S_regionFrPoly
##              653.4572293      -18.0270975

```

```

##          S_regionIndianOc          S_regionNewCal
##          156.3472742          -0.1773214
##          S_regionTaiwan          SST_a
##          84.4311630          -20.5687919
##          Tbl_binRecent          S_regionAus:LocDjib
##          12.4476175          NA
##          S_regionFrPoly:LocDjib S_regionIndianOc:LocDjib
##          NA          -139.0199280
##          S_regionNewCal:LocDjib S_regionTaiwan:LocDjib
##          NA          NA
##          S_regionAus:LocMoorea S_regionFrPoly:LocMoorea
##          NA          NA
##          S_regionIndianOc:LocMoorea S_regionNewCal:LocMoorea
##          NA          NA
##          S_regionTaiwan:LocMoorea S_regionAus:LocNewCal
##          NA          NA
##          S_regionFrPoly:LocNewCal S_regionIndianOc:LocNewCal
##          NA          NA
##          S_regionNewCal:LocNewCal S_regionTaiwan:LocNewCal
##          NA          NA
##          S_regionAus:LocOman S_regionFrPoly:LocOman
##          NA          NA
##          S_regionIndianOc:LocOman S_regionNewCal:LocOman
##          NA          NA
##          S_regionTaiwan:LocOman S_regionAus:LocRaia
##          NA          NA
##          S_regionFrPoly:LocRaia S_regionIndianOc:LocRaia
##          -1.3781712          NA
##          S_regionNewCal:LocRaia S_regionTaiwan:LocRaia
##          NA          NA
##          S_regionAus:LocTahaa S_regionFrPoly:LocTahaa
##          NA          -9.8654732
##          S_regionIndianOc:LocTahaa S_regionNewCal:LocTahaa
##          NA          NA
##          S_regionTaiwan:LocTahaa S_regionAus:LocTahiti
##          NA          NA
##          S_regionFrPoly:LocTahiti S_regionIndianOc:LocTahiti
##          NA          NA
##          S_regionNewCal:LocTahiti S_regionTaiwan:LocTahiti
##          NA          NA
##          S_regionAus:LocTaiwan S_regionFrPoly:LocTaiwan
##          NA          NA
##          S_regionIndianOc:LocTaiwan S_regionNewCal:LocTaiwan
##          NA          NA
##          S_regionTaiwan:LocTaiwan
##          NA

```

```
coef(lm(Richness ~ Loc + SST_a + Tbl_bin, data = alpha_vegan))
```

```

## (Intercept)      LocDjib      LocMoorea      LocNewCal      LocOman
## 653.4572293    17.3273461   -18.0270975   -0.1773214    156.3472742
##      LocRaia      LocTahaa      LocTahiti      LocTaiwan      SST_a
## -6.9576512   -15.4449532   -5.5794800    96.8787804   -20.5687919
## Tbl_binRecent

```

```
##          NA
```

```
coef(lm(Richness ~ S_region + SST_a + Tbl_bin, data = alpha_vegan))
```

```
##      (Intercept)  S_regionFrPoly S_regionIndianOc  S_regionNewCal
##      262.078028    -51.834461    -1.406845    -9.260449
##      S_regionTaiwan      SST_a      Tbl_binRecent
##      32.450893      -4.844437      6.038115
```

```
coef(lm(Shannon ~ S_region/Loc + SST_a + Tbl_bin, data = alpha_vegan))
```

```
##      (Intercept)          S_regionFrPoly
##      4.86890144          -0.12643332
##      S_regionIndianOc      S_regionNewCal
##      0.89017117          -0.05813085
##      S_regionTaiwan          SST_a
##      0.65306332          -0.10978413
##      Tbl_binRecent      S_regionAus:LocDjib
##      -0.16171212          NA
##      S_regionFrPoly:LocDjib  S_regionIndianOc:LocDjib
##      NA          -1.55905340
##      S_regionNewCal:LocDjib  S_regionTaiwan:LocDjib
##      NA          NA
##      S_regionAus:LocMoorea  S_regionFrPoly:LocMoorea
##      NA          NA
##      S_regionIndianOc:LocMoorea  S_regionNewCal:LocMoorea
##      NA          NA
##      S_regionTaiwan:LocMoorea  S_regionAus:LocNewCal
##      NA          NA
##      S_regionFrPoly:LocNewCal  S_regionIndianOc:LocNewCal
##      NA          NA
##      S_regionNewCal:LocNewCal  S_regionTaiwan:LocNewCal
##      NA          NA
##      S_regionAus:LocOman      S_regionFrPoly:LocOman
##      NA          NA
##      S_regionIndianOc:LocOman  S_regionNewCal:LocOman
##      NA          NA
##      S_regionTaiwan:LocOman      S_regionAus:LocRaia
##      NA          NA
##      S_regionFrPoly:LocRaia  S_regionIndianOc:LocRaia
##      0.07499862          NA
##      S_regionNewCal:LocRaia  S_regionTaiwan:LocRaia
##      NA          NA
##      S_regionAus:LocTahaa  S_regionFrPoly:LocTahaa
##      NA          0.05365529
##      S_regionIndianOc:LocTahaa  S_regionNewCal:LocTahaa
##      NA          NA
##      S_regionTaiwan:LocTahaa  S_regionAus:LocTahiti
##      NA          NA
##      S_regionFrPoly:LocTahiti  S_regionIndianOc:LocTahiti
##      NA          NA
##      S_regionNewCal:LocTahiti  S_regionTaiwan:LocTahiti
##      NA          NA
```

```
##      S_regionAus:LocTaiwan      S_regionFrPoly:LocTaiwan
##                                NA                                NA
## S_regionIndianOc:LocTaiwan      S_regionNewCal:LocTaiwan
##                                NA                                NA
##      S_regionTaiwan:LocTaiwan
##                                NA
```

```
coef(lm(log(InvSimpson) ~ S_region/Loc + SST_a + Tbl_bin, data = alpha_vegan))
```

```
##              (Intercept)              S_regionFrPoly
##              1.98072683              0.09043919
##              S_regionIndianOc              S_regionNewCal
##              0.58211886              0.27119599
##              S_regionTaiwan              SST_a
##              0.52688078              -0.02811432
##              Tbl_binRecent              S_regionAus:LocDjib
##              -0.16618725              NA
##      S_regionFrPoly:LocDjib      S_regionIndianOc:LocDjib
##              NA              -1.20174931
##      S_regionNewCal:LocDjib      S_regionTaiwan:LocDjib
##              NA              NA
##      S_regionAus:LocMoorea      S_regionFrPoly:LocMoorea
##              NA              NA
## S_regionIndianOc:LocMoorea      S_regionNewCal:LocMoorea
##              NA              NA
##      S_regionTaiwan:LocMoorea      S_regionAus:LocNewCal
##              NA              NA
##      S_regionFrPoly:LocNewCal      S_regionIndianOc:LocNewCal
##              NA              NA
##      S_regionNewCal:LocNewCal      S_regionTaiwan:LocNewCal
##              NA              NA
##              S_regionAus:LocOman      S_regionFrPoly:LocOman
##              NA              NA
##      S_regionIndianOc:LocOman      S_regionNewCal:LocOman
##              NA              NA
##      S_regionTaiwan:LocOman      S_regionAus:LocRaia
##              NA              NA
##      S_regionFrPoly:LocRaia      S_regionIndianOc:LocRaia
##              -0.09317470              NA
##      S_regionNewCal:LocRaia      S_regionTaiwan:LocRaia
##              NA              NA
##              S_regionAus:LocTahaa      S_regionFrPoly:LocTahaa
##              NA              -0.08602515
## S_regionIndianOc:LocTahaa      S_regionNewCal:LocTahaa
##              NA              NA
##      S_regionTaiwan:LocTahaa      S_regionAus:LocTahiti
##              NA              NA
##      S_regionFrPoly:LocTahiti      S_regionIndianOc:LocTahiti
##              NA              NA
##      S_regionNewCal:LocTahiti      S_regionTaiwan:LocTahiti
##              NA              NA
##              S_regionAus:LocTaiwan      S_regionFrPoly:LocTaiwan
##              NA              NA
## S_regionIndianOc:LocTaiwan      S_regionNewCal:LocTaiwan
```

```

##           NA           NA
## S_regionTaiwan:LocTaiwan
##           NA

#Step 3.2: Drop interactions = NA

dd1 <- transform(na.omit(alpha_vegan))

library(nlme)

#Step 4: build gls

rich_gls <- gls(Richness ~ S_region + SST_a + Tbl_bin, data = alpha_vegan)

rich_gls_nest <- gls(Richness ~ S_region/Loc + SST_a + Tbl_bin, data = alpha_vegan, na.action = "na.omit")

##Need to manually specify "nesting" as interaction effects
##rich_gls <- gls(Richness ~ S_region + S_region$IndianOc:Loc$Oman + S_region$IndianOc:Loc$Djib + S_region$IndianOc:Loc$Mali + S_region$IndianOc:Loc$Niger + S_region$IndianOc:Loc$Senegal + S_region$IndianOc:Loc$Togo + S_region$IndianOc:Loc$Zambia + S_region$IndianOc:Loc$Zimbabwe, data = alpha_vegan)

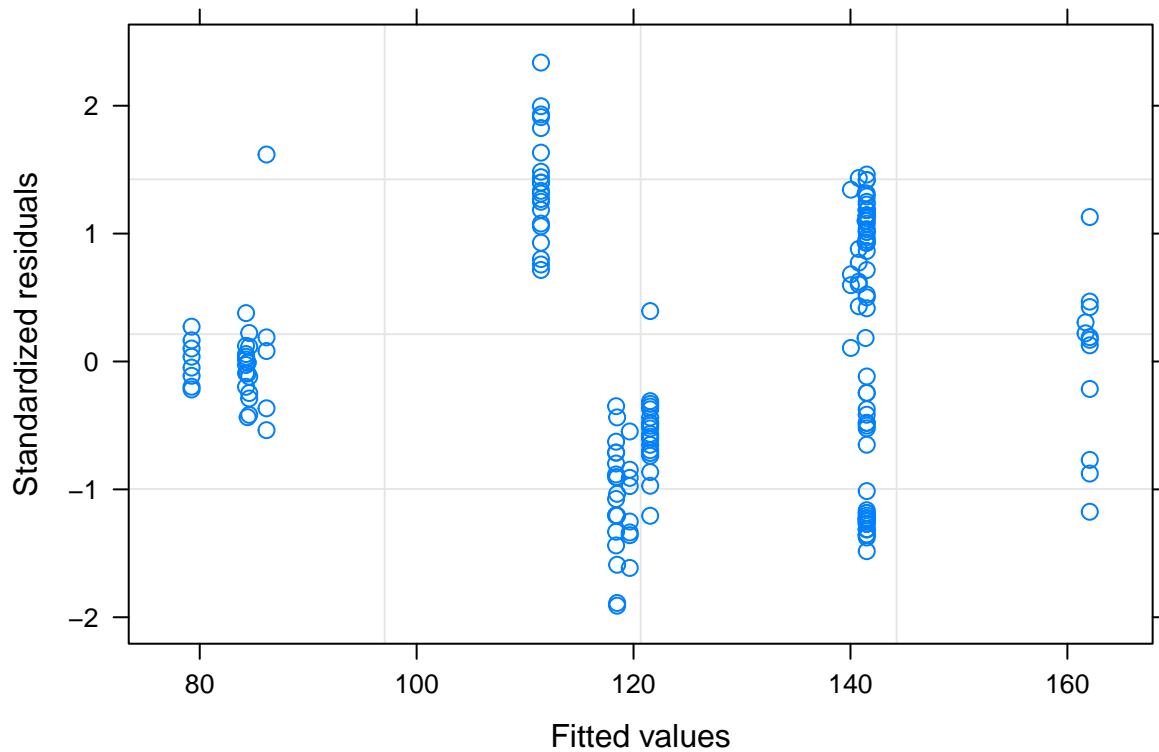
summary(rich_gls)

## Generalized least squares fit by REML
## Model: Richness ~ S_region + SST_a + Tbl_bin
## Data: alpha_vegan
##      AIC      BIC    logLik
## 1936.096 1961.639 -960.0478
##
## Coefficients:
##              Value Std.Error   t-value p-value
## (Intercept)  262.07803  93.13278   2.8140256  0.0054
## S_regionFrPoly -51.83446  19.34889  -2.6789374  0.0081
## S_regionIndianOc -1.40684  20.93064  -0.0672146  0.9465
## S_regionNewCal -9.26045  10.23044  -0.9051854  0.3666
## S_regionTaiwan  32.45089  27.74926   1.1694327  0.2438
## SST_a        -4.84444   3.73413  -1.2973411  0.1962
## Tbl_binRecent   6.03812  19.23893   0.3138488  0.7540
##
## Correlation:
##              (Intr) S_rgFP S_rgIO S_rgNC S_rgnT SST_a
## S_regionFrPoly   0.394
## S_regionIndianOc  0.883  0.462
## S_regionNewCal   0.173  0.266  0.356
## S_regionTaiwan   0.471  0.779  0.499  0.227
## SST_a           -0.998 -0.415 -0.903 -0.211 -0.486
## Tbl_binRecent    0.020 -0.729  0.018  0.004 -0.683 -0.020
##
## Standardized residuals:
##              Min      Q1      Med      Q3      Max
## -1.90958887 -0.82198659 -0.09423682  0.92977370  2.33745238
##
## Residual standard error: 46.8619
## Degrees of freedom: 187 total; 180 residual

```



```
plot(rich_gls)
```



```
##Nested
summary(rich_gls_nest)
```

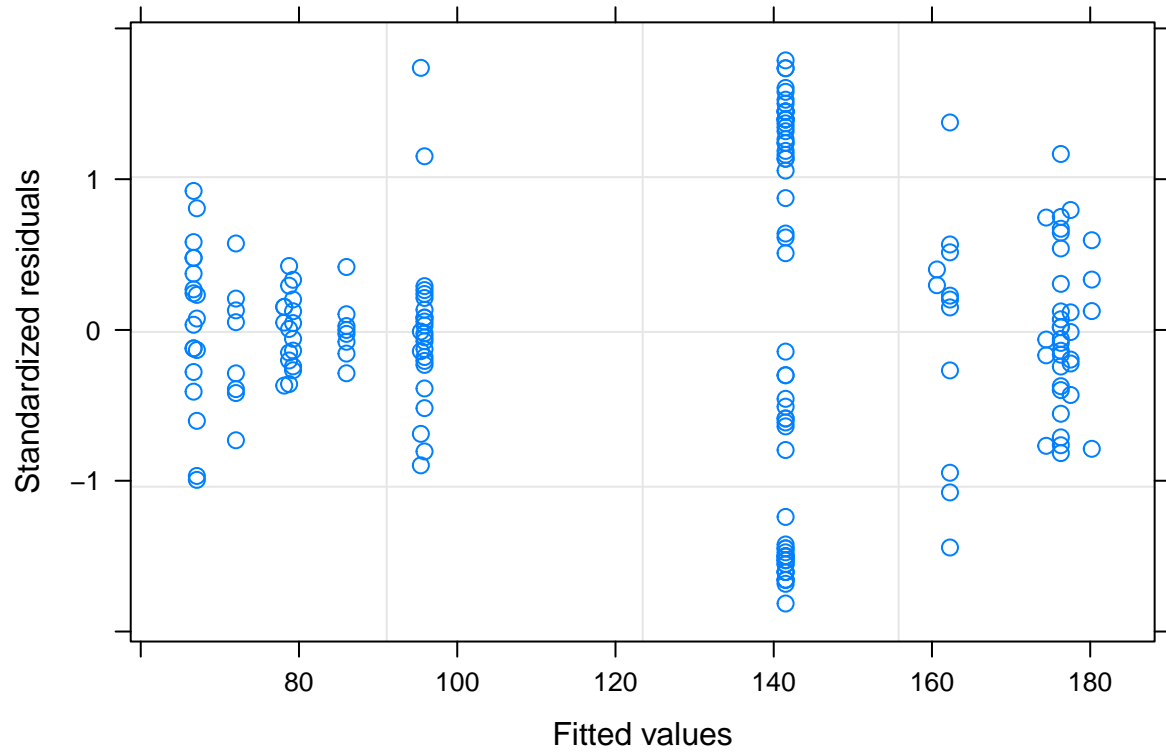
```
## Generalized least squares fit by REML
## Model: Richness ~ S_region/Loc + SST_a + Tbl_bin
## Data: alpha_vegan
##      AIC      BIC    logLik
## 1559.516 1724.248 -723.7579
##
## Variance function:
## Structure: Different standard deviations per stratum
## Formula: ~1 | Loc
## Parameter estimates:
##      NewCal      Oman      Moorea      Taiwan      Djib
##      1          1          1          1          1
##      Tahaa      Tahiti      Raia Aus_GBR_Heron
##      1          1          1          1
##
## Coefficients:
##              Value Std.Error   t-value p-value
## (Intercept)   653.4572  83.78618   7.799105  0.0000
## S_regionFrPoly -18.0271  16.10369  -1.119439  0.2649
## S_regionIndianOc 156.3473  22.09114   7.077375  0.0000
```

```

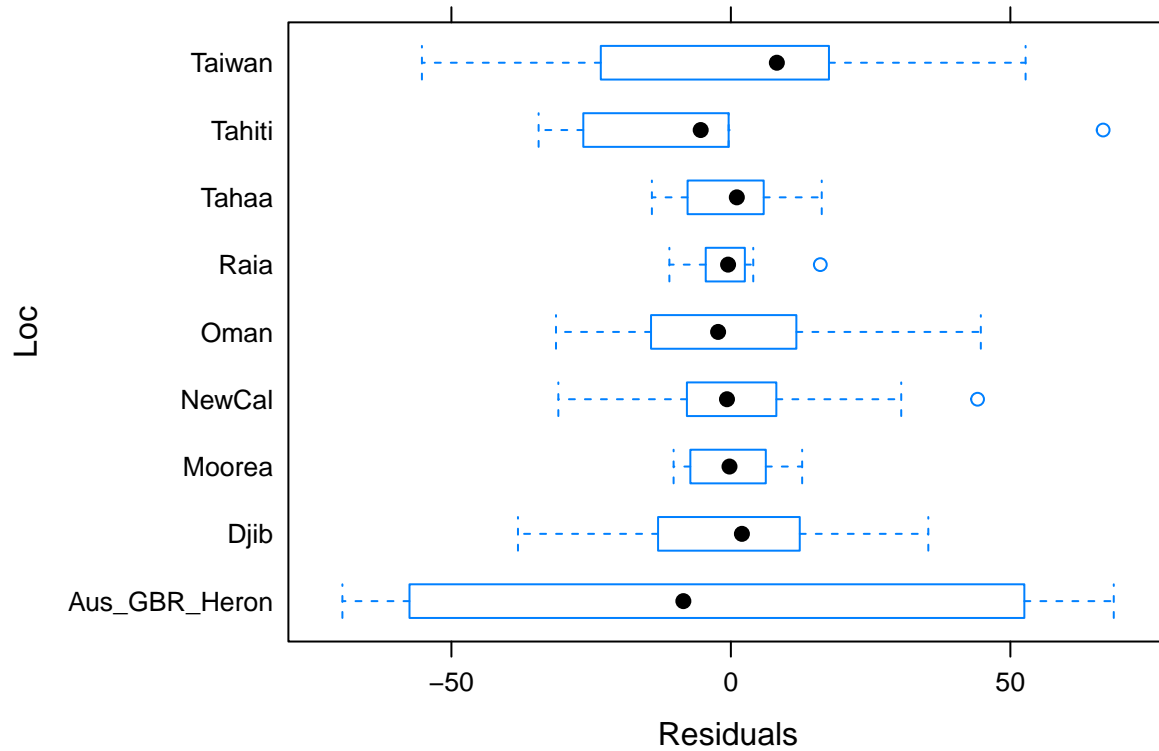
## S_regionNewCal          -0.1773   8.40269  -0.021103  0.9832
## S_regionTaiwan          84.4312  28.18048   2.996087  0.0032
## SST_a                  -20.5688   3.36058  -6.120611  0.0000
## Tbl_binRecent          12.4476  21.84821   0.569732  0.5698
## S_regionIndianOc:LocDjib -139.0199 12.25029 -11.348297  0.0000
## S_regionFrPoly:LocRaia  -1.3782  21.87913  -0.062990  0.9499
## S_regionFrPoly:LocTahaa  -9.8655  21.01449  -0.469461  0.6395
##
## Correlation:
##              (Intr) S_rgFP S_rgIO S_rgNC S_rgnT SST_a  Tbl_bR
## S_regionFrPoly      0.430
## S_regionIndianOc     0.885  0.470
## S_regionNewCal       0.197  0.279  0.335
## S_regionTaiwan       0.453  0.665  0.455  0.207
## SST_a               -0.998 -0.449 -0.899 -0.231 -0.464
## Tbl_binRecent       -0.028 -0.534 -0.025 -0.006 -0.788  0.028
## S_regionIndianOc:LocDjib -0.414 -0.186 -0.631 -0.096 -0.193  0.415  0.011
## S_regionFrPoly:LocRaia  0.060  0.027  0.054  0.014  0.504 -0.060 -0.616
## S_regionFrPoly:LocTahaa  0.055  0.025  0.049  0.013  0.521 -0.055 -0.641
##              S_IO:L S_FP:LR
## S_regionFrPoly
## S_regionIndianOc
## S_regionNewCal
## S_regionTaiwan
## SST_a
## Tbl_binRecent
## S_regionIndianOc:LocDjib
## S_regionFrPoly:LocRaia  -0.025
## S_regionFrPoly:LocTahaa -0.023  0.642
##
## Standardized residuals:
##              Min          Q1          Med          Q3          Max
## -1.814166e+00 -4.248386e-01 -4.859413e-14  4.936208e-01  1.788063e+00
##
## Residual standard error: 38.30961
## Degrees of freedom: 187 total; 140 residual

```

```
plot(rich_gls_nest) ##st residual dotplot
```



```
plot(rich_gls_nest, Loc ~ resid(.)) ##reg residual boxplot
```



##Wt per location

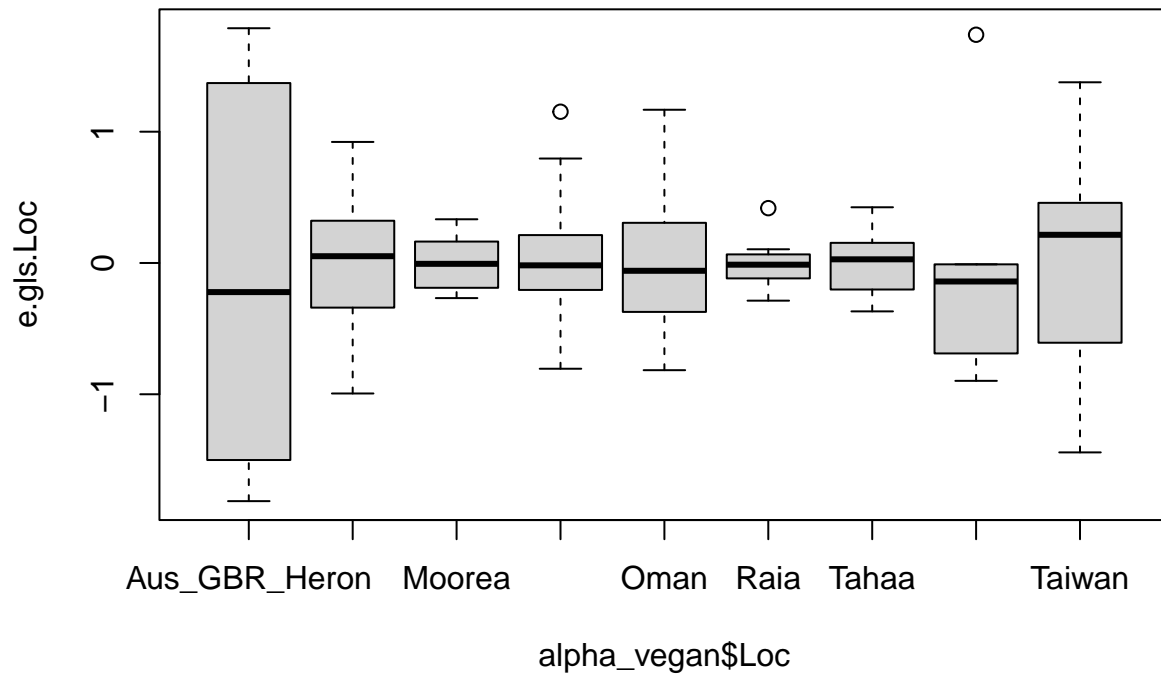
```
weight.per.loc <- attr(rich_gls_nest$model$varStruct, "weights")
weight.per.loc
```

```
##      NewCal      NewCal      Oman      Oman      Oman
##      1          1          1          1          1
##      Oman      Oman      Oman      NewCal     NewCal
##      1          1          1          1          1
##      Oman      Oman      NewCal     NewCal     NewCal
##      1          1          1          1          1
##      NewCal     NewCal     Oman      NewCal     NewCal
##      1          1          1          1          1
##      NewCal     NewCal     Oman      Oman      NewCal
##      1          1          1          1          1
##      NewCal     NewCal     NewCal     NewCal     Oman
##      1          1          1          1          1
##      Oman      Oman      Oman      NewCal     NewCal
##      1          1          1          1          1
##      Oman      Oman      Oman      Oman      Oman
##      1          1          1          1          1
##      Oman      Moorea     Taiwan     Taiwan     Taiwan
##      1          1          1          1          1
##      Taiwan     Taiwan     Djib      Tahaa      Tahaa
##      1          1          1          1          1
##      Tahaa      Tahaa     Tahiti     NewCal     Taiwan
##      1          1          1          1          1
```

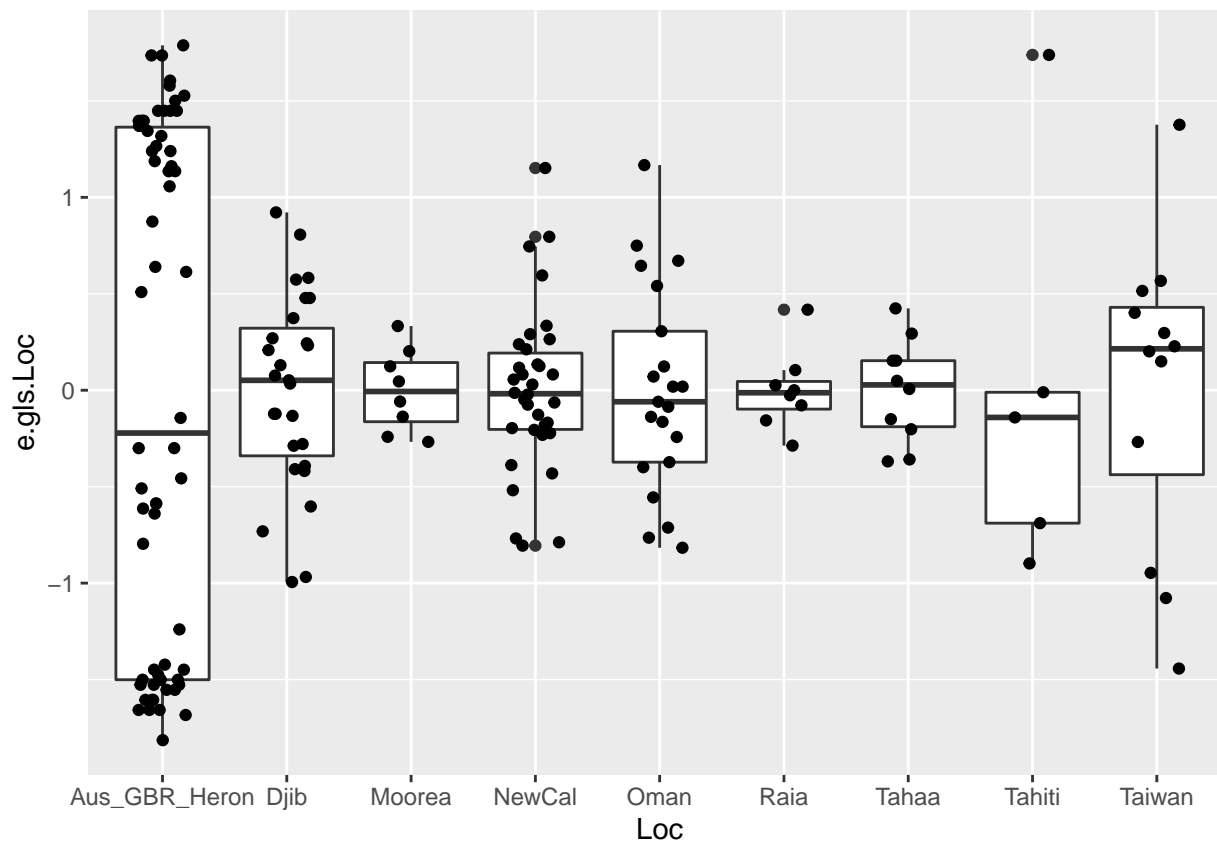
[illegible]

```
##Boxplots st residual jitterpoints
e.gls.Loc <- resid(rich_gls_nest, type = "normalized")

boxplot(e.gls.Loc~alpha_vegan$Loc)
```



```
ggplot(meta, aes(x = Loc, y = e.gls.Loc)) + geom_boxplot() +
geom_jitter(height = 0, width = .2)
```



```
qqnorm(e.gls.Loc,col="firebrick",abline = c(0,1))
```

```
## Warning in plot.window(...): "abline" is not a graphical parameter
```

```
## Warning in plot.xy(xy, type, ...): "abline" is not a graphical parameter
```

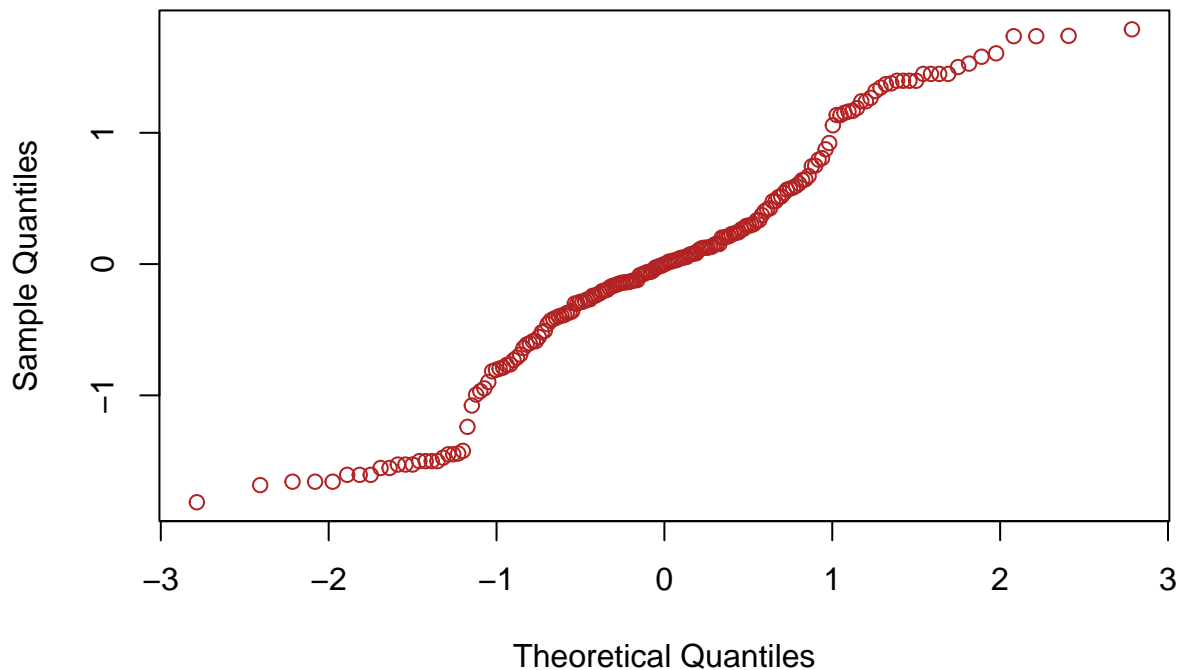
```
## Warning in axis(side = side, at = at, labels = labels, ...): "abline" is not a  
## graphical parameter
```

```
## Warning in axis(side = side, at = at, labels = labels, ...): "abline" is not a  
## graphical parameter
```

```
## Warning in box(...): "abline" is not a graphical parameter
```

```
## Warning in title(...): "abline" is not a graphical parameter
```

Normal Q-Q Plot



```
#Step 4.1: remove locations bi/trimodal distributions
##As QQ plot with regular lm and gls above do not differ significantly, the issue may lie in locations v

#Move into Vegan
set.seed(423542)

ps9 = subset_samples(ps8, Spec == "P_dam")

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'

## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
```



```

## Also defined by 'tidytree'

ps10 = subset_samples(ps9, Loc != "Aus_GBR_Heron")

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
## Also defined by 'tidytree'

ps11 = subset_samples(ps10, Season != "Summer")

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
## Also defined by 'tidytree'

ps12 = subset_samples(ps11, Tbl_bin != "Bleaching")

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
## Also defined by 'tidytree'

## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
## Also defined by 'tidytree'

```

```
ps13 = subset_samples(ps12, Loc != "NewCal")
```

```
## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
## Also defined by 'tidytree'
```

```
## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
```

```
## Also defined by 'tidytree'
```

```
## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
```

```
## Also defined by 'tidytree'
```

```
## Found more than one class "phylo" in cache; using the first, from namespace 'phyloseq'
```

```
## Also defined by 'tidytree'
```

```
colnames(sample_data(ps13))
```

```
## [1] "Loc"      "Yr"      "Spec"    "Exp_cond" "Code"
## [6] "Repro"    "Month"   "Season"  "S_region" "L_region"
## [11] "Exact.date" "Tbl_bin" "T_bleach" "DHW"      "DHW_cat"
## [16] "SST_a"    "Coord_X" "Coord_Y" "Primer"   "Pub"
## [21] "Note"
```

```
asv_css2 <- t(otu_table(ps13))
```

```
meta2 = as(sample_data(ps13), "data.frame")
```

```
Richness2 <- specnumber(asv_css2) #calculates richness from css
```

```
Shannon2 <- vegan::diversity(asv_css2, index = "shannon")
```

```
InvSimpson2 <- vegan::diversity(asv_css2, index = "invsimpson") #inverted simpson
#Beta <- vegdist(asv_css, "bray") #beta dissimilarity
```

```
alpha_vegan2 <- cbind(Richness2, Shannon2, InvSimpson2, meta2)
```

```
alpha_vegan2 <- rownames_to_column(alpha_vegan2, var = "id") %>% as_tibble()
```

```
head(alpha_vegan2)
```

```
## # A tibble: 6 x 25
```

```
##   id      Richness2 Shannon2 InvSimpson2 Loc      Yr Spec Exp_cond Code  Repro
##   <chr>      <int>    <dbl>    <dbl> <chr> <int> <chr> <chr>    <chr> <chr>
## 1 SRR5963~    173     2.39     6.46 Oman  2014 P_dam 31C    Om2    B
## 2 SRR5963~    161     2.17     4.66 Oman  2014 P_dam 31C    Om2    B
## 3 SRR5963~    155     2.32     5.33 Oman  2014 P_dam 31C    Om3    B
## 4 SRR5963~    188     2.18     4.15 Oman  2014 P_dam 31C    Om2    B
## 5 SRR5963~    177     2.18     4.52 Oman  2014 P_dam 31C    Om3    B
## 6 SRR5963~    202     2.54     7.09 Oman  2014 P_dam 31C    Om3    B
## # ... with 15 more variables: Month <chr>, Season <chr>, S_region <chr>,
## #   L_region <chr>, Exact.date <chr>, Tbl_bin <chr>, T_bleach <chr>, DHW <dbl>,
## #   DHW_cat <chr>, SST_a <dbl>, Coord_X <dbl>, Coord_Y <dbl>, Primer <chr>,
## #   Pub <chr>, Note <chr>
```

```

Locs <- c("Djib", "Oman", "Taiwan", "Moorea", "Tahiti", "Tahaa", "Raia")
alpha_vegan$Loc <- factor(alpha_vegan$Loc, levels = c("Djib", "Oman", "Taiwan", "Moorea", "Tahiti", "Tahaa", "Raia"))

DHW <- c("N", "Mod")
alpha_vegan$DHW_cat <- factor(alpha_vegan$DHW_cat, levels = c("N", "Mod"))

T_bl <- c("Recent", "Long")
alpha_vegan$Tbl_bin <- factor(alpha_vegan$Tbl_bin, levels = c("Recent", "Long"))

Season <- c("Winter", "Spring")
alpha_vegan$Season <- factor(alpha_vegan$Season, levels = c("Winter", "Spring"))

S_region <- c("IndianOc", "Taiwan", "FrPoly")
alpha_vegan$S_region <- factor(alpha_vegan$S_region, levels = c("IndianOc", "Taiwan", "Aus", "NewCal", "FrPoly"))

L_region <- c("IndianOc", "NPac", "EPac")
alpha_vegan$L_region <- factor(alpha_vegan$L_region, levels = c("IndianOc", "NPac", "EPac"))

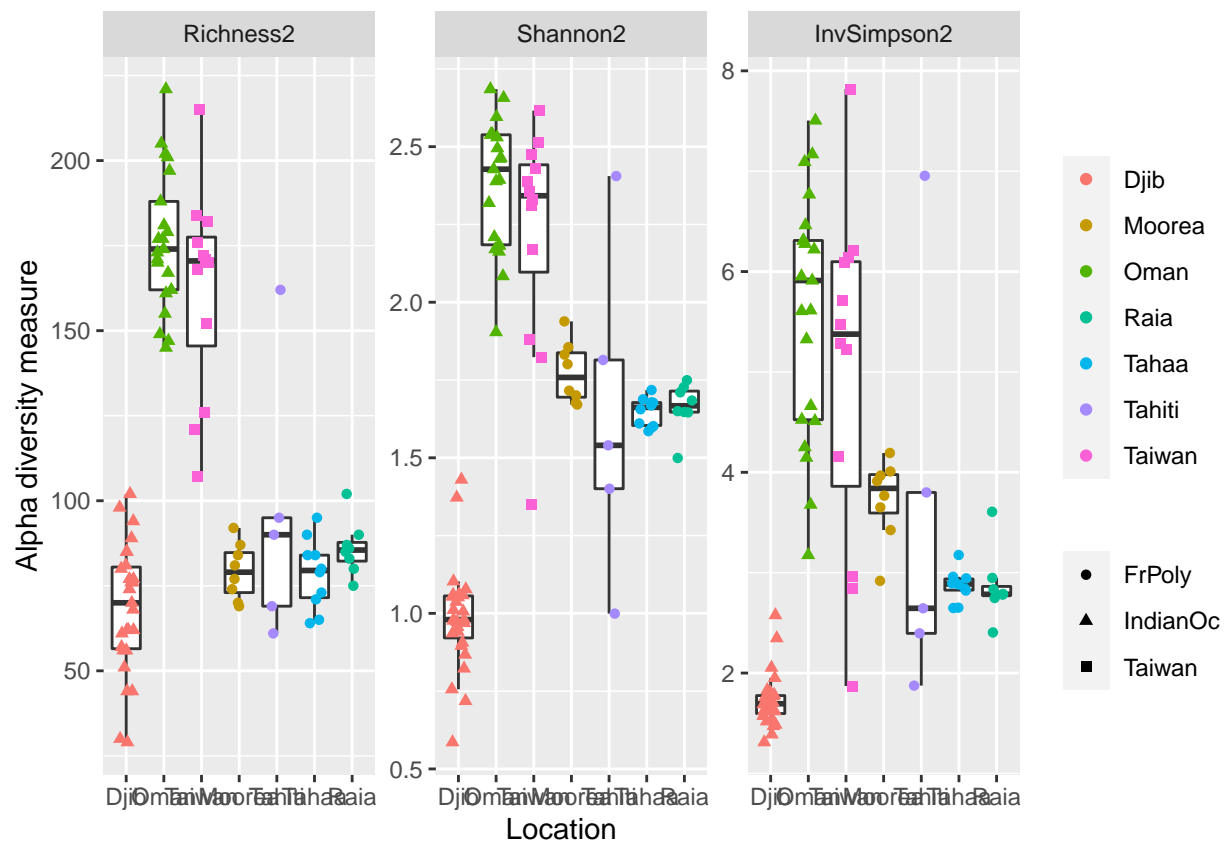
```

```
#Plot adiv measures by location
```

```

adiv_loc <- alpha_vegan2
adiv_loc %>%
gather(key = metric, value = value, c("Richness2", "Shannon2", "InvSimpson2")) %>%
mutate(metric = factor(metric, levels = c("Richness2", "Shannon2", "InvSimpson2"))) %>%
ggplot(aes(x = Loc, y = value)) +
geom_boxplot(outlier.color = NA) +
geom_jitter(aes(color = Loc, shape = S_region), height = 0, width = .2) +
labs(x = "Location", y = "Alpha diversity measure") +
facet_wrap(~ metric, scales = "free") +
theme(legend.title = element_blank()) + scale_x_discrete(limits = Locs)

```



```
ggsave("adiv_loc_region_pdam_28Nov2021.pdf")
```

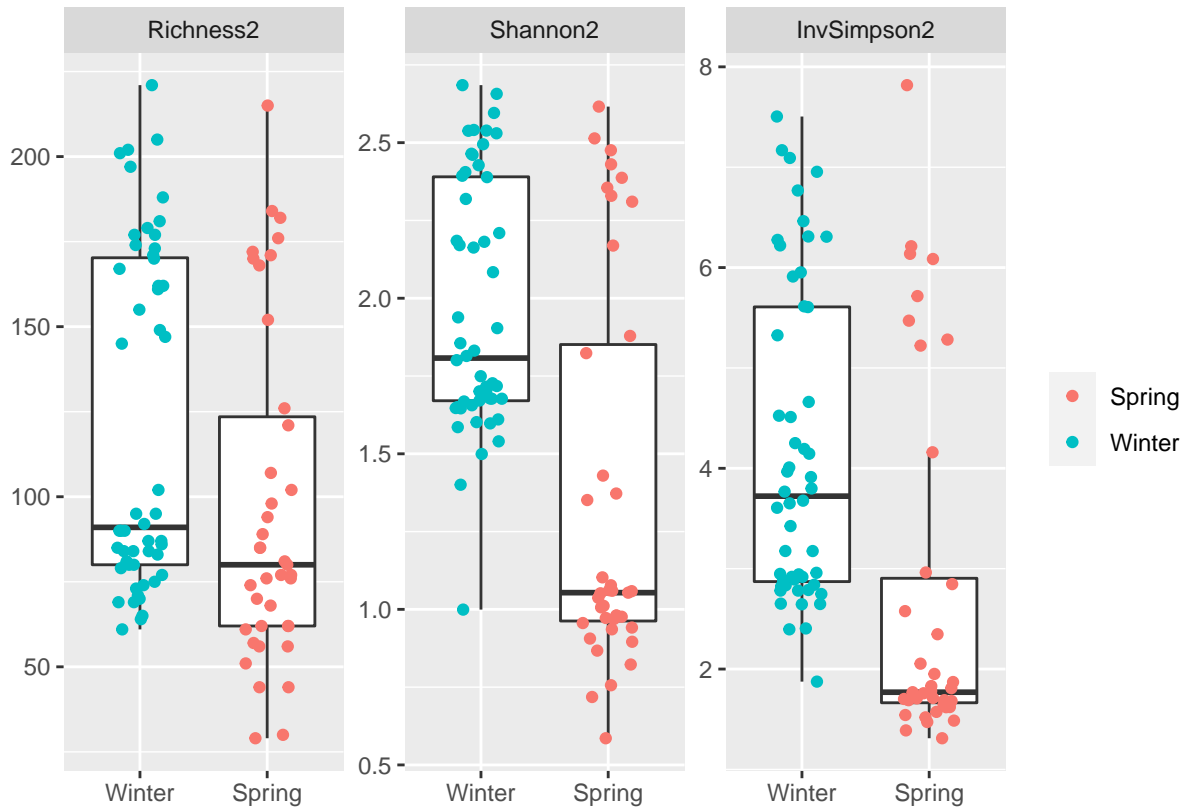
```
## Saving 6.5 x 4.5 in image
```

```
ggsave("adiv_loc_region_pdam_28Nov2021.png")
```

```
## Saving 6.5 x 4.5 in image
```

```
#Plot adiv measures by Season
```

```
adiv_sea <- alpha_vegan2
adiv_sea %>%
gather(key = metric, value = value, c("Richness2", "Shannon2", "InvSimpson2")) %>%
mutate(metric = factor(metric, levels = c("Richness2", "Shannon2", "InvSimpson2"))) %>%
ggplot(aes(x = Season, y = value)) +
geom_boxplot(outlier.color = NA) +
geom_jitter(aes(color = Season), height = 0, width = .2) +
labs(x = "", y = "") +
facet_wrap(~ metric, scales = "free") +
theme(legend.title = element_blank()) + scale_x_discrete(limits = Season)
```

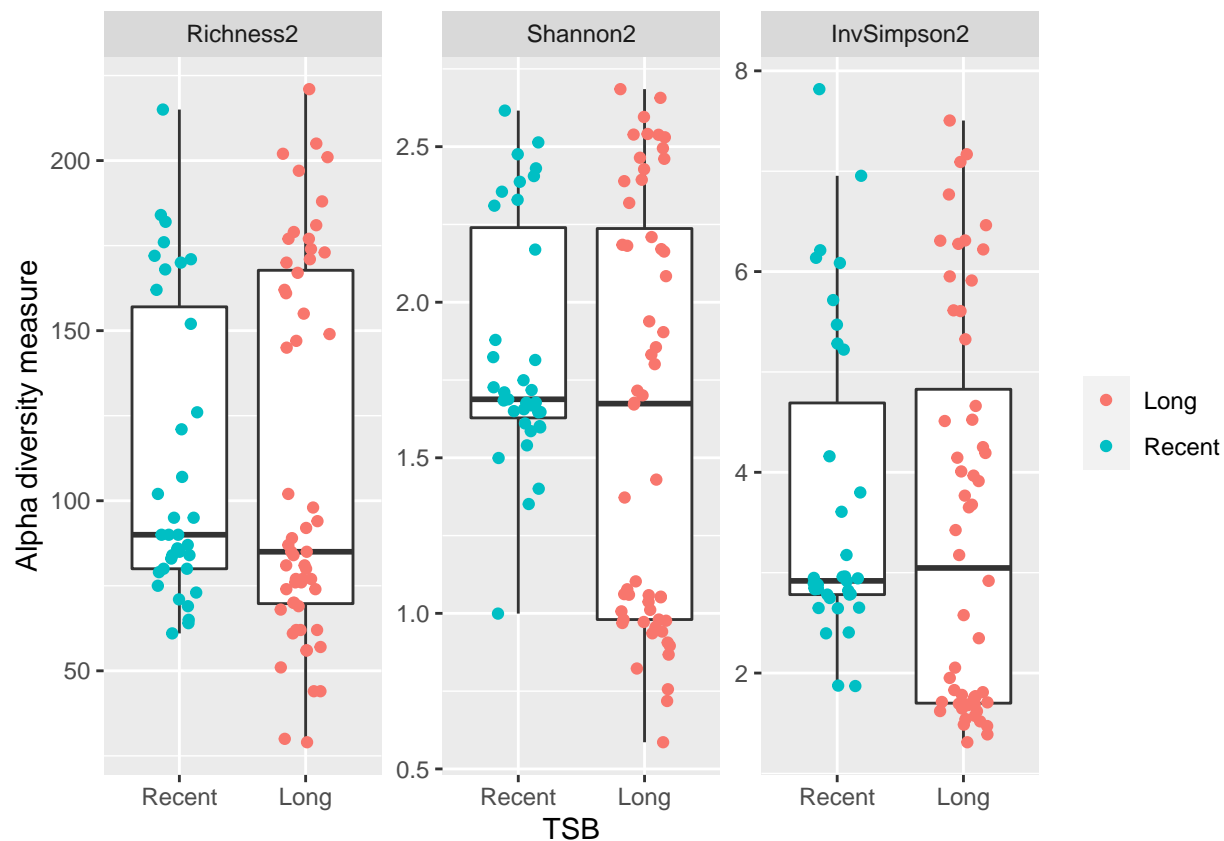


```
ggsave("adiv_sea_pdam_16Sep2021.pdf")
```

```
## Saving 6.5 x 4.5 in image
```

```
#Plot adiv measures by T_bleach categories set
```

```
adiv_tbl <- alpha_vegan2
adiv_tbl %>%
gather(key = metric, value = value, c("Richness2", "Shannon2", "InvSimpson2")) %>%
mutate(metric = factor(metric, levels = c("Richness2", "Shannon2", "InvSimpson2"))) %>%
ggplot(aes(x = Tbl_bin, y = value)) +
geom_boxplot(outlier.color = NA) +
geom_jitter(aes(color = Tbl_bin), height = 0, width = .2) +
labs(x = "TSB", y = "Alpha diversity measure") +
facet_wrap(~ metric, scales = "free") +
theme(legend.title = element_blank()) + scale_x_discrete(limits = T_bl)
```



```
ggsave("adiv_tbl_pdam_28Nov2021.pdf")
```

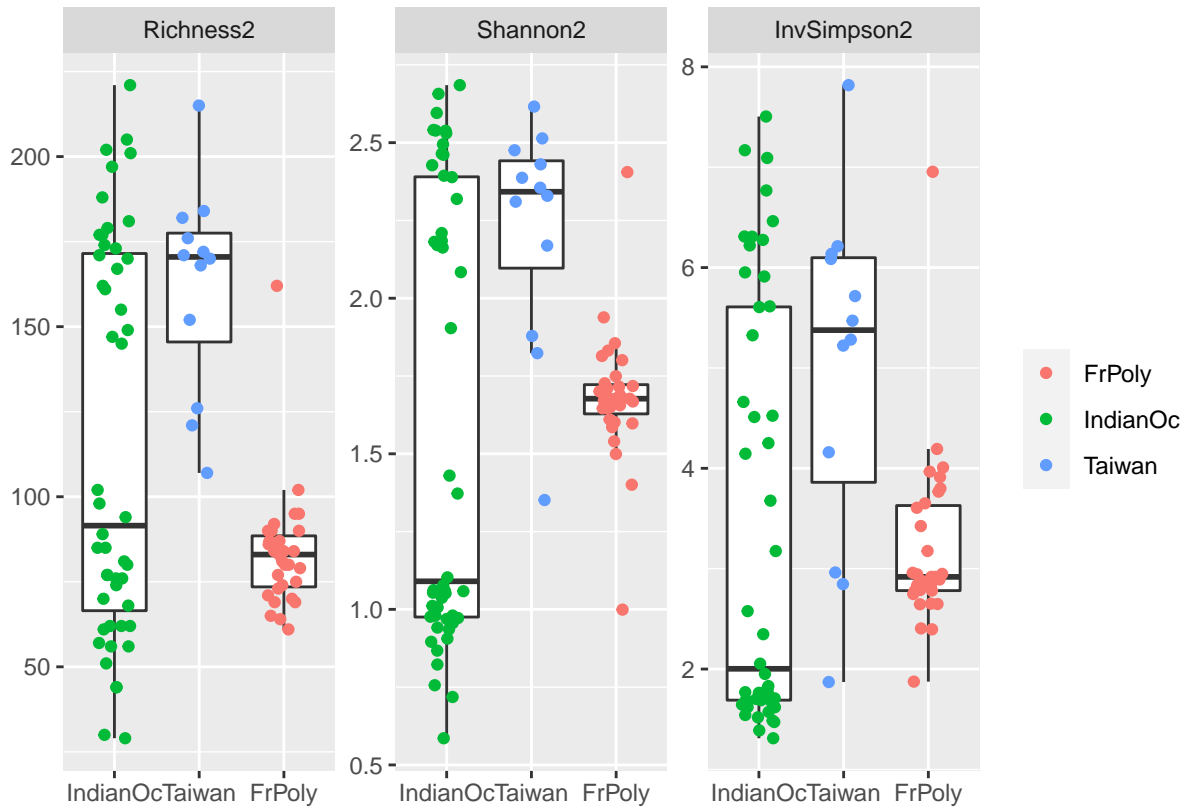
```
## Saving 6.5 x 4.5 in image
```

```
ggsave("adiv_tbl_pdam_28Nov2021.png")
```

```
## Saving 6.5 x 4.5 in image
```

```
#Plot adiv measures by small region categories set
```

```
adiv_sreg <- alpha_vegan2
adiv_sreg %>%
gather(key = metric, value = value, c("Richness2", "Shannon2", "InvSimpson2")) %>%
mutate(metric = factor(metric, levels = c("Richness2", "Shannon2", "InvSimpson2"))) %>%
ggplot(aes(x = S_region, y = value)) +
geom_boxplot(outlier.color = NA) +
geom_jitter(aes(color = S_region), height = 0, width = .2) +
labs(x = "", y = "") +
facet_wrap(~ metric, scales = "free") +
theme(legend.title = element_blank()) + scale_x_discrete(limits = S_region)
```

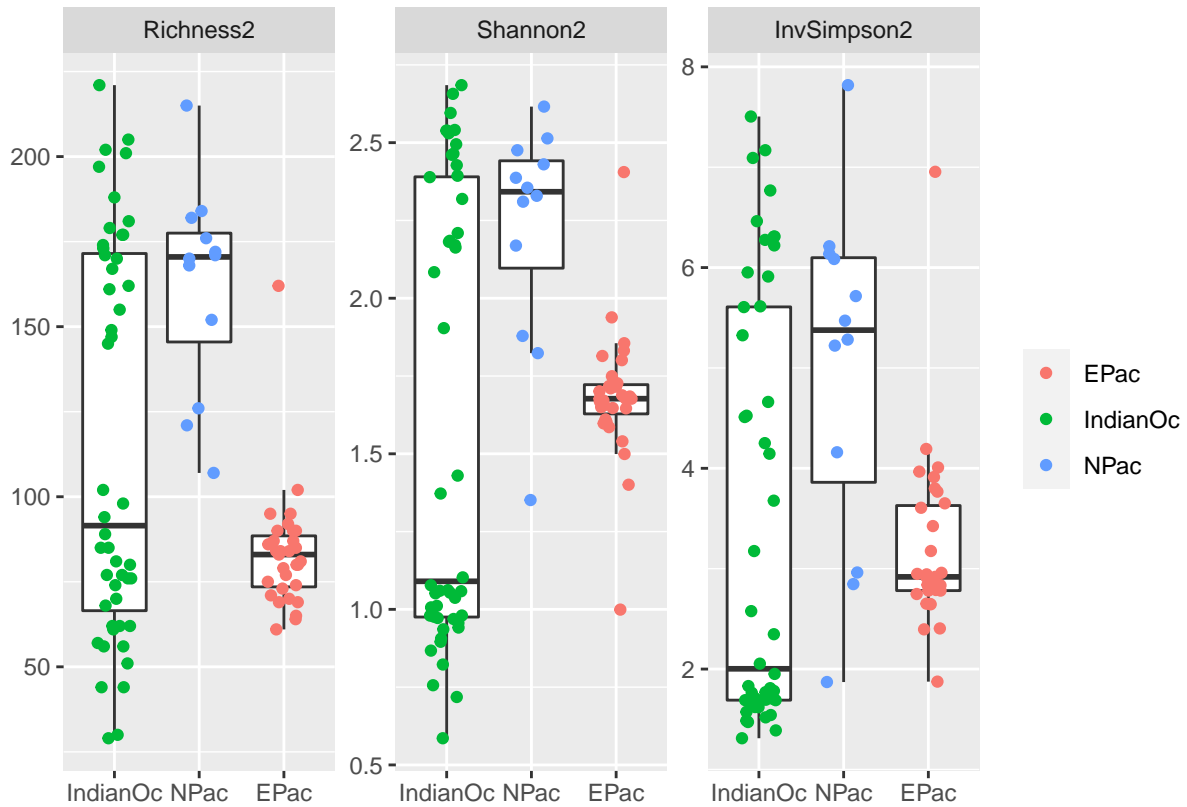


```
ggsave("adiv_sreg_pdam_1Nov2021.pdf")
```

```
## Saving 6.5 x 4.5 in image
```

```
#Plot adiv measures by large region categories set
```

```
adiv_lreg <- alpha_vegan2
adiv_lreg %>%
gather(key = metric, value = value, c("Richness2", "Shannon2", "InvSimpson2")) %>%
mutate(metric = factor(metric, levels = c("Richness2", "Shannon2", "InvSimpson2"))) %>%
ggplot(aes(x = L_region, y = value)) +
geom_boxplot(outlier.color = NA) +
geom_jitter(aes(color = L_region), height = 0, width = .2) +
labs(x = "", y = "") +
facet_wrap(~ metric, scales = "free") +
theme(legend.title = element_blank()) + scale_x_discrete(limits = L_region)
```



```
ggsave("adiv_lreg_pdam_1Nov2021.pdf")
```

```
## Saving 6.5 x 4.5 in image
```

```
##New lm models, for comparison, with amended dataset without multimodal locations
```

```
rich.lm.reduced <- lm(Richness2 ~ S_region/Loc + SST_a + Tbl_bin, na.action = "na.exclude", data = alpha_
```

```
summary(rich.lm.reduced)
```

```
##
```

```
## Call:
```

```
## lm(formula = Richness2 ~ S_region/Loc + SST_a + Tbl_bin, data = alpha_vegan2,  
##     na.action = "na.exclude")
```

```
##
```

```
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max  
## -54.847  -9.591   0.637  11.152  66.600
```

```
##
```

```
## Coefficients: (15 not defined because of singularities)
```

```
##              Estimate Std. Error t value Pr(>|t|)  
## (Intercept)    -230.44    922.57  -0.250   0.8034  
## S_regionIndianOc    -36.62    77.15  -0.475   0.6362  
## S_regionTaiwan     46.63    60.51   0.771   0.4431  
## SST_a             11.45    34.12   0.336   0.7379
```

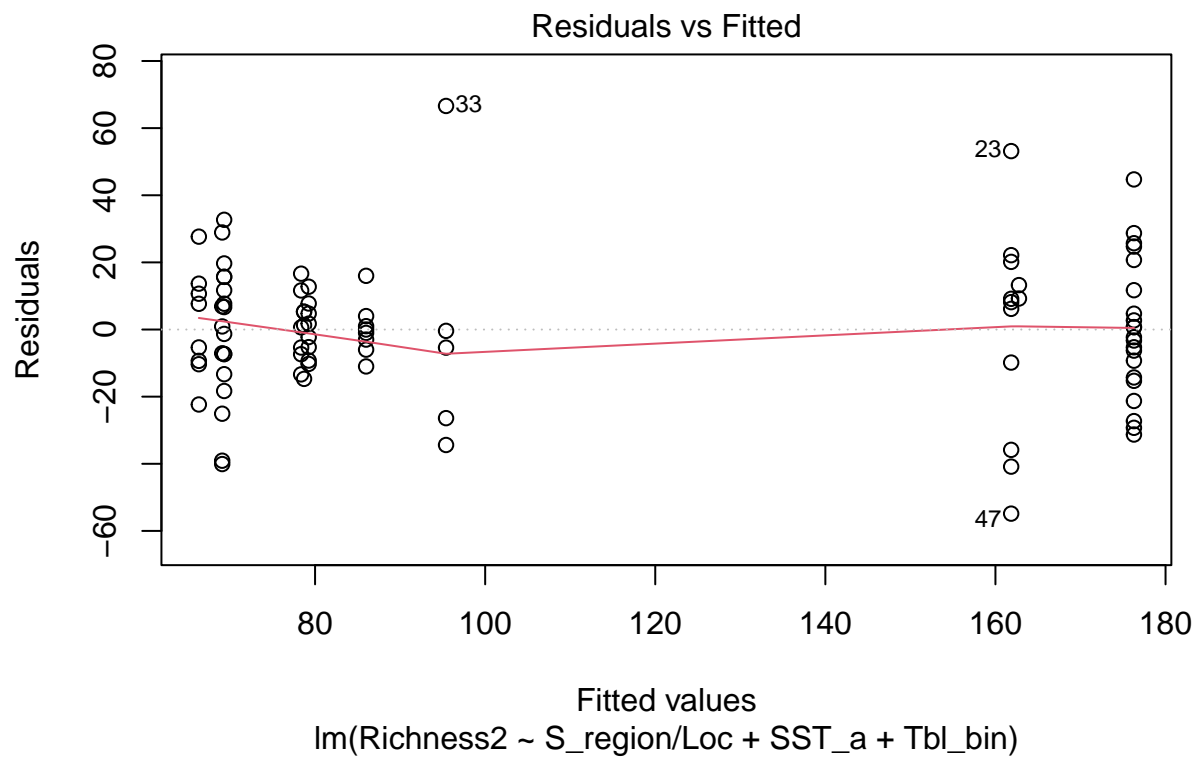


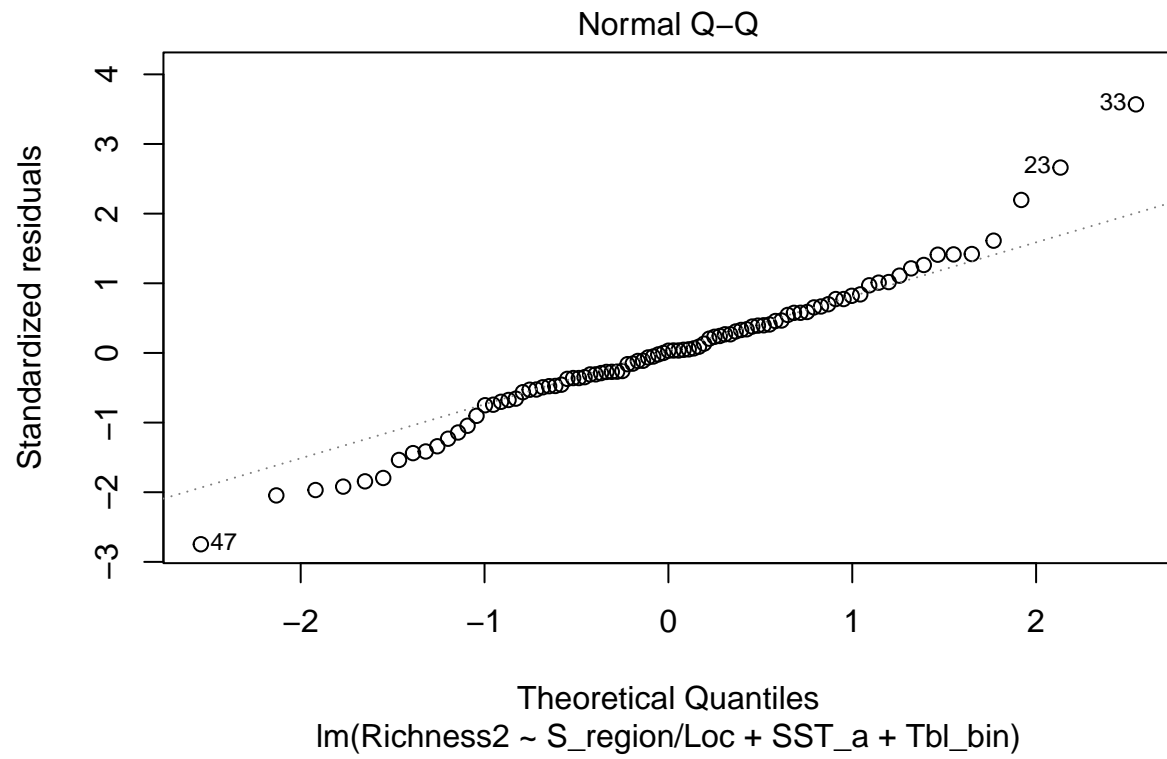
```

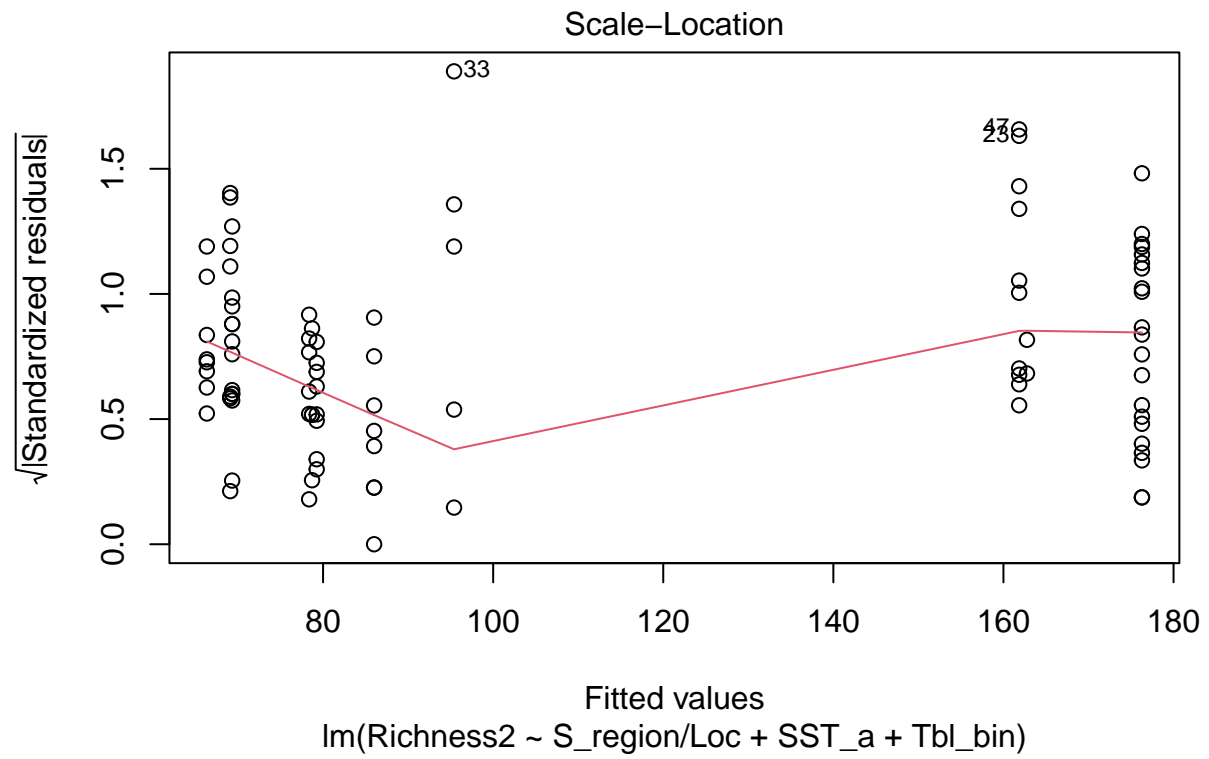
## Tbl_binRecent          18.21      13.38      1.361      0.1773
## S_regionFrPoly:LocMoorea      NA          NA          NA          NA
## S_regionIndianOc:LocMoorea     NA          NA          NA          NA
## S_regionTaiwan:LocMoorea       NA          NA          NA          NA
## S_regionFrPoly:LocOman        NA          NA          NA          NA
## S_regionIndianOc:LocOman      90.60      51.95      1.744      0.0849 .
## S_regionTaiwan:LocOman        NA          NA          NA          NA
## S_regionFrPoly:LocRaia       -13.87      17.85     -0.777      0.4394
## S_regionIndianOc:LocRaia       NA          NA          NA          NA
## S_regionTaiwan:LocRaia        NA          NA          NA          NA
## S_regionFrPoly:LocTahaa      -20.82      16.33     -1.275      0.2060
## S_regionIndianOc:LocTahaa     NA          NA          NA          NA
## S_regionTaiwan:LocTahaa       NA          NA          NA          NA
## S_regionFrPoly:LocTahiti      NA          NA          NA          NA
## S_regionIndianOc:LocTahiti     NA          NA          NA          NA
## S_regionTaiwan:LocTahiti      NA          NA          NA          NA
## S_regionFrPoly:LocTaiwan      NA          NA          NA          NA
## S_regionIndianOc:LocTaiwan     NA          NA          NA          NA
## S_regionTaiwan:LocTaiwan      NA          NA          NA          NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 20.86 on 83 degrees of freedom
## Multiple R-squared:  0.8435, Adjusted R-squared:  0.8303
## F-statistic: 63.9 on 7 and 83 DF,  p-value: < 2.2e-16

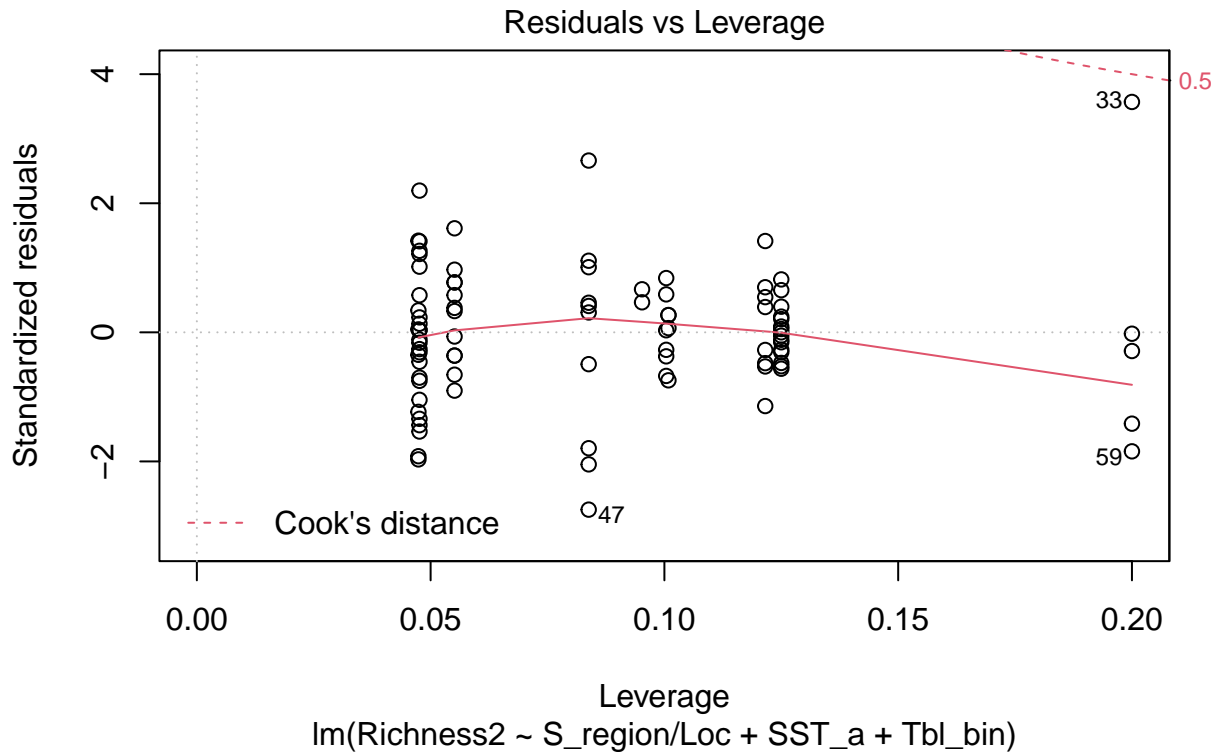
```

```
plot(rich.lm.reduced)
```









```
##New GLS with amended alpha_div dataset
```

```
###Richness
```

```
rich_gls_nest2 <- gls(Richness2 ~ S_region/Loc + SST_a + Tbl_bin, data = alpha_vegan2, na.action = "na.omit")
```

```
summary(rich_gls_nest2)
```

```
## Generalized least squares fit by REML
```

```
## Model: Richness2 ~ S_region/Loc + SST_a + Tbl_bin
```

```
## Data: alpha_vegan2
```

```
## AIC BIC logLik
```

```
## 695.5996 762.1848 -317.7998
```

```
##
```

```
## Variance function:
```

```
## Structure: Different standard deviations per stratum
```

```
## Formula: ~1 | Loc
```

```
## Parameter estimates:
```

```
## Oman Moorea Taiwan Djib Tahaa Tahiti Raia
```

```
## 1 1 1 1 1 1 1
```

```
##
```

```
## Coefficients:
```

```
## Value Std.Error t-value p-value
```

```
## (Intercept) -230.44070 1019.2544 -0.2260875 0.8218
```

```
## S_regionIndianOc -36.62356 85.2322 -0.4296916 0.6688
```

```
## S_regionTaiwan 46.63350 66.8471 0.6976145 0.4878
```

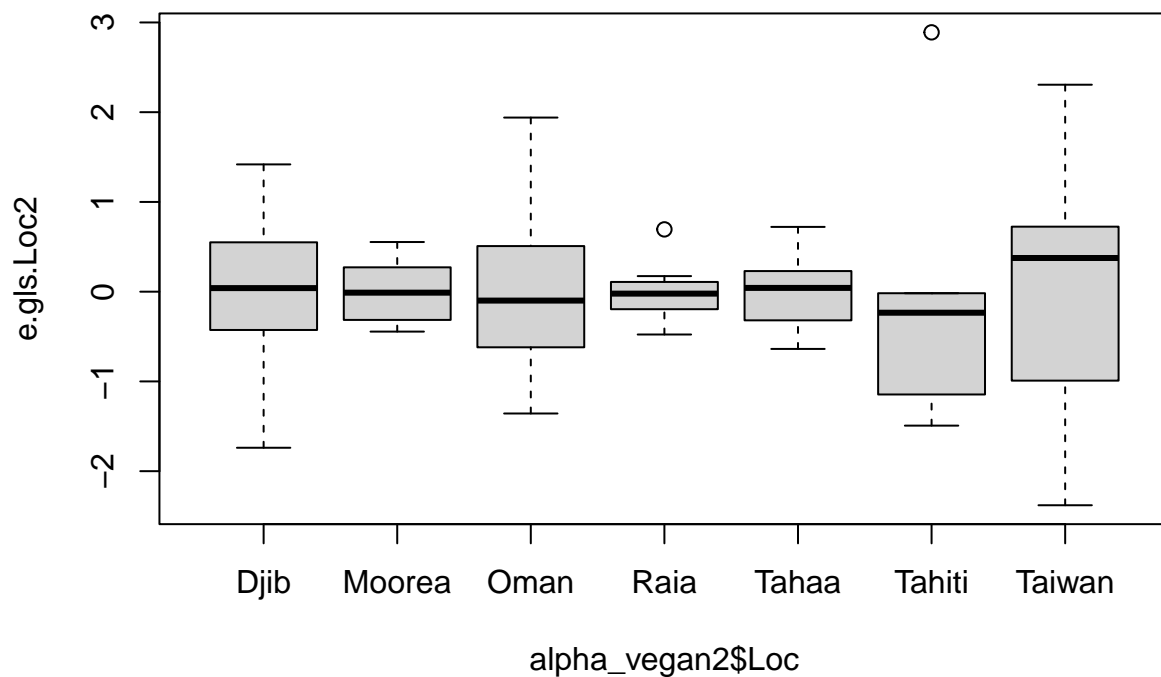
```
## SST_a 11.45306 37.6931 0.3038502 0.7622
```

```
## Tbl_binRecent      18.21155   14.7878   1.2315271   0.2224
## S_regionIndianOc:LocOman  90.59577   57.3935   1.5785027   0.1191
## S_regionFrPoly:LocRaia   -13.86669   19.7166  -0.7033001   0.4843
## S_regionFrPoly:LocTahaa  -20.81695   18.0428  -1.1537511   0.2526
##
## Correlation:
##              (Intr) S_rgIO S_rgnT SST_a  Tbl_bR S_IO:L S_FP:LR
## S_regionIndianOc    0.993
## S_regionTaiwan      0.983   0.977
## SST_a              -1.000 -0.994 -0.983
## Tbl_binRecent      -0.463 -0.403 -0.558   0.459
## S_regionIndianOc:LocOman 0.993   0.983   0.976 -0.993 -0.456
## S_regionFrPoly:LocRaia  0.746   0.741   0.814 -0.746 -0.706   0.740
## S_regionFrPoly:LocTahaa 0.714   0.710   0.790 -0.714 -0.726   0.710   0.831
##
## Standardized residuals:
##      Min      Q1      Med      Q3      Max
## -2.37968418 -0.41611957  0.02765675  0.48383874  2.88960419
##
## Residual standard error: 23.04814
## Degrees of freedom: 91 total; 68 residual
```

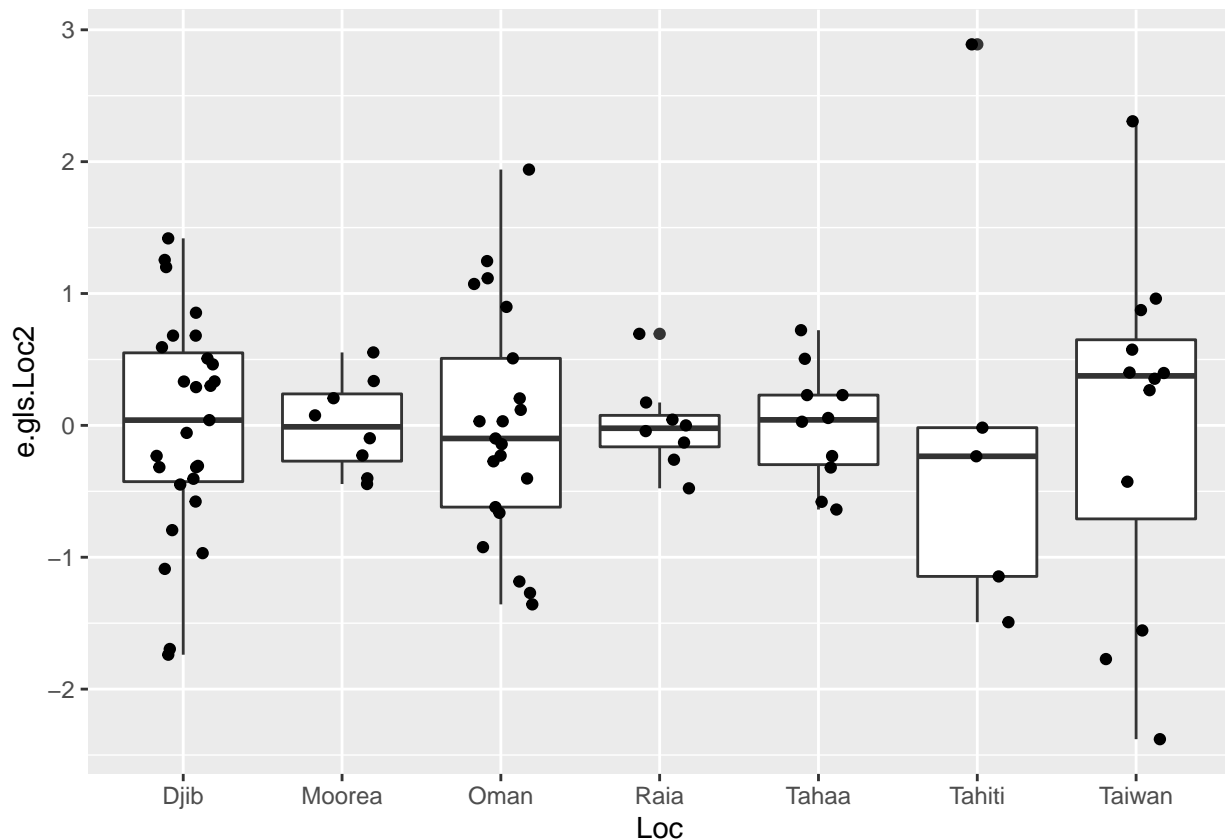
```
##Boxplots st residuals
```

```
e.gls.Loc2 <- resid(rich_gls_nest2, type = "normalized")
```

```
boxplot(e.gls.Loc2~alpha_vegan2$Loc)
```



```
##Jitter boxplot with st residuals
ggplot(meta2, aes(x = Loc, y = e.gls.Loc2)) + geom_boxplot() +
geom_jitter(height = 0, width = .2)
```



```
##qqplots with st residuals
qqnorm(e.gls.Loc2,col="firebrick",abline = c(0,1))
```

```
## Warning in plot.window(...): "abline" is not a graphical parameter
```

```
## Warning in plot.xy(xy, type, ...): "abline" is not a graphical parameter
```

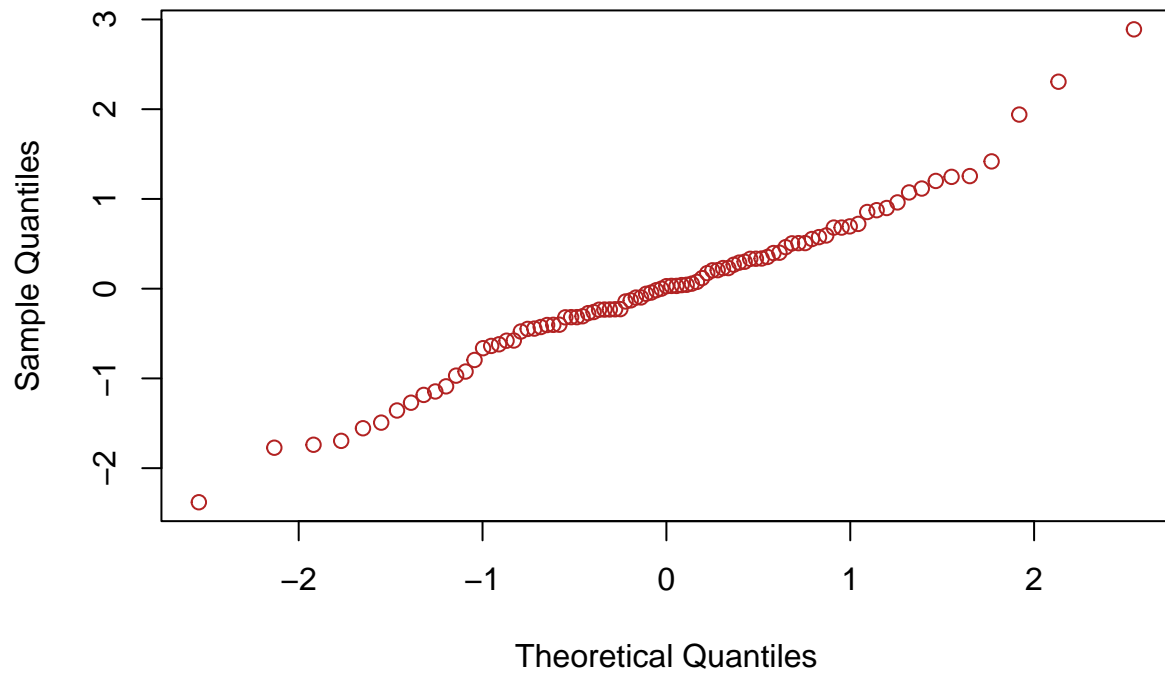
```
## Warning in axis(side = side, at = at, labels = labels, ...): "abline" is not a
## graphical parameter
```

```
## Warning in axis(side = side, at = at, labels = labels, ...): "abline" is not a
## graphical parameter
```

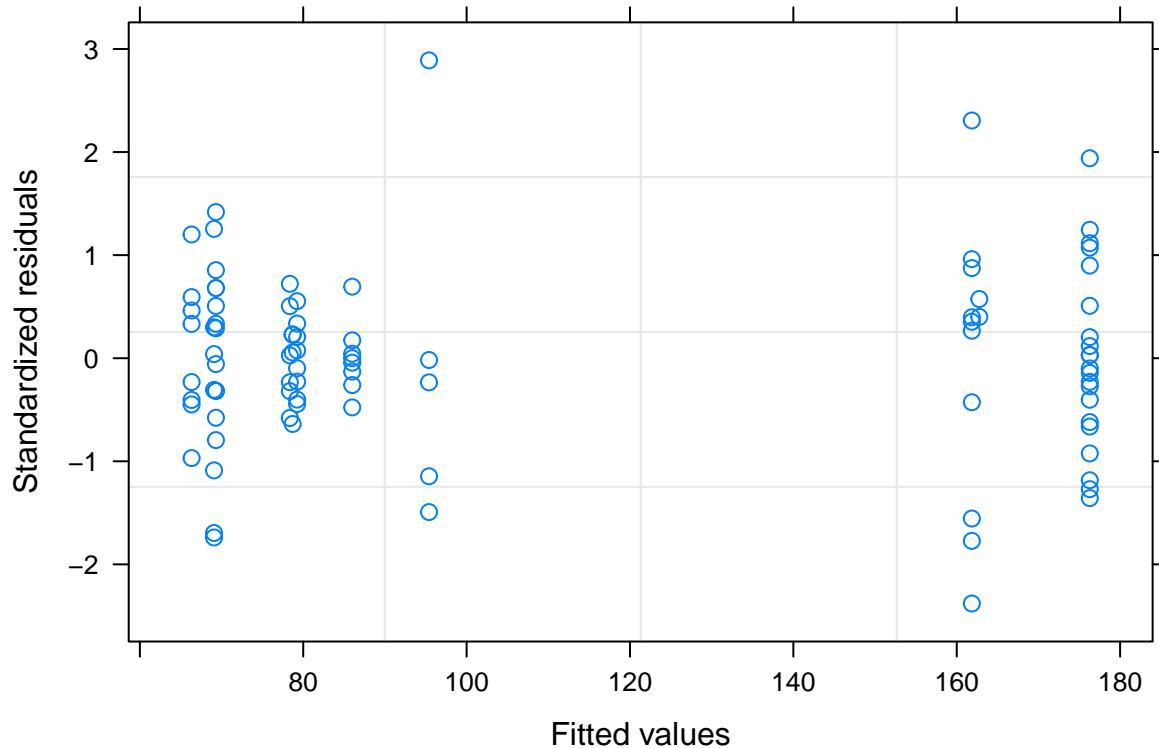
```
## Warning in box(...): "abline" is not a graphical parameter
```

```
## Warning in title(...): "abline" is not a graphical parameter
```

Normal Q-Q Plot



```
plot(rich_gls_nest2)
```

###Shannon

```
shannon_gls_nest <- gls(Shannon2 ~ S_region/Loc + SST_a + Tbl_bin, data = alpha_vegan2, na.action = "na")
summary(shannon_gls_nest)
```

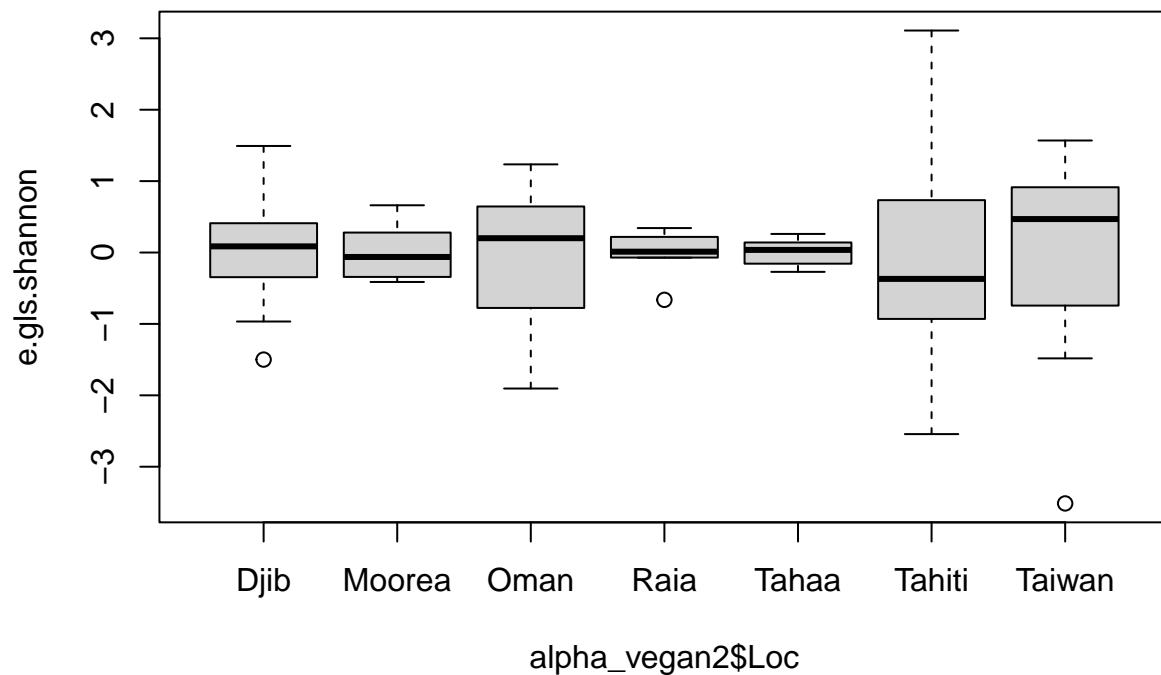
```
## Generalized least squares fit by REML
## Model: Shannon2 ~ S_region/Loc + SST_a + Tbl_bin
## Data: alpha_vegan2
##      AIC      BIC    logLik
## 79.68605 146.2713 -9.843026
##
## Variance function:
## Structure: Different standard deviations per stratum
## Formula: ~1 | Loc
## Parameter estimates:
##   Oman Moorea Taiwan  Djib  Tahaa Tahiti  Raia
##     1      1      1     1     1      1     1
##
## Coefficients:
##              Value Std.Error   t-value p-value
## (Intercept)    13.018364 11.001708   1.183304  0.2408
## S_regionIndianOc    0.145509  0.919986   0.158164  0.8748
## S_regionTaiwan     1.313063  0.721539   1.819808  0.0732
## SST_a             -0.415845  0.406855  -1.022097  0.3104
## Tbl_binRecent     -0.216803  0.159618  -1.358267  0.1789
## S_regionIndianOc:LocOman  2.021886  0.619498   3.263747  0.0017
```

```
## S_regionFrPoly:LocRaia    0.194362  0.212819  0.913277  0.3643
## S_regionFrPoly:LocTahaa   0.158328  0.194752  0.812973  0.4191
##
## Correlation:
##              (Intr) S_rgIO S_rgnT SST_a  Tbl_bR S_IO:L S_FP:LR
## S_regionIndianOc         0.993
## S_regionTaiwan           0.983  0.977
## SST_a                    -1.000 -0.994 -0.983
## Tbl_binRecent            -0.463 -0.403 -0.558  0.459
## S_regionIndianOc:LocOman  0.993  0.983  0.976 -0.993 -0.456
## S_regionFrPoly:LocRaia    0.746  0.741  0.814 -0.746 -0.706  0.740
## S_regionFrPoly:LocTahaa   0.714  0.710  0.790 -0.714 -0.726  0.710  0.831
##
## Standardized residuals:
##           Min           Q1           Med           Q3           Max
## -3.51408364 -0.34527198  0.08615284  0.41569350  3.10905272
##
## Residual standard error: 0.2487788
## Degrees of freedom: 91 total; 68 residual
```

```
##Boxplots st residuals
```

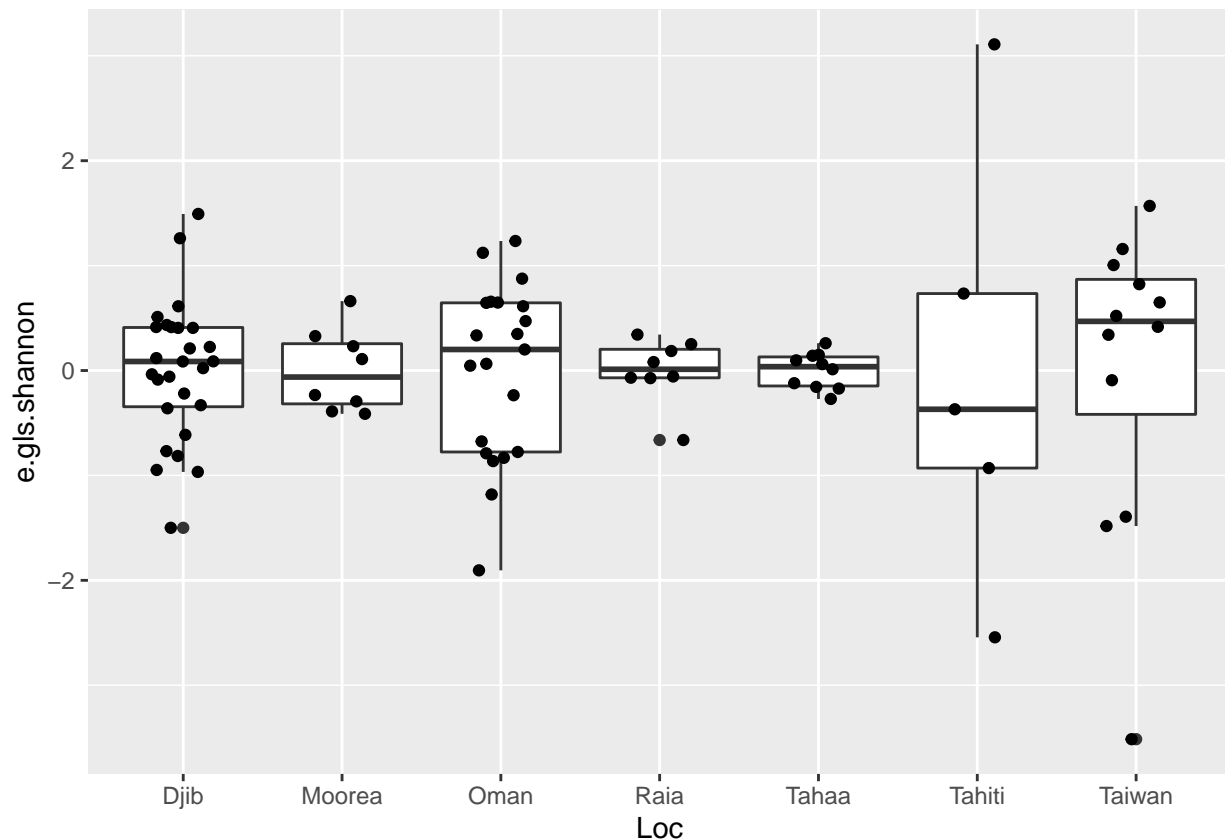
```
e.gls.shannon <- resid(shannon_gls_nest, type = "normalized")
```

```
boxplot(e.gls.shannon~alpha_vegan2$Loc)
```



```
##Jitter boxplot with st residuals
```

```
ggplot(meta2, aes(x = Loc, y = e.gls.shannon)) + geom_boxplot() +  
geom_jitter(height = 0, width = .2)
```



```
##qqplots with st residuals
```

```
qqnorm(e.gls.shannon,col="firebrick",abline = c(0,1))
```

```
## Warning in plot.window(...): "abline" is not a graphical parameter
```

```
## Warning in plot.xy(xy, type, ...): "abline" is not a graphical parameter
```

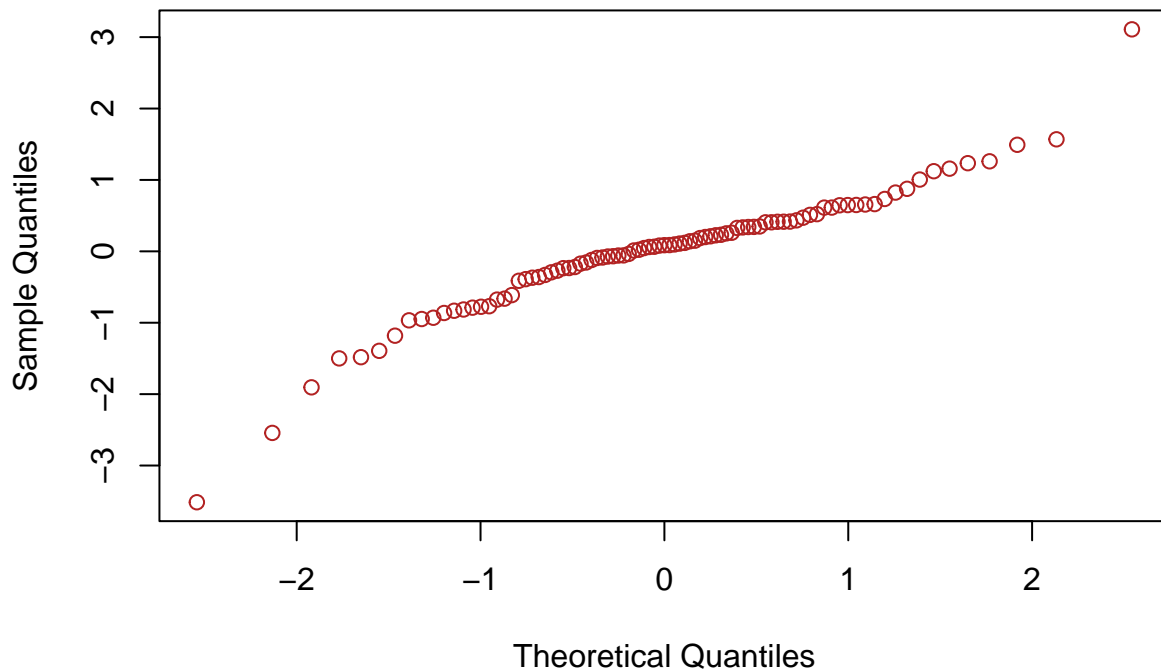
```
## Warning in axis(side = side, at = at, labels = labels, ...): "abline" is not a  
## graphical parameter
```

```
## Warning in axis(side = side, at = at, labels = labels, ...): "abline" is not a  
## graphical parameter
```

```
## Warning in box(...): "abline" is not a graphical parameter
```

```
## Warning in title(...): "abline" is not a graphical parameter
```

Normal Q-Q Plot



```
###InvSimp
invsimp_gls_nest <- gls(InvSimpson2 ~ S_region/Loc + SST_a + Tbl_bin, data = alpha_vegan2, na.action =
summary(invsimp_gls_nest)
```

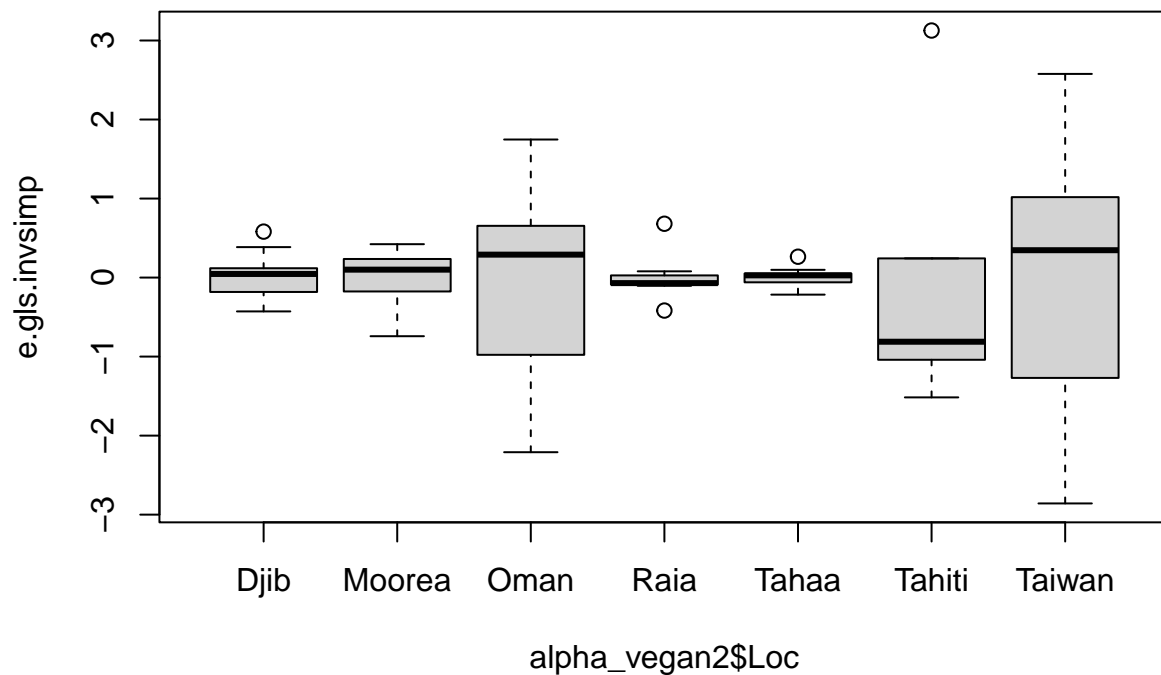
```
## Generalized least squares fit by REML
## Model: InvSimpson2 ~ S_region/Loc + SST_a + Tbl_bin
## Data: alpha_vegan2
##      AIC      BIC logLik
## 281.1201 347.7053 -110.56
##
## Variance function:
## Structure: Different standard deviations per stratum
## Formula: ~1 | Loc
## Parameter estimates:
##   Oman Moorea Taiwan  Djib  Tahaa Tahiti  Raia
##     1      1      1     1     1     1     1
##
## Coefficients:
##              Value Std.Error   t-value p-value
## (Intercept)    35.86945  48.38468   0.7413390  0.4610
## S_regionIndianOc    0.67091   4.04603   0.1658198  0.8688
## S_regionTaiwan     3.52021   3.17328   1.1093315  0.2712
## SST_a             -1.18857   1.78932  -0.6642616  0.5088
## Tbl_binRecent     -0.41047   0.70199  -0.5847217  0.5607
## S_regionIndianOc:LocOman  5.66135   2.72451   2.0779356  0.0415
```

```
## S_regionFrPoly:LocRaia   -0.20866   0.93596 -0.2229394  0.8243
## S_regionFrPoly:LocTahaa -0.25409   0.85651 -0.2966533  0.7676
##
## Correlation:
##              (Intr) S_rgIO S_rgnT SST_a  Tbl_bR S_IO:L S_FP:LR
## S_regionIndianOc       0.993
## S_regionTaiwan         0.983  0.977
## SST_a                  -1.000 -0.994 -0.983
## Tbl_binRecent          -0.463 -0.403 -0.558  0.459
## S_regionIndianOc:LocOman 0.993  0.983  0.976 -0.993 -0.456
## S_regionFrPoly:LocRaia  0.746  0.741  0.814 -0.746 -0.706  0.740
## S_regionFrPoly:LocTahaa 0.714  0.710  0.790 -0.714 -0.726  0.710  0.831
##
## Standardized residuals:
##      Min      Q1      Med      Q3      Max
## -2.85822444 -0.24458318  0.03521207  0.28015677  3.12615656
##
## Residual standard error: 1.09411
## Degrees of freedom: 91 total; 68 residual
```

```
##Boxplots st residuals
```

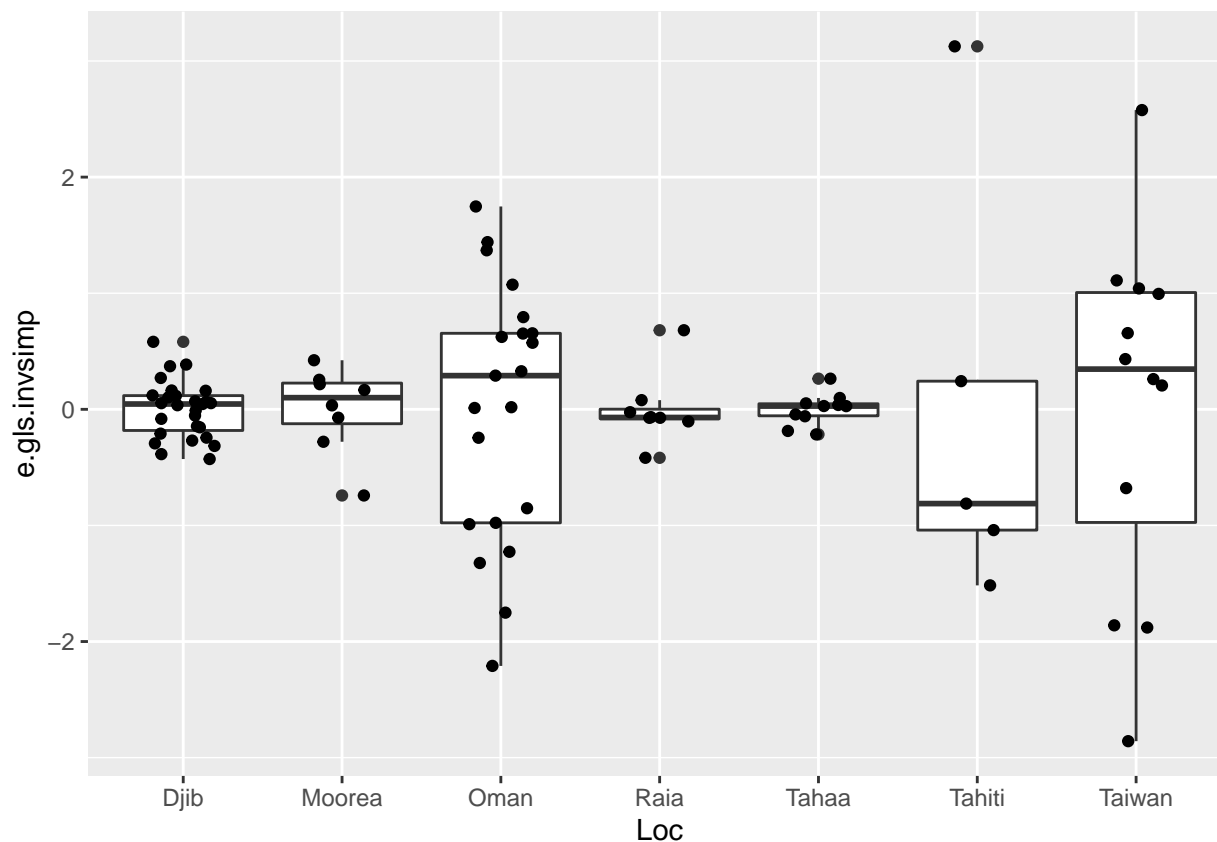
```
e.gls.invsimp <- resid(invsimp_gls_nest, type = "normalized")
```

```
boxplot(e.gls.invsimp~alpha_vegan2$Loc)
```



```
##Jitter boxplot with st residuals
```

```
ggplot(meta2, aes(x = Loc, y = e.gls.invsimp)) + geom_boxplot() +  
geom_jitter(height = 0, width = .2)
```



```
##qqplots with st residuals
```

```
qqnorm(e.gls.invsimp,col="firebrick",abline = c(0,1))
```

```
## Warning in plot.window(...): "abline" is not a graphical parameter
```

```
## Warning in plot.xy(xy, type, ...): "abline" is not a graphical parameter
```

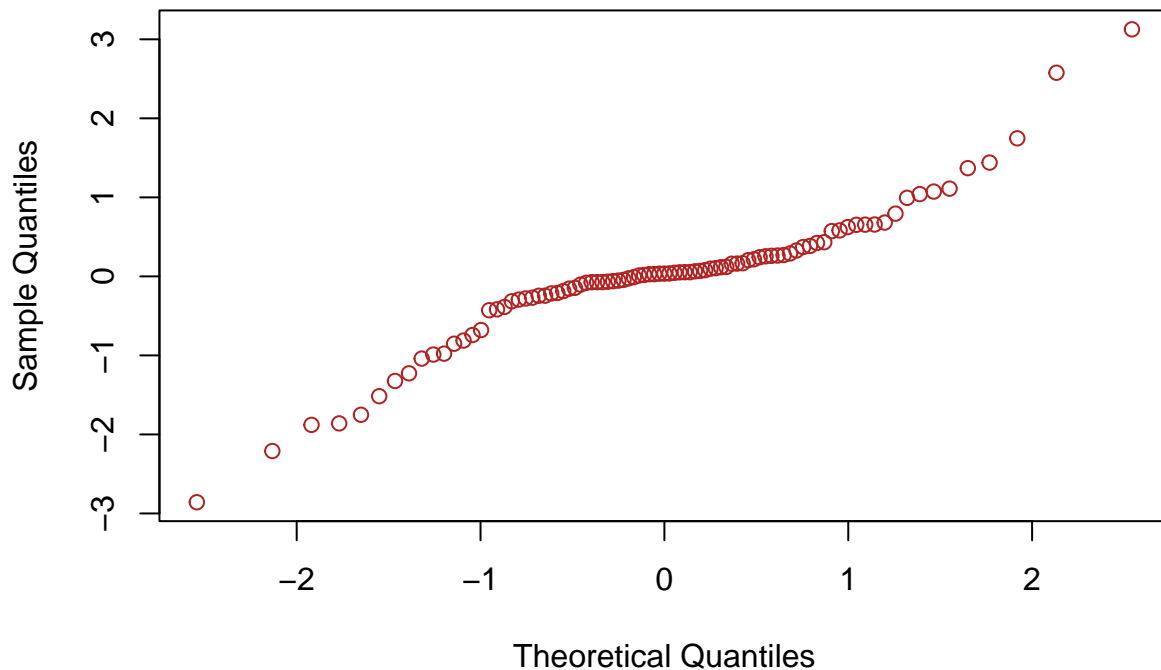
```
## Warning in axis(side = side, at = at, labels = labels, ...): "abline" is not a  
## graphical parameter
```

```
## Warning in axis(side = side, at = at, labels = labels, ...): "abline" is not a  
## graphical parameter
```

```
## Warning in box(...): "abline" is not a graphical parameter
```

```
## Warning in title(...): "abline" is not a graphical parameter
```

Normal Q-Q Plot



```
library(ggpubr)
```

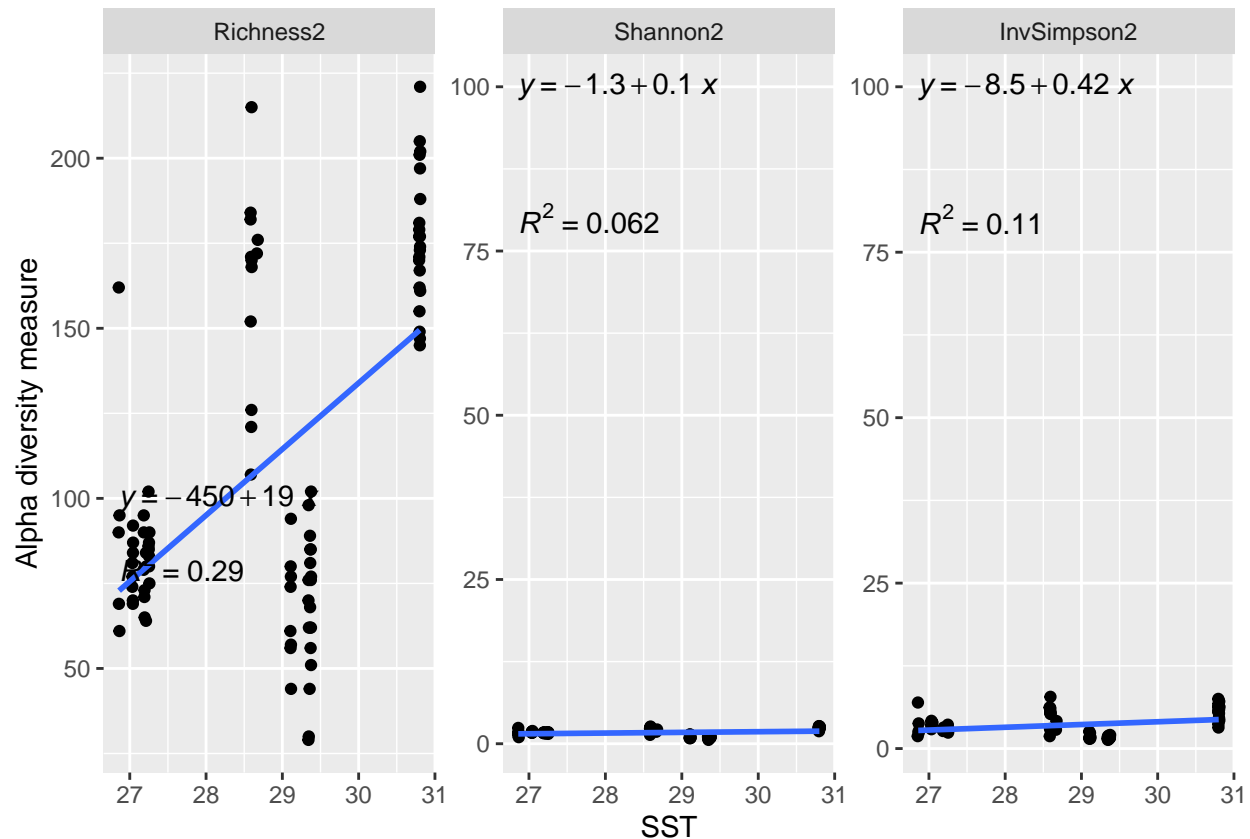
```
##
## Attaching package: 'ggpubr'
```

```
## The following object is masked from 'package:ggtree':
##
## rotate
```

```
## The following object is masked from 'package:ape':
##
## rotate
```

```
scatterplot_SST <- alpha_vegan2
scatterplot_SST %>%
  gather(key = metric, value = value, c("Richness2", "Shannon2", "InvSimpson2")) %>%
  mutate(metric = factor(metric, levels = c("Richness2", "Shannon2", "InvSimpson2"))) %>%
  ggplot(aes(x = SST_a, y = value)) +
  geom_jitter() +
  geom_smooth(method='lm' , se = FALSE) +
  stat_regline_equation(label.y = 100, aes(label = ..eq.label..)) +
  stat_regline_equation(label.y = 80, aes(label = ..rr.label..)) +
  labs(x = "SST", y = "Alpha diversity measure") +
  facet_wrap(~ metric, scales = "free") +
  theme(legend.title = element_blank())
```

```
## `geom_smooth()` using formula 'y ~ x'
```



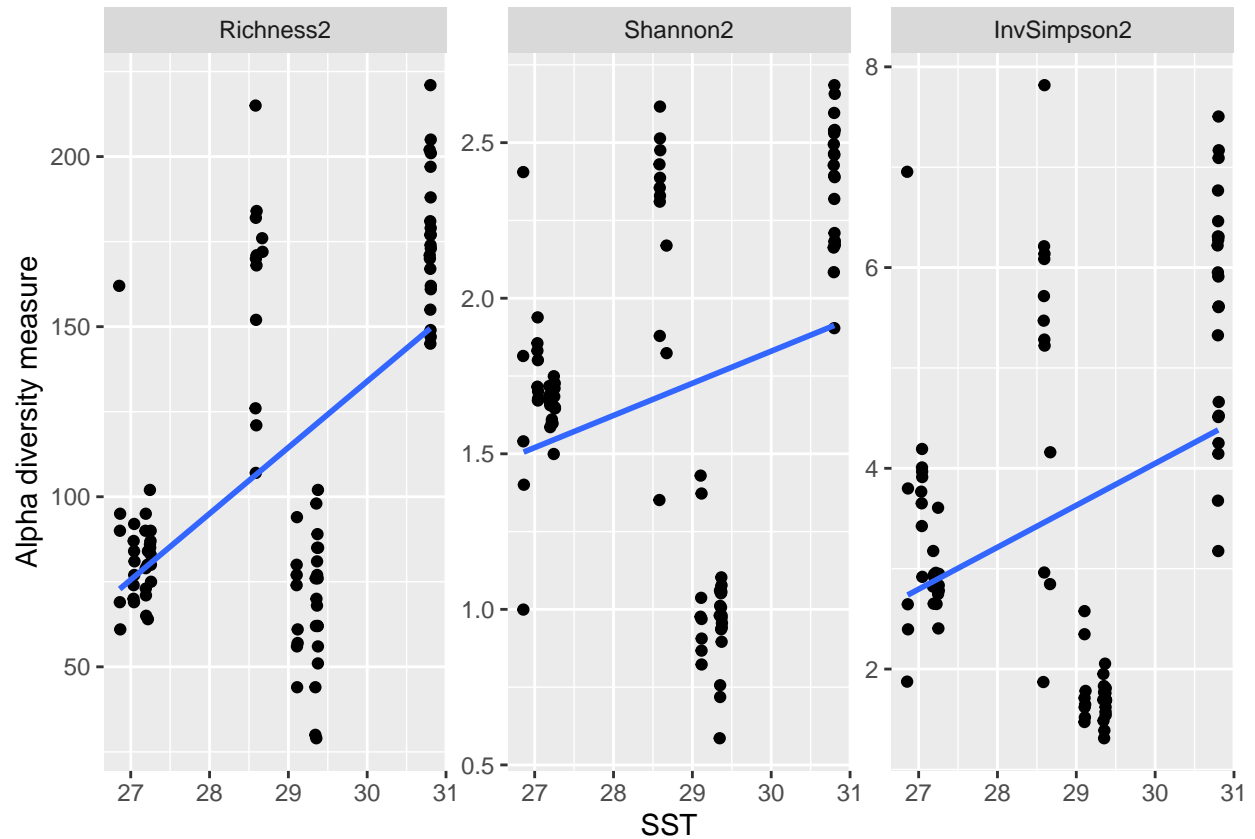
```
ggsave("scatterplot_SST_withR2overlay_28Nov2021.pdf")
```

```
## Saving 6.5 x 4.5 in image
```

```
## `geom_smooth()` using formula 'y ~ x'
```

```
scatterplot_SST2 <- alpha_vegan2
scatterplot_SST2 %>%
  gather(key = metric, value = value, c("Richness2", "Shannon2", "InvSimpson2")) %>%
  mutate(metric = factor(metric, levels = c("Richness2", "Shannon2", "InvSimpson2"))) %>%
  ggplot(aes(x = SST_a, y = value)) +
  geom_jitter() +
  geom_smooth(method='lm', se = FALSE) +
  labs(x = "SST", y = "Alpha diversity measure") +
  facet_wrap(~ metric, scales = "free") +
  theme(legend.title = element_blank())
```

```
## `geom_smooth()` using formula 'y ~ x'
```

```
ggsave("scatterplot_SST_raw_28Nov2021.pdf")
```

```
## Saving 6.5 x 4.5 in image
## `geom_smooth()` using formula 'y ~ x'
```

```
ggsave("scatterplot_SST_raw_28Nov2021.png")
```

```
## Saving 6.5 x 4.5 in image
## `geom_smooth()` using formula 'y ~ x'
```

```
## Inv simp post-hoc
## library(emmeans)

##emmeans(invsimp_gls_nest, pairwise ~ pairwise ~ Loc, weights = "outer")
```

```
##invsimp_gls_nest %>%
## emmeans(specs = pairwise ~ InvSimpson2:S_region, adjust = "BH")

##invsimp_gls_nest %>%
## emmeans(specs = pairwise ~ InvSimpson2:Tbl_bin, adjust = "BH")
```

```
## Shannon post-hoc
```

```
##library(emmeans)

##emmeans(shannon_gls_nest, pairwise ~ Loc, adjust = "BH", weights = "outer")

##shannon_gls_nest %>%
##  emmeans(pairwise ~ S_region, adjust = "BH")

##shannon_gls_nest%>%
##  emmeans(pairwise ~ Tbl_bin, adjust = "BH")
```

```
## Richness post-hoc

##library(emmeans)

##rich_gls_nest2 %>%
##  emmeans(pairwise ~ Loc, adjust = "BH")

##rich_gls_nest2 %>%
##  emmeans(pairwise ~ S_region, adjust = "BH")

##rich_gls_nest2 %>%
##  emmeans(pairwise ~ Tbl_bin, adjust = "BH")
```