Assessing source-sink dynamics in invaded habitat using metabarcoding

Script by Anna J Holmquist

2022-04-15

Read in data frames

```
# Total samples from collection
collection <- read.csv(file = "collection_data.csv")

# Dataframe including all sequences, without filtering
all_pagio <- read.csv("all_reads.csv")

# Dataframe including all prey sequences following filtering
prey_final <- read.csv("prey_reads.csv")

# Native versus non-native prey status
native_status <- read.csv("native_status.csv")

# Parasites
parasites <- read.csv("parasites.csv")
```

Summaries of results

Collection

```
# By block, transect and site
collection %>%
  group_by(site_status) %>%
  summarise(n_distinct(id)) # 82/86
```

```
## # A tibble: 2 x 2
##   site_status 'n_distinct(id)'
##   <chr>                  <int>
## 1 Ginger                    82
## 2 Native                    86
```

```
collection %>%
  group_by(site_status, transect, block) %>%
  summarise(n = n_distinct(id)) %>%
```

```
  ungroup() %>%
  summarise(mean(n), plotrix::std.error(n)) # 5.793103
```

```
## # A tibble: 1 x 2
##   `mean(n)` `plotrix::std.error(n)`
##       <dbl>                   <dbl>
## 1      5.79                   0.547
```

```
collection %>%
  group_by(transect) %>%
  summarise(n = n_distinct(id)) %>%
  ungroup() %>%
  summarise(mean(n), plotrix::std.error(n)) # 16.8
```

```
## # A tibble: 1 x 2
##   `mean(n)` `plotrix::std.error(n)`
##       <dbl>                   <dbl>
## 1      16.8                    2.38
```

Molecular

```
# Reads removed
removed_reads <- sum(all_pagio$count) - sum(prey_final$count) # 1062003

# From the spiders
all_pagio %>%
  filter(family == "Philodromidae") %>%
  summarize(sum(count) / removed_reads,
            sum(count)) # 679534,   0.6412852
```

```
##   sum(count)/removed_reads sum(count)
## 1                0.6412852     679534
```

```
# From fungi
all_pagio %>%
  filter(kingdom == "Fungi") %>%
  summarize(sum(count),
            sum(count) / removed_reads) # 234630,   0.2214234
```

```
##   sum(count) sum(count)/removed_reads
## 1     234630                0.2214234
```

```
# Hymenopterans - parisitoids
all_pagio %>%
  filter(order == "Hymenoptera") %>%
  summarise(sum(count),
            sum(count) / removed_reads) # 60060, 0.05667941
```

```
##   sum(count) sum(count)/removed_reads
## 1      60060               0.05667941
```

```r
# Summary of filtered data set

  # Reads and ASVs per spider
  prey_final %>%
    group_by(id) %>%
    summarise(n = sum(count),
              nasv = n_distinct(asv)) %>%
    ungroup() %>%
    summarise(reads_mean = mean(n), reads_stderror = plotrix::std.error(n),
              asv_mean = mean(nasv),
              asv_stderror = plotrix::std.error(nasv))
```

```
## # A tibble: 1 x 4
##   reads_mean reads_stderror asv_mean asv_stderror
##        <dbl>          <dbl>    <dbl>        <dbl>
## 1      2288.           288.     2.55        0.126
```

```r
  # ASVs and specimens by marker
  prey_final %>%
    group_by(marker) %>%
    summarise(asv = n_distinct(asv),
              sample = n_distinct(id))
```

```
## # A tibble: 3 x 3
##   marker   asv sample
##   <chr> <int>  <int>
## 1 16s     110    112
## 2 18s      10     17
## 3 28s      44     76
```

```r
  rm(removed_reads)
```

Taxonomy

```r
# Total taxonomy breakdown
prey_final %>%
  summarise(n_distinct(order), n_distinct(family),
            n_distinct(genus), n_distinct(species))
```

```
##   n_distinct(order) n_distinct(family) n_distinct(genus) n_distinct(species)
## 1                 9                 31                32                  23
```

```r
# Percent of ASVs with matches

  # Number of ASVs retained with matches
  length(unique(prey_final$asv[!is.na(prey_final$species)])) /
    length(unique(prey_final$asv)) # 0.2317073
```

```
## [1] 0.2317073
```

```
    length(unique(prey_final$asv[!is.na(prey_final$genus)])) /
        length(unique(prey_final$asv)) # 0.3780488
```

```
## [1] 0.3780488
```

```
    length(unique(prey_final$asv[!is.na(prey_final$family)])) /
        length(unique(prey_final$asv)) # 0.6158537
```

```
## [1] 0.6158537
```

```
# Percent matches by marker
  prey_final %>%
    group_by(marker) %>%
    summarise(mean(percent_match))
```

```
## # A tibble: 3 x 2
##   marker 'mean(percent_match)'
##   <chr>                  <dbl>
## 1 16s                     91.6
## 2 18s                     98.3
## 3 28s                     97.7
```

Prey diversity and abundance

Summarizing composition

```
# All ASVs
length(unique(prey_final$asv)) # 164
```

```
## [1] 164
```

```
# Shared ASVs
length(intersect(prey_final$asv[prey_final$site_status == "ginger"],
        prey_final$asv[prey_final$site_status == "native"])) # 29
```

```
## [1] 29
```

```
# Shared taxonomy
length(intersect(prey_final$order[prey_final$site_status == "ginger"],
        prey_final$order[prey_final$site_status == "native"])) # 6 orders shared
```

```
## [1] 6
```

```
length(unique(prey_final$order)) # 9
```

```
## [1] 9
```

```r
length(intersect(prey_final$family[prey_final$site_status == "ginger"],
        prey_final$family[prey_final$site_status == "native"])) # 11 families shared
```

```
## [1] 11
```

```r
length(unique(prey_final$family)) # 31
```

```
## [1] 31
```

```r
length(intersect(prey_final$genus[prey_final$site_status == "ginger"],
        prey_final$genus[prey_final$site_status == "native"])) # 10 genera shared
```

```
## [1] 10
```

```r
length(unique(prey_final$genus)) # 32
```

```
## [1] 32
```

```r
length(intersect(prey_final$species[prey_final$site_status == "ginger"],
        prey_final$species[prey_final$site_status == "native"])) # 5 species shared
```

```
## [1] 5
```

```r
length(unique(prey_final$species)) # 23
```

```
## [1] 23
```

Hill numbers

Create community matrices

```r
# Create community matrices

  # By ASVs
    asv_comm <-
      prey_final %>%
      acast(id ~ asv,
            value.var = "count", # Count as values
            fun.aggregate = sum) %>% # Sum counts for site
      as.matrix()

  # Transform
    asv_hellinger <- decostand(asv_comm, method = "hellinger")

  # Incidence
  asv_incidence <- asv_hellinger
  asv_incidence[asv_incidence > 0] <- 1
```

```r
# By taxonomy - family
family_comm <-
  prey_final %>%
  filter(!is.na(family)) %>%
  acast(id ~ family,
        value.var = "count",
        fun.aggregate = sum,
        fill = 0) %>%
  as.matrix()

family_hellinger <- decostand(family_comm, method = "hellinger")

family_incidence <- family_hellinger
family_incidence[family_incidence > 0] <- 1

# By taxonomy - order
order_comm <-
  prey_final %>%
  acast(id ~ order,
        value.var = "count",
        fun.aggregate = sum,
        fill = 0) %>%
  as.matrix()

order_hellinger <- decostand(order_comm, method = "hellinger")

order_incidence <- order_hellinger
order_incidence[order_incidence > 0] <- 1
```

Calculate Hill numbers and perform Welch t-test

```r
# ASVs

asv_hill <-
renyi(asv_hellinger,
      scales = c(0, 1, 2),
      hill = T) %>%
    cbind(collection[collection$id %in% rownames(asv_hellinger),])


  # Welch t-test - ASV
  asv_q0 <- t.test(asv_hill$`0`[asv_hill$site_status == "Ginger"],
                   asv_hill$`0`[asv_hill$site_status == "Native"])

  asv_q1 <- t.test(asv_hill$`1`[asv_hill$site_status == "Ginger"],
                   asv_hill$`1`[asv_hill$site_status == "Native"])

  asv_q2 <- t.test(asv_hill$`2`[asv_hill$site_status == "Ginger"],
                   asv_hill$`2`[asv_hill$site_status == "Native"])

  hill_asv_sum <-
```

```
        data.frame(cbind(
        rbind(asv_q0$estimate,
          asv_q1$estimate,
          asv_q2$estimate),
        rbind(asv_q0$p.value,
          asv_q1$p.value,
          asv_q2$p.value)))

    colnames(hill_asv_sum) <- c("ginger", "native", "p")
    hill_asv_sum$level <- "ASV"
    hill_asv_sum$q <- c("q = 0", "q = 1", "q = 2")

# Family

    family_hill <- renyi(family_hellinger, scales = c(0, 1, 2), hill = T) %>%
        cbind(collection[collection$id %in% rownames(family_hellinger),])

    # Welch t-test - family
    fam_q0 <- t.test(family_hill$`0`[family_hill$site_status == "Ginger"],
                    family_hill$`0`[family_hill$site_status == "Native"])

    fam_q1 <- t.test(family_hill$`1`[family_hill$site_status == "Ginger"],
                    family_hill$`1`[family_hill$site_status == "Native"])

    fam_q2 <- t.test(family_hill$`2`[family_hill$site_status == "Ginger"],
                    family_hill$`2`[family_hill$site_status == "Native"])


    hill_fam_sum <- data.frame(cbind(rbind(fam_q0$estimate,
        fam_q1$estimate,
        fam_q2$estimate),
    rbind(fam_q0$p.value,
        fam_q1$p.value,
        fam_q2$p.value)))

    colnames(hill_fam_sum) <- c("ginger", "native", "p")
    hill_fam_sum$level <- "Family"
    hill_fam_sum$q <- c("q = 0", "q = 1", "q = 2")

# Order

    order_hill <- renyi(order_hellinger, scales = c(0, 1, 2), hill = T) %>%
        cbind(collection[collection$id %in% rownames(order_hellinger),])

    # Welch t-test - order
    order_q0 <- t.test(order_hill$`0`[order_hill$site_status == "Ginger"],
                      order_hill$`0`[order_hill$site_status == "Native"])

    order_q1 <- t.test(order_hill$`1`[order_hill$site_status == "Ginger"],
                      order_hill$`1`[order_hill$site_status == "Native"])

    order_q2 <- t.test(order_hill$`2`[order_hill$site_status == "Ginger"],
                      order_hill$`2`[order_hill$site_status == "Native"])
```

```
    hill_order_sum <- data.frame(cbind(rbind(order_q0$estimate,
        order_q1$estimate,
        order_q2$estimate),
    rbind(order_q0$p.value,
        order_q1$p.value,
        order_q2$p.value)))

    colnames(hill_order_sum) <- c("ginger", "native", "p")
    hill_order_sum$q <- c("q = 0", "q = 1", "q = 2")
    hill_order_sum$level <- "Order"

# Combined table
  hill_summary <- rbind(hill_asv_sum, hill_fam_sum, hill_order_sum)
  print(hill_summary)
```

```
##     ginger   native           p  level     q
## 1 2.884615 2.103448 1.447550e-03    ASV q = 0
## 2 2.522093 1.865904 1.122441e-03    ASV q = 1
## 3 2.308503 1.733942 1.041656e-03    ASV q = 2
## 4 1.805970 1.255814 1.570216e-05 Family q = 0
## 5 1.662843 1.188902 1.143474e-05 Family q = 1
## 6 1.586544 1.152156 9.418654e-06 Family q = 2
## 7 1.846154 1.413793 3.962147e-04  Order q = 0
## 8 1.666830 1.325602 9.489919e-04  Order q = 1
## 9 1.578481 1.279930 1.463135e-03  Order q = 2
```

Differences in read abundances

```
reads <-
  prey_final %>%
  group_by(site_status, id) %>%
  summarise(n = sum(count)) %>%
  ungroup()

reads %>%
  group_by(site_status) %>%
  summarise(mean(n), plotrix::std.error(n))
```

```
## # A tibble: 2 x 3
##   site_status `mean(n)` `plotrix::std.error(n)`
##   <chr>           <dbl>                   <dbl>
## 1 ginger          3469.                    449.
## 2 native           701.                    133.
```

```
t.test(reads$n[reads$site_status == "native"],
       reads$n[reads$site_status == "ginger"])
```

```
##
##  Welch Two Sample t-test
```

```
##
## data:  reads$n[reads$site_status == "native"] and reads$n[reads$site_status == "ginger"]
## t = -5.9106, df = 90.205, p-value = 5.98e-08
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3698.444 -1837.695
## sample estimates:
## mean of x mean of y
##   700.8793 3468.9487
```

Dietary composition

Constructing community matrices, by site

```r
# Data for sites
sites_data <-
  prey_final %>%
  group_by(site) %>%
  filter(row_number() == 1) %>%
  ungroup() %>%
  select(site, site_status)

# By ASV - all
site_community_asv <-
  prey_final %>%
  acast(site ~ asv, fill = 0,
        value.var = "count", # Count as values
        fun.aggregate = sum)

site_community_asv <- decostand(site_community_asv, method = "hellinger")

# By order
site_community_order <-
  prey_final %>%
  acast(site ~ order, fill = 0,
        value.var = "count", # Count as values
        fun.aggregate = sum)

site_community_order <- decostand(site_community_order, method = "hellinger")
```

Calculate beta diversity using community matrices

```r
beta_transform_asv <- beta(site_community_asv, func = "jaccard")
beta_inc_asv <- beta(site_community_asv,
                     func = "jaccard", abund = F)

beta_transform_order <- beta(site_community_order, func = "jaccard")
beta_inc_order <- beta(site_community_order,
                       func = "jaccard", abund = F)
```

Perform NMDS

```
set.seed(200)

# ASV NMDS
site_asv_nmds1 <- metaMDS(beta_transform_asv$Btotal, trymax = 1000)
site_asv_nmds2 <- metaMDS(beta_inc_asv$Btotal, trymax = 1000)

# Order NMDS
site_order_nmds1 <- metaMDS(beta_transform_order$Btotal, trymax = 1000)
site_order_nmds2 <- metaMDS(beta_inc_order$Btotal, trymax = 1000)

# PERMANOVA
adonis2(beta_transform_asv$Btotal ~ site_status,
        data = sites_data, method = "jaccard")
```

```
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = beta_transform_asv$Btotal ~ site_status, data = sites_data, method = "jaccard")
##             Df SumOfSqs      R2      F Pr(>F)
## site_status  1   0.5205 0.04077 1.1477  0.152
## Residual    27  12.2456 0.95923
## Total       28  12.7662 1.00000
```

```
adonis2(beta_inc_asv$Btotal ~ site_status,
        data = sites_data, method = "jaccard")
```

```
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = beta_inc_asv$Btotal ~ site_status, data = sites_data, method = "jaccard")
##             Df SumOfSqs      R2      F Pr(>F)
## site_status  1    0.567 0.04574 1.2943  0.042 *
## Residual    27   11.829 0.95426
## Total       28   12.396 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# PERMANOVA
adonis2(beta_inc_order$Btotal ~ site_status,
        data = sites_data, method = "jaccard")
```

```
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
```

```
## 
## adonis2(formula = beta_inc_order$Btotal ~ site_status, data = sites_data, method = "jaccard")
##           Df SumOfSqs     R2      F Pr(>F)
## site_status  1   0.3327 0.1036 3.1205  0.036 *
## Residual    27   2.8783 0.8964
## Total       28   3.2109 1.0000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
adonis2(beta_transform_order$Btotal ~ site_status,
        data = sites_data, method = "jaccard")
```

```
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
## 
## adonis2(formula = beta_transform_order$Btotal ~ site_status, data = sites_data, method = "jaccard")
##           Df SumOfSqs      R2      F Pr(>F)
## site_status  1   0.2601 0.04704 1.3328  0.251
## Residual    27   5.2691 0.95296
## Total       28   5.5292 1.00000
```

Order-level diversity

```r
prey_final %>%
  group_by(site_status, order) %>%
  # Number of spiders in each site eating each order
  summarise(n = n_distinct(id)) %>%
  ungroup() %>%
  # Add column for total number of spiders in each habitat
  mutate(total = ifelse(site_status == "Ginger",
                        78, 58),
         # Calculate proportion of spiders eating each order
         prop = round(n / total, 3))
```

```
## # A tibble: 15 x 5
##    site_status order              n total  prop
##    <chr>       <chr>          <int> <dbl> <dbl>
##  1 ginger      Araneae            1    58 0.017
##  2 ginger      Coleoptera        15    58 0.259
##  3 ginger      Diptera           29    58 0.5
##  4 ginger      Entomobryomorpha  27    58 0.466
##  5 ginger      Hemiptera         34    58 0.586
##  6 ginger      Lepidoptera       33    58 0.569
##  7 ginger      Neuroptera         1    58 0.017
##  8 ginger      Thysanoptera       4    58 0.069
##  9 native      Coleoptera         2    58 0.034
## 10 native      Diptera           13    58 0.224
## 11 native      Entomobryomorpha   1    58 0.017
## 12 native      Hemiptera         42    58 0.724
```

```
## 13 native     Lepidoptera     21    58 0.362
## 14 native     Psocoptera       1    58 0.017
## 15 native     Thysanoptera     2    58 0.034
```

NMDS for Hemiptera and Lepidoptera

```r
# NMDS is doing the same thing because ASVs are distinct
set.seed(200)

# Hemiptera
hemi <-
  prey_final %>%
  filter(order == "Hemiptera" & marker == "16s") %>%
  acast(site ~ asv, fill = 0, value.var = "count",
        fun.aggregate = sum) %>%
  as.matrix()

hemi_phylo_16s <-
  BAT::beta(decostand(hemi, method = "hellinger"),
            tree = read.tree("16s_hemiptera_tree.newick"),
            abund = F)

hemi_nmds <- metaMDS(hemi_phylo_16s$Btotal,
                     trymax = 1000, k = 2)

# Lepidoptera
lep <-
  prey_final %>%
  filter(order == "Lepidoptera" & marker == "16s") %>%
  #filter(!site %in% problem) %>%
  acast(site ~ asv, fill = 0, value.var = "count",
        fun.aggregate = sum) %>%
  as.matrix()

lep_phylo_16s <-
  BAT::beta(decostand(lep, method = "hellinger"),
            tree = read.tree("16s_lepidoptera_tree.newick"),
            abund = F)


lep_nmds <- metaMDS(lep_phylo_16s$Btotal,
                    trymax = 1000, k = 2)
```

Native versus non-native prey

```r
# Count identified
length(unique(native_status$asv[native_status$status != "unknown"])) # 53
```

```
## [1] 53
```

```r
length(unique(prey_final$asv)) # 164
```

```
## [1] 164
```

```r
# Determining prey composition for each individual
composition <-
  prey_final %>%
  mutate(status = case_when(
    asv %in% native_status$asv[native_status$status == "Native"] ~ "Native",
    asv %in% native_status$asv[native_status$status == "Non-native"] ~ "Non-native",
    TRUE ~ "unknown"
    )
  ) %>%
  group_by(id) %>%
  mutate(prey_comp = case_when(
    any(status == "Non-native") & any(status == "Native") ~ "Both",
    any(status == "Non-native") & !any(status == "Native") ~ "Non-native",
    !any(status == "Non-native") & any(status == "Native") ~ "Native",
    !any(status == "Non-native") & !any(status == "Native") ~ "Unknown"
  )) %>%
  filter(row_number() == 1) %>%
  ungroup()
```

Parasites

```r
length(unique(parasites$id))
```

```
## [1] 61
```

```r
# First, identify number of parasite ASVs and relative reads

parasite_summary <-
    parasites %>%
    acast(id ~ asv, fill = 0, value.var = "count") %>%
    as.matrix() %>%
    decostand(method = "hellinger") %>%
    melt() %>%
    filter(value != 0) %>%
    rename(id = Var1,
           asv = Var2,
           rel_read = value) %>%
    left_join(parasites, by = c("id", "asv")) %>%
    group_by(id, order, site, site_status) %>%
    summarise(n_asv = n_distinct(asv),
              n_read = sum(rel_read))

# Add in spiders that had no parasitism
parasite_summary <-
  prey_final %>%
```

```
  filter(!id %in% parasite_summary$id) %>%
  group_by(id) %>%
  filter(row_number() == 1) %>%
  mutate(n_asv = 0,
         n_read = 0,
         order = NA) %>%
  select(id, site, site_status, order, n_asv, n_read) %>%
  rbind(parasite_summary)  %>%
  ungroup()

t.test(parasite_summary$n_asv[parasite_summary$site_status == "native"],
       parasite_summary$n_asv[parasite_summary$site_status == "ginger"]) # 0.0002957
```

```
##
##  Welch Two Sample t-test
##
## data:  parasite_summary$n_asv[parasite_summary$site_status == "native"] and parasite_summary$n_asv[pa
## t = -3.7048, df = 152.01, p-value = 0.0002957
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.6621241 -0.2015435
## sample estimates:
## mean of x mean of y
## 0.3593750 0.7912088
```

```
t.test(parasite_summary$n_read[parasite_summary$site_status == "native"],
       parasite_summary$n_read[parasite_summary$site_status == "ginger"]) # 0.00223
```

```
##
##  Welch Two Sample t-test
##
## data:  parasite_summary$n_read[parasite_summary$site_status == "native"] and parasite_summary$n_read
## t = -3.1139, df = 143.25, p-value = 0.00223
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.423945 -0.094709
## sample estimates:
## mean of x mean of y
##  0.307696  0.567023
```

```
# Wasps
wasps <-
  parasites %>%
  filter(order == "Hymenoptera")

wasps %>%
  summarise(n_distinct(id),
            n_distinct(asv))
```

```
##   n_distinct(id) n_distinct(asv)
## 1             49              29
```

```
wasps %>%
  group_by(site_status) %>%
  summarise(n_distinct(id),
            n_distinct(asv))
```

```
## # A tibble: 2 x 3
##   site_status `n_distinct(id)` `n_distinct(asv)`
##   <chr>                 <int>             <int>
## 1 ginger                   31                20
## 2 native                   18                13
```

```
wasps %>%
  filter(percent_match >= 92) %>%
  group_by(family, site_status) %>%
  summarise(n_distinct(id))
```

```
## # A tibble: 14 x 3
## # Groups:   family [11]
##     family        site_status `n_distinct(id)`
##     <chr>         <chr>                  <int>
##  1 Aphelinidae   ginger                     1
##  2 Aphelinidae   native                     3
##  3 Bethylidae    ginger                     1
##  4 Braconidae    ginger                    11
##  5 Braconidae    native                     6
##  6 Diapriidae    native                     1
##  7 Eucharitidae  ginger                     3
##  8 Eulophidae    ginger                     1
##  9 Halictidae    ginger                     1
## 10 Ichneumonidae ginger                     3
## 11 Ichneumonidae native                     6
## 12 Pteromalidae  native                     1
## 13 Signiphoridae native                     1
## 14 Sphecidae     native                     2
```

```
wasps <-
  parasite_summary %>%
  filter(order == "Hymenoptera" | is.na(order))
```

```
t.test(wasps$n_asv[wasps$site_status == "native"],
       wasps$n_asv[wasps$site_status == "ginger"]) # 0.01165
```

```
##
##  Welch Two Sample t-test
##
## data:  wasps$n_asv[wasps$site_status == "native"] and wasps$n_asv[wasps$site_status == "ginger"]
## t = -2.5617, df = 120.96, p-value = 0.01165
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.61419547 -0.07869618
## sample estimates:
## mean of x mean of y
## 0.3492063 0.6956522
```

```
t.test(wasps$n_read[wasps$site_status == "native"],
       wasps$n_read[wasps$site_status == "ginger"]) # 0.03998
```

```
##
##  Welch Two Sample t-test
##
## data:  wasps$n_read[wasps$site_status == "native"] and wasps$n_read[wasps$site_status == "ginger"]
## t = -2.0749, df = 129.37, p-value = 0.03998
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.378179954 -0.008994213
## sample estimates:
## mean of x mean of y
## 0.3034158 0.4970029
```

```
# Fungi
fungi <-
  parasites %>%
  filter(order != "Hymenoptera")

fungi %>%
  summarise(n_distinct(id),
            n_distinct(asv))
```

```
##   n_distinct(id) n_distinct(asv)
## 1             23               6
```

```
fungi %>%
  group_by(site_status) %>%
  summarise(n_distinct(id),
            n_distinct(asv))
```

```
## # A tibble: 2 x 3
##   site_status `n_distinct(id)` `n_distinct(asv)`
##   <chr>                  <int>             <int>
## 1 ginger                    22                 5
## 2 native                     1                 1
```

```
fungi <-
  parasite_summary %>%
  filter(order != "Hymenoptera" | is.na(order))

t.test(fungi$n_asv[fungi$site_status == "native"],
       fungi$n_asv[fungi$site_status == "ginger"]) # 3.777e-06
```

```
##
##  Welch Two Sample t-test
##
## data:  fungi$n_asv[fungi$site_status == "native"] and fungi$n_asv[fungi$site_status == "ginger"]
## t = -5.0233, df = 69.464, p-value = 3.777e-06
## alternative hypothesis: true difference in means is not equal to 0
```

```
## 95 percent confidence interval:
##  -0.5284656 -0.2280562
## sample estimates:
##  mean of x  mean of y
## 0.02173913 0.40000000
```

```
t.test(fungi$n_read[fungi$site_status == "native"],
       fungi$n_read[fungi$site_status == "ginger"]) # 6.097e-06
```

```
##
##  Welch Two Sample t-test
##
## data:  fungi$n_read[fungi$site_status == "native"] and fungi$n_read[fungi$site_status == "ginger"]
## t = -4.9237, df = 65.16, p-value = 6.097e-06
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.3877765 -0.1639845
## sample estimates:
##  mean of x  mean of y
## 0.01255109 0.28843163
```