

# Indicateurs de la Biodiversité

Andrew MacDonald

2020-07-06



# Contents

<b>1</b>	<b>Dependencies</b>	<b>5</b>
<b>2</b>	<b>Download data</b>	<b>7</b>
2.1	List of Indicators . . . . .	8
<b>3</b>	<b>Sites</b>	<b>11</b>
<b>4</b>	<b>Odonate communities</b>	<b>13</b>
4.1	Questions . . . . .	15
<b>5</b>	<b>Soil invertebrates</b>	<b>17</b>
5.1	Questions . . . . .	18
<b>6</b>	<b>Lepidoptera</b>	<b>19</b>
6.1	Questions . . . . .	20
<b>7</b>	<b>Histoires de la biodiversité</b>	<b>23</b>
7.1	Sampling / rarefaction . . . . .	23



## Chapter 1

# Dependencies

```
library(rcoleo)  
library(tidyverse)  
library(lubridate)
```



## Chapter 2

# Download data

The only point here is to demonstrate the workflow for downloading data

```
# On retire les cellules (classe sf) depuis l'API
cells <- rcoleo::sf_cells()
obs <- rcoleo::get_obs()
sites_dl <- rcoleo::sf_sites()

# until there is a better way to obtain the data and parse the result (perhaps
# in the form of a convenience function in rcoleo) we do this:

obs_df <- obs[[1]] %>% map("body") %>% map_df(~ select(.x, -closed_at))

all_obs <- obs_df %>%
  select(cell_code, site_code, date_obs, type,
         taxa = obs_species.taxa_name,
         var = obs_species.variable,
         val = obs_species.value) %>%
  mutate(date_obs = lubridate::ymd(date_obs),
         # convert cover into pres/abs (right?)
         count = if_else(var == "recouvrement", 1, val))

# CELLULES: On compte le nombre d'observation/nombre espece par type, année et cellule
obs_cells <- all_obs %>%
  group_by(cell_code, date_obs, type) %>%
  summarise(n = sum(count)) %>%
  ungroup

sp_cells <- all_obs %>%
  select(cell_code, date_obs, type, taxa) %>%
  distinct() %>%
```

```

group_by(cell_code, date_obs, type) %>%
summarise(n = n()) %>%
ungroup

# CAMPAGNES
sites <- sites_dl %>%
  select(site_code, off_station_code_id,
         type_milieu = type, geometry, site_id = id)

# On prépare les jeux de données pour chacun des types de campagnes

all_obs_con <- all_obs %>%
  filter(taxa != "inconnu")

microfaunes <- subset(all_obs_con, type == "insectes_sol")
papillons <- subset(all_obs_con, type == "papilionidés")
odonates <- subset(all_obs_con, type == "odonates")
vegetation <- subset(all_obs_con, type == "végétation")

```

### 2.0.1 overall summary

```

all_obs_con %>%
  group_by(type) %>%
  summarize(sites = n_distinct(site_code),
            cells = n_distinct(cell_code),
            taxa = n_distinct(taxa))

## `summarise()` ungrouping output (override with `.groups` argument)

## # A tibble: 4 x 4
##   type      sites cells taxa
##   <chr>    <int> <int> <int>
## 1 insectes_sol    13    13    37
## 2 odonates       12    10    35
## 3 papilionidés   14    11    26
## 4 végétation     26    11   170

```

## 2.1 List of Indicators

- Vitesse de décomposition de la matière organique



- Composition des communautés d'Odonates
- Phénologie végétale
- Phénologie acoustique des communautés, d'anoures et d'orthoptères
- Composition communauté d'insectes du sol
- Composition des communautés de poissons
- Présence de certains méso carnivores et grands mammifères
- Composition macroinvertébrés benthiques d'eau douce
- Composition des communautés des orthoptères



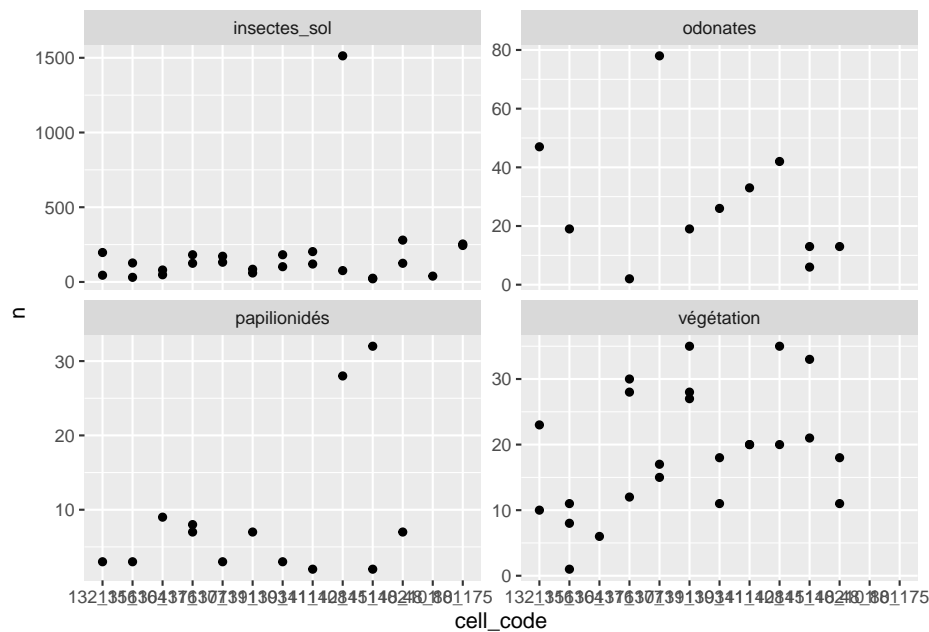
## Chapter 3

### Sites

Summary data about the different sites

```
obs_cells %>%  
  ggplot(aes(x = cell_code, y = n)) + geom_point() +  
  facet_wrap(~type, scales = "free_y")
```

## Warning: Removed 1 rows containing missing values (geom\_point).





## Chapter 4

# Odonate communities

```
# confirm that the only variable recorded here is "abondance"
odonates$var %>% unique %>% {stopifnot(. == "abondance")}

odonates %>%
  glimpse %>%
  {all.equal(.$count, .$val)}

## Rows: 151
## Columns: 8
## $ cell_code <chr> "135_104", "135_104", "135_104", "135_104", "135_104", "1...
## $ site_code <chr> "135_104_H01", "135_104_H01", "135_104_H01", "135_104_H01...
## $ date_obs <date> 2016-06-23, 2016-06-23, 2016-06-23, 2016-06-23, 2016-06-...
## $ type <chr> "odonates", "odonates", "odonates", "odonates", "odonates...
## $ taxa <chr> "Leucorrhinia hudsonica", "Nehalennia gracilis", "Gomphus...
## $ var <chr> "abondance", "abondance", "abondance", "abondance", "abon...
## $ val <dbl> 2, 1, 1, 1, 1, 1, 1, 2, 1, 1, 2, 4, 1, 10, 1, 30, 5, 20, ...
## $ count <dbl> 2, 1, 1, 1, 1, 1, 1, 2, 1, 1, 2, 4, 1, 10, 1, 30, 5, 20, ...

## [1] TRUE

## are sites nested in cells?

odonates %>%
  select(site_code, cell_code) %>% distinct %>%
  group_by(cell_code) %>% tally

## # A tibble: 10 x 2
##   cell_code      n
##   <chr>      <int>
## 1 132_116        1
```

```
## 2 135_104      1
## 3 137_107      1
## 4 137_111      1
## 5 139_103      1
## 6 139_111      1
## 7 141_108      1
## 8 142_111      2
## 9 145_102      2
## 10 148_101     1
```

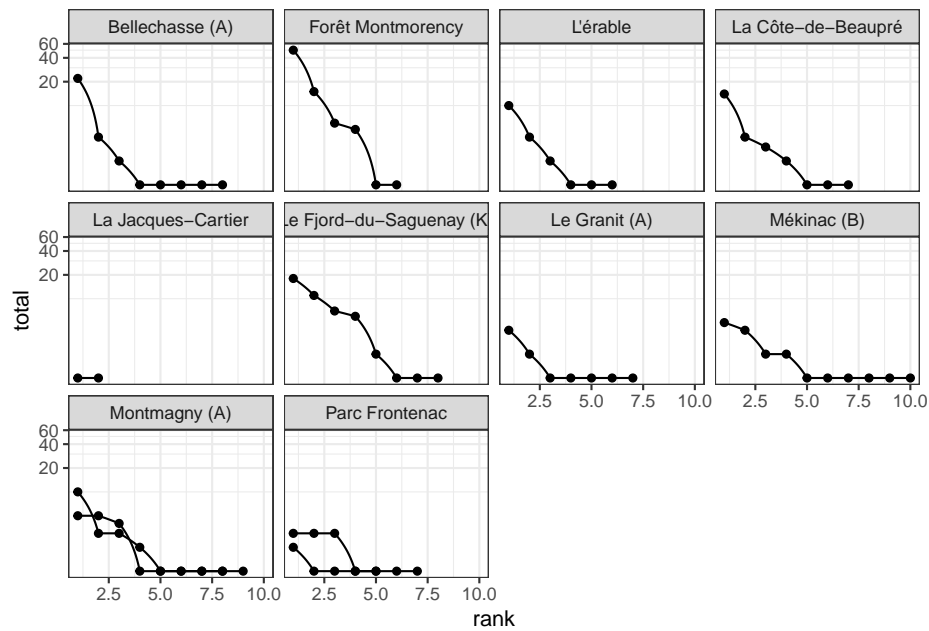
```
# yes, sometimes there are sites within cells!
```

```
rank_abundance <- odonates %>%
  left_join(select(cells, cell_code, name)) %>%
  group_by(cell_code, name, site_code, taxa) %>%
  summarize(total = sum(count)) %>%
  mutate(rank = row_number(desc(total)))
```

```
## Joining, by = "cell_code"
```

```
## `summarise()` regrouping output by 'cell_code', 'name', 'site_code' (override with `group_by()`)
```

```
rank_abundance %>%
  ggplot(aes(x = rank, y = total, group = site_code)) +
  geom_point() +
  geom_line() +
  facet_wrap(~name) +
  coord_trans(y = "log") +
  theme_bw()
```



## 4.1 Questions

Some sites appear to count **individuals**, not species? why is that?

```
odonates %>%
  left_join(select(cells, cell_code, name)) %>%
  filter(name %>% stringr::str_detect("Parc|-de-")) %>%
  select(name, site_code, taxa, val) %>% knitr::kable(.)
```

```
## Joining, by = "cell_code"
```

name	site_code	taxa	val
La Côte-de-Beaupré	139_111_H01	Leucorrhinia glacialis	1
La Côte-de-Beaupré	139_111_H01	Ladona julia	1
La Côte-de-Beaupré	139_111_H01	Ladona julia	1
La Côte-de-Beaupré	139_111_H01	Ladona julia	1
La Côte-de-Beaupré	139_111_H01	Ladona julia	1
La Côte-de-Beaupré	139_111_H01	Ladona julia	1
La Côte-de-Beaupré	139_111_H01	Ladona julia	1
La Côte-de-Beaupré	139_111_H01	Ladona julia	1
La Côte-de-Beaupré	139_111_H01	Ladona julia	1
La Côte-de-Beaupré	139_111_H01	Ladona julia	1
La Côte-de-Beaupré	139_111_H01	Ladona julia	1
La Côte-de-Beaupré	139_111_H01	Enallagma ebrium	1
La Côte-de-Beaupré	139_111_H01	Ladona julia	1
La Côte-de-Beaupré	139_111_H01	Nehalennia irene	1
La Côte-de-Beaupré	139_111_H01	Leucorrhinia proxima	1
La Côte-de-Beaupré	139_111_H01	Dorocordulia libera	1
La Côte-de-Beaupré	139_111_H01	Enallagma ebrium	1
La Côte-de-Beaupré	139_111_H01	Ladona julia	1
La Côte-de-Beaupré	139_111_H01	Ladona julia	1
La Côte-de-Beaupré	139_111_H01	Somatochlora franklini	1
La Côte-de-Beaupré	139_111_H01	Enallagma ebrium	1
La Côte-de-Beaupré	139_111_H01	Ladona julia	1
La Côte-de-Beaupré	139_111_H01	Enallagma ebrium	1
La Côte-de-Beaupré	139_111_H01	Leucorrhinia proxima	1
La Côte-de-Beaupré	139_111_H01	Leucorrhinia proxima	1
La Côte-de-Beaupré	139_111_H01	Nehalennia irene	1
Parc Frontenac	145_102_H01	Nehalennia gracilis	1
Parc Frontenac	145_102_H01	Leucorrhinia proxima	1
Parc Frontenac	145_102_H01	Nannothemis bella	1
Parc Frontenac	145_102_H01	Nannothemis bella	1
Parc Frontenac	145_102_H01	Nehalennia gracilis	1
Parc Frontenac	145_102_H01	Nannothemis bella	1
Parc Frontenac	145_102_H01	Nehalennia irene	1
Parc Frontenac	145_102_H01	Nehalennia irene	1
Parc Frontenac	145_102_H01	Nehalennia gracilis	1
Parc Frontenac	145_102_H01	Enallagma boreale	1
Parc Frontenac	145_102_H01	Leucorrhinia patricia	1
Parc Frontenac	145_102_H01	Nehalennia irene	1
Parc Frontenac	145_102_H01	Leucorrhinia glacialis	1
Parc Frontenac	145_102_H02	Nehalennia irene	1
Parc Frontenac	145_102_H02	Celithemis elisa	1
Parc Frontenac	145_102_H02	Nehalennia irene	1
Parc Frontenac	145_102_H02	Leucorrhinia proxima	1
Parc Frontenac	145_102_H02	Ischnura posita	1
Parc Frontenac	145_102_H02	Enallagma ebrium	1



## Chapter 5

# Soil invertebrates

```
# confirm that the only variable recorded here is "abondance"
microfaunes$var %>% unique %>% {stopifnot(. == "abondance")}
```

```
## are sites nested in cells?
```

```
microfaunes %>%
  select(site_code, cell_code) %>% distinct %>%
  group_by(cell_code) %>% tally
```

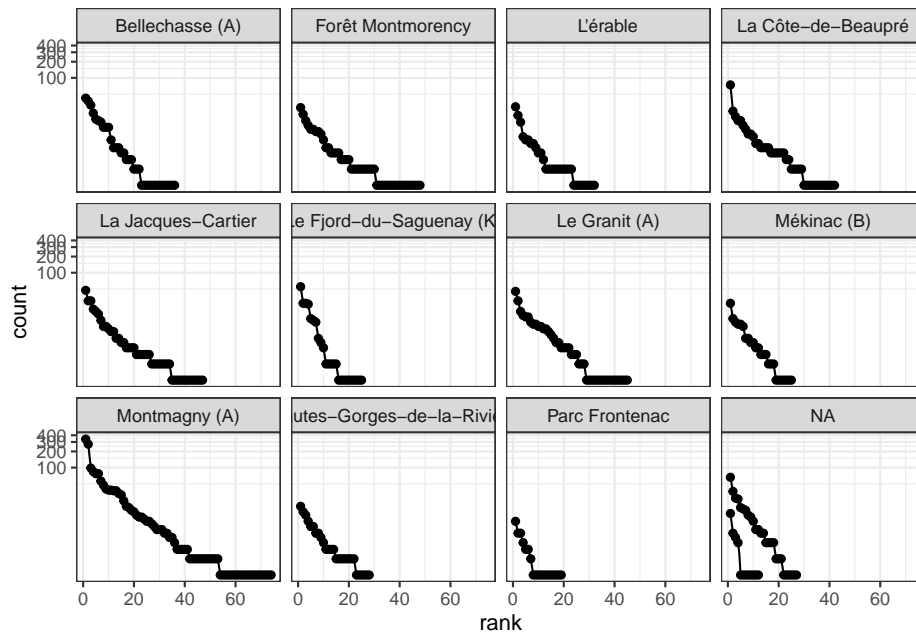
```
## # A tibble: 13 x 2
##   cell_code      n
##   <chr>      <int>
## 1 132_116         1
## 2 135_104         1
## 3 136_116         1
## 4 137_107         1
## 5 137_111         1
## 6 139_103         1
## 7 139_111         1
## 8 141_108         1
## 9 142_111         1
## 10 145_102         1
## 11 148_101         1
## 12 48_181          1
## 13 80_175          1
```

```
# no, here we have one survey per cell
```

```
rank_abundance <- microfaunes %>%
  filter(val>0) %>%
  left_join(select(cells, cell_code, name)) %>%
  group_by(cell_code, name, site_code) %>%
  mutate(rank = row_number(desc(count)))

## Joining, by = "cell_code"

rank_abundance %>%
  ggplot(aes(x = rank, y = count, group = site_code)) +
  geom_point() +
  geom_line() +
  facet_wrap(~name) +
  coord_trans(y = "log") +
  theme_bw()
```



## 5.1 Questions

Why are some site names NA? do those sites not have names?

Should 0 values be uploaded into Coleo?

## Chapter 6

# Lepidoptera

```
# confirm that the only variable recorded here is "abundance"
papillons$var %>% unique %>% {stopifnot(. == "abundance")}
```

```
## are sites nested in cells?
```

```
papillons %>%
  select(site_code, cell_code) %>% distinct %>%
  group_by(cell_code) %>% tally
```

```
## # A tibble: 11 x 2
##   cell_code      n
##   <chr>      <int>
## 1 132_116         1
## 2 135_104         1
## 3 136_116         1
## 4 137_107         2
## 5 137_111         1
## 6 139_103         1
## 7 139_111         1
## 8 141_108         1
## 9 142_111         2
## 10 145_102        2
## 11 148_101         1
```

```
# yes, cells in sites
```

```
rank_abundance <- papillons %>%
  filter(val>0) %>%
  left_join(select(cells, cell_code, name)) %>%
```

```
group_by(cell_code, name, site_code, taxa) %>%
  summarize(total = sum(count)) %>%
  mutate(rank = row_number(desc(total)))
```

```
## Joining, by = "cell_code"
```

```
## `summarise()` regrouping output by 'cell_code', 'name', 'site_code' (override with `.gr
```

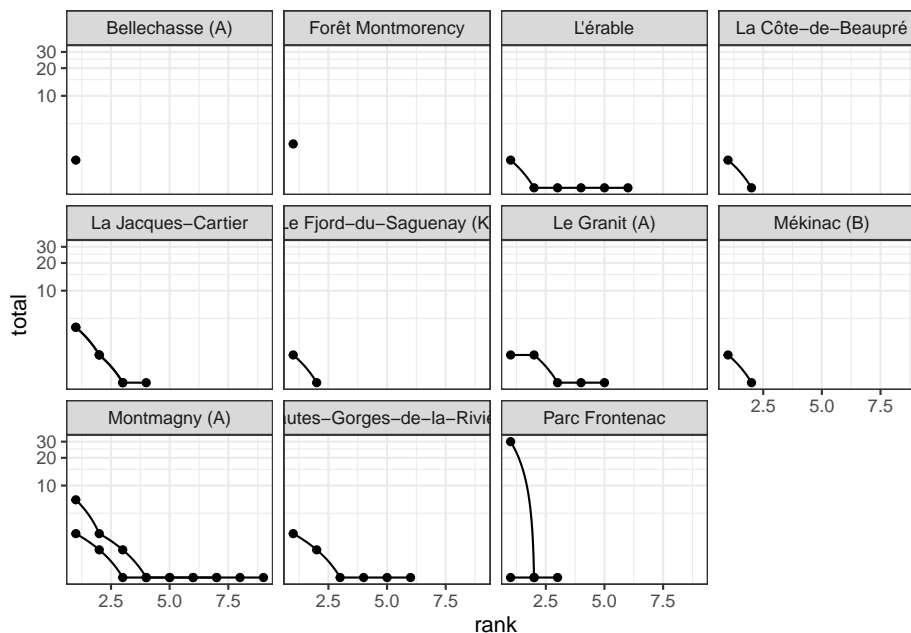
```
rank_abundance %>%
  ggplot(aes(x = rank, y = total, group = site_code)) +
  geom_point() +
  geom_line() +
  facet_wrap(~name) +
  coord_trans(y = "log") +
  theme_bw()
```

```
## geom_path: Each group consists of only one observation. Do you need to adjust
```

```
## the group aesthetic?
```

```
## geom_path: Each group consists of only one observation. Do you need to adjust
```

```
## the group aesthetic?
```



## 6.1 Questions

Why are some site names NA? do those sites not have names?

Should 0 values be uploaded into Coleo?



## Chapter 7

# Histoires de la biodiversité

1. Changement de phénologie dans le temps
2. Quelle partie de la biodiversité on connaît (sampling/rarefaction)
3. Changements de richesse au cours du temps
4. Emplacement des communautés les plus originales / uniques
5. Hotspots de biodiversité par taxa
6. Portrait de différents biomes avec des espèces emblématiques
7. Réseaux?
8. Distribution de l'âge phylogénétique des espèces
9. Distribution de l'année de première observation des espèces
10. Potentiel de human-wildlife conflict

### 7.1 Sampling / rarefaction

```
library(iNEXT)

site_spp <- papillons %>%
  select(site_code, taxa, count) %>%
  group_by(site_code, taxa) %>%
  summarize(value = sum(count)) %>% ungroup %>%
  pivot_wider(id_cols = "taxa", names_from = "site_code", values_fill = 0)

## `summarise()` regrouping output by 'site_code' (override with `.groups` argument)
pap_list <- as.list(site_spp[-1])

out <- iNEXT(pap_list, q=c(0, 1, 2), datatype="abundance", endpoint=50)

## Warning in EstiBootComm.Ind(Spec): This site has only one species. Estimation is
## not robust.
```

```
## Warning in EstiBootComm.Ind(Spec): This site has only one species. Estimation is
## not robust.
```

```
## Warning in EstiBootComm.Ind(Spec): This site has only one species. Estimation is
## not robust.
```

```
## Warning in EstiBootComm.Ind(Spec): This site has only one species. Estimation is
## not robust.
```

```
## Warning in EstiBootComm.Ind(Spec): This site has only one species. Estimation is
## not robust.
```

```
## Warning in EstiBootComm.Ind(Spec): This site has only one species. Estimation is
## not robust.
```

```
## Warning in BootstrapFun.abun(x = x, FunName, datatype, B): The Bootstrap
## community has only one species. Estimation is not robust.
```

```
## Warning in BootstrapFun.abun(x = x, FunName, datatype, B): The Bootstrap
## community has only one species. Estimation is not robust.
```

```
## Warning in BootstrapFun.abun(x = x, FunName, datatype, B): The Bootstrap
## community has only one species. Estimation is not robust.
```

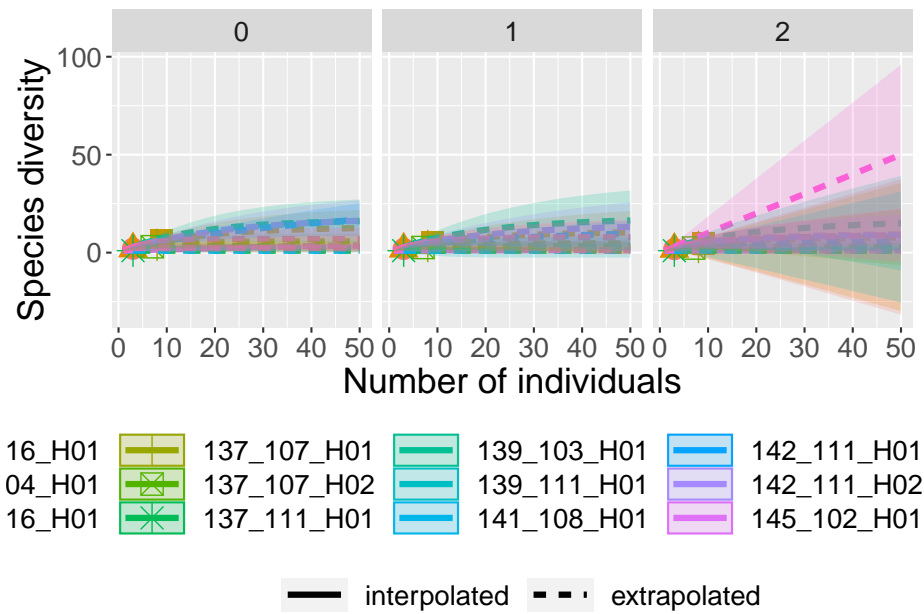
```
## Warning in BootstrapFun.abun(x = x, FunName, datatype, B): The Bootstrap
## community has only one species. Estimation is not robust.
```

```
ggiNEXT(out, type=1, facet.var="order", color.var="site")
```

```
## Warning: The shape palette can deal with a maximum of 6 discrete values because
## more than 6 becomes difficult to discriminate; you have 14. Consider
## specifying shapes manually if you must have them.
```

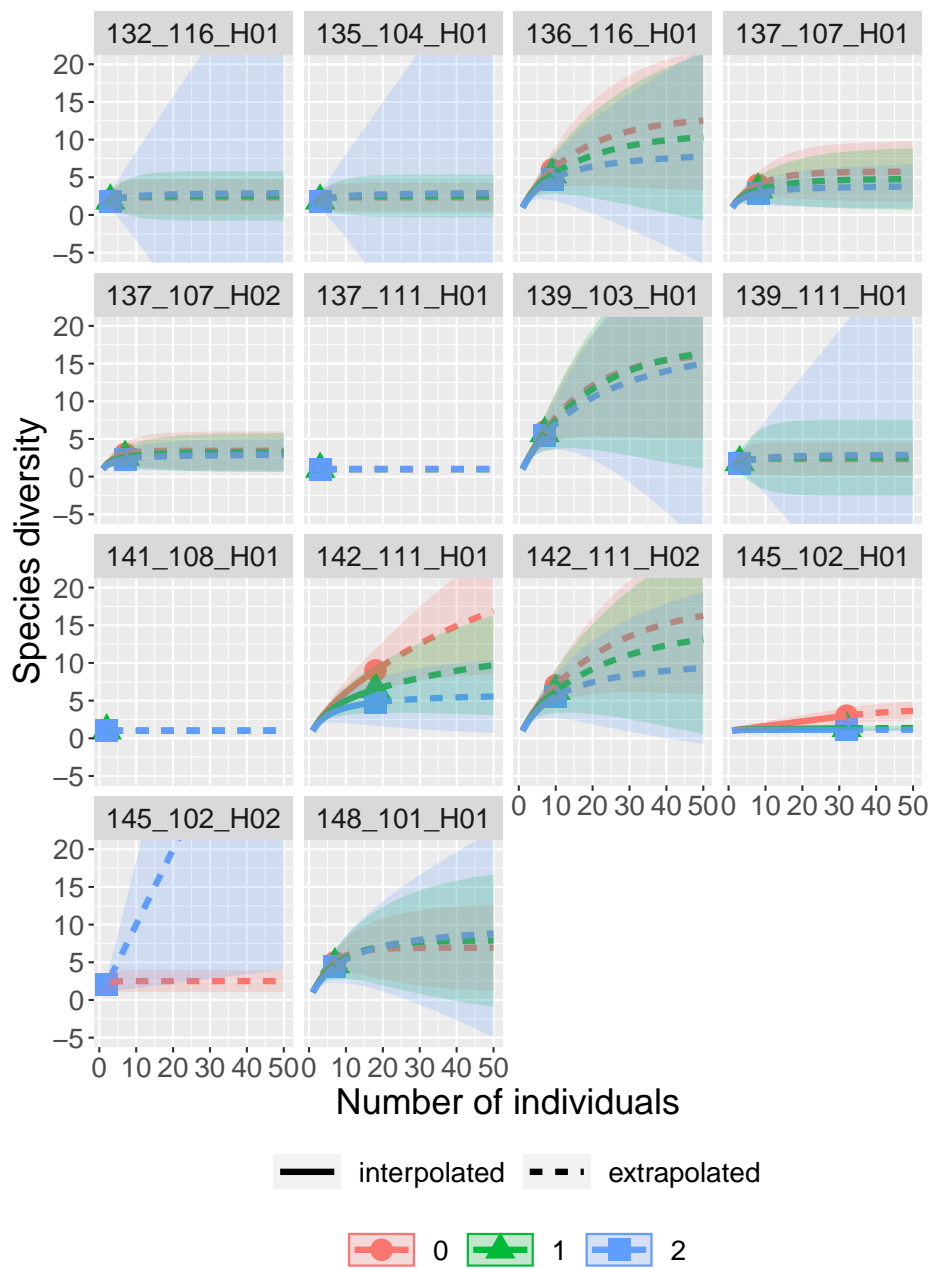
```
## Warning: Removed 24 rows containing missing values (geom_point).
```





```
ggiNEXT(out, type=1, facet.var="site") +  
  facet_wrap(~site) + coord_cartesian(ylim = c(-5,20))
```

```
## Warning: Removed 1 rows containing missing values (geom_point).
```



How many per site?

```
map_dbl(pap_list, sum)
```

```
## 132_116_H01 135_104_H01 136_116_H01 137_107_H01 137_107_H02 137_111_H01
##           3           3           9           8           7           3
```

```
## 139_103_H01 139_111_H01 141_108_H01 142_111_H01 142_111_H02 145_102_H01
##           7           3           2           18           10           32
## 145_102_H02 148_101_H01
##           2           7
```