INHERSEED

A Wiki-like document of the project

Tristan Lafont Rapnouil
Mathilde Brunel
Coralie Marais
Marie Simonin

INHERSEED's wiki

Tristan Lafont Rapnouil, Mathilde Brunel & Marie Simonin

2025-05-23

- LaTeXand its class memoir (http://www.ctan.org/pkg/memoir).

- R (http://www.r-project.org/) and RStudio (http://www.rstudio.com/)

- bookdown (http://bookdown.org/) and memoiR (https://ericmarcon.github.io/memoiR/)

Contents

Title page...

Placeholder

Data framework

Placeholder

Original seeds sources

Species list

XX plant species were used in INHERSEED. This represent a challenge both to grow and monitor such diversity as well as to sample and properly phenotype plants and seeds varying tremendously in size, shapes and looks. Acknowledging this challenge we tried to adopt protocols and data viewing/analysis allowing to compare *Arabidopsis*, carrots or beans. That being said, we recognize that critics can still be made and our results should be regarded with the diversity-challenge in mind!

The full species list:

We focused on non-woody annual species and tried to have a good spread across the Angiosperm phylogeny:

General informations

In order to interpret our results, we wanted to get as many information on our species from botany to chemistry.

###TRY First we identified a set of traits that may be found on the TRY database (Kattge et al. 2019). TRY requires manual requests and we used R to prepare our requests.

Request

First check if our species have entries in TRY.

```r
# INSTALL AND LOAD
# install.packages("rtry")
library(rtry)
library()
```

```
## Warning in library(): les bibliothèques '/usr/local/lib/R/site-library',
## '/usr/lib/R/site-library' ne contiennent aucun package
```

```r
#get list of TRY species
# download.file('https://try-db.org/dnld/TryAccSpecies.txt', destfile = "data/
```

```r
#get our species list
species_list <- readr::read_tsv("data/species_list_INHERSEED.tsv")
```

```
## Rows: 88 Columns: 1
## -- Column specification -----------------------------------------------------
## Delimiter: "\t"
## chr (1): species_for_tree
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
#check overlap and merge
try_sp <- species_list %>%
  rename(scientific_cleaned = species_for_tree)%>%
  select(scientific_cleaned) %>%
  mutate(scientific_cleaned=ifelse(scientific_cleaned=="Legousia veneris","Lego
  na.omit() %>%
  unique() %>%
  left_join(readr::read_tsv("data/TRY/TryAccSpecies.txt"),
            by = c("scientific_cleaned" = "AccSpeciesName")
  )
```

```
## Rows: 305594 Columns: 7
## -- Column specification -----------------------------------------------------
## Delimiter: "\t"
## chr (1): AccSpeciesName
## dbl (6): AccSpeciesID, ObsNum, ObsGRNum, MeasNum, MeasGRNum, TraitNum
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
#species not in try
try_sp%>%
  mutate(is_in_try=ifelse(is.na(AccSpeciesID),"Not in TRY","In Try"))%>%
  ggplot(aes(x=is_in_try,fill=is_in_try))+
  geom_histogram(stat = "count")+
  theme_minimal()+
  ylab("Number of species")+
  xlab("")+
  theme(legend.position = "none",
        text = element_text(face="bold"))+
  ggtitle(paste0(sum(!is.na(try_sp$AccSpeciesID))," species out of ", nrow(try_
```

```
## Warning in geom_histogram(stat = "count"): Ignoring unknown parameters:
## `binwidth`, `bins`, and `pad`
```

79 species out of 88 are in TRY
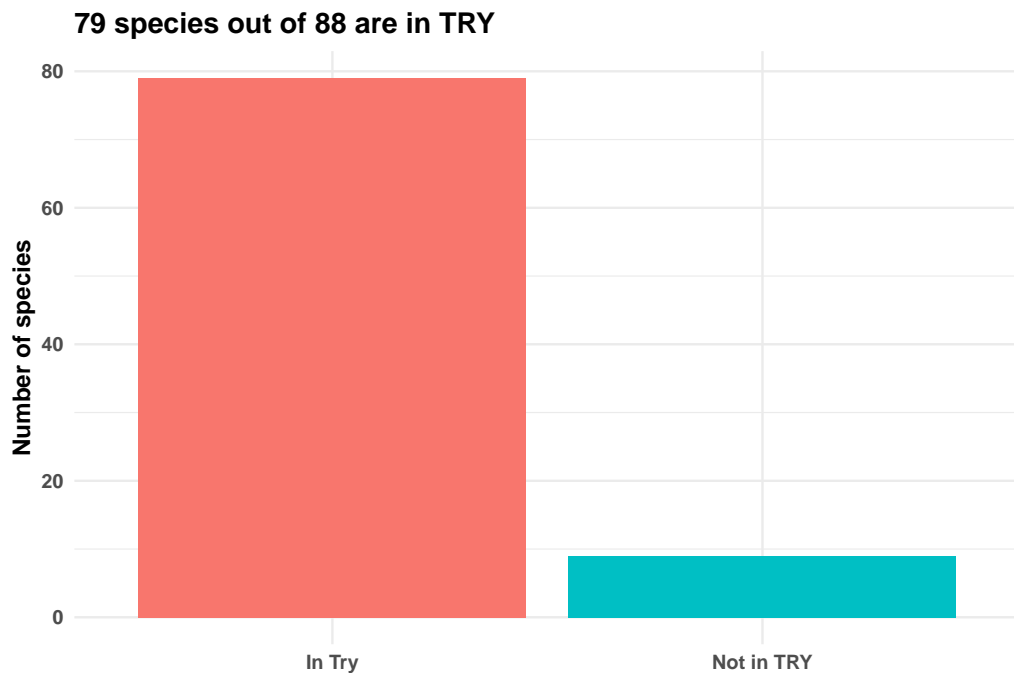
Now we can build a request for TRY with our traits of interest. We handpicked from TRY traits list XX traits that were relevant to us.

```r
try_request <- try_sp%>%
  summarise(species_list=paste0(AccSpeciesID,collapse = ", "))
try_request%>%
  readr::write_tsv("data/TRY/requests/TryReqSpecies.tsv")

traits_try <- readxl::read_xlsx("data/TRY/traits_of_interest/Traits repro TRY d
try_request_traits <- traits_try%>%
  filter(tokeep=="y")%>%
  select(TraitID)%>%
  summarise(traits_list=paste0(TraitID,collapse = ", "))
try_request_traits%>%
  readr::write_tsv("data/TRY/requests/TryReqTraits.tsv")
# 41259 - ID of the first request
```

Then we requested TRY data access using our species and traits lists.
####Release

```r
try_release <- readr::read_tsv("data/TRY/release/41259_1905202519280_TRY_relaea
```

```
## New names:
## * `` -> `...29`
```

```
## Warning: One or more parsing issues, call `problems()` on your data frame for details,
## e.g.:
##    dat <- vroom(...)
##    problems(dat)
```

```
## Rows: 417530 Columns: 29
## -- Column specification --------------------------------------------------
## Delimiter: "\t"
## chr (16): LastName, FirstName, Dataset, SpeciesName, AccSpeciesName, TraitNa...
## dbl (11): DatasetID, AccSpeciesID, ObservationID, ObsDataID, TraitID, DataID...
## lgl  (2): RelUncertaintyPercent, ...29
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
try_release %<>%
  select(AccSpeciesName, TraitName, StdValue) %>%
  rename(
    species = AccSpeciesName,
    trait = TraitName,
    value = StdValue
  ) %>%
  group_by(species, trait) %>%
  summarise(value = mean(as.numeric(value), na.rm = TRUE)) %>%
  tidyr::pivot_wider(values_from = "value", names_from = "trait")%>%
  mutate(across(everything(),~ifelse(.x=="NaN",NA,.x)))
```

```
## `summarise()` has grouped output by 'species'. You can override using the
## `.groups` argument.
```

```r
try_release%>%
  reshape2::melt(id.vars=c("species"))%>%
  group_by(variable)%>%
  mutate(is_na_var=ifelse(is.na(value),0,1))%>%
  summarise(pct_species_avail=(sum(is_na_var)/82)*100)%>%
  ggplot(aes(x=pct_species_avail,y=variable))+
  geom_point()+
  theme_minimal()
```

```r
try_release%>%
  reshape2::melt(id.vars=c("species"))%>%
  group_by(variable)%>%
  mutate(is_na_var=ifelse(is.na(value),0,1))%>%
  summarise(pct_species_avail=(sum(is_na_var)/82)*100)%>%
  filter(pct_species_avail>0)%>%
  readr::write_tsv("data/TRY/release/traits_with_some_values.tsv")
```

They helped us!

- Coralie MARAIS
- Cas
- Emile
- Claude
- Rémi GARDET
- Elodie LEMMENS

#References

Bibliography

Kattge, J. et al. (Dec. 31, 2019). "TRY plant trait database – enhanced coverage and open access." In: *Global Change Biology* 26.1, pp. 119–188. DOI: [10.1111/gcb.14904](). URL: http://dx.doi.org/10.1111/gcb.14904 (cit. on p. xi).

List of Figures

Abstract

Keywords Keyword in English, As a list.