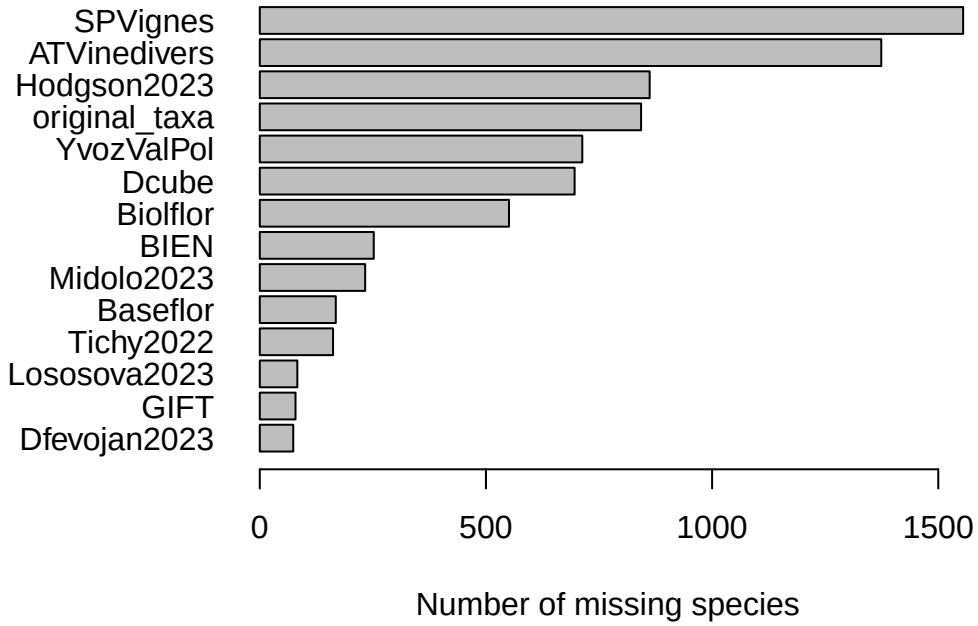
Exploration of FELLOW trait dataset

The objective of this document is to:

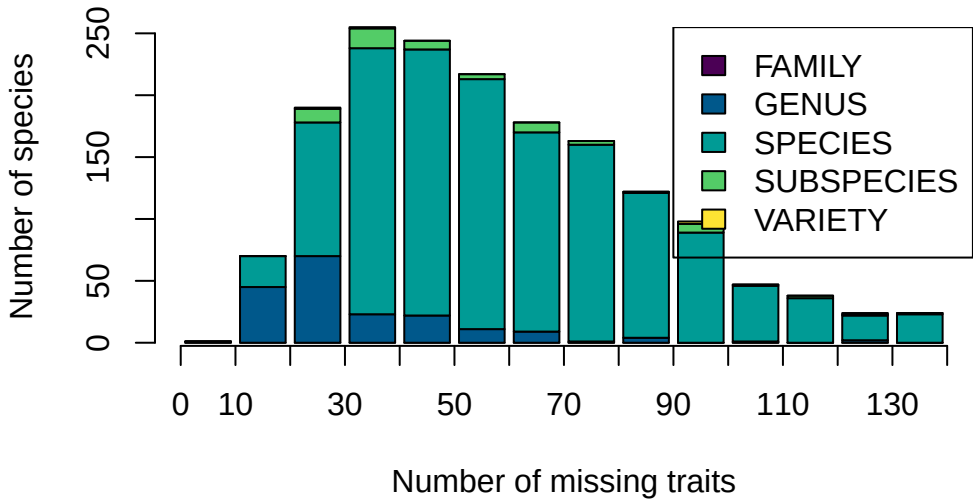
- visually explore the trait database
- understand trait coverage / data gaps
- check for possible inconsistencies

Description of trait databases

So far, we compiled 152 traits for 1706 taxa gathered from 14 trait databases. But there are many missing values.




SPVignes
ATVinedivers
Hodgson2023
original_taxa
YvozValPol
Dcube
Biolflor
BIEN
Midolo2023
Baseflor
Tichy2022
Lososova2023
GIFT
Dfevojan2023
500
1000
1500
Number of missing traits values


Number of species
0
50
150
250
FAMILY
GENUS
SPECIES
SUBSPECIES
VARIETY
0
10
30
50
70
90
110
130
Number of missing traits

Taxa with no or limited trait information (N=28).

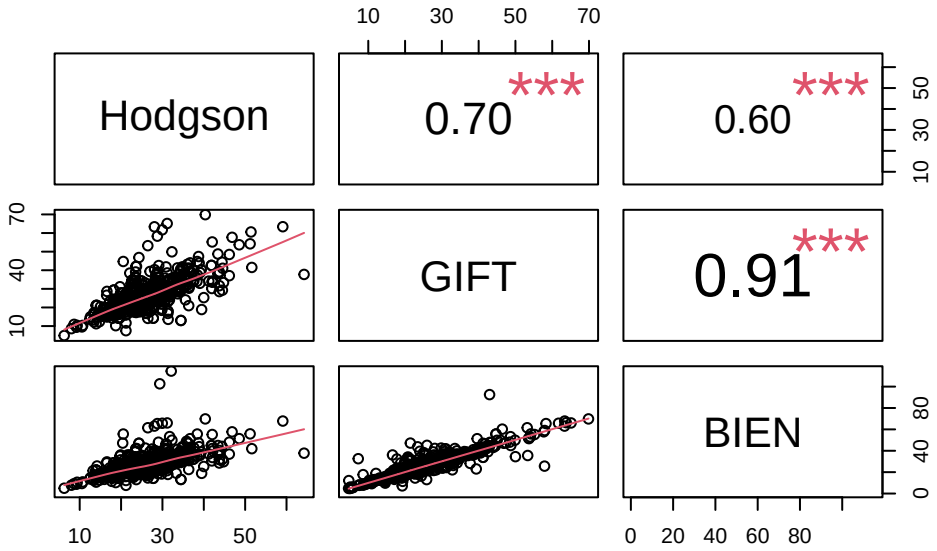
```
 [1] "Acacia"                 "Agrimonia agrimonoides" "Agropyron"
 [4] "Amaranthaceae"          "Apiaceae"               "Asparagaceae"
 [7] "Brassicaceae"           "Caryophyllaceae"        "Crambe abyssinica"
[10] "Dysphania aristata"     "Ficaria ambigua"        "Glyceria"
[13] "Lamiaceae"              "Liliaceae"              "Melomphis arabica"
[16] "Onopordum gautieri"     "Paronychia"             "Piptatherum"
[19] "Poaceae"                "Roemeria hispida"       "Rosaceae"
[22] "Rubiaceae"              "Salix"                  "Sanguisorba"
[25] "Scorpiurus subvillosus" "Viburnum"               "Vicia dasycarpa"
[28] "Vicia pseudocracca"
```

Open question: **How to deal with families taxa?**
How to deal with missing trait values? Trait imputation, discarding taxa, ...

Comparison

SLA

There are three sources of information for Specific leaf area (SLA) : Hodgson et al. 2023 (in mm2/mg), GIFT (in cm2/g) and BIEN (in m2/kg = mm2/mg).



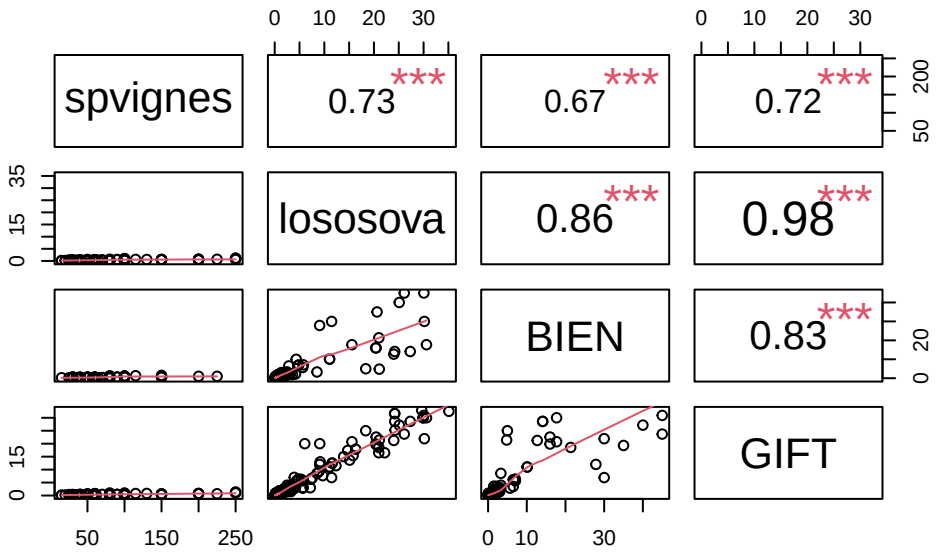
Values are highly correlated, so we could imagine filling the missing values (using preferred data sources or averaging them).

Number of NAs:

```
Hodgson    GIFT    BIEN  filled
    862     731     720     475
```

Plant height

There are four sources of information for plant height.



Not sure what is the unit of `Plant_Height` in SPVignes (cm?), and why it is limited to only small plants (<1m). Yet we could potentially fill the height trait values.

Number of NAs:

```
spvignes lososova      BIEN      GIFT    filled
    1582       129      1276       492       109
```