



IAVS-LAC

This Session

Reconciliation

taxlist

Data Shapes

From the field to the desk

[Session 2]

Miguel Alvarez

28th April 2023



This Session

IAVS-LAC

This Session

Reconciliation

taxlist

Data Shapes

- ▶ Taxonomic Names Resolution
- ▶ Taxonomic Lists in taxlist
- ▶ Data Formats and Shapes



Taxonomic Names Resolution

- ▶ Taxonomy Equalization
- ▶ Taxonomic Standardization
- ▶ Taxonomic Harmonization
- ▶ Taxonomic Normalization
- ▶ Taxonomic Reconciliation

```
library(taxlist)
summary(Easplist, "Cyclosorus interruptus")

## -----
## concept ID: 50074
## view ID: 1
## level: species
## parent: 55055 Cyclosorus Link
##
## # accepted name:
## 50074 Cyclosorus interruptus (Willd.) H. Itô
##
## # synonyms (13):
## 52002 Dryopteris gongylodes (Schkuhr) Kuntze
## 52008 Thelypteris interrupta (Willd.) K. Iwa
## 52009 Cyclosorus striatus Ching
## 53097 Pteris interrupta Willd.
## 53098 Aspidium continuum Desv.
## 53099 Aspidium ecklonii Kunze
## 53100 Aspidium gongylodes Schkuhr
## 53101 Aspidium obtusatum Sw
```



Reconciliation

IAVS-LAC

This Session

Reconciliation

taxlist

Data Shapes

Online Databases

IPNI

Tropicos

TNRS

WFO

GBIF

Flora del Conosur

African Plant Database



Reconciliation

IAVS-LAC

This Session

Reconciliation

taxlist

Data Shapes

Several R Packages

taxize

Taxonstand

Grenié et al. (2022) *Methods in Ecology and Evolution*

Daniel Laughlin, Yungas, Machu Picchu, Peru





- ▶ Object structure
- ▶ Methods for common process
 - ▶ `print()`
 - ▶ `summary()`
 - ▶ `subset()`
 - ▶ `indented_list()`

[taxlist site](#)

Alvarez & Luebert (2018) *Biodiversity Data Journal*

 **taxlist** 0.2.3 Se

taxlist

CRAN 0.2.2  Peer Reviewed  DOI 10.5281/zenodo.5107156

 R CMD-check passing  codecov 89%

downloads 465/month downloads 27K



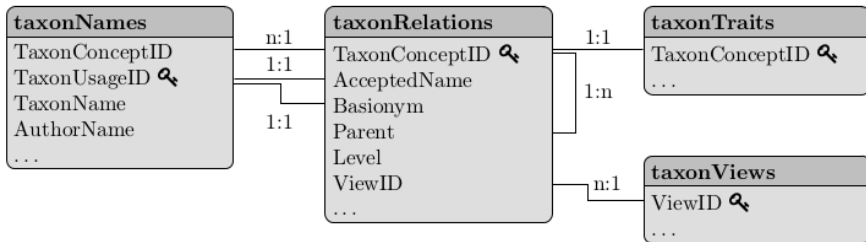
Introduction

`taxlist` is a package designed to handle and assess taxonomic lists in R, providing an object class and functions in S4 language. The homonymous object class `taxlist` was originally designed as a module for taxa recorded in vegetation-plot observations (see [vegetable](#)), but became as an independent object with the ability of contain not only lists of species but also synonymy, hierarchical taxonomy, and functional traits (attributes of taxa).

The main aim of this package is to keep consistence in taxonomic lists (a set of rules are checked by the function `validObject()`), to enable the re-arrangement of such data, and to statistically assess functional traits and other attributes, for instance taxonomy itself (function `tax2traits()` set taxonomic information as trait).

While this package only includes a function for the import of taxonomic lists from [Turboveg](#), almost any data source can be structured as `taxlist` object, so far the information is imported into data frames in an R session and the consistency rules are respected (validity).

The use of `taxlist` is recommended for people cleaning raw data before importing it to relational databases, either in the context of taxonomic work or biodiversity assessments. The other way around, people having relational databases or clean and structured taxonomic lists may



- ▶ Each slot is a data frame
- ▶ Consistency checked by `validObject()`



Target Information

- ▶ Diverse degrees of completeness
- ▶ Any hierarchical taxonomy
- ▶ Also syntaxonomy

Alvarez & Luebert (2022)
Vegetation Classification and Survey

<https://syntax.kamapu.net/>

Syntaxa of South American Vegetation

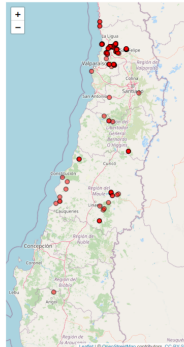
Search a syntaxon:

Lithraea causticae-Cryptocaryetea albae ▼

Lithraea causticae-Cryptocaryetea albae Oberdorfer 1960
 Cryptocaryetalia albae (Schmidt 1954) Oberdorfer 1960
 Cryptocaryon albae Schmidt 1954
 Lucuma valparadisiensis Cryptocaryetum albae (Schmidt 1954) Oberdorfer 1960
 = Lucuma valparadisiensis Schmidt 1954
 Beltschmedion miersii Amigo & Flores-Toro 2012
 Aextoxia punctata-Cryptocaryetum albae Amigo & Flores-Toro 2012
 Beltschmedion miersii Schmidt 1954
 Beltschmedion miersii osorobiotum berteri Schmidt 1954
 Beltschmedion miersii typicum Schmidt 1954
 Beltschmedion miersii-Crocodendrium patagiae Villaseñor & Serey ex Amigo & Flores-Toro 2012
 Cryptocaryon albae Oberdorfer 1960
 Boldo chilensis-Cryptocaryetum albae Oberdorfer 1960
 Kigeonekia angustifoliae-Quilajetum saponariae Amigo & Flores-Toro 2013
 Colliguajo integerrimae-Quilajetum saponariae Amigo & Flores-Toro 2013
 Mutiso latifoliae-Quilajetum saponariae Amigo & Flores-Toro 2013
 Schino montani-Austrocedretum chilensis Amigo & Flores-Toro 2013
 Lithraea causticae Schmidt 1954
 Azaro dentatae-Lithraetum causticae Amigo & Flores-Toro 2013
 Boldo chilensis-Lithraetum causticae (Schmidt 1954) Oberdorfer 1960
 = Lithraea causticae-Prunus bolusii Schmidt 1954
 Lomatium hispidum-Lithraetum causticae Amigo, San Martín & Quintanilla 2000
 Jubaea spectabilis-Lithraetum causticae Schmidt 1954
 Austrocedrus chilensis-Lithraetum causticae Oberdorfer 1960
 Acacia cavenia-Cestrum parqui Oberdorfer 1960
 Cestrum parqui-Trevoetum tinivae Oberdorfer 1960
 Azaro gillessii-Escaloniolum berteriense Oberdorfer 1960

Plot Observations

Number of records: 146







Data Shapes

IAVS-LAC

This Session

Reconciliation

taxlist

Data Shapes



Vegetation-plot observation

Phytosociological relevés

- ▶ Abundance/presence of species in plot
- ▶ Environmental variables
- ▶ Species attributes
 - ▶ Life forms
 - ▶ Chorology
 - ▶ Functional traits
 - ▶ Taxonomy



Data Shapes

IAVS-LAC

... served to the readers

Table 2. *Tessario absinthioidis*-*Baccharidetum marginalis* Oberd. 1960
Riverbank composition: R/r = rocks (> 50 cm Ø); G/g = gravel (< 50 cm Ø);
indicate abundant, lowercase scarce; (x) = xenophytic plant

Altitude (m asl)	160	225	1110	1055	505
Area (m ²)	80	180	100	250	200
Slope (°)	0	0	0	10	0
Cover E _i (%)	100	85	100	60	70
High shrub layer (m)	5-8	3-8	1-3	1-2	1-2.5
Riverbank composition	gsL	gS	rGS	RGS	rGS
Latitude	34° 03'	34° 22'	35° 02'	35° 02'	35° 42'
N. species	11	10	14	9	13
N. relevé	1	2	3	4	5
<i>Baccharidetum</i> / <i>Salicetea</i> species					
<i>Baccharis salicifolia</i>	1	3	3	3	4
<i>Otholobium glandulosum</i>	2	1	.	.	3
<i>Cortaderia</i> gr. <i>selloana</i>	.	r	1	+	1
<i>Salix humboldtiana</i>	4	4	.	.	.
<i>Tessaria absinthioides</i>	.	3	4	3	+
<i>Equisetum bogotense</i>	.	.	2	.	1
<i>Discaria trinervis</i>	.	.	3	.	.
<i>Buddleja globosa</i>	.	.	+	.	.
<i>Escallonia myrtoidea</i>	.	.	+	.	.
<i>Myrceugenia lanceolata</i>	.	.	.	.	.

... served to the statistical software

##	Adesparv	Bromtris	Cardramo
## 6216	1	1	1
## 6217	0	1	0
## 6218	0	0	0
## 6219	0	0	0
## 6220	0	0	0
## 6221	0	0	0
## 6222	0	0	0
## 6359	0	0	0
## 6360	0	1	1
## 6361	0	1	0
## 6362	0	1	0
## 6523	0	1	0
## 6524	0	1	0
## 6574	0	0	0
## 6575	0	0	0
## 6576	0	1	1
## 6577	0	0	0
## 6578	0	0	0
## 6579	0	0	0
## 6580	0	0	0
##	Crasdecu	Dichseri	
## 6216	1	1	
## 6217	0	0	



Data Shapes

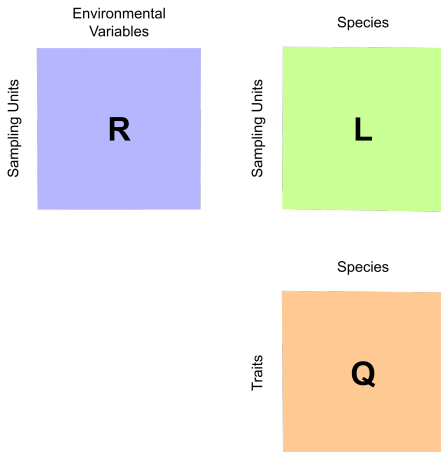
IAVS-LAC

This Session

Reconciliation

taxlist

Data Shapes



- ▶ **R** environmental table (header)
- ▶ **Q** species-trait table (vegetation matrix)
- ▶ **L** species composition table (species attributes)

Dolédec et al. (1996) Environmental and Ecological Statistics



Data Shapes

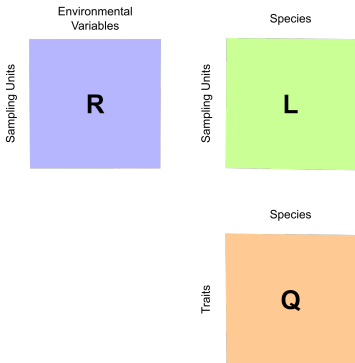
IAVS-LAC

This Session

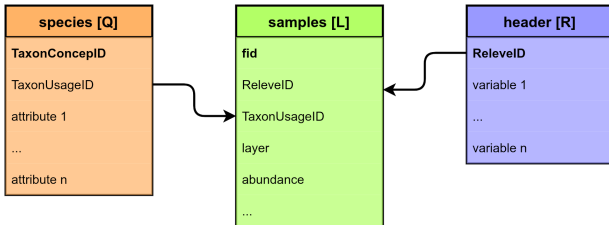
Reconciliation

taxlist

Data Shapes



Relational database schema



- ▶ Efficient storage
 - ▶ Efficient data assessment?
- ▶ Little redundancy
- ▶ Consistency



Data Shapes

IAVS-LAC

This Session

Reconciliation

taxlist

Data Shapes

Vegetation matrix cross table

```
##      Adesparv Bromtris Cardramo
## 6216      1      1      1
## 6217      0      1      0
## 6218      0      0      0
## 6219      0      0      0
## 6220      0      0      0
## 6221      0      0      0
## 6222      0      0      0
## 6359      0      0      0
## 6360      0      1      1
## 6361      0      1      0
## 6362      0      1      0
## 6523      0      1      0
## 6524      0      1      0
## 6574      0      0      0
## 6575      0      0      0
## 6576      0      1      1
## 6577      0      0      0
## 6578      0      0      0
##      Crasdecu Dichseri
## 6216      1      1
## 6217      0      0
```

Database list column-oriented table

```
## Loading required package: vegtable
##      record_id ReleveID codename cover
## 9329      137451      7993 Hypoglab      1
## 9330      137452      7993 Lythhyss      1
## 9331      137453      7993 Micrnana      1
## 9332      137454      7993 Piluamer      1
## 9333      137455      7993 Psilbrev      1
## 9335      137457      7993 Isolcern      2
## 9337      137459      7993 Plagsess      1
## 9338      137460      7994 Lythhyss      1
## 9339      137461      7994 Piluamer      1
## 9340      137462      7994 Psilbrev      1
## 9343      137465      7994 Plagsess      1
## 9344      137466      7994 Isoehier      1
## 9345      137467      7258 Agrocapi      1
## 9346      137468      7258 Bromtris      1
## 9347      137469      7258 Erodmala      1
## 9348      137470      7258 Hypoglab      1
## 9349      137471      7258 Juncbufo      1
## 9350      137472      7258 Lepuspat      1
```



Thank You!

Jesse Fleri, mangrove, Gilligan's Island, Puerto Rico