

Taxonomic Data with taxize

Cheat Sheet



Use Cases

Get taxonomic IDs

taxize::get_ids(names) /
taxize::get_colid(names)

Puma concolor → 2435099
Ursus americanus → 2433407
Quercus robur → 2878688

Get correct spellings

taxize::tnrs(names) /
taxize::gnr_resolve(names)

Pumma concolor → *Puma concolor*

Retrieve higher taxonomic classification

taxize::classification(names)

Classification for *Chironomus riparius*

	name	rank	id
1	Animalia	Kingdom	22254411
2	Arthropoda	Phylum	22254412
3	Insecta	Class	22254618
4	Diptera	Order	22254663
5	Chironomidae	Family	22254676
6	Chironomus	Genus	22350527
7	Chironomus riparius	Species	8663146

Fetch all taxonomic names downstream from a target taxonomic group

taxize::downstream(names)

Species downstream from the genus *Apis*

	childtaxa_id	childtaxa_name	childtaxa_rank
1	6971712	Apis andreniformis	Species
2	6971713	Apis cerana	Species
3	6971714	Apis dorsata	Species
4	6971715	Apis florea	Species
5	6971716	Apis koschevnikovi	Species
6	6845885	Apis mellifera	Species
7	6971717	Apis nigrocincta	Species

Data Sources (hyperlinked)

[Encyclopedia of Life](#)
[Taxonomic Name Resolution Service](#)
[Integrated Taxonomic Information Service](#)
[Phylomatic](#)
[uBio](#)
[Global Names Resolver](#)
[Global Names Index](#)
[IUCN Red List](#)
[Tropicos](#)
[Plantminer](#)
[Theplantlist](#)
[Catalogue of Life](#)
[Global Invasive Species Database](#)
[National Center for Biotechnology Information](#)
[CANADENSYS Vscan name search API](#)
[International Plant Names Index \(IPNI\)](#)
[Barcode of Life Data Systems \(BOLD\)](#)
[National Biodiversity Network \(UK\)](#)

Get taxonomic IDs

Options

- ☒ Interactively select names.
- ☒ Or, get back all names to process later.
- ☒ Or, get back certain records.

taxize::get_colid(id)

If > 1 result, interactively select by row

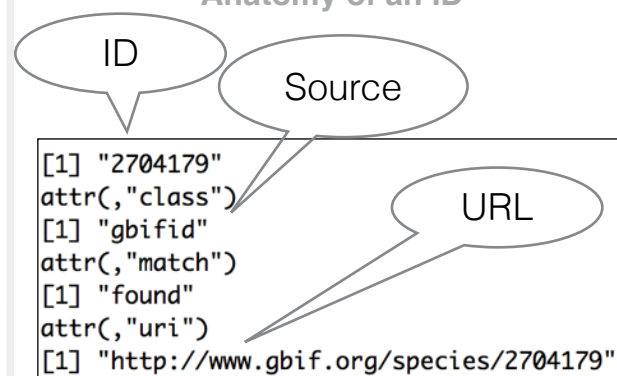
taxize::get_colid(id, rows = 1:4)

Get rows 1 through 4

taxize::get_colid_(id)

Get all data

Anatomy of an ID



Coerce to Taxonomic ID

With taxonomic IDs, you have a stable reference to a species for a specific source, and you can do many things in taxize with ids (See box to the right for examples).

From: numeric, character, list, or data.frame, e.g.:

taxize::as.colid(8663146)

Validate

Set **check=FALSE** to not check if ID is valid - if **TRUE** we ping the data source to make sure the ID is good.

taxize::as.colid(8663146, check = FALSE)

All **taxize::as.*()** take one or many IDs

Correct Names

Out of the box frameworks

taxize::gnr_resolve(names)

taxize::tnrs(names)

Queries NCBI, Mammals dataset, and iPlant, only fuzzy matching on plants

taxize::iplant_resolve(names)

iPlant has a nice name resolver - of course, it only works for plants

Bonus: **taxize::resolve(names)** is an interface to all three of the resolution functions

DIY correction

There are a lot of functions to search various taxonomic databases - they don't do name resolution per se. But you can use them to achieve the same end as the above functions. Some examples:

taxize::col_search(names)

Search Catalogue of Life

taxize::nbn_search(names)

Search UK Nat. Biodiversity Network

taxize::tp_search(names)

Search Tropicos

taxize::gni_search(names)

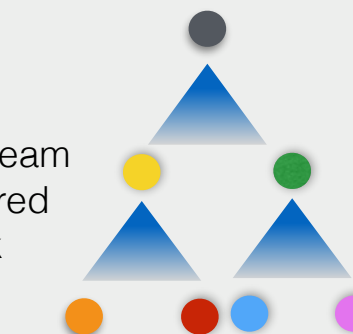
Search Global Names Index

Down-/up-stream, children, and more

● Target Taxon

taxize::downstream(names, "species")

Downstream to desired rank



taxize::children(names)

Immediate children only



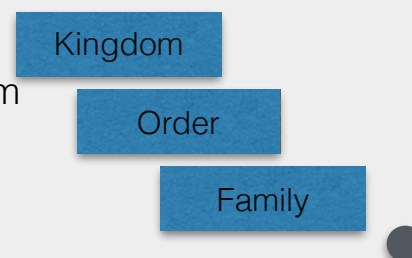
taxize::upstream(names, "genus")

Upstream to desired rank



taxize::classification(names)

All upstream names for given taxon



Combine Classification Data

Stack vertically - rbind

out <- taxize::classification(ids, db = "ncbi")

taxize::rbind(out)

	name	rank	id
1	Animalia	Kingdom	22254411
2	Arthropoda	Phylum	22254412
3	Insecta	Class	22254618

Combine horizontally - cbind

taxize::cbind(out)

kingdom	phylum	class	order
Protozoa	Euglenophycota	Euglenophyceae	Euglenales
Animalia	Platyhelminthes	Trematoda	Plagiorchiida