

Desktop analysis with Atlas data: ALA4R



ATLAS OF **LIVING**
AUSTRALIA
sharing biodiversity knowledge

Ben Raymond
Jeremy VanDerWal
Lee Belbin

What is ALA4R?

- spatial portal: value of integrated data
- ALA4R: access to ALA data from within R
 - apply own/others R tools



What is R?

- open-source statistical computing environment
 - programming language & software environment
 - statistical, mathematical, modelling, and graphical tools
 - many additional packages
- since 90s, mid-2000s
- now commonly used in ecology & biology



ALA4R core functions

- name searching
- taxon information
- occurrences and environmental/contextual data
- support functions
 - extracted data cached to local machine



ALA4R core functions

search_names

search_partial_name

search_fulltext

species_download (bulk taxonomic info)

species_info (taxon profile)



ALA4R core functions

occurrences

(with environmental/contextual data)

specieslist

species_by_site

search_layers

intersect_points

(environmental/contextual data at points)



Names and taxonomy



```
search_fulltext("penguins")
```

name	rank	score	commonName
1 SPHENISCIDAE	family	7.76	Penguins
2 <i>Eudyptula minor</i>	species	0.68	Blue Penguin, Fairy Penguin*
3 <i>Eudyptes chrysocome</i>	species	3.0e-9	Crested Penguin
4 <i>Eudyptes pachyrhynchus</i> robustus	subspecies	2.6e-9	Snares Crested Penguin
5 <i>Pteria penguin</i>	species	1.3e-9	Penguin Wing Oyster

* common name list included "Little Penguin In The Manly Point Area (being The Area On And Near The Shoreline From Canuae Point Generally Northward To The Point Near The Intersection Of Stuart Street And Oyama Cove Avenue, And Extending 100 Metres Offshore From That Shoreline)"



```

tx=species_download("family:SPHENISCIDAE",
  fields=c("guid", "parentGuid", "genus",
  "nameComplete", "rank"))

tx=tx[tx$Taxon.Rank %in%
  c("species", "subspecies"), ]

```

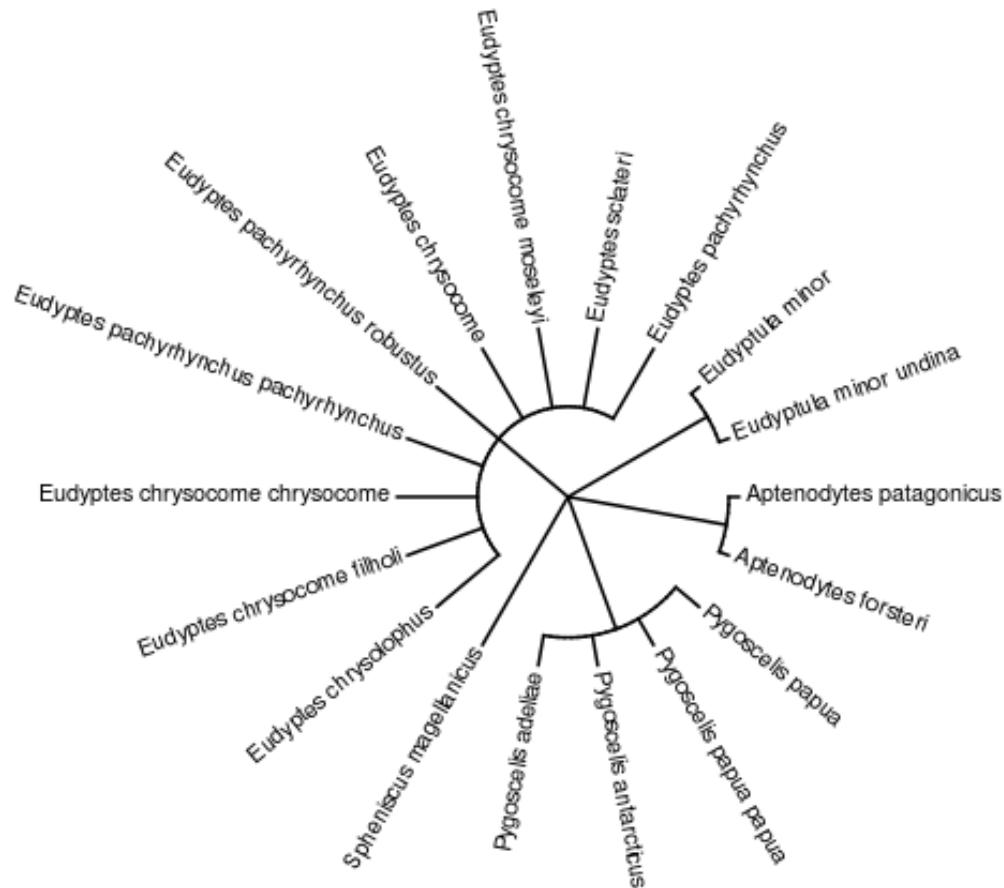
→ Table of taxonomic classifications for 18 penguin species/subspecies

guid	Genus	Scientific.Name	Taxon.Rank
1 urn:lsid:...	Eudyptula	Eudyptula minor undina	subspecies
2 urn:lsid:...	Eudyptes	Eudyptes pachyrhynchus	species
3 urn:lsid:...	Eudyptes	Eudyptes sclateri	species
4 urn:lsid:...	Spheniscus	Spheniscus magellanicus	species
5 urn:lsid:...	Eudyptes	Eudyptes chrysocome moseleyi	subspecies



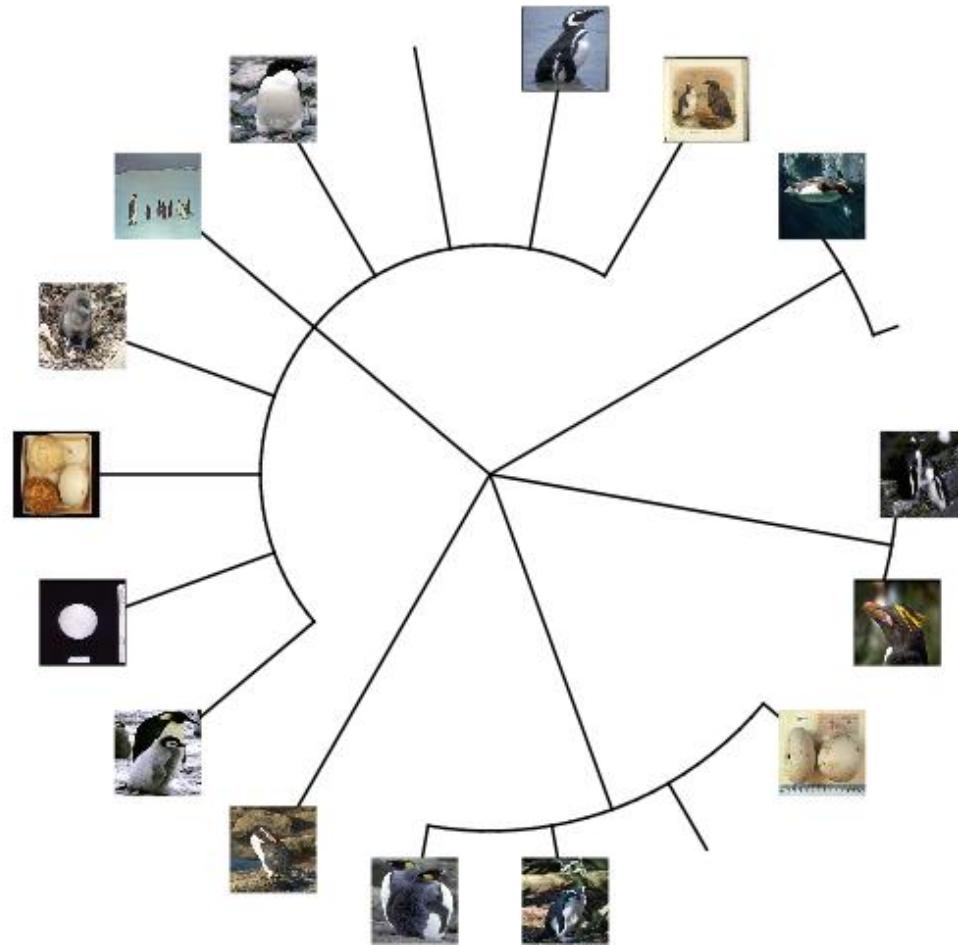
Plot taxonomic tree

```
library(ape); library(phytools)  
ax=as.phylo(~Genus/Scientific.Name, data=tx)  
plotTree(ax, type="fan", ...)
```



Plot taxonomic tree

```
s=lapply(tx$guid, function(z) {species_info(guid=z)} )  
# look for image URLs and plot
```



Listed species in an area

(see also regions.ala.org.au)



```

library(maptools); library(rgeos)
shape=readShapePoly(...)
wkt=writeWKT(shape)
x=specieslist(wkt=wkt,
fq="state_conservation:*",page_size=100)

```

Scientific.name	Common.name	N.occurrences
1 Acacia iteaphylla	Flinders Range Wattle	7
2 Adiantum capillus-veneris	Dainty Maiden-hair	6
3 Anogramma leptophylla	Annual Fern	4
4 Antechinus flavipes	Yellow-footed Antechinus	11
5 Anthocercis angustifolia	Narrow-leaf Ray-flower	25
6 Anthochaera (Xanthomyza) phrygia	Regent Honeyeater	2
7 Austrostipa densiflora	Foxtail Spear-grass	3



Assertions and data quality



```
x=occurrences(taxon="Amblyornis newtonianus",
download_reason_id=6)
summary(x)
```

number of names: **7**

number of taxonomically corrected names: **1**

number of observation records: 881

assertions checked: 22

 Invalid.collection.date: 119 records

 incomplete.Collection.Date: 159 records

 First.of.the.century: 4 records

...

```
x=occurrences(taxon="Amblyornis newtonianus",
download_reason_id=6)
summary(x)
```

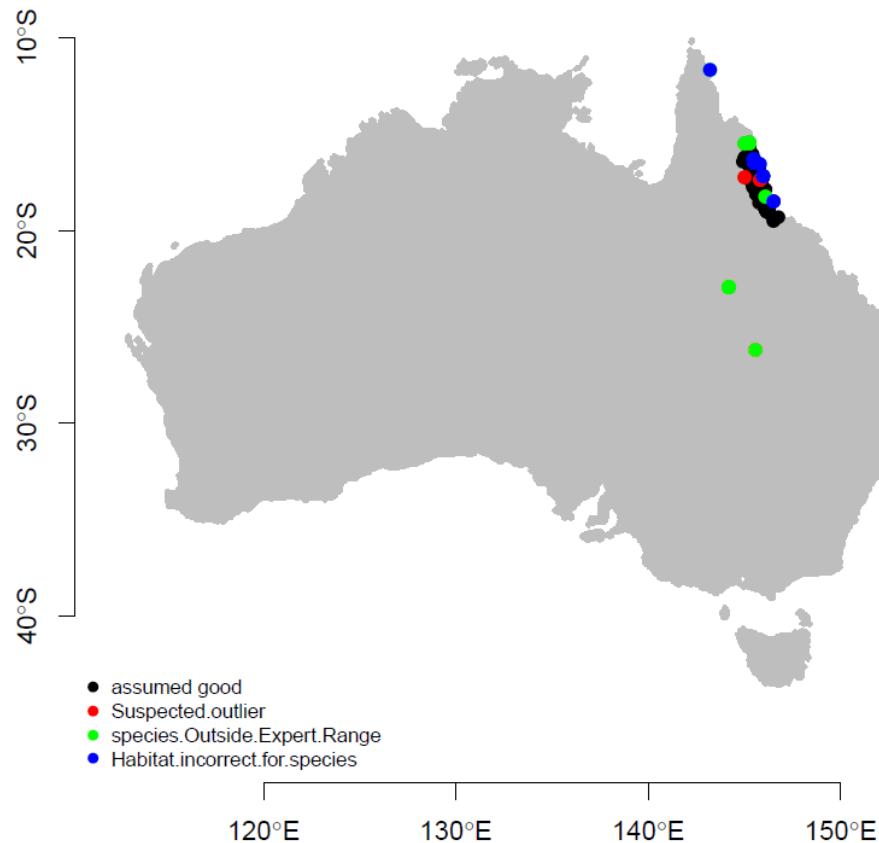
Suspected.outlier: 13 records -- **considered fatal**
Coordinate.accuracy.not.valid: 12 records
First.of.the.year: 26 records
Altitude.value.supplied.in.feet: 2 records
geodetic.Datum.Assumed.Wgs.84: 619 records
species.Outside.Expert.Range: 14 records -- **considered fatal**
decimal.Lat.Long.Converted: 5 records
Coordinate.precision.not.valid: 18 records
Country.inferred.from.coordinates: 455 records
Image.URL.invalid: 1 records
unrecognized.Geodetic.Datum: 105 records
inferred.Duplicate.Record: 104 records
Coordinates.dont.match.supplied.state: 1 records
Habitat.incorrect.for.species: 14 records -- **considered fatal**
First.of.the.month: 65 records



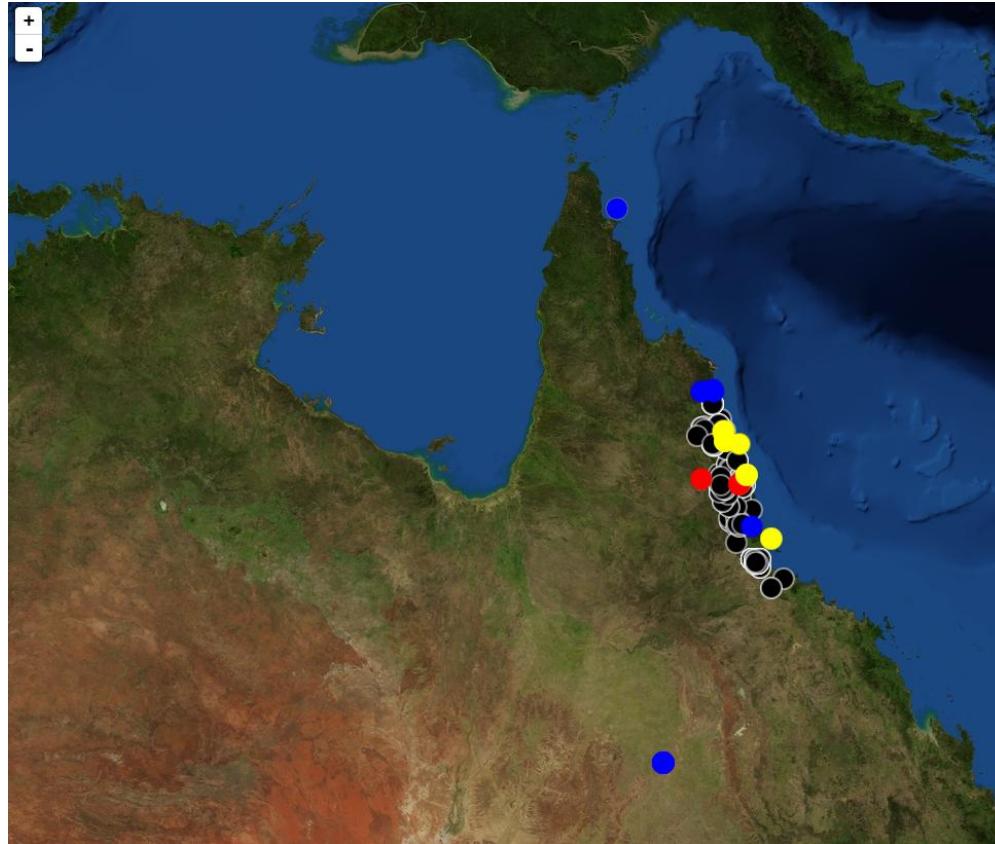
Viewing data

```
occurrences_plot(x, qa='fatal')
```

Amblyornis newtonianus



```
library(leafletR)  
dat=toGeoJSON(data=x,...)  
sty=styleCat(...)  
alamap=leaflet(data=dat,style=sty,...)
```



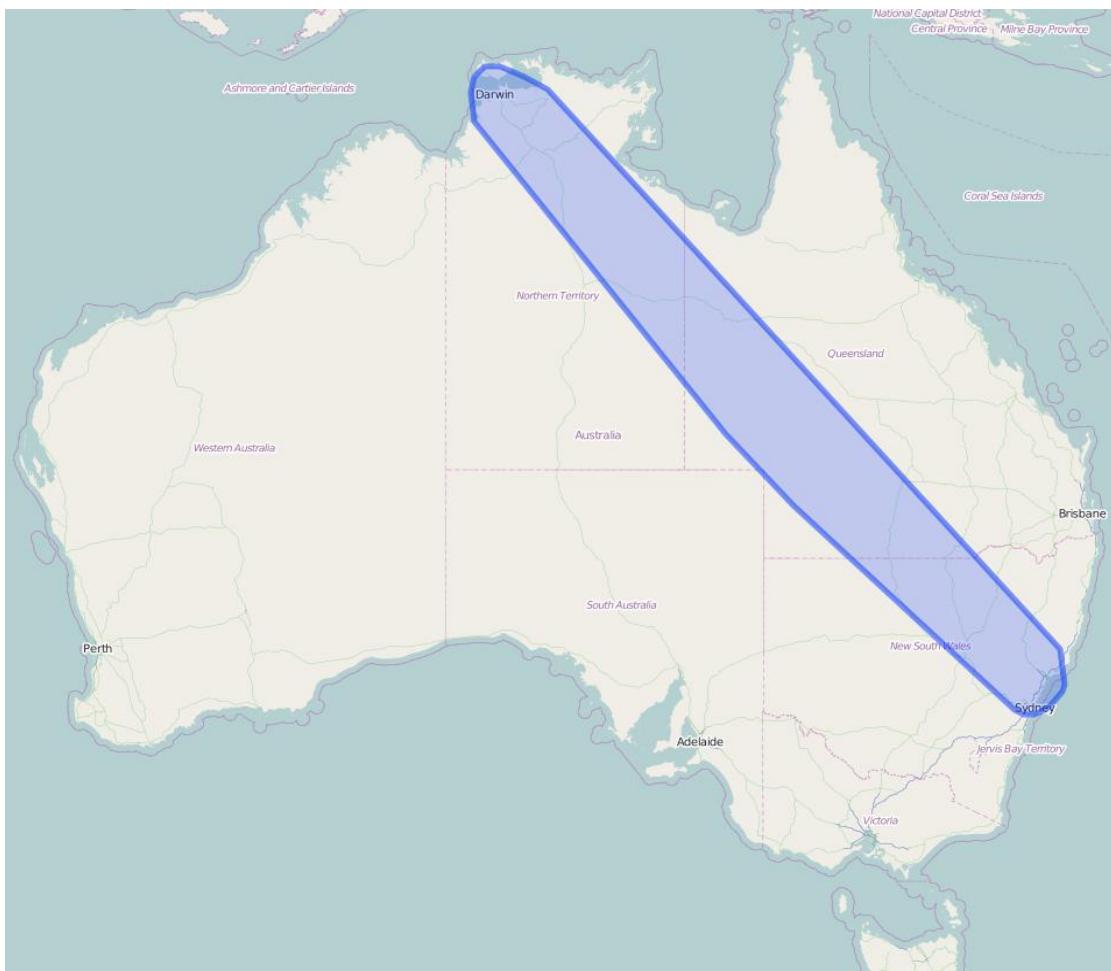
Community composition



The “Great Australian Bugger-all ... nothing but
relentlessly identical miles ...”

— Stephen Fry





Retrieve data

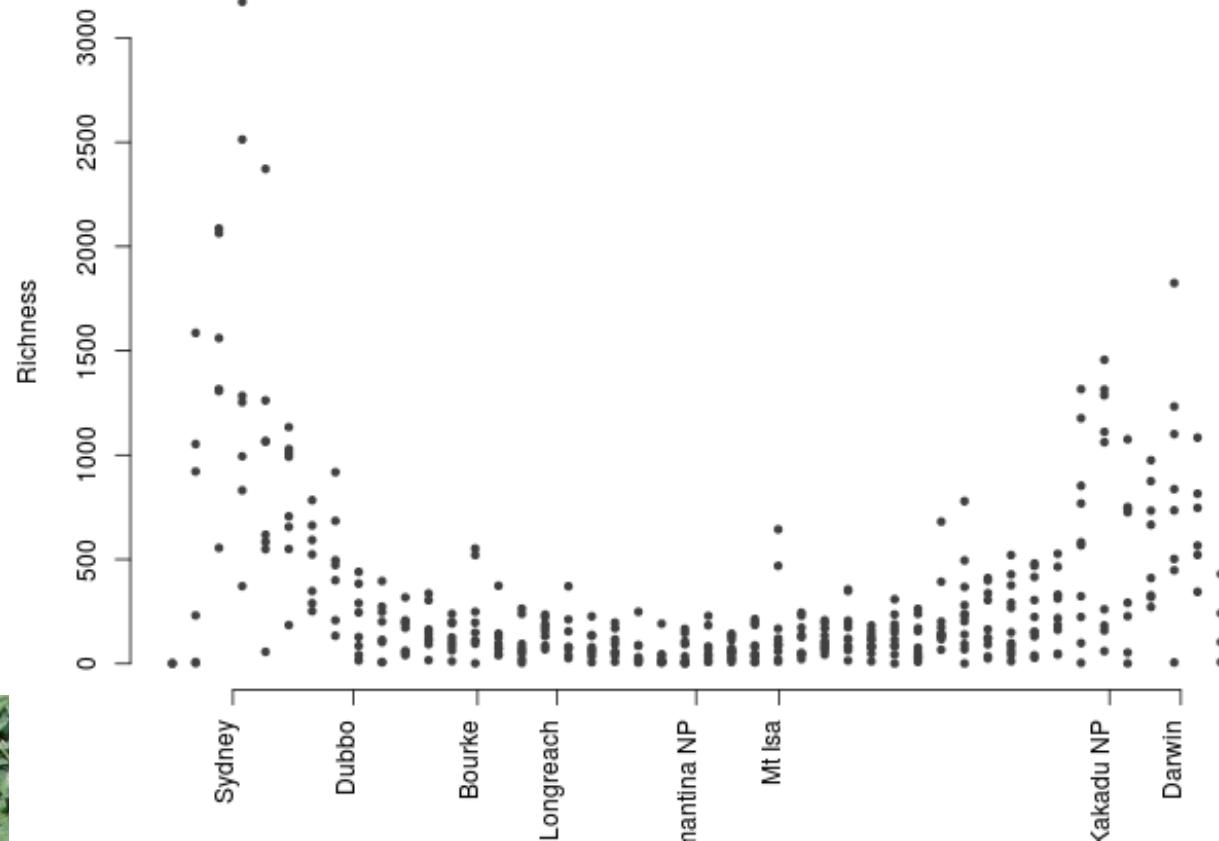
```
wkt="POLYGON( (152.2 -32.0,...) )"  
x=occurrences(taxon="kingdom:Plantae",  
wkt=wkt, qa="none", download_reason_id=6)
```

- metadata (sources)
- data (table of occurrences)
(Name, rank, taxonomic classification, location,
date, etc)



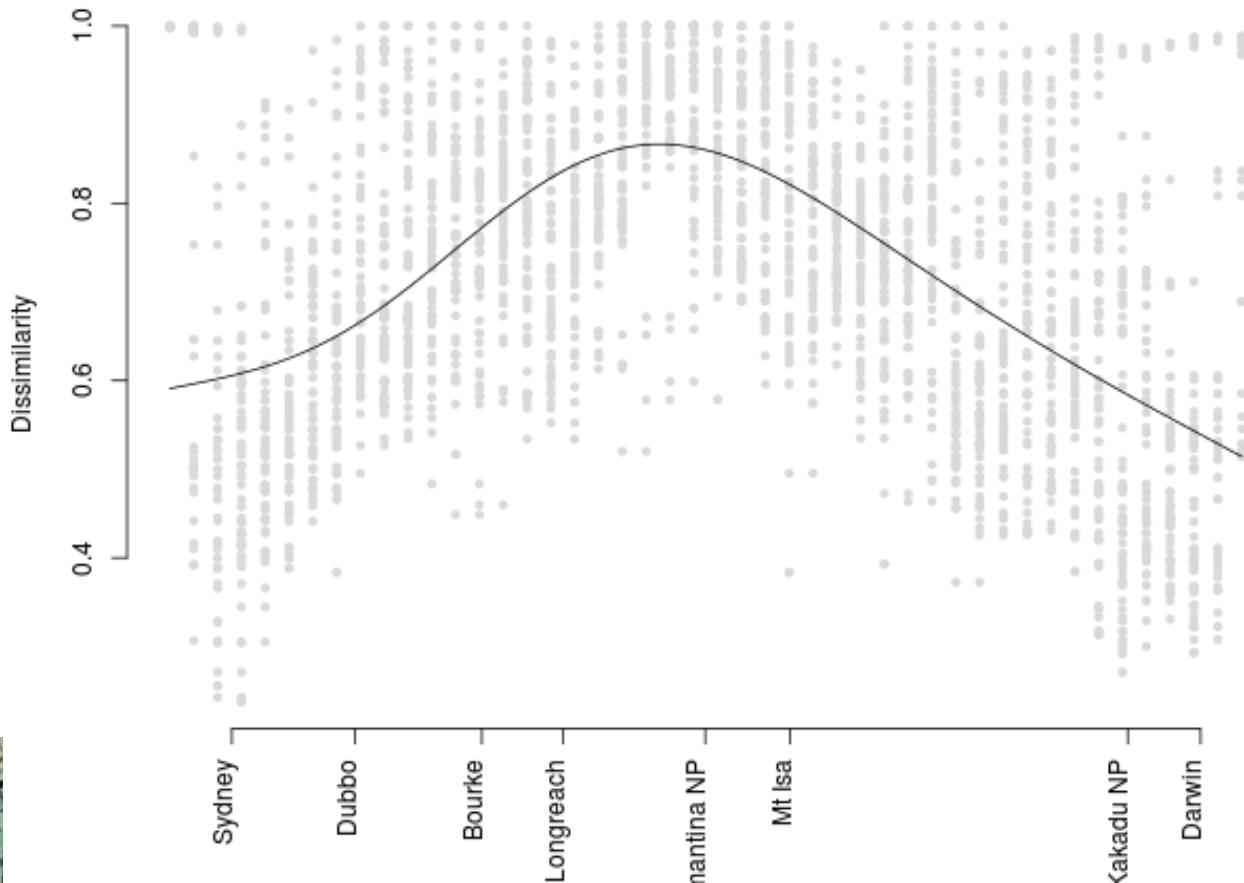
Analyses — richness

```
library(reshape)  
x=reshape(x,...) #convert to sites-by-species  
r=apply(x,1,sum) #species richness by site  
plot(...)
```



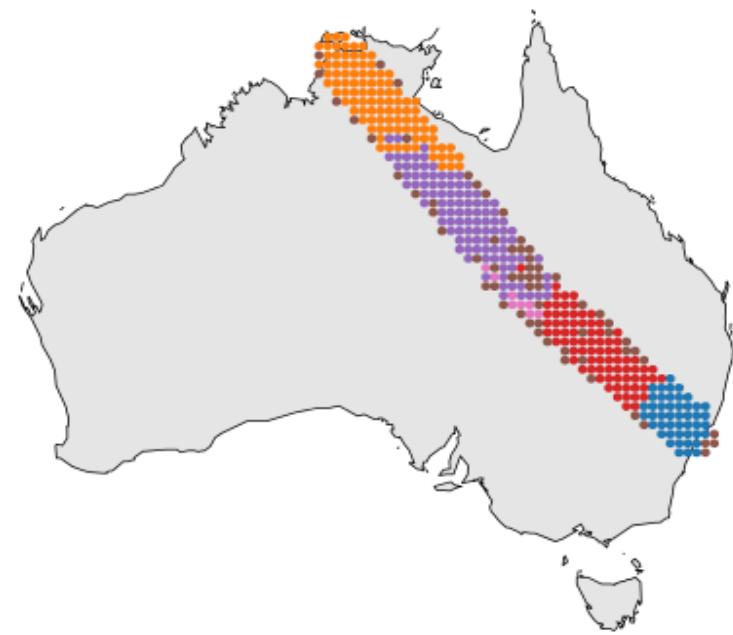
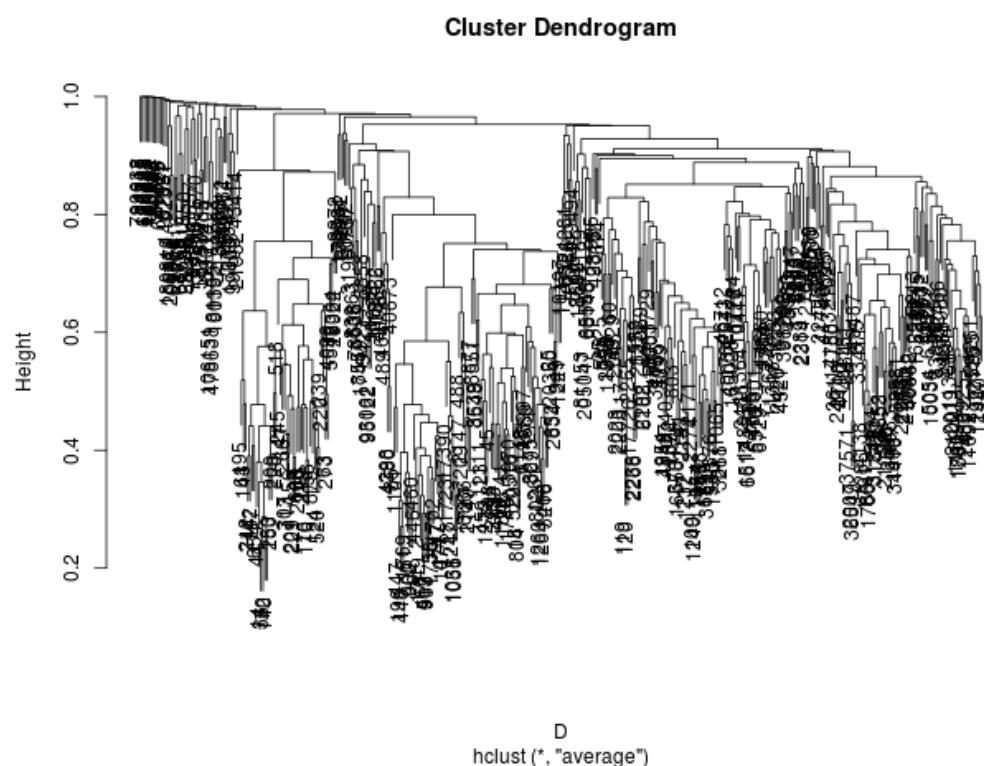
Analyses — turnover

```
library(vegan); library(geosphere)  
D=vegdist(x, "bray") #dissimilarities  
Dg=distVincentySphere(...); #geographic distances  
plot(...)
```



Analyses — clustering

```
cl=hclust(D,method="average") #UPGMA  
grp=cutree(cl,...)  
plot(...)
```



<http://github.com/jvanderwal/ALA4R>

The GitHub repository page for [jjvanderwal / ALA4R](https://github.com/jjvanderwal/ALA4R) is displayed. The repository has 348 commits, 1 branch, 0 releases, and 3 contributors. The master branch is selected. The repository description states: "The ALA4R package provides tools to enable the R programming language to access data and resources hosted at Atlas of Living Australia (ALA). It also provides functions to enable computation of distributions of species (to be queried and output in standard formats)." The repository sidebar includes links for Issues, Pull Requests, Wiki, Pulse, Graphs, Network, and a HTTPS clone URL.

ALA4R / +

fixed bug in renaming "clazz" names

raymondben authored 15 hours ago latest commit [ad2cef1e](#)

File	Description	Age
R	fixed bug in renaming "clazz" names	15 hours ago
data	interim commit for new functions	16 days ago
man	fall back to read.table if fread throws warning	17 hours ago
vignettes	vignette re-added	6 days ago
.Rbuildignore	Added other ws page references for completeness	2 months ago
ALA4R.Rproj	Lees docs	3 months ago
DESCRIPTION	version bump	15 hours ago
NAMESPACE	changed @export for s3 methods to @S3method	5 days ago
README.md	updated for sp imported rather than suggested	15 days ago

README.md

ALA4R

The Atlas of Living Australia (ALA) provides tools to enable users of biodiversity information to find, access, combine and visualise data on Australian plants and animals; these have been made available from <http://www.ala.org.au/>. Here we provide a subset of the tools to be directly used within R.

Currently in a very preliminary state: everything is liable to change.

Installing

Windows

Install the dependencies first:

```
install.packages(c("httr", "stringr", "plyr", "digest", "RCurl", "jsonlite", "assertthat", "sp"))
```

If you wish to use the `data.table` package for potentially faster loading of data matrices (optional), also do:

www.ala.org.au

ATLAS OF LIVING
AUSTRALIA
sharing biodiversity knowledge

