Desktop analysis with Atlas data: ALA4R

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	Eudyptula minor	species	0.68	Blue Penguin, Fairy Penguin*
3	Eudyptes chrysocome	species	3.0e-9	Crested Penguin
4	Eudyptes pachyrhynchus robustus	subspecies	2.6e-9	Snares Crested Penguin
5	Pteria penguin	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 Cannae Point Generally Northward To The Point Near The Intersection Of Stuart Street And Oyama Cove Avenue, And Extending 100 Metres Offshore From That Shoreline)”

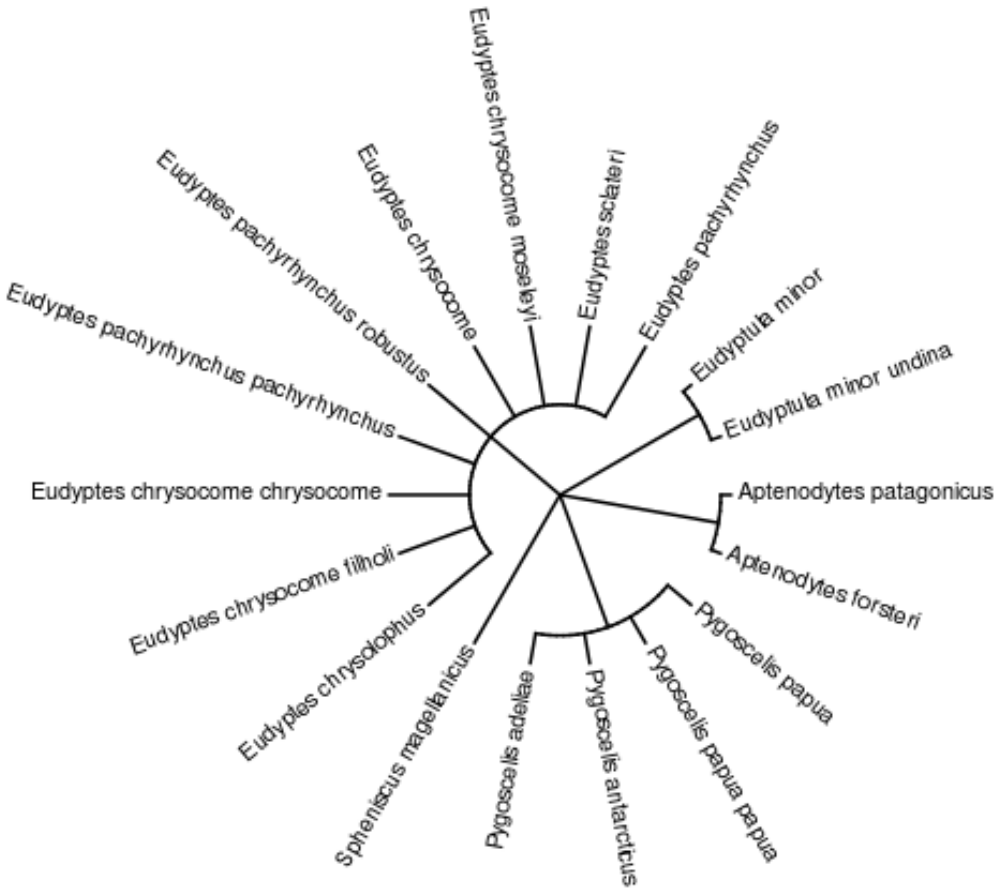
```
tx=species_download("family:SPHENISCIDAE",
  fields=c("guid","parentGuid","genus",
  "nameComplete","rank"))
tx=tx[tx$Taxon.Rank %in%
  c("species","subspecies"),]
```

→ 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

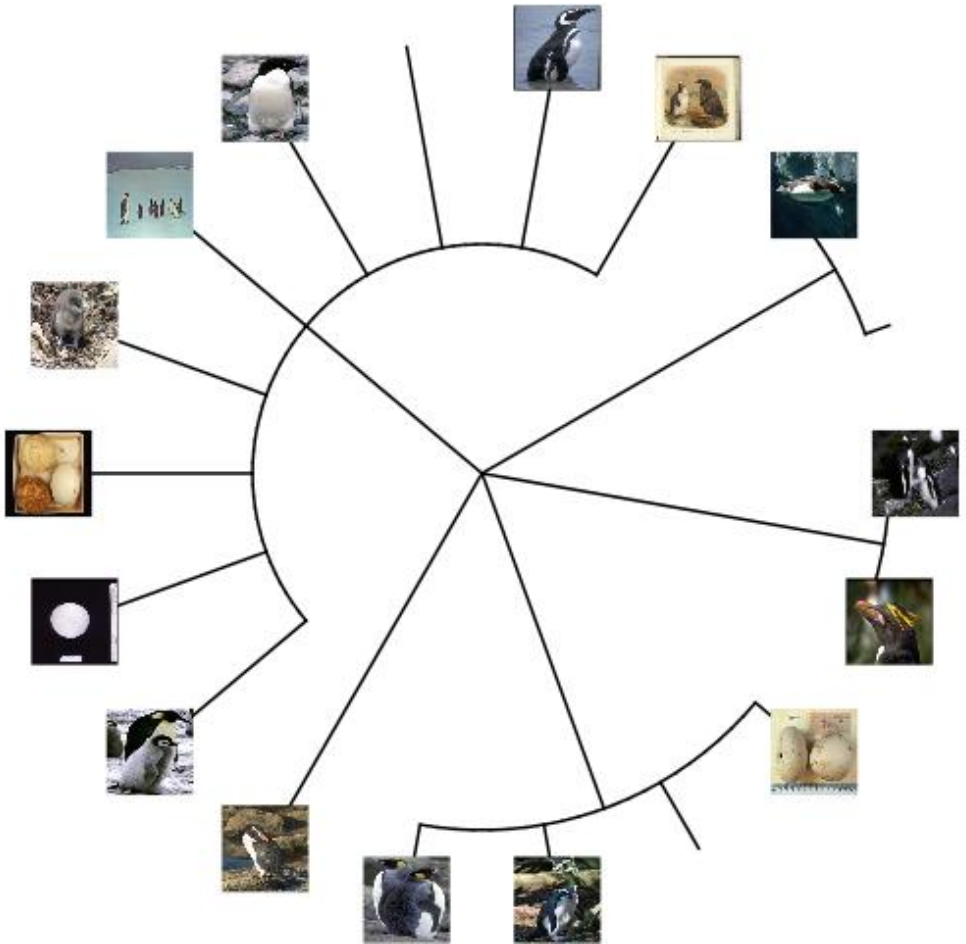
Plot taxonomic tree

```
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)
```

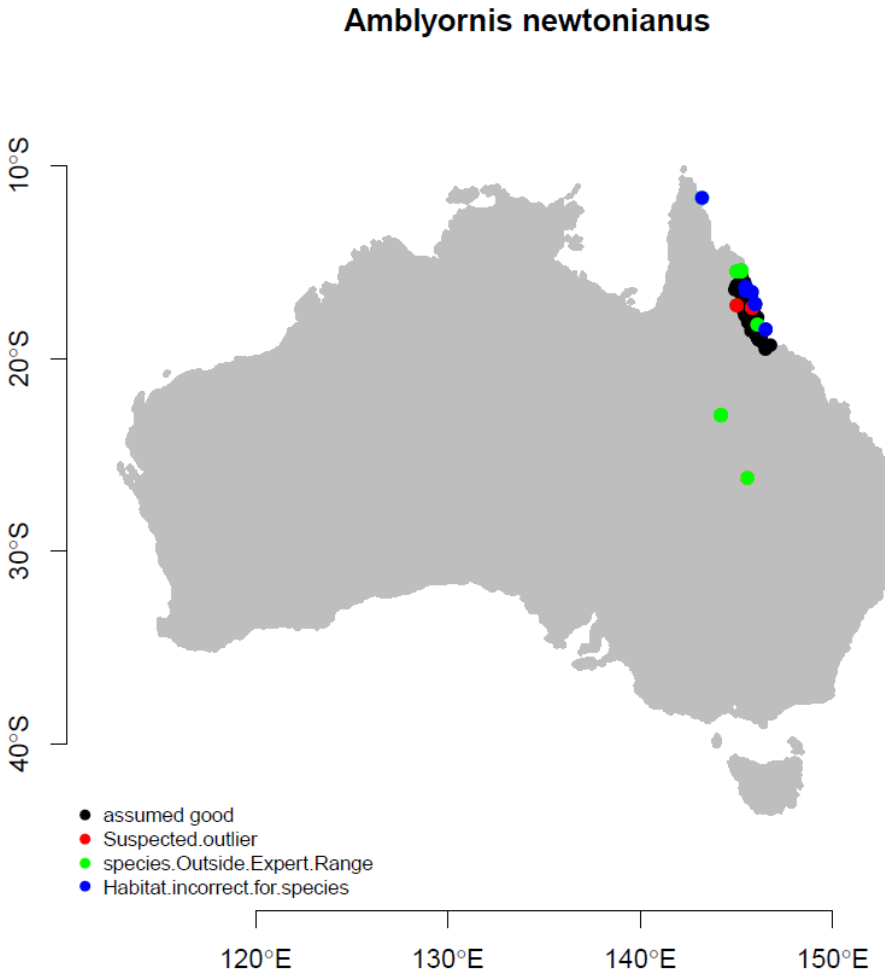
number of names: **7**
number of taxonomically corrected names: **1**
number of observation records: 881
assertions checked: 22
    Invalid.collection.date: 119 records
    incomplete.Collection.Date: 159 records
    First.of.the.century: 4 records
...

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

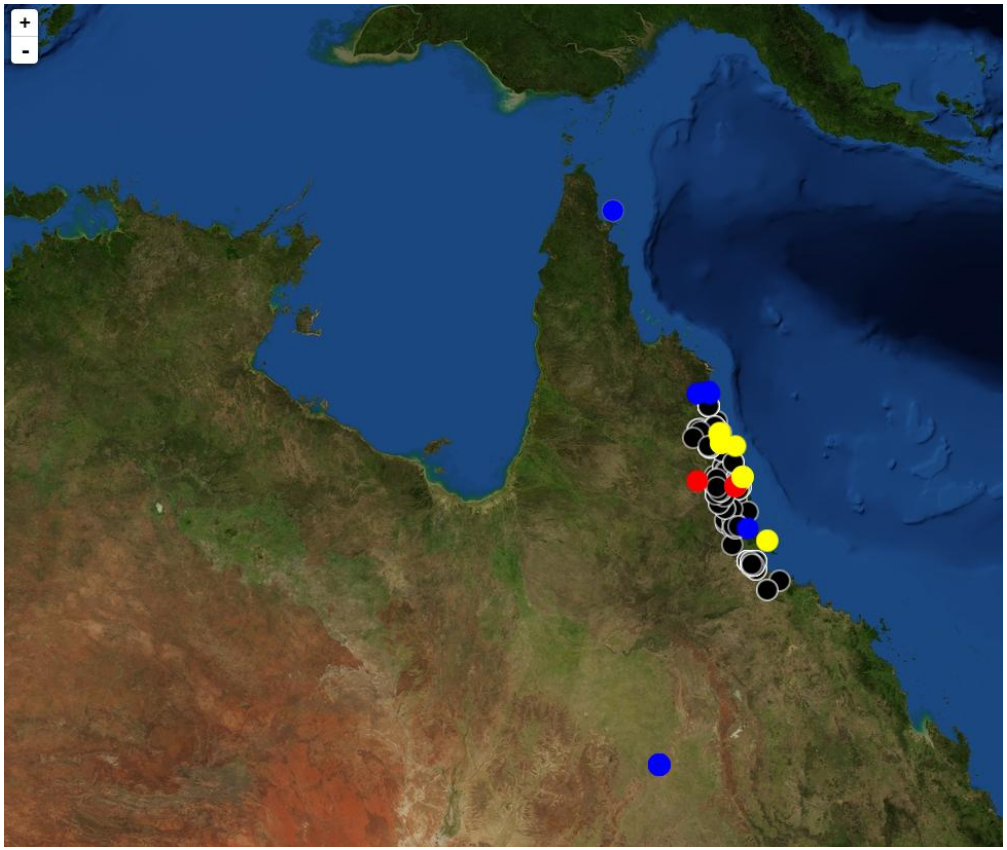
Suspected.outlier: 13 records -- considered fatal
Coordinate.accuracy.not.valid: 12 records
First.of.the.year: 26 records
Altitude.value.supplied.in.feet: 2 records
geodetic.Datum.Assumed.Wgs.84: 619 records
species.Outside.Expert.Range: 14 records -- considered fatal
decimal.Lat.Long.Converted: 5 records
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')
```



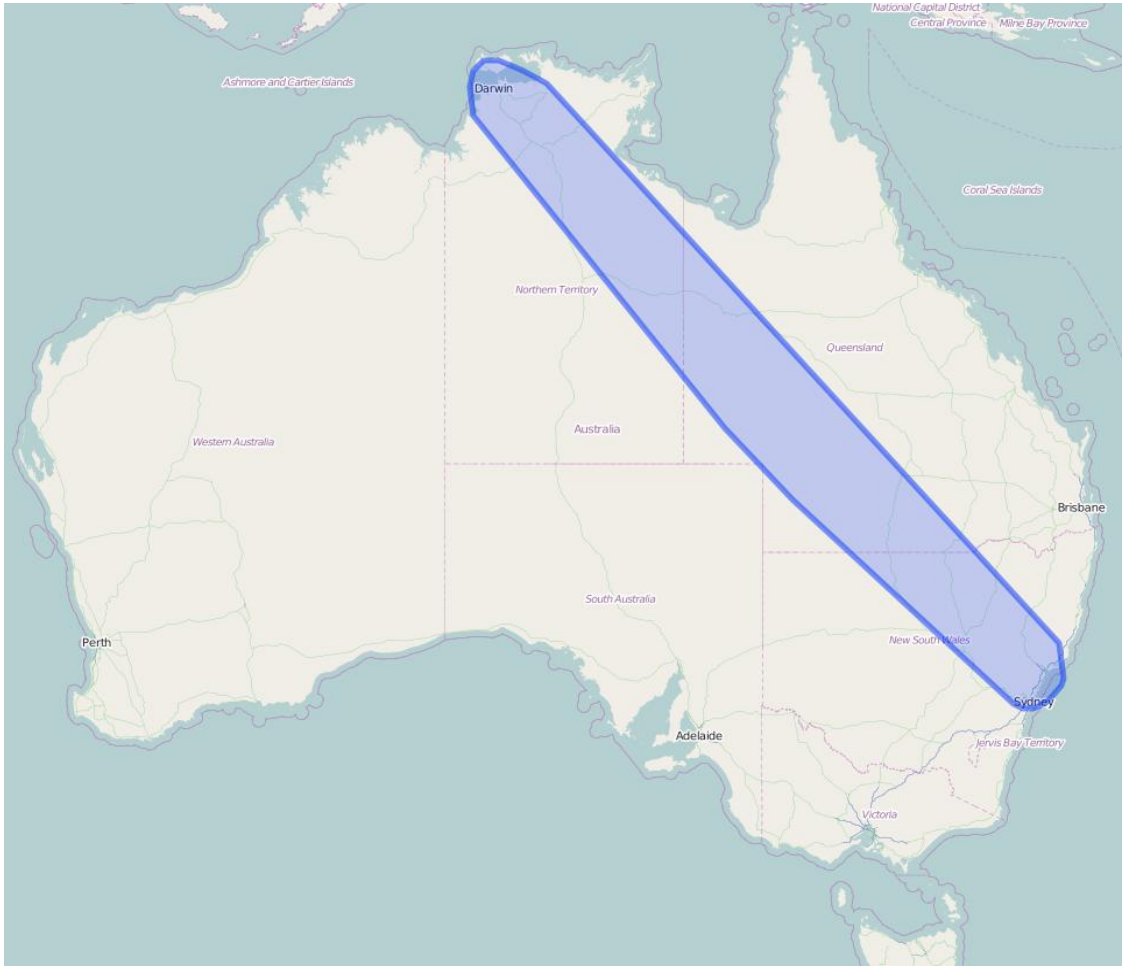
```
library(leafletR)
dat=toGeoJSON(data=x,…)
sty=styleCat(…)
alamap=leaflet(data=dat,style=sty,…)
```



Community composition

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

— Stephen Fry


National Capital District
Central Province
Milne Bay Province
Ashmore and Cartier Islands
Darwin
Coral Sea Islands
Northern Territory
Queensland
Australia
Western Australia
Brisbane
South Australia
Perth
New South Wales
Sydney
Adelaide
Jervis Bay Territory
Victoria

Retrieve data

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

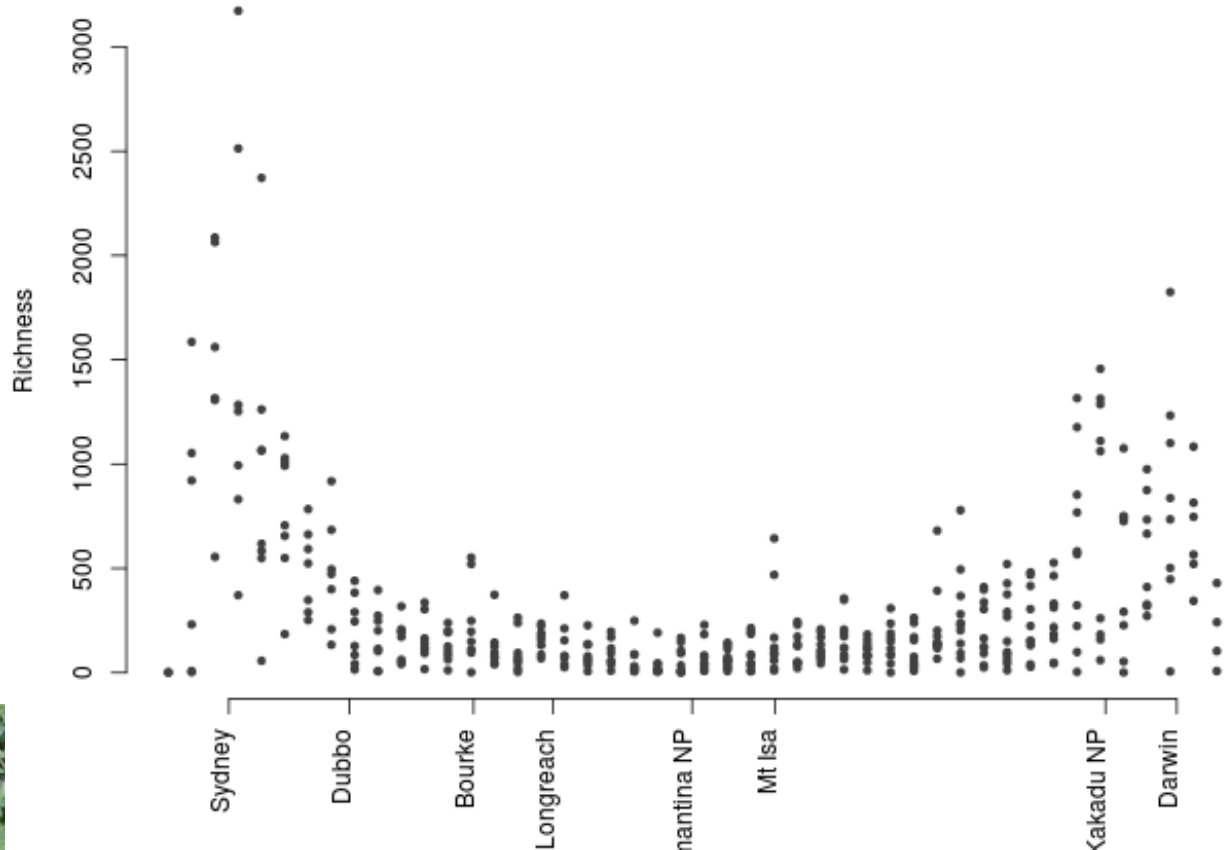
- metadata (sources)

- data (table of occurrences)

(Name, rank, taxonomic classification, location, date, etc)

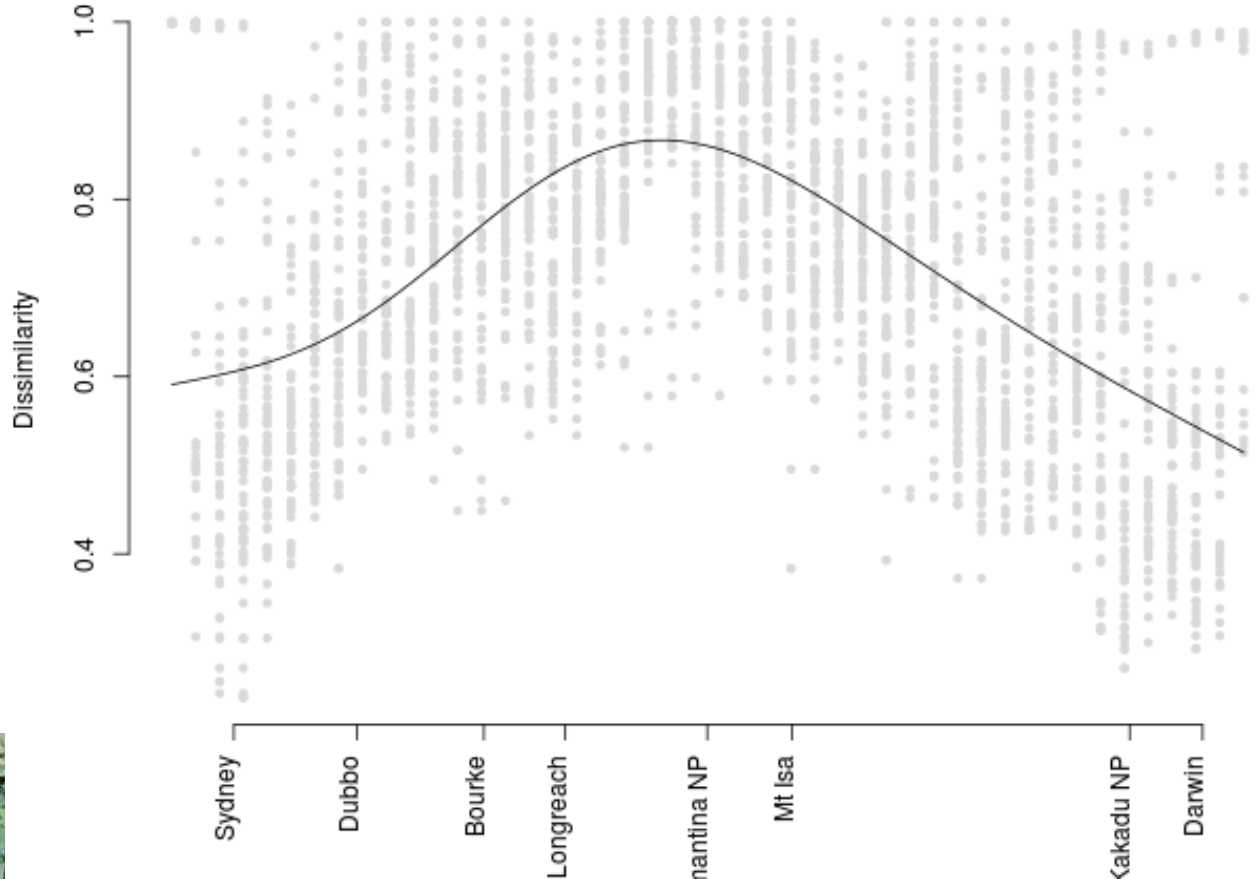
Analyses — richness

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



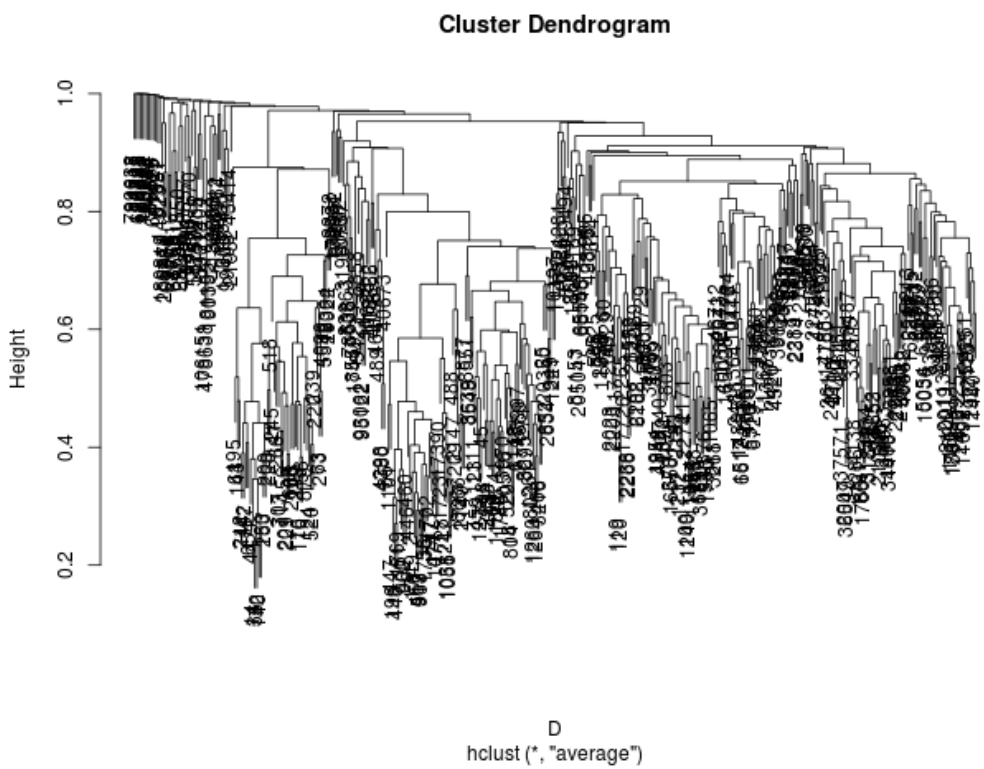
Analyses — turnover

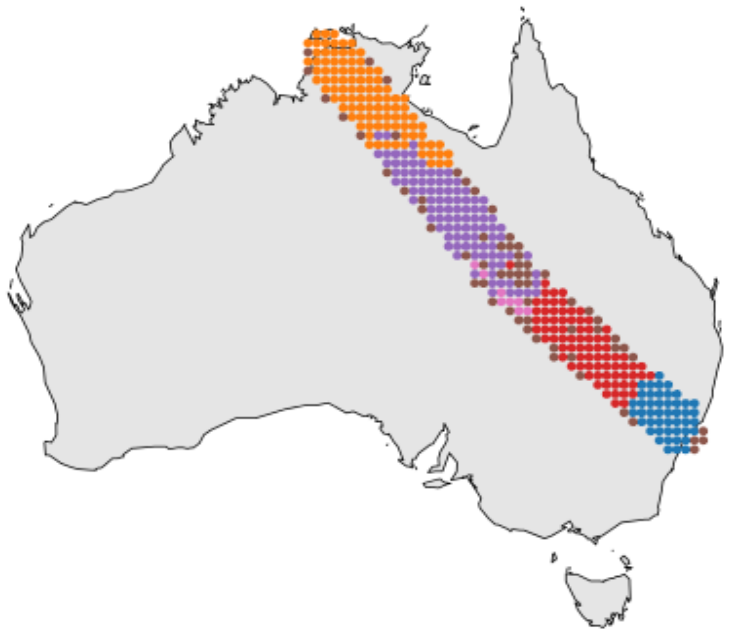
```
library(vegan); library(geosphere)
D=vegdist(x,"bray") #dissimilarities
Dg=distVincentySphere(…) #geographic distances
plot(…)
```



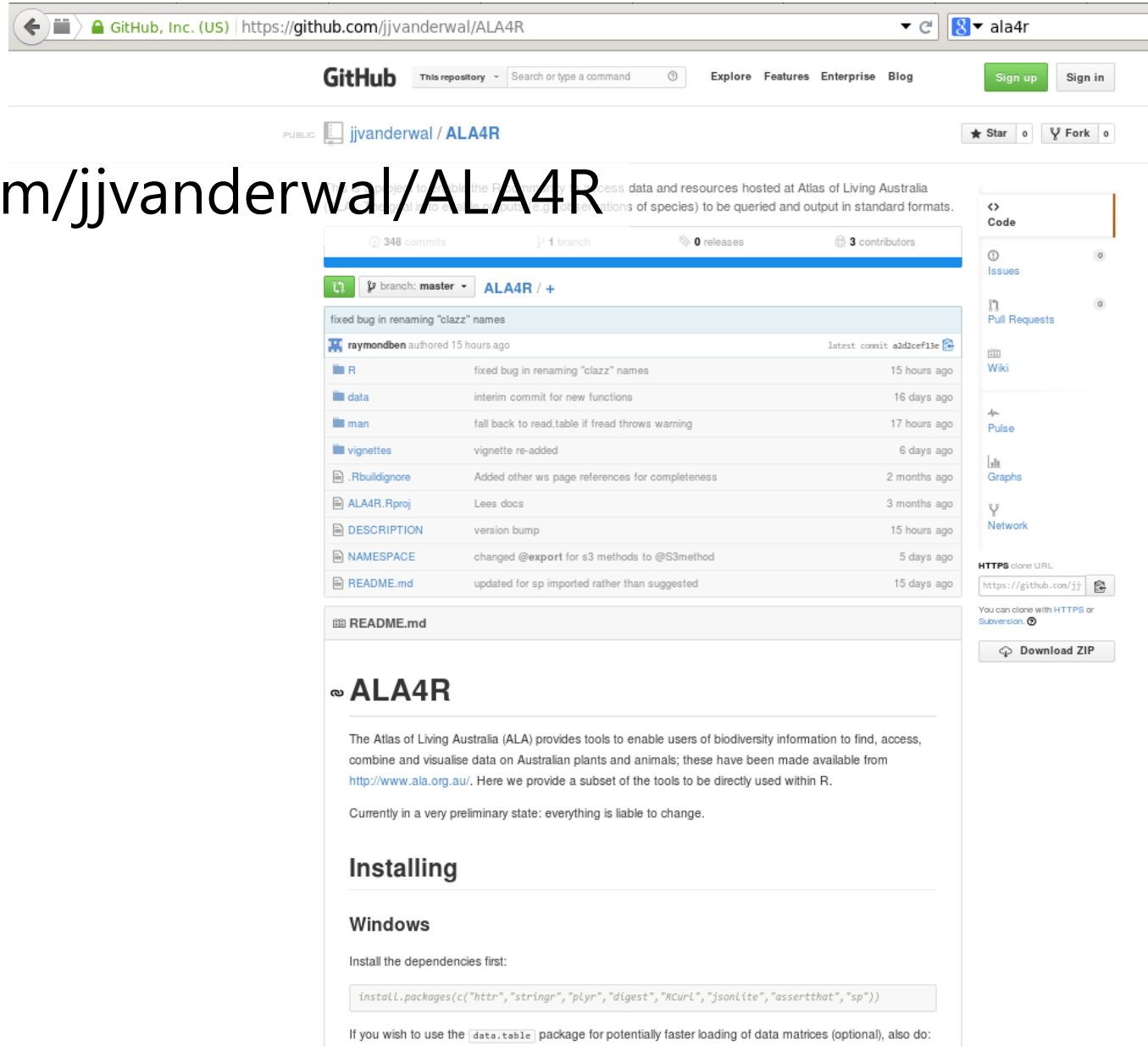
Analyses — clustering

```
cl=hclust(D,method=”average”) #UPGMA
grp=cutree(cl,…)
plot(…)
```





http://github.com/jjvanderwal/ALA4R



www.ala.org.au

ATLAS OF LIVING AUSTRALIA
sharing biodiversity knowledge