rgbif vignette - Seach and retrieve data from the Global Biodiverity Information Facilty (GBIF)

About the package

`rgbif` is an R package to search and retrieve data from the Global Biodiverity Information Facilty (GBIF). `rgbif` wraps R code around the GBIF API to allow you to talk to GBIF from R.

Install rgbif and dependencies

```
install.packages("devtools")
library(devtools)
install_github("rgbif", "ropensci", ref = "newapi")
```

Load rgbif and dependencies

```
library(rgbif)
library(XML)
library(RCurl)
library(plyr)
library(ggplot2)
library(maps)
```

Get number of occurrences for a set of search parameters

Search by type of record, all observational in this case

```
occ_count(basisOfRecord = "OBSERVATION")
```

```
[1] 286071783
```

Records for Puma concolor with lat/long data (georeferened) only

```
occ_count(nubKey = 2435099, georeferenced = TRUE)
```

```
[1] 2541
```

All georeferenced records in GBIF

```
occ_count(georeferenced = TRUE)
```

```
[1] 3.55e+08
```

Records from Denmark

```r
occ_count(country = "DENMARK")
```

```
[1] 8628822
```

Records from France

```r
occ_count(hostCountry = "FRANCE")
```

```
[1] 17272175
```

Number of records in a particular dataset

```r
occ_count(datasetKey = "9e7ea106-0bf8-4087-bb61-dfe4f29e0f17")
```

```
[1] 4591
```

All records from 2012

```r
occ_count(year = 2012)
```

```
[1] 31483292
```

Records for a particular dataset, and only for preserved specimens

```r
occ_count(datasetKey = "8626bd3a-f762-11e1-a439-00145eb45e9a", basisOfRecord = "PRESERVED_SPECIMEN")
```

```
[1] 550849
```

Get possible values to be used in taxonomic rank arguments in functions

```r
taxrank()
```

```
[1] "kingdom"      "phylum"       "class"        "order"
[5] "family"       "genus"        "species"      "infraspecific"
```

Search for taxon information

Search for a genus

```r
head(name_lookup(query = "Cnaemidophorus", rank = "genus", return = "data"))
```

```
        key  nubKey parentKey          parent  kingdom     phylum   clazz
1 116755723 1858636 110614854 Pterophoridae Animalia Arthropoda Insecta
2   1858636 1858636      8863 Pterophoridae Animalia Arthropoda Insecta
3 125802004 1858636 125793784 Pterophoridae     <NA>       <NA> Insecta
4 124531302 1858636        NA          <NA>     <NA>       <NA>    <NA>
5 126862804 1858636 126783981 Pterophoridae Animalia Arthropoda Insecta
6 115123697 1858636        NA          <NA>     <NA>       <NA>    <NA>
        order         family           genus kingdomKey phylumKey  classKey
1 Lepidoptera Pterophoridae Cnaemidophorus  116630539 116762374 116686069
2 Lepidoptera Pterophoridae Cnaemidophorus          1        54       216
3 Lepidoptera Pterophoridae Cnaemidophorus         NA        NA 125831175
4        <NA>          <NA> Cnaemidophorus         NA        NA        NA
5 Lepidoptera Pterophoridae Cnaemidophorus  126774927 126774928 126775138
6        <NA>          <NA> Cnaemidophorus         NA        NA        NA
   orderKey familyKey   genusKey  canonicalName        authorship   nameType
1 116843281 110614854 116755723 Cnaemidophorus Wallengren, 1862 WELLFORMED
2       797      8863   1858636 Cnaemidophorus Wallengren, 1862 WELLFORMED
3 125810165 125793784 125802004 Cnaemidophorus Wallengren, 1862 WELLFORMED
4        NA        NA 124531302 Cnaemidophorus                   WELLFORMED
5 126775421 126783981 126862804 Cnaemidophorus                   WELLFORMED
6        NA        NA 115123697 Cnaemidophorus                   WELLFORMED
   rank numOccurrences
1 GENUS              0
2 GENUS              0
3 GENUS              0
4 GENUS              0
5 GENUS              0
6 GENUS              0
```

Search for the class mammalia

```r
name_lookup(class = "mammalia")
```

```
Error: unused argument (class = "mammalia")
```

Look up the species Helianthus annuus

```r
name_lookup("Helianthus annuus", rank = "species")
```

```
$meta
  offset limit endOfRecords count
1      0    20        FALSE    69

$data
        key  nubKey parentKey                      parent      kingdom
1 116845199 3119195 116853573                 Helianthus       Plantae
```

```
2    3119195 3119195   3119134                Helianthus        Plantae
3  125790787 3119195 125809269                Helianthus          <NA>
4  106239436 3119195 106239325                Helianthus Viridiplantae
5  121635316 3119195 124573711                Helianthus          <NA>
6  111449704 3119195 111449703                Helianthus        Plantae
7  108157198      NA 108086589                Asteraceae        Plantae
8  108157199      NA 108086589                Asteraceae        Plantae
9  108157200      NA 108086589                Asteraceae        Plantae
10 115043868 3119195 115091988                Helianthus        Plantae
11 107290518      NA 107290513                Helianthus        Plantae
12 107001935      NA 107105089                Helianthus        Plantae
13 117214133 3119195 117208777                    Virus          Virus
14 117075019 3119195 117061550                Helianthus        Plantae
15 110853779 3119195 116128567                Helianthus        Plantae
16 125879180 3119195 126824197                Helianthus        Plantae
17 124780276 3119195 124852643                Helianthus        Plantae
18 100837541 3119195 102425010                Helianthus        Plantae
19 100019171 3119195 100009008                Helianthus          <NA>
20 125587214 3119195 106573315 unclassified phytoplasmas          <NA>
               order            family   genus kingdomKey  orderKey
1          Asterales        Asteraceae Helianthus  116668764 110610447
2          Asterales        Asteraceae Helianthus          6       414
3          Asterales        Asteraceae Helianthus         NA 125833882
4          Asterales        Asteraceae Helianthus  106147210 106237428
5              <NA>              <NA> Helianthus         NA        NA
6              <NA>        Compositae Helianthus  111449174        NA
7              <NA>        Asteraceae       <NA>  115219148        NA
8              <NA>        Asteraceae       <NA>  115219148        NA
9              <NA>        Asteraceae       <NA>  115219148        NA
10         Asterales        Asteraceae Helianthus  115107585 115106969
11         Asterales        Asteraceae Helianthus  107264512 107289189
12             <NA>        Asteraceae Helianthus  124856107        NA
13             <NA>              <NA>       <NA>  117208777        NA
14         Asterales        Compositae Helianthus  117067772 117058422
15         Asterales        Compositae Helianthus  116127234 116128510
16         Asterales        Asteraceae Helianthus  126775066 126779356
17         Asterales        Compositae Helianthus  124850847 124852488
18         Asterales        Asteraceae Helianthus  102545045 100614495
19         Asterales        Asteraceae Helianthus         NA 100023079
20 Acholeplasmatales Acholeplasmataceae Candidatus         NA 106013100
   familyKey genusKey    canonicalName authorship   nameType    rank
1  116856030 116853573 Helianthus annuus        L. WELLFORMED SPECIES
2       3065   3119134 Helianthus annuus        L. WELLFORMED SPECIES
3  125799038 125809269 Helianthus annuus        L. WELLFORMED SPECIES
4  106237535 106239325 Helianthus annuus           WELLFORMED SPECIES
5         NA 124573711 Helianthus annuus           WELLFORMED SPECIES
6  111442813 111449703 Helianthus annuus        L. WELLFORMED SPECIES
7  108086589        NA Helianthus annuus        L. WELLFORMED SPECIES
8  108086589        NA Helianthus annuus        L. WELLFORMED SPECIES
9  108086589        NA Helianthus annuus        L. WELLFORMED SPECIES
10 115105473 115091988 Helianthus annuus        L. WELLFORMED SPECIES
11 107289191 107290513 Helianthus annuus        L. WELLFORMED SPECIES
12 107079268 107105089 Helianthus annuus        L. WELLFORMED SPECIES
13        NA        NA Helianthus annuus        L. WELLFORMED SPECIES
```

```
14 117068545 117061550 Helianthus annuus        L. WELLFORMED SPECIES
15 116128511 116128567 Helianthus annuus        L. WELLFORMED SPECIES
16 126781795 126824197 Helianthus annuus        L. WELLFORMED SPECIES
17 124852489 124852643 Helianthus annuus        L. WELLFORMED SPECIES
18 102234418 102425010 Helianthus annuus   Linnaeus WELLFORMED SPECIES
19 100025154 100009008 Helianthus annuus   Linnaeus WELLFORMED SPECIES
20 106039581 106155719 Helianthus annuus            SCINAME SPECIES
   numOccurrences        phylum          clazz phylumKey  classKey
1              0          <NA>           <NA>        NA        NA
2              0 Magnoliophyta Magnoliopsida        49       220
3              0          <NA>           <NA>        NA        NA
4              0  Streptophyta           <NA> 106171079        NA
5              0          <NA>           <NA>        NA        NA
6              0 Spermatophyta Dicotyledones 111449175 111449177
7              0          <NA>           <NA>        NA        NA
8              0          <NA>           <NA>        NA        NA
9              0          <NA>           <NA>        NA        NA
10             0 Magnoliophyta Magnoliopsida 115107589 115107444
11             0 Magnoliophyta Magnoliopsida 107240291 107240313
12             0          <NA>           <NA>        NA        NA
13             0          <NA>           <NA>        NA        NA
14             0 Spermatophyta Magnoliopsida 117080124 117074792
15             0 Magnoliophyta Magnoliopsida 116127951 116128467
16             0  Tracheophyta Magnoliopsida 126775067 126775068
17             0  Tracheophyta Magnoliopsida 124851364 124852364
18             0 Magnoliophyta Magnoliopsida 102545123 101741810
19             0          <NA> Equisetopsida        NA 100023390
20             0   Tenericutes    Mollicutes 106355900 106136190

$facets
NULL
```

Get data for a single occurrence. Note that data is returned as a list, with slots for metadata and data, or as a hierarchy, or just data.

Just data

```r
occ_get(key = 773433533, return = "data")
```

```
              name       key longitude latitude
1 Helianthus annuus L. 773433533      -117    32.85
```

Just taxonomic hierarchy

```r
occ_get(key = 773433533, return = "hier")
```

```
            name    key     rank
1        Plantae      6 kingdom
2  Magnoliophyta     49  phylum
3  Magnoliopsida    220   clazz
```

```
4            Asterales     414    order
5           Asteraceae    3065   family
6           Helianthus 3119134    genus
7 Helianthus annuus L. 3119195  species
```

All data, or leave return parameter blank

```
occ_get(key = 773433533, return = "all")
```

```
$hierarch
                name     key    rank
1             Plantae       6 kingdom
2       Magnoliophyta      49  phylum
3       Magnoliopsida     220   clazz
4            Asterales     414   order
5           Asteraceae    3065  family
6           Helianthus 3119134   genus
7 Helianthus annuus L. 3119195 species
```

```
$data
                name          key longitude latitude
1 Helianthus annuus L. 773433533      -117    32.85
```

Get many occurrences. `occ_get` is vectorized

```
occ_get(key = c(773433533, 101010, 240713150, 855998194, 49819470), return = "data")
```

```
                              name        key longitude latitude
1             Helianthus annuus L. 773433533    -117.00    32.85
2  Platydoras costatus (Linnaeus, 1758)    101010     -70.07    -4.35
3                           Pelosina 240713150     163.58   -77.57
4      Sciurus vulgaris Linnaeus, 1758 855998194      12.04    58.41
5 Phlogophora meticulosa Linnaeus, 1758   49819470      13.28    55.72
```

Maps

Static map using the ggplot2 package

Make a map of **Puma concolor** occurrences

```
key <- name_backbone(name = "Puma concolor", kingdom = "plants")$speciesKey
dat <- occ_search(taxonKey = key, return = "data", limit = 300, minimal = FALSE)
gbifmap(input = dat)
```
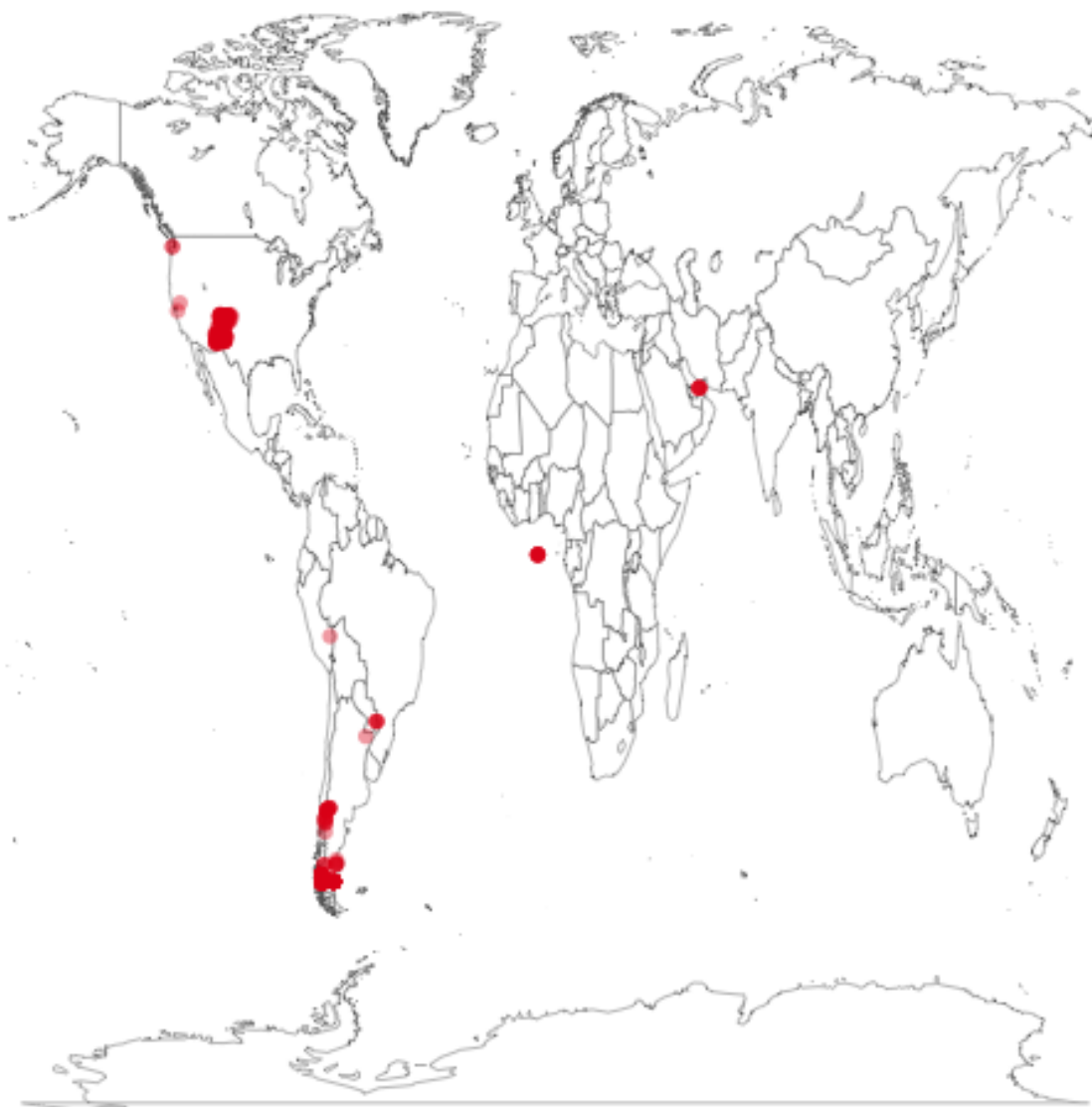
Figure 1: plot of chunk gbifmap1

Interactive map using a geoJSON file with your Github account

```r
library(plyr)
splist <- c("Accipiter erythronemius", "Junco hyemalis", "Aix sponsa")
keys <- sapply(splist, function(x) name_backbone(name = x, kingdom = "plants")$speciesKey,
    USE.NAMES = FALSE)
out <- occ_search(keys, georeferenced = TRUE, limit = 50, return = "data")
dat <- ldply(out)
datgeojson <- stylegeojson(input = dat, var = "name", color = c("#976AAE", "#6B944D",
    "#BD5945"), size = c("small", "medium", "large"))
write.csv(datgeojson, "~/my.csv")
togeojson(input = "~/my.csv", method = "web", outfilename = "my")
gist("~/my.geojson", description = "Occurrences of three bird species mapped")
```