Prepare UCNRS sites

> Author: Meixi Lin

This is a note on detailed steps for preparing UCNRS sites for comparisons with CALeDNA sites.

Download original data

The data are available from:

- Fauna: https://ucnrs.org/wp-content/uploads/images/2015/10/UC-NRS-Species-List.xlsx
  - Stored in: UC-NRS-Species-List_v0.csv
- Flora: https://ucnrs.org/reserves/flora/reserve_plant_list.xls
  - Stored in: reserve_plant_list_v0.csv
- Shapefile for UCNRS boundary: https://www.dropbox.com/sh/kwt1dvdsloe5fep/AACyPDPNtV7xuV5h3Hf0x66Wa/-%20All%20UCNRS?dl=0&preview=UC_NRS.gdb.zip&subfolder_nav_tracking=1
  - Use QGIS vector buffer tool to first create a 0.00833 degree (30s, 1km) buffer around all shapefile for UCNRS sites
  - The buffered shapefiles are stored in buffer_ucnrs_0.00833.shp

Define what CALeDNA sites were within UCNRS

A site has to be reported to be collected in UCNRS by volunteer and have matching coordinates within 1 km^2 buffer of UCNRS boundary 1_mk_phyloseq_ucnrs_sites.R

1. `loc` variable falls within UC Reserve names (127/278 sites)
2. coordinates falls within 1km^2 buffer (114/278 sites)

The final files are stored in sites_ucnrs_final.csv

The phyloseq object is stored in phydeco_uc.RData

Clean UCNRS data

1. clean up typos and spaces: 3.1_clean_ucnrs_records_format_typo.R
2. align genus names by checking the matches in the NCBI database used for building CRUX database:
   - The SQL database used: nodes.sqlite
   - For the genera that can't be found using NCBI database, manual inspections were performed with this workflow:
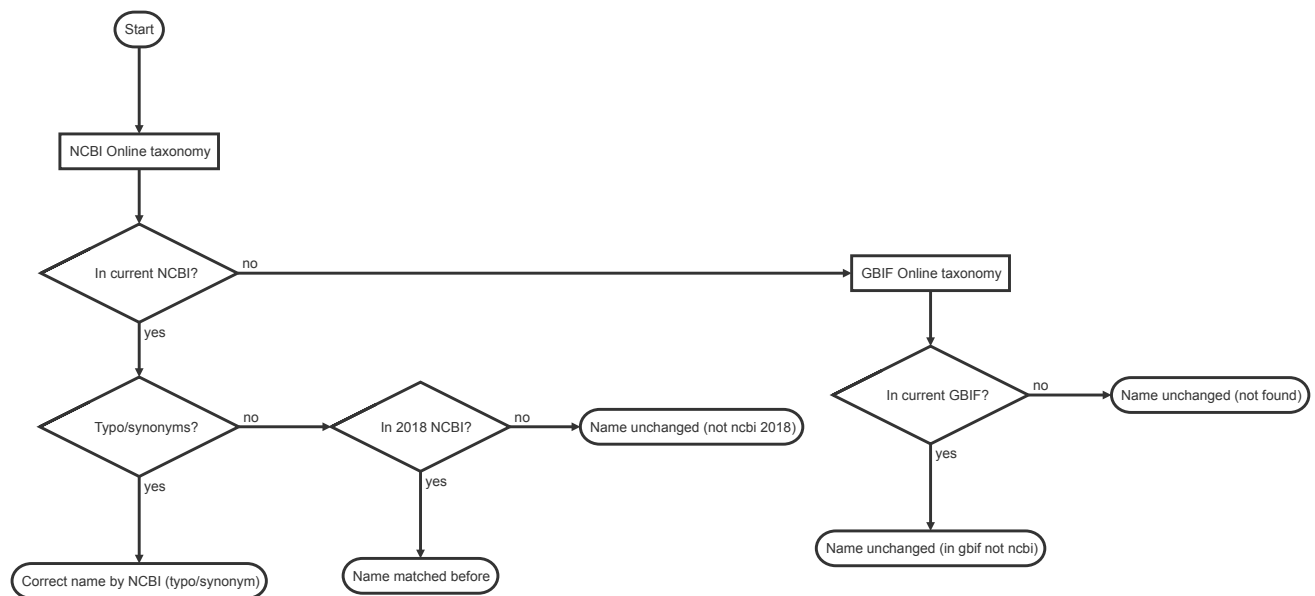
3. Species names were not corrected due to the large amount of ambiguity but if they are with the genus names, they are corrected along the way.These files after manual correction are stored in:
   1. fauna: fauna_ncbi_bind_v1.csv
   2. flora: flora_ncbi_bind_v1.csv
4. Family names were propagated using the corrected genus names with the 2018 taxonomy backbone. using 2.2_clean_ucnrs_taxlevel_genus.R and 2.3_clean_ucnrs_taxlevel_fillin.R
5. The final files are stored in the final_data folder