1. CN\_US species list compilation (stored at Users/lin/Desktop/CNUSA/Species\_list\_190715/Site\_list\_with\_code/)
2. For 6 US sites:
   1. Replace space to ‘\_’
   2. Check and delete any alien symbols like ‘?’
   3. For each site, get species list using bash ‘uniq’ command
   4. Add one column indicating site information in ${site}.csv
3. For 5 CN sites:
   1. Extract site species list from ‘CNUS15\_NW\_checked\_matrix.csv’ (R script see get\_5\_CN\_sites\_species\_list.R)
   2. Add one column indicating site information in ${site}.csv
4. Combine 11 sites to ‘11\_site\_with\_code.csv’
5. Get unique species list using ‘cut -f 1 -d ‘,’ 11\_site\_with\_code.csv |uniq > CN\_US\_species\_list.txt’
   1. With hyphen subset: CN\_US\_with\_hyphen\_species\_list.txt
   2. Without hyphen subset:CN\_US\_no\_hyphen\_species\_list.txt

Detailed notes and summary see /Users/lin/Desktop/CNUSA/Species\_list\_190715/note/clean\_sites\_note.xlsx

1. Building tree round 1

Screen log:

5495 species remained in the query for CN\_US\_no\_hyphen site

There are 99 duplicate ottids, 5036 ottids mapped to the CN\_US\_no\_hyphen\_ottid\_clean.tre

38 CN\_US\_no\_hyphen\_miss\_match\_back\_ott\_Jul162019.txt

201 CN\_US\_no\_hyphen\_query\_ottid.nomatch.txt

There are 4878 ottids mapped to the CN\_US\_no\_hyphen\_BR\_ottid\_clean\_Jul162019.tre

Gap:

5495-4878=617 species (names) off the BR tree.

To do (pwd:Users/lin/Desktop/CNUSA/Species\_list\_190715/deal617):

1. get the 617 name list

2) Modify names in species list (notes and changes are documented in Users/lin/Desktop/CNUSA/Species\_list\_190715/Site\_list\_with\_code/11\_site\_with\_code.csv)

Result:

1. place holder + 4878 = 5277 unique names to be queried.
2. Rename file should rename 382 names
3. Add\_in file contains 23 issues.
4. 185 names should be changed in site list.
5. Building tree round 2

Input: 5277 names

Results: 5268 get tips.

ott3899378 Arnoglossum\_muehlenbergii (off tree)

ott3899378 Arnoglossum\_reniforme

ott3955789 Panicum\_ovale (off tree)

ott3955789 Panicum\_villosissimum

ott3955841 Dichanthelium\_portoricense (off tree)

ott3955841 Dichanthelium\_webberianum

ott3955905 Dichanthelium\_aciculare (off tree)

ott3955905 Dichanthelium\_aciculare\_subsp.\_fusiforme

ott3955973 Dichanthelium\_sphaerocarpon (off tree)

ott3955973 Panicum\_sphaerocarpon

ott543886 Dichanthelium\_strigosum (off tree)

ott543886 Dichanthelium\_strigosum\_var.\_glabrescens

ott673012 Eubotrys\_recurva (off tree)

ott673012 Leucothoe\_recurva

ott76547 Brachyelytrum\_aristosum (off tree)

ott76547 Brachyelytrum\_septentrionale

ott791115 Quercus\_rubra (off tree)

ott791115 Quercus\_rubra\_var.\_borealis

Solution: using add in instead

Modify rename list: from 382 reduced to 377 names

Modify add in list: from 23 up to 31 issues.

common: 25

Singleton: 6

187 names to be changed: add Brachyelytrum\_aristosum (previous typo)

Leucothoe\_racemosa (previous Syn.)

After rename and addin:

After Rename : 5268 tips

After addin: 5299 tips

4. changing site species list and compile final list to test if identical to the tree tips.

5299 tips on the tree

5313of unique names

Need check 38:

Need add in Carex\_austro-caroliniana

Need add in Carex\_meritt-fernaldii

Need add in Carex\_novae-angliae

Need add in Celtis\_vandervoetiana

Need add in Dolichandra\_unguis-cati

Need add in Eupatorium\_mohrii

Need add in Hypericum\_crux-andreae

Need add in Lychnis\_flos-cuculi

Need add in Narcissus\_pseudo-narcissus

Need add in Pueraria\_montana\_var.\_lobata

Need add in Smilax\_bona-nox

Need add in Zanthoxylum\_clava-herculis

CN\_US\_5277\_speciesname\_BR\_addJul182019.txt:Acer\_linganense\_subsp.\_cappadocicum > Acer\_linganense\_subsp.\_koreanum

HARV Aster\_nemoralis > Oclemena\_nemoralis

HARV Aster\_novi-belgii > Symphyotrichum\_novi-belgii

CWHL Chimaphilia\_maculata > Chimaphila\_maculata

CWHL Luzula\_acuminata\_var.\_carolinae > Luzula\_acuminata\_subsp.\_carolinae

HARV Malva\_cf.\_moschata > Malva\_moschata

TALL Orbexilum\_psoraloides > Psoralea\_onobrychis

CWHL Panicum\_depauperatun > Panicum\_depauperatum

TALL Rhynchospora\_globulari Rhynchospora\_globularis

OSBS Rhynchospora\_grayii > Rhynchospora\_grayi

OSBS Scleria\_muelenbergii > Scleria\_muehlenbergii

CWHL Silphium\_cormpositum > Silphium\_compositum

CWHL Smilax\_rotiundifolia > Smilax\_rotundifolia

CWHL Solidago\_bicoior > Solidago\_bicolor

OSBS Solidago\_odora\_var.\_chapmanii > Solidago\_odora\_subsp.\_chapmanii

CWHL Robinia\_pseudo-acacia > Robinia\_pseudoacacia

To do

1. modify add in list
2. Minor modify

Results:

New tree tip = 5311

Identical name: 5306

Species list :5306

Replicated tips:

Ampelopsis\_delavayana (delete Ampelopsis\_aegirophylla in species list and Ampelopsis\_aegirophylla,Ampelopsis\_delavayana in rename)

Carex\_atlantica\_subsp.\_capillacea (delete Carex\_aboriginum in species list and Carex\_aboriginum,Carex\_atlantica\_subsp.\_capillacea in rename)

Eupatorium\_mohrii (delete Eupatorium\_tweedianum in species list, Eupatorium\_tweedianum,Eupatorium\_mohrii in rename, Eupatorium\_mohrii,Eupatorium\_album,Eupatorium\_aromaticum in add in)

Panicum\_dichotomum (delete Panicum\_zambesiense in species list, Panicum\_zambesiense,Panicum\_dichotomum in rename)

To do: find issues in rename

Modify original place holder list and rename list

Site = CN\_US

Results:

5306 tips matched between species list and tree!

1. generate matrics (PD\_07202019)
   1. 11 sites
   2. CN vs US
   3. Tree vs. Herbs.
   4. Disjunct vs. Non-disjunct