Dataset Issues

* Plant Associations and Plant Data have data for rows with no species name: a few papers (e.g., de Lajudie 1998, Sene et al 2012, Lammel et al 2013, others?); Banasiewics et al 2021 paper ID 302
  + De Lajudie (1998): KB will redo and extract data properly
  + Lammel at al (2013): Remove from data, the authors studied multiple plant species but did not specify which bacterial isolates were found with which plant species so cannot use for our analyses (KB)
  + Sene et al (2012): Remove from data, the authors studied multiple plant species but did not specify which bacterial isolates were found with which plant species so cannot use for our analyses (KB)
  + Sene et al (2013): the authors inoculated 40 different bacterial isolates and determined the success of those 40 bacterial isolates on 8 different plant species (5 native and 3 exotic) from a plantation. (SPL and KB decided this isn’t applicable to our work)
* paper 140 is about a moss
  + should be removed (KB)

Questions

1. Are non-native legumes utilizing the same rhizobial strains as native legumes within a given area? Within the same study (non-natives studied alongside natives)

Key Figure:

* Total number of strains associated with non-natives (anything that is non-native anywhere) compared to natives
* Percent of strains associated with non-native species that are also associated with native strains: Average percentage utilization across non-natives

What affects this relationship (growth form, lifespan, geographical range, etc.)?

1. Are non-native legumes utilizing the same rhizobial strains as they do in their native range?
   1. Home vs away (within the same study)

Key Figure:

* + - Total number of strains associated with non-natives (anything that is non-native anywhere) compared to natives
    - Percent of strains associated with one species in its non-native range that also associate with the sample species in its native range (using strains reported)

What affects this relationship (growth form, lifespan, geographical range, etc.)?

* 1. Across studies? Using genetic data

Key Figure:

* + - Percent of strains associated with one species in its non-native range that also associate with the sample species in its native range (using genetic data)

What affects this relationship (growth form, lifespan, geographical range, etc.)?

1. How many strains are associating with non-native species compared to native species? Informs if they are generalists or specialists

Key Figure(s):

1. Average number of strains per species comparing native species vs. non non-native species (non-native anywhere): richness

* + Correct for sample size (number of nodules, number of plants, number of studies?)

What affects this relationship (growth form, lifespan, geographical range, etc.)?

2. Average number of strains per species comparing native species vs. non non-native species (non-native anywhere): phylogenetic diversity

* + Correct for sample size (number of nodules, number of plants, number of studies?)

What affects this relationship (growth form, lifespan, geographical range, etc.)?