**Subject:**  **Discovering BIEN data through iPlant**

**From:**  Martha Narro <narro@iplantcollaborative.org>

**Date:**  May 22, 2014 4:13:24 PM MST

**To:**  Mark Schildhauer <schild@nceas.ucsb.edu>, Brian Enquist <benquist@email.arizona.edu>, Bob Peet <peet@unc.edu>, Brian McGill <mail@brianmcgill.org>, Brad Boyle <bboyle@email.arizona.edu>

**Cc:**  Ramona Walls <rwalls@iplantcollaborative.org>

Hi Mark, Brian E, Brad, Bob, Brian M,

Ramona Walls (copied) has been spearheading development of the plan for making data that are made public through iPlant's infrastructure easy to search and discover. A summary of the plan for the BIEN3 data is below. Please review it. The BIEN group will need to provide some information when the data are ready to be moved to iPlant. I'll create a Redmine issue, assigned to Enquist, for this so it doesn't get lost in our emails between now and then.

iPlant is in the process of becoming an issuer of DOIs so it can assign DOIs to datasets made public through its infrastructure. **For the BIEN db data, the plan is** to have a DOI for the (not publicly accessible but mirrored) BIEN3 database as well as for each publicly accessible analytical table (such as viewFullOccurrence). Because I estimate there will be only 5-10 analytical tables, cost of the BIEN DOIs will be covered by iPlant.

For the DOIs, **BIEN will need to provide** an abstract describing each analytical table. It may be useful to include the use cases (scientific questions) the data table is designed to address. In addition, we will need the metadata specified by DataCite (see the attached document). The mandatory metadata is not onerous.

**Question:** For the purposes of transparency, will general information about records that cannot be included in the public analytical tables be provided as part of the “abstract” describing each analytical table? This would inform users of the nature of records that cannot be made public and the reasons why. Some data providers do this on a per record basis, but I said this level of granularity is beyond what BIEN can provide – certainly for the first release.

iPlant will index the data in the analytical tables to make the data discoverable through iPlant’s search interfaces. This will enable iPlant users who are completely unaware of the BIEN data, but interested in, say, palms in Costa Rica, to discover the relevant data in BIEN analytical tables. Future plans under discussion include making these indexes discoverable through Google, etc.

Please let us know if you have any questions, concerns or suggestions.

Martha