Nicole A Nogoy, *GigaScience,* 2018. https://doi.org/10.6084/m9.figshare.7052075

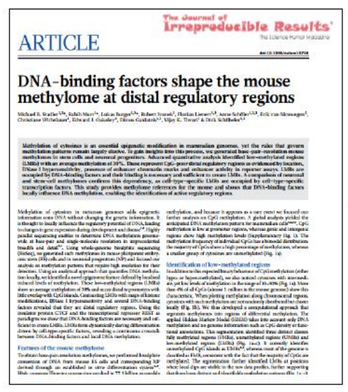
Improving Plant Data Sharing, Integration and Reproducibility

Nicole A Nogoy, Scott C Edmunds, Hans Zauner, Laurie Goodman

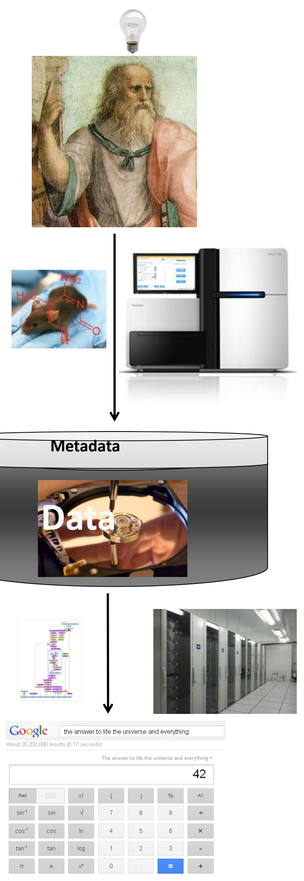
We live in an increasingly data-driven era, and the genomics community in particular has a long history of sharing data to drive this. For (mostly) technical reasons phenotyping has had a comparative lack of data sharing infrastructure and practice. For plant phenotyping, it is critical that original high-resolution scans are accessible to others. Heterogeneous datasets of multiple giga to terabytes scale are too large to include in traditional supplemental data, and repositories for high-resolution scans are often too specialized to take or will not accept such large datasets. And there has been a lack of incentives and credit to put in the effort to share these datasets.

GigaScience (http://gigasciencejournal.com), is an open access journal linked with an open repository, GigaDB (gigadb.org) and co-published by the BGI and Oxford University Press, The journal tries to address the issues of disseminating and handling large-scale data, providing resources and a forum for data producers and the open-science community. To encourage data sharing, reproducibility and re-use, research is linked with citable data DOIs directly to our associated database, GigaDB, which hosts and publishes all supporting information from our articles.

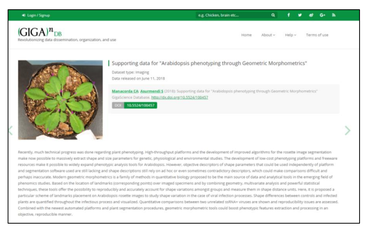
Anatomy of a traditional Publication



Briefly describe the aims, methods and draw conclusions

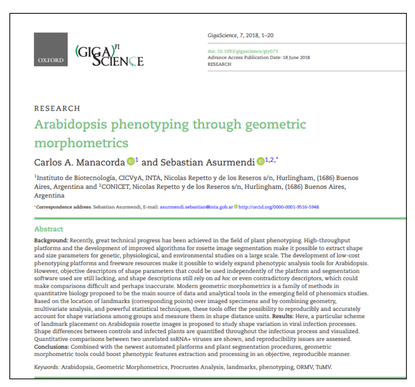


Anatomy of a GigaScience deconstructed Data Publication



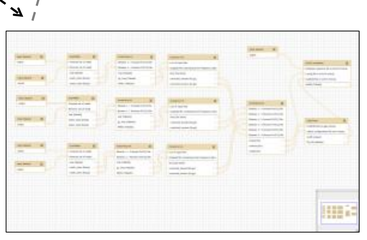
(GIGA)n DB

Share raw and intermediary data before analysis is complete in **GigaDB**



(GIGA)n SCIENCE open access

Fully describe research in ***GigaScience*** and link supporting data and workflows with a DOI



Disseminate your code and workflows via the **GigaGalaxy** bioinformatics platform

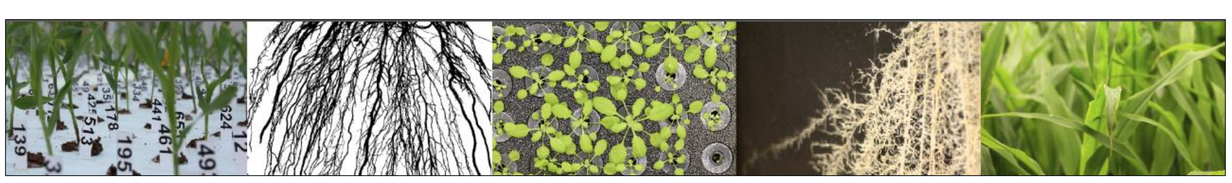
(GIGA)n Galaxy by CBIIT

open source

Plant Phenomics: Data Integration and Analyses Series

Guest Edited by Ruben Rellan Alvarez, Guillaume Lobet, Malia Gehan and Srikant Srinivasan

Our Plant Phenomics series features research and open datasets, such as “Morphometric analysis of Passiflora leaves” (http://dx.doi.org/10.5524/100251); “Hyperspectral time-series imaging of maize lines widely used in field trials” (http://dx.doi.org/10.5524/100371); “Large-scale phenomics analysis of a T-DNA tagged mutant population in rice *Oryza sativa L*.” (http://dx.doi.org/10.5524/100314). Our series also includes new methods for image-based plant phenotyping: “Deep machine learning provides state-of-the-art performance in image-based plant phenotyping” (https://doi.org/10.1093/gigascience/gix083), and analysis tools, “Phenobook: an open source software for phenotypic data collection” (https://doi.org/10.1093/gigascience/giw019).



Submit your next manuscript containing large-scale data and workflows to *GigaScience* and take full advantage of:

- No space constraints, and unlimited data and workflow hosting in *Giga*DB and *Giga*Galaxy
- Open access, open data, and highly visible work, freely available for distribution
- Collation of Data Citation Numbers by the the Data Citation Index
- Inclusion in PubMed Central, Medline, Scopus, Web of Science