**FAIR Data Management**

*…A Reminder from the MaizeGDB team*



MaizeGDB is a founding member of the AgBioData Consortium. AgBioData is a consortium of agricultural biological databases with the mission of consolidating standards and best practices for acquiring, displaying, and reusing genomic, genetic, and breeding (GGB) data.

Member databases of the AgBioData consortium follow the Findable, Accessible, Interoperable, and Reusable (FAIR) principles, enhancing the accessibility and reusability of large-scale agricultural research data. MaizeGDB consolidates vast amounts of published data, simplifying discovery and utilization for our community. Your commitment to FAIR principles ([go-fair.org](http://go-fair.org)) enables MaizeGDB and similar platforms to integrate and leverage even more data efficiently. Below are essential guidelines for applying FAIR data management practices to your generated data, as well as to the data you encounter in research papers and grant reviews.

* **Database Selection and Accession Reporting:** Ensure your data, such as DNA/RNA/Protein sequences, genome assemblies, and annotations, are submitted to long-term repositories like NCBI. Always include accession numbers in your publication. For maize SNPs, submit them to EVA at EBI. Explore more repositories at [maizegdb.org/FAIRpractices](http://maizegdb.org/FAIRpractices) and consult journal guidelines for additional instructions.
* **Data Publication:** Publish your data concurrently with your paper. For datasets not included with the article, secure a persistent identifier (e.g., DOI) from the data repository to reference in your paper. Datasets can be independently published in journals like microPublication, ensuring they are linked to the corresponding paper. Verify the presence and FAIR compliance of reported data during peer review.
* **Gene and Protein Identifier Usage:** Use established identifiers for genes, gene models, and genomes. Avoid renaming existing genes. Look up gene symbols at MaizeGDB and use precise gene model IDs. For protein data, reference the correct ID from NCBI or UniProt, submitting new sequences to these repositories as necessary. Maize nomenclature guidelines: <https://www.maizegdb.org/nomenclature>
* **Metadata and File Format Standards:** Attach comprehensive metadata to your datasets and adhere to accepted file formats. Treat metadata with the same rigor as experimental and analysis work. Incomplete or poorly described datasets compromise reusability, reproducibility, and overall research quality.
* **Machine-Readable Data Sets:** Ensure your data is machine-readable, using permanent identifiers and correct terminology (e.g., using correct case for genetic loci, incorporating GO, PO, PATO terms). Validate your data against common, established machine-readable formats.
* **Data Management Planning:** Allocate sufficient time for meticulous data management, similar to the effort dedicated to other research aspects.
* **FAIR Data Standards:** Here are some resources: [https://www.go-fair.org](https://www.go-fair.org/),<https://doi.org/10.1093/database/bay088>.

**We are always happy to answer your questions on these issues!** <https://www.maizegdb.org/contact>

**What's NEW at MaizeGDB!**

In 2024, MaizeGDB expanded its pan-genomic resources\* for the representative B73 genome, all NAM founder lines, Pan-Andropogoneae, and other historically important lines. These resources now include:

* 110 genomes, two are under the Toronto agreement, including the Pan-Andropogoneae genomes
* Over 1.5 million new gene model annotations
* New Pan-gene dataset, which includes 57 annotations
* Hundreds of new downloadable files
* 391 target databases in BLAST
* Syntenome and reelGene tracks on the B73v5 genome browser
* 400+ high-throughput sequencing data for over 80 tissues/conditions
* 300+ traits linked to over 40,000 positions in the genome
* 80+ million SNPs from EVA and Ensembl Plants
* Genotype data remapped to B73 v5 for 1,500 maize accessions
* Over 1 million predicted GO terms across 31 genomes
* Resources for 4 insertion mutation collections
* MaizeMine has been updated to include B73\_v5 and the NAM founder lines
* SNPversity 2.0, which enables users to explore and visualize extensive variant datasets with ease\*\*
* Transposable elements, structural variation, regulatory sites, and more…

**If you have questions on how to access/use these resources, contact us at** <https://www.maizegdb.org/contact>

\*Cannon, EK et al. (2024) Enhanced Pan-Genomic Resources at the Maize Genetics and Genomics Database. Genetics. doi:  [https://doi.org/10.1093/genetics/iyae036.](https://doi.org/10.1093/genetics/iyae036)

\*\*Andorf CM et al. (2024) A unified VCF data set from nearly 1,500 diverse maize accessions and resources to explore the genomic landscape of maize. G3. doi: <https://doi.org/10.1093/g3journal/jkae281>

**Thank you to the 2024 MaizeGDB Editorial Board Members!**

Amruta Bapat, Washington University St. Louis

Hank Bass, Florida State University

Diana Escamilla Sanchez, Iowa State University

Rohit Kumar, Clemson University

Songyu Liu, China Agricultural University, China

Ankita Mishra, Orissa University of Agriculture and Technology, India

Aimee Uyehara, Stanford University (2nd Year!)

****

**MaizeGDB has partnered with** ***microPublication Biology!***

***microPublication Biology* (Caltech Publishers) is a peer-reviewed, open-access journal that publishes single experiment results, which are incorporated directly into community knowledgebases like MaizeGDB**!

*microPublication Biology* gets your individual research findings, that might otherwise remain unpublished, out to the scientific community while providing credit to those who did the work. Articles are small (one figure, few pages), peer-reviewed, assigned a DOI and are discoverable on [PMC](https://www.ncbi.nlm.nih.gov/pmc/), [PubMed](https://pubmed.ncbi.nlm.nih.gov/), [EuropePMC](http://europepmc.org/), and [Google Scholar](https://scholar.google.com/).

How it works: Each maize *microPublication Biology* submission will be vetted by MaizeGDB curators at the time of peer review to ensure data meets FAIR data standards. Upon acceptance, your article is curated into MaizeGDB which couples the publication with curation and discoverability in MaizeGDB. The cost to publish is only $350.

**Here are some recent maize publications:**

Gustin JL, Zimmerman SA, and Sachs MM. (2025) Allelism of Uncharacterized Dwarf Mutants in Maize. <https://doi.org/10.17912/micropub.biology.001504>

Reneau, J; Ouslander, N; Sparks, EE. (2024) Quantification of maize brace root formation after vertical stalk displacement. microPublication Biology. <https://doi.org/10.17912/micropub.biology.001189>

Linders, KM; Santra, D; Schnable, JC; Sigmon, B. (2024) Variation in Leaf Chlorophyll Concentration in Response to Nitrogen Application Across Maize Hybrids in Contrasting Environments. microPublication Biology. <https://doi.org/10.17912/micropub.biology.001115>

**For more information:**

Visit the journal: <https://www.micropublication.org>

Questions can be sent to:

Karen Yook, Executive Editor ([karen.yook@micropublication.org](mailto:karen.yook@micropublication.org))

Carson Andorf, Maize Science Officer ([carson.andorf@usda.gov](mailto:carson.andorf@usda.gov))