Tools and Resources at the Maize Genetics and Genomics Database (MaizeGDB)

Margaret R. Woodhouse,[1,4] Ethalinda K. Cannon,[1] John L. Portwood II,[1] Jack M. Gardiner,[2] Rita K. Hayford,[1] Olivia Haley,[1] and Carson M. Andorf[1,3]

[1]*Agricultural Research Service, United States Department of Agriculture (USDA-ARS), Corn Insects and Crop Genetics Research Unit, Ames, Iowa 50011, USA;* [2]*Division of Animal Sciences, University of Missouri, Columbia, Missouri 65211, USA;* [3]*Department of Computer Science, Iowa State University, Ames, Iowa 50011, USA*

The Maize Genetics and Genomics Database (MaizeGDB) is the community resource for maize researchers, offering a suite of tools, informatics resources, and curated data sets to support maize genetics, genomics, and breeding research. Here, we provide an overview of the key resources available at MaizeGDB, including maize genomes, comparative genomics, and pan-genomics tools. This review aims to familiarize users with the range of options available for maize research and highlights the importance of MaizeGDB as a central hub for the maize research community. By providing a detailed snapshot of the database's capabilities, we hope to enable researchers to make use of MaizeGDB's resources, ultimately assisting them to better study the evolution and diversity of maize.

INTRODUCTION

Zea mays ssp. *mays* (maize or corn) has been the world's top production grain crop for over a decade (http://faostat.fao.org/). Maize's use as a resource is extensive and diverse, with its end products ranging from livestock feed to biofuels (Lawrence and Walbot 2007; Carpita and McCann 2008; Penning et al. 2009; Bosch et al. 2011). The breeding history of maize began thousands of years ago in the Americas, with traditional indigenous practices (Navarro et al. 2017; Aguirre-Liguori et al. 2019; Gonzalez-Segovia et al. 2019; Yang et al. 2023). Along the way, directed breeding, as well as the green revolution, helped shape maize into the crop most people are familiar with today (Prasanna 2012; Andorf et al. 2019).

Besides its usefulness as a grain crop, maize's role as a historic model organism for the last century has led to discoveries in plant development, evolution, domestication, and genetics and genomics that have advanced our understanding of biology across all organisms (Stadler 1928; Creighton and McClintock 1931; Beadle 1932; Rhoades 1984; Strable and Scanlon 2009). The historical collection of maize research data since the early twentieth century is vast, and includes gene lists, composite genetic maps, and germplasm accessions for use in research (Coe 2001).

The Maize Genetics and Genomics Database (MaizeGDB, https://www.maizegdb.org; Woodhouse et al. 2021a, 2023) is the community database for maize researchers, providing data curation and informatics resources to support maize genetics, genomics, and breeding research. MaizeGDB also

Advanced Online Article. Cite this overview as *Cold Spring Harb Protoc*; doi:10.1101/pdb.over108430

M.R. Woodhouse et al.

serves as the maize community's outreach and support center, facilitating data sharing, networking, and collaboration, and enforcing maize genetic nomenclature standards. MaizeGDB's first incarnation, MaizeDB, was established in 1991 to provide storage for historical maize gene and genetic data and maps (Polacco et al. 2002; Polacco and Coe 2006). In 2003, MaizeDB became MaizeGDB when it expanded to include gene sequence data (Lawrence et al. 2004), and then shifted its primary focus to genomic data in 2008 with the release of the first complete maize genome assembly for the reference genome B73 (Lawrence et al. 2004; Schnable et al. 2009). With the advancements in sequencing and assembly technology that have led to the release of dozens of maize genomes, MaizeGDB underwent an extensive revision in 2021 to include pan-genomic data for individual gene models and for genome browsers (Woodhouse et al. 2021a). A pan-genome is a comprehensive collection of genomic data from multiple individuals or lines of a species that enables researchers to study genetic variation and diversity across the species. The maize community's response to MaizeGDB's pan-genomic approach was significant, with a threefold increase in usage, particularly for the genome browsers (Woodhouse et al. 2023). As of 2024, MaizeGDB hosts 104 genome assemblies of diverse maize lines and their outgroups, ranging from the reference B73 genome to the Tripsacum outgroup *Andropogon virginicus*. Along with these assemblies, MaizeGDB hosts their associated gene model annotations, transcript and protein sequences, expression data, protein 3D structures, pathway information, variant information, and other data, where available.

In this overview, we highlight the key resources at MaizeGDB that combine genomic data sets and offer tools for comparative genomics and pan-genomics to help maize researchers study the evolution and diversity of maize more effectively. We also describe MaizeGDB's contributions to outreach resources for the maize community. The review is organized around the top eight components with the most extensive curation and support at MaizeGDB: (1) genomes, (2) browsers, (3) gene models, (4) pan-genomics, (5) diversity and phenotypes, (6) function, (7) machine learning, and (8) community.

GENOMES

Genome Data

The diversity of maize is well represented among the 104 genomes hosted by MaizeGDB. These include five iterations of the reference B73 (Schnable et al. 2009; Woodhouse et al. 2023); the important inbred lines PH207 (Hirsch et al. 2016), Mo17 (Hirsch et al. 2016; Yang et al. 2017; Sun et al. 2018), and W22 (Springer et al. 2018); four European flint lines (Springer et al. 2018; Haberer et al. 2020); sweet corn (Hu et al. 2021); 12 founder inbred lines representing heterotic groups from the United States and China (Wang et al. 2023); and a set of 25 high-quality genome assemblies of the nested associated mapping (NAM) population founder lines, which represent the broad diversity of domesticated maize (Hufford et al. 2021).

The 25 NAM founder lines were instrumental in MaizeGDB's evolution to a pan-genome-centric platform, and MaizeGDB's current framework is largely structured around these genomes. The NAM populations derived from these founder lines have been key in elucidating maize flowering time (Buckler et al. 2009), leaf architecture (Tian et al. 2011), disease resistance (Poland et al. 2011), and other important agronomic traits (Wallace et al. 2014). These high-quality NAM founder assemblies, gene model annotations, RNA-seq data, and epigenetic data were all generated using the same methods, and thus provide a robust foundation for exploring the maize pan-genome.

MaizeGDB's resources can be leveraged to investigate various aspects of the pan-genome, including presence/absence variation among genomes, copy number variation, structural variations and genome rearrangements, gene family evolution and expansion, and regulatory element conservation and divergence. By using these resources, researchers can gain a deeper understanding of the genetic and genomic diversity underlying complex traits in maize and uncover new insights into the evolution and adaptation of this important crop species.

Advanced Online Article. Cite this overview as *Cold Spring Harb Protoc*; doi:10.1101/pdb.over108430

Genome Stewardship and Curation

Since 2008, genomic data have been the cornerstone of MaizeGDB's research mission, ever since the first maize genome sequence was released (Schnable et al. 2009; Wallace et al. 2014). Initially, MaizeGDB hosted all subsequent maize genomes across all data platforms, including BLAST targets, genome browsers, gene model pages, and downloads. A nonexhaustive list of data types found on the download page includes a wide range of data types, such as genome sequences and metadata, gene model annotations, functional annotations, genome browser tracks, gene expression abundance files, variation data, protein structure models, insertions, mutant phenotype images, and metabolic pathways, which are all part of MaizeGDB's comprehensive curation efforts. As the number of maize genomes has expanded, including multiple versions of genomes of the same cultivar from different groups or projects and genomes of varying completeness, MaizeGDB now hosts maize genomes according to a tiered model:

- **Hosting level 3**. Full Support: Genomes and gene model annotations are offered as downloads, BLAST targets, individual gene pages, genome browsers, and lifted-over tracks on other browsers (which allow for the projection of annotations from one genome assembly to another), and are integrated into pan-gene data sets and other MaizeGDB tools and resources. Requirements for Full Support include (1) high-quality genome assemblies generated by long-read pipelines such as PacBio or its equivalent, particularly those that also use optical maps or Oxford Nanopore Technologies (ONT) ultra-long-read technology, and which have been assembled into pseudomolecules (i.e., complete chromosomes); (2) gene model annotations; (3) genomes of historical or contemporary significance, such as those that have played a key role in maize breeding programs or have contributed significantly to scientific breakthroughs and discoveries; and (4) genomes with at least two functional data sets associated with them, such as RNA-seq or epigenetic data.

- **Hosting level 2**. Limited Support: Genomes and gene model annotations are offered as downloads, BLAST targets, and a genome browser if two or more data sets have been aligned on the genome, in addition to gene models. If the assembly represents a maize clade that is not adequately represented in the maize pan-gene set, it will also be included in pan-gene data sets.

- **Hosting level 1**. Minimal Support: Genomes are offered as downloads and BLAST targets.

The B73 genome is the representative reference genome for maize, with the highest level of support. MaizeGDB will continue to identify and curate data onto the B73 genome and provide functional annotations for B73 genes. Additional support, integration, and functional annotation will also be provided to select high-quality genomes that represent modern heterotic groups (e.g., Mo17 and PH207) (van Heerwaarden et al. 2012), or that are used in forward genetics approaches, such as transposon insertion populations to identify mutant phenotypes (e.g., W22 and A188) (Vollbrecht et al. 2010; McCarty et al. 2013; Xiong et al. 2013; Marcon et al. 2020; Lin et al. 2021) or for *Agrobacterium*-mediated transformation (e.g., A188 and B104) (Kang et al. 2022). For all genomic data, MaizeGDB will continue to enforce nomenclature conventions, quality standards, and file formats set out by the maize nomenclature committee (see "Community" below).

B73 Genome Versions and Management

Where applicable, MaizeGDB hosts multiple versions of genome assemblies and annotations to reflect updates in genome sequencing and analysis. This is particularly evident in the case of the B73 maize genome, which has undergone five revisions since its initial release in 2009. Each new version incorporates improved sequencing technologies, assembly methods, and annotation updates, providing a more accurate and comprehensive representation of the genome. To address potential confusion, MaizeGDB has a webpage dedicated to the B73 assembly that describes the different B73 versions, provides cross-references between versions, and tracks changes between successive versions. This ensures that users can access and compare data across different versions, leveraging the strengths of each. For instance, researchers can use the latest version of B73 (v5) for its improved assembly and

annotation, while still accessing data from previous versions (v1–4) for continuity and historical context. By explicitly managing and documenting genome versions, MaizeGDB aims to facilitate a clear understanding of the relationships between different versions and to enable researchers to navigate the complexities of genome evolution with confidence.

BROWSERS

In 2010, MaizeGDB adopted the GBrowse software suite to visually represent maize genomes and their associated data sets (Sen et al. 2010). The initial versions of the reference B73 genome, along with the inbred lines PH207, W22, and the European Flint lines, are currently hosted as GBrowse instances in MaizeGDB. In 2020, MaizeGDB upgraded to the faster, more modern client-side JBrowse browser software (Sen et al. 2010; Woodhouse et al. 2021a). This upgrade enables researchers to efficiently explore and compare the genomes of diverse maize and wild relatives. Currently, MaizeGDB hosts 40 JBrowse browser instances, providing a platform for exploring the genetic diversity of *Zea*, including the NAM founder genomes and the PanAnd teosinte genomes. This extensive collection of browsers facilitates in-depth analysis of genetic variation and evolution across different maize lines and wild relatives, allowing researchers to identify genetic factors contributing to important traits and characteristics, develop new breeding strategies, and improve crop resilience. Additionally, it enhances our understanding of the complex relationships between maize and its wild relatives, driving innovation in plant breeding and genetics.

Genome browsers are the most frequently used tool in MaizeGDB (Woodhouse et al. 2023), and hence, they are given priority and prominence. The browsers are populated with gene model, expression, epigenetic, insertion, DNA binding, diversity, and other data tracks that work together to identify or validate the function of maize loci for a given genome (Fig. 1). Individual browsers are created for

FIGURE 1. Two sample JBrowse instances in MaizeGDB. (*A*) JBrowse instance of the reference genome B73, version 5. The *top* track represents B73 gene model Zm00001eb017450. *Below* it are tracks of gene model annotations lifted to B73 from other genomes, including the gene model Zm00020ab017810 from the genome CML69 (green). Notice that Zm00020ab017810 is missing two exons relative to the other tracks (red rectangle). Clicking on the gene model Zm00020ab017810 opens up a pop-up box with a link (red oval) to Zm00020ab017810 in its home genome CML69 JBrowse instance. (*B, top*) The exons missing in the B73 JBrowse are also missing in CML69, and are missing in the B73 gene model Zm00001eb017450, which has been lifted over to CML69 because those exon sequences are not present in CML69. (*Bottom*) RNA-seq tracks show that there are no RNA reads associated with the missing exon space, suggesting that the exons are indeed absent and might have been deleted and are not merely the result of a misannotation.

Advanced Online Article. Cite this overview as *Cold Spring Harb Protoc*; doi:10.1101/pdb.over108430

high-quality genomes that have gene model annotations and at least one or two functional data tracks (e.g., gene expression, epigenetic data, DNA binding, transposable elements, etc.). MaizeGDB also provides cross-genome functionality between browsers by linking gene models across the genomes using annotation liftover methods, allowing users to navigate from one genome browser to another based on gene model syntenic relationships (Woodhouse et al. 2021a). This feature assists users in comparing the function and annotation structure of syntenic orthologs across *Zea*.

GENE MODELS

Gene Model Pages

The gene model pages were developed in 2015 (Andorf et al. 2016; Woodhouse et al. 2021a) to organize gene, sequence, phenotype, and functional information together for a given gene model. The gene model record page initially had three tabs: gene model, sequence, and genetic information. The gene model tab includes information about the gene model, such as genome position, functional annotation, orthologs, protein function, browser links, expression data, and other functional evidence. The sequence tab provides FASTA sequences for the gene region, cDNA, and protein translation when available. The genetic information tab provides information such as map position, gene function, images of phenotypes associated with the gene, known variations, associated ESTs, primers, related loci, and more.

Gene model pages were initially generated for the gene models of all genomes. As MaizeGDB moved to its tiered support model, however, gene model pages were provided only for genomes that were given Full Support. In 2021, the pan-gene tab was added to show pan-gene relationships across the genomes hosted in MaizeGDB. This tab was expanded on by the creation of the pan-gene pages (see below); in the future, gene models for new genomes will simply be represented by their pan-gene relationships, if any, to the reference genome B73.

Gene Center

The Gene Center (https://maizegdb.org/gene_center) is a resource for searching, connecting, and bulk-downloading maize gene and gene model data in various formats and modalities, useful for the diverse needs of the maize research community. The Gene Center provides Search, Download, and Information features.

For the Simple Search feature, a user enters a gene or gene model ID or wildcard, selects the number of records per page, and then submits the selection. The Advanced Search feature allows a user to specify data types and chromosomal locations to be retrieved. A user can also perform a gene model search by sequence: the user may enter a list of up to five gene IDs, which will then be sent to the MaizeGDB BLAST search.

Data can be downloaded in many ways from the Gene Center at MaizeGDB, by finding the syntenic relationships between genes of interest and gene models from the genomes hosted in MaizeGDB, or by genomic coordinate region and gene model set, genetic markers, or assembly annotation positions, where the user enters a list of gene models in the search box and retrieves the gene model ID, chromosomal positions, and assembly ID for each gene model.

PAN-GENOMICS

MaizeGDB's pan-genomics resources are valuable tools for studying the genetic diversity of maize and its wild relatives. Pan-genomics involves analyzing the complete set of genes within a species or genus, focusing on gene variations, relationships, and evolutionary processes. This approach is especially important in maize research, where it facilitates the comparison of genomes across different (and often diverse) maize lines and wild relatives. It also helps connect gene models across genome

assemblies, enhancing our understanding of how genetic variations affect plant development or contribute to complex traits like drought tolerance, disease resistance, and yield potential. Using MaizeGDB's pan-genomics feature enables researchers to study the genetic basis of these important traits, assess the evolutionary relationships among maize genes, and develop strategies to enhance crop development and resilience.

As of 2023, the main framework of pan-genomics within MaizeGDB focuses on two key features: (1) the integration of pan-genomes into JBrowse for visual exploration and (2) the detailed pan-gene pages, which expand on the previously mentioned gene model pages. Pan-genome JBrowse enables visual exploration of maize genomes, while pan-gene pages provide detailed information on individual genes and their relationships across the *Zea* genus. Both features use sequence and positional similarities among maize gene models, established through distinct pipelines tailored to meet the unique needs of each tool. The next three sections describe how the pan-genomes enable the linkage of loci across genome browsers and our new pan-gene center and pan-gene pages, which make the pan-genomic data more accessible and facilitate a more comprehensive understanding of maize genetics.

Browser Crossover

The browser cross-genome functionality described above in "Browsers," which links gene models across the genomes using annotation liftover methods, allows users to navigate from one genome browser to another based on gene model syntenic relationships (Woodhouse et al. 2021a). This feature assists users in comparing the function and annotation structure of syntenic orthologs across *Zea* and, in this way, serves as a pan-genome viewer of sorts.

One of the challenges of pan-genome visualization is viewing and navigating across multiple genomes independent of a reference genome, because reference-centric pan-genomes cannot represent data that are present in some genomes but absent in the reference genome. However, some pan-genomic visualization schemes independent of a reference-genome coordinate system, such as graphical visualization with nodes and edges, can be difficult for users to navigate, and generally lack functional information. MaizeGDB overcame this challenge by connecting gene models, SNPs, or markers shared between two or more genomes across our JBrowse instances. This means that when you click on a feature in JBrowse, a pop-up window will appear with links to other maize accessions' JBrowse instances, taking you directly to the corresponding location in those genomes using pan-genome information. These linkages allow users to move from one genome to another with ease, facilitating cross-genome functional and structural comparisons of syntenic loci.

Pan-Gene Center

The pan-gene data center home page (https://maizegdb.org/pan_gene_center/pan_gene) provides a variety of means for querying and downloading pan-gene data. The page is organized similarly to the Gene Center, with a few modifications. The Simple Search section allows researchers to search pan-genes by locus association, gene model ID, transcript ID, or protein ID. The Advanced Search, on the other hand, permits more detailed searches: rather than establishing a definition for core, dispensable, and rare gene models—which refer to genes consistently present across most genomes (core), genes present in some but not all genomes (disposable), and genes unique to a single genome (rare)—for our pan-gene data sets, users can now customize their searches using various criteria. MaizeGDB has structured the pan-gene Advanced Search so that researchers can set parameters to create their own definitions of these categories, find pan-genes present or absent in a subset of annotations, limit results by pan-gene size (number of members across genomes), and other criteria. An option to download the full search results is given. Note that the results table includes the pan-gene exemplar (the gene model selected as being the standard for a pan-gene), which may not be the gene model queried. The pan-gene center also provides download options and features statistics on the pan-gene distribution based on the number of members in each pan-gene set.

Pan-Gene Pages

Gene model pan-genes were first introduced in MaizeGDB in 2020 (Woodhouse et al. 2021a) in the form of a pan-gene tab on the gene model pages, and initially included a list of all members of all genomes within the pan-gene. In 2023, MaizeGDB enhanced its pan-gene platform by developing a dedicated pan-gene record page, and through integration into MaizeGDB's comparative genomics tools like JBrowse, qTeller, and MaizeMine (Cannon et al. 2024). The pan-gene record page is divided into two or more tabs: one for the pan-gene record page and one for the Genomic Context Viewer (GCV) (https://gcv.maizegdb.org/; Woodhouse et al. 2021a; Cleary and Farmer 2018, 2023). The GCV enables exploration of pan-gene macro- and microsynteny—which refer to the large-scale (macro) and small-scale (micro) conservation of gene order and structure across different genomes—among the genomes, and provides a tab for each associated locus, if any, offering a detailed view of the genetic context and organization.

The pan-gene tab provides multiple ways of exploring the pan-gene and its members. Pan-genes are derived using the program Pandagma (https://github.com/legumeinfo/pandagma) and are based on both sequence similarity and synteny. The information in the pan-gene tab includes all data in MaizeGDB that are linked to gene models that are in the pan-gene. The majority of these data are linked to versions 3–5 of the B73 genome assemblies and gene model annotations. Therefore, if the pan-gene has B73 as a member, all data linked to it are applied to all the members of a pan-gene.

The pan-gene tab is subdivided into several sections; namely, pan-gene statistics, the gene model members within the pan-gene, CDS and protein alignments of all members, a phylogenetic tree of all pan-gene members, and CDS and protein sequence for the pan-gene exemplar or representative. The following information is also included for the B73 syntenic ortholog gene model of the pan-gene, if present: mutation insertional data; metabolic pathway information; expression information; protein information, including 3D structural visualization of the B73 gene model; SNPs and information about SNPs, such as those linked to traits by genome-wide association studies (GWASs); and functional information, including ontology terms assigned through literature curation or computational annotations.

If the pan-gene is associated with one or more loci, the locus tabs are labeled with the locus symbol (e.g., *lg1* for *liguleless1*). The locus information includes locus names, synonyms, descriptions, comments from references and personal communications, gene products, phenotypes, and other information.

DIVERSITY AND PHENOTYPES

MaizeGDB has served as a resource for maize mutant and phenotype information for maize researchers since 2003 (Lawrence et al. 2004; Cleary and Farmer 2023). The integration of variant, trait, mutant, and phenotype information into one database makes it easy for users to study important loci of interest that can contribute to breeding and other types of research. In the following sections, we describe these different classes of mutant and phenotype information and how they can be viewed in MaizeGDB.

Diversity/SNPs/Traits

The Diversity, SNP, and Trait Data section at MaizeGDB (https://maizegdb.org/diversity) provides users with multiple ways of searching variant data and offers downloads of variant information from multiple resources (described in Lawrence et al. 2004). The types of data available in this section include SNPs and indels, copy number variations (CNVs) and presence/absence variations (PAVs), complex alleles (also called "alternative loci"), sequences that do not align to the reference genome, inversion breakpoints, transcript assemblies and exome contigs alignments, and genetic maps.

Cold Spring Harbor Protocols
www.cshprotocols.org

M.R. Woodhouse et al.

Mutants and Phenotypes

The mutant and phenotype resource at MaizeGDB (https://maizegdb.org/data_center/phenotype) includes simple and advanced phenotype searches, phenotype image searches (described in detail in Portwood et al. 2019), mutant stock searches from the Maize Genetics Cooperation Stock Center, resources to obtain mutants from individual maize research laboratories, and additional links to transposon insertion and point mutation (e.g., ethylmethane sulfonate [EMS]-induced mutants) information.

Newly Characterized Genes

MaizeGDB keeps records of genes whose functions have been characterized in the literature (https://maizegdb.org/new_genes; described in Lawrence et al. 2004; Portwood et al. 2019). These records are kept up to date by MaizeGDB curators and collaborators. The newly characterized genes page provides a table of genes sorted by how recently they were added or received a new reference, gene product, and variation annotation. Information such as gene ID and links to the publications describing the function are also included.

FUNCTION

MaizeGDB hosts multiple tools and resources to study gene function, including tools to visualize RNA expression (qTeller), metabolic pathway tools (CornCyc), and platforms to compare different functions for a given gene or set of genes (MaizeMine).

qTeller

qTeller (https://qteller.maizegdb.org/) was created in 2012 as a web-hosted RNA-seq visualization platform that allows users to compare RNA-seq expression across tissues within a genomic interval or across multiple genes, or to compare expression between any two genes in a given genome (https://github.com/jschnable/qTeller). The platform displays preprocessed expression values from publicly available, published data sets. MaizeGDB adopted qTeller in 2018 (Portwood et al. 2019) and, in 2021, expanded it considerably to include the NAM founder genomes (Hufford et al. 2021) and capabilities for comparison of both mRNA and protein abundance data sets (Woodhouse et al. 2021b).

MaizeMine

MaizeMine (https://maizemine.rnet.missouri.edu/maizemine/begin.do) is a data mining resource that enables researchers to create and export customized annotation data sets that can be merged with their own data for use in downstream analyses (Shamimuzzaman et al. 2020; Woodhouse et al. 2021b). MaizeMine uses the InterMine data warehousing system (Kalderimis et al. 2014) to integrate genomic sequences and annotations from the reference maize B73 genome, as well as data from the NAM founder genomes (Kalderimis et al. 2014; Hufford et al. 2021). Data types include gene ontology annotations, SNPs, protein annotations, homologs, metabolic pathways, and precomputed gene expression levels based on RNA-seq data from MaizeGDB qTeller and the B73 Gene Expression Atlas. MaizeMine also provides a cross-reference database that enables users to link and navigate between alternative gene sets from different sources, including Gramene and NCBI. This is achieved through a lookup table that maps gene model identifiers from various annotation sources. MaizeMine includes several search tools, including a keyword search, built-in template queries, and a Query-Builder tool for creating custom queries.

Metabolic Pathways

The metabolic pathway data page at MaizeGDB (https://maizegdb.org/metabolic_pathways) is comprised of CornCyc, MaizeCyc, and various links to downloads and other resources. CornCyc is a

Cold Spring Harbor Protocols

www.cshprotocols.org

metabolic pathway resource developed using the Pathway Tools Software Suite in a collaboration between MaizeGDB and the Plant Metabolic Network (PMN) (Walsh et al. 2016). MaizeGDB curators performed literature-based manual annotation for both resources. MaizeCyc (Monaco et al. 2013) is a comparable tool that is no longer supported at MaizeGDB, but comparisons between CornCyc and MaizeCyc can be found at the metabolic pathway data page.

MACHINE LEARNING

MaizeGDB's artificial intelligence and machine learning data center (https://maizegdb.org/ai) offers a suite of tools, resources, and downloadable data to facilitate in-depth analysis and exploration of maize genetics and genomics. Leveraging machine learning methods and large-scale data sets, these tools enable researchers to visualize and compare protein structure models, predict and interpret the functional consequences of genetic variations, identify potential genetic targets for crop enhancement, and gain a deeper understanding of the complex relationships between genetic variation, protein function, and trait expression. In the following sections, we describe the features and applications of these tools, including tools for protein structures and variant effect analysis (PanEffect), and a resource to use and analyze gene-based features (Maize Feature Store).

Protein Structures

The MaizeGDB protein structure page (https://maizegdb.org/data_center/protein_structure) comprises several tools to analyze the 3D protein structure of maize genes. The Foldseek tool is based on a free and open-source program (https://github.com/steineggerlab/foldseek; van Kempen et al. 2024) that enables fast and sensitive comparisons of large structure sets. MaizeGDB used Foldseek to align the 39,299 AlphaFold-predicted protein structures from maize against eight proteomes: *Arabidopsis thaliana*, *Glycine max*, *Oryza sativa*, maize, human, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and sorghum. This resulted in a framework that allows users to easily explore and compare the structural similarities and differences between maize proteins and their orthologs in other species, facilitating a deeper understanding of protein function and evolution. Users can search for specific maize genes or proteins and visualize their predicted 3D structure, as well as browse the results of the Foldseek alignments to identify potential functional relationships and infer functional predictions based on structural similarities.

The FATCAT tool shows the structural alignments between a query maize protein with the top hits from the sequence-based tool Diamond and two structural alignment tools: Foldseek and FATCAT (https://fatcat.godziklab.org/; Li et al. 2020). Each of the 39,299 AlphaFold-predicted protein structures from maize was aligned against the following four plant proteomes: *Arabidopsis thaliana*, *Glycine max*, *Oryza sativa*, and sorghum. Notably, one of the visualizations offered is a structural overlay of the maize protein structure model with the protein model of orthologs in the other four species, enabling researchers to visualize and compare the structural conservation and divergence between maize and its orthologs. This information can be used to make functional predictions, identify potential functional relationships, and understand the evolutionary dynamics of protein structures across different species. For example, researchers can use this tool to identify structurally conserved regions between maize and its orthologs, potentially indicating functional importance, or to identify structural variations that may be associated with specific functional differences. The protein structure models are from version 3, released in July 2022. Details on the above analyses can be found in the studies by Andorf et al. (2022) and Woodhouse et al. (2023).

PanEffect

PanEffect (https://maizegdb.org/effect/maize/; Andorf et al. 2024) is a tool designed for in-depth analysis of the potential impact of missense variants on maize proteins. It provides users with an interactive interface comprising four specialized views. The Search function enables users to make

queries based on gene names and protein identifiers. The Gene Summary section presents an overview of the gene and protein annotations. In the Variant Effects within a genome view, detailed heat maps show the predicted functional impact of all possible amino acid substitutions for the B73 maize reference genome. These heat maps, which transition from blue to red to indicate impact severity, provide insights into each position of the reference protein and the possible amino acid substitutions. Additionally, the tool allows for the exploration of Variant Effects across the pan-genome, with additional heat maps that reflect the effects of natural variations on maize proteins across 50 maize genomes. In total, over 550 million effect scores for nearly 39,000 proteins across the maize pan-genome are available as downloads. These predictions can empower researchers to predict and interpret the functional consequences of genetic variations and understand how variation affects proteins involved in maize development, growth, and response to environmental stimuli.

Maize Feature Store

The Maize Feature Store (MFS) (https://mfs.maizegdb.org/; Sen et al. 2023; Andorf et al. 2024) is a centralized repository and framework that hosts a collection of gene-based features derived from multiomics data, including genomics, transcriptomics, epigenomics, variomics, and proteomics. With over 14,000 features, the MFS enables researchers to explore and model classification problems related to maize genetics and genomics. The platform integrates various machine learning algorithms, both supervised and unsupervised, to simplify the analysis and prediction of genome annotations. An initial use case demonstrated the effectiveness of the MFS in achieving high classification accuracy and identifying the contribution of different feature types, such as DNA and protein sequence, gene structure, gene expression, epigenetic marks, and genetic variation, in distinguishing core and noncore genes in the maize pan-genome.

COMMUNITY

In many respects, MaizeGDB acts as the central information hub for the global maize community. MaizeGDB team members educate users through presentations, workshops, and direct communication on the application of new tools for accessing new data. MaizeGDB considers it of the utmost importance to keep users apprised of outreach and education resources, and also plays a major role in various community and database committees, conference planning, outreach, and website development. The next sections discuss the roles of MaizeGDB in four important community-centric groups; namely, the Maize Genetics Cooperation, Maize Genetics Meeting, the AgBioData Consortium, and the Maize Nomenclature Committee.

Maize Genetics Cooperation (MGC)

The Maize Genetics Cooperation (MGC) is a 501(c)(3) nonprofit organization that exists to advance the interests of the maize genetics, genomics, and breeding community, and to train and develop the next generation of scientists (https://maizegdb.org/mgc/). MaizeGDB staff provides support for the leadership groups in the MGC, and coordinates elections for the MGC; the Committee on Outreach, Diversity, Inclusion, and Education (CODIE) (https://maizegdb.org/mgc/outreach/index.php); and the Maize Genetics Meeting Steering Committee (MGMSC). MaizeGDB team members also conduct surveys of the maize community as directed by the MGC/MGMSC and CODIE and collect nominations for community awards.

Maize Genetics Meeting

MaizeGDB team members play a critical role in facilitating the annual Maize Meeting (https://maizegdb.org/mgc/maizemeeting). Key contributions include developing software to enable the submission of abstracts, student financial aid applications, and community contests for the Annual Maize

Genetics Meeting; maintaining the conference website; creating the conference program; hosting preconference workshops pertinent to the latest research and tool development in the maize community; and providing in-person technical support at the conference such as audio and visual support.

AgBioData Consortium

MaizeGDB team members cofounded the AgBioData working group (https://www.agbiodata.org/), a collection of agriculture-related databases that work together to identify and solve common problems across agricultural databases. MaizeGDB continues to participate in and guide the goals of AgBioData to ensure the continuity of database standards between maize and other agricultural databases.

Maize Nomenclature Committee

With multiple genomes hosted at MaizeGDB, it was necessary to establish consistent naming conventions for the genome assemblies, annotation sets, and gene models that are both human- and machine-readable. MaizeGDB team members collaborate with the Maize Nomenclature Committee to create and maintain nomenclature, metadata, and format standards for genome assemblies and annotations (https://maizegdb.org/nomenclature). MaizeGDB acts as the nomenclature clearinghouse for maize genes and genome assemblies, ensuring consistency of file and format standards across the global maize community (Woodhouse et al. 2021a).

CONCLUDING REMARKS

MaizeGDB has evolved significantly since its inception, adapting to the needs of its users and the rapid advancements in technologies, from high-throughput sequencing to machine learning and artificial intelligence. Today, MaizeGDB represents a comprehensive resource for the maize research community, offering a vast collection of genomic data, comparative genomics, and pan-genomics tools, with outreach resources to facilitate collaboration and knowledge sharing. By highlighting here the key features and resources available at MaizeGDB, we aim to improve data accessibility and allow researchers to explore the diversity of maize genetics and genomics. As maize continues to play a vital role in global food security, bioenergy, and as a model organism for plant research, MaizeGDB remains committed to supporting the research community in achieving its full potential.

COMPETING INTEREST STATEMENT

The authors declare no competing interests.

ACKNOWLEDGMENTS

Cold Spring Harbor Protocols

www.cshprotocols.org

M.R. Woodhouse et al.

AUTHOR CONTRIBUTIONS

REFERENCES

Aguirre-Liguori JA, Gaut BS, Jaramillo-Correa JP, Tenaillon MI, Montes-Hernández S, García-Oliva F, Hearne SJ, Eguiarte LE. 2019. Divergence with gene flow is driven by local adaptation to temperature and soil phosphorus concentration in teosinte subspecies (*Zea mays parviglumis* and *Zea mays mexicana*). *Mol Ecol* **28:** 2814–2830. doi:10.1111/mec .15098

Andorf CM, Cannon EK, Portwood JL 2nd, Gardiner JM, Harper LC, Schaeffer ML, Braun BL, Campbell DA, Vinnakota AG, Sribalusu VV, et al. 2016. MaizeGDB update: new tools, data and interface for the maize model organism database. *Nucleic Acids Res* **44:** D1195–D1201. doi:10.1093/nar/gkv1007

Andorf C, Beavis WD, Hufford M, Smith S, Suza WP, Wang K, Woodhouse M, Yu J, Lübberstedt T. 2019. Technological advances in maize breeding: past, present and future. *Theor Appl Genet* **132:** 817–849. doi:10 .1007/s00122-019-03306-3

Andorf CM, Sen S, Hayford RK, Portwood JL II, Cannon EK, Harper LC, Gardiner JM, Sen TZ, Woodhouse MR. 2022. FASSO: an AlphaFold based method to assign functional annotations by combining sequence and structure orthology. bioRxiv doi:10.1101/2022.11.10.516002

Andorf CM, Haley OC, Hayford RK, Portwood JL 2nd, Harding S, Sen S, Cannon EK, Gardiner JM, Kim H-S, Woodhouse MR. 2024. Paneffect: a pan-genome visualization tool for variant effects in maize. *Bioinformatics* **40:** btae073. doi:10.1093/bioinformatics/btae073

Beadle GW. 1932. The relation of crossing over to chromosome association in *Zea–Euchlaena* hybrids. *Genetics* **17:** 481–501. doi:10.1093/genetics/ 17.4.481

Bosch M, Mayer C-D, Cookson A, Donnison IS. 2011. Identification of genes involved in cell wall biogenesis in grasses by differential gene expression profiling of elongating and non-elongating maize internodes. *J Exp Bot* **62:** 3545–3561. doi:10.1093/jxb/err045

Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, Browne C, Ersoz E, Flint-Garcia S, Garcia A, Glaubitz JC, et al. 2009. The genetic architecture of maize flowering time. *Science* **325:** 714–718. doi:10 .1126/science.1174276

Cannon EK, Portwood JL II, Hayford RK, Haley OC, Gardiner JM, Andorf CM, Woodhouse MR. 2024. Enhanced pan-genomic resources at the maize genetics and genomics database. *Genetics* **227:** iyae036. doi:10 .1093/genetics/iyae036

Carpita NC, McCann MC. 2008. Maize and sorghum: genetic resources for bioenergy grasses. *Trends Plant Sci* **13:** 415–420. doi:10.1016/j.tplants .2008.06.002

Cleary A, Farmer A. 2018. Genome context viewer: visual exploration of multiple annotated genomes using microsynteny. *Bioinformatics* **34:** 1562–1564. doi:10.1093/bioinformatics/btx757

Cleary AM, Farmer AD. 2023. Genome context viewer (GCV) version 2: enhanced visual exploration of multiple annotated genomes. *Nucleic Acids Res* **51:** W225–W231. doi:10.1093/nar/gkad391

Coe EH Jr. 2001. The origins of maize genetics. *Nat Rev Genet* **2:** 898–905. doi:10.1038/35098524

Creighton HB, McClintock B. 1931. A correlation of cytological and genetical crossing-over in *Zea mays*. *Proc Natl Acad Sci* **17:** 492–497. doi:10 .1073/pnas.17.8.492

Gonzalez-Segovia E, Pérez-Limon S, Cíntora-Martínez GC, Guerrero-Zavala A, Janzen GM, Hufford MB, Ross-Ibarra J, Sawers RJH. 2019. Characterization of introgression from the teosinte *Zea mays* ssp. *mexicana* to Mexican highland maize. *PeerJ* **7:** e6815. doi:10.7717/peerj .6815

Haberer G, Kamal N, Bauer E, Gundlach H, Fischer I, Seidel MA, Spannagl M, Marcon C, Ruban A, Urbany C, et al. 2020. European maize genomes highlight intraspecies variation in repeat and gene content. *Nat Genet* **52:** 950–957. doi:10.1038/s41588-020-0671-9

Hirsch CN, Hirsch CD, Brohammer AB, Bowman MJ, Soifer I, Barad O, Shem-Tov D, Baruch K, Lu F, Hernandez AG, et al. 2016. Draft assembly of elite inbred line PH207 provides insights into genomic and transcriptome diversity in maize. *Plant Cell* **28:** 2700–2714. doi:10.1105/tpc .16.00353

Hu Y, Colantonio V, Müller BSF, Leach KA, Nanni A, Finegan C, Wang B, Baseggio M, Newton CJ, Juhl EM, et al. 2021. Genome assembly and population genomic analysis provide insights into the evolution of modern sweet corn. *Nat Commun* **12:** 1227. doi:10.1038/s41467-021- 21380-4

Hufford MB, Seetharam AS, Woodhouse MR, Chougule KM, Ou S, Liu J, Ricci WA, Guo T, Olson A, Qiu Y, et al. 2021. De novo assembly, annotation, and comparative analysis of 26 diverse maize genomes. *Science* **373:** 655–662. doi:10.1126/science.abg5289

Kalderimis A, Lyne R, Butano D, Contrino S, Lyne M, Heimbach J, Hu F, Smith R, Štěpán R, Sullivan J, et al. 2014. Intermine: extensive web services for modern biology. *Nucleic Acids Res* **42:** W468–W472. doi:10.1093/nar/gku301

Kang M, Lee K, Finley T, Chappell H, Veena V, Wang K. 2022. An improved -mediated transformation and genome-editing method for maize inbred B104 using a ternary vector system and immature embryos. *Front Plant Sci* **13:** 860971. doi:10.3389/fpls.2022.860971

Lawrence CJ, Walbot V. 2007. Translational genomics for bioenergy production from fuelstock grasses: maize as the model species. *Plant Cell* **19:** 2091–2094. doi:10.1105/tpc.107.053660

Lawrence CJ, Dong Q, Polacco ML, Seigfried TE, Brendel V. 2004. MaizeGDB, the community database for maize genetics and genomics. *Nucleic Acids Res* **32:** D393–D397. doi:10.1093/nar/gkh011

Li Z, Jaroszewski L, Iyer M, Sedova M, Godzik A. 2020. FATCAT 2.0: towards a better understanding of the structural diversity of proteins. *Nucleic Acids Res* **48:** W60–W64. doi:10.1093/nar/gkaa443

Lin G, He C, Zheng J, Koo D-H, Le H, Zheng H, Tamang TM, Lin J, Liu Y, Zhao M, et al. 2021. Chromosome-level genome assembly of a regenerable maize inbred line A188. *Genome Biol* **22:** 175. doi:10.1186/ s13059-021-02396-x

Marcon C, Altrogge L, Win YN, Stöcker T, Gardiner JM, Portwood JL 2nd, Opitz N, Kortz A, Baldauf JA, Hunter CT, et al. 2020. *BonnMu*: a sequence-indexed resource of transposon-induced maize mutations for functional genomics studies. *Plant Physiol* **184:** 620–631. doi:10 .1104/pp.20.00478

McCarty DR, Suzuki M, Hunter C, Collins J, Avigne WT, Koch KE. 2013. Genetic and molecular analyses of UniformMu transposon insertion lines. *Methods Mol Biol* **1057:** 157–166. doi:10.1007/978-1-62703-568- 2_11

Monaco MK, Sen TZ, Dharmawardhana PD, Ren L, Schaeffer M, Naithani S, Amarasinghe V, Thomason J, Harper L, Gardiner J, et al. 2013. Maize metabolic network construction and transcriptome analysis. *Plant Genome* **6:** plantgenome2012.09.0025. doi:10.3835/plantgenome2012 .09.0025

Navarro JAR, Willcox M, Burgueño J, Romay C, Swarts K, Trachsel S, Preciado E, Terron A, Delgado HV, Vidal V, et al. 2017. A study of allelic diversity underlying flowering-time adaptation in maize landraces. *Nat Genet* **49:** 476–480. doi:10.1038/ng.3784

Penning BW, Hunter C3, Tayengwa R, Eveland AL, Dugard CK, Olek AT, Vermerris W, Koch KE, McCarty DR, Davis MF, et al. 2009. Genetic resources for maize cell wall biology. *Plant Physiol* **151:** 1703–1728. doi:10.1104/pp.109.136804

Polacco M, Coe E. 2006. MaizeDB: the maize genome database. In *Bioinformatics: databases and systems* (ed. Letovsky SL), pp. 151–162. Kluwer Academic Publishers, Boston.

Polacco M, Coe E, Fang Z, Hancock D, Sanchez-Villeda H, Schroeder S. 2002. MaizeDB—a functional genomics perspective. *Comp Funct Genomics* **3:** 128–131. doi:10.1002/cfg.157

Poland JA, Bradbury PJ, Buckler ES, Nelson RJ. 2011. Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize. *Proc Natl Acad Sci* **108:** 6893–6898. doi:10.1073/pnas.1010894108

Portwood JL 2nd, Woodhouse MR, Cannon EK, Gardiner JM, Harper LC, Schaeffer ML, Walsh JR, Sen TZ, Cho KT, Schott DA, et al. 2019. MaizeGDB 2018: the maize multi-genome genetics and genomics database. *Nucleic Acids Res* **47:** D1146–D1154. doi:10.1093/nar/gky1046

Prasanna BM. 2012. Diversity in global maize germplasm: characterization and utilization. *J Biosci* **37:** 843–855. doi:10.1007/s12038-012-9227-1

Rhoades MM. 1984. The early years of maize genetics. *Annu Rev Genet* **18:** 1–29. doi:10.1146/annurev.ge.18.120184.000245

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, et al. 2009. The B73 maize genome: complexity, diversity, and dynamics. *Science* **326:** 1112–1115. doi:10.1126/science.1178534

Sen TZ, Harper LC, Schaeffer ML, Andorf CM, Seigfried TE, Campbell DA, Lawrence CJ. 2010. Choosing a genome browser for a model organism database: surveying the maize community. *Database* **2010:** baq007. doi:10.1093/database/baq007

Sen S, Woodhouse MR, Portwood JL 2nd, Andorf CM. 2023. Maize feature store: a centralized resource to manage and analyze curated maize multi-omics features for machine learning applications. *Database* **2023:** baad078. doi:10.1093/database/baad078

Shamimuzzaman M, Gardiner JM, Walsh AT, Triant DA, Le Tourneau JJ, Tayal A, Unni DR, Nguyen HN, Portwood JL 2nd, Cannon EKS, et al. 2020. MaizeMine: a data mining warehouse for the maize genetics and genomics database. *Front Plant Sci* **11:** 592730. doi:10.3389/fpls.2020.592730

Springer NM, Anderson SN, Andorf CM, Ahern KR, Bai F, Barad O, Barbazuk WB, Bass HW, Baruch K, Ben-Zvi G, et al. 2018. The maize W22 genome provides a foundation for functional genomics and transposon biology. *Nat Genet* **50:** 1282–1288. doi:10.1038/s41588-018-0158-0

Stadler LJ. 1928. Genetic effects of X-rays in maize. *Proc Natl Acad Sci* **14:** 69–75. doi:10.1073/pnas.14.1.69

Strable J, Scanlon MJ. 2009. Maize (*Zea mays*): a model organism for basic and applied research in plant biology. *Cold Spring Harb Protoc* **2009:** pdb.emo132. doi:10.1101/pdb.emo132

Sun S, Zhou Y, Chen J, Shi J, Zhao H, Zhao H, Song W, Zhang M, Cui Y, Dong X, et al. 2018. Extensive intraspecific gene order and gene structural variations between Mo17 and other maize genomes. *Nat Genet* **50:** 1289–1295. doi:10.1038/s41588-018-0182-0

Tian F, Bradbury PJ, Brown PJ, Hung H, Sun Q, Flint-Garcia S, Rocheford TR, McMullen MD, Holland JB, Buckler ES. 2011. Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nat Genet* **43:** 159–162. doi:10.1038/ng.746

van Heerwaarden J, Hufford MB, Ross-Ibarra J. 2012. Historical genomics of North American maize. *Proc Natl Acad Sci* **109:** 12420–12425. doi:10.1073/pnas.1209275109

van Kempen M, Kim SS, Tumescheit C, Mirdita M, Lee J, Gilchrist CLM, Söding J, Steinegger M. 2024. Fast and accurate protein structure search with Foldseek. *Nat Biotechnol* **42:** 243–246. doi:10.1038/s41587-023-01773-0

Vollbrecht E, Duvick J, Schares JP, Ahern KR, Deewatthanawong P, Xu L, Conrad LJ, Kikuchi K, Kubinec TA, Hall BD, et al. 2010. Genome-wide distribution of transposed dissociation elements in maize. *Plant Cell* **22:** 1667–1685. doi:10.1105/tpc.109.073452

Wallace JG, Bradbury PJ, Zhang N, Gibon Y, Stitt M, Buckler ES. 2014. Association mapping across numerous traits reveals patterns of functional variation in maize. *PLoS Genet* **10:** e1004845. doi:10.1371/journal.pgen.1004845

Walsh JR, Schaeffer ML, Zhang P, Rhee SY, Dickerson JA, Sen TZ. 2016. The quality of metabolic pathway resources depends on initial enzymatic function assignments: a case for maize. *BMC Syst Biol* **10:** 129. doi:10.1186/s12918-016-0369-x

Wang B, Hou M, Shi J, Ku L, Song W, Li C, Ning Q, Li X, Li C, Zhao B, et al. 2023. De novo genome assembly and analyses of 12 founder inbred lines provide insights into maize heterosis. *Nat Genet* **55:** 312–323. doi:10.1038/s41588-022-01283-w

Woodhouse MR, Cannon EK, Portwood JL 2nd, Harper LC, Gardiner JM, Schaeffer ML, Andorf CM. 2021a. A pan-genomic approach to genome databases using maize as a model system. *BMC Plant Biol* **21:** 385. doi:10.1186/s12870-021-03173-5

Woodhouse MR, Sen S, Schott D, Portwood JL II, Freeling M, Walley JW, Andorf CM, Schnable JC. 2021b. Qteller: a tool for comparative multi-genomic gene expression analysis. *Bioinformatics* **38:** 236–242. doi:10.1093/bioinformatics/btab604

Woodhouse MR, Portwood JL 2nd, Sen S, Hayford RK, Gardiner JM, Cannon EK, Harper LC, Andorf CM. 2023. Maize protein structure resources at the maize genetics and genomics database. *Genetics* **224:** iyad016. doi:10.1093/genetics/iyad016

Xiong W, He L, Li Y, Dooner HK, Du C. 2013. InsertionMapper: a pipeline tool for the identification of targeted sequences from multidimensional high throughput sequencing data. *BMC Genomics* **14:** 679. doi:10.1186/1471-2164-14-679

Yang N, Xu X-W, Wang R-R, Peng W-L, Cai L, Song J-M, Li W, Luo X, Niu L, Wang Y, et al. 2017. Contributions of *Zea mays* subspecies *mexicana* haplotypes to modern maize. *Nat Commun* **8:** 1874. doi:10.1038/s41467-017-02063-5

Yang N, Wang Y, Liu X, Jin M, Vallebueno-Estrada M, Calfee E, Chen L, Dilkes BP, Gui S, Fan X, et al. 2023. Two teosintes made modern maize. *Science* **382:** eadg8940. doi:10.1126/science.adg8940

Cold Spring Harbor Protocols

Tools and Resources at the Maize Genetics and Genomics Database (MaizeGDB)

Margaret R. Woodhouse, Ethalinda K. Cannon, John L. Portwood II, Jack M. Gardiner, Rita K. Hayford, Olivia Haley and Carson M. Andorf

Email Alerting Service	Receive free email alerts when new articles cite this article - **click here.**
Subject Categories	Browse articles on similar topics from *Cold Spring Harbor Protocols.* Bioinformatics/Genomics, general (202 articles) Genetics, general (429 articles) Laboratory Organisms, general (933 articles) Maize (65 articles) Maize: A Laboratory Manual (65 articles) Plant Biology, general (174 articles) Zea mays L. (4 articles)

To subscribe to *Cold Spring Harbor Protocols* go to:
http://cshprotocols.cshlp.org/subscriptions