

Original article

MaizeGDB: curation and outreach go hand-in-hand

Mary L. Schaeffer^{1,*}, Lisa C. Harper^{2,3}, Jack M. Gardiner^{4,5}, Carson M. Andorf⁶, Darwin A. Campbell⁶, Ethalinda K.S. Cannon⁵, Taner Z. Sen^{5,6} and Carolyn J. Lawrence^{5,6}

¹USDA-ARS Plant Genetics Research Unit and Division of Plant Sciences, Department of Agronomy, University of Missouri, Columbia, MO 65211,

²USDA-ARS Plant Gene Expression Center, Albany, CA 94710, ³Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, ⁴Department of Plant Sciences, University of Arizona, Tucson, AZ 85721, ⁵Development and Cell Biology, Department of Genetics, Iowa State University and ⁶USDA-ARS Corn Insects and Crop Genetics Research Unit, Ames, IA 50011, USA

*Corresponding author: Tel: +1 573 884 7873; Fax: +1 573 884 7850; Email: mary.schaeffer@ars.usda.gov

Submitted 17 December 2010; Revised 31 March 2011; Accepted 9 May 2011

First released in 1991 with the name MaizeDB, the Maize Genetics and Genomics Database, now MaizeGDB, celebrates its 20th anniversary this year. MaizeGDB has transitioned from a focus on comprehensive curation of the literature, genetic maps and stocks to a paradigm that accommodates the recent release of a reference maize genome sequence, multiple diverse maize genomes and sequence-based gene expression data sets. The MaizeGDB Team is relatively small, and relies heavily on the research community to provide data, nomenclature standards and most importantly, to recommend future directions, priorities and strategies. Key aspects of MaizeGDB's intimate interaction with the community are the co-location of curators with maize research groups in multiple locations across the USA as well as coordination with MaizeGDB's close partner, the Maize Genetics Cooperation—Stock Center. In this report, we describe how the MaizeGDB Team currently interacts with the maize research community and our plan for future interactions that will support updates to the functional and structural annotation of the B73 reference genome.

A brief history of MaizeDB

In maize, the release of the B73 reference genome sequence (1), coupled with advances in sequencing technologies, continues to generate massive quantities of data. MaizeGDB, the maize research community's Model Organism Database (MOD) is charged to integrate much of these data and to provide an access point that serves data representations [e.g. the MaizeGDB Genome Browser (2), descriptions of loci and gene models, etc.] to external data resources with connections to physical entities such as genetic stocks. In addition, the Maize Genome Sequencing Consortium has requested that MaizeGDB collect, document and disseminate researcher-contributed information to aid in continued genome assembly and annotation endeavors.

In MaizeGDB's initial incarnation (as MaizeDB; 1990–2000), comprehensive literature annotation was one of the central

foci of the database resource with the main goal being to capture experimentally confirmed gene functions and trait inheritance, along with extensive community data documenting genetic maps and molecular markers (3; *aka* mapping probes). Currently, the curation focus is to facilitate data integration of very large data sets and to provide insight into development of easy-to-use interfaces and data displays. The efforts toward data integration involve gene nomenclature considerations as well as ontology development and implementation. Literature annotation is an ongoing process, but at a greatly reduced amount per guidance provided by the 2004 MaizeDB to MaizeGDB Transition Steering Committee, http://www.maizegdb.org/steering_committee.php. Priority is given to manuscripts suggested by the MaizeGDB Editorial Board (approximately five journal articles per month; http://www.maizegdb.org/cgi-bin/editorial_board.cgi) and, as time permits, to newly sequenced genes with experimentally confirmed functions.

Although a comprehensive set of tools that permits data curation by the maize community has been available for many years (4), it has been very rarely utilized; a problem consistently reported by other genome database projects (5). In an effort to increase literature curation, implementation of Textpresso (6) is currently underway. Textpresso will serve both as a tool to facilitate curatorial activities, as well as a mechanism for researchers to search publications based on biological categories such as Gene Ontology terms (7), <http://www.geneontology.org/>. Additionally, we have partnered with a research team at Truman State University. This group, comprised of faculty in biology and computer science, engages undergraduates to perform Gene Ontology annotation of maize gene models, <http://sam.truman.edu/> (8). In cases where there is experimental data in the published literature, these annotations will be reviewed, incorporated into MaizeGDB and provided to Uniprot/SwissProt, which links to MaizeGDB. We envision this to become a model for similar community-based annotation efforts.

MaizeGDB and the community

The permanent staff for MaizeGDB is relatively small (five persons), and we rely heavily on the maize genetics community and other stakeholders, notably the Maize Genetics Executive Committee, <http://www.maizegdb.org/mgec.php>, the National Corn Growers Association (9) and the MaizeGDB Working Group (described below) for guidance and assistance. The maize community has a long history of cooperation that dates to the early part of the 20th century (10) and enthusiastically participates in community surveys to help in guiding MaizeGDB to address stakeholder needs (2). The maize community has been a major contributor to the NSF's Plant Genome Research Program's accomplishments from that program's very beginnings, including laying substantial groundwork early on toward sequencing the maize genome (11). While community planning discussions are held in open forums at the Annual Maize Genetics Conference (this year in its 53rd year), smaller, more focused meetings are also held to identify community needs and set priorities. The most recent occurred in 2007 at Allerton Park, Illinois, <http://www.maizegdb.org/AllertonReport.doc>. Allerton has historic interest, as it was also the site of the first maize meeting (Figure 1). At MaizeGDB, we provide extensive outreach and other services to support continued development of the maize community, which are described in previous reports (4, 12–14) and in the accompanying *Database* (Oxford) article (Harper *et al.*)

One way that we are able to enhance our interactions with the maize community is by the physical distribution of the MaizeGDB curation staff. The software development and database management team members are centralized

at Ames, Iowa (USA), in a building where several other database groups are also located: SoyBase (15), <http://www.soybase.org/>; PLEXdb (16), <http://www.plexdb.org/>; and PlantGDB (17), <http://www.plantgdb.org/>. Ames, IA is also the home of maize researchers working in various areas of research (18–25) and the North Central Regional Plant Introduction Station where maize breeding stocks are maintained as a part of the National Plant Germplasm System, http://www.ars.usda.gov/Main/site_main.htm?modecode=36-25-12-00. MaizeGDB curators, who also serve as outreach staff, (J.M.G., L.C.H. and M.L.S.) are located in areas where many maize geneticists are stationed: at the USDA-ARS Plant Gene Expression Center near the University of California, Berkeley (26–29), at the University of Arizona (30–33) and at the USDA-ARS/University of Missouri, Columbia (34–41). A closely related and collaborative project that also serves the maize genetics research community is the Maize Genetics Cooperation–Stock Center (42), <http://maizecoop.cropsci.uiuc.edu/>, located at Urbana, Illinois. The Stock Center uses MaizeGDB as the primary interface to data about mutants, stocks and phenotypes, provides physical materials to researchers, and has close contact with other maize research groups in that area (43,44). Curators interact closely among themselves by frequent email and phone calls, conference calls, in-person meetings and postings to a wiki for the MaizeGDB team.

In addition to close contact with the maize researchers, members of the MaizeGDB team are, or have been, part of the elected Maize Genetics Executive Committee (C.J.L. and M.L.S.); the Maize Genetics Nomenclature Committee (M.L.S.); the Maize Meeting Steering Committee (C.M.A. and M.L.S.); the Editorial Board the Maize Genetics Cooperation—Newsletter (M.L.S.); and the Corn Germplasm Committee (CGC), <http://www.ars-grin.gov/npgs/cgclist.html> (C.J.L. and M.L.S.). MaizeGDB hosts community websites, <http://www.maizegdb.org/cooperators.php>, for all of these groups, save the CGC, so that activities can be made accessible for public consumption. This enables members of the MaizeGDB Team to respond appropriately and quickly to community needs and to function as a clearinghouse for maize data, as well as the enforcement of proper nomenclature.

Formal interaction with the community occurs via our working group (WG). Established in 2006, the MaizeGDB Working Group is tasked with evaluating MaizeGDB's current status and recommending a course of action that will ensure that the MaizeGDB project follows the trajectory of maize research as closely as possible, providing a robust and timely source of data and analysis tools. The WG is composed of 10–12 members of the maize community, who are active in diverse areas of research and generally serve a term of 3 years (see 'Acknowledgments' section for current members of the WG). The WG meets at least once a year,



Figure 1. The maize genetics cooperation, 1928–2009. The top timeline depicts the establishment of a Stock Center and Newsletter in 1928, followed by funding for MaizeDB in 1990, and then the 2009 release of the B73 reference genome sequence. The leftmost inset shows the participants of the first formal Maize Genetics Meeting, at Allerton, IL in 1959 [photograph courtesy Earl Patterson and the Newsletter (57)]. The color photograph [courtesy anonymous photographer, and the Newsletter (58)] depicts the subset of the current maize research community that convened in 2007, at Allerton, IL to plan infrastructure needs. This 2007 meeting included representatives of MaizeGDB (CJL, MLS, Trent Seigfried).

preferably in person at a national meeting, or when circumstances/schedules do not permit, via online conferencing. After a brief presentation by the MaizeGDB staff describing recent accomplishments, the WG provides guidance on issues of special concern to MaizeGDB. Since its creation, the WG has provided critical guidance on topics including but not limited to: content development for the supporting 2007–2012 USDA-ARS Project Plan, selection of appropriate mechanisms to visualize and interact with the maize genome sequence, outlining a specific role for MaizeGDB in hosting maize genome sequence assemblies as well as structural and functional annotations, user interface updates, sequence-based expression data representations and opportunities for interactions with other projects.

Genome annotation in the past

Over the years, about 6000 functional genes described in the literature have been curated into MaizeGDB. However,

this count is dwarfed by the 32 540 gene models predicted for the B73 genome by the Maize Genome Sequencing Consortium (1), the 25 703 RefSeq cDNA-unigenes in GenBank (45), and the approximately 10 000 gene models currently thought not to exist in B73, but to be present in other inbred lines (35,46). These new loci (predicted gene models) can be integrated with the existing functionally defined genes in many different ways. One way to integrate these data is to build a tool that relies on molecular markers often sequenced, which are shared between the genetic and physical maps and documented at MaizeGDB. Based on the requests from our users, we have created the 'Locus Lookup Tool', http://www.maizegdb.org/cgi-bin/locus_lookup.cgi?id=, to implement this idea. Locus Lookup helps researchers with genetically mapped genes to identify the chromosomal window that contains their gene of interest (47). This tool can aid positional cloning efforts and ultimately connects a theoretical gene model with a biologically defined gene. It is currently one of our

most popular tools to access maize data. In other cases, cDNA sequences aligned to the genome assembly are also assigned to locus variations by manual curation, which can be used to link classical genetic information with the genome sequence. For example, the well-studied *b1* gene has a manually curated cDNA sequence accession, X57276, associated with the mutant allele, *B1-Peru*, and has been aligned by the PlantGDB pipeline to the assembled B73 genome: http://www.maizegdb.org/cgi-bin/display_seqrecord.cgi?id=X57276 (17). Our manual curation of sequence accessions is periodically shared with the NCBI and can be found along with many gene and marker names and synonyms on the NCBI gene records; for example, see the *b1* gene record: <http://www.ncbi.nlm.nih.gov/gene/542724>.

It is useful to understand the infrastructure that has been put in place or is under development at MaizeGDB to support annotation. The MaizeGDB Genome Browser is the centerpiece for MaizeGDB's transition to a sequence-centric paradigm (48). Within the Genome Browser, tracks (Table 1) are displayed with gene model information supplied by both the Maize Genome Sequencing Consortium, maizesequence.org and PlantGDB/ZmGDB, <http://www.plantgdb.org/ZmGDB> (17). In addition, we also accept data sets from community members that are mapped to the genomic sequence and serve them as genome browser tracks. When requested, we allow community-added information to be submitted but not displayed at MaizeGDB before publication. Glyphs displayed within tracks on the MaizeGDB Genome Browser link to additional information, stored at MaizeGDB and/or offsite.

In some cases when the data need to be recomputed on an updated genome sequence assembly, we ask for action on the part of the original submitter. One of these, PlantGDB, serves community annotations via Distributed Annotation Service (DAS) wherein the changes to gene models are stored at PlantGDB, and DAS web services are used to serve the data for display within the context of the MaizeGDB Genome Browser. We have created mechanisms to recompute the alignment of genomic features that we visualize in the tracks when new versions of the assemblies (i.e. pseudomolecules) become available. Feature annotations at MaizeGDB are often provided in files associated with peer-reviewed publications, usually as Supplementary Data, where the annotation process is described. In many of these cases, the underlying supporting data used to create each browser track have been integrated systematically at MaizeGDB [e.g. UniformMu (49), ISU IBM 2009 (50), IBM2 2008 Neighbors (33)] and can be used for updating annotation of tracks in future.

More recently, we have begun a partnership with PLEXdb to update gene expression data from NimbleGen arrays used to develop a maize gene atlas (51). MaizeGDB will redefine oligo probe sets for updates to gene models; and PLEXdb will re-normalize the raw expression data for the updated probe sets. Updated associations of gene models to standard ontologies for plant anatomy and development will be supplied to the Plant Ontology site, www.plantontology.org, where we have contributed to the initial development of this ontology, and have provided associations as part of our normal operating protocols for some years (52).

Table 1. Community provided annotation tracks for the MaizeGDB Genome Browser

Tracks ^a	Source ^b	External feature link ^c
<i>Ac/Ds</i>	Brutnell & Vollbrecht, via PlantGDB (23)	PlantGDB.org
UniformMU	McCarty (49)	
ISU IBM 2009	Schnable (50)	
IBM2 2008 Neighbors	Arizona Genomics Institute (33)	
Centromere	J.M. Jiang & G. Presting, unpublished data (55)	
antiCENH3-ChiP	J.M. Jiang & G. Presting, unpublished data (55)	
MIPS repeats, Gene models	Maize Genome Sequencing Consortium (1)	maizesequence.org
PLEXdb	PLEXdb (16)	PLEXdb.org
Structural annotation-community	PlantGDB (yrGate, 54)	PlantGDB.org
cDNA, EST,GSS, unique transcripts, GSS, gene models	PlantGDB (17)	PlantGDB.org
MAGI	P. Schnable (56)	magi.plantgenomics.iastate.edu
Mo17 SNP/Indel	D.M. Rokshar (unpublished data) http://www.jgi.doe.gov/	
Leaf transcriptome	Brutnell (57)	cbsuss03.tc.cornell.edu/cgi-bin/gbrowse/c3c4_pm

^aContent of browser tracks supplied by community.

^bPerson or group providing data with literature reference for data source in parentheses.

^cThe external data source for those cases where clicking on a track feature leads directly to an external data source.

The creation of B73 RefGen_v3 is in the works with a current estimate for its delivery slated for July of 2011 (Doreen H. Ware, personal communication). This version is anticipated to be the final assembly product to be provided by the Maize Genome Sequencing Consortium, which will be made available via MaizeGDB. Thereafter, the default view of the MaizeGDB Genome Browser is scheduled to be updated annually to an incremented release of the genome assembly by the MaizeGDB Team. FTP and BLAST (53) access to builds will be made available in advance of releasing new assemblies to allow research groups to align their sequence-indexed data well in advance of their full deployment via the MaizeGDB Genome Browser. Timely access is important for our community annotators (Table 1), but also to many other researchers. For example, in microarray and RNA_seq expression data, expression values are normalized over values for oligos corresponding to the current gene models. MaizeGDB currently accepts data aligned to the B73 Reference Genome Assembly (currently B73 RefGen_v2) and is committed to working with members of the community to integrate well-documented structural and functional genome annotation and to release new assemblies annually.

Genome annotation in the future

From community surveys by the Maize Genetics Executive Committee and community discussions at the 2010 Maize Genetics Conference, the maize community has identified, as its two top priorities, to have the maize B73 sequence assembly improved and annotated. At the request of the Maize Genetics Executive Committee, we are in the process of forming a Maize Genome Annotation Consortium. The MaizeGDB Team appreciates the importance of helping people work together toward a common goal and has recently posted guidelines on its website, <http://www.maizegdb.org/assembly.php>, to groups who plan to engage in sequence assembly and annotation.

These guidelines stipulate several elements for a fully successful collaboration with the maize community and with the MaizeGDB, the ultimate disseminator of the information. A key element is transparency, whereby projects make public their plans, standards and data delivery timelines at the initial stages. As data are delivered, detailed documentation for assembly and annotation should be provided. Annotations confirmed by experimental evidence should be clearly discriminated from those based purely on *in silico* analyses. Standard evidence codes, e.g. those stipulated for GO annotations, should be employed. To facilitate data integration at MaizeGDB, and also planning by all in the research community, data delivery dates should be made known in advance, although it is understood that there may be some changes from the initial timeline. MaizeGDB will create a mechanism to display progress

with the main idea to facilitate communication with the maize research community.

In addition, there should be a mechanism to interact with the maize community directly and with a single voice for the project. Maize researchers comprise a vibrant community with researchers at all levels in both the public and private sectors and as such an annotation project means different things to different people. A bidirectional means of communicating with the maize community should be deployed at the start of the project so that the maize community can both absorb and respond to new annotation information quickly. The goal is to provide all community members with the same information at the same time so that they can plan their research activities accordingly. This can be accomplished in many ways (quarterly e-newsletters, FAQs, blogs, social media, conferences, etc.) and all options should be considered so as to reach the largest number of stakeholders (i.e. the maize researchers).

Another key element is to capture genome assembly information from the community. Currently, individual researchers are generating excellent, lab-validated genetic markers and order/orientation information of sequence fragments within BAC's. Researchers are usually willing to share this information freely, but currently, there is no robust means to capture it. There should be an easy way through a web interface for researchers to submit data. All annotation submitted by community members should be vetted manually by expert annotators, and subsequently incorporated into the assembly, with an indication of who provided the data. It is expected that while there will be comparatively little data entering the assembly process in this way, these data would be of very high quality.

As with genome assembly, researchers currently have high-quality structural and functional annotation for their genes of interest, both stored on lab computers and documented in publications. Researchers should be provided with tools to improve structural and functional annotation information that can then be integrated into the larger project's outcomes. These same tools could also be leveraged for classroom teaching. The ZmGDB/PlantGDB yrGATE (54) system and iPlant's DNA Subway, <http://dnasubway.iplantcollaborative.org>, are good examples of the sort of interface that could serve both groups.

We strongly encourage planning a workshop for education, outreach and training in all aspects of annotation and thereby increase community understanding and involvement in the annotation of the maize genome. One obvious way to involve the community would be to contact the Maize Genetics Conference Steering Committee, http://www.maizegdb.org/maize_meeting/, about getting your message out at the Maize Meeting.

Summary

The maize community is both excited about and prepared to begin work toward annotating the maize genome, both in B73 and in other inbred lines (34, 46). We at MaizeGDB are looking forward to continuing our partnership with the community to provide informatics and organizational support for ongoing research activities.

Acknowledgements

Guidance is generously provided by the *MaizeGDB Working group*: M. Pop (Chair), A. Barkin, O. Hoekenga, A. Lamblin, T. Lubberstedt, K. McGinnis, L. Mueller, M. Sachs, P. Schnable, A. Sylvester; *The Maize Genetics Executive Committee* (T. Brutnell (Chair), P. Schnable, M. Sachs, V. Walbot, W. Tracy, S. Wessler, J. Bennetzen, B. Boston, E. Buckler, and C. Lawrence); *the Maize Nomenclature Committee* (H. Dooner, (Chair), T. Brutnell, V. Chandler, C. Hannah, T. Kellogg, M. Sachs, M. Scanlon, M. Schaeffer, and P. Stinard); and the National Corn Growers Association (www.ncga.org). We thank the organizers of the Biocuration 2010 conference (<http://hin.jp/biocuration2010/index.html>) for inviting MLS to present part of this report as an oral presentation, and additionally, *Nature Precedings*, for posting PowerPoint slides from that presentation, <http://precedings.nature.com/documents/5258/version/1>.

Funding

United States Department of Agriculture Agricultural Research Service (project numbers 3625-21000-051-00, 3622-21000-034-00); the National Science Foundation (grant numbers DBI 0734804, IOS 0701405, IOS 0703273, IOS 0965380); The Center for Maize and Wheat Improvement; United States Agency for international Development. Funding for open access charge: United States Department of Agriculture Agricultural Research Service (project number 3622-21000-034-00).

Conflict of interest. None declared.

References

1. Schnable, P.S., Ware, D., Fulton, R.S. et al. (2009) The B73 maize genome: complexity, diversity, and dynamics. *Science*, **326**, 1112–1115.
2. Sen, T.Z., Harper, L.C., Schaeffer, M.L. et al. (2010) Choosing a genome browser for a Model Organism Database: surveying the Maize community. *Database*, **2010**, doi:10.1093/database/baq007.
3. Polacco, M.L. and Coe, E. (1999) MaizeDB: The maize genome database. In: Letovsky, S.I. (ed). *Bioinformatics: Databases and Systems*. Kluwer Academic Publishers, Norwell, MA, pp. 151–162.

4. Lawrence, C.J., Seigfried, T.E. and Brendel, V. (2005) The maize genetics and genomics database. The community resource for access to diverse maize data. *Plant Physiol.*, **138**, 55–58.
5. Howe, D., Costanzo, M., Fey, P. et al. (2008) Big data: The future of biocuration. *Nature*, **455**, 47–50.
6. Muller, H.M., Kenny, E.E. and Sternberg, P.W. (2004) Textpresso: an ontology-based information retrieval and extraction system for biological literature. *PLoS Biol.*, **2**, e309.
7. Ashburner, M., Ball, C.A., Blake, J.A. et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.*, **25**, 25–29.
8. Buckner, B., Beck, J., Browning, K. et al. (2007) Involving undergraduates in the annotation and analysis of global gene expression studies: creation of a maize shoot apical meristem expression database. *Genetics*, **176**, 741–747.
9. Coe, E.H. (1998) Potentials of the National Corn Genome Initiative. *Proc. Natl Acad. Sci. USA*, **95**, 2029–2032.
10. Kass, L.B., Bonneuil, C. and Coe, E.H. Jr. (2005) Cornfests, cornfabs and cooperation: the origins and beginnings of the maize genetics cooperation news letter. *Genetics*, **169**, 1787–1797.
11. Coe, E. and Schaeffer, M. (2005) Genetic, physical, maps, and database resources for maize. *Maydica*, **50**, 285–303.
12. Lawrence, C.J. (2007) MaizeGDB. *Methods Mol. Biol.*, **406**, 331–345.
13. Lawrence, C.J., Harper, L.C., Schaeffer, M.L. et al. (2008) MaizeGDB: The maize model organism database for basic, translational, and applied research. *Int. J. Plant Genomics*, **2008**, Article ID 496957.
14. Lawrence, C.J., Schaeffer, M.L., Seigfried, T.E. et al. (2007) MaizeGDB's new data types, resources and activities. *Nucleic Acids Res.*, **35**, D895–D900.
15. Grant, D., Nelson, R.T., Cannon, S.B. et al. SoyBase, the USDA-ARS soybean genetics and genomics database. *Nucleic Acids Res.*, **38**, D843–D846.
16. Wise, R.P., Caldo, R.A., Hong, L. et al. (2007) BarleyBase/PLEXdb. *Methods Mol. Biol.*, **406**, 347–363.
17. Duvick, J., Fu, A., Muppirala, U. et al. (2008) PlantGDB: a resource for comparative plant genomics. *Nucleic Acids Res.*, **36**, D959–D965.
18. Salas Fernandez, M.G., Becraft, P.W., Yin, Y. et al. (2009) From dwarves to giants? Plant height manipulation for biomass yield. *Trends Plant Sci.*, **14**, 454–461.
19. Yi, G., Luth, D., Goodman, T.D. et al. (2009) High-throughput linkage analysis of *Mutator* insertion sites in maize. *Plant J.*, **58**, 883–892.
20. Danilevskaya, O.N., Meng, X., Selinger, D.A. et al. (2008) Involvement of the MADS-box gene *ZMM4* in floral induction and inflorescence development in maize. *Plant Physiol.*, **147**, 2054–2069.
21. Hasjim, J., Srichuwong, S., Scott, M.P. et al. (2009) Kernel composition, starch structure, and enzyme digestibility of *opaque-2* maize and quality protein maize. *J. Agric. Food Chem.*, **57**, 2049–2055.
22. Jiang, H., Lio, J., Blanco, M. et al. Resistant-starch formation in high-amylose maize starch during kernel development. *J. Agric. Food Chem.*, **58**, 8043–8047.
23. Ohtsu, K., Smith, M.B., Emrich, S.J. et al. (2007) Global gene expression analysis of the shoot apical meristem of maize (*Zea mays* L.). *Plant J.*, **52**, 391–404.
24. Vollbrecht, E., Duvick, J., Schares, J.P. et al. (2010) Genome-wide distribution of transposed *Dissociation* elements in maize. *Plant Cell*, **22**, 1667–1685.
25. Yu, C., Zhang, J., Pulletikurti, V. et al. (2010) Spatial configuration of transposable element *Ac* termini affects their ability to induce chromosomal breakage in maize. *Plant Cell*, **22**, 744–754.

26. Chuck, G., Whipple, C., Jackson, D. et al. (2010) The maize SBP-box transcription factor encoded by *tasselsheath4* regulates bract development and the establishment of meristem boundaries. *Development*, **137**, 1243–1250.
27. Kelliher, T. and Walbot, V. (2011) Emergence and patterning of the five cell types of the Zea mays anther locule. *Dev. Biol.*, **350**, 32–49.
28. Pawlowski, W.P., Wang, C.J., Golubovskaya, I.N. et al. (2009) Maize *AMEIOTIC1* is essential for multiple early meiotic processes and likely required for the initiation of meiosis. *Proc. Natl Acad. Sci. USA*, **106**, 3603–3608.
29. Woodhouse, M.R., Schnable, J.C., Pedersen, B.S. et al. (2010) Following tetraploidy in maize, a short deletion mechanism removed genes preferentially from one of the two homologs. *PLoS Biol.*, **8**, e1000409.
30. Arteaga-Vazquez, M., Sidorenko, L., Rabanal, F.A. et al. RNA-mediated transcommunication can establish paramutation at the *b1* locus in maize. *Proc. Natl Acad. Sci. USA*, **107**, 12986–12991.
31. Chung, T., Wang, D., Kim, C.S. et al. (2009) Plant *SMU-1* and *SMU-2* homologues regulate pre-mRNA splicing and multiple aspects of development. *Plant Physiol.*, **151**, 1498–1512.
32. Soderlund, C., Descour, A., Kudrna, D. et al. (2009) Sequencing, mapping, and analysis of 27 455 maize full-length cDNAs. *PLoS Genet.*, **5**, e1000740.
33. Wei, F., Zhang, J., Zhou, S. et al. (2009) The physical and genetic framework of the maize B73 genome. *PLoS Genet.*, **5**, e1000715.
34. Flint-Garcia, S.A., Buckler, E.S., Tiffin, P. et al. (2009) Heterosis is prevalent for multiple traits in diverse maize germplasm. *PLoS One*, **4**, e7433.
35. Gore, M.A., Chia, J.M., Elshire, R.J. et al. (2009) A first-generation haplotype map of maize. *Science*, **326**, 1115–1117.
36. Han, F., Gao, Z. and Birchler, J.A. (2009) Reactivation of an inactive centromere reveals epigenetic and structural components for centromere specification in maize. *Plant Cell*, **21**, 1929–1939.
37. Hibbard, B.E., Meihls, L.N., Ellersieck, M.R. et al. Density-dependent and density-independent mortality of the western corn rootworm: impact on dose calculations of rootworm-resistant Bt corn. *J. Econ. Entomol.*, **103**, 77–84.
38. McMullen, M.D., Kresovich, S., Villeda, H.S. et al. (2009) Genetic properties of the maize nested association mapping population. *Science*, **325**, 737–740.
39. Slewinski, T.L., Garg, A., Johal, G.S. et al. (2010) Maize *SUT1* functions in phloem loading. *Plant Signal Behav.*, **5**, 687–690.
40. Wu, X., Skirpan, A. and McSteen, P. (2009) *Suppressor of sessile spikelets1* functions in the ramosa pathway controlling meristem determinacy in maize. *Plant Physiol.*, **149**, 205–219.
41. Yamaguchi, M. and Sharp, R.E. (2010) Complexity and coordination of root growth at low water potentials: recent advances from transcriptomic and proteomic analyses. *Plant Cell Environ.*, **33**, 590–603.
42. Scholl, R., Sachs, M.M. and Ware, D. (2003) Maintaining collections of mutants for plant functional genomics. *Methods Mol. Biol.*, **236**, 311–326.
43. Harjes, C.E., Rocheford, T.R., Bai, L. et al. (2008) Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. *Science*, **319**, 330–333.
44. Upadaya, N., da Silva, H.S., Bohn, M.O. et al. (2006) Genetic and QTL analysis of maize tassel and ear inflorescence architecture. *Theor. Appl. Genet.*, **112**, 592–606.
45. Wheeler, D.L., Smith-White, B., Chetvernin, V. et al. (2005) Plant genome resources at the national center for biotechnology information. *Plant Physiol.*, **138**, 1280–1288.
46. Lai, J., Li, R., Xu, X. et al. (2010) Genome-wide patterns of genetic variation among elite maize inbred lines. *Nat. Genet.*, **42**, 1027–1030.
47. Andorf, C.M., Lawrence, C.J., Harper, L.C. et al. (2010) The Locus Lookup tool at MaizeGDB: identification of genomic regions in maize by integrating sequence information with physical and genetic maps. *Bioinformatics*, **26**, 434–436.
48. Sen, T.Z., Andorf, C.M., Schaeffer, M.L. et al. (2009) MaizeGDB becomes 'sequence-centric'. *Database*, 2009, doi:10.1093/database/bap020.
49. McCarty, D.R., Settles, A.M., Suzuki, M. et al. (2005) Steady-state transposon mutagenesis in inbred maize. *Plant J.*, **44**, 52–61.
50. Fu, Y., Wen, T.J., Ronin, Y.I. et al. (2006) Genetic dissection of inter-mated recombinant inbred lines using a new genetic map of maize. *Genetics*, **174**, 1671–1683.
51. Sekhon, R.S., Lin, H., Childs, K.L. et al. (2011) Genome-wide atlas of transcription during maize development. *Plant J.*, March 9, 2009 (epub ahead of print; doi: 10.1111/j.1365-3113X.2011.04527.x).
52. Avraham, S., Tung, C.W., Illis, K. et al. (2008) The Plant Ontology Database: a community resource for plant structure and developmental stages controlled vocabulary and annotations. *Nucleic Acids Res.*, **36**, D449–D454.
53. Altschul, S.F., Gish, W., Miller, W. et al. (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403–410.
54. Wilkerson, M.D., Schlueter, S.D. and Brendel, V. (2006) yrGATE: a web-based gene-structure annotation tool for the identification and dissemination of eukaryotic genes. *Genome Biol.*, **7**, R58.
55. Luce, A.C., Sharma, A., Mollere, O.S. et al. (2006) Precise centromere mapping using a combination of repeat junction markers and chromatin immunoprecipitation-polymerase chain reaction. *Genetics*, **174**, 1057–1061.
56. Fu, Y., Emrich, S.J., Guo, L. et al. (2005) Quality assessment of maize assembled genomic islands (MAGIs) and large-scale experimental verification of predicted genes. *Proc. Natl Acad. Sci. USA*, **102**, 12282–12287.
57. Li, P., Ponnala, L., Gandotra, N. et al. (2010) The developmental dynamics of the maize leaf transcriptome. *Nat. Genet.*, **42**, 1060–1067.
58. Birchler, J.A., Coe, E.H., Schaeffer, M.L. (eds.) First Maize Meeting (2008) *Maize Genetics Coop. Newsl.* **82**, 39.
59. Birchler, J.A., Coe, E.H., Schaeffer, M.L. (eds.) Random photographs from past maize meetings (2008) *Maize Genetics Coop. Newsl.* **82**, 44.