

***The Future of Maize Genetics  
Planning for the Sequenced Genome Era***

**A Maize Genetics Community Retreat  
Allerton Park and Conference Center  
March 20 - 22, 2007**

***Support provided by the USDA-CSREES-NRI-Plant Genome and the NSF Plant Genome  
Research Programs and the University of Illinois.***

Executive Summary	page 2
Detailed Report	3
Introduction	3
Challenging Unanswered Questions	3
Current Resources and Future Needs	4
<i>Sequencing and annotation</i>	5
<i>Centralized databases</i>	5
<i>Transformation improvements</i>	6
<i>Reverse genetics resources</i>	7
<i>Expression profiling</i>	8
<i>Phenomics</i>	9
<i>Community membership</i>	10
Summary and Next Steps	12

**Executive Summary: *The Future of Maize Genetics - Planning for the Sequenced Genome Era* A Maize Genetics Community Retreat at Allerton, March 20 - 22, 2007**

Leaders in the maize community met for a two-day retreat to discuss the strengths, challenges, and initiatives that define the future of maize research. To guide strategic planning, the community first identified key questions in biology that can be best answered using maize as a model system. These biological questions were considered relative to the overarching goal of understanding the genetic basis of traits in maize. The research directions defined below and the plans to achieve the goals serve as the foundation for basic research and provide the tools for improving food, fuel, and crop yields in a changing environment.

**Key biological issues define our research goals and directions:**

- Maize is the pre-eminent model for studying genome evolution and trait variation due to its unsurpassed natural diversity, genome duplication history and range of adaptations.
- Because adaptation is critical to agriculture, maize research will continue to be a model for understanding the basis of genetic interactions with the environment.
- Study of maize heterosis will provide key information about how genes and alleles interact.
- Maize will continue to lead in the area of epigenetics. Imprinting, paramutation and transposons were discovered in maize and are readily studied with color markers.
- Maize is positioned as a leading model for developing cellulosic biofuels for the future.
- Maize is a model for the study of development and physiology of unique traits such as C4 photosynthesis, a persistent endosperm, inflorescence structure, etc.
- Maize cytogenetics is highly advanced and continues to provide tools for understanding mechanisms of meiosis and for developing the potential of chromosome manipulation.

**Community resources will help achieve our research goals:**

To advance these research areas, community resources must be created or strengthened. Short and long term planning will help leverage the sequenced maize genome and will position maize as a model for tool development and for hypothesis-driven and translational research.

*Short Term Planning* (Expect substantial progress in the next one to three years):

1. Current sequencing/annotation will be completed and additional map-based sequencing efforts initiated.
2. Centralized databases with increased funding are needed now.
3. Indexed reverse genetic resources need to be finalized and will accelerate many areas of research. Current mutagenesis libraries should be indexed with new technologies.
4. Expression platforms and informatic tools should be selected and developed.
5. Transformation capacity should be enhanced in the public sector. Continuous support mechanisms for public transformation need to be resolved.

*Long Term Planning* (Start now, with sustained efforts over the next decade)

1. Databases and stock center capacity will be enhanced, coordinated and supported.
2. Map-based sequences of other inbreds, races, and teosintes will be available.
3. A phenomics initiative will be underway, with large scale and multi-dimensional phenotyping capabilities for multiple environments available for the entire community.
4. The maize scientific community will be large, diverse, well-trained, and interactive.

## **Detailed Report: *The Future of Maize Genetics - Planning for the Sequenced Genome Era*, A Maize Genetics Community Retreat at Allerton, March 20 - 22, 2007**

### **Introduction**

The maize genetics community held a two-day retreat to discuss the future of maize research in the post-genomics era. The meeting included principal investigators representing approximately 60 labs from universities, colleges, industry, and USDA-ARS. Participants were primarily from the U.S., but representatives from the United Kingdom, France, and Germany were also present. Staff from ARS, NRI, NSF, DOE, and NCGA participated in discussion on the second day of the retreat.

The goal of the meeting was to develop a strategic plan for the future of maize genetics research. To guide strategic planning, the community first identified important questions in biology that can be best answered using maize as a model system. These biological questions were considered within the context of an overarching goal to understand the genetic basis of traits in maize -- traits that are the foundation for improving food, fuel, and fiber crop yields in a changing environment. Second, the community considered the current research capacity for answering these biological questions and also explored how to translate the answers to practical outcomes. It was noted that recent research accomplishments have opened many new avenues of investigation. For example, advances in genetic mapping technology have enhanced functional genomics so that gene functions can be discovered using a single population of plants and inexpensive sequencing technology. This groundwork will allow rapid establishment of productive genomics programs in related agronomically important species and will advance research in biofuels. Following the plan presented here, maize researchers will be able to accelerate the rate of discovery considerably, delivering an expanding knowledge base for the needs of breeders, for biotechnology industry and for continued basic research.

### **Challenging unanswered questions in biology best addressed by research on maize**

Discoveries in the current genomic era of biology have generated new questions and enabled new approaches to long-standing questions. The maize community discussed these issues in broad terms and then focused on the subset of questions that could clearly be addressed best in maize due to its unique development and biology, its genetic and evolutionary history and its genome architecture.

- How is genomic diversity maintained, and how does it change during evolution?
- What is the underlying molecular genetic basis for specific traits in a species?
- Can we use maize to predict what genes will regulate plant growth in related species?
- Can natural variation provide information to develop novel breeding traits?
- What drives genome evolution, and how are these processes impacted by interaction with the environment?
- What is the genetic, molecular and physiological basis of hybrid vigor (heterosis)?
- What are the impacts of sequence-independent inheritance (epigenetics) on the growth development and evolution of maize?
- How does cytogenetic variation impact genome architecture, agronomic traits, and plant breeding efforts?

Maize is particularly useful to answer questions about genome evolution, genetic diversity and selection because allelic variation in maize is greater than in any other eukaryote. Also, due to its duplication history, maize is a model monocot for understanding evolutionary mechanisms that surround genome duplication events. Gene content and order varies considerably between maize lines, reflecting local transposon-mediated rearrangements and tandem duplications. This great genetic diversity translates into differences in phenotype and variation in how maize interacts with the environment. The genetic diversity also provides a rich toolset for the study of heterosis. With the ability to compare maize alleles with those of wild *Zea* accessions (teosintes), maize also provides an excellent species to study selection. An understanding of selection will allow researchers to harness existing diversity for advancing biological understanding and crop improvement.

Maize is well suited to study epigenetics because of the accessible phenomena associated with transposition, imprinting, and paramutation, three processes that were first identified in maize. The study of epigenetics is facilitated by the separate male and female flowers, which simplifies the process of conducting controlled pollinations. Maize has a rich collection of active transposons in the genome and color markers that are simple to score. The large size of the plant allows researchers to sample specific tissues at distinct time points from a single individual. Furthermore, epigenetic events confer heritable phenotypes, and can therefore provide direct information for crop improvement, placing maize at the forefront of translating basic research for the agronomic community.

Maize is a genetic model for other grasses with its rich collection of mutants, genetic diversity and ease of moving between phenotype and genotype. Information from maize can be easily translated to other important, less tractable members of the grass family. For example, maize is a member of the Andropogoneae, and thus is closely related to other energy crops such as *Miscanthus*, switchgrass, sorghum, and sugarcane. Knowledge of cell wall synthesis and degradation can be obtained in maize and then transferred to these potential crops for which few genetic resources are available.

*Arabidopsis thaliana* has been a model for understanding principles behind growth and development, but some key developmental, cellular, and physiological processes do not occur in *Arabidopsis*. Biologically and economically important features such as C4 photosynthesis, a persistent endosperm, phase dependent epidermal differentiation, complex inflorescence structure, and sex determination are best studied in maize.

Finally, maize has been central to research in cytogenetics and continues to provide cutting edge information about genome architecture. Tools developed cytogenetically will be useful for future chromosome manipulations, which can benefit both basic and applied research. The combination of easily analyzed chromosomes, meiotic mutants, well-studied segregation phenomena, and increasingly sophisticated cytogenetic tools continues to position maize as a model system for this area of research.

### **Current Tools, Resources and Future Needs**

A major goal of the Allerton Retreat was to assess the current state of research capabilities as the maize B73 genome becomes available. It is clear that fully sequenced genomes have revolutionized the corresponding research communities. With long term planning, we can learn from these past experiences and develop the tools and capacity to optimize and fully leverage the value of a sequenced maize genome. Topics that were

considered most imperative to achieve this goal are summarized here and represent the starting point for further discussion.

***Annotation of a fully sequenced B73 genome and additional genome sequencing is the foundation for future research***

The maize B73 genome is currently being sequenced using a minimal tiling BAC approach with full display at <http://www.maizesequence.org> (a project site that will exist during the sequencing project’s funding period). The sequenced genome promises to revolutionize maize research. With annotation, it will be the foundation upon which complementary resources and activities such as reverse genetics and phenomics will be built. To achieve this promise, a fully annotated, accessible, and centralized sequence database will be essential because all additional resources depend upon robust integration of sequence information. The sequence project site must be transitioned into a community-based permanent platform that will have robust long-term support. The community expressed the desire that **MaizeGDB should become the centralized sequence resource soon after the genome is complete (2009-2010).**

The maize community recognizes that nearly all relevant questions posed here will need sequence information beyond the annotated B73 genome. **Genomic sequence from additional maize lines is essential to advance crop improvements and to exploit maize for its unparalleled strength as a model system and as a fuel, food and fiber resource worldwide.** Developing map-based sequence information of additional genomes was identified as a priority due to the unique duplication history of the maize genome and due to the exceptional haplotype variability among inbred lines. A physical map from a second inbred line, and ultimately multiple lines, was considered important. Post-Allerton follow-up discussion will allow for continued assessment about the most efficient and effective way to accomplish synthesis of sequence information from multiple genomes.

***Databases need to be centralized***

Continued assessment and coordination of data deposition is essential to all advances in maize research. Currently, various types of plant database resources exist and are utilized by maize researchers, including Model Organism Databases (MODs), Clade Oriented Databases (CODs), Automatic Annotation Shops (AA), Static Repositories, and Laboratory Information Management Systems (LIMS; a category that includes coordinated project databases). A (non-exhaustive) list of databases used by maize researchers includes:

<b>Resource Type</b>	<b>Funding Agency</b>	<b>Resource</b>	<b>Website</b>
MOD	USDA-ARS	MaizeGDB*	<a href="http://www.maizegdb.org">http://www.maizegdb.org</a>
MOD	NSF	TAIR	<a href="http://www.arabidopsis.org/">http://www.arabidopsis.org/</a>
COD	NSF, USDA-ARS	Gramene	<a href="http://www.gramene.org">http://www.gramene.org</a>
COD	USDA-ARS	GrainGenes	<a href="http://wheat.pw.usda.gov">http://wheat.pw.usda.gov</a>
COD/AA/LIMS	NSF	PlantGDB	<a href="http://www.plantgdb.org">http://www.plantgdb.org</a>
COD/AA/LIMS	NSF, USDA-ARS	PLEXdb	<a href="http://plexdb.org">http://plexdb.org</a>
AA	NSF	TIGR	<a href="http://www.tigr.org/">http://www.tigr.org/</a>
AA/LIMS	NSF/USDA/DOE	MGSC’s Maize Genome Browser*	<a href="http://www.maizesequence.org">http://www.maizesequence.org</a>
AA/LIMS	NSF	MAGI*	<a href="http://www.plantgenomics.iastate.edu/maize/">http://www.plantgenomics.iastate.edu/maize/</a>
AA/LIMS	NSF	FPC-maize*	<a href="http://www.genome.arizona.edu/fpc/maize/">http://www.genome.arizona.edu/fpc/maize/</a>

Static	NIH	NCBI	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>
Static	NIH	UniProt	<a href="http://www.pir.uniprot.org/">http://www.pir.uniprot.org/</a>
Static/LIMS	USDA-ARS	GRIN	<a href="http://www.ars-grin.gov/">http://www.ars-grin.gov/</a>
LIMS	NSF	Panzea*	<a href="http://www.panzea.org/">http://www.panzea.org/</a>
LIMS	NSF	ChromDB	<a href="http://www.chromdb.org/">http://www.chromdb.org/</a>

\* indicates a maize-specific resource

MaizeGDB is of particular interest because it is the MOD for maize. The MaizeGDB website serves biological information about the crop plant *Zea mays* ssp. *mays*. Genetic, genomic, sequence, gene product, functional characterization, literature reference, and person/organization contact information are among the datatypes accessible through MaizeGDB. Based upon community evaluation and input, MaizeGDB will continue to focus on the following areas of concentration over the next five years: 1) integration of new maize genetic and genomic data into the database, including expansion of phenotype data and tools, 2) expansion of structural and genetic map sets, 3) access to gene models calculated by leading gene structure prediction groups through the MaizeGDB interface, and 4) support of community services such as coordinating the Maize Meeting, MGEC Elections, Polls, etc.

Long-term support for MaizeGDB from USDA-ARS was recognized; however, **new and creative funding mechanisms are required now to provide sufficient resources to exploit a fully sequenced genome**. The fact that maize will become a model genome for other complex grass genomes necessitates even more careful planning. Coordination with Gramene is critical to success. With additional resources, MaizeGDB should be able to integrate project data from diverse studies, keep gene function data current, oversee community curation, and carry out gene and plant ontology as well as metabolic pathway curation. To improve access to maize sequence data, resources that integrate various gene models and annotation sets must be made available to MaizeGDB. Complex datasets from federally funded projects should be deposited into MaizeGDB. However, collaborations should be established between MaizeGDB and the researchers who develop these complex datasets to insure efficient and cost-effective data flow directly into MaizeGDB. For the other database resources listed in the above table, recommendations from this group are in agreement with those cited by the Plant Database Working Group (see <http://www.maizegdb.org/PDBNeeds.pdf>).

### ***Transformation technology needs to be advanced and costs reduced***

Improved maize transformation resources remain one of the highest priorities for the community. A sequenced maize genome will continue to drive research hypotheses that require direct testing in transgenic plants. Furthermore, transformation capabilities will bridge the gap between basic and applied research. To achieve these goals, **several critical needs were identified: an increased capacity for public sector maize transformation, improved transformation of diverse lines and reduced transformation costs**. Improved regulatory transfer would also facilitate progress and communication among researchers. The community recommends cohesive action to evaluate ways to improve regulatory compliance within current and changing Federal guidelines.

Public researchers currently produce transgenic maize primarily by outsourcing to the Plant Transformation Facility (PTF) at Iowa State University or through facilities at their own institutions. These centers constitute a critical and reliable resource. The success and demand on the PTF clearly validates the ever-increasing need for maize transformation in

the public sector. Costs remain higher than industry, however, reflecting both industry technologies that are unavailable to the public and differences in production scale. Thus, increased transformation capacity, properly implemented, will correlate with reduced costs per transgenic event.

There are several major limitations to capacity building in the public sector. First, more trained transformation experts are essential to insure quality outcomes. Second, more facilities, particularly greenhouses, are also necessary to grow transgenics to seed. Increased funding would be required to devote more resources to existing facilities, either concentrated in one main location or in multiple, collaborating centers that would allow for standardization of genotypes and transformation vectors. Third, reliable transformation of multiple genotypes is needed to reduce the time frame for post-transformation analysis by one or two years for every project.

In addition to improving the production pipeline, continued research is essential to advance methods of transformation. Ideally, this should be facilitated through dialogue with industry to address any bridgeable gaps that might exist between public and private sector methodologies. **Major investments should be made in training and facility improvement. A transformation task force that includes academic and industry representatives should be formed to facilitate this goal.**

Efforts to streamline the APHIS notification process would be beneficial. Such efforts could be accomplished within Federal guidelines. For example, the community could develop a standardized operating procedure (SOP) for transgenic lines so that users can quickly and consistently provide the required information. In particular, common notification requests for frequently used transgenic resources could be standardized by the community, then communicated to APHIS. This cooperation between the maize community and APHIS regulators would help researchers for whom regulatory compliance can be prohibitive, such as researchers at smaller institutions or researchers who experiment infrequently with transgenic maize.

Similar to the challenge of regulatory compliance, many researchers do not have the infrastructure to grow transgenic events to seed. Furthermore, in the future integrated genomics world, it is essential that all researchers should be able to navigate between *Arabidopsis* and maize to conduct transgenic experiments. Most *Arabidopsis* researchers lack both experience and facilities to carry out the intensive aspects of maize transformation. Multiple centralized field sites across the US dedicated to growing transgenic plants would both facilitate compliance and enable more researchers from diverse institutions to use this critical technology.

The community noted the importance of communication with the public sector to publicize the value of transgenic maize research. One mechanism to do this would be for qualified representatives to communicate directly with reporters or media outlets that are in place at most institutions, to widely publicize our message.

### ***Reverse genetics resources need to be expanded***

A sequence-indexed collection of mutations is essential for researchers to exploit the genome sequence fully. It was noted that multiple mutagens are necessary to insure broad coverage of the genome and generate a range of allelic lesions. These would include transposon insertions, small deletions, and point mutations. **It is imperative that these lines be accessible through a community web browser to facilitate dissemination of the**

**resource.** Training in the use of the resource should be an essential and embedded component of dissemination. This collection should be searchable by BLAST, browsable, and linked to readily available seed stocks. **The sequence-indexing of transposon collections needs to be on validated germinal alleles so that seed are available for the community to advance the study of identified mutations.**

To date, several populations have been developed for forward and reverse genetics in maize inbred lines. The use of inbred materials greatly facilitates phenotypic analysis in near-isogenic lines and should be given strong consideration in population development. This is particularly relevant for maize, where a long generation time limits most researchers to propagating at most two generations/year. Large Uniform*Mu*, *Ac/Ds* and TILLING populations have been developed in the W22 inbred. The existence of these W22 populations provides a case for sequencing the W22 genome, which potentially could be one of the choices for a second physical map. The maize community considers that further discussion is needed to come to consensus about sequencing plans after the first-stage completion of B73.

TILLING populations have also been developed in B73. *Mutator* and *Ac/Ds* populations are nearly completed in B73 as an effort to exploit the genome sequence and provide greater accessibility for researchers across a broader geographic distribution. To achieve the goal of near-saturation mutagenesis (95% chance of a disruption in any given gene) additional line development is essential.

A number of approaches were discussed for chemical, radiation and insertional mutagenesis. There was much excitement over the potential for 454 and Solexa sequencing technologies to deliver quickly a near-saturation collection of *Mutator* insertions. Several *Mutator* populations exist with high copy number germinal *Mu* insertions. It was estimated that over 300,000 *Mu* insertions could be rapidly sequenced from Uniform*Mu* from the McCarty lab, and additional populations from the Schnable and Martienssen labs could provide similar levels of coverage. It was noted that a minimal input of resources could help accomplish the task of generating these essential resources.

The possibility of increasing *Ac/Ds* populations and developing fast neutron populations was also discussed to complement the non-transgenic *Mutator* populations. For instance, over 30% of maize genes are represented in tandem duplications, suggesting that a large number of potentially redundant paralogs are present in the maize genome; because many single gene mutations cause a phenotype, detailed analysis of locally duplicated genes in maize will address a key general question in biology, namely the mechanisms that permit subfunctionalization of duplicated genes. The task of recombining single gene insertions in tightly linked paralogs to create double mutant stocks is daunting and unlikely to succeed without a strong genetic selection. *Ac/Ds* can be used to sequentially mutagenize tandem gene clusters providing a resource to define functions for a sizeable fraction of the maize genes. Similarly, fast neutron mutagenesis programs will result in a range of deletion sizes, some of which will encompass multiple adjacent genes. Detailed genetic analysis of a locus is greatly facilitated by using an allelic series of mutants wherever possible. **Increasing effort in generating, expanding and integrating the data from these populations now will pay huge dividends in the coming years.**

### ***Better access to gene expression profiling tools and datasets is needed***

A number of platforms presently exist in the maize community for expression analysis. The 44K Agilent and 46K NSF-Arizona long oligo arrays, shoot apical meristem (SAM) cDNA arrays, and a first-generation 18K Affymetrix GeneChip are publicly available. However, the fully sequenced maize genome offers the opportunity to begin conducting gene expression profiling experiments using “all genes” platforms. Moreover, recent advances in cost effective deep sequencing (e.g. Solexa, NextGen) might yet provide another alternative for expression profiling in different tissues and variants including their microRNAs and alternative splice products. The maize community will continue discussion about which platforms will be best and what toolsets need to be developed to insure long-term utility of datasets generated by expression profiling. Tools must be developed that allow datasets to be browsed, queried, visualized, meta-analyzed and linked to the physical and genetic maps of maize. Development of cost-effective platforms is also paramount. **These ambitious goals will require substantial database efforts and funding, but are absolutely critical to optimize use of these cost-intensive datasets.**

Consistent with the emergence of the genome sequence, two “all genes” platforms are currently being designed by collaborative efforts of industry, maize biologists and informaticians, including an Affymetrix 100K GeneChip and an Agilent 105K *in situ* synthesized glass slide array. **Community input into the design of expression profiling platforms continues to be a high priority to the maize community.** The Agilent arrays will allow community input via customization, which is facilitated by its flexible format. A database of sequences, customizable formats and designs will be maintained by Agilent to allow results to be compared across experiments conducted on the various versions of Agilent arrays. The Affymetrix GeneChip will be developed in consultation with the community. It is predicted to include ~70K B73 gene models, allowing the remaining ~30K sequences to be used for evaluating allele-specific expression and the abundance of sRNA, transposon and retrotransposon transcripts according to community input. Each platform provides complementary approaches that together reduce bias associated with sequence variability among alleles.

These “all genes” arrays will be valuable for annotation of the genome, particularly for those sequences that were not represented as ESTs. It will be essential that all platforms adopt a common nomenclature and should be able to retrofit with updated annotations. These features will be a key to the longevity and widespread utility of the “all genes” arrays. The group considered that both of these platforms are cost-effective choices in the current climate and that competition will accelerate the improvement of database support and will lower costs. **Efforts should focus on incorporating community input, developing data integration tools and maintaining accessibility so diverse groups of researchers can benefit from public investment in profiling experiments and database tools.** The maize community also looks toward developing profiling platforms in the future that will accommodate advances in systems biology.

### ***A major phenomics effort will contribute to basic and translational research***

Understanding the function of genes and networks is a central research goal both currently and also in the post-genomics era. Phenotyping is one of the major strengths of the maize community. The maize community envisions carrying this capacity to the next level by **developing large scale and multi-dimensional phenotyping capabilities.** The group

recognized that understanding adaptation and applying the information to improve agriculture can be best achieved through in-depth phenotyping in diverse environments for numerous traits.

The major goals of this recommended effort are to: 1) harness genetic diversity to assign biological functions to sequences, i.e., associating traits with genes, and 2) enable predictive biology via an iterative process of discovery and validation. This effort will involve broad community involvement to collect and analyze phenotypes in great depth and breadth on a common set of diverse genotypes. It likely will also require common center(s) for production and quality control of seed, shared planting locations and protocols for collecting phenotypes, and centralized quality control for experimental design and data analysis. A series of phenotypes, including agronomic, morphological, cellular and molecular traits will be measured. Genotypes will include combinations of the natural variants, transformants and mutagenized populations developed by the research community. This effort will require development of new high throughput analytical tools, *e.g.*, remote sensing and image analysis. There will be very wide dissemination of collected data and efforts to coordinate sharing of results.

**A unified phenotyping effort proposed here will require new scales of coordination within the community, will require continued advances in cyberinfrastructure, and further development of centralized databases for analysis, synthesis and dissemination of phenomics data.** At the outset, researchers will guide the effort by establishing consistency of phenotyping language and by maintaining quality of experimental design and well-designed reporting mechanisms. First and foremost, database resources must be planned to provide for the integration and dissemination of data. Planning for large-scale phenotyping efforts should include industry, if mechanisms for shared and public access of data generated can be resolved unambiguously.

### ***The maize community needs to broaden and strengthen its membership***

The maize genome can be best leveraged by increasing the diversity of participants and strengthening the depth of training. The maize research community is committed to training creative, independent and collaborative scientists who conduct hypothesis-driven research, tool development and research translatable to agriculture. This can be achieved by **strengthening and diversifying graduate education and post-doctoral training**. It is clear from the experience of other research communities that establishing effective partnerships, nationally and internationally, is also essential to exploit a genome fully. The maize community considered types of partnerships that need to be developed and strengthened in the coming years including industry, international contacts, a wider spectrum of US-based researchers involved in maize research as well as other plant and non-plant biologists in general.

### **Enhance public - private interactions**

Academic-industry partnerships have historically been strong for the maize community. A number of ideas were considered to strengthen our relationships further with industry. One idea was to provide mechanisms for industry to donate resources to the community, such as the EST database that was made accessible through an MTA, and funding of undergraduate summer internship programs. Another idea was to help young scientists make a transition to industry from academia by setting up a partnership with the

private sector to facilitate tours and visits and by providing talks for web viewing on ‘how to get trained for an industry position’. Issues related to public accessibility of data from new industry-academic partnerships need to be resolved. Adding a private sector member to the MGEC could be considered, as a mechanism to help facilitate interactions.

#### Enhance international dimension

International efforts should be coordinated to avoid duplication of effort and to foster dialogue. Ideas to enhance international collaboration included inviting one international speaker who has not been to the Annual Maize Genetics Conference to present each year. The community can also encourage and recognize greater efforts by PIs to become involved in Developing Country Collaboration supplements to NSF grants. Fellowship opportunities for graduate students abroad are currently lacking. Furthermore, to encourage the growth of maize research internationally, more meetings could be conducted outside of the US, such as in Mexico, South America, Europe, Asia, and Africa. International researchers might become more engaged in maize research if meetings are more accessible periodically. Independent or satellite workshop/short courses, such as the 2004 CIMMYT workshop conducted before the Annual Maize Genetics Conference, would further attract international participants.

#### Enhance local and national participation: Forming partnerships as outreach

The maize community is committed to improving science training in the US by reaching out to: small colleges and universities, traditionally under-funded research institutions, undergraduate institutions, minority-serving institutions, community colleges, tribal colleges, the K-12 system, and the general public. It was noted that maize genetics is an excellent “hook” for attracting new participants to science, because corn is such a familiar food item in the US and so many genetics tools are available. The goals of outreach activities should be 1) to integrate research and education and 2) to provide for a mutually beneficial partnership between members of the maize research community and new participants. These goals can be achieved by emphasizing relevance for all involved. Educational partnerships should be logical to the researcher’s expertise/interest as well as to the recipients’ needs and environment. **Communication among researchers with active and successful outreach programs should be strengthened to avoid duplication of effort.** To avoid continuous reinvention of methods, it might be useful to organize outreach advisory boards that can help guide new programs. Such advisory boards could be established and centralized through the current plant genome research outreach portal (plantgdb.pgroup). Best practices could also be highlighted at the Annual Maize Genetics Conference at the new designated poster session on Outreach and Training. The annual meeting should also be a central venue to bring new students and researchers from outside the maize field together with the current maize research community, thus enticing them to maize research. The Maize Genetics Meeting Steering Committee should continue to develop innovative ways to fund fellowships to new participants.

## Summary and Next Steps

The sequenced maize B73 genome holds great promise for contributing to basic and translational research. To take full advantage of that promise we need to 1) make MaizeGDB the centralized sequence resource, 2) make plans and implement next level map-based sequencing efforts, 3) provide increased capacity and lower costs for maize transformation technology, 4) increase sequence-based reverse genetics, 5) coordinate expression platforms so all data are easily shared, 6) conduct a major phenomics effort that is effectively integrated, and 7) increase participation in maize research. A timeline for completing these goals was discussed as follows:

Activity	Timeline for completion
Convert MaizeGDB to a sequence-oriented database	2011 (three years)
Implement additional sequencing efforts. Generate physical maps from other genomes and anchor their sequences to their maps	2013 (3-5 years)
Establish high capacity transformation facilities	2013 (5 years)
Establish a near-saturation reverse genetics resource	2013 (3-5 years)
Standardize expression platforms	2013 (5 years)
Phenomics project underway	2016 (8-10 years)
Increase and diversify the maize research community	2018 (10 years)

Allerton Retreat participants agreed on several first steps to begin to implement the long-term plan described here. **First**, this planning document will be disseminated to the broader maize research community by the MGEC via MaizeGDB. **Second**, this document, combined with an executive summary, will be presented to guests at the Allerton Retreat including representatives of the funding agencies (NSF, DOE, USDA) and the NCGA. **Third**, an article will be submitted to *The Plant Cell* to inform the broader plant biology community of the future directions of maize research in the genomic era. **Fourth**, taskforces will be formed to focus on solutions to particular research bottlenecks, including transformation, and to help shape future new research efforts, such as phenomics. Through MGEC guidance, we anticipate that taskforce action plans will help maintain dynamic assessment of progress and will guide maize research into the future.