SYMPOSIUM

GENOMICS AND PLANT BREEDING: THE EXPERIENCE OF THE INITIATIVE FOR FUTURE AGRICULTURAL AND FOOD SYSTEMS

Ann Marie Thro,* Wayne Parrott, Joshua A. Udall, and William D. Beavis, Editors

INTRODUCTION

For plant genomics to affect economic and environmental benefits, the knowledge gained must be "translated" into crop varieties having desired characteristics. The discipline of plant breeding is the "translator" of new knowledge from emerging technologies into improved crop cultivars, or "varieties." The AgGenomics section of the peer-reviewed Initiative for Future Agriculture and Food Systems [(IFAFS) offered by CSREES, USDA in 2000 and 2001] was one of the few sources of funding to date for integration of genomics and plant breeding. A symposium held 13 Nov. 2002 at the annual meetings of the Crop Science Society of America (CSSA), cosponsored by CSREES, USDA, and by CSSA Divisions C-1 and C-7, shared perspectives gained from IFAFS research on the value of genomics to plant breeding. A panel of IFAFS grantees representing a diversity of crops was invited to present summaries of their research experience and participate in a discussion. IFAFS panel members were Charles Brummer, Agronomy Dep., Iowa State Univ.; Jorge Dubcovsky, Dep. of Agronomy and Range Science, Univ. of California, Davis; Michael Havey, USDA-ARS and Horticulture Dep., Univ. of Wisconsin; Molly Jahn, Plant Breeding Dep., Cornell Univ.; Steven Knapp, Crop and Soil Sciences Dep., Oregon State Univ.; Robert Martienssen, Cold Spring Harbor Laboratories; and Andrew Paterson, Crop and Soil Sciences Dep., Univ. of Georgia. In addition, Dr. M. Goodman, Crop Science Dep., North Carolina State Univ., represented classical plant breeding, and Dr. Mark Cooper, Pioneer HiBred International, represented the private sector. Each essay in the present collection addresses the utility of molecular biology for crop improvement, from the viewpoint of an individual panelist. Following the panelist essays, are a brief report of the discussion and an overall summary of the symposium's main points, both prepared by the symposium editors.

A.M. Thro, CSREES, USDA, 800 9th St. S.W., Washington, DC 20024 USA; W. Parrott, Department of Crop and Soil Sciences, University of Georgia, Athens, GA 30602-7272; J.A. Udall, Department of Botany, Bessey Hall Iowa State University Ames, IA 50011; W. Beavis, National Center for Genome Resources, 2935 Rodeo Park Drive East, Santa Fe, NM 87505. Received 23 July 2003. *Corresponding editor (athro@csrees.usda.gov).

Application of Genomic Technologies to Crop Plants: Opportunities and Challenges

Michael J. Havey*

Enormous genomic resources have been developed for model plants such as *Arabidopsis thaliana* (L.) Heynh. and rice (*Oryza sativa* L.), including detailed genetic maps (Harushima et al., 1998), huge numbers of expressed sequence tags (ESTs) (Sasaki et al., 1994; Seki et al., 2002), deep-coverage large-insert [such as bacterial artificial chromosome (BAC)] libraries with extensive contig assemblies (Zhang et al., 1996; Mozo et al., 1998; Zhao et al., 2002), and both targeted and complete genome sequencing and annotation (Goff et al., 2002; Yu et al., 2002). These resources, coupled with the development of mutant stocks by knock-outs (Young et al., 2001) or targeted induced local lesions in genomes [TILLING (Till et al., 2003)], will allow for the efficient identification of gene(s) controlling phenotypes in model systems. However it is not clear how broadly applicable genetic associations revealed in model systems will be to economically important plants. Unquestionably, some genes identified in model plants will also condition economically important phenotypes in a crop plant. An example is the *FLC* locus that controls flowering in *Arabidopsis* (Michaels and Amasino, 1999) and in the closely related brassicas (Kole et al., 2001). Similar successes are readily communicated to the scientific community through publications; however, it will be difficult to publish, and therefore assess, the opposite scenario in which genes identified in model systems are not associated with similar traits in economically important plants. The purpose of this paper is to discuss some potential inconsistencies between model systems and economically important plants with onion (*Allium cepa* L.) as an example.

Recent studies have revealed that the Commelinanae and Asparagales are two strongly supported monophyletic sister groups within the monocots (Chase et al., 1995; Chase et al., 2000; Fay et al., 2000). The Commelinaloid monocots include the order Poales and possess the most economically important monocots, such as maize (*Zea mays* L.), rice, wheat (*Triticum aestivum* L.), etc. The Asparagales are the second most economically important monocot order and include important plants such as agave (*Agave* spp.), aloe (*Aloe* spp.), asparagus (*Asparagus officinalis* L.), chive (*Allium schoenprasum* L.), garlic (*Allium sativum* L.), iris (*Iris* spp.), leek (*Allium ampeloprasum* L.), onion, orchid (*Erycina* spp.),

USDA-ARS and Dep. of Horticulture, Univ. of Wisconsin-Madison, 1575 Linden Drive, Madison, WI 53706. Received 23 July 2003. *Corresponding author (mjhavey@facstaff.wisc.edu).

Published in Crop Sci. 44:1893–1919 (2004).

677 S. Segoe Rd., Madison, WI 53711 USA

and vanilla (*Vanilla* spp.). The "higher" Asparagales show successive microsporogensis, a cell plate is laid down after the first meiotic division, and form a well-defined group within the Asparagales (Fay et al., 2000). Economically important families in the "higher" Asparagales include the Alliaceae (chive, garlic, leek, and onion), Amaryllidaceae [various ornamentals and yucca (*Yucca* sp. L.)], and Asparagaceae (asparagus).

Genetic analyses of the Asparagales are hampered by longer generation times, severe inbreeding depression, and relative high cost of doing crosses. In addition, the nuclear genomes of the Asparagales are among the largest of all eukaryotes (Bennett and Smith, 1976; Ori et al., 1998). For example, onion has a nuclear genome of 16 415 megabasepairs per 1C, approximately equal to wheat and approximately 34 and 6 times larger than rice and maize, respectively (Arumuganathan and Earle, 1991). In contrast to wheat as a disomic hexaploid or maize as an ancient tetraploid (Celarier, 1956; Anderson, 1945), onion is a diploid ($2n = 2x = 16$) with no evidence of recent polyploidization contributing to its enormous nuclear genome. C_0t reassociation kinetics demonstrated that the onion genome consists of middle-repetitive sequences occurring in short-period interspersions among single-copy regions (Stack and Comings, 1979). Biochemical and cytological analyses, as well as genetic mapping, indicated that intrachromosomal tandem duplications may have contributed to increased chromosome sizes in onion (Jones and Rees, 1968; Ranjekar et al., 1978; King et al., 1998). The extremely large nuclear genome of onion represents a huge challenge for the development of genomic resources. For example, a BAC library of onion with 99% probability of having any single copy region would require 503 957 clones of 150 kilobases. In comparison, similar coverage libraries of rice and maize would require 15 041 and 82 000 similarly sized clones, respectively. The development of genomic resources using species with smaller nuclear genomes will be imperative for the identification and cloning of economically important genes from onion.

Conservation of the linear order of genes (synteny) on chromosomes among related species is well documented for the Poaceae (Ahn et al., 1993; Dunford et al., 1995; Devos and Gale, 2000), Solanaceae (Bonierbale et al., 1988; Tanksley et al., 1992), and between *Arabidopsis* and the brassicas (Lagercrantz 1998, Parkin et al., 2002). Significant synteny among related species will allow for the alignment of major economically important qualitative or quantitative trait loci across specific chromosome regions in major crops (Paterson et al., 1995; Maughan et al., 1996). The identification of candidate genes, either as ESTs or open-reading frames (ORFs) on genomic contigs, will be revealed by fine mapping and comparison of flanking molecular markers to the annotated sequences of model plants. These associations should augment our chances of developing efficient marker-facilitated selection of major and minor genes, significantly reducing or eliminating recombination between the marker and the desired genes. This is especially important for the application of marker-facilitated selection to open-pollinated populations at or near linkage equilibrium. We recently demonstrated linkage equilibrium between tightly linked molecular markers and the *Ms* locus in open-pollinated populations of onion (Gokce and Havey, 2002). Because economically important populations of some crop plants have been open pollinated since antiquity, genomic regions showing linkage disequilibrium may be very short and require essentially cloning of genes to tag important traits for marker-facilitated selection.

Many candidate genes will be identified by knocking out specific genes by transposon insertions or TILLING in model plants such as *Arabidopsis thaliana*. These knock-outs may affect structural genes and not reveal variation at *trans*-acting factors that control the expression of major structural genes. Pleiotropy will complicate our ability to predict the relationships between specific candidate genes and phenotypes, as well as the manipulation of the candidate genes. Finally, we may be surprised by the true gene(s) controlling a specific phenotype. An example of an unexpected relationship is nuclear restoration of male fertility in cytoplasmic-male-sterile (CMS) maize. The *Rf2* locus of maize synthesizes an aldehyde dehydrogenase that operates by an inconspicuous mechanism to restore male fertility in CMS maize (Liu et al., 2001). An onion cDNA highly homologous to the maize aldehyde dehydrogenase mapped independently of male-fertility restoration in CMS onion (Gokce et al., 2002), revealing that different genes may condition the same phenotype in different crop plants.

These challenges not withstanding, the genomic resources developed for model plants will reveal a plethora of candidate genes and provide great insights into gene expression. In some cases genes identified in model systems will condition economically important phenotypes in crop plants. However, it remains imperative that we identify, clone, and understand specific gene(s) conditioning economically important phenotypes in our major crops. In many cases, large genomic clones of specific crop plants will be imperative for the isolation of genomic regions, either up or down stream from the structural gene, interacting with important *trans*-acting factors controlling gene expression. Deep coverage genomic libraries and targeted sequencing will be required for the identification, cloning, and manipulation of specific genes affecting economically important phenotypes in major crop plants.

ACKNOWLEDGMENTS

The author acknowledges the support of an IFAFS grant from CSREES USDA.

REFERENCES

Ahn, S., J. Anderson, M. Sorrells, and S. Tanksley. 1993. Homoeologous relationships of rice, wheat and maize chromosomes. Mol. Gen. Genet. 241:483–490.

Anderson, E. 1945. What is *Zea mays*? A report of progress. Chron. Bot. 9:88–92.

Arumuganathan, K., and E. Earle. 1991. Nuclear DNA content of some important plant species. Plant Mol. Biol. Rep. 9:208–218.

Bennett, M.D., and J. Smith. 1976. Nuclear DNA amounts in angiosperms. Phil. Trans. Royal Soc. London B 274:227–274.

Bonierbale, M.W., R. Plaisted, and S. Tanksley. 1988. RFLP maps

based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. Genetics 120:1095–1103.
Celarier, R.P. 1956. Additional evidence for five as the basic chromosome number of the Andropogoneae. Rhodora 58:135–143.
Chase, M., M. Duvall, H. Hills, J. Conran, A. Cox., L. Eguiarte et al. 1995. Molecular systematics of Lilianae. p. 109–137. *In* P. Rudall et al. (ed.) Monocotyledons: Systematics and evolution. Royal Botanic Gardens, Kew.
Chase, M.W., D.E. Soltis, P.S. Soltis, P.J. Rudall, M.F. Fay, W.H. Hahn et al. 2000. Higher-level systematics of the monocotyledons: An assessment of current knowledge and a new classification. p. 3–16. *In* K.L. Wilson and S.A. Morrison (ed.) Monocots: Systematics and evolution Vol. 1. CSIRO, Melbourne, Australia.
Devos, K.M., and M.D. Gale. 2000. Genome relationships: The grass model in current research. Plant Cell 12:637–646.
Dunford, R.P., N. Kurata, D. Laurie, T. Money, Y. Minobe, and G. Moore. 1995. Conservation of fine-scale DNA marker order in the genomes of rice and the Triticeae. Nucleic Acids Res. 23:2724–2728.
Fay, M.F., P.J. Rudall, S. Sullivan, K.L. Stobart, A.Y. de Bruijn, G. Reeves, F. Qamaruz-Zaman, W-P. Hong, J. Joseph, W.J. Hahn, J.G. Conran, and M.W. Chase. 2000. Phylogenetic studies of Asparagales based on four plastid DNA regions. *In* K.L Wilson and D.A Morrison (ed.) Monocots: Systematics and evolution. CSIRO Publishing, Melbourne, Australia.
Goff, S.A., D. Ricke, T. Lan, G. Presting, R. Wang, M. Dunn et al. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science 296:92–100.
Gokce, A.F., and M.J. Havey. 2002. Linkage equilibrium among tightly linked RFLPs and the *Ms* locus in open-pollinated onion populations. J. Am. Soc. Hortic. Sci. 127:944–946.
Gokce, A.F., J. McCallum, Y. Sato, and M.J. Havey. 2002. Molecular tagging of the *Ms* locus in onion. J. Am. Soc. Hortic. Sci. 127:576–582.
Harushima, Y., M. Yano, A. Shomura, M. Sato, T. Shimano, Y. Kuboki et al. 1998. A high-density rice genetic linkage map with 2275 markers using a single F_2 population. Genetics 148:479–494.
Jones, R., and H. Rees. 1968. Nuclear DNA variation in *Allium.* Heredity 23:591–605.
King, J.J., J. Bradeen, O. Bark, J. McCallum, and M.J. Havey. 1998. A low-density genetic map of onion reveals a role for tandem duplication in the evolution of an extremely large diploid genome. Theor. Appl. Genet. 96:52–62.
Kole, C., P. Quijada, S. Michaels, R. Amasino, and T. Osborn. 2001. Evidence for homology of flowering-time genes VFR2 from *Brassica rapa* and FLC from *Arabidopsis thaliana.* Theor. Appl. Genet. 102:425–430.
Lagercrantz, U. 1998. Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that brassica genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. Genetics 150: 1217–1228.
Liu, F., X. Cui, H. Horner, H. Weiner, and P. Schnable. 2001. Mitochondrial adehyde dhydrogenase activity is required for male fertility in maize. Plant Cell 13:1063–1078.
Maughan, P., M. Saghai-Maroof, and G. Buss. 1996. Molecular-marker analysis of seed weight: Genomic location, gene action, and evidence for orthologous evolution among three legume species. Theor. Appl. Genet. 93:574–579.
Michaels, S., and R. Amasino. 1999. Flowering locus C encodes a novel MADS-domain protein that acts as a repressor of flowering. Plant Cell 11:949–956.
Mozo, T., S. Fischer, H. Shizuya, and T. Altmann. 1998. Construction and characterization of the IGF *Arabidopsis* BAC library. Mol. Gen. Genet. 258:562–570.
Ori, D., R. Fritsch, and P. Hanelt. 1998. Evolution of genome size in *Allium* (Alliaceae). Plant Syst. Evol. 210:57–86.
Parkin, I., D. Lydiate, and M. Trick. 2002. Assessing the level of collinearity between *Arabidopsis thaliana* and *Brassica napus* for *A. thaliana* chromosome 5. Genome 45:356–366.
Paterson, A., Y. Lin, Z. Li, K. Schertz, J. Doebley, S. Pinson, S. Liu, J. Stansel, and J. Irvine. 1995. Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. Science 269:1714–1718.
Ranjekar, P.K., D. Pallotta, and J. Lafontaine. 1978. Analysis of plant genomes. V. Comparative study of molecular properties of DNAs of seven *Allium* species. Biochem. Genet. 16:957–970.
Sasaki, T., J. Song, Y. Koga-Ban, E. Matsui, F. Fang, H. Higo et al. 1994. Toward cataloguing all rice genes: Large-scale sequencing of randomly chosen rice cDNAs from a callus cDNA library. Plant J. 6:615–624.
Seki, M., M. Narusaka, A. Kamiya, J. Ishida, M. Satou, T. Sakurai et al. 2002. Functional annotation of a full-length *Arabidopsis* cDNA collection. Science 296:141–145.
Stack, S.M., and D.E. Comings. 1979. The chromosomes and DNA of *Allium cepa.* Chromosoma 70:161–181.
Tanksley, S.D., M. Ganal, J. Prince, M. de Vicente, M. Bonierbale, P. Brown et al. 1992. High density molecular linkage maps of the tomato and potato genomes. Genetics 132:1141–1160.
Till, B., S. Reynolds, E. Greene, C. Codomo, L. Enns, J. Johnson et al. 2003. Large-scale discovery of induced point mutations with high-throughput TILLING. Genome Res. 13:524–530.
Young, J., P. Krysan, and M. Sussman. 2001. Efficient screening of *Arabidopsis* T-DNA insertion lines using degenerate primers. Plant Physiol. 125:513–518.
Yu, J., S. Hu, J. Wang, G. Wong, S. Li, B. Liu et al. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science 296:79–92.
Zhang, H.B., S. Choi, S.S. Woo, Z. Li, and R.A. Wing. 1996. Construction and characterization of two rice bacterial artificial chromosome libraries from the parents of a permanent recombinant inbred mapping population. Mol. Breed. 2:11–24.
Zhao, Q., Y. Zhang, Z. Cheng, M. Chen, S. Wang, Q. Feng et al. 2002. A fine physical map of the rice chromosome 4. Genome Res. 12:817–823.

Marker-Assisted Selection in Public Breeding Programs: The Wheat Experience

Jorge Dubcovsky*

It has been suggested that the recent progress in the area of plant molecular biology and plant genomics have the potential to initiate a new Green Revolution. However, these discoveries need to be implemented in new cultivars to realize that potential. The controversy about transgenic crops has delayed the incorporation of alien genes into plants and significantly increased the cost to develop and release transgenic crops. These costs are usually beyond the resources of public breeding programs and, therefore, are not currently used in most cultivated plants.

Fortunately, biotechnology has provided additional tools that do not require the use of transgenic crops to revolutionize plant breeding. Progress in molecular genetics has resulted in the development of DNA tags, which can be used in marker-assisted selection (MAS) strategies for cultivar development (Paterson et al., 1991). These

Dep. of Agronomy & Range Science, Univ. of California, One Shields Ave., Davis, CA 95616. Received 23 July 2003. *Corresponding author (jdubcovsky@ucdavis.edu).

Abbreviations: BAC, bacterial artificial chromosome; MAS, marker-assisted selection; SNP, single nucleotide polymorphism.

molecular markers can be used as chromosome landmarks to facilitate the selection of chromosome segments including useful agronomic traits during the breeding process. These markers are particularly useful for incorporating genes that are highly affected by the environment, genes for resistance to diseases that cannot be easily screened for, and to accumulate multiple genes for resistance to specific pathogens and pests within the same cultivar, a process called gene pyramiding. An additional advantage of the incorporation of MAS into breeding programs is that very different types of traits, e.g. a disease resistance gene or a gene to increase grain protein content, can be manipulated using the same technology. Dekkers and Hospital (2002) have recently reviewed some of the potential limitations of MAS strategies, and concluded that the use of MAS will be determined by the economic benefit relative to conventional selection.

The alleles that are incorporated by MAS are generally present within the gene pool of a particular crop and are transferred by meiotic chromosome recombination. One of the positive aspects of this approach is that these genes reside at their natural chromosomal locations, thereby minimizing the possibility of gene silencing. Another important aspect of cultivars developed by MAS is that they are not transgenic and therefore, do not face the public resistance against transgenic crops.

The MAS strategy is a way to capitalize on available markers and to incorporate valuable traits into elite lines that are suitable for cultivar release. In addition, release of these MAS-improved cultivars is an efficient way of demonstrating the power of these technologies to the public. However, limited funding for implementation efforts had delayed the incorporation of these powerful technologies into most public breeding programs.

Wheat Breeding in the USA: A Public Effort

Wheat (*Triticum aestivum* L.) is a self-pollinating species and therefore, growers can save seed from one harvest for the next year. This has reduced the profitability of wheat breeding for the private sector and has resulted in the continuous existence of a large, vibrant public sector involved in cultivar development. For example, the total number of cereal crop breeders in the USA in the last census was 893, with 80% being in the private sector and 20% being in the public sector (Frey, 1996). In wheat, approximately 60% of the breeders were in the public sector. By comparison, only 7% of the corn breeders were in the public sector. Public investments in wheat breeding during the past century have resulted in the development of the majority of cultivars grown by U.S. farmers. State agricultural colleges and experimental stations, USDA, or CIMMYT developed approximately 60% of the cultivars released in the USA during the 20th century. In addition, a high percentage of the area of wheat production in the USA is attributed to publicly developed cultivars (KS 62%, ND 64%, WA 88%, NE 90%) (NASS, 2001).

Fuglie et al. (1996) found a typical range of 40 to 60% return on public research investment, with public wheat breeding consistently at the top of this range. In addition, nine out of the 10 interspecific translocations involving the introgression of novel genes into cultivated germplasm that significantly affected U.S. wheat production were developed in public plant breeding programs (Mercado et al., 1996). These data provide convincing support of the broad impact of public wheat breeding efforts both in cultivar development and in germplasm enhancement.

Public wheat-breeding programs are typically supported by wheat grower associations. However, low wheat prices in the past years have resulted in a reduction of resources available to the U.S. wheat growers and a shrinking of resources for research and development in new technologies. This situation was aggravated by a limited investment of federal funding agencies during the 1990s in implementation grants for public wheat breeding programs. This limited investment in practical applications is difficult to understand in light of the large investment made by the same funding agencies in wheat molecular genetics and wheat genomics.

During the last 10 yr, public researchers constructed detailed wheat genetic maps including more than 3000 molecular markers and physical maps including more than 16 000 loci (http://wheat.pw.usda.gov/NSF/; verified 2 July 2004). In addition to mapping, U.S. federal agencies have funded the sequencing of more than 105 000 wheat ESTs, the construction of wheat Bacterial Artificial Chromosome (BAC) libraries (Cenci et al., 2003; Lijavetzky et al., 1999), the assembly of BACs into physical maps (http://wheat.pw.usda.gov/PhysicalMapping/; verified 2 July 2004), and the sequencing of large segments of wheat DNA (SanMiguel et al., 2002). These powerful genomic resources have started to yield the first successful positional cloning efforts in wheat (Faris et al., 2003; Feuillet et al., 2003; Huang et al., 2003; Yahiaoui et al., 2004; Yan et al., 2003; Yan et al., 2004). Cloning of agronomically important genes has made possible to develop "perfect markers," based directly on the allelic variation responsible for the differences in the trait. Examples of perfect markers in wheat include the glutenin genes for gluten strength (Anderson et al., 1989), the waxy genes for starch properties (Briney et al., 1998), the puroindoline genes for hardness (Beecher et al., 2002), the vernalization genes for vernalization requirement (Yan et al., 2003; Yan et al., 2004), the *Rht* genes for semi-dwarf habit (Peng et al., 1999), and the *Lr10* and *Lr21* genes for leaf rust resistance (Feuillet et al., 2003; Huang et al., 2003). Wheat researchers have also developed closely linked molecular markers to yet unidentified genes with positive effects on quality characteristics and resistance to fungi, viruses, and insects (reviewed by Dubcovsky et al., 2000; Anderson, 2000).

The most efficient way to develop a positive synergistic effect between the large research investments in wheat genomics and the growers' investment in public wheat breeding is to fund implementation research projects. The MAS programs are good examples of implementation projects that have the potential to facilitate the transfer of valuable genes identified in basic research programs into public wheat varieties.

MASwheat: A Public MAS Program

The wheat public research sector has a long tradition of collaborative projects that were initiated at the begin-

ning of the 1990s by the International Triticeae Mapping Initiative. Large multi-laboratory projects continued later in the USA under the funding of the NSF-Plant Genome Initiative (http://wheat.pw.usda.gov/NSF/; verified 2 July 2004). Many of the collaborators of these projects were wheat breeders, facilitating the integration of basic and applied wheat researchers. This integrated research community and the availability of the results from previous research efforts in marker development were instrumental in developing a successful proposal for MAS in wheat.

Wheat researchers and breeders from 12 public programs across the USA organized a national wheat MAS consortium (MASwheat) that was funded by the USDA Initiative for the Future of Agriculture and Food Systems (2001–2004). The MASwheat project structure is similar to the Australian National Wheat Molecular Marker Program (NWMMP) implemented in 1996. The main objective of both projects is to empower the breeders by implementing MAS capacities within each of the existing public breeding programs. This strategy has been successful in closing the funding gap between the development of genomic tools and the public investment in cultivar development, and in transferring the value of genomic research to the wheat growers' fields.

The MASwheat project is committed to transfer new developments in wheat genomics and biotechnology to U.S. wheat production through marker-assisted selection. Available molecular markers are being used to transfer 22 resistance genes to fungi, viruses, and insects; and 21 gene variants related to bread, pasta, and noodle quality into 75 different recurrent parents (34 whites, 33 reds and 8 durums). Eighty MAS projects have been already completed and additional 350 backcrossing programs are currently being advanced in average two generations a year by MAS.

All the information and protocols used in the MASwheat project are publicly available through the project WEB site (http://maswheat.ucdavis.edu; verified 2 July 2004). The collaborative nature of the project and the public access to the information was lauded in a recent article in *Nature* focused on the current difficulties of public breeding programs (Knight, 2003). The numerous presentations in growers' meetings, field days and symposiums by the members of the MASwheat consortium are also improving the public understanding of the potential benefits of biotechnology.

Conclusions

Marker technologies are continuously evolving. The development of 96-well DNA extraction protocols and high-throughput genotyping equipment resulted in substantial reduction of MAS costs. A new generation of molecular markers based on the detection of single nucleotide polymorphisms (SNP) promises high-throughput assays at relatively low costs, along with the potential for high levels of multiplexing. Implementation of this multiplexing technology in plant improvement strategies can provide cost-effective tools for selection of multiple traits in breeding populations.

The challenge for the public plant breeders and for the federal funding agencies will be to generate the integrated proposals and necessary funding to continue the actual MAS programs and to incorporate new marker technologies.

One important aspect of the new genomic revolution is that most of the information is publicly available. Therefore, competitiveness will not be determined by access to the information but by the speed in which these technologies are incorporated into the breeding programs. This represents both a challenge and a fantastic opportunity for the public breeding programs that have the expertise to utilize successfully MAS technologies.

ACKNOWLEDGMENTS

The authors acknowledge financial support from USDA-IFAFS competitive grant 2001-04462. The first author thanks J. Anderson, B.S. Gill, S. Kianian, N. Lapitan, J. Sherman, and M.A. Soria for a critical review of this article.

REFERENCES

Anderson, J.A. 2000. Marker-assisted selection of disease resistance genes in wheat. p. 71–84. *In* M.M. Kohli and M. Francis. (ed.) Application of biotechnologies to wheat breeding. CIMMYT-INIA.

Anderson, O.D., F.C. Greene, R.E. Yip, N.G. Halford, P.R. Shewry, and J.M. Malpica-Romero. 1989. Nucleotide sequences of the two high-molecular-weight glutenin genes from the D genome of hexaploid wheat, *Triticum aestivum* L. cv. Cheyenne. Nucleic Acid Res. 17:461–462.

Beecher, B., A. Bettge, E. Smidansky, and M.J. Giroux. 2002. Expression of wild-type pinB sequence in transgenic wheat complements a hard phenotype. Theor. Appl. Genet. 105:870–877.

Briney, A., R. Wilson, R.H. Potter, I. Barclay, G. Crosbie, R. Appels et al. 1998. A PCR-based marker for selection of starch and potential noodle quality in wheat. Mol. Breed. 4:427–433.

Cenci, A., N. Chantret, K. Xy, Y. Gu, O.D. Anderson, T. Fahima et al. 2003. Construction and characterization of a half million clones: Bacterial Artificial Chromosome (BAC) library of durum wheat. Theor. Appl. Genet. 107:931–939.

Dekkers, J.C.M., and F. Hospital. 2002. The use of molecular genetics in the improvement of agricultural populations. Nat. Rev. Genet. 3:22–32.

Dubcovsky, J., G. Tranquilli, D. Lijavetzky, I.A. Khan, A.R. Schlatter, and M.M. Manifesto. 2000. Advances in molecular markers for bread making quality. p. 57–70. *In* M.M. Kohli and M. Francis (ed.) Application of biotechnologies to wheat breeding. CIMMYT-INIA.

Faris, J.D., J.P. Fellers, S.A. Brooks, and B.S. Gill. 2003. A BAC contig spanning the major domestication locus Q in wheat and identification of a candidate gene. Genetics 164:311–321.

Feuillet, C., S. Travella, N. Stein, L. Albar, A. Nublat, and B. Keller. 2003. Map-based isolation of the leaf rust disease resistance gene Lr10 from the hexaploid wheat (*Triticum aestivum* L.) genome. Proc. Natl. Acad. Sci. USA 100:15253–15258.

Frey, K.J. 1996. National Plant Breeding Study-1. Human and financial resources devoted to plant breeding research and development in the United States in 1994. Special Report 98. Iowa State University. Iowa Agriculture and Home Economics Experiment Station, Ames.

Fuglie, K.O., K.N. Ballenger, K. Day, C. Klotz, M. Ollinger, J. Reilly et al. 1996. Agricultural research and development: Public and private investments under alternative markets and institutions 735. USDA, Economic Research Service, Washington, DC.

Huang, L., S.A. Brooks, W.L. Li, J.P. Fellers, H.N. Trick, and B.S. Gill. 2003. Map-based cloning of leaf rust resistance gene *Lr21* from the large and polyploid genome of bread wheat. Genetics 164: 655–664.

Knight, J. 2003. Crop improvement: A dying breed. Nature 421:568–570.

Lijavetzky, D., G. Muzzi, T. Wicker, B. Keller, R. Wing, and J. Dubcovsky. 1999. Construction and characterization of a bacterial artificial chromosome (BAC) library for the A genome of wheat. Genome 42:1176–1182.

Mercado, L.A., E. Souza, and K.D. Kephart. 1996. Origin and diversity of North American hard spring wheats. Theor. Appl. Genet. 93: 593–599.

NASS. 2001. National Agricultural Statistics Service, State Offices. [Online] http://www.usda.gov/nass/sso-rpts.htm; verified 2 July 2004.
Paterson, A.H., S.D. Tanksley, and M.E. Sorrells. 1991. DNA markers in plant improvement. Adv. Agron. 46:39–90.
Peng, J.R., D.E. Richards, N.M. Hartley, G.P. Murphy, K.M. Devos, J.E. Flintham et al. 1999. 'Green revolution' genes encode mutant gibberellin response modulators. Nature 400:256–261.
SanMiguel, P., W. Ramakrishna, J.L. Bennetzen, C.S. Busso, and J. Dubcovsky. 2002. Transposable elements, genes and recombination in a 215-kb contig from wheat chromosome 5A. Funct. Integr. Genom. 2:70–80.
Yahiaoui, N., P. Srichumpa, R. Dudler, and B. Keller. 2004. Genome analysis at different ploidy levels allows cloning of the powdery mildew resistance gene Pm3b from hexaploid wheat. Plant J. 37: 528–538.
Yan, L., A. Loukoianov, G. Tranquilli, M. Helguera, T. Fahima, and J. Dubcovsky. 2003. Positional cloning of wheat vernalization gene *VRN1*. Proc. Natl. Acad. Sci. USA 100:6263–6268.
Yan, L., A. Loukoianov, G. Tranquilli, A. Blechl, I.A. Khan, W. Ramakrishna et al. 2004. The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. Science 303:1640–1644.

Crop Plant Genome Sequence: What Is It Good For?

Robert A. Martienssen*

With the commitment of resources from USDA-IFAFS and NSF Plant Genome Program, it is likely that sequencing of plant genomes will be a major activity in the next few years. Nonetheless, the value of these sequences is still a matter of debate, leading to concerns that priorities need to be carefully evaluated, not just in research but also in education. It is worth therefore revisiting the largest genome project attempted so far—the human genome project—and the doubts raised at the beginning of what seemed to be an unimaginably difficult undertaking at the time.

The Human Genome Project

At the outset of the human genome project in the mid-1980s, there was heated debate over the merits of a project scheduled to take 15 to 20 yr and to cost in excess of $3000 million. Arguments against the enormous undertaking ranged from scientific, to economic, ethical, and educational. It was argued that conventional biomedical research would have to be abandoned to fund the project; that graduate education would take a back seat as students were trained in sequencing and little else; and that the sequence of our genes would breach our inalienable right to privacy. Finally, there was an underlying conviction that the human genome sequence would be of little scientific value compared to the outrageous cost.

In the event, the human genome project was completed in less than 10 yr, and cost the U.S. taxpayer less than $500 million. Technological advances halved the anticipated costs year after year, following "Moore's Law," which predicted comparable increases in computer speed and memory over the same time period. It is projected that, by the end of this decade, an entire human genome will cost less than $10 000 to sequence.

Scientifically, the human genome project is already revolutionizing our understanding of sporadic and inherited diseases, including cancer, Alzheimer's, autism, and many more. It can be argued that the first drugs designed on the basis of gene discovery were inhibitors of the novel protease found in the genome of HIV, drugs which have radically improved the prognosis for AIDS (Anon., 1996). Now that the genomes of microbes and viruses are known, as well as their human hosts, drugs that uniquely target pathogens will follow this example in large numbers.

With respect to education, the genome project raised a new generation of biologist as much at home with a computer algorithm as with a pipette, and has greatly raised the profile of biomedical sciences at campuses around the world. While the debate concerning genetic privacy has widened considerably since the sequence was announced, forensic applications have overturned hundreds of convictions and have made the "grave of the unknown soldier" a thing of the past (Williamson and Duncan, 2002).

Plant Genome Sequencing

What lessons are there to be learned from this experience for crop plant genomics and plant breeding? As with animals, model genomes (nematode, fly) have been sequenced first (*Arabidopsis* and rice, *Oryza sativa* L.). However, now that they have been completed and their impact is being felt in basic research, should we go on and sequence major crops such as maize (*Zea mays* L.), soybean [*Glycine max* (L.) Merr.], wheat (*Triticum aestivum* L.), cotton (*Gossypium* spp.), and trees? Much of the debate over crop plant genomics echoes the debate surrounding the human genome project 15 yr ago. However, while the model plant genomes have transformed basic plant biology in much the same way as animal genomes have, there are major differences between crop plant genomics and the human genome project.

For one thing, human genome research contributes to biomedical research and development, a trillion dollar activity worldwide. Crop plant genome research also underlies enormously important industries in food, feed, energy, and fiber, but here the analogy ends. First, several species must be targeted to cover agriculturally important plants, rather than one genome in the case of biomedical research. Second, the seed industry operates on far lower margins than the pharmaceutical industry, and has raised public concerns over food safety and security. Finally, the genetic information available to plant breeders is usually thought to be far less extensive than the vast array of epidemiological data collected by the biomedical community, making the sequence less useful. Each of these arguments is certainly valid, but just as plant breeders embraced the vision of genetics

Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724. Received 23 July 2003. *Corresponding author (martiens@cshl.org).

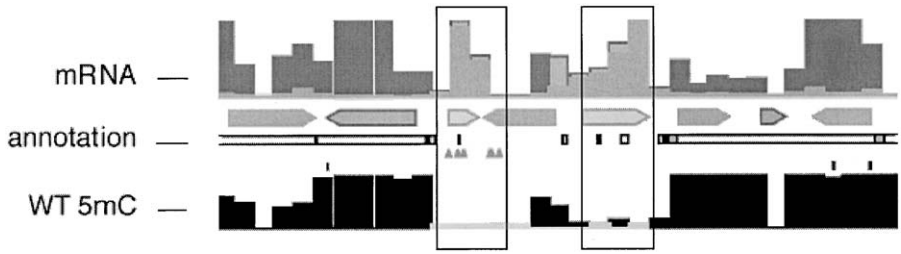

Fig. 1. The methylation profile of a 100-kb region of *Arabidopsis*. Profiling was accomplished by hybridizing a microarray with total genomic DNA and genomic DNA depleted of methylated sequences; the ratio is plotted. Annotated genes are light gray (boxed) annotated transposons are dark gray. Expression is also plotted in WT (light gray) and in DNA methylation mutants (dark gray). Only transposons are affected by loss of methylation (Lippman et al., 2004).

in the early part of the 20th century, we should not shun the promise of genomics now.

With respect to sequencing technology, plant genomes pose problems because of their size and repetitive content, as well as the number of different species required. This issue has been addressed by taking advantage of the observation that most methylation in plant genomes is restricted to transposons and high copy repeats (Fig. 1). By sequencing only the unmethylated portion of the genome, or by subtracting repeats, costs can drop by 10 fold or more, making the sequence of multiple large genomes practical (Rabinowicz et al., 2003).

Shotgun sequences can be linked to the genetic and physical maps by bacterial artificial chromosome (BAC) fingerprinting, and then ordered and oriented. Unfortunately, anchoring methods such as BAC-end sequencing, which has been widely employed in animals, is of limited use in crop plants because of the high proportion of long identical repeats. Instead, skimming methodologies can be used to anchor sequence islands to the physical map (Martienssen et al., 2004).

The Value Proposition

The value of the collective knowledge gathered by plant breeding has been underestimated, in part because of controlled pedigrees which are unavailable in human populations. Once genes underlying individual traits are known, the basis for disease resistance and stress tolerance is likely to emerge as it has in model organisms, allowing more precise "diagnosis" in breeding programs as well as genetic modification. The sequence can also be used to detect epigenetic, as well as genetic variation (Fig. 1), which likely contributes to traits such as flowering time, perennialism, apomixis, and heterotic performance. New pesticides and herbicides will also emerge from comparative genomics of crops and their pests, just as new antiviral and antimicrobial drugs are emerging in the pharmaceutical industry by selecting pathogen targets that are not found in the human genome and are less likely to be toxic. The HIV protease is one example of such a target, and protease inhibitors are some of the most successful antiviral drugs ever introduced.

Finally, the economic value of agriculture to a growing world population must not be underestimated. While the pharmaceutical industry is successful because of aging western populations, agriculture is the priority for crowded, youthful, hungry nations that make up the rest of the world. The imperative to modernize agricultural research is becoming clear.

Thus genomics can provide a road-map for the next generation of agricultural and breeding research, but it cannot replace the geneticist or the plant breeder. What it can do is open new areas of research unimagined by conventional plant breeding. Imagine, for example, small, nontoxic molecules that delay flowering or trigger apomixes. Given the importance of agricultural research for food security, energy conservation, and the environment, as well as the rich genetic resources available, genomics may yet make a greater impact on plant breeding than the human genome has had in biomedical research.

ACKNOWLEDGMENTS

The author acknowledges the support of an IFAFS grant from CSREES USDA.

REFERENCES

Anon. 1996. The disease detective. Dr. David Ho, Man of the Year. Time Magazine, Dec. 30, 1996.

Lippman, Z., A.-V. Gendrel, M. Black, M. Vaughn, N. Dedhia, W.R. McCombie, K. Lavine, V. Mittal, B. May, K. Kasschau, J.C. Carrington, R.W. Doerge, V. Colot, and R. Martienssen. 2004. Role of transposable elements in heterochromatin and epigenetic control. Nature 429 (in press).

Martienssen RA, P.D. Rabinowicz, A. O'Shaughnessy, and W.R. McCombie. 2004. Sequencing the maize genome. Curr. Opin. Plant Biol. 7:102–107.

Rabinowicz, P.D., W.R. McCombie, and R.A. Martienssen. 2003. Gene-enrichment in plant genomic shotgun libraries. Curr. Opin. Plant Bio. 6:150–156.

Williamson, R., and R. Duncan. 2002. DNA testing for all. Nature 418:585–586.

Reducing the Genetic Vulnerability of Cotton

Andrew H. Paterson,* Randal K. Boman, Steven M. Brown, Peng W. Chee, John R. Gannaway, Alan R. Gingle, O. Lloyd May, and C. Wayne Smith

THE U.S. COTTON (*Gossypium* spp.) production system exemplifies the challenges that must be met to reduce the genetic vulnerability of a major crop. As an industrial crop that sustains one of the largest U.S. industries (textiles), more than 400 000 domestic jobs are related to cotton production and processing, with an aggregate impact of over $40 000 million on the U.S. gross domestic product.

The U.S. cotton gene pool is genetically impoverished. All cottons cultivated in the USA are tetraploid, thought to have arisen in the New World about 1 to 2 million years ago, as a result of an unusual hybridization event between an invasive A-genome diploid genotype and an indigenous D-genome diploid. Polyploid cottons are thought to be monophyletic, with contributions from only two of the eight extant diploid "genome types" in the genus, representing the first genetic bottleneck. A second bottleneck was associated with the domestication of cotton from a small subset of the wild genotypes. A third bottleneck was imposed by human sampling of tetraploid cotton genotypes from their center of diversity in Mexico and Central America, and spread northward into the USA, and also to China, India, Egypt, Australia, and other countries.

Growing concern about genetic vulnerability of the cotton gene pool to a wide range of biotic and abiotic hazards is exemplified by recent investigation of trends in yield improvement. Stagnation in yield improvement of *G. hirsutum* L. recently led the National Cotton Council to form a Blue Ribbon Committee of public and private scientists to determine if there truly was a yield plateau, and if so, how it should be addressed (Helms, 2000). The committee based its findings on current public data, largely from National Cotton Variety Tests (Rayburn et al., 1999) and the National Agricultural Statistics Service, USDA (NASS, 1998). On the basis of a linear model, over the past 39 yr, cotton yields have increased about 6.7 kg ha^{-1} yr^{-1} (1.3% annual rate). A polynomial model clearly indicates that this increase has not been linear. From 1970 through 1985, the rate of change in cotton yields rose at an increasing level. However, from 1985 through 1998, the rate of change in cotton yields declined and from 1992 through 1998, the actual yields declined. Stated another way, the rate of change in cotton yields has steadily declined since 1985. By 1998, absolute cotton yields (not just the rate of change) reached a disturbing rate of decline of about 16.8 kg^{-1} ha^{-1} yr^{-1} (3.3% annual rate). Accompanying this yield decline, year-to-year variations in yield were almost four times greater in the period from 1980 to 1998 than in 1960 to 1979. This increased volatility in yield translates directly into higher risk for the grower.

The yield plateau in cotton appears to be closely associated with increasing genetic vulnerability. In addition to the genetic bottlenecks imposed by polyploid formation, domestication, and migration, the U.S. cotton gene pool has been further eroded by the over-exploitation of a few genetic backgrounds during the past 15 yr. Commercial breeding programs repeatedly employ a few closely related genotypes to generate new cultivars (May et al., 1995). Growers are planting large areas to these closely related cultivars, resulting in a high level of field genetic uniformity (Van Esbroeck et al., 1998). This has been exacerbated by the widespread (about 60% of 1999 U.S. hectares) planting of transgenic cultivars that are the result of backcross breeding with an even smaller subset of closely related genotypes.

A solution may lie in the exploration of exotic genotypes. Although there exist five tetraploid *Gossypium* species each including a huge array of feral and wild forms, at present, no exotic cottons appear in the pedigrees of *any* modern U.S. cultivars or enhanced germplasm (Bowman et al., 1996; Bowman et al., 1997; Calhoun et al., 1997). Our early explorations in a small sampling of these genotypes show many transgressive QTL alleles (P. Chee, X. Draye, A.H. Paterson, in preparation), including some cases in which different taxa appear to have evolved complementary alleles at alternative homeologous loci (Saranga et al., 2001).

However, introgressive breeding involves new challenges that require fundamentally different approaches than mainstream plant breeding, and an equally long-term effort. Particularly prominent among the challenges of introgressive breeding appear to be a high level of epistasis (nonlinear interactions between unlinked loci)—in many cases individual QTL loci in combination with different unlinked introgressed alleles affect a phenotype to very different degrees. In a few cases, the same QTL allele has even shown statistically significant opposite effects in different genetic backgrounds (P. Chee, X. Draye, A.H. Paterson, in preparation). This complication, superimposed on the effects of linkage drag, hybrid breakdown, and other well-known characteristics of wide crosses all highlight the fact that successful introgressive breeding of quantitative traits will require a dedicated effort that is designed to address such challenges, rather than being a 'side project' in a mainstream plant breeding program. Further, this will clearly be a high-risk, high-reward activity, in the sense that many promis-

Andrew H. Paterson, Plant Genome Mapping Laboratory, Univ. of Georgia, 111 Riverbend Road, Athens GA 30602; Steven M. Brown, Peng W. Chee, and O. Lloyd May, Dep. of Crop and Soil Sciences, Univ. of Georgia, Coastal Plain Exp. Station, Tifton, GA 31793; Randal K. Boman and John R. Gannaway, Dep. of Soil and Crop Sciences, Texas A&M Research and Extension Center, Lubbock TX 79403; Alan R. Gingle, Center for Applied Genetic Technologies, Univ. of Georgia, 111 Riverbend Road, Athens GA 30602; C. Wayne Smith, Dep. of Soil and Crop Sciences, Texas A&M University, College Station TX 77843. Received 23 July 2003. *Corresponding author (paterson@plantbio.uga.edu).

ing early leads will not translate into commercially viable products … but those that do may confer larger incremental gains than are available in the present cultivated gene pool.

Implementing this solution requires integration of research, education, and extension activities. Genetic vulnerability is a complex problem that results from a crop's evolutionary history, trends in breeding and biotechnology practices, and grower decisions based on inadequate information being available, all responding to the inevitable pressures of processor and consumer preferences. Gaining the partnership of stakeholders is key. Toward this end, while the introgressive breeding process proceed, we are working with both researchers and extension personnel to create a Web-based resource to provide objective information about relatedness of genotypes, as a management tool for producers to better deploy the remaining variation in the gene pool to minimize genetic vulnerability of the crop, and a research tool for scientists to glean new information useful for crop improvement.

ACKNOWLEDGMENTS

The authors acknowledge the support of an IFAFS grant from CSREES USDA.

REFERENCES

Bowman, D.T., O.L. May, and D.S. Calhoun. 1996. Genetic base of upland cotton cultivars released between 1970 and 1990. Crop Sci. 36:577–581.

Bowman, D.T., O.L. May, and D.S. Calhoun. 1997. Coefficients of parentage for 260 cotton cultivars released between 1970 and 1990. USDA Tech. Bull. 1852.

Calhoun, D.S., D.T. Bowman, and O.L. May. 1997. Pedigrees of upland and pima cotton cultivars released between 1970 and 1995. Miss. Agric. & Forestry Exp. Sta. Tech. Bull. 1069:53.

Helms, A.B. 2000. Yield study report. *In* P. Dugger and D. Richter (ed.) Proc. Beltwide Cotton Prod. Conf., San Antonio. TX. 4–9 Jan. 2000. Natl. Cotton Council, Memphis, TN.

May, O.L., D.T. Bowman, and D.S. Calhoun. 1995. Genetic diversity of U.S. Upland cotton cultivars released between 1980 and 1990. Crop Sci. 35:1570–1574.

Rayburn, S.T., R. Britton, and E. Keene. 1999. National Cotton Variety Tests. USDA-ARS, Stoneville, MS. Available at http://msa.ars.usda.gov/ms/stoneville/cgpr/rayburn.html (verified 6 July 2004).

Saranga, Y., M. Menz, C. Jiang, R. Wright, D. Yakir, and A.H. Paterson. 2001. Genomic dissection of genotype × environment adaptation conferring adaptation of cotton to arid conditions. Genome Res. 11:1988–1995.

National Agricultural Statistics Service. 1998. Anual yield report. USDA-NASS, Washington, DC.

Van Esbroeck, G., D.T. Bowman, D.S. Calhoun, and O.L. May. 1998. Changes in genetic diversity of cotton in the U.S. from 1970 to 1995. Crop Sci. 38:33–37.

The Role of Genomics Research in Improvement of "Orphan" Crops

Rebecca J. Nelson, Rosamond L. Naylor, and Molly M. Jahn*

The importance of agriculture to global food security goes beyond the need for total growth in crop yields and production. Agriculture promotes food security because it fulfills nutritional needs and/or contributes to local incomes and employment. Poverty in the developing world remains most pronounced in rural areas where agriculture is one of few sources of income and employment. The world's poorest regions are typically those where agricultural investments by the public and private sectors are extremely low. There is an urgent need for mechanisms to enhance agricultural development poor agrarian societies (Mosher, 1966).

In addition to a small number of well-known major global crops such as maize (*Zea mays* L.), rice (*Oryza sativa* L.), and wheat (*Triticum aestivum* L. em. Thell.), many more crops are regionally or locally important for nutrition and income in poor regions. Crops such as plantain and bananas (*Musa* sp. L.); root and tuber crops such as cassava (*Manihot esculenta* Cranz.), sweet potato [*Ipomoea batatas* (L.) Lam.], and yam (*Dioscorea* sp. L.); millets such as pearl millet [*Pennisetum glaucum* (L.) R. Br]), finger millet [*Eleusine coracana* (L.) Gaertn.], and foxtail millet [*Setaria italica* (L.) Beauv.]; legumes such as cowpeas [*Vigna unguiculata* (L.) Walp], groundnut (*Arachis hypogaea* L.) and Bambara groundnut [*Vigna subterranea* (L.) Verdc.]; and tree crops. Moreover, indigenous crops such as tef [*Eragrostis tef* (Zucc.) Trotter], quinoa (*Chenopodium quinoa* Willd.), and many types of vegetables are critical for food security and nutrition on a regional or local basis.

Twenty-five such "orphan" crops within developing countries total some 240 million hectares, with an additional 70 million hectares planted to fruits and vegetables (Naylor et al., 2004). In Sub-Saharan Africa, for example, sorghum [*Sorghum bicolor* (L.) Moench.] and pearl millet are more important than rice and wheat, both in area (41 million ha. vs. 9 million ha.) and in contribution to diet. Roots and tubers are essential staples in Africa, where cassava is the third most important source of calories overall. The underresearched crops are nutritious, valued culturally, adapted to harsh environments, and diverse in terms of their genetic, agroclimatic, and economic niches. Attention to locally important crops takes on added urgency given that 38% of Sub-Saharan Africa's population is undernourished, and the number of undernourished children in that region is expected to increase from present levels by 39% by 2020 (Pinstrup-Anderson et al., 1999).

A large discrepancy exists between the potential role of these crops in improving food security and livelihoods,

R.J. Nelson, Dep. of Plant Pathology, 321 Plant Science, Cornell Univ., Ithaca, NY, 14853; M.M. Jahn, Dep. of Plant Breeding, 313 Bradfield Hall, Cornell Univ., Ithaca, NY, 14853; R.L. Naylor, Center for Environmental Science and Policy, Stanford Univ., Encina Rm. 418E, Stanford, CA, 94305. Received 23 July 2003. *Corresponding author (mmj9@cornell.edu).

Abbreviations: EST, expressed sequence tag; MAS, marker-assisted selection; QTL, quantitative trait loci.

and the low levels of investment they have received. One reason for this may be that research on orphan crops may appear to have relatively low returns when measured by gross economic and welfare impacts, a view that stems in large part from inadequate measurement. The use of alternative metrics, e.g., human capital development, cropping system stability, the promotion of genetic diversity, all of which increase the capability of agricultural systems to withstand major biotic, abiotic, policy- or economic-induced shocks—provides even greater incentives to fund orphan crop germplasm improvement (Conway, 1997). While we believe these arguments offer compelling justification to enhance investment levels in crops other than wheat, rice, maize and soybean [*Glycine max* (L.) Merr.], clearly the contributions of major crops to human well-being are immense. No argument in this paper should be interpreted as suggesting that current research on them is excessive or even close to adequate.

Advances in crop genomics have resulted in a more unified understanding of the biology of the entire plant kingdom, as well as a powerful set of molecular and bioinformatic tools and methods. Such advances provide an opportunity for efficient transfer of information systems from model species and major crops to orphan crops (Naylor et al., 2004). As a result, relatively small investments in the transfer of advanced science from major crops to larger sets of orphan crops may potentially result in disproportionately high payoffs in terms of crop production, yield stability, and food security in least developed countries. It is important to emphasize that investment in genomics for a given species is only likely to be useful if a strong conventional breeding effort exists (and unfortunately, this prerequisite is too often not fulfilled).

There may also clearly be reciprocal benefits of genomics research on orphan crops for improvement of major crops, derived from insights into the genetic bases for their distinctive attributes. That is, some of the orphan crops can provide good models for traits not possessed by the model crops. Superior alleles for drought resistance, for instance, might be found in pearl millet and utilized by direct gene transfer in rice or wheat (Goodman et al., 1987). Alleles contributing tolerance to poor soils might be found in cowpea and used in other legumes.

Scientific Opportunities for Applying Advanced Technologies to Orphan Crops

Rationalizing investments in germplasm improvement for orphan crops requires a shift in investment perspective from individual crops to whole sets of crops with common genetic structures and from specific trait-crop combinations to consideration of a particular trait and its component attributes in a wide array of crops that may face similar production constraints. How important will research on models—such as rice, maize, *Arabidopsis* or *Medicago truncatula* Gaertn.—be for future improvements of orphan crop species? Will upstream research on mechanisms of plant responses to biotic and abiotic stress provide broadly applicable strategies for limiting crop loss? Will it be possible to integrate new plant traits and other findings into the ongoing, if limited, crop improvement efforts already underway in least developed countries? The benefits of transferring genomics information and techniques from model to orphan crops could take one or more of several forms: (i) improved analysis of crop biodiversity and identification of potentially useful variants, (ii) marker-assisted selection (MAS) of desired alleles and allele combinations, and (iii) cloning and direct transfer of desirable alleles among taxa.

Farmers and plant breeders have used visual selection as a fundamental tool in crop improvement for millennia. MAS has been demonstrated for a modest but increasing number of cases, and is most likely to be useful when genetic variability is obscure, phenotypes are difficult or expensive to evaluate, or where detectable variation is result of complex interactions of many genes and/or gene products. In only a few cases has a rigorous cost–benefit analysis been presented (e.g., Dreher et al., 2003).

Existing genetic variability in species can now be both identified and used in new ways for germplasm improvement. For example, any two plants from a group sharing a similar phenotype may or may not have genetic differences that would make it possible to recombine their genes to achieve a superior combination. Molecular techniques permit the visualization of molecular variation, which may allow a breeder to select the best possible parents for a crossing program. Useful gene variants may be present in plants with unpromising phenotypes, and molecular analysis of specific loci may allow cryptic but potentially useful genes to be discovered. Both these situations undoubtedly contribute to the phenomenon long apparent to plant breeders as "transgressive segregation" (Frantz and Jahn, 2004; de Vicente and Tanskley, 1993).

Imagine, for instance, that a researcher would like to improve the starch or vitamin content of a certain crop about which relatively little is known. Typically, the breeder has access to a large germplasm collection that has not been well characterized or utilized. It would make sense to analyze the collection for the phenotype of interest. Once a large group of individuals with known phenotypes has been established, it may be worthwhile to characterize the plants with a panel of markers representing the genes controlling starch and vitamin biosynthesis. Genotypes with different gene variants might be good candidates for entry into a breeding program.

To what extent is this process possible in current practice, for any crops? Progress in the area of plant genomics has been dramatic and the stage is set for efficient application of marker-assisted genetics, candidate gene analysis, and molecular breeding. Within plant families, similarities of genes and their physical organization on the chromosomes has already made it possible to use information from model species as a platform from which to pursue rapid progress on lesser-studied species. To date, however, the full impact of these technologies has yet to be felt in any crops, and it remains unclear how far-reaching results from one particular plant species will be across the whole plant kingdom.

Emerging evidence indicates that genomes for the

entire plant kingdom have much in common in terms of gene content, biochemical pathways, and chromosome organization. Genes involved in many biochemical pathways and processes are similar across the plant kingdom (Thorup et al., 2000). Functions such as gene regulation, general metabolism, nutrient acquisition, disease resistance, general defense, flowering time, and flower development are largely conserved across taxa. Comparative mapping studies reveal that gene order is conserved for chromosomal segments among grass species (Bennetzen and Freeling, 1998; Gale and Devos, 1998; Devos and Gale, 2000). Though weaker, chromosomal colinearity is detectable between monocots and dicots (Bennetzen, 2000; Devos et al., 1999; Goff et al., 2002).

Most traits of importance to farmers and consumers are governed by multiple genes of relatively small individual effects. These "quantitative traits" are the most difficult to understand and improve. Molecular genetic approaches have begun to illuminate the genetic architecture of quantitative traits (Paterson et al., 1988; Kearsey and Farquhar, 1998). Although MAS for these traits using anonymous QTL-associated markers is more challenging than was initially projected, because of the imprecise localization of QTL and by inconsistent QTL expression, recent studies have provided encouraging evidence that MAS may be useful for enhancing these traits under certain circumstances (e.g., Han et al., 1997; Bouchez et al., 2002; Villanueva et al., 2002; Mithen et al., 2003; Zhou et al., 2003).

Candidate genes, genes known or suspected to be involved in conditioning the phenotype of interest, make it possible to localize desirable variants much more precisely. Credible candidate genes have now been identified for many plant traits, including quantitative (multiple gene) disease resistance in rice (Wang et al., 2001; Ramalingam et al., 2003), wheat (Faris et al., 1999), bean (*Phaseolus vulgaris* L.; Geffroy et al., 2000), and potato (*Solanum tuberosum* L.; Trognitz et al., 2002). A number of research approaches have converged to allow genes underlying QTLs to be cloned (Frary et al., 2000; Johanson et al., 2000; El-Assal et al., 2001; Thornsberry et al., 2001). Isolation of genes controlling quantitative traits will permit both the identification of potentially useful variants of agronomically important genes and the precise selection of the most useful alleles. The availability of the isolated genes could allow natural molecular variation to be analyzed efficiently in a range of genotypes, enabling the identification of potentially useful variants for future use.

Sequence data on expressed genes and on plant and crop genomes are rapidly accumulating and present powerful tools for plant science. The increasing availability of expressed sequence tags (ESTs) puts QTL cloning within reach. EST collections also provide the basis for microarray technology that allows patterns of gene expression to be investigated in various physiological conditions, another potentially promising source of candidate genes. Combining information on mapped QTLs and ESTs provides a step toward identifying the genes that underlie quantitative trait loci. Although sequence datasets are, in themselves, imposing and cumbersome, increasingly powerful and friendly databases (e.g., Yuan et al., 2001) allow researchers to access genetic information and identify and exploit natural variation in ways previously not possible. For orphan crops, however, numbers of ESTs are meager.

While it is often possible to associate a candidate gene with a QTL, it is not so easy to actually prove that the candidate contributes to the expression of the trait of interest (Glazier et al., 2002). The number of recombination events in a mapping population is often insufficient to permit the identification of genes underlying a QTL with high resolution. QTL estimation often spans several centimorgans, and hundreds of genes will underlie a region of this size. The size of such a region can be reduced through a number of approaches, such as the use of high-resolution crosses, or the development of near-isogenic lines for small chromosomal segments across the putative QTL region. Linkage disequilibrium mapping offers another alternative, exploiting the long history of recombination and rich allelic diversity in collections of diverse germplasm (Remington et al., 2001; Buckler and Thornsberry, 2002). For example, a specific polymorphism in the *Dwarf8* gene (a gene known to affect plant height) was shown to associate with variation for flowering time in maize by this type of approach (Thornsberry et al., 2001).

Science in Context

Mass selection of landraces for desired traits generally has not kept pace with globalization or even with changes in local conditions (including population growth, changing tastes, new pest and disease pressures, and abiotic stresses). To assist poor rural communities in generating local opportunities and income, there exist great opportunities—and also major challenges—for plant breeding interventions (DeVries and Toenniessen, 2001). Insights and tools with practical utility for orphan crops can be obtained from research into both basic and applied plant biology using model species and major crops. Such transfer of technology from major or model crops to orphan crops will be cost-efficient, but will still require significant fixed costs up front in developing the basic biology of the orphan crops in question.

Success will depend on investment but also on appropriate integration of knowledge gained (Naylor et al., 2004). Integration starts with linking advanced science with plant breeding and seed programs. While the link between science and plant breeding is key, so too is the link between plant breeding, farmers, delivery systems, and consumers. Successful application of genomics is conditional on connecting the science to downstream delivery efforts. For the poorest countries, such integration may take years to achieve. Even with appropriate integration and sustained research investments, the benefits from advanced science depend critically on institutional, human capital, economic, and political contexts in regions that require agricultural growth.

ACKNOWLEDGMENTS

The authors acknowledge a USDA IFAFS Plant Genome Award No. 2001-52100-11347.

REFERENCES

Bennetzen, J.L. 2000. Comparative sequence analysis of plant nuclear genomes: Microcolinearity and its many exceptions. Plant Cell 12:1021–1029.

Bennetzen, J.L., and M. Freeling. 1998. The unified grass genome: Synergy in synteny. Genome Resour. 7:301–306.

Bouchez, A., F. Hospital, M. Causse, A. Gallais, and A. Charcosset. 2002. Marker assisted introguression of favorable alleles at quantitivative trait loci between maize elite lines. Genetics 162:1945–1959.

Buckler, E.S., and J.M. Thornsberry. 2002. Plant molecular diversity and applications to genomics. Curr. Opin. Plant Biol. 5:107–111.

Conway, G. 1997. The doubly green revolution: Food for all in the Twenty-First Century. Cornell University Press, Ithaca, NY.

Devos, K.M., J. Beales, Y. Nagamura, and T. Sasaki. 1999. *Arabidopsis*-Rice: Will colinearity allow gene prediction across eudicot-monocot divide? Genome Res. 9:825–829.

Devos, K.M., and M.D. Gale. 2000. Genome relationships: The grass model in current research. Plant Cell 12:637–646.

DeVries, J., and G. Toenniessen. 2001. Securing the harvest: Biotechnology, breeding and seed systems for African crops. CABI Publishing, New York.

Dreher, K., M. Khairallah, J.-M. Ribaut, and M. Morris. 2003. Money matters (I): Costs of field and laboratory procedures associated with conventional and marker-assisted maize breeding at CIMMYT. Mol. Breed. 11:221–234.

El-Assal, E.S., C. Alonso-Blanco, A.J.M. Peeters, V. Raz, and M. Koornneef. 2001. A QTL for flowering time in Arabidopsis reveals a novel allele of CRY2. Nat. Genet. 29:435–440.

Faris, J.D., W.L. Li, D.J. Liu, P.D. Chen, and B.S. Gill. 1999. Candidate gene analysis of quantitative disease resistance in wheat. Theor. Appl. Genet. 100:190–198.

Frary, A., T C. Nesbitt, A. Frary, S. Grandillo, E. van der Knaap, B. Cong, J. Liu, J. Meller, R. Elber, K.B. Alpert, and S.D. Tanksley. 2000. *fw2.2*: A quantitative trait locus key to the evolution of tomato fruit size. Science 289: 85–88.

Frantz, J.D., and M.M. Jahn. 2004. Five independent loci each control monogenic resistance to gummy stem blight in melon (*Cucumis melo* L.). Theor. Appl. Genet. 108:1033–1038.

Gale, M., and K. Devos. 1998. Plant comparative genetics after 10 years. Science 282:656–659.

Geffroy, V., M. Sévignac, J.C.F. De Oliveira, G. Fouilloux, P. Skroch, P. Thoquet, P. Gepts, T. Langin, and M. Dron 2000. Inheritance of partial resistance against *Colletotrichum lindemuthianum* in *Phaseolus vulgaris* and co-localization of quantitative trait loci with genes involved in specific resistance. Mol. Plant-Microb. Interact. 13:287–296.

Glazier, A.M., J.H. Nadeau, and T.J. Aitman. 2002. Finding genes that underlie complex traits. Science 298:2345–2349.

Goff, S.A., D. Ricke, T. Lan, G. Presting, R. Wang, M. Dunn et al. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science 296(5655):92–100.

Goodman, R.M., H. Hauptli, A. Crossway, and V.C. Knauf. 1987. Gene transfer in crop improvement. Science 236:48–54.

Han, F., I. Romagosa, S.E. Ullrich, B.L. Jones, P.M. Hayes, and D.M. Wesenberg. 1997. Molecular marker-assisted selection for malting quality traits in barley. Mol. Breed. 3:427–437.

Johanson, U., J. West, C. Lister, S. Michaels, R. Amasino, and C. Dean. 2000. Molecular analysis of FRIGIDA, a major determinant of natural variation in Arabidopsis flowering time. Science 290: 344–347.

Kearsey, M.J., and A.G.L. Farquhar. 1998. QTL analysis in plants; where are we now? Heredity 80:137–142.

Mithen, R., K. Faulkner, R. Magrath, P. Rose, G. Williamson, and J. Marquez. 2003. Development of isothiocyanate-enriched broccoli, and its enhanced ability to induce phase 2 detoxification enzymes in mammalian cells. Theor. Appl. Genet. 106:727–734.

Mosher, A. 1966. Getting agriculture moving. Praeger Publishers, New York.

Naylor, R.N., W.P. Falcon, R.M. Goodman, M.M. Jahn, T. Sengooba, H. Tefera et al. 2004. Biotechnology in the developing world: A case for increased investments in orphan crops. Food Policy 29:15–44.

Nelson, R.S., P.P. Abel, and R.N. Beachy. 1986. Characterization of the virus resistance in transgenic tabacco plants expressing tabacco mosaic virus coat protein. Plant Physiol. 80(4 Suppl):103.

Paterson, A.H., E.S. Lander, J.D. Hewitt, S. Peterson, S.E. Lincoln, and S.D. Tanksley. 1988. Resolution of quantitative traits into Mendelian factors using a complete linkage map of restriction fragment length polymorphisms. Nature 335:721–726.

Pinstrup-Anderson, P., R. Pandya-Lorch, and M.W. Rosegrant. 1999. World food prospects: Critical issues for the early 21st century. Food Policy Report. International Food Policy Research Institute, Washington, DC.

Ramalingam, J., C.M.V. Cruz, K. Kukreja, J.M. Chittoor, J.L. Wu, S.W. Lee et al. 2003. Candidate defense genes from rice, barley, and maize and their association with qualitative and quantitative resistance in rice. Mol. Plant-Microb. Interact. 16:14–24.

Remington, D.L., J.M. Thornsberry, Y. Matsuoka, L.M. Wilson, S.R. Whitt, J. Doebley et al. 2001. Structure of linkage disequilibrium and phenotypic associations in the maize genome. Proc. Natl. Acad. Sci. USA 98:11479–11484.

Thornsberry, J.M., M.M. Goodman, J. Doebley, S. Kresovich, D. Nielsen, and E.S. Buckler. 2001. *Dwarf8* polymorphisms associate with variation in flowering time. Nat. Genet. 28:286–289.

Thorup, T.A., B. Tanyolac, K.D. Livingstone, S. Popovsky, I. Paran, and M. Jahn. 2000. Candidate gene analysis of organ pigmentation loci in the Solanaceae. Proc. Natl. Acad. Sci. USA 97:11192–11197.

Trognitz, F., P. Manosalva, R. Gysin, D. Niño-Liu, R. Simon, M. del Rosario Herrera et al. 2002. Plant defense genes associated with quantitative resistance to potato late blight in *Solanum phureja* × *S. tuberosum* hybrids. Mol. Plant-Microb. Interact. 15:587–597.

de Vicente, M.C., and S.D. Tanskley. 1993. QTL analysis of transgressive segregation in an interspecific tomato cross. Genetics 134: 585–596.

Villanueva, B., R. Pong-Wong, and J. Wooliams. 2002. Marker-assisted selection with optimised contributions of the candidates to selection. Genet. Sel. Evol. 34:679–703.

Wang, Z., G. Taramino, D. Yang, G. Liu, S.V. Tingey, C.-H. Miao et al. 2001. Rice ESTs with disease-resistance gene or defense-response gene-like sequences mapped to regions containing major resistance genes or QTLs. Mol. Genet. Genom. 265:302–310.

Yuan, Q., J. Quackenbush, R. Sultana, M. Pertea, S.L. Salzberg, and C.R. Buell. 2001. Rice bioinformatics: Analysis of rice sequence data and leveraging the data to other plant species. Plant Physiol. 125:1166–1174.

Zhou, P.H., Y.F. Tan, Y.Q. He, C.G. Xu, and Q. Zhang. 2003. Simultaneous improvement for four quality traits of Zhenshan 97, an elite parent of hybrid rice, by marker-assisted selection. Theor. Appl. Genet. 106:326–331.

Applying Genomics to Alfalfa Breeding Programs

E. Charles Brummer*

ALFALFA, *Medicago sativa* L., is a herbaceous, perennial forage crop grown extensively throughout temperate and dry tropical regions of the world for hay, pasture, and silage. More than any other forage crop adapted to these regions, alfalfa combines high biomass productivity, optimal nutritional profiles, and adequate survival, making its cultivation ideal for dairy and livestock enterprises. Within the context of a cropping system, alfalfa controls soil erosion, improves water quality, mitigates pest outbreaks, and contributes significant

E. Charles Brummer, Raymond F. Baker Center for Plant Breeding, Dep. of Agronomy, Iowa State Univ., Ames, IA 50011. Received 23 July 2003. *Corresponding author (brummer@iastate.edu).

amounts of nitrogen to succeeding crops. Although commercial breeding programs often target marketable traits for improvement, major agronomic traits of importance are biomass yield and winter hardiness. To be profitably grown, alfalfa needs to sustain the production of high yields over a several year period. Corn silage represents the primary alternative forage to alfalfa for dairy cattle rations, and though it can produce more dry matter yield than alfalfa it requires protein supplementation. Further, its cultivation leaves cropland devoid of cover during the winter months and susceptible to significant wind and water soil erosion. Thus, improved alfalfa cultivars can have positive economic and environmental impacts on the agricultural sector.

A negative relationship exists between biomass yield and winter hardiness, complicating improvement efforts on both traits simultaneously. In response to temperature and photoperiod cues in autumn, alfalfa plants acclimate for winter, making physiological alterations that affect obvious phenotypes such as plant height and biomass production. Nondormant plants have little or no acclimation response and are typically more productive than dormant plants during autumn. Plant height in October is typically used in the upper midwestern USA as an indirect measure of dormancy response, with shorter plants being more dormant. Height is associated with winter injury when considering the entire spectrum of dormancy responses, but we have found no genetic correlation between the traits in a population derived from dormant parents but which segregated for both traits (Brummer et al., 2000). Thus, to some extent, these traits can be selected independently, although new selection methods may be needed to do it efficiently.

We are attempting to address the biomass yield and winter injury trade-off by augmenting traditional selection methods with genomics. Our efforts are focused in three areas: (i) identification of quantitative trait loci (QTL) and candidate loci associated with the traits, (ii) profiling germplasm sources to identify novel alleles and to structure breeding programs to capture heterosis, and (iii) isolation of genes associated with dormancy control to eventually hasten adaptation of genetic resources to diverse environments.

Effective selection for yield and winter hardiness requires multiple year field evaluations to ensure long-term persistence and sustained yield; consequently, genetic improvements accrue slowly. Selection based on molecular markers offers one means of decreasing the cycle time, and we are attempting to identify loci associated with these traits using standard genetic mapping and QTL detection procedures (Robins et al., 2003). In addition to mapping these ultimate phenotypes, we are dissecting both traits by concurrently mapping morphological and physiological components, such as sugar, starch, and protein content in roots during autumn (Alarcón-Zúñiga et al., 2004). Our hope is that these component traits, which may have a simpler genetic basis, will allow us to manipulate the overall complex trait in a targeted, modular manner. Finally, we are mapping candidate genes known or suspected to be associated with winter injury in these same populations to assign putative functions to QTL. These mapping studies, being conducted in both tetraploid and diploid populations, will provide a picture of the genomic landscape of these important traits and identify options for improving them.

Besides a marker-assisted selection approach, a possible means to increase yield is to capture heterosis, which currently is not being done in commercially available synthetic cultivars (Brummer, 1999). To capture heterosis, genetically distinct populations, or heterotic groups, need to be identified and improved independently. Three possible heterotic groups include *M. sativa* subsp. *falcata*, dormant subsp. *sativa*, and nondormant subsp. *sativa*. We have recently shown that crosses between falcata and dormant sativa results in significant heterosis for biomass yield (Riday and Brummer, 2002). Molecular markers have not been successful at predicting heterosis, but were more effective at differentiating the subspecies (Riday et al., 2003; Riday and Brummer, 2004). Although clear morphological differences can easily separate falcata and sativa, our results suggest that markers may not be useful to place germplasm into possible heterotic groups within each of the subspecies.

Both falcata and nondormant sativa have agronomic weaknesses for the upper Midwest: the subsp. *falcata* has slow regrowth, particularly in the late summer and autumn, and nondormant sativa is not winter hardy. Even though hybrids between falcata and elite cultivars produce high yields and exhibit heterosis for seasonal total yield, their superior performance dissipates as the growing season progresses and disappears under a harvest regime with short regrowth periods. Hybrids of dormant and nondormant sativa cannot be evaluated effectively, as winter injury obfuscates the results. These problems are associated with the dormancy response—too strong in falcata, and not strong enough in nondormant sativa. Traditional breeding, possibly aided by markers, may be useful to improve autumn growth in falcata germplasm and is clearly effective in improving winter hardiness of nondormant sativa (Weishaar et al., 2005).

Modifying the dormancy response of falcata or nondormant sativa through transgenic means might offer a way to engineer nonadapted germplasm for use directly in breeding programs outside its area of adaption. We are attempting to differentiate between genes involved in temperature and photoperiod sensing using *M. truncatula* Gaertner microarrays probed with RNA from dormant and nondormant alfalfa genotypes pre- and postacclimation in the field and in the growth chamber where one of the variables (temperature or photoperiod) was held constant. Genes identified from this screen could also be useful for marker-assisted selection aimed at altering the dormancy response of various populations.

The approaches we are using will identify loci involved in biomass yield and winter hardiness. These markers, QTL, and genes can be used as selection tools or as transgenes to aid the improvement of these genetically complex traits. Applying this information to a breeding program will not be as straightforward in alfalfa as in many other crops. Commercial alfalfa cultivars are synthetic populations, consisting of a heterogeneous mix of heterozygous genotypes (inbred lines are not available). Further, cultivated alfalfa has an auto-

tetraploid ($2n = 4x = 32$) genome. Both of these characteristics complicate the application of genomics solutions to practical breeding problems.

Identifying most of the loci within a breeding population involved in biomass yield, winter hardiness, and other complex traits will require more complicated pedigree structures than simple biparental crosses. Methods to construct and analyze pedigree based populations have been developed for diploid organisms (Yi and Xu, 2001), but extension to autotetraploids remains to be done. Regardless of the theoretical possibilities, building complex populations and accumulating phenotypic and genotypic data to successfully identify important loci will require an effort beyond the capacities of most public and private alfalfa laboratories. Selection based on markers alone will have to consider linkage disequilibrium (about which little is known) unless the markers are the genes of interest. Given the difficulties of identifying haplotypes in autotetraploid plants, all efforts to use candidate genes should be pursued. Association mapping may be a more sensible approach than linkage analysis, although it too will be complicated by tetraploidy and by relatively low marker density. Perhaps marker-assisted selection will be more useful for the identification and introgression of novel alleles from exotic germplasm into elite breeding populations. Manipulating the frequency of introduced alleles and minimizing linkage drag could be done rather expeditiously even in tetraploid populations.

The recent expansion of basic research on the model legume *M. truncatula* bodes well for the alfalfa community if the genomics tools developed in the model can be usefully applied to crop improvement. The major traits of winter survival, multiyear persistence, regrowth, biomass yield, and seed yield cannot be thoroughly assessed in the autogamous, annual, diploid model species. Thus, in order for genomics tools to be used to develop better alfalfa cultivars for farmers, breeders will need to apply the technology directly in alfalfa improvement programs. However, the infrastructure developed in *M. truncatula*, including the development of genetic markers and the analysis of certain biochemical pathways, will be directly applicable to alfalfa.

Manipulation of single gene traits by biotechnological approaches has been clearly successful, even in alfalfa. Glyphosate-tolerant (Roundup Ready, Monsanto Corp., St. Louis, MO) alfalfa will be commercialized in the near future and other value-added traits lie on the horizon (M. McCaslin, Forage Genetics, Intl., pers. comm.). Commercial emphasis on the use of genomics will lie in those traits which provide a marketing advantage, but serious efforts to improve fundamentally important but inherently complex traits such as winter hardiness or biomass yield likely will not be conducted in the private sector. For long-term improvement of alfalfa, research should be focused in four main areas: (i) constructing a comprehensive picture of both cultivated and wild genetic resources, (ii) streamlining selection procedures to shorten cycle time and increase heritabilities, (iii) developing alternative cultivar types that harness the genetic potential within and among germplasm groups, and (iv) facilitating the creation and maintenance of genetic variation for major quantitative traits in diverse breeding populations. The use of genomics methods to characterize exotic germplasm, dissect quantitative traits, and identify candidate loci could address each of these problems. The incorporation of genomics techniques into the breeding process will be challenging, but many of these problems can be overcome with enough effort.

As we consider using genomics to improve crop breeding, we need to ask whether we are using genomics tools to solve serious breeding problems, or inventing breeding problems that are solvable by genomics methods. Many essential breeding goals can continue to be met with traditional methods if effort is devoted to them. The paradox of the genomics age is that funding for plant breeding programs is decreasing at the same time that the potential of genomics is being realized. Thus, genomics initiatives have replaced, rather than augmented, breeding programs, with the result that many technological advances in genomics may not be applied to cultivar development at all! Many breeding programs in both the public and private sector have limited funding that constrains their ability to conduct extensive multienvironment selection and evaluation programs and to produce varieties. Given financial constraints, emphasis is often diverted to genomics projects for which money can be procured. This results in the somewhat ridiculous situation of high technology solutions being developed for a crop that lacks the most basic breeding capabilities. Further, minor crops, including many forage and noncommodity species, have little focused breeding effort devoted to them in the public sector and virtually none from industry. Without active and effective breeding programs, genomics will not contribute to genetic gains in any traits.

ACKNOWLEDGMENTS

This research is supported by USDA-IFAFS Grant 00-52100-9611.

REFERENCES

Alarcón-Zúñiga, B., E.C. Brummer, P. Scott, K. Moore, and D. Luth. 2004. Quantitative trait loci mapping of winter hardiness metabolites in autotetraploid alfalfa. p. 97–104. *In* A. Hopkins et al. (ed.) Molecular breeding of forage and turf. Kluwer, Dordrecht, the Netherlands.

Brummer, E.C. 1999. Capturing heterosis in forage crop cultivar development. Crop Sci. 39:943–954.

Brummer, E.C., M.M. Shah, and D. Luth. 2000. Re-examining the relationship between fall dormancy and winter hardiness in alfalfa. Crop Sci. 40:971–977.

Riday, H., and E.C. Brummer. 2002. Forage yield heterosis in alfalfa. Crop Sci. 42:716–723.

Riday, H., and E.C. Brummer. 2004. Dissection of heterosis in alfalfa hybrids. p. 315–324. *In* A. Hopkins et al. (ed.) Molecular breeding of forage and turf. Kluwer, Dordrecht, the Netherlands.

Riday, H., E.C. Brummer, T.A. Campbell, D. Luth, and P.M. Cazcarro. 2003. Comparisons of genetic and morphological distance with heterosis between *Medicago sativa* subsp. *sativa* and subsp. *falcata.* Euphytica 131:37–45.

Robins, J.G., D.R. Viands. T.A Campbell, D. Luth, and E.C. Brummer. 2003. Mapping forage biomass yield in an intersubspecific cross of alfalfa (*Medicago sativa*). Proc. 3rd Molecular Breeding of Forage

Turf Conf. Dallas, TX. 18–22 May 2003. Available at https://www.register-for.com/mbft/AbstractView.Asp?AbstractId=Abstract2b20 (verified 6 July 2004).
Weishaar, M.A., E.C. Brummer, J.J. Volenec, K.J. Moore, and S. Cunningham. 2005. Improving winter hardiness in nondormant alfalfa germplasm. Crop Sci. 45(1) in press.
Yi, N., and S. Xu. 2001. Bayesian mapping of quantitative trait loci under complicated mating designs. Genetics 157:1759–1771.

Genomics, Genetics, and Plant Breeding: A Private Sector Perspective

Mark Cooper,* Oscar S. Smith, Geoff Graham, Lane Arthur, Lizhi Feng, and Dean W. Podlich

Legacy of Phenotype-Based Pedigree Selection

Pedigree breeding strategies have been the basis for genetic improvement of corn (*Zea mays* L.) in Pioneer Hi-Bred from the foundation of the company in the 1920s through to the 1990s. Over this period, grain yield has undergone genetic improvement at a rate of around 75 kg ha^{-1} yr^{-1} (Duvick et al., 2004a, 2004b; Fig. 1a). It is widely understood that realized progress for grain yield in the U.S. corn belt has been an outcome of combining improved genetics with appropriate crop management strategies (e.g., plant populations). Systematic evaluations of the outcomes of this long-term corn breeding effort have shown that the performance phenotypes and genotypic composition of the elite germplasm pools of the breeding program can be changed by selecting directly on the trait phenotypes we seek to improve (e.g., Fig. 1; Duvick et al., 2004a, 2004b). Side-by-side phenotypic evaluations of a sequence of successful Pioneer corn hybrids, representing each decade from the 1930s to present, provides a description of the phenotypic changes for a number of the key traits that the breeders have directly or indirectly changed (Fig. 1a–c). Genetic fingerprints of the inbred parents of these hybrids provide a description of the genotypic changes that have occurred in association with the sustained breeding effort (Fig. 1d). Important phases can be identified over this period of breeding. Initially double-cross hybrids (1920s-1960s) were developed. From the 1960s there was a relatively rapid transition to the use of single-cross hybrids, the foundation of which was the organization of the corn germplasm into heterotic groups, represented in this example by the Stiff Stalk Synthetic (SS) and Non Stiff Stalk Synthetic (NSS) Groups (Fig. 1d).

Early Outcomes from Molecular Breeding

Advances in molecular genetics have reached a stage where breeding schemes can now be augmented with the use of a number of molecular technologies. Commercial breeding programs have and will continue to evaluate and invest in research that considers the prospects to either change or refine the in situ gene-to-phenotype system. From the 1990s, transgenic methods have been applied to key traits. Commercial transgenic hybrids have been developed for traits where there is a simple gene-to-phenotype relationship; e.g., Bt for insect resistance and multiple herbicide resistance genes. Some of the successful Pioneer Bt hybrids are included in the progression shown in Fig. 1. At this early stage of technology development it is tempting to conclude that transgenic solutions will only be applicable for trait targets under simple genetic control. However, this view is not accurate. Conventional genetic improvement of resistance to insect pests was viewed as a traditional complex trait problem before the widespread use of transgene sources of resistance to insect pests. Molecular marker based breeding strategies are being considered for improvement of simply inherited traits and quantitative traits that show complex inheritance in elite germplasm pools. Marker-assisted selection has been applied in breeder crosses when marker-trait associations are sufficient, the requisite marker polymorphism(s) can be identified and a high throughput system of assay is available. In most cases for complex traits the candidate regions identified by molecular markers have not been resolved to the level of candidate genes.

Organization of Genomics Efforts

Only over the last decade has the scientific community developed and had access to the range of molecular tools that provide the technological foundation that will be necessary to understand (i) the genetic architecture of the trait combinations we seek to manipulate, (ii) the nature of the genetic changes that were brought about by phenotypic selection, (iii) the power that can be attained in a breeding strategy (molecular and conventional) to achieve directed genetic changes that manipulate the trait phenotypes we seek to improve, and (iv) the limits that will ultimately be faced in using genetic technologies to make robust changes to plant phenotypes that improve the sustainability of agricultural systems.

Much of the genomic technological advancements used in plants were developed to meet the needs of the human genome effort. In most cases the application of these DNA-, RNA-, and protein-based technologies to study plant genomes has been straightforward. To take advantage of the opportunities that these genomic technologies provide to plant breeding, plant genomics efforts over the last decade have been heavily focused on plant specific gene discovery, gene function knowledge creation, and organization of the heterogeneous data sources that have emerged across the scientific community.

Creating a Molecular Breeding Focus

Today the concept of commercially successful molecular breeding is multifaceted and should be viewed as such. At its current stage of development as a proven

Pioneer Hi-Bred International Inc., 7250 N.W. 62nd Avenue, P.O. Box 552, Johnston, IA 50131, USA. Received 23 July 2003. *Corresponding author (mark.cooper@pioneer.com).

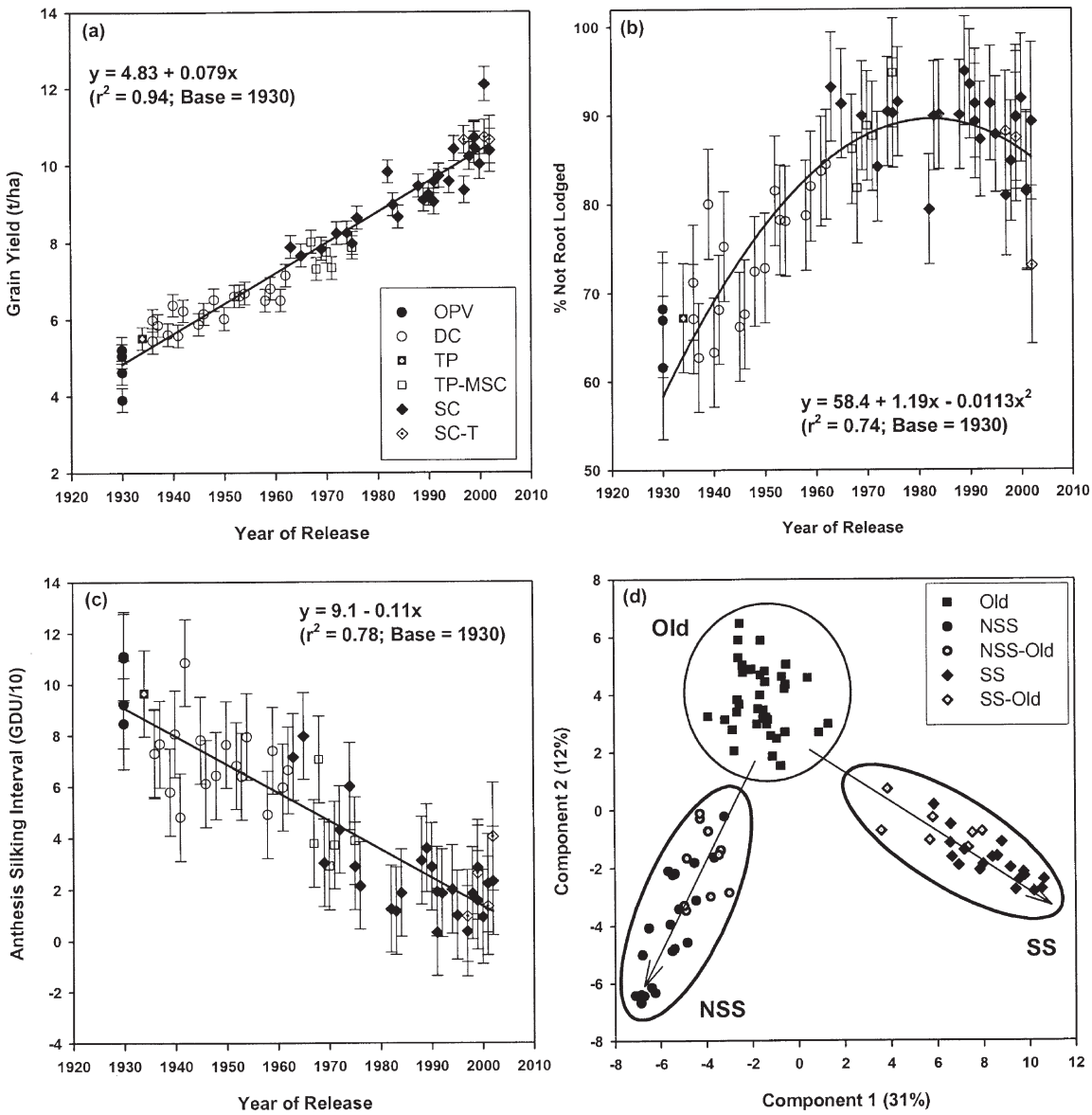

Fig. 1. Best linear unbiased predictors (BLUPs ± SE) for phenotypic performance of three traits, (a) grain yield (hybrids grown at three densities, 30, 54, and 79 thousand plants/ha, and yield per hybrid is for the density giving the highest average yield), (b) percentage of plants not root-lodged, (c) anthesis to silking interval, measured in experiments conducted from 1990 to 2002, for a sequence of commercially successful Pioneer corn hybrids taken from an ERA study (unpublished data), and (d) a plot of the inbred scores on the first two principal components from analysis of SSR molecular marker profiles of the parents of the hybrids. (OPV = open pollinated variety, DC = double cross, TC = three parent cross, TC-MSC = three parent modified single cross, SC = single cross, SC-T = single cross transgenic; SS = Stiff Stalk Synthetic inbred line, NSS = non-Stiff Stalk Synthetic inbred line). The large boundaries distinguish three main groups of lines; Old = the old inbred lines used before the formation of the heterotic groups and the other two groups represent SS and NSS inbred lines. The arrows indicate the direction of the progression of inbred improvement in the SS and NSS heterotic groups.

breeding methodology, the term "molecular breeding" is a collective descriptor of the heterogeneous efforts, challenges, and opportunities being investigated to enhance the short-term and long-term success of the systematic procedures used to improve trait phenotypes by directed manipulation of the genotype at the DNA sequence level. At this time, molecular breeding is not an identifier of a single general breeding approach in the same way that "pedigree breeding" is such an identifier. Thus, many different breeding approaches are considered under the title of molecular breeding. Two major components are in use today: (i) direct movement of genes between individuals by a range of transgenic approaches and (ii) development of associations between interindividual DNA sequence variation and trait phenotypic variation in combination with the design of DNA based prognostics that can be used in high throughput systems as a component of a breeding program. The feasibility and the range of successful outcomes from both approaches are being enhanced for a range of traits by greater fundamental knowledge of plant genome organization and the functional properties of genes.

Improving a Breeding Strategy

The concept of evaluating alternatives and building on the strengths of an incumbent strategy is not new to

plant breeding. The overriding motivation for considering molecular breeding strategies in place of conventional phenotypic–pedigree-based breeding strategies is that molecular-based selection does or with appropriate development will provide advantages over phenotype-based selection. Often many hidden assumptions are made in the theoretical discussions of the advantages that can be realized from molecular breeding strategies. One assumption that is often difficult to consider fully is the complexity of the genetics that the current strategy faces. Overly simplified genetic models can often give an associated overly optimistic assessment of the benefits, or in some cases lack of benefit, to be expected from an alternative strategy. Ultimately, validation by measuring realized benefits in situ are necessary. Because of the complex stochastic nature of the genotype–environment systems that breeding programs operate within, it has been resource intensive and difficult to demonstrate the advantage of one breeding strategy over another. A major difference between academic and commercial evaluations of molecular breeding strategies is the greater need by the commercial programs to make as many of the hidden assumptions that underlie the potential advantages and disadvantages as visible as possible for direct consideration. These advantages may come in the form of (i) reduced costs for achieving a given level of phenotypic improvement, (ii) improvements in the accuracy and precision with which we can make phenotypic changes, (iii) step–change improvements in phenotypes that were not previously accessible with comparable research investments into conventional breeding methods, and (iv) the identification of industry game-changing technologies for complex genotype–environment systems.

By emphasizing the need for a demonstrable advantage at the level of the commercial viability of breeding program outcomes, the criteria for success are set at a much higher level than would be the case if all that was required was a demonstration that genotype-based improvement of the phenotype, via manipulation of DNA sequence, was feasible. This is much the same process that was used by previous Pioneer breeders in judging the merits of alternatives to and refinements of the conventional pedigree-breeding program (Duvick et al., 2004a). Ultimately, for commercial breeding programs, the success of any alternative breeding strategy is based on the value that can be gained by all stakeholders from the improved phenotypes and the costs of attaining and maintaining these improved phenotypes.

Therefore, the challenge is to outperform the current breeding strategy for a wide range of situations. The range of approaches must work for the important traits, which will inevitably differ in genetic complexity. It is difficult to conduct comprehensive empirical evaluations of alternative breeding strategies for a large number of scenarios. An alternative approach is to use computer simulation (Cooper et al., 2002). Figure 2 provides a stochastic computer simulation comparison of a conventional phenotypic selection (PS) and marker-assisted selection (MAS) strategy for a series of putative quantitative trait models. In this figure the difference between the resulting performance of two groups of genotypes (Normalized difference in response; MAS-PS) selected after five cycles of breeding, using either MAS or phenotypic selection, is plotted against a measure of the genetic complexity of a trait [complexity here is quantified as an autocorrelation value estimated from sequences of genotypic values from random neighborhood walks in genetic space; see Cooper and Podlich (2002) for additional details]. The emphasis in this theoretical example is not on the details of the two breeding strategies, but on the average difference and variability in expected difference in response to selection between the two strategies for simple and complex genetic situations and the impact of both trait heritability and the level of knowledge of the genetic architecture of the trait that is available to the breeder to facilitate the implementation of the MAS strategy. Even though on average there is an advantage observed for MAS, this varies for different genetic architectures and also for different replicates for the same scenario.

The more effective the current breeding strategy the more difficult will be the challenge to outperform the incumbent strategy and demonstrate the advantages. As with most difficult challenges, the paths to improvement are many and any commitment to molecular breeding strategy development will be an iterative process. The commercial molecular breeding strategies we see today represent first or second cycle iterations of some of the potential paths to implementing molecular breeding strategies. These may be more accurately referred to as molecular enhanced breeding strategies, which apply molecular technologies around what are still predominantly large pedigree breeding strategies.

Foundations for Molecular Breeding: Integration of Genomics and Genetics

It is expected that the foundation for studying natural genetic variation at the DNA sequence level for traits in elite breeding germplasm pools and the design of successful molecular breeding strategies will involve integration of structural and functional genomics technologies within the framework of classical trait mapping methods. Through the latter half of the 1980s and the 1990s the development of a range of molecular marker technologies allowed mapping of traits to broad candidate regions [quantitative trait loci (QTL)] on genetic maps. In some cases continued investigation of these regions by fine mapping and ultimately direct sequencing of the resolved regions enabled identification and cloning of candidate genes and in a few cases verification of the causal gene and some knowledge of functional allelic variation. The availability of physical maps for a number of the important crop plants and the complete genome sequence for the model plant *Arabidopsis thaliana* (L.) Heynh. (The Arabidopsis Genome Initiative, 2000) and the crop plant rice (*Oryza sativa*; Goff et al., 2002; Yu et al., 2002) has enabled alignment of genetic maps with physical sequence and thus opportunity to target sequence data from genomic regions for gene discovery and gene function analysis. Further advances in technologies for the study of gene expression and protein interactions have opened up glimpses of some of

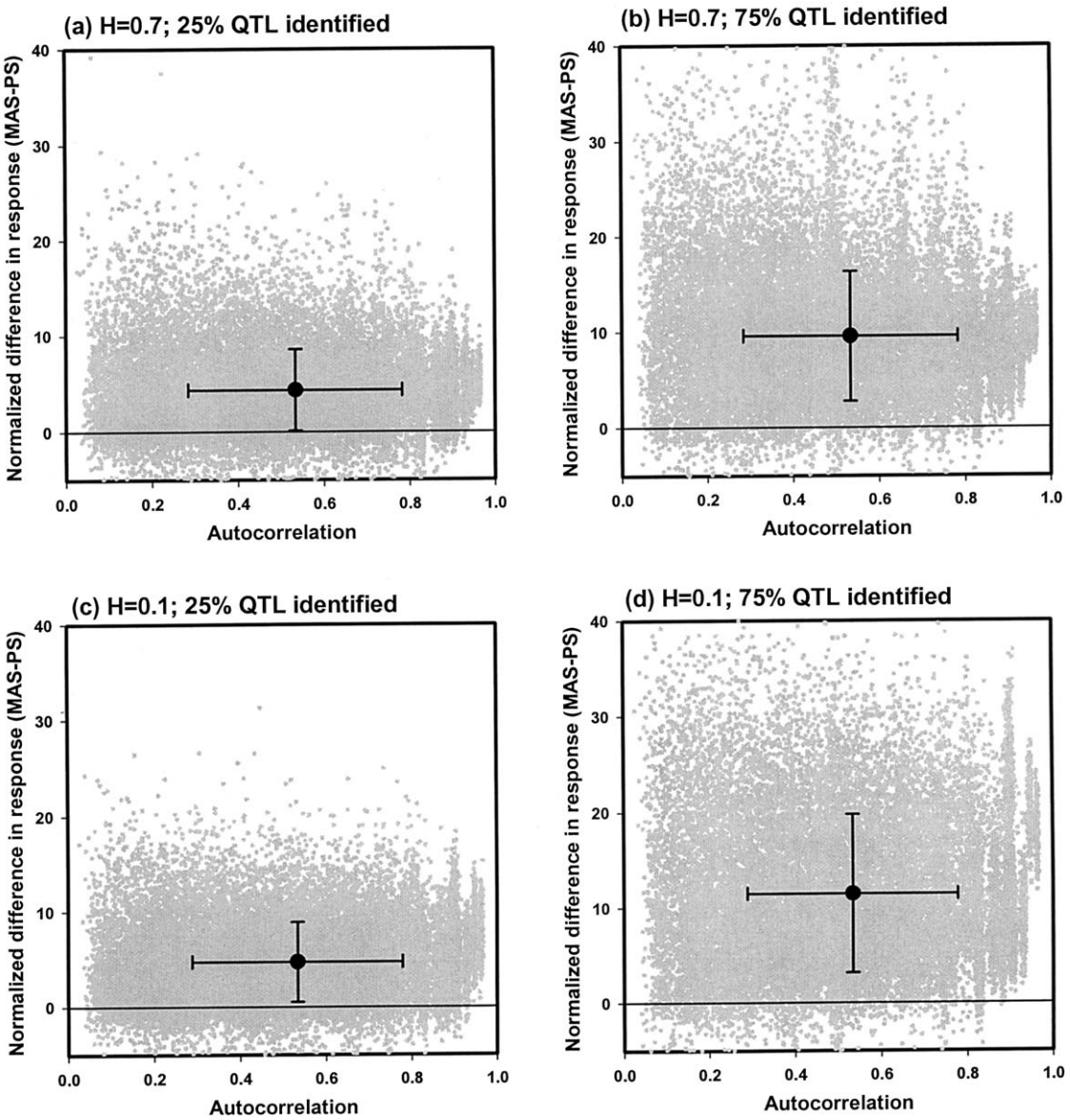

Fig. 2. Simulation comparison of a marker-assisted (MAS) and a phenotypic (PS) selection strategy following the methods developed in Cooper and Podlich (2002). The four diagrams plot the difference between the population mean performances achieved for a quantitative trait at cycle 5 by the MAS and PS recurrent selection strategies for a large number of putative genetic architectures of a quantitative trait [normalized difference in response (MAS-PS)] against the complexity of the trait measured as an autocorrelation on a performance landscape from walks in genetic space (at the extremes the autocorrelation → 1 represents the more simple additive genetic models and the autocorrelation → 0 represents the more complex genetic models of the architecture of the trait). In the four subfigures, H = broad sense heritability in the base population; the percentage of QTL identified represents the percentage of the total QTL used in the MAS strategy. The large symbols represent the grand mean of the normalized difference in response between MAS and PS and the bars represent the standard deviations of the individual estimates of the normalized difference between MAS and PS.

the gene networks that are involved in the relationships between interindividual DNA sequence variation and trait phenotypic variation in elite breeding populations. The apparent complexity of the gene networks underlying the observed gene-to-phenotype relationships for plant development, specific pathways and traits has stimulated interest in the use of a range of advanced mathematical methods, within a systems biology framework, to develop and test gene-to-phenotype models for traits that operate across scales of biological organization. Validation of these models will be a critical step in this process and will be the ultimate assessment of the potential of such modeling approaches to create new gene-to-phenotype knowledge for traits that can be used to improve the traits and design robust molecular breeding strategies.

Foundations for Molecular Breeding: Some Current Considerations

Much of the outcomes from plant genomics efforts to date have involved large-scale data generation of a narrow sample of genotypic variation and its systematic organization, combined with descriptive efforts to anno-

tate the features observed in the organized data. This process has revealed much about some important features of the structural organization of plant genomes within a broader evolutionary framework. Comparative approaches have been used to initiate candidate gene searches from model to target species. Moving from this broad comparative genomics view of gene discovery to a breeding strategy view that is focused on understanding the detailed organization of extant allelic variation for multiple traits within a selected elite germplasm pool, presents significant challenges. Nevertheless, for some traits this approach can be effective. The outcomes from experimental investigations structured around these genomic resources are being examined today and represent a logical step for testing comparative genetics hypotheses and refinement of the current genomic database annotations.

Foundations for Molecular Breeding: Power of Selection

When we consider the power that selection on phenotype has demonstrated in bringing about directed changes for traits, it is important to consider the robustness of this approach across a wide range of genetic situations where the breeder knew little about the detail of the genetic architecture of the traits. An important consideration in the design of any selection process, be it molecular or phenotypic, is that *you get what you select for*. Direct selection on the phenotype of the end-product traits of commercial significance is a relatively robust, albeit in some cases slow, approach when genetic variation exists for the target traits. Replacing or augmenting this system with a knowledge-based approach that targets selection at the level of DNA sequence variation will also rapidly bring about genetic changes. The rigor of the associations we develop between population level sequence variation and phenotypic variation will determine the robustness of this molecular breeding approach. We will be continually forced to refine our knowledge of trait genetics and gene-to-phenotype associations.

What Does a Genomics View Give Us Access to That We Did Not Previously Have?

Without appropriate investment into crop genomics research, we will always lack detailed knowledge in two areas critical for successful breeding: (i) the structural organization and functional properties of genetic variation for traits and (ii) the influences that plant breeding strategies have on the genetic variation that is widely used in agriculture. Without detailed knowledge in both of these domains, it is difficult to answer many of the important questions asked of breeding programs: (i) how sustainable is a breeding effort in the long term, (ii) how robust are the products of a given breeding program, and (iii) how important is it to and what are the appropriate procedures to maintain and utilize genetic resources? Consider two opposing views on the role of genomics in plant breeding.

The optimistic view is that much of the detail necessary for the design of molecular breeding strategies will flow from a comprehensive genomics view of our genetic resources. Broad information on the extent of colinearity and rearrangement of genomes among species is a first step. Revealing the detail of interindividual sequence colinearity within a species and within elite breeding germplasm pools is a necessary next step. Some surprises and testable hypotheses will emerge (e.g., Fu and Dooner, 2002). With a detailed knowledge of genome organization and gene function and appropriately designed experiments, it will become increasingly feasible to resolve some long-standing debates and competing genetic models of quantitative trait variation; e.g., the genetic basis of epistasis, pleiotropy, genotype $\times$ environment interactions, inbreeding depression, and heterosis. Further, a genomics view will enable the study of chromosomal recombination at the physical level. Importantly, detailed knowledge of sequence variation among individuals within a pedigree breeding structure will provide a powerful resource for understanding the distribution and genetic control of recombination and the historical importance of specific recombination events in the breeding process. There has been much debate in the scientific literature on the importance of minimizing and maximizing recombination for specific purposes. Ultimately, we would seek to be able to understand the genetic control of recombination, identify specific regions of the genome where it is important to restrict recombination and those regions where we need to create new recombinants. Collectively, knowledge generation of gene-to-phenotype relationships for traits would be greatly accelerated by access to a comprehensive view of genome organization and interindividual genomic variation. Testing the hypotheses of the organization of genes within gene networks requires experimental procedures for studying the structural and functional properties of genes, the functional nature of allelic variation and measurement of their coordinated regulation within a plant growth and development framework. Access to this genomic resource will enable the study of some of the key properties of gene networks and their involvement and roles in determination of phenotypes.

In contrast, the pessimistic view is that the biology for many of the traits targeted by a breeding program is so complex and interconnected that the context dependent knowledge generation that is required to achieve improvements in predictability of the system will be so great that it will be difficult or impossible to achieve sufficient knowledge to design a molecular breeding strategy that will consistently improve on large-scale targeted phenotypic selection (e.g., Fig. 2). This view is not void of theoretical consideration and is founded in considerations derived from complexity theory (Cooper and Podlich, 2002). However, to date there is no comprehensive experimental evidence to test such arguments. Nevertheless, experience suggests that it has been difficult to predict and that we understand little about the functional basis of many of the genetic improvements that have been achieved for quantitative traits by plant breeding programs. Within the human genetics scientific community, similar debates can be found on what are appropriate strategies for the study of and healthcare solu-

tions for complex diseases (Sing et al., 2003; Botstein and Risch, 2003).

Genomics gives us some new perspectives on genetic variation. With access to data and information at the sequence level, our views of what contributes to the natural genetic variation that resides within the germplasm pools developed by breeding programs are changing. The traditional view of genetic variation as a function of loci with fixed effects acting in a predominantly additive manner is challenged by many of the properties of genes that are observed using genomics technologies. An important component of experimental evidence indicates that gene regulation is an important source of genetic variation. What appears to be linear when examined at the phenotypic level is not necessarily linear at the level of the gene network (Peccoud et al., 2004). This creates a complex situation where many of the effects of genes can be highly context dependent. Therefore, the genetic background and environments within which the genes are studied will influence the estimates of the effects of the genes. Plant breeders have always been exposed to this phenomenon but have never had the tools to investigate its genetic basis. The predominant models used to derive the statistical estimates of genetic effects do not yet take into account the nonlinear features of many of the context-dependent properties of genes. Thus, any changes in the effects of the genes as the genetic or environmental contexts change are not accommodated in our genotype–phenotype statistical association models. The important implication of this observation is that selection at the level of the phenotype can operate and utilize all types of genotype–phenotype associations, extending from simple to highly complex genetics, and progress from selection can still be observed. However, strategies based on manipulation of the genotype at the molecular level will only be able to utilize the currently available experimental information from statistically determined genotype-phenotype associations. Reminiscent of much of the debate around the effects of epistasis on response to selection, this forces the plant breeding community to ask specific questions of the implications of molecular breeding strategies for both short-term and long-term genetic improvement of complex traits.

The reality of many of the plant breeding situations we encounter in practice is likely a mixture of and somewhere between the extremes of the knowledge-driven optimistic view and the unpredictable complexity view. A major challenge for experimental genomics is to design experiments that help to resolve components of these issues and demonstrate and define uses of genomics to enhance breeding outcomes. Here, we identify two fundamentally different, but equally critical issues, that need to be considered in experimentally demonstrating the areas where genomics will affect multitrait improvement: (i) enhancing the rate of progress of a population of individuals toward a target genotype that has already been identified and defined, e.g., by accelerating a backcross process for simple traits or focusing a pedigree effort for more complex traits and (ii) the process of predicting, defining, and creating the new gene combinations that will provide performance enhancements that have not yet been discovered. Both are important features of achieving genetic progress in both the short-term and long-term. The former is easier than the latter.

Designing Molecular Breeding Strategies: Three Themes

Today we can consider coordinated development of three themes and associated research paths arising from research over the last 10 to 15 yr.

1. There is a growing knowledge base of the genetic architecture for some traits and how genetic variation is organized within unimproved and elite germplasm pools. Further work in this area will require the integration of genomics technologies with the study of genetic variation to conduct focused gene-to-phenotype studies. This will require fundamental questioning and in some cases refinement of our models of genetic variation. Development of our bioinformatics and computational modeling tools will be necessary.
2. High throughput genetic profiling of individuals for key regions of the genome is now feasible for elite and unimproved germplasm pools. High performance management, manipulation, analysis and interpretation of molecular and phenotypic data will continue to be areas of research priority.
3. Determining the power of molecular and conventional breeding strategies to achieve directed phenotypic changes for simple to complex traits.

This last area is the least developed of the three themes we have identified here. To date, we have a lot of practical experience with conventional breeding strategies and are now gaining some practical experience with molecular enhanced breeding strategies. High performance computing and simulation is being used to complement theoretical and experimental investigations. There is a clear need for further research into the appropriate statistical and biological modeling procedures for determining and testing gene-to-phenotype associations for complex traits. Demands for advances in this area will grow as we populate and explore data rich gene-genotype-phenotype knowledge bases.

Conclusions

Current structural and functional genomics methodologies provide the foundation for studying the genetic architecture and variation for traits. Quantitative integration of interindividual molecular and phenotypic variation is a challenging step that is an area of intense research in the study of gene-to-phenotype relationships. Applying genomic methods in parallel across many genotypes is considered an important step in enabling the study of genetic variation in elite germplasm and the design of molecular enhanced breeding strategies. The design of commercially viable molecular plant breeding strategies is an experiment in progress. Genomics has and will continue to make contributions to the knowledge base of our target crops. We will continue

to improve our ability to identify the factors that have contributed to past successes in breeding and use this to identify potential new paths to improvement. As with the history of the development of conventional breeding strategies during the 20th century, it is expected that the design, evaluation, and commercial use of molecular breeding strategies will unfold in the 21st century from a rich mixture of both independent and collaborative contributions from the public and private sector research communities.

REFERENCES

Botstein, D., and N. Risch. 2003. Discovering genotypes underlying human phenotypes: Past successes for mendelian disease, future approaches for complex disease. Nature Genet. Supp. 33:228–237.

Cooper, M., and D.W. Podlich. 2002. The *E(NK)* model: Extending the *NK* model to incorporate gene-by-environment interactions and epistasis for diploid genomes. Complexity 7:31–47.

Cooper, M., D.W. Podlich, K.P. Micallef, O.S. Smith, N.M. Jensen, S.C. Chapman et al. 2002. Complexity, quantitative traits and plant breeding: A role for simulation modeling in the genetic improvement of crops. p. 143–166. *In* M.S. Kang (ed.) Quantitative genetics, genomics and plant breeding, CABI Publishing, Wallingford, UK.

Duvick, D.N., J.S.C. Smith, and M. Cooper. 2004a. Long-term selection in a commercial hybrid maize breeding program. Plant Breed. Rev. 24:109–151.

Duvick, D.N., J.S.C. Smith, and M. Cooper. 2004b. Changes in performance, parentage, and genetic diversity of successful corn hybrids, 1930–2000. p. 65–97. *In* C.W. Smith et al. (ed.) Corn: Origin, history, technology, and production. John Wiley & Sons Inc., NJ.

Fu, H., and H.K. Dooner. 2002. Intraspecific violation of genetic colinearity and its implications in maize. Proc. Natl. Acad. Sci. USA 99:9573–9578.

Goff, S.A. et al. 2002. A draft sequence of the rice genome (*Oryza sativa* L. spp. *japonica*). Science 296:92–100.

Peccoud, J., K. Vander Velden, D.W. Podlich, C.R. Winkler, W.L. Arthur, and M. Cooper. 2004. The selective values of alleles in a molecular network model are context dependent. Genetics 166: 1715–1725.

Sing, C.F., J.H. Stengård, and S.L.R. Kardia. 2003. Genes, environment and cardiovascular disease. Arterioscler Thromb. Vasc. Biol. 23:1190–1196.

The Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana.* Nature 408: 796–815.

Yu, J. et al. 2002. A draft sequence of the rice genome (*Oryza sativa* L. spp. *indica*). Science 296:79–92.

Plant Breeding Requirements for Applied Molecular Biology

Major M. Goodman*

Plant breeding is unlikely to be radically altered by genetic engineering despite progress in genomics. New traits will ultimately be added to today's breeding goals, but most are likely to require several decades of development. Many have decided that the future of plant breeding lies in genomics, relying on claims that molecular genetics has revolutionized the time frame for product development. "Seldom has it been pointed out that it is going to take as long to breed a molecular engineering gene into a successful cultivar as it takes for a natural gene" (Bingham, 1983, p. 223). Additionally, claims often suggest simple solutions to very complex problems ["Agricultural biotechnology is already having an impact" (on starvation!); Theil, 2001]. Such claims are often made with little knowledge of the problems of selecting and testing germplasm, genotype × environment interactions, or even epistasis. These claims are often accepted by management that employs breeders who certainly know that such "quick solutions" will not reach farmers' fields for well over a decade. "The public must be cautioned that the simplest advances take, on average, 10 yr from inception of breeding effort to placement on the farm in quantity" (Duvick, 1982, p. 583). It is simply untrue that a new transgenic cultivar can be routinely created, tested, and deployed within a decade (Goodman and Carson, 2000).

Transgene Utilization

Insecticidal *Bacillus thuringiensis* (*Bt*) was used by the 1950s. The first gene encoding the *Bt* toxin was cloned by Schnepf and Whiteley (1981). *Bt* gene regulation was known by 1986 (Whiteley et al., 1987). *Bt* was transformed into maize (*Zea mays* L.) in 1990 (Koziel et al., 1993). *Bt* hybrids were first sold in 1997. Because *Bt* was a well-known entity with a long history of use as an "organic" insecticide, it was relatively straightforward for regulatory agencies to assess for its initial use as a transgene, compared with less well-known genes that genomics research may make available.

Even so, its development into a commercial product took 16 yr.

Integration of Breeding with Plant Molecular Biology

Breeding progress continues to increase yield at a rate of 1 to 2% per year, with additional gains made for disease resistance, maturity, standability, and production efficiency. Virtually all gains are due to utilization of polygenic factors not readily handled by currently available molecular procedures. Molecular genetics will not add much to routine breeding practices until this is overcome. Is marker-assisted selection (MAS) an alternative? Studies by Beavis (1994), Openshaw and Frascaroli (1997), and Bernardo (2001) strongly suggest that MAS is only effective under specific circumstances. For those interested in the discouraging details, see Goodman and Carson (2000) and Melchinger (2003). Currently in place are several simply-inherited qualitative traits such as *Bt*-insect control, herbicide resistances, virus resistances, and new sources of the equivalent of cytoplasmic male sterility. What is needed by plant breeders? Traits that plant breeders can only manipulate with difficulty or traits currently unavailable. These

Dep. of Crop Science, North Carolina State Univ., Raleigh, NC 27695. Received 23 July 2003. *Corresponding author (maize_resources@ncsu.edu).

include traits like drought tolerance, fungal-toxin (aflatoxins, fumonisins) resistance, salt tolerance, heat tolerance, and general environmental stability.

It is unlikely that the minimal 15-yr lag time between gene discovery and seed sales to farmers can be reduced, but politics could effectively increase it, especially in Europe. Thus, new developments in molecular genetics must promise a 20 to 30% improvement in yield or offer a useful, novel trait without reducing yield or they are unlikely to survive the 15+ yr development curve. In addition, realism needs to accompany proposed modifications in seed characteristics. There are clear benefits to be gained from eliminating unhealthy or quality-degrading oils from soybean [*Glycine max* (L.) Merr.] or palm (*Elaeis guineensis* Jacq.). It is not clear that increasing oil or protein content in maize will be beneficial. No crop is an island unto itself. Food or rations containing oil or protein from legumes, starches from grains, and vitamins from vegetables probably make more economic sense than maize with 10% oil, a completely balanced amino acid ratio, and additional vitamins.

The several steps required to move a sequenced gene into a commercial product were outlined by Goodman and Carson (2000) and Gepts (2002). Initial estimated costs of this were as low as $5 million; current estimates are in excess of $60 million. These compare with the generally accepted, approximate cost of $1 million to develop a useful, conventionally bred inbred line. Thus, commercial development of a single gene is now roughly 50 times as costly as the development of a commercial inbred by conventional breeding. This is a formidable barrier, as *Bt* seems to have been sold at just about the break-even point for the farmer, at about 30% of seed cost (Duffy, 2001). It is unlikely that any combination of transgenes now on the horizon could greatly increase this premium while farmers are selling maize at the low price of $2.50 per bushel.

Any molecularly engineered trait of clear economic use will be rapidly utilized by plant breeders. What is lacking at present is an array of useful transgenic traits. The easy and obvious ones have been implemented. At the moment, the pipeline of molecularly engineered traits appears to be largely empty. [*Bt* for maize rootworms (*Diabrotica* spp.) has recently become available, but it has few companions.] Indeed, the question can be asked, does the pipeline exist or do we just have random bits of pipe strewn about, with rather little organization?

There is little doubt that plants (and animals) will be used to produce certain chemicals and pharmaceticals, but this is apt to be on a horticultural scale, rather than a broad-based agricultural effort. There is considerable need for fungal and bacterial protection of crop plants, but progress has been slow. Worldwide, the greatest problem that needs to be solved for most food- and feed-crops is postharvest protection against insects and vermin. That would solve far more problems than adding carotene to rice or lysine to maize.

REFERENCES

Beavis, W.D. 1994. The power and deceit of QTL experiments: Lessons from comparative QTL studies. Annu. Corn Sorghum Research Conf. Proc. 49:250–266.

Bernardo, R. 2001. What if we knew all the genes for a quantitative trait in hybrid crops? Crop Sci. 41:1–4.

Bingham, T. 1983. Molecular engineering vs plant breeding. Plant Mol. Biol. 2:222–224.

Duffy, M. 2001. Who benefits from biotechnology? Annual Corn Sorghum Research Conference Proc. 56, http://www.leopold.iastate.edu/pubs/speech/files/120501-who_benefits_from_biotechnology.pdf; verified 6 July 2004.

Duvick, D.N. 1982. Improved conventional strategies and methods for selection and utilization of germplasm. p. 577–584. *In* L.W. Schemilt (ed.) Chemistry and world food supplies: The new frontiers. Pergamon Press, Oxford, UK.

Gepts, P. 2002. A comparison between crop domestication, classical plant breeding, and genetic engineering. Crop Sci. 42:1780–1790.

Goodman, M.M., and M.L. Carson. 2000. Reality vs. myth: Corn breeding, exotics, and genetic engineering. Annu. Corn Sorghum Research Conf. Proc. 55:149–172.

Koziel, T.M., G.L. Beland, C. Bowman, N.B. Carozzi, R. Crenshaw, L. Crossland et al. 1993. Field performance of elite transgenic maize plants expressing an insecticidal protein derived from *Bacillus thuringiensis*. Bio/Technology 11:194–200.

Melchinger, A.E. 2003. Lessons from large QTL mapping experiments in maize. Annual Corn and Sorghum Research Conf. Proc. [CD-Rom]. 58:512–532.

Openshaw, S., and E. Frascaroli. 1997. QTL detection and marker-assisted selection for complex traits in maize. Annu. Corn Sorghum Research Conf. Proc. 52:44–53.

Schnepf, H.E., and H.R. Whiteley. 1981. Cloning and expression of the *Bacillus thuringiensis* crystal protein gene in *Escherichia coli.* Proc. Natl. Acad. Sci. USA 78:2893–2897.

Theil, K.A. 2001. Kernels of truth. Forbes ASAP. Feb. 19. p. 113–115.

Whiteley, H.R., H.E. Schnepf, K. Tomczak, and J.C. Lara. 1987. Structure and regulation of the crystal protein gene of *Bacillus thuringiensis*. p. 13–27. *In* K. Maramorosch (ed.) Biotechnology in invertebrate pathology and cell culture. Academic Press, San Diego.

SYMPOSIUM DISCUSSION

Panelists and the audience of this symposium were asked to address whether and how genomics will be useful in plant breeding. The most frequent themes of the discussion are summarized in this report: (i) specific genomic applications or other molecular genetic applications likely to be useful—or not—for plant breeding, (ii) investment in genomics or in plant breeding, and (iii) training and funding for genomics-assisted plant breeding.

Specific Genomics Applications and Plant Breeding

Understanding and Managing Linkage Blocks

1. Genomics may allow breeders to better manage favorable or unfavorable gene linkages. Over time, selection builds up linkage disequilibrium and creates both favorable and unfavorable linkage blocks. Previously undetected regions of the genome which seem to have no effect may be linkage blocks of mixed favorable and unfavorable alleles. Geno-

mics information will allow breeders to locate these linkage blocks and possibly provide methods for enhancing recombination in those regions. Once there is recombination, breeders can select against individual unfavorable alleles and begin to accumulate new combinations of favorable alleles.

2. Knowledge from molecular markers has already made this possible with some simply inherited traits. For example in *Capsicum*, tight linkage between an unfavorable allele conferring small fruit weight and a useful allele of disease resistance was broken. To find resistant recombinants without the small fruit gene using conventional plant breeding would have required a larger population size than any program could manage. Using molecular information, an unfavorable linkage that had not been resolved for decades was accomplished in a relatively small population. Another example is the hardness locus in wheat where several duplicated genes of the same family are closely linked. Molecular information allowed the breeders to locate the alleles and identify their individual effects. It is now possible to develop spring wheat germplasm with completely new textural qualities.

In Quest of the Perfect Marker

1. Accuracy of selection for desired traits is a central problem in plant breeding. The perfect marker would allow breeders to maximize recombination between desired and undesired genes without fear of losing the linkage between the desired trait and marker, by tracking the alleles of the gene itself. Most examples of such perfect markers available today may represent simply inherited traits, such as pungency in peppers. For some major genes, results from a one-time screening to identify the gene can be used in breeding in many different situations. Such examples are limited, and the widespread use of markers will need greater genomic, i.e., haplotypic information. Some such markers now exist, and are available to the public. For example, EST databases comprised of more than one genotype routinely provide sequence information for SSRs and SNPs. Once these markers are associated with a phenotype, nucleotide sequence divergence (i.e., polymorphism) among different alleles provides a way to track specific alleles within pedigrees and populations. In addition, these polymorphisms may yield insights into gene function. Some SSRs have been experimentally associated with a favorable phenotype
2. Public availability of these markers means that speed of transfer of the information to breeding programs will be the only competitive difference among countries.

Identification and Cloning of Candidate Genes: Comparative vs. Direct Genomics

1. Growing databases of finished genomic sequence from model species, and partial sequence for important crops, will be used in creative new ways that add information to each. For example, extrapolation of linkage order to more distantly related species may facilitate map-based cloning of genes in new crops, or in "orphan" crops, where there is little or no previous research.
2. Candidate genes can be inferred from a closely related model system, through the use of comparative linkage analyses of genes known to affect the trait in the model system. Comparative genomic analysis can identify candidate genes by predicting their location in a crop's genome. Candidate genes for economically important traits in crop plants are potentially useful in plant breeding. Candidate genes can be useful even when a gene's function is unknown, or lacking detailed information about species-specific compounds.
3. However, genetic variation in economically important species may not be conditioned by the same loci as in the model species. Moreover, many economically important crop traits are unique to the crop and not present in model species. For these traits, candidate gene prediction from model species will not be possible. Direct analysis of crop genomes will be essential to elucidate unique features such as the fibers of cotton and the underground pegs of peanuts.
4. In addition to research on expressed sequences, genomics research will provide information on regulatory elements, introns, matrix attachment regions, and other genomic features that affect gene expression, thus providing information which may be useful in transgenic research as a source for tissue-specific promoters and in plant breeding as polymorphic molecular markers.

Understanding and Using Complex Genetic Variation

1. The known horizons of heritable variation have broadened during the last several years on the basis of genomics information from DNA sequence variation. Each individual gene is part of a complex metabolic network comprised of structural genes, transcription factors, and repressors. Regulatory genes are more difficult than structural genes to identify, even with phenotype and sequence data.
2. The value of such new variation will vary from one genotype to another, e.g., in hybrids of widely differing yield potential. Like any trait currently manipulated by breeders, traits from genomics research will require re-evaluation in different genetic backgrounds and environments. Genomics-assisted plant breeding, particularly for more complex traits such as yield and adaptation, will require appropriate quantitative analyses that integrates information from genomic sequence to crop phenotype. This will require bioinformatics to integrate all relevant information from genomics to phenotype, such as is being done with The Arabidopsis Infor-

mation Resource, but also across species, to leverage information from model species to crop species.
3. Integration of genomic and phenotypic information requires, in addition, a strong framework of knowledge at the whole plant and system level. For example, it may be revealed that too little is known about whole plant physiology and plant–soil relationships. Also needed are better estimates of phenotypes for traits that underlie yield. The individual contribution of yield components to the overall phenotype—the goal of every breeding program—has not been adequately modeled in a genomics context.

Invest in Genomics? Or Invest in Plant Breeding?

1. Would plant breeding be further advanced today if the same amount of money had been spent on population improvement as on hybrid development in maize? Will the future be asking, "If the same amount of money had been spent on traditional plant breeding instead of genomics, would we be ahead of where we are now?" Particularly in developing countries, where needs are urgent, relatively less systematic breeding has been done and small investments in breeding can provide rapid gains in even a few years. It is a special concern, therefore, if international research centers, whose mandate is to assist the poorest countries, invest in genomics at the expense of plant breeding.
2. In 40 years, entire genome sequences may exist for most economically important crops. However, sequences per se do not improve yield per acre and resistance to major crop pests. To date, genetic improvements in domesticated crops have used empirically obtained phenotypic data, pedigree information, and selection. Prediction methods based on data routinely collected by plant breeding programs have enhanced the power of selection, and genetic gains from application of statistical models based on phenotypic and pedigree data are far from exhausted. In comparison to plant genomic projects, statistical approaches may provide quick and cost-effective advances in plant breeding methodology.
3. The products based on molecular techniques now available, such as transgenic varieties having a *Bt* gene, are simple improvements of high performance genotypes. The *Bt* lines were developed in 2 to 3 yr of backcrossing laboratory lines to elite parental lines. Each of these elite lines had been developed though decades of traditional plant breeding. With additional basic research, plant genomics could allow breeders to move beyond such simple improvements, to the manipulation of fundamental gene networks within high performance genotypes. So applied, plant genomic research could in theory extend the biological limits of traditional plant breeding. This will be a long-term effort—one participant estimated at least 40 yr. It may become increasingly important if transgenic approaches become unavailable because of lack of public acceptance.
4. The hybrid maize trade-off may not be an accurate representation of history. Population improvement schemes for maize were developed about 20 yr after hybrid breeding, and in fact, much of the effort to develop today's excellent theory and practice of population improvement was inspired by the early work and difficulties in breeding hybrid maize. Applying genomics to plant breeding may be similarly heuristic. Breeding hybrid corn once looked as weird, strange, and exciting as genomics does now.

Training and Funding for Genomics-Assisted Plant Breeding

Training

1. Connection between laboratory and field is essential for integration of genomics and breeding. Field observation and breeding experience may be the only way to identify the right questions, even with a plethora of genomic information at hand. However, extensive panelist experience reading research proposals presented in recent years to the National Science Foundation and the CSREES USDA National Research Initiative reveal that lab-to-field connections have not generally been in place in these proposals. Knowledge of basic agronomy and biology, population genetics, statistics, and experimental design are critical for analyzing field data; incorporating data provided by molecular markers and genomics; and then relating all to a plant phenotype. Plant breeding students are required to study statistics, plant physiology, molecular biology, and other disciplines; however, it is difficult to find a reciprocal: a molecular genetics student with enough statistical training to study quantitative plant breeding.
2. A costly outcome of reductions to programs at public universities is the loss of an educational environment that provides hands-on breeding experience. Only a handful of public universities exist with the knowledge depth and the operational field programs required to educate future plant breeders. Ironically, this plant breeding experience is in high demand in the current job market. Plant breeding students are often hired before they graduate.

Funding

1. Compared with genomics, plant breeding is inexpensive. Operational funds of $40 000 to $50 000 will sustain a medium-sized program for a year. However, in most universities today, a plant breeding program is more likely to be billed than funded. Many universities cannot fill a research position unless there are opportunities for the researcher to obtain research grants. For plant breeding, there is "nowhere to go" for grant funds. Some plant breeding is "piggybacked" on genomics research grants. Traditional resources for classical plant

breeding are continuously eroded because of depressed domestic commodity prices and declining political support as the number of growers declines. These circumstances will result in reduced numbers of field breeders. Without them as cooperators, workers in genomics will be unable to apply their new data and knowledge to plant breeding.

2. To add value, genomics must help create varieties that will be recognized and popular with consumers. Given that novelty is in demand, if public plant breeders had pursued the Baye-Dohl Act of 1980, would the resulting intellectual property rights on varieties (mostly Plant Variety Protection certificates) have provided an income stream for public plant breeding today?
3. However, because of their diverse needs, large numbers, and the uncertainty of continuing markets for any one specialty crop, how much effort can breeders, private or public, put into breeding for high-value-niche crops with small markets but relatively high margin for farmers? As fashions change, demand for specialty crops may come and go. Creation of novel types of traditional commodities, such as new classes of wheat for new products, may be a more feasible investment.
4. With regard to breeding cultivars for crops and environments that do not generate sufficient return on investment for the private sector (the public goods versus private goods conundrum): Who is in responsible for identifying the crops that are public goods and then providing public monies to support their breeding? Financial support from legislatures and granting agencies reflects priorities established by input from both scientific communities and the public. By competing for the attention of decision-makers, rather than presenting a coordinated message, the scientific community may bear part of the responsibility for not bringing money into plant breeding.

SYMPOSIUM SUMMARY

"To reach an objective, use all the information you have" (Duvick). (Undated attributions are personal communications during the symposium discussion reported on the preceding pages).

Genomics emerged recently as a term to describe investigations of the whole genome using biotechnologies. Given that substantial resources have been devoted to plant genomics and molecular genetics, it is timely to discuss how they can assist plant breeding. This symposium summary serves to emphasize and document ideas, suggestions, and recommendations made by panelists and audience participants.

The panelists were scientists who had received competitive funding for crop genomics research and had articulated a vision and a role for genomics in plant breeding from their particular perspective. Some panelists and audience participants emphasized how genomics could be applied directly to crop improvement, while others emphasized its role in understanding the fundamental biological questions of adaptation and response to selection, which one day may make breeding more efficient. Panelists generally expected new genomic applications will become available as the price of technology continues to drop and as a greater understanding of the plant genome leads to new insights into its manipulation, though views on the usefulness of genomics for crop improvement varied from enthusiasm to skepticism. Moreover, "integration of genomics and plant breeding may become increasingly important if transgenics become unavailable because of lack of public acceptance" (Knapp). Three areas were viewed as most important for application of genomics and molecular genetics: molecular-assisted breeding; gene and genome sequencing and gene networks; and use of genetic diversity.

Marker-Assisted Breeding

The use of molecular markers for marker-assisted selection (MAS) received early attention by the plant breeding community and, consequently, has been the approach most used. Validation data are still being obtained and optimal strategies to capitalize on the use of MAS are still being formulated (Cooper et al., this symposium, 1907–1913). Even so, MAS is increasingly efficient, with a steady evolution in the types of markers used. A new generation of molecular markers based on single nucleotide polymorphisms (SNPs) should permit relatively low cost, high-throughput analyses of entire breeding populations (Dubcovsky, this symposium, 1895–1898).

Successful use of MAS requires markers linked to traits of interest. Associations between markers and simply inherited traits with a strong impact on the phenotype are the easiest associations to make (the "low-hanging fruit"—Walsh), but these are precisely the type of trait with which breeders have always been the most successful. Consequently, use of MAS for simply inherited traits can be justified only when it replaces more expensive or tedious assays, or results in increased precision in the identification of desired genotypes (Cooper et al., this symposium, 1907–1913). Two examples of increased precision included the manipulation of tight linkages within *Capsicum* (Jahn) and wheat (Dubcovsky). In the future, markers linked to simply inherited traits of interest will need to be "resolved to the level of candidate genes" (Cooper et al., this symposium, 1907–1913), making the process more efficient.

Cautious optimism was voiced about MAS of complex traits. Although molecular markers have been successfully associated with quantitative trait loci (QTLs), these associations have had very limited usefulness in plant breeding programs. Complex traits are the most difficult to handle during a breeding program, but are responsible for most breeding progress in critical traits such as yield, yield stability, and adaptation (Nelson et al., this symposium, 1901–1904; Goodman, this symposium, 1913–1914). "We [plant breeders] get paid for the phenotype, yet the individual contribution of yield components to the overall phenotype has not been adequately modeled

in a genomics context" (Cooper). Long-term efforts are required to investigate the nature of complex traits at the molecular level, before selection of complex traits by molecular markers can be fully realized. "Molecular genetics will not add much to routine breeding practices until this (the inability of molecular techniques to manipulate complex traits) is overcome" (Goodman, this symposium, 1913–1914).

Gene and Genome Sequencing and Gene Networks

Most panelists agreed that more genome sequencing of crop species is essential for improved and continued application of plant genomics to plant breeding. In fact, the sequence of crop genomes may have greater eco nomic impact for developing countries than the efforts of the human genome project (Martienssen, this symposium, 1898–1899). Genome sequence for various crops would improve the quality of molecular markers used for MAS by targeting the gene of interest, rather than a nearby sequence. This limitation of linkage is being resolved as sequence data are becoming available, making it possible to use the gene itself as its own marker (Dubcovsky, this symposium, 1895–1898; Nelson et al., this symposium, 1901–1904; Martienssen, this symposium, 1898–1899). These markers are "based directly on ... variation at the gene responsible … [for] the trait. Examples of perfect markers (in wheat) include genes for gluten strength, genes for starch properties, genes for hardness, vernalization genes, and the Lr genes for leaf rust resistance" (Dubcovsky, this symposium, 1895–1898).

Polymerase chain reaction (PCR)-based markers are desirable as they can be automated, but every type of PCR-based molecular marker requires up-front DNA sequence information. In fact, SNPs require sequence knowledge of multiple alleles, i.e., sequencing genes a single time is not enough. Continuing efforts to sequence expressed genes will provide data for SNP markers for individual alleles, making MAS more efficient (Dubcovsky). In addition, "other genomic research will provide information regarding nucleotide sequences which may prove more valuable than many of the transcribed (i.e., expressed) genes. Approaches for cloning gene-rich regions of the genome will provide sequence information of for both genes and their controlling regions" (Paterson et al., this symposium, 1900–1901).

"No gene acts alone" (Walsh), yet interactions between genes—so called "gene-networks"—are little-understood at this time. Accurate gene-to-phenotype models will depend on better understanding of these intergenic interactions. Global gene expression studies will be very valuable to help address problems intractable until now. This will require, in addition, a greater understanding of whole plant physiology and yield, and the functional interactions and properties of genes (Cooper et al., this symposium, 1907–1913). Until better gene prediction models are in place, breeding programs will not be able to rely exclusively on MAS, but must supplement genomics-based efforts with analyses of phenotypic measurements from replicated field trials.

Use of Genetic Diversity

Many wild plant species are related to modern crops and contain useful traits which are not found in adapted varieties, including additional disease resistances, stress tolerance, and even genes for increased production (Brummer, this symposium, 1904–1907; Paterson et al., this symposium, 1900–1901). Use of broad-based genetic diversity in breeding will benefit from MAS to follow specific genes from unadapted relatives and/or different plant species as they are bred into elite varieties (Paterson et al., this symposium, 1900–1901; Dubcovsky, this symposium, 1895–1898; Brummer, this symposium, 1904–1907). It will be important to assay as much diversity as possible, which in turn requires that the germplasm collections be well maintained and curated (Jahn, Goodman). "Many of the answers to our questions about crop productivity, and more generally plant biology, lie in [these] germplasm resources" (Jahn).

Can Genomics be Useful to Plant Breeding?

The most frequent answer was a cautious "yes" particularly when MAS was considered. To what extent, has genomics been useful to plant breeding programs to date? The best example may be the wheat MAS project, which not only developed molecular markers but also facilitated their transfer to breeders who use them: "MAS programs are good examples of implementation projects that have the potential to facilitate the transfer of valuable genes identified in basic research programs into public varieties" (Dubcovsky, this symposium, 1895–1898).

Nearly all panelists agreed that the largest potential benefits of plant genomics are still years away. For example, "It is going to take as long to breed an engineered gene into a successful cultivar as it takes for a natural gene. If better crop performance such as yield is the ultimate goal, in forty years from now, traditional plant breeding methods would have been the best investment for today's dollar" (Goodman; Goodman, this symposium, 1913–1914). Although not all panelists predicted a forty-year time lag, all perceived that use of today's genomics information in plant breeding is limited.

What are the factors that limit current genomics applications in plant breeding? Some insights from the symposium included the following.

It will take time to develop a way to use genomics that is a practical improvement over very successful current methods. (Cooper)

Few researchers or students are conversant at an adequate professional level in both plant genomics and plant breeding. (Brummer, Duvick)

Plant genomic research has centered on model species; many critical crop traits are simply not represented in these models (Paterson; Havey, this symposium, 1893–1895; Nelson et al., this symposium, 1901–1904). "Direct analysis of crop genomes will continue to be essential to elucidate the unique features that make them important as crops (such as cotton fibers and underground peanut pegs, that is, traits not found in model species)" (Paterson).

Plant breeding infrastructure and human resources have been lost, limiting capacity to take advantage of genomics (Duvick, Goodman, Pratt).

Plant breeding capacity to use genomics is particularly lacking for "orphan" crops—that is, crops grown on small acreages or grown primarily in developing countries (Havey, this symposium, 1893–1895; Nelson et al., this symposium, 1901–1904). To the extent that the "model crops" model is valid, there is hope that genomics tools developed for any given species can be directly applied to related orphan crops.

While there were diverse opinions among the participants on whether and how to use genomics and molecular genetic approaches in plant breeding, there was a strong consensus that the plant breeding infrastructure and resources to permit genomics research to be applied to crop improvement are inadequate and declining. "The paradox of the genomics age is that funding for plant breeding programs is decreasing at the same time that the potential of genomics is being realized ... with the result that many technological advances in genomics may not be applied to cultivar development at all" (Brummer, this symposium, 1904–1907). Currently, "one costly outcome of reductions to (plant breeding) programs at public universities is the loss of an educational environment which provides hands-on breeding experience" (Jahn). Without skilled plant breeding programs constantly at work—using genomics and "all the information we have" (Duvick)—to meet ever-changing local and global challenges, plant genomics research will miss its pay-day.

ACKNOWLEDGMENTS

The editors thank the panelists and audience members for their participation; and Thomas Stalker, Crop Science Dep., North Carolina State Univ., and Edward Kaleikau, CSREES, USDA, for indispensable assistance in planning and preparations. In addition to panelists, audience participants in the discussion of the symposium topic were D. Duvick, Iowa State Univ.; S. Kaeppler, Univ. of Wisconsin; R. Pratt, The Ohio State Univ.; D. Stuthman, Univ. of Minnesota; B. Walsh, Univ. of Arizona; and two anonymous participants. The editors are responsible for any misrepresentations in the discussion report. The symposium was supported by a CSREES, USDA Innovation Grant.