

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.

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

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 \times environment adaptation conferring adaptation of cotton to arid conditions. *Genome Res.* 11:1988-1995.
- National Agricultural Statistics Service. 1998. Annual 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.

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 (Pinstrip-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 Tanksley, 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; Kearsley 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 (Frery 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 introgression of favorable alleles at quantitative 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 tobacco plants expressing tobacco 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. Tanksley. 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*

ALFAFA, *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).