DEVELOPMENT, APPLICATION AND DISTRIBUTION OF DNA MARKERS AND GENETIC INFORMATION FOR SORGHUM AND MAIZE IMPROVEMENT

M. LEE
Department of Agronomy,
Iowa State University,
Ames, Iowa, United States of America

Abstract

This final report summarizes the progress made towards the enhancement and distribution of genetic resources (e.g. genetic stocks, seed and DNA clones) used for basic and applied aspects of the genetic improvement of maize and sorghum. The genetic maps of maize and sorghum were improved through comparative mapping of RFLP loci detected by 124 maize cDNA clones and through the development of a new mapping population of maize. Comparative mapping between maize and sorghum and maize and rice, using the set of 124 maize cDNA clones (and other clones) in each study, substantiated previous observations of extensive conservation of locus order but it also provided strong evidence of numerous large-scale chromosomal rearrangements. The new mapping population for maize (intermated B73 × Mo17, 'IBM') was created by random intermating during the first segregating generation. Intermating for four generations prior to the derivation of recombinant inbred lines (RILs) increased the frequency of recombinants at many regions of the maize genome and provided better genetic resolution of locus order. Expansion of the maize genetic map was not uniform along the length of a linkage group and was less than the theoretical expectation. The 350 IBM RILs were genotyped at 512 loci detected by DNA clones, including 76 of the 124 supported by this contract. The production of the sorghum mapping population of RILs from the cross CK60 × PI229828 has been delayed by weather conditions that were not conducive to plant growth and seed development. Seed of the IBM RILs have been distributed (approximately 5000 RILs in total) to 16 research organizations in the public and private sector. The DNA clones have been distributed (1,206 in total) to nine research labs. Further distribution of the seed and clones will be managed by curators at stock centers in the public domain.

1. INTRODUCTION

The options for understanding and enhancing the genetic basis of crop improvement have increased in recent years. One source of new knowledge and technology is the ability to conduct more detailed genetic studies of important traits in one species and then compile the results in such a way that they may be cross-referenced or compared with other studies of the same species and of unrelated species. Such genetic analyses, when coupled with other approaches at the biochemical and physiological level, will reveal the biological basis of important traits and will suggest rational approaches for their modification.

One of the foundations of this new approach to basic and applied biological research in crop plants is the development and utilization of common and shared material and information. Such material and information may take many forms (e.g. data bases, seed, software) and may be used for a wide range of investigations and applications (e.g. marker-assisted selection, assessments of genetic diversity and relationships). For genetic studies, important enabling components have included standard sets of progeny for genetic mapping (mapping populations), DNA clones or sequences used to detect loci on genetic maps and data bases of phenotypic and genotypic information for several species. Subsequently, these resources have been complemented with data bases of DNA sequences of thousands of known

and unknown genes. Collectively, such resources allow researchers to treat groups of sexually isolated species as single genetic systems that may be explored and exploited for crop improvement [1,2].

The need for and utility of such genetic resources depends on many variables related to the forces of nature and actions of human society. This is especially true for the focal species of this project, maize and sorghum. The decades of research in maize and the significant industrial interests have created a vast amount of information that is often difficult to resolve. Specifically, the mapping populations typically used for genetic mapping in maize are represented by a small sample of progeny (40–60 individuals) derived from a population at maximum linkage disequilibrium [3, 4]. The limited opportunities for recovering a recombinant in a given chromosomal region hinders the ability to resolve the order of loci on a genetic map. Thus one of the objectives of this project was to produce a mapping population of maize suitable for higher resolution genetic mapping and other investigations.

In contrast to maize, the basic information and resources for sorghum research are rather limited. However, experiments in comparative genetic mapping with common sets of DNA clones have revealed that large regions of the maize and sorghum genomes have been conserved regarding their gene order and content [5, 6]. Similar analyses have been extended to include several grass species with rice (*Oryza sativa*) as the reference, model genome for the grasses. Such conservation should permit the sharing and transfer of information between the data-rich and data-poor species and research communities. Thus, another objective of the project was to improve the integration of the sorghum and maize genetic maps to facilitate such comparisons and exchange of information derived from investigations of the rice genome. The overall goal of the project was to develop and distribute genetic stocks (seed), DNA clones and sequences, and information which would enable comparative analyses of crop genomes for the purposes of crop improvement.

2. MATERIALS AND METHODS

2.1. Population development

Populations of segregating progeny were developed for maize, sorghum and rice. The maize populations were created by crossing two inbred lines followed by one or more generations of self-pollination to produce recombinant inbred lines (RILs). The maize mapping populations (inbred 1 × inbred 2) used in this project were T × 303 × CO159, T232 × CM37 and B73 × MO17 (intermated; IBM). The first two populations were developed and distributed by Ben Burr [3]. The IBM population was developed at Iowa State University. The intermated population was derived from the single cross hybrid of inbreds B73 (female in this cross) and Mo17. One F1 plant was self-pollinated to produce the F2 generation. In the F2, plants were used once, as male or female, in a cross with another plant so that 250 pairs of plants were mated. A single kernel was taken from each ear and bulked with the seed of the other ears to form the F2 Syn 1 [7]. The procedure was repeated with the F2 Syn 1 plants and for three additional generations to produce the F2 Syn 4.

A set of 370 recombinant inbred lines (RILs) was derived from the F2 Syn 4 generation of the IBM population. Production of these lines was initiated at ISU and completed at Pioneer Hi-Bred International's winter nursery at Puerto Rico. The RILs were derived through single seed descent without intentional selection. Initially, 420 F2 Syn 4 plants were self-pollinated to create F3 lines. Beginning in the F3 generation, independent lineages were

maintained through each generation of subsequent and continuous self-pollination. To advance to the F4 generation, a single F3 plant per line was self-pollinated. Kernels from that selfed ear were planted to grow the F4-generation plants. This was repeated for three successive generations to produce the F7:8 generation kernels (a.k.a. F8-generation kernels tracing to one F7-generation plant). Each of the 370 F7:8 RILs trace to a different plant in the F2 Syn 4 generation. During the derivation of the RILs, 50 lineages of the original 420 were eliminated at some stage of inbreeding for various reasons (e.g. male or female sterility, susceptibility to a disease, extreme lack of vigour) that prevented them from producing seed at some generation. Data at RFLP loci identified an additional 25 RILs that had been contaminated (e.g.. seed mixtures or nonparental pollen) during their development; thus, those RILs were eliminated. Consequently, the maize IBM populations consists of 345 RILs. Adequate seed supplies have been produced for most of those RILs.

The sorghum and rice populations used in this project have been described. The sorghum population is a group of 78 F2 generation plants of a cross between inbred lines CK60 and PI229828 [5]. The rice population is a group of BC1F1 plants of the cross, *Oryza sativa* [8, 9] (cultivar BS125)/*O.longistamata*/*O.sativa*).

2.2. Collection of segregation data at DNA marker loci

The genetic data and maps presented in this report were collected using sets of DNA clones used to detect restriction fragment length polymorphism (RFLPs). The 124 DNA clones characterized and specified under this research contract are maize cDNA clones derived from mRNA isolated from seedling roots [5]. The protocols for using these clones as probes in Southern hybridizations have been described in detail for sorghum [5], maize [10] and rice [9].

2.3. Construction of genetic maps

Construction of genetic maps was facilitated with the software MAPMAKER 3.0 [11] following the procedures described in [4] and [12].

3. RESULTS AND DISCUSSION

3.1. Maize

The 350 RILs were genotyped at 512 RFLP loci. The segregation data are being verified and edited but a preliminary analysis has been conducted. This includes 75 loci detected by the maize cDNA clones. The IBM genetic map comprises 16 linkage groups and over 3000 cM. Normally, a maize genetic map with this number of loci would consist of 10–11 linkage groups with a total of 2000 cM. Apparently, the intermating has enhanced the frequency of recombinants at many regions of the maize genome. A more detailed analysis of two groups of IBM progeny, one group of 95 individuals before intermating and a second group of 90 individuals after five generations of intermating has been conducted with 156 common RFLP loci. The analysis revealed a similar degree of expansion of the genetic map (compared to the full set of 350 RILs) after intermating. However, the degree of map expansion was not uniform along the length of a linkage group.

3.2. Sorghum

Further integration of the maize and sorghum maps was achieved. Previously, the 124 maize cDNA clones had been used to detect RFLP loci in sorghum but they had not been used for genetic mapping in maize [5]. In this project, the 124 clones were used to detect RFLP loci in the maize genome using the mapping populations Tx303 × CO159 and T232 × CM37. Generally, the data collected with the additional 124 clones strengthened the observations of the initial investigation [5]. Each maize linkage group, all of which have been clearly assigned to maize chromosomes, usually contains groups of loci that represent two linkage groups of sorghum. Within a region of a maize linkage group, locus order appears to be highly conserved relative to a given region of a sorghum linkage. However, an adjacent region of the same maize linkage group will often exhibit a high degree of colinearity with a different sorghum linkage group. These observations suggest that the maize and sorghum genomes may be distinguished by several, large-scale segmental rearrangements. Similar patterns of rearrangements have been detected between the maize and rice genetic maps.

The development of sorghum RILs for the CK60 × PI229828 population has been delayed by adverse weather conditions. Simple sequence repeats (SSRs) are being mapped in the population CK60 × PI229828 (78 F2 plants) in collaboration with the University of Milan (G. Taramino and E. Pe) and the USDA/ARS (S. Kresovich, now at Cornell University). Seven SSR loci have been mapped with the Univ. of Milan and those data have been integrated into the RFLP map for that population. Those and related results have been published [13]. Segregation data at 32 SSR loci have been collected by the USDA and those data will establish additional loci.

The CK60 × PI229828 population is being converted into a set of recombinant inbred lines suitable for widespread distribution. In 1996, self-pollinated seed in the F4 generation was produced on 180 F3 plants tracing to 180 F2 plants. The F4 generation seed has been sent to winter nursery to produce the F5 generation seed. In the summer of 1998 in Ames, we produced F8 generation seed for 110 recombinant inbred lines. This population should be suitable as a common mapping for sorghum because of its relatively high frequency of DNA polymorphism and the ease with which inbred progeny are derived. Prior to the distribution of the seed, the lines must be purified in the field on phenotypic and genotypic bases, identified by PCR-based DNA markers and genotyped with a subset of the DNA markers used to make the genetic map. Unfortunately, adverse weather conditions (cool temperatures and excess precipitation) were unsuitable for sorghum growth and development in our field nursery in Ames, Iowa in 1999. Thus, the final stage of seed production and purification has been delayed. These steps will be repeated in the year 2000 in Ames or in an environment better suited for the production of sorghum seed.

3.3. Comparative genetic mapping among maize, sorghum and rice

Of the 124 maize cDNA clones used to map RFLPs in maize and sorghum, 71 also detected RFLPs in rice. These data were added to a larger data set in collaboration with a group at Cornell University [9]. Collectively, the data sets established 182 new loci that have been mapped in the maize and rice genomes as common reference points (previously, only 146 loci had been comparatively mapped between rice and maize).

Comparative genetic analysis revealed over 20 chromosomal rearrangements in maize relative to rice. The changes included telomeric fusions between, and nested insertion of rice

linkage groups, intrachromosomal inversions and a nonreciprocal translocation. A progenitor maize genome of eight chromosomes was inferred.

3.4. Distribution of information, DNA clones and sequences, and seed

The IBM RIL mapping population is becoming a standard, widely used genetic resource in the maize research community. Seed of the IBM maize RILs have been distributed to the USDA-ARS at the University of Missouri, the USDA-ARS at North Carolina State University, the USDA-ARS at Iowa State University, University of Wisconsin, Texas Tech University, University of Georgia, Cornell University, the Maize Genetic Stock Center at the University of Illinois (USDA-ARS), University of Paris, Limagrain, DeKalb Seed Company, Monsanto, Keygene, Pioneer Hi-Bred International, DuPont and Novartis. There have been no requests for the seed of the sorghum mapping population, CK60 × PI229828.

The complete set of segregation data (RFLP loci) are being prepared for deposit into the public domain, available through the internet, at MaizeDB.

Nine sets of 134 cDNA clones have been distributed. The sets were sent to two labs in Brazil, two labs at Texas A&M University, a research institute in California, the National Center for Genome Resources (Santa Fe, New Mexico), Cornell University, Pioneer Hi-Bred and the USDA-ARS Maize Genome Center at the University of Missouri. Further distribution of the clones will be managed by the USDA-ARS at the University of Missouri.

ACKNOWLEDGEMENTS

I thank N. Sharopova, W. Woodman, B. Wilson and S. McCouch for sharing their analyses of segregation data prior to publication. This project was partially supported by IAEA contract 8494/R1 and USDA-NRI Grant 97-35300–4939.

REFERENCES

[1] GALE, M.D. DEVOS, K.M. Plant comparative genetics after 10 years. Science 282 (1998) 656–659.
[2] BENNETZEN, J.L. FREELING, M. The unified grass genome: synergy in synteny. Genome Research 7 (1997) 301–306.
[3] BURR, B. *et al.* Gene mapping with recombinant inbreds in maize. Genetics 118 (1998) 519–526.
[4] DAVIS, G.L. *et al.* A maize map standard with sequenced core markers, grass genome reference points and 932 expressed sequence tagged sites (ESTs) in a 1736-locus map. Genetics 152 (1999) 1137–1172.
[5] PEREIRA, M. G. *et al.* Construction of an RFLP map in sorghum and comparative mapping in maize. Genome 37 (1994) 236–243.
[6] BENNETZEN, J.L. *et al.* Grass Genomes. Proc. Natl. Acad. Sci. USA 95 (1998) 1975–1978.
[7] COVARRUBIAS-PRIETO, J. HALLAUER, A.R., LAMKEY, K.R. Intermating F2 populations of maize. Genetika 21 (1999) 111–126.
[8] CAUSSE, M. *et al.* 1994. Saturated molecular map of the rice genome based on an interspecific backcross population. Genetics 138 (1994) 1251–1274.

[9] WILSON, W. A. *et al.* Inferences on the genome structure of progenitor maize through comparative analysis of rice, maize, and the domesticated panicoids. Genetics 1999 (in press).
[10] VELDBOOM, L., Lee R.M., Woodman, W.L. Molecular-marker-facilitated studies in an elite maize population: I. Linkage analysis and determination of QTL for morphological traits. Theor. Appl. Genet. 88 (1994) 7–16.
[11] LANDER, E.S. *et al.* MAPMAKER: An interactive computer package for maps of experimental and natural populations. Genomics 1 (1987) 174–181.
[12] LUI, S.-C. *et al.* Genome-wide high resolution mapping by recurrent intermating using *Arabidopsis thaliani* as a model. Genetics 142 (1996) 247–258.
[13] TARAMINO, G. *et al.* Characterization and mapping of simple sequence repeats (SSRs) in *Sorghum bicolor*. Theor. Appl. Genet. 95 (1997) 66–72.