Genetic, Physical, and Informatics Resources for Maize. On the Road to an Integrated Map[1]

Karen C. Cone*, Michael D. McMullen, Irie Vroh Bi, Georgia L. Davis, Young-Sun Yim, Jack M. Gardiner, Mary L. Polacco, Hector Sanchez-Villeda, Zhiwei Fang, Steven G. Schroeder, Seth A. Havermann, John E. Bowers, Andrew H. Paterson, Carol A. Soderlund, Fred W. Engler, Rod A. Wing, and Edward H. Coe Jr.

Division of Biological Sciences (K.C.C.) and Department of Agronomy (M.D.M., I.V.B., G.L.D., Y.-S.Y., J.M.G., M.L.P., H.S.-V., Z.F., S.G.S., S.A.H., E.H.C.), University of Missouri, Columbia, Missouri 65211; Agricultural Research Service, United States Department of Agriculture, Columbia, Missouri 65211 (M.D.M., M.L.P., E.H.C.); Center for Applied Genetic Technologies, Departments of Crop and Soil Science, Botany, and Genetics, University of Georgia, Athens, Georgia 30602 (J.E.B., A.H.P.); and Arizona Genomics Computational Laboratory (C.A.S., F.W.E.) and Arizona Genomics Institute (R.A.W.), University of Arizona, Tucson, Arizona 85721

Maize (*Zea mays*) is among the most important crop plants in the world. For any crop plant, an integrated genetic and physical map serves as the foundation for numerous studies, especially those aimed at improving the agronomic characteristics of the plant. Once a phenotypically defined locus controlling a trait of interest has been mapped genetically, the integrated map facilitates isolating the underlying gene by positional cloning. Gene isolation is the prelude to studies aimed at first elucidating how the gene functions to control the targeted trait and then applying this knowledge to crop improvement.

The maize genome presents a complex challenge to the development of an integrated genetic and physical map. The genome is large, approximately 2,500 megabases (Arumuganathan and Earle, 1991); it is laden with numerous families of transposable elements, whose copy numbers can be in the tens of thousands (Bennetzen, 2000; Myers et al., 2001); and sequence information is limited. Previous large-scale mapping projects, driven by the goal of genome sequencing, were aimed at selecting a minimal tiling path of genomic clones; as clones were sequenced, they were useful in verifying and merging contigs (International Human Genome Mapping Consortium, 2001; Chen et al., 2002).

To develop an integrated genetic and physical map resource for maize, we are using a comprehensive approach that includes three core components. The first is a high-resolution genetic map that provides essential genetic anchor points for ordering the physical map and for utilizing comparative information from other smaller genome plants. The physical map

component consists of contigs assembled from clones from three deep-coverage genomic libraries. The third core component is a set of informatics tools designed to analyze, search, and display the mapping data.

GENETIC MAP

As the genetic foundation of the integrated map for maize, we are using the intermated B73/Mo17 (IBM) map with more than 1,800 markers. Details of the IBM map are provided elsewhere (Davis et al., 2001; Lee et al., 2002).

An enhanced version of the IBM map is being generated to take advantage of the many markers that were mapped on previous mapping populations, but were not mapped on the IBM because of lack of detectable polymorphisms between B73 and Mo17. We are implementing a "neighbors" map approach in which we extrapolate locations of loci from non-IBM maps to their nearest neighbors on the IBM map, such that the framework loci on the IBM serve as a fixed backbone onto which additional loci are added. To extrapolate, we look for shared loci on the two maps that define an interval containing a locus of interest, calculate the distance between the shared and target loci on the non-IBM map as a ratio of the distance for the interval, and use the ratio to estimate a map coordinate for the target locus in that interval on the IBM. In choosing which neighbors to extrapolate, we consider the depth of the genetic data and the confidence levels for locus assignment to the non-IBM map. The new map is called "IBM Neighbors."

As a trial run for extrapolating to the new map, we used two high-density simple sequence repeat (SSR) maps, the SSR Tx303 x CO159 and the SSR T218 x GT119, with 457 and 288 SSRs, respectively, that were not on the IBM (Sharopova et al., 2002). The extrapolation added 385 loci to the IBM Neighbors map.
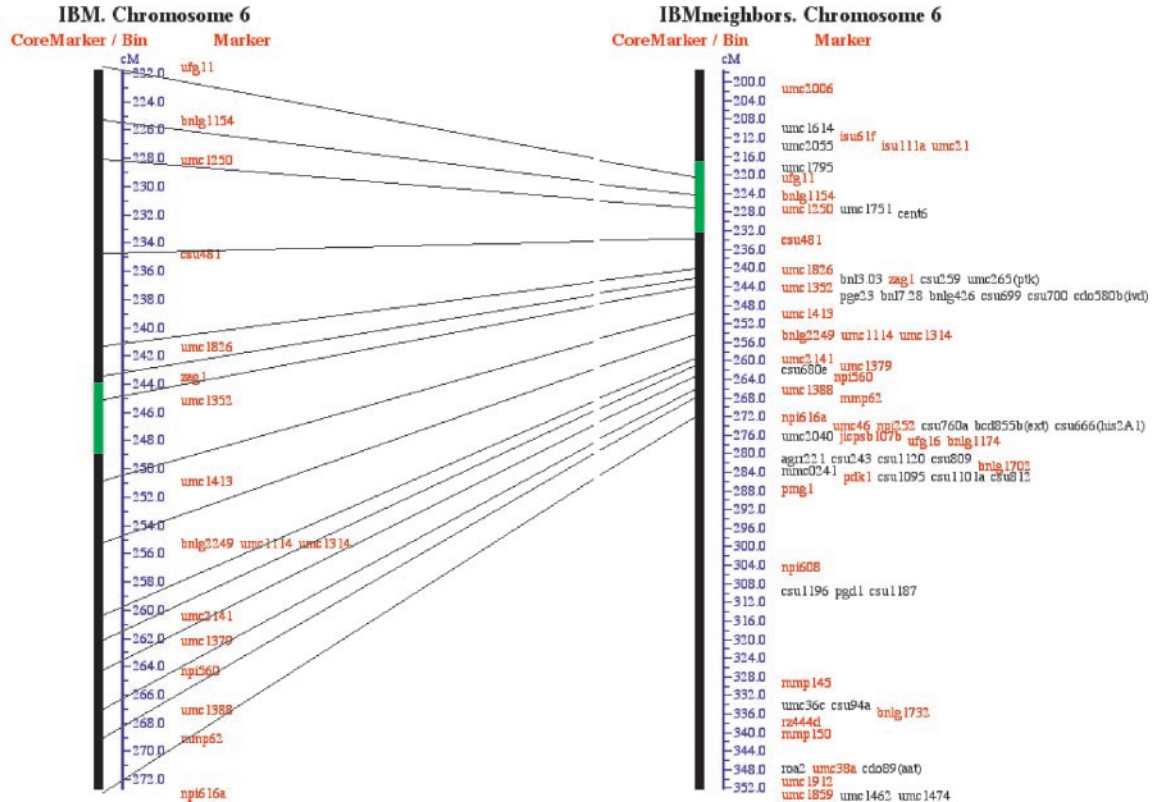
Figure 1. Comparison of a region of chromosome 6 from the IBM map (left) and the IBM Neighbors map (right). Loci in red are shared between the maps. Loci in black have been extrapolated from other maps to the IBM Neighbors map. This display was created using the Compare Maps (cMap) tool.

Extrapolations from other maps should add approximately 2,900 more markers, bringing the total to over 4,900. Figure 1 shows a comparative display of a region of chromosome 6 in the IBM and the IBM Neighbors maps. This region on the IBM has relatively few loci, but by extrapolating from other maps to the IBM Neighbors, many more loci are evident. The key distinction between the maps lies in the confidence level of locus order; the IBM has fewer well-ordered loci and IBM Neighbors has more loci, but confidence in the order is lower. For the user interested in a trait that maps to a particular interval on the IBM, having access to information for more loci in that interval should be advantageous.

The integrated map currently offers display of the IBM map, but plans are under way to present an additional view based on the IBM Neighbors map.

PHYSICAL MAP

Three genomic libraries were constructed in bacterial artificial chromosomes (BACs) using DNA from the inbred line B73. This line was used because it is one of the parents of the IBM genetic mapping population; thus, probes mapped on the IBM and used to screen the BAC libraries could provide direct anchors for integrating the genetic and physical maps. To provide for deep coverage of the genome and to minimize gaps in sequence representation, three restriction enzymes were used for library construction: HindIII, EcoRI, and MboI. Specifications of all three libraries were described previously (Coe et al., 2002) and detailed characterization of the HindIII library has been published (Tomkins et al., 2002). Together, the libraries represent 27-fold coverage of the genome with average insert sizes ranging from 137 to 167 kb.

The BACs are being assembled into contigs using the fingerprinting method of Marra et al. (1997). In brief, BACs are digested with HindIII, the digests are fractionated on high-resolution agarose gels, and banding patterns are catalogued to formulate a fingerprint using IMAGE software (Sulston et al., 1989). The HindIII banding patterns are then subjected to analysis to detect overlaps among them using the FingerPrintedContig (FPC) software package (Soder-

lund et al., 1997, 2000). Contigs are being generated automatically with a cutoff value of 1×10^{-12}. This cutoff value is similar to that used in assembly for other genomes. Contigs for the human (*Homo sapiens*) genome were assembled using a 3×10^{-12} cutoff (International Human Genome Mapping Consortium, 2001); contigs for the mouse (*Mus musculus*) genome were assembled initially using a 1×10^{-16} cutoff, but after aligning the mouse contigs with the human genome, the stringency was reduced to 1×10^{-12} for further assembly (Gregory et al., 2002; S. Gregory, personal communication). As more clones are analyzed, the chance of a false positive overlap increases as a result of association of questionable clones whose banding patterns do not align precisely with the map. A new FPC function called the de-Qer reassembles all contigs with questionable clones using a more stringent (lower) cutoff value so as to automatically remove the majority of falsely assembled clones (C. Soderlund, unpublished data). Once all fingerprints are collected, the contigs will be edited manually to merge adjacent contigs; postponing manual editing till the end of fingerprinting will make this process more efficient.

As of June 27, 2002, a total of 232,862 BAC clones, representing 12× genome coverage, had been fingerprinted and assembled into 5,318 contigs. Figure 2 demonstrates that contig numbers continue to decrease, as expected, as more BAC fingerprints are analyzed.

INTEGRATING THE GENETIC AND PHYSICAL MAPS

Screening the BAC libraries

Integrating the genetic and physical map depends on anchoring BAC contigs by association with genetically mapped probes. Two strategies have been used for screening of the BAC libraries: hybridization of radioactive probes to BAC filters and PCR-based analysis of BAC DNA pools. For the hybridization experiments, subsets of the *Hin*dIII and *Eco*RI librar-
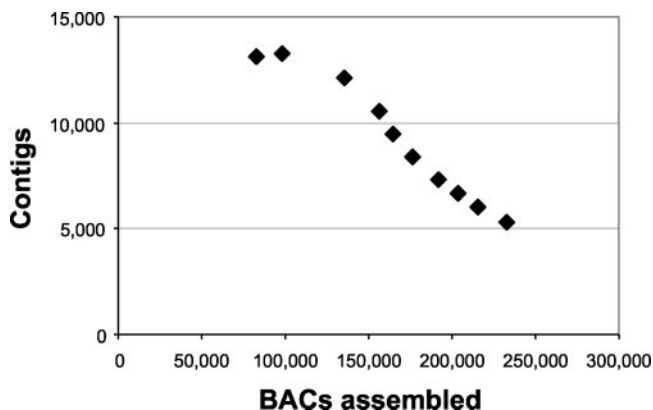
Figure 2. Decrease in contig number with increasing number of BACs fingerprinted.

ies were arrayed on high-density filters. For the PCR-based screening, DNA was isolated from a subset of the *Hin*dIII BACs that were pooled using a six-dimensional scheme (Klein et al., 2000).

In collaboration with DuPont (Wilmington, DE) and Incyte Genomics (Palo Alto, CA), overgo probes derived from a unigene set representing over 10,000 maize expressed sequence tags (ESTs) were hybridized to BAC filter arrays. The unigene set was generated by DuPont; it is called the Cornsensus, to distinguish it from other maize unigene sets. At Incyte Genomics, the unigene sequences were masked for repetitive sequences, and overgo probes were designed for each EST contig (Ross et al., 1999). Another group of overgo probes was derived from other grass sequences, including cDNA coding regions and hypomethylated genomic sequences that have been mapped genetically in sorghum (*Sorghum bicolor*) and/or other cereal genomes, including maize. The probes were radioactively labeled and hybridized to BAC filter arrays using a multiplex pooling strategy. The results allow us to assign BAC addresses to specific overgo sequences. Details of the Cornsensus development and overgo hybridization will be described elsewhere. All overgo:BAC association data are available at the Maize Mapping Project Web site (http://www.maizemap.org/resources.htm).

The BAC filters were also hybridized with probes representing mapped RFLP markers (Yim et al., 2002). Among the set of 90 core RFLP markers, most of the core probes yielded clear probe:BAC assignments, and the 22 that are single copy in the genome will serve as potential anchors linking the physical and genetic maps.

To increase the number of possible anchors, the BAC pools are being screened by PCR with primers developed from sequences corresponding to additional single-copy RFLPs and with SSR primers.

Assigning BAC Contigs to the Genetic Map

A contig can be assigned unambiguously to a specific genetic location if the contig is associated with (hit by) a probe that is mapped to one genetic location and the probe does not hit BACs in another contig. This is a simple concept, but in practice, a number of factors make contig anchoring a challenging prospect.

In an ideal world, the most direct route to anchoring the physical map to the genetic map would be to use single-copy probes derived from all mapped loci to screen the BAC libraries. This would require a completed high-resolution genetic map. In our case, genetic mapping and physical mapping were occurring simultaneously and we did not have the luxury of choosing anchoring probes based on their genetic location. Instead, our probes—mainly derived from the Cornsensus—were chosen for two reasons: They constituted a large number of genes for which se-

quence information was available and they represented expressed genes and, therefore, possible direct links to traits of interest to many researchers.

The Cornsensus unigene set contains genes that represent both mapped and unmapped sequences (Fig. 3). The unigenes corresponding to mapped sequences were identified by BLAST comparison with sequences representing SSR and RFLP markers on the IBM map (Altschul et al., 1997). This analysis provided about 800 potential anchor points. To incorporate the physical map information for the unmapped Cornsensus sequences into the integrated map, the unmapped Cornsensus unigenes have been targeted for development of SNP markers; approximately 1,000 of these will be mapped to provide additional anchors.

Many genetically mapped SSR and RFLP markers are not represented by the Cornsensus unigene set and, thus, were not used as probes in the initial BAC screening. As an additional anchoring strategy (Fig. 3), BAC addresses are being obtained for these sequences by PCR-based screening of the BAC DNA pools using SSR primers or primers generated from mapped RFLP sequences. This information is then incorporated into FPC for contig assembly.

To assign contigs to the genetic map, the most obvious rule is that one contig should have one genetic location. Unfortunately, the patterns of probe:BAC:contig associations obtained from FPC contig assembly reveal several possible types of conflict that are at odds with this rule. Although manual editing of the contigs will resolve some of these conflicts, this step is not scheduled to begin until fingerprinting is complete. In the meantime, we have analyzed the patterns of probe:BAC:contig association to develop a set of guidelines that we can use now to construct a working version of the integrated map that has minimal internal conflicts.

One very common probe:BAC pattern arises from probes that detect only a single BAC in a contig. This

pattern is unexpected because the redundancy of the BAC libraries predicts that a probe will hit multiple BACs in a contig. There are several reasons for this pattern. First, the set of BACs that was screened by filter hybridization is a subset (10×) of the BACs that are being fingerprinted (27×) and assembled into contigs. Thus, it is possible that more BACs in the contig would have been detected if all had been subjected to screening. Another possible explanation is that the sequence corresponding to the probe could be underrepresented in the libraries because of clone instability or asymmetric distribution of recognition sites for the enzymes used in library construction. An identified BAC could be a false positive, arising either by a scoring error or by spurious hybridization. However, PCR validation experiments indicate that false positives occur rarely, if at all. This is probably because of the high stringency at which the filters were washed and to the strict scoring criteria that required multiple positive hybridization signals (each double-spotted BAC clone had to hybridize to multiplexed overgos that were pooled in two or three dimensions) to establish each BAC:overgo association. These scoring demands mean that some true BAC:overgo associations may be missed. Such "false negatives" can lead to the mistaken assumption that an overgo that hits BACs that are all contained in one contig has a single genomic location. In PCR validation experiments, false negatives have been identified, but the extent to which these occur is unclear because they can only be identified by a directed search. To simplify the initial process of anchoring, we have chosen to filter out probe:contig associations if they were detected by hybridization of a probe to a single BAC in the contig. This filtering has significantly reduced the complexity of the probe:BAC: contig association (Fig. 4), allowing anchor points to be detected more readily. Single probe:contig associations made by PCR-based analysis of the BAC DNA pools are not filtered out because our experiences
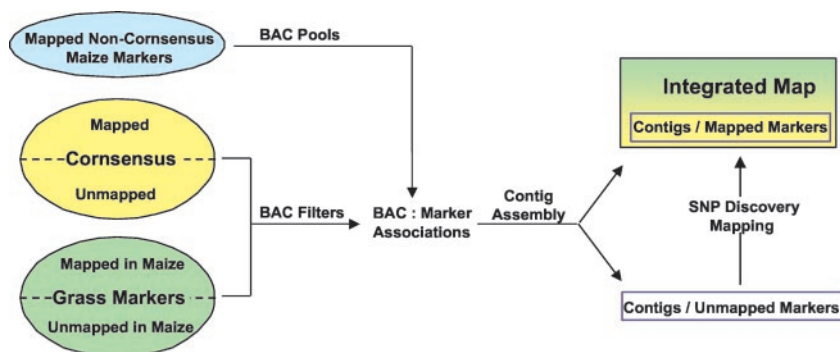
Figure 3. Anchoring strategy. Sources of markers for integrating the genetic and physical maps are shown on the left. Overgo probes from the Cornsensus unigene set and from a variety of grass markers were used to hybridize to the BAC filters, providing BAC:marker associations. Probes derived from mapped maize markers not in the Cornsensus are being screened on the BAC pools to derive additional BAC:marker associations. Contigs associated with mapped markers can serve as anchors for the integrated map. Unmapped markers associated with contigs are targets for mapping via single nucleotide polymorphisms (SNPs) to convert them to potential anchors.
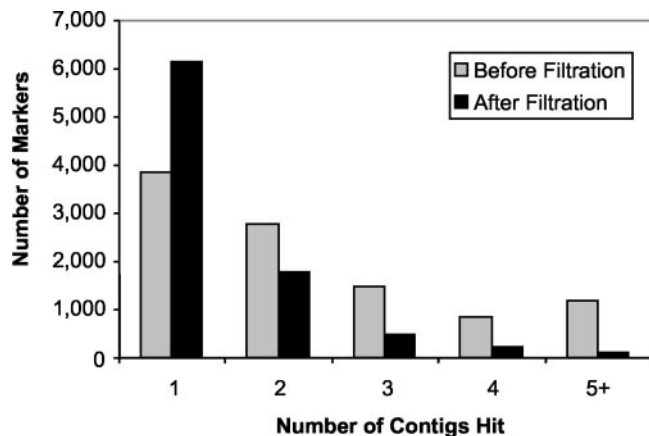
Figure 4. Effect of filtering out single probe:contig associations that are based on a single BAC hit in a contig. Graph shows the number of markers that detect one to five or more BACs. The first column in each pair represents non-filtered data. The second column represents the marker:contig associations after filtering.

indicate that fewer false positives are obtained with the PCR-based screening than with filter hybridization (Y. Yim and G. Davis, unpublished data).

Another potential source of conflict arises when a probe that maps to one genetic location hits BACs that have been assembled into more than one contig. This could occur because the probe represents a duplicated sequence in the genome and, therefore, is truly associated with more than one genetic location, perhaps only one of which has been mapped. This possibility is being addressed by developing gene-specific probes for the duplicated sequences and screening the BAC DNA pools to clarify the chromosomal assignment. Another explanation for one probe hitting two contigs could be that the two contigs have not yet been merged into one. This might be true especially if the probed BACs are at the ends of the contigs. This conflict will be resolved during manual editing by examining the fingerprints of the end clones of the contigs to ask if a merge is warranted.

An additional type of conflict occurs when two probes corresponding to distinct genetic locations associate with a single contig. This type of conflict could be because of errors in any of several steps:

BAC addressing, in silico determination of overgo: mapped marker identity, genetic mapping, or contig assembly. Resolution involves verifying results of BAC hybridization and pool analysis, double checking genetic mapping assignments, and editing contigs manually to test the possibility that the contig is chimeric and should be split apart. Finally, as contigs with genetically mapped markers are merged, a "majority" rule can be applied to assign the contig to a genetic position corresponding to the markers with the highest number of hits in the contig.

An important step in refining the integrated map will be to test its robustness. Several types of experiments are under way. For example, the BAC pools are being screened by PCR using primers derived from some of the same cDNA clones from which overgos were derived for hybridization-based screening of the BAC filters. Preliminary comparisons of the BACs identified by the two methods indicate that there is 80% agreement among the methods (Y. Yim and G. Davis, unpublished data). To verify the accuracy of anchoring, we are looking, within a contig, at the markers other than the anchoring marker to ask if there is any additional type of information available to indicate that they are associated with the same genetic location as the anchoring marker. In some cases, this confirming evidence may come from data derived from other genetic mapping experiments. In other cases, colocalization may be inferred from information gleaned from sorghum or rice (*Oryza sativa*) maps and extrapolated to maize, based on presumed syntenic relationships. Finally, as individual researchers in the maize community use the integrated map resources to isolate their genes of interest, their feedback will help validate the map.

INFORMATICS TOOLS

To retrieve and display the integrated map data, several Web-based tools have been developed. Table I lists the tools and their Web addresses. iMap provides the unambiguous contig:chromosome assignments in the context of side-by-side views of the genetic map and associated contigs. cMap allows comparison of locus order in different genetic maps

Table I. *Web tools for genetic, physical, and integrated maize map data*

| Web Tool	Web Site Address	Primary Use
iMap	http://www.maizemap.org/iMapDB/iMap.html	View unambiguous marker:BAC contig:chromosome associations in side-by-side displays of the genetic and physical maps
cMap	http://www.agron.missouri.edu/cMapDB/cMap.html	Compare locus order between pairs of maize, sorghum, or rice maps
WebFPC	http://genome.arizona.edu/fpc/maize	View BAC contigs and associated markers
WebChrom	http://genome.arizona.edu/fpc/maize/WebChrom	View all possible marker:BAC contig:chromosome associations as chromosome ideograms
Mapped Sequence Locator (MSL)	http://www.maizemap.org/cgi-bin/zmMSL.cgi	Determine if DNA sequence of interest has map location or homology to Cornsensus or other maize sequences

from maize, with extensions to rice and sorghum. WebFPC displays the maize BAC contig assemblies. WebChrom provides a chromosome-centered view of the WebFPC data. Because WebFPC and WebChrom present the data with little or no filtration, there may be ambiguity in contig:chromosome assignments. MSL searches an input sequence and returns map location and sequence homology to the Cornsensus and other maize sequences. Together these tools offer unprecedented access to the integrated genetic and physical mapping data for maize and allow the user to enter the access arena from a number of different points. The following sections describe the features of each tool and some typical applications for their use.

iMap

iMap presents a simultaneous display of the IBM genetic map and associated BAC contigs that have been unambiguously assigned to their respective genetic locations. The iMap tool was developed by adapting GIOT software originally written by the

Rice Genome Project. The iMap "front end" is part of a three-tier architecture for information retrieval in which iMap operates as an applet residing in a Web browser. Data are stored in a dedicated database populated with genetic map data and with FPC-derived physical map data. To retrieve information, requests are sent from the iMap applet (client) to a waiting servlet running on a server. The servlet retrieves the information from the database and delivers it to the client, which then displays the information. The database is updated regularly to ensure congruent connections to the most recent WebFPC data.

The display features side-by-side views of the genetic map and associated BAC contigs (Fig. 5). The database is searchable by genetic locus and plans are under way to expand the search utility to report information on map location, probe, sequence accession number, and contig number in the search results. For each locus, information about genetic marker type or contig can be displayed in popup windows. Details on genetic markers are retrieved
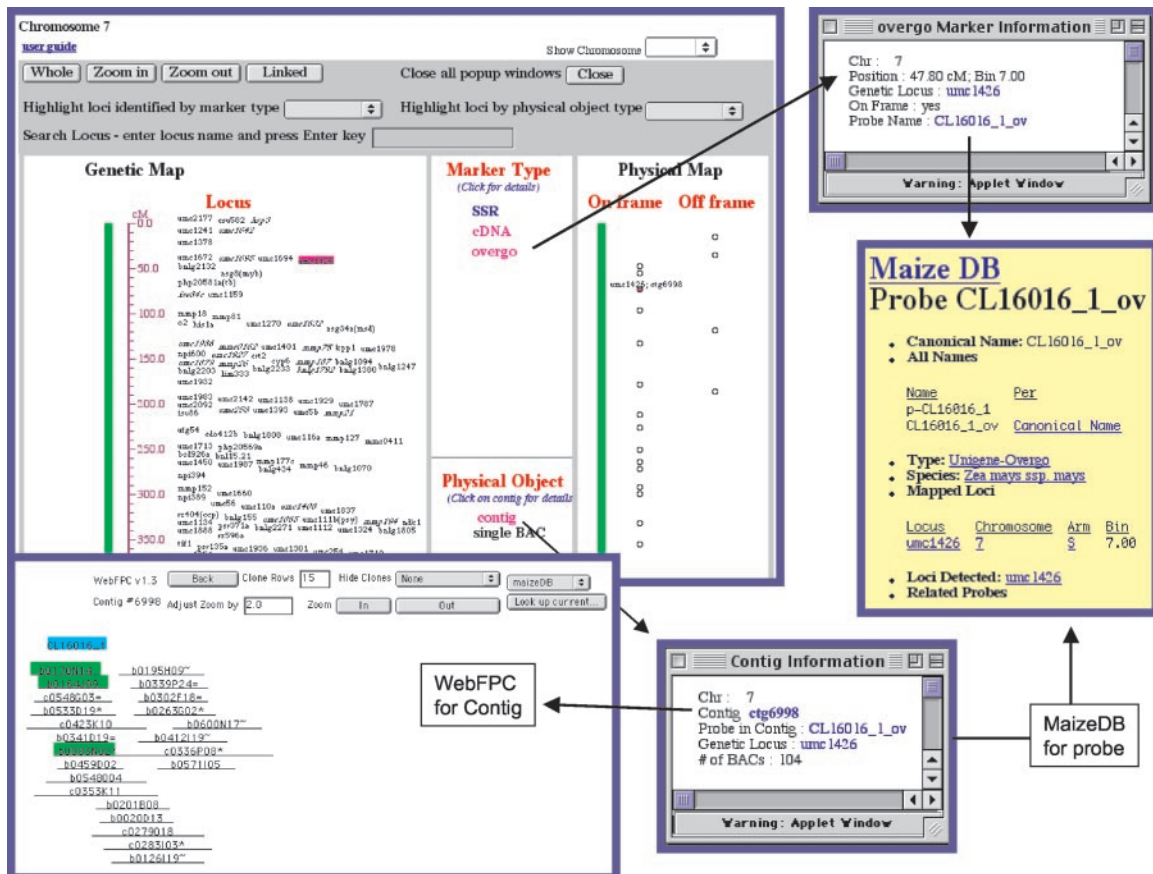
Figure 5. Representative display for iMap and linked resources. The iMap display includes side-by-side views of the genetic map on the left and associated BAC contigs (circles) on the right. Highlighting a locus on the genetic map leads to display of marker types and physical object types. Clicking on any of these words causes a pop-up window to appear with information and links to additional data. Representative links to MaizeDB for genetic marker information and to WebFPC for contig information are shown. In the WebFPC window, the marker highlighted in blue detects several BACs, which are highlighted in green.

via links to MaizeDB, TIGR, ZmDB, and GenBank; and details on BAC contigs are retrieved via a link to WebFPC.

iMap is the starting point for a user who has genetic map information and wishes to find anchored contigs associated with a favorite locus or neighboring loci. The user enters the name of the locus on the first page. The search returns a display of the genetic map of the appropriate chromosome with the locus highlighted. Following the highlighted links returns information about associated contigs.

cMap

cMap is a tool for comparing the order of loci in different maize maps. It also extends to comparisons in sorghum and rice maps. Like iMap, cMap was developed by adapting GIOT software. The display presents side-by-side views of selected pairs of genetic maps, highlights and lists the loci or probes they have in common, and shows connecting lines linking the shared loci. Figure 1 shows a portion of a cMap display. Links provide access to the underlying map data in MaizeDB.

If a user has genetic information about a rice or sorghum gene and wishes to find out if a corresponding gene in maize has been associated with BAC contigs, the starting point is cMap. Comparing rice or sorghum maps with the IBM will indicate if the gene of interest has a counterpart in maize. If so, the maize gene can be used as a query to search iMap for anchored contigs.

WebFPC

WebFPC, developed as a Web-based version of FPC, displays contigs and associated BAC clones. Figure 5 shows a representative contig display. Marker details are accessible via links to MaizeDB. The contigs in WebFPC are updated, approximately monthly, as new fingerprinting data are obtained.

WebFPC is searchable by BAC clone, probe (marker), or contig, and it is the starting point for a user who has the name of a probe (Cornsensus unigene or other) or the name of a BAC clone and wishes to find associated contigs. A search is conducted by entering the probe name (into the marker box) or BAC name (into the clone box). The search returns a list of contigs.

WebChrom

WebChrom is a companion tool to WebFPC and provides a chromosome-centric view of FPC contigs. WebChrom features a display of the 10 chromosomes of maize; clicking on any chromosome shows the chromosome with contigs and markers. Other views include distribution of probe or BAC markers along the chromosome. WebChrom uses different rules than iMap for placing contigs on chromosomes and contig:chromosome assignments may be different be-

tween the two tools. Thus, in WebChrom, it is possible that more than one contig might be displayed for a single chromosomal location based on association with a marker at that location.

If a user has genetic map information for a gene of interest and wishes to find any associated contigs, whether anchored unambiguously or not, the starting point is WebChrom. By selecting a chromosome of interest, the user can view all contigs that have been associated with that chromosome on the basis of the genetic location of probes that hit the contigs.

MSL

The MSL is a BLAST tool that uses an entered DNA sequence as a query against the Cornsensus or any other public maize sequence and returns BLAST score, sequence alignment, and map location (if any).

MSL is a good starting point to find out if a DNA sequence of interest is represented in the Cornsensus (from which overgo probes were developed to screen the BAC libraries). The sequence is pasted into the text box and a search is conducted against Cornsensus. The return will display the matches to Cornsensus sequences, sequence alignments, and genetic map location, if any. If a locus is identified, it can be used as a query in iMap to search for an anchored contig.

FUTURE PROJECTIONS

Integrated maps have never been assembled for a eukaryotic organism with the genomic structure of maize, e.g. large genome size and high content of recently evolved repetitive elements that are organized into large arrays. Mapping progress thus far suggests that the combination of robust contig assembly and anchoring with a large number of genetically mapped markers will produce a high-quality integrated map. With each FPC assembly, contig number is decreasing and we expect this trend to continue. What the final number of contigs will be is unknown, but a simple calculation provides a rough estimate: If 450,000 BACs, representing 27× genome coverage with average length of 150 kb, are fingerprinted and assembled at a stringency demanding 70% overlap among the clones, then 2,000 contigs could be expected (Lander and Waterman, 1988). The use of the de-Qer function of FPC, which minimizes false assemblies, and the incorporation of genetic anchoring information, which facilitates merging of contigs hit by adjacent markers, should ensure that the contig assemblies are robust.

To derive genetic anchors, we are taking a multidimensional approach (Fig. 3). Sources of anchors include: mapped markers (SSR and RFLP) identified in silico from the Cornsensus and from other grass genome maps or sequences, unmapped Cornsensus sequences to be targeted for SNP development and

mapping, and mapped non-Cornsensus markers (SSR and RFLP) to be placed on the physical map by screening of the BAC DNA pools. Our goal is to focus on large, gene-rich contigs to establish two anchors for each contig as a mechanism for verifying contig: genetic map assignments and determining the order of the contigs relative to the genetic map.

The genomics efforts in rice and sorghum will provide an excellent additional resource for integrating the maize map (Draye et al., 2001; Chen et al., 2002). The syntenic relationship among these grasses means that genetic:physical map associations discovered in rice or sorghum can be used as reference points for testing or confirming similar relationships in maize. As a starting point for making these comparisons, homology searches of the Cornsensus sequences have revealed that many share sequence similarity to genes in rice that have been placed on the integrated genetic and physical map. These similarities can be viewed in the context of the rice map, using a tool available at Gramene (http://www.gramene.org/perl/SeqTable), which displays an alignment of the maize Cornsensus unigenes (and other grass sequences including sorghum ESTs) with rice genomic clones. Comparisons such as this will be indispensable in translating information from rice and sorghum to maize.

DATA ACCESS

Table I presents the Web addresses for the tools developed for analyzing and searching the maize genomics resources. Additional information is available at the Maize Mapping Project Web site (http://www.maizemap.org) and at MaizeDB (http://www.agron.missouri.edu).

ACKNOWLEDGMENTS

LITERATURE CITED

Altschul SF, Madden TL, Schaffer AA, Zhange J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25**: 3389–3402

Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Rep **9**: 208–218

Bennetzen J (2000) Transposable element contributions to plant gene and genome evolution. Plant Mol Biol **42**: 251–269

Chen M, Presting G, Barbazuk W, Goicoechea J, Blackmon B, Fang G, Kim H, Frisch D, Yu Y, Sun S et al. (2002) An integrated physical and genetic map of the rice genome. Plant Cell **14**: 537–545

Coe E, Cone K, McMullen M, Chen S, Davis G, Gardiner J, Liscum E, Polacco M, Paterson A, Sanchez-Villeda H et al. (2002) Access to the maize genome: an integrated physical and genetic map. Plant Physiol **128**: 9–12

Davis G, Musket T, Melia-Hancock S, Duru N, Sharopova N, Schultz L, McMullen MD, Sanchez-Villeda H, Schroeder S, Garcia AA (2001) The intermated B73 × Mo17 genetic map: a community resource. Maize Genet Conference Abstracts **43**: W15, P62

Draye X, Lin Y, Qian X, Bowers J, Burow G, Morrell P, Peterson D, Presting GR, Wing RA, Paterson AH (2001) Toward integration of comparative genetic, physical, diversity, and cytomolecular maps for grasses and grains, using the *Sorghum* genome as a foundation. Plant Physiol **125**: 1325–1341

Gregory SG, Sekhon M, Schein J, Zhao S, Osoegawa K, Scott CE, Evens RS, Burridge PW, Cox TV, Fox CA et al. (2002) A physical map of the mouse genome. Nature **418**: 743–750

International Human Genome Mapping Consortium (2001) A physical map of the human genome. Nature **409**: 934–941

Klein P, Klein R, Cartinhour S, Ulanch P, Dong J, Obert J, Morishige D, Schlueter S, Childess K, Ale M et al. (2000) A high throughput AFLP-based method for constructing integrated genetic and physical maps: progress toward a sorghum genome map. Genome Res **10**: 789–907

Lander ES, Waterman MS (1988) Genomic mapping by fingerprinting random clones: a mathematical analysis. Genomics **2**: 231–239

Lee M, Sharopova N, Beavis W, Grant D, Katt M, Blair D, Hallauer A (2002) Expanding the genetic map of maize with the intermated B73 × Mo17 (IBM) population. Plant Mol Biol **48**: 453–461

Marra M, Kucaba T, Dietrich N, Green E, Brownstein B, Wilson R, McDonald K, Hillier L, McPherson J, Waterston R (1997) High throughput fingerprint analysis of large-insert clones. Genome Res **11**: 1072–1084

Myers BC, Tingey SV, Morgante M (2001) Abundance, distribution, and transcriptional activity of repetitive elements in the maize genome. Genome Res **11**: 1660–1676

Ross M, LaBrie S, McPherson J, Stanton V Jr (1999) Screening large-insert libraries by hybridization. *In* A Boyle, ed, Current Protocols in Human Genetics. Wiley, New York, pp 5.6.1–5.6.2.

Sharopova N, McMullen MD, Schultz L, Schroeder S, Sanchez-Villeda H, Gardiner J, Bergstrom D, Houchins K, Melia-Hancock S, Musket T et al. (2002) Development and mapping of SSR markers for maize. Plant Mol Biol **48**: 463–481

Soderlund C, Humphray S, Dunham A, French L (2000) Contigs built with fingerprints, markers and FPC v4.7. Genome Res **10**: 1772–1787

Soderlund C, Longden I, Mott R (1997) FPC: a system for building contigs from restriction fingerprinted clones. Comput Appl Biosci **13**: 523–535

Sulston J, Mallett F, Durvin R, Horsnell T (1989) Image analysis of restriction enzyme fingerprint audioradiograms. Comput Appl Biosci **13**: 101–106

Tomkins J, Davis G, Main D, Yim Y, Duru N, Musket T, Goicoechea J, Frisch D, Coe EH Jr, Wing R (2002) Construction and characterization of a deep-coverage bacterial artificial chromosome library for maize. Crop Sci **42**: 928–933

Yim Y-S, Davis G, Duru N, Musket T, Linton EW, Messing JW, McMullen MD, Soderlund C, Polacco M, Gardiner J, Coe EH Jr (2002) Characterization of three maize BAC libraries toward anchoring of the physical map to the genetic map using high density BAC filter hybrization. Plant Physiol **130**: 1686–1696