Construction and Characterization of a Deep-Coverage Bacterial Artificial Chromosome Library for Maize

J. P. Tomkins,* G. Davis, D. Main, Y. Yim, N. Duru, T. Musket, J. L. Goicoechea, D. A. Frisch, E. H. Coe, Jr., and R. A. Wing*

ABSTRACT

Modern cultivated maize (*Zea mays* L.)is one of the primary agronomic crops in the USA with an estimated genome size of 2500 megabases (Mb). To develop the resources for positional cloning and structural genomics in maize, we constructed a bacterial artificial chromosome (BAC) library for the inbred line B73 using the cloning enzyme *Hin*d III. The library contains 247 680 clones (645 384-well plates). A random sampling of 697 clones indicated an average insert size of 136 kilobase (kb) (range = 42 to 379 kb) and 0.4% empty vectors. Screening the colony filters for chloroplast DNA content indicated an exceptionally low 0.18% contamination with chloroplast DNA. Thus, the library provides 13.5 haploid genome equivalents allowing >99% probability of recovering any specific sequence of interest. High-density filters were gridded robotically using a Genetix Q-BOT (Hampshire, UK) in a 4 by 4 double-spotted array on 22.5-cm² filters. Partial screening (6× coverage) of the library with 20 single copy probes identified an average 7.1 positive signals per probe, with a range of 3 to 15 positive signals per probe. To evaluate the utility of the library for sequence tagged connector (STC) analysis, 768 BAC clones were end sequenced in both forward and reverse directions giving a total of 1415 successful reads. End sequences were queried against SWISS-PROT, Genbank NR, MIPS *Arabidopsis*, maize genomic sequence dbGSS, and maize cDNA database dbEST. Results in spreadsheet format from these searches is publicly available at the CUGI website (www.genome.clemson.edu/projects/stc/maize/ZMMBBb/).

The USA is a recognized world leader in the production and trade of agricultural commodities, and USA maize grain exports represent 80% of the total world trade in maize. While the yield of maize in the USA has steadily increased during the last 40 yr, it is beginning to level off (USDA–National Agricultural Statistics Service, 2001). With the expected increase in world population combined with concerns about the environmental effects of chemical intensive agricultural practices, it is clear that nontraditional efforts to increase maize yield and the incorporation of novel value-added traits are needed. The advent of genomics-based approaches to crop improvement provides an unprecedented opportunity to make significant genetic advances in maize.

The primary goal of the maize genomics community is to take the first step toward understanding the structure and function of the maize genome by developing and disseminating a comprehensive integrated physical and genetic map (University of Missouri Maize Genomics Center, 2001). A consolidated physical and genetic map of maize will enable gene discovery, studies of gene functions, and comparative genomics with other plant species, particularly the grasses. The information and resources that will result from this activity will be invaluable to basic research, the maize industry, and ultimately world maize consumers.

Modern cultivated maize is a diploid crop species having an estimated haploid genome size of 2500 Mb (Arumuganathan and Earle, 1991). An extensive history of genetic, mutagenesis, and cytogenetics studies has resulted in a large store of mutants, cytogenetic stocks, and knowledge about the genes that control maize functions, such as grain quality, yield, stress response and disease resistance (National Plant Germplasm System, 2000). A dense molecular marker linkage map for maize using simple sequence repeats, genomic clones, isozymes, and expressed sequence tagged sites (ESTs) has been developed (Davis et al., 1999) and is still being improved (University of Missouri Maize Genomics Center, 2001). While many important traits have been mapped in maize, physical mapping, structural analysis and map-based cloning have been limited due to the lack of a deep-coverage large-insert genomic library that would be available to the public maize research community.

An essential tool for characterizing genomes is the availability of yeast artificial chromosome (YAC) or BAC libraries containing large genomic DNA inserts. Plant YAC libraries have been constructed for several organisms, including *Arabidopsis* (Grill and Somerville, 1991), tomato (*Lycopersicon esculentum* Mill.; Martin et al., 1992), and maize (Edwards et al., 1992; Kleine et al., 1993). These libraries have been used for a number of studies, but their general use has been limited by the high frequency of chimeric and unstable clones. In contrast, BAC vectors from the mini-F plasmid allow cloning and stable maintenance of large DNA fragments in *Escherichia coli* (Shizuya et al., 1992). Bacterial artificial chromosome libraries are popular for a number of reasons, including ease of handling, relative simplicity to develop, and low frequency of chimeric clones. Bacterial artificial chromosome libraries have been developed for many plant species, including *Arabidopsis*, barley (*Hordeum vulgare* subsp. *vulgare*), cotton (*Gossypium hirsutum* L.), grape (*Vitis* spp.), rice (*Oryza sativa* L.), sorghum [*Sorghum bicolor* (L.) Moench], soybean [*Gly-*

J.P. Tomkins, D. Main, J.L. Goicoechea, and R.A. Wing, Clemson University Genomics Institute, Room 100 Jordan Hall, Clemson, SC 29634, www.genome.clemson.edu; G. Davis, N. Duru, T. Musket, E.H. Coe; Dep. of Agronomy, Univ. of Missouri-Columbia, 1-87 Agriculture, Columbia, MO 65211; D. Frisch, Genome Center of Wisconsin, 445 Henry Mall, Genetics Building B19, Madison, WI 53706. Received 6 Apr. 2001. *Corresponding authors (jtmkns@clemson.edu, rwing@clemson.edu).

Published in Crop Sci. 42:928–933 (2002).

Abbreviations: BAC, bacterial artificial chromosome; EST, expressed sequence tagged site; kb, kilobase; kbp, kilobase pairs; Mb, megabase; STC, sequence tagged connector; YAC, yeast artificial chromosome.

cine max (L.) Merr.], sugarcane (*Saccharum officinarum* L.), and tomato (Woo et al., 1994; Choi et al., 1995; Wang et al., 1995; Zhang et al., 1996; Marek and Shoemaker, 1997; Danesh et al., 1998; Mozo et al., 1998; Tomkins et al., 1999a, b, 2001, 2002; Yu et al., 2000).

Plant BAC libraries have been used for a number of structural genomics applications. In *Arabidopsis*, BAC libraries were used to develop sequence-ready physical maps (Bevan et al., 1998). In addition, BAC libraries have been used to positionally clone disease resistance genes and other genes known only by their phenotype (Song et al., 1995). Moreover, BAC libraries are useful for examining genomic structure. The genomic structure of the *a1-sh2* region of maize, sorghum, and rice has been extensively examined through the sequencing of BAC clones containing these genes (Chen et al., 1997, 1998; Tikhonov et al., 1999). Bacterial artificial chromosome libraries have also been used to develop physical maps for genomic regions containing resistance gene analog sequences (Marek and Shoemaker, 1997).

The primary objective of the present study was to establish the resources necessary to facilitate genomics research in maize. Specific objectives were to (i) develop a BAC library providing at least 10 haploid genome equivalents, (ii) characterize the library for insert size and chloroplast DNA content, (iii) screen the library with single-copy probes, and (iv) evaluate the utility of the library for STC analysis via end sequencing of BAC clones (Venter et al., 1996).

MATERIALS AND METHODS

Bacterial Artificial Chromosome Library Construction

Maize seed from the inbred line B73 were obtained from the Maize Genetics Cooperation Stock Center (Urbana, IL). Seed were germinated in the greenhouse and seedlings ≈18 cm tall were harvested just above the roots. Outer leaf layers were removed and the innermost leaf layers and the shoot meristem region were used for library construction. Bacterial artificial chromosome library construction was essentially the same as that described by Tomkins et al. (1999a, b), with the following modifications. The first size selection of *Hin*d III fragments used switch times of 1 to 40 s in a linear ramp. Fractions between 100 and 300 kilobase pairs (kbp) were cut from the gel and inserted into a second gel and run at a constant 3- to 5-s switch time to remove small trapped DNA fragments. After removing appropriate fractions from the second size selection, DNA was removed from the agarose by electroelution (Model 422 Electro-Eluter, Bio-Rad, Hercules, CA). Size-selected DNA was ligated into the pCUGI-1 BAC vector (Luo et al., 2001) and electroporated (Cell-Porator, Invitrogen, Carlsbad, CA) into *E. coli* (strain DH10B) cells using 320 to 330 V. Transformed cells were plated on 200 mL of selective medium (Luria-Bertani medium) in 24- by 24-cm plates (Genetix) with 12.5 μg mL^{-1} chloramphenicol, 0.55 m*M* isopropyl-beta-D-thiogalactopyranoside (IPTG), and 80 μg mL^{-1} X-Gal. After a 20-h incubation at 37°C, the plates were placed at room temperature in the dark for an additional 20 h to allow stronger color development of nonrecombinant colonies. Plates were either stored at 4°C or used immediately for picking. Recombinant white colonies were picked robotically using the Genetix Q-BOT and arrayed as individual clones in 384-well microtiter plates (Genetix) containing 50 μL freezing broth (Woo et al., 1994). After incubation overnight, microtiter plates were stored at −80°C. Three copies of the library were made using the replicating function of the Genetix Q-BOT and stored in separate −80°C freezers. Bacterial artificial chromosome clone characterization has been described previously by Tomkins et al. (1999a, 1999b).

Bacterial Artificial Chromosome Library Screening

High-density colony filters for hybridization based screening of the library were prepared using the Genetix Q-BOT. Clones were gridded in double spots using a 4 by 4 array with 6 fields per 22.5- by 22.5-cm nylon (Hybond N+, Amersham, Piscataway, NJ) filter. This gridding pattern allows 18 432 clones to be represented per filter. For the purpose of testing the library, screening was performed using 6 filters representing a 6× genome coverage. Filters used for this study were labeled A, B, C, D, E, and F, and are publicly available upon request. Colony filters were processed and hybridized using standard techniques (Sambrook et al., 1989). Screening for chloroplast DNA content was performed as described by Woo et al. (1994). Probes used for screening the library are described in Table 1.

Bacterial Artificial Chromosome End Sequencing

Preparation of BAC DNA for end sequencing was done in a 96-well format using standard alkaline lysis miniprep techniques. Sequencing reactions were set up according to manufacturer's instructions for the Big Dye Terminator chemistry (Applied Biosystems, Foster City, CA). Reactions were performed using both forward and reverse universal primers. Samples were loaded onto 48-lane sequencing gels in ABI377 automated sequencers. Gels (250 mL) were composed of the following: 5% (w/v) Long Ranger (FMC, Philadelphia, PA), 6 *M* urea, *N*,*N*,*N*′,*N*′-tetramethylethylenediamine (TEMED) 18 μL, 150 μL ammonium persulfate (10% stock), and 1 × 1,1,2,2-Tetrabromoethane (TBE) buffer. Reaction products were electrophoresed using a 3.5-h run. Base-calling was performed automatically using PHRED (Ewing and Green, 1998; Ewing et al., 1998), and vector sequences were removed by CROSS-MATCH (University of Washington Genome Center, 2001). High quality BAC end sequences (defined as those having >100 nonvector bases with a PHRED quality value >20) were used as queries in searches of SWISS-PROT (Bairoch and Apweiler, 2000), Genbank NR, MIPS *Arabidopsis*, maize genomic sequence dbGSS, and maize cDNA database dbEST. Results in spreadsheet format from these searches is publicly available at the CUGI website (Clemson University Genomics Institute, 2001a). All software was run locally on a Sun Ultra30 workstation using Solaris 2.6 (Sun Microsystems, Palo Alto, CA).

RESULTS

Bacterial Artificial Chromosome Library Construction and Characterization

We have constructed a BAC library for maize using the inbred line B73 that is suitable for physical mapping, map-based cloning, and high-throughput sequencing of selected genomic regions. The library consists of 247 680 clones stored in 645 384-well microtiter plates. Approximately 0.4% (3/697 samples) of the clones do not contain inserts as judged by random analysis of BACs sampled from the library. A random sampling of 697 BACs taken

Table 1. Maize B73 bacterial artificial chromosome (BAC) library hybridization results using 20 single copy restriction fragment length polymorphism (RFLP) probes. Six high-density BAC colony filter arrays were used for each probing, providing a screen of approximately six haploid genome equivalents.

Probe	Description	No. hits
agrr115	Maize RFLP core marker for bin 4.01, cDNA	9
agrr21	Maize RFLP core marker for bin 8.09, cDNA	6
asg12	Maize RFLP core marker for bin 9.07, genomic fragment (G13185, G13186)	10
asg20	Maize RFLP core marker for bin 2.08, genomic fragment (G10750, G10751)	10
asg45	Protein kinase homolog, maize RFLP core marker for bin 1.04, genomic fragment (G10757, G10756)	8
asg49	Maize RFLP core marker for bin 7.03, genomic fragment (G10758, G10759)	5
bnl2.369	Maize RFLP core marker for bin 8.05, cDNA	5
bnl4.36	Maize RFLP core marker for bin 5.04, genomic fragment	5
csu147	Maize RFLP core marker for bin 9.04, cDNA (T12740)	9
npi409	Maize RFLP core marker for bin 5.01, genomic fragment (G13178, G10788)	6
npi414	Maize RFLP core marker for bin 8.08, genomic fragment (G10789, G10790)	3
php10017	Maize RFLP core marker for bin 5.09, genomic fragment (G10791, G10792)	8
tub4	beta tubulin 4 gene, maize RFLP core marker for bin 5.03, cDNA (L10635)	10
umc109	Maize RFLP core marker for bin 9.01, genomic fragment (G13177, G10807)	10
umc161	Maize RFLP core marker for bin 1.11, genomic fragment (G10824, G10825)	5
umc15	Waxy1, maize RFLP core marker for bin 9.03, cDNA (S46940)	3
umc34	Maize RFLP core marker for bin 2.04, genomic fragment (G10839, G10840)	5
umc65	Maize RFLP core marker for bin 6.04, genomic fragment (G10860, G10861)	4
umc85	Maize RFLP core marker for bin 6.01, genomic fragment (G10869)	6
tub1	beta tubulin 1gene, maize RFLP core marker for bin 1.01, genomic fragment (X52878)	15

from the completed library indicated an average insert size of 136 kb with a range of 42 to 379 kb. Figure 1 shows 43 randomly selected clones digested with *Not*I to release the insert. The two *Not*I sites in pCUGI-1 flank the multicloning site. Because *Not*I recognizes an 8-bp GC sequence and the maize genome is relatively GC rich, digestion typically generates vector plus 2 to 6 insert bands per BAC clone.

To determine the size distribution of BAC clones in the library, the 697 BACs analyzed with *Not*I digests were grouped by insert size, and the insert size of each clone was plotted against the frequency of each group of clones represented in the library (Fig. 2). On the basis of this analysis, >88% of the clones in the library have an average insert size ≥100 kbp. On the basis of the average insert size and a haploid genome size of 2500 Mb (Arumuganathan and Earle, 1991), the coverage of the library is ≈13.5 genome equivalents, resulting in >99% probability of recovering any specific sequence of interest.

To obtain an estimate of the representation of chloroplast DNA in the library, the filters were screened with three different chloroplast genes spaced equidistantly around the 133 kbp barley chloroplast genome. Results from this screening showed that ≈0.18% of library sequences are chloroplast DNA (data not shown). This is an exceptionally low level of chloroplast sequence contamination that we attribute to the use of the inner layers of young shoots as a nuclei source.

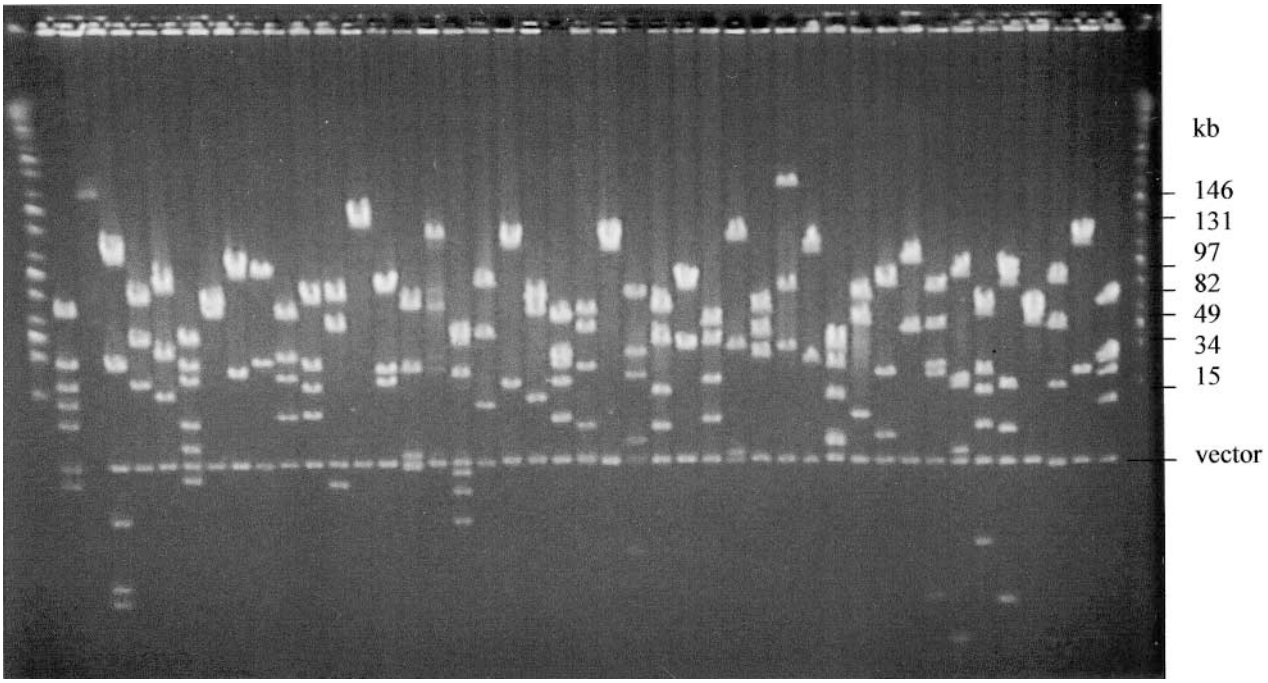

Fig. 1. Analysis of 43 randomly selected maize B73 bacterial artificial chromosome clones. Ethidium bromide stained contour-clamped homogeneous electric field gel (5–15 s switch time, 15 h) showing insert DNA above and below the common 7.5 kilobase (kb) pCUGI-1 vector band. Molecular weight marker in outside lanes is Midrange I (NEB).

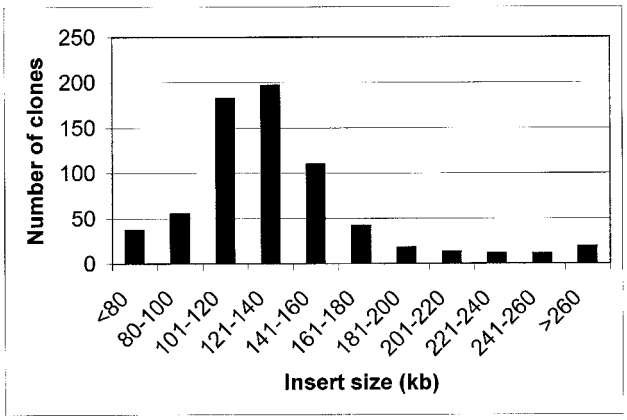

Fig. 2. Insert size distribution of clones in the maize B73 bacterial artificial chromosome (BAC) library. To estimate insert size range, BAC DNA from 697 randomly selected clones were analyzed with *Not*I digests and contour-clamped homogeneous electric field gels. Results indicate that the average insert size is 136 kilobase (kb) with >88% of the clones >100 kb.

Bacterial Artificial Chromosome Library Screening

To test the library for coverage and quality, six BAC colony filters (representing six haploid genome equivalents) were screened using 20 different probes representing single copy sequences from 8 maize chromosomes (Table 1). These gene-specific probes identified an average of 7.1 ± 3.02 positive clones with a range of 3 to 15 positive clones. Student *t*-test was performed to compare the mean of actual positives, 7.1, vs. the expected number of six. The statistical test ($P = 0.120$) indicates the difference between mean number of observed positives to the number of expectation is considered to be not statistically significant. The wide range of positive signals identified between probes may be indicative of the effects of preferential cloning obtained from the use of the *Hin*d III restriction enzyme. The hybridization data confirms adequate genomic representation of the maize BAC library throughout the whole maize genome.

Bacterial Artificial Chromosome End Sequencing

To examine the feasibility of using a STC strategy (Venter et al., 1996) to establish a framework for sequencing selected regions of the maize genome, we sequenced and analyzed the ends (forward and reverse) of the first 768 clones in the library. A total of 1536 sequencing reactions were performed, giving 1415 successful reads (success rate = 92%) with an average raw base count of 753. High-quality sequences were defined as those having >100 high-quality bases other than vector and *E. coli* sequence, and a PHRED score of 20 or greater. The number of high-quality sequences was 945, with an average high-quality base count of 290. Redundancy was evaluated by querying the high quality sequences against themselves with the result that a nonredundant data set of 873 sequences was developed. High quality nonredundant sequences were searched against the SWISS-PROT and MIPS *Arabidopsis* databases using the FASTX algorithm, and significance was determined with a probability cutoff value (E value) of at least 10^{-5}.

Table 2. Results from query of maize bacterial artificial chromosome end sequences against the SWISS-PROT database. Significant hits ($N = 206$, $E < 1 \times 10^{-5}$)† were categorized according to function. These results are based on the proportion of sequences showing hits, not reads.

Functional category	No.	Percentage
		%
Retroelement related	172	83
Enzymatic (metabolic, photosynthesis)	10	5
Regulatory (kinase, phosphatase, transcriptional, transport)	7	3
Structural	13	6
Ribosomal	2	1
Hypothetical or unclassified	2	1
Cell defense, communication, and division	0	–

† E = probability cutoff value.

As indicated above, multiple protein and DNA sequence databases were queried with the maize BAC end sequences. With the exception of the SWISS-PROT results, the database searches yielded homology hits primarily on unclassified DNA or protein sequence entries (Clemson University Genomics Institute, 2001a). For the purpose of discussion in this report, results from the SWISS-PROT search will be presented. SWISS-PROT is a curated protein sequence database that provides a high level of annotation, such as the description of protein function, domains structure, post-translational modifications, and variants. Other benefits include a minimal level of redundancy and high level of integration with other databases (Bairoch and Apweiler, 2000). The SWISS-PROT search resulted in 206 (11%) of the sequences showing similarity to genes of known function. Significant search results were then sorted into seven different functional categories (Table 2). A majority of the BAC end sequences shared sequences similar to retroelements and constituted the major component of the data set (83%). The next largest category of BAC end sequences were those involved in structural roles (6%), followed by those related to metabolism or photosynthesis (5%).

DISCUSSION

We describe the development and characterization of a high quality deep-coverage BAC library for the maize inbred line B73, a popular line for genetic studies, including mapping. This large insert library provides an important resource for map-based cloning, physical mapping, and genome sequencing. The library has been deposited in the Clemson University Genomics Institute BAC/EST Resource Center and is publicly available. Requests for the library, high-density BAC colony filter arrays, and clones can be made by accessing the Clemson University Genomics Institute web page (Clemson University Genomics Institute, 2001b).

One of the major applications of the maize BAC library will be for physical mapping of gene-rich regions of maize. There is substantial evidence that the genes in cereal genomes are not distributed evenly throughout

the genome, but rather are found in gene-rich regions separated by large blocks of retrolelements (Gill et al., 1993; Civardi et al., 1994; DeScenzo and Wise, 1996; Gill et al., 1996; Buschges et al., 1997). Examining gene distribution in three Gramineae (barley, maize, and rice) revealed similar gene distribution in these genomes (Carels et al., 1995; Barakat et al., 1997). Gene space, as these authors refer to gene-rich regions, occupies 12, 17, and 24% of the genomes of barley, maize, and rice, respectively. Conversely, in maize, intergenic regions composed primarily of retroelements has been estimated at up to 80% of the genome (San Miguel et al., 1998). Gene distribution studies at the nucleotide level for segments of the maize genome have revealed similar patterns to those described above. In two relatively large regions of the maize genome, a 225-kb section containing the *adh1* gene (Tikhonov et al., 1999) and a 78-kb segment containing a zein gene cluster (Llaca and Messing, 1998), genic regions were found to be interspersed among large blocks of repetitive DNA. In extreme cases, gene islands in maize can be very dense as Fu et al. (2001) recently described a 32-kb region containing 10 genes with an average intergenic spacing of <1 kb. Typically, genes within gene islands in maize are generally thought to be separated by 10- to 70-kb stretches of retroelements (Walbot et al., 2001). Although these studies are limited, they are consistent with previous work showing that genes are distributed unevenly across the genome into gene-rich and gene-poor regions. The availability of the maize BAC library will be an important tool for developing physical maps of the gene-rich regions. Because the maize genome is syntenous with other grass genomes, physical maps of the gene-rich regions of the maize genome will be extremely useful for examining microsynteny of particular regions across grass species. In addition, physical maps will be a starting point for positionally cloning genes from any member of the grasses.

End sequencing of 768 BAC clones (forward and reverse) produced a nonredundant high-quality set of 873 BAC end sequences. Search results against the SWISS-PROT database gave 206 significant hits ($>1 \times 10^{-5}$). Results were manually sorted according to function. Not surprisingly, a large proportion of query results were retroelement related sequences (83%). Previous BAC end sequence survey studies of a similar magnitude in grape, tomato, and polyploid cotton produced BAC retroelement contents of 41, 48, and 48%, respectively (Budiman et al., 2000; Tomkins et al., 2002a, 2002b). The high level of retroelement content for maize is due to the fact that the basic haploid genome size of maize is approximately two to five times the amount of these other plant genomes surveyed. The large number of retroelement hits and low number of gene hits will make STC database development in maize considerably less informative than in many plant species with smaller genome sizes. One must also keep in mind, however, that an STC strategy for sequencing selected genomic regions was used successfully in the human genome, which has an estimated genome size 250 to 500 Mb larger than maize (Venter et al, 1997).

ACKNOWLEDGMENTS

Appreciation for technical assistance during the process of library replication and filter production is extended to Michael Atkins, Kerry Wilson, and Stephanie Brown. This research was supported by National Science Foundation (NSF) Plant Genome grants no. 9872630 and 9872655, and the Coker Endowment to R. Wing.

REFERENCES

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

Bairoch, A., and R. Apweiler. 2000. The SWISS-PROT protein sequence database and its supplement TrEMBL. Nucleic Acids Res. 28:45–48.

Barakat, A., N. Carels, and G. Bernardi. 1997. The distribution of genes in the genomes of Gramineae. Proc. Natl. Acad. Sci. (USA) 94:6857–6861.

Bevan, M., I. Bancroft, E. Bent, K. Love, H. Goodman, C. Dean, R. Bergkamp, W. Dirkse, M. Van Staveren, W. Stiekema, L. Drost, P. Ridley, S.A. Hudson, K. Patel, G. Murphy, P. Piffanelli, H. Wedler, E. Wedler, R. Wambutt, T. Weitzenegger, T. M. Pohl, N. Terryn, J. Gielen, R. Villarroel, R. De Clerck, M. Van Montagu, A. Lecharny, S. Auburg, I. Gy, M. Kreis, N. Lao, T. Kavanagh, S. Hempel, P. Kotter, K.D. Entian, M. Rieger, M. Schaeffer, B. Funk, S. Mueller-Auer, M. Silvey, R. James, F. Montfort, A. Pons, P. Puigdomenech, A. Douka, E. Voukelatou, D. Milioni, P. Hatzopoulos, E. Piravandi, B. Obermaier, H. Hilbert, A. Dusterhoft, T. Moores, J. D. Jones, T. Eneva, K. Palme, V. Benes, S. Rechman, W. Ansorge, R. Cooke, C. Berger, M. Delseny, M. Voet, G. Volckaert, H.W. Mewes, S. Klosterman, C. Schueller, and N. Chalwatzis. 1998. Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. Nature (London) 391:485–488.

Budiman, A., L. Mao, T. Wood, and R. Wing. 2000. A deep-coverage tomato BAC library and prospects toward development of an STC framework for genome sequencing. Genome Res. 10:129–136.

Buschges, R., K. Hollricher, R. Panstruga, G. Simons, M. Wolter, A. Frijters, R. van Daelon, T. van der Lee, P. Diergaarde, J. Groenendijk, S. Topsch, P. Vos, F. Salamini, and P. Schulze-Lefert. 1997. The barley Mlo gene: A novel control element of plant pathogen resistance. Cell 88:695–705.

Carels, N., A. Barakat, and G. Bernardi. 1995. The gene distribution of the maize genome. Proc. Natl. Acad. Sci. (USA) 92:11057–11060.

Chen, M., P. SanMiguel, and J.L. Bennetzen. 1998. Sequence organization and conservation in *sh2/a1*-homologous regions of sorghum and rice. Genetics 148:435–443.

Chen, M., P. SanMiguel, A.C. De Oliviera, S.S. Woo, H. Zhang, R.A. Wing, and J.L. Bennetzen. 1997. Microcolinearity in *sh2*-homologous regions of the maize, rice and sorghum genomes. Proc. Natl. Acad. Sci. (USA) 94:3431–3435.

Choi, S., R. Creelman, J.E. Mullet, and R.A. Wing. 1995. Construction and characterization of a bacterial artificial chromosome library of *Arabidopsis thaliana*. Plant Mol. Biol. Rep. 13:124–128.

Civardi, L., Y. Xia, K.J. Edwards, P.S. Schnable, and B.J. Nikolau.. 1994. The relationship Between genetic and physical distances in the cloned *a1-sh2* interval of the *Zea mays* L. genome. Proc. Natl. Acad. Sci. (USA) 91:8268–8272.

Clemson University Genomics Institute. 2001a. Maize ZMMBBb STCs [Online]. [2 p.] Available at: http://www.genome.clemson.edu/projects/stc/maize/ZMMBBb/ [modified 22 Oct. 2001; verified 11 Dec. 2001]. CUGI, Clemson, SC.

Clemson University Genomics Institute. 2001b. CUGI home page [Online]. [1 p.] Available at: http://www.genome.clemson.edu/ [modified 30 Aug. 2001; verified 11 Dec. 2001]. GUGI, Clemson, SC.

Danesh, D., S. Penuela, J. Mudge, R. Denny, H. Nordstrom, J. Martinez, and N. Young. 1998. A bacterial artificial chromosome library for soybean and identification of clones near a major cyst nematode resistance gene. Theor. Appl. Genet. 96:196–202.

Davis, G.L., M.D. McMullen, C. Baysdorfer, T. Musket, D. Grant, M. Staebell, G. Xu, M. Polacco, L. Koster, S. Melia-Hancock, K. Houchins, S. Chao, and E.H. Coe, Jr. 1999. 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 152:1137–1172.
DeScenzo, R.A., and R.P. Wise. 1996. Variation in the ratio of physical to genetic distance in intervals adjacent to the *Mla* locus on barley chromosome 1H. Mol. Gen. Genet. 251:472–482.
Edwards, K.J., H. Thompson, D. Edwards, A. de Saizieu, C. Sparks, J.A. Thompson, A.J. Greenland, M. Eyers, and W. Schuch. 1992. Construction and characterization of a yeast artificial chromosome library containing three haploid maize genome equivalents. Plant Mol. Biol. 19:299–308.
Ewing, B., and P. Green. 1998. Base-calling of automated sequencer traces using PHRED. II. Error probabilities. Genome Res. 8:186–94.
Ewing, B., L. Hillier, M.C. Wendl, and P. Green. 1998. Base-calling of automated sequencer traces using PHRED. I. Accuracy assessment. Genome Res. 8:175–85.
Fu, H., and W. Park, X. Yan, Z. Zheng, B. Shen, and H.K. Dooner. 2001. The highly recombinogenic *bz* locus lies in an unusually gene-rich region of the maize genome. Proc. Natl. Acad. Sci. (USA) 98: 8903–8908.
Gill, K.S., B.S. Gill, and T.R. Endo. 1993. A chromosome region-specific mapping strategy reveals gene-rich telomeric ends in wheat. Chromosoma 102:374–381.
Gill, K.S., B.S. Gill, T.R. Endo, and E.V. Boyko. 1996. Identification and high-density mapping of gene-rich regions in chromosome Group 5 of wheat. Genetics 143:1001–1012.
Grill, E., and C. Somerville. 1991. Construction and characterization of a yeast artificial chromosome library of *Arabidopsis* which is suitable for chromosome walking. Theor. Appl. Genet. 226:484–490.
Kleine, M., W. Michalek, A. Graner, R.G. Herrmann, and C. Jung. 1993. Construction of a barley (*Hordeum vulgare* L.) YAC library and isolation of a Hor1-specific clone. Mol. Gen. Genet. 240:265–272.
Llaca, V., and J. Messing. 1998. Amplicons of maize in genes are conserved within genic, but expanded and constricted in intergenic regions. Plant J. 15:211–220.
Luo, M., Y. Wang, D. Frisch, T. Joobeur, R. Wing, and R. Dean. 2001. Melon BAC library construction using improved methods and identification of clones linked to the locus conferring resistance to melon Fusarium Wilt (*Fom*-2). Genome 44:154–162.
Marek, L.F., and R.C. Shoemaker. 1997. BAC contig development by fingerprint analysis in soybean. Genome 40:420–427.
Martin, G.B., M.W. Ganal, and S.D. Tanksley. 1992. Construction of a yeast artificial chromosome library of tomato and identification of cloned segments linked to two disease resistance loci. Mol. Gen. Genet. 233:25–32.
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.
National Plant Germplasm System. 2000. Maize Genetics COOP Stock Center's home page [Online]. [1 p.] Available at: http://w3.aces.uiuc.edu/maize-coop/ [modified 22 May 2000; verified 11 Dec. 2001]. NPGS, Dep. of Crop Sciences, Univ. of Illinois, Ubana, IL.
Sambrook, J., E. Frisch, and T. Maniatus. 1989. Molecular cloning: A laboratory manual. 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
SanMiguel, P., B.S. Gaut, A. Tihonov, Y. Nakajima, and J.L. Bennetzen. 1998. The paleontology fo intergene retrotransposons of maize. Nat. Genet. 20:43–45.
Shizuya, H., B. Birren, U.J. Kim, V. Manicino, T. Slepak, Y. Tashiiri, and M. Simon. 1992. Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Eschericihia coli* using an F-factor based vector. Proc. Natl. Acad. Sci. (USA) 89:8794–8797.
Song, W.Y., G.L. Wang, L.L. Chen, H.S. Kim, L.Y. Pi, T. Holsten, J. Gardner, B. Wang, W.X. Zhai, L.H. Zhu, C. Fauquet, and P. Ronald. 1995. A receptor kinase protein encoded by the rice disease resistance gene *Xa21*. Science (Washington, DC) 270:1804–1806.
Tikhonov, A.P., P.J. San Miguel, Y. Nakajima, N.M. Gorenstein, J.L. Bennetzen, and Z. Avramova. 1999. Colinearity and its exceptions in orthologous *adh* regions of maize and sorghum. Proc. Natl. Acad. Sci. (USA) 96:7409–7414.
Tomkins, J.P., R. Mahalingham, H. Miller-Smith, J.L. Goicoechea, H.T., Knapp, and R.A. Wing. 1999a. A soybean bacterial artificial chromosome library for PI 437654 and the identification of clones associated with cyst nematode resistance. Plant Mol. Biol. 41:25–32.
Tomkins, J.P., D.G. Peterson, T.J. Yang, D. Main, E.F. Ablett, R.J. Henry, L.S. Lee, T.A. Holton, D. Waters, and R.A. Wing. 2002. Tools for genome analysis in Grape (*Vitis vinifera* L.): BAC library construction, preliminary STC analysis, and identification of clones associated with flavonoid and stilbene biosynthesis. Am. J. Enol. Vitic. (In press).
Tomkins, J.P., D.G. Peterson, T.J. Yang, D. Main, T.A. Wilkins, A.H. Paterson, and R.A. Wing. 2001. Development of genomic resources for cotton (*Gosypium hirsutum*): BAC library development, preliminary STC analysis, and identification of clones associated with fiber development. Mol. Breed. 8:255–261.
Tomkins, J.P., Y. Yu, H. Miller-Smith, D.A. Frisch, S.S. Woo, and R.A. Wing. 1999b. A bacterial artificial chromosome library for sugarcane. Theor. Appl. Genet. 99:419–424.
University of Missouri Maize Genomics Center. 2001. Maize mapping project [Online]. [1 p.] Available at http://cafnr.missouri.edu/mmp/index.htm [modified 1 Nov. 2001; verified 11 Dec. 2001.] Univ. of Missouri, Clemson Univ., and Univ. of Georgia, Columbia, MO.
University of Washington Genome Center. 2001. University of Washington Genome Center [Online]. [2 p.] Available at: http://www.genome.washington.edu/UWGC/index.cfm [modified 8 Aug. 2001; verified 11 Dec. 2001]. UWGC, Seattle, WA.
USDA–National Agricultural Statistics Service. 2001. National Agricultural Statistics Service for U.S. agriculture statistical information and graphics [Online]. [1 p.] Available at: http://www.usda.gov/nass/ [modified 7 Dec. 2001; verified 11 Dec. 2001]. USDA–NASS, Washington, DC.
Venter, C., H.O. Smith, and L. Hood, 1996.A new strategy for genome sequencing. Science (Washington, DC) 381:364–366.
Walbot, V. and D.A. Petrov. 2001. Gene galaxies in the maize genome. Proc. Natl. Acad. Sci. (USA) 98:8163–8164.
Wang, G.L., T.E. Holsten, W.Y Song, H.P. Wang, and P.C. Ronald. 1995. Construction of a rice bacterial artificial chromosome library and identification of clones linked to the Xa-21 disease resistance locus. Plant J. 7:525–533.
Woo, S.S., J. Jiang, B.S. Gill, A.H. Paterson, and R.A. Wing. 1994. Construction and characterization of a bacterial artificial chromosome library of Sorghum bicolor. Nucleic Acids Res. 22:4922–4931.
Yu, Y., J.P. Tomkins, R. Waugh, D.A. Frisch, D. Kudrna, A. Kleinhofs, R. Brueggeman, G. Muehlbauer, R. Wise, and R.A. Wing. 2000. A bacterial artificial chromosome library for barley (*Hordeum vulgare* L.) and the identification of clones containing putative resistance genes. Theor. Appl. Genet. 101:1093–1099.
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.