Copyright © 2001 by the Genetics Society of America

An Integrated Map of *Arabidopsis thaliana* for Functional Analysis of Its Genome Sequence

Yueh-Long Chang,*[,1] Quanzhou Tao,*[,1] Chantel Scheuring,* Kejiao Ding,* Khalid Meksem† and Hong-Bin Zhang*

**Department of Soil and Crop Sciences and Crop Biotechnology Center, Texas A&M University, College Station, Texas 77843-2123 and †Department of Plant Soil and General Agriculture, Southern Illinois University, Carbondale, Illinois 62901-4415*

Manuscript received July 11, 2001
Accepted for publication August 13, 2001

ABSTRACT

The genome of the model plant species *Arabidopsis thaliana* has recently been sequenced. To accelerate its current genome research, we developed a whole-genome, BAC/BIBAC-based, integrated physical, genetic, and sequence map of the *A. thaliana* ecotype Columbia. This new map was constructed from the clones of a new plant-transformation-competent BIBAC library and is integrated with the existing sequence map. The clones were restriction fingerprinted by DNA sequencing gel-based electrophoresis, assembled into contigs, and anchored to an existing genetic map. The map consists of 194 BAC/BIBAC contigs, spanning 126 Mb of the 130-Mb Arabidopsis genome. A total of 120 contigs, spanning 114 Mb, were anchored to the chromosomes of Arabidopsis. Accuracy of the integrated map was verified using the existing physical and sequence maps and numerous DNA markers. Integration of the new map with the sequence map has enabled gap closure of the sequence map and will facilitate functional analysis of the genome sequence. The method used here has been demonstrated to be sufficient for whole-genome physical mapping from large-insert random bacterial clones and thus is applicable to rapid development of whole-genome physical maps for other species.

ARABIDOPSIS thaliana is a model system for genomic studies of plant species (Meinke *et al.* 1998). To facilitate genome research of the species, chromosome- and genome-wide physical maps were developed from large-insert yeast artificial chromosome or bacterial artificial chromosome (BAC) libraries (Hwang *et al.* 1991; Schmidt *et al.* 1995; Zachgo *et al.* 1996; Camilleri *et al.* 1998; Marra *et al.* 1999; Mozo *et al.* 1999). Using these maps as frameworks, the genome of *A. thaliana* was sequenced (Arabidopsis Genome Initiative 2000). However, a number of gaps still exist in the sequence map, especially in the heterochromatic regions surrounding the centromeres of its chromosomes. These gaps are intractable to closure by the conventional chromosome walking approach using the ends of BACs adjacent to the gaps as probes because of repetitive sequences. This piecemeal approach is also time-consuming. Therefore, to rapidly close most, if not all, of the gaps in the sequence map, it is necessary to construct a whole-genome physical map from a new DNA library that complements the Texas A&M University (TAMU) and the Institut für Genbiologische Forschung (IGF) BAC libraries of the sequence map and to integrate the new map with the existing sequence map.

Sequence analysis has indicated that the genome of *A. thaliana* contains ~25,498 genes. However, the functions of >90% of the predicted genes remain to be characterized experimentally (Arabidopsis Genome Initiative 2000). Experimental determination of the functions of these genes and related sequences has been targeted as a goal for the coming decade (Somerville and Dangl 2000). To this end, several methods have been developed, including T-DNA-based (Azpiroz-Leehan and Feldmann 1997; Sussman *et al.* 2000) or transposon-based (Martienssen 1998) gene tagging, DNA microarray or gene chip analysis (Schena *et al.* 1995; Desprez *et al.* 1998; Ruan *et al.* 1998), and genetic transformation (Feldmann and Marks 1987; Chang *et al.* 1994; Liu *et al.* 1999). Because transformation of *A. thaliana* via Agrobacterium is efficient and can be accomplished without tissue culture procedures (Feldmann and Marks 1987; Koncz *et al.* 1989; Bechtold *et al.* 1993), this method facilitates functional analysis of the genome sequence. Therefore, a whole-genome, binary, clone-based map that is integrated with the existing sequence map will be significant for accelerated experimental determination of the functions of every segment of the genome sequence. However, none of the Arabidopsis physical and sequence maps developed to date contains clones that can be directly transformed in plants. The TAMU and IGF BAC clones of the existing physical and sequence maps were cloned in general DNA cloning BAC vectors (Choi *et al.* 1995; Mozo *et al.* 1998), which are incompetent for direct transforma-

Corresponding author: Hong-Bin Zhang, Department of Soil and Crop Sciences and Crop Biotechnology Center, 2123 TAMUS, Texas A&M University, College Station, TX 77843-2123.
E-mail: hbz7049@pop.tamu.edu

[1]These authors contributed equally to this study.

Downloaded from https://academic.oup.com/genetics/article/159/3/1231/6049583 by guest on 19 April 2024

tion in plants via *A. tumefaciens*. For functional analysis of the genome sequence by genetic transformation, these clones must be subcloned into a plant-transformation-competent binary vector. However, the process of subcloning is often tedious. Furthermore, the goals of Arabidopsis genome research are to identify every gene and determine the function(s) of every gene of this model species. Although individual clones for functional analysis by genetic transformation could be isolated from a plant-transformation-competent binary bacterial artificial chromosome (BIBAC) library using the corresponding sequences as probes, this process would be inefficient for isolation of a large number of BIBAC clones for functional analysis of different segments of the genome sequence. Therefore, it is desirable to develop a physical map from a large-insert BIBAC library that is competent for plant transformation and to integrate it with the existing sequence map.

In this study, we developed a whole-genome integrated physical and genetic map of the *A. thaliana* ecotype Columbia from a new plant-transformation-competent BIBAC library and integrated it with the existing sequence map of the species. The integration of the new map with the sequence map will significantly accelerate genome research of the model species in many aspects. The complementarity of the new BIBAC library to the source libraries of the sequence map and the competency of the BIBACs for plant transformation will facilitate gap closure of the sequence map and large-scale functional analysis of the genome sequence. Furthermore, because the new map was constructed using a DNA sequencing gel-based fingerprinting method (Tao *et al.* 2001) that differs from those used for construction of the existing physical maps of Arabidopsis (Marra *et al.* 1999; Mozo *et al.* 1999), it may provide a tool to further verify the existing physical and sequence maps in which errors have been recently reported (Stupar *et al.* 2001). This study has also demonstrated that the DNA sequencing gel-based fingerprinting method is powerful for rapid development of whole-genome physical maps from large-insert random bacterial clones.

MATERIALS AND METHODS

BAC and BIBAC libraries: A new plant-transformation-competent binary library (the clones are hereafter referred as to BIBACs; Y.-L. Chang, K. Meksem, H.-W. Chuang, C. Scheuring and H.-B. Zhang, unpublished data) and the TAMU (Choi *et al.* 1995) and IGF (Mozo *et al.* 1998) BAC libraries of the *A. thaliana* ecotype Columbia were used to develop the integrated physical map. The average insert sizes of the BIBAC library and the TAMU and IGF BAC libraries are 110, 100, and 100 kb, respectively. These BAC and BIBAC libraries are publicly available at the GENE*finder* Genomic Resources (formerly the Texas A&M BAC Center; http://hbz.tamu.edu).

The TAMU and IGF BAC libraries are the source libraries of the existing physical and sequence maps of *A. thaliana* (Marra *et al.* 1999; Mozo *et al.* 1999; Arabidopsis Genome Initiative 2000). The BIBAC library was cloned in the bacterial P1-based binary vector pCLD04541, which was designed for Agrobacterium-mediated transformation in plants (Jones *et al.* 1992; Tao and Zhang 1998). The ability of BIBACs to transform plants via Agrobacterium (Jones *et al.* 1992; Bechtold *et al.* 1993; Bent *et al.* 1994; Wu *et al.* 2000; Y.-L. Chang, K. Meksem, H.-W. Chuang, C. Scheuring and H.-B. Zhang, unpublished data) will facilitate functional analysis of the genome sequence by genetic transformation. Furthermore, the BIBAC library was constructed with a restriction enzyme (*Bam*HI) differing from those used for the TAMU (*Hin*dIII; Choi *et al.* 1995) and IGF (*Eco*RI; Mozo *et al.* 1998) BAC libraries. The *Bam*HI sites are G/C-rich, whereas the *Hin*dIII and *Eco*RI sites are A/T rich. These differences were expected to allow cloning of the regions that are not represented in the existing BAC libraries and thus to close the gaps in the existing Arabidopsis sequence map.

Fingerprinting and contig assembly: BAC and BIBAC DNA were isolated and fingerprinted according to Tao *et al.* (2001) with modifications in which [^{33}P]dATP was used to label the digested BAC DNA fragments and 3.5% (w/v) denaturing DNA sequencing gels were used to fractionate the DNA fragments. The fingerprints were scanned into image files using a UMAX Mirage D-16L scanner and edited using Image 4.0 of the FingerPrinted Contig (FPC) package (Soderlund *et al.* 1997). The resolvable bands of each fingerprint ranged from 58 to 2225 nucleotides in size. Considering the lower resolution (probably >1 nucleotide) of the bands in the higher-molecular-weight portion at the top of each autoradiograph, only the fragments ranging from 58 to 773 bases were used for contig assembly, on average, 36 bands per fingerprint. The bands derived from the BAC vectors were deleted from the data files, while the bands of the BIBAC vector pCLD04541 were not present in the fingerprint range (*e.g.*, see Figure 1). The clones that had no inserts or produced four or fewer bands were excluded during fingerprint image editing because the number of bands per fingerprint was insufficient to be included for contig assembly. Consequently, 9389 of the clones, being equivalent to $7.2 \times$ Arabidopsis haploid genomes, were used to assemble the physical map contigs. To construct the physical map, we first assembled automatic contigs using the FPC3.8 program (Soderlund *et al.* 1997) and then merged the automatic contigs or singletons with the automatic contigs using a lower comparison stringency. The automatic contigs were assembled at tolerance 2, cutoff 10^{-12}, DiffBury 0.1, and MinBands 8. The mergence of contig-contig and singleton-contig was conducted at tolerance 2, cutoff 10^{-7}, DiffBury 0.1, and MinBands 8.

Library screening: The source clones of the physical map were double spotted on Hybond N+ membrane in a format of 3×3 by using the Biomek 2000 Robotic Workstation (Beckman, Fullerton, CA), and the high-density colony filters were prepared according to Zhang *et al.* (1996). To determine the genome origin of the BAC/BIBAC contigs, three chloroplast DNA probes, *adhA*, *psbA*, and *rbcL*, which are ~50 kb apart on the chloroplast genome, were used to screen the source BAC and BIBAC libraries, and the IGF BACs derived from mitochondrial DNA were used to search the database of the new physical map. To verify the accuracy of the contigs and anchor them to Arabidopsis genetic maps, 77 restriction fragment length polymorphism (RFLP) markers were selected from the Arabidopsis genetic map (Liu *et al.* 1996; see Table 2), obtained from the Arabidopsis Biological Resource Center (Ohio State University), and used as probes to hybridize the high-density BAC and BIBAC filters. The probes were prepared using the PCR Dig-Probe synthesis kit as described by its manufacturer (Roche Molecular Biochemicals, Indianapolis). In addition, nine Arabidopsis cDNAs were selected from the Arabidopsis expressed sequence tagged (EST) set obtained from the Arabidopsis Biological Resource Center (Ohio State University) and positioned to the integrated map by colony hybridization.

Downloaded from https://academic.oup.com/genetics/article/159/3/1231/6049583 by guest on 19 April 2024

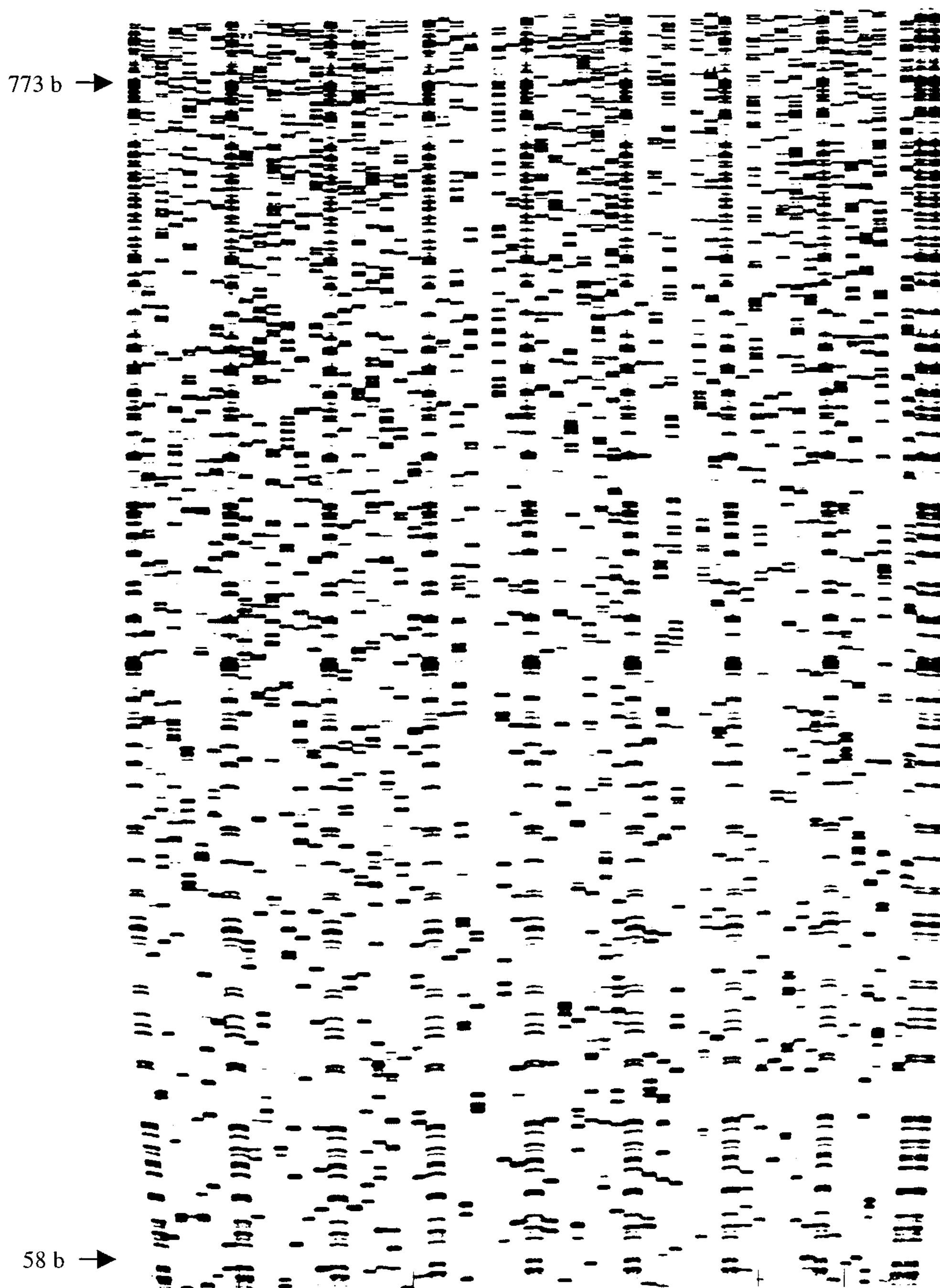


Figure 1.—Autoradiograph of the Arabidopsis BIBAC fingerprints used for contig assembly of the new BAC/BIBAC-based map. DNA was isolated, double digested with *Hin*dIII/*Hae*III, end labeled with [^{33}P]dATP, electrophoresed on a 3.5% (w/v) denaturing DNA sequencing gel, and exposed to an X-ray film. λDNA/*Sau* 3AI markers labeled with [^{33}P]dATP were used in the first lane and every seventh lane thereafter. Note that the bands derived from the cloning vector of the new BIBAC library, pCLD04541 (Jones *et al.* 1992; Tao and Zhang 1998), were not present in the fingerprint range, which facilitated the map contig assembly. However, one or two bands derived from the cloning vectors of TAMU and IGF BACs were present in the range and were deleted manually from the fingerprint image files. The fragments ranging from 58 to 773 bases were used for contig assembly.

Downloaded from https://academic.oup.com/genetics/article/159/3/1231/6049583 by guest on 19 April 2024

RESULTS AND DISCUSSION

Fingerprinting the BAC and BIBAC clones: To construct the whole-genome physical map that is integrated with the existing sequence map (Arabidopsis Genome Initiative 2000), we selected 1536 IGF BACs (4 384-well plates), 3072 TAMU BACs (8 384-well plates), and 6144 BIBACs (16 384-well plates) from their libraries. The 10,752 BAC and BIBAC clones were fingerprinted on 224 autoradiographs using the DNA sequencing gel-based restriction fingerprinting method (Tao *et al.* 2001). Figure 1 shows an autoradiograph of the BAC and BIBAC fingerprints.

Assembling the BAC/BIBAC map: We scanned the clone fingerprints into image files and edited with the Image program of the FPC package (Sulston *et al.* 1988; Soderlund *et al.* 1997). During editing (see materials and methods), 1363 clones were deleted from the data files because they had no inserts or produced four or fewer bands in their fingerprints, which were insufficient to be included in contig assembly. As a result, the data from 9389 clones with an average of 36 bands per clone were used to assemble contigs with the FPC program (Soderlund *et al.* 1997). Of the 9389 clones (7.2 × genome equivalents), the clones equivalent to 2.1 × haploid genomes were from the TAMU BAC library, the clones equivalent to 1.0 × haploid genomes

TABLE 1

BAC/BIBAC contigs of *A. thaliana*

Contig	Clone	kb	Contig	Clone	kb	Contig	Clone	kb	Contig	Clone	kb
ctg1	161	162	ctg51	5	177	ctg1025	63	990	ctg4006	10	159
ctg2	108	564	ctg52	5	120	ctg1026	148	2,220	ctg4007	20	324
ctg3	87	939	ctg53	5	48	ctg1027	119	1,665	ctg4008	3	111
ctg4	32	318	ctg54	5	96	ctg2001	91	1,371	ctg4009	45	747
ctg5	27	369	ctg55	5	171	ctg2002	39	579	ctg4010	12	174
ctg6	24	270	ctg56	5	159	ctg2003	42	429	ctg4011	230	3,255
ctg7	22	201	ctg57	4	69	ctg2004	37	441	ctg4012	31	522
ctg8	22	219	ctg58	4	99	ctg2005	89	1,029	ctg4013	37	555
ctg9	22	252	ctg59	4	135	ctg2006	13	237	ctg4014	148	2,202
ctg10	20	297	ctg60	4	99	ctg2007	55	747	ctg4015	16	309
ctg11	19	390	ctg61	4	75	ctg2008	29	357	ctg4016	150	2,325
ctg12	16	219	ctg62	4	150	ctg2009	12	246	ctg4017	128	1,764
ctg13	14	186	ctg63	3	132	ctg2010	200	2,292	ctg5001	37	597
ctg14	13	252	ctg64	3	111	ctg2011	45	585	ctg5002	181	2,412
ctg15	13	240	ctg65	3	120	ctg2012	72	924	ctg5003	15	297
ctg16	13	177	ctg66	3	57	ctg2013	26	321	ctg5004	41	654
ctg17	13	99	ctg67	3	90	ctg2014	51	639	ctg5005	99	1,452
ctg18	12	171	ctg68	3	108	ctg2015	66	852	ctg5006	16	315
ctg19	12	141	ctg69	2	114	ctg2016	13	180	ctg5007	48	828
ctg20	12	222	ctg70	2	126	ctg2017	21	348	ctg5008	40	630
ctg21	11	81	ctg71	2	84	ctg2018	63	972	ctg5009	62	582
ctg22	11	243	ctg72	2	93	ctg2019	97	1,086	ctg5010	74	1,056
ctg23	10	108	ctg73	2	75	ctg2020	294	4,581	ctg5011	32	402
ctg24	10	90	ctg74	2	102	ctg3001	31	714	ctg5012	9	225
ctg25	9	96	ctg75	2	84	ctg3002	55	954	ctg5013	15	303
ctg26	9	138	ctg76	2	132	ctg3003	25	528	ctg5014	25	498
ctg27	9	108	ctg1001	205	3,009	ctg3004	395	4,899	ctg5015	27	381
ctg28	9	150	ctg1002	99	1,338	ctg3005	7	129	ctg5016	31	459
ctg29	8	51	ctg1003	123	1,800	ctg3006	155	2,112	ctg5017	85	951
ctg30	8	105	ctg1004	29	360	ctg3007	97	1,626	ctg5018	67	711
ctg31	8	141	ctg1005	50	774	ctg3008	49	642	ctg5019	45	657
ctg32	8	192	ctg1006	66	675	ctg3009	11	294	ctg5020	24	327
ctg33	7	57	ctg1007	148	2,166	ctg3010	77	810	ctg5021	9	201
ctg34	7	156	ctg1008	10	159	ctg3011	39	345	ctg5022	8	183
ctg35	7	72	ctg1009	48	573	ctg3012	112	1,308	ctg5023	63	870
ctg36	7	81	ctg1010	111	1,668	ctg3013	24	354	ctg5024	8	159
ctg37	7	105	ctg1011	75	990	ctg3014	52	651	ctg5025	157	2,157
ctg38	6	75	ctg1012	55	882	ctg3015	97	1,152	ctg5026	59	900
ctg39	6	135	ctg1013	71	864	ctg3016	25	342	ctg5027	132	1,977
ctg40	6	156	ctg1014	114	1,335	ctg3017	33	510	ctg5028	72	1,131
ctg41	6	81	ctg1015	11	144	ctg3018	79	1,164	ctg5029	39	612
ctg42	6	81	ctg1016	91	1,071	ctg3019	140	2,118	ctg5030	104	1,362
ctg43	6	75	ctg1017	18	234	ctg3020	88	1,086	ctg5031	29	483
ctg44	6	129	ctg1018	116	1,668	ctg3021	9	213	ctg5032	44	642
ctg45	6	93	ctg1019	30	375	ctg3022	26	492	ctg5033	16	300
ctg46	6	93	ctg1020	130	1,749	ctg4001	73	1,116	ctg5034	22	489
ctg47	6	75	ctg1021	25	414	ctg4002	80	942			
ctg48	5	120	ctg1022	212	3,243	ctg4003	38	570			
ctg49	5	126	ctg1023	10	237	ctg4004	29	507			
ctg50	5	63	ctg1024	84	909	ctg4005	113	1,401			

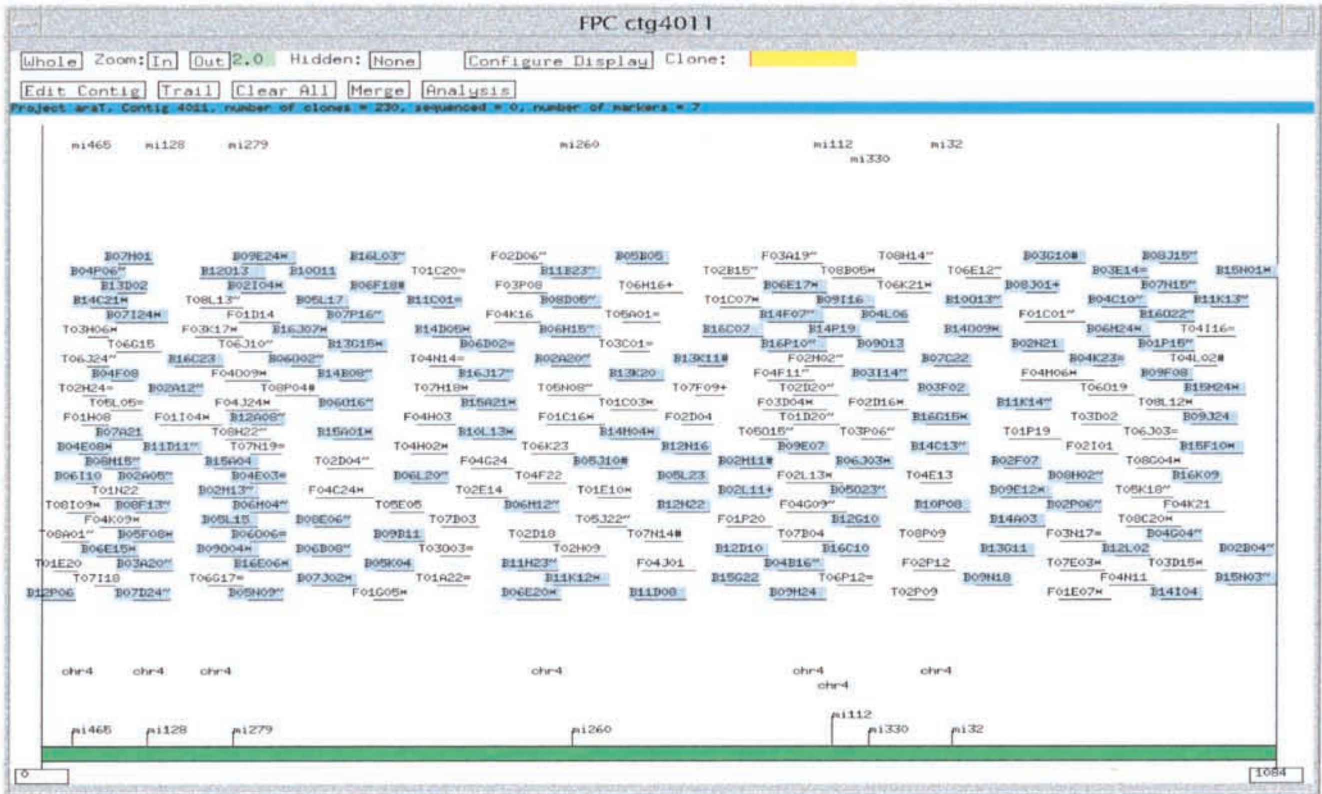
Contig 1 (ctg1) was derived from chloroplast DNA, contig 2 (ctg2) from the mitochondrial DNA, and the remaining 194 contigs from nuclear DNA. Of the 194 nuclear DNA clone contigs, ctg3 through ctg76 had not been anchored to chromosomes and the 120 remaining contigs were anchored to the chromosomes of their origin. The anchored contigs are numbered using four-digit numbers, with the first one identifying the chromosome to which the contig was anchored and the last two specifying the contig.

Downloaded from https://academic.oup.com/genetics/article/159/3/1231/6049583 by guest on 19 April 2024

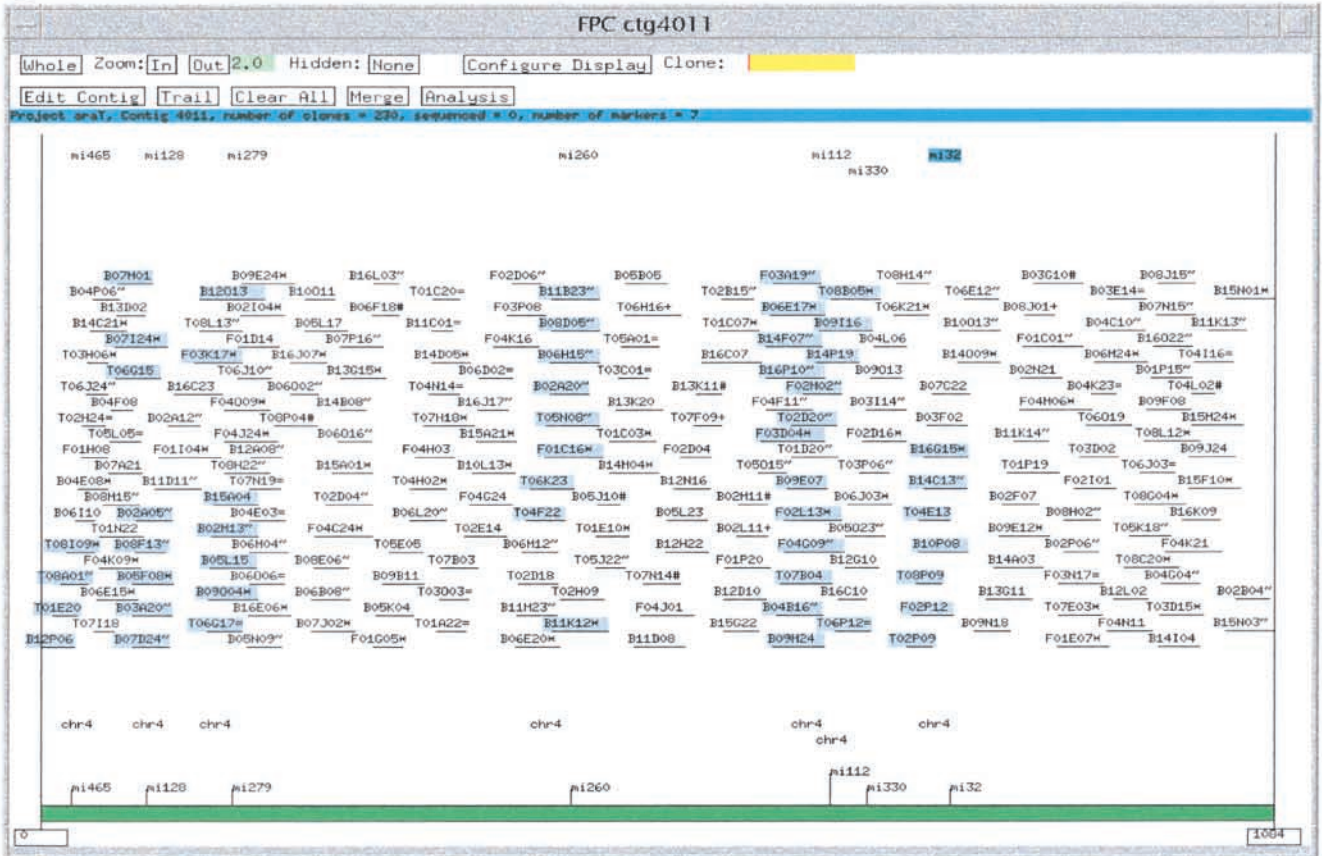
Figure 2.—Comparison of a BAC/BIBAC contig (ctg1026) of the integrated map developed in this study (A) *vs.* the corresponding BAC contig of the AGI sequence map (B; Arabidopsis Genome Initiative 2000) anchored to chromosome 1. The clones highlighted in green are the positive clones of the DNA marker mi425. The IGF and TAMU BAC clones shared by these two contigs are highlighted in blue. Note that the content and order of the clones in the two contigs are identical, indicating that the contigs assembled from the fingerprints generated by the DNA sequencing gel-based restriction fingerprinting method are accurate.

Downloaded from https://academic.oup.com/genetics/article/159/3/1231/6049583 by guest on 19 April 2024

A



B



Downloaded from https://academic.oup.com/genetics/article/159/3/1231/6049583 by guest on 19 April 2024

were from the IGF BAC library, and the clones equivalent to 4.1 × haploid genomes were from the new BIBAC library. A total of 196 contigs were assembled, each consisting of at least 2 clones (Table 1 and Figure 4), whereas 279 clones remained as singletons due to the insufficient numbers of bands (<10 bands) in their fingerprints.

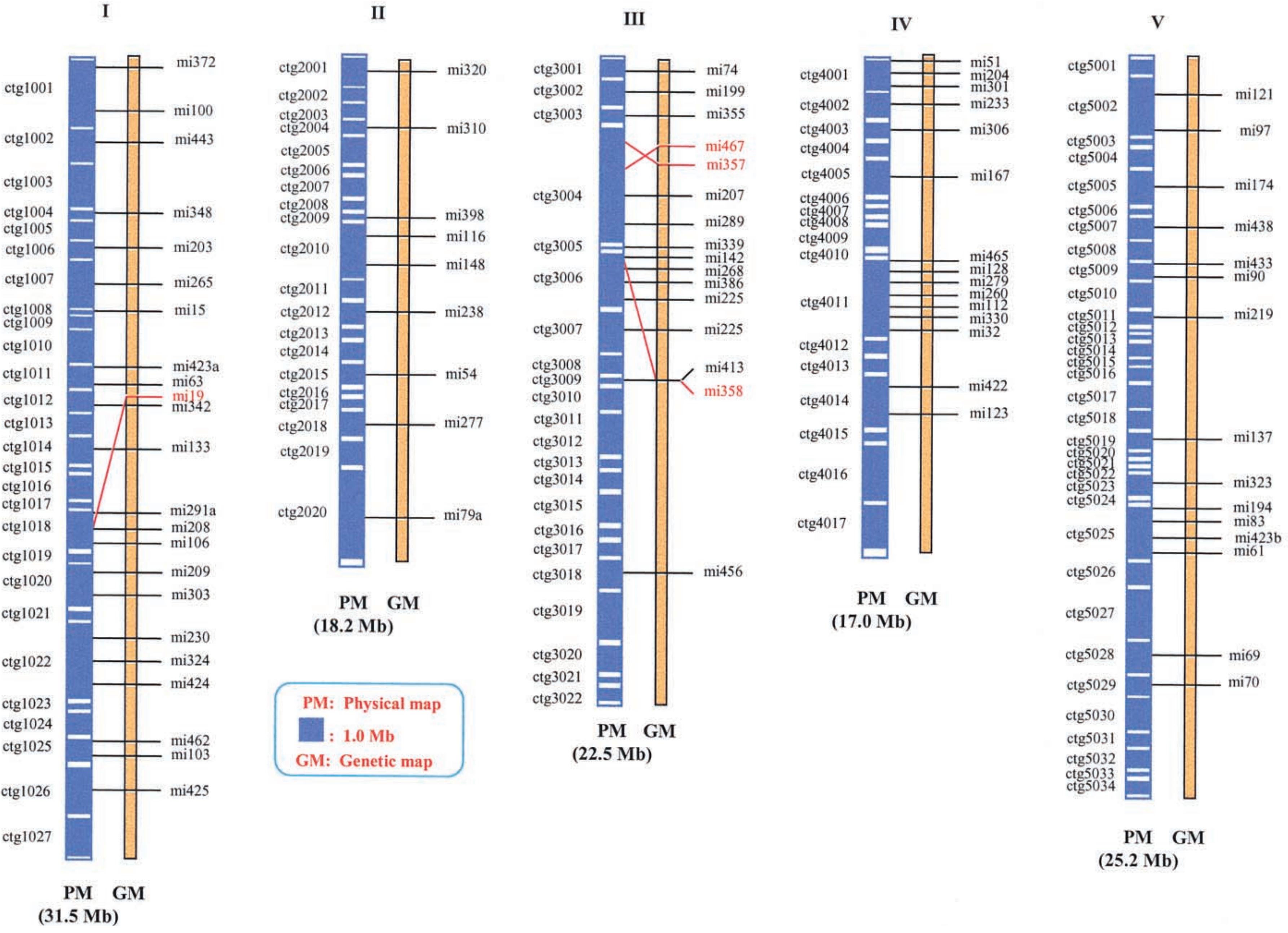
Origin of the 196 contigs was investigated by colony hybridization using chloroplast DNA as probes and by analyzing the mitochondrial DNA-derived clones of the IGF library (Mozo *et al.* 1998) against the BAC/BIBAC contigs. Contig 1 (ctg1 in Table 1) was shown to derive from chloroplast DNA. It consists of 161 clones and was estimated to be 162 kb in length. The size of the contig is close to the 153-kb chloroplast genome size (Palmer *et al.* 1994). Contig 2 (ctg2 in Table 1) was shown to derive from the mitochondrial DNA. It consists of 108 clones and was estimated to be 564 kb in length. The length of the contig is much greater than the size of either the cytoplasmic mitochondrial genome (372 kb; Klein *et al.* 1994) or the mtDNA inserted into the nuclear genome revealed by sequencing (270 kb; Lin *et al.* 1999), but close to the size of the mtDNA inserted into the nuclear genome revealed by fiber-fluorescence *in situ* hybridization (618 ± 42 kb; Stupar *et al.* 2001). Stupar *et al.* (2001) demonstrated that the size of the mtDNA inserted into the nuclear genome was underestimated in the sequence map (Lin *et al.* 1999) due to errors of the physical and sequence map assembly. This result indicated that it is necessary to further verify the existing physical and sequence maps using an approach differing from those used in the development of the existing maps. The 194 remaining contigs were derived from nuclear DNA (Table 1). The FPC program showed that these 194 contigs consisted of 42,119 unique bands, each band representing a 3.0-kb fragment. Therefore, the nuclear DNA clone contigs were estimated to span >126 Mb in length.

Verifying the map: To test the accuracy of the new map, we compared the BAC/BIBAC contigs constructed in this study with the BAC contigs of the existing physical (Mozo *et al.* 1999) and sequence (Arabidopsis Genome Initiative 2000; http://www.arabidopsis.org) maps using the TAMU and IGF BACs shared between the contigs. The comparison showed that ∼95% of the new BAC/BIBAC contigs were consistent with the contigs of the existing physical and sequence maps in both clone content and order (Figure 2), but 5% of the contigs were different. To further verify the accuracy of the new BAC/BIBAC contigs, we screened them with 77 DNA markers selected from the genetic map (Liu *et al.* 1996) and nine cDNA clones selected from the Arabidopsis EST set. The results showed that all positive clones of every marker or cDNA probe used were located at a restricted fragment of a single contig (see Table 2 and Figure 3). This indicates that the contigs assembled in this study are accurate. The cause of the 5% difference between the contigs of the new map and the existing physical and sequence maps needs to be further investigated as errors were recently reported in the existing physical and sequence maps (Stupar *et al.* 2001). We have also investigated the orders of markers in the contigs that contained three or more of the DNA markers. The result showed that they were consistent with those in the genetic map (see Figure 4), except for four DNA markers (mi19, mi467, mi357, and mi358). To explain this inconsistency, we further investigated the orders of the four DNA markers in the existing physical and sequence maps (Mozo *et al.* 1999; http://www.arabidopsis.org). The investigation showed that the orders of the four DNA markers in the new physical map were consistent with their orders in the existing physical and sequence maps, indicating that these DNA markers might be mismapped in the genetic map. All above results strongly indicate that the BAC/BIBAC contigs constructed in this study are accurate and that the fingerprinting method used is reliable for whole-genome physical mapping from random BACs and/or BIBACs.

Anchoring the physical map contigs to the genetic map: To anchor the BAC/BIBAC contigs to a public genetic map of *A. thaliana*, we used the above screening results of the contig BACs and BIBACs with the 77 mapped DNA markers and the database of the TAMU and IGF BAC-based maps of the IGF (Mozo *et al.* 1999) and the Arabidopsis Genome Initiative (AGI; Arabidopsis Genome Initiative 2000). Of the 194 nuclear DNA clone contigs, 120 were anchored to the five chromosomes of Arabidopsis

Downloaded from https://academic.oup.com/genetics/article/159/3/1231/6049583 by guest on 19 April 2024

Figure 3.—BAC/BIBAC contig of the Arabidopsis BAC/BIBAC-based, integrated genetic, physical, and sequence map (ctg4011 in Table 1). (A) Integration of the new map with the existing physical and sequence maps of *A. thaliana* (http://www.arabidopsis.org) and distribution of TAMU BACs, IGF BACs, and BIBACs in the map. The contig (ctg4011) contains 230 clones and has a length of 1084 unique bands, being equivalent to 3255 kb (see Table 1). The clones prefixed with letter T were from the TAMU BAC library, the clones prefixed with F from the IGF BAC library, and the clones prefixed with B from the new BIBAC library (see http://hbz.tamu.edu). Each clone is specified with a library letter (T, F, or B), plate number (two-digit number), row letter, and column number (two-digit number). For instance, the clone B07M01 was from the new BIBAC library, plate 7, row M, and column 1. (B) Accuracy verification and anchoring of the BAC/BIBAC contigs to the Arabidopsis chromosomes. The contig 4011 was anchored to chromosome 4 by seven linked-DNA markers (mi465, mi128, mi279, mi260, mi112, mi330, and mi32) mapped to linkage group 4 of the Arabidopsis genetic map (Liu *et al.* 1996). The clones highlighted in blue indicate the positive clones of each marker below and above them. Note that all positive clones of each marker were located to a single location of the contig and the order of the seven markers in the contig is completely the same as that in the genetic map (see Figure 4). Asterisk (*) indicates a parent clone that covers one or more clones.



Downloaded from https://academic.oup.com/genetics/article/159/3/1231/6049583 by guest on 19 April 2024

(Figure 4 and Table 2). These 120 contigs collectively span 114 Mb in length. Of them, 27 were anchored to chromosome 1, spanning 31.5 Mb in length; 20 to chromosome 2, spanning 18.2 Mb; 22 to chromosome 3, spanning 22.5 Mb; 17 to chromosome 4, spanning 17.0 Mb; and 34 to chromosome 5, spanning 25.2 Mb (see Table 1 and Figure 4).

Potential applications of the new integrated map for accelerated genome research of Arabidopsis: This new integrated map has provided a platform for accelerated genome research of Arabidopsis in many aspects. To test the utility of the integrated map for gap closure in the sequence map, we attempted to close four gaps in the sequence map using the new map. One of the gaps was between clones T2P3 and F2G19 in the sequence map of chromosome 1, and three gaps were between ctg714-ctg719-ctg731-ctg11 in the sequence map of Arabidopsis chromosome 3 (http://genome.wustl.edu/gsc/arab/arabidopsis.html; communicated with Dr. Christopher Town, The Institute for Genome Research). The TAMU and IGF BACs in the contigs were used to search the new BAC/BIBAC contigs. One contig (ctg1024) was identified from the new map to span the gap between T2P3 and F2G19 by two BIBACs (B05G22 and B09C04). The two BIBACs have been used to close the gap of the existing sequence map (C. Town, personal communication). Similarly, we searched the contigs of the new map that span the three gaps between ctg714-ctg719-ctg731-ctg11 in chromosome 3. As a result, from the new map we identified three contigs, ctg3010 (810 kb), ctg3011 (345 kb), and ctg3012 (1308 kb). To further determine whether ctg3011 and ctg3012 overlap, we further analyzed the fingerprints at the contig ends under less comparison stringency and screened the BACs and BIBACs of the new map using one IGF BAC clone (F03O21) at a ctg3012 end as a probe. As a result, eight positive BIBAC clones at one of the ctg3011 ends were identified. The hybridization and fingerprint analysis results suggested that the ctg3011 and ctg3012 overlapped and thus were merged. The merged contig (1618 kb) seems to span the gaps between ctg719, ctg731, and ctg11. Therefore, of the four gaps in the AGI sequence map, three were likely to be closed using the new map. Whether the gap between ctg714 and ctg719 in the sequence map can be closed by ctg3010 and ctg3011 remains to be determined. Similarly, other gaps in the existing sequence map could also be closed using the new map.

Experimental determination of the function of the genes and related sequences predicted by the genome sequence analysis of *A. thaliana* will be a significant challenge. Since *A. thaliana* can be readily transformed via Agrobacterium (Feldmann and Marks 1987; Koncz *et al.* 1989; Bechtold *et al.* 1993), genetic transformation and subsequently transgenic plant analysis will provide an alternative tool for experimental determination of the functions of the predicted genes and related sequences in the genome sequence map. The binary vector pCLD04541 used for the new BIBAC library was designed for plant transformation via Agrobacterium (Jones *et al.* 1992). It has been widely used in Arabidopsis genetic complementation studies (*e.g.*, Bent *et al.* 1994; http://www.jic.bbsrc.ac.uk/staff/ian-bancroft/vectorspage/htm). Although further investigation will be needed to transform large DNA fragments using this vector in Arabidopsis and other plant species, it has been shown recently that a 135-kb clone of Brassica DNA in pCLD04541 was stable in Agrobacterium (Wu *et al.* 2000) and transformed into Brassica (Y.-Z. Wu and Y.-P. Zhang, personal communication). We transformed a 120-kb clone of soybean DNA in the vector into *A. thaliana* by the vacuum-infiltration method (Bechtold *et al.* 1993; Y.-L. Chang, K. Meksem, H.-W. Chuang, C. Scheuring and H.-B. Zhang, unpublished data). Furthermore, using the same vacuum-infiltration method and a similar vector, Liu *et al.* (1999) successfully transformed a DNA fragment of 80 kb into *A. thaliana* and showed that the transformation efficiency was not affected substantially by the sizes of introduced T-DNA. By transforming two clones of 75 and 80 kb carrying the *SGR1* locus into Arabidopsis, Liu *et al.* (1999) were able to complement the *sgr1* mutant and thus identify the functional sequence of the *SGR1* gene. Transformation of DNA fragments of 150 kb via Agrobacterium was also documented in tobacco (Hamilton *et al.* 1996) and tomato (Hamilton *et al.* 1999). These studies have indicated that high-molecular-weight DNA

Downloaded from https://academic.oup.com/genetics/article/159/3/1231/6049583 by guest on 19 April 2024

Figure 4.—The BAC/BIBAC-based, integrated genetic, physical, and sequence map of *A. thaliana* ecotype Columbia. Since the clones of the map equivalent to 3.1 × Arabidopsis haploid genomes were from the TAMU and IGF BAC libraries used in the sequence map (http://www.arabidopsis.org), the new map is integrated with the sequence map (see Figure 3). The contigs containing DNA markers were anchored to the chromosomes of their origin using the screening results of the contig BACs and BIBACs with the DNA markers. The contigs not containing DNA markers were anchored to the chromosomes of their origin by using the database of the TAMU and IGF BAC-based maps of the IGF (Mozo *et al.* 1999) and the AGI (Arabidopsis Genome Initiative 2000). The collective length of the contigs anchored to each chromosome is given below its physical map (PM) in mega base pairs (Mb) in parentheses. The sizes of gaps between neighboring contigs were not determined in this study; most of the neighboring contigs may overlap even though the extent of their overlaps could not be detected by the FPC program under the conditions used in this study. Although the orders of the four DNA markers mi19, mi467, mi357, and mi358, highlighted in red on the BAC/BIBAC-based physical map, are different from those on the genetic map (GM; Liu *et al.* 1996), they are consistent with their orders on the physical map of the IGF and the AGI (http://www.arabidopsis.org).

TABLE 2

Positive BAC and BIBAC clones of the integrated map selected with RFLP marker and nine cDNA probes

Probe	Clone	Contig[a]
mi15	B08O03, B09B02, B14O12, F03C06, F04M13, T01H23, T05M04	1008
mi19	B04P08, T04P06	1018
mi32	B10P08, B14C13, B16G15, F02P12, T02P09, T04E13, T08P09	4011
mi51	B13B12, T07N12	4001
mi54	B05C06, B08A08, B14E08, B16L19, F03G21	2015
mi61	B01P12, B02C13, B03H10, B05H09, B08A17, B14O01, B14E21, F02F17, F04O21, T02A13, T04P10	5025
mi63	B03L12, B04M03, B09C18, B15D09, T02E10, T08K04	1011
mi69	B01D23, B12A21, B12P09, B13D03, B14G05, F01L09, T06O12	5028
mi70	B08I09, B09G04, B11A10, B15N07, T04K24, T05J16, T08F08	5029
mi74	B01A10, B06O10, B06L08, B07D06, F01C09, F01J06, F02P08, T01E08	3001
mi79a	B06H12, B14O14, F03G08, T06G23	2020
mi83	B01G03, B13N03, T04H07, T07N23	5025
mi90	B05P13, B11B07, T04L22	5009
mi97	B06O24, B15A03, B16L13, F01G19, F04I15, T04L12	5002
mi100	B10O23, B10P23, F01A17, F03E19, T08D18	1001
mi103	B01M16, B02L04, B15I24, B16D23, F02N11, T04G12	1025
mi106a	B02E01, B04H04, B05E22, B06F14, B08F21, B12L16, F02J10, F03F06, T05N18, T06G11	1018
mi106b	B06D07, F04J19	1001
mi112	B04B16, B06E17, B09E07, B09M24, B14F07, B16P10, F02M02, F02L13, F03A19, F03D04, F04G09, T02D20, T07B04	4011
mi116	B06J02, B06P14, B12F02, B12H05, B12N23, F01P15, F02P06, F03C16	2010
mi121	B04K04, B08K11, B10K15, B10F02, F01H15, T01E03, T02J20	5002
mi123	F02L16, F03P01, F04G17	4014
mi128	B02A05, B03A20, B05F08, B07I24, B07D24, B08F13, T06G15	4011
mi133	B02J11, B16N18, F02G10, T02A12, T04M23, T05L02, T06M18, T06F11, T07L04, T08D08	1014
mi137	B04I14, B06C18, B09A17, T08D21	5019
mi142	B04E22, B12M12, B12A16, B12H10, F03B08, F03P10, F03J21, T02C08, T03C23, T04K14, T05G03, T06K16	3006
mi148	B04E16, B08E08, B16G12, T02C02, T02G12, T05D24	2010
mi167	B03L03, B09O05, B10E11, B13P09, B16D13, T01H20, T04P21	4005
mi174	B14M01, F01A23, T04J18	5005
mi194	B01M09, B02E20, B03L21, B08E21, B11D20, B13E17, B13D08, B14D19, B14D20, F04J06, T05J02	5025
mi199	B07I02, B08N19, B09M17, B10M17, T04P02, T05N20, F04J14, T08J15	3002
mi203	B02O04, B04L19, B06G01, B09J17, B10F21, B14M17, F03D08, F04C20, F04E19, F04N04, T06A18, T07J22, T08O20	1006
mi204	B07D03, B08A18, B12K16, F03C15, T08M24	4001
mi207	B05L07, B08G16, B08F05, B14L03, B15D14, T01A09, T03I04, T03O05	3004
mi208	B02N17, B03P13, B04K24, B08B01, T06G18	1018
mi209	B07F14, B12F05, B14E04, B16O23, T05A14	1020
mi219	B02C02, B03D21, B08B10, B11C06, B16P06	5011
mi225	B04C03, B13A04, F02A23, F02P22, F03M05	3007
mi230	B08G24, B09G21, B11L05, B12L09, B16O01, F02M24, T01N23, T02H02	1022
mi233	B02E06, T03B24, B14I12, B16M08, F04F14, T07P21	4002
mi238	B07P02, B08D20, B09H01, B15E01, B16M17, F02H05, F02P11, T07N05, T08A08	2012
mi260	B02A20, B06H15, B08D05, B11K12, B11B23, F01C16, T04F22, T05N08, T06K23	4011
mi265	B14J09, F01C02, T06M11, T07G13	1007
mi268	B13G21, B15O19, F01M24, F03K10, T04M20	3006
mi277	B02K08, B06G11, B06L19, B12C14, B14I07, T04E23	2018
mi279	B02M13, B05L15, B09O04, B12O13, B15A04, F03K17, T06G17	4011
mi289	B02I12, B09A21, B10F12, T05D07	3004
mi291a	B02B10, B11O16, B15O17, B15N19, B16H16, T03F24, T05I04	1018
mi301	B14E10, B14I01, T05A19, T07B07	4001
mi303	B04O17, B07G22, B08K24, B08L14, B10C11, B10E01, B13A13, B14B04, B15A06, B16F06, F01K04, F03G17, T06H22	1020
mi306	B05G06, B07O07, T02B23, T06F14, T06F18, T07G22, T07N06	4003
mi310	B04K11, B09I20, B15N02, F04M17, T01C04, T01F17, T02J01, T03K03, T03J06, T04E14, T05B06, T06P09	2004
mi320	B01D22, B02N13, B15C12	2001

(*continued*)

Downloaded from https://academic.oup.com/genetics/article/159/3/1231/6049583 by guest on 19 April 2024

TABLE 2

(Continued)

Probe	Clone	Contig[a]
mi323	B01A09, B06J14, B14A23, F04O02, T05G18	5023
mi324	B02K22, B08N24, B09G12, B11H02, B14O24, F02K11, F04L13, T01N14, T03K11	1022
mi330	B09I16, B14P19, T06P12, T08B05	4011
mi339	B01H13, T04L24, T07H05	3005
mi342	B01I17, T03I20, T07A21, T08A20	1012
mi348	B10B09, B16J06, F02C14, F02O23, T02E17	1004
mi355	B14G08	3003
mi357	B02K21, B04B09, B04D04, B05C23, B09G02, B12D12, F01I03, F03K15, T01O12, T04C01, T08A07	3004
mi358	B04E22, B12M12, B12A16, B12H10, F03B08, F03P10, T02C08, T03C23, T04K14, T05G03, T06K16	3006
mi372	B01H03, B02O17, B04I08, B04K19, B06C08, T03F10, T06L08, T08I17	1001
mi386	B07H15, B13K17, T05D03	3006
mi398	B02P18, B08L05, B08N21	2009
mi413	B03B12, B03J05, B08I07, B11C22, T07I19	3009
mi422	B07L03, B14B15, T03P20, T04J24, T06I01	4014
mi423a	B09L02, B12O24, F04L20, T07F05	1011
mi423b	B10A02, B11D07, B12I11, B14I23, B14H05, B14F13, B15M08, F04A18	5025
mi424	B01P05, B02K05, B04N05, B13A17, B14C01, F02A01, T06J21, T08L05	1022
mi425	B04I16, B05N20, B07D11, F02H22, F03E06	1026
mi433	B03P15, B04A01, B05P13, B07F18, B11B07, T04L22	5009
mi438	B01K12, B04B10, B07C05, B08I04, B08F11, T02B06, T03C20, B16K16	5007
mi443	B04F07, B04D19, B05A19, B09C16, T04A12, T04F13, B14F16, B15F22, B16O12, F02J24	1002
mi456	B02E14, B07A08, B10A17, B15K10, F03C22	3018
mi462	B15P17, F02N03, F03G06, F03G19, T02K22	1025
mi465	B12P06, T01E20, T08A01, T08I09	4011
mi467	B02E19, B06K16, B08O01, F01P10, F03B23, F03L24, F04N16, T01A06, T01J13, T05L22	3004
AtAF1	B02O23, B03E03, B11M24, F02H17, F03K08, F04P07, T08A06	2019
AtNAM	B05G19, B05B02, B07P05, B09D02, B09H20, B11O22, B14P18, B16D19, F03H05, F04F05, T02M02, T03O02, T03O22	1022
AtS20-1	B01F01, B10F03, B11G08, B13B05, T01I07, T02C06, T02A24, T08K19	1020
AtS20-2	B02H12, B02J24, B03I18, B07A24, B07E19, B11M13, B11M19	3004
AtS478	B01F23, B03M15, B03B22, B04K21, B05L13, B09C01, B11H15, B12J02, B12J05, B14G15	3014
Oleosin	B06I14, B09L01, T03K09, T03N19, T08K24	2020
PEI1	B03M16, B08A24, B10N07, F03M10, F03J19, T08M16	3019
C11	T04O08, B15I06, B15I07, B15F24	1022
C12	B02K06, B02J04, B05J08, B06J17, B07M22, B10O15, B14G01, F01G19, T01C21, T02I01, T06I09, T07E22, T08I24	5002
C13	B01O10, B06H05, B08I18, B08J17, B13L08, B16F07	3008

[a] Each contig is named using a four-digit number, with the first one identifying the chromosome to which the contig was anchored and the last two specifying the contig. The clones were named as described in Figure 3.

could be transformed into plants via Agrobacterium. Since the BAC/BIBAC-based map developed here is integrated with the existing sequence map, the BIBAC sequences could be deduced from their overlapping BACs and used for large-scale functional analysis of the genome sequence by genetic transformation.

In addition to their utility in gap closure and functional analysis of the Arabidopsis genome sequence, the new integrated map and fingerprint database have provided a platform for numerous other studies of not only *A. thaliana*, but also of many other plant species. We have already received a number of inquiries, including those from the laboratories of the AGI. These studies include gap closure in the sequence map (see above), isolation of genes by positional cloning, identification of the functional sequence of centromeres, studies of gene regulation, engineering of a cluster of genes at a locus, and comparative genomics research between *A. thaliana* and crop plants. Since errors have recently been identified in the existing sequence map (Stupar *et al.* 2001), this new integrated map may provide a tool to identify and correct the errors in the sequence map because it was developed with a method different from those used for development of the existing maps. Furthermore, this study, along with the recent development of the whole-genome physical map of rice (Tao *et al.* 2001), has demonstrated that DNA sequence electrophoresis-based restriction fingerprint analysis is a reliable and high-throughput method for rapid genome-wide physical mapping of large, complex genomes from large-insert random bacterial clones.

Accessing the integrated map: The integrated map

Downloaded from https://academic.oup.com/genetics/article/159/3/1231/6049583 by guest on 19 April 2024

and the new BIBAC library have been posted at http://hbz.tamu.edu (Physical Mapping-Arab Map) and made available to the public. Users can access the map using any of the following approaches: clone-FPC hitting; clone-graphic contig map; clone-fingerprint map; contig no.-graphic contig map; or marker/EST-positive clones-contig/PFC hit/ fingerprint matches. The contigs, clones, and libraries can be requested at http://hbz.tamu.edu-BAC Library-Library List.

We thank Dr. Christopher Town for kindly providing the gap information in the Arabidopsis sequence map of AGI. We thank Drs. Huey-Wen Chuang and David A. Lightfoot for critically reading the manuscript. This study was supported in part by Texas Agricultural Experiment Station (8536-203104) and the Texas Higher Education Coordinating Board (999902-042).

LITERATURE CITED

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

Azpiroz-Leehan, R., and K. A. Feldmann, 1997 T-DNA insertion mutagenesis in *Arabidopsis*: going back and forth. Trends Genet. **13:** 152–156.

Bechtold, N., J. Ellis and G. Pelletier, 1993 *In planta Agrobacterium*-mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. C. R. Acad. Sci. Paris Life Sci. **316:** 1194–1199.

Bent, A. F., B. N. Kunkel, D. Dahlbeck, K. L. Brown, R. Schmidt *et al.*, 1994 RPS2 of *Arabidopsis thaliana*: a leucine-rich repeated class of plant disease resistance genes. Science **265:** 1856–1860.

Canilleri, C., J. Lafleuriel, C. Macadre, F. Varoquaux, Y. Parmentier *et al.*, 1998 A YAC contig map of *Arabidopsis thaliana* chromosome 3. Plant J. **15:** 633–642.

Chang, S. S., S. K. Park, B. C. Kim, B. J. Kang, D. U. Kim *et al.*, 1994 Stable genetic transformation of *Arabidopsis thaliana* by *Agrobacterium* inoculation in planta. Plant J. **5:** 551–558.

Choi, S. D., R. Creelman, J. Mullet and R. A. Wing, 1995 Construction and characterization of a bacterial artificial chromosome library from *Arabidopsis thaliana*. Weeds World **2:** 17–20.

Desprez, T., J. Amselem, M. Caboche and H. Hofte, 1998 Differential gene expression in *Arabidopsis* seedlings monitored using cDNA arrays. Plant J. **14:** 643–652.

Feldmann, K. A., and M. D. Marks, 1987 *Agrobacterium*-mediated transformation of germinating seeds of *Arabibdopsis thaliana*: a non-tissue culture approach. Mol. Gen. Genet. **208:** 1–9.

Hamilton, C. M., A. Frary, C. Lewis and S. D. Tanksley, 1996 Stable transfer of intact high molecular weight DNA into plant chromosomes. Proc. Natl. Acad. Sci. USA **93:** 9975–9979.

Hamilton, C. M., A. Frary, Y. Xu, S. D. Tanksley and H.-B. Zhang, 1999 Construction of tomato genomic DNA libraries in a binary-BAC (BIBAC) vector. Plant J. **18:** 223–229.

Hwang, I., T. Kohchi, B. M. Hauge and H. M. Goodman, 1991 Identification and map position of YAC clones comprising one-third of the *Arabidopsis* genome. Plant J. **1:** 367–374.

Jones, J. G., L. Shlumukov, F. Carland, J. English, S. R. Scofield *et al.*, 1992 Effective vectors for transformation, expression of heterologous genes, and assaying transposon excision in transgenic plants. Transgenic Res. **1:** 285–297.

Klein, M., U. Eckert-Ossenkopp, I. Schmiedeberg, P. Brandt, M. Unseld *et al.*, 1994 Physical mapping of the mitochondrial genome of *Arabidopsis thaliana* by cosmid and YAC clones. Plant J. **6:** 447–455.

Koncz, C., N. Martini, R. Mayerhofer, Z. Koncz-Kalman, H. Korber *et al.*, 1989 High-frequency T-DNA-mediated gene tagging in plants. Proc. Natl. Acad. Sci. USA **86:** 8467–8471.

Lin, X., S. Kaul, S. Rounsley, T. P. Shea, M.-I. Benito *et al.*, 1999 Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. Nature **402:** 761–768.

Liu, Y.-G., N. Mitsukawa, C. Lister, C. Dean and R. F. Whittier, 1996 Isolation and mapping of a new set of 129 RFLP markers in *Arabidopsis thaliana* using recombinant lines. Plant J. **10:** 733–736.

Liu, Y.-G., Y. Shirano, H. Fukaki, Y. Yanai, M. Tasaka *et al.*, 1999 Complementation of plant mutants with large genomic DNA fragments by a transformation-competent artificial chromosome vector accelerates positional cloning. Proc. Natl. Acad. Sci. USA **96:** 6535–6540.

Marra, M., T. Kucaba, M. Sekhon, L. Hiller, R. Martienssen *et al.*, 1999 A map for sequence analysis of the *Arabidopsis thaliana* genome. Nat. Genet. **22:** 265–270.

Martienssen, R. A., 1998 Functional genomics: probing plant gene function and expression with transposons. Proc. Natl. Acad. Sci. USA **95:** 2021–2026.

Meinke, D. W., J. M. Cherry, C. Dean, S. D. Rounsley and M. Koornneef, 1998 *Arabidopsis thaliana*: a model plant for genome analysis. Science **282:** 662–682.

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.

Mozo, T., K. Dewar, P. Dunn, J. R. Ecker, S. Fischer *et al.*, 1999 A complete BAC-based physical map of the *Arabidopsis thaliana* genome. Nat. Genet. **22:** 271–275.

Palmer, J. D., S. R. Downie, J. M. Nugent, P. Brandt, M. Unseld *et al.*, 1994 Chloroplast and mitochondrial DNAs of *Arabidopsis thaliana*: conventional genomes in an unconventional plant, pp. 37–62 in *Arabidopsis*, edited by E. M. Meyerowitz and C. R. Somerville. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Ruan, Y., J. Gilmore and T. Conner, 1998 Towards Arabidopsis genome analysis: monitoring expression profiles of 1400 genes using cDNA microarrays. Plant J. **15:** 821–833.

Schena, M., D. Shalon, R. W. Davis and P. O. Brown, 1995 Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science **270:** 467–470.

Schmidt, R., J. West, K. Love, Z. Lenehan, C. Lister *et al.*, 1995 Physical map and organization of *Arabidopsis thaliana* chromosome 4. Science **270:** 480–483.

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

Somerville, C., and J. Dangl, 2000 Plant biology in 2010. Science **290:** 2077–2078.

Stupar, R. M., J. W. Lilly, C. D. Town, Z. Cheng, S. Kaul *et al.*, 2001 Complex mtDNA constitutes an approximate 620-kb insertion on *Arabidopsis thaliana* chromosome 2: implication of potential sequencing errors caused by large-unit repeats. Proc. Natl. Acad. Sci. USA **98:** 5088–5103.

Sulston, J., F. Mallett, R. Staden, R. Durbin, T. Horsnell *et al.*, 1988 Software for genome mapping by fingerprinting techniques. Comput. Appl. Biosci. **4:** 125–132.

Sussman, M. R., R. M. Amasino, J. C. Young, P. J. Krysan and S. Austin-Phillips, 2000 The Arabidopsis knockout facility at the University of Wisconsin-Madison. Plant Physiol. **124:** 1465–1467.

Tao, Q.-Z., and H.-B. Zhang, 1998 Cloning and stable maintenance of DNA fragments over 300 kb in *Escherichia coli* with conventional plasmid-based vectors. Nucleic Acids Res. **26:** 4901–4909.

Tao, Q.-Z., Y.-L. Chang, J. Wang, H. Chen, C. Scheuring *et al.*, 2001 Bacterial artificial chromosome-based physical map of the rice genome constructed by restriction fingerprint analysis. Genetics **158:** 1711–1724.

Wu, Y., L. Tulsieram, Q. Tao, H.-B. Zhang and S. J. Rothstein, 2000 A binary vector-based large insert library for *Brassica napus* and identification of clones linked to a fertility restorer locus for *Ogura* cytoplasmic male sterility (CMS). Genome **43:** 102–109.

Zachgo, E. A., M. L. Wang, J. Dewney, D. Bouchez, C. Camilleri *et al.*, 1996 A physical map of chromosome 2 of *Arabidopsis thaliana*. Genome Res. **6:** 19–25.

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.

Communicating editor: C. S. Gasser

Downloaded from https://academic.oup.com/genetics/article/159/3/1231/6049583 by guest on 19 April 2024