Theor Appl Genet (2005) 111: 150–161
DOI 10.1007/s00122-005-2007-7

ORIGINAL PAPER

Y.-W. Nam · J.-R. Lee · K.-H. Song · M.-K. Lee
M. D. Robbins · S.-M. Chung · J. E. Staub
H.-B. Zhang

Construction of two BAC libraries from cucumber (*Cucumis sativus* L.) and identification of clones linked to yield component quantitative trait loci

Received: 2 February 2005 / Accepted: 17 March 2005 / Published online: 30 April 2005

Abstract Two bacterial artificial chromosome (BAC) libraries were constructed from an inbred line derived from a cultivar of cucumber (*Cucumis sativus* L.). Intact nuclei were isolated and embedded in agarose plugs, and high-molecular-weight DNA was subsequently partially digested with *Bam*HI or *Eco*RI. Ligation of double size-selected DNA fragments with the pECBAC1 vector yielded two libraries containing 23,040 *Bam*HI and 18,432 *Eco*RI clones. The average *Bam*HI and *Eco*RI insert sizes were estimated to be 107.0 kb and 100.8 kb, respectively, and BAC clones lacking inserts were 1.3% and 14.5% in the *Bam*HI and *Eco*RI libraries, respectively. The two libraries together represent approximately 10.8 haploid cucumber genomes. Hybridization with a C_0t-1 DNA probe revealed that approximately 36% of BAC clones likely carried repetitive sequence-enriched DNA. The frequencies of BAC clones that carry chloroplast or mitochondrial DNA range from 0.20% to 0.47%. Four sequence-characterized amplified region (SCAR), four simple sequence repeat, and an randomly amplified polymorphic DNA marker linked with yield component quantitative trait loci were used either as probes to hybridize high-density colony filters prepared from both libraries or as primers to screen an ordered array of pooled BAC DNA prepared from the *Bam*HI library. Positive BAC clones were identified in predicted numbers, as screening by polymerase chain reaction amplification effectively overcame the problems associated with an overabundance of positives from hybridization with two SCAR markers. The BAC clones identified herein that are linked to the *de* (determinate habit) and *F* (gynoecy) locus will be useful for positional cloning of these economically important genes. These BAC libraries will also facilitate physical mapping of the cucumber genome and comparative genome analyses with other plant species.

Keywords BAC libraries · Cucumber · Molecular markers · PCR screening · Quantitative trait loci · Yield components

Communicated by I. Paran

Y.-W. Nam (✉) · J.-R. Lee · K.-H. Song
Department of Life Science,
Sogang University, Shinsoo-dong,
Mapo-gu, Seoul 121-742, Korea
E-mail: ywnam@ccs.sogang.ac.kr
Fax: +82-2-7043601

M.-K. Lee · H.-B. Zhang
Department of Soil and Crop Sciences,
Institute for Plant Genomics and Biotechnology,
Texas A&M University, College Station,
TX 77843-2123, USA

M. D. Robbins · S.-M. Chung · J. E. Staub
Department of Horticulture and the
United States Department of Agriculture,
Agricultural Research Service,
University of Wisconsin, Madison,
WI 53706-1590, USA

Introduction

Cucumber (*Cucumis sativus* L., $2n = 2x = 14$) is one of the ten most economically important vegetable crops worldwide (FAO 1993). Cucumber is also a facile experimental model plant species for genetic studies (Malepszy and Niemirowicz-Szczytt 1991) because of its relatively small haploid genome of 367 Mbp/C (Arumuganathan and Earle 1991) and its diverse array of unisexual or bisexual flowering sex phenotypes. Cucumber has seven pairs of chromosomes (Dane 1991) and belongs to the dicot family Cucurbitaceae, which includes melon (*C. melo* L.), watermelon [*Citrullus lanatus* (Thunburg) Matsumura and Nakai] and zucchini (*Cucurbita pepo* L.) (Jeffrey 1980). Although cucumber plants exhibit an array of sex phenotypes from monoecious to dioecious, the expression of such phenotypes can shift from gynoecious (all pistillate) to androecious (all staminate) depending on genetic and environmental factors (Cantliffe 1981; Grant et al. 1994). The *F* (gynoecy) locus, for instance, controls the expression of

pistillate flowers when plants develop under non-stressful conditions (e.g., controlled greenhouse), but its control is altered by the action of modifying genes potentiated under stressful growing environments typical of open-field conditions (e.g., low soil moisture and high temperature) (Serquen et al. 1997b; Fazio et al. 2003b). Such alterations in sex phenotype are stimulated by environmental conditions that elicit modified patterns of gene expression through the action of ethylene (Kamachi et al. 2000). In fact, a 1-aminocyclopropane-1-carboxylic acid (ACC) synthase gene (*CsACS1G*) that encodes a key enzyme in the ethylene biosynthetic pathway is tightly linked with the *F* locus (Trebitsh et al. 1997).

Sex expression is an important component of yield, and a major objective of cucumber improvement programs has been to increase yield (Serquen et al. 1997a). However, average yield of US processing cucumber has remained virtually unchanged for the past 20 years (USDA-NASS 1999) despite early efforts (1950–1980) to pyramid specific yield, disease resistance, and sex (gynoecy) genes in elite breeding lines (Lower and Edwards 1986). Consequently, breeding to incorporate architectural genes such as *de* (determinate habit) and quantitative trait loci (QTL) for multiple lateral branching (MLB), days to anthesis, female nodes on the mainstem and lateral branches, and fruit number and weight have been the focus of some plant improvement efforts (Serquen et al. 1997a, 1997b; Fazio et al. 2003a, 2003b). As a result of such efforts, economically important molecular marker-trait associations have been identified, and a moderately saturated map spanning 706 cM with a mean marker interval of 5.6 cM has been constructed using polymerase chain reaction (PCR)-based randomly amplified polymorphic DNA (RAPD), sequence-characterized amplified region (SCAR), amplified fragment length polymorphisms (AFLP), single nucleotide polymorphisms (SNP), and simple sequence repeat (SSR) markers (Bradeen et al. 2001; Fazio et al. 2002, 2003b). The development of recombinant inbred lines (RILs) that are essentially homozygous at all loci has facilitated these mapping efforts (Staub et al. 1996; Fazio et al. 2003b). Many important yield component QTL are located in linkage groups 1 and 6, wherein QTL exhibiting relatively high LOD scores (2.7–13.0) overlap at a limited number of genomic regions (designated as loci by Fazio et al. 2003b). The construction of physical contigs that span these valuable genomic regions would likely provide for an increased understanding of the genetic factors controlling yield components at the molecular level. Moreover, the identification of molecular markers tightly linked to important architectural genes, such as *de*, *F*, and *ll* (little leaf) in linkage group 1 has increased opportunities for their isolation by map-based cloning for potential use in plant improvement.

To facilitate positional cloning of yield component genes of cucumber, a genomic library is required that enables chromosome walking to targeted genes. The bacterial artificial chromosome (BAC) system has been effective for constructing such a large-insert library. The BAC libraries have been proven capable of carrying and stably maintaining inserts of 100–300 kb in size (Shizuya et al. 1992; Tao and Zhang 1998). Unlike the previously introduced yeast artificial chromosome (YAC) system (Burke et al. 1987), however, the BAC system has a significantly low rate of chimera formation (Woo et al. 1994). Importantly, the DNA of BACs is much easier to manipulate than that of YACs. The BAC libraries have been constructed in a number of plant species including *Arabidopsis* (Choi et al. 1995; Mozo et al. 1998), grasses (Woo et al. 1994; Tao et al. 2002), legumes (Danesh et al. 1998; Nam et al. 1999), the solanaceous plants (Budiman et al. 2000; Yoo et al. 2003; Chen et al. 2004), and melon in the Cucurbitaceae family (Luo et al. 2001; van Leeuwen et al. 2003). The BAC libraries have facilitated map-based cloning of important plant genes (Song et al. 1995; Li et al. 1996; Frary et al. 2000), and have provided unique opportunities for whole-genome physical mapping (Chang et al. 2001; Wu et al. 2004b). Additionally, the ability to directly sequence BAC ends has facilitated large-scale local- or whole-genome sequencing (Venter et al. 1996; Zhang and Wu 2001).

The construction and characterization of cucumber BAC libraries is an important step leading to a more complete understanding of its genome. Therefore, experiments were designed for BAC library construction that resulted in the establishment of two such libraries in an inbred line of this species and the derivation of a sequentially ordered array of BAC DNA pools. During this process, we identified BAC clones linked to a previously mapped architectural gene (*de*) and the *F* locus that is associated with an ACC synthase gene. These clones are expected to serve as the physical landmarks from which chromosome walking toward the *F* and *de* loci (physical mapping) could be executed for eventual use in plant improvement and comparative genomic analyses (e.g., melon vs cucumber).

Materials and methods

Plant material and preparation of high-molecular-weight DNA

An inbred line (F_6) derived from the cucumber cultivar Suiseifushinari 2-go was used as the source of high-molecular-weight (HMW) DNA to create the libraries. Suiseifushinari 2-go, a release from Kurume Vegetable Breeding, Japan, was obtained from the Cucumber Crops Experiment Station (http://gurecum.jares.go.kr) of Jeonnam Agricultural Research and Extension Services, Gure, Jeonnam, Korea. This cultivar was chosen because of its resistance to downy mildew [*Pseudoperonospora cubensis* (Berk. & Curt) Rostow], and its broad use as an inbred parent for hybrid seed production in Korea and Japan. Additionally, it is similar in characteristics such as large leaves, gynoecious sex expression, and determinate growth habit to G421 (synom.

GY7), which has been used extensively as an inbred parent for genetic mapping of morphological traits (Serquen et al. 1997b; Fazio et al. 2003b). Seeds germinated in flats produced seedlings that were grown in a growth chamber with a daily cycle of 14 h of light at 25°C, and 10 h of dark at 20°C for a week, before transplanting into pots. After 3 weeks in pots, plants were kept in the dark for 2 days before nuclei isolation. The nuclei were isolated from 90 g of leaves and embedded in low-melting-point agarose plugs as described by Zhang et al. (1995). The HMW DNA was obtained by lysing the nuclear envelope followed by repeated washing according to Zhang (2000).

Partial digestion and double size selection

Each plug of HMW DNA released from nuclei was sliced into nine small pieces with a slide cover slip and incubated in the *Bam*HI or *Eco*RI reaction buffer containing 2 mM spermidine and 1 mM dithiothreitol on ice for 30 min. Three slices were then transferred into a 1.5-ml Eppendorf tube, and 170 µl of digestion buffer was added. To each tube, 0.9, 1.2, or 1.5 U of *Bam*HI or 1.6, 2.0, or 2.4 U of *Eco*RI were added, and then the mixture was pre-incubated on ice for 1 h, and subsequently incubated at 37°C for 8 min. The resulting DNA fragments were combined and loaded in a 1% agarose gel for pulsed field gel electrophoresis (PFGE). The PFGE was carried out with the CHEF DRIII system (Bio-Rad, Hercules, Calif., USA) under the conditions of 120° angle, 6 V/cm, and 90 s of initial and final pulse times at 12.5°C for 14 h. After PFGE, a zone containing DNA fragments of 100–400 kb in size was excised from the gel and divided into two sections horizontally. Each gel section was placed in a dialysis tube and submerged in 0.5× 45 mM Tris–borate and 1 mM EDTA, pH 8.0, in the CHEF chamber. The DNA was electroeluted under the conditions of 120° angle, 6 V/cm, and 30 s of initial and final pulse times at 12.5°C for 4 h. The eluted DNA was subjected to second size selection. The DNA fragments from the 100- to 250-kb and 250- to 400-kb zones were loaded separately in a new 1% agarose gel, and PFGE was carried out at 120° angle, 4 V/cm, and 5 s of initial and final pulse times at 12.5°C for 8 h. After PFGE, compressed zones containing DNA fragments of over 100 kb in size were excised from the gel and subjected to electroelution again. Finally, the DNA samples were recovered and kept on ice before use in ligation experiments.

Library construction

The pECBAC1 vector (Frijters et al. 1997) was prepared as previously described (Tao et al. 2002). Size-selected cucumber DNA fragments were mixed with the *Bam*HI-digested or *Eco*RI-digested and dephosphorylated vector at the molar ratio of approximately 5 (vector) to 1 (insert) in a volume of ~100 µl. Ligation was performed in the presence of T4 DNA ligase (Invitrogen, UK) at 16°C for 12 h. After ligation, the mixture was electroporated into *Escherichia coli* DH10B electrocompetent cells (Invitrogen) by using the Cell-Porator system (Invitrogen) under the conditions of 375 V, 330 μF, low ohms, and the fast charge rate. Cells were immediately transferred to 1 ml of 2% Bacto-tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 20 mM $MgSO_4$/$MgCl_2$, and 10 mM glucose, and shaken at 37°C for 1 h. Cells were subsequently spread on LB plates that contained 12.5 µg/µl of chloramphenicol, 50 µg/ml of 5-bromo-4-chloro-3-indolyl-β-galactoside and 25 µg/ml of isopropyl-β-D-thiogalactopyranoside. After incubation at 37°C for 24 h, white colonies were transferred into 384-well microtiter plates (Nunc). Each well of the plates contained 60 µl of LB plus freezing buffer [36 mM K_2HPO_4, 13.2 mM KH_2PO_4, 1.7 mM Na citrate, 0.4 mM $MgSO_4 \cdot 7H_2O$, 6.8 mM $(NH_4)_2SO_4$, 4.4% (v/v) glycerol]. The plates were then incubated at 37°C for 24 h, and stored at −80°C.

Preparation of specific probes and C_0t-1 DNA

The *F* locus-linked ACC synthase gene (*CsACS1G*) probe was amplified from genomic DNA of the cucumber inbred line using two primers: 5′-CAA CCA GCT TTA GAA CAA GC-3′ (ACS3F) and 5′-ACT TCA ATC TTC GGA TAG CG-3′ (ACS3R). For *CTR1* that encodes a negative regulator of the ethylene signaling pathway (Kieber et al. 1993), two degenerate primers were designed from conserved regions of CTR1 proteins available at the National Center for Biotechnology Information GenBank database: 5′-GMT GYH TRT CRT ACT WTG AYA AAG TBC CTG-3′ (forward) and 5'-ADA RYA AAG CWC GAT GYC TRC ARA GBC CAA-3′ (reverse), wherein letters other than A, C, G, and T denote nucleotide mixtures (B: C, G, or T; D: A, G, or T; H: A, T, or C; K: G or T; M: A or C; R: A or G; S: C or G; V: A, C, or G; W: A or T; and Y: C or T). The amplified *CTR1* fragment was subcloned into the pBS SK(−) vector (Stratagene) before being used as a hybridization probe. Six restriction fragment length polymorphism (RFLP) markers that belong to the linkage groups as designated by Bradeen et al. (2001) were used: CsP094, group G; CsP130, group J; CsP280, group E; CsP314, group K; CsP347, group H; and CsC558, group E/F/H. A peroxidase cDNA clone of zucchini (*Cucurbita pepo*) was a gift from Dr. C. Penel (University of Geneva, Geneva, Switzerland). Two barley (*Hordeum vulgare*) chloroplast cDNA (*rbcL* and *psbA*) and four pepper (*Capsicum annuum*) mitochondrial cDNA (*cox*II, *cox*III, *atp*6, and *atp*9) were gifts from Dr. M. Kim (Seoul National University, Seoul, Korea) and Dr. S. Kim (FnP, Suwon, Korea), respectively. To prepare C_0t-1 DNA (Zwick et al. 1997), total genomic DNA was isolated from cucumber leaves, and then dissolved in one time 10 mM Tris-HCl, 1 mM

EDTA, pH 8.0, (TE) containing 1.5 M NaCl. Subsequent preparation of C_0t-1 DNA was carried out as previously described by Nam et al. (1999).

QTL-linked molecular markers

Nine QTL-linked markers, including four SCAR, four SSR, and one RAPD, were employed in this study. Seven (BC523SCAR, CSWCT28, CSWCTT14, CSWTAAA01, L19-2-SCAR, NR60, and OP-W7-1) were described previously (Fazio et al. 2003b), and the remaining SCARs (AJ6SCAR and M8SCAR) were derived from their corresponding RAPD markers, OP-AJ6 and OP-M8 (Fazio et al. 2003b), respectively (Table 1). Their primer sequences are 5′-GAT GGC AGT CTG ATA ACT ATG TGA-3′ (AJ6SCARF), 5′-GAT GGC AGT CGG GAA GGT CAG TTG-3′ (AJ6SCARR), 5′-TCT GTT CCC CAT ACA AGA ATT AAA-3′ (M8SCARF), and 5′-TCT GTT CCC CAT GAT GTA GAC TTC-3′ (M8SCARR). Using these primers, probe DNA was amplified from genomic DNA of the inbred lines G421 or H19. Alternatively, the primers for the AJ6SCAR and BC523SCAR markers were directly used for PCR screening of the BAC library DNA pools.

Filter hybridization and BAC clone analysis

Three sets of double-spotted, high-density BAC clone filters in a 4×4 format were prepared from 48 plates (plates 1–48) of the *Bam*HI library constructed herein using the GeneTAC G3 Robotic Workstation (GenomicSolutions, Ann Arbor, Mich., USA) at the GENEfinder Genomic Resources (Texas A&M University, College Station, Tex., USA). Six additional sets of filters in a 4×4 format were prepared from all the plates of the *Bam*HI and *Eco*RI libraries using the Microgrid II robot (Biorobotics) at FnP. Each filter set was hybridized with a gene-specific probe. Probe labeling was carried out with [α-^{32}P]dCTP (New England Nuclear) using the Ready-to-Go labeling kit (Amersham Pharmacia, Piscataway, N.J., USA). Filters were prehybridized in a buffer [0.5 M sodium phosphate, 7% (w/v) SDS, 1% (w/v) BSA, 1 mM EDTA, pH 8.0] for 2 h. Hybridization was carried out with labeled probes in the HIR12 hybridization oven (Grant/Boekel, Cambridgeshire, UK) at 65°C for 12 h. After hybridization, filters were washed successively with two times SSC/0.1% SDS, one time SSC/0.1% SDS, and 0.5 times SSC/0.1% SDS, 30 min each at 65°C. Washed filters were then exposed to Kodak MXG-1 films (Kodak) at −80°C for ~48 h and developed to visualize results.

For randomly selected BAC clones or positive BAC clones obtained from hybridization, DNA was isolated by an alkaline mini-preparation method according to Zhang (2000). DNA was digested with *Not*I and analyzed by PFGE with the CHEF DRIII system under the conditions of 120° angle, 6 V/cm, and 5s and 15 s of initial and final pulse times with a linear ramp at 12.5°C for 16 h. Gel-separated DNA fragments were then transferred to Hybond-N$^+$ nylon membrane (Amersham Pharmacia), and the blot was used for hybridization.

Pooling of the library for PCR screening

The 23,040 *Bam*HI library clones stored in 60 384-well microtiter plates were organized into 30 sets, two plates per set. All the clones from two consecutive plates were pooled together and designated as set pools. The two plates from a given set were vertically placed adjacent to each other, and clones belonging to the same column

Table 1 Characteristics of gene-specific and quantitative trait loci (*QTL*) markers used to screen the cucumber bacterial artificial chromosome (BAC) libraries

Marker name	Type[a]	Parent[b]	MW[c]	Linkage group[d]	Linked QTL[e] (LOD score)[f]	Source or reference
AJ6SCAR	SCAR	G	541	1	B (3.3), E (2.7)	This study
BC523SCAR	SCAR	G	879	1	E (2.7)	Fazio et al. (2003b)
CSWCT28	SSR	G& H	207	1	E (7.1), M (3.6)	Fazio et al. (2003b)
CSWCTT14	SSR	G& H	263	1	B (11.6), E (6.4), Fl (3.3), Fm (7.3), M (4.6)	Fazio et al. (2003b)
CSWTAAA01	SSR	G& H	306	4	B (4.6)	Fazio et al. (2003b)
F	Gene[g]	G	754	1	Fl (3.8), Fm (13.0)	Fazio et al. (2003b)
L19-2-SCAR	SCAR	H	1009	6	B (4.2)	Fazio et al. (2003b)
M8SCAR	SCAR	H	1400	6	B (2.7)	This study
NR60	SSR	G& H	170	6		Fazio et al. (2003b)
OP-W7-1	RAPD	H	1150	6		Fazio et al. (2003b)

[a] *SCAR* Sequence-characterized amplified region, *SSR* simple sequence repeat, *RAPD* randomly amplified polymorphic DNA
[b] *G* present in G421, *H* present in H19, *G& H* present in G421 and H19 (codominant marker)
[c] Estimated molecular weight in base pairs
[d] Linkage groups as designated by Fazio et al. (2003b)
[e] *B* number of lateral branches, *E* earliness, *Fl* female nodes (lateral), *Fm* female nodes (mainstem), *M* mean fruit length/diameter ratio
[f] Highest LOD score shown from values obtained from multiple field trials (Serquen et al 1997b; Fazio et al. 2003a, 2003b)
[g] *CsACS1G* Genomic DNA fragment

(across both plates) or row were pooled together. This gave 24 column pools and 32 (16 from each plate) row pools that contained 25-µl aliquots from each clone. Together, 30 set pools and 1,680 [30×(24 + 32)] primary column or row pools were produced. The DNA was isolated from each pool by the alkaline lysis method and dissolved in 40 µl of one-time TE. To reduce the number of PCR reactions necessary for identifying positive BACs, the primary column and row pools from a given set were pooled again two-dimensionally (see Fig. 4). Aliquots were taken from either horizontally or vertically arranged tubes and combined. This gave ten (six horizontal and four vertical)-column and 12 (eight horizontal and four vertical)-row superpools per set, wherein the final clone complexity was calculated to be 1/192 to 1/96. The resulting superpools contained BAC DNA at approximately 20 ng/µl per tube.

DNA in each superpool was diluted 1/10 in one time TE and used as template for PCR screening of the library. For PCR, 1 µl of diluted BAC DNA template was added to the reaction mixture (2.5 mM $MgCl_2$, 50 µM dNTPs, 0.25 pmol primers, and 0.1 U *Taq* polymerase). A typical amplification condition was 94°C, 30 s; 55°C, 20 s; and 72°C, 1 min, for 35 cycles. After PCR, the products were analyzed on a 2.0% agarose gel using the wide-format electrophoresis system (Owl Scientific, Woburn, Mass., USA).

Results

Construction of BAC libraries

Two cucumber BAC libraries were constructed from HMW DNA that was isolated from nuclei embedded in agarose plugs. When prepared, the agarose plugs were white and opaque, indicating the presence of a relatively high amount of polysaccharides. After lysis of the nuclear envelope, however, large quantities of partially digested HMW DNA were observed to migrate in the pulsed field gel. The optimal concentration range of *Bam*HI for producing the maximum number of DNA fragments of 100–400 kb in size was 0.9–1.5 U/tube.

To increase the average insert size, two size selections were carried out. Partially digested DNA fragments were gel-separated and size-selected once, and then the resulting electroeluted fragments were gel-separated again to remove fragments smaller than ~80 kb in size. Two preparations were consequently obtained from 100- to 250-kb (Cs8B1) and 250- to 400-kb regions (Cs8B2). These samples contained DNA of at least ~1 ng/µl, which was suitable for ligation with the dephosphorylated pECBAC1 vector. Ligation mixtures from Cs8B1 and Cs8B2 gave 145 and 25 white colonies per microliter transformation, respectively. The PFGE analysis of randomly selected clones from these mixtures showed that almost every clone carried inserts (Fig. 1a). Another ligation mixture prepared with DNA fragments from the 250- to 400-kb region after only one round of size selection also yielded white colonies (160 μl^{-1}). However, more than 25% of the clones lacked inserts, which, when present, were significantly small (< 50 kb, data not shown). Therefore, the two-ligation mixtures from Cs8B1 and Cs8B2 were used exclusively for mass electroporation. Finally, a total of 23,040 clones were arrayed in 60 384-well microtiter plates.

A similar strategy was used to obtain partially digested fragments for the *Eco*RI library, except that the optimal enzyme concentration range was 1.6–2.4 U/tube. Two ligation reactions were carried out with Cs12E1 (100–250 kb) and Cs12E2 (250–400 kb) to the dephosphorylated pECBAC1 vector. Mass electroporation produced a library containing 18,432 clones arrayed in 48 384-well microtiter plates.

Estimation of the average insert sizes and repetitive DNA content

Seventy-seven *Bam*HI and 117 *Eco*RI BAC clones were randomly selected, and their plasmids were isolated.

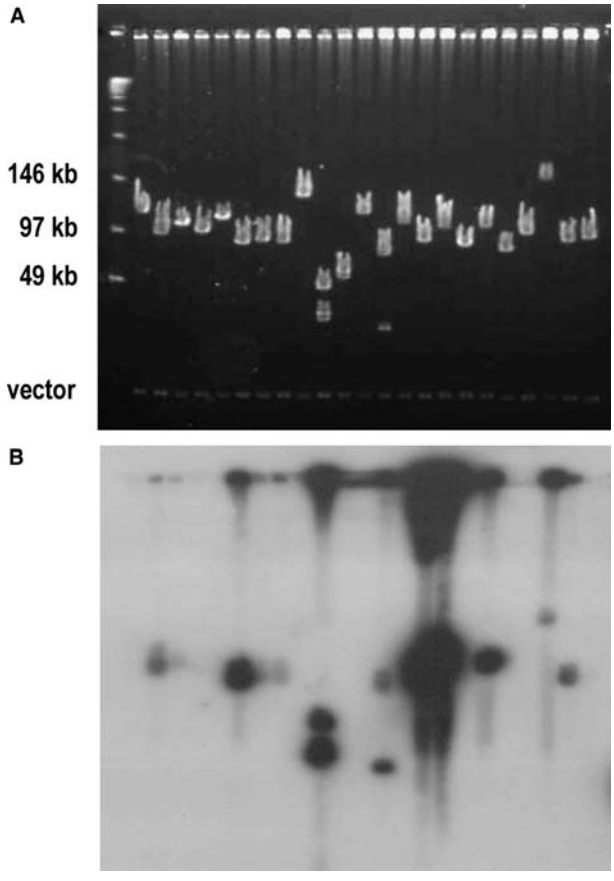

Fig. 1 The pulsed field gel electrophoresis (PFGE) analysis of bacterial artificial chromosome (BAC) clones of the *Bam*HI library. **a** Twenty-three BAC clones were randomly selected, and their plasmids were analyzed by *Not*I digestion and PFGE. Lambda concatemer markers are shown on the *left*. **b** Southern hybridization with C_0t-1 DNA. Gel-separated DNA fragments in **a** were transferred to nylon filter and hybridized with the cucumber C_0t-1 DNA probe

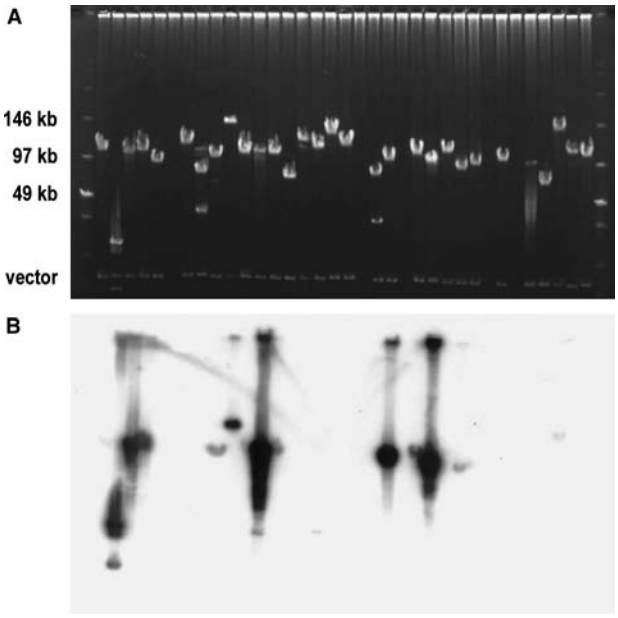

Fig. 2 The PFGE analysis of BAC clones of the *Eco*RI library. **a** Thirty-five randomly selected BAC clones were analyzed by *Not*I digestion and PFGE. Lambda concatemer markers are shown on the *left*. **b** Southern blot hybridization of the gel in **a** with the C_0t-1 DNA probe

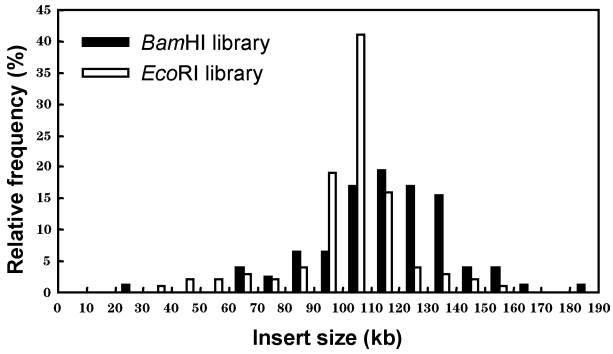

Fig. 3 Distribution of insert sizes of the cucumber BAC library clones. The BACs isolated from 76 and 100 randomly selected and insert-carrying clones from the *Bam*HI and *Eco*RI libraries, respectively, were analyzed by *Not*I digestion and PFGE. Estimated insert sizes are plotted vs relative frequencies (percentage)

When DNA was analyzed by *Not*I digestion and PFGE, 76 *Bam*HI and 100 *Eco*RI BACs carried inserts, and only one *Bam*HI (1.3%) and 17 *Eco*RI (14.5%) BACs had no inserts. As shown in Figs. 1a and 2a, the majority of DNA inserts fell into the 90- to130-kb range. The average insert sizes were calculated to be 107.0 kb for the *Bam*HI library and 100.8 kb for the *Eco*RI library (Fig. 3).

To examine the repetitive DNA content of the library clones, C_0t-1 DNA (Zwick et al. 1997) was prepared from cucumber. The 23 *Bam*HI and 30 *Eco*RI BAC DNA fragments separated by PFGE (Figs. 1a, 2a) were transferred to nylon filters, and the blots were hybridized with the radiolabeled C_0t-1 DNA probe. A total of 25 (12 *Bam*HI and 13 *Eco*RI) clones hybridized with this repetitive sequence-enriched DNA (Figs. 1b, 2b). Of these, 11 (five *Bam*HI and six *Eco*RI) clones hybridized strongly, eight (four *Bam*HI and four *Eco*RI) moderately, and six (three *Bam*HI and three *Eco*RI) weakly with the C_0t-1 DNA. Simple calculation from these results indicates that approximately 47% of the library clones examined contains repetitive sequences. Likewise, approximately 36% of the clones are likely to carry inserts that are derived from moderately or highly repeated DNA elements in the cucumber genome.

Library screening with specific probes and RFLP markers

To further characterize these libraries, high-density colony filters were prepared and used for hybridization with specific probes. When a ^{32}P-labeled mixture of chloroplast *rbcL* and *psbA* cDNA was used as the probe, 72 *Bam*HI- and 32 *Eco*RI-positive clones were identified. The two genes used as probes are located approximately 60 kb apart from each other in the chloroplast genome and hybridized with at least two-thirds of cucumber BAC clones examined that contain chloroplast-derived DNA inserts of ≥60 kb (Woo et al. 1994). It is therefore estimated that 108 *Bam*HI clones (0.47% of library) and 48 *Eco*RI clones (0.26% of library) carry chloroplast-originated DNA sequences. Hybridization with the mitochondrial DNA probe containing a radiolabeled mixture of *cox*II, *cox*III, *atp*6, and *atp*9 genes (Yoo et al. 2003) identified 93 *Bam*HI- and 36 *Eco*RI-positive clones, respectively. Thus, approximately 0.40% of the *Bam*HI and 0.20% of the *Eco*RI libraries likely carry inserts derived from mitochondrial DNA.

Three gene-specific fragments and six RFLP markers were used as nuclear DNA-specific probes. A probe derived from the *F* locus-linked *CsACS1G* gene (Trebitsh et al. 1997) hybridized to 79 *Bam*HI and 46 *Eco*RI library clones. However, the intensities of the radioactive signals varied, indicating that the probe likely hybridized to DNA sequences with different levels of relatedness. Further screening with gene-specific primers by PCR reduced the number of positives to three (one *Bam*HI and two *Eco*RI clones, Table 2). In a subsequent experiment, an ethylene-signaling pathway gene, *CTR1*, was used as the probe, and five *Bam*HI- and three *Eco*RI-positive BACs were identified (Table 2). A peroxidase cDNA of zucchini was then used as a probe to verify the genomic relationship between cucumber and zucchini. This probe hybridized to six *Bam*HI and six *Eco*RI BACs (Table 2). Finally, hybridization with six RFLP clones belonging to six linkage groups (Bradeen et al. 2001) resulted in zero to ten positive BACs from each library (data not shown). In total, three gene-specific and six RFLP probes identified 43 *Bam*HI BACs and 33 *Eco*RI BACs, averaging 4.8 (*Bam*HI library) and 3.7 (*Eco*RI library) clones per probe. The only notable exception was the RFLP marker, CsC558, with which

Table 2 BAC clones of the *Bam*HI (prefixed with *B*-) and *Eco*RI (prefixed with *E*-) libraries identified with gene-specific and molecular marker probes

Probe	Number of hits	Positive clones (insert size in kb)[a]
Gene-specific		
CsACS1G[b]	3[c]	B35J22 (180), E25E11 (100), E34K19 (90)
CTR1[d]	7	B09H24 (145), B12F24 (100), B19D01 (90), B56D04 (95), E01E20 (120) E25D21 (105), E29A10 (100)
Peroxidase[e]	12	B03F13 (100), B09D03 (120), B18D19 (100), B36G13 (125), B46H01 (105) B48G15 (100), E08D01 (110), E16P04 (115), E17J01 (115), E19M15 (105) E27G11 (125), E42P07 (120)
Molecular marker		
AJ6SCAR	3[f]	B07A24 (105), B18A07 (100), B50O22 (170)
BC523SCAR	9[f]	B12P24 (110), B14K20 (90), B19M17 (100), B22O22 (90), B44D16 (120) B47L04 (85), B56O12 (95), B58D06 (110), B60N01 (155)
L19-2-SCAR	2	B08C06 (105), B22B20 (105)
M8SCAR	20	B01E13 (105), B07J14 (105), B09I04 (105), B14F06 (100), B22B19 (100) B22O04 (155), B24B17 (125), B28O19 (110), B29B20 (105), B32F15 (105) B35D22 (100), B37K17 (110), B42I09 (105), B43B20 (115), B50D07 (130) B54F13 (120), B60J19 (120), E06N15 (105), E15E16 (155), E47G15 (95)
CSWCT28	9	B10B02 (120), B12J22 (115), B21C18 (115), B22B18 (110), B23I05 (105) B44J23 (105), E14M13 (95), E36P12 (100), E41O16 (95)
CSWCTT14	25	B10P11 (85), B18P23 (120), B31C15 (105), E03L13 (105), E05C10 (115) E10G23 (105), E11J19 (115), E11P14 (105), E12P12 (130), E13I01 (110) E18J09 (120), E20E11 (100), E22B02 (95), E27J19 (105), E29I11 (100) E31G01 (95), E33A02 (105), E34M19 (90), E37D08 (105), E38L03 (100) E40N14 (95), E41P03 (105), E44D24 (95), E46F09 (95), E47P10 (100)
CSWTAAA01	8	B05E10 (110), B15M07 (110), B17C05 (110), B41H18 (110), B42P04 (105) E11B08 (125), E13K10 (140), E28F03 (125)
NR60	10	B11D14 (120), B11F12 (170), B12A06 (40), B35A18 (120), E02D10 (50) E19A08 (110), E19N10 (115), E24D02 (105), E35O12 (95), E41O10 (115)
OP-W7-1	15	B10B15(125), B23B20(115), E03K04 (110), E13A08 (50), E17N15 (100) E22B12 (100), E29G06 (105), E34A09 (95), E38F05 (90), E40B02 (105) E40G16 (100), E41B24 (100), E42M07 (105), E43H11 (145), E44J08 (130)

[a]Positive-BAC DNA was isolated and analyzed by *Not*I digestion and pulsed field gel electrophoresis
[b]An *F* locus marker probe
[c]Determined by confirmative polymerase chain reaction (PCR) amplification of the initial positive BACs that resulted from hybridization
[d]Cucumber DNA fragment amplified by degenerate PCR
[e]Zucchini cDNA clone (Carpin et al. 1999)
[f]Determined by PCR amplification of the multiplex BAC DNA pools

no positive BAC clones were identified from the *Eco*RI library.

Library screening with QTL-linked markers

Four SCAR, four SSR, and an RAPD marker were used to identify BACs linked to mapped QTL conditioning traits important to cucumber yield. These traits include MLB, earliness, mainstem, and lateral female flowering nodes, and mean fruit length to diameter ratio (Table 1). When each PCR-amplified fragment was used as a probe for filter hybridization, multiple BAC clones hybridized with predicted numbers of positives (2–20) for seven of the nine markers (CSWCT28, CSWCTT14, CSWTAAA01, L19-2-SCAR, M8SCAR, NR60, and OP-W7-1; Table 2). However, no positive BAC clone was identified with L19-2-SCAR from the *Eco*RI library. For the AJ6SCAR marker, 160 positives were identified from both libraries (data not shown). Similarly, BC523SCAR hybridized with approximately 1,500 *Bam*HI and 600 *Eco*RI BACs. Such observations were, however, not considered unusual, because hybridization with three other RAPD markers (BC515, OP-AD12-1, and OP-AI10) also provided thousands of positives (data not shown). Thus, the identification of a finite number of positives with these markers through hybridization screening was not possible.

PCR screening of BAC library DNA pools with QTL markers

To facilitate PCR-based screening of the BAC libraries, a multiplex DNA pool system was developed from the *Bam*HI library (see "Materials and methods"). Using this pool system, three rounds of PCR analysis were required to identify individual BAC clones. The first round of 30 reactions on the set pools identified PCR-positive sets, the second round of PCR on the super-pools identified the specific column and row intersects, and the final round of PCR on non-pooled, specific clones identified positive BAC clones in the original library. The multiplex system was tested with primers

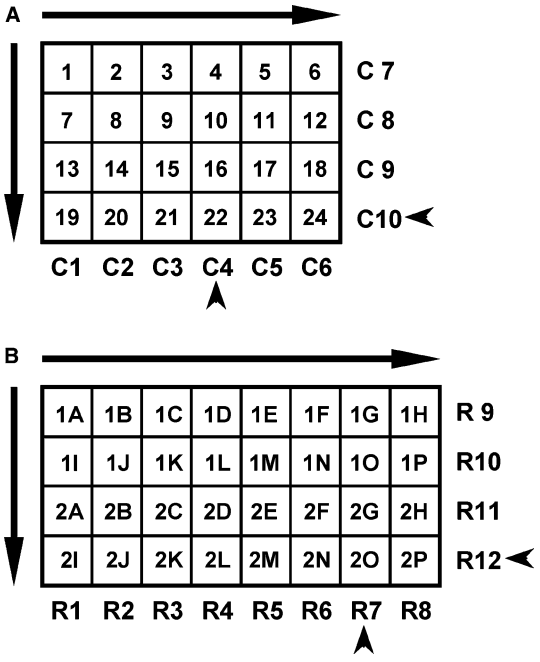

Fig. 4 Schematic representation of the pooling strategy used to construct the BAC DNA superpools from the *Bam*HI library. **a** Two-dimensional pooling of 24-column DNA pools (1–24 *in squares*). **b** Two-dimensional pooling of 32-row DNA pools (1A–1P and 2A–2P *in squares*). By these strategies, ten-column and 12-row superpools were constructed from each of the 30 set pools. *Arrowheads* indicate the pools from set no. 11, from which positive fragments were amplified as shown in Fig. 5

specific to *CsACS1G* and peroxidase cDNA. As predicted, the same single and six clones, respectively, as detected from hybridization screening (Table 2) were identified from the *Bam*HI library (data not shown). Subsequently, use of the primers for the AJ6SCAR marker identified three positives (B07A24, B18A07, and B50O22), all of which matched those detected from hybridization (data now shown). Screening of the multiplex system with yet another marker, BC523SCAR, with which 1,000 hybridization signals had been detected, identified nine positives (Table 2). The procedure for screening DNA pools for these positives is shown in Figs. 4 and 5. These results demonstrate that PCR screening of the BAC library DNA pools effectively reduced the number of positive BACs that presumably resulted from repeat element hybridization with the two SCAR markers. However, PCR screening with three other RAPD markers (BC515, OP-AD12-1, and OP-AI10) whose hybridization had also identified an abundance of positive BAC clones was unsuccessful.

Discussion

Two cucumber BAC libraries were constructed in this study. These two libraries comprise a total of 2,411 Mbp (*Bam*HI library) and 1,578 Mbp (*Eco*RI library) of nuclear DNA, equivalent to approximately 6.5 (*Bam*HI library) and 4.3 (*Eco*RI library) haploid genomes of cucumber. The probability of finding any particular sequence from the cucumber genome represented at least once in the *Bam*HI or *Eco*RI library is greater than 99% and 95%, respectively (Zhang et al. 1996). Screening of the libraries with nuclear gene-specific or RFLP probes resulted in the identification of 4.8 (*Bam*HI library) and 3.7 (*Eco*RI library) clones per probe on average. Screening the libraries with QTL-linked markers gave averages of 5.7 *Bam*HI and 7.1 *Eco*RI clones per probe, when considering results from seven of the nine markers tested (excluding AJ6SCAR and BC523SCAR). Hybridization and/or PCR screening taken collectively identified 5.2 clones per probe for each library, which was slightly smaller (*Bam*HI library) or larger (*Eco*RI library) than the values obtained from empirical characterization of these libraries.

One of the most technically difficult parts of BAC library construction is the isolation of high-quality HMW DNA. For this purpose, we isolated the nuclei and embedded them in agarose plugs. Compared with the previously used agarose microbead method (Nam et al. 1999), this method has been effective in embedding abundant amounts of DNA to a unit quantity of agarose (Zhang et al. 1996). We found the plug method to be particularly suitable for cucumber, because its relatively small genome size frequently resulted in a concentration of final size-selected DNA fragments too low for a successful ligation reaction. We carried out double size selection of partially digested HMW DNA to overcome the problem of small DNA fragments dominating the ligation step, which results in many BAC clones with short inserts. Our results demonstrate that DNA fragments obtained from single size-selection alone resulted in a substantial proportion of ligation products with inserts smaller than 50 kb in size. The fact that the *Eco*RI library contains a relatively high percentage (14.5%) of false-positive clones when compared to the *Bam*HI library (1.3%) is likely a reflection of over-digestion or over-dephosphorylation of the cloning vector.

The majority of our BAC clones contained single *Not*I inserts (Figs. 1, 2). This characteristic is shared by BAC libraries made from dicotyledonous species (Mozo et al. 1998; Budiman et al. 2000; Wu et al. 2004a) and is contrary to those from monocotyledonous species (Woo et al. 1994; Tao et al. 2002). When a blot prepared from gel-separated BAC DNA was hybridized with the C_0t-1 DNA probe, approximately 36% of the library clones were shown to carry moderately or highly repetitive sequences. These estimates are consistent with the relative amounts of repetitive sequences predicted in small diploid genomes (The Arabidopsis Genome Initiative 2000) and with estimates from other plant BAC libraries of comparable genome sizes (Woo et al. 1994; Nam et al. 1999). Accordingly, additional observations suggest that the amount of repetitive DNA sequences is comparatively high in the cucumber genome. Filter hybridization with a SCAR and three RAPD probes gave thousands

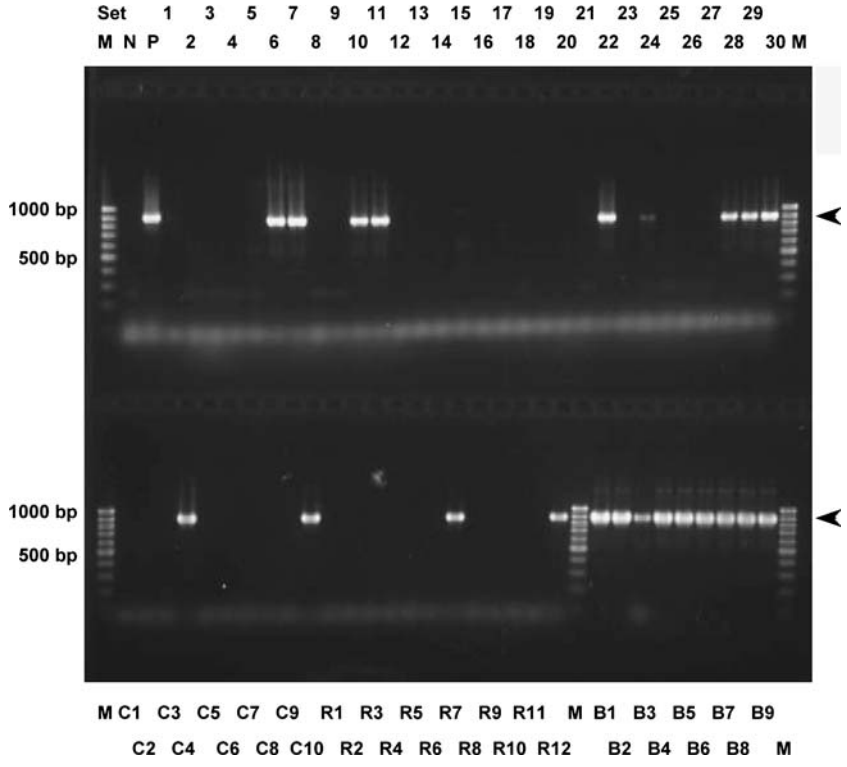

Fig. 5 The polymerase chain reaction (PCR) amplification of BAC DNA pools with primers specific to the BC523SCAR marker. Shown in the agarose gel are amplified fragments from 30 set pools (*upper lanes* 1–30), ten-column (C1–C10) and 12-row (R1–R12) superpools (*lower left lanes*) as well as the final individual BACs (B1–B9, *lower right lanes*). The nine positives from set pools (*upper lanes*) were examined further with the corresponding column and row superpools, of which samples from set no. 11 are shown as an example (*lower left lanes*). Positives from the two-dimensional superpools cross at the 22nd column and the 20th row (see Fig. 4). Because the row pools are numbered consecutively, starting with set no. 1 (1A–1P and 2A–2P), the row pools of set no. 11 are 21A–21P and 22A–22P. Therefore, the final address of the BAC clone is identified as B22O22 (*B* from the *Bam*HI library, *22O* from row 2O in set no. 11, and *22* from column 22). *M* 100-bp DNA size standards, *N* negative control without template DNA, *P* PCR fragment amplified from genomic DNA. *B1* B12P24, *B2* B14K20, *B3* B19M17, *B4* B22O22, *B5* B44D16, *B6* B47L04, *B7* B56O12, *B8* B58D06, *B9* B60N01. *Arrowheads* indicate the expected size (~880 bp) of the PCR products

of positives, presumably due to the presence of repetitive sequences in the probes. Sequence analysis of 281 ends of the BAC clones identified as positive for various probes revealed that approximately 12.5% of the BAC ends had retroelement-like sequences, one of the first types of repetitive DNA recognized in the cucumber genome (K.-H. Song and Y.-W. Nam, unpublished results). These observations support the hypothesis that the cucumber genome is composed of a significant amount of repetitive DNA.

The *F* locus-linked ACC synthase gene was chosen as a marker associated with sex phenotype and yield components (i.e., early flowering and first-harvest yield, Fazio et al. 2003a, 2003b). Earlier studies established that an ACC synthase gene enhanced female sex expression, and no recombinant genotype was found between it and the *F* locus in the segregating population (Trebitsh et al. 1997). More recently, alleles of the *F* locus-linked *CsACS1G* were isolated and compared between different genotypes to show that *CsACS1G* exclusively amplified from the gynoecious phenotype was the most likely candidate for the dominant *F* allele (Mibus and Tatlioglu 2004). Despite these findings, however, it is worth verifying whether the *F*-linked ACC synthase gene and *CsACS1G* indeed represent an identical locus in the genomic context. Initially, filter hybridization with a genomic fragment amplified from a long 3′ exon of *CsACS1G* as the probe identified 125 positive BAC clones from both libraries (data not presented). This result was not surprising, because a family consisting of 10–12 ACC synthase genes is present in other plant species. In *Arabidopsis*, for instance, 12 members of this family are located at separate loci on five chromosomes and share significantly high sequence homology at the amino acid and even at the nucleotide levels (Yamagami et al. 2003). Therefore, the hypothesis that the *CsACS1G* probe most likely hybridized with putative members of the ACC synthase gene family in cucumber is supported by the closeness of the number of positives to the expected genome coverage of the two libraries. Indeed, subsequent PCR analysis identified three BAC clones that carried *CsACS1G* (Table 2). These three clones are considered as physical landmarks linked with the *F* locus on linkage group 1. An effort toward positional cloning of the *F* locus will provide sequence information around the genomic region spanning the putative position of *CsACS1G* to unequivocally confirm the genetic location and identity of the *F* locus.

The effects of the *F* locus and the surrounding region are actually pleiotropic, exerting substantial influences on several yield-associated QTL such as earliness (*ear1.2*), female nodes on the mainstem (*sex1.1*) or primary lateral branches (*fnl1.1*), and mean fruit length:diameter ratio (*ldr1.1*) (Fazio et al. 2003b). A molecular marker linked simultaneously with two of these QTL (*ear1.2* and *ldr1.1*) is CSWCT28, which is 5.0 cM from the *F* locus. Together, *F* and CSWCT28 delimit a chromosomal region on linkage group 1 enriched with genes having various effects on yield components. The BACs isolated with these two markers are expected to provide the physical landmarks with which to initiate a more in-depth molecular analysis of these QTL.

Another notable molecular marker on linkage group 1 is CSWCTT14, which is located 0.8 cM from the *de* locus. This plant-habit locus has been investigated extensively due to its potential impact on early yield and sex stabilization of gynoecy (Serquen et al. 1997b). The region surrounding *de* exerts large pleiotropic effects on MLB (*mlb1.1*), sex expression (*sex1.2* and *fnl1.3*), earliness (*ear1.3*), and comparatively smaller effects on length/diameter ratio (*ldr1.2*) (Fazio et al. 2003a). As a genetic distance of 1 cM equals approximately 200 kb in *Arabidopsis* (The Arabidopsis Genome Initiative 2000), an 0.8-cM distance in cucumber might predictably be equivalent to several hundred kilobases in physical length. This suggests that our BAC clones are positioned at a chromosomal location from which a walk towards the *de* locus could be implemented. Fingerprinting of the positive BACs (Chang et al. 2001) for the CSWCTT14 marker will create overlapping clones spanning this locus. These contigs will assist in more accurately defining this unique genomic region by facilitating development of additional specific markers (Chen et al. 2004) and enhancing the physical mapping and map-based cloning of *F* and *de*.

Screening with markers linked to yield component QTL further characterized these libraries. Filter hybridization of four SSR, two SCAR, and an RAPD marker as probes identified positive BAC clones from both libraries. The relative abundance of positive *Bam*HI or *Eco*RI clones varied, depending on the marker used. For instance, whereas the M8SCAR marker detected more positive *Bam*HI than *Eco*RI BACs, the CSWCTT14, NR60, and OP-W7-1 markers revealed fewer *Bam*HI than *Eco*RI BACs. In contrast, the *Bam*HI- and *Eco*RI-positive BACs for the CSWCT28 and CSWTAAA01 markers were obtained in predicted ratios. This apparent discrepancy likely reflects the different accessibility of restriction enzymes to a particular genomic region (Wu et al. 2004a) and demonstrates the advantages of employing two complementary libraries when attempting to cover the cucumber genome. As the amounts of contaminating chloroplast and mitochondrial DNA were lower than 1%, the combined coverage of approximately 10.8 haploid genomes is essentially equivalent to the minimal size (10–12 haploid genomes) of a BAC library required for successful physical mapping of a genome (Zhang et al. 1996; Xu et al. 2004). Previous studies demonstrated that alternative DNA fragments increased the capacity of genome coverage significantly (Tao et al. 2002; Chen et al. 2004; Wu et al. 2004a). Therefore, the two libraries described herein are expected to adequately cover the cucumber genome as they are applied to fine mapping of QTL and physical mapping for chromosome localization.

Most of the recently developed molecular markers of cucumber are PCR-based such as RAPD, SCAR, SNP, AFLP, and SSR (Bradeen et al. 2001; Fazio et al. 2002, 2003b). Thus, a capability for efficient screening of BAC clones with these markers would be crucial to positional cloning strategies. Therefore, the development of the PCR-based screening system described herein has accommodated the needs for rapid and reliable screening of the cucumber BAC libraries. When the system was used to screen the *Bam*HI library with two primer pairs for *CsACS1G* and peroxidase genes per se, results were obtained consistent with those from filter hybridization (data not presented). The true value of the BAC library DNA pool system was, nevertheless, demonstrated when BACs were amplified with BC523SCAR, a marker closely linked to QTL for earliness (*ear1.4*) and mean fruit length:diameter ratio (*ldr1.3*), with which thousands of BACs had been identified by hybridization. The PCR screening of the BAC library that obviates the use of a lengthy DNA probe could effectively circumvent this problem, because one of the likely reasons, presumably, for the emergence of such a large number of positives is the presence of unknown repetitive sequences included in the probe DNA. However, PCR screening for BACs with three other RAPD markers was unsuccessful. These results appear to be due to low specificity of the amplification conditions used for the RAPD markers. We found it difficult to identify polymorphic bands from several unspecific amplification fragments co-migrating on an agarose gel. In addition, the sizes of the polymorphic fragments excised from the gel and used as probes were relatively large (650–1,150 bp) compared with other markers such as SSRs (170–306 bp) (Table 1). This might have aggravated the probability that repetitive sequences are included in the amplified DNA fragments. Therefore, continued conversion of important RAPD markers to sequence-specific SCARs will be necessary not only for obtaining codominant markers for mapping, but also for effectively screening the BAC libraries with the multiplex system.

The five markers, AJ6SCAR, CSWCTT14, CSWTAAA01, L19-2-SCAR, and M8SCAR, are individually linked with QTL for the number of lateral branches in three different linkage groups with different degrees of effects (Table 1). Besides CSWCTT14 that is linked to *mlb1.1* with a comparatively high LOD (11.6), linkage group 1 has three additional QTL for MLB (*mlb1.2*, *mlb1.3*, and *mlb1.4*), of which *mlb1.3* is linked with the AJ6SCAR marker at approximately 5 cM. The CSWTAAA01 marker is linked tightly with *mlb4.4* in

linkage group 4, whereas the M8SCAR and L19-2-SCAR markers are individually linked with *mlb6.1* and *mlb6.2* in linkage group 6, respectively. However, OP-AD12-1, an RAPD marker closely linked with *mlb1.4*, which showed the highest LOD (32.9) (Fazio et al. 2003b), was unsuccessfully used in the BAC screening described herein. As an important metric trait positively correlated with yield increase in cucumber, MLB has been a central target of marker-assisted selection (MAS) (Fazio et al. 2003a). The end sequences of the BACs identified as linked with these QTL are currently being analyzed in order to produce PCR-amplified markers such as SSR (Cregan et al. 1999; Lichtenzveig et al. 2005) and cleaved amplified polymorphic sequences (Michaels and Amasino 1998). This will provide a relatively simple and rapid means to develop specific, strategically based markers for use in MAS for plant improvement.

Acknowledgements We thank Dr. Claude Penel (University of Geneva), Dr. Minkyun Kim (Seoul National University), and Dr. Shinje Kim (FnP) for gene-specific probes. This work was supported by a grant from Korea Research Foundation (KRF-2000-DP0401) to Y.-W.N. and a grant from Korea Science and Engineering Foundation (KOSEF) to the Agricultural Plant Stress Research Center (APSRC) at Chonnam National University (Y.-W.N.). Purchase of key experimental equipment and software programs was supported by Sogang University Research Grants (20021508 and 20041037).

References

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

Bradeen JM, Staub JE, Wye C, Antonise R, Peleman J (2001) Towards an expanded and integrated linkage map of cucumber (*Cucumis sativus* L.). Genome 44:111–119

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

Burke DT, Carle GF, Olson MV (1987) Cloning of large fragments of exogenous DNA into yeast by means of artificial chromosome vectors. Science 236:801–811

Cantliffe DJ (1981) Alteration of sex expression in cucumber due to changes in temperature, light intensity, and photoperiod. J Am Soc Hort Sci 106:133–136

Carpin S, Crevecoeur M, Greppin H, Penel C (1999) Molecular cloning and tissue-specific expression of an anionic peroxidase in zucchini. Plant Physiol 120:799–810

Chang Y-L, Tao Q, Scheuring C, Meksem K, Zhang H-B (2001) An integrated map of *Arabidopsis thaliana* for functional analysis of its genome sequence. Genetics 159:1231–1242

Chen Q, Sun S, Ye Q, McCuine S, Huff E, Zhang H-B (2004) Construction of two BAC libraries from the wild Mexican diploid potato, *Solanum pinnatisectum*, and the identification of clones near the late blight and Colorado potato beetle resistance loci. Theor Appl Genet 108:1002–1009

Choi S, Creelman RA, Mullet JE, Wing RA (1995) Construction and characterization of a bacterial artificial chromosome library of *Arabidopsis thaliana*. Plant Mol Biol Rep 13:124–128

Cregan PB, Mudge J, Fickus EW, Marek LF, Danesh D, Denny R, Shoemaker RC, Matthews BF, Jarvik T, Young ND (1999) Targeted isolation of simple sequence repeat markers through the use of bacterial artificial chromosomes. Theor Appl Genet 98:919–928

Dane F (1991) Cytogenetics in the genus *Cucumis*. In: Tsuchiya T, Gupta PK (eds) Chromosome engineering in plants: genetics, breeding, and evolution. Elsevier, Amsterdam, pp 201–214

Danesh D, Penuela S, Mudge J, Denny RL, Nordstorm H, Martinez JP, Young ND (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

FAO (1993) Year book production 1992. Food and Agriculture Organization of the United Nations, Rome, Italy

Fazio G, Staub JE, Chung S-M (2002) Development and characterization of PCR markers in cucumber. J Am Soc Hort Sci 127:545–557

Fazio G, Chung SM, Staub JE (2003a) Comparative analysis of response to phenotypic and marker-assisted selection for multiple lateral branching in cucumber (*Cucumis sativus* L.). Theor Appl Genet 107:875–883

Fazio G, Staub JE, Stevens MR (2003b) Genetic mapping and QTL analysis of horticultural traits in cucumber (*Cucumis sativus* L.) using recombinant inbred lines. Theor Appl Genet 107:864–874

Frary A, Nesbitt TC, Grandillo S, Knaap E, Cong B, Liu J, Meller J, Elber R, Alpert KB, Tanksley SD (2000) *fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size. Science 289:85–88

Frijters ACJ, Zhang Z, van Damme M, Wang G-L, Roland PC, Michelmore RW (1997) Construction of a bacterial chromosome library containing large *Eco*RI and *Hin*dIII genomic fragments of lettuce. Theor Appl Genet 94:390–399

Grant S, Houben A, Vyskot B, Siroky J, Pan WH, Macas J, Saedler H (1994) Genetics of sex determination in flowering plants. Dev Genet 15:214–230

Jeffrey C (1980) A review of the Cucurbitaceae. Bot J Linn Soc 81:233–247

Kamachi S, Mizusawa H, Matsuura S, Sakai S (2000) Expression of two 1-aminocyclopropane-1-carboxylate synthase genes, *CS-ACS1* and *CS-ACS2*, correlated with sex phenotypes in cucumber plants (*Cucumis sativus* L.). Plant Biotechnol 17:69–74

Kieber JJ, Rothenberg M, Roman G, Feldmann KA, Ecker JR (1993) *CTR1*, a negative regulator of the ethylene response pathway in *Arabidopsis*, encodes a member of the raf family of protein kinases. Cell 72:427–441

van Leeuwen H, Monfort A, Zhang H-B, Puigdoménech P (2003) Identification and characterisation of a melon genomic region containing a resistance gene cluster from a constructed BAC library. Microcolinearity between *Cucumis melo* and *Arabidopsis thaliana*. Plant Mol Biol 51:703–718

Li J, Nagpal P, Vitart V, MacMorris TC, Chory J (1996) A role for brassinosteroids in light-dependent development of *Arabidopsis*. Science 272:398–401

Lichtenzveig J, Scheuring C, Dodge J, Abbo S, Zhang H-B (2005) Construction of BAC and BIBAC libraries and their applications for generation of SSR markers for genome analysis of chickpea, *Cicer arietinum* L. Theor Appl Genet 110:492–510

Lower RL, Edwards MD (1986) Cucumber breeding. In: Bassett MJ (ed) Breeding vegetable crops. AVI, Westport

Luo M, Wang Y, Frisch D, Joobeur T, Wing RA, Dean RA (2001) Melon bacterial artificial chromosome (BAC) library construction using an improved method and identification of clones linked to the locus conferring resistance to melon Fusarium wilt (*Fom-2*). Genome 44:154–162

Malepszy S, Niemirowicz-Szczytt K (1991) Sex determination in cucumber (*Cucumis sativus*) as a model system for molecular biology. Plant Sci 80:39–47

Mibus H, Tatlioglu T (2004) Molecular characterization and isolation of the *F/f* gene for femaleness in cucumber (*Cucumis sativus* L.) Theor Appl Genet 109:1669–1676

Michaels SD, Amasino RM (1998) A robust method for detecting single nucleotide changes as polymorphic markers by PCR. Plant J 14:381–385

Mozo T, Fischer S, Shizuya H, Altmann T (1998) Construction and characterization of the IGF *Arabidopsis* BAC library. Mol Gen Genet 258:562–570

Nam YW, Penmetsa RV, Endre G, Uribe P, Kim D, Cook DR (1999) Construction of a bacterial artificial chromosome library of *Medicago truncatula* and identification of clones containing ethylene-response genes. Theor Appl Genet 98:638–646

Serquen FC, Bacher J, Staub JE (1997a) Genetic analysis of yield components in cucumber at low plant density. J Am Soc Hort Sci 122:522–528

Serquen FC, Bacher J, Staub JE (1997b) Mapping and QTL analysis of horticultural traits in a narrow cross in cucumber (*Cucumis sativus* L.) using random-amplified polymorphic DNA markers. Mol Breed 3:257–268

Shizuya H, Birren B, Kim U-J, Mancino V, Slepak T, Tachini Y, Simon M (1992) Cloning and stable maintenance of 300 kilo base pair fragment of human DNA in *Escherichia coli* using an *F*-factor based vector. Proc Natl Acad Sci USA 89:8794–8797

Song WY, Wang GL, Chen LL, Kim HS, Pi LY, Holsten T, Gardner J, Wang B, Zhai WX, Zhu LH, Franquet C, Ronald PC (1995) A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. Science 270:1804–1806

Staub JE, Serquen FC, Gupta M (1996) Genetic markers, map construction, and their application in plant breeding. Hort Sci 31:729–741

Tao Q, Zhang H-B (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, Wang A, Zhang H-B (2002) One large-insert plant-transformation-competent BIBAC library and three BAC libraries of Japonica rice for genome research in rice and other grasses. Theor Appl Genet 105:1058–1066

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

Trebitsh T, Staub JE, O'Neill SD (1997) Identification of a 1-aminocyclopropane-1-carboxylic acid synthase gene linked to the *Female* (*F*) locus that enhances female sex expression in cucumber. Plant Physiol 113:987–995

USDA-NASS (1999) Vegetables: final estimates by States, 1992–97. US Department of Agriculture, Agricultural Statistics Service, Statistical Bulletin No. 946c

Venter JC, Smith HO, Hood L (1996) A new strategy for genome sequencing. Nature 381:364–366

Woo S-S, Jiang J, Grill BS, Paterson AH, Wing RA (1994) Construction and characterization of a bacterial artificial chromosome library of *Sorghum bicolor*. Nucleic Acids Res 22:4922–4931

Wu C, Nimmakayala P, Santos FA, Springman R, Tao Q, Meksem K, Lightfoot DA, Zhang H-B (2004a) Construction and characterization of a soybean bacterial artificial chromosome library and use of multiple complementary libraries for genome physical mapping. Theor Appl Genet 109:1041–1050

Wu C, Sun S, Padmavathi N, Santos FA, Springman R, Meksem K, Lightfoot D, Zhang H-B (2004b) A BAC and BIBAC-based physical map of the soybean genome. Genome Res 14:319–326

Xu ZY, Sun S, Covaleda L, Ding K, Zhang A, Scheuring C, Zhang H-B (2004) Genome physical mapping with large-insert bacterial clones by fingerprint analysis: methodologies, source clone genome coverage and contig map quality. Genomics 84:941–951

Yamagami T, Tsuchisaka A, Yamada K, Haddon WF, Harden LA, Theologis A (2003) Biochemical diversity among the 1-amino-cyclopropane-1-carboxylate synthase isozymes encoded by the Arabidopsis gene family. J Biol Chem 278:49102–49112

Yoo EY, Kim S, Kim YH, Lee CJ, Kim B-D (2003) Construction of a deep coverage BAC library from *Capsicum annuum*, 'CM334'. Theor Appl Genet 107:540–543

Zhang H-B (2000) Construction and manipulation of large-insert bacterial clone libraries manual. Texas A&M University, College Station (available at http://hbz.tamu.edu)

Zhang H-B, Zhao X, Ding X, Peterson AH, Wing RA (1995) Preparation of megabase-size DNA from plant nuclei. Plant J 7:175–184

Zhang H-B, Woo S-S, Wing RA (1996) BAC, YAC and cosmid library construction. In: Foster G, Twell D (eds) Plant gene isolation: principles and practices. Wiley, UK, pp 75–99

Zhang H-B, Wu C (2001) BAC as tools for genome sequencing. Plant Physiol Biochem 39:195–209

Zwick MS, Hanson RE, McKnight TD, Islam-Faridi MN, Stelly DM, Wing RA, Price HJ (1997) A rapid procedure for the isolation of C_0t-1 DNA from plants. Genome 40:138–142