

Research article

THE CONSTRUCTION AND CHARACTERISTICS OF A BAC LIBRARY FOR *Cucumis sativus* L. 'B10'

WOJCIECH GUTMAN^{1*}, MAGDALENA PAWEŁKOWICZ², RAFAŁ WOYCICKI², EWA PISZCZEK¹ and ZBIGNIEW PRZYBECKI²

¹Department of Biochemistry, Warsaw Agricultural University, 02-776 Warsaw

²Department of Genetics, Breeding and Biotechnology of Plant, Warsaw Agricultural University, 02-776 Warsaw

Abstract: Cloning using bacterial artificial chromosomes (BACs) can yield high quality genomic libraries, which are used for the physical mapping, identification and isolation of genes, and for gene sequencing. A BAC genomic library was constructed from high molecular weight DNA (HMW DNA) obtained from nuclei of the cucumber (*Cucumis sativus* L. cv. Borszczagowski; B10 line). The DNA was digested with the *Hind*III restriction enzyme and ligated into the pCC1BAC vector. The library consists of 34,560 BAC clones with an average insert size of 135 kb, and 12.7x genome coverage. Screening the library for chloroplast and mitochondrial DNA content indicated an exceptionally low 0.26% contamination with chloroplast DNA and 0.3% with mitochondrial DNA.

Key words: *Cucumis sativus* L., Bacterial artificial chromosome, High molecular weight DNA, Pulsed field gel electrophoresis

* Author for correspondence; e-mail: wojciech_gutman@sggw.pl

Abbreviations used: BAC – bacterial artificial chromosome; BSA – albumin from bovine serum; DAPI – 4'-6'-Diamidino-2-phenylindole dihydrochloride; HMW DNA – high molecular weight DNA; LMP – low melting agarose; PCR – polymerase chain reaction; PFGE – pulsed field gel electrophoresis; PVP – polyvinylpyrrolidone; SCAR – sequence-characterised amplified regions

INTRODUCTION

Cucumber (*Cucumis sativus* L.) belongs to the cucurbit family (*Cucurbitaceae*), which also includes melon (*C. melo* L.), watermelon (*Citrullus lanatus*), field pumpkin (*Cucurbita pepo* L.) and pumpkin (*Cucurbita maxima*). From a genetic point of view, cucumber shows the most similarity to melon (*C. melo* L.), which is in the same genus, although the two species differ in their chromosome number (*C. sativus* L. $2n = 14$; *C. melo* L. $2n = 24$). The commonly accepted cucumber genome size is 367 Mkb/C and 1000 cM [1], and the karyotype consists of seven relatively small chromosomes [2]. Cucumber can be monoecious or dioecious, as determined by a simple configuration of four non-allelic cooperating main sex-determining genes [3].

The ongoing molecular tests on cucumber being conducted in various centres worldwide are mainly focused on the determination of the genome structure (including the nuclear and organelle genomes), and more recently on the determination of the transcriptome [1-2]. There are two existing constructed BAC libraries for cucumber: *Bam*HI and *Eco*RI [4].

Libraries containing long DNA fragments (ranging from several hundred to two thousand kbp) are the tools used to conduct genomics research, especially in the field of structural genomics. The artificial bacterial chromosome (BAC) constructed in 1992 by Shizuya [5] was the first bacterial vector able to maintain and reproduce long inserts (> 100 kb) of DNA in *E. coli* cells. The BAC vector cloning system has become an invaluable tool in genomics studies thanks to its ability to stably maintain large DNA fragments and its ease of manipulation. BAC vectors allow high quality genomic libraries with long inserts to be obtained, which can then be used for different purposes, like physical mapping, DNA fingerprinting, sequencing, and gene identification or isolation [6-7].

BACs are characterised by high stability (a low percentage of recombinations), and are therefore becoming an increasingly popular tool in the construction of microbiological libraries, and plant and animal genomes. This paper deals with the construction and characteristics of a cucumber (B10 line) BAC library.

MATERIALS AND METHODS

Plant material

To construct the library, we used DNA from the monoecious, highly homozygotic cucumber line B10, Borszczagowski variant. It had the following genotype with regard to sex-determining genes: *ff/MM/GyGy*. Seeds were obtained in 2003, from the collection of the Department of Genetics, Breeding and Biotechnology of Plant SGGW in Warsaw.

In 2004, the plants were grown in a greenhouse under controlled photoperiod conditions: 16 h/8 h day/night, at 25-27°C during the day and 18-20°C at night. The light intensity was 1500 μmol (quantum) $\text{m}^{-2}\text{s}^{-1}$. In order to isolate DNA, leaves from 4- to 6-week old plants were collected, packed in aluminium foil, and then frozen in liquid nitrogen and stored at -75°C.

Library construction

Isolation of nucleic and high molecular weight DNA

The library was constructed in several stages in accordance with a highly modified version of Peterson's procedure [8]. The first stage was the isolation of cell nuclei from the leaf tissue, and their melting and immobilisation in LMP agarose blocks. The solidified blocks were incubated in a lysis buffer (1% w/v N-lauroylsarcosine sodium salt, 0.1 mg/ml proteinase K, 0.1% w/v ascorbic acid, 2% w/v PVP-40, 0.13% w/v sodium diethyldithiocarbamate) and transferred to -20°C.

In order to do that, 30 g of frozen leaf tissue was homogenised in liquid nitrogen in a mortar and pestle until a fine powder was obtained. The powder was transferred to a flask containing 300 ml of prechilled extraction buffer (500 mM sucrose, 4 mM spermidine, 1 mM spermine tetrachloride, 0.1% w/v ascorbic acid, 2% w/v PVP-40, 0.13% w/v sodium diethyldithiocarbamate). The mixture was incubated on ice for 40 minutes while being gently stirred, and then filtered through four layers of sterile cheesecloth and a nylon mesh filter with a pore diameter of approximately 40 µm (Medlab, PL). Fifteen milliliters of prechilled extraction buffer with 10% v/v Triton X-100 was added to the filtrate and the mixture was incubated on ice for 20 minutes. The whole volume (approx. 300 ml) was poured into 50-ml test tubes and centrifuged for 15 minutes at 2250 rpm (Beckman CPR 90E2891) and 4°C. The supernatant was carefully poured off and the remaining pellets were suspended with a sterile brush in 10 ml of prechilled extraction buffer. Then, with gentle stirring, prechilled extraction buffer was added to each sample up to a volume of 50 ml, and the mixture was again centrifuged for 15 minutes at 2250 rpm and 4°C. The supernatant was discarded and the remaining pellets were suspended with a brush in 10 ml of prechilled extraction buffer. The resuspensions were combined into four tubes. The mixture was supplemented with prechilled extraction buffer to a volume of 50 ml, and centrifuged for 15 minutes at 2250 rpm and 4°C. The procedure of centrifugation was confirmed until the whole pellet could be placed in one test tube. After the last centrifugation, the supernatant was carefully discarded, leaving 2 cm of the solution over the cell nuclei. The pellet was gently resuspended with a brush in the remaining buffer (homogenous suspension was attained) and stored on ice. In order to check the isolate quality, 50-100 µl of the mixture was collected and placed on a microscope slide, and mixed with an equal volume of DAPI staining agent. The observations were performed using a fluorescence microscope.

The isolated cell nuclei were immobilised in LMP-agarose blocks (Bio-Rad, USA). In order to do that, the nucleus suspension was mixed in an equal volume with 1.5% LMP-agarose melted at 70°C, and then the wells in the template were filled in (approx. 100 µl/well) (Bio-Rad, USA). The filled templates were incubated for 30 minutes at a temperature of 4°C, and solidified LMP blocks with the cell nuclei were taken from the template and transferred to the lysis buffer

(1% w/v N-lauroylsarcosine (sodium salt), 0.1 mg/ml proteinase K, 0.1% w/v ascorbic acid, 2% w/v PVP-40, 0.13% w/v sodium diethyldithiocarbamate dissolved in 0.5 M EDTA, pH 9.1) and incubated at 50°C for 72 hours, exchanging the buffer every 24 hours. The blocks were then placed in fresh lysis buffer and incubated at 4°C overnight. After the incubation, the blocks were put into 70% v/v ethanol, incubated for 90 minutes on ice, and then stored at -20°C.

Digestion of genomic DNA

Isolated HMW DNA was partially digested by the *Hind*III restriction enzyme. The following enzyme concentrations were used (20,000 U/ml): 0.5 U/5 μ l; 1 U/5 μ l; 2 U/5 μ l; 4 U/5 μ l; and 6 U/5 μ l. Blocks with DNA were divided into halves and then each half (approx. 50 μ l) was macerated with a sterile microscope slide and digested with the *Hind*III restriction enzyme. The composition of the reaction mixture for one sample was 10X buffer *Hind*III, 10 mg/ml BSA, 40 mM spermidine, ½ of DNA block, and *Hind*III restriction enzyme [2 U]. The mixtures (without the restriction enzyme) were incubated on ice for 1 hour. After the incubation, 2 U of the *Hind*III restriction enzyme was added to each sample. These were then incubated at 37°C and digested for 10 minutes. The reaction was stopped by adding 10 μ l of 0.5 M EDTA, pH 8.0. Partially digested DNA was separated with PFGE electrophoresis. The Lambda Ladder (New England Biolabs, USA) was used as the size marker. The electrophoresis was run under the following conditions: voltage 6 V/cm, initial switch time 1.0 s, final switch time 40.0 s, angle 120°, temperature 12°C, separation time 18 hours [8].

First size selection

As the DNA fragments for the next stages of library construction should be 120-140 kb long, a DNA selection was performed after the separation was performed. After electrophoresis, the gel was placed on a sterile tray and gel strips were cut from the left and right so that each strip would contain a size marker and a narrow strip with plant DNA. These gel strips were stained in a solution of ethidium bromide (10 mg/ml) for 40 minutes and washed with sterile water. Stained gel fragments were viewed with a transilluminator emitting UV light of wavelength: $\lambda = 312$ nm, and the region of gel containing the DNA of the required size (120 kb to 400 kb) was marked. Then the gel was reconstructed and the marked region from the non-stained gel was cut out. The cut fragment was used for the second size selection.

Second size selection

The gel fragments obtained during the first selection were placed in 1% w/v agarose gel in 0.5xTBE and immersed in 1.5% w/v LMP agarose, and air bubbles from the connection sites between the gel edges were eliminated. The Lambda Ladder (New England Biolabs, USA) marker was placed in the lateral wells. The gel was placed in a pulsed field gel electrophoresis chamber containing fresh buffer 0.5xTBE, and the electrophoresis was run under the

following conditions: voltage 6 V/cm, initial switch time 3.0 s, final switch time 5.0 s, angle 120°, temperature 12°C, separation time 18 hours [8]. After the electrophoresis, the selection was conducted as described for the first size selection. The obtained gel fragments were placed in sterile buffer 0.5xTBE and stored at 4°C.

Isolation of DNA from the gel

DNA was isolated from the gel via electroelution. Prior to electroelution, dialysis bags (Sigma, USA) were prepared as in Strong [29]. The bags were cut into 10-cm fragments and placed in a solution containing 1 mM EDTA, pH 8.0, and 2% NaHCO₃, and heated at 90°C in a water bath for 10 minutes. Then they were placed in sterile water and boiled for another 10 minutes. After boiling, they were washed five times in sterile water, placed in 50% ethanol and stored at 4°C.

The gel fragments were placed in a dialysis bag from which all the air bubbles had been eliminated carefully. 500 µl of 1xTAE buffer was added, and the bags were closed with plastic clamps. The bags were placed in a horizontal electrophoresis apparatus with the gel parallel to the electrodes, and immersed in 1xTAE buffer. The electroelution was run for 2 hours at 135 V (4.5 V/cm). After 2 hours, the direction of the electrical field was reversed for 2 minutes in order to unstick DNA from the walls of the dialysis bags. After electroelution, the DNA was transferred to 1.5-ml tubes and placed on ice. To make sure no DNA was left in the gel, it was stained in a solution of ethidium bromide and checked under UV light.

DNA concentration after elution

10% w/v PEG solution (Sigma, USA) and nitrocellulose membranes with 0.025-µm pores (Millipore, USA) were used to condense DNA. 10% w/v PEG (Sigma, USA) was poured into a sterile Petri dish and then the membrane was placed on the surface of the solution with the glossy side upwards. With a wide tip pipette, a DNA drop (500 µl) was placed on the membrane. The dish was covered and left for about 5 hours at a temperature of 4°C. Post-concentration, 200-300 µl DNA was collected with a wide tip pipette and the concentration efficiency was checked with agarose gel electrophoresis. DNA prepared in this way was used for ligation.

DNA ligation to the vector and transfection of E. coli cells

The next stage in the process of library construction was DNA ligation to the pCC1BAC vector (Epicentre, USA) and electroporation to *E. coli* cells of the strain ElectroMAXDH10B (Invitrogen, USA). The ligation was performed in accordance with the procedure of Epicentre, USA, with some modifications. They included a correct adjustment of the (molar) ratio between the insert and the vector, and the period of ligation. For optimisation, different amounts of the insert DNA were used (100 ng/5 µl – 20 ng/µl): 30 ng, 100 ng, 120 ng, 200 ng and 250 ng (the amount of the vector was constant – 25 ng). The volume of the

reaction mixture was lowered to 50 μ l. Ligation was performed at 15°C overnight in a water bath, and then the mixture was transferred for 1 hour to a temperature of 4°C. After incubation, the ligase was inactivated for 15 minutes at 65°C. Finally, the ligation mixture was desalted in accordance with Peterson [8]. In order to do that, agarose blocks containing 0.45% w/v of glucose were used. 50 ml sterile water containing 0.9 g glucose and 0.5 g agarose were heated up to 100°C in order to dissolve the ingredients. Then the mixture was cooled down to 50°C and transferred in 800- μ l aliquots to 1.5-ml tubes. Another set of smaller (0.5 ml) tubes was placed in the filled tubes in order to make wells during agarose solidification. The ligation mixture (50 μ l) was placed in the formed wells and then incubated for 90 minutes on ice. After incubation, the mixture was transferred to new tubes and placed at 4°C. The mixture was then ready for the transformation of *E. coli* cells.

Before starting electroporation, the competent cells of *E. coli* DH10B (Invitrogen, USA) were placed on ice. 2 μ l of the ligation mixture was pipetted into tubes and placed on ice, then 20 μ l of *E. coli* cells were added and gently mixed with a wide tip pipette. The mixture was transferred to prechilled electroporation cuvettes with a distance between the electrodes of 0.1 cm (Bio-Rad, USA) and stored on ice. The cuvettes were placed in an *E. coli* pulser (Bio-Rad, USA). Electroporation was performed at 1.8 kV. After electroporation, 1 ml of the SOC medium (2% w/v Peptone, 1.5% w/v yeast extract, 100 mM NaCl, 25 mM KCl, 100 mM MgSO₄, 200 mM glucose) was added to the mixture, which was then gently mixed and transferred to 15-ml tubes and incubated at 37°C for one hour. After incubation, the transformed bacteria were cultured in 100 μ l on Petri dishes with solid LB medium containing 12.5 μ g/ml of chloramphenicol, 50 μ g/ml X-Gal (5-bromo-4-chloro-3-indolyl- β -D-galactoside) and 25 μ g/ml IPTG (Isopropyl β -D-1-thiogalactopyranoside). The dishes were stored in an incubator at 37°C overnight.

Selection and replication of the library

After the overnight incubation, the bacterial colonies (clones) were individually transferred to 384-well microplates (Nunc, USA) containing FM medium (LB, 13 mM KH₂PO₄, 36 mM K₂HPO₄, 1.7 mM sodium citrate, 6.8 mM (NH₄)₂SO₄, 4.4% v/v glycerol) with sterile toothpicks. Then, the plates with clones were placed in an incubator at 37°C and incubated overnight. After incubation, replicas of the plates were done using a manual (sterile) 384-needle replicator (Nunc, USA) and incubated as above. The plates were stored in a deep freezer at -85°C. By this method, three copies of the library were created: the master copy, QI, and two copies, QII and QIII.

Library characteristics

The characteristics determined were: the number of clones in the library, the average clone length, and the participation of clones containing organellar DNA, and the number of clones with known cucumber marker sequences. Pools of the library were collected and screened by PCR, and hybridisation with radioactively labelled probes was performed. For pool preparation, each of the microplates was grown on a 14-cm plate with LB agarose medium supplemented with chloramphenicol. The grown colonies of each plate were collected and used for DNA isolation.

Determination of the length of inserts in the library

In order to estimate the length of inserts in the vectors, recombinant BAC DNA was isolated from a bacterial culture and then digested with the *NotI* restriction enzyme. A pulse-field gel electrophoresis was run.

Determination of the number of BAC clones containing organellar DNA by PCR

In order to find the BAC clones containing chloroplast or mitochondrial DNA, PCR was performed with primers for the chloroplast and mitochondrial genes. The primers for the chloroplast gene *ycf4* (Acc. AJ970307 – forward *30rf10:5'*CCGTTTTATCCCTCTTGCTAC3' and reverse *40rf10:5'*CAGTTATCTTGGTAGGGATTTGATA3) and the forward/reverse primers of the *Cox* mitochondrial gene (Acc. AJ223416) were used. The template for the reaction was DNA isolated from the pools of the library and 100 randomly chosen clones of the BAC library. The clones were collected with sterile toothpicks and each was placed in a tube containing the reaction mixture. After 5 minutes, the toothpicks were removed and PCR was performed.

BAC library screening for SCAR markers

PCR with DNA from the collective pools and chosen bacterial clones as the template was performed with the *Cucumis sativus* SCAR marker primers from *Gy* gene (*CSMSG_7PSE700a*).

Screening the library by hybridisation

Preparing the membranes for dot-blot hybridisation

One microliter (10 ng) of DNA from the collective pools of BAC was placed on Hybond N⁺ membranes (Amersham, USA) and denaturated by the addition of 1 µl 0.6 M NaOH. Then the mixture was neutralized in 0.5 M Tris-HCL (pH 7.5) for 4 minutes and washed in sterile water. The membranes were blocked for 2 hours at 80°C according to [9].

Preparing the membranes for colony hybridisation

Imprints of the microplates were done on solid LB medium with the addition of 12.5 µg/ml chloramphenicol and with a manual replicator (Nunc, USA). The next day, as per [9], Hybond N⁺ (Amersham, USA) membranes were placed on

the plates containing bacterial clones for 10 minutes. After the membranes had been taken from the plates, they were treated with the following solutions: the denaturing solution (0.5 M NaOH; 1 M NaCl) for 10 minutes; the neutralising solution (0.5 M Tris-HCL; 3 M NaCl) for 10 minutes; and 2xSSPE for the next 10 minutes. After this treatment, the membranes were backed at 80°C for 2 hours.

Probe preparation

The SCAR markers from the cucumber lines B10 and 2 gg were used to construct the probes. Probe labelling was performed according to the procedure of Ready-To-Go DNA Labelling Beads (-dCTP) (Amersham, USA). 25 ng DNA and 50 μ Ci [α -³²P]dCTP (6,000 Ci/mmol) of the isotope were taken for labelling.

Hybridisation

Prehybridisation was performed [9] for 2 hours in a solution of 5xSSC, 1% w/v SDS and 5x Denhardt with the addition of 100 μ g/ml of denaturated DNA from herring sperm (Sigma, USA). Hybridisation was overnight at 68°C in a solution of the same composition but with the addition of a probe. After hybridisation, the membranes were washed out for 10 minutes at 68°C in 2xSSC and 0.1% w/v SDS, and then twice at room temperature for 15 minutes [9]. Visualisations of the signals after hybridisation were documented with the FX Molecular Imager system and Quantity One software (BioRad, USA). The exposure time was 12 hours.

RESULTS AND DISCUSSION

Isolation of high molecular weight DNA

The construction of a BAC library requires good quality high molecular weight DNA (HMW DNA). However, working with fragments longer than 30 kb requires extreme care. Pipetting and phenolisation cannot be employed, as both may cause mechanical fragmentation of such long DNA molecules. Therefore, DNA is isolated from nuclei or protoplasts stabilised in agarose [10]. In this study, we chose to isolate the cell nuclei from leaf tissue, instead of isolation from protoplasts since the last scale is time-consuming, expensive and not always effective. Moreover, HMW DNA isolated from protoplasts contains a significantly higher amount of DNA of chloroplast or mitochondrial origin [11-13], which may have a negative effect on the library quality.

The method designed for the isolation of high purity nuclear DNA from tomato [14, 15] was used. This method was also used for the isolation of HMW DNA during the construction of BAC libraries for vine (*Vitis vinifera* L.) and cotton (*Gossypium hirsutum* L.) [16]. The first stage of isolation is the homogenisation of fresh tissues pre-treated with ether in order to eliminate plant waxes which disturb the process of DNA isolation. Treatment of fresh leaf tissue with ether not only eliminates waxes from the leaf surfaces but also makes the cell nuclei more fragile, increasing the isolation efficiency [17]. In the case of cucumber

leaf tissue, treatment with ether did result in the elimination of plant waxes: there was a characteristic wax ring on the ether layer. The tissue could then be homogenised with a laboratory homogeniser in the presence of the nuclear extraction buffer, containing 2-methyl-2,4-pentenediol to stabilise the nuclei and protect them against degradation. However, the obtained nucleus preparations were contaminated with cellular fragments from cell degradation (Fig. 1A). Moreover, it turned out that the DNA was significantly degraded, and that DNA fragments ranging from 50 to 100 kb were the most abundant (Fig. 2A).

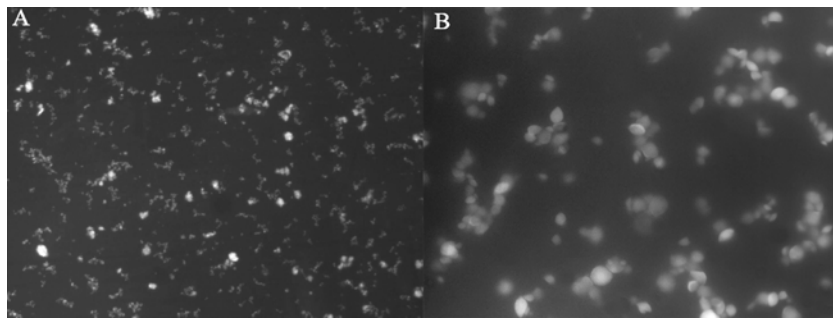


Fig. 1. The cell nuclei of cucumber, isolated after homogenisation. A – isolated from leaves that were mechanically homogenised in TE/ether, B – isolated from leaves that were homogenised with a mortar and pestle in liquid nitrogen without the use of TE/ether. The preparations were stained with DAPI.

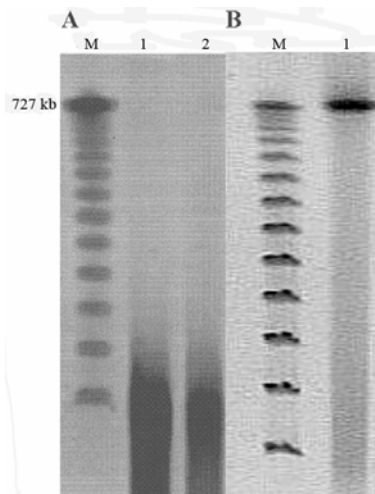


Fig. 2. Pulse-field gel electrophoresis of the isolated cucumber HMW DNA. A – HMW DNA isolated from fresh leaf tissue treated with TE/ether and mechanically homogenised (lane M – Marker Lambda Ladder PFGE, New England Biolabs, USA, lanes 1 and 2 HMW DNA B10). B – HMW DNA isolated from frozen leaf tissue that was homogenised (lane M – Marker Lambda Ladder PFGE, New England Biolabs, USA, lane 1 HMW DNA B10).

This method turned out not to be very effective for the isolation of HMW DNA from cucumber. DNA degradation could have occurred when the collected tissue was placed in TE buffer and during the treatment with ether. The mechanical homogenisation of tissue might also have caused the fragmentation of larger DNA (> 100 kb).

Taking this into account, we made an attempt to isolate the nuclei with Peterson's method [8]. This method has been used to construct the libraries of rye [18], sorgo [13], *Arabidopsis* [19], wheat [20], and barley [21]. The collected leaf tissues are frozen in liquid nitrogen, and then blended in order to obtain a homogenate, which is then placed in the nuclear isolation buffer. A white suspension is obtained, indicating the presence of cell nuclei. The isolated nuclei were checked with a fluorescence microscope (Olympus, USA). Fig. 1B shows a sample microscopic image.

The round structures visible in the photographs are the nuclei, which did differ in their shapes and sizes, with the smaller nuclei being rounder and the larger ones more lenticular. Characteristic aggregates were also observed (Fig. 1B). Thirty grams of cucumber leaf tissue yielded 5 ml of nuclei (approx. 1 000 000 nuclei). This confirms that Peterson's method of isolation can yield a great number of undamaged nuclei (Fig. 1B). Pulse-field gel electrophoresis performed after DNA release from the nuclei and deproteinisation yielded an intensive band at 700 kb, representing cucumber genomic DNA, and a weak streak of degraded DNA below (Fig. 2B). This proves the good quality of the isolated HMW DNA. The DNA isolation efficiency was estimated at ~300 µg from 30 g of leaves.

Restriction digestion and selection of HMW DNA size

It is thought that restriction cutting of high molecular weight DNA is the most crucial stage in the construction of a BAC library. The subsequent stages depend on the quality and size of the fragments. The optimal length of DNA for ligation is fragments from 200 to 350 kb in length [5]. The initial stage for restriction cutting is the optimisation of its conditions. Therefore, the following parameters were evaluated: washing out blocks, the concentration of the *Hind*III restriction enzyme [22] and the incubation time of samples with the enzyme [23]. It was determined that the optimal concentration of the *Hind*III restriction enzyme for cucumber DNA is 2 U of the enzyme per sample (Fig. 3, lane 4). Moreover, changes to the incubation time were also introduced: macerated blocks in the digestion buffer were incubated for 1 hour on ice, and after the addition of the restriction enzyme, they were incubated on ice again, for 30 minutes, whereas the proper digestion was performed at 37°C for 10 minutes. The DNA fragments obtained as the result of digestion were almost evenly distributed in the whole range of 120-400 kb on the gel (Fig. 3, lane 4). In order to obtain an adequate amount of DNA to construct the library, the DNA from 20 blocks was digested.

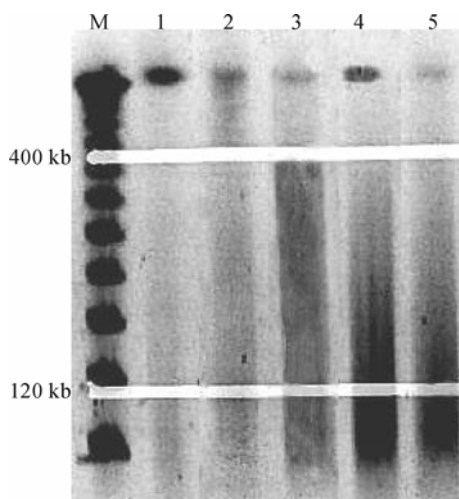


Fig. 3. The optimisation of DNA digestion with the *Hind*III restriction enzyme. The first lane (M) contains the Lambda PFGE marker, and the other lanes contain cucumber (*Cucumis sativus* L.) HMW DNA digested with the *Hind*III restriction enzyme at the following concentrations: lane 1 – 0.5 U/5 µl, lane 2 – 1 U/5 µl, lane 3 – 2 U/5 µl, lane 4 – 4 U/5 µl and lane 5 – 6 U/5 µl.

After digestion, the DNA was submitted to two rounds of size selection, which is required in order to eliminate as many DNA fragments of less than 100 kb as possible; many such fragments are tangled up with longer fragments, and during the electrophoresis separation, they migrate together [8]. As a result of the first selection, a large amount of DNA fragments ranging from 120 to 400 kb were obtained. However, the initial ligation after the first selection and transformation of *E. coli* cells yielded a few clones which included inserts that were too short (approx. 10-20 kb) when checked with the *Not*I restriction enzyme. Therefore, the second size selection was performed. It seems that such procedures were appropriate, as no short fragments were found during the quality analysis of the library clones.

Isolation of DNA from the gel

The ligation efficiency depends to a great extent on the quality of the DNA obtained as a result of the isolation from the gel. We used isolation with electroelution, which seems to be appropriate for isolating long fragments, as it is characterised by the purity of the isolated DNA and by a significantly lower risk of mechanical fragmentation. As a result of electroelution, we obtained high quality DNA in the volume of 500 µl from the first attempt (HMW DNA I) and 500 µl from the second attempt (HMW DNA II). The quality and concentration were checked with typical electrophoresis on a 1% w/v agarose gel (Fig. 4).

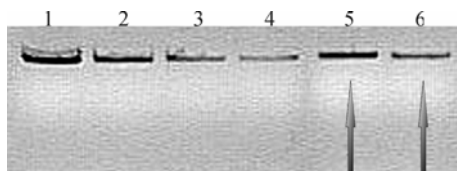


Fig. 4. DNA from cucumber (*Cucumis sativus* L.). Lane 1 – 200 ng/5 μ l, lane 2 – 100 ng/5 μ l, lane 3 – 50 ng/5 μ l and lane 4 – 25 ng/5 μ l. Lanes 5 and 6 contain cucumber DNA after electroelution. Full description in the text.

Lanes 1, 2, 3 and 4 in Fig. 4 each contain the Lambda marker (New England Biolabs, USA) of the determined concentration (respectively 200 ng/5 μ l, 100 ng/5 μ l, 50 ng/5 μ l and 25 ng/5 μ l) allowing the calculation of the amount of isolated DNA. Lanes 5 and 6 (marked with arrows) contain DNA samples after isolation from the second selection. The first sample (lane 5) contained 10 ng/ μ l, and as a result 5 μ g DNA was obtained. The second sample (lane 6) contained 5 ng/ μ l DNA, and as result, 2.5 μ g DNA was obtained. The final DNA concentration was 7.5 μ g/ml.

Ligation of the DNA to the BAC vector and transformation of the bacteria

Before ligation, the BAC vector was linearised with the same restriction enzyme which was used to cut the genomic DNA, and was dephosphorylated. Dephosphorylation prevents self-ligation of the vector. In order to check the cloning system, we performed a series of ligations with the control insert/*Hind*III (33 kb). It turned out that the BAC cloning system acts appropriately, as 70% of the clones obtained were white, and the remaining 30% were blue clones, containing an empty vector. This indicates that the following results were achieved: > 50% white clones and < 40% blue clones, proving that the vector had been prepared correctly and that the ligation conditions were optimised [8]. However, for the ligation of plant DNA, the conditions had to be optimised once more due to the difference in the DNA, which is above all a few times longer than the control. A series of DNA concentrations was tested. The best results were achieved when 200 ng of plant DNA was used with 25 ng of the vector, ligated at 15°C overnight.

After ligation and desalting, the electroporation of the *E. coli* cells was performed. Electroporation was optimised by changing the parameters such as the distances between the electrodes in the cuvette (0.1 and 0.2 cm) and the electrical voltage (1.80 kV and 2.50 kV). In the 0.2 cm cuvettes and at 2.5 kV, the cells were not transfected, possibly due to the electrical voltage being too high. Using the cuvettes, the electrodes 0.1 cm apart and the voltage at 1.80 kV was effective: transformants were obtained. After electroporation, a small amount of the mixture was cultured on Petri dishes in order to check the quality of transformation. If the result was satisfactory, the remainder was cultured on large plates (500 cm²) and incubated at 37°C overnight. Bacterial colonies were

obtained that were then individually transferred to 384-well microplates. 34,560 BAC clones were transferred, and the master copy (QI) of the library was obtained.

In order to determine the size of the library necessary to cover the genome, the Clark-Carbon equation [24] was used: $(N = \ln(1-P)/\ln(1-[I/GS]))$, where: N is the number of clones of the library necessary to cover the genome; P is the probability (99%); GS is the genome size (cucumber 367 Mbp [2]); and I is the insert length. The average insert length in our BAC library is 135 kb.

For this library:

$$N = -4.61/\ln(1-[135000/367000000]) = -4.61/\ln(1-135/367000) = -4.61/\ln 0.99963 = -4.61/-1.607e^{-4} = 28,812.$$

This means that 28,812 BAC clones should be obtained for the library to be 99% representative. For this number of clones, the genome coverage would be about 10x. In reality, the library contains 34,560 clones, giving a genome coverage of 12.7x.

The characteristics of the BAC library and isolation of clones containing molecular markers

The characteristics of the library include parameters such as the number of clones, the number of empty vector (without inserts), the insert length, the range of the insert length, the genome coverage, and the determination of the number of clones containing chloroplast and mitochondrial inserts.

In order to determine the insert length, the isolation of BAC DNA from 100 individual randomly chosen clones was performed, and DNA was digested with the restriction enzyme *NotI* and separated by pulse-field gel electrophoresis (Fig. 5). The first lane contains the size marker and the subsequent lanes contain BAC DNA/*NotI*. At 8 kb, a vector band is visible. The insert length was calculated by adding the lengths of separate bands of a given clone. About 60% of the cucumber BAC library inserts contain two fragments, meaning there is less than one *NotI* cutting site per insert, on average.

However, the lengths of the inserts of 100 randomly chosen clones (Fig. 6) range from < 50 kb to 260 kb. Despite the double selection, there is quite a large number of clones shorter than 50 kb, which correlates with the results obtained by other authors [21, 25-26]. The average length of clones in the library is 135 kb.

In order to make library screening easier, collective pools were formed by combining clones of each 384-well plate and isolations of DNA from the achieved mixtures. Ninety collective pools were obtained (as many as the library plates). Organelle DNA was screened by PCR with primers to the chloroplast gene *ycf4* (Acc. AJ970307). All 90 collective pools were screened, and no signal was identified in any of them. For that reason, the method of hybridisation of collective pools and two single 384-well plates with a radioactively labelled chloroplast probe was used. Two weak signals (+) were observed on the

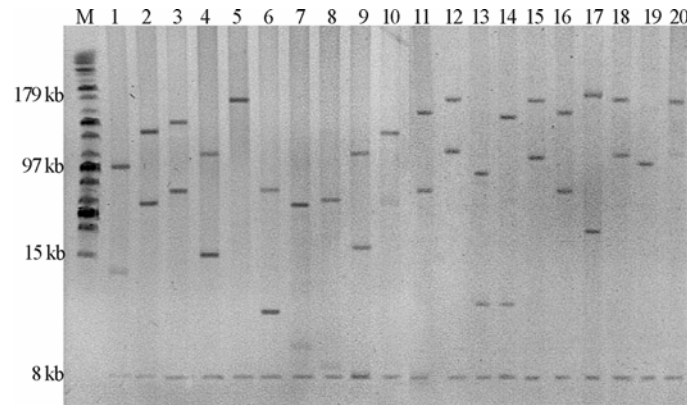


Fig. 5. An analysis of 20 randomly selected cucumber BAC clones after *NotI* restriction enzyme digestion. Lane M – Marker Lambda Ladder PFGE, MidRange I, New England Biolabs, USA. Lanes 1-20 contain cucumber BAC DNA after digestion with the *NotI* restriction enzyme.

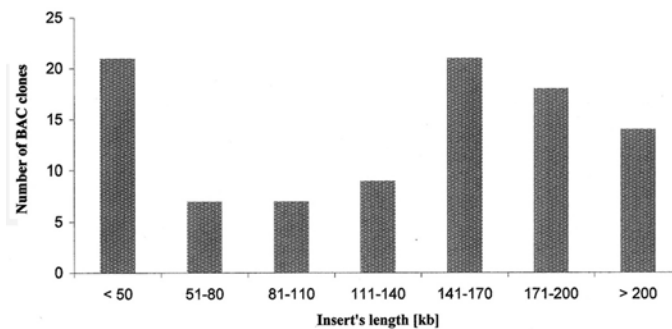


Fig. 6. The participation of separate length classes of inserts from the BAC library of cucumber (*Cucumis sativus* L.) line B10.

membrane containing individual clones, which constitutes just 0.26% of all the clones transferred to the membranes. In a similar way, the number of clones containing mitochondrial DNA was determined. Initially, PCR was used to screen all the collective pools of the library with primers of the mitochondrial gene *Cox* (Acc. AJ223416). The result of PCR was the identification of a signal in a collective pool corresponding to library plate number two. The presence of this band proves that the second plate contained at least one clone with mitochondrial DNA. Hybridisations with a radioactively labelled mitochondrial probe were also performed. Three weak signals (+) were observed on the membrane containing individual clones, which constitutes 0.30% of all the clones transferred to the membranes. However, it is not known whether these sequences are contaminations with organelle DNA during the construction of the library, or whether they were transferred to the nucleus in the process of

organelle-to-nucleus transfer, which frequently occurs [27], and were cloned together with nuclear DNA. The organelle-to-nucleus transfer seems to occur in cucumber [28]. The lower number of clones found with PCR than with hybridisation seems to indicate that a large part of organelle DNA was present in the nucleus. These conclusions may be substantiated by the fact that there are quite short organelle DNA fragments in the nucleus, and therefore some amplicons are cracked and are not amplified.

An attempt to find clones for some sequences previously isolated in cucumber was undertaken. The process started with the screening of clones bearing signals of two SCAR markers present in the vicinity of the sex-determining gene *Gy* at a distance of 3 and 21 cM. The library was screened by PCR with a pair of primers to the SCAR markers. DNA from the collective pools was used as a template. PCR revealed a weak band at 900 bp in the collective pool of plates 55, 64 and 80. Additionally, hybridisation with a radioactively labelled SCAR probe (700a) was performed. 8 membranes with randomly chosen plate replicas (3072 BAC clones altogether) were hybridised; however, as a result of hybridisation, no single signals were observed. The second marker (SCAR 1500a) was only hybridised with collective pools from the library. Strong signals in pools 22 (9B) and 24 (11B) were obtained (Fig. 7).

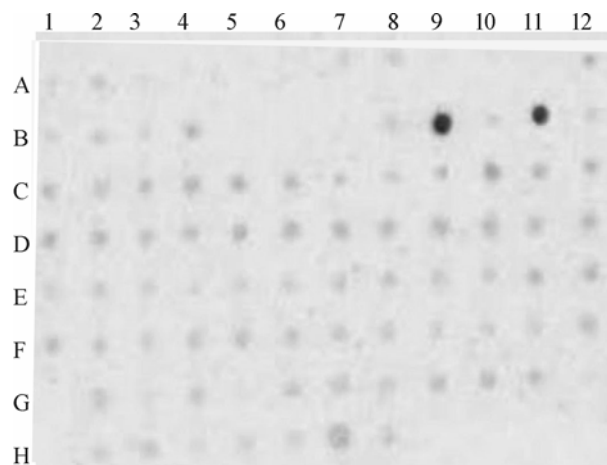


Fig. 7. Hybridization of collective pools with the radioactively labelled probe CSMRG_1500a.

SUMMARY

The constructed BAC library of cucumber is characterised by a length of clones which may be determined as average; however, it is the longest of all the known cucumber libraries, with an average amount of organelle DNA, a lack of clones without inserts and a large genome coverage.

Acknowledgements. We would like to thank Dr. W. Płader (Department of Genetics, Breeding and Biotechnology of Plant, SGGW in Warsaw) for the chloroplast gene *ycf4* primers. We also would like to thank Dr. G. Bartoszewski (Department of Genetics, Breeding and Biotechnology of Plant, SGGW in Warsaw) for the mitochondrial *Cox* gene primers. This work was supported by grants from the Polish Committee for Scientific Research (KBN) No. 2P06A00226 and No. 2P06A02328.

REFERENCES

1. Staub, J.E. and Meglic, V. Molecular genetic markers and their legal relevance for cultigen discrimination: A case study in cucumber. **Hort. Technol.** 3 (1993) 291-300.
2. Arumuganathan, K. and Earle, E.D. Nuclear DNA content of some important plant species. **Plant Mol. Biol. Rep.** 9 (1991) 208-219.
3. Malepszy, S. and Niemirowicz-Szczytt, K. Sex determination in cucumber (*Cucumis sativus* L.) as a model system for molecular biology. **Plant Sci.** 80 (1991) 39-47.
4. Nam, Y.W., Lee, J.R., Song, K.H., Lee, M.K., Robbins, M.D., Chung, S.M., Staub, J.E. and Zhang, H.B. Construction of two BAC libraries from cucumber (*Cucumis sativus* L.) and identification of clones linked to yield component quantitative trait loci. **Theor. Appl. Genet.** 111 (2005) 150-160.
5. Shizuya, H., Birren, B., Kim, U.J., Mancino, V., Slepak, T., Tachiiri, Y. and Simon, M. Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. **Proc. Natl. Acad. Sci. USA** 89 (1992) 8794-8797.
6. Shultz, J., LaMontagne, C.H., Zobrist, K., Meksem, K., Wu, C.H., Tao, Q., Nimmakayala, P., Santos, F., Zhang, H. and Lightfoot, D.A. Integrated physical mapping of the soybean genome: A tool for rapid identification of economically important genes. **Soyb. Genet. Newsl.** Vol. 27 (2000) 1-7.
7. Zhang, P., Li, W., Fellers, J., Friebe, B. and Gill, B.S. BAC-FISH in wheat identifies chromosome landmarks consisting of different types of transposable elements. **Chromosoma** 112 (2004) 288-299.
8. Peterson, D.G., Tomkins, J.P., Frisch, D.A., Wing, R.A. and Paterson, A.H. Construction of plant bacterial artificial chromosome (BAC) libraries: An illustrated guide. **J. Agricult. Genomics** 5 (2000) 1-100.
9. Sambrook, J., Fritsh, E.F. and Maniatis, T. Molecular cloning-A laboratory manual. (N Ford, C Nolan, M Ferguson Ed.) 2th edition, **Cold Spring Harbor Lab. Pres** (1989) New York.
10. Woo, S.S., Rastogi, V.K., Zhang, H.B., Paterson, A.H., Schertz, K.F. and Wing, R. Isolation of megabase-size DNA from *sorghum* and applications for physical mapping and bacterial and yeast artificial chromosome library construction. **Plant Mol. Biol.** 13 (1995) 82-94.

11. Martin, G.B., Ganal, M.W. and Tanksley, S.D. Construction of a yeast artificial chromosome library of tomato and identification of cloned segments linked to two disease resistance loci. **Mol. Gen. Genet.** 233 (1992) 25-32.
12. Wing, R.A., Rastogi, V.K., Zhang, H.B., Paterson, A.H. and Tanksley, S.D. An improved method of plant megabase DNA isolation in agarose microbeads suitable for physical mapping and YAC cloning. **Plant J.** 4 (1993) 893-898.
13. Woo, S.S., Jiang, J., Gill, B.S., Paterson, A.H. and Wing, R.A. Construction and characterization of a bacterial artificial chromosome library of *Sorghum bicolor*. **Nucleic Acid Res.** 22 (1994) 4922-4931.
14. Peterson, D.G., Boehm, K.S. and Stack, S.M. Isolation of milligram quantities of DNA from tomato (*Lycopersicon esculentum*), a plant containing high levels of polyphenolic compounds. **Plant Mol. Biol.** 15 (1997) 148-153.
15. Peterson, D.G., Pearson, W.R. and Stack, S.M. Characterization of the tomato (*Lycopersicon esculentum*) genome using *in vitro* and *in situ* DNA reassociation. **Genome** 41 (1998) 346-356.
16. Tomkins, J.P., Peterson, D.G., Yang, T.J., Main, D., Wilkins, T.A., Paterson, A.H. and Wing, R.A. Development of genomic resources for cotton (*Gossypium hirsutum* L.): BAC library construction, preliminary STC analysis, and identification of clones associated with fiber development. **Mol. Breed.** 8 (2001) 255-261.
17. Watson, J.C. and Thompson, W.F. Purification and restriction endonuclease analysis of plant nuclear DNA. **Meth. Enzymol.** 118 (1986) 57-79.
18. Song, B.K., Nadarajah, K., Romanov, M.N. and Ratnam, W. Application of overgo-based hybridization method in screening oryza rufipogon bacterial artificial chromosome (BAC) library. **Cell. Mol. Biol. Lett.** 10 (2005) 425-437.
19. Choi, S., Creelman, R.A., Mullet, J.E., Wing, R.A. Construction and characterization of a bacterial artificial chromosome library of *Arabidopsis thaliana*. **Plant Mol. Biol.** 13 (1995) 124-128.
20. Cenci, A., Chantret, N., Kong, X., Gu, Y., Anderson, O.D., Fahima, T., Distelfeld, A. and Dubcovsky, J. Construction and characterization of a half million clone BAC library of durum wheat (*Triticum turgidum* ssp. *durum*). **Theor. Appl. Genet.** 107 (2003) 931-939.
21. Yu, Y., Tomkins, J.P., Waugh, R., Frich, D.A., Kudrna, D., Kleinhofs, A., Bruggeman, R.S., Muehlbauer, G.J., Wise, R.P. and Wing, R.A. A bacterial artificial chromosome library for barley (*Hordeum vulgare* L.) and the identification of clones containing putative resistance genes. **Theor. Appl. Genet.** 101 (2000) 1093-1099.
22. Burke, D.T. and Olsen, M.V. Preparation of clone libraries in yeast artificial chromosome vectors. **Meth. Enzymol.** 194 (1991) 251-270.

23. Anand, R., Viasante, A. and Tyler-Smith, C. Construction of yeast artificial chromosome libraries with large insert using fractionation by pulsed-field gel electrophoresis. **Nucleic Acids Res.** 17 (1989) 3425-3433.
24. Clarke, L. and Carbon, J. A colony bank containing synthetic CoIEI hybrid plasmids representative of the entire *E. coli* genes. **Cell** 9 (1976) 91-99.
25. Hamilton, C.M., Frary, A., Xu, Y., Tanksley, S.D. and Zhang, H.B. Construction of tomato genomic DNA library in a binary-BAC (BIBAC) vector. **Plant J.** 18 (1999) 223-229.
26. Wu, C.C., Nimmakayala, P., Santos, F.A., Springman, R., Scheuring, C., Meksem, K., Lightfoot, D.A. and Zhang, H.B. Construction and characterization of a soybean bacterial artificial chromosome library and use of multiple complementary libraries for genome physical mapping. **Theor. Appl. Genet.** 109 (2004) 1041-1050.
27. Blanchard, J.L. and Schmidt, G.W. Pervasive migration of organellar DNA to the nucleus in plants. **J. Mol. Evol.** 4 (1995) 397-406.
28. Przybecki, Z., Kowalczyk, M.E., Witkowicz, J., Filipecki, M. and Siedlecka, E. Polymorphism of sexually different cucumber (*Cucumis sativus* L.) NIL lines. **Cell. Mol. Biol. Lett.** 9 (2004) 919-933.
29. Strong, S.J., Ohta, Y., Litman, G.W. and Amemiya, C.H.T. Marked improvement of PAC and BAC cloning is achieved using electroelution of pulsed-field gel-separated partial digests of genomic DNA. **Nucleic Acid Res.** 25 (1997) 3959-3961.