

Application of fiber-FISH in physical mapping of *Arabidopsis thaliana*

Scott A. Jackson, Ming Li Wang, Howard M. Goodman, and Jiming Jiang

Abstract: *Arabidopsis thaliana* has become a model plant species for genetic studies because of its small genome and short juvenility period. However, the small chromosomes of this species are not suitable for classical cytogenetic studies. Here we demonstrate that the fluorescence in situ hybridization (FISH) technique using extended DNA fibers can be a powerful tool in the physical mapping of the *A. thaliana* genome. Using a refined fiber-FISH technique we were able to measure DNA clusters as long as 1.71 Mb, more than 1% of the *A. thaliana* genome. Several small DNA loci, including the telomeres and a dispersed repetitive DNA sequence, mi167, were also analyzed with this technique. The results show that without known adjacent DNA markers such small DNA loci cannot be mapped precisely using fiber-FISH. One of the most difficult obstacles in physical mapping by contig assembly is closing the gaps that are present between adjacent contigs. Currently available molecular techniques are not sufficient to accurately estimate the physical sizes of these gaps. We isolated bacterial artificial chromosome (BAC) clones bordering gaps 2 and 3 on the physical contig map of *A. thaliana* chromosome II. The BAC clones were used in fiber-FISH analysis and the physical sizes of the two gaps were estimated as 31 kb and more than 500 kb, respectively. Thus, we have demonstrated that fiber-FISH is an efficient technique for determining the physical size of gaps on molecular contig maps.

Key words: fluorescence in situ hybridization, DNA fibers, physical mapping, genome analysis.

Résumé : *Arabidopsis thaliana* est devenu une espèce végétale modèle pour les études génétiques en raison de la petite taille de son génome et son cycle vital court. Cependant, la petite taille des chromosomes chez cette espèce constitue un obstacle aux études de cytogénétique classique. Ici, les auteurs montrent que la technique d'hybridation in situ à fluorescence (FISH) utilisée sur des fibres d'ADN en extension peut constituer un outil puissant dans la cartographie physique du génome d'*A. thaliana*. Suite à des raffinements de la technique fiber-FISH, les auteurs ont été capables de mesurer des segments d'ADN allant jusqu'à 1,7 Mb soit plus de 1% du génome d'*A. thaliana*. Plusieurs petits loci d'ADN, incluant les télomères et mi167, une séquence répétitive dispersée, ont été analysés à l'aide de cette technique. Les résultats font ressortir qu'en absence de marqueurs adjacents connus, de tels petits loci ne peuvent être localisés précisément par la technique fiber-FISH. Un des défis les plus importants en cartographie physique par assemblage de contigs est de combler les trous qui séparent des contigs adjacents. Les méthodes moléculaires disponibles présentement sont insuffisantes pour évaluer de façon précise la taille de ces trous. Des clones de chromosomes artificiels de bactéries (BAC) bordant les trous 2 et 3 de la carte de contigs du chromosome II d'*A. thaliana* ont été isolés. Ces clones BAC ont été employés en analyse fiber-FISH et la taille des deux trous a pu être estimée à 31 kb et plus de 500 kb respectivement. Ainsi, les auteurs ont montré que la technique fiber-FISH est un outil efficace pour déterminer la taille des trous qui séparent les divers contigs d'une carte physique.

Mots clés : hybridation in situ à fluorescence, fibres d'ADN, cartographie physique, analyse de génomes.

[Traduit par la Rédaction]

Introduction

Arabidopsis thaliana has become the most important model species for plant genome studies (Goodman et al. 1995). The small genome size and low proportion of repeti-

tive DNA sequences in *A. thaliana* make it ideal for various types of molecular analyses. However, because of its small chromosomes, *A. thaliana* has not been as extensively utilized for cytogenetic analysis as other plant species with much larger chromosomes. There have been only a few cytogenetic reports involving the mapping of repetitive DNA sequences on *A. thaliana* metaphase chromosomes using fluorescence in situ hybridization (FISH) (Maluszynska and Heslop-Harrison 1991; Murata et al. 1994, 1997; Brandes et al. 1997). Recent development of new FISH techniques using extended DNA fibers (or extended chromatin) (Heng et al. 1992; Wiegant et al. 1992; Parra and Windle 1993) has dramatically changed the suitability of *A. thaliana* for molecular cytogenetic studies. Because extended DNA fibers are the cytological target in fiber-FISH, the small size of the *A. thaliana* chromosomes is no longer an obstacle to cytological analysis. Most of the *A. thaliana* genomic DNA

Corresponding Editor: G. Fedak.

Received May 4, 1998. Accepted July 7, 1998.

S.A. Jackson and J. Jiang.¹ Department of Horticulture, University of Wisconsin—Madison, Madison, WI 53706, U.S.A.

M.L. Wang and H.M. Goodman. Department of Genetics, Harvard Medical School and Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114, U.S.A.

¹ Author to whom all correspondence should be addressed.

clones, such as cosmid and bacterial artificial chromosome (BAC) clones, contain few or no repetitive DNA sequences. Such clones can be used directly for fiber-FISH analysis without an extra procedure to remove or block the repeats in the clones. Therefore, *A. thaliana* is a better target for fiber-FISH analysis than other plant species with large complex genomes.

Fiber-FISH has been applied in various ways in genome research in humans. This technique can be used to analyze overlapping clones (Heiskanen et al. 1994), to detect chromosomal rearrangements (Heiskanen et al. 1995b), to determine the physical distance between genes and their 5' to 3' orientation (Heiskanen et al. 1995a), to measure sizes of long DNA loci (Shiels et al. 1997), and eventually, to expedite positional cloning (Laan et al. 1996). Fransz et al. (1996) successfully adopted the fiber-FISH technique for plant species. The technique was calibrated using three *A. thaliana* cosmid clones and used to analyze several classes of repetitive DNA sequences in tomato (Fransz et al. 1996). Here we report the application of the fiber-FISH technique to estimate the physical sizes of DNA loci and gauge the sizes of gaps on the physical contig map of *A. thaliana* chromosome II.

Materials and methods

Plant materials

Arabidopsis thaliana ecotype Columbia, *Oryza sativa* ssp. *javanica* line DV86 (rice), and *Sorghum* sp. line NC+ 270 (obtained from Dr. G.H. Liang, Kansas State University) were used as source materials for all cytological preparations.

DNA probes

Clone mi167, containing a 980-bp dispersed repetitive DNA sequence, was obtained from the Arabidopsis Biological Resource Center (Ohio State University, Columbus, Ohio). Dr. M. Murata (Okatama University, Kurahiki, Japan) is gratefully acknowledged for providing clones pAtMr1 and pAtHr1, which contain the 180-bp repeat. Cosmid clones pOCA18-C, pOCA18-E, and pOCA18-H were obtained from Dr. M. Estelle (Indiana University, Bloomington, Indiana). BAC clones T01O24, T06D20, T03K09, T02P04, T07M07, T02F01, and T03A03 were obtained from Dr. R.A. Wing (Clemson University, Clemson, S.C.). Probe pTa71 contains a 9-kb *EcoRI* fragment from wheat (Gerlach and Bedbrook 1979); the insert of this clone is a unit consisting of the sequences for the 5.8S, 18S, and 26S rRNA genes and the spacer regions.

Preparation of interphase nuclei and chromosomes

Seeds were germinated on moist filter paper in petri dishes at room temperature. When seedlings were several millimetres in size, they were placed in the dark at 4°C for 24 h. The seedlings were then allowed to grow at room temperature for 24 h in the light before being placed in ice-cold water for an additional 20–24 h. They were then fixed in 100% ethanol – glacial acetic acid (3:1) for at least 48 h before being stained in acetocarmine. Root tips were excised and squashed on microscope slides and preparations with good chromosome and (or) nuclei morphology were immediately placed in a freezer at –80°C until used for FISH.

Preparation of DNA fibers

Extended DNA fibers were obtained by first isolating leaf nuclei according to Liu and Whittier (1994). In short, 2 g of fresh leaf tissue were ground to a fine powder in liquid nitrogen, using a

precooled mortar and pestle. The powder was then transferred to a 50-mL tube with 20 mL of cold nuclei isolation buffer (NIB; 10 mM Tris-HCl (pH 9.5), 10 mM EDTA, 100 mM KCl, 0.5 M sucrose, 4 mM spermidine, plus 1.0 mM spermine, 0.1% (v/v) mercaptoethanol) and gently shaken on ice for 5 min. Subsequently, the solution was filtered sequentially through 180, 120, and 47 µm mesh nylon membranes, while on ice. We omitted the 22-µm mesh filtration step in order to ensure a high concentration of nuclei. The filtrate was supplemented with 1 mL of 10% (v/v) Triton X-100 in NIB and then centrifuged at 2000 × *g* for 10 min at 4°C. The pellet was resuspended in 200-µL NIB, supplemented with an equal volume of 100% glycerol, and stored at –20°C.

Extension of DNA fibers followed the method of Fransz et al. (1996) with several modifications. Instead of tilting the slide and allowing the fibers to run down it, the DNA fibers were extended by dragging with a cover slip, which gave longer and cleaner DNA fibers. Poly-L-lysine (Sigma) slides were used in preference to silitated slides, which generated too much background, or untreated slides, on which the DNA fibers tended to clump and in consequence were too short for our purposes. For 10 slides, 10 µL of the nuclei suspension was centrifuged at 3600 rpm for 5 min, and the pellet resuspended in 15–30 µL of PBS (10 mM sodium phosphate (pH 7.0) plus 140 mM NaCl); to optimize the quality of the DNA fibers, the number of nuclei that are deposited on the slide can be increased or decreased by adjusting the amount of PBS used to resuspend the pellet. Then, 2 µL of the suspension was deposited on one end of a microscope slide (poly-L-lysine slides, Sigma) and dried briefly. After this, 10 µL of lysis buffer (0.5% (w/v) SDS, 5 mM EDTA, plus 100 mM Tris (pH 7.0)) was placed on the nuclei and incubated at room temperature for 4 min; the DNA fibers were then extended using a clean cover slip. The slides were completely air-dried, fixed in 100% ethanol – glacial acetic acid (3:1) for 2 min, and finally, baked at 60°C for 30 min. Slides were stored in a slide box for up to several weeks at room temperature.

Probe labeling and hybridization

Probes were labeled with either digoxigenin-11-dUTP or biotin-11-dUTP by nick translation. Preparation of the hybridization mixture was according to Jiang et al. (1996). Approximately 10 µL of hybridization mixture was added to the slide and covered with a cover slip, and the cover slip was then sealed with rubber cement. The probe and chromosomal DNA were denatured by placing in an oven at 80°C in for 3–5 min, followed by incubation at 37°C in a hybridization chamber overnight.

The washes, buffers, and protocol used for detection followed Zhong et al. (1996). The biotin-labeled probes were detected with avidin DN (Vector Laboratories), followed by biotinylated anti-avidin D (Vector Laboratories), and finally with fluorescein – avidin DN (Vector Laboratories). Digoxigenin-labeled probes were detected by mouse anti-digoxigenin (Boehringer Mannheim), followed by digoxigenin conjugated sheep – anti-mouse IgG antibodies (Boehringer Mannheim), and finally with rhodamine anti-digoxigenin (Boehringer Mannheim). The slides were counterstained with 4',6-diamidino-2-phenylindole (DAPI) or propidium iodide (PI) in Vectashield mounting medium (Vector Laboratories).

Image capture and analysis

Detection and analysis of FISH signals were accomplished using an Olympus BX60 microscope with an external Olympus BH2-RFL-T3 epifluorescence source and 60× Olympus PlanApo and 100× Olympus UPlanFl lenses. Images were captured with a Photometrics SenSys CCD (charge coupled device) camera coupled to a Macintosh computer. Fluorescent signals from different probes were captured individually and merged using IPLab Spectrum v3.1 software. IPLab Spectrum software was also used to

Fig. 1. (a) Representative fiber-FISH signals from two BAC clones. The average sizes of signals from T02P04 (green) and T03K09 (red) are 28.74 and 29.54 μm , respectively, and they are separated by 43.62 μm . Scale bar = 20 μm . (b) Fiber-FISH signals derived from the 180-bp repeat (red) and the mi167 probe (green). The putative signals from the mi167 sequences are marked by arrows. Scale bar = 20 μm . (c and d) The longest fiber-FISH signals derived from the 180-bp repeat and the NOR probes. (c) A 392.4 μm long signal, corresponding to approximately 1126 kb, from the 180-bp repeat probe pAtMr1. Scale bar = 50 μm . (d) A 595.9 μm long signal, corresponding to approximately 1710 kb, from the NOR probe pTa71. Scale bar = 50 μm . (e) Fiber-FISH signals, detected by pAtT4, derived from the telomeres of *A. thaliana*, rice, and sorghum chromosomes. Scale bar = 5 μm . (f) Fiber-FISH signals derived from BACs T03A03 (red) and T02F01 (green) flanking gap 2 on chromosome II. These two BACs are separated by 10.72 μm , corresponding to 31 kb. Scale bar = 20 μm . (g) BACs T06D20 (red) and T01O24 (green), flanking gap 3, overlap on a pair of metaphase chromosomes. Scale bar = 5 μm . (h) BACs T06D20 (red) and T01O24 (green) are separated by an average of 1.82 μm on interphase nuclei. Scale bar = 5 μm . (i) BACs T03K09 (red) and T07M07 (green), 220 kb apart, are separated by an average of 0.99 μm on interphase nuclei. Scale bar = 5 μm . (j) Fiber-FISH analysis of BAC clones T06D20 (red) and T01O24 (green) that flank gap 3. No physical association between these two clones was observed. Scale bar = 50 μm . (k) Fiber-FISH analysis of two BAC clones, T07M07 (green) and T03K09 (red), that are separated by 220 kb. The physical connection and distance between these two BACs were consistently identified on extended DNA fibers. Scale bar = 20 μm .

make measurements on digitized images. Standard deviations were calculated for all measurement means, except where 95% confidence intervals (CI) are noted.

Results and discussion

Calibration of the fiber-FISH technique

Three cosmid clones and two BAC clones with known insert sizes were hybridized to extended DNA fibers in order to calibrate and standardize the measuring process. The fiber-FISH signals resembled the typical "beads-on-a-string" pattern described in previous reports (Parra and Windle 1993; Bengtsson et al. 1994; Haaf and Ward 1994; Houseal et al. 1994; Senger et al. 1994; Franz et al. 1996). The insert size (kilobases) and fiber-FISH signals (micrometres) from the five clones are summarized in Table 1. The average resolution based on the data from cosmids C, E, and H and BAC clones T02P04 and T03K09 was 2.87 kb/ μm (Table 1). BAC clones T02P04 and T03K09 have insert sizes of 85 and 90 kb, respectively, and are separated by 126 kb based on sequencing data (M.L. Wang and H.M. Goodman, unpublished results). Fiber-FISH results from T02P04 gave 28.74 μm (\pm 3.87 μm , n = 15 (15 signals)), corresponding to 82 kb (\pm 11 kb). The average size of the signals with T03K09 was 29.54 μm (\pm 4.60 μm , n = 15), corresponding to 85 kb (\pm 13 kb). The two clones were separated by 43.62 μm (\pm 4.25 μm , n = 13), corresponding to 125 kb (\pm 12 kb) (Fig. 1a). These data confirmed that the estimated sizes based on fiber-FISH were very close to those derived from sequencing analysis. All the cosmid and BAC clones consist almost exclusively of unique sequences, because no significant cross-hybridization signals were detected in the absence of blocking DNA.

According to the Watson-Crick DNA model, double-stranded DNA molecules can be maximally stretched to 2.97 kb/ μm . However, in humans, Parra and Windle (1993) reported that fibers can actually be extended up to twice the expected length. Some of the variability in previous reports may be due to the different methods of slide preparation: for example, slide pretreatments may or may not have been applied to promote DNA adhesion, or various methods for extending the DNA fibers may have been used. Several different methods have been used to extend DNA fibers, these include tilting the slide and permitting the DNA to run down it, physically dragging the DNA across the slide with a

Table 1. Determination of fiber-FISH resolution using three cosmid clones and two BAC clones.

Clone	Insert size (kb)	Fiber-FISH			Resolution ^b (kb/ μm)
		n^a	Length (μm)	Standard deviation	
pOCA18-H	20	30	7.89	1.22	2.53
pOCA18-C	22	19	7.28	1.31	3.02
pOCA18-E	22	10	6.83	0.68	3.22
T02P04	85	15	28.74	3.87	2.96
T03K09	90	15	29.54	4.6	3.05

^a n . Number of measurements.

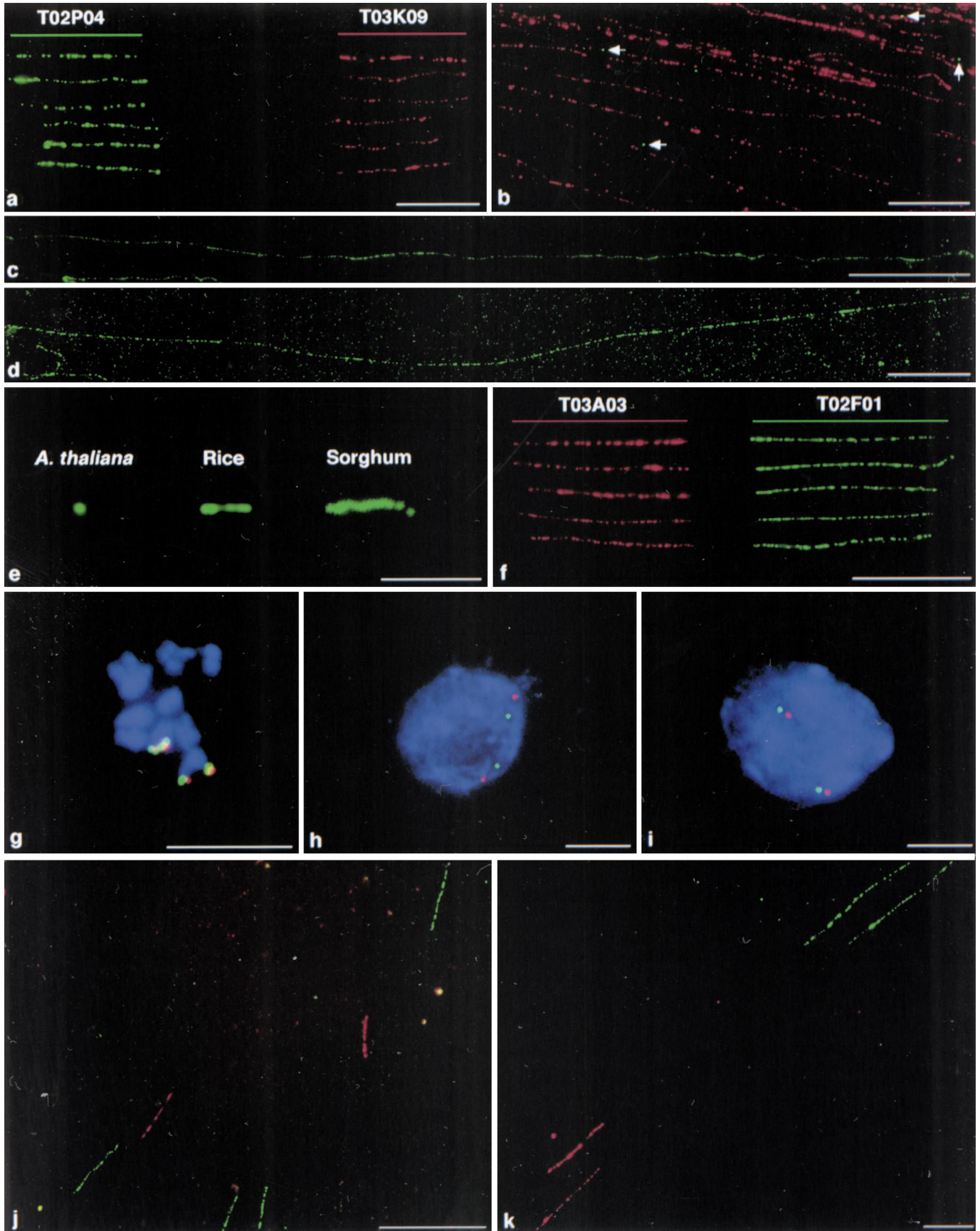
^bWeighted average = 2.87 kb/ μm .

cover slip, or placing a cover slip on top of the DNA and allowing it to dry, thereby causing the DNA molecules to extend as the meniscus recedes. It has been reported that measurements in the centromeric region are less reliable (up to 60% variability), whereas measurements from other regions only vary by about 20% (Nishio et al. 1996). Previous reports of fiber-FISH resolution range from 2.78 to 3.3 kb/ μm in both mammalian and plant species (Heiskanen et al. 1995b; Franz et al. 1996; Shiels et al. 1997; Sjöberg et al. 1997). Therefore it is important that the fiber-FISH technique be calibrated in individual laboratories.

Physical mapping of large (>1 Mb) and small (<3 kb) DNA loci from *A. thaliana*

One direct application of the fiber-FISH technique is in measuring the physical sizes of DNA loci by converting microscopic sizes (micrometre) into linear DNA sizes (kilobases). Using this technique, Franz et al. (1996) was able to estimate the 5S ribosomal RNA gene cluster in tomato to be 660 kb. Our first goal was to determine the limits of fiber-FISH for measuring variously sized DNA loci in *A. thaliana*.

The 180-bp repeat (pAL1 family) is the most abundant repetitive DNA family, except for the NOR (nucleolus organizer region) sequences, in the *A. thaliana* genome and is located in the centromeric region in all five chromosomes (Martinez-Zapater et al. 1986; Pruitt and Meyerowitz 1986; Maluszynska and Heslop-Harrison 1991; Murata et al. 1994). Pulsed-field gel electrophoresis (PFGE) analysis indicated that the 180-bp repeat is organized into long uninterrupted arrays, and that some of the arrays are as long as 1000 kb



(Round et al. 1997). We used two DNA probes, pAtMr1 and pAtHr, both containing the 180-bp repeat, for fiber-FISH analysis. Long stretches of hybridization signals were observed (Fig. 1*b*). The longest signal measured on a single fiber was 392 μm , corresponding to approximately 1126 kb (Fig. 1*c*). Extended DNA fibers of *A. thaliana* were also hybridized to probe pTa71 that contains coding sequences for the 5.8S, 18S, and 26S rRNA genes of wheat (Gerlach and Bedbrook 1979). The longest signal obtained with this probe was 596 μm , corresponding to approximately 1710 kb, more than 1% of the *A. thaliana* genome (Fig. 1*d*). Based on molecular analysis, the two NOR loci on chromosomes 2 and 4 in *A. thaliana* are in the order of 3.5–4 Mb each (Copenhaver and Pikaard 1996*b*). We were not able to obtain DNA fibers that were long enough to measure a complete NOR locus.

Given a 2.87 kb/ μm resolution, the fiber-FISH signals generated from a small DNA probe (<3 kb) will be $\sim 1 \mu\text{m}$ long and may consist of only one or two spots. To determine the feasibility of mapping such probes using fiber-FISH, we used several short DNA probes from *A. thaliana* for analysis.

Probe mi167 is an interspersed middle repetitive DNA element and contains a 980-bp *Pst*I fragment. We observed individual bright fluorescent spots in fiber-FISH. Because images of single DNA fibers cannot be generated by staining with DAPI or PI, it is not possible to determine whether a single fluorescent spot is on a DNA fiber if neighboring reference signals from other probes are not available. Therefore, most putative signals cannot be distinguished unambiguously from the hybridization background, even though most true signals were probably stronger than the background spots (Fig. 1*b*). Although both the 180-bp repeat and the mi167 sequence are located in the centromeric regions (Thompson et al. 1996), the physical relationship of these two probes on the same fibers is difficult to determine, mainly because of the difficulty of localizing true mi167 signals (Fig. 1*b*). Most of the putative mi167 signals were not localized on the same fibers as those derived from the 180-bp repeat.

The telomeres of *A. thaliana* chromosomes are approximately 2.5 kb long (Richards and Ausubel 1988). Fiber-FISH signals generated from the telomeric DNA probe pAtT4 resembled those from the mi167 probe. Without reference signals, the true telomeric signals were difficult to confirm. We also analyzed the telomeres of rice and sorghum chromosomes using the same probe, pAtT4. Short stretches of signals were consistently observed on both rice and sorghum DNA fibers. The largest measured telomeres in rice and sorghum were 2.51 and 5.40 μm , corresponding to approximately 7.2 and 15.5 kb, respectively (Fig. 1*e*).

The results from the 180-bp repeat and the NOR probe demonstrated that with the current technique, we could define a DNA locus as long as 1.71 Mb, which corresponds to more than 1% of the *A. thaliana* genome. Thus, this technique can be used to map most of the cloned genomic DNA segments, including megabase-sized YAC (yeast artificial chromosome) clones, and to analyze large repetitive regions of the genome. Recently, repetitive DNA clusters as long as 2.6 Mb were measured by fiber-FISH in humans (Shiels et al. 1997). We expect that DNA loci longer than 1.71 Mb

could be measured by fiber-FISH in plants, if the technique of fiber preparation can be improved in the future.

The results from probes pAtT4 and mi167 showed that fiber-FISH signals derived from short DNA loci (<3 kb) are often single spots on extended DNA fibers. This type of hybridization pattern will be problematic in distinguishing true signals from background noise. Based on molecular analysis, the telomeres on chromosomes 2 and 4 in *A. thaliana* are directly attached to the NOR loci (Copenhaver and Pikaard 1996*a*; Haberer et al. 1996). However, we were unable to find the junctions between the telomeres and the NOR clusters, because it was difficult to confirm true telomeric signals. Although there have been reports of ordering DNA probes a few kilobases or even a few hundred base pairs long (Heiskanen et al. 1996; Fransz et al. 1996; Palotie et al. 1996), these analyses would be very difficult if there was no pre-existing knowledge about the sizes and distances between these probes.

Analysis of gaps in the physical contig of chromosome II of *A. thaliana*

A physical map of *A. thaliana* chromosome II based on YAC contigs has been constructed (Zachgo et al. 1996). This physical map contains three gaps that could not be closed, owing either to the absence of YACs containing these sequences in the CIC YAC library (Creusot et al. 1995) or to the repetitive nature of the YAC end probes. Gap 1 may represent the centromere, because several YACs in this region contain the 180-bp repeat. The sizes of gaps 2 and 3, both on the long arm of the chromosome, were estimated to be 340 and 720 kb, respectively, based on the genetic distances between the markers bordering the gaps and a conversion factor of 200 kb per cM (Zachgo et al. 1996). However, these estimations can be faulty, as recombinational events are not evenly distributed along chromosomes in higher eukaryotes.

Two BAC clones, T02F01 and T03A03, were isolated using DNA markers bordering gap 2. Fiber-FISH analysis revealed that these two BACs were separated by only 10.74 μm ($\pm 0.95 \mu\text{m}$, $n = 16$), corresponding to approximately 31 kb (29.3–32.3 kb, 95% CI) (Fig. 1*f*). Thus, gap 2 was significantly smaller than expected (340 kb). It should not be difficult to close this gap, if this chromosomal region is well covered by the *A. thaliana* BAC library (Choi et al. 1995).

Recently gap 3 was spanned by a 120 kb YAC clone, EW16G04 (Wang et al. 1997), suggesting that this gap could be smaller than 120 kb. Two BAC clones, T06D20 and T01O24, were isolated using DNA markers bordering gap 3 and were used in FISH analysis. These two clones overlapped on metaphase chromosomes (Fig. 1*g*), indicating that they are located in the same chromosomal region. However, no physical connection between these two clones was observed on extended DNA fibers (Fig. 1*j*). To ensure the feasibility of fiber-FISH for analyzing physical gaps as large as several hundred kilobases, two BAC clones, T03K09 and T07M07, separated by 220 kb from each other, were analyzed by fiber-FISH. Consistent physical association between these two clones was observed on extended DNA fibers (Fig. 1*k*) and the size of the gap that separates the two clones was close to the expected value based on the fiber-FISH measurements (data not shown). DNA fibers from

500 kb to 1.5 Mb were consistently generated using our current technique. If gap 3 is smaller than 500 kb, it should be readily defined by fiber-FISH. Therefore, although the physical size of gap 3 is not determined, based on the fiber-FISH results, it is probably larger than 500 kb. The YAC clone EW16G04 that spans gap 3 is most likely either a chimeric clone or a rearranged clone with an original size larger than 500 kb.

BACs T06D20 and T01O24 were also analyzed on interphase nuclei. They were separated by an average of 1.82 μm ($\pm 0.49 \mu\text{m}$, $n = 14$) (Fig. 1h). As a control, BACs T03K09 and T07M07 had an average distance of 0.99 μm ($\pm 0.57 \mu\text{m}$, $n = 16$) on interphase nuclei (Fig. 1i). Thus the interphase distance between T06D20 and T01O24 is almost twice as long as that between T03K09 and T07M07, which are separated by 220 kb. In the only plant interphase FISH mapping report, Jiang et al. (1996) demonstrated that, in maize, two DNA clones separated by 140 kb had an interphase distance of 0.5 μm . In humans, the interphase distance measured in micrometres is linearly related to the physical distance measured in kilobases, if the probes are separated by less than 1 Mb and the physical distance can be estimated by the equation $R = Ln^{-2}$ (where R is the average interphase distance, n is the physical distance, and L is a constant) (Van Den Engh et al. 1992). However, it is not known whether the pattern of interphase chromatin organization is the same in plants as in humans. Thus, a standard calibration between interphase and physical distances has to be established in *A. thaliana* before the interphase distances can be converted directly into DNA distances. Nevertheless, the present interphase FISH data confirm that the physical size of gap 3 is larger than the distance separating BACs T03K09 and T07M07.

The assembly of genomic contigs using BAC clones will be the primary strategy for sequencing large complex genomes. Large genomes currently being sequenced, or targeted for sequencing, include humans, *A. thaliana*, and recently, rice (Sasaki 1997). One of the most difficult obstacles in contig assembly is closing the gaps separating linearly related contigs along a chromosome. Several phenomena may account for these gaps, such as the presence of repetitive DNA sequences, insufficient coverage of the BAC libraries in certain chromosomal regions, or not enough DNA markers in certain genomic regions because of low recombination rates. These gaps have been present in almost every assembled contig reported (Schmidt et al. 1995; Saji et al. 1996; Zachgo et al. 1996; Umehara et al. 1997). It is almost impossible to size such gaps using standard fingerprinting or other molecular techniques. Here we demonstrated that FISH techniques using different targets, including metaphase chromosomes, interphase nuclei, and extended DNA fibers, provide valuable tools for estimating the physical sizes of such gaps. Although FISH on metaphase chromosomes has a low resolution, it can be used to efficiently determine whether two clones are located in the same chromosomal region (Fig. 1g). Fiber-FISH can be used to orientate probes that are separated by a few kilobases up to possibly 500 kb. Currently it is not known whether fiber-FISH will be effective in orienting probes separated by more than 1 Mb, as long DNA fibers tend to bend. Not only is it difficult to find the connection between probes on bent

fibers, but the accuracy of fiber-FISH measurements will be decreased. However, once a standard calibration system for interphase mapping is established in *A. thaliana*, interphase FISH will be an effective method for estimating physical distances between clones separated by a few hundred kilobases up to 1 Mb.

Acknowledgement

This research is supported by Funds 135-0534 and 135-0528 from the Graduate School of the University of Wisconsin—Madison.

References

- Bengtsson, U., Altherr, M.R., Wasmuth, J.J., and Winokur, S.T. 1994. High resolution fluorescence in situ hybridization to linearly extended DNA visually maps a tandem repeat associated with facioscapulohumeral muscular dystrophy immediately adjacent to the telomere of 4q. *Hum. Mol. Genet.* **3**: 1801–1805.
- Brandes, A., Thompson, H., Dean, C., and Heslop-Harrison, J.S. 1997. Multiple repetitive DNA sequences in the paracentric regions of *Arabidopsis thaliana* L. *Chromosome Res.* **5**: 238–246.
- Choi, S., Creelman, R.A., Mullet, J.E., and Wing, R.A. 1995. Construction and characterization of bacterial artificial chromosome library of *Arabidopsis thaliana*. *Plant Mol. Biol. Rep.* **13**: 124–128.
- Copenhaver, G.P., and Pikaard, C.S. 1996a. RFLP and physical mapping with an rDNA-specific endonuclease reveals that nucleolus organizer regions of *Arabidopsis thaliana* adjoin the telomeres on chromosomes 2 and 4. *Plant J.* **9**: 259–272.
- Copenhaver, G.P., and Pikaard, C.S. 1996b. Two-dimensional RFLP analyses reveal megabase-sized clusters of rRNA gene variants in *Arabidopsis thaliana*, suggesting local spreading of variants as the mode for gene homogenization during concerted evolution. *Plant J.* **9**: 273–282.
- Creusot, F., Fouilloux, E., Dron, J., Lafleurriel, J., Picar, G., Billaut, A., Le Paslier, D., Chaboute, M.-E., Durr, A., Fleck, J., Gigot, C., Camilleri, C., Bellini, C., Caboche, M., and Bouchez, D. 1995. The CIC library: a large insert YAC library for genome mapping in *Arabidopsis thaliana*. *Plant J.* **8**: 763–770.
- Fransz, P.F., Alonso-Blanco, C., Liharska, T.B., Peeters, A.J.M., Zabel, P., and de Jong, J.H. 1996. High-resolution physical mapping in *Arabidopsis thaliana* and tomato by fluorescence in situ hybridization to extended DNA fibres. *Plant J.* **9**: 421–430.
- Gerlach, W.L., and Bedbrook, J.R. 1979. Cloning and characterization of ribosomal RNA genes from wheat and barley. *Nucleic Acids Res.* **7**: 1869–1885.
- Goodman, H.M., Ecker, J.R., and Dean, C. 1995. The genome of *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* **92**: 10 831 – 10 835.
- Haaf, T., and Ward, D.C. 1994. High resolution ordering of YAC contigs using extended chromatin and chromosomes. *Hum. Mol. Genet.* **3**: 629–633.
- Haberer, G., Fischer, T.C., and Torres-Ruiz, R.A. 1996. Mapping of the nucleolus organizer on chromosome 4 in *Arabidopsis thaliana*. *Mol. Gen. Genet.* **250**: 123–128.
- Heiskanen, M., Karhu, R., Hellsten, E., Peltonen, L., Kallioniemi, O.P., and Palotie, A. 1994. High resolution mapping using fluorescence in situ hybridization to extended DNA fibers prepared from agarose-embedded cells. *BioTechniques*, **17**: 928–933.

- Heiskanen, M., Saitta, B., Palotie, A., and Chu, M. 1995a. Head to tail organization of the human COL6A1 and COL6A2 genes by fibre-FISH. *Genomics*, **29**: 801–803.
- Heiskanen, M., Hellsten, E., Kallioniemi, O., Mäkelä, T.P., Alitalo, K., Peltonen, L., and Palotie, A. 1995b. Visual mapping by fibre-FISH. *Genomics*, **30**: 31–36.
- Heiskanen, M., Kallioneimi, O., and Palotie, A. 1996. Fiber-FISH: experiences and a refined protocol. *Genet. Anal. Biomol. Eng.* **12**: 179–184.
- Heng, H.H.Q., Squire, J., and Tsui, L.-C. 1992. High-resolution mapping of mammalian genes by in situ hybridization to free chromatin. *Proc. Natl. Acad. Sci. U.S.A.* **89**: 9509–9513.
- Houseal, T.W., Dackowski, W.R., Landes, G.M., and Klinger, K.W. 1994. High resolution mapping of overlapping cosmids by fluorescence in situ hybridization. *Cytometry*, **15**: 193–198.
- Jiang, J., Hulbert, S.H., Gill, B.S., and Ward, D.C. 1996. Interphase fluorescence in situ mapping: a physical mapping strategy for plant species with large complex genomes. *Mol. Gen. Genet.* **252**: 497–502.
- Laan, M., Isosomppi, J., Klockars, T., Peltonen, L., and Palotie, A. 1996. Utilization of FISH in positional cloning: an example on 13q22. *Genome Res.* **6**: 1002–1012.
- Liu, Y., and Whittier, R.F. 1994. Rapid preparation of megabase plant DNA from nuclei in agarose plugs and microbeads. *Nucleic Acids Res.* **22**: 2168–2169.
- Maluszynska, J., and Heslop-Harrison, J.S. 1991. Localization of tandemly repeated DNA sequences in *Arabidopsis thaliana*. *Plant J.* **1**: 159–166.
- Martinez-Zapater, J.M., Estelle, M.A., and Somerville, C.R. 1986. A highly repeated DNA sequence in *Arabidopsis thaliana*. *Mol. Gen. Genet.* **204**: 417–423.
- Murata, M., Ogura, Y., and Motoyoshi, F. 1994. Centromeric repetitive sequences in *Arabidopsis thaliana*. *Jpn. J. Genet.* **69**: 361–370.
- Murata, M., Heslop-Harrison, J.S., and Motoyoshi, F. 1997. Physical mapping of the 5S ribosomal RNA genes in *Arabidopsis thaliana* by multi-color fluorescence in situ hybridization. *Plant J.* **12**: 31–37.
- Nishio, H., Heiskanen, M., Palotie, A., Belanger, L., and Cugaiczkyk. 1996. Tandem arrangement of the human serum albumin multigene family in the sub-centromeric region of 4q: evolution and chromosomal direction of transcription. *J. Mol. Biol.* **259**: 113–119.
- Palotie, A., Heiskanen, M., Lann, M., and Horelli-Kuitunen, N. 1996. High-resolution fluorescence in situ hybridization: a new approach in genome mapping. *Ann. Med.* **28**: 101–106.
- Parra, I., and Windle, B. 1993. High resolution visual mapping of stretched DNA by fluorescent hybridization. *Nature Genet.* **5**: 17–21.
- Pruitt, R.E., and Meyerowitz, E.M. 1986. Characterization of the genome of *Arabidopsis thaliana*. *J. Mol. Biol.* **187**: 169–183.
- Richards, E.J., and Ausubel, F.M. 1988. Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. *Cell*, **53**: 127–136.
- Round, E.K., Flowers, S.K., and Richards, E.J. 1997. *Arabidopsis thaliana* centromere regions: genetic map positions and repetitive DNA structure. *Genome Res.* **7**: 1045–1053.
- Saji, S., Umehara, Y., Kurata, N., Ashikawa, I., and Sasaki, T. 1996. Construction of YAC contigs on rice chromosome 5. *DNA Res.* **3**: 297–302.
- Sasaki, T. 1997. 5W 1H for rice whole genome sequencing. *Rice Genome*, **6**: 1.
- Schmidt, R., West, J., Love, K., Lenehan, Z., Lister, C., Thompson, H., Bouchez, D., and Dean, C. 1995. Physical map and organization of *Arabidopsis thaliana* chromosome 4. *Science (Washington, D.C.)*, **270**: 480–483.
- Senger, G., Jones, T.A., Fidlerová, H., Sanséau, P., Trowsdale, J., Duff, M., and Sheer, D. 1994. Released chromatin: linearized DNA for high resolution fluorescence in situ hybridization. *Hum. Mol. Genet.* **3**: 1275–1280.
- Shiels, C., Coutelle, C., and Huxley, C. 1997. Analysis of ribosomal and alpha repetitive DNA by fiber-FISH. *Cytogenet. Cell Genet.* **76**: 20–22.
- Sjöberg, A., Peelman, L.J., and Chowdhary, B.P. 1997. Application of three different methods to analyze fibre-FISH results obtained using four lambda clones from the porcine region MHC III. *Chromosome Res.* **5**: 247–253.
- Thompson, H., Schmidt, R., Brandes, A., Heslop-Harrison, J.S., and Dean, C. 1996. A novel repetitive sequence associated with the centromeric regions of *Arabidopsis thaliana* chromosomes. *Mol. Gen. Genet.* **253**: 247–252.
- Umehara, Y., Kurata, N., Ashikawa, I., and Sasaki, T. 1997. Yeast artificial chromosome clones of rice chromosome 2 ordered using DNA markers. *DNA Res.* **4**: 127–131.
- Van Den Engh, G., Sachs, R., and Trask, B.J. 1992. Estimating genomic distance from DNA sequence location in cell nuclei by a random walk model. *Science (Washington, D.C.)*, **257**: 1410–1412.
- Wang, M.L., Huang, L., Bongard-Pierce, D.K., Belmonte, S., Zachgo, E.A., Morris, J.W., Dolan, M., and Goodman, H.M. 1997. Construction of an ~2 Mb contig in the region around 80 cM of *Arabidopsis thaliana* chromosome 2. *Plant J.* **12**: 711–730.
- Wiegant, J., Kalle, W., Mullenders, L., Brookes, S., Hoovers, J.M.N., Dauwerse, J.G., van Ommen, G.J.B., and Raap, A.K. 1992. High-resolution in situ hybridization using DNA halo preparations. *Hum. Mol. Genet.* **8**: 587–591.
- Zachgo, E.A., Wang, M.L., Dewdney, J., Bouchez, D., Camilleri, C., Belmonte, S., Huang, L., and Goodman, H.M. 1996. A physical map of chromosome 2 of *Arabidopsis thaliana*. *Genome Res.* **6**: 19–25.
- Zhong, X., Frasz, P.F., Wennekes-Van Eden, J., Zabel, P., Van Kammen, A., and Hans de Jong, J. 1996. High-resolution mapping on pachytene chromosomes and extended DNA fibres by fluorescence in situ hybridisation. *Plant Mol. Biol. Rep.* **14**: 232–242.