Arabidopsis genome의 physical map 구축과 그 이용

Construction of a physical map of *Arabidopsis* genome and its application

황 인 환, H.G. Goodman

Inhwan Hwang[1,2] and Howard M. Goodman[2]

[1]Plant Molecular Biology and Biotechnology Research Center, Gyeongsang National University, Chinju, 660-701, Korea and [2]Department of genetics, Harvard Medical School and Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114,USA

Introduction

Many genes have been identified simply by mutations without information about the biological activity of the encoded proteins. Chromosome walking to clone the genes at these loci is a laborious and time consuming process (Bender et al., 1983). As a more general approach, there are attempts to construct physical maps of entire genomes of *C. elegans* (Coulson et al., 1986) and *Arabidopsis* (Hauge and Goodman, 1991) using cosmid clones by a fingerprinting method (Cross and Little, 1986). The physical map of an entire genome which consists of overlapping clones or DNA fragments will allow immediate access to any given locus in the genome, therefore greatly facilitating the cloning process by chromosome walking. Physical maps can also serve as a platform for other studies such as genome organization. However, it has not been possible to construct a complete physical map using cosmid clones alone. Efforts are now being made to complete the physical maps of *C. elegans* (Coulson et al., 1088) and *Arabidopsis* genomes (Hwang et al., 1991) using yeast artificial chromosome (YAC) libraries. The YAC cloning system permits the cloning of large DNA fragments and provides reduced cloning bias compared to other conventional cloning systems in *E. coli* (Burke et al., 1987). However, the large insert size and low copy number of YACs make conventional techniques unsuitable for linking YAC clones. Several approaches have been developed in order to

use the YAC clone for these purposes. One of them was the polymerase chain reaction (PCR) screening method (Heard et al., 1989). This is a two step approach to identify positive clones. First pools of YAC clones are screened by PCR and then positive clones are identified from the pools by hybridization. Primers for the PCR screening were obtained from either known sequences or the sequences of the end fragments of YAC inserts isolated by two PCR approaches, inverse PCR (IPCR) (Ochman et al., 1988) or the PCR using a specialized adaptor (bubble structured adaptor) (Riley et al., 1990). The sequence of the end fragments were then obtained to generate new primers for the next round of screening. Repeating of such a process eventually allows isolation of contiguous clones to generate a physical map. Since the Arabidopsis has relatively small genome compared to other known higher organisms the screening process could be simplified to a one step process. Thus we have developed a new simplified method and used to construct an *Arabidopsis* physical map of chromosome 2.

Here we describe construction of a physical map of *Arabidopsis* using YAC clones and its application in various aspects of biological studies such as cloning of genetically identified genetic loci and construction of a fine genetic map.

Main Contents

Strategy for the construction of a physical map

To construct a physical map of Arabidopsis we evaluated advantages and disadvantages of Arabidopsis genome in the construction of a physical map. It has been known that the *Arabidopsis* genome is relatively small, the size being approximately 100 Mb and has limited repetitive sequences (Leutwiler et al., 1984). Thus approximately 2,000 clones of the available YAC clones could represent 7-fold genomic equivalents. Also there are over 300 RFLP markers available to use probes for screening (Chang et al., 1988; Nam et al., 1989). Therefore we decided to do direct screening of YAC libraries using the cosmid RFLP markers as hybridization probes and then do chromosome walking bidirectionally from the initial YAC clones corresponding to the cosmid markers (Fig.1).

Development of a new method for the construction of a physical map

The overall strategy for the identification of overlapping YAC clone is to use direct hybridization method using end fragments of the YAC inserts as probes. This end fragment could maximize the walking distance and also

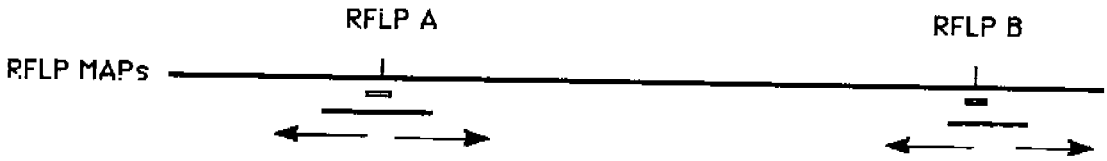


Figure 1. The strategy for the construction of a physical map
The small open bars and closed bars indicate cosmid RFLP markers and corresponding YAC clones, respectively. The arrows indicate chromosome walking directions.

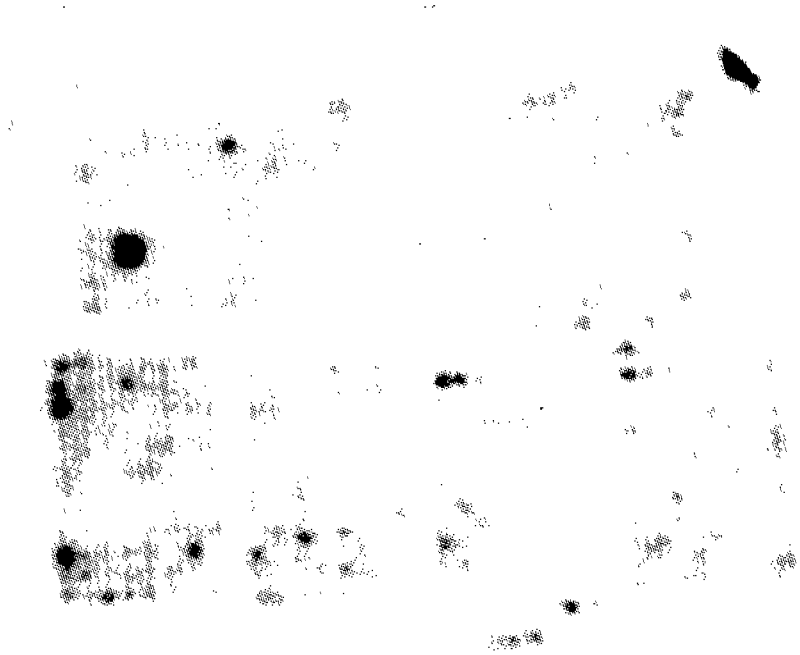


Figure 2. Screening of YAC libraries
2000 YAC clones were spotted on 4 nylon membranes and used for hybridization.

reduce cross-hybridization. The end fragments were isolated by inverse PCR (IPCR).

Isolation of end fragments of YAC inserts by the IPCR method has been described previously (Ochman et al., 1988). To prepare end fragments of YAC inserts by PCR methods the YAC DNA was purified by CHEF gel electrophoresis from total yeast genomic DNA. YAC DNA was digested with multiple restriction enzymes to obtain DNA fragments which are suitable in size for PCR amplification. After digestion, the 5′ overhang was repaired to flush the ends with DNA polymerase Klenow in the presence of deoxynucleotide triphosphates. The digested DNA fragments were self-ligated to produce circular molecules. The ligated DNA is then used as a substrate for PCR using the specific primers derived from each arm of vector sequence. Thus fragments specific to each end of the YAC inserts are produced in separate PCRs.

The amplified products were subsequently labeled by PCR to give a high specific activity. Since the PCR product has a portion of the vector sequence which gave significant background at the condition we used for hybridization (data not shown), it is necessary to remove the vector portion from the labeled probe to obtain high signal to background ratio. For this purpose specific labeling primers are used for the PCR labeling. These primers are 25-mers composed of 15 nucleotides derived from the vector sequence flanking the cloning site and 10 nucleotides containing a Fok I restriction site. The Fok I site in the labeling primers is at the 5′ to the vector sequence as shown in Fig. 2. Since the Fok I restriction endonuclease cleaves DNA 9 to 12 bases away from the recognition site 9 to 12 base pairs of vector portion from the labeled probe will be removed by digestion with Fok I restriction endonuclease after labeling reaction.

This method was tested using an arbitrary YAC clone, EW12G4. YAC DNA band was cut out from a low melting point agarose gel after CHEF gel electrophoresis (Bancroft and Wolk, 1988). From this clone YAC DNA was purified and used for PCR amplification. The PCR products were gel purified and subjected to reamplification in the presence of 32-P-dATP to label resulting products to use as hybridization probes.

Screening of YAC libraries

Two *Arabidopsis* YAC libraries (EG and EW libraries) (Ward and Jen, 1990; Grill and Somerville, 1991) were screened using these left and right end probes. As shown in Fig. 1, clear positive signals were easily identified on filters containing grids of about 800 YAC clones. The left and right end probes identified YAC clones EG19C10 and EW12H5, respectively. Also the probe hybridized to its own YAC colony, EW 12G4,

serving as a positive control for the hybridization. However, it is necessary to confirm the overlap by southern blot hybridization due to the possibility of cross-hybridization to sequences contained in the probes. To establish linkage Southern blot hybridization was carried out with total yeast DNA from the clones EG19C10 and EW12H5. The results are shown in Fig. 3. The overlapping YAC clones (EG19C10 for left end and EW12H5 for right end) have identical hybridization pattern with total genomic DNA. This process was repeated with all the initial YAC clones corresponding to the cosmid RFLP markers. The results of the chromosome walking to each locus are shown in Fig. 4.

Application of the physical map for other studies

As a first application of the physical map we decided to isolate a genetic locus in the chromosome 2 of *Arabidopsis* (Loornneef et al., 1983). The genetic loci, *er* and *hyl*, are located near one of the RFLP marker 6842 on the chromosome 2 (Nam et al., 1989) and we constructed a YAC contig around the maker 6842. The YAC contig, shown in Fig 5, is a 1.3 Mb in size and covers the *er* and *hyl* loci and two known RFLP markers, 6842 and Gpa on chromosome 2 of *Arabidopsis*. YAC clones were identified by the RFLP marker 6842 first and used as starting points for construction of the YAC contig. Overlapping YAC clones have been isolated bi-directionally from two YAC libraries by a hybridization method by using end fragments of YAC inserts as describe above. A detailed Bam H1 restriction map has been constructed throughout the 1.3 Mb region (data not shown). By using recombinant plants from a cross between Lan, *er/hyl* and Col. and end fragments of YAC inserts which show RFLPs we have investigated the correlation between genetic map and physical map around the *er* and *hyl* region at the theoretical resolution of 0.1 cM in the 1.3 Mb region. The recombination events are very much uneven throughout the 1.3Mb region we have investigated: some regions have very rare recombination events and some have one recombination every 25 kb. Using these fine genetic map we located the *er* and *hyl* loci within 20 kb regions and attempted to complementation of the *er* and *hyl* genetic loci as shown in Fig 6.

At present the work is in progress to pin point the exact location of the *er* and *hyl* genetic loci in a DNA fragment and to know the identity of the genetic loci.

Conclusion

Recently developed YAC cloning system offers great advantages for the purposes of the construction of physical maps or chromosome walking. The YAC cloning system permits cloning of large fragments of DNA and provides

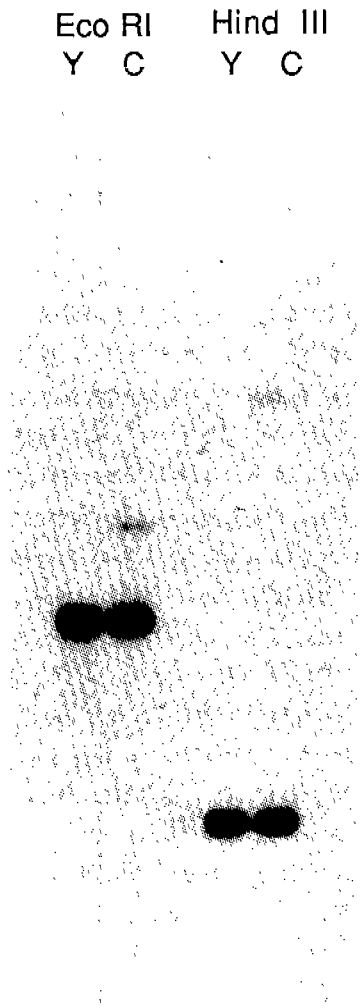


Figure 3. Southern blot hybridization
Genomic and yeast DNAs were digested by restriction enzymes indicated and separated by agarose gel electrophoresis. The DNA was transferred to a nylon membrane and used for hybridization. The end products were labeled by PCR and used for hybridization at a strigent condition. Y and C indicate yeast and *Arabidopsis* Columbia DNA.

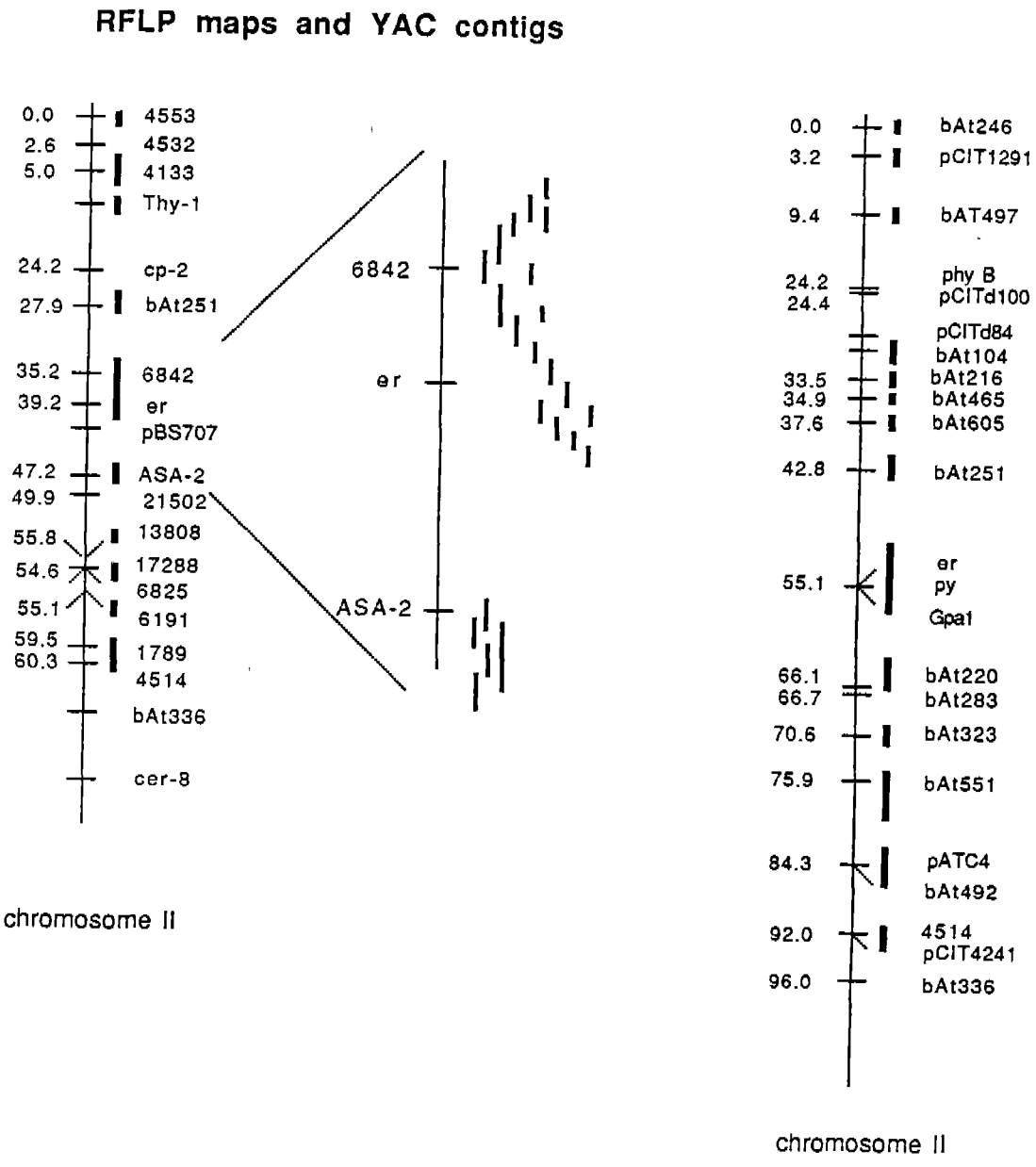
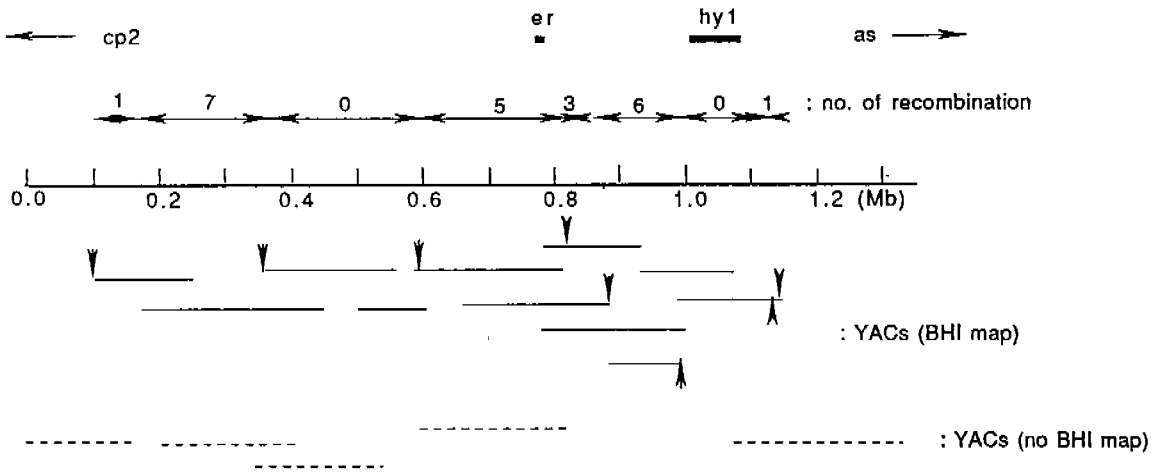


Figure 4. YAC contigs on chromosome 2

The long thin lines and small bars next to them represent chromosome and contigs on the chromosome 2. The numbers on the left of the chromosome indicate distances from the top of chromosome in cM unit and the numbers or English letters on the right of the chromosome indicate RFLP markers.



er to cp2 : 95 recombination/10.5 cM - expected resolution : 21.5 Kb
er to hy1 : 9 recombination/ 1.1 cM - expected resolution : 24 Kb
hy1 to as : 92 recombination/ 16.3 cM -expected resolution : 35.4 Kb

Figuree 5. A fine map around the *er* and *hy1* region
The contig and physical distance are shown schematically. The directions of the chromosome are indicated by *cp2* and *as* which are on the upper and lower arms of chromosome 2, respectively. The numbers within the arrows indicate recombination events among 90 recombinant plants. The thin lines represent YAC clones and the arrown heads pointing downward indicate end clones which give polymorphisms between Columbia and Landsberg ecotypes of *Arabidopsis*.

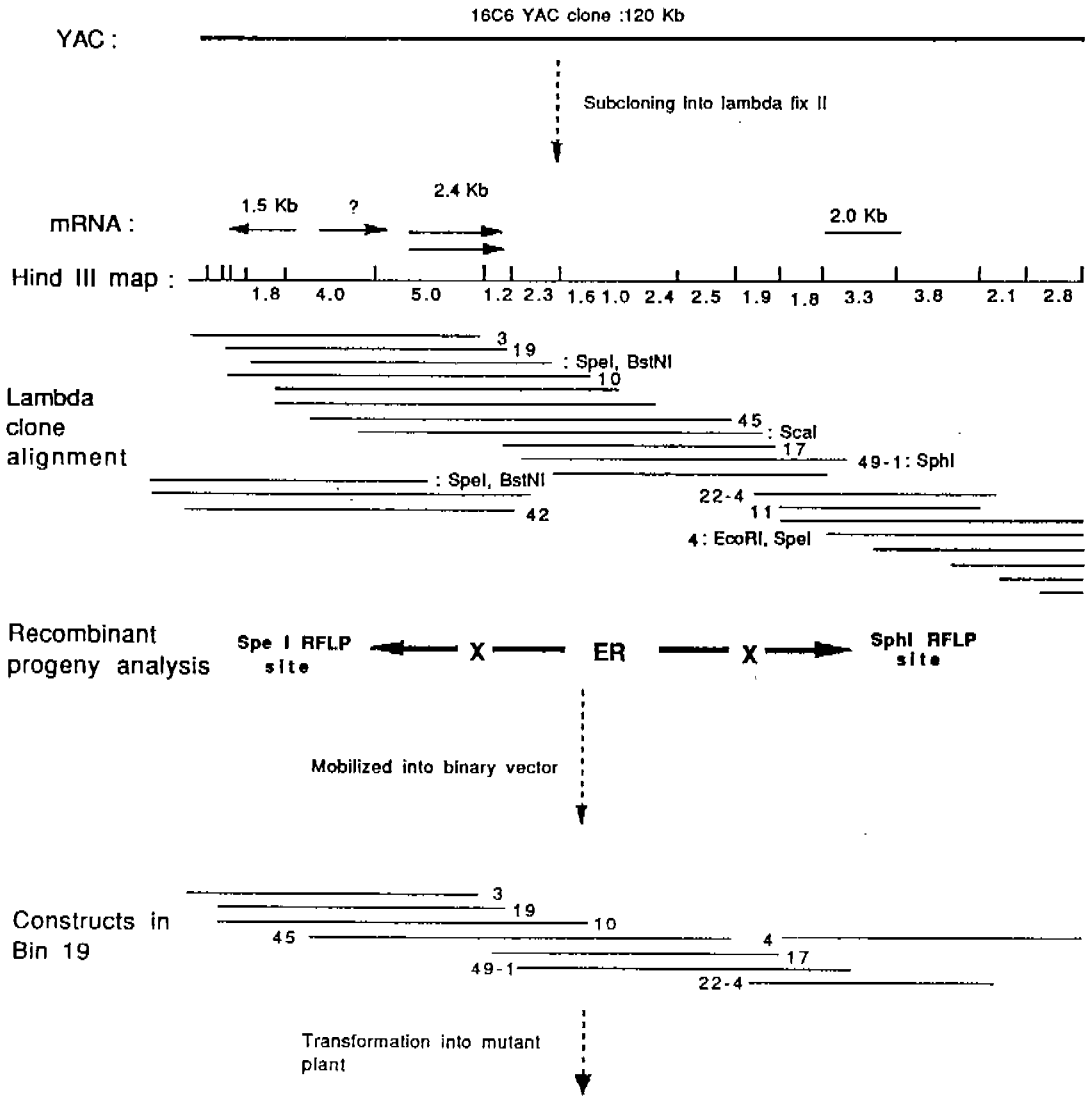


Figure 6. Location of the *er* locus and complementation of the *er* locus
One of the YAC clones containing the *er* locus was subcloned into lambda fix II vector and used to locate the *er* locus. The lambda subclones were aligned by the fingerprinting method and each clone were tested for polymorphism. The restriction enzymes and clones which give polymorphism are indicated. The representative clones were mobilized into pBIN19 binary vector and used for complementation experiments.

better representation of clones compared to other conventional cloning systems based on *E. Coli* (Burke et al., 1987; Schlessinger, 1990). There have been a few attempts to take advantage of the YAC system for physical mapping(Coulson et al., 1988; Hwang et al., 1991). We developed a new simple method for chromosome walking using the YAC clones and its successful amplification to the construction of YAC physical map of *Arabidopsis* genome. The approach we have taken is to use end fragments of YAC inserts as hybridization probes to screen for overlapping clones. The end fragments which could maximize walking distance each step were isolated IPCR (Ochman et al., 1988) and directly PCR labeled to high specific activity. The probes were then used to hybridize YAC colony blots to identify overlapping clones. Also we optimized condition in order to eliminate purification steps. Therefore, by this approach, it minimizes the necessary steps to identify overlapping clones. Especially this direct hybridization approach to screen for overlapping clones using end fragments as probes will be very attractive for chromosome walking or physical mapping of reasonably small genome in size or a subgenomic library of large genome such as chromosome-specific library where several genomic equivalents of YAC clones can be easily plated on two or three membranes for the hybridization. By this approach we were able to walk over 1 Mb in the *Arabidopsis* genome.

Using this method we constructed 15 YAC contigs which may cover approximately 50% of chromosome 2 of *Arabidopsis* genome. We also attempted to use the YAC contigs to isolate known genetic loci such as er and hyl. From this attempt we have located the *er* and *hyl* loci within 20 kb regions and identified few transcripts within the candidate regions.

References

Bancroft, I. and Wolk, C.P. (1988). Pulsed homogeneous orthogonal field gel electrophoresis (PHOGE). Nuclei Acids Res. 16, 7405-7418.

Bender, W., Spierer, P., and Hogness, D.S. (1983). Chromosome walking and jumping to isolate DNA from the Ace and rosy loci and Bithorex complex in Drosophila melanogaster. J Mol. biol. 168, 17-33.

Burke, D.T., Carle, G.F., and Olson, M.V. (1987). Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. Science 236, 806-812.

Chang, C., Bowman, J.L., DeJohn, A.W., Lander, E.S., and Meyerowitz, E.M. (1988). Restriction fragment length polymorphism linkage map for

Arabidopsis thaliana. Proc. Natl. Acad. Sci. USA 85, 6856-6860.
Coulson, A., Suston, J., Brenner, S., and Karn, J. (1986). Toward a physical map of the nematode Caenorhabditis elegans. Proc. Natl. Acad. Sci. USA 83. 7821-7825.
Coulson, A., Waterson, R., Kiff, J., Sulston, J., and Kohara, Y. (1988). Genome linking with yeast artificial chromosomes. Nature 335, 184-186.
Cross, S. H., and Little, P.F.R. (1986). A cosmid vector for systematic chromosome walking. Gene 49, 9-22.
Grill, E., and Somerville, C. (1991). Construction and characterization of a yeast artificial chromosome library of Arabidopsis which is suitable for chromosome walking. Mol. Gen. Genet 226, 484-490.
Hauge, B.M., and Goodman, H.M. (1991). Mapping the Arabidopsis genome. In: Koncz, C., Chua, N.-H., Schell, J. (eds) Methods in Arabidopsis research. World Scientific, Singapore, pp 191-224.
Heard, E., Davies, B., Feo, S., and Fried, M. (1989). An improved method for the screening of YAC libraries. Nucleic Acids Res. 17, 5861-5864.
Hwang, I., Takayuki, K., Hauge, B.M., Goodman, H.M., Schmidt, R., Cnops, G. Dean, C., Gibson, S., Iba, K., Lemieux, B., Arondel, V., Danhoff, L., and Somerville, C. (1991). Identification and map position of YAC clones comprising one-third of the Arabidopsis genome. Plant J. 1, 367-374.
Koornneef, M., van Eden, J., Hanhart, C.J., Stam, P., Braaksma, F.J., and Feenstra, W.J. (1983). Linkage map of Arabidopsis thaliana. J. Hered. 74, 265-272.
Leutwiler, L.S., Hough-Evans, B.R., Meyerowitz, E.M. (1984). The DNA of Arabidopsis thaliana. Mol. Gen. Genet. 194,15-23.
Nam, H-G., Giraudat, J., den Boer, B., Moonan, F., Loss, W..D.B., Hauge, B.M., and Goodman, H.M. (1989). Restriction fragment length polymorphism linkage map of Arabidopsis thaliana. Plant Cell 1, 699-705.
Ochman, H., Gerber, A.S., and Hartl, D.L. (1988). Genetic application of an inverse polymerase chain reaction. Genetics 120, 621-624.
Riley, J., Ogilvie, D., Finniear, R., Jenner, D., Powell, S., Anand, R., Smith, J.C., and Markham, A.F. (1990). A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. Nucleic Acids Res. 18, 2887-2890.
Schlessinger D. (1990). Yeast artificial chromosomes : Tolls for mapping and analysis of complex genomes. Trends Genet. 6, 248-259.
Ward, E.R. and Jen, J.C. (1990). Isolation of single-copy-sequence clones from a yeast artificial chromosome library of randomly sheared Arabidopsis thaliana DNA. Plant. Mol. Biol. 14, 561-568.