J. Plant Biol. 40 (3) : 141~148 (1997)

This paper was presented at the 11th Symposium on Plant Biotechnology entitled "Plant Genes and Genetic Resources" organized by Gynheung An, held July 4-5, 1997, by the Botanical Society of Korea

Current Progress in the *Arabidopsis* Genome Project at Home and Abroad

Hwang, Inhwan*

Plant Molecular Biology and Biotechnology Research Center,
Department of Plant Molecular Biology, Gyeongsang National University, Chinju, 660-701, Korea

Arabidopsis, a small flowering plant, has become the model system in which to decipher the genetic information contained in a plant genome. After physical maps of the *Arabidopsis* genome have been constructed, now the genome project of the *Arabidopsis* is entered its final stage, i.e. the sequencing of the entire *Arabidopsis* genome. An international consortium, the Arabidopsis Genome Initiative, has been set up to coordinate the large scale sequencing of the *Arabidopsis* genome. This review article describes the progress of the *Arabidopsis* genome project and summarizes the current status of the sequencing effort.

ARABIDOPSIS AS A MODEL SYSTEM FOR THE PLANT GENOME PROJECT

Although plants have been favored subjects for various aspects in classical biology for many decades, they were neglected in the 70's and early 80's, when there was a great deal of advancement achieved towards understanding the molecular biology in animal and bacterial systems. Interest in the plant as a molecular biology system was rejuvenated in the mid 80's. With the renewed interest in plant biology, researchers started looking for a model system for a new field of study, the molecular biology of plants, in analogy to the animal system, where *Drosophila* has been a powerful model in which to study and elucidate the complex biological mechanisms that are spelled out in the blueprint of the genome.

Arabidopsis thaliana is a small flowering plant of the family of Cruciferae with a world-wide distribution (Redei, 1969). This small weed has gained much appreciation in classical and molecular genetic studies of plants because, unlike traditional crop species, it has many features that make it well suited for molecular genetics. It has a short life cycle of approximately 6 weeks and a large seed output. Owing to its small size, a large number of M2 plants of mutagenized parent plants can easily be screened in a small space, even in a petri dish (reviewed in Meyerowitz, 1984; 1987). Furthermore, *Arabidopsis* has the smallest genome size, approximately 100 Mb, known for flowering plants with little repetitive DNA (Redei and Hirono, 1964; Meyerowitz, 1994). In addition, by the mid 80's numerous morphological and biochemical mutants had been isolated and mapped on the 5 chromosomes (Koornneef *et al.*, 1983). With these advantages, it is not surprising that *Arabidopsis* became very popular as a model system for molecular genetics in plants.

CONSTRUCTION OF RFLP MAPS

As an initial step in the genome project of *Arabidopsis*, two groups, Meyerowitz's group at CalTech and Goodman's group at Harvard Medical School, have been engaged in the construction of a new kind of a map, the restriction fragment length polymorphism (RFLP) map. Lambda or cosmid clones were used to search for RFLPs between two ecotypes, Columbia and Lan,er. More than 300 RFLP markers have been mapped in three different RFLP maps (Chang *et al.*, 1988; Nam *et al.*, 1992; Lister and Dean, 1993). These maps also include a few genetic markers that provide contact points between genetic and RFLP maps. The maps have been widely used to clone genes identified by mutations (Arondel *et al.*, 1991; Giraudat *et al.*, 1992). Later, these markers have been utilized as probes to isolate YAC clones in the construction of a physical map (Hwang *et al.*, 1991; Schmidt *et al.*, 1995; Zachgo *et al.*, 1996). An example of an RFLP map is shown in Fig. 1. By now, the RFLP maps and the classical genetic map have been integrated into a unified map, making it

*Corresponding author: Fax +82-591-759-9363

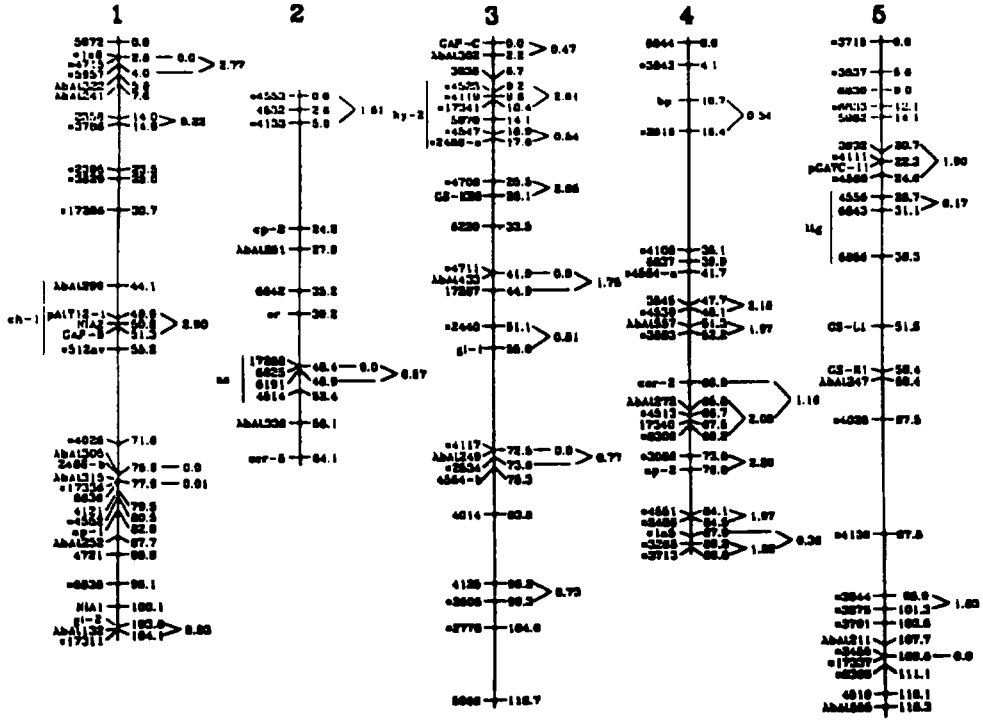

Fig. 1. RFLP linkage map of *Arabidopsis*. The linkage map was generated by Nam *et al.* (1992).

much more convenient to use in a map-based cloning approach (Hauge *et al.*, 1993). In addition to the RFLP maps a variety of other maps have been constructed using PCR-based markers such as simple sequence length polymorphisms (SSLPs) (Bell and Ecker, 1994), cleaved amplified polymorphic sequences (CAPS) (Konieczny and Ausubel, 1993), and random amplified polymorphic DNAs (RAPDs) (Reiter *et al.*, 1992).

CONSTRUCTION OF A PHYSICAL MAP OF THE *ARABIDOPSIS* GENOME

Another major effort in the Arabidopsis genome project was given to the construction of a physical map of the entire genome. Goodman's group at Harvard Medical School has been engaged in the construction of a complete physical map of the *Arabidopsis* genome using cosmid clones and the fingerprinting method (Coulson *et al.*, 1986; Hauge and Goodman, 1992). Cosmid clones were isolated, digested with *Hin*dIII restriction endonuclease, labeled by reverse transcriptase in the presence of ^{35}S-α-dATP, and finally digested again with *Sau*3AI restriction endonuclease as shown in Fig. 2. The digestion products were then separated on a sequencing type polyacrylamide gel. The fingerprints of the cosmid clones

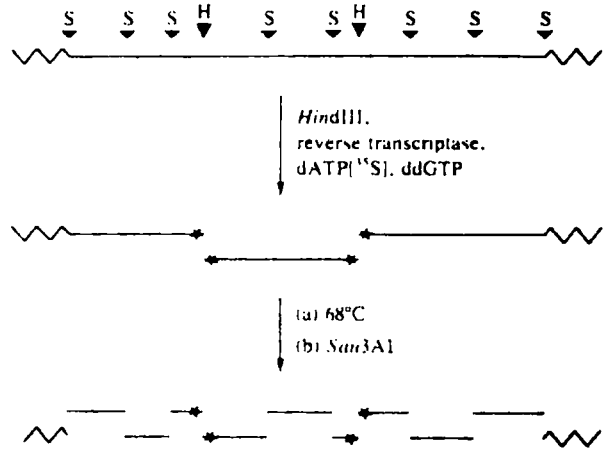

Fig. 2. Fingerprinting of a cosmid DNA. Cosmid DNA was isolated from *E. coli* and digested with *Hin*dIII in the presence of ^{35}S-α-dATP and reversetrascriptase. The enzyme was heat inactivated, and the DNA was subsequently digested with *Sau*3AI. The resulting fragments were then size separated in a sequencing type polyacrylamide gel.

were scored by scanning the autoradiograms and analyzing them with a computer program. Subsequently the patterns of the individual cosmid clones were assembled into 'contigs' using overlapping characteristic banding patterns of overlapping cosmid clones as shown in Fig. 3. More than 20,000 clones were

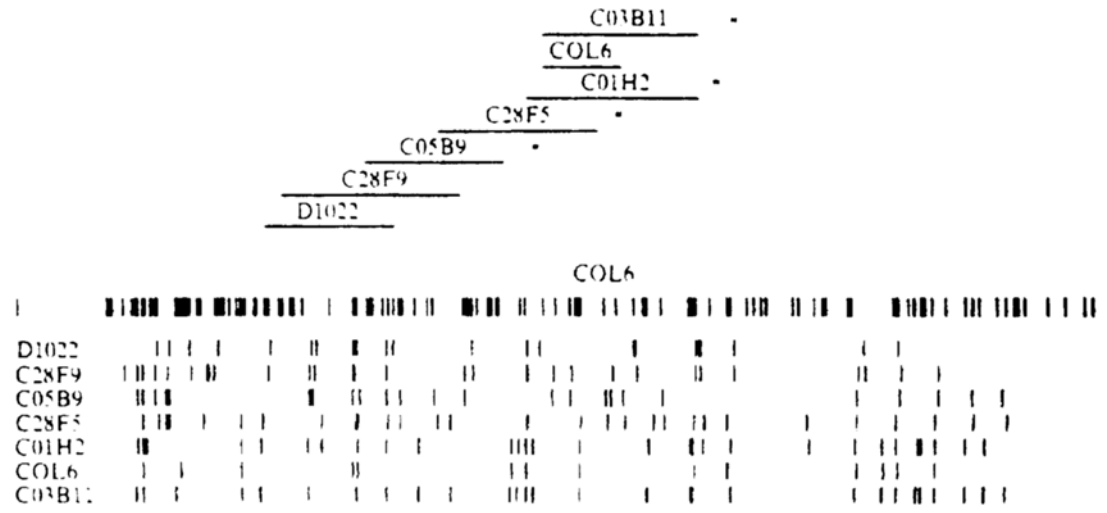

Fig. 3. Assembly of fingerprinted cosmid clones. The autoradiograms of cosmid clones were scanned and assembled into a contig using a computer program.

fingerprinted and assembled into approximately 750 contigs which encompass about 90,000 kb representing 90% to 95% of the *Arabidopsis* genome (reviewed by Goodman *et al.*, 1995). However, using cosmid clones in the construction of the physical map had some limitations. In fact, such limitations were anticipated because of problems inherent to a cosmid library such as bias in the representation of clones and recombination events that can take place within the *E. coli* host, although care has been taken to eliminate this problem by carefully selecting for a host, an *E. coli* strain lacking recombinational activity. Also, a significant number of clones had internal deletions. These deletion clones were a big problem, because they resulted in false linkages. Burke *et al.* (1987) have developed a new cloning system that can be used as an alternative for physical mapping. This system utilizes yeast as a host for library construction. The cloning vector, the yeast artificial chromosome (YAC) vector, has of two selectable markers, two teleomeres, a centromere, and an ARS. Thus, a YAC clone is a linear molecule much like the natural chromosomes in yeast cells. The new cloning system offered enormous advantages for the construction of a physical map. The YAC vector, in theory, does not have an upper limit for the insert size. Also, it was believed that repetitive regions of eukaryotic DNA would be more stable in yeast than in *E. coli* and therefore less likely to suffer internal deletions. It was also expected that a YAC library would result in a better representation of all cloned DNA fragments. Because of these advantages, the new cloning system has quickly gained popularity and has been adopted as the library of choice for the construction of a physical map for a variety of organisms ranging from human to plant (Coulson *et al.*, 1988; Garza *et al.*, 1989; Hwang *et al.*, 1991). For *Arabidopsis* two YAC libraries were constructed, one by Ward and Jen (1990) and the other by Grill *et al.* (1991). With an average insert size of 150 kb, only 1,000 YAC clones are needed to cover the entire *Arabidopsis* genome taking 100 Mb to 150 Mb for its genome size. Since the two libraries have about 4,500 clones, they represent approximately six genomic equivalents. Recently two additional YAC libraries, yUP and CIC, have been constructed that have even bigger inserts than the two previous libraries. The characteristics of the libraries are summarized in Table 1.

Table 1. Caracteristics of *Arabidopsis* YAC libraries

Library name	Average insert size	No. of clone	Ecotype
EG	150 kb	2,300	Columbia
EW	150 kb	2,200	Columbia
yUP	250 kb	2,300	Columbia
CIC	450 kb	1,150	Columbia

The efforts to construct a physical map for the entire *Arabidopsis* genome has been internationally coordinated to minimize redundancy. The strategy for the construction of the physical map using the YAC clones is shown in Fig. 4. Initially YAC clones corresponding to RFLP markers were isolated by the colony hybridization method as shown in Fig. 5. Overlapping clones were rescreened using the end fragment of the YAC clones rescued by the inverse PCR method (Ochman *et al.*, 1988). As a first int-

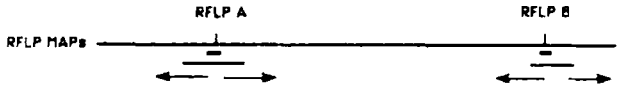

Fig. 4. Strategy of physical map construction using YAC clones. RFLP markers are used as a hybridization probe to screen a YAC library. Subsequently the end fragments of YAC clones are isolated by IPCR and used to isolated overlapping YAC clones to construct a physical map.

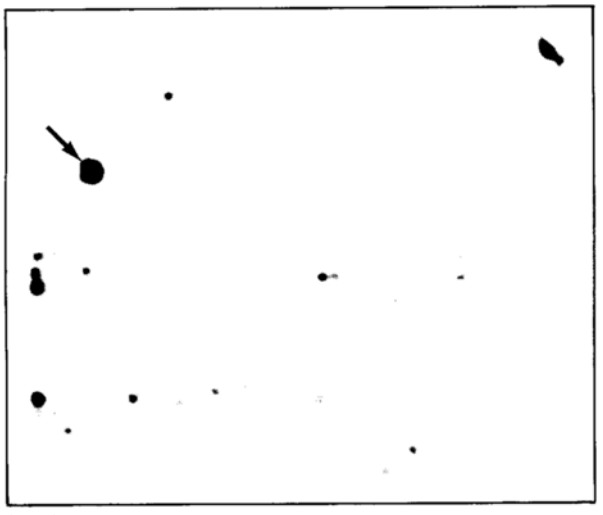

Fig. 5. Screening of a YAC library. The YAC clones were grown on a nylon membrane and processed to lyse the cells. The DNA was UV cross-linked to the membrane which was then used in hybridization to a ^{32}P-labeled probe such as an RFLP marker.

ernational effort to construct a physical map of the *Arabidopsis* genome, YAC clones corresponding to approximately 300 RFLP markers have been isolated from the EG and EW YAC libraries applying the strategy described above (Hwang *et al.*, 1991). It has been estimated that these YAC clones cover one third of the entire genome. Subsequently the effort for the construction of the physical map has been divided in such a way that individual groups focus on particular chromosomes instead of taking a random approach. Three groups have been involved in the construction of the physical maps. Chromosomes 1 and 3 were to be constructed by Ecker's group at the University of Pennsylvania, chromosome 2 by Goodman's group at Harvard Medical School, and chromosomes 4 and 5 by Dean's group at the John Inners Center. Initially EG and EW YAC libraries were widely used for the construction of the physical maps. However, because of the smaller insert size and the chimeric nature of the YAC clones in the EG and EW libraries, the other two YAC libraries, CIC and yUP, have become more useful. By now,

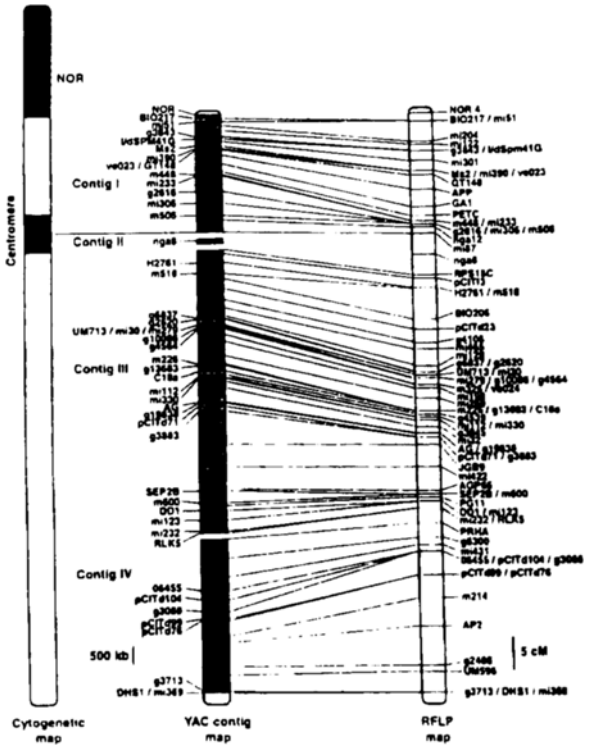

Fig. 6. Physical map of chromosome 4 of *Arabidopsis*. The picture was obtained from Schmidt *et al.* (1995).

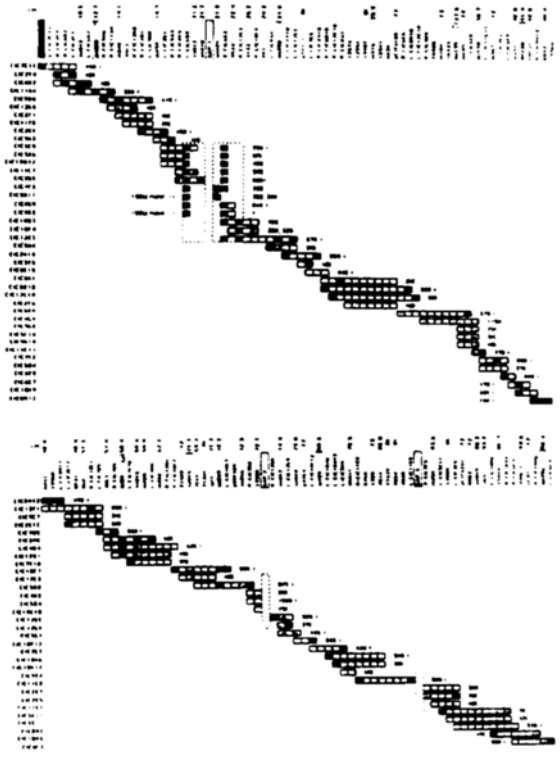

Fig. 7. Physical map of chromosome 2 of *Arabidopsis*. The picture was obtained from Zachgo *et al.* (1996).

Table 2. Physical map of the *Arabidopsis* genome

Chromosome	Research Team	Current status
1	Joe Ecker	42 contigs (65%)
2	Howard Goodman	4 contigs (90%)
3	Joe Ecker	42 contigs (60%)
4	Carolin Dean	4 contigs (95%)
5	Carolin Dean	35 contigs (85%)

the physical maps for chromosome 2 and chromosome 4 have been published (Schmidt *et al.*, 1995; Zachgo *et al.*, 1996) (Fig. 6 and 7). The physical map of chromosome 4, constructed by Dean's group in John Inners Center, is nearly complete (95%) with only 3 three small gaps. The physical map of chromosome 2, constructed by Goodman's group at Harvard Medical School, is also nearly complete (90%) with only 3 small gaps (Fig. 7). Construction of the physical maps of the other chromosomes is under way. At this moment, more than 60% of chromosomes 1,3, and 5 have been linked. Table 2 summarizes the current status of the physical maps of *Arabidopsis* chromosomes. In the meantime, new bacterial artificial chromosome (BAC) libraries became available and are now used to construct detailed maps. The advantage of the BAC library lies in the easy of manipulation, but the insert size is smaller than it is in a YAC library (Shizuya *et al.*, 1992). The physical maps contained in YAC libraries are now refined with BAC clones proceeding to the next stage of the genome project, the sequencing of the entire genome.

SEQUENCING OF THE *ARABIDOPSIS* GENOME

With progress in the construction of the physical map of the *Arabidopsis* genome and with the technological advancements in sequencing, the focus of the genome project has now shifted to the sequencing of the entire genome. In Europe, a concerted effort has been made to sequence a large segment of the genomic DNA in 17 laboratories. This project, termed ESSA (European Scientific Sequencing *Arabidopsis*), is considered a pilot project and is supported by the European Union with $7.5 million over 2 years as shown in Table 3. The sequencing project of the entire *Arabidopsis* genome has gained momentum, when the National Science Foundation (USA) decided to support the sequencing of the *Arabidopsis* genome as a model system for plants as a part of the human genome project in 1996 (Table 3) (Kaiser, 1996). The US government provided grants to 3 consortia, SPP, CSH-WU, and TIGR, for sequencing the *Arabidopsis* genome. The SPP consortium is led by Ron Davis and encompasses the Human Genome Center at Stanford University, the Plant Gene Expression Center of the University of California at Berkeley, and the *Arabidopsis* Genome Center at the University of Pennsylvania. The CHS-WU-ABA consortium led by Richard McCombie consists of the Cold Spring Harbor Laboratories, Washington University, and Applied Biosystems (USA). Finally, the third team is led by Craig Venter at The Institute of Genomic Research. These US consortia bring the total number of sequencing groups to five teams sequencing the *Arabidopsis* genome including a group at the Kazusa Institute in Japan headed by Satoshi Tabatat and the aforementioned European consortium headed by Mike Bevan at the John Inners Center of UK. This massive sequencing effort brought up the need for coordination to minimize redundancy in the sequencing, which resulted in an organization for international collaboration in sequencing the *Arabido-*

Table 3. International *Arabidopsis* Sequencing Initiatives

Lead scientist	Laboratories	Funding	Chromosome	Rate (kb/Mo)
Richard McCombie (CHL)	Cold Spring Harbor Lab Washington Univ. Applied Biosystem (USA)	$4.2 million/3yr	4,5	150
Ron Davis (Stanford Univ.)	Stanford Univ. USDA/U C-Berkely U. Penn	$3.8 million/3yr	1	150
Craig Venter (TIGR)	The Institute of Genome Research (USA)	$4.7 million/3yr	2	220
Satoshi Tabata (Kazusa Institute)	Kazusa DNA research Institute (Japan)	$4.5 million/yr	3,5	500
Mike Bevan (John Inners Center)	17-lab consortium of European Union	$7.5 million/2yr	4,5	200

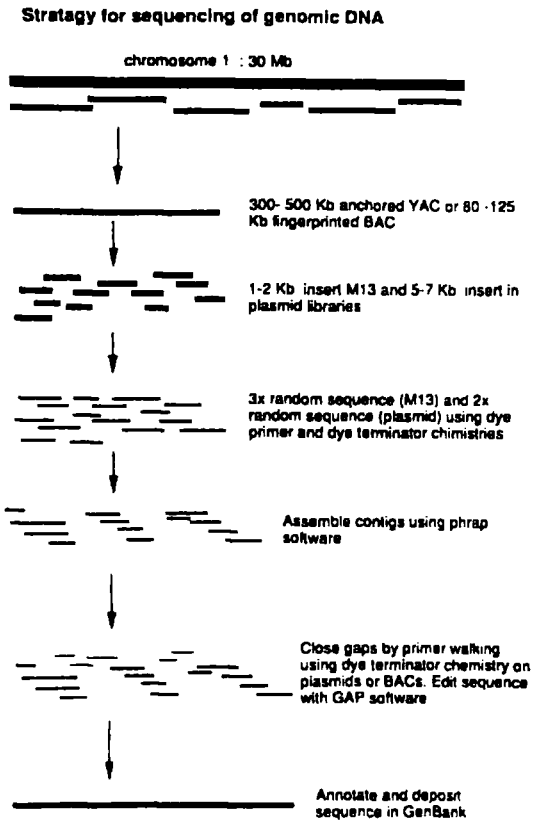

Fig. 8. Sequencing strategy for the *Arabidopsis* genome. The inserts of YAC or BAC clones are subcloned into pBluescript or M13 vector. The clones are then randomly sequenced (2 times for pBluescript clones and 3 to 4 times for M13 clones) by the automatic sequencing method and sequences are assembled into contigs. The remaining gaps are to be filled in by primer-directed sequencing. This picture is from the stratagy of the SPP consortium.

psis genome, the *Arabidopsis* genome initiative (AGI). Thus, the international AGI consists of five sequencing teams: SPP, CHS-WU-ABA, TIGR, Kazusa Institute of Japan, European consortium (Table 3). These groups have divided and distributed the sequencing task in order to avoid overlapping: chromosome 1 to SPP, chromosome 2 to TIGR, chromosome 3 and part of chromosome 5 to the Kazusa Institute, parts of chromosomes 4 and 5 to CSH-WU-ABI, and parts of chromosomes 4 and 5 to the European consortium (Kaiser, 1996). The sequencing project aims to finish 40% of the genome during the first three years. The strategies employed in the sequencing by the different groups are similar. An example of the sequencing strategy is shown in Fig. 8. From a physical map, the inserts of YAC clones or BAC clones are to be subcloned into M13 or pBluescript. Then the subclones are randomly sequenced 2 times for a pBluescript clone and 3 to 4 times for M13 clones. The sequences are then assembled into contigs. The gaps are filled in by sequencing using specific primers. The small sequences are finally assembled into a contiguous sequence. The research team expects 99.9% accuracy. Currently each consortium has sequenced approximately 1-2 Mb of genomic DNA.

ARABIDOPSIS EXPRESSED SEQUENCE TAGS

Although the entire sequence of the genome will be an invaluable resource in the investigation of the biology of an organism, it requires a large amount of money as well as time. In an alternative strategy to speed up the process, there has been an effort to isolate as many genes as possible by randomly sequencing cDNA libraries. Expressed sequence tags (ESTs) of *Arabidopsis* have been generated by two groups (Table 4). One of them is a French consortium and the other is at Michigan State University/USDA. Both groups have generated a variety of cDNA libraries from various tissues of *Arabidopsis* as well as *Arabidopsis* treated with pathogens. Random sequencing has produced now more than 30,000 ESTs which are thought to cover approximately 70% of all *Arabidopsis* genes (Höfte *et al.*, 1993; Newman *et al.*,

Table 4. Current status of Arabidopsis ESTs

Research Teams	Tissues for cDNA library	No. of ESTs
French Consortium	Developing siligues	30,000 ESTs
Michigan State Univ./USDA	Flower Buds	(70% of total Arabidopsis genes)
	Cultured leaf strips	
	Etiolated seeding	
	Tissue culture cells	
	Green shoots	
	Flowering tips	
	Dry seeds	
	Cell suspension cultures treated with pathogen	

1994). Analyses of these ESTs revealed that approximately 40% have significant sequence similarity to known genes (reviewed in Goodman *et al.*, 1995). The ESTs are an invaluable tool when trying to identify a new gene as well as to get an overall picture of the expression patterns of genes. Also, the ESTs will be important in analyzing the genomic sequence obtained from the massive sequencing effort of the genome project.

INFORMATION SHARING AND AVAILABILITY

From the beginning of the genome project there has been concern about the free access to sequence information. There was fear that valuable information may be withheld or screened first by the sequencing group for the purpose of commercial exploitation. Also, the patenting of a sequence has been an issue. In response to such concerns, the sequencing groups have adopted a policy that requires that the sequence will be posted on the Internet as quickly as possible in a form easily usable by all plant scientists. At the moment the sequence information is deposited in the public databases as well as various websites. However, the issue has not entirely been resolved, since each group appears to still practice under a different policy regarding the release of sequence information. For example, there have been complaints by US groups about groups in Japan and Europe not sharing their sequence information as freely as possible. Also, the groups expressed concern about the patenting of simple sequences. At the moment, though, it appears that the sequence information is not going to be patented by the research teams.

FUTURE PROSPECTIVE

It has been proposed that the *Arabidopsis* genome will be sequenced completely by 2004 (Kaiser, 1996). However, current progress of the sequencing forecasts that the *Arabidopsis* genome could be fully sequenced earlier than scheduled. The entire sequence of the *Arabidopsis* genome will be extremely important for the identification of new genes as well as for understanding the biological functioning of the whole genome. We will be able to answer questions pertaining to the organization of the genome and the relationship between structure and function of genes in the context of the genome. In addition, there will be changes in the way we conduct our research. When the complete sequence of the *Arabidopsis* genome is available, we will be able to spend more time to decipher the code with regard to the biological role of the encoded proteins.

The sequencing of the entire genome of *Arabidopsis* abroad will inevitably affect our research at home. It will be necessary to prepare for the entirely new ways of conducting research in molecular biology. Several approaches can be envisioned. One of them would be to obtain as many genes as possible as valuable genetic resources. Already several research groups are actively generating ESTs from various plant systems such as rice, Brassica, and *Arabidopsis* (Park *et al.*, 1993; Lim *et al.*, 1996; Pih *et al.*, 1997). We have been isolating a large number of genes that are specifically induced by osmotic stress (Pih *et al.*, 1997). We believe that our efforts will be more successful, when they are focused on a certain aspect of biology. At the same time we have to actively participate and use the available sequence information as quickly as possible.

ACKNOWLEDGMENTS

The author would like to acknowledge the people at Howard M. Goodman's laboratory who were involved in the genome project. This work was supported by a grant from KOSEF to PMBBRC.

LITERATURE CITED

Arondel, V., B. Lemieux, I. Hwang, S. Gibson, H.M. Goodman, and C.R. Somerville. 1991. Map-based cloning of a gene controlling omega-3 fatty acid desaturation in *Arabidopsis*. *Science* **258**: 1353-1355.

Bell, C.J. and J.R. Ecker. 1994. Assignment of 30 microsatlellite loci to the linkag map of *Arabidopsis*. *Genomics* **19**: 137-144.

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

Chang, C., J.L. Bowman, A.W. De John, E.S. Lander, and E.M. Meyerowitz. 1988. Restriction fragment length polymorphism linkage map for *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **85**: 6856-6860.

Coulson A.R., R.H. Waterson, J.E. Kiff, J.E. Sulston, and Y. Kohara. 1988. Genome linking with yeast artificial chromosomes. *Nature* **335**: 184-186.

Coulson, A., J. Sulston, S. Brenner, and J. Karn. 1986. Toward a physical map of the genome of nematode *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. USA* **83**: 7821-7825.

Creusot, F., E. Fouilloux, M. Dron, J. Lafleuriel, G. Picard, A. Billault, D. Le Paslier, C.D. , M.-E. Chaboute, A. Durr, J. Fleck, C. Camilleri, C. Bellini,

M.Caboche, and D. Bouchez. 1995. The CIC library: A large insert YAC library for genome mapping in *Arabidopsis thaliana. Plant J.* **8**: 763-770.

Garza, D., J.W. Ajioka, D.T. Burke, D.L. Hartl. 1989. Mapping the *Drosophila* genome with yeast artificial chromosomes. *Science* **246**: 641-646.

Giraudat, J., B.M. Hauge, C. Valon, J. Smalle, F. Parcy, and H.M. Goodman. 1992. Isolation of the *Arabidopsis* ABI3 gene by positional cloning. *Plant Cell* **4**: 1251-1261.

Goodman, H.M., J.R. Ecker, and C. Dean. 1995. The genome of *Arabidopsis thaliana. Proc. Nat. Acad. Sci. USA* **92**: 10831-10835.

Grill, E. and C. Somerville. 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 H.M. Goodman. 1992. *In* Method in Arabidopsis Research. C. Koncz, N.-H. Chua, J. Schell. (eds), World Science, Singapore, pp 191-223.

Hauge, B.M., S.M. Hanley, S. Cartinhour, J.M. Cherry, H.M. Goodman, M. Koornneef, P. Stam, C. Chang, S. Kempin, and L. Medrano. 1993. An integrated genetic/RFLP map of the *Arabidopsis genome. Plant J.* **3**: 745-754.

Höfte, H., T. Desprez, J. Amselem, H. Chiapello, M Chaboche, A. Moisan, M.F. Jourjon, J.L. Charpenteau, P. Berthomieu, D. Guerrier, J. Giraudat, F. Quigley, F. Thomas, D.Y. Yu, R. Mache, M. Raynal, R. Cook, F. Grellet, M. Delseny, Y. Parmetier, G.D. Matcillac, C. Gigot, J. Fleck, M. Axelos, C. Bardet, D. Tremousaygue, and B. Lescure. 1993. An inventory of 1,152 expressed sequence tags obtained by partial sequencing of cDNAs form *Arabidopsis thaliana. Plant J.* **4**: 1051-1061.

Hwang, I., T. Kouchi, B.M. Hauge, H.M. Goodman, R. Schmidt, G. Cnops, C. Dean, S. Gibson, K. Iba, B. Lemieux, and C. Somerville. 1991. Identification and map position of YAC clones comprising one-third of the *Arabidopsis* genome. *Plant J.* **1**: 367-374.

Kaiser, J. 1996. First global sequencing effort begins. *Science* **274**: 30.

Koniczny, A. and F.M. Ausubel. 1993. A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. *Plant J.* **4**: 403-410.

Koornneef, M., J. Van Eden, C.J. Hanhart, P. Stam, F. J. Braakasma, and W.J. Feenstra. 1983. Linkage map of *Arabidopsis thaliana. J. Hered.* **74**: 265-272.

Lim, C.O., H.Y. Kim, M.G. Kim, S.I. Lee, W.S. Chung, S.H. Park, I Hwang, and M.J. Cho. 1996. Expressed sequence tags of chinese cabbage flower bud cDNA. *Plant Physiol.* **111**: 577-588.

Lister, C. and C. Dean. 1993. Recombinant inbred lines for mapping RFLP and phenotypic markers in *Arabidopsis thaliana. Plant J.* **4**: 745-750.

Meyerowitz, E.M. 1987. *Arabidopsis thaliana. Annu. Rev. Genet.* **21**: 93-111.

Meyerowitz, E.M. 1994. *In Arabidopsis.* E.M. Meyerowitz and C.R. Somerville (eds.). Cold Sprng Harbor Laboratory Press, Cold Spring Harbor, NY, pp 21-36.

Nam, H.G., J. Giraudat, B. Den Boer, F. Moonan, W. D.B. Loos, B.M. Hauge, and H.M. Goodman. 1989. Restriction fragment length polymorphism linkage map of *Arabidopsis thaliana. Plant Cell* **1**: 699-705.

Newman, T., F.J. de Bruijin, P. Green, K. Keegstra, H. Kende, L. McIntosh, J. Ohlrogge, N. Raikhel, S. Somverville, M. Thomashow, E. Ratzel, and C. Somerville. 1994. Genes galore. A summary of methods for accessing results from large-scale partial sequencing of anonymous *Arabidopsis* cDNA clones. *Plant Physiol.* **106**: 1241-1255.

Ochman, H., A.S. Gerber, and D.L. Hartl. 1988. Genetic application of an iverse polymerase chain reactrion. *Genetics* **120**: 621-623.

Park, Y.S., J.M. Kwak, O.Y. Kwon, Y.S. Kim, D.S. Lee, M.J. Cho, H.H. Lee, and H.G. Nam. 1993. Generation of expressed sequence tags of random root cDNA clones of *Brassica napus* by single-run partial sequencing. *Plant Physiol.* **103**: 359-370.

Pih, K.T., H.J. Jang, S.G. Kang, H.L. Piao, and I. Hwang. 1997. Isolation of Molecular Markers for Salt Stress Responses in *Arabidopsis thaliana. Mol. and Cell* (Korea) (In press).

Redei, G.P. 1969. *Arabidopsis thaliana* (L.) Heynh. A review of the genetics and biology. *Bibliogr. Genet.* **21**: pp 1-151.

Redei, G.P. and Y. Hirono. 1964. Linkage studies. *Arabidopsis Info. Serv.* **1**: 9-10.

Reiter, R.S., J.G.K. Williams, K.A. Feldman, J.A. Rafalski, S.V. Tangey, and P.A. Scolnik. 1992. Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNAs. *Proc. Natl. Acad. Sci. USA* **89**: 1477-1481.

Schimdt, R., J. West, K. Love, Z. Lenehan, C. Lister, H. Thompson, D. Bouchez, and C. Dean. 1995. Physical map and organization of *Arabidopsis thaliana* chromosome 4. *Science* **270**: 480-483.

Shizuya H, B. Birren, U.J. Kim, V. Mancino, T. Slepak, Y. Tachiiri, and M. Simon. 1992. 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**: 8794-8797.

Ward, E.R. and G.C. Jen. 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.

Zachgo, E.A., M.L. Wang, J. Dewdney, D. Bouchez, C. Camilleri, S. Belmonte, L. Huang, M. Dolan, and H.M. Goodman. 1996. A physical map of chromosome 2 of *Arabidopsis thaliana. Genome Research* **6**: 19-25.