TOWARDS ISOLATION OF THE TOMATO ROOT-KNOT NEMATODE RESISTANCE GENE *MI* VIA POSITIONAL CLONING

Raymond A. J. J. VAN DAELEN

Promotor:	Dr. A. van Kammen, hoogleraar in de moleculaire biologie
Co-promotor:	Dr. W.J.T. Zabel, universitair hoofddocent moleculaire biologie

Raymond Antoine Jean Joseph van Daelen

TOWARDS ISOLATION OF THE TOMATO ROOT-KNOT NEMATODE RESISTANCE GENE *MI* VIA POSITIONAL CLONING

Proefschrift
ter verkrijging van de graad van
doctor in de landbouw- en milieuwetenschappen,
op gezag van de rector magnificus,
dr. C. M. Karssen,
in het openbaar te verdedigen
op vrijdag 7 april 1995
des namiddags te half twee in de aula
van de Landbouwuniversiteit te Wageningen

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

van Daelen, Raymond A.J.J.
Towards isolation of the tomato root-knot nematode resistance
gene *Mi* via positional cloning/
Raymond Antoine Jean Joseph van Daelen./[S.1.:s.n.]
Thesis Wageningen. - With summary in Dutch
ISBN 90-5485-372-7
Subject headings: tomato; yeast artificial chromosomes; long range restriction maps.

NN08201, 1911

STELLINGEN

1. Aangezien de stellingen bij proefschriften als enige volledig worden gelezen, mogen deze niet worden afgeschaft maar moet het gebruik veeleer worden gestimuleerd.

2. De veronderstelling, dat tijdens gelijktijdige transfectie van gistcellen met twee moleculen, het molecuul waarop uiteindelijk wordt geselekteerd, het andere molecuul tijdens transfectie bevoordeelt, is onhoudbaar.

Larionov, V., *et al*, Nucleic Acids Res. (1994) 22: 4154-4162.

3. Het gemak waarmee, voor het kloneren van grote DNA fragmenten van planten, van YACs ("yeast artificial chromosomes") naar BACs ("bacterial artificial chromosomes") wordt overgestapt, getuigt niet van een kritische instelling.

Third International Meeting on the Plant Genome, San Diego 1995.

4. De veronderstelling van Gnirke *et al*, dat voor transfectie van YACs naar hamster cellen, fusie met gistsferoplasten te prefereren is boven behandeling met gezuiverd gist DNA, wordt niet door experimentele resultaten ondersteund.

Gnirke, A., *et al*, EMBO J (1991) 10: 1629-1634.

5. Er is niet aangetoond dat "shearing" van DNA leidt tot een willekeurige grootte verdeling van de fragmenten en derhalve beter zou zijn voor constructie van genomische DNA banken dan partiële restrictie digestie van DNA.

Arriata, R., *et al*, Genomics (1991) 11: 806-827.

6. Bij de claim dat deYAC bank van humaan DNA slechts 4 % chimere YACs zou bevatten, gaan Deloukas *et al* geheel voorbij aan het feit dat er naast ligatie-chimeren ook chimeren voorkomen die door *in vivo* recombinatie zijn ontstaan.

Deloukas, P., *et al* Gene (1992) 110: 189-195.
Green, E., *et al* Genomics (1991) 11: 658-669.

7. Omdat het onderkennen en identificeren van genetische faktoren bij het ontstaan van erfelijke ziekten het risico van het krijgen van dergelijke ziekten eerder kleiner dan groter maakt, is het niet correct dergelijke ziekten buiten de ziektekosten verzekeringen te houden.

8. Wanneer we geen, delen van, transgene planten willen eten op grond van de aanwezigheid van het kanamycineresistentie gen moeten we maar vlug stoppen met het eten van sla.

Flavell, R.B., et al, Bio/technology (1992) 10: 141-144.

9. De methoden die gebruikt worden voor het opsporen van criminele handelingen, maakt de aard van de criminele handelingen niet meer of minder crimineel en deze methoden mogen dan ook absoluut geen invloed hebben op de mate en aard van strafvervolging.

Stellingen behorende bij het proefschrift "Towards isolation of the root-knot nematode resistance gene *Mi* via positional cloning".

Raymond van Daelen, Wageningen, 7 april 1995.

aan mijn ouders

CONTENTS

SCOPE OF THIS THESIS

Root-knot nematodes of the genus *Meloidogyne* are severe pathogens of plants and world-wide they cause damage to many economically important crops like potato, rice, cotton, and tomato. So the control of nematodes and the protection of plants against nematode damage are matters of major concern. Some plants carry resistance genes which prevent damage by nematodes. If these resistance genes would become available for introducing into other plant species, this could be of considerable economic importance. From a fundamental point of view, nematode resistance genes are also very interesting since they control the interaction between plants and parasites in a very specific manner. The *Mi*-gene of tomato is a single dominant gene, located on chromosome 6, that confers resistance against several species of root-knot nematodes and may serve as a model for the study of plant-nematode interactions. Such studies, to date, are severely hampered by a lack of knowledge of the proteins and the functions of the proteins encoded by the resistance genes. If the product of the gene is not known, the isolation of such a gene is very difficult. But nowadays several approaches have been developed to isolate genes that are only characterized by phenotype and genetic position. In chapter 1 an overview is presented of the different approaches, their relevance to the isolation of the nematode resistance gene *Mi* is evaluated.

The most relevant procedure for isolating the *Mi*-gene seemed to be positional cloning. This approach involves the identification of tightly linked molecular markers followed by a chromosomal walk which starts at these markers. The walk results in cloning of all sequences between two markers at either side of the gene. The gene itself can subsequently be identified from these sequences using the information encoded in the sequence. In order to proceed with the positonal cloning approach several technical requirements have to be met which were, at the onset of the present work, not available. The development and application of these new techniques is described in this thesis.

First of all, an efficient method for isolating and handling of megabase-sized plant DNA had to be developed, which is described in chapter 2. Furthermore, the technical means to electrophoretically separate large DNA fragments, by so called pulsed field gel electrophoresis, were becoming commercially available at the beginning of this work. In chapter 3, the application of these techniques to physically characterize large genomic regions and the construction of long range restriction maps for two *Mi*-flanking markers, GP79 and *Aps-1*, are described. These maps provided the first physical characterization of parts of the *Mi*-region that was sofar only genetically characterized. The work described in this thesis has focused on *Aps-1* and GP79 since those were the earliest available tightly linked markers. By now many other markers, not available at the onset of the work, are known.

The yeast artificial chromosome (YAC) cloning system allows for the cloning of fragments many hundred kb's in size. Such very large DNA fragments are an essential ingredient for chromosomal walking. The characteristics of YAC cloning are reviewed in chapter 4. For YAC cloning of plant DNA several technical adaptations were necessary. These were worked out in detail in the construction and characterization of a tomato *Cla*I YAC library, as described in chapter 5, and of a partial *Eco*RI YAC library, as described in chapter 6. Finally, in chapter 7, the lessons learned from the excercises described in the chapters 2, 3, 5 and 6 are evaluated, along with a discussion of how to proceed further now the techniques required for positional cloning the *Mi*-gene are available.

CHAPTER 1

GENERAL INTRODUCTION

1.1 NEMATODES AND THEIR ECONOMICAL IMPORTANCE

1.1.1 Introduction

Root-knot nematodes (genus *Meloidogyne*) are small worm-like animals about 1 mm in length which live in soil and feed on plant roots. They occur world-wide and more than 50 species have been found and described to date [89]. The different species have a very wide host range and attack together over 700 plant species of many families. Some species, notably *M. incognita, M. arenaria, M. javanica* and *M. hapla*, infest economically important crop plants and cause major damage. World-wide crop losses have been estimated to accumulate to 5% of the total agricultural production [89]. Particularly in tropical regions, the root-knot nematodes can cause severe damage. The broad host range, wide distribution and large economic losses [87, 88] make nematodes very serious subjects of research.

1.1.2 Pathology

Juvenile root-knot nematodes in their 2nd larval stage can penetrate the roots of the host plants and migrate intercellularly towards the stele to inject a glycoprotein-like substance [53] into cells of, or adjacent to, the vascular tissue, using their stylet. These cells subsequently develop into the so-called "giant cells" which are several times their original volume. The giant cells are formed by a stimulus of mitosis without subsequent cell divisions leading to large multi-nucleated cells, which are metabolically very active and function as food source for the developing nematode. Cells of tissue surrounding the giant cells start to divide and grow to form the galls or root-knots within which the nematodes develop. The nematodes then feed for 3 - 8 weeks on the giant cells thereby undergoing a second molding. After this molding they develop without feeding into the 5th, adult stage. Adult females then resume feeding and remain sedentary, whereas the adult males refrain from feeding, become mobile and leave the roots in search for a female. The entire life cycle can be completed within 6 weeks in which each female can produce between 300 and 1000 offspring [32].

The feeding of nematodes results in a distortion of the functioning of the root system by which plants become more susceptible to stress. Heavily infected plants suffer from a severely impaired root-development that results in a very shallow and knotted root system. This leads to mineral deficiencies and reduced growth, resulting in large crop losses [67].

1.1.3 Control of the nematodes

Over the past forty years nematodes have mainly been controlled by chemical means [32, 83]. Initially, nematicide compounds were obtained as very cheap byproducts from the petrochemical industry and widely applied. However, these compounds turned out to be harmful to man and to have deleterious effects on the environment. Consequently, their use has gradually been abandoned. Currently, only a few nematicides are still allowed to be used, but these are very expensive. As alternatives, non-chemical means are employed to control damage caused by root-knot nematodes. These include tillages such as soil heating of farm land, which is, however, very expensive and submerging of agricultural lands, which is less efficient. Crop rotation, on the other hand, keeps damage by nematodes limited to a large extent. Because of the wide host range of the major root-knot nematodes there are few crops completely resistant and crop rotation by itself is therefore not sufficient. The most economical way to prevent damage by root-knot nematodes is the use of resistant cultivars which, in combination with crop rotation, has proven very effective in keeping the nematode populations under control [68]. Unfortunately, many nematode resistance genes known to date are not effective at elevated temperatures (higher than 25 - 30 °C) and hardly offer a solution to nematode problems in tropical regions [23]. For instance, the resistance of tomato varieties which carry the *Mi*-gene is lost at 34 °C [27].

1.1.4 Properties of resistance

Several defense mechanisms are associated with the resistance against root-knot nematodes. Resistance to the invading nematodes is commonly based on a hypersensitive response (HR). Upon injection of the stylet in the feeding cells, necrosis of these cells occurs within a few hours after penetration of the roots [27, 28, 32]. The nematode then dies either of starvation or through the effects of accumulating toxic compounds. Other responses correlated with the resistance reaction include: accumulation of phenolic compounds [80, 94], induction of the enzyme phenylalanine ammonia lyase (PAL) [14], increased content of hydroxyproline-rich proteins within the cell walls [113], and production of phytoalexins [54].

Specific recognition of the nematodes, that leads to the hypersensitive response, is supposed to be triggered by the interaction of the product encoded by the resistance gene of the plant and a factor encoded by the avirulence gene of the nematode. This interaction, in which a resistance gene product interacts with an avirulence gene product, follows the model described by Flor [35] and is called the "gene for gene" concept. To date, the study of plant nematode resistance mechanisms is greatly hampered by the lack of knowledge of the products of the resistance genes. Cloning of these genes and identification of the resistance gene products will therefore be of great help to elucidate the mechanisms that are involved in resistance.

1.1.5. Nematode resistance breeding

Resistance against root-knot nematodes in tomato was first discovered in 1941 by Bailey *et al* [6] in the wild tomato species *Lycopersicon peruvianum*. To introduce the resistance from *L. peruvianum* into the cultivated tomato *L. esculentum*, the existing compatibility barrier between both species had to be overcome. This was achieved in 1944 by *in vitro* rescue of embryo's, resulting in a hybrid line [95]. From this line, after repeated backcrossings, the first nematode resistant tomato cultivars (Anahu and VFN8) were obtained. These cultivars have subsequently functioned as founder lines for all nematode resistant tomato cultivars currently available [36, 45, 67].

The nematode resistance in tomato was found to be based on a single dominant gene *Mi*, an acronym taken from the first letters of the species *Meloidogyne incognita*. The gene confers very effective resistance to the species *M. incognita, M. javanica* and *M. arenaria* [7]. Yet, the narrow base of only one resistance gene poses a risk and crop rotation in conjunction with stringent sanitation measures remains important to minimize the development of new pathotypes. Some *Meloidogyne* populations capable of breaking the resistance have, in fact, been found [85] but sofar these have not been spreading.

Classical plant breeding for nematode resistance, which involves crosses with resistant wild tomato species, is laborious and very time consuming. Conceivably, once resistance genes have been identified and cloned, recombinant DNA technologies can be used to introduce these genes into existing, nematode sensitive cultivars so as to speed up breeding programs [41]. Furthermore, it is an attractive idea that by these methods incompatibility barriers can be circumvented and that specific genes can be introduced without additional properties [62]. Moreover, resistance might be introduced into crops for which no such genes are known [62]. Genetic engineering, although its feasibility for introducing resistance against organisms other than viruses remains to be proven as yet, seems very promising. The study of resistance mechanisms and isolation of the responsible genes is instrumental to this approach.

1.2 POSSIBLE STRATEGIES TO CLONE THE MI-GENE

Cloning of the *Mi*-gene is severely hampered by the lack of knowledge of the product or function of the gene. Thus, conventional approaches that are based on information of a protein or mRNA product cannot be applied. Similarly, cloning of *Mi* through the complementation of mutations in bacteria and yeast [43] is not possible since the function of the *Mi* gene is unknown. Application of sequences from similar genes as probes in genomic library screening [20, 81] is neither feasible, since no such genes have been identified.

Several methods have been devised for cloning a gene which is only characterized by its phenotype and genetic position. The phenotype of *Mi* is well known, since plants possessing the gene are resistant to root-knot nematodes, whereas plants without *Mi* are not. Furthermore, *Mi* has been mapped to chromosome 6, at a location close to the centromere and in tight linkage to an allele of the acid phosphatase isozyme *Aps-1*^{1} [44]. In the following, some possible strategies and their relevance for cloning the *Mi*-gene will be discussed.

1.2.1 Positional cloning.

The procedure of positional cloning [103] is in principle applicable to every gene of which the phenotype can be identified and the genetic position is known. The procedure involves the determination of the position of the gene on a genetic map as precisely as possible, followed by the identification of molecular markers, such as restriction fragment length polymorfisms (RFLPs), that are tightly linked to the gene. Starting from these markers a chromosomal walk is performed to clone the region in which the gene of interest is located. Finally, the gene itself must be identified among the candidate sequences by complementation of the corresponding mutant phenotype.

The first genes to be cloned following chromosomal walking were *Drosophila* homeotic genes from the bithorax complex [9]. The first human gene to be cloned by positional cloning was the human retinoblastoma (RB) gene involved in childhood tumors [60]. Since then many other human disease genes and genes from many other organisms, have been isolated by this approach. Recently, several plant genes have been cloned by positional cloning, including the *Arabidopsis* abscisic acid insensitive locus *ABI3* [47], a gene from *Brassica napus*, controlling omega-3 fatty acid desaturation in *Arabidopsis thaliana* [5], the *Arabidopsis thaliana* auxin-resistance gene *AXR1* [77], and the *Pto*-gene of tomato which confers resistance against races of *Pseudomonas syringae* [65].

Isolation of the *Mi*-gene via positional cloning seemed feasible as it is tightly linked to the isozyme marker *Aps-1*^{1}, at a genetic distance which appeared to be less than 1 cM [67]. Besides, the availability of nearly isogenic lines for *Mi* should allow the isolation of additional markers. Furthermore, the experimental means required for positional cloning were rapidly evolving. In section 1.5 the procedure of positional cloning and its application to *Mi* will be discussed in more detail.

1.2.2 Genomic substraction

Genomic substraction [8] can be used to clone sequences corresponding to deletions. The procedure involves the mixing of wild type DNA with an excess of DNA from a deletion

mutant that lacks the gene sequences in question, followed by denaturation and reassociation of the DNA mixture. Reannealed molecules containing one or both strands from the mutant are removed. As the relevant DNA has no complement in the mutant DNA it will be enriched in the remaining DNA. The process is repeated several times to obtain a large enrichment (up to 1000 times) of the target DNA. This approach has successfully been applied by Sun *et al* [98] in cloning the GA1 locus of giberellin-responsive dwarfs of *Arabidopsis thaliana*. The method can only be applied if deletion mutants are available. As such mutants are not known for *Mi*, the genomic substraction strategy did not seem a feasible approach to clone the gene.

1.2.3 Differential cDNA screening

When mRNAs can be isolated from two sources differing only in expression of the target gene differential screening of a cDNA library can be applied. As to *Mi*, this would imply the comparison of mRNAs made from two nearly isogenic lines with and without the *Mi* gene, respectively. A cDNA library from the *Mi*-resistant line is then probed successively with the mRNAs from the susceptible and the resistant line. Clones that are identified with cDNA from the resistant, but not with cDNA from the susceptible plant, will represent sequences only present in resistant plants and are thus candidates for representing *Mi*. Although nearly isogenic lines of tomato differing in the *Mi*-locus are available, some technical aspects of differential cDNA screening make it doubtful whether this method can be applied with any chance of success. Firstly, the approach is only effective in isolating highly expressed sequences which occur in the cDNA libraries in a high copy number. The level of expression of *Mi* is unknown and *Mi* might even not be expressed constitutively. Experiments to achieve cloning *Mi* via differential cDNA screening was unsuccesfull, suggesting that *Mi* is not highly expressed [Hulsebos & Zabel, unpublished results]. If expression of the gene is induced by nematodes, plants attacked by nematodes must be used with the disadvantage that many other disease-related genes will also be expressed and may be detected in the differential screening of the library. Secondly, since the susceptible and the resistant lines are not perfectly isogenic other genes are likely to be detected.

1.2.4 Gene tagging by insertional mutagenesis

The method of gene tagging by insertional mutagenesis utilizes the insertion of a defined DNA sequence -the tag- into a target gene. Insertion of a tag into a gene, or in regulatory sequences, can alter its expression, resulting in the appearance of a changed phenotype. The gene itself can then be identified with techniques that rely on the detection of the introduced, well

characterized, tag sequences. Essentially two types of elements have been used for tagging strategies in plants: transposons and T-DNA.

Transposon tagging. Transposon tagging was first applied in plants for which endogenous transposons were known, such as maize and *Antirrhinum majus* [33, 64, 66]. In tomato, no endogenous transposable elements are known [108] and this approach cannot be used. The heterologous element Ac-Ds (*Activator-Dissociator*) from maize has been introduced into tomato and shown to be transpositionally active [111]. However, activity of transposable elements may differ greatly per genotype, as was shown for several *Arabidopsis thaliana* genotypes [25] and the activity, if any, in various tomato lines was not known. Furthermore, at the onset of our study little was known on the preference, if any, of the integration sites, nor on the frequency and distances of the jumps of active transposons. Thus, the effectiveness for tagging genes in tomato was unknown.

T-DNA tagging. *Agrobacterium tumefaciens* T-DNA has also been used as an insertional mutagen and, thus, as a tag for genes. The procedure involves the introduction of T-DNA through *Agrobacterium*-mediated transformation. In transgenic plants carrying a mutation in the desired gene complete linkage between the mutant phenotype and inserted T-DNA has to be established. Subsequently, the inserted T-DNA can be used to isolate flanking sequences including the gene. In *Arabidopsis* various genes have been isolated by this approach [26, 58].

Both aforementioned tagging strategies require the screening of large numbers of plants and are only applicable to genes of which the mutation can be easily scored. Feldman *et al* [34] calculated that the identification of a specific gene by T-DNA tagging in *Arabidopsis thaliana,* would require the generation of more than 75,000 individual transformants. Since the tomato genome is about 10 times the size of *Arabidopsis thaliana* the isolation of *Mi* through T-DNA tagging would require an impractical large number of transformed tomato plants. Using transposons the same problems arise, although possibly somewhat smaller numbers of transformed plants will be needed if the jumping frequency of the transposons is high. Furthermore, the screening for nematode resistance requires each transformant to be infected with nematodes which is rather laborious and involves an enormous greenhouse capacity. An additional setback is the somaclonal variation which may be introduced during the transformation steps and which is very difficult to separate from the effects of an insertion [102].

1.2.5 Conclusion

Considering all possible options, positional cloning seemed the most appropriate strategy to isolate *Mi*. At the onset of our work, the other approaches either lacked the required prerequisites or were insufficiently characterized to be applicable.

1.3 POSITIONAL CLONING STRATEGY AND APPLICATION TO MI

1.3.1 Marker identification

To identify a gene by positional cloning, at first its chromosomal position has to be established. This is usually accomplished by genetic mapping studies using populations segregating for the gene. When the position on the genetic map has been determined, molecular markers have to be identified. In plants, restriction fragment length polymorphisms (RFLPs) [13, 100, 110] and random amplified polymorphic DNA (RAPDs) [105, 106] are now widely used as molecular markers.

For a chromosomal walk, the global RFLP maps, which cover the entire genome, do usually not provide a sufficiently high marker density [73, 99]. At present, for tomato an average marker density of 1 RFLP per 950 kb has been obtained [99], which will often mean that a chromosomal walk covering sveral megabases has to be made. This is still a long distance and therefore additional markers need to be identified, specifically for the region of interest. To identify markers very tightly linked to the target gene, a variety of approaches has been developed. Young *et al* [112] have used nearly isogenic lines (NILs) to isolate RFLP markers linked to the tobaco mosaic virus resistance gene (*Tm2-a*) of tomato. These lines differ only in a small chromosomal segment in the region of the target gene and are screened by hybridization with anonymous genomic probes. Only probes linked to the target gene will provide a different banding pattern on Southern blots of the two lines. Alternatively, substitution lines [84] in which a complete or a very large part of an entire chromosome is exchanged for a chromosome from another species, may be used to screen for additional markers [57, 104].

The construction of nearly isogenic lines and substitution lines is time consuming and such lines are seldom available. A PCR-based approach, called bulked segregant analysis (BSA, [46, 70]) has been developed as an alternative of isolating markers linked to the gene of interest. In this approach two pools of DNA from plants of a population originating from a single cross but differing in a specific trait, are screened with a set of random oligonucleotide primers (usually 10-mers) in a PCR assay. Markers specific for the region of interest can be identified since PCR products from regions not tightly linked to the target gene will yield an identical banding pattern.

Additionally, markers may be identified from region-specific genomic libraries constructed of DNA from microdissected chromosomes [90]. Clones from these libraries can be used in a screening of NILs or substitution lines to detect polymorphisms. The microdissection technique is rather demanding and has, as yet, hardly found application.

The first available marker linked to *Mi* was the isozyme *Aps-1*[1] cloned by Aarts *et al* [1] following "classical" cloning protocols starting from the information of the aminoacid

sequence of the protein. Several additional markers, including RFLPs, RAPDs and cDNAs, have subsequently been identified (see figure 1) for the *Aps-1/Mi* region by Tanksley *et al* [99], Klein-Lankhorst *et al* [56] and Ho *et al* [51]. Tanksley *et al* [99] identified five markers for the region in his global mapping of the tomato genome. Klein-Lankhorst *et al* [56] and Ho *et al* [51] identified five additional RFLP markers by screening substitution lines and nearly isogenic lines with anonymous genomic clones. To date the RFLP marker most tightly linked to the *Mi* gene is the potato genomic clone GP79 [42, 56].

RAPD markers are relatively new to the marker scene and so far only a few have been identified. Klein-Lankhorst *et al* [57] screened 11 primers on near isogenic lines and substitution lines and found 6 markers located on chromosome 6, one of which was linked to *Mi*. Williamson *et al* [107] identified one RAPD marker, REX-1, which is currently the marker most tightly linked to *Mi*.

Another valuable source for single copy markers are cDNA clones and screening for these markers was done by Ho *et al* [51]. Their research led to the identification of four chromosome 6 markers, one of which (LC379) was found to be very tightly linked to *Mi*. Unfortunately, this marker recognized a gene family consisting of 30 members, only one of which was specific for the *Mi*-region. Therefore its applicability for a chromosomal walk is limited. Tanksley *et al* [99] also identified two cDNA markers linked to *Aps-1*.

The markers linked to *Mi*, that are currently available, have been ordered into integrated genetic and molecular linkage maps (Messenguer *et al* [69] and Ho *et al* [51]) as shown in figure 1. Assigning exact genetic distances on these integrated maps is not possible since various crosses have been applied which differ in genetic background and have only a limited number of markers in common (see figure 1 b, c, d, e). Thus, the genetic distance between the two *Mi*-flanking markers *Aps-1* and GP79, varies from 0.2 cM to 1.7 cM, depending on which cross was used for mapping [51, 69]. Recently Weide *et al* [104] constructed a linkage map of tomato chromosome 6, integrating the classical and the molecular linkage map, on which the distance between *Aps-1* and GP79 was even larger: 9.2 cM.

1.3.2 Physical mapping

Once markers have been identified that are genetically tightly linked to the target gene a chromosomal walk can be started. However, many examples have been described in which genetic distances led to underestimating [37], or overestimating [50, 72] the physical distances. For instance, it was shown for yeast chromosome III that the ratio cM/kb varies up to 10-fold for different regions along the chromosome [76]. This ratio was highest halfway the arms and lowest in the centromeric region. Apparently regions occur along the chromosome that have high and low recombination frequencies, giving high and low ratios of cM/kb respectively.

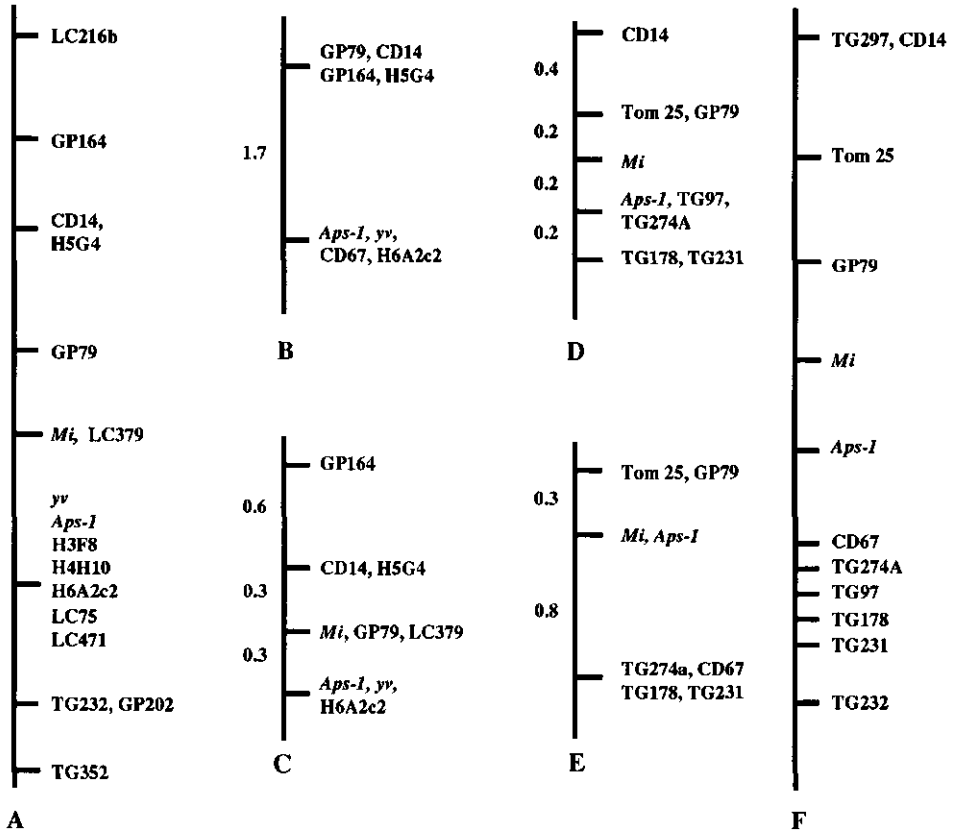


Figure 1. Maps of the *Mi*-region of tomato chromosome 6.

Integrated maps constructed by Ho *et al* [51], are shown at the left (A) and by Messenguer *et al* [69] shown at the right (F). The basis of these integrated maps were various crosses. The maps resulting of some of these crosses are shown, along with the genetic distances in cM between the markers. The crosses involved are: (B) W601 x WSL-6, and (C) VFNT-Cherry x LA55 (by Ho *et al* [51]) and (D) Moneymaker x VFNT-Cherry and (E) VFNT-Cherry x *L. cheesemannii* (LA483) by Messenguer *et al* [69]). The distances are not drawn to scale.

In tomato 1 cM corresponds on average to 750 kb (based upon a haploid genome size of 950 Mbp [4] and a genetic size of 1620 cM [99]), but also for tomato the length in kb per cM varies considerably along the chromosome. It was already shown by Khush and Rick [55] and Rick [84] that regions close to the centromere, which consist of heterochromatic DNA, show a low recombination frequency.

The relation between genetic and physical distances was studied in more detail by Segal *et al* [97], who showed that in the middle of the short arm of chromosome 11, the region where the *Fusarium* resistance gene, I_2, is located, 1 cM corresponds on average to 40 kb, almost 20 fold less than expected. Ganal *et al* [38] showed that for the region around *Tm2-a*, which is close to the centromere of chromsome 9, 1 cM corresponds to more than 4 Mbp. Although both studies involved different crosses, the 100-fold difference in ratio cM/kb does provide

additional, physical, evidence for a severe suppression of recombination in centromeric regions.

Thus, taking in account the large variation in the ratio cM/kb and the reported variable genetic distances (see section 1.3.1) it is impossible to determine the physical distance between *Mi* and its flanking markers merely on a genetic basis. As *Mi* is located near the centromere of chomosome 6 [44] the physical distance between *Mi*, *Aps-1* and GP79 is probably very large and may well be thousands of kilobases in size. Therefore, the exact distances in kb have to be determined by physical means. This involves the construction of so-called "long range restriction maps" which cover hundreds or thousands of kb and provide a physical linkage between the various markers linked to *Mi*. To construct such maps, methods are required for the analysis of DNA molecules of hundreds or even thousands of kb in size. Only with the introduction of pulsed field gel electrophoresis (PFGE) this became possible. Using conventional electrophoresis fragments up to 50 kb can be separated and larger fragments only with very great difficulties [31]. The development of PFGE by Schwartz and Cantor in 1984 [91] enabled the separation of DNA molecules several megabases in size.

While in conventional electrophoresis a single constant electric field is employed, separation in PFGE is achieved by applying two perpendicularly orientated electric fields in an alternating mode. Schwartz and Cantor succeeded in separating the 16 individual yeast chromosomes, ranging in size from 230 kb to 2000 kb by PFGE. In the original Schwartz and Cantor PFGE apparatus the speed and migration direction of DNA molecules is dependent on the position at which DNA is loaded on the gel. Consequently, DNA migrates in a complex path and skewes towards the edges of the gel. These problems were solved by modifications of the electrode configuration, resulting in the development of several new systems. By using the Orthogonal Field Alternation Gel Electrophoresis (OFAGE [18, 19]), separation up to 2000 kb could be achieved while maintaining straight lanes at the middle part of the gel. With the introduction of Field Inversion Gel Electrophoresis (FIGE [17]) straight lanes over the entire width of the gel were obtained, but the separation range and resolving power was somewhat less. Finally, by Contour Clamped Homogeneous field gel Electrophoresis (CHEF), as developed by Chu *et al* [21], a large separation range, good resolution and straight lanes were achieved.

In studies using PFGE it is essential to have available preparations of DNA molecules that are several million base pairs in size. These large DNA molecules are extremely sensitive to shearing and nuclease activity, therefore special precautions must be taken. The technique for isolating very high molecular weight DNA, as developed by Schwartz and Cantor [91], implies preparing DNA from cells that have been embedded in agarose plugs so as to avoid shearing by mechanical stress. In subsequent DNA isolation steps a high concentration of a chelating agent is needed to prevent nuclease activity. Unfortunately, not all cells do readily liberate their DNA content. In plant cells the presence of a rigid cell wall interferes with the

liberation of the DNA. Therefore, protoplasts need to be prepared before embedding cells in agarose.

Being able to isolate and separate large DNA fragments, one further needs the appropriate restriction enzymes to generate very large restriction fragments. Currently, a large set of these enzymes, the so-called "rare-cutters", is available. Some of these enzymes recognize either 7 bp (*Rsr*II) or 8 bp (eg *Sfi*I and *Not*I) sequences and therefore cut DNA infrequently. However, the majority of the "rare-cutters" are "6 bp-cutters" which cut plant [2] and animal [10] DNA much less frequently than predicted from the base composition of the source DNA. A common feature of rare-cutters is the presence of either CpG or CpNpG sequences in their recognition sites and their sensitivity for methylation of DNA at these sites. In plants these CpG and CpNpG sites are often methylated [2, 10] and, if so, can not be cut by these enzymes, resulting in relatively few digestable sites.

By now long range restriction maps covering hundreds of kb's using the rare-cutting enzymes have been constructed for several plant species such as rice [109], tomato [41], barley [86] and pea [101], showing the feasibility of these large scale physical mapping procedures.

1.3.3 Chromosome walking

From the above it is clear that distances between markers can vary widely, from a few hundred kb to megabases. Usually distances will be quite large, even when there is a tight linkage between markers. The process to cover these large distances is known as "chromosome walking" [9]. In this technique, the ends of a cloned fragment, containing a marker linked to the gene of interest, are used as probes to detect clones containing overlapping fragments. The fragments identified in this way provide again ends that can be used to detect a new set of overlapping clones. This process of identifying clones containing common parts of DNA is continued untill all the sequences between two markers flanking the gene of interest have been arranged in a continuous array.

Chromosome walking using conventional cloning vehicles like cosmid or phage lambda vectors has been performed, for example, in the isolation of *Drosophila* homeotic genes [9b] and the human retinoblastoma (RB) gene [60], in which relatively short distances of 150 and 130 kb had to be covered, respectively. However, a chromosomal walk using conventional cloning vehicles does not allow, on average, steps of more than 25 kb, resulting in a very time consuming undertaking when megabase distances have are involved. A further complicating factor can be the presence of repetitive sequences which interfere with the walking proces. If, during the walk, clones containing a repetitive sequence are identified, it should be realized that such clones may originate from other -unlinked- parts of the genome. Continuing the walk using these clones would lead away into many different directions. Therefore, clones or

subclones used for walking need to be of single copy nature. The use of very large clones can overcome both disadvantages. Firstly, due to the larger clone sizes the speed of a chromosomal walk increases and, secondly, the chances for single copy sequences at the ends of a clone are higher. The stable cloning of fragments of many hundreds of kb's in a single clone has become possible by the development of the yeast artificial chromosome (YAC) cloning technique [16]. The positional cloning approach would not have been so successful without the development of YAC cloning.

The YAC cloning system involves the ligation of all the elements required for a functional yeast chromosome -telomeres, centromere, autonomously replicating sequence (ARS)- to large restriction fragments. The ligation products are introduced into yeast cells where they are maintained as stable chromosomes. Using YACs it is possible to clone fragments megabases in size, which is a major improvement over all other known cloning systems. Furthermore, YAC cloning is an eukaryote cloning system and therefore more sequences are clonable, because the sequence organisation of eukaryotic DNA is more compatible with the yeast replication system than with *E. coli* [16]. The early success of cloning nearly all *Caenorhabditis elegans* sequences in a YAC library [24], soon followed by a number of other organisms, showed that YACs fullfilled their promises. The YAC cloning system will be discussed extensively in chapter 4.

Besides the YAC cloning system other techniques have been developed which also allow the cloning of fragments larger than can be inserted into cosmids. For example, the P1 cloning strategy using the bacteriophage P1 as a vector allows fragments of approximately 95 kb -but not much smaller or larger!- to be cloned as large circular molecules in *E. coli* [97]. So far it has been used only for the construction of a *Drosophila* library [96] and no experience is available using plant DNA. Another example is the bacterial artificial chromosome (BAC) cloning system, which is based on the F-factor plasmid in *E. coli* [75, 93]. Large restriction fragments are ligated into pBAC vectors after which the plasmids with recombinant DNA are transformed into the bacteria. The pBAC vectors can maintain fragments up to 300 kb in one copy per cell as shown by Shizaya *et al* [93] for the cloning of human DNA fragments.

The advantage of prokaryotic systems is that, once cloned, DNA can be isolated in larger quantities due to the usually higher copy number of recombinant DNA in bacteria and the smaller cell volume of bacteria. The P1 system only allows the cloning of DNA fragments twice the size of cosmids and is, therefore, not such a great improvement, but the BAC system may be a welcome additional cloning technique.

At the onset of our work both the P1 and BAC cloning were not available and thus the choice for YACs to chromosome walk was obvious. Currently, however, the choice would still be the same since only the YAC system offers the posibility of cloning the very large fragments necessary to bridge the gap between genetic and molecular analysis. If analysis in more detail is required, such as detailed restriction mapping or sequencing, cosmids, phage lambda or P1 clones may be preferable since they allow a more easy handling of DNA.

An alternative approach to chromosome walking is chromosome jumping [22, 82]. In the chromosome jumping strategy only the ends of large restriction fragments are cloned into conventional phage lambda vectors. This is accomplished by circularizing large restriction fragments, generated by a partial digestion, or a complete rare-cutter digestion, by which the two ends which were originally a few hundred kb apart are linked. After ligation the large circular fragments are cut with a frequently cutting restriction enzyme so that a small fragment, containing only the ends of the large restriction fragment, is obtained. This small fragment can then be cloned into conventional phage lambda vectors. To be able to select for the phage clone containing the ends of the large restriction fragment a phage selectable marker is introduced during circularization. These phage libraries, containing the ends of large restriction fragments, are called jumping libaries since they allow to find with each clone DNA fragments that are separated in the chromosome by a few hundred kb. By using both parts of a clone as individual probes in the screening of the jumping library, one can "jump" to an adjacent large restriction fragment without the need of cloning all intervening sequences. After a jump has been made and positioned, cosmids or phage libraries can be screened and a chromosomal walk can be made to obtain the intervening sequences. The disadvantage of jumping libraries is that only small parts are actually cloned and that one always has to go back to a genomic library and start a conventional chromosomal walk to obtain the sequences containing the gene of interest. Currently, no jumping libraries have been described for plants and with the other cloning systems now available, the value of chromosome jumping is questionable.

1.3.4 Identification of putative coding sequences

When the chromosomal walk is completed, the isolated clones from the relevant region must be screened for coding sequences for the sake of gene identification. There are several methods available to detect coding sequences which are usually applied in combination [43]. Once coding sequences have been identified the target gene must finally be identified through complementation of a null mutant via transformation (section 1.3.5).

In the following, a short overview of the most important methods utilized in positional cloning to detect coding sequences from a chromosomal region is presented.

Expression analysis. Cloned genomic DNA in phage lambda, cosmids or YACs, can be used as hybridization probes in a Northern blot analysis or a cDNA library screening to detect whether coding sequences are located on the clone [30, 48, 63]. These screening assays are mainly effective in detecting highly expressed genes [43, 79]. However, with the advent of several PCR-based modifications of the conventional screening protocols [78, 79] the detection of cDNAs present in very low numbers has been improved and the chance of identifying all genes located within a specific chromosomal segment has increased.

In plants, screening of both Northern blots [77] and cDNA libraries [5, 47] appears at present the best choice to find coding sequences within a genomic clone obtained after chromosome walking.

Evolutionary conservation of sequences. Coding sequences tend to be more conserved during evolution than non-coding sequences [61]. Therefore, genomic DNA containing coding sequences will hybridize at higher stringency than less conserved, putatively non-coding sequences, when used as hybridization probes on blots of genomic DNA from other, related, species. Monaco *et al* [71] were first to use this approach and dubbed the blots, which contained DNAs from several animal species, "Zoo-blots". In plants, a similar approach may be followed to detect coding sequences using a collection of plants on "Flora-blots". Alternatively, a set of tomato cultivars that differ in nematode resistance, may be analyzed in order to select for resistance linked sequences. Currently, no report on the use of "Flora-blots" is available and a good evaluation of the method is not possible.

CpG islands. An important feature of mammalian as well as plant DNA is the occurrence of so-called CpG islands which are regions, varying in size from 0.5 to 2 kb, where CpG dinucleotides are present at 10 to 20 times their average density while the C-residue is usually non-methylated [12]. The islands occur at the 5'-side of all housekeeping and many other, non-housekeeping genes [3, 12]. Due to the GC-richness and the unmethylated character of the CpG's, sites for rare-cutting enzymes are predominantly located within these islands. As a result, at least in mammals, a clustering of rare-cutter sites may indicate the presence of a CpG-island and thus a gene [11]. To date the method has been used succesfully only in mammalian systems [52]. Since island characteristics between mammals and dicotyledonous plants differ substantially with respect to average GC-content [39, 40], the possibility to detect coding sequences by rare-cutter restriction analysis is as yet uncertain.

Exon trapping. A very promising method for identifying coding sequences in animal DNA developed by Buckler *et al* [15] and improved by Duyck *et al* [29] and Hamagucchi *et al* [49], is exon trapping. It is based on the detection of splice site sequences which are highly conserved in eukaryotic organisms. North *et al* [74] found that almost 90 % of the genes present within a set of cosmids could be detected by this method. The procedure involves the subcloning of the selected fragment that is to be screened for coding sequences in "exon trap" vectors, which contain an intron with the conserved 5' and 3' splice sites of a tester gene. Between these splice sites a multiple cloning site is present into which the target DNA is ligated. After *in vivo* expression and processing, using monkey fibroblast cells, an intron/exon boundary present within the insert will result in a RNA molecule in which the exon is retained or trapped. If the genomic fragment does not contain any splice sites the intron of the tester gene will be removed completely, resulting in a smaller RNA molecule. Reverse transcription followed by PCR is then used to amplify the RNA products. The presence of a trapped exon is revealed by reaction products larger than the ones that would be obtained if no coding sequences were present. The method has been developed for mammalian gene detection,

application to plants requires a suitable transcription and processing system. Possibly the use of transient expression in protoplasts can be used in conjunction with exon-trap vectors. Such systems are currently not available and it may require a large effort to develop such a system for plants.

In conclusion, the method of choice for detecting coding sequences in plants seems to be screening for expressed sequences on the basis of cDNA or mRNA detection. All other methods described have not been tested, or require special developments, for use in plants.

1.4.5 Complementation studies

The ultimate proof that the candidate sequences identified after a chromosomal walk indeed contain the gene of interest, is to complement a null mutant with these sequences by direct gene transfer or *Agrobacterium* T-DNA mediated transformation. For example, identification of the gene conferring resistance against races of *Pseudomonas syringae* pv *tomato* in tomato (*Pto*) was accomplished by transferring cDNAs, selected with a YAC clone encompassing the *Pto*-region, to sensitive plants using the *Agrobacterium* Ti-plasmid [65].

Alternatively, the complementation assay may precede the identification of coding sequences within the candidate region, as has been done in cloning the *Arabidopsis thaliana* genes *ABI3* [47] and *AXR1* [77]. In these cases, the DNA from several cosmid clones from the candidate region was transformed into a mutant plant. After selecting the plants that had acquired the desired phenotype only the clone that achieved this transformation was further analyzed for the gene of interest. This approach is manageable only when the candidate region, carrying the gene of interest, is small as the procedure requires the generation and characterization of a large number of transgenic plants.

To identify *Mi* among putative coding sequences a susceptible plant should become resistant against the root-knot nematodes, after transformation with the proper clone. This step should not pose a serious problem as efficient protocols for transformation of tomato are available [59].

1.5 REFERENCES

1. Aarts JMMJG, Hontelez JGJ, Fischer P, Verkerk R, van Kammen A, Zabel P (1991) Acid phosphatase-1[1], a tightly linked molecular marker for root-knot nematode resistance in tomato: from protein to gene, using PCR and degenerate primers containing deoxyinosine. Plant Mol Biol 16: 647-61.
2. Anand R (1986) Pulsed field gel electrophoresis: a technique for fractionating large DNA molecules. Trends Genet 2: 278-283.
3. Antequera F, Bird AP (1988) Unmethylated CpG islands associated with genes in higher plant DNA. EMBO J 7: 2295-2299.

4. Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Rept 9: 208-218.
5. Arondel V, Lemieux B, Hwang I, Gibson S, Goodman H, Somerville CR (1992) Map-based cloning of a gene controlling omega-3 fatty acid desaturation in *Arabidopsis*. Science 258: 1353-1355.
6. Bailey DM (1940) The seedling test method for root-knot nematode resistance. Proc Am Soc Hortic Sci 38: 573-575.
7. Barham WS, Winstead NN (1957) Inheritance of resistance to root-knot nematodes. Tomato Genet Coop Rep 7:3.
8. Bautz EKF, Reilly E (1966) Gene specific messenger RNA: isolation by the deletion method. Science 151: 328-330.
9. Bender W, Spierer P, Hogness D (1979) Gene isolation by chromosomal walking. J Supra Mol Struct 10 (supl): 32.
9b. Bender W, Akam M, Karch F, Beachy PA, Peifer M, Spierer P, Lewis EB, Hogness DS (1983) Molecular genetics of the bithorx complex in *Drosophila melanogaster*. Science 221: 23-29.
10. Beverly SM (1988) Characterization of the 'unusual' mobility of large circular DNAs in pulsed field-gradient electrophoresis. Nucleic Acids Res 16: 925-939.
11. Bickmore WA, Bird AP (1992) The use of restriction endonucleases to detect and isolate genes from mammalian cells. Methods Enzymol 216: 224-244.
12. Bird AP (1987) CpG islands as gene markers in the vertebrate nucleus. Trends Genet 3: 342-347.
13. Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32: 314-331.
14. Brueske CH (1980) Phenylalanine ammonia lyase activity in tomato roots infected and resistant to the root-knot nematode, *Meloidogyne incognita*. Physiol Pl Pathol 16: 409-414.
15. Buckler AJ, Chang DD, Graw SL, Brook JD, Haber DA, Sharp PA, Housman DE (1991) Exon amplification: a strategy to isolate mammalian genes based on RNA splicing. Proc Natl Acad Sci USA 88: 4005-4009.
16. Burke DT, Carle GF, Olson MV (1987) Cloning of large segments of exogeneous DNA into yeast by means of artificial chromosome vectors. Science 236: 806-812.
17. Carle GF, Frank M, Olson MV (1986) Electrophoretic separation of large DNA molecules by periodic inversion of the electric field. Science 232: 65-68.
18. Carle GF, Olson MV (1984) Separation of chromosomal DNA molecules from yeast by orthogonal-field-alternation-gel electrophoresis. Nucleic Acids Res 12: 5647-5664.
19. Carle GF, Olson MV (1985) An electrophoretic karyotype for yeast. Proc Natl Acad Sci USA 82: 3756-3760.
20. Carrasco AE, McGinnis W, Gehring WJ, De Robertis EM (1984) Cloning of an *X. levis* gene expressed during early embryogenesis coding for a peptide region homologous to *Drosophila* homeotic genes. Cell 37: 409-414.
21. Chu G, Vollrath D, Davis RW (1986) Separation of large DNA molecules by contour-clamped homogeneous electric fields. Science 234: 1582-1585.
22. Collins FS, Weissman SM (1984). Directional cloning of DNA fragments at a large distance from an initial probe: a circularization method. Proc Natl Acad Sci USA 81: 6812-6816.
23. Cook R (1991) Resistance in plants to cyst and root-knot nematodes. Agricult Zool Rev 4: 213-240.
24. Coulson A, Waterston R, Kiff J, Sulston J, Kohara Y (1988) Genome linking with yeast artificial chromosomes. Nature 335: 184-186.
25. Dean C, Sjodin C, Page T, Jones J, Lister C (1992) Behaviour of the maize transposable element *Ac* in *Arabidopsis thaliana*. Plant J 2: 69-81.
26. Deng XW, Matsui M, Wei N, Wagner D, Chu AM, Feldmann KA, Quail PH (1992) *COP1*, an *Arabidopsis* regulatory gene, encodes a protein with both a zinc-binding motif and a Gß homologous domain. Cell 71: 791-801.
27. Dropkin VH (1969) The necrotic reaction of tomatoes and other hosts resistant to *Meloidogyne*: reversal by temperature. Phytopath 59: 1632-1637.
28. Dropkin VH (1972) Pathology of *Meloidogyne* - Galling, giant cell formation, effects on host physiology. OEPP/EPPO Bull 6: 23-32.
29. Duyck GM, Kim S, Myers RM, Cox DR (1990) Exon trapping: a genetic screen to identify candidate transcribed sequences in cloned mammalian genomic DNA. Proc Natl Acad Sci USA 87: 8995-8999.
30. Elvin P, Slynn G, Black D, Graham A, Butler R, Riley J, Anand R, Markham AF (1990) Isolation of cDNA clones using yeast artificial chromosome probes. Nucleic Acids Res 18: 3913-3917.
31. Fangman WL (1978). Separation of very large DNA molecules by gel electrophoresis Nucleic Acids Res 5: 653-665.
32. Fassuliotis G (1991) Root-knot resistance in tomato. In: Genetic improvement of tomato. Kalloo G (Ed.) Springer, Heidelberg. 143-152.

33. Federoff NV, Furtek DB, Nelson OE (1984) Cloning of the *bronze* locus in maize by a simple and generalizable procedure using the transposable controlling element *Activator* (*Ac*). Proc Natl Acad Sci USA 81: 3825-3829.
34. Feldman KA, marks MD, Christianson ML, Quatrano RS (1989) Dwarf mutant of *Arabidopsis* generated by T-DNA insertion mutants. Science 243: 1351-1354.
35. Flor HH (1955) Host-parasite interaction in flax rust - its genetics and other implications. Phytopath 45: 680-685.
36. Frazier WA, Dennett RK (1949) Isolation of *Lycopersicon esclentum* type tomato lines essentially homozygous resistant to root-knot. Proc Am Soc Hortic Sci 54: 225-236.
37. Fulton TR, Bowlook AM, Smith DR, Daneshvar L, Green P, Cavalli-Sforza LL, Donis-Keller H (1989) Nucleic Acids Res 17: 271-284.
38. Ganal MW, Young ND, Tanksley SD (1989) Pulsed field gel electrophoresis and physical mapping of large DNA fragments in the *Tm-2a* region of chromosome 9 in tomato. Mol Gen Genet 215: 395-400.
39. Gardiner-Garden M, Frommer M (1992) Significant CpG-rich regions in angiosperm genes. J Mol Evol 34: 231-245.
40. Gardiner-Garden M, Sved JA, Frommer M (1992) Methylation sites in Angiosperm genes. J Mol Evol 34: 219-230.
41. Gasser CS, Fraley RT (1989) Genetically engineering plants for crop improvement. Science 244: 1293-1299.
42. Gebhardt C, Ritter E, Barone A, Debener T, Walkemeier B, Schachtschabel U, Kaufman H, Thompson RD, Bonierbale MW, Ganal MW, Tanksley SD, Salamini F (1991) RFLP maps of potato and their alignment with the homoeologous tomato genome. Theor Appl Genet 83: 49-57.
43. Gibson S, Sommerville C (1993) Isolating plant genes. Trends Biotechnol 11: 306-313.
44. Gilbert JC (1958) Some linkage studies with the *Mi* gene for resistance to root-knot. Tomato Genet Coop Rep 8: 15-17.
45. Gilbert JC, McGuire DC (1956) Inheritance of resistance to severe root-knot from *Meloidogyne incognita* in commercial type tomatoes. Proc Am Soc Hortic Sci 68: 437-442.
46. Giovannoni JJ, Wing RA, Ganal MW, Tanksley SD (1991) Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations. Nucleic Acids Res 19: 6553-6558.
47. Giraudat J, Hauge BM, Valon C, Smalle J, Parcy F, Goodman HM (1992) Isolation of the Arabidopsis *ABI3* gene by positional cloning. The plant cell 4: 1251-1261.
48. Goldberg YP, Lin BY, Andrew SE, Nasir J, Graham R, Glaves ML, Hutchinson G, Theilmann J, Ginzinger DG, Schappert K, Clarke L, Rommens JM, Hayden MR (1992) Cloning and mapping of the α-adducin gene close to D4S95 and assessment of its relationship to Huntington disease. Hum Mol Genet 1: 669-675.
49. Hamagucchi M, Sakamoto H, Tsuruta H, Sasaki H, Muto T, Sugimura T, Terada M (1992) Establishment of a highly sensitive and specific exon-trapping system. Proc Natl Acad Sci 89: 9779-9783.
50. Higgins MJ, Turmel C, Noolandi J, Neumann PE (1990) Construction of the physical map for three loci in chromosome band 13q14: comparison to the genetic map. Proc Natl Acad Sci USA 87: 3415-3419.
51. Ho JY, Weide R, van Wordragen MF, Lambert K, Koornneef M, Zabel P, Williamson VM (1992) The root-knot nematode resistance gene (*Mi*) in tomato: construction of a molecular linkage map and identification of dominant cDNA markers in resistant genotypes. The Plant J 2: 971-982.
52. Hochgeschwender U, Brennan MB (1991) Identifying genes within the genome: new ways for finding the needle in a haystack. BioEssays 13: 139-144.
53. Hussey RS, Paguio OR, Seabury F (1990) Localization and purification of a secretory protein from the oesophageal glands of *Meloidogyne incognita* with a monoclonal antibody. Phytopath 80: 709-714.
54. Kaplan DT, Keen NT, Thomasson IJ (1980) Association of glyceollin with the incompatible response of soybean roots to *Meloidogyne incognita*. Physiol Pl Pathol 16: 309-318.
55. Khush GS, Rick CM (1967) Studies on the linkage map of chromosome 4 of the tomato and on the transmission of induced deficiencies. Genetica 38: 74-94.
56. Klein-Lankhorst R, Rietveld P, Machiels B, Verkerk R, Weide R, Gebhardt C, Koornneef M, Zabel P (1991) RFLP markers linked to the root-knot nematode resistance gene *Mi* in tomato. Their Appl Genet 81: 661-667.
57. Klein-Lankhorst RM, Vermunt A, Weide R, Liharska T, Zabel P (1991) Isolation of molecular markers for tomato (*L. esculentum*) using random amplified polymorphic DNA (RAPD). Theor Appl Genet 83: 108-114.
58. Koncz C, Mayerhofer R, Koncz-Kalman Z, Nawrath C, Reiss B, Redei G, Schell J (1990) Isolation of a gene encoding a novel chloroplast protein by T-DNA tagging in *Arabidopsis thaliana*. EMBO J 9: 1337-1346.

59. Koornneef M, Jongsma M, Weide R, Zabel P, Hille J (1987) Transformation of tomato. In: Tomato biotechnology, Nevins DJ, Jones RA (Eds) New York, pp 169-178
60. Lee WH, Bookstein R, Hong F, Young LJ, Shew JY, Lee EYHP (1987) Human retinoblastoma susceptibility gene: cloning, identification, and sequence. Science 235: 1394-1399.
61. Lewin B (1990) Genes IV. Oxford University Press, Cambridge, Mass.
62. Lindsey K (1992) Genetic manipulation of crop plants. J Biotechnol 26: 1-28.
63. Lovett M, Kere J, Hinton LM (1991) Direct selection: A method for the isolation of cDNAs encoded by large genomic regions. Proc Natl Acad Sci USA 88: 9628-9632.
64. Martin C, Carpenter R, Sommer H, Saedler H, Coen E (1985) Molecular analysis of instability in flower pigmentation of *Antirrhinum majus*, following isolation of the *Pallida* locus by transposon tagging. EMBO J 4: 1625-1630.
65. Martin G, Brommonschenkel SH, Chunwongse J, Frary A, Ganal M, Spivey R, Wu T, Earle ED, Tanksley SD (1993) Map-based cloning of a protein kinase gene conferring disease resistance in tomato. Science 262: 1432-1436.
66. McCarthy DR, Hattori T, Carson CB, Vasil V, Lazar M, Vasil IK (1991) The *viviparous-1* developmental gene of maize encodes a novel transcriptional activator. Cell 66: 895-905.
67. Medina Filho HP, Stevens MA (1980) Tomato breeding for nematode resistance: survey of resistant varieties for horticultural characteristics and genotype of acid phosphatases. Acta Hortic 100: 383-393.
68. Medina-Filho H, Tanksley SD (1983) Breeding for nematode resistance. In: Evans DA, Sharp WR, Ammirato PV, Yamada Y (Eds) Handbook of plant cell culture, Vol 1, Techniques for propagation and breeding, Mac Millan, New York, pp 904-923.
69. Messenguer R, Ganal M, de Vicente MC, Young ND, Bolkan H, Tanksley SD (1991) High resolution RFLP map around the root-knot nematode resistance gene (*Mi*) in tomato. Theor Appl Genet 82: 529-536.
70. Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci 88: 9828-9832.
71. Monaco AP, Neve RL, Colletti-Feener C, Bertelson CJ, Kurnit DM, Kunkel LM (1986) Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. Nature 323: 646-650.
72. Nakamura Y, Lathrop M, O'Connell P, Leppert M, Kamboh MI, Lalouel JM, White R (1989) Frequent recombination is observed in the distal end of the long arm of chromosome 14. Genomics 4: 76-81.
73. Nam HG, Giraudat J, den Boer B, Moonan F, Loos WDB, Hauge BM, Goodman HW (1989) Restriction fragment length polymorphism linkage map of *Arabidopsis thaliana*. Plant Cell 1: 699-705.
74. North MA, Sanseau P, Buckler AJ, Church D, Jackson A, Patel K, Trowsdale J, Lehrach H (1993) Efficiency and specificy of gene isolation by exon amplification. Mammalian Genome 4: 466-474.
75. O'Conner M, Peifer M, Bender W (1989) Construction of large DNA segments in *Escherischia coli*. Science 244: 1307-1312.
76. Oliver SG *et al* (1992) The complete sequence of yeast chromosome III. Nature 357: 40-47.
77. Ottoline-Leyser HM, Lincoln CA, Timpte C, Lammer D, Turner J, Estelle M (1993) *Arabidopsis* auxin-resistance gene *AXR1* encodes a protein related to ubiquitin-activating enzyme E1. Nature 364: 161-164.
78. Parimoo S, Patanjali SR, Shukla H, Chaplin DC, Weissman SM (1991) cDNA selection: Efficient PCR approach for the selection of cDNAs encoded in large chromosomal DNA fragments. Proc Natl Acad Sci USA 88: 9623-9627.
79. Parrish JE, Nelson DL (1993) Methods for finding genes a major rate limiting step in positional cloning. Gene Anal Techn Appl. 12: 29-41.
80. Pi CL, Rohde RA (1967) Phenolic compounds and host reaction in tomato to injury caused by root-knot nematodes (abstr). Phytopathology 57: 344
81. Pnueli L, Abu-Abeid M, Zamir D, Nacken W, Schwarz-Sommer ZS, Lifschitz E (1991) The MADS box gene family in tomato: temporal expression during floral development, conserved secondary structures and homology with homeotic genes from *Antirrhinum* and *Arabidopsis*. Plant J 1: 255-266.
82. Poustka A, Lehrach H (1986) Jumping libraries and linking libraries: the next generation of molecular tools in mammalian genetics. Trends Genet. 2: 174-179
83. Powers TO (1992) Molecular diagnostics for plant nematodes. Parasitol Today 8: 177-179.
84. Rick CM (1969) Controlled introgression of chromosomes of *Solanum pennellii* into *Lycopersicon esculentum*: segregation and recombination. Genetics 62: 753-768.
85. Roberts PA, Dalmasso A, Cap GB, Castagnone-Sereno P (1990) Resistance in *Lycopersicon peruvianum* to isolates of *Mi* gene-compatible *Meloidogyne* populations. J Nematol 22: 585-589.
86. Roeder MS, Lapitan NLV, Sorrels ME, Tanksley SD (1993) Genetic and physical mapping of barley telomeres. Mol Gen genet 238: 294-303.

87. Sasser JN (1977) Worldwide dessimination and importance of the root-knot nematodes, *Meloidogyne* spp. J Nematol 9: 26-29.
88. Sasser JN (1980) Root-knot nematodes: a global menace to crop production. Plant Dis 64: 36-41.
89. Sasser JN, Eisenback JD, Carter CC, Triantaphyllou (1983) The international *Meloidogyne* project - its goals and accomplishments. Ann Rev Phytopathol. 21: 271-288.
90. Saunders RDC (1990) Short cuts for genomic walking: chromosome microdissection and the polymerase chain reaction. BioEssays 12: 245-248.
91. Schwartz DC, Cantor CR (1984) Separation of yeast chromosome-sized DNA by pulsed field gradient gel electrophoresis. Cell 37: 67-75.
92. Segal G, Sarfatti M, Schaffer MA, Ori N, Zamir D, Fluhr R (1993) Correlation of genetic and physical structure in the region surrounding the I_2 *Fusarium oxysporum* resistance locus in tomato. Mol Gen Genet 231: 179-185.
93. Shizuya H, Birren B, Kim UJ, Mancino V, Slepak T, Tachiiri Y, Simon M (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherischia coli* using an F-factor-based vector. Proc Natl Acad Sci USA 89: 8794-8797.
94. Singh B, Choudhury B (1973) The chemical characteristics of tomato cultivars resistant to root-knot nematode (*Meloidogyne* spp). Nematologica 19: 443-448.
95. Smith PG (1944) Embryo culture of a tomato species hybrid. Proc Am Soc Hortic Sci 44: 413-416.
96. Smoller DA, Petrov D, Hartl D (1991) Characterization of bacteriophage P1 library containing inserts of *Drosophila* DNA of 75-100 kilobase pairs. Chromosoma 100: 487-494.
97. Sternberg N (1990) Bacteriophage P1 cloning sustem for the isolation, amplification, and recovery of DNA fragments as large as 100 kilobase pairs. Proc Natl Acad Sci USA 87: 103-107.627.
98. Sun TP, Goodman HM, Ausubel FM (1992) Cloning the *Arabidopsis GA1* locus by genomic substraction. Plant Cell 4: 119-128.
99. Tanksley SD, Ganal MW, Prince JP, de Vincente MC, Bonierbale MW, Broun P, Fulton TM, Giovanonni JJ, Grandillo S, Martin GB, Messeguer R, Miller JC, Miller L, Paterson AH, Pineda O, Röder MS, Wing RA, Wu W, Young ND (1992) High density molecular linkage maps of the tomato and potato genomes. Genetics 132: 1141-1160.
100. Tanksley SD, Young ND, Paterson AH, Bonierbale MW (1989) RFLP mapping in plant breeding: new tools for old science. Bio/technology 7: 257-264.
101. Turner L, Hellens RP, Lee D, Ellis THN (1993) Genetic aspects of the organization of legume genes in pea. Plant Mol Biol 22: 101-112.
102. Walden R, Hayashi H, Schell J (1991) T-DNA as a gene tag. Plant J 1: 281-288.
103. Wallace MR, Marchuck DA, Andersen LB, Letcher R, Odeh HM, Saulino AM, Fountain JW, Brereton A, Nicholson J, Mitchell AL, Brownstein BH, Collins FS (1990) Type 1 neurofobromatosis gene: identification of a large transcript disrupted in three NF1 patients. Science: 249: 181-186.
104. Weide R, van Wordragen M, Klein-Lankhorst R, Verkerk R, Hanhart C, Liharska T, Pap E, Stam P, Zabel P, Koornneef M (1993) Integration of the classical and molecular linkage maps of tomato chromosome 6. Genetics 135: 1175-1186.
105. Welsh J, McClelland M (1990) Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res 18: 7213-7218.
106. Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingy SV (1990) DNA polymorphisms amplified by arbitrary primers are usefull as genetic markers. Nucleic Acids Res 18: 6531-6535.
107. Williamson VM, Ho JY, Wu FF, Miller N, Kaloshian I (1994) A PCR-based marker tightly linked to the nematode resistance gene, *Mi*, in tomato. Theor Appl Genet 87: 757-763.
108. Wisman E (1993) Genetic (in)stability in tomato. Thesis, Agricultural University Wageningen.
109. Wu KS, Tanksley SD (1993) Genetic and physical mapping of telomeres and macrosatellites of rice. Plant Mol Biol 22: 861-872.
110. Wyman AR, White R (1980) A highly polymorphic locus in human DNA. Proc Natl Acad Sci USA 77: 6754-6758.
111. Yoder JI, Palys K, Lassner M (1988) *Ac* transposition in transgenic tomato plants. Mol Gen Genet 213: 291-296.
112. Young ND, Zamir D, Ganal MW, Tanksley SD (1988) Use of isogenic lines and simultaneous probing to identify DNA markers tightly linked to the *Tm-2a* gene in tomato. Genetics 120: 579-585.
113. Zacheo G, Molinari S, Dacoda D (1988) Hydroxyproline rich proteins and peroxidases in tomato roots infested by root-knot nematode. Nematol Mediter 16: 235-237.

CHAPTER 2

PREPARATION OF MEGABASE-SIZED TOMATO DNA AND SEPARATION OF LARGE RESTRICTION FRAGMENTS BY FIELD INVERSION GEL ELECTROPHORESIS.

Abstract

The Schwartz and Cantor technique for releasing and fractionating megabase-sized DNA from agarose-embedded cells is beginning to bridge the gap between classical genetics and current molecular DNA techniques, particularly in mammalian systems. As yet no conditions have been described for preparing plant DNA that is of sufficient length to allow similar long-range restriction mapping experiments in plant systems. In this report we describe the application of the Schwartz and Cantor technique for preparing high molecular weight DNA from embedded tomato leaf protoplasts, as well as conditions for generating and fractionating large restriction fragments, by field inversion gel electrophoresis (FIGE). The bulk of DNA released from lysed protoplasts was at least 2 Mb in size and amenable to restriction digestion as shown by hybridizing Southern blots with, among others, a probe for the *Adh-2* gene of tomato. Restriction fragments as large as 700 kb were detected. Chloroplast DNA is isolated intact, amenable to restriction analysis and, in its native form, not mobile in FIGE.

This chapter is published: Raymond A. J. van Daelen, Jos J. Jonkers and Pim Zabel (1989) Preparation of megabase-sized tomato DNA and separation of large restriction fragments by field inversion gel electrophoresis (FIGE). Plant Mol Biol 12: 341-352.

Introduction

The development of a technique for releasing intact chromosomes from agarose-embedded yeast cells and fractionating them by pulsed field gel (PFG) electrophoresis (Schwartz and Cantor [38, 39]) has added a new dimension to the molecular analysis of chromosomal DNA from both lower [6, 7, 17, 41, 42] and higher eukaryotes [1, 2, 10, 19, 20, 30, 35, 36, 43, 44]. While conventional procedures allowed the preparation and separation of DNA fragments up to 50-100 kb, the agarose-embedded cell lysis procedures in conjunction with novel electrophoresis techniques [8, 9, 39, 41] extended this range to the megabase level. As a result of these techniques, the molecular analysis of mammalian genetic diseases already has made spectacular progress [5, 7, 13, 29, 30]. Genetic loci, which were only known for their mutant phenotype, became amenable to molecular analysis and cloning following the construction of physical maps with 100-2000 kb resolution [2, 23]. In principle, application of these novel electrophoresis and mapping techniques to plant chromosomal DNA could have a similar impact on plant genetics, provided megabase-sized DNA can be isolated from plant cells and digested by restricition enzymes.

Disease resistance loci are obvious candidates for such a positional cloning approach, since their protein products are as yet unknown and identification of coding sequences corresponding to the locus remains otherwise elusive. Genetic mapping has provided information on the genomic position of many disease resistance loci and, in some cases, on their linkage to molecular markers.

In tomato, the gene *Mi*, which confers resistance against the major root-knot nematode species of the genus *Meloidogyne*, offers a typical example. This gene is located at position 35 on chromosome 6 in tight linkage (< 1.0 cM) to an allele of the acid phosphatase-1 gene (*Aps-1*^{1}), which has thus far served as a convenient molecular (isozyme) marker in breeding programs [27]. In view of this tight linkage, the molecular identification of the gene *Mi* should be attainable by using the *Aps-1*^{1} gene as a starting position in jumping and linking libraries [11, 34]. The isolation of the *Aps-1*^{1} gene is currently in progress at our laboratory.

As a prelude to constructing a long-range restriction map of the *Aps-1/Mi* region on chromosome 6 of tomato we have developed a procedure for isolating plant DNA from agarose-embedded leaf cell protoplasts, which is greater than 2 Mb in size and is amenable to restriction digestion with a variety of restriction enzymes including various 'rare cutters'.

Materials and methods

Plant material

We have used a tomato (*Lycopersicon esculentum*) breeding line (83M), obtained from De Ruiter Seeds (Bleiswijk, Netherlands), homozygous for the nematode resistance gene *Mi*.

Protoplast isolation

Protoplasts were isolated from the leaves of *in vitro* shoot cultures which were two to three weeks old. To obtain shoot cultures, seeds were surface-sterilized (incubation in 1% (w/v) NaOCl, 0.05% (v/v) Tween-80 for 8 minutes, followed by extensive washing in sterile distilled water) and placed on R3B medium [28] without hormones. Shoot tips were excised from 10-day-old seedlings and transferred to fresh medium. The cultures were grown under an 18 hour photoperiod at 25 °C. Protoplasts were isolated as described by Tan *et al* [40], with some modifications. The leaves were cut into small pieces and incubated overnight at 25 °C in the dark in enzyme medium containing CPW-salts [15], 9% (w/v) mannitol, 3 mM 2-(N-morpholino)-ethane-sulfonic acid (MES)-KOH (pH 5.8), 1% (w/v) cellulase (Onozuka R-10) and 0.2% (w/v) macerozyme (Onozuka R-10). After incubation, protoplasts were filtered through a 50 µm nylon membrane to remove undigested leaf material and washed once by centrifugation (100x**g**, 5 min) and resuspension in wash medium (CPW-salts, 3 mM MES-KOH, pH 5.8 and 2% (w/v) KCl). Protoplasts were purified by underlayering the suspension with a sucrose solution (CPW-salts, 3 mM MES-KOH (pH 5.8) and 18% (w/v) sucrose) and by centrifugation for 5 min at 120x**g**. Viable protoplasts were collected from the interface, resuspended in whash medium, counted, collected and taken up in wash medium at a concentration of 50 million protoplasts per ml. Viability of the protoplast preparation was determined by staining with fluoresceine diacetate [22] and subsequent examination using fluorescence microscopy.

'Conventional' DNA isolation

Total tomato DNA was isolated from leaf tissue using the CTAB-procedure of Murray and Thompson [31].

Preparation of high molecular weight DNA in agarose plugs

The protoplast suspension was warmed to 37 °C and mixed with an equal volume of a 1% (w/v) LMP-agarose (Seaplaque, FMC) solution in protoplast wash medium at 37 °C. The mixture was pipetted into a perspex mold (van Ommen and Verkerk [43]) to obtain slot-sized plugs (130 µl). The mold was precooled on ice for 30-60 min and kept on ice for an additional 30 min to allow the agarose to solidify. Solidified plugs were removed from the mold, collected in lysis mix (0.5 M EDTA, 1% (w/v) N-Lauroylsarcosine and 1 mg/ml proteinase-K, pH 8.0; 2 ml per plug) and incubated for 48 h at 50 °C. The lysis mix was changed once, after 24 h. After lysis, the plugs were stored in fresh lysis mix at 4 °C. Under those conditions, plugs could be stored without noticable degradation of DNA for several months. Approximately 3 million protoplasts per plug were applied to obtain about 4.5 µg of chromosomal DNA (assuming a DNA content of 0.74 pg per haploid genome [16]).

Preparation of size markers

Phage lambda DNA multimers.
Wild-type phage lambda particles were isolated from an infected *Escherichia coli* culture by PEG precipitation as described by Maniatis *et al* [25]. The PEG precipitate was taken up in 4 volumes of SM (0.1 M NaCl, 8 mM $MgSO_4$, 50 mM TRIS-HCl (pH 7.5) and 0.01% (w/v) gelatin), warmed to 40 °C and subsequently mixed with an equal volume of a 1% (w/v) LMP-agarose solution in SM at 40 °C. The mixture was pipetted into the mold, as described above. The plugs were incubated in lysismix for 4 h at 50 °C and stored in lysismix at 4 °C to allow annealing of the phage DNA molecules. Within 2 days multimers of up to 1000 kb were formed. In accordance with Anand [1] and van Ommen and Verkerk [43] we found that oligomers of only 200-300 kb were formed when commercially available phage DNA was applied in the annealing reaction.

Yeast chromosomes.
Yeast chromosomes were prepared as described by van Ommen and Verkerk [43]. Cells of *Saccharomyces cerevisiae* strain YT6-2-1L were grown overnight in YPD (1% (w/v) yeast extract, 2% (w/v) bactopepton and 2% (w/v) dextrose) at 30 °C, to early stationary phase and, after dilution (10 times), grown to an A_{600} of 1.0. Cells were collected by centrifugation (3000x**g**, 5 min) and pellets were taken up in 4 volumes of SE (75 mM NaCl, 25 mM EDTA, pH 7.5). At 40 °C the suspension was mixed with an equal volume of a 1% (w/v) LMP-agarose solution in SE containing dithiotreitol (DTT, 20 mM) and Zymolyase-100T (Kirin Breweries, 6 u/ml) and poored into a plug-former. Plugs were incubated for 2 h in an equal

volume of SE, DTT, Zymolyase-100T at 37 °C and subsequently lysed and stored as described for plant DNA.

Restriction digestion of embedded DNA

Prior to restriction digestion, plugs were washed eight times with a large volume (50 times the plug volume) of $T_{10}E_{10}$ (10 mM TRIS-HCL pH 8.0, 10 mM EDTA) containing 0.1 mM phenylmehylsulfonylfluoride (PMSF), for a total period of 48 h. After washing, each plug was equilibrated, twice for 30 min, with 1 ml of the appropriate digestion buffer, as recommended by the manufacturer. Each plug was then incubated in 250 µl digestion buffer which was supplemented with DTT (1mM), spermidine (8 mM) and bovine serum albumine (BSA, 100 µg/ml). Enzyme (100 units total) was added in two or three portions during the incubation period. Digestion was performed at the recommended temperature and allowed to proceed overnight. After digestion, plugs were directly transferred to the gel and subjected to electrophoresis.

Field inversion gel electrophoresis (FIGE)

Agarose gels of 1% in TBE (0.089 M TRIS-borate, 0.089 M boric acid and 0.002 M EDTA, pH 8.3) were run in a BRL horizontal slab gel (20 cm x 25 cm x 0.8 cm) electrophoresis unit (Bethesda Research Laboratories, series 1025, model H4). Although this system has not been specifically designed for FIGE or other pulsed-field electrophoresis systems, satisfactory results were obtained. Buffer (TBE) was cooled to approximately 11 °C by circulating it through a reflux cooling spiral which was kept at 4 °C with a cooler (LKB, 2219 multitemp II). A problem associated with the BRL electrophoresis tank, when used in its original configuration, concerns the temperature gradient formed between the surface of the gel, which is in direct contact with the circulating buffer, and the bottom of the gel, which is isolated from effective cooling by the perspex gel tray (6 mm thick) and tray support, resulting in bands that are bending towards the slots. This problem was solved by raising the gel tray so as to allow the circulating buffer to pass between the gel tray and the tray support.

Switching of the electric field was performed by a switchbox (field inversion unit Mark 4; Biocent BV, P.O. Box 280, 2160 AG LISSE, Netherlands) and control of the switch times of the applied fields was done by computer (Olivetti M19) equipped with a program developed by G.J.B. van Ommen (Dept. of Human Genetics, Sylvius Laboratory, University of Leiden, Netherlands) and marketed by Biocent BV. Power supply was by a BioRad unit (model 500/200). Field strengths for forward and reverse electrophoresis were the same.

The computer program makes it possible to easily adjust the electrophoresis conditions to specific needs. Parameters that can be varied include duration of the run, ratio between forward and reverse pulse times, intervals between pulses, the time ramp (gradual increase of applied pulse times throughout the duration of the electrophoresis, e.g. from 1 to 60 s) and mode (linear, logarithmic or exponential) of pulse time increment. In later experiments, the program was extended such that a sequence of cycles, differing in length and time ramp, could be used in a simple run. (On request the modified version of the original program is available.) We found that by using a sequence of different time ramps, within one electrophoresis run, a much higher resolution could be achieved (e.g. fig. 5) without the occurence of a compression or minimum mobility zone [12, 21] within the range of the separated yeast chromosomes (data not shown).

According to G.J.B. van Ommen (pers. comm.) sharper bands are obtained with intervals between pulses that are 2% of the forward pulse times and also when pulse time increases mainly during the first half of the run (viz. logarithmically). We routinely used a logarithmic increase of the pulse, by which after 40% of the total running time the pulse time had increased to half its maximum value. The applied reverse pulse times were 33% of the forward pulse time and pause intervals between pulses were 2% of the forward pulse time. Time ramps used were as indicated in the figure legends. Gels were usually run for 16 h (at 7 V/cm), sometimes however for 22, 24 or 42 h (at 5.8 V/cm).

Gels were stained after electrophoresis in 0.5 x TBE with ethidium bromide (EtBr) (1.5 mg/l) for 2 to 6 h and destained for approximately 16 h in 0.5 x TBE or water.

Southern blot hybridization

DNA fragments were transferred onto Gene-Screen Plus membranes (New England Nuclear) using the alkaline blotting procedure as described [37]. Initially we used acid depurination of DNA (0.25 M HCl for 20 min) prior to blotting, to facilitate transfer of large fragments. In agreement with van Ommen and Verkerk [43], we found, however, that complete transfer of DNA could only be achieved by irradiating the EtBr-stained gel for 5 min with 302 nm ultraviolet light before blotting.

Probes were labelled with ^{32}P-dATP, either by M13 primer extension [25], nick translation [25] or random primer labeling [14], to a specific activity of approximately 10^8 cpm/µg. Hybridization was carried out for 16 h under conditions recommended by the manufacturer, using 50% (v/v) formamide and 10% (w/v) dextran sulfate at 42 °C. After hybridization, blots were washed as recommended by the manufacturer.

Hybridization probes

The following probes for genomic DNA were used: (a) the genomic clone TG54 (kindly supplied by Dr. S. Tanksley, Cornell University, USA) which contains a 2.4 kb *Pst*I tomato genomic DNA fragment in pUC8, assigned to chromosome 6 [45] and (b) two tomato alcohol dehydrogenase-2 (*Adh-2*) cDNA clones (pCB25E6 and pAdh31, a generous gift of Dr. Th. Chase Jr., Rutgers University, USA). The *Adh-2* clones encompass, along with a third clone (pC3-7), the entire *Adh-2* coding region. Clone pC3-7 contains the first 197 bases of the coding region as well as 65 bases 5' to the starting codon. Clone pCB25E6 carries a 435 bp insert from the central region and pAdh31 a 688 bp segment of the 3'-terminal region [8a].

Purified total chloroplast DNA of tomato was a gift from Dr. E. Wisman (Wageningen, Netherlands).

Results

To determine the best source for isolating high molecular weight plant DNA, tomato nuclei isolated from plant tissue, suspension cells and leaf protoplasts were embedded in agarose. Embedded nuclei and protoplasts were then subjected to the *in situ* lysis procedure and the released DNA molecules fractionated by FIGE. Embedded suspension cells were treated with macerozyme and cellulase prior to incubation with the lysis mix. DNA released from embedded nuclei was comparable in size to conventionally isolated DNA (50-150 kb). DNA from the embedded suspension cells was somewhat larger (150-250 kb) but still not of the desired size (data not shown). Protoplasts clearly gave the best results (>2000 kb) and only the use of these will be discussed further.

To determine whether the lysis and restriction endonuclease digestion conditions employed for isolating high molecular weight DNA from embedded human and animal cells were applicable to tomato protoplasts, two series of embedded protoplasts were incubated with lysis mix for different periods, ranging from 1 to 48 h. One series was subjected to electrophoresis to determine the size of the released DNA (fig. 1A), whereas the other series was incubated with *Eco*RI prior to electrophoresis (fig. 1B).

In the absence of restriction endonuclease, the bulk of DNA released from the protoplasts was found to be retained in the slots, irrespective of the length of the lysis period (fig. 1A). This result strongly suggests the bulk of released DNA to be at least 2 Mb in size (which is the size of the largest yeast chromosome). Remarkably, one hour incubation with lysis mix was already sufficient to render the DNA accessible to digestion with *Eco*RI (fig. 1B). Longer incubation had an effect only on the amount of DNA released and not on the specificity of the digestion. After 48 h of lysis and subsequent treatment with *Eco*RI the slot was virtually devoid of DNA and the smear of *Eco*RI fragments was most intense, whereas

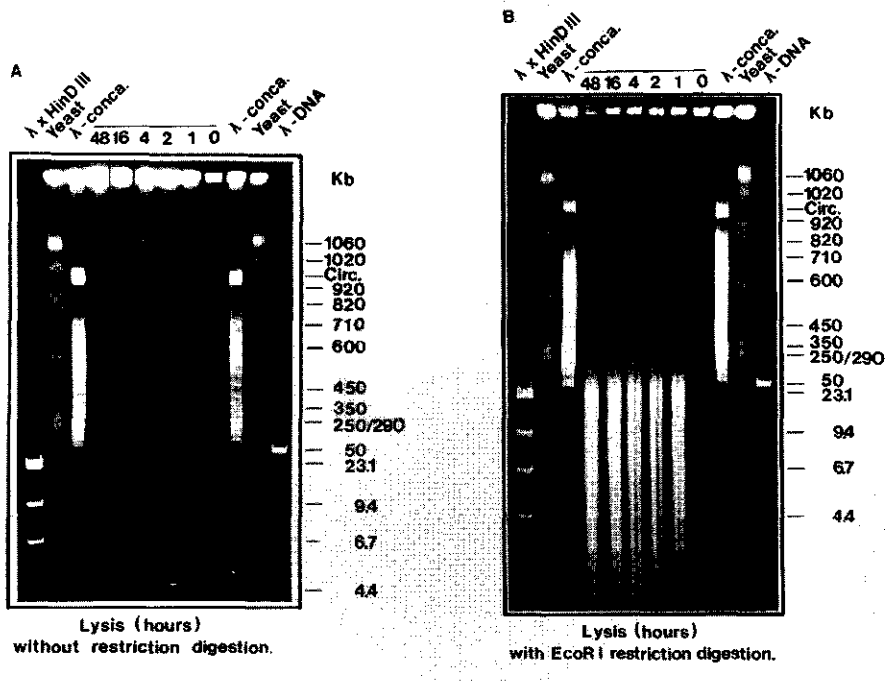


Figure 1. Release and restriction digestion of DNA from agarose-embedded tomato protoplasts following incubation in lysis mix for various times. Two series of agarose-embedded protoplasts were incubated with lysis mix for the periods indicated. One series (A) was then subjected directly to electrophoresis. The other series (B) was incubated with *Eco*RI and subsequently analysed by electrophoresis. Electrophoresis conditions were (for both gels): time ramp 3 to 65 s, field strength 7 V/cm, 16 h run. The ethidium bromide-stained DNA patterns are shown.

the pattern of the repeat bands within the digestion smear was the same, irrespective of lysis time (fig. 1B).

In addition to the bulk of high molecular weight DNA, which remained in the slot after electrophoresis, there was always a small amount of DNA entering the gel (100-1000 kb; fig 1A). As the amount of this 'small' DNA released from embedded protoplasts was found to vary among different batches, it seemed plausible that this degraded DNA was, to a large extent,

derived from non-viable protoplasts present to a variable proportion in the respectivepreparations. Two batches of protoplasts were chosen for comparison, which contained 90% and 60% viable cells, respectively. The protoplasts were embedded and subjected to the lysis and electrophoresis procedure. Subsequent staining with EtBr indeed revealed the presence of a much higher proportion of degraded DNA in the protoplast preparation containing only 60% viable cells (fig. 2).

The results presented above demonstrate that *in situ* lysis of viable tomato protoplasts yields DNA that is far larger than conventionally isolated DNA and, in addition, is accesible to restriction digestion, as judged from the EtBr staining patterns. To assess whether restriction

digestion in agarose proceeds to completion and at the desired specificity, the *Eco*RI and *Hin*dIII restriction patterns of conventionally isolated DNA and *in situ* digested protoplast DNA were compared by Southern blot hybridization analysis, using a probe for the *Adh-2* gene (pCB25E6). For each of the enzymes used the expected fragments (1.0 kb for *Eco*RI, 8.0 kb for *Hin*dIII; fig. 3) were found, demonstrating that both *in situ* digestion and digestion in solution are comparable.

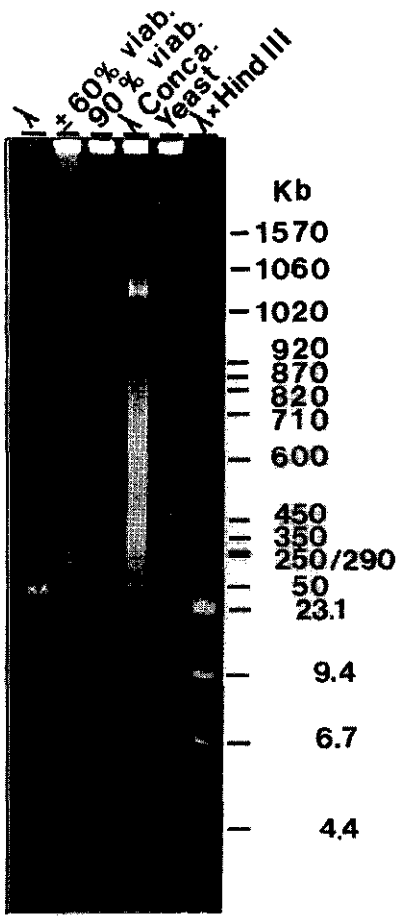


Figure 2. Effect of the proportion of viable cells in a batch of protoplasts on the size of DNA released after lysis of embedded protplasts. Protoplasts batches (50 million epr ml) containing approximately 60% and 90% viable cells, as determined with FDA staining, were embedded in agarose, lysed for 48 h and subjected to FIGE. Electrophoresis conditions were: 24 h run, time ramp 4 to 100 s, field strength of 5.8 V/cm. After electrophoresis gels were stained with ethidium bromide.

Having established the proper *in situ* digestion conditions, we tested a series of restriction enzymes, including *Kpn*I, *Pst*I and the 'rare cutters' *Mlu*I, *Sfi*I, *Not*I, *Sal*I and *Cla*I (fig. 4), *Pvu*I, *Sma*I and *Sst*II (fig. 5), and *Nru*I and *Xho*I (fig. 6) known to generate high molecular weight fragments [2]. In all those cases, digestion resulted in the appearance of a smear of DNA fragments ranging in size from 20 kb to at least 1600 kb. Subsequent blotting and hybridization using an *Adh-2* probe (pCB25E6) revealed the presence of distinct high molecular weight fragments in the *Mlu*I (300 and 700 kb), *Sal*I (100 and 500 kb), *Cla*I (150 kb), *Pst*I (25 kb), *Pvu*I (500 kb) and *Sst*II (700 kb) digests (fig. 4 and 5) and in the *Nru*I digest (170 kb, not shown).

Incubations with larger amounts of enzyme did not alter the *Mlu*I and *Sal*I hybridization patterns, suggesting that the presence of two *Adh-2* containing fragments in each digest was not due to partial digestion. Rehybridization of the blot with an *Adh-2* cDNA probe (pAdh31) which contains the 3'-terminal coding region absent in pCB25E6, showed the same two high molecular weight *Mlu*I and *Sal*I fragments, indicating that the presence of two fragments is not due to a *Mlu*I and *Sal*I site within the gene (data not shown).

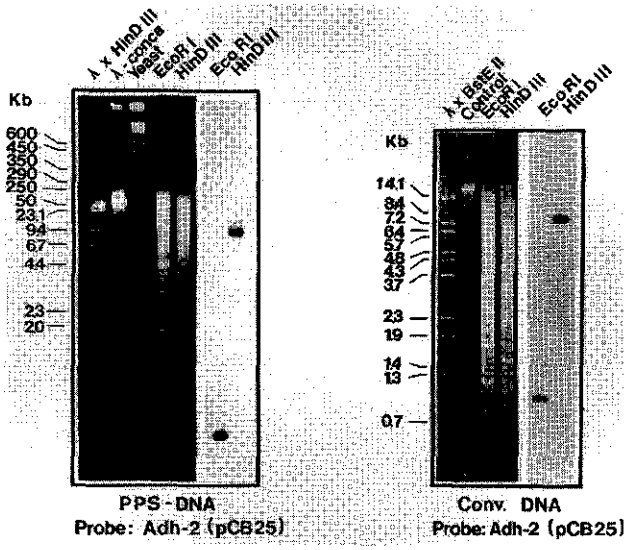


Figure 3. Digestion of tomato DNA in solution and in agarose. *Left panel.* Agarose embedded tomato protoplasts (PPS) were lysed and the released DNA was digested with *Eco*RI and *Hin*dIII. Restriction fragments were resolved by FIGE (time ramp 1 to 60 s, 24 h run, field strength 3.5 V/cm). After staining with ethidium bromide, fragments were blotted and hybridized with an *Adh-2* probe of tomato (pCB25E6). The two lanes at the right show the autoradiograph after 3 days of exposure with one intensifying screen. *Right panel.* Total DNA was isolated from tomato leaf tissue, using the CTAB procedure of Murray and Thompson [31]) (Conv. DNA) and 5 µg was digested with *Eco*RI and *Hin*DIII. Restriction fragments were separated by agarose gel electrophoresis (0.8% (w/v) gel). After staining with ethidium bromide, fragments were blotted and hybridized with a probe (pCB25E6) for the *Adh-2* gene of tomato. Autoradiographs (two lanes on the right) were exposed for 2 days with one intensifying screen.

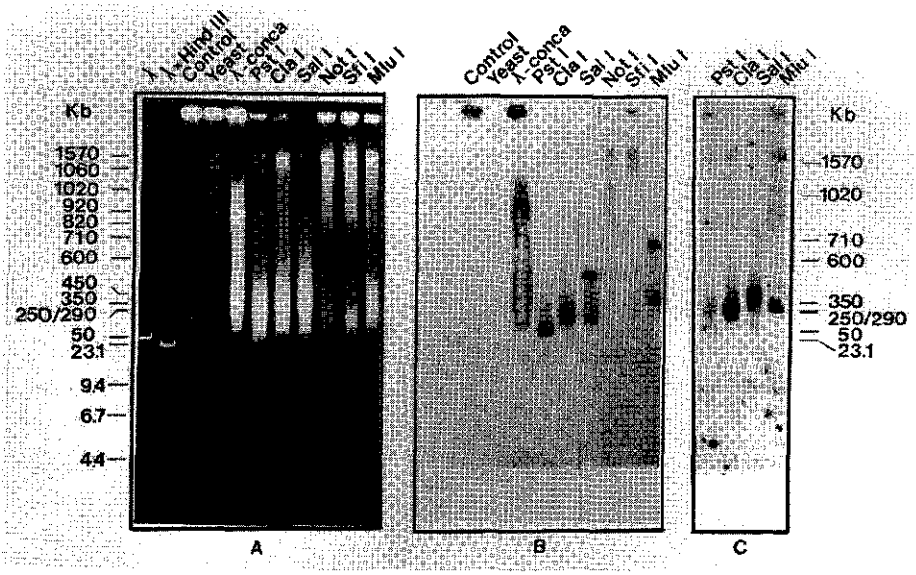


Figure 4. Restriction digestion of agarose-embedded tomato protoplasts and hybridization analysis using an *Adh-2* and TG54 probe. A. Agarose-embedded protoplasts were lysed for 48 h and the released DNA was digested with the restriction enzymes indicated and electrophoresed using a time ramp of 1 to 60 s 17 h run, 7 V/cm). The ethidium bromide stained FIGE gel is shown. B. Autoradiograph (3 days exposure, one intensifying screen) of the blotted gel, after hybridization with an *Adh-2* probe of tomato (pCB25E6). C. Autoradiograph (3 days exposure, one intensifying screen) of the blotted gel after hybridization with the TG54 probe of tomato.

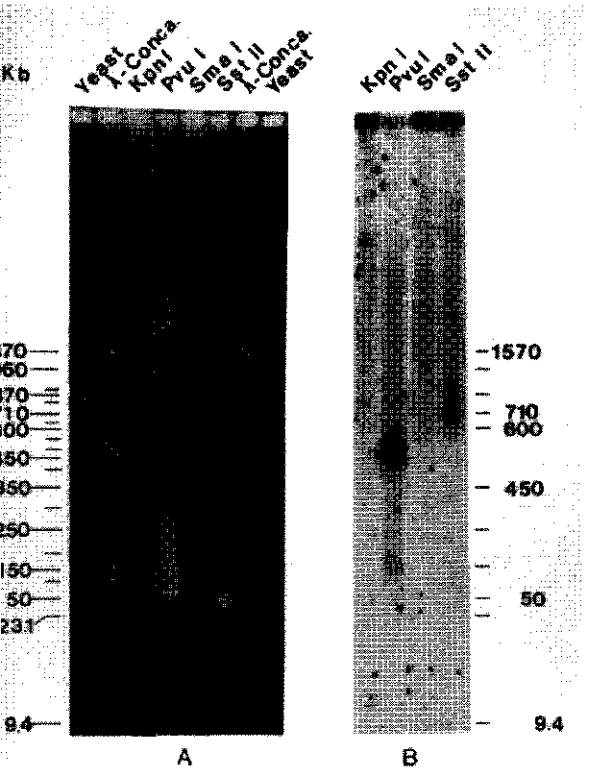


Figure 5. Restriction digestion of agarose embedded tomato protoplasts and hybridization analysis using an *Adh-2* probe. A. Agarose embedded protoplast were lysed (48 h) and the released DNA was digested with the restriction enzymes indicated and electrophoresed using a sequence of time ramps of 1-20 s (8 h), 3-65 s (14 h) and 5-150 s (20 h) respectively, at 5.8 V/cm). The ethidium bromide-stained FIGE gel is shown. B. Autoradiograph (1 day exposure, one intensifying screen) of the blotted gel after hybridization with an Adh-2 probe (pCB25E6).

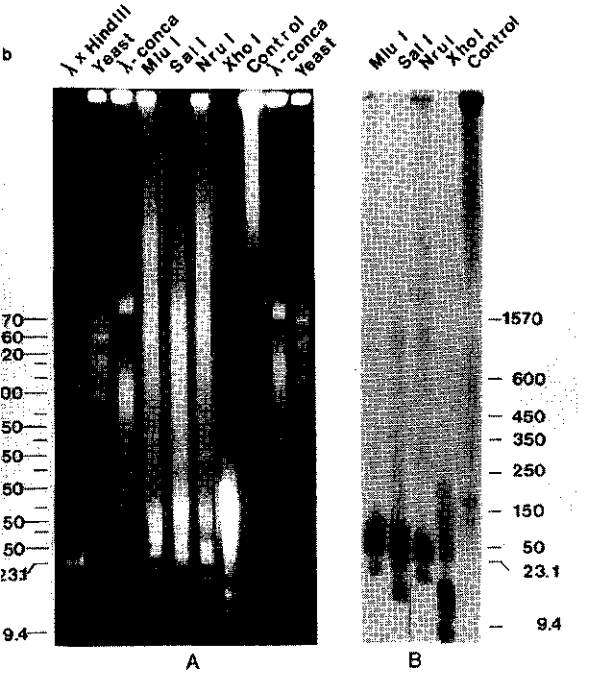


Figure 6. Restriction digestion of agarose embedded tomato protoplasts and hybridization analysis using total chloroplast DNA as a probe. A. Agarose embedded protoplast were lysed for 48 h and the released DNA was digested, with the restriction enzymes indicated and electrophoresed using a sequence of time ramps of 1-20 s (8 h), 3-65 s (14 h) and 5-150 s (20 h) respectively, at 5.8 V/cm. The ethidium bromide-stained FIGE gel is shown. B. Autoradiograph (15 min exposure, one intensifying screen) of the blotted gel, following hybridization with total nick-translated chloroplast DNA as probe.

The specificity of the hybridization signals was further demonstrated after removal of the probe and subsequent rehybridization with the genomic clone TG54 [45]. Unlike the Adh-2 hybridization patterns, a single band was now detected in both the *Sal*I and *Mlu*I digestions, approximately 400 kb and 250 kb in size, respectively (fig. 4C).

A smear at the high molecular weight region of the gel, rather than specific hybridization signals, was observed in the *Sfi*I and *Not*I digests. Since both enzymes were active *in situ*, as shown by the appearance of distinct, low molecular weight chloroplast DNA (see below) fragments (fig. 4), apparently no sites for these enzymes are present within the probed region. With the enzymes *Kpn*I and *Sma*I (fig. 5) we never obtained specific hybridization signals, nor any chloroplast bands.

In summary, these results indicate that *in situ* digestion by *Eco*RI, *Hin*dIII and *Pst*I and the 'rarely cutting' enzymes *Mlu*I, *Sal*I *Cla*I, *Pvu*I, *Sst*II, *Xho*I and *Nru*I indeed proceeds to completion and is highly useful in generating restriction fragments of plant DNA in the 100-1000 kb range.

As leaf cell protoplasts were used as source of nuclear DNA it was of interest to follow the fate of chloroplast DNA during FIGE after its release from embedded protoplasts. Since large circular molecules (both supercoiled and relaxed) are known to migrate in FIGE [24], albeit at a much slower rate than linear molecules of the same size, tomato chloroplast DNA (156-159 kb [33]) was expected to enter the gel. Surprisingly, unrestricted chloroplast DNA remained within the slot after lysis of the protoplasts, as shown after hybridization using total nick-translated chloroplast DNA as a probe (fig. 6). On the other hand, after restriction digestion with *Xho*I, *Nru*I, *Sal*I and *Mlu*I, specific bands could be detected (in the range from 9 to 100 kb, fig 6B), which corresponded to the bands staining intensely with EtBr, indicating that the chloroplast DNA molecules were released from the chloroplasts. These results show that intact circular chloroplast DNA, as it is released from the embedded protoplasts, is not mobile under the electrophoresis conditions used.

Discussion

As far as we know, no conditions have been described for isolating plant DNA, sufficient in size to allow long-range restriction mapping. We have shown here that the procedure used for releasing high molecular weight DNA from agarose-embedded yeast and animal cells can be applied to tomato protoplasts. Provided protoplasts preparation containing a high proportion of viable cells were embedded, the bulk of the released DNA was at least 2 Mb in size and amenable to digestion with a variety of restriction enzymes, including the commonly used 'rare cutters'. Using approximately 3 million protoplasts per agarose plug, sufficient DNA was obtained to allow the detection, by Southern blot analysis, of a single copy sequence

(*Adh-2* and TG54) within large restriction fragments. Digestion of the agarose-embedded tomato DNA proceeds to completion and is similar to digestion in solution.

Since we found similar results when using protoplasts from pea and soybean (data not shown), we feel confident that this procedure can be of general use for isolating megabase-sized plant DNA.

A possible explanation for the detection of mulitple bands in the *Mlu*I and *Sal*I digests, after hybridization with the Adh-2 cDNA probe, might be partial methylation of the respective sites. The recognition sites for *Mlu*I (ACGCGT) and *Sal*I (GTCGAC) contain the dinucleotide CpG, which is known to be heavily methylated at the cytosine residue in higher-plant DNA [18, 26]. On the other hand, partial digestion due to methylation can be very helpfull in preparing long-range restriction maps [3, 10, 13, 23].

The conditions described in this paper for generating and identifying large restriction fragments of plant DNA should facilitate the construction of physical maps of plant chromosomal regions of interest, using 'high-altitude chromosome walking' techniques [4, 32] or chromosome jumping and linking libraries [11, 34].

Acknowledgements

We thank Dr. Th. Chase Jr. for the *Adh-2* clones, Dr. S. Tanksley for the clone TG54, Jacques Kirch for his help in modifying the computer program, Dr. A. van Kammen and O. Richards for critically reading the manuscript, the other members of the tomato group for stimulating discussion and P. van Druten for photography.

References

1. Anand R (1986) Pulsed field gel electrophoresis: a technique for fractionating large DNA molecules. Trends Genet 2: 278-283
2. Barlow DP, Lehrach H (1987) Genetics by gel electrophoresis: the impact of pulsed field gel electrophoresis on mammalian genetics. Trends Genet 3: 167-171.
3. Brown WRA, Bird AP (1986) Long-range restriction site mapping of mammalian genomic DNA. Nature 322: 477-481.
4. Burke DT, Carle GF, Olson MV (1987) Cloning of large segments of exogeneous DNA into yeast by means of artificial chromosome vectors. Science 236: 806-812.
5. Burmeister M, Lehrach H (1986) Long-range restriction map around the Duchenne dystrophy gene. Nature 324: 582-585.
6. Carle GF, Olson MV (1985) An electrophoretic karyotype for yeast. Proc Natl Acad Sci USA 82: 3756-3760.
7. Carle GF, Olson MV (1984) Separation of chromosomal DNA molecules from yeast by orthogonal-field-alternation-gel electrophoresis. Nucleic Acids Res 12: 5647-5664.
8. Carle GF, Frank M, Olson MV (1986) Electrophoretic separation of large DNA molecules by periodic inversion of the electric field. Science 232: 65-68.
8b. Chase T, Williams BG (1986) Characterizatioon of complete cDNA for tomato Adh-2. Abstract Tomato Biotechnology Symposium, University of California, Davis, august 20-22.

9. Chu G, Vollrath D, Davis RW (1986) Separation of large DNA molecules by contour-clamped homogeneous electric fields. Science 234: 1582-1585.
10. Collins FS, Drumm ML, Cole JL, Lockwood WK, Van den Woude CF, Ianuzzi MC (1987). Construction of a general human chromosome jumping library, with application to cystic fibrosis. Science 235: 1046-1049.
11. Collins FS, Weissman SM (1984). Directional cloning of DNA fragments at a large distance from an initial probe: a circularization method. Proc Natl Acad Sci USA 81: 6812-6816.
12. Ellis THN, Cleary WG, Burcham KWG, Bowen BA (1987) Ramped field inversion gel electrophoresis: a cautionary note. Nucl Acids Res 15: 5489.
13. Estivill X, Farrall M, Scambler PJ, Bell GM, Hawley KMF, Lench NJ, Bates GP, Kruyer HC, Frederick PA, Stanier P, Watson EK, Williamson R, Wainwright BJ (1987) A candidate for the cystic fibrosis locus isolated by selection for methylation-free islands. Nature 326: 840-845.
14. Feinberg AP, Vogelstein B. (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 132: 6-10.
15. Frearson EM, Power JB, Cooking EC, (1973). The isolation, culture and regeneration of *Petunia* leaf protoplasts. Dev Biol 33: 130-137.
16. Galbraith DW, Harkins KR, Maddox JM, Ayres NM, Sharma DP, Firoozabady E. (1983) Rapid flow cytometric analysis of the cell cycle in intact plant tissues. Science 220: 1049-1051.
17. Garvey EP, Santi RK (1986) Stable amplified DNA in drug resistant *Leishmania* exists as extrachromosomal circles. Science 233: 535-540.
18. Gruenbaum Y, Naveh-Maney T, Cedar H, Razin A (1981) Sequence specificity of methylation in higher plant DNA. Nature 292: 860-862.
19. Hermann BG, Barlow DP, Lehrach H (1987) A large inverted duplication allows homologous recombination between chromosomes heterologous for the proximal t-complex inversion. Cell 48: 813-825.
20. Kenwrick S, Patterson M, Speer A, Fischbeck K, Davies K. (1987) Molecular analysis of the Duchenne muscular dystrophy region using pulsed field gel electrophoresis. Cell 48: 351-357.
21. Lalande M, Noolandi J, Turmel C, Rousseau J, Slater GW (1987) Pulsed-field electrophoresis: application of a computer model to the separation of large DNA molecules. Proc Natl Acad Sci USA 84: 8011-8015.
22. Larkin RH (1976) Purification and viability determinations of plant protoplasts. Planta 128: 213-216.
23. Lawrence SK, Srivastava R, Rigas B, Chorney MJ, Gillespie GA, Smith CL, Cantor CR, Collins FS, Weissman SM (1986) Molecular approaches to the characterization of megabase regions of DNA: applications to the human major histocompatibility complex. Cold Spring Harbor Symp Quant Biol 51: 123-130.
24. Levene SD, Zimm BH (1987) Separation of open-circular DNA using pulsed-field electrophoresis. Proc Natl Acad Sci USA 84: 4054-4057
25. Maniatis T, Fritsch EF, Sambrook J (1983) Molecular cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
26. McClelland M (1983) The frequency and distribution of methylatable DNA sequences in leguminous plant protein coding genes. J Mol Evol 19: 346-354.
27. Medina-Filho HP, Stevens MA (1980) Tomato breeding for nematode resistance: survey of resistant varieties for horticultural chraracteristics and genotype of acid phosphatases. Acta Hort 100: 383-391.
28. Meredith CP (1979) Shoot development in established callus cultures of cultivated tomato (*Lycopersicon esculentum* Mill.) Z Pflanzenphysiol 95: 405-411.
29. Michiels F, Burmeister M, Lehrach H (1987) Derivation of clones close to *met* by preparative field inversion gel electrophoresis. Science 236:1305-1308.
30. Monaco AP, Kunkel L (1986) A giant locus for the Duchenne and Becker muscular dystrophy gene. Trends Genet 3: 33-37.
31. Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8: 4321-4332.
32. Nasmyth K, Sulston J (1987) High-altitude walking with YAC's. Nature 328: 380-381.
33. Philips AL (1985) Restriction map and clone bank of tomato plastid DNA. Current Genet 2: 174-179.
34. Poustka A, Lehrach H (1986) Jumping libraries and linking libraries: the next generation of molecular tools in mammalian genetics. Trends Genet. 2: 174-179
35. Poustka A, Pohl T, Barlow DP, Zehetner G, Craig A, Michiels F, Ehrich E, Frischauf AM, Lehrach H (1986) Molecular appoaches to mammalian genetics. Cold Spring Harbor Symp Quant Biol Vol 51: 131-139.
36. Poustka A, Pohl T, Barlow DP, Frischauf AM, Lehrach H (1987) Construction and use of human chromosome jumping libraries from *Not*I-digested DNA. Nature 325: 353-355.
37. Reed KC, Mann DA (1985) Rapid transfer of DNA from agarose gels to nylon membranes. Nucleic Acids Res 13: 7207-7221.

38. Schwartz DC, Cantor CR (1984) Separation of yeast chromosome-sized DNA by pulsed field gradient gel electrophoresis. Cell 37: 67-75.
39. Schwartz DC, Saffran W, Welsh J, Haas R, Goldenberg M, Cantor CR (1982) New techniques for purifying large DNAs and studying their properties and packaging. Cold Spring Harbor Symp Quant Biol 47: 189-195.
40. Tan MMC, Boerrigter HS, Kool AJ (1987) A rapid procedure for plant regeneration from protoplasts isolated from suspension cultures and leaf mesophyl cells of wild *Solanum* species and *Lycopersicon pennellii*. Plant Sci 49: 63-72.
41. Van der Ploeg LHT, Schwartz DC, Cantor CR, Borst P (1984) Antigenic variation in *Trypanosoma brucei* analyzed by electrophoretic separation of chromosomal DNA molecules. Cell 37: 77-84.
42. Van de Ploeg LHT, Smits M, Ponnuburai T, Vermeulen A, Meuwissen JHTh, Langsley G (1985) Chromosome-sized DNA molecules of *Plasmodium falciparum*. Science 229: 658-661.
43. Van Ommen GJB, Verkerk JMH (1986) Restriction analysis of chromosomal DNA in a size range up to two million base pairs by pulsed field gradient electrophoresis. In: Davis KE (Ed) Human Genetic Disease, A Practical Approach, IRL Press, Oxford pp 113-133.
44. Van Ommen GJB, Verkerk JMH, Hofker MH, Monaco AP, Kunkel LM, Ray R. Worton R. Wieringa B, Bakker E, Pearson PL (1986) A physical map of 4 million bp around the Duchenne muscular dystrophy gene on the human X-chromosome. Cell 47: 499-504.
45. Young ND, Miller JC, Tanksley SD (1987) Rapid chromosomal assignment of multiple genomic clones in tomato using primary trisomics. Nucleic Acids Res 15: 9339-9348.

CHAPTER 3

LONG-RANGE RESTRICTION MAPS OF TWO LOCI (*APS-1* AND GP79) FLANKING THE ROOT-KNOT NEMATODE RESISTANCE GENE (*Mi*) NEAR THE CENTROMERE OF TOMATO CHROMOSOME 6.

Abstract

The root-knot nematode resistance gene *Mi* in tomato has been mapped in the pericentromeric region of chromosome 6. With the objective of isolating *Mi* through a map-based cloning approach, we have previously identified and odered into a high-resolution genetic linkage map a variety of tightly linked molecular markers. Using pulsed-field gel electrophoresis and various rarely cuttting restriction enzymes in single, double, and partial digestions, we now report long range restriction maps of the two closest flanking markers, acid phosphatase-1 (*Aps-1*) and GP79, which span over 400 and 800 kb, respectively. It is concluded that the physical distance between both markers is much larger than predicted on the basis of genetic linkage analysis. Furthermore, two RFLP markers (H3F8 and H4H10) which map genetically to the same locus as *Aps-1* do not show physical linkage, indicating severe suppression of recombination in this region of the chromosome. Finally, no evidence was obtained showing the presence of a CpG island near *Aps-1*.

This chapter is published: Raymond A.J.J. van Daelen, Frans Gerbens, Fred van Ruisen, Jac Aarts, Jan Hontelez, and Pim Zabel (1993) Long range physical maps of two loci (*Aps-1* and GP79) flanking the root-knot nematode resistance gene (*Mi*) near the centromere of tomato chromosome 6. Plant Mol Biol 23: 185-192.

Introduction

Infection of tomato (*Lycopersicon esculentum* Mill.) with root-knot nematodes (*Meloidogyne* spp) leads worldwide to large crop losses [32]. Resistance against the major root-knot nematodes has been identified in the wild tomato species *L. peruvianum* (L.) Mill. and introduced into the cultivated tomato through breeding programs. The resistance is based upon a single dominant gene (*Mi*), located near the centromere of chromosome 6 [9]. To date, nothing is known about the mode of action or the product of the gene. Since the genetic position of *Mi* is known, molecular identification of the gene may be achieved through positional cloning [37, 38].

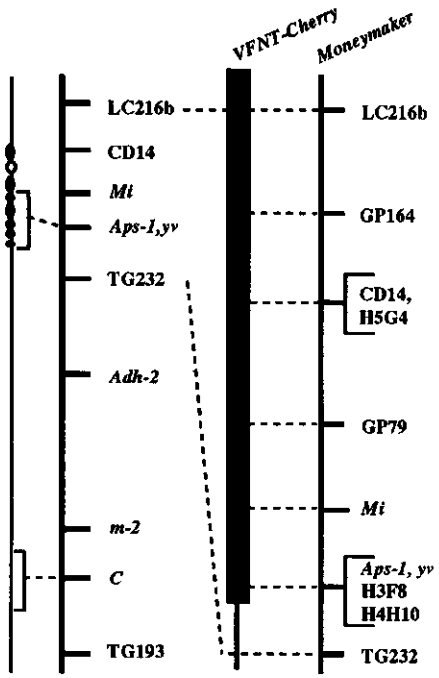


Figure 1. Map of the *Mi*-region of chromosome 6. A cytological (pachytene) map of chromosome 6 is shown at the left with the centromere and heterochromatin indicated by open and black circles, respectively Chromosomal regions of *L. esculentum* are represented as vertical lines on the molecular linkage map, the region of *L. peruvianum* origin as a vertical black bar (see also Ho, *et al* [21]).

So far, most examples of genes isolated by a positional cloning approach are to be found among human hereditary diseases including cystic fibrosis [31] and neurofibromatosis type 1 [37] but more recently two plant genes have been identified: the *ABI3* gene of *Arabidopsis* [20] and a gene involved in fatty acid desaturation identified [4]. Instrumental to this approach are a highly saturated molecular linkage map and techniques for physical mapping and cloning large DNA fragments. As these requirements have been met for tomato [27, 35] cloning of the root-knot nematode resistance gene *Mi* should be feasible by this approach.

Recently a variety of tightly linked molecular markers have been identified including the acid phosphatase-1 (*Aps-1*) gene [1, 12, 33, 39] and various RFLP [21, 23] and RAPD [24] markers. These markers have been ordered into a high resolution molecular linkage map of the *Mi* region [21, 28], as shown in figure 1. Due to strong suppression of recombination accurate genetic mapping in the *Mi* region proved

rather cumbersome, with genetic distances ranging from 0.3 to 2 cM between the closest *Mi* flanking markers, *Aps-1* and GP79 [21, 28]. To translate the genetic distances of the markers into physical distances we have constructed physical maps of the two closest markers bracketing *Mi*, GP79 and *Aps-1*, using a set of rarely cutting restriction enzymes and pulsed field gel electrophoresis. It is shown that these two markers and two other markers (H3F8 and H4H10) which map genetically to the same locus as *Aps-1* [21, see Fig. 1] are physically unlinked.

Materials and methods

Isolation of high-molecular-weight DNA and restriction digestion.

Isolation and restriction digestion of megabase-sized tomato DNA from nematode-resistant (VFNT Cherry) and susceptible (Moneymaker) genotypes was performed as described [35]. For double restriction digestions, the enzyme used in the first reaction was removed by washing the plugs overnight at 4° C in a large volume (50 - 100 ml) of 10 mM Tris-HCl pH 8.0, 10 mM EDTA before addition of the second enzyme. For partial digestions, the enzyme was added in amounts ranging from 0.01 to 20 units and allowed to diffuse into the plugs by keeping them on ice for one hour prior to incubation for two hours at the appropriate temperature. The appearance of chloroplast DNA bands in ethidium bromide-stained pulsed field electrophoresis gels (except for *Not*I) was used as an indication for proper digestion [35].

Pulsed Field Gel electrophoresis (PFGE), Southern blotting, hybridization analysis.

Agarose gels of 1% (w/v) were run in 0.5 x TBE [26] at 15 °C using a field inversion (FIGE)[10] or a CHEF [11] (Biorad DRII) apparatus. Electrophoresis conditions and markers are specified in figure legends.

Gels were blotted as described [36]. Hybridizations were performed as described [23] using probes radiolabeled according to the random hexamer priming procedure [13]. As the apparent sizes of the restriction fragments could differ substantially between different runs (in some extreme cases up to 30 kb) the values presented were always based on at least two, usually three to four, independent experiments.

Blots were deprobed by boiling them for 15 min in 0.1 x SSC (Standard Saline Citrate [26]) and 1 % (w/v) SDS and checked for complete removal of the probe by exposing them for 2 to 3 days to X-ray film.

Probes.

A genomic clone of the *Aps-1* gene was prepared as follows. The 2.4 kb genomic *Aps-1*[1] fragment amplified by PCR under direction of primers as described [1] was ligated into the *Sma*I site of PTZ18 (Pharmacia) and propagated in *Escherichia coli* DH5αF' after transformation. The cloned insert was released from the recombinant plasmid by restriction digestion, resolved by gelelectrophoresis and purified by freeze squeeze. The isolation and characterization of the RFLP markers GP79, H3F8 and H4H10 has been described previously [19, 21, 23]. Hybridization probes of these markers were also made from gel-purified inserts using the random hexamer priming procedure.

Results

To construct physical maps of the loci *Aps-1* and GP79 tomato DNA from the nematode resistant cultivar VFNT Cherry and from the sensitive cultivar Moneymaker was digested to completion with a large set of rare-cutting restriction endonucleases (*Bss*HII, *Cla*I, *Eag*I, *Fsp*I, *Mlu*I, *Not*I, *Pvu*I, *Rsr*II, *Sal*I, *Sfi*I, *Sma*I, *Sst*II and *Xho*I) and subjected to Southern analysis following PFGE at several separation regimes. In all digests, except for *Not*I and *Rsr*II, distinct bands were observed, ranging from 50 kb to several megabases (fig. 2 and table 1). In some cases (*Cla*I, *Fsp*I and *Mlu*I) the GP79 probe hybridized to multiple bands as did the *Aps-1* probe. These bands probably represent "natural partials" [29], originating from partially methylated recognition sites, since digestion using an excess of enzyme did not alter the pattern.

Previously, it has been shown by conventional Southern analysis that the introgressed region containing *Mi* in the cultivar VFNT Cherry is large and highly polymorphic [21, 28] (because of its *L. peruvianum* origin) when compared to susceptible lines. Using PFGE and rare-cutting restriction endonucleases a similar result has now been obtained. Virtually all the bands detected by both the GP79 and *Aps-1* probes in the VFNT Cherry differed in size from those in the susceptible cultivar Moneymaker (table 1).

In view of the very tight genetic linkage of GP79 and *Aps-1* [21, 28] it would have been expected to find a (large) restriction fragment hybridizing to both probes. However, in none of the digests were common fragments detected. In an attempt to extend the covered range for both probes a series of partial digestions was carried out using the enzymes *Cla*I, *Eag*I, *Fsp*I, *Mlu*I, *Pvu*I and *Sma*I (table 2). In figure 3 examples are shown for *Eag*I and *Fsp*I. In addition to the major fragments generated upon complete digestion several larger, partial restriction fragments

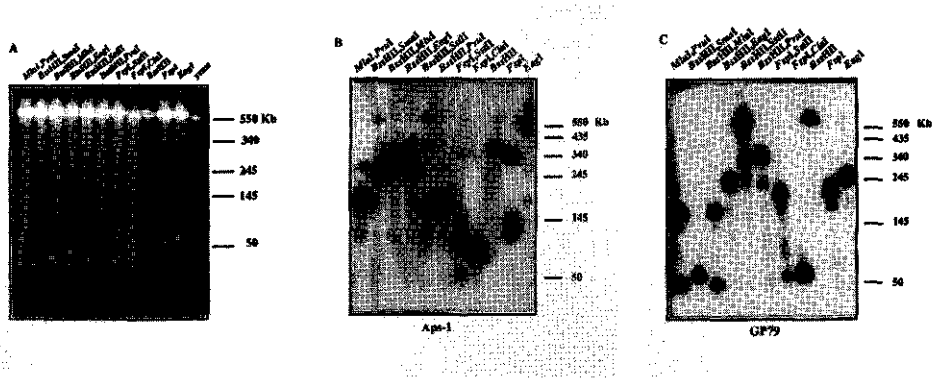


Figure 2. Southern blot hybridization analysis of restricted DNA from the cv VFNT Cherry using the tightly linked markers GP79 and *Aps-1* as probe. Agarose-embedded protoplasts of VFNT Cherry were lysed and the released DNA was digested with the restriction enzymes as indicated. Restriction fragments were resolved on a CHEF gel (electrophoresis conditions: 16 h ramp with pulses from 1 to 30 s followed by 9 hours of 30 s pulses at 185 V). After staining with ethidium bromide (A; note the chloroplast DNA banding pattern in the lower part of the gel, indicating complete digestion) fragments were blotted and hybridized to a probe corresponding to the *Aps-1* (B) and GP79 (C) locus, respectively. Autoradiographs were exposed for one day using one intensifying screen. DNA fragment sizes (in kb) are shown at the side.

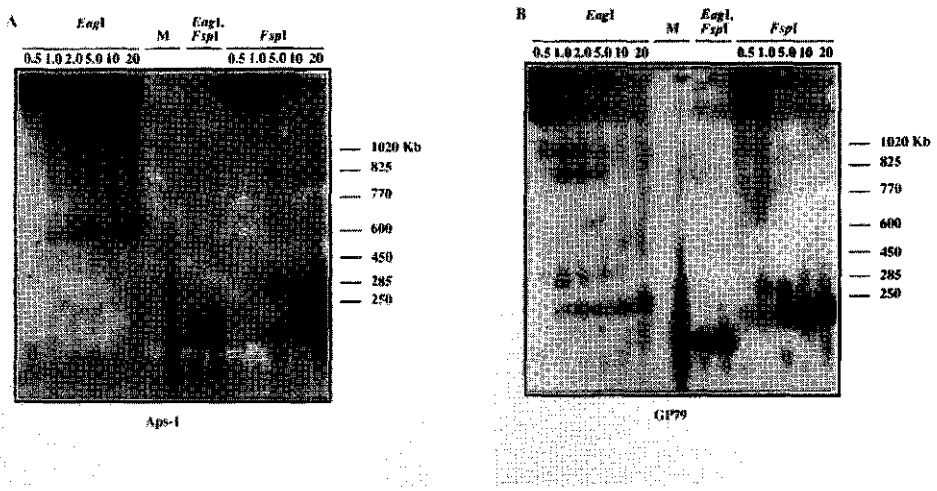


Figure 3. Southern blot hybridization analysis of partially restricted DNA from the cv VFNT Cherry using the tightly linked markers GP79 and *Aps-1* as probe. Agarose-embedded protoplasts of VFNT Cherry were lysed and the released DNA was digested with *Eag*I and *Fsp*I using the amounts (units) as indicated. The lanes denoted with M contained the size markers. Restriction fragments were resolved on a CHEF gel (16 hour ramp with pulses from 1 to 30 s followed by 9 hours of 30 s pulses, at 185 V). After staining with ethidium bromide, fragments were blotted and hybridized to a probe corresponding to the *Aps-1* (A) and GP79 (B) locus, respectively. Autoradiographs were exposed for two days using one intensifying screen. DNA fragment sizes (kb) are shown at the side.

Table 1. Restriction fragments (in kb) of digested tomato DNA (cv's VFNT Cherry and Moneymaker) after hybridization with the indicated probes.

	GP79		*Aps-1*		H3F8	H4H10
	VFNT	Moneymaker	VFNT	Moneymaker	VFNT	VFNT
*Bss*HII	> 700	> 600	390	> 600	> 2000	nd[2]
*Cla*I	65/135/205	nd[1]	90	nd[1]	nd[2]	> 2000
*Eag*I	250	260	720	170/270	nd[2]	> 2000
*Fsp*I	(180)/220	200	105/130/320	110/150	500	> 1500
*Mlu*I	45/160	40/120	280	370	> 2000	> 2000
*Not*I	> 6000	nd[1]	> 6000	nd[1]	> 6000	nd[2]
*Pvu*I	360	> 600	340	270/400	1500	nd[2]
*Rsr*II	> 6000	nd[1]	> 6000	nd[1]	nd[1]	nd[1]
*Sal*I	200/400	300	450	150	nd[2]	700
*Sfi*I	4200	nd[1]	> 6000	nd[1]	> 6000	nd[2]
*Sma*I	60	> 600	400	370	> 2000	> 2000
*Sst*II	> 700	> 600	330	> 600	nd[2]	nd[1]
*Xho*I	< 50	< 50	< 50	< 50	nd[1]	nd[1]

[1]: Not determined.
[2]: No clear signal obtained.
>: Indicates that a strong signal was detected in the compression zone, which, depending on the pulse regime applied, contains fragments larger in size than the number following the sign (see also Fig. 5 [10])

Table 2. Restriction fragments of partially digested digested tomato DNA (cv VFNT Cherry) after hybridization with *Aps-1* and GP79. Sizes of the fragments are in kb and estimates are based on one, sometimes two, observations.

enzyme	probe	
	Aps-1	GP79
*Eag*I	720/>1500	250/290/850/1000
*Fsp*I	105/130/320	180/220/650
*Pvu*I	340/>1500	360/450/760/800
*Cla*I	90/150/200/370/400	65/135/205
*Mlu*I	nd 45/180/350/470/500	
*Sma*I	400/590	60/610/650

nd: Fragment sizes not determined.

Table 3. Restriction fragments (in kb) of double digested tomato DNA (cv VFNT Cherry) after hybridization with Aps11(lower left panel) and GP79 (upper right panel).

	*Bss*HII	*Cla*I	*Eag*I	*Fsp*I	*Mlu*I	*Not*I	*Pvu*I	*Sma*I	*Sst*II	GP79
*Bss*HII	•	65/135/205	250	180/120	45/160	nd	360	60	>700	*Bss*HII
*Cla*I	nd	•	65/135/205	50	nd	65/135/205	nd	55	65/135/205	*Cla*I
*Eag*I	270	90	•	70	45/160	250	nd	nd	250	*Eag*I
*Fsp*I	nd	55/80	105/130/200	•	25	nd	160	nd	180/220	*Fsp*I
*Mlu*I	280	nd	270	105/130/230	•	45/160	45/160	nd	45/160	*Mlu*I
*Not*I	nd	90	720	nd	280	•	nd	nd	>700	*Not*I
*Pvu*I	180	nd	340	nd	180	nd	•	nd	360	*Pvu*I
*Sma*I	245	90	nd	nd	nd	nd	nd	•	nd	*Sma*I
*Sst*II	160	40	330	50/95/140	160	330	nd	nd	•	*Sst*II
Aps-11	*Bss*HII	*Cla*I	*Eag*I	*Fsp*I	*Mlu*I	*Not*I	*Pvu*I	*Sma*I	*Sst*II	

nd: fragment sizes were not determined

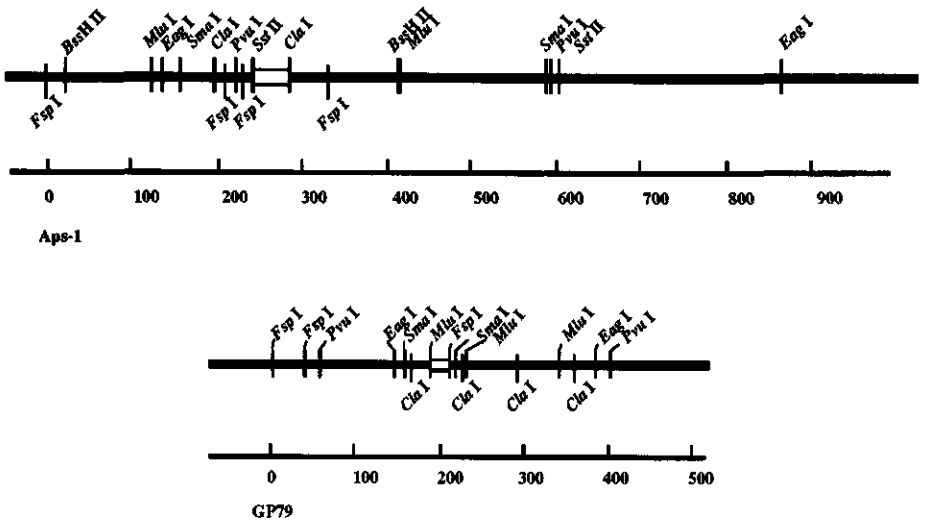
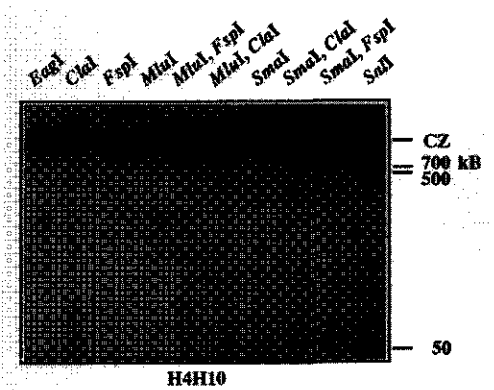


Figure 4. Consensus long range restriction maps of the *Mi* flanking markers *Aps-1* and GP79 for the nematode resistant cultivar VFNT Cherry. The regions where *Aps-1* and GP79 are located are indicated by an open box.

were produced at limiting amounts of enzyme. This allowed the *Cla*I, *Eag*I, *Fsp*I, *Mlu*I, *Sma*I and *Pvu*I sites around both markers to be surveyed for over 1000 kb of DNA surrounding the region corresponding to these probes. By combining these data with those obtained from the single and double digestions (fig 2 and table 3) we were able to construct restriction maps of *Aps-1* and GP79 which encompass 870 kb and 400 kb respectively (Fig 4). The maps do not show any overlap, indicating a distance between both markers of at least 550 kb.

In a search for additional molecular markers tightly linked to *Mi*, recently two other genomic

Figure 5. Southern blot hybridization analysis of restricted DNA from cv. VFNT Cherry using as a probe H4H10 that maps genetically to the same locus as *Aps-1*. Agarose-embedded protoplasts of VFNT Cherry were lysed and the released DNA was digested with the restriction enzymes as indicated. Restriction fragments were resolved on A CHEF gel (electrophoresis conditions: 20 h ramp with pulses from 35 to 70 s at 6.0 V/cm). After staining, fragments were blotted and hybridized to a probe corresponding to the H4H10 locus. Autoradiographs were exposed for two days using one intensifying screen. DNA fragmet sizes (kb) are shown at the side. Note that, except for *Sal*I, all fragments hybridizing appear in the compression zone which, under the electrophoresis conditions applied, contains fragments larger in size than 1500 kb.



DNA markers, H3F8 and H4H10 (see fig.1) were identified [21], that could not be resolved genetically from the *Aps-1* locus. To determine whether these markers could be mapped physically relative to *Aps-1*, pulsed-field gels of VFNT Cherry DNA digested with a variety of restriction enzymes were subjected to Southern analysis. The fragments hybridizing to H3F8 and H4H10 are listed in table 1. Again no common fragments were detected, neither for these markers nor for these markers in regard to *Aps-1* and GP79, respectively. Remarkably, all the fragments detected with probes H3F8 and H4H10 were much larger than those corresponding to *Aps-1* and GP 79 (Figure 5, table 1).

Discussion

Using several rare-cutting restriction enzymes long-range restriction maps have been constructed of the regions surrounding two markers (*Aps-1* and GP79) flanking the nematode resistance gene *Mi* in the cultivar VFNT Cherry. The maps encompass 870 kb for *Aps-1* and ca. 400 kb for GP79. Since no overlap between both maps was detected it is concluded that the loci *Aps-1* and GP79 are at least 500 kb apart. Taking into account that for none of the enzymes, neither in the single nor in the partial digests, were common fragments observed, the physical distances must be much larger.

Genetically the distance between *Aps-1* and GP79 has been found to vary from 0.3 cM to ca. 2.0 cM, depending on the genetic background [21, 28]. These genetic distances would, on average, correspond to physical distances ranging from 250 kb to 1500 kb, assuming a DNA content of $9.5 \cdot 10^5$ kb [3] and a genetic size of 1276 cM [34]. As suggested previously [21, 28] the small genetic distance of 0.3 cM is likely to be an underestimate as a result of suppressed recombination around *Mi* in crosses involving resistant varieties which carry introgressed chromosomal segments derived from *L. peruvianum*. This hypothesis is now supported by the present finding showing that GP79 and *Aps-1* are a least 550 kb and most likely much farther apart. Besides the contribution to suppression by the chromosomal variation differentiating susceptible and resistant varieties (*L. peruvianum* vs *L. esculentum* DNA [30]), the location of the *Aps-1* and GP79 loci in the heterochromatic region around the centromere (see fig. 1) may exert an additional negative effect on recombination [18, 22, 34]. In mapping a large set of molecular markers tightly linked to *Mi*, we have found that the majority cluster at the *Aps-1* locus [21]. From the high density molecular linkage maps, constructed by Tanksley and coworkers [34], it is apparent that clustering of markers is typical for centromeric areas of all tomato chromosomes. Our physical mapping data for the markers H3F8, H4H10 and *Aps-1*, which all map genetically to the same locus, but show completely different long range restriction maps with no fragments in common, are in agreement with those of Ganal *et al* [15] regarding the *Tm2-a* region of chromosome 9 and support the idea of centromeric heterochromatin-induced suppression of recombination. The long range restriction maps presented here provide a first step towards the molecular

dissection of a plant centromeric region with high marker density. Work is now in progress to connect the restriction maps of GP79 and *Aps-1* through a chromosomal walk in the region contained between these two markers.

In both vertebrates [6, 7, 16] and higher plants [2] many housekeeping genes are marked at their 5'-end by a CpG island, a relatively short region (0.5 - 2 kb) of the genome that is not depleted of CpG dinucleotides and is generally unmethylated. The presence of a CpG island in vertebrates is often indicated by the clustering of restriction sites for a variety of rare cutting enzymes [5, 25]. Unlike the CpG islands in vertebrates, which are generally G+C-rich and contain clusters of cleavage sites for methylation sensitive enzymes with G+C-rich recognition sequences (such as *Bss*HII, *Eag*I, *Not*I and *Sst*II [5]), the dicotyledonous CpG islands tend to be rather A+T rich [17]. These islands can therefore be expected to contain sites for methylation sensitive enzymes containing A's and T's in their recognition site. As to the *Aps-1* gene of tomato, which is constitutively expressed [14], we have not been able to detect a CpG island on the basis of the above mentioned criteria (fig 4). Whether CpG islands are too small to be detected through the represent approach or are methylated, remains to be investigated.

Methylation, at least in part, of CpG dinucleotides, was apparent from the multiple restriction fragments often generated by *Cla*I, *Fsp*I, *Mlu*I and *Sal*I, which each contain only one CpG and A and T in their recognition sites. In vertebrate DNAs sites of cleavage for such enzymes are outwith islands and subject to variable degrees of methylation [8]. Whatever their location in tomato DNA, the partial methylated sites were helpfull in constructing extended restriction maps around both *Aps-1* and GP79.

Acknowledgments.

We thank Dr. C. Gebhardt for the GP79 clone and Ruud Verkerk for the RFLP markers H3F8 and H4H10. This research was supported in part by a grant (No. CI1-CT91-0932) from the Commission of the European Communities.

References

1. Aarts JMMJG, Hontelez JGJ, Fischer P, Verkerk R, van Kammen A, Zabel P (1991) Acid phosphatase-1[1], a tightly linked molecular marker for root-knot nematode resistance in tomato: from protein to gene, using PCR and degenerate primers containing deoxyinosine. Plant Mol Biol 16: 647-661.
2. Antequera F, Bird AP (1988) Unmethylated CpG islands associated with genes in higher plant DNA. EMBO J 7: 2295-2299.
3. Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Rept 9: 208-218.
4. Arondel V, Lemieux B, Hwang I, Gibson S, Goodman H, Somerville CR (1992) Map-based cloning of a gene controlling omega-3 fatty acid desaturation in *Arabidopsis*. Science 258: 1353-1355.

5. Bickmore WA, Bird AP (1993) The use of restriction endonucleases to detect and isolate genes from mammalian cells. Methods Enzymol 216: 224-244.
6. Bird AP (1980) DNA methylation and the frequency of CpG in animal DNA. Nucleic Acids Res 8: 1499-1504.
7. Bird AP (1986) CpG-rich islands and the function of DNA methylation. Nature 321: 209-213.
8. Bird AP (1989) Two classes of observed frequency for rare-cutter sites in CpG islands. Nucleic Acids Res 17: 9485.
9. Braham WS, Winstead NN (1957) Inheritance of resistance to root-knot nematodes in tomatoes. Proc Am Soc Hort Sci 69: 372-377.
10. Carle GF, Frank M, Olson MV (1986) Electrophoretic separation of large DNA molecules by periodic inversion of the electric field. Science 232: 65-68.
11. Chu G, Vollrath D, Davis RW (1986) Separation of large DNA molecules by contour-clamped homogeneous electric fields. Science 234: 1582-1585.
12. Erion JL, Ballo B, Fox TW, May LA (1992) Tomato acid phosphatase-1 gene from a nematode resistant cultivar. Plant Physiol. 98: 1535-1537.
13. Feinberg AP, Vogelstein B. (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 132: 6-10
14. Fobes (1980) Trisomic analysis of isozymic loci in tomato species: segregation and dosage effects. Biochem Gen 18: 401-421.
15. Ganal MW, Young ND, Tanksley SD (1989) Pulsed field gel electrophoresis and physical mapping of large DNA fragments in the *Tm2-a* region of chromosome 9 in tomato. Mol Gen Genet 215: 395-400.
16. Gardiner-Garden M, Frommer M (1987) CpG islands in vertebrate genomes. J Mol Biol 196: 261-282.
17. Gardiner-Garden M, Frommer M (1992) Significant CpG-rich regions in angiosperm genes. J Mol Evol 34: 231-245.
18. Gaudet AM, Fitzgerald-Hayes M (1990) The function of centromeres in chromosome segregation. In: Strauss PR, Wilson SH (eds), The Eukaryotic Nucleus: Molecular Biochemistry and Macromolecular Assemblies. Telford Press, Caldwell, New Jersey pp 845-881.
19. Gebhardt C, Ritter E, Barone A, Debener T, Walkemeier B, Schachtschabel U, Kaufman H, Thompson RD, Bonierbale MW, Ganal MW, Tanksley SD, Salamini F (1991) RFLP maps of potato and their alignment with the homoeologous tomato genome. Theor Appl Genet 83: 49-57.
20. Giraudat J, Hauge BM, Valon C, Smalle J, Parcy F, Goodman HM (1992) Isolation of the Arabidopsis *ABI3* gene by positional cloning. The Plant Cell 4: 1251-1261.
21. Ho JY, Weide R, van Wordragen MF, Lambert K, Koornneef M, Zabel P, Williamson VM (1992) The root-knot nematode resistance gene (*Mi*) in tomato: construction of a molecular linkage map and identification of dominant cDNA markers in resistant genotypes. The Plant J 2: 971-982.
22. Khush GS, Rick CM (1967) Studies on the linkage map of chromosome 4 of the tomato and on the transmission of induced deficiencies. Genetica 38: 74-94.
23. Klein-Lankhorst R, Rietveld P, Machiels B, Verkerk R, Weide R, Gebhardt C, Koornneef M, Zabel P (1991) RFLP markers linked to the root-knot nematode resistance gene *Mi* in tomato. Theor Appl Genet 81: 661-667.
24. Klein-Lankhorst RM, Vermunt A, Weide R, Liharska T, Zabel P (1991) Isolation of molecular markers for tomato (*L. esculentum*) using random amplified polymorphic DNA (RAPD). Theor Appl Genet 83: 108-114.
25. Larsen FL, Gundersen G, Prydz H (1992) Choice of enzymes for mapping based on CpG islands in the human genome. GATA 9: 80-85
26. Maniatis T, Fritsch EF, Sambrook J (1983) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
27. Martin GB, Ganal MW, Tanksley SD (1992) Construction of a yeast artificial chromosome library of tomato and identification of cloned segments linked to two disease resistance loci. Mol Gen Genet 233: 25-32.
28. Messenguer R, Ganal M, de Vicente MC, Young ND, Bolkan H, Tanksley SD (1991) High resolution RFLP map around the root-knot nematode resistance gene (*Mi*) in tomato. Theor Appl Genet 82: 529-536.
29. Rappold G, Lehrach H (1988) A long range restriction map of the pseudoautosomal region by partial digest PFGE analysis from the telomere. Nucleic Acids Res. 16: 5361-5377.
30. Rick CM (1969) Controlled introgression of chromosomes of *Solanum pennellii* into *Lycopersicon esculentum*: segregation and recombination. Genetics 62: 753-768.
31. Rommens JM, Ianuzzi MC, Kerem B, Drumm ML, Melmer G, Dean M, Rozmahel R, Cole JL, Kennedy D, Hidaka N, Zsiga M, Buchwald M, Riordan JR, Tsui LC, Collins FS (1989) Identification of the cystic fibrosis gene: chromosome walking and jumping. Science 245: 1059-1065.
32. Sasser JN (1980) Root-knot nematodes: a global menace to crop production. Plant Dis 64: 36-41.

33. Tanaka H, Hoshi J, Nakata K, Chung K, Ito T, Ohta A, Takagi M (1992) Partial sequence of acid phosphatase-1^1 gene (*aps-1*^{1}) linked to nematode resistance gene (*Mi*) of tomato. Biosci Biotech Biochem 56: 583-587.
34. Tanksley SD, Ganal MW, Prince JP, de Vincente MC, Bonierbale MW, Broun P, Fulton TM, Giovanonni JJ, Grandillo S, Martin GB, Messeguer R, Miller JC, Miller L, Paterson AH, Pineda O, Röder MS, Wing RA, Wu W, Young ND (1992) High density molecular linkage maps of the tomato and potato genomes. Genetics 132: 1141-1160.
35. Van Daelen RAJ, Jonkers JJ, Zabel P (1989) Preparation of megabase-sized tomato DNA and separation of large restriction fragments by field inversion gel electrophoresis (FIGE). Plant Mol Biol 12: 341-352.
36. Van Daelen RAJJ, Zabel P (1991) Preparation of high molecular weight plant DNA and analysis by pulsed-field gel electrophoresis. In: Gelvin SB, Schilperoort RA (Eds) Plant Mol Biol Manual, Kluwer, Dordrecht, **A15**: 1-25.
37. Wallace MR, Marchuck DA, Andersen LB, Letcher R, Odeh HM, Saulino AM, Fountain JW, Brereton A, Nicholson J, Mitchell AL, Brownstein BH, Collins FS (1990) Type 1 neurofobromatosis gene: identification of a large transcript disrupted in three NF1 patients Science 249: 181-186.
38. Wicking C, Williamson B (1991) From linked marker to gene. Trends Genet 7: 288-293.
39. Williamson VM, Colwell G (1991) Acid phosphatase-1 from nematode resistant tomato; isolation and characterization of its gene. Plant Physiol 97: 139-146.

CHAPTER 4

YAC CLONING

4.1 INTRODUCTION

Cloning of very long stretches of DNA as large artificial chromosomes in yeast (YACs [35]) has bridged the gap between molecular biology, cytogenetics and genetics [7]. The ability of cloning large fragments has culminated in the construction of contigs encompassing the entire human chromosomes 21q [47] and Y [76], an achievement that would not have been possible with any other cloning vehicle.

Based upon the work of, among others, Murray and Szostack [147], who made long linear constructs which were able to function as chromosomes in yeast, the YAC cloning system was at first developed by David Burke in Olsons lab [35], and than has been optimized in several laboratories to such an extent that fragments of up to 2000 kb can now be cloned. Apart from the seemingly unlimited large sizes of fragments that can be cloned, the YAC cloning technique has some other features that makes it a very valuable system. Firstly, it is a eukaryotic system in which more sequences are expected to be clonable than in *E. coli* or any other prokaryotic system. For several human YAC libraries it has indeed been shown that the cloning of at least euchromatic regions is completely random [47, 76, 173], covers nearly all sequences in stable clones, and that such libraries are a genuine representation of genomic sequences [1, 30, 50, 75, 88, 124, 161, 181, 196]. The same holds true for *Drosophila* [2, 59, 81, 138] of which currently more than 75 % of the euchromatin has been cloned and mapped in 965 contigs each measuring up to 2.5 Mbp in size [94]. Secondly, homologous recombination happens in yeast cells efficiently and this allows the manipulation and alteration of YAC sequences in any conceivable manner, as will be discussed in paragraph 4.7. Thirdly, the wide range of organisms for which YAC libraries have been constructed clearly reflects the versability of this cloning system.

There are now many human genomic libraries [1, 35, 46, 105, 119, 125, 136, 191] as well as genomic libraries of various other species including: the fly *Drosophila* [2, 59, 81, 138]; the mouse *Mus musculus* [119, 166]; the nematode *Caenorhabditis elegans* [266], of which extensive tracts of DNA, that could not be cloned in cosmids, were succesfully cloned in YACs [55]; *Schizosaccharomyces pombe* for which a contig encompassing the entire genome has been constructed [128]; the amoeba *Myxococcus xanthus* for which a contig encompassing the entire genome of approximately 10 Mbp has been constructed [113]; the malaria parasite *Plasmodium falciparum* [62, 117, 192], the DNA of which was found to be very difficult to clone in *E. coli* due to its high A/T content; and the slime mold *Dictyostelium discoideum* [112]. Last, but certainly not least, several plant YAC libraries are currently available. They are listed in Table 1 with some characteristics. There are as yet no reports of extensive studies concerning the reliability of cloning of plant sequences in YACs, but results so far look promising. In screening the *Arabidopsis* libraries, for example, it was found that

Table 1. Characteristics of the various constructed plant YAC libraries.

Species	number of clones	average insert size (kb)	genome equivalents	proportion chloroplast YACs (%)	vector	reference
Arabidopsis thaliana	2185	150	4.4	6	pYAC3	[198]
Arabidopsis thaliana	2400	150	5	4	pYAC4	[91]
Arabidopsis thaliana	10-16,000	75	±10	57	pYAC4	[94]
Beta vulgaris	15,000	140	2	2-3	pYAC4	[73]
Daucus carota	16,000	150	5	-	pYAC4	[94]
Hordeum vulgare	18,000	160	0.5	1	pYAC4	[67]
Lycopersicon esculentum	22,000	140	3	10	pYAC4	[133]
Oryza sativa	8000	350	6.4	-	pYAC4	[15]
Oryza sativa	4000	350	3.2	-	pYAC5	[15]
Zea mays	79,000	145	3	4-5	pYAC4	[70]

for every RFLP marker used, a corresponding clone was detected indicating that most regions of the genome were represented in the libraries. For *Arabidopsis* it is estimated that at present 30 % of the genome is covered in mapped contigs [104].

Surprisingly, the introduction of very large pieces of DNA appears to have little effect on the yeast cells, even though the foreign DNA may amount to more than 6% of the total DNA content. Introduction of YACs may sometimes lead to an increased spontaneous reversion of genes involved in the adenine biosynthesis pathway [13], but only few data are available on this or possible other effects.

The introduction of megabase techniques, like YAC-cloning and Pulsed Field Gel Electrophoresis (PFGE), has enormously advanced the identification of genes via positional cloning [197]. It should be feasible now to clone any gene of any organism provided the gene can be genetically mapped and identified by its phenotype. Several human genes for which a phenotype and linked molecular markers were available have now been cloned, including the gene of Cystic Fibrosis [165], the gene of Myotonic Dystrophy [99], the Fragile X-locus [64, 97], the Neurofibromatosis type 1 gene [197], the gene involved in Myeloid Lymphoid Leukemia [169, 200] and the APC-gene involved in familial adenomatous polyposis [95, 109].

Recently, the isolation of some plant genes has been achieved via positional cloning, including the *ABI3* gene of *Arabidopsis* [85], a gene from *Brassica napus* controlling omega-3 fatty acid desaturation in *Arabidopsis* [14], and the tomato resistance gene *Pto* which confers

resistance against strains of the bacterium *Pseudomonas syringae* [132 . Actually, the *ABI3* gene has been identified by a cosmid walk, but in cloning the tomato *Pto*-gene YACs played a major role, and they undoubtedly will do in cloning genes from other plants with large genomes.

The YAC system has for the first time allowed the integral cloning of large genes of which the coding sequences with intervening intron sequences extend over large stretches of DNA. Examples of such large genes, of which the sizes range from 40 to 250 kb, that have now been cloned, include the Cystic Fibrosis gene [8], the gene for human factor VIII [1], and human factor IX [30], the glucose-6-phosphate dehydrogenase gene [43], the gene of cineamide ribonucleotide transformylase (GART [184]), and the gene of the leukocyte common antigen CD45 [75]. Furthermore, by homologous recombination of two or more YACs in yeast cells, continuous tracts of very large genes have been obtained for BCL2 (230 kb [181]), cystic fibrosis (250 kb, [88]) and even the Duchenne muscular dystrophy gene which measures 2.4 Mbp [63].

The YAC system also allows the cloning of genes alongside large tracts of surrounding DNA, containing regulatory sequences, in their native configuration. This will greatly facilitate studies on the mode of action and regulation of these genes [78]. The cloning of DNA surrounding the coding region of the human β-globin gene has enabled to show that regulatory elements 50 kb upstream and 20 kb downstream of the gene itself are essential for the natural level of expression of β-globin [93, 189].

Making use of YACs, the hitherto unknown physical organisation of some large gene families has been revealed. These include the the human β-globin gene family of 60 kbp [78], complete human ribosomal DNA repeat units with a total size of 44 kbp [115], and the MHC complex which is over 4 Mbp in size [82, 111, 161, 163]. Such physical maps are important tools in revealing the spatial organization and functional analysis of gene complexes.

Finally, it should be noted that the YAC cloning system has facilitated studies on the functioning of yeast centromeres during meiosis. These studies were possible because aneuploidy for a human YAC has no effect on spore viability and thus allows the scoring of every kind of meiotic missegregation [100, 176].

Important specific applications of YAC cloning are the isolation of human telomeres with surrounding sequences (section 4.9.1), construction of chromosome specific libraries (section 4.4), and transfer of entire YACs to mammalian cells (section 4.8), which will be discussed separately.

In the following, the construction of YAC libraries, some characteristics of YACs and ways to characterize them, manipulation of YACs *in vivo*, and the identification of coding sequences on YACs, will be discussed. Since plant YAC libraries have become available relatively recently, they have not yet been characterized in depth and thus most studies to

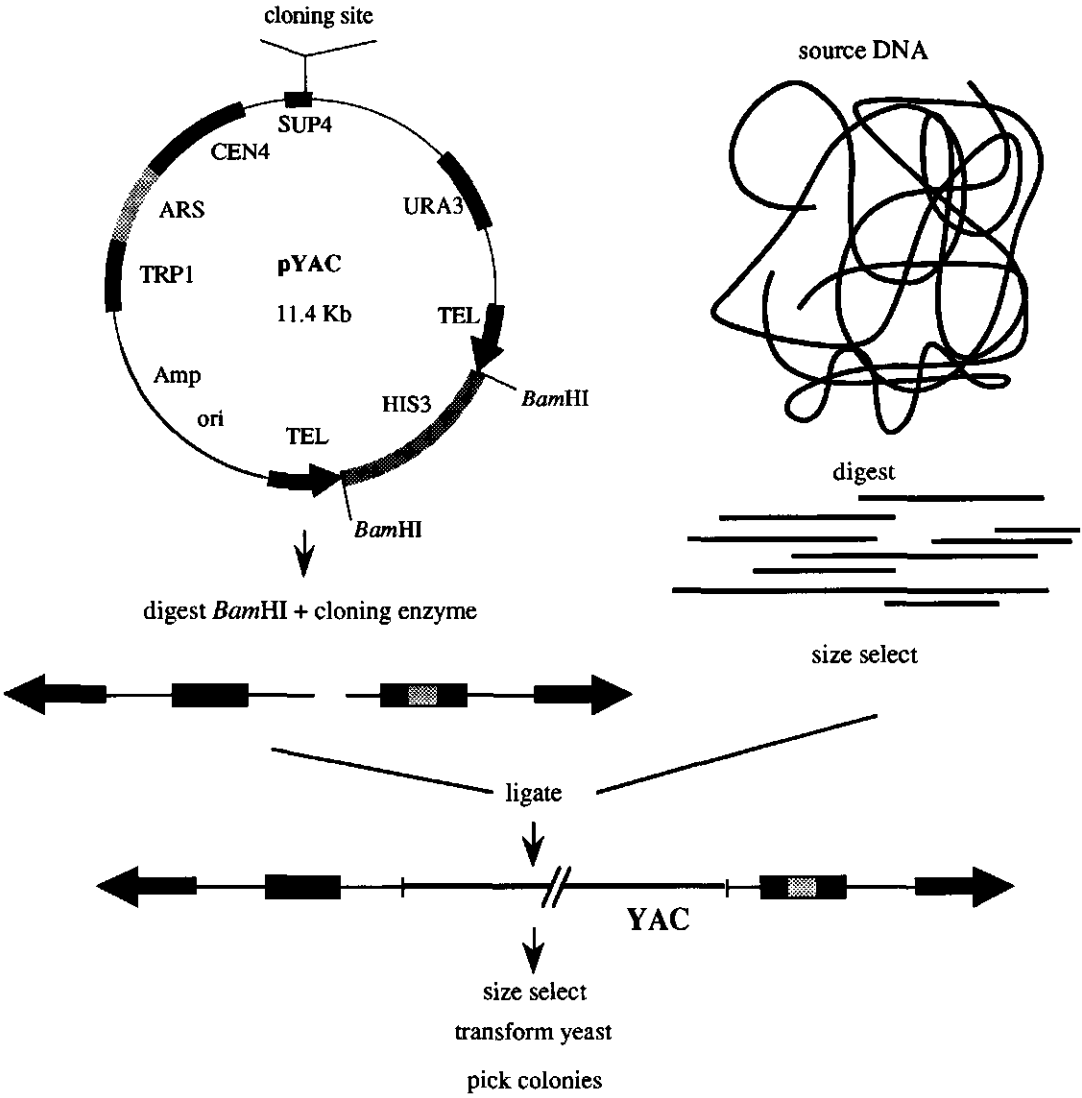


Figure 1. Schematic representation of the construction of a yeast artificial chromosome. For explanation see text.

which we shall refer concern human YACs. However, many of the results and methods appear universally applicable and will similarly be of interest for plant YAC studies.

4.2 CONSTRUCTION OF YEAST ARTIFICIAL CHROMOSOMES.

Figure 1 shows the different steps involved in the construction of YACs. In brief, the procedure is as follows. Source DNA is digested to generate large restriction fragments which are ligated to the YAC-vector DNA. This vector consists of two segments which together contain all the elements necessary to function as a stable chromosome in yeast, namely a centromere, an autonomously replicating sequence (ARS) and two telomeres, one on each arm. For selection of YACs within the yeast cells, there are two yeast selectable markers present (*ura* and *trp*), one on each vector arm. The vectors also contain an origin of replication and an ampicillin resistance gene to allow manipulation and amplification in *E. coli*. The *HIS3* fragment present in the vector is a throwaway fragment, not functional during cloning, and is merely present as an artifact of the telomere cloning. The cloning site is located within a suppressor gene which facilitates the discrimination between transformants containing only vector sequences and those containing YACs (see section 4.2.3). To obtain a functional linear chromosome, the two vector arms are ligated to the insert DNA which is subsequently transferred into yeast (*Saccharomyces cerevisiae*). The successive steps involved will each be discussed in more detail in the following.

4.2.1 Source DNA

DNA to be used for YAC cloning has to be very large, preferably many megabases in size (see also chapter 6). This quality is best achieved from agarose embedded living cells followed by *in situ* lysis of the embedded cells. In the case of plants, it is best to start from protoplasts embedded in agarose [194]. The isolation of high molecular weight plant DNA in large quantities is anyhow much more cumbersome than the isolation of mammalian DNA since the cellwall has to be removed prior to embedding the plant cells. This leads to some degradation of DNA which is reflected by the, on average, smaller size of plant YACs. A further interfering factor is the presence of many copies of chloroplast DNA in plant cells. This may lead to a considerable percentage of chloroplast YACs in the libraries (see table 1).

DNA may also be isolated in solution, via sucrose gradient centrifugation, as has been applied in the construction of some human [1, 35, 36], *Drosophila* [81], and plant [94, 198] YAC libraries. This has the advantage that DNA can be obtained in higher concentrations as compared to DNA from agarose embedded cells, which is advantageous for the transformation of yeast. However, DNA in liquid medium is very fragile [1, 36] and extremely sensitive to shearing, easily resulting in small YACs. Only when DNA can readily be prepared in large amounts, sucrose gradient centrifugation may be worthwhile, since, in that case, batches of DNA containing small fragments can be discarded.

For the construction of a library covering an entire genome, one needs a set of large overlapping fragments. This can be achieved by partial digestion with a frequently cutting restriction enzyme. The enzyme *Eco*RI is very often used for this purpose. Partial digests are obtained by incubating with limiting amounts of enzyme [36, 105], and for brief periods of time [120], using limiting amounts of Mg^{2+} [4], or in a competition reaction with *Eco*RI methylase [119, 142]. Alternatively, randomly sheared DNA can be used to obtain overlapping fragments [1, 94, 198]. The latter then requires filling in of ends and ligation of adaptors and this additional handling of the DNA often causes a degradation of the DNA. The claimed advantage of using sheared rather than restricted DNA, is that it provides a more random representation of sequences, since restriction sites may not be completely random distributed over the genome .

4.2.2 The YAC vectors

The original pYAC2 vector, described by Burke [35], is a pBR322 derivative and contains all the necessary elements for a functional chromosome: a centromere (CEN4 derived from yeast chromosome 4 [129]), an Autonomously Replicating Sequence (ARS [186]) derived from yeast, and two telomere sequences derived from *Tetrahymena* [145]. The *Tetrahymena* telomere sequences are gradually replaced by yeast telomeric sequences during replication of the YACs in the yeast cells. On each arm there is a yeast selectable marker: *URA3* for the left arm and *TRP1* for the right arm. The cloning site (a *Sma*I site in pYAC2 or *Eco*RI in pYAC4, see fig 1) resides within the *SUP4* gene (an ochre suppressing allele of a $tRNA^{TYR}$ gene) allowing selection of recombinant clones containing an insert.

For cloning purposes the telomeres are released by restriction digestion, with *Bam*HI, thereby removing the stuffer fragment that contains the *HIS3* gene which is located between the telomeres and does not participate in cloning. The vector is also cut at the cloning site, thereby disrupting the suppressor gene. If an insert is ligated into the cloning site no functional suppressor can be formed and yeast colonies with recombinant YACs will become red (see section 4.2.3). After digestion the arms are dephosphorylated to avoid religation, which otherwise is a serious problem, since vector is always added in a large molar excess (up to 100 times).

Many modifications of the vector have been constructed. The adaptations initially included the introduction of additional cloning sites: pYAC3 with a *Sna*BI site generating blunt ends, pYAC4 with an *Eco*RI site, and pYAC55, previously named pYAC5, with a *Not*I site for cloning large *Not*I restriction fragments, all developed by Burke *et al* [35]. Several cloning sites for rarely cutting enzymes were introduced into pYAC4 to make pYAC-RC [130]. Many modifications were made to facilitate the recovery of sequences of the ends of the YAC inserts. This has resulted in pYAC41 and pYAC45 containing two T3, and a T3 and T7

promotor respectively, to prepare riboprobes from the ends of the insert [91]. The two vector system pJS has the chromosomal functions for both arms located on two different plasmids (pJS97 and pJS98 representing respectively the centric and acentric arm [178]), both of which contain *E. coli* selectable markers, allowing the subcloning of the ends as plasmids in *E. coli*, and T7 promotors.

Smith *et al* [183] constructed a vector that allows copy number control of the YAC in the yeast cells. In this vector (pCGS966) the CEN4 sequence is placed after a *gal* promotor and is not functional, due to transcription, when yeast cells are grown on a medium containing galactose. Selection for the heterologous thymidine gene, also present on the pCGS966 vector, in the abscence of centromere function, results in an increase of copy number of the YAC.

Some vectors were specifically constructed for the isolation of human telomeres: pYAC-*neo-not* [57] and pTYAC [163] both of which contain only one telomere (see section 9.1). A further modification of the cloning vectors involved the introduction of the neomycine gene allowing for selection for G418 resistance in mammalian cells (pYACneo [191], pYAC4-neo [54] and pYAC151 [188]). Vectors which contain plant selectable marker genes, such as kanamycine, are currently being constructed [26].

4.2.3 Yeast strains used for YAC cloning

The yeast strain most widely applied for YAC cloning is AB1380, that was originally used by Burke *et al* [35], and has the following genotype: Mat a, ψ^+, *ura3-52, trp1, ade2-1, can1-100, lys2-1, his5*. Apart from the two auxotrophic markers *ura3* and *trp1* (both amber mutations of orotidine-5'-P-decarboxylase and N-(5-phosphoribosyl)-anthranilate isomerase respectively) used for selection of the presence of introduced YACs, several other phenotypic markers are present. *Ade2-1*, is an ochre mutation in the gene for aminoimidazole carboxylase, which leads to the accumulation of a red pigment (p-ribosylaminoimidazole; AIR), when grown under limiting amounts of adenine. In the presence of a suppressor, like SUP4 encoded by the vector, no accumulation of AIR occurs and the colonies appear white. If the suppressor is interrupted by an introduced sequence the colonies will appear red thus allowing the selection of clones containing YACs. The *can1-100*, an ochre mutation in the gene for arginine permease, *lys2-1* an ochre mutation in 2-amino-adipate-reductase and *his5* genes provide additional selectable alleles.

The strains YPH252 and YPH274 [180], used in combination with the pJS vectors, contain, apart from *ura3* and *trp1*, the *his3* and *leu2* alleles. The strains 3a [42], 814 series [148], and LIV2 [68] all harbor the *rad52* mutation which renders them deficient in homologous recombination, thereby tolerating YACs which are unstable in AB1380 (see section 5.2).

4.2.4 Ligation

The ligation of insert DNA to the vector arms is very straightforward. In most cases DNA is embedded in agarose and, before adding the vector arms, the agarose has to be melted to allow mixing of the components. During this step DNA is very sensitive to shearing and special precautions must be taken to prevent DNA from breaking (see section 4.2.5). Once both components are mixed the ligation can proceed in solidified agarose.

4.2.5 Size selection

When DNA with ligated vector arms is directly used for transformation, YAC clones of rather small size, on average 150 kb or less, are obtained [10, 19, 187]. To get larger YACs a size selection of the DNA is necessary [3, 9, 10] and then YACs with an average insert size of 300 to 400 kb can be obtained [3, 9, 48, 119]. For average insert sizes up to 1000 kb, two size selections, both before and after ligation, have been found necessary [3, 9, 47, 48, 105, 119].

The size selections are performed by pulsed field gel electrophoresis (PFGE) [3, 10, 119] or by sucrose gradient centrifugation [36, 105, 198]. In the latter approach, the DNA is continuously in solution and liable to degradation. Consequently, the obtained YACs are, on average, smaller [105, 166], even after double size selection [198]. Moreover, due to the large volumes, concentration steps have to be introduced leading to an additional loss of DNA [36]. The best results were obtained with size selections using PFGE in which DNA is isolated from the limiting mobility or compression zone (CZ). As size selection usually leads to a loss of DNA often only one size selection is performed. Notably, in the construction of plant YAC libraries this is often done.

The large proportion of small YAC clones in the first constructed libraries was found to be due to a bias in transformation efficiency for smaller molecules and the fact that large molecules are more sensitive to shearing. Connely *et al* [53] showed that a 330 kb molecule has a tranformation efficiency 37 times less than a 110 kb molecule. Similar results were obtained by Albertsen *et al* [3] who showed that a 600 kb YAC has a eight times, and a 490 kb YAC a five times, lower transformation efficiency than a 180 kb YAC. According to Connely *et al* [53] this size bias could be decreased by using the polyamines, spermidine trihydrochloride and spermine tetrahydrochloride, in all solutions with DNA, including the transformation steps. Larin *et al* [119] and Lee *et al* [120] showed that DNA was very sensitive to shearing during and after melting of the agarose plugs and they observed that this often resulted in degradation of DNA. They found that both polyamines [119] or salt (125 mM NaCl [120]) in the solutions containing DNA prevented shearing of DNA. According to these authors inclusion of polyamines during transformation appeared not necessary and, should be

dissuaded as it resulted in a markedly decreased transformation efficiency [3, 119]. The latter effect, for that matter, was not noticed by Connely *et al* [53].

4.2.6 Transformation

DNA to be used for transformation of yeast cells must be in solution to allow efficient uptake into the cells. This is already the case when sucrose gradient centrifugation is used for preparing DNA and it will be sufficient to concentrate the DNA solution prior to transformation. If, on the other hand, PFGE is used for size selection the DNA has to be released from the agarose matrix. Initially, the agarose was simply melted [53, 91, 136] and, to avoid setting of the agarose, mixed and diluted with the spheroplast suspension for transformation. However, the presence of agarose appears to lower the transformation efficiency considerably and therefore removal of agarose by agarase is preferred [74].

The method most widely used for yeast transformation is the spheroplast procedure as developed by Hinnen *et al* [101] and modified by Burgers and Percival [34]. With that, spheroplast are prepared by incubating yeast cells, in mid- or late-log growth phase, with cell wall degrading enzymes such as lyticase, zymolyase, and novozyme. It is important that the amount of enzymes used results in 80 - 90 % of the cells becoming spheroplasts within 15 to 20 minutes [3, 34, 36]. Spheroplasts are then mixed with the DNA sample, incubated for 10 minutes whereafter uptake into the cells is mediated by adding poly-ethylene glycol (PEG). This is followed by a recovery step in which cells are kept in a rich medium for 30 minutes, after which the spheroplasts are plated out on a selective medium without uracil.

Some authors recommend the use of polyamines during transformation for protection of the DNA, as mentioned in the previous paragraph. On the other hand it is reported that the presence of polyamines during transformation results in a decrease in efficiency by a factor of 10 to 50 [3, 53, 119]. The inclusion of lipofectin during transformation was found to enhance the uptake of the plasmids YCp50 and pYAC4 into yeast and also the uptake of large linear fragments of DNA into *Schizosaccharomyces pombe* [5]. It is as yet not clear whether the transformation of yeast using YAC-DNA is also improved by using lipofectin.

After recovery, the spheroplasts are plated out on selective medium lacking uracil. Since the spheroplasts are very fragile they are plated out embedded in agarose. This results in colonies that are enclosed within a thin layer of agarose from which they must be removed prior to storage and screening. The picking of colonies from the agarose matrix is very labourious and so, not surprisingly, methods have been developed to pick up, and transfer, all colonies in one step by using multipin transfer devices containing 3000 [1] or 40,000 [119, 140, 142] steel pins.

To select the clones containing YACs with both vector arms the colonies are preferably first cultured on double selective medium, lacking tryptophane and uracil, prior to storage in

glycerol. But some authors omit this selection step and have grown colonies directly in YPD in the wells of microtiter dishes [166], without the loss of YACs.

It is worth noting that double selection cannot be applied on transformation plates as the deleted version of the promotor of the *trp1* gene in pYAC4 works inefficiently and does not allow immediate double selection [Burke, pers. comm.]. The cells have first to be grown on regeneration plates and then a double selection can be applied.

Several alternatives have been developed to avoid the picking of individual colonies. The simplest method is to scrape of the topagar layer, wash out the cells and, after amplification by growing on double selective liquid medium, to store the whole mixture [166]. However, due to the, sometimes large, differences in colony size and there will be a chance that a subset of clones becomes over- or underrepresented in the amplified mixture of clones. The spheroplasts can also be plated out in alginate rather than agarose as was done by Traver *et al* [191]. In this case the transformed cells are mixed with an alginate containing medium and plated out on top of a membrane placed on a medium containing Ca^{2+} which brings about the setting of the alginate. When colonies have developed, the filter with the alginate layer is placed on a medium containing EDTA after which the alginate liquifies and the colonies can be replica plated. It should be added that some authors have reported that the transformation efficiency decreases, if spheroplasts are plated out in alginate [1, 36, 127].

4.2.7 Storage of the library

Colonies that grow on double selective medium and show a red pigmentation, contain YACs that have both vector arms and are qualified for storage. The most widely used way is to store each clone separately in a well of a microtiter dish in rich medium (YPD). The rich medium is necessary to increase recovery of the cells after freezing. Glycerol is added at a final concentration of 15-20 % (w/v) and the microtiter plates are stored at -80°C. Sometimes 7% DMSO is applied in stead of glycerol [120].

The YAC libraries can also be stored as replica filters which are placed on 3MM paper soaked in a YPD/glycerol mixture [181, 220].

4.3 SCREENING OF YAC LIBRARIES

Several methods have been used for screening YAC libraries, two of which are most widely applied: colony filter hybridization and PCR screening. Both methods have their specific merits and will be described below. It is noteworthy that proponents of both methods claim their method to be the fastest [118, 119].

4.3.1 High density colony filter hybridization

Screening of YAC libraries by colony filter hybridization involves growing the colonies either directly on a membrane, or on agar plates followed by transfer of the colonies to a membrane by a filter lift, after which the cells are lysed. The advantage of using filter lifts is that several replicas of the same plate can be made [56]. For efficient screening it is essential that many colonies are brought together on a small surface. This has been achieved by using robotics, as was described by Bentley *et al* [23] who applied the Biomek 1000 (Beckman instruments) to transfer 1536 clones (from 16 microtiter dishes) onto a membrane of 8 by 12 cm. Schmidt *et al* [174] obtained similar results by manually transferring colonies using a 96 pin replicator. Ross *et al* [167] managed to transfer automatically the even higher number of 20,000 clones, to filters measuring 22 x 22 cm. It is important to obtain an even growth of all colonies so that each clone will be represented by similar amounts of DNA. Picking false positive or missing false negative ones is then prevented as much as possible. Furthermore, an equal low background radiation level over the membrane is necessary to identify the position of a positive clone within the grid pattern. Such constant background radiation has sometimes been provided by adding a small amount of ^{35}S labelled vector DNA during hybridization [127].

Especially if high density colony filters are used, colony filter hybridization is fast and cost effective, since the right clone can be identified in one hybridization. Moreover, if desired the filters can be screened with complex probes by which YACs from a large number of loci, or a large chromosomal region, can be identified in one step [168].

4.3.2 PCR screening

Screening YAC libraries by PCR, requires a series of PCR's with appropriate oligonucleotide primers on DNA from pools of clones. If, in that way, the microtiter plate containing the YAC sought after has been identified, the last step, to identify the specific clone on the microtiter plate, can be achieved by colony filter hybridization, or, alternatively, by screening pools of clones from the 8 rows and 12 columns of the microtiter plate which will reveal the position of the positive clone with 20 PCR reactions. A variety of pooling strategies have been described [6, 51, 89, 114] and a theoretical study to identify the most efficient pooling strategy was performed by Barillot *et al* [18]. A simple procedure, for identifying a clone, based solely on PCR, is to pool all clones per plate and all clones from the 8 rows and 12 columns of the microtiter plates. If 10 plates are used only 30 reactions are required to identify the positive clone [47, 123]. The number of plates used should be kept small enough to minimize the chances of finding two positives within one pool but large enough to avoid screening a large number of pools.

PCR screening of libraries is mainly used for quickly identifying single copy sequences in pools of clones, and has the advantage of not using radioactivity. However, each screening requires sequence information, primer synthesis, and multiple screening rounds. For example, Chumakov *et al* [47] reported that per STS (sequence tagged site) marker, on average, 355 tests were necessary to identify the positive YAC clone from a 9.4 genome equivalent library.

4.3.3 Alternative screening procedures

Alternative procedures for screening YAC libraries include the preparation of chromosomal DNA from pools of clones. Chromosomal DNA from pools containing up to 384 clones, from 4 microtiter dishes, has been used for separation by pulsed field gel electrophoresis [80, 137]. Mendez *et al* [137] used in that way only three CHEF gels to separate 150 pools of 384 clones, representing five human genome equivalents. After identification of positive pools by hybridization of the Southern blots of the separated chromosomal DNA, individual clones were identified by colony filter hybridization. The advantage of this procedure is that no sequence information is required and that Southern blots can better be reused than colony filters. However, since two hybridization rounds are required to identify the YAC of interest, the procedure is less efficient than colony filter hybridization.

For organisms with small genomes chromosomal DNA from every individual YAC clone has been prepared and separated on CHEF gels. For instance, the entire genome of *Dictyostelium* (40 Mbp) was covered, with a five-fold redundancy, by 1016 clones for which three CHEF gels were sufficient to enable the identification of YACs sought after [112].

4.4 SPECIFIC YAC LIBRARIES

Several YAC libraries have been constructed that contain only part of the genome of the organism under study. These include chromosome specific libraries, made from human/rodent somatic cell hybrid lines or from flow sorted human chromosomes. The so-called rare-cutter libraries also contain only part of the genome but are easier to construct than chromosome specific libraries. The major advantage of such libraries is that they contain less clones and, therefore, they are easier to screen and allow to determine chromosome specific contigs more quickly.

4.4.1 Chromosome specific libraries from human/rodent somatic cell hybrids

Chromosome specific YAC libraries have been constructed from human/rodent somatic cell hybrid lines for the human chromosomes 17 [66], 21 [22, 136, 159], and the X-chromosome [120, 125, 149, 196]. Apart from providing a source of clones specific for one chromosome, or a part of it, YAC libraries prepared from DNA isolated from these somatic hybrid cell lines have the additional advantage of containing less chimeric YAC clones, as will be discussed in paragraph 4.5.2.

From the total libraries, the human chromosome specific YAC clones are distinguished from the non-human YACs by colonyfilter hybridization using as a probe the human specific *Alu*-repeat, that is present every 2 kb in human, but absent in rodent, DNA. The human YAC clones can then be separated from the rodent clones. An alternative approach to separate the human YACs from the rodent YAC clones was developed by Pavan *et al* [157] who used integrative selection vectors, or ISVs. The clones from the total human/rodent YAC library were pooled and the yeast cells were transformed with the ISV. The ISV specifically recombines with the *Alu* repeat sequence and, by the recombination, the selectable HIS3 marker is introduced into the human YAC. By screening for clones containing the HIS3 marker a 50 fold enrichment of human YACs has been obtained. A disadvantage of this method is that mutation activity proved to be introduced by the integration of the ISV. It was shown that in 18% of the YAC clones deletions had occurred.

The construction of these chromosome specific libraries is very labour intensive since the human YACs represent less than 2.5 % of all the generated YACs [168]. With the efficient screening protocols currently available, chromosome specific YACs may just as easily be isolated from total genomic libraries using chromosome specific probes in a colony filter hybridization screening [48, 141, 168]. Human chromosome specific probes can be generated via *Alu-Alu*-PCR on total DNA from somatic human/rodent cell hybrids, containing human chromosomes, as described by Monaco *et al* [144] and Chumakov *et al* [48]. The generated *Alu*-PCR products are used as a probe on high density colony filtes. Alternatively, DNA from phage libraries, made from flow sorted chromosomes, has been used as probe to identify chromosome specific YACs on colony filters [168].

4.4.2 Flow-sorted chromosome specific libraries

Chromosome specific libraries can also be made directly from DNA obtained from flow sorted chromosomes, as was shown for human chromosome 6 by McCormick *et al* [136]. Unfortunately, the library contained only 21 YACs, with an average insert size of 150 kb. A better result was recently obtained by McCormick *et al* [135], who succeeded in constructing rare-cutter libraries from flow sorted human chromosomes 16 and 21, containing 2150 clones

with an average size of 200 kb. Unfortunately, the technological requirements to isolate individual chromosomes hamper a widespread application [135].

4.4.3 Rare cutter libraries

Libraries made from large DNA fragments obtained after restriction digestion with rare-cutting enzymes are referred to as rare-cutter libraries [18, 19, 46, 59, 84, 118, 187, 188]. These libraries cover only part of the genome, since only a subset of all restriction fragments will be represented. The advantage is that smaller numbers of clones need to be generated and screened, as compared to a complete genomic library based on a partial *Eco*RI digest. An important application of rare-cutter libraries is found in the isolation of specific large restriction fragments from a region of interest on long range restriction maps [19].

4.5 CHARACTERIZATION AND ANALYSIS OF YACs

4.5.1. General characterization

The general characterization of YAC libraries involves the determination of (1) the average insert size, (2) the presence of genomic DNA and (3), for plant YACs, the proportion of chloroplast DNA containing YACs. Usually the initial screening does not involve the determination of YAC stability or the number of chimeric YACs, because this requires a rather large effort. In the course of the use of a library the latter characteristics are revealed.

Size determination is done by preparing total chromosomal DNA from a subset of clones (usually 100 - 200) and determine the size of the YACs on a pulsed field gel. The percentage of chloroplast DNA containing YACs is determined by examining a subset of clones by colony filter hybridization using a chloroplast DNA specific probe. For most plant YAC libraries between 5 and 15 % of the YACs appear to have chloroplast DNA inserts (see table 1). A quick way of screening for the presence of genomic DNA is to determine for a number of clones the presence of genomic repetitive sequences. The applied repeat should be present in a high frequency in the genome so that many of the clones from the library will contain it. Good examples are the *Alu* repeat for human DNA, that is present in every human YAC clone and, for tomato, the TGRII repeat, that occurs every 130 - 150 kb [79].

The detailed characterization of the YACs involves restriction mapping and fingerprinting, which will be delt with in the following paragraphs.

Restriction mapping. A complete restriction map allows a YAC to be placed on the genomic restriction map, to determine the overlap with other YACs, to identify CpG islands which may point to putative coding regions, as will discussed in section 4.6, and it may aid to determine whether the cloned sequence is chimeric (see section 4.5.2). The method to establish all restriction sites for a specific enzyme involves a partial digestion of the YAC DNA, separation of the resulting restriction fragments on a pulsed field gel, and hybridization of the Southern blot with a probe for one of the two YAC ends [35]. From the lengths of the fragments that hybridize with the end probe the location of the restriction sites can be determined. Differences between restriction maps of genomic DNA and DNA cloned in YACs are often found, since the methylation of the DNA cloned in yeast differs from that of the organisms from which the cloned DNA originates. In yeast, DNA is not methylated at all, whereas in plants and humans the cytosine residue in CpG's are often methylated [12]. Especially when enzymes are used that contain CpG's in their recognition sites and are methylation sensitive, which applies to nearly all rare-cutting enzymes, such differences will often be found. It also implies that on the DNA cloned in the YAC more restriction sites for a specific enzyme should be found than are already known for the genomic DNA.

Fingerprinting. Construction of YAC contigs that cover large chromosomal regions requires the systematic analysis of the overlap of many clones. A quick way to determine the overlap, which does not require an extensive characterization of the YACs, is to generate fingerprints of each clone and to compare the patterns. Fingerprints of the YACs are obtained by digesting with a restriction enzyme, followed by electrophoretic separations of the fragments, and Southern analysis using as a probe a repetitive sequence that occurs abundantly in the genome [39, 185]. Several authors have applied this approach to characterize human YACs using the high copy repeats *Alu* and L1 as probes [21, 50, 66, 77, 196]. For specific chromosomal regions the human low copy repeats LF1 and pTR5 have been used [201]. Cangiano *et al* [39] have applied this approach to order *Caenorhabditis* YACs into contigs.

4.5.2 Generation of end fragments

In order to perform a chromosomal walk one needs to have the ends of a YAC insert available to identify overlapping YACs. Apart from their application in chromosome walking, end probes are important means to determine whether a YAC is chimeric. A variety of methods has been developed to obtain end fragments, either cloned or through PCR. None of the methods is applicable to all YAC clones and, usually, several methods should be tried for generating end fragments.

End fragments can be obtained by subcloning the entire YAC into plasmid, phage lambda, or cosmid vectors followed by isolating the clones containing the ends [41, 49, 80, 158, 200]. Alternatively, only the YAC-ends can be cloned by plasmid rescue. This implies cloning of

part of he vector arm, containing the origin of replication and the ampicilline resistance gene, alongside a part of the insert [35, 82, 99, 111, 124, 131, 136, 166, 191].

A variety of procedures have been developed to obtain end fragments without cloning, including inverse PCR developed by Ochman *et al* [150] and now widely used [8, 37, 50, 134, 166, 182], the vectorette method developed by Riley *et al* [8, 37, 50, 134, 164, 166], the generation of riboprobes using T3 or T7 promotors flanking the cloning site [91], the *Alu*-vector PCR applicable to human YACs [28, 77, 91, 149, 152] and, finally, capture PCR [116, 144].

4.5.2. Chimeric YACs

Often, so called scrambled or chimeric YACs are found in libraries. These are clones in which two (or rarely three) fragments from different chromosomal regions are linked and form one insert. Such YACs are usually larger than non chimeric YACs [120, 177]. Several human YAC libraries have been shown to contain a high proportion of chimeric YACs. About 30% of the clones in the CEPH MarkI library was found to be chimeric [3, 33], the "St Louis" YAC library [30, 125] contained between 60 -70% chimeric clones, while 40% of the clones used for the construction of the chromosome 21q contig were chimeric [30, 47, 125]. Unfortunately, plant YAC libraries have so far not been examined in detail regarding the frequency of occurrence of chimeric clones.

Chimeric clones can arise either by coligation of two different fragments, or by *in vivo* homologous recombination between two YACs, after co-cloning into one cell. Which mechanism is the major cause is difficult to assess and both may be important. To date, the only human chimeric YAC extensively studied was found to have originated from homologous recombination at an *Alu* repeat [90].

An important observation is that YAC libraries made from human/rodent somatic cell hybrids contain less human/human chimeric clones than libraries made from conventional genomic DNA. The proportions of human/human chimeric clones in human/rodent libraries vary from 0 - 20 % [1, 120, 124, 125]. Since the *Alu* repeat is present on every human YAC, there is ample opportunity for homologous recombination between two human YACs. The fact that only a low number of human/rodent chimeric YAC clones are present in libraries from somatic cell hybrids indicates that there are few ligation chimeres [1, 149]. Furthermore, it suggests that homologous recombination is an important cause for chimerism, since homologies between the distantly related rodent and human DNA do not frequently occur (there are no *Alu* repeats in rodents). This suggestion is further supported by the fact that in libraries from *Drosophila* and *Schizosaccharomyces pombe,* two species containing only little repetitive DNA in their genome, low numbers of chimeric YACs are found [128, 138].

The formation of chimeric YACs via homologous recombination implies that *Arabidopsis* YAC libraries should contain relatively few chimeric YACs since the *Arabidopsis* genome contains only few repetitive sequences [139]. However in the construction of a 1.7 Mb contig of *Arabidopsis* DNA, made by Putterill *et al* [160], 35 % of the YACs used for contig building were chimeres. Furthermore, some of the YACs contained both genomic and chloroplast DNA, indicating that ligation chimeras represent at least a part of all chimeric clones.

The high proportion of chimeric YAC clones in human libraries has not prevented the construction of contigs covering entire chromosome arms [47, 76]. The consequence of frequent occurence of chimeric clones is that large libraries, of up to 9 genome equivalents, are required to enable the isolation of a sufficient number of non-chimeric clones [90]. For chromosome walking, which requires the generation of end probes to identify overlapping clones, non-chimeric YACs are essential as chromosome walking is impossible with chimeric YACs.

4.5.3 Unstable YACs

Many studies show that YAC cloning of most euchromatic regions of human DNA is reliable and this has been confirmed by several in depth studies of specific regions, like β-globin gene [78], G6PD region [43], and the HLA complex [29]. In *Drosophila* the same holds true: euchromatic regions are almost all stably represented in YAC libraries [59, 81]. Heterochromatic regions, however, are generally underrepresented in YAC libraries [47, 59, 81] and also some euchromatic regions appear unclonable, as they are either absent from the libraries or present as unstable YACs.

Clones with unstable YACs produce a series of fragments decreasing in size [148], or they contain a deleted YAC which is smaller than the one in the original clone, or something in between these two extremes.

The reasons for unstability or unclonability of some DNA sequences are gradually being elucidated. Neil *et al* [148] performed extensive studies on unstable YACs and found that the presence of tandemly repeated sequences often results in unstability. When he transferred YACs to recombination deficient yeast strains (eg the *rad52* yeast strains 814-9/2d and 814-7/4c) the stability of many YACs increased. This indicates that within the yeast cell YACs recombine at homologous tracts of DNA, leading to loss of intervening sequences. In some cases the unstable YAC clones contain DNA originating from regions that have a highly recombinogenic character such as the end of chromosome 4 [20]. Vilageliu and Tyler-Smith [195], who constructed the recombination deficient yeast strains LIV1 and LIV2, reported that barley YACs, that are unstable in RAD52, strains show increased stability upon transfer to these *rad52* strains [68]. Unfortunately the transformation efficiency of these strains is about

four-fold less compared to the AB1380 strain and therefore these strains are less suitable for library construction.

Until now, only one complete library using a recombination deficient strain has been made by Chartier *et al* [42] who used the *rad52* strain 814-7/3a to construct a mouse YAC library. Although the transformation efficiency of this strain was lower, it did not seriously impart the construction of the library.

On the other hand, the instability of some YAC clones of the MD locus have not prevented the cloning of sequences from this region, for which cosmid walks could not be completed. Despite the unstable YACs, the chromosome walking could be continued as the deleted YACs provided enough sequences to bridge the ustable regions and closing the gaps [16, 179]. Similar results were obtained by Palmieri *et al* [152].

4.6 IDENTIFICATION OF CODING SEQUENCES ON YACs

An important goal of the analysis of genomes is to identify the coding sequences representing genes of interest. A number of strategies has been developed to identify coding seqences and to construct so-called transcription maps, on which all coding sequences are mapped. So far, no methods are available that will detect all coding sequences, with the exception perhaps of sequencing the entire YAC. Here we will discuss some of the methods developed to detect coding sequences on YACs.

4.6.1 Restriction analysis

CpG islands are present at the 5' region of most housekeeping genes [25, 32] and many other genes and they can be used as landmarks to detect the location of a gene. These islands are, in mammals, characterized by a stretch of DNA ranging in size from 0.5 to 2 kb, that is GC-rich, usually unmethylated, and with a proportion of CpG dinucleotides which is higher than might be expected from the overall nucleotide composition of the DNA. Due to these characteristics, the islands contain, on average, a large number of sites for rare-cutting enzymes like *Not*I, *Bss*HII, *Sst*II, *Eag*I and *Nae*I [24]. The islands are detected on a restriction map by a clustering of sites for these rare-cutting enzymes. Although islands can be detected on a genomic restriction map, YACs allow the construction of much more detailed and accurate maps which advances the identification of CpG-islands. With the help of CpG islands 18 new transcripts in the MHC complex class II locus were identified [108, 170], as was the location of genes involved in the fragile X syndrome [64, 97], and in the RCA (regulators of complement activation) gene cluster [102]. After identification of an island, probes from the

region within and downstream of the islands are often used on "zoo-blots" (see section 4.6.3) to detect conserved sequences [43] and, thereafter, for screening of cDNA libraries.

4.6.2 cDNA analysis

Human genes have been identified by probing cDNA libraries with entire YACs [44, 61, 72, 109]. By screening a cDNA library using a complete, gel purified, and radiolabeled YAC of 270 kb, containing the neurofibromatosis type 1 gene, the corresponding cDNA clone was identified by Wallace *et al* [197]. Recently this approach has been used in plants by Arondel *et al* [14] who used an *Arabidopsis* YAC to screen a *Brassica* cDNA library in order to isolate a gene involved in fatty-acid desaturation.

Unfortunately, the method appears rather insensitive since only 10% of the clones identified by conventional screening assays, with phage lambda or plasmid derived probes, were identified by screening with YACs [72, 197]. This means that genes with low expression levels and, therefore, present in low percentages in the cDNA libraries, may be missed.

The reversed approach using the cDNA library as a probe on the YACs was followed by Lanzer *et al* [127]. They prepared total DNA from a YAC clone, which was restriction digested and then hybridized to the blot with radiolabelled cDNA. In that way they were able to detect three transcripts, two of which were previously not recognized, which could then be mapped onto the YAC.

As the sensitivity of the direct cDNA library screening with YACs is rather low, methods have been developed to enhance the sensitivity (Parimo *et al* [153], Lovett *et al* [126]). In these procedures the YAC is immobilized on nylon membranes followed by hybridization with the cDNA libraries, which may be amplified by PCR first . After washing away non-specifically bound cDNAs, specifically bound cDNAs are eluted, amplified by PCR, and recloned. The obtained libraries were enriched by a factor of 800 to 2000 for cDNAs present on the YAC and were then used in a second screening-round to identify cDNAs mapping on the YAC. To avoid a disturbing of repetitive sequences in the hybridization reactions these sequences were blocked by hybridizing either the YAC DNA [153] or cDNA [126], with total human DNA.

4.6.3 Cross-hybridization analysis

Coding regions are often more conserved than non-coding regions [122] and therefore zoo-blots, which contain genomic DNAs from a variety of species, can be used to identify coding sequences [143]. Clones, either cosmid-, phage lambda-, or plasmid clones, representing coding regions will detect hybridizing fragments in DNA of most or all species at higher

stringency hybridization conditions then when they contain non-coding sequencesare detected. For this aproach, YACs have to be first subcloned in cosmid, phage or plasmid vectors. Kinzler *et al* [109] have been able to identify genes from the Familial Adenomatous Polyposis locus, using the procedure.

4.6.4. Exon trapping

The detection of human coding sequences by exon trapping was developed by Buckler *et al* [32b] for human DNA and was improved by Duyck *et al* [69] and Hamagucchi *et al* [94b]. The method is based on the detection of splice site sequences, which are highly conserved in eukaryotic organisms, and present in interrupted genes. In outline, the procedure involves the subcloning of the YAC that is to be screened for coding sequences in "exon trap" vectors, wich contain an intron with the conserved 5' and 3' splice sites of a tester gene. Between these splice sites a multiple cloning site is present into which the YAC fragments are ligated. Following *in vivo* expression and processing, using monkey fibroblast cells, an intron/exon boundary present within the insert will result in a RNA molecule in which the exon is retained or trapped. If an exon is present, this results in RNA molecules larger than the ones that would be obtained if no coding sequences were present.

4.7 MANIPULATION OF YACs IN VIVO

The efficient homologous recombination in yeast cells allows specific manipulation of YACs *in vivo* [154]. This includes the linking of two or more YACs to form one large YAC, the addition of new sequences into the YAC, introduction of specific deletions to the YAC, and specifically altering sequences in any conceivable way, all of which will briefly be discussed in the following. Many of these applications rely on the use of the so called Yeast Integrating plasmids (YIps) which, after transfer to yeast cells, integrate into the yeast genome or into the YAC at homologous regions. Efficient integration of the plasmid at the homologous site is obtained when the plasmid is linearized at this site which results in so called targeted integration.

4.7.1 Combination of YACs

Homologous recombination has allowed the reconstruction of very large genes or gene complexes as a single YAC, which contains the entire gene as a continuous stretch of DNA in its natural configuration. Den Dunnen *et al* [60, 63] have used meiotic homologous

recombination to construct, in a number of steps, a YAC of 2.3 Mb that covers the entire Duchenne Muscular Dystrophy gene. The procedure requires YACs with an overlapping sequence in the proper orientation and yeast cells of compatible mating types. After mating, the diploids are induced to undergo meiosis by plating them out on sporulation agar followed by subculturing of the spores. During meiosis, recombination occurs between the overlapping segments of the cloned YACs, resulting in one large YAC. By similar approaches the CF-gene [88], the BCL2 gene [181] and NF1 gene locus [131] were reconstructed into a single YAC.

4.7.2 Deletion analysis

Integrating plasmids provide means to delete parts of a YAC. Pavan *et al* [155] devised a vector system to introduce a telomere and a selectable marker into a human YAC by means of homologous recombination, in the course of which the original YAC was fragmented into smaller subYACs. The method uses the so called "fragmentation vectors" which are usually introduced via lithium acetate transformation [106] and integrate at the *Alu*-repeats present within the YAC. In that way, a telomere was introduced and, at the same time, the rest of the YAC was deleted. Selection of transformants that have the phenotype Ura^+ (original left arm), Trp^- (due to deletion of right arm) and His^+ (introduced with new right arm), resulted in a set of clones containing deleted YACs.

An approach to construct a set of deletions was developed by Campbell *et al* [38]. The fragmentation vector, which they used, contained both the L1 repeat, and a sequence present on the vector part of the YAC. After introduction into yeast cells the vector integrated into the YAC by the so called omega recombination, which involves a cross-over at the L1 repeat and at the vector sequences, through which all intervening YAC sequences between the repeat and the YAC vector arm are replaced with fragmentation vector sequences. By including the appropriate homologous sequences into the fragmentation vectors, they can be targeted to any required sequence.

This type of vectors can be usefull in identifying genes on YACs by loss of function or in studying regulatory elements, provided means are available to test for the biological function of the gene.

4.7.3 Introduction of mutations

To study the effect of site specific mutations, it is also possible to apply the YIp derived vectors in a so called two step gene replacement. The desired mutations are introduced *in vitro* in a cloned segment. The mutated version of the sequence is inserted into the integrating vector and then introduced into yeast cells with the relevant YACs, where it integrates into the YAC.

In the yeast cells the mutated version of the gene is present on the YAC, alongside the non-mutated sequence. The second step involves a recombination between the nonmutated part of the inserted gene and the original gene of the YAC, thereby removing the vector sequences. The end result is a YAC with a mutated version of the desired gene [154].

4.7.4 Introduction of new sequences

Integrating vecors can also be used to introduce new sequences into YACs. This proces of adding sequences to existing YACs has been named "retrofitting" [48, 71, 184]. It has been applied to introduce selectable markers, such as the neomycine resistance gene, into YACs used for transfection studies in mammalian cells (section 4.8). Any other sequence may be introduced in a similar fashion. Several vectors have been constructed for the introduction of new sequences into YACs and can be used for targeted insertion either into the YAC-vector arms or into the insert part of YACs [71, 156, 184].

4.8 TRANSFER OF ENTIRE YACs

To determine whether a gene, retaining its function together with regulatory elements, has been cloned it is necessary to transfer the cloned DNA back to cells from which it originates. Several reports have now appeared which describe how entire YACs have been transferred to mammalian cells and found to be stably integrated into the genome, while sequences present on the YACs were correctly transcribed and translated [71, 75, 86, 188]. The advantages of YAC as opposed to other cloning systems, are several. Firstly, very large genes can be transferred in their entirety. Secondly, regulatory elements up- or downstream of the coding region can be included. Thirdly, the chances that YACs integrate into the genome at homologous sites are higher due to the presence of longer stretches of homologous sequences [40]. Finally, the chances that the introduced DNA adopt the correct higher order chromatin structure will be increased [154]. Recently, YACs have been introduced, stable and intact, into mouse oocytes resulting in transgenic mice in which the genes present on the YACs were expressed [87, 171]. Even though, the transferred YAC was rather small (35 kb), this opens the way to gene therapy by repairing large genes. Although all the reported cases concern transfer of YACs to mammalian cells, the transfer of plant YACs into plant cells should be feasible following similar approaches.

Several techniques have been used to transfer YACs to mammalian cells. Fusion of yeast spheroplasts with mammalian cells, mediated by PEG, resulted in the integration of YAC sequences into the mammalian genome [86, 103, 151, 156, 191]. Total DNA preparations

from the YAC containing yeast cells have been used for transfection and transfer of the DNA has been achieved by applying lipofectin [86], coprecipitation with calciumphosphate [71, 199] or electroporation [75]. Large fragments (up to 500 kb) have also been transferred by microinjection [87, 171, 172]. In all cases, yeast DNA was transferred together with the YAC DNA and, as yeast DNA also integrated into the recipient genome, mutations may be introduced. However, it was found recently that the introduction of the entire yeast genome into mouse embryo stem (ES) cells did not disturb the functioning nor the growth of the resulting germline chimeras [107, 172]. To avoid this transfer of yeast DNA together with the YAC, purified YAC DNA, obtained from pulsed field gels, can be used [58, 188].

To select for cells containing YAC-DNA, marker genes are required. For that purpose the antibiotic resistance gene *Neo*^{R}, conferring resistance against G418, is most widely used [151, 156, 188, 191]. Alternatively, a gene complementing a host cell deficiency, such as the thymidine kinase (Tk) gene can be used in conjunction with Tk$^-$ cells [71]. Most YAC libraries currently available were made with vectors that do not contain selectable marker genes for mammalian cells, but these markers can be easily introduced into existing YACs by retrofitting (see section 7.4) [71, 151, 156].

4.9 SPECIFIC APPLICATIONS AND TECHNIQUES

4.9.1 Isolation of telomeres

Cloning of telomeres is of considerable interest for several reasons. Firstly, the isolation of telomeric DNA will enable to study their organization and functioning. Secondly, for genetic maps of chromosomes to be of full value they have to be closed, which implies that loci at the end of the chromsome have to be included. The telomeres define the ends of chromosomes and, therefore, also the farthest point of genetic maps. Single copy sequences very close to the telomeres may provide molecular markers like sequence tagged sites (STS's) or RFLP's which can then be used in genetic mapping studies. The genetic maps of several human chromosomes have already been closed by the use of probes derived from YACs containing the proper telomeres [11, 98, 163]. Thirdly, some important genes (eg Huntingtons disease (HD) [83]) are located at the end of a chromosome, very close to the telomeres and may be present on YACs containing the telomeres.

The principle of telomere cloning is simple and relies on the fact that a linear sequence needs telomeres at both ends to be stably maintained within a cell. Furthermore, telomeres are structurally and functionally conserved during evolution allowing the cloning by functional complementation in yeast cells. Telomere cloning vectors provide only one telomeric end, and after ligation to genomic restriction fragments and transfer to yeast cells only those fragments

that contain two telomeres, one from the vector and one from the source DNA, will be maintained in the yeast cells. Clones containing the telomeres are then identified by screening with telomere specific probes. Currently, YACs provide the best way to isolate telomeres since telomeric and subtelomeric repeats are difficult to clone in *E. coli* [65, 163]. Several studies have appeared in which the isolation of human telomeres is described, all of which follow the scheme as described above [31, 45, 57, 163].

4.9.2 Fish mapping

An increasingly important tool is the Fluorescence *In Situ* Hybridization (FISH) technique. Using this method DNA sequences cloned in YACs can be very quickly assigned to specific chromosomal regions [66, 196]. The technique is also very usfull to determine whether a YAC is chimeric, as in that case the YAC sequence will hybridize to two or more different chromosomal locations.

The principle of the procedure is that probes labeled with hapten-like receptor molecules, such as biotin, are hybridized to denatured chromosome preparations fixed on a microscope slide. The specific hybridization is detected by using appropriate fluorescence affinity reagents and a UV microscope. Non-specific hybridization, due to the presence of repetitive sequences, is suppressed by pre-hybridization of the fixed chromosomes with total genomic DNA preparations, or by addition of chromosomal DNA during the hybridization step.

Initially, total DNA preparations of YAC clones, including both yeast and YAC DNA, were used for hybridization but this resulted in rather weak signals [66, 196]. Improvements in FISH hybridization of human chromsomes with YAC DNA were obtained by using DNA fragments that were obtained after an *Alu-Alu* PCR amplification of the YAC-DNA [17, 27, 121]. In this way DNA released from cells from a single YAC clone, after *Alu-Alu* PCR amplification, provided enough material to be used in a FISH assay and, thus, allowed to rapidly determine whether a YAC is chimeric or not [17].

So far FISH mapping of YACs has not yet been applied to plant YACs but, the FISH technique has found application in many plant species to map repetitive elements [eg 162, 175] and should be applicable to map plant YACs as well.

References

1. Abidi FE, Wada M, Little RD, Schlessinger D (1990) Yeast artificial chromosomes containing human Xq24-Xq28 DNA: library construction and representation of probe sequences. Genomics 7: 363-376.
2. Ajioka JW, Smoller DA, Jones RW, Carulli JP, Vellek AEC, Garza D, Link AJ, Duncan IW, Hartl DL (1991) *Drosophila* genome project: one-hit coverage in yeast artificial chromosomes. Chromosoma 100: 495-509.

3. Albertsen HM, Abderrahim H, Cann HM, Dausset J, Le Paslier D, Cohen D (1990) Construction and characterization of a yeast artificial chromosome library containing seven haploid human genome equivalents. Proc Natl Acad Sci USA 87: 4256-4260.
4. Albertsen HM, Le Paslier D, Abderrahim H, Dausset J, Cann H, Cohen D (1989) Improved control of partial DNA restriction digestion in agarose using limiting concentrations of Mg^{2+}. Nucleic Acids Res 17: 808.
5. Allshire RC (1990) Introduction of large linear minichromosomes into *Schizosaccharomyces pombe* by an improved transformation procedure. Proc Natl Acad Sci USA 87: 4043-4047.
6. Amemiya CT, Alegria-Hartman MJ, Aslanidis C, Chen C, Nikolic J, Gingrich JC, de Jong PJ (1992) A two-dimensional YAC pooling strategy for library screening via STS and *Alu*-PCR methods. Nucleic Acids Res 20: 2559-2563.
7. Anand R (Ed) (1992) Techniques for the Analysis of Complex genomes. Academic press, New York.
8. Anand R, Ogilvie DJ, Butler R, Riley JH, Finniear RS, Powell SJ, Smith JC, Markham AF (1991) A yeast artificial chromosome contig encompassing the cystic fibrosis locus. Genomics 9:124-130.
9. Anand R, Riley JH, Butler R, Smith JC, Markham AF (1990) A 3.5 genome equivalent multi access YAC library: construction, characterization, screening and storage. Nucleic Acids Res 18: 1951-1955.
10. Anand R, Villasante A, Tyler-Smith C (1989) Construction of yeast artificial chromosome libraries with large inserts using fractionation by pulsed field gel electrophoresis. Nucleic Acids Res 17: 3425-3433.
11. Anonymous (1992) A comprehensive genetic linkage map of the human genome. Science 258:67-86.
12. Antequera F, Bird AP (1988) Unmethylated CpG islands associated with genes in higher plant DNA. EMBO J 7: 2295-2299.
13. Arman IP, Legchilina SP, Yefimko IG, Devin AB (1992) Genetic effects of heterologous DNA cloned in yeast artificial chromosomes. Abstracts 16th Int conf yeast genetics and molecular biology, p S696.
14. Arondel V, Lemieux B, Hwang I, Gibson S, Goodman H, Somerville CR (1992) Map-based cloning of a gene controlling omega-3 fatty acid desaturation in *Arabidopsis*. Science 258: 1353-1355.
15. Ashikawa I, Fujimura T, Katayose Y, Miyao A, Miyazaki A, Momma T, Saji S, Tanoue H, Umehara Y, Yasukochi Y, Yoshino K, Kurata N (1992) Physical mapping and chromosome mapping. Rice Genome 1: 8-9.
16. Aslanidis C, Jansen G, Amemiya C, Shutler G, Mahadevan M, Tsilfidis C, Chen C, Alleman J, Wormskamp NGM, Vooijs M, Buxton J, Johnson K, Smeets HJM, Lennon GG, Carrano AV, Korneluk RG, Wieringa B, de Jong PJ (1992) Cloning of the essential myotonic dystrophy region and mapping of the putative defect. Nature 355: 548-551.
17. Baldini A, Ross M, Nizetic D, Vatcheva R, Lindsay EA, Lehrach H, Siniscalco M (1992) Chromosomal assignment of human YAC clones by fluorescence *in situ* hybridization: use of single-yeast colony PCR and multiple labeling. Genomics 14: 181-184.
18. Barillot E, Lacroix B, Cohen D (1991) Theoretical analysis of library screening using a N-dimensional pooling strategy. Nucleic Acids Res 19: 6241-6247.
19. Bates GP, MacDonald ME, Baxendale S, Sedlacek Z, Youngman S, Romano D, Whaley WL, Allitto BA, Poustka A, Gusella JF, Lehrach H (1990) A yeast artificial chromosome telomere clone spanning a possible location of the Huntington disease gene. Am J Hum Genet 46: 762-775.
20. Bates GP, Valdes J, Hummerich H, Baxendale S, Le Paslier DL, Monaco AP, Tagle D, MacDonald ME, Altherr M, Ross M, Brownstein BH, Bentley D, Wasmuth JJ, Gusella JF, Cohen D, Collins F, Lehrach H (1992) Characterization of a yeast artificial chromosome contig spanning the Huntington's disease gene candidate region. Nature Genet 1: 180-187.
21. Bellanné-Chantelot C, Lacroix B, Ougen P, Billault A, Beaufils S, Bertrand S, Geordges I, Glibert F, Gros I, Lucotte G, Susini L, Codani JJ, Gesnouin P, Pook S, Vaysseix G, Lu-Kuo J, Ried T, Ward D, Chumakov I, Le Paslier D, Barillot E, Cohen D (1992) Mapping the whole human genome by fingerprinting yeast artificial chromosomes. Cell 70: 1059-1068.
22. Bellis M, Gerard A, Charlieu JP, Marcais B, Brun ME, Viegas-Péquignot E, Carter DA, Roizes (1991) Construction and characterization of a partial library of yeast artificial chromosomes from human chromosome 21. DNA Cell Biol 10: 301-310.
23. Bentley DR, Todd C, Collins J, Holland J, Dunham I, Hassock S, Bankier A, Gianelli F (1992) The development and application of automated gridding for efficient screening of yeast and bacterial ordered libraries. Genomics 12: 534-541.
24. Bickmore WA, Bird AP (1992) The use of restriction endonucleases to detect and isolate genes from mammalian cells. Methods Enzymol 216: 224-244.
25. Bird AP (1987) CpG islands as gene markers in the vertebrate nucleus. Trends Genet 3: 342-347.
26. Bonnema AB, Peytavi R, van Daelen RAJ, Zabel P, Grimsley N (1992) Development of an *in vivo* complementation system for identification of plant genes using yeast artificial chromosomes (YACs). 10th John Innes symposium, The Chromosome, Norwich, p 16.
27. Breen M, Arveiler B, Murray I, Gosden JR, Porteous DJ (1992) YAC mapping by FISH using *Alu*-PCR-generated probes. Genomics 13: 726-730.

28. Breukel C, Wijnen J, Tops C, v/d Klift H, Dauwerse H, Meera Khan P (1990) Vector-*Alu* PCR: a rapid step in mapping cosmids and YACs. Nucleic Acids Res 18: 3097.
29. Bronson SK, Pei J, Taillon-Miller P, Chorney MJ, Geraghty DE, Chaplin DD (1992) Isolation and characterization of yeast artificial chromosome clones linking the HLA-B and HLA-C loci. Proc Natl Acad Sci USA 88: 1676-1680.
30. Brownstein BH, Silverman GA, Little RD, Burke DT, Korsmeyer SJ, Schlessinger D, Olson MV (1989) Isolation of single-copy human genes from a library of yeast artificial chromosome clones. Science 244: 1348-1351.
31. Brown WRA (1989) Molecular cloning of human telomeres in yeast. Nature 338: 774-776.
32. Brown WRA, Bird AP (1986) Long-range restriction site mapping of mammalian genomic DNA. Nature 322: 477-481.
33. Buonavista N, Balzano C, Pontarotti P, Le Paslier D, Golstein P (1992) Molecular linkage of the human CTLA4 and CD28 Ig-superfamily genes in yeast artificial chromosomes. Genomics 13: 856-861.
34. Burgers PMJ, Percival KJ (1987) Transformation of yeast spheroplasts without cell fusion. Anal Biochem 163: 391-397.
35. Burke DT, Carle GF, Olson MV (1987) Cloning of large segments of exogeneous DNA into yeast by means of artificial chromosome vectors. Science 236: 806-812.
36. Burke DT, Olson MV (1991) Preparation of clone libraries in yeast artificial chromosome vectors. Methods Enzymol 194: 251-270.
37. Butler R, Ogilvie DJ, Elvin P, Riley JH, Finniear RS, Slynn G, Morten JEN, Markham AF, Anand R (1992) Walking, cloning and mapping with yeast artificial chromosomes: a contig encompassing D21S13 and D21S16. Genomics 12: 42-51.
38. Campbell C, Gulati R, Nandi AK, Floy K, Hieter P, Kucherlapati RS (1991) Generation of a nested series of interstitial deletions in yeast artificial chromosomes carrying human DNA. Proc Natl Acad Sci USA 88: 5744-5748.
39. Cangiano G, Ameer H, Waterston R, La Volpa A (1990) Use of repetitive DNA probes as physical mapping strategy in *Caenorhabditis elegans*. Nucleic Acids Res 18: 5077-5081.
40. Cappecchi MR (1989) Altering the genome by homologous recombination. Science 244: 1288-1292.
41. Carrozzo R, Ellison J, Yen P, Taillon-Miller P, Brownstein B, Persico G, Ballabio A, Shapiro L (1992) Isolation and characterization of a yeast artificial chromosome (YAC) contig around the human steroid sulfatase gene. Genomics 12: 7-12.
42. Chartier FL, Keer JT, Sutcliffe MJ, Henriques DA, Mileham P, Brown SDM (1992) Construction of a mouse yeast artificial chromosome library in a recombination-deficient strain of yeast. Nature Genetics 1: 132-136.
43. Chen EY, Cheng A, Lee A, Kuang WJ, Hillier L, Green P, Schlessinger D, Ciccodicola A, D'Urso M (1991) Sequence of human glucose-6-phosphate dehydrogenase cloned in plasmids and a yeast artificial chromosome. Genomics 10: 792-800.
44. Chen ZY, Hendriks RW, Jobling MA, Powell JF, Breakefield XO, Sims KB, Craig IW (1992) Isolation and characterization of a candidate gene for Norrie disease. Nature Genet 1: 204-208.
45. Cheng JF, Smith CL, Cantor CR (1989) Isolation and characterization of a human telomere. Nucleic Acids Res 17: 6109-6127.
46. Choi TK, Hollenbach PW, Pearson BE, Ueda RM, Weddell GN, Kurahara CG, Woodhouse CS, Kay RM, Loring JF (1993) Transgenic mice containing a human heavy chain immunoglobulin gene fragment cloned in a yeast artificial chromosome. Nature Genet 4: 117-123.
47. Chumakov, I, Rigault P, Guillou S, Ougen P, Billaut A, Guasconi G, Gervy P, LeGall I, Soularue P, Grinas L, Bougueleret L, Bellané-Chantelot C, Lacroix B, Barillot E, Gesnouin P, Pook S, Vaysseix G, Frelat G, Schmitz A, Sambucy J, Bosch A, Estivill X, Weissenbach J, Vignal A, Riethman H, Cox D, Patterson D, Gardiner K, Hattori M, Sakaki Y, Ichikawa H, Ohki M, Le Paslier D, Heilig R, Antonarakis S, Cohen D (1992) Continuum of overlapping clones spanning the entire human chromosome 21q. Nature 359: 380-387.
48. Chumakov IM, Le Gall I, Billault A, Ougen P, Soularue P, Guillou S, Rigault P, Bui H, De Tand MF, Barillot E, Abderrahim H, Cherif D, Berger R, Le Paslier D, Cohen D (1992) Isolation of chromosome 21-specific yeast artificial chromosomes from a total human genome library. Nature Genet 1: 222-225.
49. Cimino G, Moir DT, Canaani O, Williams K, Crist WM, Katzav S, Cannizzaro L, Lange B, Nowell PC, Croce CM, Canaani E (1991) Cloning the *All-1*, the locus involved in leukemias with the t(4;11)(q21;q23), t(9;11)(p22;q23), and t(11;19)(q23;p13) chromosome translocations. Cancer Res 51: 6712-6714.
50. Coffey AJ, Roberts RG, Green ED, Cole C, Butler R, Anand R, Gianelli F, Bentley DR (1992) Construction of a 2.6-Mb contig in yeast artificial chromosomes spanning the human dystrophy gene using an STS-based approach. Genomics 12: 474-484.

51. Cole CG, Dunham I, Coffey AJ, Ross MT, Meier-Ewert S, Bobrow M, Bentley DR (1992) A random strategy for construction of YAC contigs spanning defined chromosomal regions. Genomics 14: 256-262.
53. Connely C, McCormick MK, Shero J, Hieter P (1991) Polyamines eliminate an extreme size bias against transformation of large yeast artificial chromosome DNA. Genomics 10: 10-16.
54. Cooke H, Cross S (1988) pYAC-4 Neo, a yeast artificial chromosome vector which codes for G418 resistance in mammalian cells. Nucleic Acids Res 16: 11817.
55. Coulson A, Kozono Y, Lutterbach B, Shownkeen R, Sulston J, Waterston R (1991) YACs and the *C. elegans* genome. BioEssays 13: 413-417.
56. Coulson A, Waterston R, Kiff J, Sulston J, Kohara Y (1988) Genome linking with yeast artificial chromosomes. Nature 335: 184-186.
57. Cross SH, Allshire RC, McKay SJ, McGill NI, Cooke HJ (1989) Cloning of human telomeres by complementation in yeast. Nature 338: 771-774.
58. D'Urso M, Zucchi I, Ciccodicola A, Palmieri G, Abidi FE, Schlessinger D (1990) Human Glucose-6-phosphate dehydrogenase gene carried on a yeast artificial chromosome encodes enzyme in monkey cells. Genomics 7: 531-534.
59. Danilevskaya ON, Kurenova EV, Pavlova MN, Bebehov DV, Link AJ, Koga A, Vellek A, Hartl DL (1991) He-T family DNA sequences in the Y chromosome of *Drosophila melanogastor* share homology with the X-linked stellate genes. Chromosoma 100: 118-124.
60. Den Dunnen JT, Grootscholten PM, Dauwerse JD, Monaco AP, Walker AP, Anand R, Coffey AJ, Bentley DR, Steensma HY, Van Ommen GJB (1992) Reconstruction of the 2.4 Mb human DMD-gene by homologous YAC recombination. Hum Mol Genet 1: 19-28.
61. Denizot F, Mattei MG, Vernet C, Pontarotti P, Chimini G (1992) YAC-assisted cloning of a putative G-protein mapping to the MHC class I region. Genomics 14: 857-862.
62. De Bruin D, Lanzer M, Ravetch JV (1992) Characterization of yeast artificial chromosomes from *Plasmodium falciparum*: construction of a stable, representative library and cloning of telomeric DNA fragments. Genomics 14: 332-339.
63. Den Dunnen JT, Steensma HY, Grootscholten PM, Van Ommen GJB (1992) Reconstruction of megabase-sized genomic regions using overlapping yeast artificial chromosomes. In: Techniques for the analysis of complex genomes, Anand R (Ed) Academic Press, London, pp 197-214.
64. Dietrich A, Kioschis P, Monaco AP, Gross B, Korn B, Williams SV, Sheer D, Heitz D, Oberle I, Toniolo D, Warren ST, Lehrach H, Poustka A (1991) Molecular cloning and analysis of the fragile X region in man. Nucleic Acids Res 19: 2567-2572.
65. Dobson MJ, Brown WRA (1992) Cloning human telomeres in yeast artificial chromosomes. In: Techniques for the analysis of complex genomes, Anand R (Ed) Academic Press, London, pp 81-92.
66. Driesen MS, Dauwerse JG, Wapenarr MC, Meershoek EJ, Mollevanger P, Chen KL, Fischbeck KH, van Ommen GJB (1991) Generation and fluorescent *in situ* hybridization mapping of yeast artificial chromosomes of 1p, 17p, 17q and 19q from a hybrid cell line by high-density screening of an amplified library. Genomics 11: 1079-1087.
67. Dunford R, Rogner U (1991) Construction of yeast artificial chromosomes containing barley DNA and the identification of clones carrying copies of the repeated element BIS-1. Hereditas 115: 133-138.
68. Dunford R, Vilageliu L, Moore G (1993) Stabilisation of a yeast artificial chromosome containing plant DNA using a recombinant-deficient host. Plant Mol Biol 21: 1187-1189.
69. Duyck GM, Kim S, Myers RM, Cox DR (1990) Exon trapping: a genetic screen to identify candidate transcribed sequences in cloned mammalian genomic DNA. Proc Natl Acad Sci USA 87: 8995-8999.
70. Edwards K, Thompson H, Edwards D, de Saizieu A, Sparks C (1992) Construction and characterisation of a yeast artificial chromosome library containing three haploid maize genome equivalents. Plant Mol Biol 19: 299-308.
71. Eliceiri B, Labella T, Hagino Y, Srivastava A, Schlessinger D, Pilia G, Palmieri G, D'Urso M (1991) Stable integration and expression in mouse cells of yeast artificial chromosomes harboring human genes. Proc Natl Acad Sci USA 88: 2179-2183.
72. Elvin P, Slynn G, Black D, Graham A, Butler R, Riley J, Anand R, Markham AF (1990) Isolation of cDNA clones using yeast artificial chromosome probes. Nucleic Acids Res 18: 3913-3917.
73. Eyers M, Edwards K, Schuch W (1992) Construction and characterization of a yeast artificial chromosome library containing two haploid *Beta vulgaris* L. genome equivalents. Gene 121: 195-201.
74. Feingold JM, Ogden SD, Denny CT (1990) Streamlined approach to creating yeast artificial chromosome libraries from specialized cell sources. Proc Natl Acad Sci USA 87: 8637-8641.
75. Fernandez-Luna JL, Matthews RJ, Brownstein BH, Schreiber RD, Thomas ML (1991) Characterization and expression of the human leukocyte-common antigen (CD45) gene contained in yeast artificial chromosomes. Genomics 10: 756-764.
76. Foote S, Vollrath D, Hilton A, Page DC (1992) The human Y chromosome: overlapping DNA clones spanning the euchromatic region. Science 258: 60-66.

77. Freije D, Schlessinger D (1992) A 1.6-Mb contig of yeast artificial chromosomes around the human factor VIII gene reveals three regions homologous to probes for the DXS115 locus and two for the DXYS64 locus. Am J Hum Genet 51: 66-80.
78. Gaensler KM, Burmeister M, Brownstein BH, Taillon-Miller P, Myers R (1991) Physical mapping of yeast artificial chromosomes containing sequences from the human ß-globin gene region. Genomics 10: 976-984.
79. Ganal MW, Lapitan NLV, Tanksley SD (1988) A molecular and cytogenetic survey of major repeated DNA sequences in tomato (*Lycopersicon esculentum*). Mol Gen Genet 213: 262-268.
80. Gao J, Erickson P, Gardiner K, LeBeau M, Diaz M, Patterson D, Rowley JD, Drabkin HA (1991) Isolation of a yeast artificial chromosome spanning the 8;21 translocation breakpoint t(8;21)(q22;q22.3) in acute myelogenous leukemia. Proc Natl Acad Sci USA 88: 4882-4886.
81. Garza D, Ajioka JW, Burke DT, Hartl DL (1989) Mapping the *Drosophila* genome with yeast artificial chromosomes. Science 246: 641-646.
82. Geraghty DE, Pei J, Lipsky B, Hansen JA, Taillon-Miller P, Bronson SK, Chaplin DD (1992) Cloning and physical mapping of the *HLA* class I region spanning the *HLA-E*-to-*HLA-F* interval by using yeast artificial chromosomes. Proc Natl Acad Sci USA 89: 2669-2673.
83. Gilliam TC, Tanzi RE, Haines JL, Bonner TI, Faryniarz AG, Hobbs WJ, MacDonald ME, Cheng SV, Folstein SE, Conneally PM, Wexler NS, Gusella JF (1987) Localization of Huntington's disease gene to a small segment of chromosome 4 flanked by *D4S10* and the telomere. Cell 50: 565-571.
84. Gingrich JC, Lowry SR, Kuo WL, Gray J, Smith CL, Cantor CR (1993) Cloning and characterization of *Eag*I YACs from human chromosome 21. Genomics 15: 228-230
85. Giraudat J, Hauge BM, Valon C, Smalle J, Parcy F, Goodman HM (1992) Isolation of the *Arabidopsis ABI3* gene by positional cloning. The Plant Cell 4: 1251-1261.
86. Gnirke A, Barnes TS, Patterson D, Schild D, Featherstone T, Olson MV (1991) Cloning and *in vivo* expression of the human GART gene using yeast artificial chromosomes. EMBO J 10: 1629-1634.
87. Gnirke A, Huxley C (1991) Transfer of the human HPRT and GART genes from yeast to mammalian cells by microinjection of YAC DNA. Somat Cell Mol Genet 17: 573-580.
88. Green ED, Olson MV (1990) Chromosomal region of the cystic fibrosis gene in yeast artificial chromosomes: a model for human genome mapping. Science 250: 94-98.
89. Green ED, Olson MV (1990) Systematic screening of yeast artificial chromosome libraries by the use of the polymerase chain reaction. Proc Natl Acad Sci USA 87: 1213-1217.
90. Green ED, Riethman HC, Dutchik JE, Olson MV (1991) Detection and characterization of chimeric yeast artificial-chromosome clones. Genomics 11: 658-669.
91. Grill E, 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.
92. Groden J, Thliveris A, Samowitz W, Carlson M, Gelbert L, Albertsen H, Joslyn G, Stevens J, Spiro L, Robertson M, Sargeant L, Krapcho K, Wolff E, Burt R, Hughes J, Warrington J, McPhersonJ, Wasmuth J, Le Paslier D, Abderrahim H, Cohen D, Leppert M, White R (1991) Identification and characterization of the familial adenomatous polyposis *coli* gene. Cell 66: 589-600.
93. Grosveld F, Blom van Assendelft J, Greaves DR, Kollias G (1987) Position-independent, high-level expression of the human β-globin gene in transgenic mice. Cell 51: 975-985
94. Guzman P, Ecker JR (1988) Development of large DNA methods for plants: molecular cloning of large segments of *Arabidopsis* and carrot DNA into yeast. Nucleic Acids Res 16: 11091-11105.
94b. Hamagucchi M, Sakamoto H, Tsuruta H, Sasaki H, Muto T, Sugimura T, Terada M (1992) Establishment of a highly sensitive and specific exon-trapping system. Proc Natl Acad Sci 89: 9779-9783.
95. Hampton GM, Tristan JR, Ward J, Cottrell S, Howe K, Thomas HJW, Ballhausen WG, Jones T, Sheer D, Solomon E, Frischauf AM, Bodmer H (1992) Yeast artificial chromosomes for the molecular analysis of the familial polyposis *APC* gene region. Proc Natl Acad Sci USA 89: 8249-8253.
96. Hartl DL, Ajioka JW, Cai H, Lohe AR, Lozovskaya ER, Smoller DA, Duncan IW (1992) Towards a *Drosophila* genome map. Trends Genet 8: 70-75.
97. Heitz D, Rousseau F, Devys D, Saccone S, Abderrahim H, Le Paslier D, Cohen D, Vincent A, Toniolo D, Della Valle G, Johnson S, Schlessinger D, Oberlé I, Mandel JL (1991) Isolation of sequences that span the fragile X and identification of a fragile X-related CpG island. Science 251: 1236-1239.
98. Helms C, Mishra SK, Riethman H, Burgess AK, Ramachandra S, Tierney C, Dorsey D, Donis-Keller H (1992) Closure of a genetic linkage map of human chromosome 7q with centromere and telomere polymorphisms. Genomics 14: 1041-1054.
99. Hermanson GG, Hoekstra MF, McElligott DL, Evans GA (1991) Rescue of end fragments of yeast artificial chromosomes by homologous recombination in yeast. Nucleic Acids Res 19: 4943-4948.
100. Hieter P, Floy K, Puziss J, Sears D, Spencer F (1992) *Saccharomyces cerevisiae* centromere genetics. Abstracts 16th Int conf yeast genetics and molecular biology, p S22.
101. Hinnen A, Hicks J, Fink G (1978) Transformation of yeast. Proc Natl Acad Sci USA 75: 1929-1933.

102. Hourcade D, Garcia AD, Post TW, Taillon-Miller P, Holers VM, Wagner LM, Borsa NS, Atkinson JP (1992) Analysis of the human regulators of complement activation (RCA) gene cluster with yeast artificial chromosomes. Genomics 12: 289-300.
103. Huxley C, Hagino Y, Schlessinger D, Olson MV (1991) The human HPRT gene on a yeast artificial chromosome is functional when transferred to mouse cells by cell fusion. Genomics 9: 742-750.
104. Hwang I, Kohchi T, Hauge B, Goodman HM, Schmidt R, Cnops G, Dean C, Gibson S, Iba K, Lemieux B, Arondel V, Danhoff L, Somerville C (1991) Identification and map position of YAC clones comprising one-third of the *Arabidopsis* genome. The Plant J 1: 367-374.
105. Imai T, Olson MV (1990) Second-generation approach to the construction of yeast artificial chromosome libraries. Genomics 8: 297-303.
106. Ito H, Fukuda Y, Murata K, Kimura A (1983) Transformation of intact yeast cells treated with alkali cations. J Bacteriol 153: 163-168.
107. Jakobovits A, Moore A, Green LL, Vergara GJ, Maynard-Currie CE, Austin HA, Klapholz S (1993) Germ-line transmission and expression of a human-derived yeast artificial chromosome. Nature 362: 255-258.
108. Kendall E, Sargent CA, Campbell RD (1990) Human major histocompatibility complex contains a new cluster of genes between the HLA-D and complement C4 loci. Nucleic Acids Res 18: 7251-7257.
109. Kinzler KW, Nilbert MC, Su LK, Vogelstrein B, Bryan TM, Levy DB, Smith KJ, Preisinger AC, Hedge P, McKechnie D, Finniear M, Markham A, Groffen J, Boguski MS, Altschul SF, Horii A, Miyoshi Y, Miki Y, Nishisho I, Nakamura Y (1991) Identification of FAP locus genes from chromosome 5q21. Science 253: 661-665.
110. Kleine M, Michalek W, Graner A, Herrmann RG, Jung C (1993) Construction of a Barley (*Hordeum vulgare* L.) YAC library and isolation of a *Hor1*-specific clone. Mol Gen Genet 240: 265-272.
111. Kozono H, Bronson SK, Taillon-Miller P, Kavita Moorti M, Jamry I, Chaplin DD (1991) Molecular linkage of the HLA-DR, HLA-DQ and HLA-DO genes in yeast artificial chromosomes. Genomics 11: 577-586.
112. Kuspa A, Maghakian D, Bergesch P, Loomis WF (1992) Physical mapping of genes to specific chromosomes in *Dictyostelium discoideum*. Genomics 13: 49-61.
113. Kuspa A, Vollrath D, Cheng Y, Kaiser D (1989) Physical mapping of the *Myxococcus xanthus* genome by random cloning in yeast artificial chromosomes. Proc Natl Acad Sci USA 86: 8917-8921.
114. Kwiatkowsky Jr TJ, Zoghby HY, Ledbetter SA, Ellison KA, Craig Chinault A (1990) Rapid identification of yeast artificial chromosome clones by matrix pooling and crude lysate PCR. Nucleic Acids Res. 18: 7191.
115. Labella T, Schlessinger D (1989) Complete human rDNA repeat units isolated in yeast artificial chromosomes. Genomics 5: 752-760.
116. Lagerström M, Parik J, Malmgren H, Stewart J, Pettersson U, Landegren U (1991) Capture PCR: efficient amplification of DNA fragments adjacent to a known sequence in human and YAC DNA. PCR Meth Appl 1:111-119.
117. Lanzer M, de Bruin D, Ravetch JV (1992) Transcription mapping of a 100 kb locus of *Plasmodium falciparum* identifies an intergenic region in which transcription terminates and reinitiates. EMBO J 11: 1949-1955.
118. Larin Z, Lehrach H (1990) Yeast artificial chromosomes: an alternative approach to the molecular analysis of mouse developmental mutations. Genet Res Camb 56: 203-208.
119. Larin Z, Monaco A, Lehrach H (1991) Yeast artificial chromosome libraries containing large inserts from mouse and human DNA. Proc Natl Acad Sci USA 88: 4123-4127.
120. Lee JT, Murgia A, Sosnoski DM, Olivos I, Nussbaum RL (1992) Construction and characterization of a yeast artificial chromosome library for Xpter-Xq27.3: A systematic determination of cocloning rate and X-chromosome representation. Genomics12: 526-533.
121. Lengauer C, Green ED, Cremer T (1992) Fluorescence *in situ* hybridization of YAC clones after *Alu*-PCR amplification. Genomics 13: 826-828.
122. Lewin B (1990) Genes IV. Oxford University Press, Cambridge, Mass.
123. Libert F, Lefort A, Okimoto R, Womack J, Georgest M (1993) Construction of a bovine genomic library of large yeast artificial chromosome clones. Genomics 18: 270-276.
124. Little RD, Pilia G, Johnson S, D'Urso M (1992) Yeast artificial chromosomes spanning 8 megabases and 10-15 centimorgans of human cytogenetic band Xq26. Proc Natl Acad Sci USA 89; 177-181.
125. Little RD, Porta G, Carle G, Schlessinger D (1989) Yeast artificial chromosomes with 200- to 800-kilobase inserts of human DNA containing HLA, V_K, 5S, and Xq24-Xq28 sequences. Proc Natl Acad Sci USA : 86: 1598-1602.
126. Lovett M, Kere J, Hinton LM (1991) Direct selection: A method for the isolation of cDNAs encoded by large genomic regions. Proc Natl Acad Sci USA 88: 9628-9632.
127. Macina RA, Riethman HC (1992) Direct DNA hybridization screening of primary yeast transformants in the construction of targeted yeast artificial chromosome (YAC) libraries. GATA 9: 58-63.

128. Maier E, Hoheisel JD, McCarthy L, Mott R, Grigiriev AV, Monaco AP, Larin Z, Lehrach H (1992) Complete coverage of the *Schizosaccharomyces pombe* genome in yeast artificial chromosomes. Nature Genetics 1: 273-277
129. Mann C, Davis RW (1986) Structure and sequence of the centromeric DNA of chromosome 4 in *Saccharomyces cerevisiae*. Mol Cell Biol 6: 241-245.
130. Marchuk D, Collins FS (1988) pYAC-RC a yeast artificial chromosome vector for cloning DNA cut with infrequently cutting restriction endonucleases. Nucleic Acids Res 16: 7743.
131. Marchuk DA, Tavakkol R, Wallace MR, Brownstein BH, Taillon-Miller P, Fong CT, Legius E, Andersen LB, Glover TW, Collins FS (1992) A yeast artificial chromosome contig encompassing the type 1 neurofibromatosis gene. Genomics 13: 672-680.
132. Martin G, Brommonschenkel SH, Chungwongse J, Frary A, Ganal M, Spivey R, Wu T, Earle ED, Tanksley SD (1993) Map-based cloning of a protein kinase gene conferring disease resistance in tomato. Science 262: 1432-1436.
133. Martin GB, Ganal MW, Tanksley SD (1992) Construction of a yeast artificial chromosome library of tomato and identification of cloned segments linked to two disease resistance loci. Mol Gen Genet 233: 25-32.
134. Matallana E, Bell CJ, Dunn P, Lu M, Ecker JR (1991) Genetic and physical linkage of the *Arabidopsis* genome: methods for anchoring yeast artificial chromosomes EMBO/EEC advanced laboratory course on *Arabidopsis* molecular genetics. pp 144-169.
135. McCormick MK, Campbell E, Deaven L, Moyzis R (1993) Low-frequency chimeric yeast artificial chromosome libraries from flow sorted human chromosome 16 and 21. Proc Natl Acad Sci 90: 1063-1067.
136. McCormick MK, Shero JH, Cheung MC, Kan YW, Hieter PA, Antonarakis SE (1989) Construction of human chromosome 21-specific yeast artificial chromosomes. Proc Natl Acad Sci USA 86: 9991-9995.
137. Mendez MJ, Klapholz S, Brownstein BH, Gemmill RM (1991) Rapid screening of a YAC library by pulsed field gel southern blot analysis of pooled YAC clones. Genomics 10: 661-665.
138. Merriam J, Ashburner M, Hartl DL, Kafatos FC (1991) Toward cloning and mapping the genome of *Drosophila* Science 254: 221-225.
139. Meyerowitz EM (1987) *Arabidopsis thaliana*. Annu Rev Genet 21: 93-111.
140. Monaco A, Walker AP, Millwood I, Larin Z, Lehrach H (1992) A yeast artificial chromosome contig containing the complete Duchenne muscular dystrophy gene. Genomics 12: 465-473.
141. Monaco AP, Lam VMS, Zehetner G, Lennon GG, Douglas C, Nizetic D, Goodfellow PN, Lehrach H (1991) Mapping irradiation hybrids to cosmid and yeast artificial chromosome libraries by direct hybridization of *ALU*-PCR products. Nucleic Acids Res 19: 3315-3318.
142. Monaco AP, Müller U, Larin Z, Meier-Ewert S, Lehrach H (1991) Isolation of the human sex determining region from an Y-enriched yeast artificial chromosome library. Genomics 11: 1049-1053.
143. Monaco AP, Neve RL, Colletti-Feener C, Bertelson CJ, Kurnit DM, Kunkel LM (1986) Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. Nature 323: 646-650.
144. Mueller PR, Wold B (1989) In vivo footprinting of a muscle specific enhancer by ligation mediated PCR. Science 246: 780-786.
145. Murray AW, Claus TE, Szostak JW (1988) Characterization of two telomeric DNA processing reactions in *Saccharomyces cerevisiae*. Mol Cell Biol 8: 4642-4650.
146. Murray AW, Claus TE, Szostak JW (1988) Characterization of two telomeric DNA processing reactions in *Saccharomyces cerevisiae*. Mol Cell Biol 8: 4642-4650.
147. Murray AW, Szostak JW (1983) Construction of artificial chromosomes in yeast. Nature 305: 189-193.
148. Neil DL, Villasante A, Fisher RB, Vetrie D, Cox B, Tyler-Smith C (1990) Structural instability of human tandemly repeated DNA sequences cloned in yeast artificial chromosome vectors. Nucleic Acids Res 18: 1421-1428.
149. Nelson DL, Ballabio A, Victoria MF, Pieretti M, Bies RD, Gibbs RA, Maley JA, Craig Chinault A, Webster TD, Caskey CT (1991) *Alu*-primed polymerase chain reaction for regional assignment of 110 yeast artificial chromosome clones from the human X chromosome: Identification of clones associated with a disease locus. Proc Natl Acad Sci USA 88: 6157-6161.
150. Ochman H, Gerber AS, Hartl DL (1988) Genetic application of an inverse polymerase chain reaction. Genetics 120: 621-623.
151. Pachnis V, Pevny L, Rothstein R, Constantini F (1990) Transfer of a yeast artificial chromosome carrying human DNA from *Saccharomyces cerevisiae* into mammalian cells. Proc Natl Acad Sci USA 87: 5109-5113.
152. Palmieri G, Capra V, Romano G, D'Urso M, Johnson S, Morris P, Hopwood J, Natale P, Gatti R, Ballabio A (1991) The iduronate sulfatase gene: isolation of a 1 Mb YAC contig containing the entire gene and identification of heterogeneous deletions in patients with Hunter syndrome. Cold Spring Harbor Genome mapping and sequencing abstracts. p 168. {ref 42 in 142}

153. Parimoo S, Patanjali SR, Shukla H, Chaplin DC, Weissman SM (1991) cDNA selection: Efficient PCR approach for the selection of cDNAs encoded in large chromosomal DNA fragments. Proc Natl Acad Sci USA 88: 9623-9627.
154. Pavan WJ, Hieter P, Griffin CA, Hawkins AL, Reeves RH (1992) Yeast artificial chromosome modifying vectors: application for the analysis of complex genomes.In: Techniques for the analysis of complex genomes, Anand R (Ed) Academic Press, London, pp 173-196.
155. Pavan WJ, Hieter P, Reeves R (1990) Generation of deletion derivatives by targeted transformation of human-derived yeast artificial chromosomes. Proc Natl Acad Sci USA 87: 1300-1304.
156. Pavan WJ, Hieter P, Reeves RH (1990) Modification and transfer into an embryonal carcinoma cell line of a 360-kilobase human derived yeast artificial chromosome. Mol Cel Biol 10: 4163-4169.
157. Pavan WJ, Reeves RH (1991) Integrative selection of human chromosome-specific yeast artificial chromosomes. Proc Natl Acad Sci USA 88: 7788-7791.
158. Pieretti M, Tonlorenzi R, Ballabio A (1991) Rapid assembly of l phage contigs within YAC clones. Nucleic Acids Res 19: 2795-2796.
159. Potier MC, Kuo WL, Dutriaux A, Gray J, Goedert M (1992) Construction and characterization of a yeast artificial chromosome library containing 1.5 equivalents of human chromosome 21. Genomics 14: 481-483.
160. Putterill J, Robson F, Lee K, Coupland G (1993) Chromosome walking with YAC clones in *Arabidopsis*: isolation of 1700 kb of contiguous DNA on chromosome 5, including a 300 kb region containing the flowering-time gene *CO*. Mol Gen Genet 239: 145-157.
161. Ragoussis J, Monaco A, Mockridge I, Kendall E, Duncan Campbell R, Trowsdale J (1991) Cloning of the HLA class II region in yeast artificial chromosomes. Proc Natl Acad Sci USA 88: 3753-3757.
162. Ricroch A, Peffley EB, Baker RJ, Robert T, Yakovlev S, Sarr A (1992) Chromosome organization using FISH and fluorochrome banding techniques in *Allium cepa*. Abstract Int Conf Plant Genome, November 1992, San Diego, USA, no 118.
163. Riethman HC, Moyzis RK, Meyne J, Burke DT, Olson MV (1989) Cloning human telomeric DNA fragments into *Saccharomyces cerevisiae* using a yeast-artificial-chromosome vector. Proc Natl Acad Sci USA 86: 6240-6244.
164. Riley J, Butler R, Ogilvie D, Finniear R, Jenner D, Powell S, Anand R, Smith JC, Markham AF (1990) A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. Nucleic Acids Res 18: 1887-2890.
165. Rommens JM, Ianuzzi MC, Kerem B, Drumm ML, Melmer G, Dean M, Rozmahel R, Cole JL, Kennedy D, Hidaka N, Zsiga M, Buchwald M, Riordan JR, Tsui LC, Collins FS (1989) Identification of the cystic fibrosis gene: chromosome walking and jumping. Science 245: 1059-1065.
166. Rossi JM, Burke TD, Leung JCM, Koos DS, Chen H, Tilghman SM (1992) Genomic analysis using a yeast artificial chromosome library with mouse DNA inserts. Proc Natl Acad Sci USA 89: 2456-2460.
167. Ross MT, Hoheisel JD, Monaco AP, Larin Z, Zehetner G, Lehrach H (1992) High-density gridded YAC filters: their potential as genome mapping tools. In: Techniques for the analysis of complex genomes, Anand R (Ed) Academic Press, London, pp 137-154.
168. Ross MT, Nizetic D, Nguyen C, Knughts C, Vatcheva R, Burden N, Douglas C, Zehetner G, Ward DC, Baldini A, Lehrach H (1992) Selection of a human chromosome 21 enriched YAC sub-library using a chromosome-specific composite probe. Nature Genetics 1: 284-290.
169. Rowley JD, Diaz MO, Espinoza III R, Patel YD, van Melle E, Ziemin S, Taillon-Miller P, Lichters P, Evans GA, Kersey JH, Ward DC, Domer PH, Le Beau MM (1990) Mapping chromosome band 11q23 in human acute leukemia with biotinylated probes: Identification of 11q23 translocation breakpoints with a yeast artificial chromosome. Proc Natl Acad Sci USA 87: 9358-9362.
170. Sargent CA, Dunham I, Duncan Campbell R (1989) Identification of multiple HTF-islands associated genes in the human major histocompatibility complex class III region. EMBO J 8: 2305-2312.
171. Schedl A, Beerman F, Thies E, Montoliu L, Kelsey G, Schütz G (1992) Transgenic mice generated by pronuclear injection of a yeast artificial chromosome. Nucleic Acids Res 20: 3073-3077.
172. Schedl A, Montoliu L, Kelsey G, Schütz G (1993) A yeast artificial chromosome covering the tyrosinase gene confers copy number-dependent expression in transgenic mice. Nature 362: 258-261.
173. Schlessinger D, Little RD, Freije D, Abidi F, Zucchi I, Porta G, Pilia G, Nagaraja R, Johnson SK, Yoon JY, Srivastava A, Kere J, Palmieri G, Ciccodicola A, Montanaro V, Romano G, Casamassimi A, D'Urso M (1991) Yeast artificial chromosome-based genome mapping: some lessons from Xq24-Xq28. Genomics 11: 783-793.
174. Schmidt R, Cnops G, Bancroft I, Dean C (1991) Development of an overlapping set of YAC clones for the top halves of chromosome 4 and 5 of *Arabidopsis thaliana*. Abstract 1693, ISPMB 3rd Int Congress Plant Mol Biol, Tucson, Arizona, USA.

175. Schwarzacher T, Anamthawat-Jonson K, Harrison GE, Islam AKMR, Jia JZ, King IP, Leitch AR, Miller TE, Reader SM, Rogers WJ, Shi M, Heslop-Harrison JS (1992) Genomic *in situ* hybridization to identify alien chromosome segments in wheat. Theor Appl Genet 8: 778-786.
176. Sears DD, Hegemann JH, Hieter P (1992) Meiotic recombination and segregation of human-derived artificial chromosomes in *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 89: 5296-5300.
177. Selleri L, Eubanks JH, Giovannini M, Hermanson GG, Romo A, Djabali M, Maurer S, McElligot DL, Smith MW, Evans GA (1992) Detection and characterization of "chimeric" yeast artificial chromosome clones by fluorescent *in situ* suppression hybridization. Genomics 14: 536-541.
178. Shero JH, McCormick MK, Antonarakis SE, Hieter P (1991) Yeast artificial chromsome vectors for efficient clone manipulation and mapping. Genomics 10: 505-508.
179. Shutler G, Korneluk RG, Tsilfidis C, Mahadevan M, Bailly J, Smeets H, Jansen G, Wieringa B, Lohman F, Aslanidis C, de Jong PJ (1992) Physical maping and cloning of the proximal segment of the myotonic dystrophy gene region. Genomics 13: 518-525.
180. Sikorski RS, Hieter P (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *S. cerevisiae*. Genetics 122: 19-27.
181. Silverman GA, Green ED, Young RL, Jockel JI, Dohmer PH, Korsmeyer SJ (1990) Meiotic recombination between yeast artificial chromosomes yields a single clone containing the entire *BCL2* protooncogene. Proc Natl Acad Sci USA 87: 9913-9917.
182. Silverman GA, Ye RD, Pollock KM, Evan Sadler J, Korsmeyer SJ (1989) Use of yeast artificial chromosome clones for mapping and walking within human chromosome segment 18q21.3. Proc. Natl. Acad. Sci. USA 86: 7485-7489.
183. Smith DR, Smyth A, Moir DT (1990) Amplification of large artificial chromosomes. Proc Natl Acad Sci USA 87: 8242-8246.
184. Srivastava AK, Schlessinger D (1991) Vectors for inserting selectable markers in vector arms and human DNA inserts of yeast artificial chromosomes (YACs) Gene 103: 53-59.
185. Stallings RL, Torney DC, Hildebrand CE, Longmire JL, Deavan LL, Jett JH, Doggett NA, Moyzis RK (1990) Physical mapping of human chromosomes by repetitive fingerprinting. Proc Natl Acad Sci USA 87: 6218-6222.
186. Stinchcomb DT, Struhl K, Davis RW (1979) Isolation and characterisation of a yeast chromosomal replicator. Nature 282: 39-43.
187. Strauss WM, Jaenisch E, Jaenisch R (1992) A strategy for rapid production and screening of yeast artificial chromosome libraries. Mammalian genome 2: 150-157.
188. Strauss WM, Jaenisch R (1992) Molecular complementation of a collagen mutation in mammalian cells using yeast artificial chromsomes. EMBO J 11:417-422.
189. Strouboulis J, Dillon N, Grosveld F (1992) Developmental regulation of a complete 70-kb human ß-globin locus in transgenic mice. Genes Dev 6: 1857-1864.
190. Trask B (1991) Fluorescence *in situ* hybridization: applications in cytogenetics and gene mapping. Trends Genet 7: 149-152.
191. Traver CN, Klapholz S, Hyman RW, Davis RW (1989) Rapid screening of a human genomic library in yeast artificial chromosomes for single-copy sequences. Proc Natl Acad Sci USA 86: 5898-5902.
192. Triglia T, Kemp DJ (1991) Large fragments of *Plasmodium falciparum* DNA can be stable when cloned in yeast artificial chromosomes. Mol Biochem Parasitol 44: 207-212.
193. Tschumper G, Carbon J (1980) Sequence of a yeast DNA fragment containing a chromosomal replicator and the *TRP1* gene. Gene 10: 157-166.
194. Van Daelen RAJJ, Zabel P (1991) Preparation of high molecular weight plant DNA and analysis by pulsed-field gel electrophoresis. Plant Mol Biol Manual **A15**: 1-25
195. Vilageliu L, Tyler-Smith C (1992) Structural instability of YAC clones and the use of recombination-deficient yeast host strains. In: Techniques for the analysis of complex genomes, Anand R (Ed) Academic Press, London, pp 93-112.
196. Wada M, Little RD, Abidi F, Porta G, Labella T, Coper T, Della Valle G, D'Urso M, Schlessinger D (1990) Human Xq24-Xq28: approaches to mapping with yeast artificial chromosomes. Am J Hum Genet 46: 95-106.
197. Wallace MR, Marchuck DA, Andersen LB, Letcher R, Odeh HM, Saulino AM, Fountain JW, Brereton A, Nicholson J, Mitchell AL, Brownstein BH, Collins FS (1990) Type 1 neurofobromatosis gene: identification of a large transcript disrupted in three NF1 patients Science: 249: 181-186.
198. Ward ER, Jen GC (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.
199. Wigler M, Pellicer A, Silverstein S, Axel R, Urlaub G, Chasin L (1979) DNA-mediated transfer of the adenine phosphoribosyltransferase locus into mammalian cells. Proc Natl Acad Sci USA 76: 1373-1376.
200. Ziemin-van der Poel S, McCabe NR, Gill HJ, EspinosaIII R, Patel Y, Harden A, Rubinelli P, Smith SD, LeBeau MM, Rowley JD, Diaz MO (1991) Identification of a gene, *MLL*, that spans the

breakpoint in 11q23 translocations associated with human leukemias. Proc Natl Acad Sci USA 88: 10735-10739.
201. Zucchi I, Schlessinger D (1992) Distribution of moderately repetitive sequences pTR5 and LF1 in Xq24-q28 human DNA and their use in assembling YAC contigs. Genomics 12: 264-275.

CHAPTER 5

CONSTRUCTION AND SCREENING OF A RARE-CUTTER TOMATO YAC-LIBRARY

Abstract

Yeast artificial chromosomes (YACs) allow the cloning of very large DNA fragments and are thus a major tool in the isolation of genes through a positional cloning approach. Currently, several plant YAC libraries covering the genomes of tomato, *Arabidopsis*, carrot, sugar beet, barley, rice and maize have been constructed. The construction of these whole genome libraries that consist of a set of overlapping clones generated from partially digested genomic DNA is very laborious and time consuming. Here, we describe a relatively simple procedure for constructing a YAC library from a rare-cutter digest that serves in the isolation of YAC clones specific for the tomato *Aps-1/Mi*-region. The choice of the rarely cutting restriction enzyme is directed by the long range restriction maps available for this region. A small amount of DNA (25 µg) was digested with *Cla*I and after ligation to pYAC-RC transformed to yeast strain AB1380. Of the 40,000 clones obtained, 6000 were picked and characterized. The library was screened by high density colony filter hybridization with *Aps-1* and GP79, two markers flanking the root-knot nematode resistance gene *Mi*. A GP79 YAC-clone of 200 kb was isolated and characterized.

INTRODUCTION

The recent development of cloning large DNA fragments in yeast as artificial chromosomes (YACs [8]) allows large chromosomal regions to be traversed and, thus, bridges the gap between molecular and genetic analysis. This makes YAC cloning an essential tool in positional cloning approaches [39] as examplified by the cloning of many human and plant genes, for which only a phenotype and genetic position was known [4, 16, 27b, 32].

To date, plant YAC libraries have been constructed for *Arabidopsis* [17, 18, 21, 40], tomato [27], carrot [18], barley [10], sugar beet [12], rice [4b] and maize [11]. All these libraries have been constructed so as to cover the entire genome, using total genomic DNA partially digested with a frequently cutting enzyme as a source for cloning. The construction of such libraries is very laborious and time consuming. Furthermore, many clones are required to cover the entire genome. As to tomato, this would require 22,000 to 36,000 clones in a 3 to 5 genome equivalent library, assuming an average YAC size of 140 kb [27]. To avoid the construction and handling of such large libraries we followed an alternative approach. Based on the information from long range restriction maps we used a YAC vector to clone specific large restriction fragments generated by a rarely cutting restriction enzyme.

Our interest lies in the positional cloning of the root-knot nematode resistance gene *Mi* which is located on chromosome 6 of tomato. The gene confers resistance against several species of root-knot nematodes of the genus *Meloidogyne* and is tightly linked to the *Aps-1*[1] gene [1, 31]. Several molecular markers have been identified that are tightly linked to *Mi* [15, 20, 22, 29]. These markers have been ordered into high resolution linkage maps (Messenguer *et al* [29], Ho *et al* [20]). For two closely flanking markers, GP79 and *Aps-1*, long range restriction maps have been constructed [37]. Here we describe the construction and screening of a *Cla*I YAC library, designed specifically to isolate GP79 and *Aps-1* YACs. The enzyme *Cla*I was chosen since both GP79, the closest single copy marker, and *Aps-1*, a marker on the opposite site of *Mi*, recognize restricion fragments of similar sizes. GP79 recognizes three restriction fragments of 65, 135 and 200 kb while *Aps-1* resides on a 90 kb *Cla* I fragment. A library, enriched for fragments of these sizes, was made of DNA isolated from the nematode resistant cultivar VFNT-Cherry. The library encompasses 6000 clones with an average size of 110 kb. Here, the isolation and physical characterization of a GP79 YAC of 200 kb, isolated by high density colony filter hybridization, is reported.

Materials and methods

DNA methodology

All standard DNA manipulations were carried out as described in Sambrook *et al* [33]. Isolation of high molecular weight DNA, restriction digestion, pulsed field gel electrophoresis and hybridization of Southern blots are described elsewhere [36]. Probes were labelled according to the "random primer" method of Feinberg and Vogelstein [13]

YAC library construction.

High molecular weight DNA, isolated from the nematode resistant tomato cultivar VFNT cherry, was digested to completion with the rarely cutting restriction enzyme *Cla*I. Appearance of the distinct, low molecular weight chloroplast DNA banding pattern on a pulsed field gel [36] was used as an indication for completeness of the restriction digestion. A total of 5 plugs was used, containing approximately 25 µg of DNA in a total volume of 500 µl. After digestion, the enzyme was removed by two washes of 50 ml TE (10 mM TRIS-HCl pH 8.0, 1 mM EDTA). The plugs were equilibrated with ligation buffer (50 mM TRIS-HCl pH 7.5, 10 mM $MgCl_2$ and 125 mM NaCl) by incubating each plug with 1 ml, twice for 30 min. After equilibration, the plugs were collected in one tube and 200 µg vector DNA (pYAC-RC [26], digested with *Bam*HI and *Cla*I and dephosphorylated) was added. The plugs were melted (5 min at 68 °C) and after cooling to 37 °C ATP (1 mM final concentration), DTT (1 mM final concentration) and ligase (40 U, Pharmacia) were added and only slightly mixed by a few stirs with a yellow pipet tip. After incubation at 37 °C for 4 hours, the mixture was pipetted into a mold (kept at 0 °C) to produce 5 plugs of 100 µl each. The ligation was allowed to proceed within the plugs at 15 °C after transfer of the plug into an equal volume of ligation buffer containing an extra 20 u of ligase.

In order to select ligation products of the appropriate size, the plugs were loaded on a 1% (w/v) low melting point agarose gel (made in 0.5 x TBE) and subjected to CHEF [9] electrophoresis (Biorad CHEF DRII) using 5 s pulses for 16 hours, at 100 Volts. Using this protocol a compression zone at approximately 100 kb was obtained. The compression zone was identified by aligning the ethidium bromide-stained sides of the gel, which included the markers and part of the ligation products, to the unstained part, as described by Albertsen *et al* [2] and Anand *et al* [3]. A slice of approximately 250 µl, containing the compression zone, was cut out and stored in 0.5 M EDTA, 10 mM TRIS-HCl pH 8.0 until use. Prior to transformation, the slice was equilibrated with TE containing 125 mM NaCl, melted at 68 °C, cooled to 40 °C, and incubated with agarase (NEB; 1u per 20 µl) for 2 to 4 hours to

accomplish complete liquifaction. Transformation, according to the protocol of Burgers and Percival [7], was carried out by adding 20 µl ligation mixture to 150 µl of spheroplasts.

Colonies appeared after two days and the red color was fully developed after five days. Red colonies (over 95 %) were picked and transferred to the wells of a microtiter dish, containing 75 µl YPD (1% (w/v) yeast extract, 2% (w/v) bacto pepton, 2% (w/v) dextrose [33]). The plates were incubated for 1 day at 30 °C, under continuous rotation (210 rpm). After growth, glycerol was added to obtain a final concentration of 15-20 % (w/v). YPD (4 times concentrated) was added to obtain a final concentration of 1 x YPD to enhance recovery after storage. The plates were stored at -50 °C

Preparation of DNA from YAC clones

Low molecular weight yeast DNA was isolated following the protocol of Hoffmann and Winston [19]. Intact yeast chromosomes for small scale isolations were prepared from 600 µl cultures (YPD), grown overnight at 30 °C, as described elsewhere [36]. From a 600 µl culture one agarose plug was made for PFGE analysis. For large scale isolations either 50 or 100 ml cultures in AHC [6] were used.

Screening of the YAC library.

In order to screen the library by colony filter hybridization the colonies were plated with a 96 pin replicator onto double selective medium (AHC [6]). After growth for 3 days, the colonies of 16 plates were all transferred to one YPD plate with 3% (w/v) agar. The selective plates were freshly used to obtain relatively wet colonies which adhere evenly to the pins of the replicator and thus allow the transfer of equal amounts of cells per colony to the YPD plates. Gridding of the colonies was done manually, using a (self made) replicator with 96 pins 0.2 mm in diameter. After growth for 1 or 2 days, lifts (normally four, but up to 10 were possible) were made by placing a dry Gene Screen Plus membrane on the colonies. After lifting, the membranes were placed, colony side up, on YPD medium and incubated at 30 °C for one night to achieve an even growth of all colonies. The cells were spheroplasted by placing the filters onto 3MM paper soaked with SCEM (1 M sorbitol, 0.1 M Na-Citrate pH 5.8, 10 mM EDTA, 144 mM β-mercaptoethanol) containing lyticase (200 u/ml) and incubated overnight at 30 °C. The spheroplasts were lysed by incubating the filters with 10 % (w/v) SDS for 10 min, followed by two incubations with 0.5 N NaOH for 10 min. Neutralization was achieved by floating the filters for 5 min on a 0.5 M Tris-HCL pH 7.6, 1.5 M NaCl solution followed by incubation for 5 min in 0.05 M Tris-HCl pH 7.6, 0.15 M NaCl. Colony debris was gently wiped off using a Kleenex tissue after which the filters were submerged in

2 x SSC. Finally, the filters were air dried and baked (80 °C for 45 min). Hybridization of the filters was performed as described for Southern blots [22]. Usually hybridizations were carried out in duplicate, sometimes in triplicate, to avoid picking false positives.

To screen the library by PCR all colonies were cultured in the wells of microtiter dishes, containing 50 µl YPD, for 2 days at 30 °C. In the first screening round all the cells from each microtiter plate were pooled and DNA was isolated, to identify the plate containing positive clones. After identification of a positive plate, the positive clone was identified by screening DNAs prepared from pools of cells from each individual row and column. For a PCR reaction approximately 50 ng of total yeast DNA was used. Amplification during PCR was obtained by an initial denaturation of 5 min at 94 °C followed by 30 cycles consisting of 1 min denaturation at 94 °C, 1 min annealing at 60 °C, and 1 min elongation at 72 °C and the run was terminated by a final incubation for 10 min at 72 °C. Reaction conditions during PCR were as described elsewhere [23]. Primers used for the amplification of *Aps-1*[1] were 5-ATGGTGGGTCCAGGTTATAAG and 5-CAGAATGAGCTTCTGCCAATC. For GP79 no primers were available. As a positive control for the PCR screening assay, the genomic *Hin*DIII [41] repeat was used. Primers for this repeat were 5-GACTACTGCACCAATT-CTCACC and 5-GCTCCCTGCTTTTCTACTTAAGG.

Used probes

Gel purified fragments of *Aps-1*[1] [37], the genomic *Hin*DIII repeat (pTHG2 [41]) and GP79 [14, 20] were used as probes. For detection of chloroplast DNA a probe for the large subunit of Rubisco [38] was used.

Results

YAC library construction

From long range restriction mapping studies of the Aps/*Mi* region it was known that the marker GP79 recognized three *Cla*I restriction fragments of 65, 135 and 205 kb, respectively, and *Aps-1* a restriction fragment of approximately 90 kb [37]. Since fragments of these sizes are easily cloned in YAC vectors we set out the construction of a small *Cla*I tomato YAC library. To this end, DNA from the cultivar VFNT Cherry was digested to completion with *Cla*I and cloned into the yeast host strain AB1380 using the vector pYAC-RC [26]. The bulk of clones obtained after transformation of a ligation mixture without size selection, was expected to be in the range up to 150 kb [3]. Initial experiments resulted, however, in clones carrying, on average, small inserts (far below 75 kb) with only very few (< 5%) YACs larger

than 100 kb. In order to obtain larger YACs, size selection by pulsed field gel electrophoresis was included in the cloning protocol. Upon size selection both before and after ligation, only few clones were obtained per transformation experiment (25 to 100 clones per μg of ligated DNA). We, therefore, applied only size selection after ligation, as this has the advantage of removing the excess of vector DNA, resulting in a low number of white colonies and the absence of false positives (ic red colonies which do not contain a YAC). There was no difference in the average size of the YACs obtained from DNA ligations with size selection before or after ligation.

To protect the YACs from shearing in a liquid medium either polyamines [24, 28] or a high concentration of NaCl [25] can be used. The disadvantage of using polyamines is the apparent decrease in transformation efficiency [2, 24]. With polyamines in all steps but not during transformation, as described by Larin *et al* [24], we found a lower transformation efficiency (2 to 5 times less colonies) as compared to using NaCl. The additional presence of polyamines during transformation led to an even stronger decrease in transformation efficiency (up to 50 times less transformants using YCp50). We did not find any difference in the sizes of the YACs obtained by the two methods. Using the method described in the materials and methods section 1000 to 3000 clones per μg of starting DNA were obtained. Of the total number of 40,000 transformants, 6000 were picked and transferred to microtiter dishes for further analysis.

Virtually all YACs (> 98 %) appeared to contain both vector arms, as shown by hybridization with a probe specific for the centromeric arm (fig 1b), containing the TRP1-gene, and by growing the clones on double selective medium AHC [6]. Selection of the clones for the presence of both vector arms on double selective medium was, therefore, not necessary and the colonies were transferred directly to the wells of a microtiter dish containing YPD and allowed to grow for 16 hours at 30 °C. No YACs were lost during this step. In this manner 2000 clones could be easily picked per day by one person.

Characterization of the library

Of a subset of 125 clones, randomly chosen from the library, the sizes of the YACs were determined by Southern analysis of a CHEF gel of total chromosomal yeast DNA (fig 1a) using either pBR322 or a part of the centromeric arm of pYAC-4 as probes (an example of the latter is given in fig 1b). The average YAC size was 110 kb with a size distribution as shown in figure 2. Less than 5% of the clones in the wells of the microtiter plates had multiple YACs. We did not determine whether this was due to cocloning or picking two colonies into one well of the microtiter dish.

The library was screened for the presence of repetitive genomic sequences using as probe the dispersed *Hin*DIII repeat pTHG2 [41]. This repeat is present in a high copy number of

approximately 15.000 copies in the *L. esculentum* genome (on average one repeat per 30 kb [41]). Of the 1536 clones analyzed by colony filter hybridization (fig 3a) only 7.7 % of the clones hybridized to the probe pTHG2, far less than the expected 90-100%. The low number of YACs containing pTHG2 may be due to an uneven distribution of the repeated sequences or to nonrandom distribution of *Cla*I sites over the genome. A similar result was obtained by Martin *et al* [27] upon screening their *Eco*RI tomato YAC library with a comparable repeat (TGRII). They found that YACs up to 150 kb in size contained the TGRII repeat less frequently than expected. Interestingly, larger YACs contained the repeat at the expected frequency.

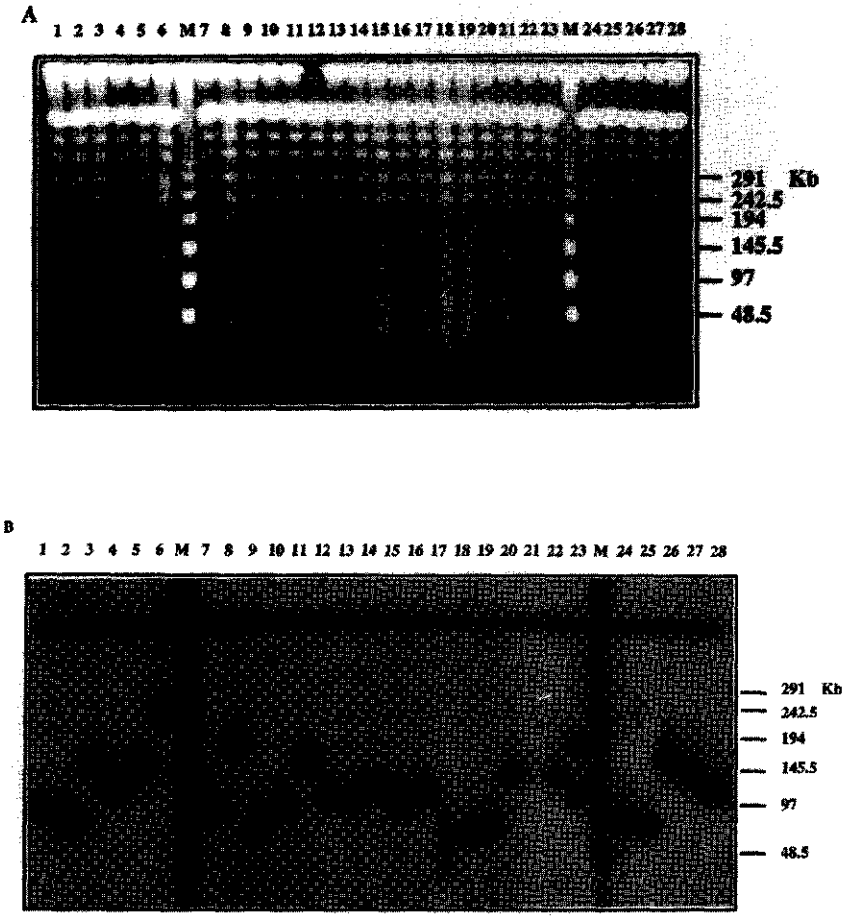


Figure 1. Chromosomal DNA from 28 random *Cla*I YAC clones. A. Ethidium bromide stained CHEF gel (25 hr run, 16 hr ramp of 1 - 20 s followed by 9 hr ramp of 30 - 40 s, at 185 V). B. Southern blot of the gel shown in A hybridized with a probe for the centromeric vector arm. Marker lanes, containing phage lamda concatemers, are denoted with M.

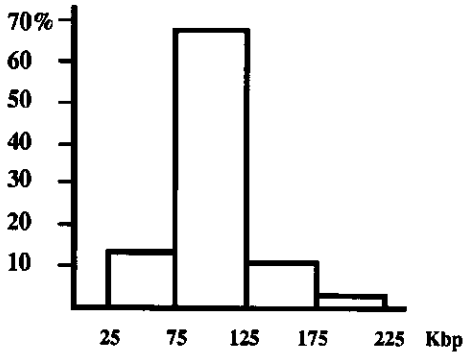
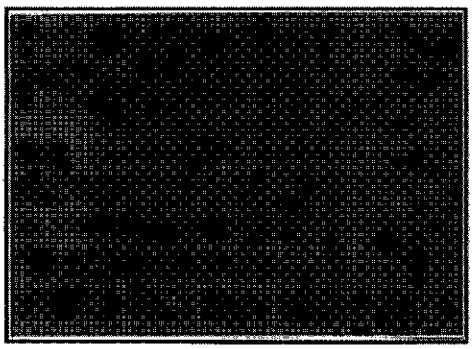
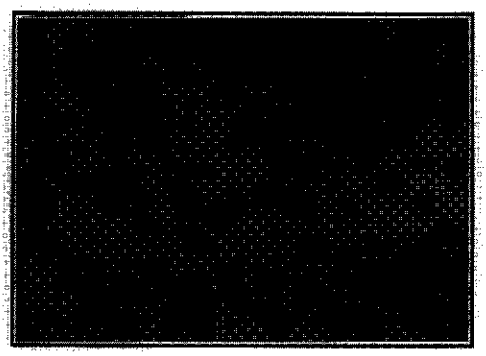


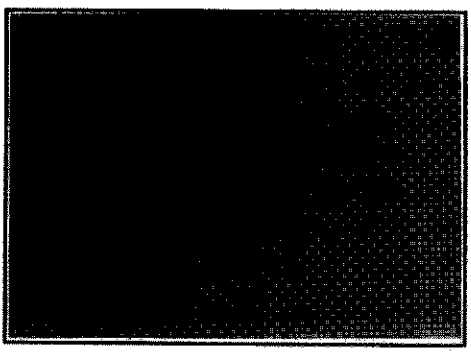
Figure 2. Size distribution (in percentage) of the clones in the *Cla*I YAC library. Results are based on size determinations of 125 individual clones.



A



B



C

Figure 3. High density colony filter hybridization of 1536 clones of the *Cla*I YAC library using (A) a genomic repeat probe (pTHG2), (B) a chloroplast DNA specific probe (Rubisco) and (C) the RFLP marker GP79, which identifies two positive clones.

To determine the fraction of chloroplast YACs the library was screened for the presence of Rubisco clones. Of the 1536 clones screened by high density colony filter hybridization, 125 clones, or approximately 8%, appeared to contain chloroplast DNA sequences (fig 3b). Since Rubisco detects only part of the chloroplast genome, this means that at least 8 % of all the clones represent YACs with chloroplast DNA.

Screening the library for GP79 and Aps-1

In order to screen the library for specific sequences, high density colony filters (8 x 12 cm) were prepared each containing clones from 16 microtiter dishes (1536 in total). The entire library was thus represented on only four filters. Using GP79 as a probe two positive colonies were detected (211-C12, 212-C12, as shown in figure 3). Upon colony purification of both positives, in which 10 individual clones were screened by analysis of chromosomal yeast DNA, only clone 212-C12 gave a hybridization signal for each of the 10 clones. The size of the YAC in all 10 clones was about 200 kb, corresponding to the largest genomic *Cla*I fragment containing the GP79 sequences ([37] see chapter 3, table 1). The validity of the YAC was further confirmed by comparing the restriction pattern of genomic and YAC DNA using seven different restriction enzymes (figure 4). All the enzymes used generated the same restriction fragments in both genomic and YAC-DNA, indicating that the YAC contains genuine GP79 sequences [37]. A restriction map of the YAC (figure 6) was generated using a set of rare-cutting restriction enzymes in partial digestions of the DNA. The location of the sites was determined in a Southern blot hybridization analysis using a probe for the centromeric arm of the YAC and a probe for GP79, examples of which are shown in figure 5. Comparison of the YAC-map with the genomic long range restriction map of GP79 [37] shows that all genomic restriction sites are recovered on the YAC, at the expected positions. This allows the positioning of the YAC on the genomic long range restriction map and, furthermore, indicates that the clone is unlikely to be chimeric. As expected, a large number of additional restriction sites was detected on the YAC due to its unmethylated character.

Screening the library for *Aps-1*^{1} was unsuccessful, both by colony filter hybridization as well as by PCR. Apparently no YAC containing *Aps-1*^{1} sequences was present in the library.

Discussion

The goal of our work was to isolate YAC clones, from a rare-cutter YAC library, that corresponded to the *Aps-1-Mi* region of tomato chromosome 6, taking advantage of the long range restriction maps available for two markers from this region [37]. To this end, genomic tomato DNA (cv VFNT Cherry) was digested to completion with *Cla*I and cloned in AB1380 using pYAC-RC. The enzyme *Cla*I was chosen because the GP79 containing fragments of 65, 135 and 205 kb in a complete digest, and the *Aps-1*^{1} fragment of 90 kb, fall in a size range which is easily cloned into YAC vectors. Even though the required YACs were relatively small, size selection of the ligated fragments was found to be necessary. Omission of the size selection resulted in YACs that were much smaller than 75 kb, a result that differs from several human libraries where YACs with an average size of 120 kb were generated [3]. On the other hand, Strauss *et al* [35] described the construction of a *Sal* YAC library without

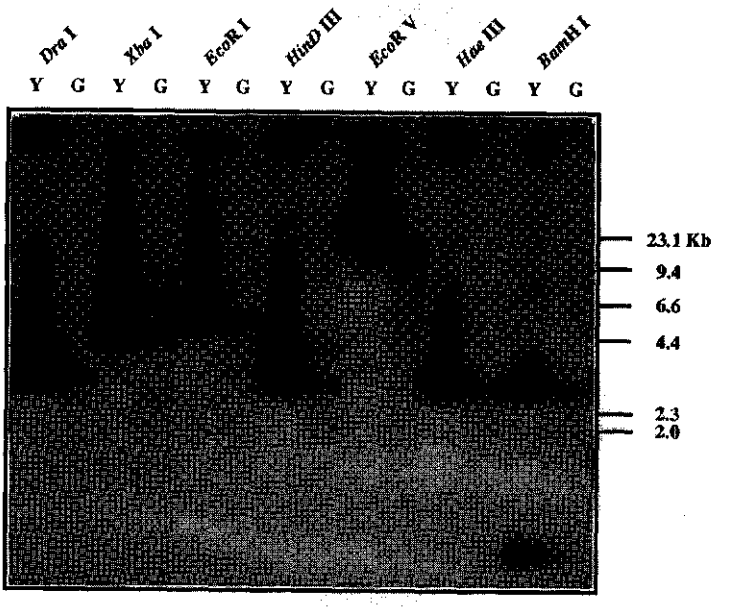


Figure 4. Southern blot hybridization analysis of GP79 YAC 212-C12 (Y) and genomic (G) DNA. DNA from both the GP79 YAC 212-C12 and genomic DNA (cv VFNT-Cherry) was digested with a set of enzymes as indicated. The observed restriction fragments for the YAC and genomic DNA are identical, indicating that a genuine GP79-containing YAC was isolated.

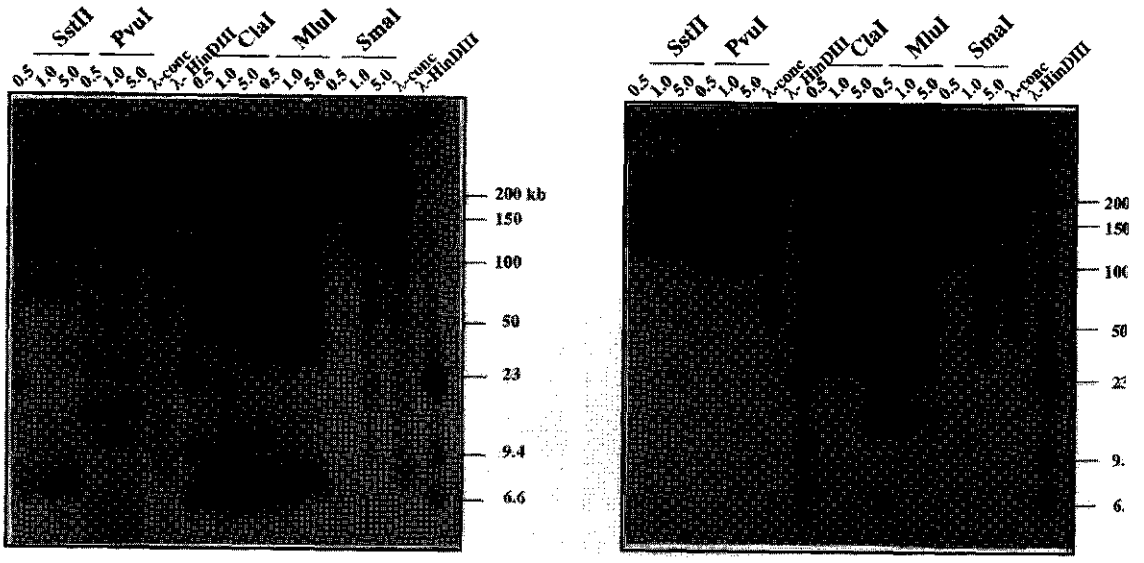


Figure 5. Southern blot of partially digested YAC 212-C12, using the enzymes as indicated, following separation on a CHEF gel (16 hr run, timeramp 1-10 s, 185 V). In A the hybidization results with a centromeric arm specific probe are shown and in B the results for GP79.

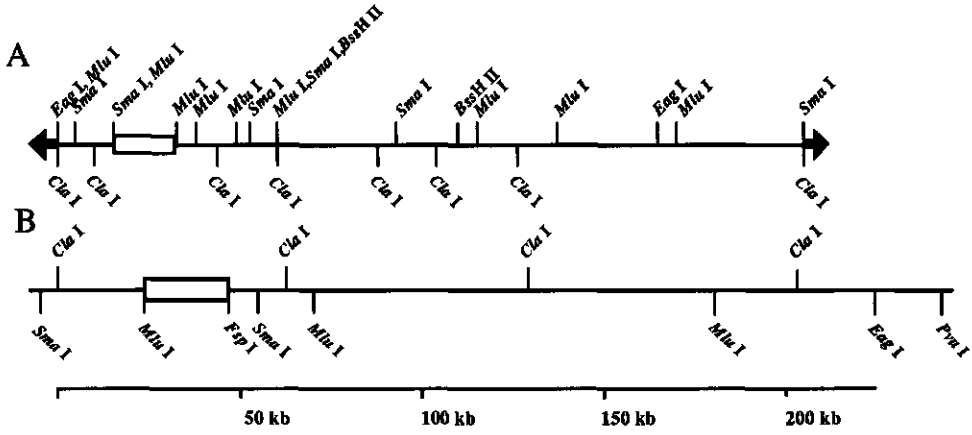


Figure 6. Restriction map of YAC 212-C12 (A) compared to the genomic map (B). The position of GP79 is indicated with a open box and vector sequences with an arrow. The centromeric arm is positioned here on the left side

size selection with most YACs (well over 55%) smaller than 75 kb. Besides, chloroplast fragments in the size range up to 50 kb are present in large amounts in a *Cla*I digest and these also transform very efficiently. Thus, to prevent the cloning of many small YACs, including the chloroplast YACs, a size selection was applied. The size distribution of the YACs (figure 2) indicates that this step was efficient since more than 65% of the YACs in the library were in the range of 75 - 125 kb. The fact that 40,000 clones were obtained from a starting amount of DNA of only 25 µg, indicates that the procedure is highly efficient.

Upon screening the library with the RFLP probe GP79, one YAC clone was detected that corresponded to the 205 kb genomic fragment and appeared to contain the restriction sites as expected on basis of the genomic map. This shows that, in principle, the procedure of isolating clones from a rare-cutter library designed on the basis of a restriction map is feasible. Surprisingly, however, no *Aps-1*[1] containing clones were present in the library. As *Cla*I fragments in the range from 75 to 125 kb represent less than 10 % of all fragments in a genomic digest, as determined from several ethidium bromide stained CHEF gels of complete *Cla*I digests we estimate that the *Cla*I YAC library library, of 6000 clones, represents 3 to 5 genome equivalents. So either we were unlucky, or Aps-1 containing YACs are unstable.

Since in tomato 85% of all CpNpG and 55% of all CpG sequences are methylated at their C-residue [30] we expected to find many additional restriction sites on the YAC for the rare-cutting restriction enzymes used. Indeed, for the enzymes *Cla*I, *Mlu*I and *Sma*I two, three and six times as many sites were present, respectively, in YAC DNA as compared to genomic plant DNA.

DNA preparations from pools of clones from the library have been screened by PCR for the tomato sequences TG359, TG242, TG284 and Cab11 [35b], by the group of Hille (VU Amsterdam The Netherlands). They identified several positive clones indicating that our *Cla*I library provides a valuable supplement to the existing *Eco*RI library constructed by Martin *et al* [27].

In conclusion, the results presented show that library construction based on long range restriction maps provides an alternative to the construction of whole genome YAC libraries, especially when clones from specific genotypes are needed for which no library is available.

Acknowledgements

We thank Edwin Cuppen for part of the experimental work described in this chapter.

References

1. Aarts JMMJG, , Hontelez JGJ, Fischer P, Verkerk R, van Kammen A, Zabel P (1991) Acid phosphatase-1^1, a tightly linked molecular marker for root-knot nematode resistance in tomato: from protein to gene, using PCR and degenerate primers containing deoxyinosine. Plant Mol Biol 16: 647-61.
2. Albertsen HM, Abderrahim H, Cann HM, Dausset J, Le Paslier D, Cohen D (1990) Construction and characterization of a yeast artificial chromosome library containing seven haploid human genome equivalents. Proc. Natl. Acad. Sci. USA 87: 4256-4260.
3. Anand R, Villasante A, Tyler-Smith C (1989) Construction of yeast artificial chromosome libraries with large inserts using fractionation by pulsed field gel electrophoresis. Nucleic Acids Res 17: 3425-3433.
4. Arondel V, Lemieux B, Hwang I, Gibson S, Goodman H, Somerville CR (1992) Map-based cloning of a gene controlling omega-3 fatty acid desaturation in *Arabidopsis*. Science 258: 1353-1355.

4b. Ashikawa I, Fujimura T, Katayose Y, Miyao A, Miyazaki A, Momma T, Saji S, Tanoue H, Umehara Y, Yasukochi Y, Yoshino K, Kurata N (1992) Physical mapping and chromosome mapping. Rice Genome 1: 8-9.

5. Bates GP, MacDonald ME, Baxendale S, Sedlacek Z, Youngman S, Romano D, Whaley WL, Allitto BA, Poustka A, Gusella JF, Lehrach H (1990) A yeast artificial chromosome telomere clone spanning a possible location of the Huntington disease gene. Am J Hum Genet 46: 762-775.
6. Brownstein BH, Silverman GA, Little RD, Burke DT, Korsmeyer SJ, Schlessinger D, Olson MV (1989) Isolation of single-copy human genes from a library of yeast artificial chromosome clones. Science 244: 1348-1351.
7. Burgers PMJ, Percival KJ (1987) Transformation of yeast spheroplasts without cell fusion. Anal Biochem 163: 391-397.
8. Burke DT, Carle GF, Olson MV (1987) Cloning of large segments of exogeneous DNA into yeast by means of artificial chromosome vectors. Science 236: 806-812.
9. Chu G, Vollrath D, Davis RW (1986) Separation of large DNA molecules by contour-clamped homogeneous electric fields. Science 234: 1582-1585.
10. Dunford R, Rogner U (1991) Construction of yeast artificial chromosomes containing barley DNA and the identification of clones carrying copies of the repeated element BIS-1. Hereditas 115: 133-138.
11. Edwards K, Thompson H, Edwards D, de Saizieu A, Sparks C (1992) Construction and characterisation of a yeast artificial chromosome library containing three haploid maize genome equivalents. Plant Mol Biol 19: 299-308.
12. Eyers M, Edwards K, Schuch W (1992) Construction and characterization of a yeast artificial chromosome library containing two haploid *Beta vulgaris* L. genome equivalents. Gene 121: 195-201.
13. Feinberg AP, Vogelstein B. (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 132: 6-10

14. Gebhardt C, Ritter E, Barone A, Debener T, Walkemeier B, Schachtschabel U, Kaufman H, Thompson RD, Bonierbale MW, Ganal MW, Tanksley SD, Salamini F (1991) RFLP maps of potato and their alignment with the homoeologous tomato genome. Theor Appl Genet 83: 49-57.
15. Gilbert JC (1958) Some linkage studies with the *Mi* gene for resistance to root knot. Tomato Genet Coop Rep 8: 15-17.
16. Giraudat J, Hauge BM, Valon C, Smalle J, Parcy F, Goodman HM (1992) Isolation of the Arabidopsis *ABI3* gene by positional cloning. The Plant Cell 4: 1251-1261.
17. Grill E, 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.
18. Guzman P, Ecker JR (1988) Development of large DNA methods for plants: molecular cloning of large segments of *Arabidopsis* and carrot DNA into yeast. Nucleic Acids Res 16: 11091-11105.
19. Hoffman CS, Winston F (1987) A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. Gene 57: 267-272.
20. Ho JY, Weide R, van Wordragen MF, Lambert K, Koornneef M, Zabel P, Williamson VM (1992) The root-knot nematode resistance gene (*Mi*) in tomato: construction of a molecular linkage map and identification of dominant cDNA markers in resistant genotypes. The Plant J 2: 971-982.
21. Hwang I, Kohchi T, Hauge B, Goodman HM, Schmidt R, Cnops G, Dean C, Gibson S, Iba K, Lemieux B, Arondel V, Danhoff L, Somerville C (1991) Identification and map position of YAC clones comprising one-third of the *Arabidopsis* genome. The Plant J 1: 367-374.
22. Klein-Lankhorst R, Rietveld P, Machiels B, Verkerk R, Weide R, Gebhardt C, Koornneef M, Zabel P (1991) RFLP markers linked to the root knot nematode resistance gene *Mi* in tomato. Theor Appl Genet 81: 661-667.
23. Klein-Lankhorst RM, Vermunt A, Weide R, Liharska T, Zabel P (1991) Isolation of molecular markers for tomato (*L. esculentum*) using random amplified polymorphic DNA (RAPD). Theor Appl Genet 83:108-114.
24. Larin Z, Monaco A, Lehrach H (1991) Yeast artificial chromosome libraries containing large inserts from mouse and human DNA. Proc Natl Acad Sci USA 88: 4123-4127.
25. Lee JT, Murgia A, Sosnoski DM, Olivos I, Nussbaum RL (1992) Construction and characterization of a yeast artificial chromosome library for Xpter-Xq27.3: A systematic determination of cocloning rate and X-chromosome representation. Genomics 12: 526-533.
26. Marchuk D, Collins FS (1988) pYAC-RC a yeast artificial chromosome vector for cloning DNA cut with infrequently cutting restriction endonucleases. Nucleic Acids Res 16: 7743.
27. Martin GB, Ganal MW, Tanksley SD (1992) Construction of a yeast artificial chromosome library of tomato and identification of cloned segments linked to two disease resistance loci. Mol Gen Genet 233: 25-32.
27b. Martin G, Brommonschenkel SH, Chungwongse J, Frary A, Ganal M, Spivey R, Wu T, Earle ED, Tanksley SD (1993) Map-based cloning of a protein kinase gene conferring disease resistance in tomato. Science 262: 1432-1436.
28. McCormick MK, Shero JH, Cheung MC, Kan YW, Hieter PA, Antonarakis SE (1989) Construction of human chromosome 21-specific yeast artificial chromosomes. Proc. Natl. Acad. Sci. USA 86: 9991-9995.
29. Messenguer R, Ganal M, de Vicente MC, Young ND, Bolkan H, Tanksley SD (1991) High resolution RFLP map around the root-knot nematode resistance gene (*Mi*) in tomato. Theor Appl Genet 82: 529-536.
30. Messenguer R, Ganal MW, Steffens JC, Tanksley SD (1991) Characterization of the level, target sites and inheritance of cytosine methylation in tomato nuclear DNA. Plant Mol Biol 16: 753-770.
31. Rick CM, Fobes JA (1974) Association of an allozyme with nematode resistance. Rep Tomato Genet Coop 24: 25.
32. Rommens JM, Ianuzzi MC, Kerem B, Drumm ML, Melmer G, Dean M, Rozmahel R, Cole JL, Kennedy D, Hidaka N, Zsiga M, Buchwald M, Riordan JR, Tsui LC, Collins FS (1989) Identification of the cystic fibrosis gene: chromosome walking and jumping. Science 245: 1059-1065.
33. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning, a laboratory manual. Cold Spring Harbor laboratory press, Cold Spring Harbor, New York.
34. Sherman F, Fink GR, Lawrence CW (1974) Methods in yeast genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
35. Strauss WM, Jaenisch E, Jaenisch R (1992) A strategy for rapid production and screening of yeast artificial chromosome libraries. Mammalian genome 2: 150-157.
35b. Tanksley SD, Ganal MW, Prince JP, de Vincente MC, Bonierbale MW, Broun P, Fulton TM, Giovanonni JJ, Grandillo S, Martin GB, Messeguer R, Miller JC, Miller L, Paterson AH, Pineda O, Röder MS, Wing RA, Wu W, Young ND (1992) High density molecular linkage maps of the tomato and potato genomes. Genetics 132: 1141-1160.

36. van Daelen RAJJ, Zabel P (1991) Preparation of high molecular weight plant DNA and analysis by pulsed-field gel electrophoresis. In: Gelvin SB, Schilperoort RA (Eds) Plant Molecular Biology Manual pp A15: 1-25 Kluwer, Dordrecht.
37. van Daelen RAJJ, Gerbens F, van Ruisen F, Aarts J, Hontelez J, Zabel P (1993) Long range physical maps of two loci (*Aps-1* and GP79) flanking the root-knot nematode resistance gene (*Mi*) near the centromere of tomato chromosome 6. Plant Mol Biol 23: 185-192.
38. van Grinsven MQJM, Gielen JJL, Zethof JLA, Nijkamp HJJ, Kool AJ (1986) Transcriptional and post-transcriptional regulation of chloroplast gene expression in *Petunia hybrida*. Theor Appl Genet 73: 94-101.
39. Wallace MR, Marchuck DA, Andersen LB, Letcher R, Odeh HM, Saulino AM, Fountain JW, Brereton A, Nicholson J, Mitchell AL, Brownstein BH, Collins FS (1990) Type 1 neurofobromatosis gene: identification of a large transcript disrupted in three NF1 patients. Science: 249: 181-186.
40. Ward ER, Jen GC (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.
41. Zabel P, Meyer D, van der Stolpe O, van der Zaal B, Ramanna MS, Koornneef M, Krens F, Hille J (1985) Towards the construction of artificial chromosomes of tomato. In: van Vloten-Doting L, Groot GSP, Hall TC (Eds) Molecular form and function of the plant genome. NATO ASI Series, Plenum New York, pp 609-624.

CHAPTER 6

CONSTRUCTION AND SCREENING OF A YAC LIBRARY FROM PARTIALLY *ECO*RI DIGESTED TOMATO DNA.

Introduction

Currently, several markers linked to the nematode resistance gene *Mi* have been identified [11, 12, 9], including the acid phosphatase gene (*Aps-1*, [7]) and the potato derived clone GP79 [11] at either side of *Mi*. These markers function as anchors in the physical characterization of the *Mi*-region. In the previous chapter a procedure is described to isolate the sequences between these markers as YAC clones. In that case, long range restriction maps of the *Aps-1* and GP79 regions were made, prior to the isolation of the YAC clones. Based upon these restriction maps, a rare-cutter restriction enzyme was chosen for the complete digestion of tomato DNA. The restricted DNA was subsequently cloned into YAC vectors and a segment specific for the GP79 region was isolated. In this chapter an alternative approach for YAC library construction is described. Tomato DNA is partially digested with *Eco*RI to generate a set of overlapping DNA fragments which are subsequently cloned into YAC vectors. After isolation of overlapping YAC clones, specific for the *Mi*-region, these can be used to generate a restriction map of the region. Here the construction of a YAC library, based on a partial *Eco*RI digest, and the isolation and characterization of an *Aps-1*^{1} containing YAC is described.

Materials and methods

Standard DNA methodology

All standard DNA manipulations were carried out as described in Sambrook *et al* [14]. High molecular weight DNA isolation, restriction digestion, pulsed field gel electrophoresis (PFGE) and hybridization of Southern blots, are described in chapter 2 and [15]. Chromosomal yeast DNA was isolated as described (chapter 5). Probes were labeled according to the "random primer" method of Feinberg and Vogelstein [6]

Partial restriction digestion of high molecular weight DNA

Prior to restriction digestion, the agarose plugs (100 µl) were washed in $T_{10}E_{10}$ (10 mM TRIS-HCl pH 8.0 and 10 mM EDTA) containing phenylmethylsulfonylfluoride (PMSF) as described previously [15]. After washing, the plugs were cut into 6 pieces (circa 15 µl) each containing about 1 µg of DNA. These were equilibrated (twice for 30 min on ice) in digestion buffer consisting of 100 mM TRIS-HCl 8.0, 100 mM NaCl, 2 mM $MgCl_2$, 1 mM DTT, 500 µg/ml BSA, 2.6 mM spermidine trihydrochloride and 80 µM S-Adenosylmethionine. After

equilibration, 100 µl of fresh buffer containing 1 u *Eco*RI and varying amounts of *Eco*RI methylase (M. *Eco*RI), was added. Usually two or three concentrations of methylase, as determined from a titration using a range of methylase from 10 to 80 units, were used. The plugs were kept on ice for 30 min and subsequently incubated at 37 °C for 4 hours. Reactions were terminated by adding EDTA and proteinase-K to final concentrations of 20 mM and 0.25 µg/ml respectively, and incubating at 37 °C for 30 min. The degree of partial digestion was determined by PFGE analysis of the DNA in a few plugs. The remainder of the plugs, in which the bulk of DNA was larger than 200 kb, were pooled and washed (twice for 30 min at 37 °C) with 50 ml $T_{10}E_{10}$ containing 0.25 mM PMSF. For large scale isolations, larger plugs (100 µl) were used and all incubation volumes were increased accordingly, while maintaining the same concentrations.

Ligation of vector arms to genomic DNA

Plugs containing the partially digested DNA, 10 - 20 µg in total, were equilibrated with 10 ml ligation buffer (50 mM TRIS-HCl, pH 7.5, 10 mM $MgCl_2$, 30 mM NaCl, 0.75 mM spermidine trihydrochloride and 0.30 mM spermine tetrahydrochloride) at 4 °C. Vector DNA, 100 - 150 µg of pYAC4 [4], digested with *Eco*RI and *Bam*HI and dephosphorylated, was added to the plugs. The plugs were melted at 68 °C for about 5 min, until they were completely liquified and subsequently cooled to 37 °C. Ligase (50 units), in 100 µL ligation buffer supplemented with ATP and DTT to final concentrations of 1 mM each in the ligation mixture, was added. The molten mixture was not stirred and kept at 37 °C for 2 to 4 hours, after which it was pipetted into a precooled mold to make plugs of 100 µl. These plugs were further incubated in 1 ml ligation mixture at 15 °C for 16 hours.

Size selection by PFGE

The plugs with ligated DNA were placed in two rows into a large slot of a 1% (w/v) LMP agarose gel, made in 0.5 x TBE. The gel was transferred to a CHEF apparatus (BIORAD DR II) and subjected to electrophoresis. Gels were run for 18 hours at 140 volts and 20 s pulsetimes to obtain a compression zone (CZ) at 300 kb. For CZ's at 400 and 500 kb, gels were run at 140 V with 25 s pulsetimes and at 160 V with 50 s pulsetimes, respectively. The compression zone was identified and excised as described by Larin *et al* [13]. A large agarose block of 200 to 500 µl was obtained and stored at 4 °C in 0.5 M EDTA until use.

Transformation of the yeast host AB1380

The yeast host strain AB1380 [4] was used for transformation following the protocol of Burgers and Percival [3]. A part (± 20 µl) of the agarose block containing the size-selected ligated YAC DNA was equilibrated for 1 hr at 4°C in 50 ml TE (10 mM TRIS-HCL pH 8.0, 1 mM EDTA) containing polyamines [13]. The plug was melted at 68 °C, cooled to 40 °C and incubated with agarase (2 u, NEB) for 4 hours. To the liquified sample 10 µg of salmon sperm DNA in a volume of 1 µl and spheroplasts (100 µl containing 5 x 10^7 cells) were added. The mixture was incubated for 10 min at room temperature after which 1 ml of a 20 % (w/v) PEG-8000 solution made in 10 mM TRIS-HCl pH 7.5 and 10 mM $CaCl_2$, was added. The incubation was continued for another 10 minutes and than cells were collected by centrifugation and taken up for recovery in SOS-medium (1 M Sorbitol, 0.25 % (v/v) YPD, 6.5 mM $CaCl_2$, 10 µg/ml tryptophane, 1 µg/ml uracil). After recovery for 1 hour at 30 °C cells were plated out on selective medium (lacking uracil) with limiting amounts af adenine (2.5 µg/ml adeninehemisulfate). After growth for 5 days at 30 °C red colonies were picked.

Growth and storage of transformants

The red colonies were transferred into the wells of microtiter dishes containing YPD (75 µl per well). The plates were incubated for 1 day at 30 °C with vigorous shaking, then supplemented with 50 µl of glycerol (50%) and 40 µl of 4 times concentrated YPD per well and stored at -30 °C for 2 to 5 days followed by long term storage at -80 °C.

PCR screening of the library

In order to screen the library, clones were grown for 2 days at 30 °C in the wells of microtiter dishes containing 50 µl of YPD. After growth, the cells were pooled per plate and, from each plate, all cells per row and column were pooled. DNA was isolated from the pooled cells according to the protocol of Hoffman and Winston [10]. For a PCR reaction approximately 50 ng of DNA was used. The PCR mixture contained: 10 mM TRIS-HCl 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.01 % gelatin, 100 µM dNTP's, 2 ng/µL primers and 0.4 units Supertaq DNA polymerase in a total volume of 50 µL. The amplification conditions were: an initial denaturation for 5 min at 94 °C followed by 30 cycles of 1 min at 94 °C, 1 min at 60 °C, and 1 min at 72 °C and finally 10 min incubation at 72 °C. Reaction products were analyzed on a 1.5 % agarose gel run in TAE [14]. The *Aps-1*[1] primers were 5'-ATGGTGGGTCCAG-GTTATAAG and 5'-CAGAATGAGCTTCTGCCAATC.

Probes.

Gel purified inserts from the plasmid pTZ18R containing part of the *Aps-1* gene [16] and the genomic *Hin*dIII repeat (pTHG2 [19]) were used as probes. A probe for the large subunit of Rubisco [17] was used for detection of chloroplast DNA.

Results and discussion

Isolation of high molecular weight DNA.

The construction of large YACs requires source DNA of very high (> 2 Mb) molecular weight. To obtain a high percentage of YAC clones with an average size of 250 kb the initial size of the DNA should be larger than 1500 kb and to obtain a high percentage of YACs of 400 kb an initial size of 3000 kb is required [14]. If the DNA is of low molecular weight (less than 1000 kb) poor partial digests (see next section) and many false YAC clones, which grow on regeneration plates but fail to do so on double selective medium, will be obtained.

In a good batch of source DNA the bulk of the molecules remains within the slot of a pulsed field gel (PFG), if separation conditions are used that allow all yeast chromosomes to be resolved. Some degradation during DNA isolation is unavoidable resulting in a very slight staining of the gel at the compression zone and sometimes at the 100 kb region (see fig 1A in chapter 2). To obtain high quality DNA, vital tomato protoplasts must be used for embedding in agarose. For this purpose the protoplasts should be isolated from *in vitro* tomato shoot cultures. The growth conditions of the plants from which protoplasts are isolated can then be carefully controlled which is very important as the quality of the protoplasts is greatly determined by the conditions of the plants. A very important phenomenon in protoplast isolations, to which I will refer to as "wintereffect", is that during the winterseason protoplast isolations yield very low numbers of protoplasts which often are of poor quality and produce DNA of inferior quality. The "wintereffect" cannot be overcome by using *in vitro* shoot cultures and the season suitable for isolating DNA is therefore roughly limited from April to November.

Determination of viability of protoplasts was initially done by staining vital cells with FDA (fluoresceïnediacetate) [18]. In later experiments other characteristics proved equally effective. Usually the quality of protoplasts was very good, if the leaves were almost completely digested by incubation with cellulase and macerozyme. The resulting protoplast suspension contained only a small amount of debris, the shape of the protoplasts was globular with chloroplasts evenly distributed over the entire cell volume, and the number of protoplasts obtained from a fixed amount of tissue was high.

Preparation of partially digested tomato DNA

Several protocols for partial restriction digestion have been described (see section 4.2) but in our hands the use of a combination of the restriction endonuclease (*Eco*RI) with its methylating counterpart (M. *Eco*RI) [13] gave the most reproducible results and was the least sensitive to the amount or concentration of the DNA. Prior to the preparative restriction digestion a sample of each batch of DNA was subjected to digestion using increasing amonts of methylase, at a constant amount of restriction enzyme, and assayed by CHEF gel analysis. Apart from providing the required amounts of methylase for preparative digestion, the titration was also used to determine the quality of the DNA in the batch. In figure 1 the results of two such tests are shown. Digestion of a DNA batch shown in figure 1B with 50 u of M. *Eco*RI resulted in a large amount of DNA visible as a smear in the 100 - 600 kb range in the ethidium bromide stained PFG gel. Increasing the amount of methylase up to 200 u did not improve the size distribution as a similar smear of fragments was obtained (not shown), indicating the presence of a large proportion of degraded DNA. Using DNA of this quality in cloning experiments resulted in many false YAC clones, that grow on regeneration plates but fail to do so when they are transferred to double selective plates. On the other hand digestion of a different batch of DNA (figure 1A), using 50 u of methylase, yielded DNA remaining mostly within the compression zone and slot, indicating DNA molecules of a very large size (figure 1A, compare the lanes showing the use of 50 u methylase). For cloning experiments using DNA as shown in figure 1A approximately 25 u of M. *Eco*RI was used to prepare partial digestions of DNA.

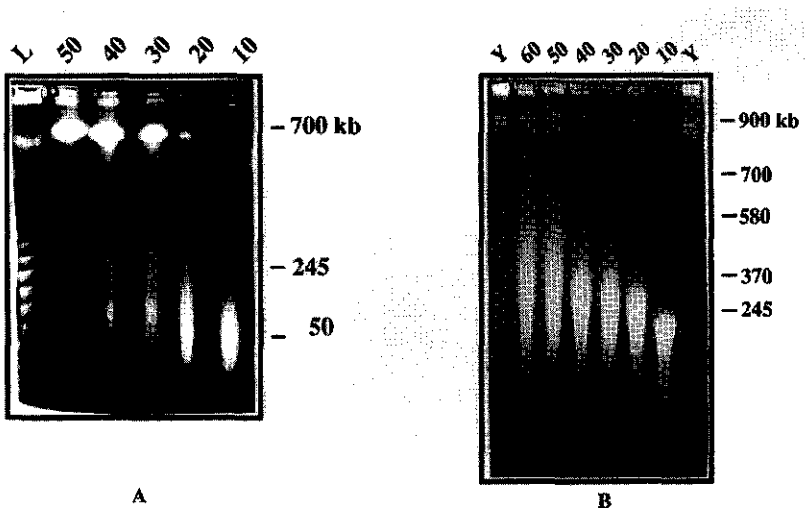


Figure 1. Partial digestion of tomato DNA using a combination of the restriction enzyme *Eco*RI and its methylating counterpart M. *Eco*RI. The amount of *Eco*RI was kept constant at 1 unit and the amount of M. *Eco*RI varied (as indicated above the lanes in the figures). In left panel A, a igestion of a good quality DNA batch is shown, and in the right panel B, the digestion of a poor quality DNA batch is represented.

TABLE 1. Number of clones and the average YAC sizes obtained from DNA size selected for various molecular weight classes. The results were based on transformation of 2 ligation mixtures for the 300 and 400 kb and on one ligation mixture using DNA from the 500 kb compression zone.

Size selection range (CZ) size[1] (kb)	number of clones obtained per µg starting DNAs	average YAC (kb)
-	> 4000	< 75 kb
300 kb	300 - 500	145
400	200 - 400	150
500	200 - 400	160

[1]: Determined from at least 60 clones.

Size selection

As has been reported extensively in the literature (see chapter 4.2), size selection of the partially digested DNA used for transformation is crucial to obtain YACs with an average insert size of more than 100 kb. This is clearly illustrated in figure 2A in which a sample of 28 YAC clones, obtained upon transformation with a ligation mixture to which no size selection was applied, was analyzed by PFGE and Southern blot hybridization analysis. Almost all YACs were less than 100 kb and too small for a YAC library to be used in chromosomal walking. Applying a size selection at either 300, 400 or 500 kb resulted in much larger YACs with average sizes of 150 kb (table 1). Unfortunately, the size selection led to a loss in the total number of transformants by a factor 10 or more (see Table 1). Without size selection tens of thousands of clones were easily obtained from a few µg of starting DNA whereas only a few hundreds were obtained after size selection.

Remarkably, there was no large difference between the average size of YACs obtained from 300, 400 or 500 kb compression zones (see table 1 and figures 2, 3 and 4). Apparently, the DNA isolated from the compression zone and used for transformation still contains smaller YACs trapped in the agarose. Taking into account the observations of a much higher transformation efficiency for smaller YACs made by Connely *et al* [5] and Albertsen *et al* [1], it appears that the contamination with the smaller, trapped YACs partly undoes the result of the relative enrichment for larger YACs of 300, 400 and 500 kb in the compressionzones. Alternatively, the low number of transformants with large YACs may be due to shearing of the large DNA fragments eluted from the compression zone. If an additional size selection was applied by fractionating the DNA, prior to ligation to the vector arms, much larger YACs were obtained [13]. In our hands, however, virtually no transformants were obtained after an additional size selection. This was possibly due to a low DNA concentration and we did not further persue this strategy.

Characterization of the library

All the transformation experiments added up to a small library of approximately 3000 clones with an average insert size between 150 and 180 kb (see table 1). This library is thus incomplete and contains approximately 0.5 genome equivalents. A subset of 90 clones from the library was screened for the presence of chloroplast YACs using the Rubisco probe (results not shown), which resulted in the detection of only 4 clones. Although Rubisco may not detect all the YACs containing chloroplast DNA, the result indicates that less than 10% of the clones in the library

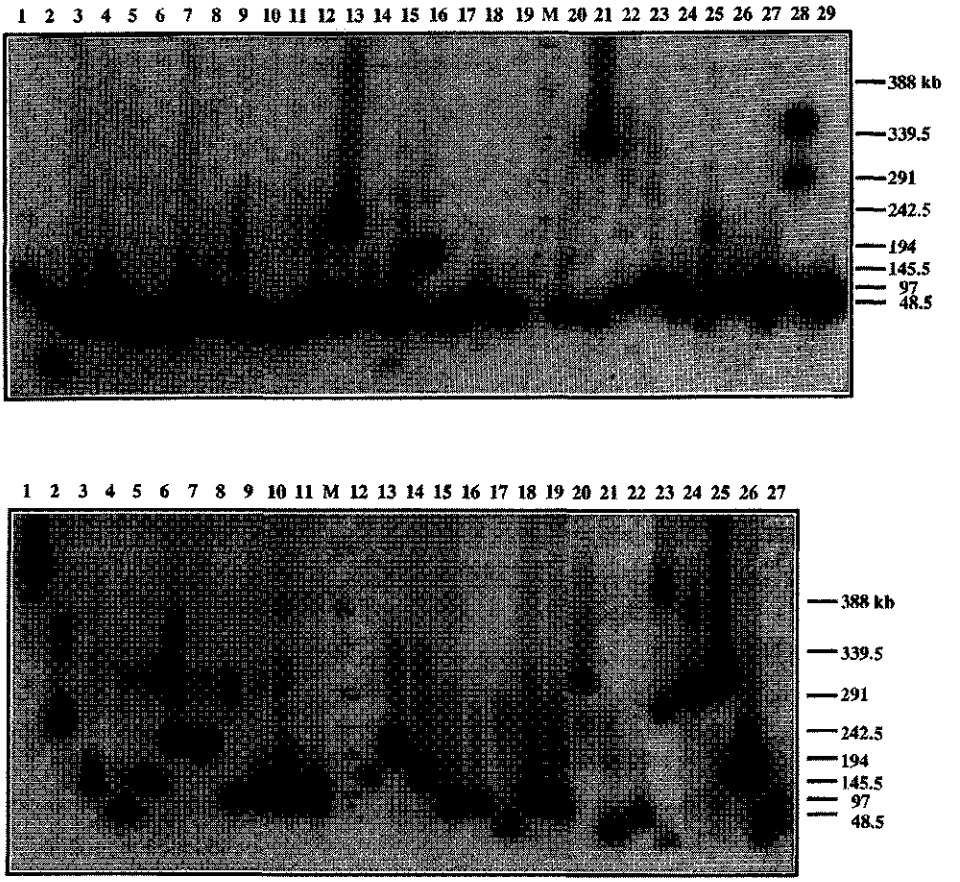


Figure 2. Chromosomal DNA from random *Eco*RI YAC clones. A. Southern blot of a CHEF gel (25 hr run, 16 hr ramp of 1 - 20 s followed by 9 hr ramp of 30 - 40 s, at 185 V) hybridized with pBR322 showing the presence of the YACs. The DNA used for transformation was not size selected, which resulted in very small YACs. Due to the high density of colonies on the regeneration plates, it was not possible to pick individual clones and thus many lanes contain two or even three or four copicked clones. B. Southern blot of a CHEF gel (18 hr run, 20 s pulsetimes, at 140 V) hybridized with pBR322. The DNA used for transformation was size selected for 300 kb.

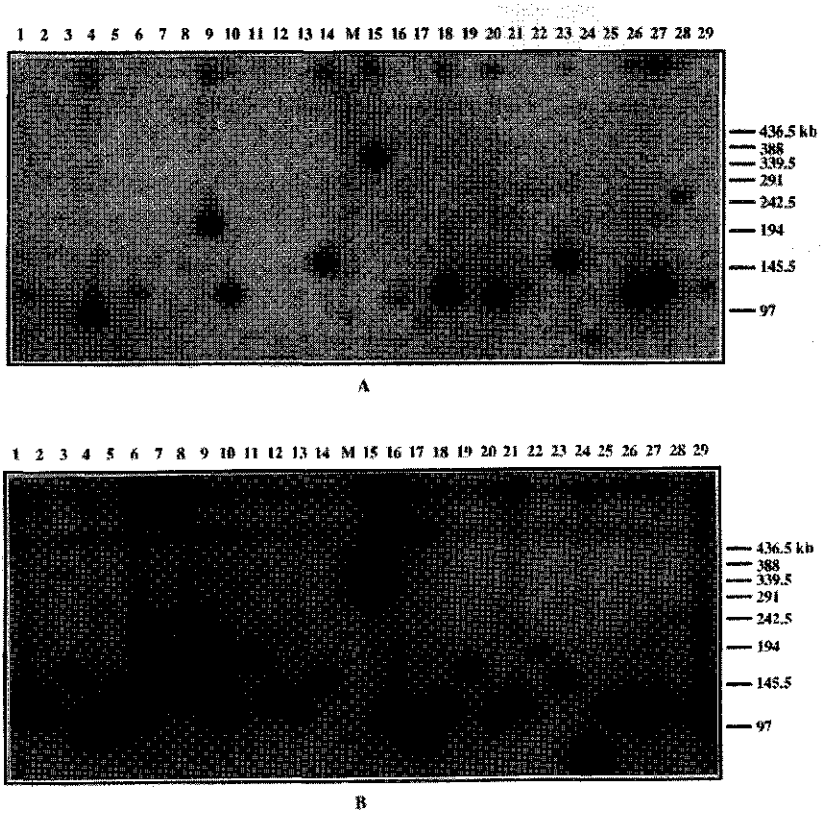


Figure 3. Southern blot hybridization analysis of chromosomal DNA of YAC clones. The blot of a CHEF gel (22 hr run, 185 V, 1 - 20 s ramp for 16 hrs followed by 30 s pulsetimes for 6 hrs) hybridized with (A) the genomic repeat probe pTHG2, indicating the presence of genomic tomato DNA and (B) with pBR322. The ligated DNA used for transformation was size selected for 400 kb.

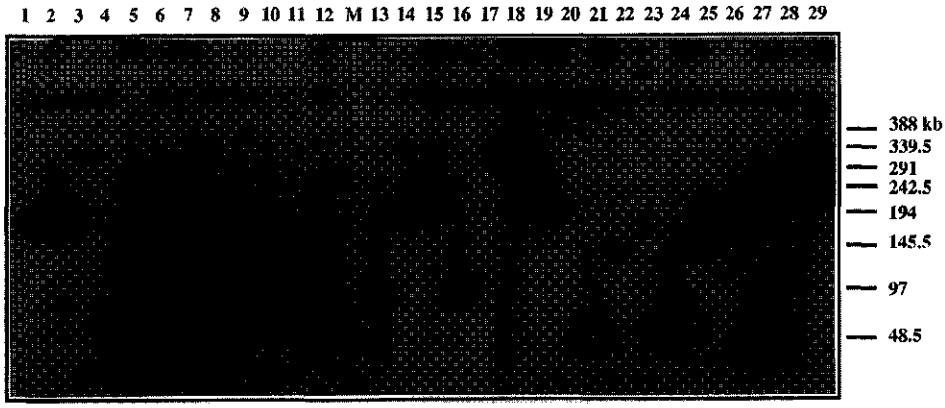


Figure 4. Southern blot hybridization analysis of chromosomal DNA from random *Eco*RI YAC clones. The Southern blot of a CHEF gel (25 hr run, 16 hr ramp of 1 - 20 s followed by 9 hr ramp of 30 - 40 s, at 185 V) was hybridized with a probe for the right YAC vector arm (trp/cen arm) showing the presence of the YACs. The DNA was transferred to the yeast strain AB1380 after size selection of the ligation mixture at a cut-off size of 500 kb.

contain chloroplast DNA, which is similar to most of the plant YAC libraries described so far (see chapter 4).

Part of the library was screened for tomato genomic sequences using the dispersed genomic repeat pTHG2 as a probe (figure 3). This *Hin*dIII repeat is present in about 30,000 copies in the tomato genome [41] and expected to recognize the majority of the YACs. As shown in figure 3 several YACs contained the repeat sequence, indicating the presence of genomic tomato DNA. Only twenty clones, out of 58 tested, were found to contain the repeat sequence which is less than expected. Possibly, the repeat is not evenly dispersed over the genome. Alternatively, the tested sample from the library may have been too small.

Screening for the single copy marker Aps-1^1 and characterization of an Aps-1^1 YAC clone

Part of the library was screened for the presence of *Aps-1^1* sequences by PCR. DNA was isolated from a total of 768 YAC clones (8 microtiter plates) and tested in 8 pools of 96 clones each. On each of the 8 DNA pools a PCR for *Aps-1^1* was performed and one positive plate (no 134) was identified. The position of the clone among the 96 clones from the plate was determined by 20 PCR reactions on DNA pools originating from the rows and columns and was found to be A1 (see fig 5). The validity of the isolated YAC was confirmed by comparing restriction fragment patterns of YAC and total genomic DNA by Southern blot analysis using an *Aps-1* probe (fig 6). The similar restriction patterns for 6 restriction enzymes indicated that indeed a *Aps-1^1* containing YAC had been isolated. The size of the YAC was small, 45 kb, which amounts to an insert of only 35 kb of genomic DNA and will therefore not be very useful for a genomic walk. For further characterization the YAC was cut with the following rare-cutter restriction enzymes: *Cla*I, *Bss*HII, *Mlu*I, *Sst*II, *Eag*I, *Pvu*I and *Sma*I and analyzed

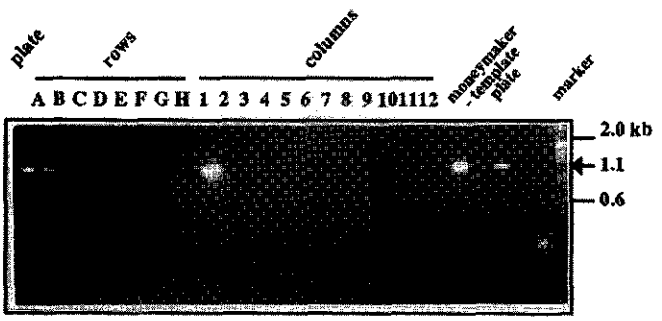


Figure 5. PCR screening of part of the YAC library. DNA from the identified positive plate was pooled by rows and columns revealing the location of the positive YAC clone.

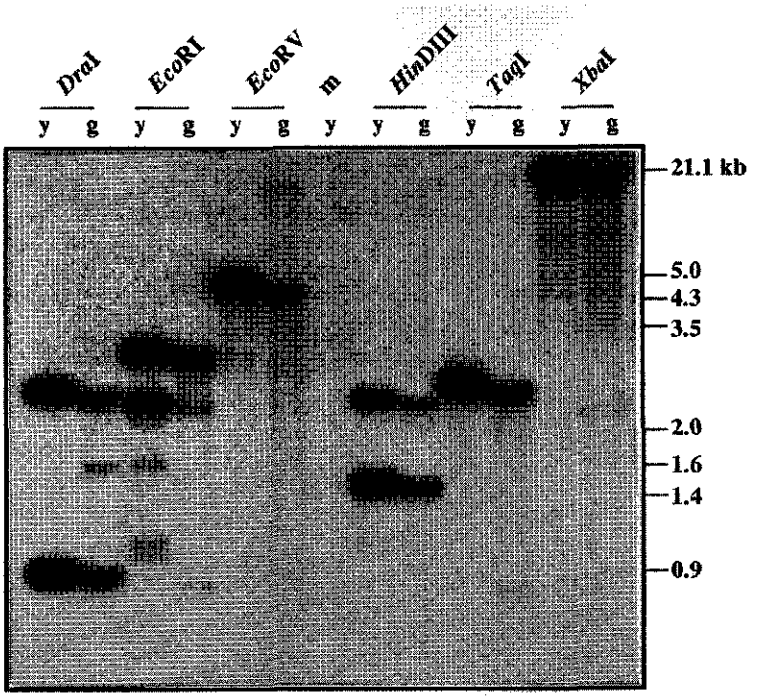


Figure 6. Southern blot analysis of yeast and tomato genomic DNA. DNA from both the YAC and genomic DNA (cv VFNT-Cherry) was digested with a set of indicated enzymes. The restriction fragments hybridizing with the *Aps-1* probe in both the YAC and genomic DNA were identical for every digest, indicating that a genuine *Aps-1*^{1} YAC was isolated.

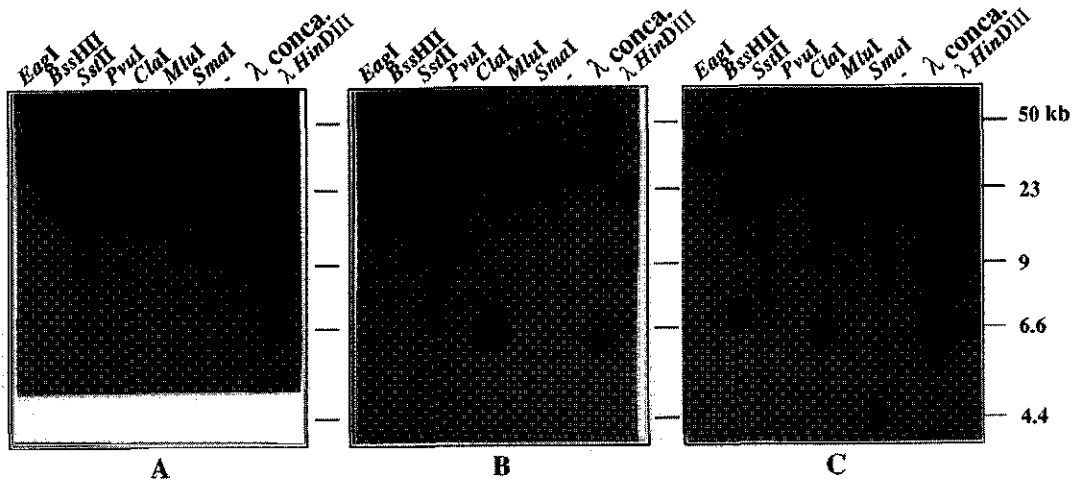


Figure 7. Southern blot hybridization analysis of the *Aps-1* containing YAC 134A1. Chromosomal yeast DNA was digested with the enzymes indicated and the restriction fragments were separated on a FIGE gel (1 - 5 s ramp, length of reverse pulsetimes 1/3 from forward, 18 hr run at 200 V). The gel was blotted and hybridized with (A) *Aps-1*^{1}, and, after deprobing, with a right arm specific probe (B), and finally with pBR322 (C).

by PFGE Southern blot hybridization analysis using *Aps-1*, a right arm specific fragment, and pBR322 as probes (shown in figure 7). The blot was deprobed between each hybridization. A restriction map of the YAC with the approximate location of *Aps-1* was constructed and is shown in figure 8. Eventhough the YAC is very small several rare-cutter sites were detected, but the small size precluded it to be placed on the available genomic long range restriction maps [16]. On the restriction map of the YAC only two rare-cutting sites, *Pvu*I and *Eag*I, mapped to the same location. This site might indicate the presence of a CpG-island, and the proximity of the 5'-side of the *Aps-1* gene. A more detailed analysis remains necessary to unequivocably determine the presence of an CpG-island and the possible position of the *Aps-1* gene.

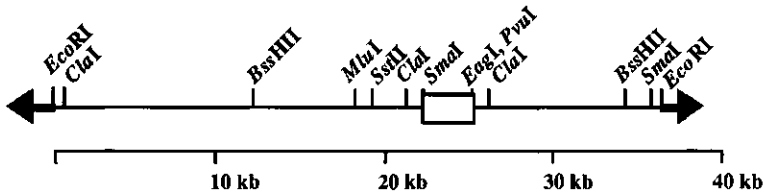


Figure 8. Restriction map of YAC 134 A1. The putative position of *Aps-1* is indicated with an open box.

Concluding remarks

It will be clear that the construction of a complete *Eco*RI library requires a large effort. Many precautions have to be taken to obtain a YAC library with large inserts, which is crucial for an efficient chromosomal walk. It has been shown by both theoretical analysis [2] and actual walking [8] that the required effort to perform a walk decreases substantially with increasing YAC size. Therefore, if the goal is to isolate a specific gene, like *Mi* in our case, it can be argued whether the effort of making such a library is worthwhile, especially compared to the relative ease with which a rare-cutter library can be constructed.

References

1. Albertsen HM, Abderrahim H, Cann HM, Dausset J, Le Paslier D, Cohen D (1990) Construction and characterization of a yeast artificial chromosome library containing seven haploid human genome equivalents. Proc Natl Acad Sci USA 87: 4256-4260.
2. Arratia R, Lander ES, Tavaré S, Waterman MS (1991) Genomic mapping by anchoring random clones: a mathematical analysis. Genomics 11: 806-827.
3. Burgers PMJ, Percival KJ (1987) Transformation of yeast sphaeroplasts without cell fusion. Anal Biochem 163: 391-397.
4. Burke DT, Carle GF, Olson MV (1987) Cloning of large segments of exogeneous DNA into yeast by means of artificial chromosome vectors. Science 236: 806-812.

5. Connely C, McCormick MK, Shero J, Hieter P (1991) Polyamines eliminate an extreme size bias against transformation of large yeast artificial chromosome DNA. Genomics 10: 10-16.
6. Feinberg AP, Vogelstein B. (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal Biochem. 132: 6-10
7. Gilbert JC (1958) Some linkage studies with the *Mi* gene for resistance to root-knot. Tomato Genet Coop Rep 8: 15-17.
8. Grigoriev AV (1993) Theoretical predictions and experimental observations of genomic mapping by anchoring rando clones. Genomics 15: 311-316.
9. Ho JY, Weide R, van Wordragen MF, Lambert K, Koornneef M, Zabel P, Williamson VM (1992) The root-knot nematode resistance gene (*Mi*) in tomato: construction of a molecular linkage map and identification of dominant cDNA markers in resistant genotypes. The plant J., 2: 971-982.
10. Hoffman CS, Winston F (1987) A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherischia coli*. Gene 57: 267-272.
11. Klein-Lankhorst R, Rietveld P, Machiels B, Verkerk R, Weide R, Gebhardt C, Koornneef M, Zabel P (1991) RFLP markers linked to the root knot nematode resistance gene *Mi* in tomato. Theor Appl Genet 81: 661-667.
12. Klein-Lankhorst RM, Vermunt A, Weide R, Liharska T, Zabel P (1991) Isolation of molecular markers for tomato (*L. esculentum*) using random amplified polymorphic DNA (RAPD). Theor Appl Genet 83: 108-114.
13. Larin Z, Monaco A, Lehrach H (1991) Yeast artificial chromosome libraries containing large inserts from mouse and human DNA. Proc Natl Acad Sci USA 88: 4123-4127.
14. Maniatis T, Fritsch EF, Sambrook J (1983) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
15. van Daelen RAJJ, Zabel P (1991) Preparation of high molecular weight plant DNA and analysis by pulsed-field gel electrophoresis. In: Gelvin SB, Schilperoort RA (Eds) Plant Mol Biol Manual, Kluwer, Dordrecht, A15: 1-25
16. van Daelen RAJJ, Gerbens F, van Ruissen F, Aarts J, Hontelez J, Zabel P (1993) Long-range physical maps of two loci (*Aps-1* and GP79) flanking the root-knot nematode resistance gene (*Mi*) near the centromere of tomato chromosome 6. Plant Mol Biol 23: 185-192.
17. van Grinsven MQJM, Gielen JJL, Zethof JLA, Nijkamp HJJ, Kool AJ (1986) Transcriptional and post-transcriptional regulation of chloroplast gene expression in *Petunia hybrida*. Theor Appl Genet 73: 94-101.
18. Widholm LM (1972) The use of fluorescein diacetate and phenosafranine for detemining viability of cultured plant cells. Stain Technol. 47: 189-194.
19. Zabel P, Meyer D, van der Stolpe O, van der Zaal B, Ramanna MS, Koornneef M, Krens F, Hille J (1985) Towards the construction of artificial chromosomes of tomato. In: van Vloten-Doting L, Groot GSP, Hall TC (Eds) Molecular form and function of the plant genome. NATO ASI Series, Plenum New York, pp 609-624.

CHAPTER 7

SUMMARY AND CONCLUDING REMARKS

The work described in this thesis is part of the ongoing research project which aims at the isolation of the tomato root-knot nematode resistance gene *Mi*. The *Mi*-gene is a typical example of a resistance gene that is of monogenic nature. Cloning of such genes will greatly advance the understanding of the mechanisms underlying disease resistance in plants. Furthermore, plant breeding might benefit from the isolation of resistance genes [22]. As no product of the *Mi*-gene is known, cloning is only possible via indirect approaches as discussed in chapter 1. Since *Mi* is a monogenic trait, accurately located on a genetic map, and tightly linked to an isozyme marker (*Aps-1*), it was, in principle, a good candidate for isolation by positional cloning. The basic idea of positional cloning, which is described in full in chapter 1, is to perform a chromosome walk, starting from genetically tightly linked markers towards the gene to be isolated. Ultimately, this will result in cloning of the relevant gene.

The first human gene to be isolated by a positional cloning approach was the human retinoblastoma gene (RB) in 1983 by Lee *et al* [21]. This was a rather fortuitous case, as after a chromosomal walk of only 30 kb from the marker linked to RB-1, a cosmid clone identified a cDNA which turned out to represent the RB-1 gene. Such a close linkage between marker and gene of interest is rare for, more often, distances which must be covered are in the range of megabases. Using conventional cloning vectors such as cosmids, which have a maximum cloning capacity of 40 kb, an impracticle large and time consuming number of walking steps must then be performed. An additional problem, often encountered in chromosome walking experiments, which also caused some trouble in theanalysis of the RB-1 region, is the presence of repetitive DNA. The walking experiments performed by Lee *et al* [21] started initially from another marker linked to RB-1, but were terminated due to the presence of highly repetitive sequences. At that time both the large distances which had to be walked through and the presence of repetitive DNA sequences held up a large scale application of positional cloning.

The efficiency of chromosomal walking and, therefore, the feasibility of positional cloning, was boosted enormously with the development of techniques to handle and analyze large DNA molecules. These include methods for the preparation of megabase-sized DNA (Schwartz and Cantor [26]), electrophoretic procedures like Pulsed Field Gel Electrophoresis (PFGE) [26] to separate such large molecules and cloning systems as Yeast Artificial Chromosomes (YACs), developed by Burke *et al* [6] to isolate very large DNA fragments.

At the onset of our work the techniques to isolate megabase sized DNA, PFGE, and YAC cloning, were just emerging and positional cloning of *Mi* seemed a feasible approach. However, neither of these techniques had yet been applied to plant systems. The work described in this thesis involves the application and adjustment of these techniqes to their use in plants. The first technique to be tackled was isolation of megabase-sized plant DNA, which was described in chapter 2.

For the isolation of high molecular weight DNA, living cells have to be lysed as quickly as possible and in such a way that nuclease activity and shearing are kept to a minimum. As for mammalian cells this was achieved by embedding them within an agarose matrix followed by lysis in the presence of a high concentration of a chelating agent, as described in chapters 1 and 2. Mammalian cells are readily available from rapidly growing tissue cultures and can be easily harvested and embedded within agarose. For plants, similar tissue cultures are not available. Furthermore, the plant cell wall needs to be removed to render the DNA accessible for isolation. Therefore, protoplasts seemed a good source for preparing DNA and this appeared indeed to be the case. With the use of protoplasts, isolated from tomato leaves, DNA could be prepared that was several megabases in size and suitable for long range restriction mapping and YAC cloning. Unfortunately, the disruption of plant tissues unavoidably leads to a certain loss of living cells and, thus, to degradation of DNA. Preparations with a high percentage of viable protoplast appeared to be crucial to obtain high molecular weight DNA, as discussed in chapters 2 and 5. An important phenomenon that gave trouble each year, is the so called wintereffect resulting in poor protoplast isolations during the winterseason. The cause of the wintereffect is not understood but during the winter season it was virtually impossible to obtain good protoplasts from which large-sized DNA could be prepared.

Once high molecular weight DNA was available, the next step was to generate large restriction fragments and to separate these using available PFG electrophoresis equipment. It turned out that the rare-cutting restriction enzymes often used in mammalian studies could similarly be applied for tomato and these were accordingly used. Although only a few long range restriction maps have been constructed in tomato, and generalizations on the restriction patterns of the rare-cutter enzymes can, as yet, not be made, some differences to mammalian DNA are apparent. Enzymes containing A, T, and C and G's in their recognition site, such as *Mlu*I and *Nru*I, generate on average smaller fragments in tomato as compared to human DNA whereas enzymes, such as *Sst*II and *Bss*HII, containing only G and C in their recognition site, generate on average larger fragments. This may be due to the higher GC-content of human DNA, which is 41% as compared to 37% for tomato [23].

During the course of our studies the electrophoretic techniques for separating large DNA molecules were rapidly improving. Initially, a field inversion gel electrophoresis (FIGE [7]) apparatus was used as this was the first, relatively cheap, system that generated straight lanes. Later on this could also be achieved by, among others, the CHEF [8] system at a much better resolution and larger separation range.

The application of the YAC cloning technique turned out rather cumbersome and protocols, described in chapters 5 and 6, are based on the experience gained from many experiments in which numerous problems were encountered. An aspect, that already seemed important at an early stage, was the concentration of DNA used during yeast transformation. According to Burke (personal communication) a concentration of 50 to 500 ng per µL was preferable and

much effort was put into attaining this goal. This can be achieved by starting with higher DNA concentrations, or by concentrating DNA after the restriction digestion, ligation and size selection steps. Several approaches were followed to obtain high starting concentrations. Embedding a large number of protoplasts resulted in some improvement, but the number of protoplasts is less than the number of human cells that can be embedded due to the relatively large cell volume of protoplasts as compared to human cells. Moreover, embedding a very large number of protoplasts led to increased degradation of DNA. When nuclei, which have a very small volume and can be embedded at a high concentration, were used to obtain higher concentrations of DNA, it appeared that the DNA was strongly degraded, irrespective of the nuclei isolation method followed. The use of sucrose gradient centrifugation allows isolation of DNA of high molecular weight in high concentrations [17, 28]. Unfortunately, such DNA was very difficult to handle due to a high viscosity and inhomogeneous character of the preparations. Partial digestions could not be accurately controlled and, furthermore, since the DNA is very sensitive to shearing, much degradation was unavoidable. It was therefore decided to use only DNA from agarose embedded protoplasts. The concentration of DNA prepared in this way was 50 µg/ml at the maximum.

Initially the agarose embedded DNA was partially digested using either limiting amounts of enzyme or a limiting amount of Mg^{2+}. Although both methods were effective, they were not as reproducible as partial digestion using *Eco*RI in combination with *Eco*RI methylase, as discussed in chapter 6.

After partial digestion and ligation to the vector arms, the recombinant DNA had to be size selected in order to obtain YACs larger than 100 kb. For this size selection the phenomenon of limiting mobility, that is typical for pulsed field gel electrophoresis, was exploited as all molecules above a certain size, depending on the electrophoresis conditions, become concentrated in the limiting mobility, or compression zone, located just below the slots of a pulsed field gel. This zone can be cut out of the gel and the DNA liberated from the agarose gel matrix can be used for transformation of yeast cells. When agarases free of nucleases was brought on the market this could be achieved efficiently. To protect DNA against shearing, polyamines or NaCl had to be included in the buffers used in the liquification step (see chapter 5).

For the transformation of yeast an efficient protocol was available (Burgers and Percival [5]). The combination of the yeast strain vector appeared important to obtain a high transformation efficiency. The yeast strain AB1380 and vector pYAC4 turned out to be a poor combination to test the transformation procedures as it yielded low transformation efficiencies. Therefore, other yeast strains and vectors were tested and it was found that the combination of AB1380 with plasmid YCp50 was the most efficient, yielding transformation efficiencies of $1 - 5 \times 10^5$.

In retrospect the most crucial aspects in YAC cloning appeared to be the quality of tomato DNA and size selection of the recombinant DNA used for transformation (as discussed in

chapter 6). Considering the difficulties in isolating large sized plant DNA and the, consequently, rather poor DNA preparations, as compared to mammalian DNA, it is not strange that nearly all published plant YAC libraries, including ours, consist of clones with inserts much smaller than those of human YAC libraries (see chapter 4).

With the techniques available for analyzing large size DNA molecules, a start was made with the long range restriction mapping of specific chromosomal regions. Initially, when the markers *Aps-1*[1] and GP79 were not available, the RFLP marker TG54 and the *Adh-2* gene were used. These loci, which are located close to each other on the available genetic maps, served as a model for the choice of rarely cutting restriction enzymes and for determination of digestion conditions (data not shown). As markers linked to *Mi* became available, *Aps-1*[1] and GP79 being the first, long range restriction mapping of the *Mi*-region was initiated in an attempt to link both markers physically, as described in chapter 3. Using several rare-cutting enzymes two maps of 800 and 400 kb were obtained, for the loci *Aps-1* and GP79 respectively, but no physical linkage could be established between both markers. In conjunction with the PFGE data on the markers H3F8 and H4H10, both genetically linked to the *Aps-1* locus, but showing no physical linkage either, it became apparent that physical distances within the *Aps-1*/*Mi*-region are very large. Thus, the small genetic distances within the *Mi*-region do not correspond to small physical distances.

The explanation for this underestimation of physical distances can be found in the chromosomal location of both *Aps-1* and GP79, which is in the heterochromatic region surrounding the centromere of chromosome 6. These regions are prone to lower recombination frequencies than other parts of the chromosome which results in small genetic distances [14, 19, 25, 27]. In the case of the *Mi*-region an additional negative effect on recombination frequency is exerted by the chromosomal variation between the cultivars used in many genetic mapping studies. Crosses involving plants containing sequences originating from *L. peruvianum* and *L. esculentum* at specific chromosomal regions, show a large reduction in recombination within these regions [18].

It is noteworthy that during physical mapping of the region containing the *Tm-2a* gene, which confers resistance against tobacco mosaic virus in tomato, similar results have been obtained [11]. This gene is located on the long arm of chromosome 9 close to the centromere. Several tightly linked RFLP markers have been genetically mapped within a segment of only 1.2 cM, but are distributed over 4000 kb. This represents a physical distance which is 5 times larger than the average distance of 750 kb per cM for the tomato genome [11, 29].

From the above it is clear that a physical characterization of the target region must be made, before a chromosomal walk is started. Apart from the now available technical means to analyze the target region, as described above, it is important to have enough markers at hand so that they can be linked by long range restriction mapping allowing determination of the physical distances within the region. Although RFLP and RAPD technology provided

additional markers, currently the AFLP® technology is the method of choice to generate many markers within a short time span. Especially with the resistant cultivar "Motelle", that carries the smallest known introgression segment (estimated to be less than 600 kb, Zabeau, unpublished results), many tightly linked markers can easily be identified. Following the identification of additional markers the *Mi*-region can be characterized physically in more detail.

Yeast Artificial Chromosomes were used to clone sequences from the *Mi*-region for the reasons discussed in chapter 1. With the ultimate aim to obtain YAC clones from the *Mi*-region two strategies were tested. First, based on long range restriction maps for the specific region, so-called "rare cutter libraries", enriched for restriction fragments of the appropriate size class, were generated. The second approach was to construct a "complete library", consisting of a set of overlapping clones that encompass the entire genome using partially digested DNA.

The advantage of a rare-cutter library is that only a small number of clones are needed since all restriction fragments larger and smaller than the identified fragment are not included and, therefore, both the construction and screening of the library is enormously less labor-intensive. The disadvantage is that the library contains only a fraction of the entire genome and is, most likely, only useful for the isolation of specific restriction fragments. Construction of such a library is only undertaken if a physical characterization of the region has been peformed, after which the appropriate fragments are isolated.

The advantage of a complete library is that the entire genome will be represented as YAC clones allowing chromosomal walking. For efficient chromosomal walking clones need to be large, preferably with an average size of several hundreds of kb, and available in a large number. To cover 95% of the genome, a library has to consist of 3 genome equivalents, although for a chromosomal walk 5 genome equivalents are preferred, covering 99% of the genome [20]. Taking into account the relative abundance of chimeric YACs, a size of 9 to 10 genome equivalents may be necessary [2, 9]. These large numbers make the construction, handling, and screening of a complete library very laborious.

With respect to the number of chimeric YACs it must be mentioned that their proportion in plant libraries may be somewhat smaller compared to human YAC libraries, as *in vivo* recombination of YAC DNA, a major cause for chimerism of human YACs [15], may be less important. The reason for this is that repetitive elements with an omnipresence such as the human *Alu* repeat, seem to be absent from plant genomes.

For the construction of the rare-cutter library (chapter 5), the enzyme *Cla*I was used as both GP79 and *Aps-1* recognize restriction fragments of similar sizes. A YAC library of 6000 clones, roughly 10 % of which contained chloroplast DNA, with an average size of 110 kb was constructed and screened. Although, the YACs carried predominantly fragments between

75 and 125 kb, encompassing the *Aps-1*^{1} fragment, no clone containing *Aps-1*^{1} could be detected. Strikingly, a GP79 containing clone with an insert of the appropriate size was found, even though chances for the latter were substantially lower since only a small fraction (< 5 %) of the clones in the library was of the appropriate size class.

Unfortunately, the size of a library with a chance of 95% to find an *Aps-1*^{1} clone, is difficult to determine since no data are available on the distribution, nor on the average size, of tomato *Cla*I restriction fragments. However, from several ethidium bromide-stained CHEF gels, containing complete *Cla*I digests, it was estimated that the size class from 75 - 125 kb encompasses approximately 10 % of all the generated restriction fragments. This corresponds to an average *Cla*I fragment size of 150 kb, assuming a normal size distribution. Similar experiments, performed with human rare-cutter YAC libraries, [3] based on *Bss*HII, which generates fragments with an average size of 100 kb [3, 24], showed that a library of 2500 clones represents 1 genome equivalent [3]. Therefore, the approximately 4000 *Cla*I YACs, with inserts between 75 and 125 kb, may correspond to approximately 5 genome equivalents (taking into account the three times larger human genome). Data from Tanksley and coworkers [10, 11], show an average of 350 kb regarding the size of *Cla*I restriction fragments of tomato. With these data at hand, the entire library of 6000 clones was expected to contain several genome equivalents and thus large enough to hold an *Aps-1*^{1} clone. Therefore, the abscence of an *Aps-1*^{1} containing YAC in the library was either a matter of bad luck or the *Aps-1*^{1} sequence is difficult to clone.

The identification of a GP79 containing clone in the *Cla*I library shows that the library is virtually satisfactory, but more clones should be generated to make the library all embracing. In this respect the picking and screening protocols can, and should be, further optimized. This might be achieved by replica plating of all clones simultaneously using the multipin transfer device as described by Larin *et al* [20]. One set of replica plates can then be used for a colony filter screening of the whole library while another set can provide the source of cells for the identified positive clones. The complete library can be stored as pools of clones rather than as individual clones in microtiter dishes. The entire screening procedure may be completed within several weeks.

The construction of a complete library from a partial *Eco*RI digest, consisting of a set of overlapping clones, is technically much more demanding. In chapter 6, the construction and analysis of a partial *Eco*R1 YAC library consisting of approximately 3000 clones, with an average insert size of 150 kb, is described. To allow chromosomal walking towards *Mi* the library should still be enlarged considerably. However, chromosome walking, with clones that are, on average, 100 - 200 kb in size, is very laborious and clones with a larger insert size are to be preferred [2, 16]. This will require a major effort. An alternative could be, to make a physical map of the region, following the identification of additional (AFLP®) markers, and to localize *Mi* on this map. Subsequently, a relatively small rare-cutter library can be constructed to secure the isolation of the appropriate restriction fragment. Nowadays,

however, the above is purely hypothetical as currently a cosmid contig covering almost the entire Motelle region, containing the *Mi*-gene, has been made (Zabeau, unpublished results).

Only a part of the *Eco*RI library could be screened due to a bacterial infection, but from the 800 clones screened -representing less than 1/8 genome equivalent !- an *Aps-1*^{1} containing YAC was isolated and characterized. Unfortunately this clone is not very usefull for chromosome walking as it contains an inser of only 35 kb.

For the screening of the YAC libraries two approaches were followed: colony filter hybridization and PCR screening. The screening for GP79 was performed by high density colony filter hybridization since no sequence data were available. Moreover, the potato derived sequences might not be sufficiently homologous to tomato sequences for a specific PCR. Using high density colony filter hybridization, the clone was quickly identified in only one screening round. The PCR screening approach was applied for the identification of the *Aps-1*^{1} containing YAC. It turned out to be easier to get the PCR screening procedures working than colony filter hybridization. However, the results indicate that both methods work well and the approach to be followed will strongly depend on specific conditions. As PCR screening requires rather elaborate DNA isolation and pooling protocols, it is most usefull if the isolated DNAs are used repeatedly, as will most likely be the case with a complete (*Eco*RI) library. Colony filter hybridization, on the other hand, is preferable for the screening of a rare-cutter library, especially in combination with high density replicators, as the library will only be screened for a very limited number of clones.

It should be mentioned here that the methylation status of the YAC-DNA introduced into yeast is lost, since in yeast DNA is not methylated. This is clearly reflected in the much higher number of rare cutter sites detected on genomic DNA cloned in YACs as compared to the original genomic DNA (see chapters 5 and 6). In the analysis of the 200 kb GP79 containing YAC, an increase in the number of rare-cutter restriction sites was found, which varied by a factor two to six, depending on the enzyme used. This is in rough agreement with the methylation data presented by Messenguer *et al* [23], describing that 55% of all CpG's and 85% of all CpNpG's are methylated.

For the identification of putative *Mi* coding sequences in the last step of the positional cloning approach, use can be made of the presence of CpG-islands. These unmethylated regions, with a CpG content higher than the surrounding DNA, are present at the 5' end of all housekeeping genes and at the 5' end of many tissue specific genes in both vertebrates and plants [1, 4]. They can be detected, at least in vertebrate systems, by a cluster of restriction sites of rarely cutting restriction enzymes. Based on the long range restriction maps of genomic tomato DNA, we could not detect an island linked to the *Aps-1*^{1} gene, which is constitutively expressed [9b], using the clustering of rare-cutter restriction sites as criterium. Neither was a clear clustering of restriction sites found on the insert of the *Aps-1*^{1} containing

YAC. Whether the CpG-island, if present at all, is too small to be detected remains to be investigated. However, CpG-islands present in plants differ in some characteristics from mammalian islands. A major difference between human and dicotyledonous plant CpG-islands is their AT content, which is much higher in plants. Furthermore, is was found that the ratio of observed/expected CpG-dinucleotide frequency in plant genomes, is higher than in vertebrate genomes [12, 13], indicating that in plants islands may be more difficult to detect. These characteristics may explain the abscence of a cluster of rare-cutter restriction sites close to *Aps-1*.

Currently, locating CpG islands by restriction analysis, to identify putative coding sequences, has not been used in plants and its value in this respect remains to be proven. So far detection of coding sequences, in positional cloning of plant genes, have mostly been achieved by identifying cDNAs mapping to the region of interest.

In conclusion, all procedures for positional cloning have been developed and cloning of *Mi* should be feasible in the near future.

References

1. Antequera F, Bird AP (1988) Unmethylated CpG islands associated with genes in higher plant DNA. EMBO J 7: 2295-2299.
2. Arratia R, Lander ES, Tavaré S, Waterman MS (1991) Genomic mapping by anchoring random clones: a mathematical analysis. Genomics 11: 806-827.
3. Bates GP, MacDonald ME, Baxendale S, Sedlacek Z, Youngman S, Romano D, Whaley WL, Allitto BA, Poustka A, Gusella JF, Lehrach H (1990) A yeast artificial chromosome telomere clone spanning a possible location of the Huntington disease gene. Am J Hum Genet 46: 762-775.
4. Bird AP (1986) CpG-rich islands and the function of DNA methylation. Nature 321: 209-213.
5. Burgers PMJ, Percival KJ (1987) Transformation of yeast sphaeroplasts without cell fusion. Anal Biochem 163: 391-397.
6. Burke DT, Carle GF, Olson MV (1987) Cloning of large segments of exogeneous DNA into yeast by means of artificial chromosome vectors. Science 236: 806-812.
7. Carle GF, Frank M, Olson MV (1986) Electrophoretic separation of large DNA molecules by periodic inversion of the electric field. Science 232: 65-68.
8. Chu G, Vollrath D, Davis RW (1986) Separation of large DNA molecules by contour-clamped homogeneous electric fields. Science 234: 1582-1585.
9. Ewens WJ, Bell CJ, Donnely PJ, Dunn P, Matellana E, Ecker JR (1991) Genome mapping with anchored clones: theoretical aspects. Genomics 11: 799-805.

9b. Fobes (1980) Trisomic analysis of isozymic loci in tomato species: segregation and dosage effects. Biochem Gen 18: 401-421.

10. Ganal MW, Tanksley SD (1989) Analysis of tomato DNA by Pulsed Field Gel Electrophoresis. Plant Mol Biol Rep 7: 17-27
11. Ganal MW, Young ND, Tanksley SD (1989) Pulsed field gel electrophoresis and physical mapping of large DNA fragments in the *Tm-2a* region of chromosome 9 in tomato. Mol Gen Genet 215: 395-400.
12. Gardiner-Garden M, Frommer M (1992) Significant CpG-rich regions in angiosperm genes. J Mol Evol 34: 231-245.
13. Gardiner-Garden M, Sved JA, Frommer M (1992) Methylation sites in Angiosperm genes. J Mol Evol 34: 219-230.
14. Gaudet AM, Fitzgerald-Hayes M (1990) The function of centromeres in chromosome segregation. In: *Strauss PR, Wilson SH (eds), The eukaryotic nucleus. Molecular biochemistry and macromolecular* assemblies. Telford Press, Caldwell, New Jersey pp 845-881.

15. Green ED, Riethman HC, Dutchik JE, Olson MV (1991) Detection and characterization of chimaeric yeast artificial-chromosome clones. Genomics 11: 658-669.
16. Grigoriev AV (1993) Theoretical predictions and experimental observations of genomic mapping by anchoring random clones. Genomics 15: 311-316.
17. Guzman P, Ecker JR (1988) Development of large DNA methods for plants: molecular cloning of large segments of *Arabidopsis* and carrot DNA into yeast. Nucleic Acids Res 16: 11091-11105.
18. Ho JY, Weide R, van Wordragen MF, Lambert K, Koornneef M, Zabel P, Williamson VM (1992) The root-knot nematode resistance gene (*Mi*) in tomato: construction of a molecular linkage map and identification of dominant cDNA markers in resistant genotypes. The plant J 2: 971-982.
19. Khush GS, Rick CM (1967) Studies on the linkage map of chromosome 4 of the tomato and on the transmission of induced deficiencies. Genetica 38: 74-94.
20. Larin Z, Monaco A, Lehrach H (1994) Large insert yeast artificial chromosome libraries. In: Nelson DL, Brownstein BH (Eds) YAC libraries, A user's guide. UWBC, Freeman and Co, New York, pp 43-56.
21. Lee WH, Bookstein R, Hong F, Young LJ, Shew JY, Lee EYHP (1987) Human retinoblastoma susceptibility gene: cloning, identification, and sequence. Science 235: 1394-1399.
22. Lindsey K (1992) Genetic manipulation of crop plants. J Biotechnol 26: 1-28.
23. Messenguer R, Ganal MW, Steffens JC, Tanksley SD (1991) Characterization of the level, target sites and inheritance of cytosine methylation in tomato nuclear DNA. Plant Mol Biol 16: 753-770.
24. New England Biolabs (1988) Catalog. USA.
25. Rick CM (1969) Controlled introgression of chromosomes of *Solanum pennellii* into *Lycopersicon esculentum*: segregation and recombination. Genetics 62: 753-768.
26. Schwartz DC, Cantor CR (1984) Separation of yeast chromosome-sized DNA by pulsed field gradient gel electrophoresis. Cell 37: 67-75.
27. Tanksley SD, Ganal MW, Prince JP, de Vincente MC, Bonierbale MW, Broun P, Fulton TM, Giovanonni JJ, Grandillo S, Martin GB, Messenguer R, Miller JC, Miller L, Paterson AH, Pineda O, Röder MS, Wing RA, Wu W, Young ND (1992) High density molecular linkage maps of the tomato and potato genomes. Genetics 132: 1141-1160.
28. Ward ER, Jen GC (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.
29. Young ND, Zamir D, Ganal MW, Tanksley SD (1988) Use of isogenic lines and simultaneous probing to identify DNA markers tightly linked to the *Tm-2a* gene in tomato. Genetics 120: 579-585.

SAMENVATTING

Wortelknobbelaaltjes, behorende tot het genus *Meloidogyne*, zijn pathogene organismen die grote schade kunnen toebrengen aan economisch belangrijke gewassen zoals tomaat, aardappel en katoen. Resistentie tegen wortelknobbelaaltjes is onder meer aanwezig in de wilde tomaat *Lycopersicon peruvianum* en berust op een enkel dominant gen: *Mi*. Dit gen, dat in de cultuurtomaat is ingekruist, verleent tomaten planten resistentie tegen diverse nematoden species behorende tot het genus *Meloidogyne*. Ofschoon de funktie van het gen bekend is, het verlenen van resistentie, is de wijze waarop deze resistentie wordt bereikt onbekend. Om de werking van het gen te kunnen bestuderen is de klonering van het *Mi*-gen zeer belangrijk. Daarnaast kan de plantenverdeling er voordeel van hebben wanneer geïsoleerde resistentie genen beschikbaar zijn. De klonering van resistentie genen, en dus ook van *Mi*, is echter zeer omslachtig omdat geen produkt van het gen, noch de aard van funktioneren van het gen, bekend zijn. Klonering kan daarom alleen via indirekte benaderingen geschieden, zoals de positionele klonering (hoofdstuk 1).

Als eerste stap in een positionele klonerings strategie worden zeer nauw aan het te kloneren gen gekoppelde moleculaire merkers geisoleerd. Vervolgens kan een chromosoom wandeling uitgevoerd worden die begint bij deze nauw gekoppelde merkers. Dit resulteert in de klonering van alle sequenties tussen twee het gen flankerende merkers en uit die gekloneerde sequenties kan vervolgens het gen worden geïsoleerd.

De efficiënte toepassing van positionele klonering is lange tijd ernstig belemmerd door de grote afstanden tussen genen en merkers. Deze afstanden zijn veelal in de orde grootte van enkele miljoenen basen paren (megabasen paren). Een chromosoom wandeling uitgevoerd met de klassieke klonerings vectoren zoals plasmiden, faag lambda, en cosmiden, waarmee DNA fragmenten van 10 to 40 kilobasen paren (kb) gecloneerd kunnen worden, is zeer lastig. Daarnaast kan ook het voorkomen van repetitief DNA een groot struikelblok zijn bij het uitvoeren van een chromosoom wandeling. De efficiëntie van chromosoom wandelen, en daarmee de mogelijkheid om via positionele klonering genen te isoleren, werd sterk verbeterd met het beschikbaar komen van technieken om zeer grote DNA moleculen te hanteren en analyseren. Schwartz en Cantor slaagden er in 1984 als eersten in om DNA moleculen met een lengte van meerdere megabasen te isoleren en electroforetisch te scheiden. Deze electroforetische scheidings methode wordt "Pulsed Field Gel Electrophoresis" (PFGE) genoemd. Daarnaast is van doorslaggevende betekenis geweest dat er methoden werden ontwikkeld om zeer grote DNA fragmenten -tot een lengte van een miljoen basenparen- stabiel te kloneren in gistcellen als extra chromosomen, de zogenaamde YACs ("Yeast Artificial Chromosomes"). Met deze methode (het YAC systeem) kunnen nu zeer grote fragmenten van het planten genoom moleculair gekarakteriseerd worden.

Bij de aanvang van het in dit proefschrift beschreven onderzoek was er nog geen ervaring met bovengenoemde technieken bij het onderzoek van het genoom van planten. In dit proefschrift wordt beschreven welke aanpassingen van deze technieken nodig waren voor planten moleculair biologisch werk, en hoe deze methoden gebruikt worden voor de positionele klonering van het *Mi*-gen van de tomaat.

Voor de isolatie van hoog moleculair DNA is het noodzakelijk dat de levende cellen waaruit het DNA wordt geisoleerd snel worden gelyseerd, onder condities waarbij breuken in het DNA, door wrijfkrachten en nuclease aktiviteiten, zo veel mogelijk worden vermeden. Dit kan bereikt worden door tomaten protoplasten in te bedden in een agarose matrix, waardoor het DNA, wanneer het vrij komt enigzins beschermd wordt tegen breken, en bovendien de cellen te lyseren in aanwezigheid van een hoge concentratie EDTA om de Mg^{2+} ionen weg te vangen die (endo)nucleases nodig hebben voor hun aktiviteit (hoofdstuk 2).

Nadat hoog moleculair planten DNA beschikbaar was gekomen kon een begin worden gemaakt met de fysische kartering van grote chromosomale gebieden. Hiervoor zijn restrictie enzymen nodig die het DNA maar op weining plaatsen knippen, zogenaamde "rare cutters". Dit zijn veelal restrictie enzymen die een of meerdere CpG doubletten in hun herkenningsequentie bezitten. Het C-residu van zo'n doublet is in DNA vaak gemethyleerd en het enzym kan niet knippen in DNA met een gemethyleerd CpG doublet. De rare-cutters knippen alleen die plaatsen waar de herkenningsplaats niet gemethyleerd is. Deze enzymen genereren grote tot zeer grote restrictie fragmenten met dierlijk en humaan DNA. Met tomaten DNA bleken in grote lijnen vergelijkbare resultaten verkregen te worden.

De nauw aan *Mi* gekoppelde merkers aan weerszijde van het gen, die het eerst beschikbaar waren, zijn *Aps-1*[1] en GP79 en deze zijn gebruikt voor een eerste fysische karakterisering van het *Mi*-gebied. Vanuit deze merkers zijn "lange afstands restrictie kaarten" vervaardigd, waarop restrictie plaatsen van zeldzame knippers zijn aangegeven (hoofdstuk 3). Uit deze kaarten bleek dat de fysische afstanden in het *Aps-1/Mi* gebied veel groter waren dan in eerste instantie, op grond van genetische analyse, was geschat. Een verklaring hiervoor kan gevonden worden in de ligging van *Mi* in het heterochromatine rijke gebied rond het centromeer van chromosoom 6. Deze heterochromatine rijke gebieden vertonen een lagere overkruisings frequentie dan overige delen van het chromosoom, hetgeen resulteert in relatief kleinere genetische afstanden in deze gebieden. Bij het onderschatten van de afstanden tussen genetische merkers op grond van genetische kartering speelt de herkomst van de planten die zijn gebruikt bij de kruisingen, waarmee het *Mi*-gebied genetisch is gekarakteriseerd is, ook een rol. Voor die kruisingen zijn planten gebruikt waarin het DNA in het *Mi*-gebied afkomstig was van, oftewel *L. esculentum* (de sensitieve planten) of van *L. peruvianum* (de resistente planten); en met DNA van verschillende soorten treedt minder frequent overkruising op dan tussen DNA van dezelfde soorten.

Uit de grote fysische afstanden tussen de merkers in het *Aps-1/Mi*-gebied volgt dat een chromosoom wandeling alleen doenlijk zal zijn, wanneer er merkers beschikbaar komen die dichter bij het *Mi*-gen liggen. Belangrijk zal dan ook zijn om meer, zeer nauw gekoppelde, merkers te isoleren en het gebied daarmee te karakteriseren. Recentelijk zijn andere onderzoekers erin geslaagd om, gebruik makend van de cultivar Motelle, die een zeer klein introgressie gebied bezit, en de AFLP® fingerprinting techniek, en, tot zekere hoogte, ook met RAPDs, een groot aantal additionele merkers te genereren.

Voor het kloneren van sequenties uit het *Aps-1/Mi* gebied is het YAC systeem gekozen (hoofdstuk 1 en 4) en er zijn twee strategien gevolgd voor de constructie van YAC banken met DNA uit het *Mi* gebied. Ten eerste, kunnen grote restrictie fragmenten, verkregen door knippen met een zeldzame knipper, in YACs gekloneerd worden, waarna, op basis van de aanwezige restrictie kaarten, de YACs met DNA uit een bepaald gebied geïdentificeerd kunnen worden. Na isolatie van deze YACs kan het DNA verder gekarakteriseerd worden. Voordat de constructie van een bank zinvol is, moet dus eerst het betreffende gebied fysisch gekarteerd worden. In de tweede strategie, kan er een bank worden gemaakt die bestaat uit een set YAC klonen met overlappende fragmenten, gegenereerd door een partieel restrictie digest van het DNA, die met elkaar het gehele genoom omvatten.

De constructie van een zeldzame knipper bank, waarbij het enzym *Cla*I is gebruikt, staat beschreven in hoofdstuk 5. Uit deze bank, bestaande uit circa 6000 kloons, is, middels koloniefilter hybridisatie, een 200 kb YAC geisoleerd die GP79 sequenties bevat. Dit geeft aan dat de methode, in principe, bruikbaar is. Echter, een *Aps-1* kloon, waarvan verwacht werd dat hij aanwezig zou zijn, is niet gevonden. De reden hiervoor is niet duidelijk.

Het voordeel van een zeldzame knipper bank is dat deze een relatief gering aantal kloons behoeft te omvatten, omdat slechts een subset van alle restrictie fragmenten gekloneerd hoeft te worden. Dit maakt de constructie en screening van een dergelijke bank aanzienlijk eenvoudiger dan van een bank met het totale genoom. Het nadeel is dat zo'n bank slechts een deel van alle restrictie fragmenten van het genoom omvat en niet zonder meer gebruikt kan worden voor chromosoom wandelen.

Naast de constructie van een zeldzame knipper bank, wordt in hoofdstuk 6 beschreven hoe een begin is gemaakt met het construeren van een bank van het volledige genoom. Het voordeel van zo'n bank is dat een chromosoom wandeling kan worden uitgevoerd en fysische karakterisering van het beoogde gebied vooraf, niet perse noodzakelijk is. Echter, een complete genomische bank die zich leent voor chromosoom wandelen moet een omvang hebben van circa 5 genoomequivalenten, hetgeen de isolatie van verschillende tienduizenden kloons betekend. Dit vereist een veel grotere inspanning dan de constructie van een zelzame knipper bank. Uit de *Eco*RI YAC bank, die is gemaakt en circa 3000 kloons omvat, kon, middels een PCR screening, een *Aps-1*^{1} bevattende YAC van 45 kb geisoleerd worden.

Een belangrijk facet van YAC klonering is dat DNA in gist niet is gemethyleerd, waardoor de methylerings-status van het tomaten DNA, dat in gist gekloneerd wordt, verloren gaat. Dit kwam duidelijk naar voren bij de constructie van restrictiekaarten van de geisoleerdeYACs waarbij een groot aantal extra knip plaatsen werd gevonden voor de diverse gebruikte zeldzame knippers.

De laatste stap bij de positioneel kloneren betreft het identificeren van coderende gebieden die het gen kunnen bevatten. Een hulpmiddel om zulke gebieden te herkennen, kunnen de zogenaamde CpG eilanden zijn. Dit zijn stukken DNA van circa 2 kb groot, gesitueerd aan de 5'-zijde van vele genen, met een CpG inhoud groter dan het omringende DNA, waarin de C-residuen niet gemethyleerd zijn. Deze gebieden kunnen in dierlijk en humaan DNA gedetecteerd worden doordat zich in een dergelijk gebied een ophoping van restrictie plaatsen van zeldzame knippers voordoet. In onderzoek met dierlijk en humaan DNA is deze methode meermalen succesvol toegepast. Helaas kon op de *Aps-1*[1] bevattende YAC geen CpG-eiland worden geidentificeerd op grond van de restrictie kaart. De reden hiervoor is nog niet duidelijk. Voor de toepassing van CpG-eiland detectie, als middel om genen te identificeren, is nader onderzoek nodig. Recentelijk zijn bij planten enkele malen cDNA's gebruikt voor het detecteren van coderende sequenties en op het ogenblik lijkt dit bij planten de meest aangewezen weg.

Samenvattend kan gezegd worden dat alle ingredienten vereist om het *Mi*-gen via positioneel kloneren te isoleren beschikbaar zijn en dat, derhalve, het gen binnen afzienbare tijd gekloneerd moet kunnen worden.

NAWOORD

Nu, woensdag avond 23 februari 1995, exact 2580 dagen na aanvang van mijn promotie-onderzoek kan ik eindelijk het nawoord bij mijn proefschrift schrijven.

Tijdens mijn studie bosbouw leek het er niet op dat ik ooit een promotieonderzoek zou doen in de moleculaire biologie. De bosbouwstudie echter bleek zich meer te richtten op politiek en beleid dan op bosteelt en fundamenteel onderzoek. Mijn afstudeervak op "Molbi" vormde dan ook een groot contrast met de bosbouw, maar vormde wel een zeer prettige en leerzame tijd van mijn studie. Gezien deze achtergrond lag AIO worden bij Molbi dan ook niet zo voor de hand. Daarom wil ik dan ook mijn dank uitdrukken aan mijn promotor Ab van Kammen en co-promotor Pim Zabel voor het vertrouwen, en de mogelijkheid, dit promotie onderzoek uit te voeren.

Een promotie onderzoek in de moleculaire biologie vereist bij aanvang het aanleren van vele "lab"technieken en hiervoor wordt nagenoeg een ieder die op het lab werkzaam is ingeschakeld. Daarom wil ik dan ook alle collega's van de tomaten groep, Jac, Ruud, Jan, Ellen, Rob, Els en Rene en de semi-tomaten groepsleden, Maarten en Jelle, bedanken voor de prettige samenwerking. Ook alle overige collega's van de vakgroep, zowel "lab" als "niet lab" horen bij deze dankzegging thuis, al was het maar voor de vrijdag middag borrels, die op het einde van mijn tijd helaas niet meer werden gehouden. Echter, niet alleen de "vaste medewerkers" van molbi, waren belangrijk, ook de vele studenten speelden een grote rol. Ook met hen was het contact intensief omdat in een korte tijd een redelijke hoeveelheid praktische vaardigheden moesten worden overgedragen en als daarnaast de werkbesprekingen ook nog in de kroeg werden voortgezet maakte het tot een zeer plezierige periode. Ik wil hier dan ook met name noemen John Groenendijk en Wilbert van Workum zonder de anderen, Jos Jonkers, Edwin Cuppen, Joke Orsel, Ronald Roepman, Frans Gerbens, Fred van Ruisen, Jaap Scholtens en Nicolette Klijn, tekort te willen doen.

Na een periode van labwerk komt gewoonlijk een periode van het opschrijven van de resultaten en het verwerken ervan tot een publiceerbaar geheel. Hier wil ik dan Pim nogmaals bedanken voor het doornemen en becommentarieren van de vele versies van een manuscript. In een wat latere fase was het aan Ab van Kammen de zaken nogmaals kritisch door te nemen en met name de houding, zoals Peter Sterk al uitdrukt: "ik weet wel wat je bedoelt, maar het staat er niet" maakte mijzelf kritischer op hetgeen ikzelf had opgeschreven en uiteindelijk het geheel tot een beter stuk, waarvoor mijn dank.

Daarnaast hebben ook regelmatig mensen van buiten de vakgroep alle perikelen rond de promotie moeten aanhoren. Gelukkig geschiedde dit altijd onder toeziend oog van glazen gevuld met alcoholhoudende dranken. Matthieu en Jacques, bedankt, nu alleen nog even als paranimf fungeren.

Inmiddels ben ik reeds enkele jaren bij Keygene werkzaam en ook hier is een woord van dank op zijn plaats gezien de grote vrijheid van tijdindeling om regelmatig tijdens werktijd weer eens iets te doen aan mijn proefschrift.

En als laatste genoemd, maar in feite op de eerste plaats, Natascha, waarvan ik een "heeeeeeleboelie veel" houd en zonder jou was en is het allemaal veel minder goed.

Raymond

CURRICULUM VITAE

Raymond Antoine Jean Joseph van Daelen werd geboren op 15 maart 1961 te Geldrop. Na diverse scholen (HAVO en VWO) en het leger te hebben bezocht werd de studie Bosbouw (nieuwe stijl) in september 1982 aan Landbouw Hogeschool aangevangen. Gedurende de doctoraal fase zijn afstudeervakken uitgevoerd aan de vakgroepen Bosteelt (Prof. Dr. R. Oldeman), Phytopathologie (Prof. Dr. P. de Wit), Plantenveredeling (Dr. R. Niks) en Moleculaire Biologie (Prof. Dr. A. van Kammen). Na het behalen van het ingenieursdiploma in januari 1988, is in hetzelfde jaar (1 januari 1988) het promotie onderzoek, beschreven in dit proefschrift, aangevangen bij de vakgroep Moleculaire Biologie van de Landbouwuniversiteit Wageningen. Sinds 1 Februari 1993 is hij werkzaam bij Keygene NV te Wageningen.