Mol Gen Genet (1999) 261: 1021–1031

ORIGINAL PAPER

J. N. A. M. Rouppe van der Voort · H. J. van Eck
P. M. van Zandvoort · H. Overmars · J. Helder
J. Bakker

Linkage analysis by genotyping of sibling populations: a genetic map for the potato cyst nematode constructed using a "pseudo-F2" mapping strategy

Received: 7 January 1999 / Accepted: 16 April 1999

Abstract A mapping strategy is described for the construction of a linkage map of a non-inbred species in which individual offspring genotypes are not amenable to marker analysis. After one extra generation of random mating, the segregating progeny was propagated, and bulked populations of offspring were analyzed. Although the resulting population structure is different from that of commonly used mapping populations, we show that the maximum likelihood formula for a normal F2 is applicable for the estimation of recombination. This "pseudo-F2" mapping strategy, in combination with the development of an AFLP assay for single cysts, facilitated the construction of a linkage map for the potato cyst nematode *Globodera rostochiensis*. Using 12 pre-selected AFLP primer combinations, a total of 66 segregating markers were identified, 62 of which were mapped to nine linkage groups. These 62 AFLP markers are randomly distributed and cover about 65% of the genome. An estimate of the physical size of the *Globodera* genome was obtained from comparisons of the number of AFLP fragments obtained with the values for *Caenorhabditis elegans*. The methodology presented here resulted in the first genomic map for a cyst nematode. The low value of the kilobase/centimorgan (kb/cM) ratio for the *Globodera* genome will facilitate map-based cloning of genes that mediate the interaction between the nematode and its host plant.

Communicated by R. Hagemann

J. N. A. M. Rouppe van der Voort · P. M. van Zandvoort
H. Overmars · J. Helder (✉) · J. Bakker
Graduate School of Experimental Plant Sciences,
Department of Nematology,
Wageningen Agricultural University, P.O. Box 8123,
6700 ES Wageningen, The Netherlands
e-mail: hans.helder@medew.nema.wau.nl.
Tel.: +31-317-485085; Fax: +31-317-484254

H. J. van Eck
Department of Plant Breeding,
Wageningen Agricultural
University, P.O. Box 386,
6700 AJ Wageningen, The Netherlands

Key words AFLP marker · *Globodera*

Introduction

Linkage maps are important tools in genetic research. High-density maps that include numerous molecular markers facilitate the genomic localization of genes responsible for monogenic and polygenic traits, and have facilitated the cloning of genes whose products were initially unknown. In studies on evolution and population structure, linkage maps have been used to identify genomic regions involved in selection processes. Moreover, linkage maps are of practical relevance for diagnostics, breeding and germplasm conservation.

During the past decade, detailed linkage maps based on restriction fragment length polymorphism (RFLP) markers have been developed for a number of species (reviewed in Tanksley et al. 1989; O'Brien 1993). However, RFLP analysis involves labor-intensive procedures and requires large quantities of DNA. With the advent of PCR-based molecular marker techniques, such as randomly amplified polymorphic DNA (RAPD; Williams et al. 1990), sequence-tagged micro satellites (STMS; Beckmann and Soller 1990) and AFLP markers (Vos et al. 1995), more sensitive and efficient methods have become available for the generation of large numbers of genetic markers. High-density maps (> 500 markers) have been published for – among others – *Arabidopsis thaliana* (Reiter et al. 1992; Hauge et al. 1993), rice (Harushima et al. 1998), tomato and potato (Tanksley et al. 1992), barley (Qi et al. 1998), mouse (Dietrich et al. 1996) and human (Murray et al. 1994).

Linkage maps are preferably constructed by analysis of inbred or immortalized populations derived from the F1 of a cross between two homozygous inbred parents, i.e. RILs (recombinant inbred lines) or IF2 (immortalized F2; Gardiner et al. 1993). RILs and IF2s have the advantage that the population can be propagated indefinitely, thereby allowing an unlimited number of

markers and traits to be analyzed repeatedly at different times and in different environments. Where litter sizes are small, data from several families can be combined and maps can be refined by collecting information from common databases. In plants, F2 and backcross (BC) populations are useful because many crops can be propagated clonally and backcrosses can be carried out without severe inbreeding depression. The advantage of molecular markers is that individual members of a population can be genotyped using minute amounts of tissue. Among the DNA marker systems available, the AFLP technique has the highest sensitivity. As little as 25 pg of genomic DNA is sufficient for the generation of a virtually unlimited number of markers (Vos et al. 1995). Nevertheless, the isolation of even such a quantity can be an insurmountable problem in the case of microscopic organisms that reproduce only by outbreeding.

We set out to generate a genetic linkage map of the genome of the potato cyst nematode, a tiny roundworm which has so far only been analyzed at the population level. The potato cyst nematodes (PCNs) *Globodera rostochiensis* and *G. pallida* are serious pests of potato, accounting for estimated yield losses of 10% annually (Oerke et al. 1994). For growth and development, PCNs depend on a feeding site; a conglomerate of metabolically highly active plant cells. In resistant host plants, feeding site induction is hampered, resulting in a substantial reduction in the size of the nematode population. As with many other plant parasites, a gene-for-gene relationship is postulated as a model for the interaction between cyst nematode genotypes and potato clones (Bakker et al. 1993; Thompson 1994). Several potato resistance genes have been identified and will probably be cloned in the near future (Leister et al. 1997; Rouppe van der Voort et al. 1999). However, very few data are available on the genetics of virulence in cyst nematodes. The ability to construct a genetic map is a prerequisite for the analysis of virulence traits in potato cyst nematodes.

By far the best characterized nematode genome is that of *Caenorhabditis elegans*: its genetic map contains more than 1000 mutationally defined genes (Brenner 1974; Waterston et al. 1997). *C. elegans* has many attractive features for genetic research; it has a 3-day generation time, it reproduces either by self- or by cross-fertilization, and strains can be grown on agar plates containing *Escherichia coli* as the sole food source. These features do not apply to the potato cyst nematode. Under natural conditions *G. rostochiensis* has one life cycle per year. It is an obligate biotrophic and amphimictic plant parasite. Inbred lines cannot be raised and a diapause seriously hinders mass culture under laboratory conditions and thus makes the development of homozygous populations difficult.

To circumvent our inability to analyze individual nematodes, a mapping procedure was developed which is based on segregation analyses of a series of sibling populations. These sibling populations consisted of the encysted progenies formed as a result of random mating. The strict physical separation of the progenies in the cysts and the ability to analyze single cysts allowed us to follow the segregation patterns of molecular marker alleles.

As a consequence of the non-inbred nature of the offspring, the family of sibling populations is not comparable to the commonly used F2, BC1 and RIL families. Nevertheless, the maximum likelihood estimator of the recombination frequency was shown to be identical to the estimator normally used to analyze F2 populations. This novel strategy enables the generation of linkage maps of non-inbred species in cases where individual genotypes are not amenable to marker analysis. Estimates of the physical and genetic sizes of the *Globodera* genome revealed a low kilobase/centimorgan (kb/cM) ratio. These attributes are both necessary and sufficient for the isolation of genes involved in the intimate interaction between cyst nematode and host plant.

Materials and methods

Nematode crosses

The parental populations Ro_1-19 and Ro_5-22 were selected from controlled matings using individuals from the natural populations Ro_1-Mierenbos and Ro_5-Harmerz, respectively (Janssen et al. 1990). Females were obtained by inoculating pre-parasitic juveniles singly onto the roots of *Solanum tubersosum* cv. Eigenheimer grown in petri dishes on water agar. In PCNs, sex is determined epigenetically, and females will develop if the inoculum density is low, e.g., one second-stage juvenile per root tip. Males were reared in pots and extracted from the soil with an Oostenbrink elutriator (Oostenbrink 1960). Controlled matings were performed in petri dishes as described by Mugniery (1982). The plates were kept at 21° C. Four to six weeks after inoculation, bright, white-colored females were induced to mate by placing one male on the gelatinous matrix of the female. Two months later dark-brown cysts were collected and stored at 4° C for at least 9 months. Although the diapause can be circumvented by avoiding desiccation of the cysts (Janssen et al. 1987), we observed a reduction in the reproduction rate if Ro_1-19/Ro_5-22 hybrids were used immediately for the production of F1 females.

The mating scheme (Fig. 1) involved two generations of controlled crosses and one generation of random mating. The first generation was obtained from a controlled cross performed between a male of population Ro_5-22 and a female of population Ro_1-19 (generation P1). For the second generation, an F1 hybrid female was backcrossed with a Ro_1-19 male. As individual nematodes cannot be analyzed separately, the BC1 population hatched from the cyst was propagated in a pot containing the susceptible potato cultivar Eigenheimer. Random mating (RM) was assumed between the individuals of the BC1 population during this multiplication step. The result of each mating was a cyst. Thus, the family of sibling cyst populations consists of BC1RM1 cysts.

The BC1RM1 family used in this study consisted of 107 cysts; the highest number found among a set of 36 such families examined. Before production of the BC1RM1 progeny, the numbers of eggs within BC1 cysts were counted to determine the multiplication rate for each BC1 cyst. Out of a BC1RM1 progeny of 107 cysts, 56 individual cysts allowed successful preparation of AFLP templates. Hence, the mapping population is represented by 56 BC1RM1 cysts.

Estimates of recombination frequencies

To estimate the frequency of recombination between the hypothetical dominant marker alleles M_1 and M_2 in the meiosis of the F1 hybrid female based on the distribution of phenotypic markers in

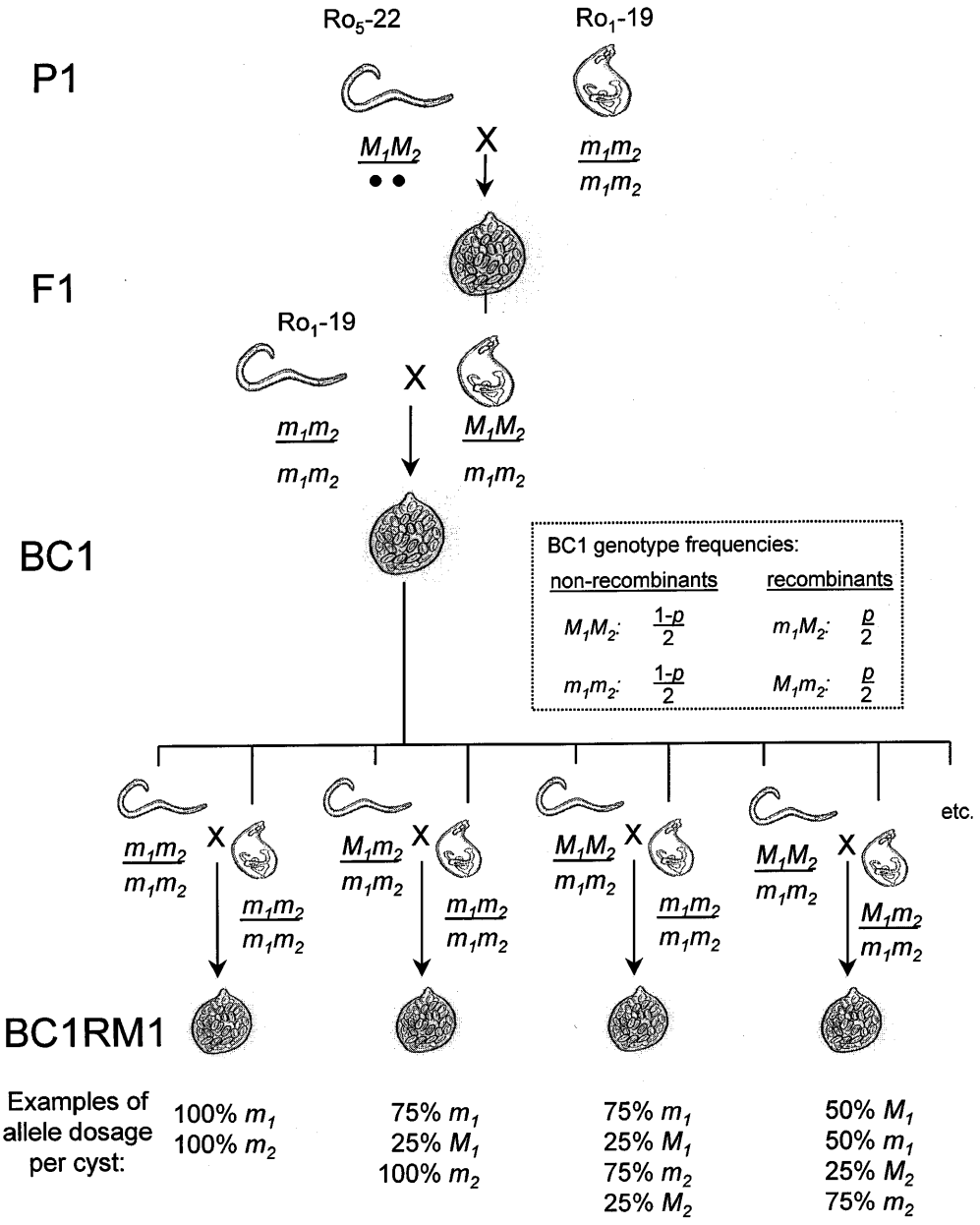

Fig. 1 Crossing scheme used to generate a BC1RM1 progeny. Males are represented as *vermiform objects*, females are *pear-shaped*, whereas egg-containing cysts are depicted as *ovals*. M_1 and M_2 are hypothetical Ro_5-22-specific markers that are linked in coupling phase. Throughout this scheme, the genotypes for the hypothetical loci M_1 and M_2 are indicated. After a controlled mating between a Ro_1-19 female and a Ro_5-22 male (generation P1), the F1 females (hatched from the F1 cyst) are backcrossed with a Ro_1-19 male. Random mating between the BC1 nematodes generates the BC1RM1 family of sibling populations. Recombination between markers M_1 and M_2 is estimated on the basis of the marker phenotypes observed in the BC1RM1 cysts

the mapping population (Table 1), a mating table was constructed (Table 2) in which only the Ro_5-22 specific markers were considered. As a consequence of the fact that a non-inbred species was used, the linkage phase of the markers is variable. A backcross was therefore included in the mating scheme, to ensure that all AFLP alleles segregating in the BC1RM1 family are in coupling phase. Knowledge of the linkage phase is essential for proper estimation of the recombination frequency. One result of the backcross is that the Ro_5-22 specific markers alone account the mating table.

Due to this backcross, the maternal alleles M_1 and M_2 segregate in the BC1 genotypes. These BC1 genotypes mate randomly and, depending on the BC1 genotype, up to four different types of gamete are produced (Table 1). The phenotype of the BC1RM1 cyst is the result of a single random mating and is determined by the combination of the paternal and maternal gamete pools involved (Table 2). As a consequence, the frequencies of the bulked gamete phenotypes are the same as the genotype frequencies of these (BC1) parents. By pairwise multiplication of the BC1 genotype frequencies (Table 1), the formulas for expected frequencies of the BC1RM1 phenotypes were obtained (Table 3). These formulas are similar to the formulas that describe a cross of the type $M_1M_2/m_1m_2 \times M_1M_2/m_1m_2$ with markers M_1 and M_2 in coupling phase (Allard 1956). Solving the maximum likelihood equation for the recombination frequency p (Fisher 1921) results in an estimator of p (Allard 1956; Ritter et al. 1990; Ritter and Salamini 1996). It is concluded that despite the aberrant design of the crosses, recombination frequencies can be estimated as described for an F2 population.

DNA manipulations

DNA was extracted from single cysts as follows. A clean, acetone-washed cyst was transferred to a fixation dish (Seinhorst 1962) in a humid chamber and soaked in 75 µl of deionized, double distilled water overnight. The cyst was crushed to release the eggs or J2 juveniles and the cyst wall was discarded. The mixture was centrifuged for 90 s and the pellet was washed twice with 100 µl of distilled water. The pellet was frozen for 30 min at −80° C and thawed at room temperature. This freeze-thaw procedure was repeated twice. Subsequently, 200 µl of Proteinase K buffer (0.1 M

Table 1 Parameters of the BC1RM1 mapping population used to derive the estimator of the recombination frequency p for a family of BC1RM cysts

BC1 genotypes[a]	BC1 genotype frequency	BC1 phenotypes[b]	Gametes
M_1M_2/m_1m_2	$(1-p)/2$	M1M2	M_1M_2, m_1m_2, M_1m_2, m_1M_2
m_1m_2/m_1m_2	$(1-p)/2$	m1m2	m_1m_2
M_1m_2/m_1m_2	$p/2$	M1m2	M_1m_2, m_1m_2
m_1M_2/m_1m_2	$p/2$	m1M2	m_1M_2, m_1m_2

[a] M_1 and M_2 are visible alleles of loci *M1* and *M2* for which m_1 and m_2 are null alleles
[b] Because the BC1RM1 phenotypes represent the bulked offspring of a single mating, the BC1 genotype frequency is similar to the (pooled) gamete frequency

Table 2 Mating table for the BC1RM1 population

Gametes	BC1RM cyst phenotypes[a]			
	M_1M_2, m_1m_2, M_1m_2, m_1M_2	m_1m_2	M_1m_2, m_1m_2	m_1M_2, m_1m_2
M_1M_2, m_1m_2, M_1m_2, m_1M_2	M1M2	M1M2	M1M2	M1M2
m_1m_2	M1M2	m1m2	M1m2	m1M2
M_1m_2, m_1m_2	M1M2	M1m2	M1m2	M1M2
m_1M_2, m_1m_2	M1M2	m1M2	M1M2	m1M2

[a] By multiplying the BC1 genotype frequencies for each parental genotype combination (see Table 1), the probabilities for the BC1RM1 cyst phenotypes can be obtained (see Table 3).

Table 3 Expected and observed frequencies for the occurrence of BC1RM1 cyst phenotypes

Recombination frequency	BC1RM1 cyst phenotypes			
	M1M2	m1m2	M1m2	m1M2
p_j[a]	$(3-2p+p^2)/4$	$(1-p)^2/4$	$p(2-p)/4$	$p(2-p)/4$
Complete linkage ($p = 0$)	3	1	–	–
Absence of linkage ($p = 0.5$)	9	1	3	3

[a] p_j is the probability for the occurrence of the marker phenotype j where p_j is a function of the recombination frequency p. The formulas for p_j can be obtained by multiplying the corresponding BC1 genotype frequencies given in Table 1. These formulas are similar as obtained for the expected frequencies of marker phenotypes in an F2 population with alleles M_1 and M_2 in coupling phase (Allard 1956)

TRIS-HCl pH 8.5, 0.1 M EDTA, 0.5% SDS) and 5 µl of Proteinase K (20 ng/µl) was added and the mixture was incubated at 50° C for 4 h or overnight. After extraction first with phenol and then with chloroform/isoamylalcohol, the DNA was precipitated by adding 480 µl of 96% ethanol and 40 µl of 0.1 M ammonium acetate and incubating at −20° C for 1 h. The mixture was centrifuged for 10 min at 13,000 rpm. The pellet was washed in 70% ethanol, dried and dissolved in 20 µl of TE (Sambrook et al. 1989). For AFLP template preparation, 10 µl of DNA was used, and 3 µl of fresh 5 × restriction-ligation buffer (Vos et al. 1995), 0.1 µl (2.5 U) of *Eco*RI, 0.12 µl (2.5 U) of *Mse*I and 1.78 µl of distilled water was added. After incubation at 37° C for 1 h, an adapter-ligation mix was added containing 0.5 µl of the *Eco*RI (5 pmol/µl) and the *Mse*I adapter (50 pmol/µl), and the AFLP procedure was continued as described (Vos et al. 1995).

For the selective amplification of restriction fragments, "*Eco*RI-primers" (E-primers) and "*Mse*I-primers" (M-primers) were used (Vos et al. 1995). A total of four or five selective nucleotides for each primer pair was used – indicated by the primer codes "E + 2/M + 2" and "E + 2/M + 3", respectively.

Genomic DNA from the parental populations was isolated from approximately 6000 young females, harvested 5–6 weeks after inoculation (Roosien et al. 1993). AFLP template preparation was performed using 50 ng of this DNA (Vos et al. 1995).

Construction and analysis of the genetic map

Linkage analyses of pairwise recombination frequencies between the markers were carried out using the computer program JoinMap version 1.4 (Stam 1993). The data were analyzed according to an F2-type segregation model and marker genotypes were encoded as either "C" (presence of a marker) or "A" (absence of a marker). Markers which were linked to other markers at LOD values between 1.0 and 3.0 were subjected to a Chi-square test of independence by means of 2 × 2 contingency tables. These tests were applied using the program Linkage-1 (Suiter et al. 1983). Markers were considered to be linked if the Chi-square test criterion exceeded 8.0 (Ellis et al. 1992). To correct for spurious linkage due to skewed segregation, this test was also applied for linked markers which deviated from the expected segregation ratio of 3:1. A Chi-square test is more conservative than the log-likelihood method since the production of 2 × 2 contingency tables results in two additional degrees of freedom to test for significant linkage. Recombination frequencies were converted into map units (cM) using the Kosambi function. Graphic representations of the maps were made by the computer program Drawmap (Van Ooijen 1994).

Clustering of markers was tested by a Chi-square test for goodness of fit at $P \geq 0.05$ (Rouppe van der Voort et al. 1997).

Estimation of genome coverage

An estimate of the total length of the genetic map of *G. rostochiensis* was obtained by the method of Hulbert et al. (1988). This method is based on calculation of the probability P that a pair of linked markers randomly chosen from a genome of size G has a LOD score ≥ a threshold value T. An estimate of P is given by the number of marker pairs obtained at a LOD value equal to T and the theoretical number of marker pairs expected for a given number of marker loci. The genome size G follows from the expression $2x/P$, where x is the genetic distance obtained from the expected LOD

score. An MS Excel spreadsheet for obtaining map length estimates by this method is available upon request.

Estimation of the physical parameters of the *Globodera* genome

Estimates of the physical characteristics of the *G. rostochiensis* genome were obtained by determining the number of AFLP fragments amplified with a given primer combination and comparing this with the number of AFLP fragments obtained from the *C. elegans* genome. For *C. elegans*, AFLP fragments were generated from the strain N2 Bristol (kindly provided from H. van Luenen, The Netherlands Cancer Institute, Amsterdam, The Netherlands) using 64 primer combinations. Thirty-two primer combinations were obtained with AFLP primers E + GA or E + AA and *Mse*I primers with two selective nucleotides at their 3′ end. Another 32 combinations were made between primer E + GA and all possible M + 3 primers. The effects of species, number of selective nucleotides and G + C content of the primer extension on the total number of AFLP fragments amplified were analyzed by multiple linear regression (Genstat 5 release 3.2, Payne et al. 1987).

Results

Selection of AFLP markers

AFLP markers specific for parental population Ro_5-22 were selected by testing 70 E + 2/M + 2 and E + 2/M + 3 primer combinations on template DNA of populations Ro_5-22 and Ro_1-19. Because the AFLP template is based on a population of nematodes in which markers can have an allele frequency as low as $p = 0.25$ (Fig. 1), a third template – a 9:1 mixture of Ro_1-19/Ro_5-22 template DNA – was included to permit us to discard Ro_5-22 specific AFLP markers that might not give a clear amplification product in the genetic background of Ro_1-19 alleles. An example of a primer combination prescreen and AFLP genotyping of the progeny is shown in Fig. 2.

Out of a total of 7104 DNA fragments amplified from population Ro_5-22, about 7% were absent in Ro_1-19; i.e. on average, 7.6 Ro_5-22 markers per primer combination (SD = 3.6). About 60% of these Ro_5-22-specific markers could either not be readily amplified from the 9:1 template mixture or could not be reliably scored due to co-migrating fragments. Taking this into account, 12 AFLP primer combinations were selected which amplified at least three Ro_5-22 markers per primer combination.

Using these primer combinations, 66 segregating AFLP markers were reproducibly amplified from the BC1RM1 progeny. Fifteen AFLP markers did not segregate – they were present in all BC1RM1 cysts. This may be the result of the coincidental presence of the same Ro_5-22 marker allele in the Ro_1-19 F1 male, or indicate that the segregation ratio is strongly skewed towards Ro_5-22 alleles. Eighteen AFLP markers could not be recovered from the progeny cyst fingerprints, despite their presence in the 9:1 mixtures. This can be explained if it is assumed that these markers (1) were not transmitted by the P1 Ro_5-22 male (Fig. 1), (2) were subject to strong negative selection, or (3) could not reliably be amplified from the BC1RM1 cysts. The last case is not inconceivable, because below an absolute threshold of starting material (approximately 25 pg of DNA) amplification products may remain below the detection limit (Vos et al. 1995). The appearance of template-independent amplification products in some of the AFLP profiles of single cysts indicates that the protocol for single-cyst AFLP is close to the limit of the

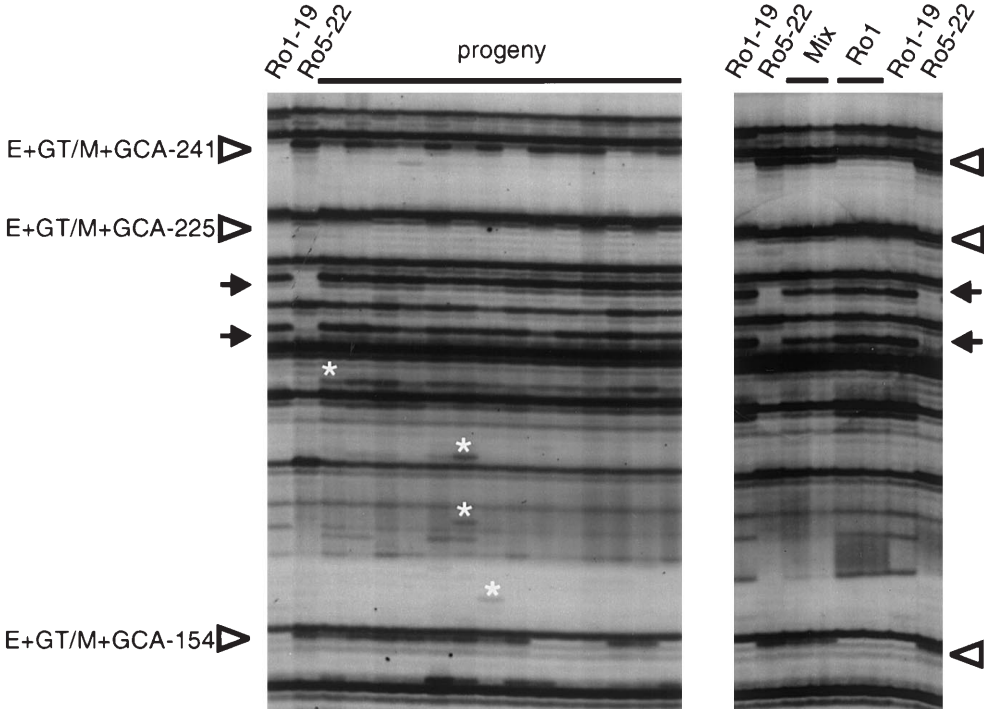

Fig. 2 Part of an autoradiogram showing AFLP fingerprints generated using the primer combination E + GT/M + GCA. On the left are AFLP markers amplified from the parental lines Ro_1-19 and Ro_5-22 and a subset of the BC1RM1 progeny. Ro_5-22 specific markers E + GT/M + GCA-241, E + GT/M + GCA-225 and E + GT/M + GCA-154 are indicated. Ro_1-19 specific markers that are present in all BC1RM1 progeny cysts are indicated by *small arrows*. Template-independent fragments are marked by an *asterisk*. On the right are AFLP profiles generated from the parental lines, two 9:1 Ro_1-19/Ro_5-22 template mixtures, and two templates prepared from field population Ro_1-Mierenbos to screen for reproducible amplification at low allele frequencies of Ro_5-22 specific markers

sensitivity of the AFLP technique (Fig. 2). AFLP fingerprints which involve templates prepared from such minute amounts of genomic DNA, may contain template-independent amplification products (Vos et al. 1995). The occurrence of such fragments was both rare and unpredictable and did not compromise marker analysis.

Construction and analysis of the genetic map

At first glance the BC1RM1 mapping population would seem to require the use of different estimates of recombination frequencies from those commonly used to analyze F2 or BC1 mapping populations derived from inbred parents (Fig. 1). However, evaluation of the frequencies at which the different marker phenotypes can occur (Table 1) revealed that recombination frequencies can be estimated using the maximum likelihood estimator for two markers in coupling phase segregating in an F2 population. Therefore, the data set could be analyzed like an F2-type population using JoinMap.

Out of the 66 segregating markers, 62 (94%) were incorporated into nine linkage groups. Four markers could not be assigned to any linkage group. This number of linkage groups corresponds to the haploid number of chromosomes in *G. rostochiensis* (Riley and Chapman 1957; Rouppe van der Voort et al. 1996). The linkage groups were established by stepwise lowering of the LOD score from 8.0 to 2.0. Seven linkage groups could be discriminated at a LOD of 6.0. Lowering the linkage threshold to 3.0 resulted in the addition of two linkage groups (Fig. 3). Expansion of the map was observed without agglomeration of linkage groups.

The nine linkage groups of the *G. rostochiensis* map range in size from 20 cM to 85 cM, with a total map distance of 431 cM. The average distance between the markers is 8 cM, as determined by the dividing the total map distance by the number of markers minus the number of linkage groups. As tested by a Chi-square test for goodness of fit, no clustering of markers was observed.

Proper analysis of the genetic constitution of the BC1 generation presupposes monogamy. This was tested by investigating the level of segregation distortion of the markers. Since 75% of the progeny was expected to contain a Ro_5-22 marker (Fig. 1), insemination of one female by multiple males would tend to increase the fraction of the progeny bearing such a Ro_5-22 marker. Deviations from the expected Mendelian ratio were observed for 19 (29%) of the segregating markers (tested by a Chi-square test at $P \leq 0.05$). However only 8/19 of the aberrantly segregating marker loci showed a bias towards the presence of Ro_5-22 alleles. These markers mapped primarily to two regions of the genome. On linkage group *A*, Ro_5-22 alleles were under-represented, possibly due to selection. On this linkage group, segregation ratios varied between 1:1 and 8:1, where a 1:3 ratio for absence vs. presence of an AFLP marker was expected. For three markers on linkage group *G*, Ro_5-22 markers were over-represented; with segregation ratios varying between 1:6 and 1:8. Hence, no genetic evidence was obtained for the fertilization of single PCN females by multiple males.

Genome coverage

The finding that virtually all markers mapped to one or other of the linkage groups indicated that this set of markers covers a significant proportion of the genome of the potato cyst nematode. An estimate of the maximum genetic length of the genome was obtained using the method of Hulbert et al. (1988). Since no clustering was observed on the map all 66 markers were taken into consideration. With a mapping population of 56 genotypes, 253 marker pairs were found within 34 cM at LOD 1.0, whereas 170 marker pairs were linked within a distance of 26 cM at LOD 3.0. This corresponds to estimates for *G* of 629 cM and 676 cM, respectively. Taking these data together, the genome size of *G. rostochiensis* is likely to be approximately 650 cM. This implies that the markers presented here cover about 65% of the genome of the potato cyst nematode.

Estimates of genome size and G + C base composition

Estimates of the physical parameters of the *G. rostochiensis* genome were obtained by a comparative analysis with the *C. elegans* genome. Theoretically, the total number of AFLP fragments obtained per primer combination depends on the number of selective nucleotides and the number of GC residues in the extension of the primers. Since the physical size and the G + C base composition of the *C. elegans* genome are known (Sulston and Brenner 1974), a comparative AFLP analysis may provide a relative measure of these parameters for the *G. rostochiensis* genome.

In Fig. 4, numbers of amplification products obtained with varying G + C contents in the primer extensions are presented for both nematode species. For *C. elegans* strain N2 Bristol, a total of 6773 AFLP fragments were amplified using a similar set of primers to that used to analyze the *G. rostochiensis* genome. AFLP fingerprints were not taken into account when fragment numbers differed by more than three standard deviations from the mean number of fragments for a given G + C content of the primer extension. Such extreme differences were observed infrequently and are presumably caused by the relative paucity or abundance of particular sequence motifs. A multiple linear regression analysis showed that both the number of selective nucleotides and the G + C content of the extension had a significant effect on the numbers of AFLP fragments amplified (at $P \leq 0.001$, determined by a *t*-test). Similar intercepts (genome size) and slopes (G + C content) for the regression lines of both species were obtained for

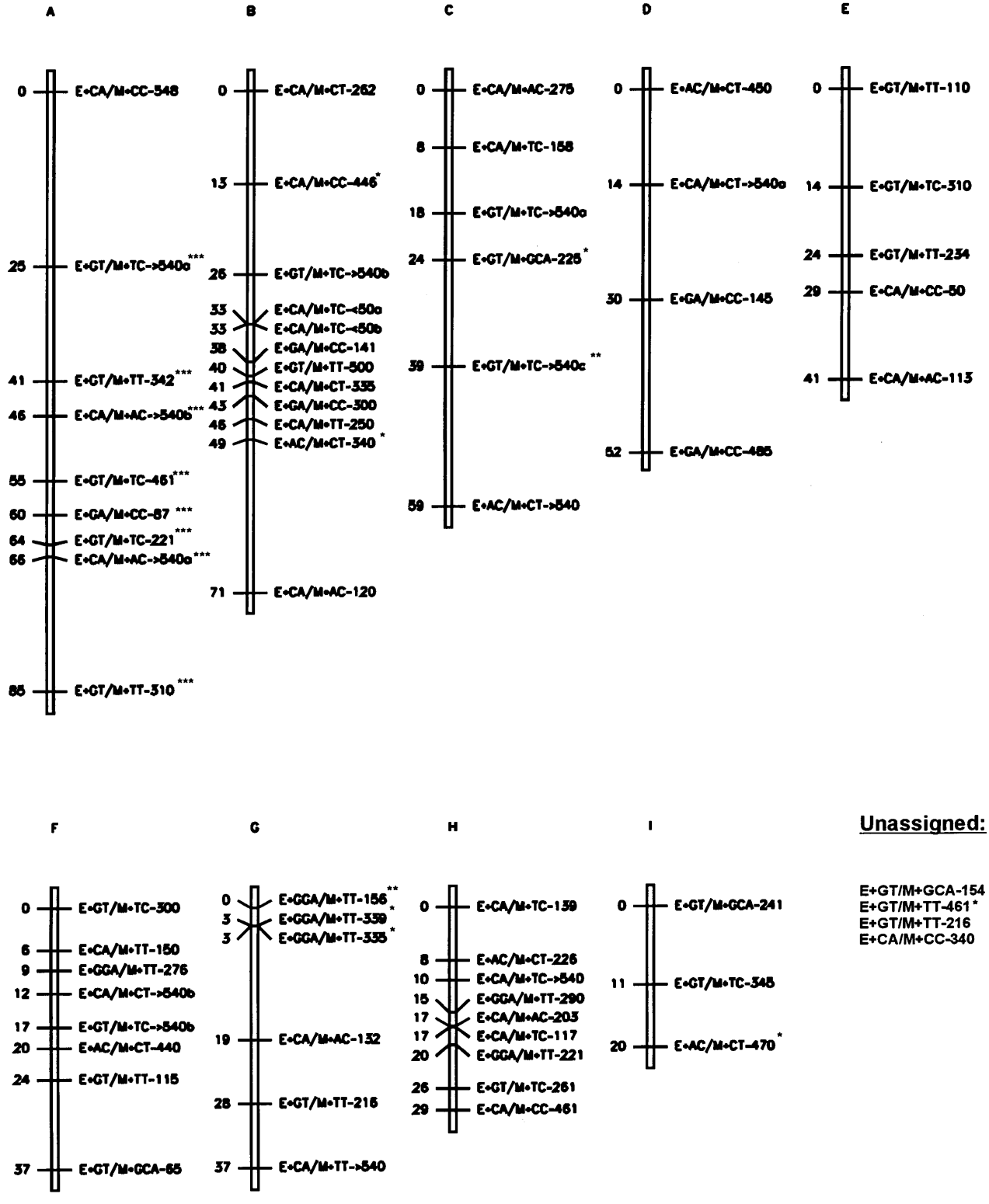

Fig. 3 Genetic linkage map of *G. rostochiensis*. Markers are designated by the enzyme combination (E and M), followed by the selective nucleotides used (+NNN) and a relative measure for the molecular weight of the AFLP fragment. All markers are mapped with a LOD score of greater than 3.0, with the exception of markers E + CA/M + AC-120 (linkage group B), E + GA/M + CC-145 and E + GA/M + CC-485 (linkage group D), E + GT/M + GCA-65 (linkage group F) and E + AC/M + CT-470 (linkage group I). These markers were mapped with LOD scores between 1.0 and 3.0 and assigned to linkage groups by means of a Chi-square test for independence with a Chi-square value above 8 (Ellis et al. 1992). Markers that deviated from the expected 3:1 segregation are marked with *asterisks* (* indicates deviation at $P = 0.01–0.05$; **, deviation at $P = 0.001 – 0.01$; ***, deviation at P much less than 0.001). Four markers remained unassigned

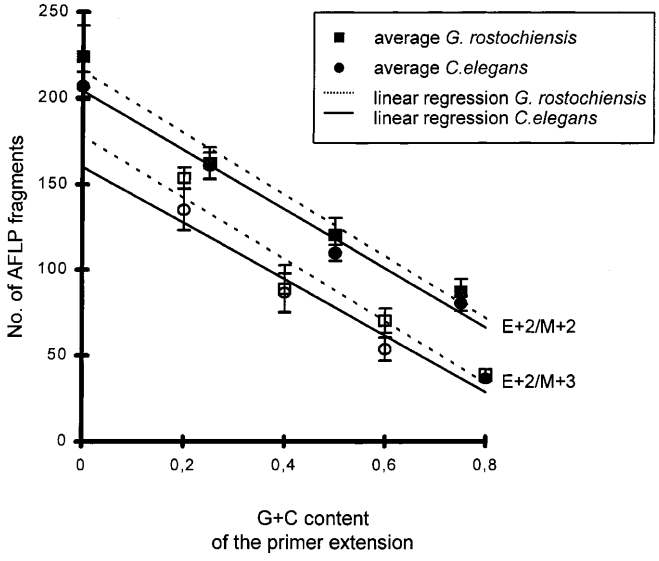

Fig. 4 Multiple linear regression of the species effect, the G+C contents and the numbers of selective nucleotides in the AFLP primer on the numbers of AFLP fragments obtained. The regression lines can be considered to be parallel as there is no interaction between species effect and G + C content. The similarity of the intercepts and slopes of the regression lines for both types of primer combinations indicate that the size and G + C content of the *G. rostochiensis* genome is similar to that of *C. elegans*

both primer extension lengths (Fig. 4). In addition, no statistical interaction ($P \geq 0.05$) was found between G + C content and species effect, which shows that the regression lines for both species can be considered to be parallel. This indicates that the size and average G + C composition of the *G. rostochiensis* genome are similar to the estimates obtained for the *C. elegans* genome -8×10^7 bp for haploid genome size and a G + C content of 36% (Sulston and Brenner 1974).

Discussion

A linkage map for the potato cyst nematode was constructed by means of a "pseudo F2" mapping strategy. This novel approach enabled us to generate a linkage map for a non-inbred species in which individual genotypes are not amenable to extensive marker analysis. We have chosen the name "pseudo F2" because the experimental design for analysis of a family of sibling populations allowed for the use of the same estimator of recombination frequencies as is applicable to a normal F2. The mapping procedure differed from that for an F2 in at least three aspects. Firstly, the mapping was not based on recombination in both male and female meiosis, but only in the meiosis of the female F1 hybrid. Secondly, a backcross was introduced to ensure that all AFLP alleles segregating in the BC1RM1 were in the same linkage phase. Thirdly, data on marker segregation were derived from the analysis of sibling populations instead of individual genotypes.

Segregation analysis of markers in families of sibling populations can be performed by use of F2 estimators if: (1) random mating is assumed, (2) parental-specific markers are identified and (3) the linkage phase of these markers is known. Linkage analysis between dominant markers in coupling phase in an F2 is straightforward. A simulation study on the informativeness of an F2 population has shown that for two dominantly segregating markers in coupling phase, linkage can reliably be detected for recombination frequencies up to 0.2, even within a progeny of about 50 in size (Maliepaard et al. 1998). Moreover, the advantage of an F2 estimator of recombination frequencies lies in the fact that it is then possible to use the algorithms in existing mapping software.

The level of polymorphism between the parental populations appeared to be sufficient for construction of a genetic map. An average of seven Ro_5-22 specific amplification products were observed per AFLP primer combination. However, since only one linkage phase could be analyzed for segregation, a substantial proportion of the Ro_5-22 specific AFLP markers could not be used in the experimental design. In addition, the amplification of Ro_5-22 AFLP markers from the offspring bulks may be hampered by the occurrence of competing priming sites derived from the Ro_1-19 genome and/or differences in allele frequencies of the marker in the bulks. At segregating marker loci, allele frequencies of 0.25 and 0.5 (Fig. 1) should be clearly detectable. To ensure reliable discrimination between presence and absence of a Ro_5-22 marker allele in a bulked BC1RM1 progeny, a control sample equivalent to an allele frequency of 0.1 was included. This pre-screen resulted in the selection of 12 AFLP primer combinations which identified 66 segregating markers in the BC1RM1 family. In retrospect, it is clear that because they exhibit stronger selectivity in the amplification reaction, the use of more E + 2/M + 3 primer combinations would make detection of informative primer combinations more efficient. Hence, attempts to generate a saturated map of *G. rostochiensis* will primarily be based on the use of E + 2/M + 3 primer combinations.

The initial presumption of monogamy was justified, since no overall bias for the occurrence of Ro_5-22 alleles in the population was detected. A feature which may prevent insemination of one female by multiple males is the formation of a so-called copulatory plug which prevents multiple inseminations (Barker 1994). "Plugging" and "non-plugging" races have been found for *C. elegans* in which males of a plugging race cause the formation of gelatinous material over the vulva of the recipient female after mating (Hodgkin and Doniach 1997). For other genera of nematodes, copulatory plugs have also been described (Baird et al. 1994). However, the presence of a gelatinous matrix on the vulva of virgin PCN females makes it difficult to discern whether PCN females produce a copulatory plug after mating.

The instances of segregation distortion encountered may be explained by the differential viability of the

segregating genotypes. Before production of the BC1RM1 progeny, the numbers of eggs in 36 BC1 cysts were counted (average 420 eggs/cyst; SD = 171) and compared with the numbers of new BC1RM1 cysts. The BC1RM1 progeny sizes ranged between 0 and 114 with an average of 35 (SD = 35). Differences in cyst contents were also observed within a BC1RM1 population, which largely determined the success rate of the AFLP template preparation. The largest, healthiest BC1RM1 family was selected for further analysis. Out of a total of 107 cysts, only 56 cysts allowed the preparation of AFLP template. Cysts with the lowest numbers of eggs invariably failed to yield a usable AFLP template. Differences in viability between BC1RM1 cysts may be the result of inbreeding depression; an inbreeding coefficient of 5/16 applies for the BC1RM1 family. Janssen et al. (1991) observed a similar reduction in the vitality of the selected populations Ro_1-19 and Ro_5-22, as compared to F1 juveniles derived from controlled matings between these populations. Inbreeding depression may be caused by the combination of lethal (recessive) alleles. Likely genomic positions for such loci are located on linkage group *A* which harbors the largest number of marker loci showing skewed segregation. Selection processes may also occur during the life cycle of the nematode that favor the growth of certain genotypes under the artificial conditions of laboratory culture.

The linkage map of the *G. rostochiensis* genome includes 62 markers which are randomly distributed over the genome. Four markers could not be assigned to any genomic region. The 62 markers fall into nine linkage groups, a number similar to the haploid chromosome number of the potato cyst nematode (Riley and Chapman 1957; Grisi et al. 1995; Rouppe van der Voort et al. 1996). The random distribution of markers on the map indicates a homogenous distribution of recombination events along the nine chromosomes. This random distribution of markers on the map is in contrast to the severe clustering of markers observed on the maps of many plant genomes (e.g. Tanksley et al. 1992; Kesseli et al. 1994; Rouppe van der Voort et al. 1997). Clustering of markers is often attributed to suppression of recombination in centromeric regions (Sherman and Stack 1995). The mode of chromatid separation during meiosis suggests that the *G. rostochiensis* chromosomes lack a localized centromere (Rouppe van der Voort et al. 1996), which may explain the random distribution of the markers along the map.

Based on the marker segregation data, a rough estimate of 650 cM was obtained for the total genetic length of the *G. rostochiensis* genome. The physical size of the *G. rostochiensis* genome was estimated by a comparative analysis of the number of AFLP fragments produced, using the *C. elegans* genome as a reference. The size and G + C content of the *C. elegans* genome has been determined by DNA renaturation kinetics (Sulston and Brenner 1974) and will be known precisely when the current genome sequencing project is finished (The *C. elegans* Sequencing Consortium 1998). Similar numbers of bands per fingerprint were obtained for both species, irrespective of the experimental conditions applied. This suggests that the haploid genome sizes and G + C contents for both nematodes are highly comparable. These estimates of the physical parameters of the *G. rostochiensis* genome do not differ much from estimates obtained for the genome of the related plant parasitic nematode species *Meloidogyne* (Pableo et al. 1988; Pableo and Triantaphyllou 1989). The estimates of the physical and genetic sizes of the *Globodera* genome yield an overall ratio of 120 kb of DNA per cM genetic distance.

This type of comparative analysis appears to be a convenient method for estimating the physical parameters of the *Globodera* genome. When large AFLP data sets are available for a particular organism, the generation of fingerprints with the same primer combinations for a reference organism is straightforward. As the thousands of AFLP fragments generated are distributed over the genome, these numbers are directly correlated with the genome size and G + C base composition of the genome. Moreover, the use of many different primer combinations provides insights into the relative abundance or scarcity of particular sequence motifs. For example, using primer combination E + GA/M + CT only 39 AFLP fragments were produced for *G. rostochiensis*, where an average number of 120 AFLP fragments is produced by E + 2/M + 2 combinations with a G + C content of 0.5 (Fig. 2). At the other extreme is primer combination E + GA/M + ACC, with which 170 AFLP fragments were generated for *G. rostochiensis* (the average number of AFLP fragments produced for this type of primer combination is 90). The comparative method provides a relatively rapid estimate of the physical properties of the genome. However, in addition to information on size and base composition, renaturation kinetics provides an estimate of the proportion of repetitive DNA present in the genome.

The ability to construct a genetic map for the potato cyst nematode opens the way for map-based cloning in this organism. The genome size is small, and sufficient recombination occurs to allow fine-scale mapping of any gene of interest. Although the generation of sufficiently large mapping populations for *Globodera* is a difficult task, generation of marker data from multiple BC1RM1 populations and subsequent analysis using JoinMap offers the opportunity to merge segregation data from different populations. This should allow the construction of a map with a resolution in the sub-centimorgan range. As *Globodera* has an estimated kilobase/centimorgan ratio of 120 kb/cM, this implies that one unit of recombination is equivalent to a DNA fragment length which can now be readily cloned in bacterial artificial chromosomes (BACs; Shizuya et al. 1992; Zhang et al. 1996). Simple extrapolation from units of recombination to distances in base pairs can be misleading, due to regional differences in recombination frequencies. However, the random distribution of markers on the *G. rostochiensis* map indicates that the majority of the genome is accessible for isolating genomic clones with which to construct overlapping contigs.

The coding capacity of nematode DNA and the fact that the majority of nematode genomes analyzed so far contains a small proportion of repetitive DNA (Hammond and Bianco 1992) should facilitate the construction of BAC contigs. Potato cyst nematodes are an important pest of potato, and a number of resistance genes are now in the process of being cloned from the plant (Ganal et al. 1995; Leister et al. 1997; Rouppe van der Voort et al. 1999). The approach described in this paper opens the way to identifying the nematode counterparts with which the products of resistance gene interact. For example, a similar set-up would in principle allow for the identification of e.g. *AvrHs-1*$^{pro-1}$ in the beet cyst nematode *Heterodera schachtii* whose counterpart, the sugar beet resistance gene *Hs1*$^{pro-1}$, was recently cloned (Cai et al. 1997).

Acknowledgements We thank J. Verhaaf for technical assistance and T. van der Lee, P. Prins and T. Schipper for helpful advice. The research was funded in part by the European Community grant AIR3 92.0062.

References

Allard RW (1956) Formulas and table to facilitate the calculation of recombination values in heredity. Hilgardia 24:235–278

Baird SE, Sutherlin ME, Emmons SW (1992) Reproductive isolation in Rhabditidae (Nematoda: Secernentea): mechanisms that isolate six species of three genera. Evolution 46:585–594

Bakker J, Folkertsma RT, Rouppe van der Voort JNAM, De Boer JM, Gommers FJ (1993) Changing concepts and molecular approaches in the management of virulence genes in potato cyst nematodes. Annu Rev Phytopathol 31:169–190

Barker DM (1994) Copulatory plugs and paternity assurance in the nematode *Caenorhabditis elegans*. Anim Behav 48:147–156

Beckmann JS, Soller M (1990) Toward a unified approach to genetic mapping of eukaryotes based on sequence-tagged microsatellite sites. Biotechnology 8:930–932

Brenner S (1974) The genetics of *Caenorhabditis elegans*. Genetics 77:71–94

Cai D, Kleine M, Kifle S, Harloff HJ, Sandal NN, Marcker KA, Klein LR, Salentijn EJ, Lange W, Stiekema WJ, Wyss U, Grundler FW, Jung C (1997) Positional cloning of a gene for nematode resistance in sugar beet. Science 275:832–834

Dietrich WF, Miller J, Steen R, Merchant MA, Damron BD, Husain Z, Dredge R, Daly MJ, Ingalls KA, O' CT, Evans CA, DeAngelis MM, Levinson DM, Kruglyak L, Goodman N, Copeland NG, Jenkins NA, Hawkins TL, Stein L, Page DC, Lander ES (1996) A comprehensive genetic map of the mouse genome. Nature 380:149–152

Ellis TN, Turner L, Hellens RP, Lee D, Harker CL, Enard C, Domoney C, Davies DR (1992) Linkage maps in pea. Genetics 130:649–663

Fisher RA (1921) On the mathematical foundations of theoretical statistics. Philos Trans R Soc London A 122:309–368

Ganal MW, Simon R, Brommonschenkel S, Arndt M, Phillips MS, Tanksley SD, Kumar A (1995) Genetic mapping of a wide spectrum nematode resistance gene (*Hero*) against *Globodera rostochiensis* in tomato. Mol Plant-Microbe Interact 8:886–891

Gardiner EH, Coe S, Melia-Hancock D, Hoisington A, Chao S (1993) Development of a core RFLP map in maize using an immortalized F_2 population. Genetics 134:917–930

Grisi E, Burrows PR, Perry RN, Hominick WM (1995) The genome size and chromosome complement of the potato cyst nematode *Globodera pallida*. Fundam Appl Nematol 18:67–70

Hammond MP, Bianco AE (1992) Genes and genomes of parasitic nematodes. Parasitol Today 8:299–305

Harushima Y, Yano M, Shomura A, Sato M, Shimano T, Kuboki Y, Yamamoto T, Lin SY, Antonio BA, Parco A, Kajiya H, Huang N, Yamamoto K, Nagamura Y, Kurata N, Khush GS (1998) A high-density rice genetic linkage map with 2275 markers using a single F2 population. Genetics 148:479–494

Hauge BM, Hanley S, Cartinhour S, Cherry JM, Goodman HM, Koornneef M, Stam P, Chang C, Kempin S, Medrano L, Meyerowitz EM (1993) An integrated genetic/RFLP map of the *Arabidopsis thaliana* genome. Plant J 3:745–754

Hodgkin J, Doniach T (1997) Natural variation and copulatory plug formation in *Caenorhabditis elegans*. Genetics 146:149–164

Hulbert SH, Ilott TW, Legg EJ, Lincoln SE, Lander ES, Michelmore RW (1988) Genetic analysis of the fungus *Bremia lactucae*, using restriction fragment length polymorphisms. Genetics 120:947–958

Janssen R, Bakker J, Gommers FJ (1987) Circumventing the diapause of potato cyst nematodes. Neth J Plant Pathol 93:107–113

Janssen R, Bakker J, Gommers FJ (1990) Selection of virulent and avirulent lines of *Globodera rostochiensis* for the *H1* resistance gene in *Solanum tuberosum* ssp. *andigena* CPC 1673. Rev Nematol 13:265–268

Janssen R, Bakker J, Gommers FJ (1991) Mendelian proof for a gene-for-gene relationship between virulence of *Globodera rostochiensis* and the *H1* resistance gene in *Solanum tuberosum* ssp. *andigena* CPC 1673. Rev Nematol 14:213–219

Kesseli RV, Paran I, Michelmore RW (1994) Analysis of a detailed genetic linkage map of *Lactuca sativa* (lettuce) constructed from RFLP and RAPD markers. Genetics 136:1435–1446

Leister D, Berger A, Thelen H, Lehmann W, Salamini F, Gebhardt C (1997) Construction of a potato YAC library and identification of clones linked to the disease resistance loci *R1* and *Gro1*. Theor Appl Genet 95:954–960

Maliepaard C, Jansen J, Van Ooijen JW (1998) Linkage analysis in a full-sib family of an outbreeding species: overview and consequences for applications. Genet Res 70:237–250

Mugniery D (1982) A method for screening potatoes for resistance to *Globodera* spp. under laboratory conditions. In: Hooker WJ (ed) Research for the potato in the year 2000. CIP, Lima, pp 137–138

Murray JC, et al. (1994) A comprehensive human linkage map with centimorgan density. Science 265:2049–2054

O'Brien SJ (1993) Genetic maps: locus maps of complex genomes. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Oerke EC, Dehne HW, Schönbeck F, Weber A (1994) Crop production and crop protection: estimated losses in major food and cash crops. Elsevier, Amsterdam

Oostenbrink M (1960) Estimating nematode populations by some selected methods. In: Sasser JN, Jenkins WR (eds) Nematology. University of North Carolina Press, Chapel Hill, pp 85–102

Pableo EC, Triantaphyllou AC (1989) DNA complexity of the root-knot nematode (*Meloidogyne* spp.) genome. J Nematol 21:260–263

Pableo EC, Triantaphyllou AC, Kloos WE (1988) DNA isolation and GC base composition of four root-knot nematode (*Meloidogyne* spp.) genomes. J Nematol 20:1–8

Payne RW, Lande PW, Ainsley AE, Bicknell KE, Digbey PGN (1987) Genstat 5 Reference Manual. Clarendon Press, Oxford, UK

Qi X, Stam P, Lindhout P (1998) Use of locus-specific AFLP markers to construct a high-density molecular map in barley. Theor Appl Genet 96:376–384

Reiter RS, Williams JK, Feldmann KA, Rafalski JA, Tingey SV, Scolnik PA (1992) Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNAs. Proc Natl Acad Sci USA 89:1477–1481

Riley R, Chapman V (1957) Chromosome of the potato root eelworm. Nature 180:662

Ritter E, Salamini F (1996) The calculation of recombination frequencies in crosses of allogamous plant species with applications to linkage mapping. Genet Res 67:55–65

Ritter E, Gebhardt C, Salamini F (1990) Estimation of recombination frequencies and construction of RFLP linkage maps in plants from crosses between heterozygous parents. Genetics 125:645–654

Roosien J, Van Zandvoort P, Folkertsma RT, Rouppe van der Voort JNAM, Goverse A, Gommers FJ, Bakker J (1993) Single juveniles of the potato cyst nematodes *Globodera rostochiensis* and *G. pallida* differentiated by randomly amplified polymorphic DNA. Parasitology 107:567–572

Rouppe van der Voort JNAM, Van Enckevort E, Pijnacker LP, Helder J, Gommers FJ, Bakker J (1996) Chromosome number of the potato cyst nematode *Globodera rostochiensis*. Fundam Appl Nematol 19:369–374

Rouppe van der Voort JNAM, Van Zandvoort P, Van Eck H, Folkertsma RT, Hutten RB, Draaistra J, Gommers FJ, Jacobsen E, Helder J, Bakker J (1997) Use of allele specificity of comigrating AFLP markers to align genetic maps from different potato genotypes. Mol Gen Gen 255:438–447

Rouppe van der Voort J, Kanyuka K, Van der Vossen E, Bendahmane A, Klein-Lankhorst R, Stiekema W, Baulcombe D and Bakker J (1999) Tight linkage of the nematode resistance gene *Gpa2* and the virus resistance gene *Rx1* on a single segment introgressed from the wild species *Solanum tuberosum* ssp. *andigena* CPC1673 into cultivated potato. Mol Plant Microbe Interact 12:197–206

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual (2nd edn). Cold Spring Laboratory Press, Cold Spring Harbor, NY

Seinhorst JW (1962) On the killing, fixation and transferring to glycerin of nematodes. Nematologica 8:29–32

Sherman JD, Stack SM (1995) Two-dimensional spreads of synaptonemal complexes from solanaceous plants. VI. High-resolution recombination nodule map for tomato (*Lycopersicon esculentum*). Genetics 141:683–708

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 *Escherichia coli* using an F-factor-based vector. Proc Natl Acad Sci USA 89:8794–8797

Stam P (1993) Construction of integrated genetic linkage maps by means of a new computer package: JOINMAP. Plant J 3:739–744

Suiter KA, Wendel JF, Case JS (1983) LINKAGE-1: a PASCAL computer program for the detection and analysis of genetic linkage. J Hered 74:203–204

Sulston JE, Brenner S (1974) The DNA of *Caenorhabditis elegans*. Genetics 77:95–104

Tanksley SD, Young ND, Paterson AH, Bonierbale MW (1989) RFLP mapping in plant breeding: new tools for an old science. Biotechnology 7:257–264

Tanksley SD, Ganal MW, Prince JP, Vicente de MC, Bonierbale MW, Broun P, Fulton TM, Giovannoni JJ, Grandillo S, Martin GB, Messeguer R, Miller JC, Miller L, Paterson AH, Pineda O, Roder MS, Wing RA, Wu W, Young ND (1992) High density molecular linkage maps of the tomato and potato genomes. Genetics 132:1141–1160

The *C. elegans* sequencing consortium (1998) Genome sequence of the nematode *C. elegans*: a platform for investigating biology. Science 282:2012–2018

Thompson J (1994) The coevolutionary process. University of Chicago Press, Chicago

Van Ooijen J (1994) DrawMap: a computer program for drawing genetic linkage maps. J Hered 85:66

Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Waterston RH, Sulston JE, Coulson AR (1997) The Genome. In: Riddle DL, Blumenthal T, Meyer BJ, Priess JR (Eds) *C. elegans* II. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Williams JK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535

Zhang HB, Choi SD, Woo SS, Li ZK, Wing RA (1996) Construction and characterization of two rice bacterial artificial chromosome libraries from the parents of a permanent recombinant inbred mapping population. Mol Breeding 2:11–24