Theor Appl Genet (1995) 90:841–846

M. Hühn

Determining the linkage of disease-resistance genes to molecular markers: the LOD-SCORE method revisited with regard to necessary sample sizes

Received: 3 September 1994 / Accepted: 18 October 1994

Abstract Some approaches to molecular marker-assisted linkage detection for a dominant disease-resistance trait based on a segregating F_2 population are discussed. Analysis of two-point linkage is carried out by the traditional measure of maximum lod score. It depends on (1) the maximum-likelihood estimate of the recombination fraction between the marker and the disease-resistance gene locus, (2) the observed absolute frequencies, and (3) the unknown number of tested individuals. If one replaces the absolute frequencies by expressions depending on the unknown sample size and the maximum-likelihood estimate of recombination value, the conventional rule for significant linkage (maximum lod score exceeds a given linkage threshold) can be resolved for the sample size. For each sub-population used for linkage analysis [susceptible (= recessive) individuals, resistant (= dominant) individuals, complete F_2] this approach gives a lower bound for the necessary number of individuals required for the detection of significant two-point linkage by the lod-score method.

Key words Linkage analysis · Disease resistance · Molecular markers · Lod-score method · Sample size

Introduction

Molecular markers, such as restriction fragment length polymorphisms (RFLPs), could provide an abundant supply of co-dominant genetic markers. Because of their almost unlimited number and their independence on environmental factors, as well as on dominance and epistatic effects, they are highly superior to protein (isozyme) markers and to morphological markers. RFLP linkage maps, therefore, have been constructed for many economically important organisms such as tomato, potato, maize, rice, lettuce, cabbage, sugar beet, barley, soybean and lentil (see, for example: Helentjaris 1987; Landry et al. 1987; Coe et al. 1988; Mc Couch et al. 1988; Tanksley 1988; Gebhardt et al. 1989, 1991; Havey and Muehlbauer 1989; Slocum et al. 1990; Graner et al. 1991; Tingey et al. 1991; Barzen et al. 1992; Pillen et al. 1992, 1993).

In the field of practical plant breeding, selection for a dominant disease-resistance trait is a common breeding objective; for example, nematode resistance in sugar beet (Jung et al. 1992). Application of RFLP-assisted selection schemes can greatly accelerate the breeding process.

RFLP linkage maps of crop species are often constructed with segregating populations, i.e., F_2 populations or backcrosses. In this paper, some comments on molecular marker-assisted linkage detection for a dominant disease-resistance trait based on a segregating F_2 population will be presented. The investigations, however, will be restricted to the analysis of two-point linkage by the traditional measure of maximum LOD SCORE (Ott 1991).

The problem

Assume a diploid segregating F_2 population co-segregating for RFLP markers and disease-resistance genes, as is the case, for example, for nematode resistance in sugar beet (Jung et al. 1992; Pillen et al. 1992). The two alleles at the resistance-gene locus are denoted by A (= resistant) and a (= susceptible) with A dominant over a. The marker alleles with co-dominant expression are B_1 and B_2 with a recombination value R between the marker and the disease-resistance gene locus. Selfing or intercrossing the F_1 generation AaB_1B_2 of an initial cross of homozygous parents creates a segregating F_2 population. Linkage analysis is based on this F_2. Three sub-populations of the F_2 can then be used for linkage analysis:

I. Linkage analysis based on susceptible (= recessive) individuals.
II. Linkage analysis based on resistant (= dominant) individuals.
III. Linkage analysis based on the complete F_2.

Communicated by P. M. A. Tigerstedt

M. Hühn
Institute of Crop Science and Plant Breeding, Christian-Albrechts-University Kiel, Olshausenstrasse 40, 24118 Kiel, Germany

In the following investigations the recombination value is assumed to be equal in both sexes. The double heterozygote AaB_1B_2 produces the gametes AB_1, aB_1, AB_2, and aB_2 with frequencies $\frac{1}{2}(1-R)$, $\frac{1}{2}R$, $\frac{1}{2}R$ and $\frac{1}{2}(1-R)$, respectively. The composition of the segregating F_2 is given in Table 1.

In this paper, all investigations are restricted to the analysis of the most simple situation of two-point linkage by the traditional approach of maximum LOD SCORE. This is defined as the logarithm to base 10 of the ratio of the likelihoods when the loci are at their maximum-likelihood recombination fraction and when the loci are taken to be unlinked.

We denote

N = number of tested individuals,

k = number of phenotypically distinct classes,

f_i = expected relative frequency of class i, $i = 1, 2, \ldots, k$,

z_i = observed absolute frequency of class i, $i = 1, 2, \ldots k$,

$L(R)$ = likelihood function dependent on the recombination fraction R, $Z(R)$ = LOD SCORE for recombination value R which is defined as

$$Z(R) = \log\left[\frac{L(R)}{L(0.50)}\right] \tag{1}$$

where log denotes the logarithm to base 10.

The maximum-likelihood estimate $\hat{R}$ of the recombination fraction R is that value for which the LOD SCORE is maximized. A conventional rule is to conclude that autosomal loci are linked whenever the maximum LOD SCORE exceeds 3 (Ott 1991). In the following numerical calculations, however, the linkage thresholds 4.0 and 2.0 will be additionally applied.

The LOD SCORE, of course, depends on the number N of tested individuals. The condition for significant linkage

$$Z(\hat{R}) = \log\left[\frac{L(\hat{R})}{L(0.50)}\right] \geq \left\{\begin{matrix}\text{linkage}\\ \text{threshold}\end{matrix}\right\} \tag{2}$$

can be resolved for N. This approach, therefore, provides an inequality (= lower bound) for the necessary number N of individuals required for a detection of significant two-point linkage by the LOD-SCORE method.

Theory, results and discussion

For a sub-population of susceptible individuals

The expected relative frequencies f_i, $i = 1, 2, \ldots, k$, of the three phenotypically distinct classes ($k = 3$) are (Table 1):

$f_1 = R^2$ (for aaB_1B_1)

$f_2 = 2R(1-R)$ (for aaB_1B_2) and

$f_3 = (1-R)^2$ (for aaB_2B_2).

The observed absolute frequencies are z_1, z_2 and z_3, respectively, with $z_1 + z_2 + z_3 = N$ = number of tested susceptible individuals. The likelihood function $L(R)$ is

$$L(R) = R^{2z_1+z_2}\cdot(1-R)^{z_2+2z_3}\cdot 2^{z_2}. \tag{3}$$

With $L(0.50) = 2^{z_2-2N}$ one obtains for the LOD SCORE:

$$\text{LOD SCORE} = Z(R) = \log\left[R^{2z_1+z_2}\cdot(1-R)^{z_2+2z_3}\cdot 2^{2N}\right]. \tag{4}$$

The maximum likelihood estimate $\hat{R}$ is found by setting to zero the first derivative of the LOD SCORE and verifying that this maximizes the LOD SCORE:

$$\hat{R} = \frac{2z_1 + z_2}{2(z_1 + z_2 + z_3)}. \tag{5}$$

By (5) one obtains:

$$2z_1 + z_2 = 2N\hat{R} \text{ and } z_2 + 2z_3 = 2N(1-\hat{R}). \tag{6}$$

The value $Z(\hat{R})$ of the LOD SCORE for $R = \hat{R}$ depends on $\hat{R}$, z_1, z_2, z_3, and N. If we replace the frequencies z_1, z_2 and z_3 by the expressions (6) depending on N and $\hat{R}$, the condition for significant linkage $Z(\hat{R}) \geq \{\text{linkage threshold}\}$ can be resolved for N:

$$F(N, \hat{R}) = 2N\log\left[2\hat{R}^{\hat{R}}(1-\hat{R})^{1-\hat{R}}\right] \geq \left\{\begin{matrix}\text{linkage}\\ \text{threshold}\end{matrix}\right\}. \tag{7}$$

This inequality (7) can be interpreted by the following point of view: if we assume an existing linkage with a 'true' recombination fraction R, and provided that it is well estimated by the maximum likelihood estimate $\hat{R}$, then (7) gives a lower bound for the necessary number N of individuals required for the detection of significant two-point linkage by the LOD-SCORE method.

Some numerical values for these lower bounds are presented in Table 2; N decreases, of course, with increasing deviations from $R = 0.50$ (Table 2).

For a sub-population of resistant individuals

The expected relative frequencies f_i, $i = 1, 2, \ldots, k$, of the three phenotypically distinct classes ($k = 3$) are

Table 1 The composition of a segregating F_2 by combining male and female gametes

Female gametes (with frequencies)	♂ / ♀	Male gametes (with frequencies)			
		AB_1 $\frac{1}{2}(1-R)$	AB_2 $\frac{1}{2}R$	aB_1 $\frac{1}{2}R$	aB_2 $\frac{1}{2}(1-R)$
AB_1	$\frac{1}{2}(1-R)$	AAB_1B_1 $\frac{1}{4}(1-R)^2$	AAB_1B_2 $\frac{1}{4}R(1-R)$	AaB_1B_1 $\frac{1}{4}R(1-R)$	AaB_1B_2 $\frac{1}{4}(1-R)^2$
AB_2	$\frac{1}{2}R$	AAB_1B_2 $\frac{1}{4}R(1-R)$	AAB_2B_2 $\frac{1}{4}R^2$	AaB_1B_2 $\frac{1}{4}R^2$	AaB_2B_2 $\frac{1}{4}R(1-R)$
aB_1	$\frac{1}{2}R$	AaB_1B_1 $\frac{1}{4}R(1-R)$	AaB_1B_2 $\frac{1}{4}R^2$	aaB_1B_1 $\frac{1}{4}R^2$	aaB_1B_2 $\frac{1}{4}R(1-R)$
aB_2	$\frac{1}{2}(1-R)$	AaB_1B_2 $\frac{1}{4}(1-R)^2$	AaB_2B_2 $\frac{1}{4}R(1-R)$	aaB_1B_2 $\frac{1}{4}R(1-R)$	aaB_2B_2 $\frac{1}{4}(1-R)^2$

Table 2 Lower bounds of necessary sample sizes required for linkage detection by the LOD-SCORE method

Sub-population	Susceptible			Resistant			Susceptible + resistant		
linkage threshold	2.0	3.0	4.0	2.0	3.0	4.0	2.0	3.0	4.0
Recombination fraction 0.05	5	7	10	26	39	52	13	19	25
0.10	7	10	13	38	57	76	17	26	34
0.15	9	13	18	56	84	112	24	36	47
0.20	12	18	24	85	128	170	34	51	68
0.25	18	27	36	134	201	268	51	76	101
0.30	28	42	56	226	338	451	82	123	164
0.35	51	76	101	426	638	851	149	223	298
0.40	115	172	229	1000	1500	2000	341	511	682
0.45	460	690	920	4110	6165	8220	1376	2064	2752

(Table 1):

$f_1 = (1-R^2)/3$ (for $A\cdot B_1B_1$)

$f_2 = 2(1+R^2-R)/3$ (for $A\cdot B_1B_2$) and

$f_3 = (2R-R^2)/3$ (for $A\cdot B_2B_2$)

where the dot notation indicates that the respective allele may be the dominant one or the recessive one.

The observed absolute frequencies are again z_1, z_2 and z_3, respectively, with $z_1+z_2+z_3 = N$ = number of tested resistant individuals. The likelihood function $L(R)$ is

$$L(R) = (1-R^2)^{Z_1}(1+R^2-R)^{Z_2}(2R-R^2)^{Z_3}\cdot 2^{Z_2}\cdot 3^{-N}. \quad (8)$$

With $L(0.50) = 2^{Z_2-2N}$ one obtains for the LOD SCORE:

$$Z(R) = \log\left[(1-R^2)^{Z_1}(1+R^2-R)^{Z_2}(2R-R^2)^{Z_3}\left(\tfrac{4}{3}\right)^N\right]. \quad (9)$$

The maximum likelihood estimate $\hat{R}$ is found by setting to zero the first derivative of the LOD SCORE and verifying that this maximizes the LOD SCORE. One then obtains the condition:

$$\frac{2\hat{R}z_1}{1-\hat{R}^2} = \frac{z_2(2\hat{R}-1)}{1+\hat{R}^2-\hat{R}} + \frac{z_3(2-2\hat{R})}{2\hat{R}-\hat{R}^2}. \quad (10)$$

To apply a similar approach as in the previous analysis for the sub-population of susceptible individuals, the condition for significant linkage $Z(\hat{R}) \geq$ {linkage threshold} must be transformed into an inequality for N. $Z(\hat{R})$ depends on $\hat{R}, z_1, z_2, z_3$, and N. If we are able to replace the frequencies z_1, z_2, and z_3 by expressions depending on $\hat{R}$ and N, then the condition $Z(\hat{R}) \geq$ {linkage threshold} can be resolved for N. But, for the determination of the three 'unknowns', z_1, z_2, and z_3, there are only two equations available: extreme condition (10) and $z_1+z_2+z_3 = N$. To obtain an additional equation, we replace one of the three absolute frequencies z_1 or z_2 or z_3 by its theoretical frequency Nf_i calculated at $R = \hat{R}$ (Replacement of z_1 or z_2 or z_3 leads to the same results). By this procedure, the lower bounds of necessary numbers of tested plants can be easily calculated and one obtains:

$$F(N,\hat{R}) = \frac{N}{3}\log\left[\left(\frac{4}{3}\right)^3 \cdot(1-\hat{R}^2)^{1-\hat{R}^2}\cdot(1+\hat{R}^2-\hat{R})^{2(1+\hat{R}^2-\hat{R})} \cdot(2\hat{R}-\hat{R}^2)^{2\hat{R}-\hat{R}^2}\right] \geq \left\{\begin{matrix}\text{linkage}\\ \text{threshold}\end{matrix}\right\}. \quad (11)$$

Some numerical results are presented in Table 2.

For the complete population of susceptible and resistant individuals

The expected relative frequencies f_i, $i = 1,2,\ldots,k$, of the six phenotypically distinct classes ($k = 6$) are (Table 1):

$f_1 = (1-R^2)/4$ (for $A\cdot B_1B_1$)

$f_2 = (1+R^2-R)/2$ (for $A\cdot B_1B_2$)

$f_3 = (2R-R^2)/4$ (for $A\cdot B_2B_2$)

$f_4 = R^2/4$ (for aaB_1B_1)

$f_5 = (R-R^2)/2$ (for aaB_1B_2) and

$f_6 = (1-2R+R^2)/4$ (for aaB_2B_2).

The observed absolute frequencies are z_i, $i = 1,2,\ldots,6$, with $z_1+z_2+z_3+z_4+z_5+z_6 = N$ = number of tested individuals. The likelihood function $L(R)$ is

$$L(R) = R^{Z_3+Z_5+2Z_4}(1-R)^{Z_1+Z_5+2Z_6}\cdot(1+R)^{Z_1}(2-R)^{Z_3}(1+R^2-R)^{Z_2}\cdot 2^{Z_2+Z_5-2N}. \quad (12)$$

With $L(0.50) = 2^{z_2+z_5-4N} \cdot 3^{z_1+z_2+z_3}$ one obtains for the LOD SCORE:

$$Z(R) = \log\left[R^{z_3+z_5+2z_4} \cdot (1-R)^{z_1+z_5+2z_6} \cdot (1+R)^{z_1} (2-R)^{z_3} (1+R^2-R)^{z_2} \cdot \frac{4^N}{3^{z_1+z_2+z_3}} \right]. \quad (13)$$

The maximum likelihood estimate $\hat{R}$ is found by setting to zero the first derivative of the LOD SCORE and verifying that this maximizes the LOD SCORE. One then obtains the condition:

$$\frac{z_2(2\hat{R}-1)}{(1+\hat{R}^2-\hat{R})} + \frac{2z_3(1-\hat{R})}{\hat{R}(2-\hat{R})} + \frac{2z_4}{\hat{R}} + \frac{z_5(1-2\hat{R})}{\hat{R}(1-\hat{R})} = \frac{2\hat{R}z_1}{1-\hat{R}^2} + \frac{2z_6}{1-\hat{R}}. \quad (14)$$

Again, the condition for significant linkage $Z(\hat{R}) \geq \{\text{linkage threshold}\}$ must be transformed into an inequality for N by replacing the frequencies z_i, $i = 1, 2, \ldots, 6$, by expressions depending on $\hat{R}$ and N. But, there are only two equations available: extreme condition (14) and $z_1 + z_2 + z_3 + z_4 + z_5 + z_6 = N$. To obtain additional equations, we replace four of the six absolute frequencies z_i by their theoretical frequencies Nf_i calculated at $R = \hat{R}$ (for each choice of these four replacements one obtains the same results). By this procedure, lower bounds of the necessary numbers of tested plants can be easily calculated. One obtains:

$$F(N, \hat{R}) = \frac{N}{4} \log\left[\hat{R}^{4\hat{R}-\hat{R}^2} \cdot (1-\hat{R})^{3-2\hat{R}-\hat{R}^2} \cdot (1+\hat{R})^{1-\hat{R}^2} \cdot (2-\hat{R})^{2\hat{R}-\hat{R}^2} (1+\hat{R}^2-\hat{R})^{2(\hat{R}^2-\hat{R}+1)} \cdot \frac{256}{27} \right] \geq \left\{ \begin{matrix} \text{linkage} \\ \text{threshold} \end{matrix} \right\}. \quad (15)$$

Some numerical results are presented in Table 2. These results need no further discussion and interpretation. They are highly self-explanatory.

The necessary sample sizes are lowest for the sub-population of susceptible individuals and largest for the sub-population of resistant individuals, while the sample sizes for the complete F_2 are intermediate (Table 2).

With regard to their practical relevance, the numerical sample sizes from Table 2 may be criticised since they have been calculated for the special case $R = \hat{R}$. But, a generalization and improvement can be obtained by the construction of a two-sided central α%-confidence interval on the true recombination fraction R: since $\hat{R}$ is asymptotically normally distributed and unbiased (Ott 1991), the lower (U_1) and upper (U_2) limits of this confidence interval are $R \pm u_{1-\alpha/2}\sqrt{V(\hat{R})}$ where $u_{1-\alpha/2}$ is the $(1-\alpha/2)$-quantile from the normal distribution and $V(\hat{R})$ denotes the variance of $\hat{R}$ which depends on R and N:

$$V(\hat{R}) = \frac{R(1-R)}{2N} \quad (16)$$

(for the sub-population of susceptible individuals).

$$V(\hat{R}) = \frac{3}{N\left\{ \frac{4}{1-R^2} + \frac{4}{2R-R^2} - \frac{6}{1+R^2-R} \right\}} \quad (17)$$

(for the sub-population of resistant individuals).

$$V(\hat{R}) = \frac{1}{N\left\{ \frac{1}{1-R^2} + \frac{1}{2R-R^2} + \frac{1}{2R-2R^2} - \frac{3}{2+2R^2-2R} \right\}} \quad (18)$$

(for the complete F_2).

The limits of the confidence interval depend on R, α and N: $U_1 = U_1(N, \alpha, R)$ and $U_2 = U_2(N, \alpha, R)$. For the true recombination fraction R and a required confidence level α, the confidence interval depends only on N.

Replacement of $\hat{R}$ in (7), (11), and (15), respectively, by U_1 and by U_2 leads to inequalities which cannot be solved analytically for N. Lower bounds for necessary sample sizes must, therefore, be derived by numerical methods, i.e., the solution is the smallest integer N with $F(N, U_1) \geq \{\text{linkage threshold}\}$ and $F(N, U_2) \geq \{\text{linkage threshold}\}$.

Numerical example

(sub-population of susceptible individuals; $R = 0.25$; linkage threshold = 3.0; confidence level $\alpha = 0.10$). The lower (U_1) and upper (U_2) limits of the confidence interval are

$$U_{1,2} = R \pm 1.64\sqrt{\frac{R(1-R)}{2N}}. \quad (19)$$

By (7) one obtains

$$N \geq \frac{3}{2\log[2U_1^{U_1}(1-U_1)^{1-U_1}]} = f_1(N) \quad (20)$$

$$N \geq \frac{3}{2\log[2U_2^{U_2}(1-U_2)^{1-U_2}]} = f_2(N). \quad (21)$$

The smallest integer N which fulfills both inequalities (20) and (21) can be derived from Fig. 1.

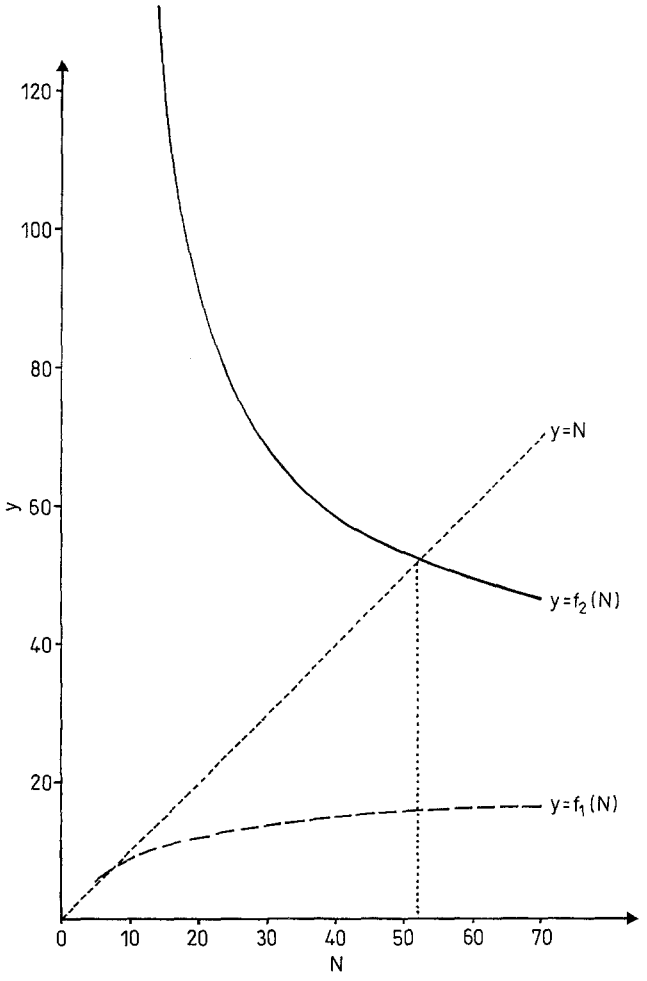

Fig. 1 Determination of necessary sample sizes by numerical methods [smallest N with $N \geq f_1(N)$ and $N \geq f_2(N)$] (see text)

Table 3 Necessary sample sizes for different values of R and v

R / v	0.05	0.15	0.25	0.35
Susceptible individuals				
0.05	7–8	13–14	24–30	61–98
0.10	7–8	12–15	22–33	50–130
0.20	7–8	11–16	18–43	35–268
0.30	7–8	10–18	16–56	26–883
Resistant individuals				
0.05	38–40	80–90	178–227	503–826
0.10	37–40	75–95	159–258	404–1112
0.20	36–42	67–108	127–338	272–2369
0.30	34–44	59–122	103–455	192–8182
Susceptible and resistant individuals				
0.05	18–19	34–37	69–85	178–288
0.10	18–19	32–39	62–96	145–386
0.20	17–20	29–44	51–123	100–811
0.30	17–20	26–49	42–162	53–2609

For this numerical example, the necessary sample size is $N = 52$. The previous result for $R = \hat{R}$, i.e., without allowing any stochastic variation of $\hat{R}$, was $N = 27$ (Table 2). A consideration of the variability of $\hat{R}$ increases the necessary sample size considerably. This procedure, however, can only be applied for sufficiently large sample sizes so that the approximations based on the normal distribution are reliably valid.

For small samples [with unknown variance $V(\hat{R})$] an application of the limits $R \pm \sqrt{V(\hat{R})}$ with $\sqrt{V(\hat{R})} = vR$ and v = coefficient of variation for $\hat{R}$, i.e., $U_1 = (1 - v)R$ and $U_2 = (1 + v)R$, may provide some rough numerical results on necessary sample sizes for linkage detection. This approach presumes unbiasedness of $\hat{R}$ and independence of v on the sample size N. Maximum-likelihood estimates, however, are often biased. In linkage analysis, the statistical bias of the recombination fraction has not reveived much attention. One reason may be the fact that this bias tends to vanish with increasing sample size. Some numerical results for the latter approach with limits $U_1 = (1 - v)R$ and $U_2 = (1 + v)R$ are presented in Table 3. For varying $\hat{R}$ within the limits from U_1 up to U_2, the range of necessary sample sizes increases, of course, with increasing v and with increasing R.

An essential improvement and generalization of the previous approaches can be attained if the exact distribution of $\hat{R}$ for small sample sizes N is known. For each sample size, exact confidence intervals, bias, variance etc., can be calculated exactly. These exact distributions of $\hat{R}$ will be derived by simulation.

References

Barzen E, Mechelke W, Ritter E, Seitzer JF, Salamini F (1992) RFLP markers for sugar beet breeding – chromosomal linkage maps and location of major genes for rhizomania resistance, monogermy and hypocotyl colour. Plant J 2:601–611

Coe EH, Neuffer MG, Hoisington DA (1988) The genetics of corn. In: Sprague GF, Dudley JW (eds) Corn and corn improvement, 3rd edn. American Society of Agronomy, Madison Wisconsin, pp 81–258

Gebhardt C, Ritter E, Debener T, Schachtschabel U, Walkemeier B, Uhrig H, Salamini F (1989) RFLP-analysis and linkage mapping in *Solanum tuberosum*. Theor Appl Genet 78:65–77

Gebhardt C, Ritter E, Barone A, Debener T, Walkemeier B, Schachtschabel U, Kaufmann 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

Graner A, Jahoor A, Siedler H, Schondelmaier J, Pillen K, Fischbeck G, Wenzel G, Herrmann RG (1991) Construction of an RFLP map of barley. Theor Appl Genet 83:250–256

Havey MJ, Muehlbauer FJ (1989) Variability for restriction fragment lengths and phylogenies in lentil. Theor Appl Genet 77:839–843

Helentjaris T (1987) A genetic linkage map for maize based on RFLPs. Trends Genet 3:217–221

Jung C (1992) Utilization of RFLP markers in plant breeding. In: Kahl G, Appelhans H, Kämpf J, Driesel AJ (eds). Advances in molecular genetics (Bio Tech Forum) 5. Hüthig, Heidelberg, pp 101–117

Jung C, Koch R, Fischer F, Brandes A, Wricke G, Herrmann RG (1992) DNA markers closely linked to nematode resistance genes in sugar beet (*Beta vulgaris* L.) using chromosome additions and translocations originating from wild beets of the Procumbentes species. Mol Gen Genet 232:271–278

Landry BS, Kesseli RV, Farrara B, Michelmore RW (1987) A genetic map of lettuce (*Lactuca sativa* L.) with restriction fragment length

polymorphism, isozyme, disease resistance and morphological markers. Genetics 116:331–337
McCouch SR, Kockert G, Yu ZH, Wang ZY, Khush GS, Coffman WR, Tanksley SD (1988) Molecular mapping of rice chromosomes. Theor Appl Genet 76:815–829
Ott J (1991) Analysis of human genetic linkage. Revised edition. The Johns Hopkins University Press, Baltimore London
Pillen K, Steinrücken G, Wricke G, Herrmann RG, Jung C (1992) A linkage map of sugar beet (*Beta vulgaris* L.). Theor Appl Genet 84:129–135
Pillen K, Steinrücken G, Herrmann RG, Jung C (1993) An extended linkage map of sugar beet (*Beta vulgaris L.*) including nine putative lethal genes and the restorer gene X. Plant Breed 111:265–272
Slocum MK, Figdore SS, Kennard WC, Suzuki JY, Osborn TC (1990) Linkage arrangement of restriction fragment length polymorphism loci in *Brassica oleracea*. Theor Appl Genet 80: 57–64
Tanksley SD (1988) Molecular mapping of plant chromosomes. In: Gustafson RAP (ed) Chromosome structure and function. Plenum Press, New York, pp 157–173
Tingey SV, Rafalski JA, Williams JGK, Sebastian S (1991) Soybean genome analysis: DNA polymorphisms are identified by oligonucleotide primers of arbitrary sequence. In: Herrmann RG, Larkins BA (eds) Plant molecular biology 2, NATO ASI Series, Series A: Life Sciences vol. 212. Plenum Press, New York London, pp 263–268