Theor Appl Genet (1987) 74:646–653

Comparison of restriction endonucleases and sources of probes for their efficiency in detecting restriction fragment length polymorphisms in lettuce (*Lactuca sativa* L.)

B. S. Landry *, R. Kesseli, Hei Leung ** and R. W. Michelmore

Department of Vegetable Crops, University of California, Davis, CA 95616, USA

Received April 12, 1987; Accepted May 22, 1987
Communicated by R. Riley

Summary. As a first step in developing a detailed genetic map of lettuce (*Lactuca sativa* L.), 156 cDNA and 123 genomic DNA clones of lettuce were compared for their efficiency to detect restriction fragment length polymorphism (RFLP) between four lines of lettuce. Polymorphism was detected 2.5 times more frequently with cDNA probes than random genomic probes. Less polymorphism was detected with cDNA clones homologous to single copy than with cDNA clones homologous to multiple copy DNA sequences. A lower percentage of polymorphism was detected with genomic DNA clones homologous to repetitive sequences than with other types of probes. Digests with each of nine restriction endonucleases were compared; increased polymorphism was not correlated with the presence of a CpG dimer in the recognition sequence of the restriction endonuclease. Digests with enzymes recognizing four base pairs, however, displayed RFLPs less frequently. The six pairwise comparisons of the four lettuce lines showed different frequencies of polymorphism which only approximately corresponded to genetic distances obtained from previous isozyme analyses.

Key words: Polymorphism – Restriction endonuclease – Mapping – Lettuce – *Lactuca sativa* – Resistance gene

Introduction

Analysis of restriction fragment length polymorphisms (RFLP) permits the rapid construction of detailed genetic maps (Botstein et al. 1980). This approach is being used extensively to map the human genome (Gusella 1986) and is being extended to a wide variety of organisms including plant species such as tomato, maize and lettuce (Bernatzky and Tanksley 1986; Evola et al. 1986; Helentjaris et al. 1985; Landry et al. 1987). There is little information, however, regarding the most efficient methodologies for the construction of genetic maps. The ability of different restriction endonucleases or genomic verses cDNA probes to detect polymorphisms has not been studied in detail. In one study, however, restriction endonucleases containing CpG dimers in their recognition sequences revealed more polymorphisms in humans, possibly because methylated cytosine at this position has a high mutation rate (Barker et al. 1984).

A detailed genetic map of lettuce is part of our strategy to clone genes for disease resistance. As a first step in developing a genetic map using RFLP analysis, cDNA and genomic DNA clones were compared for their efficiency to detect RFLPs between four lettuce lines. Nine restriction endonucleases with and without a CpG dimer in their recognition sequences were used and compared.

Present address:
* Agriculture-Canada, Research, P.O. Box 457, St-Jean-sur-Richelieu, Quebec J3B 6Z8, Canada
** International Rice Research Institute, Los Banos, Phillipines

Materials and methods

Cloning of lettuce genomic DNA and isolation of single/low copy DNA probes for RFLP analysis

The construction and selection of single/low copy DNA probes for RFLP analysis in lettuce has been described (Landry and Michelmore 1985). Genomic DNA was digested with *Mbo*I and short, random fragments (0.4–1.0 kb) were cloned into the *Bam*HI site of the vector, pUC13. Plasmid miniscreens of positive clones were denatured and placed on Zetaprobe membrane (Bio-Rad) using a dot-blot apparatus

(Bio-Rad). Lettuce genomic DNA was labelled with ^{32}P and hybridized to the membranes using 5×10^7 cpm (specific activity $2–5 \times 10^8$ cpm/μg DNA). Clones showing hybridization were assumed to contain repetitive DNA and therefore discarded. The remaining clones contained mainly single and low copy DNA sequences as shown in subsequent hybridizations to Southern blots of genomic DNA.

Isolation of mRNA, synthesis and cloning of cDNA

Total lettuce RNA was isolated from lettuce line PI221936 using the method described by Cathala et al. (1983). Polyadenylated RNA was selected using two cycles of selection on oligo (dT)-cellulose columns (Maniatis et al. 1982). The quality of the mRNA preparation was determined by sizing peptides produced by *in vitro* translation on polyacrylamide gels. The lettuce cDNA library was constructed following the cDNA cloning procedure described by Alexander et al. (1984). Approximately 2×10^6 lettuce cDNA clones were obtained from 7 μg of lettuce mRNA after transforming *E. coli* strain MM294. Individual colonies (1,200) were picked, incubated overnight at 37 °C in phosphate buffered LB media containing 15% glycerol in 96 – well microtest plates (Falcon) and stored at −70 °C.

Isolation of plasmid DNA inserts

Plasmids were isolated from lettuce clones using a miniscreen procedure (Rodriguez and Tait 1983). Lettuce inserts were separated from the vectors by digesting with either *Eco*RI plus *Hin*dIII (genomic DNA clones) or *Pst*I (cDNA clones) followed by electrophoresis on 1.2% TAE-agarose gels. The insert DNA was purified from the gel by the freeze-squeeze method (Tautz and Renz 1983).

Lettuce genomic DNA isolation

Lettuce DNA was extracted using a procedure modified from Timberlake (1978) and Fisher and Goldberg (1982). Frozen leaves (10–30 g) were blended in a stainless steel Waring blender in 160 ml ice-cold H buffer (4 mM spermidine, 1 mM spermine, 10 mM EDTA, 10 mM Tris, 80 mM KCl, 1 mM PMSF, 500 mM sucrose, 0.2% 2-mercaptoethanol, pH 9.5) at maximum rpm. The homogenate was filtered through four layers of cheesecloth and one layer of Miracloth (Calbiochem) into a 250 ml centrifuge bottle. After centrifuging at 2,000 g, 4 °C for 20 min, the supernatant was discarded and the pellet resuspended in 40 ml of ice cold HT buffer (1× H buffer, 0.5% triton X-100). The suspension was again centrifuged at 2,000 g, 4 °C for 10 min, in a 50 ml teflon tube and the supernatant discarded. Resuspension of the pellet in HT buffer and centrifugation was repeated until the pellet of nuclei was grey to white (1–3 times). The pellet was then resuspended in 12 ml of HT buffer and 12 ml of lysis buffer (100 mM Tris, 40 mM EDTA, 2% Na-sarcosyl, pH 9.5) was added. Immediately, 23.28 g of CsCl were added and the tubes were incubated at 55–60 °C for 1 h with occasional inversions. After the CsCl had dissolved, the tubes were centrifuged at 28,000 g, 15 °C for 30 min. The supernatant was filtered through two layers of cheesecloth into a 38 ml quick-seal ultracentrifuge tube (Beckman) containing 1.47 ml of ethidium bromide (10 mg/ml) with a 50 ml syringe and a 16 gauge needle acting as a funnel. Each tube was made up to volume with a solution of CsCl (97 g of CsCL in 100 ml TE), sealed and centrifuged at $242{,}000 \times g$, 15 °C for 18 h in a vertical rotor (Vti50, Beckman). DNA was recovered using standard procedures (Maniatis et al. 1982) and resuspended in TE at a final concentration of 0.5 μg/ml.

Lettuce stocks

L. sativa cv. Calmar (crisphead type), Kordaat (butterhead type), Galleaga (latin type) and PI251245 and PI221936 (landraces of *L. sativa*) were all obtained from the UCD lettuce germplasm collection. Initially, PI251245 and Galleaga were selected as parents for a wide cross and the frequency of RFLPs was assessed using genomic clones (data not given). Simultaneously, a diverse group of *Lactuca* accessions was analysed for isozyme polymorphism (Kesseli and Michelmore 1986) and variation for resistance genes (Farrara and Michelmore 1987); the differences between PI251245 and Galleaga was shown to be less than that for other crosses. In particular, a cross of Calmar and Kordaat would segregate for five genes for resistance to downy mildew and five isozyme loci. Therefore, Calmar and Kordaat were included in the analysis described in this paper and, subsequently, were used to develop the genetic map (Landry et al. 1987)

Restriction digests, electrophoresis, blotting, nick-translation and hybridization

The DNA from Calmar, Kordaat, Galleaga and PI251245 was individually digested according to manufacturer's recommendations with each of the following restriction endonucleases: *Bam*HI, *Eco*RI, *Eco*RV, *Hin*dIII, *Xba*I (BRL) and *Dra*I, *Map*I, *Pal*I, *Taq*I (P-L Biochemicals). Digestions proceeded for over 20 h x units/μg to ensure completion. Digested DNAs (7 μg/lane) were electrophoresed in 1.0% TAE agarose gels containing ethidium bromide (0.5 μg/ml). DNA was depurinated (0.25 M HCl, 25 min), denatured (0.4 M NaOH, 0.8 M NaCl, 30 min), neutralized (0.5 M Tris-HCl, 1.5 M NaCl, pH 7.6, 30 min) and transferred from the gel onto Zetaprobe membrane (Bio-Rad), using 5× or 10× SSC as transfer buffer (Southern 1975). Alternatively, DNA was depurinated (0.25 M HCl, 15 min) and directly transferred to the membrane using 0.4 N NaOH as the transfer solution (Reed and Mann 1985). Prehybridization and hybridization reactions were carried out in sealed plastic bags at 42 °C with 50% deionized formamide (BRL) according to membrane manufacturer's recommendations for single copy hybridization. DNA inserts from genomic and cDNA clones were labeled with ^{32}P by nick-translation ($1–15 \times 10^7$ cpm) (Rigby et al. 1977) and added to the bag after unincorporated nucleotides had been removed by passing the sample through bio-gel P-60 (Bio-Rad). Hybridizations were incubated for 24–48 h. Filters were washed in 2×SSC, 1% SDS at room temperature for 10–20 min then 0.1×SSC, 1% SDS and 0.1% sodium pyrophosphate at 65 °C for 20 min. The washed filters were exposed to XAR-5 X-ray film (Kodak) for one hour to four days at −70 °C using Cronex Lightening-Plus intensifier screens (DuPont). Probes were removed from the membrane by washing in 0.1×SSC, 0.1% SDS at 85 °C for 30 min. The membranes were rehybridized up to 10 times.

Analysis of polymorphism

Probes were grouped into one of four classes: 1) cDNA clones, which hybridized to only one or two bands larger than the size of the probe with all restriction endonucleases, were designated putative "single copy clones"; 2) cDNA clones which produced three or more hybridization bands with at least one restriction endonuclease were designated putative "multiple copy clones"; 3) genomic DNA clones were designated "single/low copy clones" when an exposure time longer than one day was necessary to reveal hybridization to a low number (<4) of bands; 4) genomic DNA clones were classified as "repetitive clones" when exposure for less than one day

revealed clear hybridization. Data from each probe with each of the nine different restriction endonucleases were recorded separately for each of the six pairwise comparisons between the four lines; this gave 36 sets of data. These were grouped as required to carry out the desired comparisons. The lack of hybridization was considered as missing data. Pearson's χ^2 tests of independence and heterogeneity were calculated from the raw data (Feinberg 1977). Percentages are given only to aid comparison.

Results

Comparison of cDNA and genomic DNA probes

Polymorphism was detected less frequently with genomic DNA probes than with cDNA probes (Table 1). Of 156 cDNA clones analysed, 103 (66%) hybridized as single copy; 53 (34%) detected three or more bands and were, therefore, classified as multiple copy cDNA sequences. Single copy cDNA probes detected significantly less polymorphism than multiple copy cDNA (Table 1).

Our selection procedures resulted in two classes of genomic clones. In preliminary experiments, 23 genomic clones were hybridized to Southern blots without being preselected for low copy number sequences; 19 clones contained repetitive sequences and four contained single/low copy sequences. Random genomic clones were then assayed for presence of repetitive sequences by hybridizing with ^{32}P-labelled total genomic DNA (Landry and Michelmore 1985). When 100 genomic clones, which had been selected as putative single/low copy sequences, were hybridized to Southern blots of digested genomic DNA, 22 clones still were found to contain highly repeated DNA. Our approach to select single/low copy DNA probes, therefore, allowed approximately 22% of the genomic clones containing repetitive DNA to escape detection. Polymorphism was detected less frequently with this nonrandom subset of clones containing repeated DNA compared to other classes of probes (Table 1).

Table 1. Classes of probes compared for their ability to detect RFLPs. Includes the pooled data of all accessions for all restriction endonucleases

Class of probes	No. of clones	No. pairwise comparisons	% polymorphic	χ^2 (independence)[a]
Genomic DNA				
Single/Low copy	82	1,951	12.5	9.65***
Repetitive	41	1,075	8.7	
cDNA				
Single copy	103	1,684	24.8	13.65***
Multiple copy	53	1,275	30.9	
ALL gDNA	123	3,026	11.1	255.50***
ALL cDNA	156	2,959	27.4	
Total	279	5,985	19.2	

[a] The total number of polymorphic pairwise comparisons was compared to the total number of non polymorphic ones with Pearson's χ^2 test for independence. This was done between classes of probes within genomic DNA and cDNA. Genomic DNA probes were also compared to cDNA probes

*** Significant χ^2 for independence with $P < 0.001$

Efficiencies of restriction endonucleases

DNA from the four lines was individually digested with each of nine restriction endonucleases and then hybridized to each probe (Fig. 1). The pooled data showed significant heterogeneity (χ^2 test) among endonucleases for the ability to detect polymorphisms (Table 2). DNA digested with *Dra*I, *Pal*I and *Taq*I displayed significantly less polymorphism, while DNA digested with *Eco*RI displayed significantly more polymorphism. When the frequencies of polymorphism detected by each class of probe were analysed separately, data from single/low copy genomic probes were heterogeneous mainly because of the low percentage of polymorphism detected by *Pal*I and *Taq*I. The heterogeneity in the data for the single copy cDNA probes was the consequence of the high and low polymorphism levels detected with *Xba*I and *Dra*I, respectively. Occasionally hybridization could not be detected; enzymes with four base pair recognition sequences provided three of the four highest values for this missing data (Table 2).

Restriction endonucleases containing a CpG dimer in their recognition sequences (*Msp*I and *Taq*I) did not display significantly different frequencies of polymorphism when compared to the other enzymes. Digests with *Msp*I ($^{5'}$CCGG$^{3'}$) and *Pal*I ($^{5'}$GGCC$^{3'}$) were compared to *Dra*I ($^{5'}$TTTAAA$^{3'}$) to indicate whether genomic regions rich in cytosine and guanine were more polymorphic than regions rich in adenine and thymine; no significant difference (χ^2(independence) = 0.085) for the frequency of polymorphism was found.

The average size of the fragments detected were determined for each restriction endonuclease in a random sample of hybridizations (Table 3). The standard errors on the observed values were large as the observed fragment sizes ranged from 0.2 to 50.0 kb. The expected average size of the fragments produced by each restriction endonuclease was calculated on the basis of an approximate G+C content of 38% (35% determined by analytical ultracentrifugation [Unpublished data] plus 3% added to take methylation into account [Green 1972]). In all cases, the observed fragment sizes were consistently larger than expected.

Table 2. The percent polymorphism detected by the nine restriction endonucleases for each class of probe. All pairwise combinations of the four accessions are pooled

Restriction endonucleases	Genomic DNA		cDNA		All classes	No. pairwise comparisons	Missing data
	Single/low copy	Repetitive	Single copy	Multiple copy			
*Bam*HI	13.5	6.0	22.3	34.4	19.5	728	8.2
*Dra*I	13.4	9.0	15.8**	26.7	16.2*	730	6.9
*Eco*RI	16.7	9.2	27.9	37.7	23.1*	750	7.2
*Eco*RV	9.4	8.6	25.8	30.2	19.9	715	9.4
*Hin*dIII	13.7	14.3*	27.6	33.1	21.7	770	8.6
*Msp*I	11.9	9.5	25.9	26.3	18.5	676	16.9
*Pal*I	7.7*	8.7	23.4	22.2*	14.6**	602	21.8
*Taq*I	7.2*	9.7	18.4	31.3	15.5*	491	29.1
*Xba*I	16.6	2.7*	37.7**	34.9	22.4	523	24.1
χ^2 (heterogeneity)[a]	16.28*	10.44	23.62**	11.44	28.06***		

[a] χ^2 tests for heterogeneity were performed on the raw data
Significant χ^2 test for heterogeneity at: * $0.01 < P < 0.05$; ** $0.001 < P < 0.01$; *** $P < 0.001$

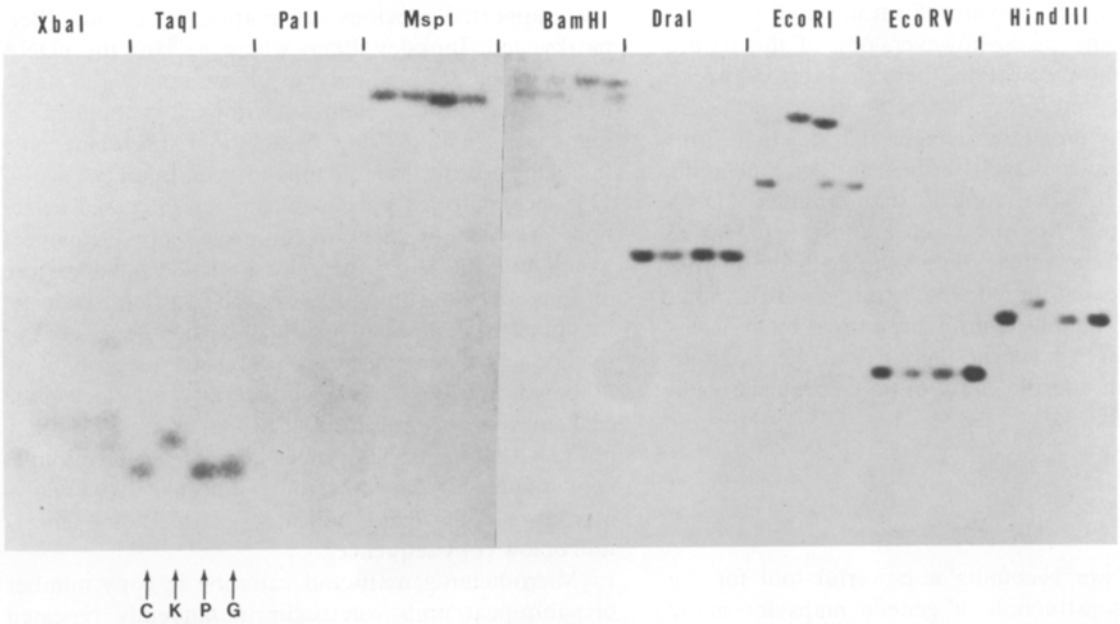

Fig. 1. Restriction fragment length polymorphism detected with probe *CL162*. Southern blots were exposed for 2 days at −70 °C with intensifier screens (DuPont Cronex II). C = Calmar; K = Kordaat; P = PI251245; G = Galleaga. Restriction endonucleases *Xba*I, *Taq*I, *Bam*HI, *Eco*RI and *Hin*dIII detected one or more RFLPs. *Pal*I did not produce any detectable hybridization and *Xba*I produced diffuse and faint bands

Polymorphism between lettuce lines

The genetic distances based on isozyme data (Kesseli and Michelmore 1986) showed more divergence between the Calmar and Kordaat (Nei's D = 0.282) than the other five pairwise combinations of this study. We expected this RFLP study to show similar relative values, yet a different pair, PI251245 and Kordaat had the greatest difference (22.3% polymorphism between lines, Table 4). We also expected the differences among accessions to show the same relative rankings, regardless of which class of probe was used. This proved incorrect since, for example among the comparisons PI251245 and Kordaat had the greatest difference (44.7% polymorphism between accessions) for the multiple copy cDNA class but the least (5.5%) for the repetitive genomic DNA class (Table 4).

Table 3. Recognition sequences and comparison of the observed and expected sizes of the fragments produced with nine restriction endonucleases

Restriction endonuclease	Recognition sequence	No. of fragments measured	Fragment size (kb)	
			observed	expected
*Bam*HI	'5GGATCC3'	124	11.4	8.0
*Exo*RV	GATATC	107	11.2	3.0
*Xba*I	TCTAGA	121	8.1	3.0
*Eco*RI	GAATTC	200	6.9	3.0
*Hin*dIII	AAGCTT	192	6.3	3.0
*Dra*I	TTTAAA	107	3.8	1.1
*Msp*I	CCGG	183	6.2	0.8
*Pal*I	GGCC	184	3.4	0.8
*Taq*I	TCGA	135	1.3	0.3

[a] Expected fragment sizes are based on 38% GC

Types of polymorphism

It is generally difficult to determine the types of mutation causing the polymorphism detected in RFLP analyses. In some cases, however, one of the lettuce lines did not show detectable hybridization in digests with any of the nine restriction endonucleases (Fig. 2); the probe was therefore detecting a deletion. This occurred with two cDNA probes and one genomic DNA probe. In later studies, with another cDNA probe, one lettuce line displayed a consistently larger hybridization fragment in digests with all five restriction endonucleases tested (data not shown); polymorphism was, therefore, probably caused by an insertion. In most cases, however, there was no clear pattern of hybridization and the cause of polymorphism could not be identified.

Discussion

RFLP studies are becoming a powerful tool for the analysis and construction of genetic maps for many organisms. The paucity of useful genetic markers in lettuce (Robinson et al. 1983) led us to develop a large number of RFLP probes and analyse their segregation in a single cross (Landry et al. 1987). In the process, we investigated the effect of the following parameters on the frequency of polymorphism detected: 1) source and type of probes; 2) restriction endonucleases; and, 3) diversity of potential parental lettuce lines.

Contrary to our expectations, cDNA clones detected 2.5 times more polymorphism than genomic DNA probes. Since transcribed regions of DNA are more conserved than noncoding regions, we expected the genomic probes, which included both transcribed and non transcribed sequences, to detect higher levels of variation than cDNA clones. Our data contrast with data from humans in which genomic clones detected polymorphism more frequently than cDNA clones (Helentjaris and Gesteland 1983). The reasons for this inconsistency are unknown. Variability could have gone undetected with repeated genomic DNA clones if the short exposure time was not long enough to detect faint single copy polymorphic bands. Also, low copy, repeated genomic sequences (dispersed or tandem) could be under concerted evolutionary mechanisms similar to those that operate on higher copy sequences (e.g. ribosomal RNA genes, Lassner and Dvorak 1986). Finally, cDNA clones represent the 3′ end of transcribed sequences and may therefore be more prone to detect polymorphism in regions flanking transcribed sequences.

Polymorphism detected with cDNA probes homologous to multiple copy sequences was higher than with cDNAs homologous to single copy sequences. cDNA clones homologous to single copy as well as multiple copy sequences frequently detected polymorphism. This supported previous observations in tomato (Bernatzky and Tanksley 1986) where 53% of the cDNA clones which detected polymorphism, mapped to single loci. These results contrast with data obtained in humans (Helentjaris and Gesteland 1983) where only the cDNA clones homologous to gene families (two of 12 clones) detected polymorphism. An increased detection of polymorphism in multiple copy sequences would be expected if the release of selection pressure on one of the copies allowed the accumulation of mutations. If 24.8% is assumed to be the basic frequency of polymorphism (Table 1) and the probability of polymorphism is independent for each locus, the frequency of polymorphism at two loci should be 43.5% $(1-(1-0.248)^2)$. cDNA probes homologous to multiple copy sequences, however, only detected 30.9% polymorphism. This, again, indicates possible homogenization of low copy sequences.

Micro-heterogeneity and variation in copy number of sub-repeat units can occur in tandemly repeated sequences (e.g. Lassner and Dvorak 1986; Saghai-Maroof et al. 1984); also, some repetitive sequences are species or chromosome specific (Liao et al. 1982; Oberle et al. 1985). Some repeated DNA sequences, therefore, could be valuable as genetic markers. Clustered sequences, such as the alpha satellite, have been utilized to trace the inheritance of particular regions of the human genome (Willard 1985). Dispersed repetitive sequences, such as the hypervariable "minisatellite" simultaneously revealed many polymorphic loci from throughout the human genome (Jeffreys et al. 1985; Wong et al. 1986); the level of polymorphism detected was so high that it provided "fingerprints" of individuals since the probability that two unrelated in-

Table 4. Relative divergence among the four accessions. Data from the nine endonucleases are pooled to show the percent polymorphism within the four probe classes

Pairwise[a] comparison	Genomic DNA		cDNA		All classes	No. pairwise comparisons	Nei's[b] genetic distance
	Single/low copy	Repetitive	Single copy	Multiple copy			
P vs G	8.3**	9.0	20.7	24.6*	14.9***	1,459	0.160
P vs C	11.0	11.3	27.9	34.4	20.4	840	0.125
P vs K	16.6	5.5	23.4	44.7***	22.3*	882	0.196
G vs C	14.2	10.2	36.7***	25.0	21.4	826	0.121
G vs K	17.5	5.9	25.7	38.0	21.8	866	0.138
C vs K	10.0	10.9	21.5	24.0*	17.6	1,112	0.282
χ^2 (heterogeneity)[c]	18.77**	6.00	21.75***	30.91***	26.76***		

[a] P = PI251245; G = Galleaga; C = Calmar; K = Kordaat
[b] Kesseli and Michelmore 1986
[c] χ^2 tests for heterogeneity were performed on the raw data
Significant χ^2 test for heterogeneity at: * $0.01 < P < 0.05$; ** $0.001 < P < 0.01$;*** $P < 0.001$

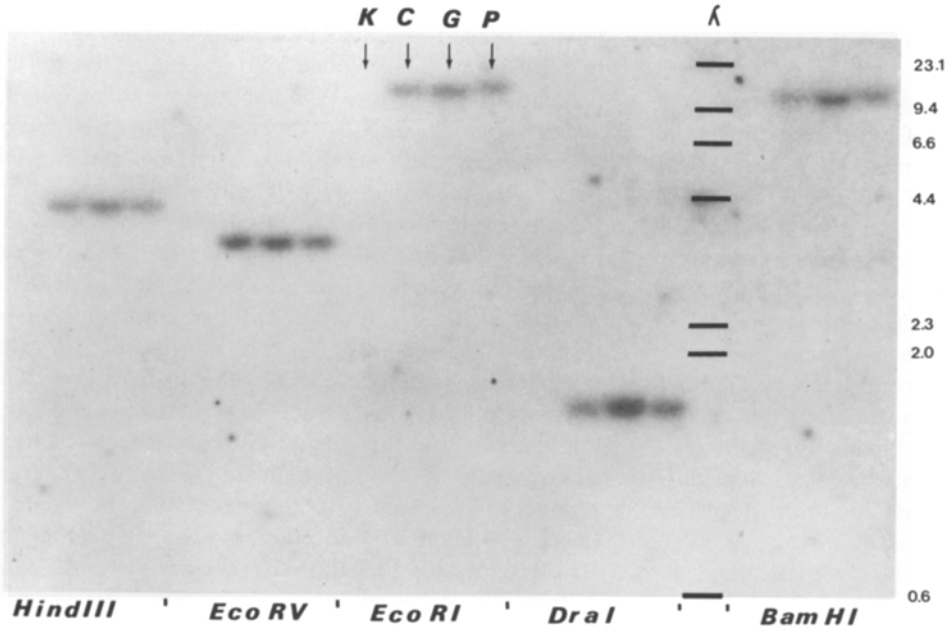

Fig. 2. Identification of a deletion with probe *GL358*. Southern blots were exposed for 3 days at −70 °C with intensifier screens (DuPont Cronex II). C = Calmar; K = Kordaat; P = PI251245; G = Galleaga; DNA size marker listed in the right is Lambda DNA digested with *Hin*dIII. Only five of the nine restriction endonucleases used to digest genomic DNAs are shown (listed above). No hybridization was detected to DNA from cv. Kordaat digested with any of the nine restriction endonucleases

dividuals share fragments at all loci is exceedingly low. Such hypervariable sequences were not, however, found within the subset of repeated DNA sequences in our study with lettuce and have not been reported for any plant species.

In humans, the efficiency of 16 restriction endonucleases for detecting RFLPs between nine unrelated individuals was compared (Barker et al. 1984). Thirty one single copy genomic DNA probes were tested. *Msp*I ($^{5'}$CCGG$^{3'}$) and *Taq*I ($^{5'}$TCGA$^{3'}$) revealed nine of the ten polymorphisms detected. The frequently methylated cytosine in CpG dimers was postulated to be hotspots for mutation in mammalian DNA as in prokaryotes (Barker et al. 1984). In a limited study of tomato and maize, however, polymorphism was not detected more frequently on DNA digested with restriction endonucleases containing a CpG dimer than with other enzymes (Helentjaris et al. 1985). In our study,

polymorphism was also not detected more frequently in digests with *Msp*I and *Taq*I than with the other restriction endonucleases. We concluded that the presence of a CpG dimer in the recognition sequence does not lead to increased detection of RFLPs and we found no evidence for CpG as a hotspot for mutations in lettuce.

The observed fragment sizes were consistently larger than expected (Table 3). This may be a consequence of DNA methylation (4 to 6% in plants; Shapiro 1975). Also, this might reflect the nonrandom distribution of restriction sites throughout the lettuce genome. In humans, the distribution of 54 restriction endonucleases in the mitochondrial DNA was analysed and 45 were found less frequently than expected (Adams and Rothman 1982). Also, the examination of nearest neighbour base sequences in eukaryote DNA established that the four bases occur next to one another in the 16 possible arrangements, not at random but in a pattern unique for that DNA (Setlow 1975). In lettuce, fragments produced by *Msp*I ($^{5'}$CCGG$^{3'}$) are on the average twice as large as those produced by *Pal*I ($^{5'}$GGCC$^{3'}$) indicating that some restriction endonuclease sites are not distributed randomly and that CpG dimers are under represented in the lettuce genome. Our size estimates may also have been biased; of the 1,353 DNA fragments measured none were smaller than 200 bp. The difficulty in obtaining clear hybridization signals from small DNA fragments might be responsible. Consistent with this, all restriction endonucleases recognizing four base pair sequences had a high frequency of missing data. Under our experimental conditions, the optimal size range for detecting polymorphic fragments was approximately 250 bp to 15 kb. Small fragments were not detected efficiently; differences in sizes were difficult to detect with large fragments due to the decreased resolution as fragment size increased.

In tomato, the choice of the parental lines, for genetic mapping studies was of critical importance (Helentjaris et al. 1985); an interspecific cross had to be analysed to ensure sufficient polymorphism. In maize, the choice of the parents was less important as extensive polymorphism was detected even between commercial inbred lines (Helentjaris et al. 1985; Johns et al. 1983; Evola et al. 1986). The lettuce accessions used in this study represented a wide range of morphological types, thus we expected to detect frequent RFLPs between accessions. The genetic distance between these lines had been previously calculated from isozyme data (Kesseli and Michelmore 1986) but these values did not correlate to the frequencies of polymorphism detected with RFLP probes (Table 4). Two possible reasons for this discrepancy are: 1) the sample size of only six pairwise comparisons makes a significant correlation difficult to detect; and, 2) of the 45 isozyme loci surveyed, at most five were polymorphic for any pairwise comparison of the accessions in this study. The latter explanation is particularly pertinent since it is now known that two of the isozyme loci, *Est*8 and *Dia*3 are linked to different downy mildew resistance genes (*Dm*1 and *Dm*5/8 respectively, Farrara and Michelmore 1987). Since Calmar and Kordaat are fixed for different resistance genes and isozyme loci, it is likely that their genetic distance (Nei's D = 0.28) is inflated by artificial selection for the different resistance genes in different breeding programs and hitchhiking of the linked isozyme loci.

In conclusion, it is technically easier to obtain low copy genomic sequences for probes than to make a cDNA library. The latter is, however, becoming increasingly routine and the former requires the only partially efficient selection against clones carrying repeated sequences. cDNA clones detected more polymorphism than genomic DNA probes and, consequently, were preferred for the initial development of a linkage map. Genomic DNA probes might, however, be required to complete the map if some regions of the genome are not detected by cDNA probes. The criteria which should be used to choose the restriction endonucleases for RFLP analysis are: 1) Does the enzyme detect an acceptable level of polymorphism? and, 2) Does it result in only a low frequency of missing data? *Bam*HI, *Eco*RI, *Eco*RV and *Hin*dIII best fulfilled these criteria and were the most efficient in detecting RFLPs. *Dra*I, *Msp*I, *Pal*I, *Taq*I and *Xba*I were unpredictable or displayed significantly less polymorphism. The frequencies of RFLPs detected in the six pairwise comparisons of the four lettuce lines corresponded only approximately to the genetic distances calculated from isozyme analyses. Phylogenetic data based on a limited number of markers can, therefore, only be used to make a preliminary selection of diverse parents for RFLP analysis. RFLP data are needed to confirm the choice.

Acknowledgements. We thank Dr. William E. Timberlake for advice, Dr. Danny Alexander (ARCO PCRI, now Calgene Inc.) for helping in the construction of the lettuce cDNA library and Mike Lassner (UC Davis) for his comments on the manuscript. This research was supported by United States Department of Agriculture grant number 85-CRCR-1-1561. B. Landry was sponsored by Agriculture-Canada, research division.

References

Adams J, Rothman ED (1982) Estimation of phylogenetic relationships from DNA restriction patterns and selection of endonuclease sites. Proc Natl Acad Sci USA 79:3560–3564

Alexander DC, McKnight TD, Williams BG (1984) A simplified and efficient vector-primer cDNA cloning system. Gene 31:79–89

Barker D, Schafer M, White R (1984) Restriction sites containing CpG show a higher frequency of polymorphism in human DNA. Cell 36:131–138

Bernatzky R, Tanksley SD (1986) Majority of random cDNA clones correspond to single loci in the tomato genome. Mol Gen Genet 203:8–14

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphism. Am J Hum Genet 32: 314–331

Cathala G, Savouret JF, Mendez B, West BL, Karin M, Martial JA, Baxter JD (1983) A method for isolation of intact translationally active ribonucleic acid. DNA 2:329–335

Evola SV, Burr FA, Burr B (1986) The suitability of restriction fragment length polymorphisms as genetic markers in maize. Theor Appl Genet 71:765–771

Farrara B, Michelmore RW (1987) Genetic analysis of factor for resistance to downy mildew (*Bremia lactucae*) in lettuce (*Lactuca sativa*). Plant Pathol (in press)

Feinberg SE (1977) The analysis of cross-classified categorical data. MIT Press, Massachusetts

Fisher RL, Goldberg RB (1982) Structure and flanking regions of soybean seed protein genes. Cell 29:651–660

Green BR (1972) Isolation and base composition of DNAs of primitive land plants. Biochim Biophys Acta 277:29–34

Gusella JF (1986) DNA polymorphism and human disease. Ann Rev Biochem 55:831–854

Helentjaris T, Gesteland R (1983) Evaluation of random cDNA clones as probes for human restriction fragment polymorphism. J Mol Appl Genet 2:237–247

Helentjaris T, King G, Slocum M, Siedenstrang C, Wegman S (1985) Restriction fragment polymorphisms as probes for plant diversity and their development as tools for plant breeding. Plant Mol Biol 5:109–118

Jeffreys AJ, Wilson V, Thein SL (1985) Individual-specific fingerprints of human DNA. Nature 316:76–79

Johns MA, Strommer JN, Freeling M (1983) Exceptionally high levels of restriction site polymorphism in DNA near the maize AdhI gene. Genetics 105:733–743

Kesseli RV, Michelmore RW (1986) Genetic analysis and phylogenies detected from isozymes markers among species of *Lactuca*. J Hered 77:324–331

Landry BS, Michelmore RW (1985) Selection of probes for restriction fragment length analysis from plant genomic clones. Plant Mol Biol Rep 3:174–179

Landry BS, Kesseli RV, Farrara B, Michelmore RW (1987) A genetic map of lettuce (*Lactuca sativa* L.) with restriction fragment length polymorphisms, isozymes, disease resistance genes and morphological markers. Genetics 116: 331–337

Lassner M, Dvorak J (1986) Preferential homogenization between adjacent and alternate subrepeats in wheat rDNA. Nucleic Acids Res 14:5499–5512

Liao Law M, Davidson JN, Kao FT (1982) Isolation of a human repetitive sequence and its application to regional chromosome mapping. Proc Natl Acad Sci USA 79: 7390–7394

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning. Cold Spring Harbor, New York

Oberle I, Drayna D, Camerino G, White R, Mandel JL (1985) The telomeric region of the human X chromosome long arm: Presence of a highly polymorphic DNA marker and analysis of recombination frequency. Proc Natl Acad Sci USA 82:2824–2828

Reed KC, Mann DA (1985) Rapid transfer of DNA from agarose gels to nylon membranes. Nucleic Acids Res 13:7207–7221

Rigby PWJ, Dieckmann M, Rhodes C, Berg P (1977) Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J Mol Biol 113: 237–251

Robinson RW, McCreight JD, Ryder ED (1983) The genes of lettuce and closely related species. Plant Breed Rev 1: 267–293

Rodriguez RL, Tait RC (1983) Recombinant DNA techniques, an introduction. Addison-Wesley, Massachusetts

Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer length polymorphism in barley: Mendelian inheritance, chromosomal location and population dynamics. Proc Natl Acad Sci USA 81:8014–8018

Setlow P (1975) Nearest neighbour frequencies in deoxyribonucleic acids. In: Fasman GD (ed) Handbook of biochemistry and molecular biology. Nucleic acids 2. CRC press, Cleveland, Ohio, pp 312–318

Shapiro HS (1975) Distribution of purines and pyrimidines in deoxyribonucleic acids. In: Fasman GD (ed) Handbook of biochemistry and molecular biology. Nucleic acids 2. CRC press, Cleveland, Ohio, pp 241–311

Southern EM (1975) Detection of specific sequences among DNA fragments separated by electrophoresis. J Mol Biol 98:503–517

Tautz D, Renz M (1983) An optimized freeze-squeeze method for the recovery of DNA fragments from agarose gels. Anal Biochem 132:14–19

Timberlake WE (1978) Low repetitive DNA content in *Aspergillus nidulans*. Science 202:973–975

Willard HF (1985) Polymorphic alpha satellite repeated DNA at the centromere of the X chromosome. Cyto Cell Genet 40:778

Wong Z, Wilson V, Jeffreys AJ, Thein SL (1986) Cloning a selected fragment from a human DNA fingerprint: Isolation of an extremely polymorphic minisatellite. Nucleic Acids Res 14:4605–4616