Theor Appl Genet (1996) 93:1202–1210

M. Hill · H. Witsenboer · M. Zabeau · P. Vos
R. Kesseli · R. Michelmore

PCR-based fingerprinting using AFLPs as a tool for studying genetic relationships in *Lactuca* spp.

Received: 16 November 1995 / Accepted: 23 February 1996

Abstract AFLP markers were evaluated for determining the phylogenetic relationships *Lactuca* spp. Genetic distances based on AFLP data were estimated for 44 morphologically diverse lines of cultivated *L. sativa* and 13 accessions of the wild species *L. serriola*, *L. saligna*, *L. virosa*, *L. perennis*, and *L. indica*. The same genotypes were analyzed as in a previous study that had utilized RFLP markers. The phenetic tree based on AFLP data was consistent with known taxonomic relationships and similar to a tree developed with RFLP data. The genetic distance matrices derived from AFLP and RFLP data were compared using least squares regression analysis and, for the cultivar data, by principal component analysis. There was also a positive linear relationship between distance estimates based on AFLP data and kinship coefficients calculated from pedigree data. AFLPs represent reliable PCR-based markers for studies of genetic relationships at a variety of taxonomic levels.

Key words AFLP marker · *Lactuca* · Lettuce · Phylogenetic analysis

Introduction

Molecular markers have been used for a variety of studies of diversity in crop species. These investigations have provided insights into genome evolution (Bonierbale et al. 1988; Song et al. 1990; Tanksley et al. 1988), the phyletic origins of cultivated species (Kesseli et al. 1991; Skroch et al. 1993), and the current levels of diversity in modern agricultural crops (Cox et al. 1986; Kazan et al. 1993; Vierling and Nguyen 1992). Genetic markers can be useful for several aspects of breeding programs including acceleration of backcross programs and the study of quantitative traits (Tanksley 1989; Melchinger et al. 1990; Smith et al. 1990). Markers are also critical for the preservation and analysis of germplasm diversity (Kresovich et al. 1992).

Several types of markers have been used to assess genetic diversity. Hybridization-based assays, particularly restriction fragment length polymorphism (RFLP) analysis (Botstein et al. 1980), have been valuable tools for detecting patterns of DNA polymorphism among and within plant species. While RFLPs have been analyzed in a number of systematic studies (Akopyanz et al. 1992; dos Santos et al. 1994; Kesseli et al. 1991; Song et al. 1990), the procedure is laborious, expensive, and few loci are detected per assay. Random amplified polymorphic DNA (RAPD) analysis allows large numbers of markers to be assayed inexpensively using the polymerase chain reaction (PCR) technique) and single oligonucleotide primers of arbitrary sequence (Welsh and McClelland 1990; Williams et al. 1990). RAPD analysis is technically straight-forward compared to RFLP analysis and requires only nanograms of genomic DNA.; however, RAPD markers are usually dominant and the results can be sensitive to subtle changes in reaction conditions. RAPDs have been successfully used to analyze genetic relationships in a number of taxonomic groups (Chalmers et al. 1992; Demeke et al. 1992; Fukuoka et al. 1992; Koller et al. 1993; Vierling and Nguyen 1992; Yang and Quiros 1993; Williams and StClair 1993). Microsatellites or simple sequence repeats offer the potential for codominant, PCR-based markers; however, the development of locus-specific oligonucleotide primers is time-consuming as well as expensive, and few loci are identified per reaction. Therefore, despite the range of techniques available, there is still the need for a reliable, inexpensive marker technology.

Recently, a new multiplex PCR-based method [AFLP (trademark) analysis] has been developed (Zabeau and Vos 1993; Vos et al. 1995). A subset of restriction fragments are selectively amplified using oligonucleotide primers complementary to sequences that have been ligated to each end. Polymorphisms are usually detected as the presence or absence of an amplified restriction fragment and are therefore dominant. The procedure first involves the digestion of genomic DNA with two restricition enzymes. Oligonucleotide

Communicated by H.F. Linskens

M. Hill · R. Kesseli
Department of Biology, University of Massachusetts, Boston, MA 02125-3393, USA

H. Witsenboer · R. Michelmore (✉)
Department of Vegetable Crops, University of California, Davis, CA 95616, USA

M. Zabeau · P. Vos
Keygene BV, AgroBusiness Park 90, P.O. Box 216, NL-6700 AE Wageningen, The Netherlands

adaptors of known sequence are ligated to each end. Restriction fragments are then amplified using a pair of primers that are complementary to the adaptor sequence and the remains of the restriction site plus up to four random nucleotides at the 3′ end. The number of restriction fragments amplified is determined by the complexity of the template and the number of selective nucleotides at the 3′ end of the primer. Depending on the genomic complexity of the template, one or two sets of amplification and primer pairs are used. The radiolabelled PCR products are size-fractionated on a polyacrylamide gel. AFLP analysis allows the reliable identification of over 50 loci in a single reaction (Zabeau and Vos 1993; Vos et al. 1995).

Lactuca L. is a widely-distributed genus of the Compositae and contains the cultivated species *L. sativa* (lettuce). *L. sativa* belongs to a subsection of Lactuca that also includes three well-defined wild species, *L. serriola*, *L. saligna*, and *L. virosa*. Genetic relationships among accessions of this subsection and two other wild species, *L. indica* and *L. perennis*, have been investigated using RFLP markers (Kesseli et al. 1991). The resultant dendrogram was consistent with earlier estimates of inter-specific relationships that had been based on morphological, physiological, cytological, or isozyme evidence (Linqvist 1960; Robinson et al. 1976; Kesseli and Michelmore 1986). We have been developing a genetic map for lettuce using RFLP, RAPD, and isozyme markers based on an intra-specific cross (Landry et al. 1987; Kesseli et al. 1994). Most recently, we have been constructing inter-specific and intra-specific genetic maps that include AFLP markers (H. Witsenboer et al. unpublished). Both mapping and systematic analyses are part of our investigations of the evolution of disease resistance genes within the genus (Farrara et al. 1987; Hulbert and Michelmore 1985; Witsenboer et al. 1995; Michelmore et al. 1994).

In the present study, we evaluated the use of AFLP markers for determining phylogenetic relationships in *Lactuca* using the same genotypes as had been analyzed in the previous RFLP analysis. Genetic distances based on AFLP data were estimated for 44 morphologically diverse lines of cultivated *L. sativa* and 13 accessions of the wild species *L. serriola*, *L. saligna*, *L. virosa*, *L. perennis*, and *L. indica*. The AFLP distance matrix was compared with the original RFLP data set using a least squares regression analysis. A principle component analysis was also performed to contrast AFLP and RFLP distance matrices for the cultivars. In addition, a comparison was made between AFLP distance estimates and kinship coefficients calculated from pedigree data.

Materials and methods

Plant material

A total of 57 accessions of morphologically diverse inbred lines of *sativa L.*, and five wild species of *Lactuca* were examined. DNA was extracted from leaves as described previously (Landry et al. 1987). DNA of each accession represents a bulked sample from at least 20 plants. As all the species, except *perennis* L., are self-pollinators and accessions are likely founded on a narrow genetic base (due to selective breeding in cultivars and small population sizes for wild species), heterogeneity within samples is low (Kesseli et al. 1986,1991)

AFLP analysis

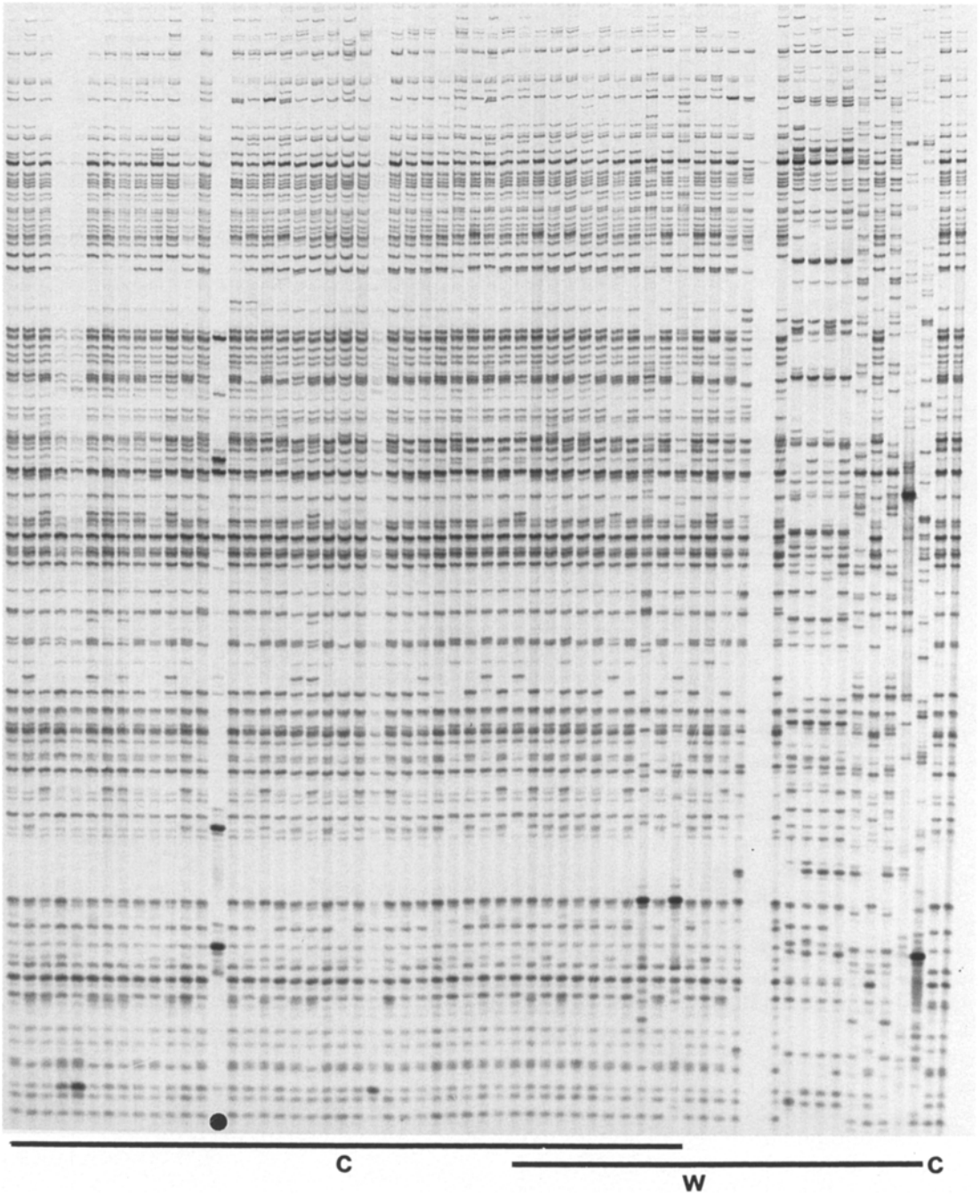
AFLP protocol was performed essentially as developed by Keygene (Zabeau and Vos 1993; Vos et al. 1995) with minor modifications. All primer and adaptor sequences were designed by Keygene (Table 1). A simpler version of this protocol is now available that omits the selection of a subset of fragments on streptavidin beads (Zabeau and Vos 1995). Genomic DNA (0.5 µg aliquots) was digested with 5 units of both *Eco*RI and *Mse*I. *Mse*I adaptors and biotinylated *Eco*RI adaptors were ligated to the restriction fragments. Biotinylated fragments were selected using streptavidin beads (Dynal, Lake Success, N.Y.) using the binding and washing buffer (10 m*M* Tris-HCL pH 7.5, 1.0 m*M* EDTA, 2.0 *M* NaCl) recommended by Dynal. Magnetic beads were resuspended in 200 µl of a modified TE buffer (10 m*M* Tris HCl, 0.1 m*M* EDTA, pH 8.0) and stored at (− 20°C). PCR amplification was done in a 9600 thermocycler (Perkin Elmer Cetus) in two steps. A selective preamplification was performed using 2.5 µl of template DNA on magnetic beads, 37.5 ng each of two primers with one selective nucleotide (one for the *Eco*RI side and for the *Mse*I side of the fragments; Tabel 1), 0.2 m*M* of ultrapure dNTPs (Pharmacia), 0.5 units of *Taq* polymerase (Perkin Elmer Cetus), in PCR buffer (10 m*M* Tris pH 8.3, 1.5 m*M* $MgCl_2$, 50 m*M* KCl) in a final volume of 25 µl. The cycle profile was 30 s at 94°C, 30 s at 60°C, 60 s at 72 °C for 30 cycles. After preamplification, the reaction mixture was diluted eight times in the modified TE buffer and stored at − 20°C. The second amplification was performed using 5 µl of the diluted preamplification mixture, 30 ng each of two primers with three selective nucleotides (Table 1) [6.25 ng of one of the primers was end-labeled with [^{33}P] or [^{32}P]γ-ATP using T_4 polynucleotide kinase (New England Biolabs)], 0.2 m*M* of ultrapure dNTPs, 0.4 units of *Taq* polymerase, in PCR buffer in a final volume of 20 µl. The touch-down cycle profile was 30 at 94°C, 30s at 65°C, 60s at 72°C for 1 cycle; during the next 11 cycles the annealing temperature was lowered 0.7°C in each cycle, followed by 24 cycles with an annealing temperature of 56°C. After amplification, samples were mixed with an equal volume of 98% formamide, 10 m*M* EDTA and bromophenol blue and xylene cyanol as tracking dyes. The resulting mixtures were heated for 3 min at 90°C and run on a 4.5% polyacrylamide gel under

Table 1 Oligonucleotide adaptors and primers used for AFLP analysis

*Eco*RI adaptors[a]	CTCGTAGACTGCGTACC
	CTGACGCATGGTTAA
*Mse*I adaptors[a]	GACGATGAGTCCTGAG
	TACTCAGGACTCAT
AFLP primers[b]	
*Eco*RI + 1:	AGACTGCGTACCAATTC+A / T
*Mse*I + 1:	GACGATGAGTCCTGAGTAA+G/A/T/C
*Eco*RI + 3:	
SO3	GACTGCGTACCAATTC+AAG
SO5	GACTGCGTACCAATTC+ACA
B12	GACTGCGTACCAATTC+AGG
*Mse*I + 3	
F44	GATGAGTCCTGAGTAA+GGA
G17	GATGAGTCCTGAGTAA+GCA
G18	GATGAGTCCTGAGTAA+GCT

[a] *Eco*RI and *Mse*I adaptors were ligated onto the ends of genomic restriction fragments

[b] *Eco*RI + 1 and *Mse*I + 1 primers were used in the preamplification of template DNA. The AFLP fingerprint was generated using pairs of *Eco*RI + 3 and *Mse*I + 3 primers



standard conditions for sequencing gels (Sambrook et al. 1989). As a size marker, a sequencing reaction of pUC19 was run on the same gel. The gel was fixed in 10% methanol, 10% glacial acetic acid for 12 min, then dried in a vacuum dryer at 80°C and exposed to film for 1 [^{32}P-labeled primer) to 4 ([^{33}P]-labeled primer) days to produce autoradiographs. Polymorphic loci were identified as bands in the autoradiographs and scored as either present (1) or absent (0).

Fig. 1 AFLP analysis of *Lactuca* spp. The DNA fingerprints were generated using primer pair B12 + F44 to amplify genomic fragments from *L. sativa* (*C*) or wild *Lactuca* spp. (*W*). Due to the consequences of loading 12 lanes simultaneously from a microtitre plate, the *C* and *W* lanes alternate in the overlap region. The lane marked *O* resulted from a degraded DNA sample that was excluded from the analysis. The empty lane was not loaded

Genetic analysis

AFLP markers were generated using four pairs of primers, B12 + F44, S03 + G17, S05 + G18, and B12 + G18 (Table 2). Each locus was treated as a separate character and scored as either present (1) or absent (0) across all genotypes. Genetic similarities between pairs were estimated using the formula of Nei and Li (1979) as

$$Gs(ij) = C(ij)/N(ij)$$

where *Gs* is the measure of genetic similarity between the ith and jth accession, $C(ij)$ is the number of loci in i and j that both have a band present, and $N(ij)$ is the total sum of scored bands (Nei and Li 1979). Genetic distance between lines were calculated as $-\ln(Gs)$ (Nei and Li 1979). A dendrogram was constructed using the unweighted pair

Table 2 DNA markers identified using four pairs of AFLP primers

Primer pair	Number markers of all lines	Number of markers for *L. sativa*
B12, F44	151	40
SO3, G17	88	21
SO5, G18	81	17
B12, G18[a]	–	19
Total	320	97

[a] Markers detected by B12+G18 were only used in the principle component analysis of *L. sativa*

group method average (UPGMA) clustering procedure. Calculation of the distance matrix and cluster analysis were carried out using the program NTSYS-pc (Rohlf 1992).

The genetic data obtained with AFLPs was assessed by making comparisons with our previous RFLP-based phylogenetic study (Kesseli et al. 1991). To determine the level of correlation between elements of the AFLP and RFLP distance matrices, we performed a least square linear regression analysis, and a Pearson product-moment correlation coefficient (r_p) was calculated. A direct comparison between the two marker techniques was also made using a principle coordinate analysis (PCO)(Gower 1966) to compare their abilities to distinguish between the four cultivated types of *L. sativa*. The PCO was performed using the DCENTER and EIGEN procedures in NTSYS-pc (Rohlf 1992)

Estimations of kinship coefficient (k) for 13 cultivars and 28 pairwise comparisons were obtained from breeding records and personal communications with breeders. The value of k for two genotypes X and Y is defined as the probability that a random allele at a locus from X is identical by descent with a random allele at that same locus from Y. The assumptions are that each allele has equal probabilty of being transmitted from a parent and that there is no selection (Cox et al. 1986; Martin et al. 1991). Basal ancestors of a pedigree were assigned similarity values based on AFLP data (the ancestral lines in all pedigrees were always inbred cultivars). A least square regression analysis was used to compare $(1-k)$ values, and distance values (d) were estimated with AFLP markers for accessions with known pedigrees.

Results

Each pair of primers resulted in the amplification of large numbers of AFLP fragments from the 57 accessions of *Lactuca* L. spp. (Fig. 1). All bands that could be reliably read on each autoradiograph were treated as individual dominant loci and scored as either present (1) or absent (0) across all gen.otypes. Fragments of the same size in different genotypes were considered to be the same allele. This is difficult to confirm without extensive genetic analysis or hybridization studies; however, the correspondence of the AFLP data with analyses based on other types of markers (see below) indicates that the scoring of multiple fragments as one locus is not common. A total of 320 polymorphic AFLP loci were identified using only three pairs of primers, B12+F44, S03+G17, and S05+G18. Only 5 fragments were monomorphic across all genotypes tested. Many fragments were restricted to a single species. The number of scored loci amplified by each pair of primers

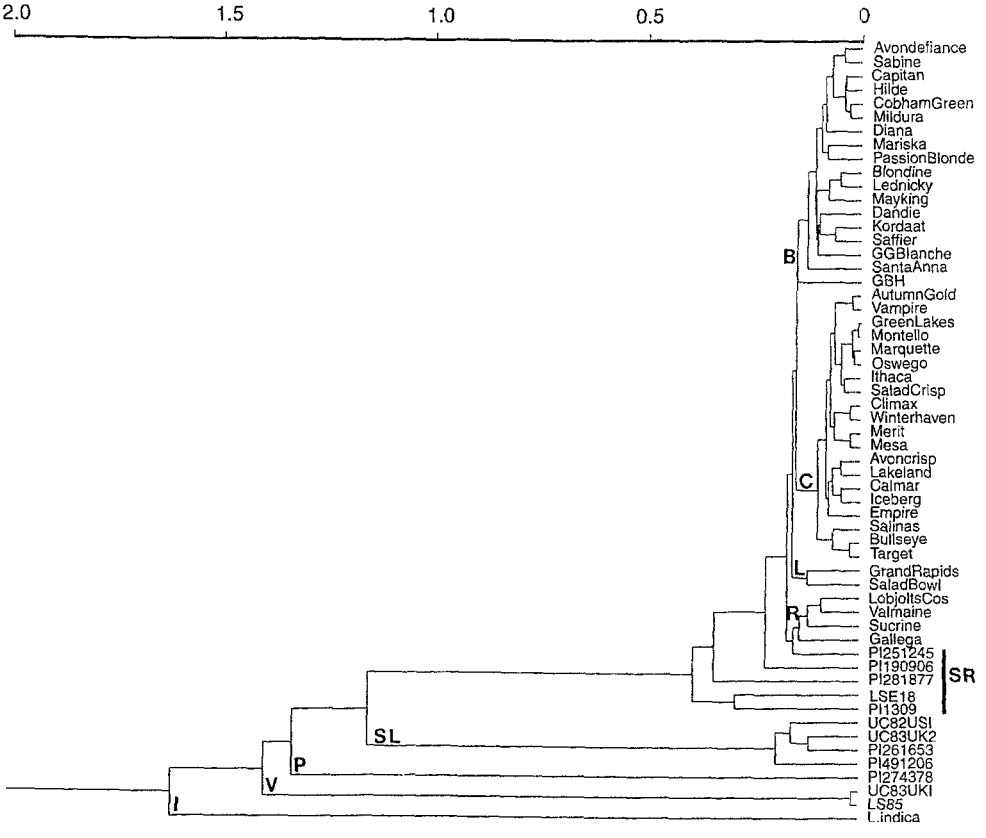


Fig. 2 Phenetic tree of *Lactuca* spp. based on AFLP polymorphisms. Phenetic relationships were derived from pairwise genetic distance estimates between 57 genotypes. Cluster analysis was performed using the UPGMA method. Different plant types formed distinct branches: *B* butterhead, *C* crisphead, *L.* looseleaf, *R* romaine, *SR L. serriola, SL L. saligna, L. virosa, P L. perennis, I L. indica.*

set varied from 81 to 151 with an average of 107 per reaction. The size of amplified fragments (not including the adapters ligated on the ends) ranged from 30 to 350 bp.

The AFLP genotypic data were used to calculate pairwise genetic distances according to the formula of Nei and Li (1979), and a cluster analysis was performed using the UPGMA method. The resultant dendrogram (Fig. 2) grouped all species as distinct taxa and was consistent with previously defined inter-species relationships (Kesseli et al. 1991). The accessions of *L. serriola* L. had the highest mean intra-specific distance of all the seven species, 0.35 ± 0.06, and clustered on a sister branch of the *L. sativa* complex. This is consistent with *L. serriola* L. as the likely progenitor species of *L. sativa*. The four accessions of *L. saligna* and two accessions of *L. virosa* L. each clustered as distinct units with a relatively low mean intra-specific genetic distance of 0.16 ± 0.05 and 0.02 respectively; greater intra-specific variation may have been observed if larger sample sizes had been analyzed. Except for *L. serriola* that clustered with *L. sativa*, all species formed their own major branch that was distinct from other species by a distance value greater than 1.0 *L. perennis* L. represented the outlying species and was separated from all other species by a distance of 1.74. About 20% of the AFLP markers identified in *L. perennis* were unique.

The 44 accessions of *L. sativa* were subdivided as discrete branches according to plant type: butterhead, crisphead, romaine, and looseleaf (Fig. 2). Each type can be separated from the others by a phenoline drawn in the dendrogram at 0.15. The accession, PI251245, is a landrace type that is intermediate between *L. sativa* and *L. serriola*; this was located on a sister branch to the 4 romaine lettuces that lay just outside the phenoline of 0.15. Of the 119 bands identified within the *L. sativa* group, 36 were monomorphic across all cultivars. Ten bands were uniquely associated with a single cultivar. Six of these bands were unique to the looseleaf type, 2 were unique to the butterhead type, while the crisphead and romaine types each contained a single distinct band.

Comparison of AFLP and RFLP data

The AFLP data were compared to RFLP data derived from analysis of the same 56 accessions (AFLP data for cv 'Lakeland' was not included in this comparison because 'Lakeland' was not analyzed in the RFLP study by Kesseli et al. 1991). Although AFLP and RFLP data resulted in similar dendrograms, the overall genetic distance between taxa was generally higher with RFLP markers. The Pearson product-moment correlation coefficient (r_p) between the elements of the AFLP and RFLP distance matrices is 0.86 ($P < 0.001$). A least squares linear regression for 1540 pairwise distance estimates indicates that AFLPs predict about 60% of the genetic distance found with RFLPs ($r^2 = 0.88$, $P <$ 0.001) (Fig. 3). Some coordinates deviate considerably from the regression line especially at high distance values where there are many points that exceed the 95% confidence level of the slope. Points within the circled region are comparisons of distance estimates between species, while the remaining points clustered at the lower end of the graph contrast estimates within species (Fig. 3).

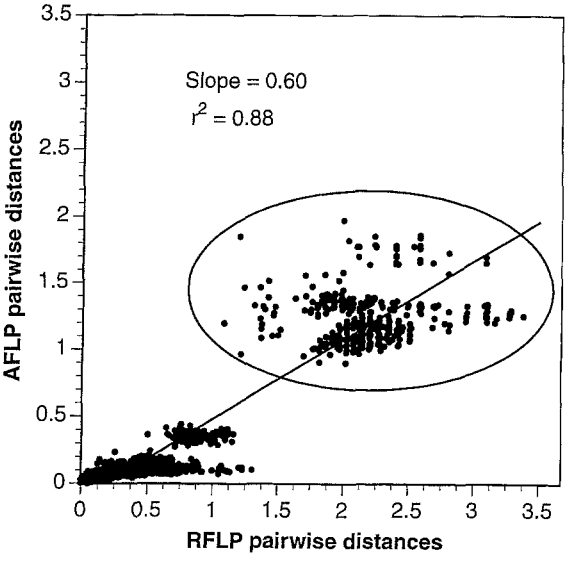


Fig. 3 Comparison of pairwise genetic distance estimates based on AFLP and RFLP data. Estimates of 1540 pairwise distances calculated from AFLP and RFLP data were compared by least squares linear regression analysis. *Circled* data points indicate pairwise distance comparisons between species

A second comparison between AFLP and RFLP distance matrices was made using a principle coordinate analysis (PCO) to test the ability of each method to distinguish between the 43 lines of *L. sativa* that has a narrow genetic base (Fig. 4). The AFLP distance matrix for this analysis was calculated using 97 bands that were polymorphic within the accessions of *L. sativa* (monomorphic loci were excluded). Fifteen of these fragments were identified using a fourth pair of primers, B12 + G18. The first two eigenvalues accounted for 60% and 94% of the variation among accessions for AFLP and RFLP markers, respectively. AFLP analysis resulted in a more definitive grouping of accessions; three of the four plant types clustered in separate quadrants of the plots (Fig. 4a). The 2 looseleaf cultivars were also separated from the other 3 cultivar types but are not obviously clustered, partially due to the small sample size and the heterogeneity of this plant type. A plot of the first two principal components of RFLP distance estimates shows less differentiation of the lettuce groups with all four cultivar types overlapping near the center of the plot (Fig. 4b).

Correlation between AFLP distances and kinship coefficients

Estimations of kinship coefficient (k) were made from pedigree data obtained from breeding records and personal communication with breeders. The k statistic estimates the probability that two alleles at a given locus are identical by descent (Cox et al. 1986; Martin et al. 1991) and gives a measurement of the similarity between breeding lines. The distance between accessions can therefore be estimated as $1-k$ (Tinker et al. 1993). The relationship between ancestors at the basal end of a pedigree were assigned a value based on similarity estimates obtained from the AFLP data set. This method is biased in that estimates of $1-k$ and genetic distance (d) will fall within the same range; however, the relatedness between accessions along various breeding lines should be a function of the pedigree itself and not the original level of similarity assigned to ancestral lines.

Values of $1-k$ were calculated using 13 different cultivars, for a total of 28 pairwise comparisons. The distribution of values ranged from 0.001 to 0.212 with an average value of 0.068. Genetic distance estimates for the same pairs of individuals were compared with pedigree data by graphing d as a function of $1-k$ (Fig. 5). A least squares liner regression indicated a highly linear relationship between methods with a slope of 0.85 ± 0.068 ($r^2 = 0.86$; $P < 0.001$). The coordinates of the graph deviated considerably from the linear relationship at lower values where accessions are most closely related and may reflect the selection criteria used in the production of newly created lines. The Spearman ranked correlation (r_s) coefficient between $1-k$ and d for the 28 observations was highly significant at 0.86 ($P < 0.01$).

Discussion

AFLP markers are useful for measuring genetic diversity and determining genetic relationships within and among species as judged by several criteria. AFLPs distinguished between previously established taxonomic units at both the species and cultivar level in *Lactuca* spp. The close relationship between *L. sativa* and *L. serriola* is consistent with previous studies indicating *L. serriola* as the progenitor species of cultivated lettuce (Linquist 1960; Kesseli et al. 1991). The more distant placement of both *L. saligna* and *L. virosa* from *L. sativa* parallels earlier work; 75% of the RFLP loci in *L. sativa* have no alleles in common with either wild species (Kesseli et al. 1991). The dendrogram obtained from AFLP fingerprints is similar to those derived from other marker types. Furthermore, there was a highly significant correlation ($r_p = 0.86$, $P < 0.001$) between elements of the AFLP and RFLP distance matrices. The clear clustering of *L. sativa* accessions by principle com-

Fig. 4 Principle component analysis (PCO) of pairwise genetic distances of 44 cultivated varieties of *L. sativa*. **a**. PCO based on AFLP data **b** PCO based on RFLP data. See text for details

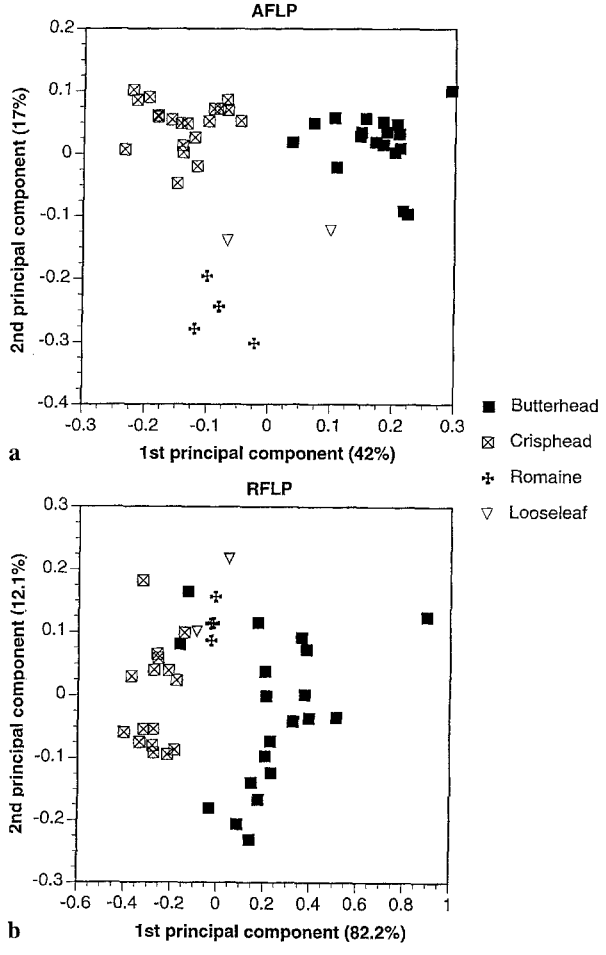
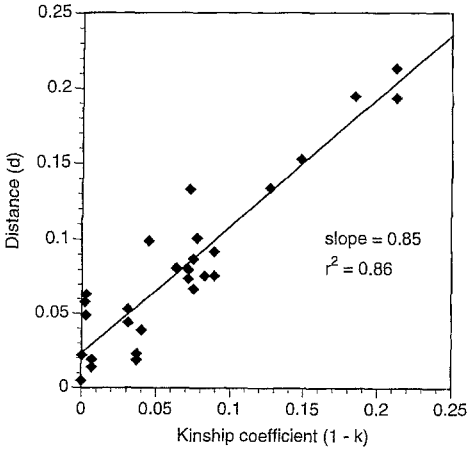


Fig. 5 Comparison of AFLP genetic distance estimates and kinship coefficients calculated rom pedigree data. AFLP genetic distance estimates (d) and kinship coefficients (k) were compared using least squares regression analysis. Kinship coefficients were calculated from breeding records describing 13 cultivars of *L. sativa* (Table 3). Using measurements of k, we estimated the distance between accessions as $(1-k)$. A total of 28 pairwise comparisons were made. The slope of the best fit line is 0.85



ponent analysis indicated that AFLPs were capable of distinguishing between the major plant types at least as well as RFLP analysis. This result was interesting because *L. sativa* has a narrow genetic base; of the AFLP markers scored, 30% of those that were present in *L. sativa* were monomorphic. Therefore, AFLP analysis provides an independent estimate of genetic relationships that is consistent with other molecular techniques.

The regression of AFLP and RFLP distance elements revealed that AFLPs generally detected a lower level of genetic diversity among the genotypes. This is especially true for cultivated accessions of *L. sativa* where the mean intra-specific genetic distances were estimated at $0.40 + 0.22$ for RFLPs (Kesseli et al. 1991) and $0.18 + 0.06$ for AFLPs. This difference is almost certainly due to the method used to select probes for RFLP analysis. As usually only a single locus is identified per RFLP assay and as we wished to analyze genetically defined loci, probes were selected that were known to detect polymorphism in the intra-specific mapping population of *L. sativa*, cvs 'Calmar' (crisphead) × 'Kordaat' (butterhead). Random probes that may have identified monomorphic loci were not analyzed. Consequently, the previous distance estimates based on RFLP data are inflated.

There was a positive linear relationship between kinship coefficients (k) and AFLP distance estimates (d) based on 28 pairs of accesions ($r_s = 0.86$). Since k values were calculated by assigning the basal ancestors of each pedigree a genetic distance estimated from the AFLP data set, a plot of d as a function of $1 - k$ should yield a line with a slope of 1. The regression line for the observed coordinates of $1 - k$ and d produced a slope of 0.85 ± 0.068; the expected value of 1 lies just outside the upper 95% confidence limit (0.99). This may be a consequence of selection violating the assumption of neutrality inherent in these calculations and may account for the deviation from the line at low values of d and $1 - k$ (Fig. 5). The overall Spearman rank correlation coefficient is higher than reported in previous studies of inbred maize ($r_s = 0.74$), barley ($r_s = 0.73$), and modern soybean cultivars ($r_s = 0.60$) (Melchinger et al. 1991; Tinker et al. 1993; Cox et al. 1985). Because the genetic diversity between the 28 accessions used in this study was low (d varied between 0.014 and 0.214), no conclusions can be drawn as to the relationship of kinship coefficients and estimates of d for more distantly related genotypes.

AFLP analysis has several advantages over other marker techniques for the analysis of genetic diversity. A major advantage of AFLP analysis is the short time required to assay large numbers of DNA loci. Compared to RAPDs, AFLP analysis detected about 12 times the number of polymorphic loci per assay in soybean varieties (Vogel et al. 1994). The 320 markers used in our dataset were generated in about 1 week, although scoring and analysis took considerably longer. This represents an almost 50-fold decrease in production time compared with the time required to generate data for the 143 RFLP markers used in the previous study. The speed of analysis would be considerably further increased by the automated scoring of AFLP loci. Another advantage of AFLP analysis is the effectively unlimited number of loci that can be assayed with different combinations of a relatively small number of oligonucleotide primers. The ability of AFLPs to distinguish among genotypes is not hindered by their bi-allelic nature (presence or absence). PCO and ordination of similarity matrices constructed with RAPD markers have resulted in similar conclusions in other plant species (Demeke et al. 1992; Heun et al. 1994). However, because it is straightforward to assay many AFLP loci reliably, polymorphisms can be identified between closely related genotypes. Therefore, AFLP analysis is a more powerful approach than RAPD analysis. AFLP analysis does require more preparation of template DNA than RAPD analysis. Consequently, if genotypes of only a few loci are to be assayed for many accessions, then RAPD markers may still be advantageous.

AFLP loci can be analyzed more reliably than RAPD markers and therefore may provide more accurate markers for the measurements of genetic diversity. The AFLP procedure is robust and relatively insensitive to reaction conditions, unlike RAPD reactions (Lamboy 1994). The most critical step in AFLP analysis is ensuring complete digestion of the genomic DNA; however, with experience and by reference to the usually monomorphic bands, the products of partial digestions are obvious. The number of fragments amplified can be controlled by adjusting the number and base compositon of the 3′ selective bases (Vos et al. 1995). In combination with the accurate separation of fragments on a polyacrylamide gel, this will reduce bias in distance estimates due to the miscoring of fragments of similar length from different regions of the genome as identical. Biases in diversity calculations can also be introduced by false positives (bands that do appear in a lane but should not) due to differential competition for primer sites in different genetic backgrounds; RAPD fragments that contain base pair mismatches near the 5′ end of the primer site may sometimes amplify in one genetic background but not in another, thereby producing artifactual polymorphisms (Lamboy 1994). AFLP primers are always fully complementary except at the bases at the 3′ end where correct primer base pairing is more critical (Williams et al. 1990). Therefore, genetic background is less likely to result in artifactual polymorphisms.

Another application of the AFLP technique resides in its ability to detect rare monomorphic bands across a large number of divergent genotypes. The large numbers of fragments that are generated in a single reaction allow efficient searches for loci that are monomorphic between distantly related taxa. Of the 320 AFLP loci scored, 5 (1.6%) were monomorphic across all six *Lactuca* species. Some of these fragments may be located in highly conserved regions of the genome, making them interesting templates for sequence analysis. Sequence

analysis of multiple nuclear fragments from potentially conserved regions of the genome would be a powerful tool for establishing phenetic trees above the species level as well as characterizing patterns of divergence and phylum-specific genetic clocks.

Acknowledgements We thank Dean Lavelle and Oswaldo Ochoa for technical assistance. We gratefully acknowledge the financial support of USDA NRICGP 94-37300-0284 and the University of California Systemwide Biotechnology Research and Education Program.

References

Akopyanz N, Bukanov NO, Westblom TU, Berg DE (1992) PCR-based RFLP analysis of DNA sequence diversity in the gastric pathogen *Helicobacter pylori*. Nucleic Acids Res 20:6221–6225

Bonierbale MW, Plaisted RL, Tanksley SD (1988) RFLP maps based on a common set of clones reveal modes of chromosome evolution in potato and tomato. Genetics 120:1095–1103

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

Chalmers KJ, Waugh R, Sprent JI, Simons AJ, Powell W (1992) Detection of genetic variation between and within populations of *Gliricidia sepium and G. maculata* using RAPD markers. Genetics 69:465–472

Cox TS, Yiang YT, Gorman MB, Rogers DM (1985) Relationship between coefficient of parentage and genetic similarity indices in soybean. Crop Sci 25:529–532.

Cox TS, Murphy JP, Rodgers DM (1986) Changes in genetic diversity in the red winter wheat regions of the United States. Proc Natl Acad Sci USA, 83:5583–5586

Demeke T, Adams RP, Chibbar R (1992) Potential taxonomic use of random amplified polymorphic DNA (RAPD): a case study in *Brassica*. Theor Appl Genet 84:990–994

dos Santos JB, Nienhuis J, Skroch P, Tivang J, Slocum MK (1994) Comparison of RAPD and RFLP genetic markers in determining genetic similarity among *Brassica oleracea* L. genotype. Theor Appl Genet 87:909–915

Farrara BF, Ilott TW, Michelmore RW (1987) Genetic analysis of factors for resistance to downy mildew (*Bremia lactucae*) in species of lettuce (*Lactuca sativa* and *L. serriola*). Plant Pathol 36:499–514

Fukuoka S, Hosoka K, Kamijima O (1992) Use of random amplified polymorphic DNAs (RAPDs) for identification of rice accessions. Jpn J Genet 67:243–252

Gower JC (1966) Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53:325–338

Heun M, Murphy JP, Phillips TD (1994) A comparison of RAPD and isozyme analyses for determining the genetic relationships among *Avena sterilis* L. accessions. Theor Appl Genet 87:689–696

Hulbert SH, Michelmore RW (1985) Linkage analysis of genes for resistance to downy mildew (*Bremia lactucae*) in lettuce (*Lactuca sativa*). Theor Appl Genet 70:520–528

Kazan K, Manners JM, Cameron DF (1993) Genetic relationships and variation in the *Stylosanthes guianensis* species complex assessed by random amplified polymorphic DNA. Genome 36:43–49

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

Kesseli R, Ochoa O, Michelmore R (1991) Variation at RFLP loci in *Lactuca* spp. and origin of cultivated lettuce (*Lactuca sativa*). Genome 34:430–436

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

Koller B, Lehmann A, Gessler C (1993) Identification of apple cultivars using RAPD markers. Theor Appl Genet 85:901–904

Kresovich S, Williams JGK, McFerson JR, Routman EJ, Schaal BA (1992) Characterization of genetic identities and relationships of *Brassica oleracea* L. via a random amplified polymorphic DNA assay. Theor Appl Genet 85:190–196

Lamboy WF (1994) Computing genetic similarity coefficients from RAPD data: the effects of PCR artifacts. PCR Meth Applic 4:31–37

Landry BS, Kesseli R, Leung H, Michelmore RW (1987) Comparison of restriction endonucleases and sources of probes for their efficiency in detecting restriction fragment length polymorphisms in lettuce (*Lactuca sativa* L.). Theor Appl Genet 74:646–653

Linquist K (1960) On the origin of cultivated lettuce. Hereditas 46:19–350

Martin JM, Blake TK, Hockett EA (1991) Diversity among North American spring barley cultivars based on coefficients of parentage. Crop Sci 31:1131–1137

Melchinger AE, Messmer MM, Lee M, Woodman WL, Lamkey KR (1991) Diversity and relationship among US maize inbreds revealed by restriction fragment length polymorphisms. Crop Sci 31:669–678

Michelmore RW, Anderson P, Okubara P, Witsenboer H (1994) Clusters of resistance genes in lettuce: map-based cloning in non-intensively studied species, In: Coruzzi G, Puigdomenech P (eds) NATO ASI series, vol H 81: plant molecular biology. Springer, Berlin, Hiedelberg New York, pp 501–509

Nei M, Li WH (1979) Mathematical models for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci, USA 76:5269–5273

Robinson RW, McCreight JD, Ryder EJ (1976) The genes of lettuce and closely related species. Plant Breed Rev 1:267–293

Rohlf FJ (1992) NTSYS-pc. Numerical taxonomy and multivariate analysis system. Version 1.70. Exeter Software, Setauket, N.Y.

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning, a laboratory manual, 2nd edn. Cold Spring Harbour Press, Cold Spring Harbour, N.Y.

Skroch P, Tivang J, Nienhuis J (1993) Analysis of genetic relationships using RAPD marker data. Bio/Technology 11:26–30

Smith OS, Smith JSC, Bowen SL, Tenborg RA, Wall SJ (1990) Similarities among a groups of elite maize inbreds as measured by pedigree F_1 grain yield, heterosis, and RFLPs. Theor Appl Genet 80:833–840

Song K, Osborn TC, Williams PH (1990) *Brassica* taxonomy based on nuclear restriction fragment length polymorphisms (RFLPs) 3. Genome relationships in *Brassica* and related genera and the origin of *B. oleracea* and *B. rapa* (syn. *campestris*). Theor Appl Genet 79:497–506

Tanksley SD (1989) RFLP mapping in plant breeding: new tools for an old science. Biotechnology 7:257–264

Tanksley SD, Bernatzky R, Lapitan NL, Prince JP (1988) Conservation of gene repertoire but not gene order in pepper and tomato. Proc Natl Acad Sci USA 85:6419–6423

Tinker NA, Fortin MG, Mather DE (1993) Random polymorphic DNA and pedigree relationships in spring barley. Theor Appl Genet 85:976–984

Vierling RA, Nguyen HT (1992) Use of RAPD markers to determine the genetic diversity of diploid, wheat genotypes. Theor Appl Genet 84:835–838

Vogel JM, Powell W, Rafalski A, Morgante M, Tudo JD, Taramino G, Biddle P, Hanafey M, Tingley SV (1994) Application of genetic diagnosis to plant genome analysis: comparison of marker systems. Appl Biotechnol Tree Cult 1:119–124

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

Welsh J, McClelland M (1990) Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res 18:7213–7218

Williams CE, StClair DA (1993) Phenetic relationships and levels of variability detected by restriction fragment length polymorphism and random amplified polymorphic DNA analysis of cultivated and wild accessions of *Lycopersicon esculentum*. Genome 36:619–630

Williams JGK, 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

Witsenboer H, Kesseli RV, Fortin MG, Stanghellini M, Michelmore RW (1995) Sources and genetic structure of a cluster of genes for resistance to three pathogens in lettuce. Theor Appl Genet 91: 1787–188

Yang X, Quiros C (1993) Identification and classification of celery cultivars with RAPD markers. Theor Appl Genet 86: 205–212

Zabeau M, Vos P (1993) Selective restriction fragment amplification: a general method for DNA fingerprinting. 924026297 (Publ number 0 534 858 Al)