Theor Appl Genet (1992) 83:528–532

THEORETICAL AND APPLIED GENETICS



# Restriction fragment length polymorphism and allozyme linkage map of *Cuphea lanceolata*

**D.M. Webb *, S.J. Knapp **, and L.A. Tagliani**

Department of Crop and Soil Science, Oregon State University, Corvallis, OR 97331, USA

Received May 15, 1991; Accepted May 29, 1991
Communicated by F. Salamini

**Summary.** *Cuphea lanceolata* Ait. has had a significant role in the domestication of *Cuphea* and is a useful experimental organism for investigating how medium-chain lipids are synthesized in developing seeds. To expand the genetics of this species, a linkage map of the *C. lanceolata* genome was constructed using five allozyme and 32 restriction-fragment-length-polymorphism (RFLP) marker loci. These loci were assigned to six linkage groups that correspond to the six chromosomes of this species. Map length is 288 cM. Levels of polymorphism were estimated for three inbred lines of *C. lanceolata* and an inbred line of *C. viscosissima* using 84 random genomic clones and two restriction enzymes, *Eco*RI and *Hin*dIII. Of the probes 29% detected RFLPs between *C. lanceolata* and *C. viscosissima* lines. Crosses between these species can be exploited to expand the map.

**Key words:** Allozymes – Restriction fragment length polymorphisms – Linkage map – *Cuphea*

## Introduction

The constituents of the seed storage lipids of many *Cuphea* species are medium-chain fatty acids (MCFAs) (Thompson 1984; Hirsinger 1985). The effort to domesticate *Cuphea* uses *Cuphea lanceolata* Ait. and *Cuphea viscosissima* Jacq. The oils of wild populations of *C. lanceolata* and *C. viscosissima* average 83.2% and 70.1% capric acid (C10:0), respectively (Graham 1989a; Knapp *Cuphea* 1991). In addition, *C. viscosissima* lines have been developed with as much as 28.4% caprylic acid (C8:0), 14.0% lauric acid (C12:0), and 29.4% myristic acid (C14:0) (Knapp and Tagliani 1991).

Caprylic and capric acid-rich medium-chain triglycerides (MCTs) have been used to treat disorders of lipid metabolism, epilepsy, gallstones, and other diseases, and are a source of rapidly absorbed energy for critically ill patients (Bach and Babayan 1982; Babayan 1987). In addition, caprylic and capric acids may have nutritional benefits not yet exploited for the general public (Babayan 1987). Presently, the primary commercial application of MCTs is the use of lauric and myristic acids in the manufacture of soaps and detergents (Knaut and Richtler 1985).

Coconut (*Cocos nucifera* L.) and palm kernel (*Elaeis guineensis* Jacq.) oils are the primary commercial sources of MCTs (Arkcoll 1988). However, every medium-chain fatty acid (MCFA) except lauric acid is more concentrated in seed oils of *C. lanceolata* and *C. viscosissima* lines than in oils of coconut and palm kernel. These higher concentrations of individual fatty acids would reduce the processing needed for their purification. In addition, because *Cuphea* is adapted to temperate climates, the production sources of MCTs would be diversified and help ensure longterm supplies and stable prices.

*C. lanceolata* and *C. viscosissima* are closely related (Graham 1988), and certain interspecific populations of these species are fertile (unpublished data). This greatly increases the genetic variation that can be exploited to breed *Cuphea*. In addition, these species have characteristics that make them useful as experimental organisms. They are diploids with $x=6$ chromosomes (Graham 1989b), have genomes approximately twice the size of *Arabidopsis thaliana* (unpublished data), are annual dicots, are amenable to transformation using *Agrobacterium tumefaciens*, and are regenerable by tissue culture (unpublished data).

* Current address: Pioneer Hi-Bred International, P.O. Box 1004, 7250 NW 62nd Ave., Johnston, IA 50131, USA
** To whom correspondence should be addressed

Highly autogamous species, e.g., tomato (*Lycopersicon esculentum* Mill.), tend to have few restriction fragment length polymorphisms (RFLPs) compared to highly allogamous species, e.g., maize (*Zea mays* L.) (Helentjaris et al. 1985). This has been observed in *Cuphea* as well. Estimated outcrossing rates for *C. viscosissima* (unpublished data) and *C. lanceolata* (Knapp et al. 1991; unpublished data) are 7–14% and 61–94%, respectively. Allozyme polymorphisms are abundant among *C. lanceolata* accessions (Knapp and Tagliani 1989), but rare among *C. viscosissima* accessions (unpublished data). Based on this, we presumed that restriction fragment length polymorphisms (RFLPs) were more prevalent within *C. lanceolata* than within *C. viscosissima*. Consequently, we selected *C. lanceolata* for constructing the initial linkage map. In this paper, we describe an RFLP and allozyme linkage map of *C. lanceolata* and RFLP levels between *C. lanceolata* and *C. viscosissima* inbred lines.

## Materials and methods

### *C. lanceolata plant materials*

We inbred 50 *C. lanceolata* lines to the $S_5$ generation by single-seed descent in 1987 and 1988. These lines were randomly derived from 24 accessions. LN43/1 and LN68/1 were selected as parent lines from this collection because of phenotypic differences in capric acid and seed oil content. RFLPs between these lines could not be considered because surveys had not been done. The parent lines were crossed and the $F_1$ was selfed pollinated. A population of 140 $F_2$ plants were grown in the greenhouse. These plants were propagated from stem-cuttings of field plants. Cuttings were dipped in 0.1% indole-3-butyric acid and 0.05% 1-naphthaleneacetic acid and placed in potting soil. They were covered with plastic for 7 days and transplanted to larger pots after 14 days.

### *Genomic DNA library*

The RFLP probes used in this study came from a genomic DNA library of *C. lanceolata*. DNA was extracted from leaf tissue (Webb and Knapp 1990), digested with *Pst*I, and size-fractionated on a 0.8% agarose gel. DNA fragments between 0.5 and 2 kb were electroeluted from the gel and ligated into the plasmid, pTZ18R, which was then transformed into the *E. coli* strain, DH5α. The library was plated on agar containing ampicillin and X-gal. Colonies having recombinant plasmids were isolated. We extracted the plasmids from the bacteria using the alkaline method of Birnboim and Doly (1979) followed by a phenol/chloroform and chloroform wash (Maniatis et al. 1982).

### *DNA extraction, digestion, and blotting*

Leaf tissue was harvested from 4- to 8-week-old greenhouse-grown $F_2$ plants. The tissue was lyophilized, ground to a powder in a coffee grinder, and stored at −60 °C. Total genomic DNA was extracted as described in Webb and Knapp (1990). Approximately 11 µg of DNA was digested with 50 units of *Eco*RI or *Hin*dIII for approximately 3 h at 37 °C. The digested DNA was size fractionated on a 0.8% agarose gel at 0.5–1.4 V/cm for 15–18 h. The running buffer was 0.4 *M* TRIS base and 20 m*M* EDTA with the pH adjusted to 8.0 with glacial acetic acid. The DNA was denatured in 0.4 *M* NaOH and transferred in the same solution to charged nylon membranes (CUNO Zetabind) by vacuum blotting. The membranes were neutralized in 6 × SSC for 10 min, baked at 90 °C for 3 h, and washed in 0.1 × SSC and 0.5% SDS at 65 °C for 30 min.

### *Hybridizations and autoradiography*

Membranes were prehybridized overnight at 65 °C in 6 × SSC, 0.5% SDS, 5 × Denhardt's solution, and 100 µg/ml denatured salmon sperm DNA (Maniatis et al. 1982). Whole plasmids or inserts from the genomic library were labeled by random primer extension (Feinberg and Vogelstein 1983) with [$\alpha^{32}$]-dCTP. Unincorporated nucleotides were removed from the labeling mixture using Sephadex G-50 spin columns. The labeled DNA was denatured at 95 °C and added to 5–8 ml of prehybridization solution and one to three membranes. The membranes were hybridized overnight at 65 °C with gentle shaking, then washed from one to four times for 30 min at 65 °C in 2 × SSC and 0.1% SDS with vigorous shaking.

X-ray film (Kodak XAR-5) was exposed to the hybridized membranes with one intensifying screen at −60 °C for 2 h to 3 days. Radioactivity was removed from the membranes by gentle shaking in 0.4 *M* NaOH for 20 min at 65 °C. The membranes were reconditioned for reuse by gentle shaking in 0.1 × SSC, 0.5% SDS, and 0.2 *M* TRIS pH 8.0 for 20 min at 65 °C.

### *RFLP probes*

The first 35 of 119 clones isolated from the genomic library were screended for high-copy-number sequences. Plasmids from these clones were dot blotted (Maniatis et al. 1982) with 500 ng DNA per clone and probed with 200 ng labeled total *C. lanceolata* DNA digested with *Pst*I. Clones that had detectable hybridizations were considered to have high-copy-number sequences and were not used. High-copy-number probes from the nuclear genome were expected to produce hybridization bands too numerous to resolve.

The *C. lanceolata* inbred lines LN61/1, LN43/1, and LN68/1 and the *C. viscosissima* line VS55/1 were screened for RFLPs when digested with *Eco*RI or *Hin*dIII restriction enzymes. Probes polymorphic between LN43/1 and LN68/1 were subsequently hybridized to digested DNA from the $F_2$ population.

### *Allozyme assays*

Methods for *C. lanceolata* allozyme assays have been described (Knapp and Tagliani 1989). Allozyme extracts from the inbred lines and $F_2$ population in this experiment were from cotyledons and young leaves of 3- to 4-week-old seedlings, respectively. Allozyme variation between LN43/1 and LN68/1 was assayed for aconitase (ACO), diaphorase (DIA), esterase (EST), fluorescent esterase transaminase (FES), glutamine oxaloacetate transaminase (GOT), menadione reductase (MNR), phosphoglucomutase (PGM), phosphoglucose isomerase (PGI), and shikimate dehydrogenase (SKD) activity.

### *Segregation and linkage analyses*

Log-likelihood ratios (G-statistics) were used to test hypotheses about the segregation and linkage of markers using GMENDEL (Liu and Knapp 1991). Multipoint linkage analysis and map distances (Kosambi 1944) were determined using MAPMAKER II (Lander et al. 1987). Loci were considered linked in 2-point and 3-point analyses when LOD scores exceeded 4.0 and recombination frequencies were less than 0.30. Linkage group map orders were determined by multipoint analyses

with minimum LOD scores from 1.0 to 4.0. When more than one map sequence was possible, the sequence having the highest LOD score was used.

## Results and discussion

### *Probe copy number*

Two of the 35 clones screened on dot blots had detectable hybridizations. In addition, the 117 clones used to probe the inbred lines hybridized at one to ten bands using a low stringency wash (2 × SSC). The high frequency of low copy DNA probes was probably due to the small size of the *C. lanceolata* genome and to the nature of the genomic library that was constructed using the methylation-sensitive restriction enzyme, *Pst*I, to avoid repetitive sequences.

### *Allozyme polymorphisms*

Allozyme polymorphisms between LN43/1 and LN68/1 were observed for *Fes-1*, *Fes-2*, *Mnr-1*, *Pgm-1*, *Pgm-2*, *Aco-1*, *Aco-3*, *Est-1*, and *Est-2*. *Aco-1* and *Aco-3* activities were not observed in $F_2$ leaf extracts, and the alleles for *Est-1* and *Est-2* could not be reliably scored; thus, these loci were not mapped. Loci for the DIA, GOT, and PGI enzymes and *Mnr-2*, *Aco-2*, and *Est-3* were monomorphic.

### *Restriction fragment length polymorphisms*

RFLPs between *C. lanceolata* lines were observed for 14.3–20.2% of the clones per restriction enzyme and 29.0% of the clones for both enzymes (Table 1). Polymorphisms between *C. viscosissima* and *C. lanceolata*, however, were approximately twice as great as those within *C. lanceolata* (Table 1).

Greater frequencies of RFLPs between species compared to within species have been observed for tomato (Helentjaris et al. 1985), *Brassica* spp. (Figdore et al. 1988), lentil (Havey and Muehlbauer 1989), and soybean (Keim et al. 1989). Because of this, interspecific populations have been widely used to construct linkage maps (Bernatzky and Tanksley 1986; Havey and Muehlbauer 1989; Keim et al. 1990b) and to map QTL (Osborn et al. 1987; Nienhuis et al. 1987; Keim et al. 1990a, b). Interspecific mapping is useful for *Cuphea* as well. Of the 53 clones showing polymorphisms between *C. lanceolata* and *C. viscosissima* lines, 26 were polymorphic for this cross and were mapped (Fig. 1). The remaining clones can be mapped using interspecific populations, thereby increasing the density and utility of the map.

Digests with *Eco*RI and *Hin*dIII restriction enzymes resulted in roughly the same polymorphism rates (Table 1). Of the polymorphisms 83.0% occurred with one but

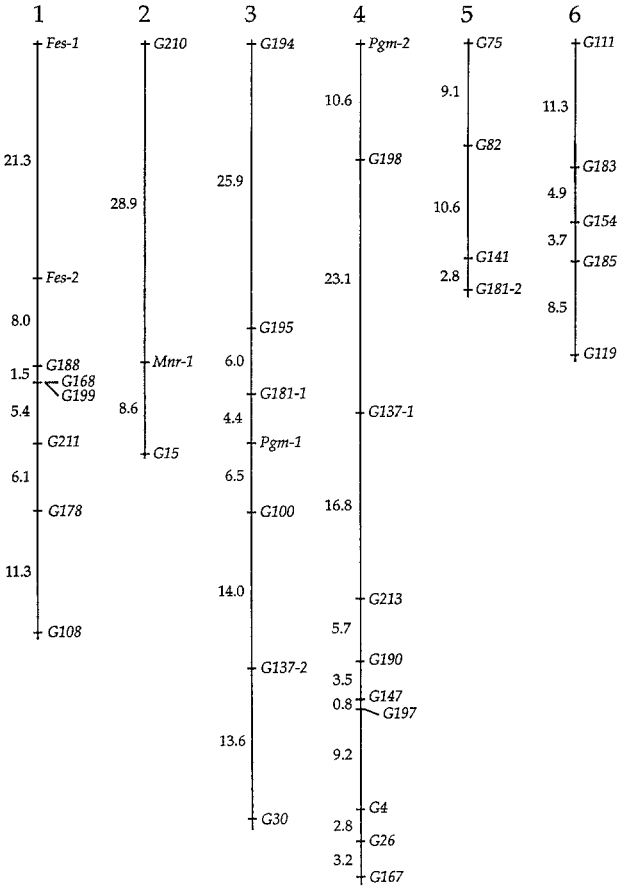


**Fig. 1.** A restriction fragment length polymorphism and allozyme linkage map of *Cuphea lanceolata*. The chromosomes are represented by *vertical lines*. The relative positions of marker loci are shown by *crossbars* with names on the right and Kosambi (1944) map distances (cM) between markers on the left

**Table 1.** Percentage of restriction fragment length polymorphism observed among three *Cuphea lanceolata* (LN) inbred lines and between those *C. lanceolata* lines and one *Cuphea viscosissima* (VS) inbred line using 84 genomic DNA clones as probes against genomic DNA digested with EcoRI or HindIII

| Comparison | Restriction enzymes | | |
|---|---|---|---|
| | EcoRI | HindIII | EcoRI or HindIII |
| Between LN lines | | | |
| LN43/1 and LN68/1 | 14.3 | 17.9 | 31.0 |
| LN61/1 and LN68/1 | 17.9 | 17.9 | 31.0 |
| LN43/1 and LN61/1 | 15.5 | 20.2 | 25.0 |
| Mean | 15.9 | 18.7 | 29.0 |
| Between LN and VS lines | | | |
| LN43/1 and VS55/1 | 36.9 | 38.1 | 61.9 |
| LN68/1 and VS55/1 | 38.1 | 35.7 | 63.1 |
| LN61/1 and VS55/1 | 38.1 | 36.9 | 63.1 |
| Mean | 37.7 | 36.9 | 62.7 |

not the other enzyme. This indicates gains or losses in enzyme recognition sequences due to base pair changes. Insertions or deletions are more likely to cause polymorphisms with multiple enzymes. *Eco*RI and *Hind*III were used because they were inexpensive and digested *Cuphea* DNA well. Other enzymes would undoubtedly yield additional polymorphisms to increase the percentage of clones that can be used as markers.

### *Marker segregation*

The observed segregation ratios for the markers with null alleles, *Fes-2*, G181-1, and G181-2 were not significantly different from the expected (3:1). The observed segregation ratios for the codominant markers, except *Fes-1* and G141, were not significantly different from the expected (1:2:1). The distortion for *Fes-1* (27:95:45, $P=0.02$) was due primarily to lower than expected LN43/1 homozygous genotypes and higher than expected heterozygous genotypes. The distortion for G141 (30:64:39, $P=0.05$) was due primarily to higher than expected LN68/1 homozygous genotypes. Distortion at 2 of 37 markers is expected from random deviation. No pattern of genotypic selection was observed.

### *Linkage map*

The map we made of the *C. lanceolata* genome has five allozyme and 32 RFLP markers in six linkage groups (Fig. 1). The six linkage groups correspond to the six chromosomes of *C. lanceolata*. Every markers was assigned to a linkage group. If large segments of the genome were excluded from this map, we would expect to have unlinked markers or more than six linkage groups. Although this map cover a significant fraction of the genome, we expect it to increases as more loci are added.

The size of the map is 288 c*M*. The *C. lanceolata* genome is approximately twice the physical size of the *Arabidopsis* genome, yet the total map distance is significantly less than the approximately 500 c*M* *Arabidopsis* maps (Chang et al. 1988; Nam et al. 1989). Several factors may account for this. First, we expect the size of the map to increase as additional markers map outside the present linkage groups. Second, as the map becomes more dense, we expect map distances between distantly linked markers to increase. Third, the rate of chiasmata formation of *Cuphea* and *Arabidopsis* may not be similar.

The *Cuphea* linkage map will be expanded using *C. lanceolata* × *C. viscosissima* populations. The interspecific map should have great utility since interspecific crosses are extensively used to breed *Cuphea*. Interspecific populations between certain lines of these species are fully fertile and cytogenetically normal. As a consequence, they behave and are bred like intraspecific populations.

*Acknowledgements*. This research was partially supported by grants from the USDA and the Soap and Detergent Association. Oregon Agric. Exp. Stn. Tech. Paper No. 9469

## References

Arkcoll D (1988) Lauric oil resources. Econ Bot 42:195–205

Babayan VK (1987) Medium-chain triglycerides and structured lipids. Lipids 22:417–420

Bach AC, Babayan VK (1982) Medium-chain triglycerides: an update. Am J Clin Nutr 36:950–962

Bernatzky R, Tanksley SD (1986) Toward a saturated linkage map in tomato based on isozymes and random cDNA sequences. Genetics 112:887–898

Birnboim HC, Doly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7:1513–1523

Chang C, Bowman JL, DeJohn AW, Lander ES, Meyerowitz EM (1988) Restriction fragment length polymorphism linkage map for *Arabidopsis thaliana*. Proc Natl Acad Sci USA 85:6856–6860

Feinberg AP, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Figdore SS, Kennard WC, Song KM, Slocum MK, Osborn TC (1988) Assessment of the degree of restriction fragment length polymorphism in *Brassica*. Theor Appl Genet 75:833–840

Graham SA (1988) Revision of *Cuphea* section Heterodon (Lythraceae). System Bot Monogr, vol 20. Am Soc Plant Taxon

Graham SA (1989a) *Cuphea*: A new plant source of medium-chain fatty acids. In: Critical Reviews Food Science Nutrition vol 28. CRC Press, Boca Raton, Fla, pp 139–173

Graham SA (1989b) Chromosome numbers in *Cuphea* (Lythraceae): new counts and a summary. Am J Bot 76:1530–1540

Havey MJ, Muehlbauer FJ (1989) Linkages between restriction fragment length, isozyme, and morphological markers in lentil. Theor Appl Genet 77:395–401

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 applied plant breeding. Plant Mol Biol 5:109–118

Hirsinger F (1985) Agronomic potential and seed composition of *Cuphea*, an annual crop for lauric and capric seed oils. JAOCS 62:76–80

Keim P, Shoemaker RC, Palmer RG (1989) Restriction fragment length polymorphism diversity in soybean. Theor Appl Genet 77:786–792

Keim P, Diers B, Shoemaker RC (1990a) Genetic analysis of soybean hard seededness with molecular markers. Theor Appl Genet 79:465–469

Keim P, Diers B, Olson TC, Shoemaker RC (1990b) RFLP mapping in soybean: Association between marker loci and variation in quantitative traits. Genetics 126:735–742

Knapp SJ, Tagliani LA (1989) Genetics of allozyme variation in *Cuphea lanceolata* Ait. Genome 32:57–63

Knapp SJ, Tagliani LA (1991) Mutations of *Cuphea viscisissima* affecting the synthesis of medium-chain fatty acids. Plant Breed 106:338–341

Knapp SJ, Tagliani LA, Liu BH (1991) Outcrossing rates of experimental populations of *Cuphea lanceolata*. Plant Breed 106:334–337

Knapp SJ, Tagliani LA, Roath WN (1991) Fatty acid and oil diversity of *Cuphea viscosissima*: a source of medium-chain fatty acids. JAOCS 68:515–517

Knaut J, Richtler HJ (1985) Trends in industrial uses of palm and lauric oils. JAOCS 62:317–327

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12:172–175

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newberg L (1987) MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Liu B, Knapp SJ (1991) A program for Mendelian segregation and linkage analysis of individual or multiple progeny populations using log-likelihood ratios. J Hered 81:407

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Nam H, Giraudat J, den Boer B, Moonan F, Loos WDB, Hauge BM, Goodman HM (1989) Restriction fragment length polymorphism linkage map of *Arabidopsis thalliana*. Plant Cell 1:699–705

Nienhuis J, Helentjaris T, Slocum M, Ruggero B, Schaefer A (1987) Restriction fragment length polymorphism analysis of loci associated with insect resistance in tomato. Crop Sci 27:797–803

Osborn TC, Alexander DC, Fobes JF (1987) Identification of restriction fragment length polymorphisms linked to genes controlling soluble solids content in tomato fruit. Theor Appl Genet 73:350–356

Thompson AE (1984) *Cuphea* – a potential new crop. Hort Sci 19:352–354

Webb DM, Knapp SJ (1990) DNA extraction from a previously recalcitrant plant genus. Plant Mol Biol Rep 8:180–185