AN ABSTRACT OF THE THESIS OF

David M. Webb for the degree of Doctor of Philosophy in Crop
Science presented on November 13, 1990.

Title: Genetic Mapping of *Cuphea lanceolata*: Molecular-Marker
           Linkage to Quantitative-Trait Loci Affecting Seed Capric
           Acid, Seed Oil, and Embryo Development

Abstract approved: Redacted for Privacy

Steven J. Knapp

Cuphea is an herbaceous genus having species whose seed
storage lipids are predominantly medium-chain triglycerides (MCTs).
Cuphea lanceolata Ait. and *Cuphea viscosissima* Jacq. are central to
the breeding of *Cuphea* as a new commercial source of MCTs. The
objectives of this thesis were to develop a reliable method to
extract DNA from *Cuphea*, to construct a molecular-marker linkage
map of *C. lanceolata*, and to evaluate quantitative-trait loci (QTL)
affecting seed capric acid content, seed oil content, and embryo
development in *C. lanceolata*. We used allozyme and restriction-
fragment-length-polymorphism (RFLP) markers to study *Cuphea* because
molecular markers are the best method available to investigate the
Mendelian genetics underlying quantitative traits, and are useful
in breeding. Adequate yield of clean DNA is essential for RFLP
mapping. We tried numerous DNA extraction methods that failed to
remove contaminants that interfere with restriction digests of

Cuphea DNA. The method described here was developed to remove those contaminants and maintain relatively high DNA yields. In this method, the critical step in purification consists of washing the DNA with phenol while it is complexed with CTAB and dissolved in 1 M NaCl. An RFLP and allozyme linkage map of *C. lanceolata* was constructed having 37 markers in six linkage groups with a total distance of 288 cM. Levels of polymorphism were estimated for three lines of *C. lanceolata* and one line of *C. viscosissima* using 84 random genomic clones and two restriction enzymes, *EcoRI* and *HindIII*. Twenty-nine percent of the probes detected RFLPs between *C. lanceolata* lines, whereas 63% of the probes detected RFLPs between *C. lanceolata* and *C. viscosissima* lines. Thirty RFLP and four allozyme markers were used to locate on the *C. lanceolata* linkage map QTL affecting seed capric acid content, seed oil content, and embryo development. Three unlinked QTL explained 19.4% of the phenotypic variation in capric acid content in F_2 seed. Seed oil content and seed weight were measured on seed from field-grown F_2 plants. Seed weight was indicative of embryo size (development). Four unlinked QTL explained 33.9% of the phenotypic variation in embryo size. One of these QTL, which explained 20.3% of the variation, may have been a chromosomal deletion detected by a marker having a null allele. Three QTL pleiotropically affected seed oil by affecting embryo size. The one QTL that only affected seed oil accounted for 2.8% of the phenotypic variation.

Genetic Mapping of *Cuphea lanceolata*:
Molecular-Marker Linkage to Quantitative-Trait Loci
Affecting Seed Capric Acid, Seed Oil, and Embryo Development

by

David M. Webb

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Completed November 13, 1990

Commencement June 1991

APPROVED:

Redacted for Privacy

Professor of Crop and Soil Science in charge of major

Redacted for Privacy

Head of Department of Crop and Soil Science

Redacted for Privacy

Dean of Graduate School   ∨

Date thesis is presented _____ November 13, 1990

Presented by _____ David M. Webb

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GENETIC MAPPING OF *CUPHEA LANCEOLATA*:
MOLECULAR-MARKER LINKAGE TO QUANTITATIVE-TRAIT LOCI
AFFECTING SEED CAPRIC ACID, SEED OIL, AND EMBRYO DEVELOPMENT

INTRODUCTION

Use of genetic markers to detect and map genes affecting
quantitative traits has been theorized at least since 1923 when the
segregation of several pigmentation genes was associated with seed
size in bean (Sax 1923). Extensive linkage maps based on morpho-
logical markers have been developed for some important crops, e.g.
tomato (Rick and Yoder 1988). However, morphological markers have
had limited use in plant breeding because of their often dominant
or deleterious effects and their relative scarcity. For example,
MacArthur (1934) made 158 crosses and evaluated 48,000 segregants
in 10 years to map just 21 markers in tomato.

Allozyme markers became widely used in plant breeding (Peirce
and Brewbaker 1973) after esterase isomorphs from mouse tissue were
first observed by starch gel electrophoresis (Hunter and Markert
1957). Allozymes are generally codominantly expressed, not delete-
rious, more abundant than morphological markers in most germplasm,
and relatively easy to assay. Allozymes have been used to identify
quantitative-trait loci (QTL) in tomato (Tanksley et al. 1982) and
maize (Stuber et al. 1987) using 12 and 20 markers, respectively.
Allozyme markers improved the applicability of marker-based linkage
analysis but usable markers were still fewer than needed for
complete genome coverage.

Recombinant-DNA technologies now enable the use of DNA re-
striction-fragment-length polymorphisms (RFLPs) as genetic markers.

RFLP markers are genetically neutral (are not genes per se), usually codominant, and almost unlimited in number. Genetic linkage maps based on RFLPs can be made more dense and thus more useful to breeding and genetics.

Dense RFLP linkage maps are potentially important in plant breeding for varietal identification, detection of quantitative-trait loci (QTL), and marker-facilitated backcross breeding to minimize linkage drag during introgression (Soller and Beckmann 1983; Tanksley et al. 1989). Relatively dense RFLP linkage maps have been constructed for tomato (Bernatzky and Tanksley 1986; Helentjaris et al. 1986), maize (Helentjaris 1987), lettuce (Landry et al. 1987), *Arabidopsis thaliana* (Chang et al. 1988; Nam et al. 1989), pepper (Tanksley et al. 1988), rice (McCouch et al. 1988), potato (Bonierbale et al. 1988; Gebhardt et al. 1989), lentil (Havey and Muehlbauer 1989), *Brassica oleracea* (Slocum et al. 1990), and soybean (Keim et al. 1990). Some of these maps have already been used to detect QTL affecting agronomic traits (Nienhuis et al. 1987; Osborn et al. 1987; Tanksley and Hewitt 1988; Keim et al. 1990).

A molecular-marker linkage map could also be used for the breeding and genetics of *Cuphea*, a genus being investigated as a new commercial source of medium-chain triglycerides (MCTs) (Thompson 1984; Hirsinger 1985). Oils with high concentrations of lauric (C12:0) and myristic (C14:0) acids are used in foods, soaps, detergents, cosmetics, and other products (Knaut and Richtler 1985). MCT formulations comprised of caprylic (C8:0) and capric (C10:0) acids are used to treat disorders of lipid metabolism and

other diseases, and as a source of rapidly absorbed, high-energy fuel for critically ill patients (Bach and Babayan 1982; Babayan 1987). MCTs have beneficial nutritional characteristics (Babayan 1987) and may be used as a nontoxic, antimicrobial preservative in food and other perishables (Kabara 1984).

Research to domesticate *Cuphea* is concentrating on *Cuphea lanceolata* Ait. and *Cuphea viscosissima* Jacq., whose seed oils average 83.5% and 75.5% capric acid (C10:0), respectively (Graham 1989a). In addition to capric acid, *C. viscosissima* is a source of other medium-chain fatty acids (MCFAs). *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) (unpublished data).

Coconut (*Cocos nucifera* L.) and palm kernel (*Elaeis guineensis* Jacq.) oils are currently the primary commercial sources of MCFAs (Arkcoll 1988). However, every MCFA except lauric acid is more concentrated in the seed oils of *C. lanceolata* and *C. viscosissima* lines than in oils of coconut and palm kernel. 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 for MCFAs could be diversified and help ensure longterm supplies at 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 ex-

perimental organisms for genetic studies. They are diploids with x
— 6 chromosomes (Graham 1989b); have relatively small genomes, only
about twice the size of *Arabidopsis thaliana* (unpublished data);
are annual dicots; are amenable to transformation using *Agrobac-
terium tumefaciens*; and are regenerable by tissue culture (un-
published data).

Highly self-pollinated species such as tomato (*Lycopersicon
esculentum* Mill.) tend to have few restriction-fragment-length
polymorphisms (RFLPs) compared with highly cross-pollinated species
such as maize (*Zea mays* L.) (Helentjaris et al. 1985). This has
been observed in *Cuphea* as well. Estimated outcrossing rates were
14% for *C. viscosissima* (unpublished data) and 61 to 94% for *C.
lanceolata* (Knapp et al. 1991). Allozyme polymorphisms are abun-
dant among *C. lanceolata* accessions (Knapp and Tagliani 1989), but
rare among *C. viscosissima* accessions (unpublished data). Based on
this, restriction-fragment-length polymorphisms (RFLPs) are expect-
ed to be more prevalent within *C. lanceolata* than within *C. visco-
sissima*, and *C. lanceolata* should be the better species for experi-
ments involving molecular markers. Consequently, we selected *C.
lanceolata* to make the initial linkage map.

This thesis includes three closely related manuscripts as
chapters and a fourth manuscript as an appendix. The first chapter
describes a protocol we developed for extracting high yields of
digestable DNA from *Cuphea* tissue. The second chapter presents a
C. lanceolata allozyme and RFLP linkage map and estimated restric-
tion-fragment-length polymorphism rates within *C. lanceolata* and
between *C. lanceolata* and *C. viscosissima*. The third chapter is

the identification of QTL affecting seed capric acid content, seed oil content, and embryo development in *C. lanceolata*. The appendix presents estimates of genetic parameters for oil yield and related traits using half-sib families derived from a population of *C. lanceolata*.

CHAPTER 1.

DNA EXTRACTION FROM *CUPHEA*

ABSTRACT

Numerous DNA extraction methods failed to remove contaminants that interfere with restriction digests of *Cuphea* DNA. The method described here removes those contaminants and maintains relatively high DNA yields. The critical purification process is to wash the DNA with phenol while it is complexed with CTAB and dissolved in 1 *M* NaCl.

INTRODUCTION

DNA extraction needs to be simple, rapid, and inexpensive when many samples must be used, such as in breeding and population studies. We tried numerous plant DNA-extraction methods (Dellaporta et al., 1983; Keim et al., 1988; Rogers and Bendich, 1985; Saghai-Maroof et al., 1984; and Wagner et al., 1987), including many modifications, on *Cuphea lanceolata* Ait. and *Cuphea viscosissima* Jacq., new seed-oil crops high in capric acid (Graham, 1989a). These methods failed because contaminants that interfere with restriction enzyme digestion of the DNA were not removed. We modified a method by Liechtenstein and Draper (1985), which is based on that of Murray and Thompson (1980), and obtained DNA from greenhouse-grown *C. lanceolata* and *C. viscosissima* plants suitable for complete restriction digestion and restriction-fragment-length-polymorphism (RFLP) analysis. The essential modifications are as follows: 1) The concentration of CTAB in the extraction and precipitation buffers is reduced from 1% to 0.5%, 2) tissue is mixed in the extraction buffer using a mechanical tissue homogenizer, such as a polytron, instead of a mortar and pestle, 3) the second CTAB/chloroform wash is deleted, 4) centrifugation speed to pellet the CTAB/DNA precipitant is increased from 1500 g to 4500 g, and 5) a phenol/chloroform wash is incorporated before dissolving the DNA in TE buffer.

MATERIALS AND METHODS

Solutions

Extraction Buffer: 0.5% (w/v) CTAB (hexadecyltrimethylammonium bromide), 50 mM Tris-HCl pH 8.0, 0.7 M NaCl, 10 mM EDTA (ethylenediaminetetraacetic acid, disodium salt) pH 8.0, 1% (v/v) 2- mercaptoethanol

Precipitation Buffer: 0.5% (w/v) CTAB, 50 mM Tris-HCl pH 8.0, 10 mM EDTA pH 8.0

Chloroform: chloroform/octanol 24:1 (v/v)

Phenol: equilibrated to pH > 7.8 with equal volume 1 M Tris-HCl pH 8.0, then equal volume 0.1 M Tris-HCl pH 8.0

TE: 10 mM Tris-HCl pH 8.0, 0.1 mM EDTA pH 8.0

Tissue preparation

Collect tissue on ice, lyophilize, and grind it to a fine powder using a coffee grinder. Store the tissue at -20°C.

Cell disruption

Add 900 mg ground tissue and 15 mL extraction buffer to a 50 mL disposable centrifuge tube. Mix the tissue and buffer with a polytron or other tissue homogenizer at a moderate speed for 45-60 sec. Pour the mixture into a 50 mL teflon-coated oak ridge tube. Rinse the first tube and polytron head with 9 mL extraction buffer and add this to the first mixture. Set the mixture at 60°C for 60 minutes. Mix the sample occasionally during this incubation.

Chloroform wash

Add 15 mL chloroform and mix by gently inverting the tube for 4 min or until a nearly complete emulsion forms. Spin at 14,500 g

for 10 min to separate the phases.

CTAB/DNA precipitation

Add 19 mL precipitation buffer to a 50 mL polysulfone oak ridge tube. Transfer the supernatant to the same tube using a 25 mL pipette. Mix the sample by several gentle inversions and set it at room temperature for 20 min. Spin at 4500 g for 10 min using a swinging bucket rotor. Pour off the supernatant and drain the tubes for several minutes on a paper towel.

CTAB/DNA dissolving

Add 400 μL 1 M NaCl to the pellet and set the sample at 45-65°C with gentle shaking for 10-30 min or at room temperature overnight without shaking. Dissolve as much material as possible. Pour the solution into a 2.0 mL microfuge tube. Rinse the 50 mL tube with 200 μL 1 M NaCl and add this to the same microfuge tube.

Phenol wash

Add 700 μL phenol to the sample and mix by gentle inversions. Spin at high speed for 4 min to separate phases. Transfer the aqueous (top) phase to a 2.0 mL tube. Do not discard the interface/phenol.

Back extraction

Add 300 μL TE to the interface/phenol and mix by gentle inversions. Spin at high speed for 2 min. Add this aqueous phase to the 2.0 mL tube with the first aqueous phase of that sample. Discard the interface and phenol.

Chloroform wash

Add 900 μL of chloroform to the sample and mix by gentle shaking. Spin at high speed for 2 min. Transfer the aqueous phase

to a 2.0 mL tube.

DNA precipitation

Add 950 μL cold isopropanol to the sample and mix by gentle inversions. Place the sample at -20°C for 30 min or overnight. Spin at high speed for 4 min. Pour off the supernatant and drain the tube on a paper towel.

Pellet washing

Wash the salt from the DNA by adding 1 mL of 65% (v/v) ethanol. Invert the tube several times and pour off the ethanol. Add 1 mL 85% (v/v) ethanol and pour off. Drain the tube on a paper towel, remove residual alcohol from the tube bottom using a micropipettor, then vacuum or air dry the sample. Add 200-300 uL TE to dissolve the DNA.

RESULTS AND DISCUSSION

Typical yields using this method range from 40 to 70 μg DNA per gram of fresh tissue. Sufficient quantities of DNA are obtained from 6 grams of fresh leaf tissue for large-scale RFLP analysis. DNA obtained by this method dissolves easily and is completely digestable with *Eco*RI, *Hind*III, *Bam*HI, and *Pst*I.

The polytron effectively disrupts cells without significant DNA shearing, possibly because the highly viscous contaminants cushion the DNA against shearing forces. DNA yields are four to six times higher when the tissue and buffer are polytroned rather than hand mixed.

The effectiveness of CTAB as a detergent and precipitant is not diminished by lowering its concentration from 1.0% to 0.5%. The benefits are that problematic foaming during mixing with a polytron is eliminated and the solution viscosity is reduced. Having less viscosity, the centrifuge speed and time required to pellet the precipitated CTAB/DNA is reduced by half that otherwise needed for *Cuphea*.

A moderately tight CTAB/DNA pellet is made to minimize DNA loss when the supernatant is removed and to minimize supernatant carry-over into the 1 M NaCl. Thorough drainage of the supernatant from the pellet improves the dissolving in 1 M NaCl and keeps the subsequent volumes low for using microfuge tubes.

We deleted the second CTAB and chloroform wash used in the methods by Murray and Thompson (1980) and Liechtenstein and Draper (1985) because it was unnecessary with *Cuphea*. We added a

phenol/chloroform wash to effectively remove most contaminants. Usually a phenol/chloroform wash occurs after the DNA is dissolved in TE. By incorporating the phenol/chloroform wash earlier, the DNA dissolves in TE much easier and substantial time is saved because a second alcohol precipitation and dissolving in TE is eliminated.

Phenol is most effective in removing contaminants from *Cuphea* tissue when not mixed with chloroform and when equilibrated once with 1.0 M Tris-HCl pH 8.0 then once with 0.1 M Tris-HCl pH 8.0. More extensive equilibration methods substantially reduce DNA yields for a reason we cannot explain. Back-extracting the interface/phenol with TE almost doubles the DNA yield.

Using this protocol, we extracted *Cuphea* DNA that is sufficiently clean for complete digestion with restriction enzymes. Other methods and their modifications yielded much less DNA or DNA that is highly contaminated and thus unusable. This protocol is relatively simple and inexpensive. Twenty-four samples can be processed in one normal workday. The DNA dissolves quickly and can be used the next day. This protocol may be useful with other plant species where DNA extraction is made difficult by contaminants.

REFERENCES CITED

Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA minipreparation: Version II. Pl. Mol. Biol. Rep. 1:19-21.

Graham SA (1989a) *Cuphea*: A new plant source of medium-chain fatty acids. CRC Crit. Rev. Food Sci. Nutr. 28:139-173.

Keim P, Olson TC, Shoemaker RC (1988) A rapid protocol for isolating soybean DNA. Soybean Genet. Newsl. 15:150-152.

Liechtenstein C, Draper J (1985) Genetic engineering of plants. p.67-120. In: Glover, D.M. (ed.) DNA cloning. Vol.2., IRL Press, Oxford.

Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight DNA. Nucleic Acids Res. 8:4321-4325.

Rogers SO, Bendich AJ (1985) Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. Pl. Mol. Biol. 5:69-76.

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

Wagner DB, Furnier GR, Saghai-Maroof MA, Williams SM, Dancik BP, Allard RW (1987) Chloroplast DNA polymorphisms in lodgepole and jack pines and their hybrids. Proc. Natl. Acad. Sci. USA 84:2097-2100.

CHAPTER 2.

RESTRICTION-FRAGMENT-LENGTH-POLYMORPHISM AND ALLOZYME
LINKAGE MAP OF *CUPHEA LANCEOLATA*

ABSTRACT

Cuphea lanceolata Ait. has a significant role in the domestication of *Cuphea* and is a useful organism for investigating how medium-chain lipids are synthesized in developing seeds. To expand the genetics of this species, we constructed a linkage map of the *C. lanceolata* genome using five allozyme and 32 restriction-fragment-length-polymorphism (RFLP) markers. These markers were assigned to six linkage groups which correspond to the six chromosomes of this species. The map length is 288 cM. We estimated levels of polymorphism among three inbred lines of *C. lanceolata* and one inbred line of *C. viscosissima* using 84 random genomic clones and two restriction enzymes, *Eco*RI and *Hind*III. Twenty-nine percent of the probes detected RFLPs between *C. lanceolata* lines, whereas 63% of the probes detected RFLPs between *C. lanceolata* and *C. viscosissima* lines. Interspecific crosses may be exploited to expand the map and further its utility in the domestication of *Cuphea*.

INTRODUCTION

The genus *Cuphea* is being investigated as a new commercial source of medium-chain triglycerides (MCTs) (Thompson 1984; Hirsinger 1985). The United States effort to domesticate *Cuphea* is concentrating on *Cuphea lanceolata* Ait. and *Cuphea viscosissima* Jacq., whose seed oils average 83.2% and 75.5% capric acid (C10:0), respectively (Graham 1989a). 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) (unpublished data).

Caprylic and capric acid 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* than 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 sources of MCTs could be diversified to help ensure longterm supplies at 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 for genetic studies. 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).

Highly self-pollinated species such as tomato (*Lycopersicon esculentum* Mill.) tend to have few restriction-fragment-length polymorphisms (RFLPs) compared with highly cross-pollinated species such as maize (*Zea mays* L.) (Helentjaris et al. 1985). This has been observed in *Cuphea* as well. Estimated outcrossing rates were 14% for *C. viscosissima* (unpublished data) and 61 to 94% for *C. lanceolata* (Knapp et al. 1991). Allozyme polymorphisms are abundant among *C. lanceolata* accessions (Knapp and Tagliani 1989), but rare among *C. viscosissima* accessions (unpublished data). Based on this, we expect restriction-fragment-length polymorphisms (RFLPs) to be more prevalent within *C. lanceolata* than within *C. viscosissima*. Consequently, we selected *C. lanceolata* for the construction of this linkage map.

In this paper, we describe an RFLP and allozyme linkage map of *C. lanceolata* and RFLP levels among *C. lanceolata* and *C. viscosis-*

sima inbred lines.

MATERIALS AND METHODS

C. lanceolata plant materials

We inbred fifty *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 based on 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 self-pollinated. A population of 140 F_2 plants was grown in the greenhouse. These plants were propagated from stem-cuttings of field plants. Those cuttings were dipped in 0.1% indole-3-butyric acid and 0.05% 1-naphthaleneacetic acid to induce rooting and placed in potting soil. They were covered with plastic for seven days and trans-planted 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 (Chapter 1; 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 used to transform 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 Birn-boim 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 4-8 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 Chapter 1 and Webb and Knapp (1990). Approximately 11 μg of DNA was digested with 50 U of *Eco*RI or *Hind*III 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 mM EDTA with pH adjusted to 8.0 with glacial acetic acid. The DNA was denatured in 0.4 M NaOH and transfered in the same solution to charged nylon membranes (CUNO Zetabind) by vacuum blotting. The membranes were neutralized in 6X SSC for 10 min, baked at 90°C for 3 h, and washed in 0.1X SSC and 0.5% SDS at 65°C for 30 min.

Hybridizations and autoradiography

Membranes were prehybridized overnight at 65°C in 6X SSC, 0.5% SDS, 5X 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 [α^{32}P]-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 2X SSC and 0.1% SDS with vigorous shaking.

X-ray film (Kodak XAR-5) was exposed to the hybridized mem-branes with one intensifying screen at -60°C for two hours to three 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 0.1X 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 screened 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 *Hind*III restriction enzymes. Probes poly-morphic 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 3-4 week old seedlings, respectively. Allozyme variation between LN43/1 and LN68/1 was assayed for aconitase (ACO), diaphor-

ase (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 10 bands using a low stringency wash (2X SSC). This 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, which was constructed using the methylation-sensitive restriction enzyme, *PstI*, 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 enzyme systems and *Mnr-2*, *Aco-2*, and *Est-3* were monomorphic.

Restriction-fragment-length polymorphisms

The RFLP rates between *C. lanceolata* lines were 14.3 to 20.2% per restriction enzyme and 29.0% for both enzymes (Table 1). The polymorphism rates between *C. viscosissima* and *C. lanceolata*, however, were approximately twice those within *C. lanceolata* (Table 1).

Greater frequencies of RFLPs between species compared to within species have also been observed for tomato (Helentjaris et

Table 1. Percentages 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 *Eco*RI or *Hind*III.

Comparison	Restriction enzyme		
	*Eco*RI	*Hind*III	*Eco*RI or *Hind*III
Between LN lines			
LN43/1 & LN68/1	14.3	17.9	31.0
LN61/1 & LN68/1	17.9	17.9	31.0
LN43/1 & LN61/1	15.5	20.2	25.0
Mean	15.9	18.7	29.0
Between LN and VS lines			
LN43/1 & VS55/1	36.9	38.1	61.9
LN68/1 & VS55/1	38.1	35.7	63.1
LN61/1 & VS55/1	38.1	36.9	63.1
Mean	37.7	36.9	62.7

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 molecular-marker linkage maps (Bernatzky and Tanksley 1986; Havey and Muehlbauer 1989; Keim et al. in press) and to map QTL (Osborn et al. 1987; Nienhuis et al. 1987; Keim et al. 1990a; Keim et al. 1990b). Interspecific mapping may also be useful for *Cuphea*. Of the 53 clones showing polymorphisms between *C. lanceolata* and *C. viscosissima* lines, 26 were polymorphic for this cross and placed on the *C. lanceolata* map (Fig. 1). The remaining clones can be mapped using interspecific populations, thereby increasing the size, density, and utility of the map.

Digests with *Eco*RI and *Hind*III restriction enzymes resulted in roughly the same polymorphism rates (Table 1). 83.0% of the poly-morphisms occurred with one but not the other enzyme. This indi-
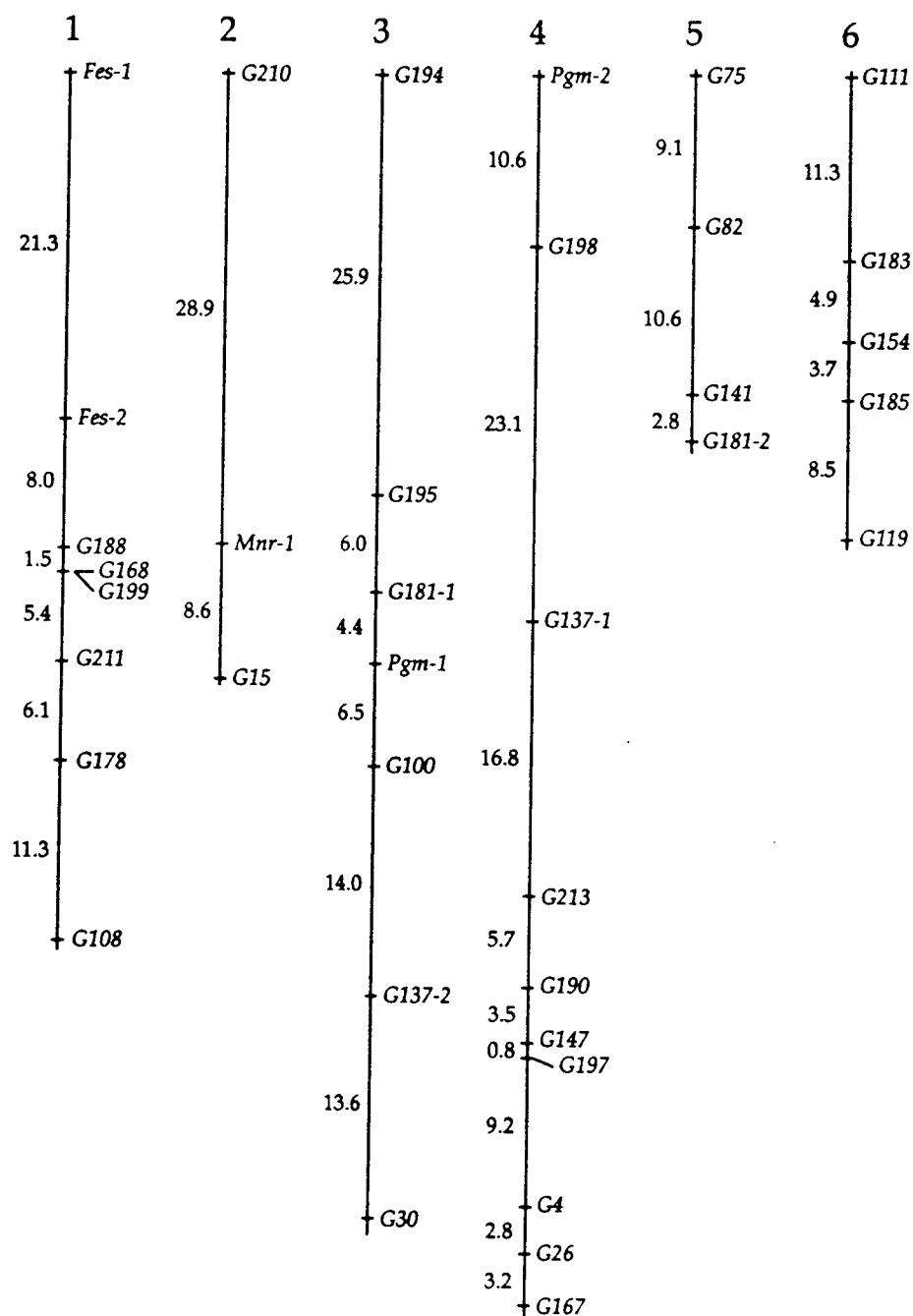
Fig. 1. A restriction-fragment-length-polymorphism and allozyme linkage map of *Cuphea lanceolata*. The chromosomes are represented by vertical lines. The positions of marker loci are shown by crossbars with names on the right and Kosambi (1944) map distances (centimorgans) between markers on the left.

cates gains or losses in enzyme recognition sequences due to basepair 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 polymorphic clones.

Marker segregation

The observed segregation ratios for the null-allele markers, *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 two 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*. All of the markers were assigned to linkage groups. If large segments of the genome were excluded from this map, we would expect to have unlinked markers or more than six linkage groups; thus, this map probably covers a significant fraction of the genome.

The size of this map is 288 cM. 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 cM *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.

This map will be expanded using *C. lanceolata* and *C. viscosissima* populations, and should assist the domestication of a *Cuphea* crop having the best characteristics of both *C. lanceolata* and *C. viscosissima*.

REFERENCES CITED

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. Amer. 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. Amer. Soc. Plant Taxon.

Graham SA (1989a) *Cuphea*: A new plant source of medium-chain fatty acids. p.139-173. Crit. Rev. Food Sci. Nutr., Vol.28. CRC Press, Boca Raton, Fla.

Graham SA (1989b) Chromosome numbers in *Cuphea* (Lythraceae): new counts and a summary. Amer. 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, 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.

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

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

Knapp SJ, Tagliani LA, Liu BH (1991) Outcrossing rates of experimental populations of Cuphea lanceolata. Plant Breeding (in press)

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. (in press)

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

Nam H, Giraudat J, den Boer B, Moonan F, Loos WDB, Hauge BM, Goodman HM (1989) Restriction fragment length polymorphism linkage map of *Arabidopsis thaliana*. 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. HortScience 19:352-354.

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

CHAPTER 3.

MAPPING QUANTITATIVE-TRAIT LOCI AFFECTING SEED CAPRIC ACID,
SEED OIL, AND EMBRYO DEVELOPMENT IN *CUPHEA LANCEOLATA*

ABSTRACT

Cuphea lanceolata Ait. is important for breeding and genetic
research to develop a new commercial source of medium-chain fatty
acids and triglycerides. Molecular markers are the best method
available to investigate the Mendelian genetics underlying quan-
titative traits, and are useful in breeding. Thirty restriction-
fragment-length-polymorphism (RFLP) and four allozyme markers were
used to locate on the *C. lanceolata* linkage map quantitative-trait
loci (QTL) affecting seed capric acid content, seed oil content,
and embryo development. Three unlinked QTL explained 19.4% of the
phenotypic variation in capric acid content in F_2 seed. Seed oil
content and seed weight were measured on seed from field-grown F_2
plants. Seed weight was indicative of embryo size (development).
Four unlinked QTL explained 33.9% of the phenotypic variation in
embryo size. One of these QTL, which explained 20.3% of the
variation, may have been a chromosomal deletion detected by a
marker having a null allele. Three QTL pleiotropically affected
seed oil by affecting embryo size. The one QTL that only affected
seed oil accounted for 2.8% of the phenotypic variation.

Key words: *Cuphea* - Quantitative-trait loci - Linkage map - Re-
striction-fragment-length polymorphisms - Allozymes

INTRODUCTION

Cuphea is a herbaceous genus having species whose seed storage lipids are predominantly medium-chain triglycerides (MCTs) (Graham 1989a). MCTs are also the storage lipids of some tree and shrub families, e.g., Lauraceae, Ulmaceae, and Myristicaceae (Shorland 1963). However, the primary commercial sources of MCTs are coconut (*Cocos nucifera* L.) and African palm (*Elaeis guineensis* Jacq.) (Arkcoll 1988).

Cuphea is being used as a model system to study medium-chain fatty acid (MCFA) and MCT biosynthesis in developing seeds because of its short life cycle and high concentrations of MCFAs (Slabas et al. 1982; Singh et al. 1986; Cao and Huang 1987). Similarly, *Cuphea* is being used as a model organism to study the genetics underlying MCFA and MCT biosynthesis.

Cuphea breeding and genetic research are presently focused on two species, *Cuphea lanceolata* Ait. and *Cuphea viscosissima* Jacq., that have predominantly capric acid seed oil. These species are conducive to genetic manipulation and experimentation. They are adapted to temperate climates (Graham 1988), produce multiple generations per year in controlled environments, regenerate from tissue culture, and can be transformed by *Agrobacterium tumefaciens* (unpublished data). They are diploids with x = 6 chromosomes (Graham 1989b), and have genome sizes approximately twice that of *Arabidopsis thaliana* (unpublished data). In addition, progeny derived from certain crosses between these species are fertile.

An RFLP and allozyme linkage map of *C. lanceolata* was con-

structed (Chapter 2) for use in breeding and genetic studies of both *C. lanceolata* and *C. viscosissima*. Molecular-marker linkage maps have been used to identify the genomic location and gene action of quantitative-trait loci (QTL) (Nienhuis et al. 1987; Edwards et al. 1987; Osborn et al. 1987; Tanksley and Hewitt 1988; Keim et al. 1990). Such studies are the best available means to understand the Mendelian inheritance of quantitative traits. In this paper, we used molecular markers to map QTL affecting seed capric acid content, seed oil content, and embryo size in a *C. lanceolata* F_2 population. We used novel statistical methods to assign QTL genotypes to recombinant marker classes and to map QTL with a dominant marker.

MATERIALS AND METHODS

Inbred-parents, F_2 population, and measurements of quantitative traits

Fifty *C. lanceolata* inbred lines were screened for seed capric acid and oil contents. Gas chromatographic (GC) measurements of fatty acid methyl esters (percentages of total fatty acid content) and nuclear magnetic resonance (NMR) measurements of seed oil (percentages of total seed mass) of the inbred lines were as described by Knapp et al. (1991). Fatty acid measurements were performed on bulk samples of 5-10 seeds from one greenhouse-grown plant per line. Oil was measured using approximately 1 g seed from each of two open-pollinated plants per line grown at the Lewis Brown Experiment Farm near Corvallis, Oregon in 1988. LN43/1 and LN68/1 were selected to maximize variation for seed capric acid and oil contents. No RFLP data were available for consideration at the time the lines were selected.

The *C. lanceolata* F_2 population used in this experiment was produced by selfing an F_1 hybrid between LN43/1 and LN68/1. Seeds for the F_2 population were imbibed and about one third of each seed's cotyledons was cut off and used for fatty acid measurements. Solution volumes to prepare the fatty acid methyl esters from these small samples were 20% of the volumes used for the inbred lines. The remainder of each seed was planted into pots, and the resulting seedlings were grown in the greenhouse for six weeks. The seed-lings were transplanted and grown 1.0 m apart at the Lewis Brown Experiment Farm near Corvallis, Oregon in 1989. Stem-cuttings from

each F_2 plant were rooted and grown in the greenhouse as the source of tissue for DNA. Open-pollinated seed was hand harvested from each F_2 plant in the field. Seed was also harvested and bulked by line from three plants of each of the parent lines, LN43/1 and LN68/1, grown in the field for comparison to F_2 phenotypes. Oil content of this seed was measured as previously described (Knapp et al. 1991).

Many of the embryos produced by LN68/1 and the F_2 population did not develop fully; however, the parental plants and seed coats of these embryos developed normally. Fifty seeds from each field-plant were weighed as a measurement of relative embryo size. To test whether underdeveloped embryos had proportionally less oil than full embryos, we measured oil content on embryos without seed coats from F_2 plants #240 and #246. Fifty embryos per plant were excised from their seed coats, dried overnight at 50°C, and weighed. The oil was extracted by grinding the embryos in 10 mL hexane using a polytron, incubating at 50°C for 1 hr, and pelleting the cellular debris at 2000 g for 5 min. The hexane was decanted to a pre-weighed pan and evaporated at 100°C. The pan was weighed again to obtain the weight of oil. Percentage oil per embryos (w/w) was calculated.

RFLP and allozyme markers

We used 29 RFLP and four allozyme codominant-markers and one RFLP dominant-marker to detect QTL in this experiment. The procedures we used to detect restriction-fragment-length and allozyme polymorphisms and the linkage map of this population have been described (Chapter 2).

Statistical analyses for each trait

The F_2 population sizes used in the analyses of seed capric acid content, seed oil content, and seed weight were 137, 131, and 131, respectively. We used a prototype of QTL-STAT, which uses linear models and estimation methods essentially equivalent to those of MAPMAKER-QTL (Lander and Botstein 1989), except that the QTL parameters for each trait were simultaneously estimated. As Lander and Botstein (1989) showed, the problem of estimating QTL parameters can be viewed as one of estimating unknown QTL genotypes.

We used an algorithm that searches the parameter space between two linked markers for estimates of recombination frequency and QTL genotypes that maximize the likelihood function for normally distributed observations. Presuming a QTL existed between linked marker loci, the algorithm estimated QTL genotypes for each recombinant flanking-marker genotype based on the expected frequencies of genotypes and the initial estimates of the means of those QTL-genotype classes (Knapp et al. 1990; Knapp and Bridges 1990). The initial estimates of μ_{11}, μ_{12}, and μ_{22} for an F_2 population are the means of the $A_1A_1B_1B_1$, $A_1A_2B_1B_2$, and $A_2A_2B_2B_2$ marker classes. QTL genotype vectors were estimated for every position between the markers at intervals of 1.0% recombination. Likelihood ratios were estimated for each of these positions. The position that had the largest likelihood-ratio estimate was taken as the position of a QTL, assuming one QTL between flanking markers. This process was repeated until the entire linkage map was searched.

Linked flanking-marker segments often detect the same QTL;

therefore, when two or more putative QTL were estimated to be within 40 recombination units of each other, only the most significant of them was tested in the multilocus model. QTL that had significant ($P \leq 0.01$) effects when estimated simultaneously were retained in the model and reestimated. This was repeated until all QTL in the model were significant.

All possible two-way interactions were tested independently of the multilocus main-effects model. Interactions were first tested using QTL genotypes of nonrecombinant flanking-marker genotype classes. Segments having significant ($P \leq 0.05$) interaction were then assigned QTL genotypes for the recombinant flanking-marker genotype classes at intervals of 1.0% recombination, and their parameters were reestimated. When two or more putative QTL were estimated within 40 recombination units of each other, only the most significant of them was tested for interactions.

One marker (G181-2) having a null allele was used to extend the genome coverage. QTL effects for the segment G181-2/G141 were tested for seed capric acid content, seed oil content, and seed weight using non-linear regression and the QTL were mapped using maximum likelihood methods (Knapp unpublished).

Additive (a) and dominance (d) effects were estimated where $\hat{a}$ = $\bar{x}_{11} - \bar{x}_{22}$ and $\hat{d}$ = $\bar{x}_{11} + \bar{x}_{22} - 2\bar{x}_{12}$. $\bar{x}_{11}$, $\bar{x}_{12}$, and $\bar{x}_{22}$ were the trait means for the QTL-genotype classes. The degree of dominance (gene action) for each QTL was estimated using $\hat{d}/\hat{a}$. Gene action was additive, partial dominance, complete dominance, or overdominance when $\hat{d}/\hat{a}$ equaled zero, more than zero but less than one, one, or more than one, respectively (Falconer 1981).

RESULTS AND DISCUSSION

The phenotypic frequency distributions (Fig. 2) were nega-
tively skewed for all three traits, most notably for seed oil
content. LN43/1, which had full embryo development, probably had
relatively low seed weight compared to much of the F_2 population
because of inbreeding. Small seed size has been found associated
with inbreeding in other *C. lanceolata* lines (unpublished data).

The inbred line LN68/1 and much of the F_2 population produced
underdeveloped embryos of varying sizes within fully developed seed
coats. This abnormality affected the measurement of seed oil by
NMR. NMR measures oil content on the mass of whole seeds, though
seed oil is mainly deposited in the embryos. Consequently, seeds
with underdeveloped embryos had low oil, and QTL on segments
G210/*Mnr*-1, G82/G141, and G111/G183 had nearly identical effects
(R^2 and d/a) on seed weight and oil content (Table 2). F_2 plants
#240 and #246, which had 10.4 and 27.5% seed oil, had 53.2 and
58.7% embryo oil by the hexane extraction method. Singh et al.
(1986) measured 53.0% oil content in fully developed embryos of *C.
lutea*, approximately that amount we found for plant #240. Because
the underdeveloped embryos of plant #240 did not have low oil
content, we conclude that the three QTL that affected embryo size
and seed oil content did not affect the synthesis of seed oil per
se.

The QTL in segment *Pgm*-1/G100 was the only QTL affecting seed
oil that did not affect seed weight; however, it only accounted for
2.8% of the phenotypic variation (Table 2). A heritability
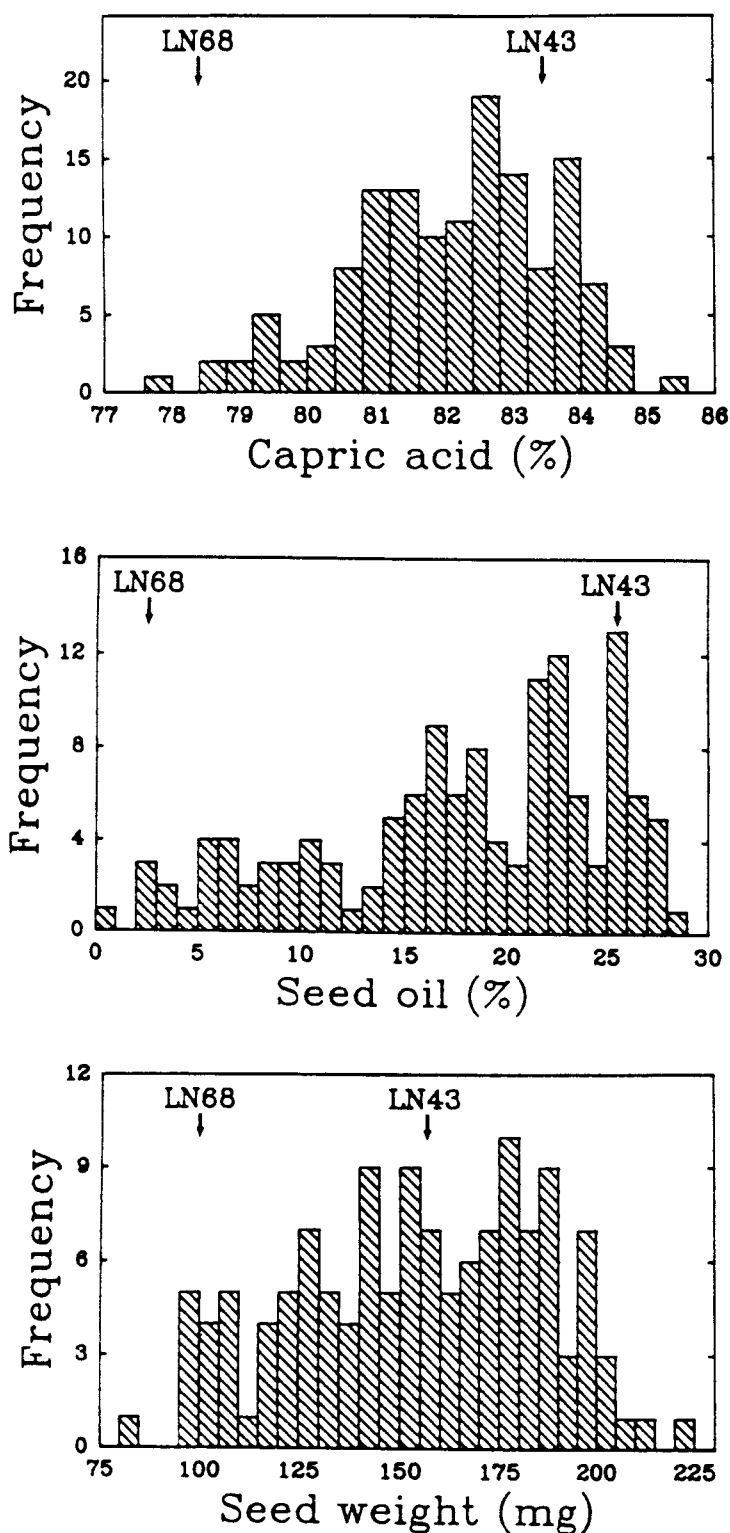
Fig. 2. Histograms showing the frequency distributions among *Cuphea lanceolata* LN43/1 x LN68/1 F₂ seed for seed capric acid content, and LN43/1 x LN68/1 F₂ plants for seed oil content and seed weight. LN43/1 and LN68/1 phenotypes are shown for comparison.

Table 2. QTL-genotypic class means and other statistics of QTL affecting seed capric acid content (% of fatty acids), seed oil content (% of seed mass), and seed weight (mg 50-seed^{-1}) in a *Cuphea lanceolata* LN43/1 x LN68/1 F_2 population. Number of individuals per QTL class are in parentheses.

	QTL class				R^2		Gene
QTL segment	43/43	43/68	68/68	P-value[†]	x100[‡]	d/a[§]	action[#]
Seed capric acid content							
G210/*Mnr*-1	81.2 (29)	82.1 (72)	82.9 (36)	≤0.001	8.2	0.04	PD
G100/G137-2	83.0 (29)	82.0 (79)	81.6 (29)	0.002	5.3	0.39	PD
G213/G190	82.7 (32)	82.3 (69)	81.2 (36)	≤0.001	5.9	0.47	PD
Seed oil content							
G210/*Mnr*-1	12.9 (26)	17.8 (70)	21.5 (35)	≤0.001	8.0	0.14	PD
Pgm-1/G100	20.8 (34)	17.7 (63)	15.2 (34)	0.003	2.8	0.11	PD
G82/G141	24.1 (31)	18.1 (72)	10.2 (28)	≤0.001	27.2	0.14	PD
G111/G183	21.3 (34)	18.0 (73)	12.6 (24)	0.010	2.2	0.24	PD
Seed weight							
Fes-1/G188	144 (26)	166 (71)	136 (34)	0.010	2.3	6.50	OD
G210/*Mnr*-1	131 (26)	154 (70)	171 (35)	≤0.001	8.9	0.15	PD
G82/G141	181 (31)	155 (72)	123 (28)	≤0.001	20.3	0.10	PD
G111/G183	170 (33)	154 (73)	133 (25)	0.008	2.4	0.14	PD

[†]Significance level of QTL in the multilocus main-effects model.
[‡]Percentage of phenotypic variance explained by QTL.
[§]Dominance ratio of QTL.
[#]PD is partial dominance and OD is overdominance.

estimate for seed oil content in a *C. lanceolata* population based

on analysis of half-sib families at two locations was 0.46 with

interval estimates of 0.29 and 0.59 (Appendix; Webb and Knapp

1991). Because these estimates were relatively high, we expected QTL to account for more of the variation for oil than we observed. In the future, a population having normal embryo development should be used to study QTL affecting seed oil.

The QTL detected by segment G82/G141 accounted for 20.3% of the phenotypic variation (59.9% of the detected genotypic variation) for seed weight, and was the primary cause of the abnormal embryo development and, consequently, seed oil content. Of the three other QTL affecting seed weight, the QTL in segment Fes-1/G188 affected seed weight without affecting seed oil content. Altogether, these QTL explained 33.9% of the phenotypic variation for seed weight (Table 2).

The segment G141/G181-2, linked to segment G82/G141 in linkage-group five (Fig. 3), was tested independently of the multilocus model because G181-2 was a dominant marker. Both segments G141/G181-2 and G82/G141 detected the major QTL affecting seed weight and seed oil content, and, for seed weight, the QTL class means and the dominance ratios were nearly identical (Tables 2 and 3). To determine which of these segments contained that QTL, we conducted a nonlinear analysis using the entire linkage group, segment G75/G181-2. That analysis located the QTL near to G181-2 (Fig. 3). The null allele for G181-2 came from LN68/1, the parent that had the abnormal embryo size. This leads us to speculate that the cause of that abnormality was a chromosomal deletion.

Embryo size as determined by seed weight in this experiment was under maternal (not embryo) nuclear genetic control. All seed produced by the F_1 plant (F_2 seed) had fully developed embryos, so
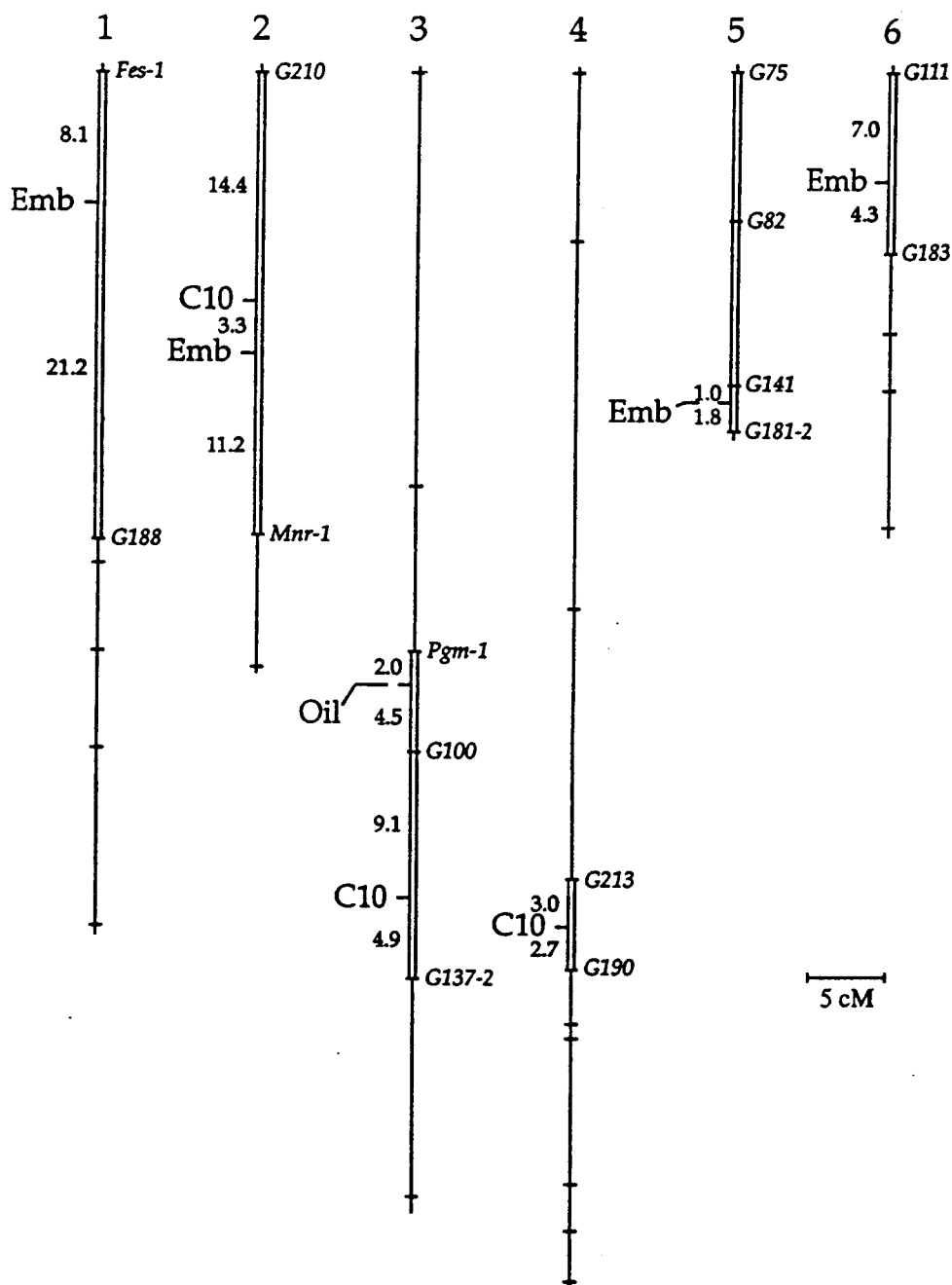
Fig. 3. Maximum-likelihood locations on the *C. lanceolata* linkage
map for QTL associated with variation in LN43/1 x LN68/1 F_2 seed
capric acid content (C10), seed oil content (Oil), and embryo size
(Emb). Marker-segments most likely to have QTL are boxed. The
positions of marker loci are shown by crossbars with the names of
markers associated with QTL on the right. Estimates of Kosambi
(1944) map distances (centimorgans) between QTL and their respec-
tive flanking markers are on the left.

Table 3. QTL-genotypic class means and other statistics of a QTL in the segment G141/G181-2 affecting seed oil content and seed weight (mg 50-seed^{-1}) in a *Cuphea lanceolata* LN43/1 x LN68/1 F_2 population.

Trait	QTL class			P-value[†]	d/a[‡]	Gene action[§]
	43/43	43/68	68/68			
Oil	23.8	18.2	10.4	≤0.001	0.17	PD
Seed weight	180.5	155.8	122.1	≤0.001	0.15	PD

[†]Significance level of QTL.
[‡]Dominance ratio of QTL.
[§]PD is partial dominance.

variation in embryo size was only evident in seed produced by the F_2 plants. The embryo-size trait, therefore, did not cause segregation distortion in the F_2 population which might have occurred from differential germination of F_2 seed, and did not affect measurements of capric acid in the F_2 seed.

Three QTL explained 19.4% of the phenotypic variation for capric acid content (Table 2).

LN43/1 had more capric acid and heavier seed than LN68/1 (Fig. 2). In the F_2 population, the QTL class means for the homozygous LN43/1 genotype were higher than those for the homozygous LN68/1 genotype for every capric-acid and embryo-size QTL except those found in segment G210/*Mnr*-1 (Table 2). LN43/1 was selected as a parent for this experiment partly because of its relatively high seed capric acid content. However, because higher concentrations of capric acid have been found in *C. lanceolata* (unpublished data), the negative LN43/1 allele at the capric-acid QTL in the segment G210/*Mnr*-1 was a reasonable occurrence. The negative LN43/1 allele at the embryo-size QTL in that same segment may be associated with

normal variation in seed weight found in *C. lanceolata* (Appendix; Webb and Knapp 1991). Genetic improvement for these traits could be made in LN43/1, perhaps by using marker-assisted selection for LN68/1 or other superior alleles at those loci.

Given a progeny population of sufficient size for estimability, simultaneous multiple-loci estimates of QTL parameters are better than single-locus estimates. Often, fewer loci are significant when estimated simultaneously because each locus effect is considered relative to the effects of all other loci. In our case, the simultaneous test for QTL affecting seed weight reduced the number of significant loci from six to four. The same is true for testing interactions with main effects. However, a simultaneous test for interactions and main effects was not done because the statistical technology has not been developed for this. We therefore tested two-way interactions individually (Table 4).

Assessment of QTL affecting medium-chain fatty acid and oil contents would be improved using populations having greater variation than found here. Phenotypic variation for fatty acid and oil content is being sought among newly developed *C. lanceolata* and *C. viscosissima* breeding lines, and the interfertility between these species offers opportunities for wide crosses. The seed fatty acid composition of several *C. viscosissima* lines have been greatly altered by chemical mutagenesis (unpublished data). Use of these *C. viscosissima* mutants crossed with *C. lanceolata* should enable us to map those mutations and better detect and estimate the effects of genes controlling medium-chain fatty acids.

Table 4. Significant two-way interactions between QTL segments for seed capric acid content, seed oil content, and seed weight in a *Cuphea lanceolata* LN43/1 x LN68/1 F_2 population.

| Trait | QTL segments | | Interaction | P-value |
	1	2		
Capric acid	G178/G108	G190/G147	Add x Add	0.008
			Dom x Add	0.003
Oil	G82/G141	G185/G119	Add x Dom	0.005
Seed weight	G82/G141	G100/G137-2	Add x Add	0.003

REFERENCES CITED

Cao YZ, Huang AHC (1987) Acyl coenzyme A preference of diacylgly-cerol acyltransferase from the maturing seeds of Cuphea, maize, rapeseed, and canola. Plant Physiol. 84:762-765.

Edwards MD, Stuber CW, Wendel JF (1987) Molecular-marker-facili-tated investigations of quantitative-trait loci in maize. I. Numbers, genomic distribution and types of gene action. Genetics 116:113-125.

Falconer DS (1981) Introduction to quantitative genetics. 2nd ed. John Wiley & Sons, Inc., New York.

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

Graham SA (1989a) Cuphea: A new plant source of medium-chain fatty acids. p.139-173. Crit. Rev. Food Sci. Nutr., Vol.28. CRC Press, Boca Raton, Fla.

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

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, Diers BW, Olson TC, Shoemaker RC (1990) RFLP mapping in soybean: Association between marker loci and variation in quantitative traits. Genetics 126:735-742.

Knapp SJ, Bridges WC (1990) Using molecular markers to estimate quantitative trait locus parameters: Power and genetic vari-ances for unreplicated and replicated progeny. Genetics 126:769-777.

Knapp SJ, Bridges WC, Jr, Birkes D (1990) Mapping quantitative trait loci using molecular marker linkage maps. Theor. Appl. Genet. 79:583-592.

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

Knapp SJ, Tagliani LA, Roath WW (1991) Fatty acid and oil diversity of Cuphea viscosissima: A source of medium-chain fatty acids. JAOCS (in press)

Kosambi DD (1944) The estimation of map distances from recombina-tion values. Ann. Eugen. 12:172-175.

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying

quantitative traits using RFLP linkage maps. Genetics 121:185-199.

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.

Singh SS, Nee TY, Pollard MR (1986) Acetate and mevalonate labeling studies with developing *Cuphea lutea* seeds. Lipids 21:143-149.

Slabas AR, Roberts PA, Ormesher J, Hammond EW (1982) *Cuphea procumbens* - a model system for studying the mechanism of medium-chain fatty acid biosynthesis in plants. Biochem. Biophys. Acta. 711:411-420.

Tanksley SD, Hewitt J (1988) Use of molecular markers in breeding for soluble solids content in tomato - a re-examination. Theor. Appl. Genet. 75:811-823.

Thompson AE (1984) Cuphea - A potential new crop. HortScience 19:352-354.

Webb DM, Knapp SJ (1991) Genetic parameters for oil yield in a population of *Cuphea lanceolata*. Crop Sci. (in press)

CONCLUSIONS

We used molecular markers to develop a linkage map and identify quantitative-trait loci in *C. lanceolata*. Molecular-marker linkage is the best method available to elucidate the Mendelian genetics underlying quantitative traits. This thesis contributes new knowledge to this rapidly expanding area of research.

Technology developments

Members of the *Cuphea* project developed many of the techniques used in this study. The DNA extraction method developed for *Cuphea* was vital to the continuation and completion of the project. Other lab procedures, including vacuum transfer of DNA to membranes, radiolabeling, and DNA hybridization were simplified for efficiency. Rapid inbred-line development, vegetative propagation and tissue production, single-seed fatty acid measurement, and embryo-oil extraction methods were perfected for *Cuphea* during this project. The statistical software, GMENDEL and QTL-STAT, were used on real data for the first time, and a nonlinear regression method to evaluate QTL with both dominant and codominant markers was developed as a direct result of this research.

It was important to develop a reliable "miniprep" method of extracting high yields of digestible *Cuphea* DNA because of the large number of extractions needed for this and future population studies. Five published and several unpublished protocols were tried, including many modifications of each. The modifications consisted in changing detergents, nuclease inhibitors, reducing agents, chelators, osmotica, tissue homogenization methods, vol-

umes, concentrations, times, temperatures, and centrifugation speeds. DNase activity was never a problem. The problem was due entirely to the physical interference of a contaminant in high concentration. Those properties affected the physical disruption of cells, the viscosities of solutions, and the isolation of DNA. The contaminant and DNA were finally separated by a combination of physical and chemical procedures. Much of the contaminant was removed with the supernatant after precipitating the DNA with CTAB and pelleting. Water saturated and buffered phenol effectively removed the remainder of the contaminant. However, most of the DNA was lost with the contaminant unless the DNA was bound to CTAB and dissolved in 1 M NaCl. Perhaps the CTAB blocked the physical or chemical association of the contaminant with the DNA.

Segregation distortion

Segregation distortion caused by genotypic selection or scoring mistakes can cause map distances, map orders, and QTL effects to be miscalculated. In this experiment, *Fes*-1 had segregation distortion, and the QTL affecting embryo development detected by segment *Fes*-1/G188 had strongly overdominant expression (Table 2). This degree of overdominance was unusual and may have resulted from incorrect scoring of bands or from epigenetic changes at *Fes*-1. The importance of segregation distortion, especially when many loci are affected, has not been addressed in the literature.

Marker-assisted selection

An important use of linkage maps is for marker-assisted

selection. In this study, we simultaneously estimated QTL parameters for each trait. These simultaneous estimates are more realistic than single-locus estimates. Similarly, simultaneous selection of QTL genotypes is more likely to give genetic improvement than selection for only one of several QTL affecting a trait. In this experiment, the LN43/1 homozygous genotype was associated with higher levels of capric acid content at two of three QTL (Table 2). Ideally, selection of segregants to maximize capric acid content would be for the LN43/1 homozygous genotype at those two loci and for the LN68/1 homozygous genotype at the third locus. Positive QTL alleles from both parents recombined in an individual progeny represents transgressive segregation for that trait.

Linkage between QTL is an important consideration in marker-assisted selection. The statistical technology to detect and differentiate closely linked QTL affecting the same trait has not been adequately developed. In this experiment, only one QTL could be detected within each 40 recombination units. If positive alleles for linked QTL are in repulsion, their effects may offset each other in the nonrecombinant classes so that a QTL will not be detected. If positive alleles for linked QTL are in coupling, a QTL may be identified but its map location may not be accurately assigned. Without accurate positioning, the wrong markers may be used in selection and QTL may be incorrectly selected. The closer the markers are to the QTL and the farther the QTL are from each other, the less of a problem this presents. These data can be re-examined when a statistical test for linked QTL has been developed.

BIBLIOGRAPHY

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. Amer. 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.

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

Cao YZ, Huang AHC (1987) Acyl coenzyme A preference of diacylglycerol acyltransferase from the maturing seeds of Cuphea, maize, rapeseed, and canola. Plant Physiol. 84:762-765.

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.

Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA minipreparation: Version II. Pl. Mol. Biol. Rep. 1:19-21.

Edwards MD, Stuber CW, Wendel JF (1987) Molecular-marker-facilitated investigations of quantitative-trait loci in maize. I. Numbers, genomic distribution and types of gene action. Genetics 116:113-125.

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 *Bras   sica*. Theor. Appl. Genet. 75:833-840.

Gebhardt C, Ritter E, Debener T, Schachtschabel U, Walkemeier B, Uhrig H, Salamini F (1989) RFLP analysis and linkage mapping in *Solanum tuberosum*. Theor. Appl. Genet. 78:65-75.

Graham SA (1988) Revision of Cuphea section Heterodon (Lythraceae). Systematic Botony Monographs, vol 20. Amer. Soc. Plant Taxon.

Graham SA (1989a) *Cuphea*: A new plant source of medium-chain fatty acids. CRC Crit. Rev. Food Sci. Nutr. 28:139-173.

Graham SA (1989b) Chromosome numbers in *Cuphea* (Lythraceae): new counts and a summary. Amer. 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 (1987) A genetic linkage map for maize based on RFLPs. TIG 3:217-221.

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.

Helentjaris T, Slocum M, Wright S, Schaefer A, Nienhuis J (1986) Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. Theor. Appl. Genet. 72:761-769.

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

Hunter RL, Markert CL (1957) Histochemical demonstration of enzymes separated by zone electrophoesis in starch gels. Science 125:1294-1295.

Kabara JJ (1984) Antimicrobial agents derived from fatty acids. J. Am. Oil Chem. Soc. 61:397-403.

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.

Keim P, Olson TC, Shoemaker RC (1988) A rapid protocol for isolating soybean DNA. Soybean Genet. Newsl. 15:150-152.

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

Knapp SJ, Bridges WC (1990) Using molecular markers to estimate quantitative trait locus parameters: Power and genetic variances for unreplicated and replicated progeny. Genetics 126:769-777.

Knapp SJ, Bridges WC, Jr, Birkes D (1990) Mapping quantitative trait loci using molecular marker linkage maps. Theor. Appl. Genet. 79:583-592.

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

Knapp SJ, Tagliani LA, Liu BH (1991) Outcrossing rates of experimental populations of Cuphea lanceolata. Plant Breeding (in press)

Knapp SJ, Tagliani LA, Roath WW (1991) Fatty acid and oil diversity of *Cuphea viscosissima*: A source of medium-chain fatty acids. JAOCS (in press)

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.

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

Liechtenstein C, Draper J (1985) Genetic engineering of plants. p.67-120. In: Glover, D.M. (ed.) DNA cloning. Vol.2., IRL Press, Oxford.

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

MacArthur JW (1934) Linkage groups in tomato. J. Genet. 29:123-133.

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

McCouch SR, Kochert G, Yu ZH, Wang ZY, Khush GS, Coffman WR, Tanksley SD (1988) Molecular mapping of rice chromosomes. Theor. Appl. Genet. 76:815-829.

Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight DNA. Nucleic Acids Res. 8:4321-4325.

Nam H, Giraudat J, den Boer B, Moonan F, Loos WDB, Hauge BM, Goodman HM (1989) Restriction fragment length polymorphism

linkage map of *Arabidopsis thaliana*. 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) Identificaiton of restriction fragment length polymorphisms linked to genes controlling soluble solids content in tomato fruit. Theor. Appl. Genet. 73:350-356.

Peirce LC, Brewbaker JL (1973) Applications of isozyme analysis in horticultural science. HortScience 8:17-22.

Rick CM, Yoder JI (1988) Classical and molecular genetics of tomato: Highlights and perspectives. Annu. Rev. Genet. 22:281-300.

Rogers SO, Bendich AJ (1985) Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. Pl. Mol. Biol. 5:69-76.

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

Sax K (1923) The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. Genetics 8:552-560.

Singh SS, Nee TY, Pollard MR (1986) Acetate and Mevalonate labeling studies with developing *Cuphea lutea* seeds. Lipids 21:143-149.

Slabas AR, Roberts PA, Ormesher J, Hammond EW (1982) *Cuphea procumbens* - a model system for studying the mechanism of medium-chain fatty acid biosynthesis in plants. Biochem. Biophys. Acta. 711:411-420.

Slocum MK, Figdore SS, Kennard WC, Suzuki JY, Osborn TC (1990) Linkage arrangement of restriction fragment length polymorphism loci in *Brassica oleracea*. Theor. Appl. Genet. 80:57-64.

Soller M, Beckmann JS (1983) Genetic polymorphism in varietal identification and genetic improvement. Theor. Appl. Genet. 67:25-33.

Tanksley SD, Bernatsky 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.

Tanksley SD, Hewitt J (1988) Use of molecular markers in breeding for soluble solids content in tomato - a re-examination. Theor. Appl. Genet. 75:811-823.

Tanksley SD, Medina-Filho H, Rick CM (1982) Use of naturally-occurring enzyme variation to detect and map genes controlling quantitative traits in an interspecific backcross of tomato. Heredity 49:11-25.

Tanksley SD, Young ND, Paterson AH, Bonierbale MW (1989) RFLP mapping in plant breeding: New tools for an old science. Bio/Technology 7:257-264.

Thompson AE (1984) *Cuphea* - A potential new crop. HortScience 19:352-354.

Wagner DB, Furnier GR, Saghai-Maroof MA, Williams SM, Dancik BP, Allard RW (1987) Chloroplast DNA polymorphisms in lodgepole and jack pines and their hybrids. Proc. Natl. Acad. Sci. USA 84:2097-2100.

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

Webb DM, Knapp SJ (1991) Genetic parameters for oil yield in a population of *Cuphea lanceolata*. Crop Sci. (in press)

APPENDIX

GENETIC PARAMETERS FOR OIL YIELD IN A POPULATION
OF *CUPHEA LANCEOLATA*

ABSTRACT

Cuphea lanceolata Ait. is a new oilseed crop being developed as a source of capric acid. Little is known about genetic variation for economically important traits in this or other *Cuphea* species. Our objectives were to estimate genetic parameters for oil yield, seed yield, seed oil content, and seed weight using half-sib families in two diverse environments and to devise a selection scheme for oil-yield improvement in *C. lanceolata*. In 1987, we derived 160 half-sib families from a broad-based synthetic population (LN-43). These families were grown in a replications-in-incomplete blocks experiment design at Corvallis and Medford, Oregon in 1988. There were significant additive genetic variances for all traits based on individual-location analyses and for all traits except seed yield based on combined-location analyses. Family X location variances were significant for oil and seed yields. Family-mean heritabilities for oil yield, seed yield, oil content, and seed weight were 0.24, 0.20, 0.46, and 0.58, respectively. Additive genetic correlations of oil yield with seed yield, oil content, and seed weight were 0.93, 0.91, and 0.85, respectively. The significant additive genetic variances, moderate heritabilities, and positive expected selection responses observed in LN-43 for oil yield suggest this population has adequate genetic variation for increasing the mean oil yield per unit area of *C. lanceolata*. Selections for oil-yield improvement should be by direct selection based on results at Medford where the environment was appropriate for the commercial production of *C. lanceolata*.

INTRODUCTION

Cuphea lanceolata Ait. is being investigated as a new commercial source of capric acid (Hirsinger 1985; Hirsinger and Knowles 1984). It is a diploid species with a haploid number of n = 6 (Graham 1989a), and is highly allogamous with multilocus outcrossing estimates from 0.61 to 0.94 (unpublished data).

Caprylic (C8:0), capric (C10:0), and lauric (C12:0) acids are the main fatty acids in seed oils of *Cuphea* species in Section Heterodon of the Lythraceae (Graham et al. 1981; Graham 1989b). These fatty acids are commercially derived from coconut (*Cocos nucifera* L.) and palm kernel (*Elaeis guineensis* Jacq.) oils and petrochemicals. *C. lanceolata* capric acid content in seed from open-pollinated plants of the population, LN-43, ranged from 780 to 860 g kg^{-1} of oil (unpublished data) compared with 32 to 150 g kg^{-1} and 26 to 70 g kg^{-1} in coconut and palm kernel oils, respectively (Young 1983). Medium-chain triglycerides rich in caprylic and capric acids have been used to treat dietetic disorders of lipid metabolism, epilepsy, gallstones, and other diseases; and as a source of rapidly absorbed, high energy fuel for critically ill patients (Bach and Babayan 1982; Babayan 1987). In addition, medium-chain triglycerides may have nutritional benefits not yet exploited for the general public (Babayan 1987).

The harvestable oil yield from *C. lanceolata* needs to be increased if the species is to be profitable. Seed shattering greatly reduces *C. lanceolata* oil yields but this problem may be overcome by optimizing crop management and harvesting techniques

and identifying mutants that retain seed. Oil yield may also be improved by breeding for oil yield, but no estimates of genetic variances and other genetic parameters for oil yield and related traits are available.

The germplasm resources of *C. lanceolata* presently consist of 24 accessions from European botanical garden collections. The wild origin of these accessions is unknown, though the species is native to Mexico. A population (LN-43) was created by randomly inter-mating all the accessions (Knapp and Tagliani 1990). Our objectives were to estimate genetic parameters in LN-43 for oil yield, seed yield, seed oil content, and seed weight using half-sib families in two diverse environments and to devise a selection scheme to improve oil-yield.

MATERIALS AND METHODS

In 1986, seeds from the 24 available *C. lanceolata* accessions were bulked and grown in a dense block at a Corvallis, Oregon isolation to produce the LN-43 population. In 1987, LN-43 plants spaced 1.0 m apart at a Corvallis, Oregon isolation produced 160 half-sib families. Pollination in both years was primarily by bumblebees *(Bombus* spp.). The half-sib families were grown in a balanced replications-in-incomplete blocks experiment design (Comstock and Robinson 1948) with 10 families and two replications in each of 16 blocks at Medford and Corvallis, Oregon in 1988. The experimental unit was a single 2.0 m row plot. The row spacing used was 1.3 m.

We planted 140 seeds per plot 1.0 cm deep on 3 May and 14 June at Medford and Corvallis, respectively. The soil type at Corvallis was Malabon (silty clay loam, fine, mixed, mesic Pachic Ultic Argixerolls) and at Medford was Central Point (sandy loam, mixed, mesic Pachic Haploxerol). The fields were irrigated using sprin-klers. Weeds were controlled by using the preemergence herbicide benefin (N-butyl-N-ethyl-a,a,a-trifluoro-2,6-dinitro-p-toluidine) and by mechanical methods. The primary pollinators were bumble-bees. Leafcutter bees *(Megachile rotundata)* were stocked in domiciles at Medford but were not effective pollinators. Leafcut-ter bees were not stocked at Corvallis because summer temperatures were too cool for them to be effective.

We combine-harvested on 21 Sept. and 30 Sept. at Medford and Corvallis, respectively, one week after applying diquat dibromide

[6,7-dihydrodipyrido(1,2-a:2',1'-c)pyrazinediium dibromide] to dessicate and defoliate the plants. The seed was then dried, threshed, and screened. We measured seed yield per plot, oil content, oil yield per plot (seed yield per plot x oil content), and seed weight. A 5.0 g sample of seed from each plot was used to measure oil content using broad band nuclear magnetic resonance (Oxford, Inc.).

There were no missing observations. Family, incomplete block, location, and replication effects were defined as random effects. Analyses were computed across and within locations. For combined location analyses, the expected mean squares for family, family X location, and residual sources of variation were $E(M_f) = \sigma^2_e + r\sigma^2_{fl} + rl\sigma^2_f$, $E(M_{fl}) = \sigma^2_e + r\sigma^2_{fl}$, and $E(M_e) = \sigma^2_e$, respectively, where r and l are the numbers of replications and locations, respectively, and σ^2_e, σ^2_{fl}, σ^2_f are residual, family X location, and family variances. Family and family X location variances were estimated using $\hat{\sigma}^2_f = (M_f - M_{fl})/rl$ and $\hat{\sigma}^2_{fl} = (M_{fl} - M_e)/r$, respectively, where M_f, M_{fl}, and M_e are mean squares corresponding to $E(M_f)$, $E(M_{fl})$, and $E(M_e)$, respectively. Approximate $(1 - \alpha)$ confidence intervals were estimated for σ^2_f and σ^2_{fl} (Graybill 1976).

Family-mean heritabilities were estimated using $H = 1 - (M_{fl}/M_f)$ across locations and $H = 1 - (M_e/M_f)$ within locations. Their $(1 - \alpha)$ confidence intervals were estimated according to Knapp et al. (1985). Expected response to half-sib family selection (R) was estimated using $\hat{R} = i\hat{H}\hat{\sigma}_p$ where i is the standardized selection differential and $\hat{\sigma}_p = (M_f/rl)^{\frac{1}{2}}$ is an estimate of the family-mean phenotypic standard deviation. An approximate estimate

of i — 1.663 was used (Lindgren 1986) based on 10% selection intensity. Normal-approximation intervals for R were estimated using the exact standard error of $\hat{R}$ (Bridges et al. 1990).

Additive genetic correlations were estimated using

$$\hat{\rho}_g = \hat{\sigma}_{a12}/(\hat{\sigma}^2_{a1}\hat{\sigma}^2_{a2})^{\frac{1}{2}}$$

where $\hat{\sigma}_{a12} = 4\hat{\sigma}_{f12}$, $\hat{\sigma}^2_{a1} = 4\hat{\sigma}^2_{f1}$, and $\hat{\sigma}^2_{a2} = 4\hat{\sigma}^2_{f2}$ and $\hat{\sigma}_{a12}$ is the additive genetic covariance and $\hat{\sigma}^2_{a1}$ and $\hat{\sigma}^2_{a2}$ are the additive genetic variances for the traits. Phenotypic correlations were estimated using

$$\hat{\rho}_p = \hat{\sigma}_{p12}/(\hat{\sigma}^2_{p1}\hat{\sigma}^2_{p2})^x$$

where $\hat{\sigma}_{p12} = M_{f12}$, $\hat{\sigma}^2_{p1} = M_{f1}$, and $\hat{\sigma}^2_{p2} = M_{f2}$, and where M_{f12} is the mean cross-product between the traits and M_{f1} and M_{f2} are the mean squares for the traits. Correlated responses to half-sib family selection were estimated using $i\hat{H}^{\frac{1}{2}}\hat{\rho}_g\hat{\sigma}_f$ where $\hat{H}^{\frac{1}{2}}$ is the square root of the heritability of the directly-selected trait and $\hat{\sigma}_f$ is the family phenotypic standard deviation for the indirectly selected trait (Falconer 1981).

RESULTS

The mean oil yield at Medford was approximately three times greater than the mean oil yield at Corvallis (Table 5). Significant variation among families was observed for oil yield, oil content, and seed weight but not seed yield across locations (Table 6). All traits had significant variation among families within both locations (data not shown). There was significant family X location interaction for oil yield and seed yield but not for oil content and seed weight (Table 6).

Family X location variances for oil content and seed weight were negative. These variances were set to zero (Table 7) and the model was redefined to $\hat{\sigma}^2_f = (M_f - M_e)/rl$. The adjustment reduced the across-location estimates of σ^2_f, H, R, and correlated responses for oil content but did not markedly change estimates for seed weight.

Heritability estimates across locations were 0.24, 0.20, 0.46, and 0.58 for oil yield, seed yield, oil content, and seed weight (Table 7). The across-location estimates were less than within-location estimates for oil yield and seed yield and greater than within-location estimates for oil content and seed weight. Consequently, estimates of across-location expected selection responses were less than within-location estimates for oil yield and seed yield and greater than within-location estimates for oil content and seed weight (Table 7). The probability is 90.0% that oil yield will increase by 0.16 to 1.22 g plot^{-1} at Corvallis, 0.75 to 2.60 g plot^{-1} at Medford, and 0.02 to 1.27 g plot^{-1} across loca-

Table 5. Oil yield, seed yield, oil content, and seed weight means, standard deviations (SD), and class limits for *Cuphea lanceolata* half-sib families grown at Corvallis and Medford, Oregon in 1988.

Location	Oil yield		Seed yield		Oil content		Seed weight	
	Mean±SD	Limits	Mean±SD	Limits	Mean±SD	Limits	Mean±SD	Limits
	——— g plot⁻1 ———		——— g plot⁻¹ ———		——— g kg⁻¹ ———		——— mg ———	
Corvallis	4.61±1.45	1.74, 9.04	22.4± 6.1	8.5, 38.1	200±17	160, 251	1268±110	1090, 1587
Medford	13.75±2.92	7.42, 23.24	53.6±10.1	31.5, 93.9	256±19	206, 296	1466±147	1226, 1730
Overall	9.18±1.77	5.56, 14.65	38.0± 6.4	24.6, 59.	229±15	186, 268	1367±163	1190, 1645

Table 6. Analyses of variance for traits of *Cuphea lanceolata* half-sib families grown at Corvallis and Medford, Oregon in 1988.

Source	df	Oil yield Mean square	P-value	Seed yield Mean square	P-value	Oil content Mean square	P-value	Seed weight Mean square	P-value
Location (L)	1	13366	≤0.001	155286	≤0.001	454980	≤0.001	6241593	≤0.001
Block (B)	15	28	0.06	471	0.02	1065	0.36	15609	0.54
L x B	15	12	0.68	148	0.20	878	0.98	16429	0.99
Replication:LB	32	15		105		2412		51700	
Family:B (F:B)	144	11	0.05	135	0.09	805	≤0.001	25918	≤0.001
F:B x L	144	8	0.02	108	0.05	397	0.74	10976	0.62
Residual	288	6		86		436		11479	

67

Table 7. Point and interval[†] estimates (C.I.) of family ($\hat\sigma^2_f$) and family X location ($\hat\sigma^2_{fl}$) variances, half-sib family-mean heritability (H), and expected response (R) to half-sib family selection for oil yield, seed yield, oil content, and seed weight of *Cuphea lanceolata* half-sib families grown at Corvallis and Medford, Oregon in 1988.

Trait	Location	$\hat\sigma^2_f$	($\hat\sigma^2_f$ C.I.)	$\hat\sigma^2_{fl}$	($\hat\sigma^2_{fl}$ C.I.)	Ĥ	(Ĥ C.I.)	Ř	(Ř C.I.)
								g plot⁻¹	
Oil yield	Corvallis	0.60	(0.10, 1.19)			0.28	(0.06, 0.46)	0.69	(0.16, 1.22)
	Medford	2.76	(1.05, 4.80)			0.37	(0.17, 0.52)	1.68	(0.75, 2.60)
	Combined	0.64	(-0.14, 1.41)	1.04	(0.18, 2.11)	0.24	(-0.01, 0.42)	0.65	(0.02, 1.27)
								g plot⁻¹	
Seed yield	Corvallis	8.6	(-0.2, 18.7)			0.24	(-0.01, 0.42)	2.37	(0.08, 4.66)
	Medford	26.9	(7.1, 50.2)			0.32	(0.10, 0.48)	4.86	(1.60, 8.11)
	Combined	6.7	(-1.5, 16.2)	11.05	(-0.14, 24.94)	0.20	(-0.05, 0.39)	1.92	(-0.36, 4.20)
								g kg⁻¹	
Oil content	Corvallis	79.0	(10.0, 160.0)			0.27	(0.04, 0.45)	7.7	(1.5, 14.0)
	Medford	98.0	(22.0, 187.0)			0.30	(0.08, 0.47)	9.1	(2.7, 15.5)
	Combined	92.3	(48.1, 145.6)	0.0		0.46	(0.29, 0.59)	10.8	(6.5, 15.2)
								mg	
Seed weight	Corvallis	3141	(1495, 5112)			0.43	(0.25, 0.57)	60.8	(33.4, 88.1)
	Medford	3827	(1281, 6835)			0.34	(0.14, 0.50)	60.5	(24.4, 96.7)
	Combined	3736	(2386, 5397)	0.0		0.58	(0.44, 0.68)	77.2	(55.5, 98.9)

[†]The confidence coefficient used was $(1 - \alpha) - 0.90$.

tions through direct selection for oil yield.

Additive genetic correlations between traits across locations were positive and comparatively high (Table 8). The high correlations of oil yield with seed yield (0.93) and oil content (0.91) were expected given that oil yield was the product of those two traits. However, oil yield was also highly correlated with seed weight (0.85).

Table 8. Additive genetic and phenotypic (in parentheses) correlations between traits of *Cuphea lanceolata* half-sib families grown at Corvallis and Medford, Oregon in 1988.

	Seed yield	Oil content	Seed weight
Oil yield	0.93	0.91	0.85
	(0.94)	(0.58)	(0.49)
Seed yield		0.55	0.71
		(0.30)	(0.30)
Oil content			0.69
			(0.70)

The expected correlated selection responses were positive but most were less than the corresponding expected direct selection responses (Table 9). Selection for high oil content or seed weight across locations was expected to increase oil yield more than direct selection for oil yield across locations. However, expected responses from direct and indirect selections across locations were less than from selections for Medford alone.

Table 9. Expected correlated response to half-sib family selection for oil yield, seed yield, oil content, and seed weight of *Cuphea lanceolata* half-sib families grown at Corvallis and Medford, Oregon in 1988.

Directly selected trait	Location	Response trait			
		Oil yield	Seed yield	Oil content	Seed weight
		— g plot^{-1} —		g kg^{-1}	mg
Oil yield	Corvallis	**0.69**[†]	2.49	5.24	24.3
	Medford	**1.68**	5.01	6.92	46.8
	Combined	**0.65**	1.96	7.02	40.9
Seed yield	Corvallis	0.61	**2.37**	3.16	18.8
	Medford	1.49	**4.86**	4.00	38.7
	Combined	0.55	**1.92**	3.89	29.3
Oil content	Corvallis	0.45	1.11	**7.74**	30.4
	Medford	1.04	2.04	**9.08**	31.5
	Combined	0.82	1.59	**10.82**	46.3
Seed weight	Corvallis	0.42	1.35	6.19	**60.8**
	Medford	1.18	3.26	5.18	**60.5**
	Combined	0.85	2.19	8.33	**77.2**

[†]Bold-face values indicate response to direct selection.

DISCUSSION

The significant additive genetic variances, moderate herita-
bilities, and positive expected selection responses observed in
LN-43 for oil yield (Tables 6 and 7) suggest this population has
adequate genetic variation for increasing the mean oil yield per
unit area of *C. lanceolata*.

Initial selections for oil yield improvement should be directly
for oil yield based on results at Medford where the enviroment was
more appropriate for commercial production of this species. We
attribute the significantly higher yields at Medford (Table 5) to
larger plant size resulting from higher temperatures and a longer
growth period. The mean daily-high-temperatures between 1 June and
30 Sept. 1988 were 28.7°C and 26.0°C at Medford and Corvallis,
respectively, and the growth period at Medford was 33 days longer
than at Corvallis. *C. lanceolata* has indeterminate growth, there-
fore larger plants have more branches and consequently more flow-
ers. Assuming adequate pollination and growth conditions, more
flowers produce more seed per plant. Since these temperatures and
growth periods were not unusual, we expect lower *C. lanceolata*
yields at Corvallis than at warmer locations in most years.

Indirect selection for oil yield by selection for oil content
or seed weight should be evaluated in future *C. lanceolata* experi-
ments. The high genetic correlations between oil yield and the
other traits (Table 8) and the higher heritabilities of oil content
and seed weight compared with oil yield (Table 7) indicate indirect
selection for oil yield may be advantageous for some environments.

The 0.55 correlation (Table 8) observed between seed yield and oil content for *C. lanceolata* LN-43 is notable for being relatively high and positive. With other seed crops, increasing the yield of high energy compounds such as lipids or proteins often concomitantly decreases overall seed yield by decreasing carbohydrate content. Unlike most seed crops, however, *C. lanceolata* has indeterminate growth. Increasing its seed oil content may be at the expense of vegetative rather than seed development. Higher seed oil content may contribute to higher total seed yield but additional data are needed to assess this relationship.

REFERENCES CITED

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

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

Bridges WC, Knapp SJ, Cornelius PL (1991) Standard errors and confidence interval estimators for expected selection response. Crop Sci. (in press)

Comstock RE, Robinson HF (1948) The components of genetic variance in populations of biparental progenies and their use in estimating the average degree of dominance. Biometrics 4: 254-266.

Falconer DS (1981) Introduction to quantitative genetics. 2nd ed. John Wiley & Sons, Inc., New York.

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

Graham SA (1989b) Cuphea: A new plant source of medium-chain fatty acids. CRC Crit. Rev. Food Sci. Nutr. 28:139-173.

Graham SA, Hirsinger F, Robbelen G (1981) Fatty acids of *Cuphea* (Lythraceae) seed lipids and their systematic significance. Amer. J. Bot. 68:908-917.

Graybill FA (1976) Theory and application of the linear model. Duxbury Press, Boston.

Hirsinger F (1985) Agronomic potential and seed composition of cuphea, an annual crop for lauric and capric seed oils. J. Amer. Oil Chem. Soc. 62:76-80.

Hirsinger F, Knowles PF (1984) Morphological and agronomic description of selected *Cuphea* germplasm. Econ. Bot. 38:439-451.

Knapp SJ, Stroup WW, Ross WM (1985) Exact confidence intervals for heritability on a progeny mean basis. Crop Sci. 25:192-194.

Knapp SJ, Tagliani L (1990) Genetic variation for seed dormancy in *Cuphea laminuligera* and *Cuphea lanceolata*. Euphytica 47:65-70.

Lindgren D (1986) An approximate formula for selection intensity. Silvae Genetica 35:249.

Young FVK (1983) Palm kernel and coconut oils: Analytical characteristics, process technology, and uses. J. Amer. Oil Chem. Soc. 60:374-379.