AN ABSTRACT OF THE THESIS OF

Carol A. Powers for the degree of Master of Science in Crop Science presented on August 30, 2005.
Title: Breeding Tools for *Cuphea*: An Updated Genetic Linkage Map and Fatty Acid Mutant Candidate Gene Analysis

Abstract approved:
Redacted for privacy

Steven J. Knapp

Medium chain fatty acids (MCFAs) are important industrial feedstocks for soaps, detergents and surfactants. Interspecific lines derived from *Cuphea lanceolata* and *Cuphea viscosissima* are being domesticated as a temperate source of MCFAs. These species also serve as model organisms for understanding the biosynthesis of MCFAs in seeds. Various markers (176 SSRs, 22 fatty acid synthesis loci, and 27 converted RFLP markers) were screened for polymorphism among at least four genotypes (VS55, LN185, LN193-S_1, and PSR23). The SSRs had heterozygosities ranging from 0.278 to 0.889. An F_2 population ((LN185 x LN193-S_1, 91 individuals) of *C. lanceolata* (x=6) was genotyped at 90 loci for constructing a genetic linkage map. Eighty eight loci (48 SSRs, 23 SSCPs, 7 SNPs, 10 fatty acid synthesis loci) coalesced into eight linkage groups (two markers segregated independently). The map was 340.8 cM long with a mean density of 3.9 cM per locus. This linkage map will be a useful tool for the continued breeding and domestication of *Cuphea.*

Two available EMS mutants of *C. viscosissima*, *cei-2* and *cpr-1*, that reduce C8:0 and C10:0 respectively, were characterized. The *CvFatB3* and *CvKASIV-2*

candidate genes were sequenced in wild type and mutant lines to identify single nucleotide polymorphisms (SNP). A SNP was identified in each gene and SNP assays were developed. An F_2 population of 92 individuals segregating for the two mutations was phenotyped for fatty acids by quarter seed analysis on a gas chromatograph and genotyped for the two SNPs using DNA prepared from remnant seed tissue. The SNP in the *cei-2* candidate gene, *CvFatB3*, correlated with a decrease in C8:0. The SNP in the *cpr-1* candidate gene, *CvKASIV-2*, correlated with a decrease in C10:0. The double homozygous mutant increased C14:0 >30-fold relative to wild type *C. viscosissima*. The combination of these two mutations produced a novel oil profile in domesticated *Cuphea*.

©Copyright by Carol A. Powers
August 30, 2005
All Rights Reserved

Breeding Tools for *Cuphea*: An Updated Genetic Linkage Map of and Fatty Acid Mutant Candidate Gene Analysis

by

Carol A. Powers

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented August 30, 2005
Commencement June 2006

Master of Science thesis of Carol A. Powers presented on August 30, 2005.

APPROVED:

Redacted for privacy

Major Professor, representing Crop Science

Redacted for privacy

Head of the Department of Crop and Soil Sciences

Redacted for privacy

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Redacted for privacy

Carol A. Powers, Author

ACKNOWLEDGEMENTS

I thank my major professor, Dr. Steven Knapp, for introducing me to the world of *Cuphea* and guiding and supporting this research project. I will be forever grateful to Dr. Mary Slabaugh, for her advice, troubleshooting expertise, and help in so many aspects of my research, as well as serving on my committee. I thank Dr. Isabel Vales and Dr. Greg Thompson for serving on my committee and giving generously of their time to help make the completion of my graduate program a smooth process.

I thank the members of the lab who made my time here at Oregon State University enjoyable and provided much needed insight so many times: Mary Slabaugh, Bob Brunick, Shaun Bushman, Sonali Ghandi, Adam Heesacker, Cate Hass, Felix Schuppert, Robin Hawley, and Guillermo Pizzaro. I particularly thank my fellow graduate students Cate, Felix, Robin and Guillermo, for all their support during this research. My time here would not have been nearly so enjoyable nor productive without you.

I thank all the graduate students in Crop Science for their friendship and support during my time here. I am appreciative of my friends, both near and far, who supported me during this research.

I am grateful to my family, for their love, support, cheerleading, and advice during my time here. Without them I would not be able to accomplish all that I have in life.

CONTRIBUTION OF AUTHORS

Dr. Steven J. Knapp proposed and guided this research. He significantly assisted in data analysis, writing and polishing of this thesis. Dr. Mary B. Slabaugh greatly contributed to the marker development, particularly the conversion of the RFLP markers and additional development of fatty acid synthesis markers. She also developed the *KASIV-2* SNP assay and greatly contributed to the writing and polishing of this thesis.

TABLE OF CONTENTS

	Page
Chapter 1: Introduction	1
Works Cited	7
Chapter 2: An Updated Genetic Linkage Map of *Cuphea*	11
Abstract	11
Introduction	12
Materials and Methods	14
Results	22
Discussion	40
Works Cited	44
Chapter 3: Fatty Acid Mutant Candidate Gene Analysis in *Cuphea*	48
Abstract	48
Introduction	49
Materials and Methods	52
Results	58
Discussion	69
Works Cited	75
Chapter 4: Conclusions	77
Bibliography	80
Appendix	86
Amplification of SSR Markers	87

LIST OF FIGURES

Figure		Page
2-1	Heterozygosities of SSR markers for which allele fragment sizes were available for at least four of the screening panel genotypes..	26
2-2	Primers designed to span the RFLP insert fragments were screened across four genotypes VS55, PSR23, LN185 and LN193 and assayed under SSCP conditions..........................	32
2-3	Scatter plots from six SNP FP-TDI assays...........................	35
2-4	LN185 allele frequency and heterozygote frequency among LN185 x LN193-S_1 F_2 progeny for ordered loci on the linkage map of *C. lanceolata*...	37
2-5	Linkage map of LN185 x LN193-S_1..................................	39
3-1	Cladogram based upon multiple sequence alignment of available *Cuphea FatB* amino acid sequences with the three clades and substrate specificities noted...	59
3-2	Amino acid alignment of *CvFatB3-WT*, *CvFatB3-cei-2* and *ClFatB3*..	60
3-3	C8:0/C14:0 fatty acid phenotypes and *CvFatB3* (A) and *CvKasIV-2* (B) SNP genotypes among 93 F_2 progeny from VS342-6 x VL198..	62
3-4	The *CvFatB3* SNP (candidate for the *cei-2* mutation) assay results..	65
3-5	The *CvKasIV-2* SNP (candidate for the *cpr-1* mutation) assay results..	66

LIST OF TABLES

Table		Page
2-1	Fragment lengths for SSRs screened on a panel of six genotypes	23
2-2	Fatty acid synthesis primers used to detect polymorphism between LN185 and LN193-S_1	27
2-3	Amplification of *FatB* and *KasIV* loci in LN185 and LN193-S_1	29
2-4	RFLP inserts were sequenced and primers designed to each end of the insert	30
2-5	SNP markers used in mapping LN185 x LN193-S_1	34
3-1	Fatty acid profiles of wild type *C. viscosissima* and *cpr-1* and *cei-2* mutants	53
3-2	The LSMeans for fatty acids of the nine genotypic classes of *cei-2* and *cpr-1*	67
3-3	Additive and dominant effects of the candidate *cpr-1* and *cei-2* SNPs	68

Breeding Tools for *Cuphea*: An updated genetic linkage map of *Cuphea* and fatty acid mutant candidate gene analysis in *Cuphea*

CHAPTER 1

Introduction

Medium chain fatty acids (MCFAs) are used in various non-food and food products, including soaps, detergents, surfactants, special diets for lipid metabolism disorders, dietary margarines, confectionary, cosmetics, and laundries (Topfer et al. 1995). The US relies heavily on imported palm kernel (*Elaeis guinenes* Jacq.) and coconut (*Cocus nucifera* L.) oils for natural MCFAs, with petroleum as the other major source (Thompson et al. 1990, Murphy 2002). In 2002, nearly 700,000 Mt of coconut and palm kernel oil were imported into the US (FAOSTAT, 2005). Species of the genus *Cuphea* are known for their oil profiles rich in MCFAs and are potential sources of MCFAs for temperate regions and thus are subject to domestication efforts in the US (Thompson 1984, Graham 1988, Webb et al. 1992, Knapp 1993). The domestication of *Cuphea* will diversify the supply of MCFAs and should help stabilize longterm supplies and prices of MCFAs (Webb et al. 1992). *Cuphea lanceolata* Ait. and *Cuphea viscosissima* Jacq., which contain ~83% and ~70% capric acid (C10:0) respectively, have been the primary focus of the US domestication efforts, in part because they produce fertile hybrids, expanding the genetic base that can be utilized in domestication and breeding efforts (Webb et al. 1992, Brandt and Knapp 1993, Knapp

1993). *C. viscosissima* is native to the southeastern United States and *C. lanceolata* is native to the Sierra Madre region of Mexico (Graham 1988).

Most plants accumulate longer fatty acids for energy storage in their seeds; *Cuphea* species are among few species that store primarily MCFAs (Earle et al. 1960, Miller et al. 1964, Thompson 1984, Graham 1988). The interest in *Cuphea* is threefold. First, *Cuphea* species are model organisms for understanding MCFA synthesis in seeds (Webb et al. 1992, Kopka et al. 1993, Dehesh 2001). Second, this understanding can lead to targeted breeding efforts that result in novel and desirable oil profiles in *Cuphea*. Third, genes in *Cuphea* species can, and have, be utilized to alter other crop profiles through genetic transformation (Jones et al. 1995, Dehesh et al. 1996a and 1998,). *C. lanceolata* and *C. viscosissima* are good models for elucidating these mechanisms as they have small diploid genomes, in addition to being important in the breeding of *Cuphea* (Webb et al. 1992, Kopka et al. 1993).

Two major classes of enzymes are known to affect MCFA profiles in *Cuphea* seeds, medium chain specific fatty acyl-ACP thioesterases (FATB) and β-ketoacyl-ACP synthase (KAS) (Fuhrmann and Heise 1993, Dehesh et al. 1996a, b and 1998, Leonard et al. 1998). KAS enzymes extend the acyl chain by two carbons. Four classes of KAS enzymes are known in plants, distinguished by substrate specificity and cerulenin sensitivity (Schutt et al. 2002). In *C. lanceolata*, KASIV lengthens C4:0- to C10:0-ACP preferentially, though C12:0- and C14:0-ACP can also be substrates (Schutt et al. 2002). In various C8:0 to C12:0 producing *Cuphea* species (*C. wrightii, C. pulcherrima, C. hookeriana, C. viscosissima, C. lanceolata,* and *C. tolucana*) Slabaugh et al. (1998) found significant expression of KASIV, while *C.*

denticulata, a species lacking C8:0-C14:0, had minimal expression of KASIV. Thus, KASIV appears to be necessary to produce MCFAs. FatB thioesterases hydrolyze the bond between the acyl group and ACP, releasing the acyl group so it can be transported to the cytosol in the ER to be incorporated into triacylglycerol (Topfer et al. 1995). FatBs tend to prefer saturated acyl-ACPs as substrate (Salas and Ohlrogge 2002). In transgenic assays, FatBs were shown to be necessary, but not sufficient to yield the unique oil profile (30% capric and 54% lauric acid) of *C. wrightii* (Leonard et al. 1997). In *C. palustris* two FatBs whose substrate specificity and kinetic efficiency clearly correspond to the oil profile high in myristic (64%) and caprylic (20%) acids (Dehesh et al. 1996). Substrate specificity and kinetic efficiency of FatB enzymes play a role in determining the oil profile of *Cuphea* seeds.

Cuphea species are also of interest as sources of genes to alter domestic crop oil profiles through genetic modification. Much effort has gone into increasing MCFA in canola, with limited success, resulting in the release of Laurical®, which has approximately 45% lauric acid (C12:0) (Voelker et al. 1992, Dehesh 2001, Murphy 2002). High copy number of a C12:0-acyl-carrier protein (ACP) specific thioesterase gene from *Umbellularia californica* (California Bay) was required to reach these concentrations (Voelker et al. 1996). The goal of 80-90% concentration of a single MCFA has not been realized through genetic modification, likely because other enzymes, such as specialized acyltransferases, in the fatty acid synthesis and storage pathways appear to be necessary for the MCFA produced to be incorporated into triacylglycerols in the seed storage organs (Dehesh 2001, Larson et al 2002). *Cuphea*

species are already capable of high concentrations of one or two MCFAs which is one reason for continued interest in their domestication.

A significant tool for modern-day plant breeders is genetic linkage maps. Not only are genetic maps prerequisites for genetic studies such as quantitative trait loci (QTL) analysis and map based cloning, but they are key to marker assisted selection (MAS) (Liu et al. 1996). The current linkage map in *C. lanceolata* is comprised of 32 restriction fragment length polymorphism (RFLP) and five allozyme markers (Webb et al. 1992). Subsequently, 22 fatty acid synthesis loci were mapped as allele-specific PCR or single strand conformational polymorphism (SSCP) markers using a *C. viscosissima* x *C. lanceolata* interspecific F_2 population (Slabaugh et al. 1997). A second generation map consisting of PCR-based markers, notably with the addition of recently developed simple sequence repeat (SSR, Karmarkar 2004) and the existing fatty acid synthesis markers, will increase the density, utility and ease of use of the map. Ultimately, as a map of *C. viscosissima* and an interspecific map of *C. viscosissima* x *C. lanceolata* are constructed, it should be possible to detect duplications, rearrangements and other characteristics of the genomes. The allogamous *C. lanceolata* was selected to construct a linkage map in first because it showed more allozyme polymorphism than the autogamous *C. viscosissima* (Knapp and Tagliani 1989, Webb et al. 1992). Karyotype analysis has indicated that *C. lanceolata* and *C. viscosissima* share a close phylogenetic relationship, as five of the six homologous chromosomes have the same overall structure and the C-banding patterns are similar (Chen and Roath 1995).

A screen of fatty acid content in available *C. viscosissima* accessions showed a very narrow range for each MCFA (Knapp et al. 1991). The limited variation in *C. viscosissima* fatty acid profiles led researchers to perform chemical mutagenesis with ethyl methanesulfonate (EMS), to induce mutations affecting fatty acid synthesis in seeds (Knapp and Tagliani 1991). Various mutants that exhibit reduced capric or caprylic acid have been characterized and have aided in the understanding of MCFA synthesis in *Cuphea* seeds (Knapp and Tagliani 1991, Knapp et al. 1997). The characterization of mutations and their interactions allows for the effective utilization of mutations in breeding *Cuphea*. Fatty acid mutants also are useful in understanding the mechanisms of MCFA synthesis and their role in controlling the fatty acid profile of *C. viscosissima* (Knapp and Tagliani 1991). *C. lanceolata* and *C. viscosissima* have very different C8:0 content, with *C. lanceolata* containing ~1% C8:0 and *C. viscosissima* containing ~18% C8:0 (Knapp 1993). The locus controlling this difference has been designated *cei-1*. One mutant, *cei-2*, which reduces C8:0 from 18% to ~2%, has been combined with another mutant, *cpr-1*, which reduces C10:0 from 71% to ~38%. Together these two mutations result in both reduced C8:0 and C10:0 and increased C14:0 and C16:0. The FatB enzyme with C8:0-C10:0-ACP specificity is a candidate for the *cei-2* mutation. KasIV-2 extends C10:0-ACP is a candidate for the *cpr-1* mutation. Identification of single nucleotide polymorphisms (SNPs) in these candidate genes that correlate to the mutant phenotype will allow for the screening of seedlings, rather than the seed off the plant, saving substantial amounts of time (Stuber et al. 1999).

The currently cultivated form of *Cuphea* is an interspecific of *C. viscosissima* and *C. lanceolata*, PSR 23 (Knapp and Crane 2000). PSR23 is a partially seed retaining line, that has reduced seed shattering, one of the major hurdles to the domestication of *Cuphea* (Knapp 1993). Research into the best management practices for this semi-domesticated *Cuphea* is currently underway (Gesch et al. 2003). While the incorporation of *C. lanceolata* background in the domestication process has resulted in some variation to the fatty acid profiles, mutants are still an integral part of breeding for the economically significant MCFAs (Knapp and Tagliani 1991).

The objectives of the studies presented here were to:

- Identify polymorphic SSR markers for 6 *Cuphea* genotypes
- Develop PCR-based markers from existing *Cuphea* RFLP markers
- Construct a genetic linkage map in *C. lanceolata* utilizing SSR, fatty acid markers (SSCP), and RFLP markers converted to PCR-based markers
- Characterize the *cei-2* and *cpr-1* mutants of *C. viscosissima*
- Identify a single nucleotide polymorphisms (SNPs) that may correlate with the *cei-2* and *cpr-1* mutations
- Develop a SNP assay to detect the *cei-2* and *cpr-1* mutations
- Determine the additive and dominant effects of *cei-2* and *cpr-1* and their interactions

Works Cited

Brandt, T., and S.J. Knapp. 1993. Gametic selection at fatty acid and allozyme marker loci and meiosis within *Cuphea viscosissima* x *Cuphea lanceolata* populations. Crop Science 33:1138-1143.

Chen, W., and W.W. Roath. 1995. Karyotype of *Cuphea lanceolata* Ait. and *Cuphea viscosissima* Jacq. Crop Science 35:246-250.

Dehesh, K. 2001. How can we genetically engineer oilseed crops to produce high levels of medium-chain fatty acids? European Journal of Lipid Science and Technology 103:688-697.

Dehesh, K., P. Edwards, J. Fillatti, M. Slabaugh, and J. Byrne. 1998. KAS IV: a 3-ketoacyl-ACP synthase from *Cuphea* sp. is a medium chain specific condensing enzyme. The Plant Journal 15:383-390.

Dehesh, K., P. Edwards, T. Hayes, A.M. Cranmer, and J. Fillatti. 1996. Two novel thioesterases are key determinants of the bimodal distribution of acyl chain length of *Cuphea palustris* seed oil. Plant Physiol. 110:203-210.

Dehesh, K., A. Jones, D.S. Knutzon, and T.A. Voelker. 1996. Production of high levels of 8:0 and 10:0 fatty acids in transgenic canola by overexpression of Ch FatB2, a thioesterase cDNA from *Cuphea hookeriana*. Plant J 9:167-172.

Earle, F.R., C.A. Glass, G.C. Geisinger, I.A. Wolff, and Q. Jones. 1960. Search for new industrial oils IV. Journal of American Oil Chemists Society 67:440-447.

FAOSTAT data 2005. http://faostat.fao.org/faostat/collections?version=ext&hasbulk=0&subset=agriculture

Fuhrmann, J., and K. Heise. 1993. Factors controlling medium-chain fatty acid synthesis in plastids from maturing *Cuphea* embryos. Z. Naturforsch 48c:616-622.

Gesch, R.W., F. Forcella, N.W. Barbour, W.B. Voorheees, and B. Phillips. 2003. Growth and yield response of *Cuphea* to row spacing. Field Crops Research 81:193-199.

Graham, S. 1988. Revision of *Cuphea* section *Heterodon* (Lythraceae). Syst. Bot. Mono. 20:1-168.

Jones, A., H.M. Davies, and T.A. Voelker. 1995. Palmitoyl-acyl carrier protein (ACP) thioesterase and the evolutionary origin of plant acyl-ACP thioesterases. Plant Cell 7:359-371.

Karmarkar, V.M. 2004. Simple sequence repeat marker development and fingerprinting in *Cuphea lanceolata* Ait. and *C. viscosissima* Jacq. Masters, Oregon State University, Corvallis.

Knapp, S.J. 1993. Modifying the seed storage lipids of *Cuphea*: a source of medium-chain tiglycerides, p. 142-154, In S. MacKenzie and D. Taylor, eds. Seed oils for the future. American Oil Chemistr' Society, Champaign IL.

Knapp, S.J., J. Crane, L. Tagliani, and M. Slabaugh. 1997. *Cuphea viscosissima* mutants with decreased capric acid. Crop Science 37:352-357.

Knapp, S.J., and J.M. Crane. 2000. Registration of reduced seed-shattering *Cuphea* germplasm PSR23. Crop Science 40:299-300.

Knapp, S.J., and L. Tagliani. 1991. Two medium-chain fatty acid muntants of *C. viscosissima*. Plant Breeding 106:338-341.

Knapp, S.J., L. Tagliani, and W.R. Roath. 1991. Fatty acid and oil diversity of *Cuphea viscosissima*: a source of medium-chain fatty acids. Journal of American Oil Chemists Society 68:515-517.

Knapp, S.J., and L.A. Tagliani. 1989. Genetics of allozyme variation in *Cuphea lanceolata* Ait. Genome 32:57-63.

Kopka, J., M. Robers, R. Schuch, and F. Spener. 1993. Acyl carrier proteins from developing seeds of *Cuphea lanceolata* Ait. Planta 191:102-111.

Larson, T.R., T. Edgell, J. Byrne, K. Dehesh, and I.A. Graham. 2002. Acyl CoA profiles of transgenic plants that accumulate medium-chain fatty acids indicate inefficient storage lipid synthesis in developing oilseeds. Plant J 32:519-527.

Leonard, J.M., S.J. Knapp, and M.B. Slabaugh. 1998. A *Cuphea* β-ketoacyl-ACP synthase shifts the synthesis of fatty acids towards shorter chains in *Arabidopsis* seeds expressing *Cuphea* FatB thioesterases. The Plant Journal 13:621-628.

Leonard, J.M., M.B. Slabaugh, and S.J. Knapp. 1997. *Cuphea wrightii* thioesterases have unexpected broad specificities on saturated fatty acids. Plant Molecular Biology 34:669-679.

Liu, Z., R. Biyashev, and M. Saghai Maroof. 1996. Development of simple sequence repeat DNA markers and their integration into a barley linkage map. Theoretical and Applied Genetics 93:869-876.

Miller, R., F.R. Earle, I.A. Wolff, and Q. Jones. 1964. Search for new industrial oils IX *Cuphea*, a versatile source of fatty acids. Journal of American Oil Chemists Society 41:279-280.

Murphy, D.J. 2002. Biotechnology and the improvement of oils crops—genes, dreams and realities. Phytochemistry Reviews 1:67-77.

Salas, J.J., and J.B. Ohlrogge. 2002. Characterization of substrate specificity of plant FatA and FatB acyl-ACP thioesterases. Archives of Biochemistry and Biophysics 403:25-34.

Schutt, B., A. Abbadi, B. Loddenkotter, M. Brummel, and F. Spener. 2002. β-Ketoacyl-acyl carrier protein synthase IV: a key enzyme for regulation of medium-chain fatty acid synthesis in *Cuphea lanceolata* seeds. Planta 215:847-854.

Slabaugh, M.B., G.M. Huestis, J. Leonard, J.L. Holloway, C. Rosato, V. Hongtrakul, N. Martini, R. Toepfer, M. Voetz, J. Schell, and S.J. Knapp. 1997. Sequence-based genetic markers for genes and gene families: single-strand conformational polymorphisms for the fatty acid synthesis genes of *Cuphea*. Theoretical and Applied Genetics 94:400-408.

Slabaugh, M.B., J.M. Leonard, and S.J. Knapp. 1998. Condensing enzymes from *Cuphea wrightii* associated with medium chain fatty acid biosynthesis. Plant J 13:611-620.

Stuber, C.W., M. Polacco, and M.L. Senior. 1999. Synergy of empirical breeding, marker-assisted selction, and genomics to increase crop yield potential. Crop Science 39:1571-1583.

Thompson, A.E. 1984. *Cuphea*-a poltential new crop. HortScience 19:352-354.

Thompson, A.E., D. Dierig, S.J. Knapp, and R. Kleiman. 1990. Variation in fatty acid content and seed weight in some lauric acid rich *Cuphea* species. Journal of American Oil Chemists Society 67:611-617.

Topfer, R., N. Martini, and J. Schell. 1995. Modification of Plant Lipid Synthesis. Science 268:681-686.

Voelker, T.A., T.R. Hayes, A.M. Cranmer, J.C. Turner, and H.M. Davies. 1996. Genetic engineering of a quantitative trait: metabolic and genetic parameters influencing the accumulation of laurate in rapeseed. The Plant Journal 9:229-241.

Voelker, T.A., A.C. Worrell, L. Anderson, J. Bleibaum, C. Fan, D.J. Hawkins, S.E. Radke, and H.M. Davies. 1992. Fatty acid biosynthesis redirected to medium chains in transgenic oilseed plants. Science 257:72-74.

Webb, D., S.J. Knapp, and L. Tagliani. 1992. An RFLP and allozyme linkage map of *Cuphea lanceolata*. Theoretical and Applied Genetics 83:528-532.

An updated genetic linkage map of *Cuphea*

CHAPTER 2

Abstract

Cuphea lanceolata is a temperate source of medium chain fatty acids (MCFAs), as well as a model organism for understanding the biosynthesis of MCFAs in seeds. The goal of this study was to construct an updated genetic linkage map of *C. lanceolata* (2n=2x=12) from recently developed SSRs, converted RFLPs from the previous map, and fatty acid synthesis loci. One hundred seventy six SSRs were screened for polymorphism on a panel of six genotypes (VS55, LN185, LN193-S_1, LN237, PSR23 and PSR108-2). Twenty two fatty acid synthesis loci and 27 RFLPs converted to PCR-based markers (20 SSCPs and 7 SNPs) were screened on four genotypes (VS55, LN185, LN193-S_1, and PSR23). Of 117 SSRs which produced product in most genotypes screened, the mean heterozygosity was 0.73, with a range of 0.278 to 0.889. Eighty eight loci (48 SSRs, 23 SSCPs, 7 SNPs, 10 fatty acid synthesis loci) coalesced into eight linkage groups, with two markers segregating independently. The map was 340.8 cM long with a mean density of 3.9 cM per locus. This map will be a useful tool for *Cuphea* breeders and geneticists. Genome rearrangements and duplications may be detected when maps in *C. viscosissima* and interspecific populations are constructed.

Introduction

Cuphea species have unique seed oil profiles consisting almost entirely of medium chain fatty acids (MCFAs) (Earle et al. 1960, Miller et al. 1964, Thompson 1984, Graham 1988). *Cuphea lanceolata* ssp. *silenoides* and *Cuphea viscosissima* Jacq. are the focus of domestication efforts of *Cuphea* (Earle et al. 1960, Knapp et al. 1991a). The oil of these species ranges from 70 to 83% capric acid (C10:0) (Earle et al. 1960, Knapp et al. 1991a). Chemical mutagenesis has led to the development of lines with reduced capric acid and increased caprylic (C8:0), lauric (C12:0) and myristic (C14:0) acids (Knapp and Tagliani 1991, Knapp 1993).

Commercial uses of MCFAs primarily utilize lauric and myristic acids in products such as soaps, detergents, and surfactants; other uses of MCFAs include as part of a special diet for lipid metabolism disorders, confectionary, and cosmetics (Topfer et al. 1995). Coconut (*Cocus nucifera* L.) and palm kernel (*Elaeis guinenes* Jacq.) oils are the primary natural sources of MCFAs (Thompson et al. 1990) and must be imported by most developed countries. Both *C. lanceolata* and *C. viscosissima* have higher concentrations of all MCFAs, except lauric acid, than coconut and palm kernel oils (Webb et al. 1992). Not only would these higher concentrations reduce the need for purification of individual MCFAs, but *Cuphea* is adapted to temperate regions and will diversify the sources of MCFAs (Webb et al. 1992). Diversification of MCFA supplies should increase the stability of longterm supplies and prices (Webb et al. 1992).

C. lanceolata and *C. viscosissima* are able to produce fertile hybrids, significantly increasing the genetic variation available for exploitation in breeding (Knapp et al. 1991b, Webb et al. 1992). These *Cuphea* species are diploid (x=6, Graham 1989), with a small genome, approximately twice the size of *Arabidopsis thaliana* (unpublished data) making them good model organisms for understanding the mechanism of MCFA biosynthesis in developing seeds (Webb et al. 1992).

The current genetic linkage map of *C. lanceolata* consists of 32 restriction fragment length polymorphism (RFLP) and five allozyme markers (Webb et al. 1992). After this map was constructed, 22 fatty acid synthesis loci were mapped using allele-specific PCR or single strand conformational polymorphism (SSCP) markers in a *C. viscosissima* x *C. lanceolata* F_2 population (Slabaugh et al. 1997). Simple sequence repeat (SSR) markers have recently been developed for *Cuphea* and provide the opportunity to significantly increase the density of the genetic linkage map (Karmarkar 2004). The RFLP markers from the existing map were successfully converted to PCR-based SSCP or single nucleotide polymorphism (SNP) markers and act as anchors for limited comparison of the first and second generation maps. Increasing the density of the genetic linkage map will provide a tool that will assist in the further domestication and breeding of *Cuphea*. Presented here are 1) the identification of polymorphic SSR markers, 2) the development of PCR-based markers from existing RFLP markers, and 3) the construction of a genetic linkage map in *C. lanceolata* utilizing SSR, fatty acid synthesis loci, and RFLP markers converted to PCR-based markers.

Materials and Methods

SSR Marker development

SSR markers were developed by Karmarkar (2004). CA, GA, and GAA repeat enriched libraries were constructed by Genetic Identification Services (GIS, Chatsworth, CA) from DNA isolated from the *C. viscosissima* x *C. lanceolata* interspecific line PSR23 as specified in Karmarkar (2004) using procedures described by Karagyozov et al. (1993) and Edwards et al. (1996). Inserts were sequenced and analyzed for the presence of SSRs using SSRIT software (http://brie2.cshl.org:8082/gramene/searches/ssrtool). Unique sequences were identified and primers (Tm of 59 to 63°C) designed to flank SSRs of $\geq$5 repeats. Expected fragment lengths ranged from 100-500bp for each primer pair. At the 5' end of each forward primer was a 'universal' sequence of 20nt that was complimentary to a 20nt labeled primer for use in nested (three primer) PCR reactions (Schuelke 2000).

SSR Screen Panel of Genotypes

A panel of six genotypes was screened for polymorphism with the SSR markers; the panel included the following genotypes. VS55 (PI534911 a *C. viscosissima* line to gauge the utility of the SSR markers in *C. viscosissima* and is the line which was mutagenized and analyzed for fatty acid mutants), LN185 (PI594934 a *C. lanceolata* line that exhibits partial self-pollination and is a parent of the mapping population), LN193-S_1 (PI594940, is a wild *C. lanceolata* that was selected as the other parent of the mapping population as it is from a different subcluster than LN185

based on a diversity analysis), LN237 (the F_1 of LN185 x LN193-S_1), PSR 23 (an interspecific line that is partially seed retaining (Knapp and Crane 2000)), and PSR108-2 (the F_1 of LN185 x PSR23).

SSR Screening by Polymerase Chain Reaction

Nested PCR was used to screen the SSR markers. The 'universal' primer was fluorescently labeled with HEX, FAM or NED (MWG, High Point, NC) and was complementary to 20 nt at the 5' end of each forward primer. Each 10 µl reaction contained 1.5x buffer, 2.25 mM $MgCl_2$, 0.1 units Taq DNA polymerase (New England Biolabs, Ipswich, MA), 1 µM each of reverse and labeled universal primer, 0.125 µM forward primer, 200 µM dNTPs, and 10 ng of template DNA. A touchdown program was used on an MJ Research PTC-200 thermocycler (Waltham, MA) (Don et al. 1991). After an initial denaturation step at 95˚C for 4 min; a program of 40 cycles was used, consisteng of 7 cycles of touchdown PCR (20 sec at 95˚C, 45 s at 63˚C to 57˚C, 45 s at 72˚C) followed by 33 cycles at a fixed annealing temperature of 57˚C, and a final extension time of 5 min. PCR products were combined with loading buffer and electrophoresed on a 1% agarose gel in 0.5 x TBE.

SSR Screening by Fragment Analysis

Primer pairs that produced amplicons from at least one of the six genotypes were selected for further analysis. One to six amplicons were multiplexed. GS 500 ROX was used as an internal lane standard. The multiplexed samples were run on an ABI377 Genetic Analyzer, filter set D, and analyzed with GeneScan Analysis Version 2.1 (Applied Biosystems Perkin Elmer, Foster City, CA).

Conversion of RFLP Markers to PCR-based Markers

C. lanceolata genomic *Pst*I fragments ranging from 0.5 and 2.0 kb were inserted into pTZ18R plasmids (Webb et al 1992). The inserts were amplified using T3 and T7 primers which flanked the inserts. The amplified insert was bidirectionally sequenced by ABI BigDye Terminator Cycle Sequencing Ready Reaction Kit v3.1 chemistry at the Genomics Center at the University of Nevada at Reno, Nevada. The sequences were compiled into contigs if sequence overlapped sufficiently using VectorNTI. Primers were designed toward the right and left ends of each sequence. PCR amplicon were screened for polymorphism on a panel of four genotypes: LN193-S_1, LN185, VS55 and PSR23 using a single strand conformational polymorphism assay (see below).

Single Strand Conformational Polymorphisms

The fragments from the four member panel (see above) were screened by SSCP in a modification of Slabaugh et al. (1997). Each acrylamide gel contained 0.5X MDE acrylamide (Cambrex Bio Science Rockland Inc., Rockalnd, ME), 0.6X TBE, 0.8 ml 10% APS and 160 μl TEMED, to a final volume of 110 ml. One plate was treated with BindSilane (Sigma, St Louis, MO) and other plate was treated with Rainx® (UNELKO Corp., Scottsdale, AZ). Gels (20cm x 50cm x 1mm) were allowed to polymerize for 2 hr before a pre-run of 30 min at 3.5 watts. Gels were run at constant 3.5 watts at room temperature for 16^{+}hr as described in Slabaugh et al. (1997). Gels were stained with silver nitrate (Sanguinetti et al. 1994). One glass plate was removed after electrophoresis; the second glass plate with affixed gel was soaked

in 1 l fixer solution (210 ml 95% ethanol, 10 ml glacial acetic acid and water) with gentle oscillation for 10 min. The gel was then moved to stain solution (1.5g silver nitrate, 750ml fixer) for 5-10 min. The gel was rinsed with water 2X and then put in 1 l developer (75 ml 10N sodium hydroxide, 1 ml 37% formaldehyde, water) until bands appeared, approximately 15-20 min.

Single Nucleotide Polymorphism Marker Development

For those primer pairs that did not produce polymorphic bands in the SSCP screening the fragments from LN193-S_1 and LN185 were sequenced as above. The sequences were aligned and ChromasPro (Technelysium Pty Ltd, Australia) used to identify single nucleotide polymorphisms. SNP detection primers were designed immediately upstream and downstream of a SNP for each RFLP marker, such that when extended the polymorphic nucleotide was the first nucleotide that would be added, the so called template-directed primer extension described by Chen et al. (1999). Target primers were designed to produce fragments of 200-300 nt with the SNP toward the middle of the fragment (Primer3, Whitehead Institute). Fragments were first amplified with the target primers in a 10 μl reaction containing 1 X PCR buffer, 1.5 mM $MgCl_2$, 50 μM dNTPs, 0.1 μmol each forward and reverse primers, 0.5 U Taq polymerase and 10 ng DNA template. A touchdown PCR program was used to amplify the target fragment. After an initial denaturation step at 94°C for 4 min; a program of 40 cycles was used, consisteng of 5 cycles of touchdown PCR (20 sec at 95°C, 30 s at 63°C to 59°C, 45 s at 72°C) followed by 33 cycles at a fixed annealing temperature of 58°C, and a final extension time of 10 min. Five μl of product was

transferred to a black 384 well plate for PCR cleanup. PCR cleanup was performed according to Xiao et al. (2004). The cleanup included 2 µl 1x Cleanup Reagent (exonuclease and alkalinephosphatase; PerkinElmer, Shelton, CT) and 11% pyrophosphatase (Roche, Indianapolis, IN). The PCR product and cleanup reagent was incubated at 37°C for 60 min, followed by 15 min at 80°C. A fluorescence polarization-template directed incorporation (FP-TDI) assay was used to extend the SNP with a fluorescently labeled acyclo-dNTP (R110 or TAMRA) terminator (Chen et al. 1999). The reaction (20 µl) contained 1x reaction buffer, 1 µl Acyclo Terminator mix, 0.25 µM SNP primer, 0.05 µl AcycloPolymerase (PerkinElmer, Shelton, CT) and 7 µl of processed PCR product. The mixture was brought to 95°C for 2 min, then ~25 cycles of 15 s 95°C and 30 s 55°C. A Victor^2 1420 Multilabel Counter, using the Wallac-1420 Workstation program (PerkinElmer), was used to analyze the polarization ratio of the samples. The fluorescence polarization value (P) is calculated by the equation:

$$P = [Ivv - Ivh]/[Ivv + Ivh]$$

where Ivv is the emission intensity detected when the emission and excitation polarizers are parallel, Ivh is the emission intensity detected when the polarizers are perpendicular to each other (Chen et al. 1999). An Excel macro provided by PerkinElmer was used to evaluate the FP results.

Fatty Acid Synthesis Marker Development

Primers were designed to flank introns of *FatB* based on available *C. lanceolata* and *C. wrightii FatB* cDNA and genomic sequences in the database. The

F178/R179 amplicon included intron 1, the F-B15/R-B13 and F-B20/R-B21 amplicons included introns 2 and 3, and the F180/R181 amplicon included introns 4 and 5. Primers were screened on a panel of *C. lanceolata* and *C. visosissima* genotypes, including LN185, LN237-1 [LN185 x LN193-S_1] F_2 bulk, PSR108-2, PSR23, VS342-6, and [VS342-6 x VL198] F_1. Four aligned *KasIV* genomic sequences from *C. viscosissima* (Slabaugh, unpublished data) were used to design primers, such that one of the primers was gene-specific. R66 and R68 were located in intron 8; R93 and R95 were located in the 3'UTR region of the gene sequence. F22/R8 are non-gene specific primers designed from *C. viscosissima* cDNA sequences (KasII22 and KasII8 in Table 1, Slabaugh et al. 1998). PCR products were combined with loading buffer and electrophoresed on a 1% agarose gel. SSCP analysis was used to identify primer combinations able to detect polymorphism between LN185 and LN193-S_1.

Plant Materials: Mapping Population Development

The genetic linkage map was constructed using 91 LN185 x LN193-S_1 F_2 progeny. LN185 was manually emasculated and hand pollinated to produce the F_1. F_2 seeds were germinated on moistened blotter paper in 11x11x3 clear plastic boxes. Seedlings were transplanted to potting soil (pumice: peat moss: sandy loam) in 1 gallon plastic pots and grown in a greenhouse at 22˚C with 16 hr light per day in Corvallis Oregon in 2003 and 2004. Leaf and bud tissue was harvested from 30 to 80 day old plants and put on ice. Collected tissue was then frozen in a -80˚C freezer before lyophyllization.

DNA Extraction

DNA was isolated using an adapted method after Porebski et al. (1997). Lyophilized leaf and bud material was homogenized with a 2% CTAB lysis buffer using a Polytron (Brinkmann Instruments, Westbury NY). The homogenate was incubated for 30-60 min at 60°C. An equal volume of chloroform:octanol was added, the phases mixed and separated by centrifugation. The aqueous phase was removed to a clean tube and the chloroform:octanol extraction repeated. The DNA was precipitated with 0.2 vol 5 M NaCl and 2 vol 95% cold ethanol and stored overnight at -20°C. DNA pellets were washed with 70% ethanol, the DNA dried and then resuspended in TE. The DNA was treated with RNase A for 1 hr and Proteinase K for 1 hr to remove the RNase. An equal volume of phenol:chloroform was mixed with the DNA solution and the phases separated by centrifugation. The organic phase was back-extracted with TE, the phases separated and the aqueous phases combined. DNA was precipitated with 0.1 vol 3 M Na acetate and 2 vol 100% ethanol and stored at -20°C overnight. After centrifugation, DNA pellets were washed with cold 70% ethanol and air dried. The DNA was dissolved in TE and then quantitated.

Fragment Amplification and Analysis: F_2 Progeny

The forward primer of primer pairs that showed clear polymorphism between LN185 and LN193-S_1 in the screening of SSRs was labeled with HEX or FAM (MWG, High Point NC). These markers were then used to amplify 91 F_2 progeny. Each 10 µl PCR reaction contained 1x buffer, 2.25 mM $MgCl_2$, 0.5 units Taq DNA polymerase, 1 mM each forward and reverse primer, 200 µM dNTPs, and 6ng

template DNA. Fragments were then multiplexed with up to five markers and run on an ABI 3100 DNA Analyzer and analyzed as above (Applied Biosystems Perkin Elmer, Foster City, CA).

Map Construction and Statistics

The genetic linkage map was constructed using MAPMAKER (Lander et al. 1987). The log-likelihood ratio (G) test for segregation distortion was calculated for each locus. The distortion was significantly different from the expected 1:2:1 when $G>\chi^2_{(2,\ 0.01)}$, where G is the log-likelihood ratio test statistic and $\chi^2_{(2,\ 0.01)}$ is a random variable from the χ^2 distribution with two degrees of freedom. A likelihood odds (LOD) threshold of 7.0 and recombination frequency of 0.35 were used to group loci. The MAP function, which compares multipoint likelihoods to estimate locus orders, was used to order the loci. This estimates locus orders based upon the comparison of multipoint likelihoods. The RIPPLE function, which compares the multipoint likelihood for groups of 5 loci, was used to select final marker order. The mean locus density was estimated by summing the total length of each linkage group and dividing by the total number of loci.

Results

SSR, SSCP, and INDEL Marker Polymorphisms

We screened 176 SSR markers (Karmarkar 2004) for length polymorphisms among *C. viscosissima* (VS55), *C. lanceolata* ssp. *silenoides* (LN185 and LN193), and *C. viscosissima* x *C. lanceolata* ssp. *silenoides* (PSR23) germplasm accessions and hybrid individuals used to produce two F_2 mapping population (LN185 x LN193 and LN185 x PSR23). Slightly less than half of the SSR markers (85/176) amplified alleles from VS55, 83 to 95 amplified alleles from LN185 and LN193, and 117 markers which amplified alleles in at least one genotype were subsequently screened for length polymorphisms using fluorescent DNA fragment analysis methods (Ziegle et al. 1992, Schuelke 2000). Of the latter, 50 codominant and 22 dominant SSR markers were polymorphic in the LN185 x LN193 mapping population (61.5%), 56 codominant and 24 dominant SSR markers were polymorphic in the LN185 x PSR23 mapping population (68.4%), 92 were polymorphic in one or both mapping populations (Table 2-1). Heterozygosity (*H*) was calculated for all markers which had fragment size data for at least four of the panel genotypes and ranged from 0.278 to 0.889 (Figure 2-1). Forty-five of the co-dominant markers were fluorescently labeled for use in the construction of a genetic linkage map.

SSCP and INDEL markers developed for genes encoding fatty acid synthesis enzymes (Slabaugh et al. 1997) were screened for polymorphisms among LN185 and LN193 (Table 2-2). Our efforts focused on developing new primer combinations that

Table 2-1. Fragment lengths for SSRs screened on a panel of six genotypes. Fragments were labeled with a tailed universal primer complementary to the 5' tail of each forward primer. Markers in bold are on the linkage map. No amplification indicated by N, M indicates missing data.

Marker	VS55	LN193 S1	LN237-1	LN185	PSR108-2	PSR23
ORC14	129	M	141 146	141	141 144	144
ORC16	138	137 183	137 183	137	136 146	146
ORC17	129	141 146	141 146	141	141 144	144
ORC18	174	146	146 176	155 176	121 155 164	164
ORC19	138	138 141	138	138	138	138 142
ORC20	M	M	M	M	M	M
ORC21	138	136 183	136 183	136	136 146	146
ORC22	138	137 183 190	137 183 190	137	137 146	146
ORC23	M	151	151	151	149 151	149
ORC27	N	144	144	145	145 150	150
ORC31	N	135	120 135	120	120 154	154
ORC34	151	145	145 151	151	151 157	162
ORC37	162	162	162	162	162	166
ORC38	166	152 155	152 169	169	155 169	155
ORC39	146	159	160	138	168	N
ORC40	M	137 148 161	137 148	137 163	137 171	138 163 171
ORC41	169 257	169 255	169 255	169 255	169 255	169 255
ORC43	160	176	172 176	172	172 174	174
ORC47	178	182	178 182	178	178	178
ORC54	N	170	160 170	160	160 189	158
ORC55	154 165	165	165 205 213	165	M	165
ORC58	179	193	173 193	173	173 189	189
ORC60	200 221	206	186 206 210	186 210	186 192	192
ORC61	191	183 185	181 185	181	181 189	189
ORC62	196	N	N	N	196	196
ORC64	M	M	M	M	M	M
ORC65	189	184	M	189	184 189 198	M
ORC70	194	194 241	192 241	192	192 203	203
ORC73	204	194 200	194 205	205	200 205	200
ORC76	170	174 208	174 208	174	174 203	203
ORC78	214	203 210	203 215	215	210 215	210
ORC80	203	193 232	193 232	232	207 232	207
ORC82	191	144 191 205	144 188 191	145 188	188	150
ORC87	211	213	213	211	213	213
ORC89	184	184 198	184	N	211	M
ORC90	204	198	M	192	192 215	215
ORC91	226	226	M	234	216 234	216
ORC92	215	205 303	205 213	213	213 215	215
ORC95	212	243 251	184 251	N	220	220
ORC98	210	212	212	N	222	222

Table 2-1. (Continued)

Marker	**VS55**	**LN193 S1**	**LN237-1**	**LN185**	**PSR108-2**	**PSR23**
ORC100	217	217	202 217	202	202 220	221
ORC102	233	220 229	217 220	217 223	217 222	222
ORC103	206	206 232 238	206 217 232	206 217	206 217 225	206 225
ORC108	146	159	M	N	168	N
ORC110	208	M	206	206	206 236	236
ORC112	220	236	215 236	215	215 232	232
ORC124	252	244	244 246	246	M	258
ORC126	239	243 266	239 243	239	M	252
ORC127	188	193	193	N	251	251
ORC137	260	M	N	M	N	M
ORC143	246 380	245	170 245 250	170 250	250 265	265
ORC145	259 285	262 321	262 271 289 321	271 289	271 277 289 317	277 317
ORC146	264	255	255	255	255 272	272
ORC147	M	M	M	M	262 270	270
ORC153	N	275	242	242	242	M
ORC159	248	M	M	246	M	M
ORC162	242 265	N	242 272	M	242	M
ORC163	M	N	M	N	N	N
ORC164	262	270 272	270 284	284	274 284	274
ORC165	N	313 325	321 325	271 321	272 282	N
ORC166	275	348	293	M	275	275 348
ORC171	266	256 260	254 256	254	254 284	284 287
ORC172	294	282	282 297	M	M	284
ORC173	196	227	227 279	M	196	196
ORC175	N	232	232	N	286	286
ORC176	N	264 270	264 267	267	267 287	287
ORC181	281	283	283	293	293	293
ORC182	242	268	260 268	260	260 292	292
ORC183	296	287	287 297	297	290 297	290 292
ORC187	296	291	288 291	288	288 296	296
ORC189	N	N	M	M	317	317
ORC192	348	301 311	301	N	317	317
ORC196	M	M	N	N	N	N
ORC199	N	317	315 317	315	317 322	317 322
ORC202	N	M	M	N	335	335
ORC204	N	N	N	N	346	346
ORC205	M	315	M	M	M	M
ORC207	M	348	M	M	M	M
ORC208	340	340	336 340	313 336	336 341	341
ORC210	264 334	231 321	231 265 312 321	265 312	312 346	346
ORC212	M	348 371	M	348 366	M	348
ORC213	334	231 321	312 321	312	312 346	346
ORC214	344	329 333	255 329 351 405	351	345	345
ORC215	M	M	M	M	M	M

Table 2-1. (Continued)

Marker	**VS55**	**LN193 S1**	**LN237-1**	**LN185**	**PSR108-2**	**PSR23**
ORC216	335	N	326	326	326 346	346
ORC218	336	351 357	339 351	339	339 354	354
ORC219	N	329	329 433	433	M	433
ORC221	351	383 390	323	323	323 359	359
ORC222	346	349	346	N	355	355
ORC223	386	358	358 382	370 382	350 382	350
ORC224	426	375 386	380 385	367 380	345 357 380	345 357
ORC225	319	329	301 329	301	301 364	364
ORC226	345	354	354	N	M	M
ORC227	343	N	347	N	307 371	307 371
ORC228	418	N	N	N	N	M
ORC229	N	363	363	N	368	368
ORC231	N	372	363 372	363	N	377
ORC239	389	393 417	390 394	390	M	404
ORC241	384	N	389	M	409	409
ORC242	382	346 381	346 381	346 380	346 380 408	346 408
ORC244	N	M	M	N	M	M
ORC245	375	411	411 418	408 416	408 411 416	411 416
ORC246	402	391	382 391	382	382 415	415
ORC247	418	420	414 420	414	414 422	422
ORC251	378	382	378 382	378	378 418	418
ORC252	M	432	422 432	M	422	422
ORC253	362 388	365 425	365 392 425	392	380 392 421	380 421
ORC254	N	398 404	N	407	407 421	421
ORC255	411	415 420	404 415	404	404 422	422
ORC256	414	402	393 402	393	393 426	426
ORC257	N	406 435 441	406 419 435	419	M	428
ORC258	422	387 392 401	401	398	398 431	431
ORC259	N	N	N	N	N	432
ORC260	N	M	N	N	430	430
ORC263	382	387 437	387	387	M	438
ORC264	M	452 460	444 452	444	444 448	448
ORC271	318	315 322	315 318 322	318	315 318 360	315 360

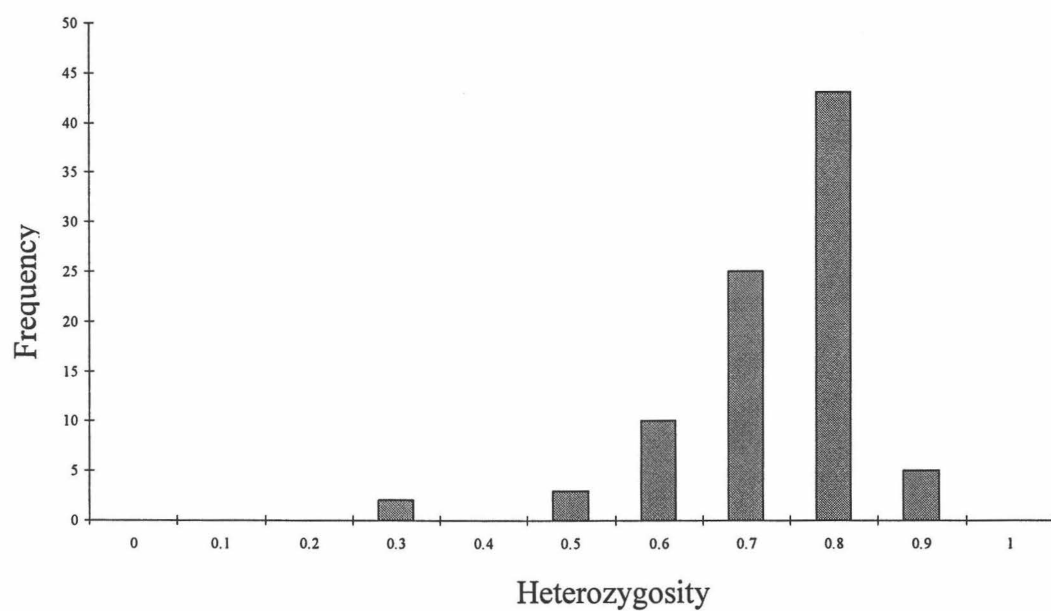


Figure 2-1. Heterozygosities of SSR markers for which allele fragment sizes were available for at least four of the screening panel genotypes.

Table 2-2. Fatty acid synthesis primers used to detect polymorphism between LN185 and LN193.

Name	Sequence	Loci
F-B15	TGACGGCTTTGGTCGTACTCCTG	*FatB1, FatB2, FatB3*
R-B13	TCGGAGTTAGACCCTTGCG	*FatB1, FatB2, FatB3*
F-B20	GACGGCTTTGGTCGTACYCC	*FatB3, FatB4*
R-B21	GTSTTTTGATTCATCATNRCCC	*FatB3, FatB4*
F178	GATCGAACWGCCTCDATAGAGAC	*FatB1, FatB3*
R179	CAGGAGTACGACCAAAGCCGT	*FatB1, FatB3*
F180	TGATGTGAAGACNGGTGATTCCATT	*FatB3*
R181	GGGTCTCWAAAACTTCTRTTGGMA	*FatB3*
F22	TGAATCCCTTTTGTGTACC	*KasIV1, KasIV3, KasIV5*
R8	GTGATAGGCATCGCAAGTG	*KasIV1, KasIV3, KasIV5*
R66	AGCTCATGCAAGAAGATGCA	*KasIV1*
F86	TCVTCYATACTCWTYGCCCCTT	*KasIV2, KasIV4*
R93	ATTGATCAAAGGATCAGCACATTG	*KasIV2*
F84	AATGATGCCATTGAAGCCCTAAGGA	*KasIV3*
R68	GATACATTTAGAAAGCTATGGA	*KasIV3*
R95	CGGCCAATAAGCGGTTTTCAAGAT	*KasIV4*
F88	GGAGCTGGAGTGCTACTACTA	*KasIV4*
R89	TTCCACAGGAAACGCCCTAC	*KasIV4*
F-KasIII37	CCCAACACAAATCCACTGC	*KasIII1*
R-KasIII38	ACCAATGATGAATGGATTTC	*KasIII1*
F-Kar3	GAAATCTCAGTGGCAGGAAG	*Kar27*
R-Kar5	AAACACAAACGTCTGCTCAC	*Kar27*

detected polymorphisms in fatty acyl-ACP thioesterases (FatB) and β-ketoacyl-ACP synthase (KAS) genes, as these genes have a significant effect on the seed oil profile (Dormann et al. 1993, Fuhrmann and Heise 1993, Dehesh et al. 1996a, b and 1998, Leonard et al. 1998). Of the 9 SSCP and INDEL markers for *FatB* and *KasIV* we screened, 4 amplified alleles from a single locus, 5 amplified alleles from two to four loci, and 15 were polymorphic between LN185 and LN193 (Table 2-3). Primers F-KasIII37/R-KasIII38 (Slabaugh et al. 1998) amplified two non-polymorphic loci (*KasIII2* and *KasIII3*) and one polymorphic (*KasIII1*) locus. Primers F-Kar2/R-Kar6 amplified a non-polymorphic *Kar10* locus. Primers F-Kar3/R-Kar5 amplified a polymorphic *Kar27* locus.

SSCP and SNP Markers for Previously Mapped RFLP Markers

We sequenced 30 *Pst*I-genomic DNA clones originally isolated from *C. lanceolata* and previously used as probes for mapping RFLP loci (Webb et al. 1992). The clone inserts were bidirectionally sequenced and aligned and ranged in length from 394 to 1377 bp. DNA sequences for the RFLP probes were used as templates for SSCP marker development and allele resequencing. SSCP markers were developed by targeting forward and reverse primers to opposite ends of the reference allele sequences (Table 2-4). The alleles amplified from each locus ranged in length from 250 to 1208 bp and were monomorphic on agarose. SSCP analyses were used to check the locus specificities of the primers.Five amplified multiple loci and five markers had null alleles (Figure 2-2). Of the 38 SSCP primer sets tested 20 were polymorphic between LN185 and LN193 (Table 2-4). One to three SSCP alleles/locus were

Table 2-3. Amplification of *FatB* and *KasIV* loci in LN185 and LN193-S_1.Bolded fragments are polymorphic in the mapping population. Primers with an * were designed to be gene-specific.

Primer combination	*FatB1*	*FatB2*	*FatB3*	*FatB4*
F178/R179	**600 bp**		**470 bp**	370 bp
B15/B13	**750 bp**	**750 bp**	**600 bp**	
B20/B21	550 bp	510 bp	**460 bp**	**450 bp**
F180/R181			**430 bp**	400 bp

	KasIV1	*KasIV2*	*KasIV3*	*KasIV4*	*KasIV5*
F22/R66*	**1300 bp**				
F86/R93*		**1500 bp**			
F84/R84*			**1000 bp**		
F86/R95*				**600 bp**	
F22/R8	**1350 bp** upper bands		**1200 bp** middle bands		**1000 bp** lower bands

Table 2-4. RFLP inserts were sequenced and primers designed to each end of the insert. Primer combinations were screened on a panel of four genotypes, VS55, PSR23, LN185, and LN193 under SSCP conditions. LN185 and LN193 fragments were sequenced for monomorphic primer combinations and SNP assays developed.

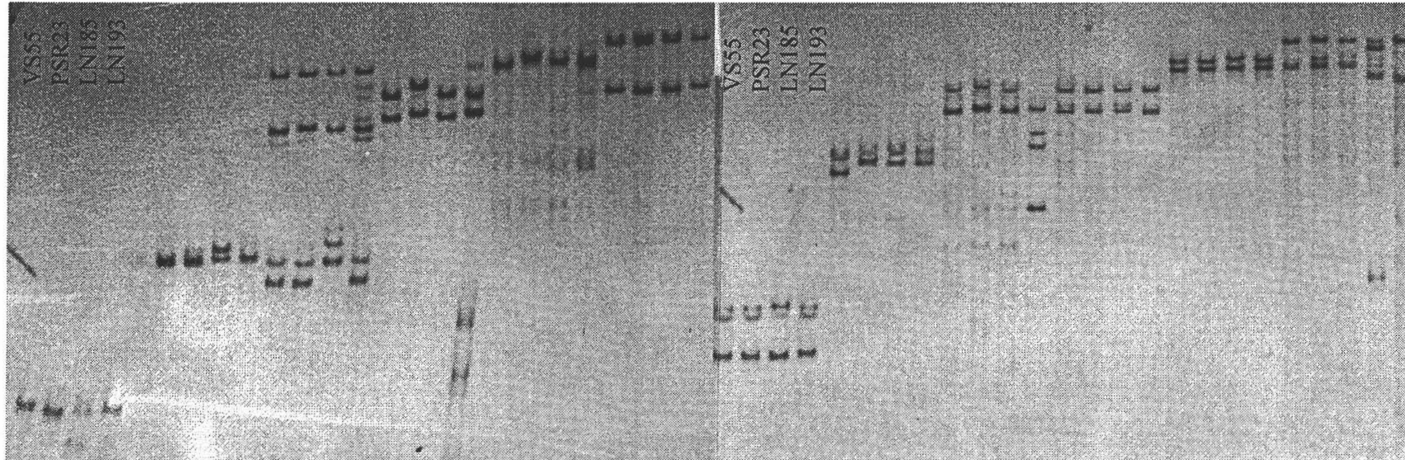
Marker	Insert size*	Primer name	F primer sequence, 5'-3'	R primer sequence, 5'-3'	PCR amplicon size	VS 55	PSR 23	LN 185	LN 193	SNP marker	SSCP marker	Total SSCP loci	SSCP Polymorphic loci
G4	1068	G4F/R	GGACAATTGACGGGCTTAGA	CCGAGTTTCCGACCATGATA	923	Y	Y	Y	Y		x	2	1
G15	966	G15F/R	TGATGCAATCACTGCCATTT	TGAAGGTGGACTGCTCTTCC	776	Y	Y	Y	Y		x	2	2
G26	900*	G26F/R	AGCAGGTTGAAGAAGCCAAA	ACACTATGGCCTCAGCCTTC	731	Y	Y	Y	Y		x	1	1
G30	555	G30F/R	TGCAGCTGTGATATCTTGCTC	TCGACCTGCAGAATCAACC	555	Y	Y	Y	Y	x			
G75	1111	G75F/R	ACTCCACCAAGCAAAAGCAC	GCCTCAATTCGTTCAAATGC	998	Y	Y	Y	Y		x	1	1
G82	1072	G82F/R	TCACCGTACCATTTCAAGCA	AGATTCAAGACACCGGCAAC	840	Y	Y	Y	Y		x	3	2
G100	1000*	G100F/R	TGTGCCAATAAAAGCAGAGTG	TTGGGAAAAGGAATATCAGCA	No amp	N	N	N	N				
G108	1100*	G108F/R	GTCGGGATCCGACAAGCTAT	TGATTATGAATGCGCCATGT	689	Y	Y	Y	Y		x	1	1
G111	1150*	G111fF/R	CCAACAGATTGGATGCATAA	CGCGCTGTCTCTTCTTCTTC	700	Y	Y	Y	Y		x	1	1
		G111rF/R	TGTGTTTATTTTGGAGTGTCCATT	AAAGAAAGGCCCATGAAAGC	250	Y	Y	Y	Y				
G119	483	G119F/R	ACCCCGAATCCCAAATAAAG	GACTTTAGGGCCGTGATGAA	471	Y	Y	Y	Y	x			
G137	1193	G137F/R	GGAGGATCGACCAGTTGCT	CCATTGAAGCCTTTAAGGTGAA	987	Y	Y	Y	Y		x	1	1
G141	1377	G141F/R	TGAGTAGCCAGGACCCTTCA	TCGGGATTTCTCGGTTTGTA	1208	Y	Y	Y	N				
G147	645	G147F/R	TGGCCTCAATCAGGAACAAT	TTGGATTTAAAGGCCGATCA	630	Y	Y	Y	Y		x	1	1
G154	1195	G154F/R	GCCGTGGCTCACAACTCTAC	CAGAGATCGTGCGATGATGA	1078	Y	Y	Y	Y	x			
G167	669	G167F/R	CCGATAAGAGCCCCTTTCA	AAGAACGCTGCAACTTCTCACT	631	Y	Y	Y	Y		x	3	2
G168	555	G168F/R	CGAACTCAAATGGCAACAGAT	TAACTGAGTCTGCCGGCTTC	467	Y	Y	Y	Y	x			
G178	1100*	G178fF/R	TGATGGATGATGAGCTTTATGA	TTTCCAACCTCAGCGAGTTC	687	Y	Y	Y	Y		x	1	1
		G178rF/R	GGGACATGGAATGGTTCTTC	TTTCCTGATCCCGATACCTT	700	Y	Y	Y	Y				
G181	1300*	G181rF/R	AACCATCGAAAAGACACGTT	CAGAACTTCCCACGGTAATC	252	Y	Y	Y	Y		x	1	1
G183	528	G183F/R	ACATGCATAATATATTGCCACCTG	CATAGGGAACCAGCTGAGGA	447	Y	Y	Y	Y	x			
G185		G185fF/R**	AAGGATGCCCAAACAGAATG	TTTTTACCCCCGTTTTCGAT	724	Y	N	Y	Y		x	1	1
		G185rF/R**	TTTATTACAACAAAAGGATTAATGTGC	ACAGCTCACCAGCCCTGA	701	Y	N	Y	Y				

Table 2-4. (Continued)

Marker	Insert size*	Primer name	F primer sequence, 5'-3'	R primer sequence, 5'-3'	PCR amplicon size	VS 55	PSR 23	LN 185	LN 193	SNP marker	SSCP marker	Total SSCP loci	SSCP Polymorphic loci
G188	1431	G188F/R	GTCGGGCAAACTATGAATGC	GAAAAGGATGCCCAAACAGA	1189	Y	Y	Y	Y		x	1	1
G190	1300*	G190fF/R	GAGGAATCTTGCTGCTGGTT	GGCTAAGACAAAAATCAGAACAAAC	500	Y	Y	Y	Y		x	1	1
		G190rF/R	CATTGGCTACCGTGTTTCCT	GATTAATTGACGGCATCTTATCG	659	N	N	N	N				
G194	469	G194F/R	TGCCACAGAAAGTGGAAACC	TCACGAGGGCAAATCATAGAC	431	Y	Y	Y	Y		x	3	1
G195	538	G195F/R	CAAACCCAACGAAATGGAAG	ATTGGTCGTGACCGAAAGAG	443	Y	Y	Y	Y		x	1	1
G197	1149	G197F/R	TTCTTCCTCCCCCTTAGTCG	CAAGAGGCTTCTGCTGCTG	1041	Y	Y	Y	Y				
G198	1012	G198F/R	TGCATCGATTAGGCACGAC	CCCATTTGGATGTTCATCGT	898	Y	Y	Y	Y		x	1	1
G199	1100*	G199fF/R	GCCTGAGGCTGAGGTTATTG	TCTCTCTCAATTAGAGGATAGGTGA	600	Y	Y	Y	Y	x			
		G199rF/R	AGTGCAAAGGACCACGTCA	ACGGTATGTGCATTCCTGGT	726	N	Y	N	Y				
G210	1200*	G210fF/R	GCTGCTTCAGACCTTTCAGC	TGCCGCTTGACACATTAAGA	727	Y	N	Y	Y				
		G210rF/R	TGCAGTGGAAGTAAAGATTTGG	CGACTCAGCATTGGAGGAAC	478	Y	Y	Y	Y	x			
G211	1100*	G211fF/R	CGACTGTCCTTGCAGGTTAAG	GAGCTCGTTTGGGAACTCTG	716	Y	Y	Y	Y				
		G211rF/R	TGCTCAGGGAGATATGCTTTT	GAGTTCAGCCCGAAGGAAG	695	Y	Y	Y	Y		x	1	1
G213	394	G213F/R	GCCTTAGCGTCTCCTTGACA	GGGCTTCGACAGAGAAAACA	328	Y	Y	Y	Y		x	1	1
					Totals	34	32	35	34	7	20	28	23

*Insert sizes followed by * are estimates from gels

**Each primer set was dominant for one parent



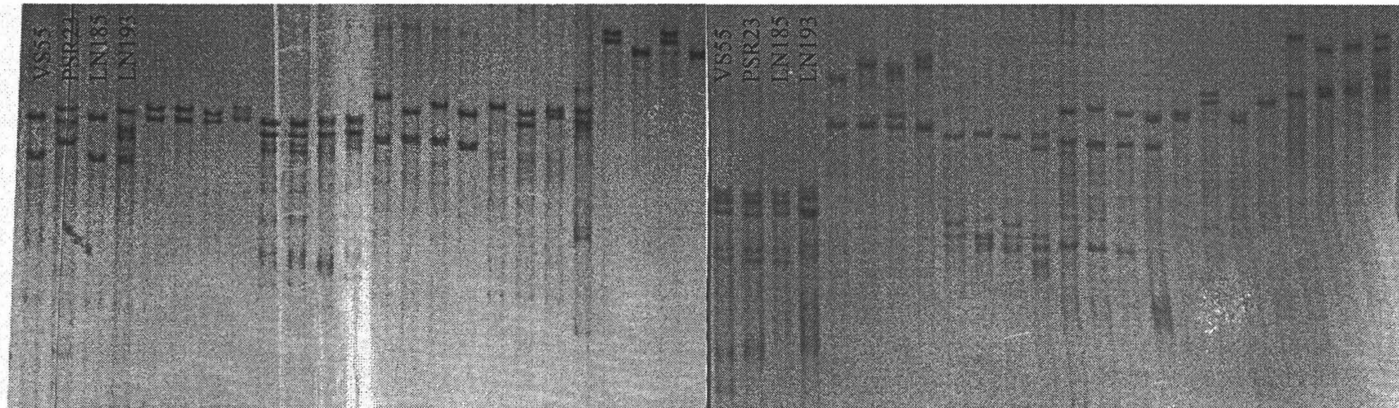


Figure 2-2. Primers designed to span the RFLP insert fragments were screened across four genotypes VS55, PSR23, LN185 and LN193 and assayed under SSCP conditions.

observed among the four germplasm accessions. SSCP markers for seven of the original 32 loci were monomorphic. We amplified and sequenced LN185 and LN193 alleles from the latter using reference allele sequences as templates. The allele sequences ranged in length from 447 to 726 bp and were aligned with the reference allele sequence to identify SNPs. We surveyed 3278 bp across the seven sequence-tagged-sites and identified 64 SNPs between LN185 and LN193 alleles. LN193 was heterozygous at 34 of the SNP sites, while LN185 was heterozygous at only one. The average SNP frequency was 1/60 bp. We developed FP-TDI SNP assays (Chen et al. 1999) by targeting SNPs identified between LN185 and LN193 alleles (Figure 2-3, Table 2-5).

Genetic Mapping

We genotyped 52 codominant SSR markers, 11 SSCP markers for previously mapped fatty acid synthesis loci (Slabaugh et al. 1997), and 20 SSCP and seven SNP markers for previously mapped RFLP marker loci (Webb et al. 1992) in the LN185 x LN193 F_2 mapping population. The segregation ratios for 36 of the 90 loci were significantly distorted (Figure 2-4). The LN185 allele frequency was significantly less than 0.5 for 34 loci on linkage groups 1, 2, 4, and 5, whereas the LN193 allele frequency was significantly less than 0.5 for 2 loci on linkage group 3. Heterozygote frequencies were significantly greater than 0.5 for 22 loci on linkage groups 1, 2, 4, and 5 and significantly less than 0.5 for 1 locus on linkage group 3. Using a LOD threshold of 7 and a maximum recombination frequency of 0.35, 88 of 90 DNA marker loci coalesced into eight linkage groups (Figure 2-5). Two linkage groups (LGs) only had two loci each (LG 7 and 8), a third linkage group (LG 6) had only three loci, and two

Table 2-5. SNP markers used in mapping LN185 x LN193.

Marker	Target length	Target primers	SNP primer	SNP target	Acyclo mix	Primer sequence
G30	305	G30_12p1				TACACTGAACTACCAGTTTCTCC
		G30_12p2				AATTTTCAGTAAGGACTTCATCA
			G30_1p4	snp286	G/A	CCATGCTTTGTTCTTCAAGTGTT
G119	471	G119F				ACCCCGAATCCCAAATAAAG
		G119R				GACTTTAGGGCCGTGATGAA
			G119_1p4	snp93	G/A	TGCTGCCAGTACTGTCGATC
G154	293	G154snp_p1				TGCAACATGTGATGTTTGGT
		G154snp_p2				TTCGTACATGTGAACGGATG
			G154snp389_p4	snp389	G/A	CAGGAGTTTGCCCTGCCTC
G168	248	G168_2p3				CCTTATCTTCCGGTGCTCC
		G168snp439p4				CTGCCGGCTTCCACTTGAC
			G168snp430p3	snp430	G/A	GTCTTTGTGAGGGGAAGGACA
G178	293	G178_2p1				TAGGAAGGTAAATCCAAAATGAT
		G178_2p2				CTGCACAGAAAGAACATAATACC
			G178_2p3	snp374	G/C	TATCGGGTCTGCAAATCAGGT
G183	447	G183F				ACATGCATAATATATTGCCACCTG
		G183R				CATAGGGAACCAGCTGAGGA
			G183snp250p4	snp250	C/A	AGACCATAATCTTACCACTGATTC
G199	726	G199fF				GCCTGAGGCTGAGGTTATTG
		G199fR				TCTCTCTCAATTAGAGGATAGGTGA
			G199f_snp346p3	snp346	G/A	CTCTGAAGGCTCAAGAAAAGCCA
G210	478	G210rF				TGCAGTGGAAGTAAAGATTTGG
		G210rR				CGACTCAGCATTGGAGGAAC
			G210snp150_p4	snp150	G/A	CTAACAATCAGATCAGATCACATC

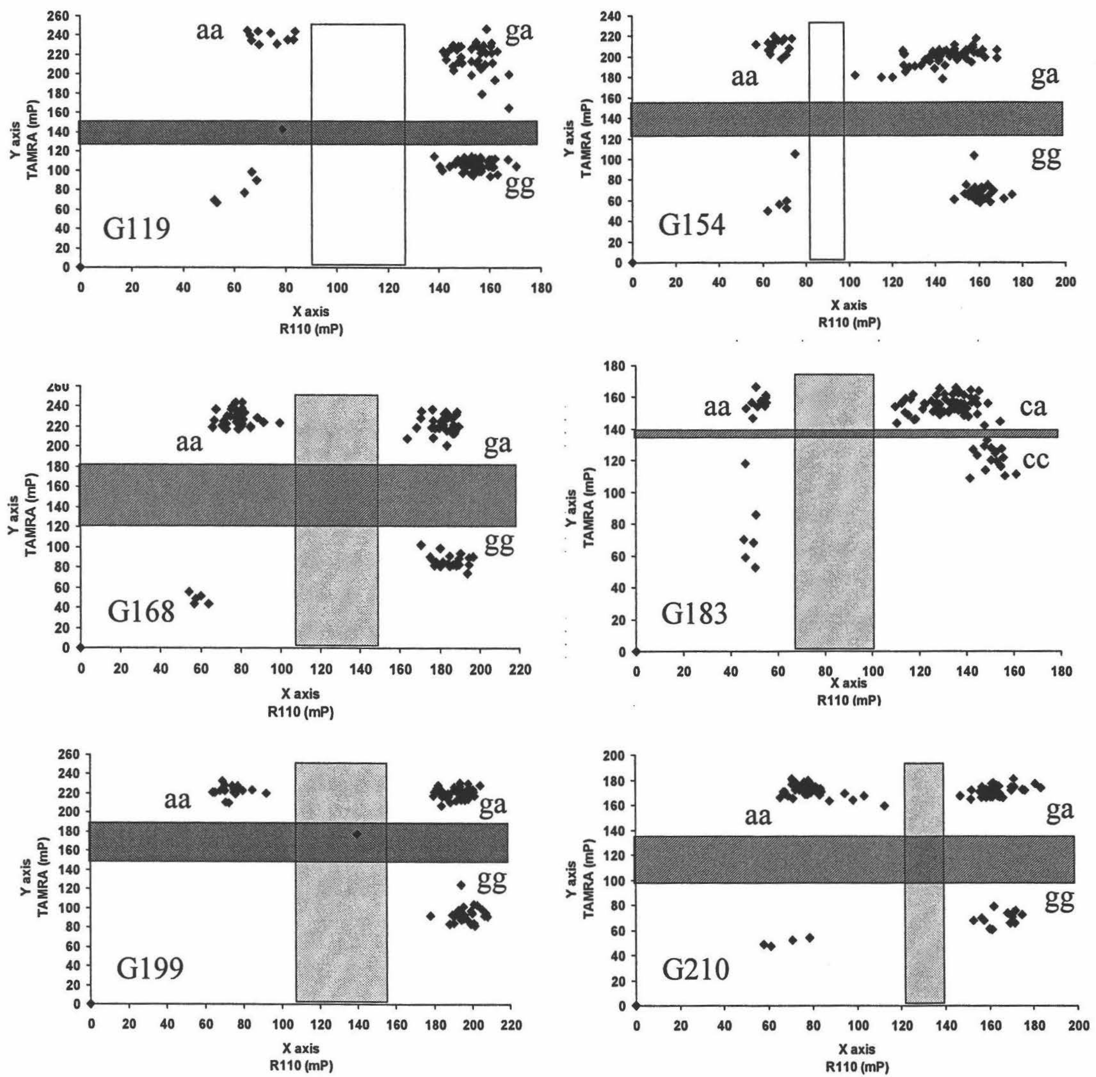


Figure 2-3. Scatter plots from six SNP FP-TDI assays. Genotypes are indicated in the appropriate quadrant. Negative controls and PCR failures are in the lower left quadrant.

of the 90 DNA marker loci (*KASIII1* and ORC252) were unlinked. Six linkage groups were populated with 3 to 20 DNA marker loci each and ranged in length from 1.1 cM (LG 6) to 99.8 cM (LG 3). The mean number of loci/LG was 10.9. The map was 340.8 cM long and had a mean density of 3.9 cM/locus. The marker densities ranged from 0.4 cM/locus (LG 6) to 10.2 cM/locus (LG 8). The largest gap was 26.7 cM between ORC21 and ORC14 on LG 3.

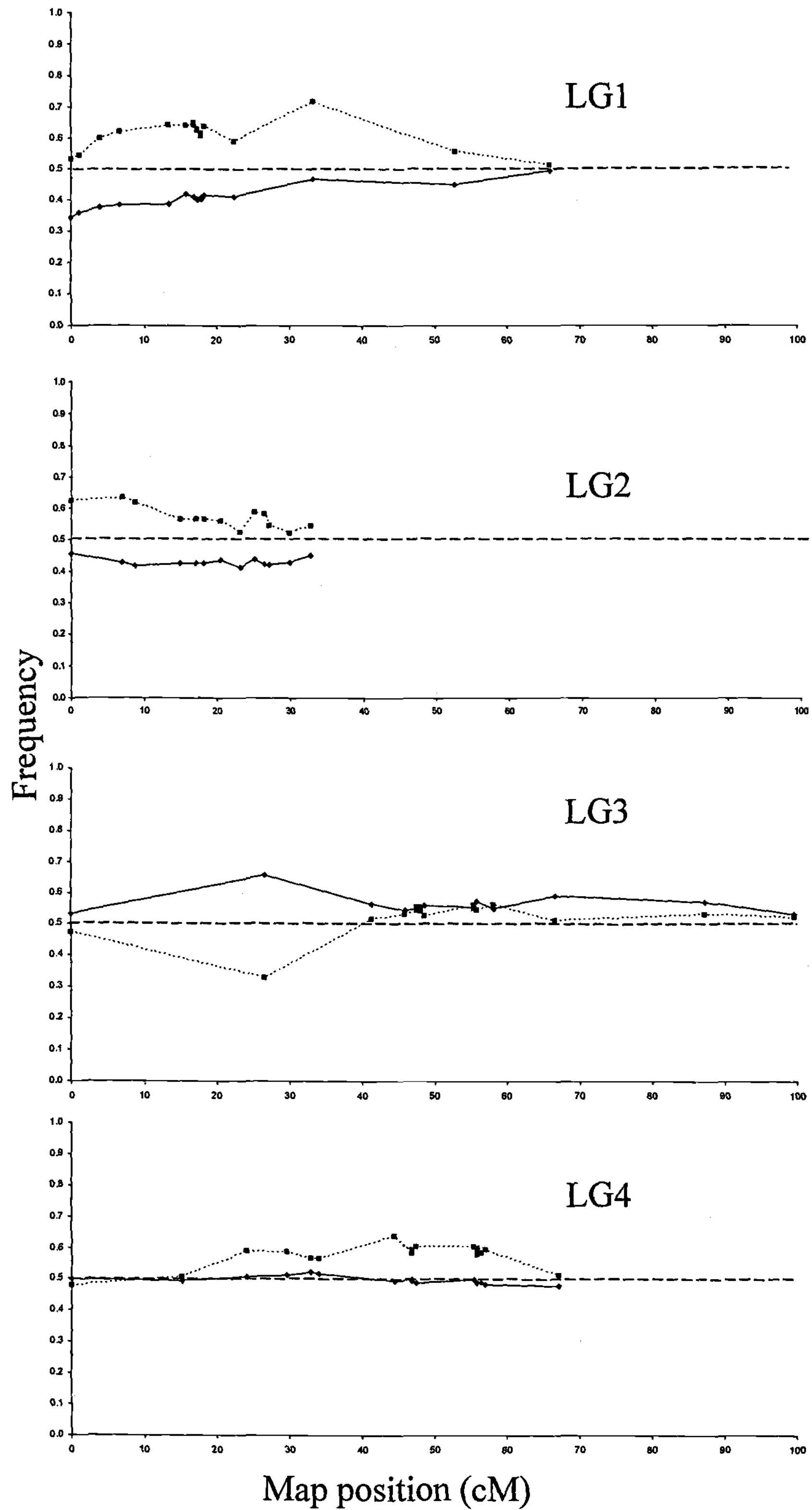


Figure 2-4. LN185 allele frequency (solid line) and heterozygote frequency (dashed line) among LN185 x LN193-S_1 F_2 progeny for ordered loci on the linkage map of *C. lanceolata*.

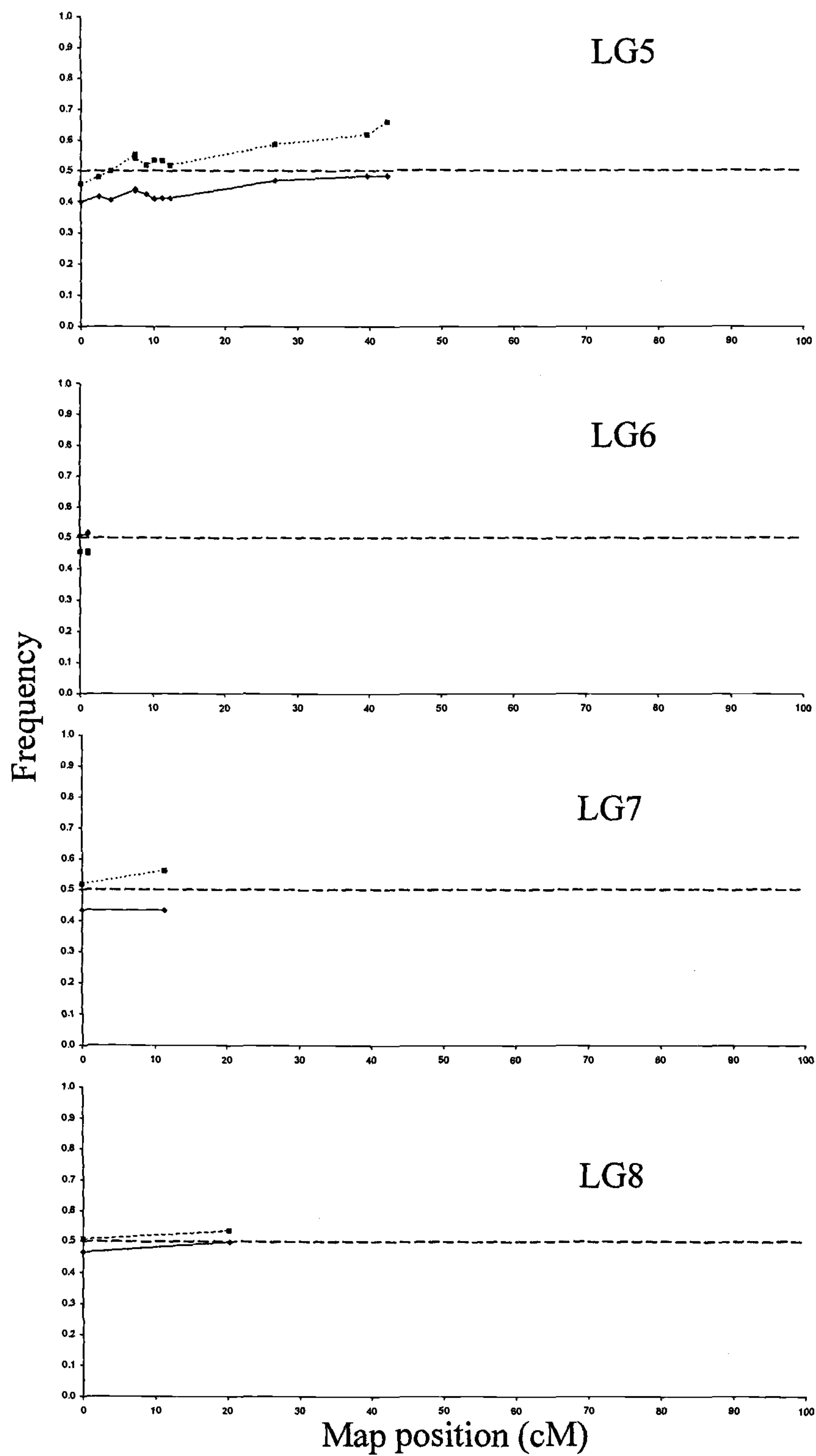
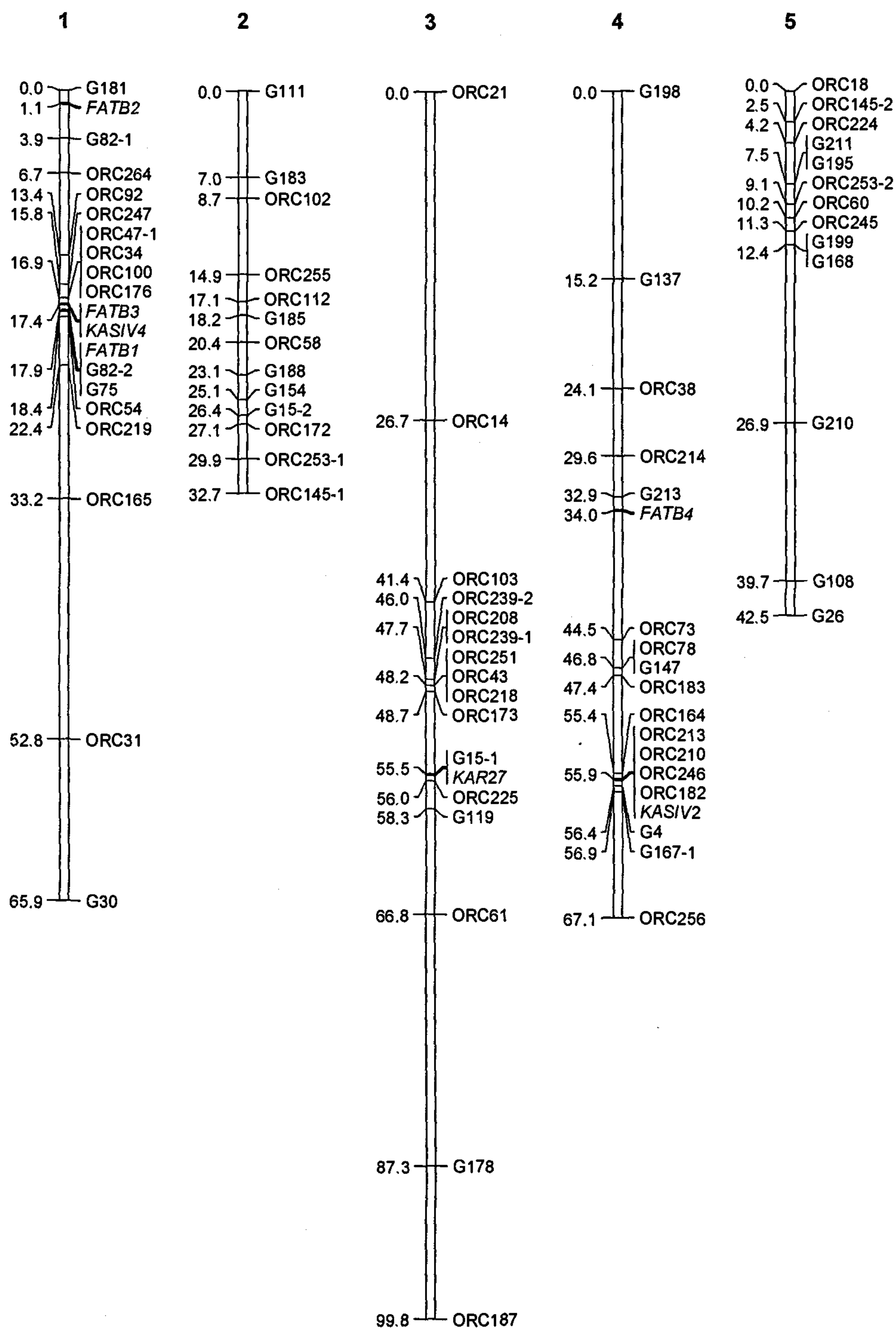


Figure 2-4. (Continued)

Figure 2-5. Linkage map of LN185 x LN193-S_1. ORC markers are SSRs, G markers are derived from previously mapped RFLPs, and italicized markers are fatty acid synthesis genes.



Discussion

The utility of SSR markers developed for *Cuphea* were assessed on a panel of six genotypes, which included one *C. viscosissima*, four *C. lanceolata* lines, and one interspecific line. The panel included one *C. lanceolata* F_1 genotype, LN237-1 (LN185 x LN193-S_1) and an interspecific F_1, PSR108-2 (LN185 x PSR23). Of the 176 markers screened for amplification of product, 117 (66.4%) produced amplicons in at least one genotype and were further assessed for labeled fragment size (Table 2-1). More than half these 117 SSR markers were able to detect alleles from each parent in the F_1 genotypes LN237-S_1 (56.9%) and PSR108-2 (51.8%) (Table 2-1). SSRs showed cross taxa amplification, as 74% (87) of the 117 markers that amplified alleles amplified *C. viscosissima* alleles (Table 2-1). SSR markers proved to be highly polymorphic and informative in *C. lanceolata* and *C. viscosissima*, as the mean heterozygosity was 0.73 for those markers for which at least four of the panel genotypes produced fragment size data (Figure 2-1). The hypervariability of SSRs makes them very informative, particularly in comparison to other markers such as RFLPs (Wu and Tanksley 1993). As SSRs are PCR-based markers and the fragments can be multiplexed for analysis, they are 'ideal' markers for constructing genetic linkage maps and then be utilized for marker assisted selection (MAS) or quantitative trait loci (QTL) analysis (Liu et al. 1996). Second generation linkage maps have been constructed with a backbone of SSR markers in a variety of species, including barley (Ramsay et al. 2000), sunflower (Tang et al. 2002), potato (Milbourne et al. 1998), and

rice (Temnykh et al. 2000). SSRs were thus used to increase the density of the existing *C. lanceolata* map.

The cross taxa utility of these SSR markers will be critical in future mapping efforts in both *C. viscosissima* and interspecific populations, as these markers will be a source of common markers among the three maps, which will allow for genome rearrangements and duplication to be assessed. As the currently cultivated form of *Cuphea* is the result of an interspecific cross (Knapp and Crane 2000), these maps will allow for more targeted and rapid domestication and breeding efforts. Of the 48 SSRs genotyped on the mapping population, 41 (87.5%) amplified *C. viscosissima* as well (Table 2-1). These markers will provide a starting point for the future *C. viscosissima* map and *C. viscosissima* x *C lanceolata* interspecific map.

The map presented here is 340.8 cM long, 52.8 cM longer than the previous map, and the mean density per locus decreased from 7.8 cM per locus to 3.9 cM per locus. This indicates an increase in genome coverage. We report here the mapping of fatty acid synthesis loci in a *C. lanceolata* population for the first time, as previous work in our lab utilized an interspecific F_2 population (Slabaugh et al. 1997). The inclusion of these loci on the map will be helpful to breeding efforts, as the fatty acid synthesis loci may be an indicator to the phenotype. This would allow for genetic screening at an earlier stage than the currently required phenotypic screening, expediting the breeding process. Additionally, the mapping of converted RFLP markers from the first *C. lanceolata* map shows that not all the linkage groups held together in the new map. RFLP markers can be successfully converted to PCR-based markers, here either SSCP or SNP markers. Conversion to more user-friendly PCR-

based markers is preferable. Marker order was conserved on LG1 (old LG5: G75, G82, G181-2), LG4 (G198, G137, G213, G147, G4, G167-1, two markers G190 and G26 moved to other linkage groups). The PCR-based markers did not necessarily map to the same location as the RFLP markers. The markers on the previous linkage groups 2 and 3 did not remain linked on the present map, while some of the markers on linkage groups 1 (to LG5) and 4 (LG4) did remain linked on the present map. The conservation of some marker order is reassuring, as correct marker order is key to effective MAS and QTL work. The converted markers may map to different location for a variety of reasons. SNPs assess a single nucleotide difference, while the RFLP markers they were derived from assess a polymorphism based on fragment size and restriction sites. There is no guarantee that a SNP will score the same locus as the RFLP it was derived from, thus movement of these markers is not unexpected. The RFLP map is also quite sparse, increasing the number of markers nearly threefold, one would expect that linkage groups may change substantially.

Fatty acid synthesis loci were previously mapped in an interspecific population (Slabaugh et al. 1997) making polymorphism detection more likely; thus, additional primer development was necessary to detect polymorphisms between *C. lanceolata* lines (Tables 2-2 and 2-3). We focused our efforts on detecting polymorphisms in *FatB* and *KasIV* genes, as they are both implicated in determination of MCFA seed oil profiles (Dormann et al. 1993 Fuhrmann and Heise 1993, Dehesh et al. 1996a, b and 1998, Voelker 1996, Leonard et al. 1997 and 1998), as well as being candidate genes for mutations in *C. viscosissima.* Three of the *FatB* loci (1, 2 and 3) are found on linkage group one (Figure 2-5), similar to the interspecific mapping of these loci,

though the order was not preserved (Slabaugh et al. 1997). *Kar27* was in the same linkage group as these *FatB*s previously, but here was the only fatty acid synthesis locus on its linkage group (LG3, Figure 2-5). The *KasIV* loci (designated *KasII* in Table 1 in Slabaugh et al. 1997) segregated independently in previous work, while here *KasIV-4* is tightly linked to *FatB3* and *FatB1* (Figure 2-5). Clustering of fatty acid synthesis loci was observed in the interspecific population, with all but one gene family being represented on one linkage group, suggesting all the necessary genes for MCFA synthesis are in one region of the genome (Slabaugh et al. 1997).

Six linkage groups that would correspond to the six chromosomes of *C. lanceolata* were expected at the outset of map construction. Instead, five large linkage groups (LG1 to LG5) and three small ones (LG6 to LG8) coalesced (Figure 2-5). The five large linkage groups likely correspond to five of *C. lanceolata*'s six chromosomes. The remaining three linkage groups indicate there are regions of the genome not yet mapped. A single physically small chromosome was not detected by karyotypic analysis (Chen and Roath 1995), suggesting that other causes must be responsible. The remaining chromosome may be highly conserved and monomorphic, thus not mapped as linkage mapping depends on the detection of polymorphisms.

SSRs helped to greatly increase the density of the *Cuphea* map. They also amplified products in both the species used for domestication, *C. lanceolata* and *C. viscosissima*. The map presented here represents a new tool for *Cuphea* researchers and breeders. The increased density and addition of fatty acid synthesis loci make for a higher utility map. The markers on this map are all PCR-based, making them superior for use in MAS and QTL work.

Works Cited

Chen, W., and W.W. Roath. 1995. Karyotype of *Cuphea lanceolata* Ait. and *Cuphea viscosissima* Jacq. Crop Science 35:246-250.

Chen, X., L. Levine, and P.-Y. Kwok. 1999. Fluorescence polarization in homogeneous nucleic acid analysis. Genome Res. 9:492-498.

Dehesh, K., A. Jones, D.S. Knutzon, and T.A. Voelker. 1996a. Production of high levels of 8:0 and 10:0 fatty acids in transgenic canola by overexpression of Ch FatB2, a thioesterase cDNA from *Cuphea hookeriana*. Plant J 9:167-172.

Dehesh, K., P. Edwards, T. Hayes, A.M. Cranmer, and J. Fillatti. 1996b. Two novel thioesterases are key determinants of the bimodal distribution of acyl chain length of *Cuphea palustris* seed oil. Plant Physiol. 110:203-210.

Dehesh, K., P. Edwards, J. Fillatti, M. Slabaugh, and J. Byrne. 1998. KAS IV: a β-ketoacyl-ACP synthase from*Cuphea* sp. is a medium chain specific condensing enzyme. The Plant Journal 15:383-390.

Don, R., P. Cox, B. Wainwright, K. Baker, and J. Mattick. 1991. 'Touchdown' PCR to circumvent spurious priming during gene amplification. Nucleic Acid Research 19:4008.

Dormann, P., F. Spener, and O.J. B. 1993. Characterization of two acyl-acyl carrier protein thioesterases from developing *Cuphe*a seeds specific for medium-chain- and oleoyl-acyl carrier protein. Planta 189:425-432.

Earle, F.R., C.A. Glass, G.C. Geisinger, I.A. Wolff, and Q. Jones. 1960. Search for new industrial oils IV. Journal of the American Oil Chemists' Society 67:440-447.

Edwards, K., J. Barker, A. Daly, C. Jones, and A. Karp. 1996. Microsatellite libraries enriched for several microsatellite sequences in plants. BioTechniques 20:758-760.

Fuhrmann, J., and K. Heise. 1993. Factors controlling medium-chain fatty acid synthesis in plastids from maturing *Cuphea* embryos. Z. Naturforsch 48c:616-622.

Graham, S. 1988. Revision of *Cuphea* section *Heterodon* (Lythraceae). Systematic Botany Monographs 20:1-168.

Graham, S. 1989. Chromosome numbers in *Cuphea* (Lythraceae): new counts and a summary. American Journal of Botany 76:1530-1540.

Karagyosov, L., I. Kalcheva, and V. Chapman. 1993. Construction of random small-insert genomic libraries highly enriched for simple sequence repeats. Nucleic Acid Research 21:3911-3912.

Karmarkar, V.M. 2004. Simple sequence repeat marker development and fingerprinting in *Cuphea lanceolata* Ait. and *C. viscosissima* Jacq. Masters, Oregon State University, Corvallis.

Knapp, S.J. 1993. Modifying the seed storage lipids of *Cuphea*: a source of medium-chain triglycerides, p. 142-154, In S. MacKenzie and D. Taylor, eds. Seed oils for the future. American Oil Chemists' Society, Champaign IL.

Knapp, S.J., and J.M. Crane. 2000. Registration of reduced seed-shattering *Cuphea* germplasm PSR23. Crop Science 40:299-300.

Knapp, S.J., and L. Tagliani. 1991. Two medium-chain fatty acid muntants of *C. viscosissima*. Plant Breeding 106:338-341.

Knapp, S.J., L. Tagliani, and B. Liu. 1991a. Outcrossing rates of experimental population of *Cuphea lanceolata*. Plant Breeding 106:334-337.

Knapp, S.J., L. Tagliani, and W.R. Roath. 1991b. Fatty acid and oil diversity of *Cuphea viscosissima*: a source of medium-chain fatty acids. Journal of the American Oil Chemists Society 68:515-517.

Lander, E., P. Green, J. Abrahamson, A. Barlow, M. Daly, S. Lincoln, and L. Newverg. 1987. MAPMAKER: an interactive computer package for constructing primary linkage maps of experimental and natural populations. Genomics 1:174-181.

Leonard, J.M., S.J. Knapp, and M.B. Slabaugh. 1998. A *Cuphea* β-ketoacyl-ACP synthase shigts the synthesis of fatty acids towards shorter chains in *Arabidopsis* seeds expressing *Cuphea* FatB thioesterases. The Plant Journal 13:621-628.

Leonard, J.M., M.B. Slabaugh, and S.J. Knapp. 1997. *Cuphea wrightii* thioesterases have unexpected broad specificities on saturated fatty acids. Plant Molecular Biology 34:669-679.

Liu, Z., R. Biyashev, and M. Saghai Maroof. 1996. Development of simple sequence repeat DNA markers and their integration into a barley linkage map. Theoretical and Applied Genetics 93:869-876.

Milbourne, D., R.C. Meyer, A.J. Collins, L.D. Ramsay, C. Gebhardt, and R. Waugh. 1998. Isolation, characterization and mapping of simple sequence repeat loci in potato. Molecular Genetics and Genomics 259:233-245.

Miller, R., F.R. Earle, I.A. Wolff, and Q. Jones. 1964. Search for new industrial oils IX *Cuphea*, a versatile source of fatty acids. Journal of the American Oil Chemists Society 41:279-280.

Porebski, S., L.G. Bailey, and B.R. Baum. 1997. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. Plant Molecular Biology Reporter 15:8-15.

Ramsay, L.D., M. Macaulay, S. degli Ivanissevich, K. MacLean, L. Cardle, J. Fuller, K. Edwards, S. Tuvesson, M. Morgante, A. Massari, E. Maestri, N. Marmiroli, T. Sjakste, M. Ganal, W. Powell, and R. Waugh. 2000. A simple sequence repeat-based linkage map of barley. Genetics 156:1997-2005.

Sanguinetti, D., E. Neto, and A. Simpson. 1994. Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. BioTechniques 17:915-919.

Schuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology 18:233-234.

Slabaugh, M.B., G.M. Huestis, J. Leonard, J.L. Holloway, C. Rosato, V. Hongtrakul, N. Martini, R. Toepfer, M. Voetz, J. Schell, and S.J. Knapp. 1997. Sequence-based genetic markers for genes and gene families: single-strand conformational polymorphisms for the fatty acid synthesis genes of *Cuphea.* Theoretical and Applied Genetics 94:400-408.

Slabaugh, M.B., J.M. Leonard, and S.J. Knapp. 1998. Condensing enzymes from *Cuphea wrightii* associated with medium chain fatty acid biosynthesis. Plant J 13:611-620.

Tang, S., J.-K. Yu, M.B. Slabaugh, D. Shintani, and S.J. Knapp. 2002. Simple sequence repeat map of the sunflower genome. Theoretical and Applied Genetics 105:1124-1136.

Temnykh, S., W.D. Park, N. Ayres, S. Cartinhour, N. Hauck, L. Lipovich, Y.G. Cho, T. Ishii, and S.R. McCouch. 2000. Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa*). Theoretical and Applied Genetics 100:697-712.

Thompson, A.E. 1984. *Cuphea*-a potential new crop. HortScience 19:352-354.

Thompson, A.E., D. Dierig, S.J. Knapp, and R. Kleiman. 1990. Variation in fatty acid content and seed weight in some lauric acid rich *Cuphea* species. Journal of the American Oil Chemists Society 67:611-617.

Topfer, R., N. Martini, and J. Schell. 1995. Modification of Plant Lipid Synthesis. Science 268:681-686.

Voelker, T.A. 1996. Plant acyl-ACP thioesterases: chain-length determining enzymes in plant fatty acid biosynthesis, p. 111-133, In J. Setlow, ed. Genetic Engineering, Vol. 18. Plenum Press, New York.

Webb, D., S.J. Knapp, and L. Tagliani. 1992. An RFLP and allozyme linkage map of *Cuphea lanceolata*. Theoretical and Applied Genetics 83:528-532.

Wu, K., and S. Tanksley. 1993. Abundane, polymorphism and genetic mapping of microsatellites in rice. Molecular Genetics and Genomics 245:187-194.

Xiao, M., A. Phong, K.L. Lum, R.A. Greene, P.R. Buzby, and P.-Y. Kwok. 2004. Role of excess inorganic pyrophosphate in primer-extension genotyping assays. Genome Research 14:1749-1755.

Ziegle, J.S., Y. Su, K.P. Corcoran, L. Nie, P.E. Mayrand, L.B. Hoff, L.J. McBride, M.N. Kronick, and S.R. Diehl. 1992. Application of automated DNA sizing technology for genotyping microsatellite loci. Genomics 14:1026-1031.

Fatty acid mutant candidate gene analysis in *Cuphea*

CHAPTER 3

Abstract

Medium chain fatty acids (MCFAs) are important components of soaps, detergents and surfactants. An interspecific *Cuphea* line derived from *C. lanceolata* x *C. viscosissima* is being domesticated as a temperate source of MCFAs. Both of these *Cuphea* species are rich in capric acid (C10:0) and there is little natural variation in the fatty acid phenotypes in *C. viscosissima.* EMS mutagenesis has been used to create novel fatty acid profiles. Two mutants, *cpr-1* and *cei-2*, found to reduce C10:0 and C8:0, respectively, were characterized. Single nucleotide polymorphisms (SNPs) in candidate genes were identified based on comparison of DNA sequences from mutant and wild type *C. viscosissima.* An F_2 population of 92 individuals segregating for the two mutations was phenotyped for fatty acids by quarter seed analysis on a gas chromatograph and genotyped for the two SNPs using DNA prepared from remnant seed tissue. The SNP in the candidate gene for *cpr-1*, *CvKasIV-2*, correlated with a decrease in C10:0. The SNP in the candidate gene for *cei-2*, *CvFatB3*, correlated with a decrease in C8:0. The double homozygous mutant increased myristic acid (C14:0) >30-fold relative to wild type *C. viscosissima.* The combination of these two mutations produced a novel oil profile in domesticated *Cuphea.*

Introduction

Medium chain fatty acids (MCFAs) are used in surfactants, soaps, detergents, pharmaceuticals, cosmetics and confectionary (Topfer et al. 1995). Lauric (C12:0) and myristic (C14:0) acids are the primary MCFAs used by industry. Palm kernel (*Elaeis guinenes* Jacq.) and coconut (*Cocus nucifera* L.) are the currently utilized renewable sources of MCFAs and must be imported by temperate countries. *Cuphea* (Lythraceae) species have been identified as temperate sources of MCFAs and are currently being domesticated (Thompson 1984, Graham 1988, Webb et al. 1992, Knapp 1993, Gesch et al. 2003). *Cuphea viscosissima* and *Cuphea lanceolata* are the focus of US breeding and domestication efforts, as the two species produce fertile interspecific progeny (Webb et al. 1992, Brandt and Knapp 1993, Knapp 1993).

VS335 is a *C. viscosissima* line with a fatty acid profile that deviates dramatically from wild type, in that myristic acid is increased from 1% to >30% while the normally predominant C8:0 and C10:0 fatty acids are reduced. The high myristic phenotype is an important advance in generating fatty acid diversity in domesticated *Cuphea* as C14:0 is an important MCFA feedstock for industrial purposes. The ability to breed for high C14:0 phenotypes would be facilitated by DNA based markers tightly linked to the genes controlling this phenotype.

VS335 was identified in a mutagenized population derived from ethyl methanesulfonate (EMS) treatment of VS-CPR-1 (Crane, unpublished data). VS-CPR-1 is a line derived from EMS treatment of wild type *C. viscosissima* (Knapp and Tagliani 1991, Knapp 1993). VS-CPR-1 is homozygous for the *cpr-1* mutation that

reduces capric acid (C10:0) from ~70% to ~40% and increases caprylic acid (C8:0) from 18% to ~25% (Knapp 1993). The *cpr-1* mutation co-segregated with a G to A splice-site mutation at the beginning of the 11th intron in the *CvKasIV-2* gene in 100/100 F_2 progeny from VS-CPR-1 (*cpr-1/cpr-1*) X LN43 (*CPR-1/CPR-1*) (Slabaugh, unpublished data). Thus, *KasIV-2* is a candidate gene for *cpr-1*. VS335 differs from VS-CPR-1 in that both C8:0 and C10:0 are decreased. VS335 is hypothesized to be homozygous for the *cpr-1* mutation and also for a new mutation, designated *cei-2*, responsible for the decrease in C8:0.

Medium chain specific fatty acyl-ACP thioesterase (FatB) and β-ketoacyl-ACP synthase (KAS) enzyme activities control fatty acid profiles in *Cuphea* (Fuhrmann and Heise 1993, Dehesh et al. 1996a, b and 1998, Leonard et al. 1998). KAS enzymes extend fatty acyl-ACP chains by adding two-carbon units and FatB thioesterases terminate extension by hydrolyzing the acyl-ACP bond. Gene families encoding these two activities have been identified in several *Cuphea* species (Voelker 1996). Gene duplication and subspecialization within one or both of these gene families may be the mechanism by which *Cuphea* has amplified its MCFA-producing capacity and diversified its unique MCFA profiles.

Cuphea FatB genes cluster into three clades by multiple amino acid sequence analysis (Voelker 1996). In vitro and transgenic analyses indicate the clades correlate with different substrate specificities (Voelker 1996). Clade BI enzymes prefer longer chain substrates such as C16:0, Clade BII enzymes prefer C12:0/C14:0 substrates and Clade BIII enzymes prefer short chain substrates such as C8:0 or C10:0 (Voelker

1996). Based on the reduction of C8:0 in VS335, the *C. viscosissima FatB* Clade BIII homologue is a proposed candidate gene for *cei-2*.

I present here (1) the characterization of the *cei-2* and *cpr-1* mutants of *C. viscosissima*, (2) the identification of single nucleotide polymorphisms (SNPs) in the candidate genes *FatB3* and *KasIV-2*, respectively, (3) the development of a SNP assay to detect the *cei-2* and *cpr-1* mutations, and (4) the additive and dominant effects of *cei-2* and *cpr-1* and their interactions.

Materials and Methods

Plants Materials

VS55 (wild type) was mutagenized with 0.04 M EMS for 3 hr. The resulting population was screened for fatty acid variants, one of which had low C10:0 and was designated VS101-*cpr-1* (designated VS-CPR-1 in Knapp 1993) (Table 3-1). Subsequently, this line was re-mutagenized under the same conditions and a mutant with low C8:0 and low C10:0 (Table 3-1) was identified and designated VS335. VS335 was crossed to VS55 and an F_2 segregating population developed. VS342-6 was selected from this population as it displayed only the low C8:0 phenotype (Table 3-1) and thus the putative *cei-2* mutation. To develop VL198, VS101-*cpr-1* was backcrossed to VL90, a non-dormant interspecific line, until a non-dormant *cpr-1* phenotype was recovered (Table 3-1). VS342-6 was manually emasculated and fertilized with pollen from VL198. F_2 seeds from a single F_1 plant were collected for fatty acid analysis and DNA extraction.

Quarter Seed Fatty Acid Extraction

Samples were analyzed in a method similar to Brandt and Knapp (1993). A quarter of a seed, from the cotyledon end, was placed in 0.5 ml hexane in a glass test tube. The seed was crushed with a Teflon rod, which was cleaned between samples. The mixture was incubated at 50°C for 15 min. Samples were evaporated to dryness under a stream of nitrogen gas, then 0.1 ml ethyl ether and 0.1 ml KOH in methanol

Table 3-1. Fatty acid profiles of wild type *C. viscosissima* and *cpr-1*, *cei-1* and *cei-2* mutants. Fatty acids in bold are those decreased by the mutation indicated after the line name.

Line	*CEI* genotype	*CPR-1* genotype	C8:0	C10:0	C12:0	C14:0	C16:0	C18:0	C18:1	C18:2	C18:3
VS55-WT[1]	*CEI-2/CEI-2*	*CPR-1/CPR-1*	17.5	71.1	2.2	0.9	1.8	0.2	2.5	2.8	0.3
VS101-*CPR-1*[1]	*CEI-2/CEI-2*	*cpr-1/cpr-1*	24.5	**38.3**	6.3	14.2	8.2	0.2	2.5	3.7	0.3
VL198-*CPR-1*[2]	*CEI-2/CEI-2*	*cpr-1/cpr-1*	27.3	**35.6**	5.5	13.5	7.4	0.4	3.2	5	0.4
LN43-*cei-1*[2]	*cei-1/cei-1*	*CPR-1/CPR-1*	**0.9**	81.8	2.2	2.7	3.9	0.6	5.2	2.9	0.3
VS342-6-*cei-2*[2]	*cei-2/cei-2*	*CPR-1/CPR-1*	**2.7**	69.6	2.7	8.1	6.1	1.2	3.5	5.2	0.5
VS335[2]	*cei-2/cei-2*	*cpr-1/cpr-1*	**2.6**	**25.1**	5.1	38.5	14.1	0.5	4.4	9.1	0.4

Sources: 1) Knapp 1993; 2) unpublished data

were added to each sample. The samples were mixed by shaking and incubated at 50˚C for 5 min. Then 0.1 ml 0.15 M HCl and 1 ml hexane were added and swirled to mix. About 0.5 ml of the top layer containing the fatty acid methyl esters was transferred to a glass vial using a clean Pasteur pipette for each sample. The samples were capped and loaded on to an autosampler carousel and injected into a Hewlett Packard 6890 gas chromatograph system (Palo Alto, CA) fitted with a JW DB-23 (30.0m by 0.25 mm) column. The helium gas flow was set at 17 mL min^{-1}. The column oven temperature was programmed to increase from 50˚C to 185˚C at 30˚C min^{-1}. The injector and flame ionization detector were set at 220˚C. Peak area was used to quantify each fatty acid-methyl ester as a percentage of total fatty acid-methyl esters.

Seed DNA Extraction

Remnant seed tissue was removed from the seed coat and lyophilized for three days in a 96-deep well plate. The seeds were powdered using a Mixer Mill (Retsch, Haan, Germany) and titanium beads for two 30 sec cycles. CTAB lysis buffer (100 µl) that contained 0.1% ascorbic acid and 0.1% diethyldithiocarbamic acid (DIECA) was added to each tube and homogenized using a Mixer Mill for two 30 sec cycles. The slurry was incubated at 65˚C for 20 min, then 0.3 ml phenol:chloroform was added to each tube. The solution was mixed by inversion. The samples were centrifuged for 10 min at 7000 rpm to separate the organic and aqueous phases. The aqueous phase from each tube was applied to a P30 spin column equilibrated with TE (BioRad, Madison,

WI) and centrifuged for 3 min at 3500 rpm. The eluate containing the DNA was collected in a clean 1.5 ml tube.

Sequencing FatB3

I designed a series of 12 forward and 11 reverse primers targeting the *C. lanceolata FatB3* genomic sequence (GenBank AJ131740). Primers were designed such that multiple fragments would overlap and provide high quality sequence for mutant detection. *C. viscosissima FatB3* sequences were amplified from VS55 and VS342-6 in 10 μl reactions containing 1X PCR buffer, 1.5 mM $MgCl_2$, 0.5 μM each forward and reverse primer, 0.2 mM dNTPs, 0.1 units Taq polymerase (New England Biolabs, Ipswich, MA) and 10 ng template DNA. A touchdown program was used on an MJ Research PTC-200 thermocycler (Waltham, MA) to amplify the DNA (Don et al. 1991). After an initial denaturation step at 95°C for 4 min, a program of 44 cycles was used, consisting of 7 cycles of touch-down PCR (95°C for 20 s, 62°C to 56°C for 45 s, 72°C for 90 s) followed by 33 cycles at a fixed annealing temperature of 56°C, and a final extension time at 72°C for 10 min. PCR products were checked on a 1% agarose gel and screened for gene-specific amplification by SSCP prior to purification using QIAquick PCR purification kits (QIAGEN Inc., Valencia, CA). Fragments were sequenced from both the forward and reverse directions with the ABI BigDye Terminator Cycle Sequencing Ready Reaction Kit v3.1 chemistry at the Genomics Center at the University of Nevada at Reno, Nevada. Overlapping sequences were compiled into a contig and examined for single nucleotide polymorphisms between VS55 and VS342-6 using CodonCode Aligner (CodonCode Corp.).

Single Nucleotide Polymorphism Marker Development

Markers for detection and scoring of *CvFatB3* and *CvKasIV-2* SNPs were designed for use with a commercial fluorescence polarization template-directed incorporation (FP-TDI) kit (PerkinElmer, Shelton, CT). SNP detection primers were designed immediately upstream and downstream of a SNP, such that when extended the polymorphic nucleotide was the first nucleotide to be added. Fragments were first amplified with the target primers in a 10 μl reaction containing 1 X PCR buffer, 1.5 mM $MgCl_2$, 50 μM dNTPs, 0.125 μM each forward and reverse primer, 0.35 U Taq polymerase and 10 ng DNA template. A touchdown PCR program was used to amplify the target fragment as described above. Five μl of product was transferred to a black 384 well plate. PCR cleanup was performed according to Xiao et al. (2004). The PCR product was combined with 2 μl of PCR Clean-up Reagent (exonuclease and alkaline phosphatase, PerkinElmer, Shelton, CT) and 1/10 vol pyrophosphatase (Roche, Indianapolis, IN). The PCR product and cleanup reagents were incubated at 37˚C for 60 min, followed by 15 min at 80˚C. An FP-TDI assay was used to extend the SNP detection primer with a labeled acyclo-dNTP (R110 or TAMRA) (Chen et al. 1999). The reaction (20 μl) contained 1x reaction buffer, 1 μl G/A or C/T (as appropriate) Acyclo Terminator mix, 0.25 μM SNP primer, 0.05 μl AcycloPolymerase (PerkinElmer, Shelton, CT) and 7 μl of processed PCR product. The mixture was brought to 95˚C for 2 min, followed by ~25 cycles of 15 s 95˚C and 30 s 55˚C. A Victor^2 1420 Multilabel Counter, using the Wallac-1420 Workstation program (PerkinElmer), was used to analyze the polarization ratio of the samples. The fluorescence polarization value (P) is calculated by the equation:

$$P = [Ivv - Ivh]/[Ivv + Ivh]$$

where *Ivv* is the emission intensity detected when the emission and excitation polarizers are parallel, *Ivh* is the emission intensity detected when the polarizers are perpendicular to each other (Chen et al. 1999). An Excel macro provided by PerkinElmer was used to evaluate the FP results.

Statistical Analysis

PROC GLM in SAS (SAS Institute Inc., Cary, NC) was used to analyze general linear models of fatty acids as quantitative traits and the SNP genotypes as fixed effects. The main effects (intralocus effects) for *cei-2* and *cpr-1* were assessed, as well as the interaction effect (interlocus effect). Thus additive and dominant effects for both intralocus and interlocus effects were estimated.

Results

The *CvFatB3* gene in a line homozygous for the *cei-2* mutation contains a G to A base change that alters the amino acid sequence.

In order to test the hypothesis that the low C8:0 phenotype in the VS335 line was due to a mutation in a short chain-cleaving thioesterase (*FatB3*) gene, we first selected an F_2 segregant from the cross VS335 (*cei-2/cei-2 cpr-1/cpr-1*) x VS55 (*CEI-2/CEI-2 CPR-1/CPR-1*) that exhibited low C8:0 but wild type levels of C10:0 in seed lipids. The F_2 segregant, designated VS342-6, was presumed to carry wild type alleles at the *cpr-1* locus and mutant alleles at the *cei-2* locus. To scan the *cei-2* candidate gene, *CvFatB3*, for mutations, we designed primers to exclusively amplify Clade BIII *FatB* genes. The *CvFatB3* genes were PCRed from VS342-6 and VS55 DNA and sequenced.

As expected, the *C. viscosissima FatB3* sequences were highly homologous to *FatB3* from the closely-related species *C. lanceolata,* as well as to other Clade BIII *FatB* genes (Figures 3-1 and 3-2). The wild type *C. viscosissima* and *C. lanceolata* FatB3 amino acid sequences were 98.3% identical, with differences at seven sites (Figure 3-2). When the DNA sequence amplified from VS342-6 was compared to that of wild type VS55, two single nucleotide polymorphisms (SNPs) were noted. One SNP was an additional T in a run of polyT within an intron. This difference would not affect the coding sequence and is not expected to alter splicing of the precursor mRNA. The other SNP was a G to A change that resulted in a serine (aa 231) to asparagine amino acid change in exon three (Figure 3-2).

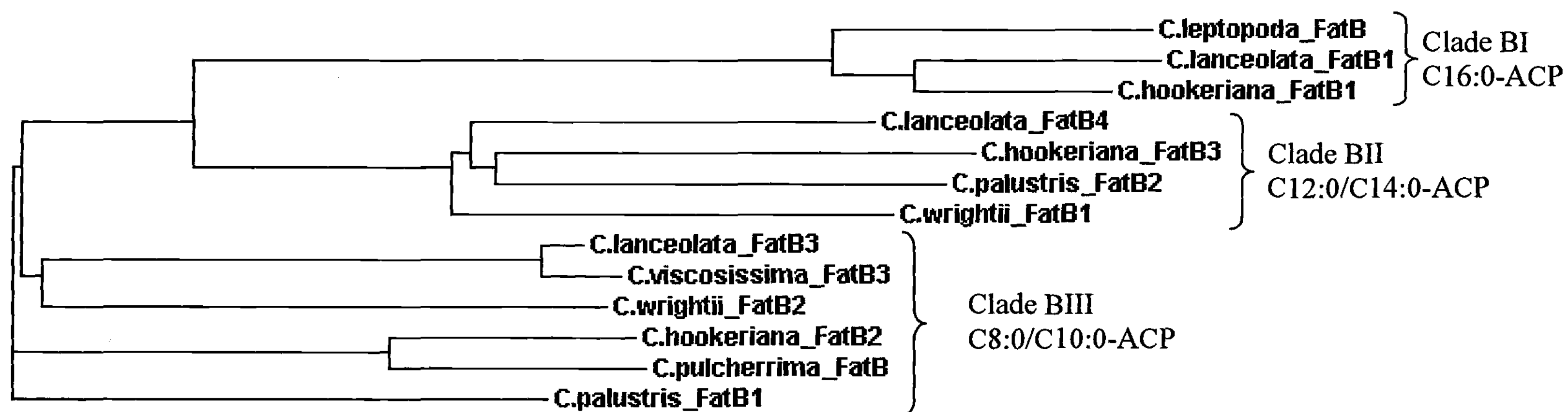


Figure 3-1. Cladogram based upon multiple sequence alignment of available *Cuphea FatB* amino acid sequences with the three clades and substrate specificities noted.

```
                      *        20         *        40         *        60         *        8
CvFatB3-WT    : MVAAAATTAFFPVPAPGTSPKPGKSGNWPSSLSPTFKPKSIPNGGFQVKANASAHPKANGSAVNLKSGSLNTQEDTSSS :  79
CvFatB3-cei2  : MVAAAATTAFFPVPAPGTSPKPGKSGNWPSSLSPTFKPKSIPNGGFQVKANASAHPKANGSAVNLKSGSLNTQEDTSSS :  79
ClFatB3       : MVAAAATSAFFPVPAPGTSPKPGKSGNWPSSLSPTFKPKSIPNAGFQVKANASAHPKANGSAVNLKSGSLNTQEDTSSS :  79

                0         *       100         *       120         *       140         *        1
CvFatB3-WT    : PPPRAFLNQLPDWSMLLTAITTVFVAAEKQWTMLDRKSKRPDMLVDSVGLKSIVRDGLVSRHSFSIRSYEIGADRTASI : 158
CvFatB3-cei2  : PPPRAFLNQLPDWSMLLTAITTVFVAAEKQWTMLDRKSKRPDMLVDSVGLKSIVRDGLVSRHSFSIRSYEIGADRTASI : 158
ClFatB3       : PPPRAFLNQLPDWSMLLTAITTVFVAAEKQWTMLDRKSKRPDMLVDSVGLKSIVRDGLVSRQSFLIRSYEIGADRTASI : 158

                60        *       180         *       200         *       220         *
CvFatB3-WT    : ETLMNHLQETTINHCKSLGLHNDGFGRTPGMCKNDLIWVLTKMQIMVNRYPTWGDTVEINTWFSQSGKIGMASDWLISD : 237
CvFatB3-cei2  : ETLMNHLQETTINHCKSLGLHNDGFGRTPGMCKNDLIWVLTKMQIMVNRYPTWGDTVEINTWFSQSGKIGMANDWLISD : 237
ClFatB3       : ETLMNHLQETSINHCKSLGLLNDGFGRTPGMCKNDLIWVLTKMQIMVNRYPTWGDTVEINTWFSQSGKIGMASDWLISD : 237

                240       *       260         *       280         *       300         *
CvFatB3-WT    : CNTGEILIRATSVWAMMNQKTRRFSRLPYEVRQELTPHFVDSPHVIEDNDQKLRKFDVKTGDSIRKGLTPRWNDLDVNQ : 316
CvFatB3-cei2  : CNTGEILIRATSVWAMMNQKTRRFSRLPYEVRQELTPHFVDSPHVIEDNDQKLRKFDVKTGDSIRKGLTPRWNDLDVNQ : 316
ClFatB3       : CNTGEILIRATSVWAMMNQKTRRFSRLPYEVRQELTPHFVDSPHVIEDNDQKLHKFDVKTGDSIRKGLTPRWNDLDVNQ : 316

                 320      *       340         *       360         *       380         *
CvFatB3-WT    : HVSNVKYIGWILESMPIEVLETQELCSLTVEYRRECGMDSVLESVTAVDPSENGGRSQYKHLLRLEDGTDIVKSRTEWR : 395
CvFatB3-cei2  : HVSNVKYIGWILESMPIEVLETQELCSLTVEYRRECGMDSVLESVTAVDPSENGGRSQYKHLLRLEDGTDIVKSRTEWR : 395
ClFatB3       : HVSNVKYIGWILESMPIEVLETQELCSLTVEYRRECGMDSVLESVTAVDPSENGGRSQYKHLLRLEDGTDIVKSRTEWR : 395

                 400      *
CvFatB3-WT    : PKNAGTNGAISTSTAKTSNGNSVS : 419
CvFatB3-cei2  : PKNAGTNGAISTSTAKTSNGNSVS : 419
ClFatB3       : PKNAGTNGAISTSTAKTSNGNSAS : 419
```

Figure 3-2. Amino acid alignment of *CvFatB3-WT*, *CvFatB3-cei-2* and *ClFatB3*. The serine to asparagine amino acid change caused by the SNP in the mutant VS342-6 is highlighted in black. Differences between the *CvFatB3* and the closely related *ClFatB3* are highlighted in grey.

SNP markers were developed for the *CvFatB3* and *CvKasIV-2* mutations.

To determine if the identified SNP mutation correlated with the mutant phenotype, markers were developed to detect the SNPs in the candidate genes *CvFat3* and *CvKasIV-2*. For the *CvFatB3* mutation, primers used previously to sequence the gene were selected for target amplification. The forward primer, III-E2F (GTACTCCTGGGATGTGTAAAAAC) and the reverse primer, III-i3R (TCTGCTCACTGATCGAGCATCA) produced a 356 bp fragment. The SNP primer which annealed immediately 3' of the SNP, FatB3SNP-D (GCAATCACTTATTAGCCAATCG) was used in the FP-TDI assay to score the F_2 population. Paralog-specific primers were designed to span the region of the *CvKasIV-2* SNP. The forward primer, F182 (TGCAGTTAAGAGTGAATTCCACC) and the reverse primer, R183 (CAAGGGGCAAAGAGTATGGAT) produced a 505 bp fragment. The SNP detection primer CPRsnp-U (GTAGAAGCAGTTACCGTAGTTCA) that annealed immediately 5' of the SNP was used to score the F_2 population.

Combining the *cei-2* and *cpr-1* mutations by crossing lines homozygous for each mutation recreated the high myristic VS335 fatty acid phenotype in a subset of the F_2 progeny.

A new F_2 population segregating for the *cei-2* and *cpr-1* mutations was developed by crossing VS342-6 (*cei-2/cei-2 CPR-1/CPR-1*) with VL198 (*CEI-2/CEI-2 cpr-1/cpr-1*) to determine if the VS335 fatty acid phenotype could be reproduced from individuals carrying the *cei-2* and *cpr-1* mutations. Fatty acid profiles were determined for 92 F_2 individuals from quarter seeds. Four individuals exhibited the VS335-like fatty acid profile of low C8:0 with high C14:0 (Figure 3-3).

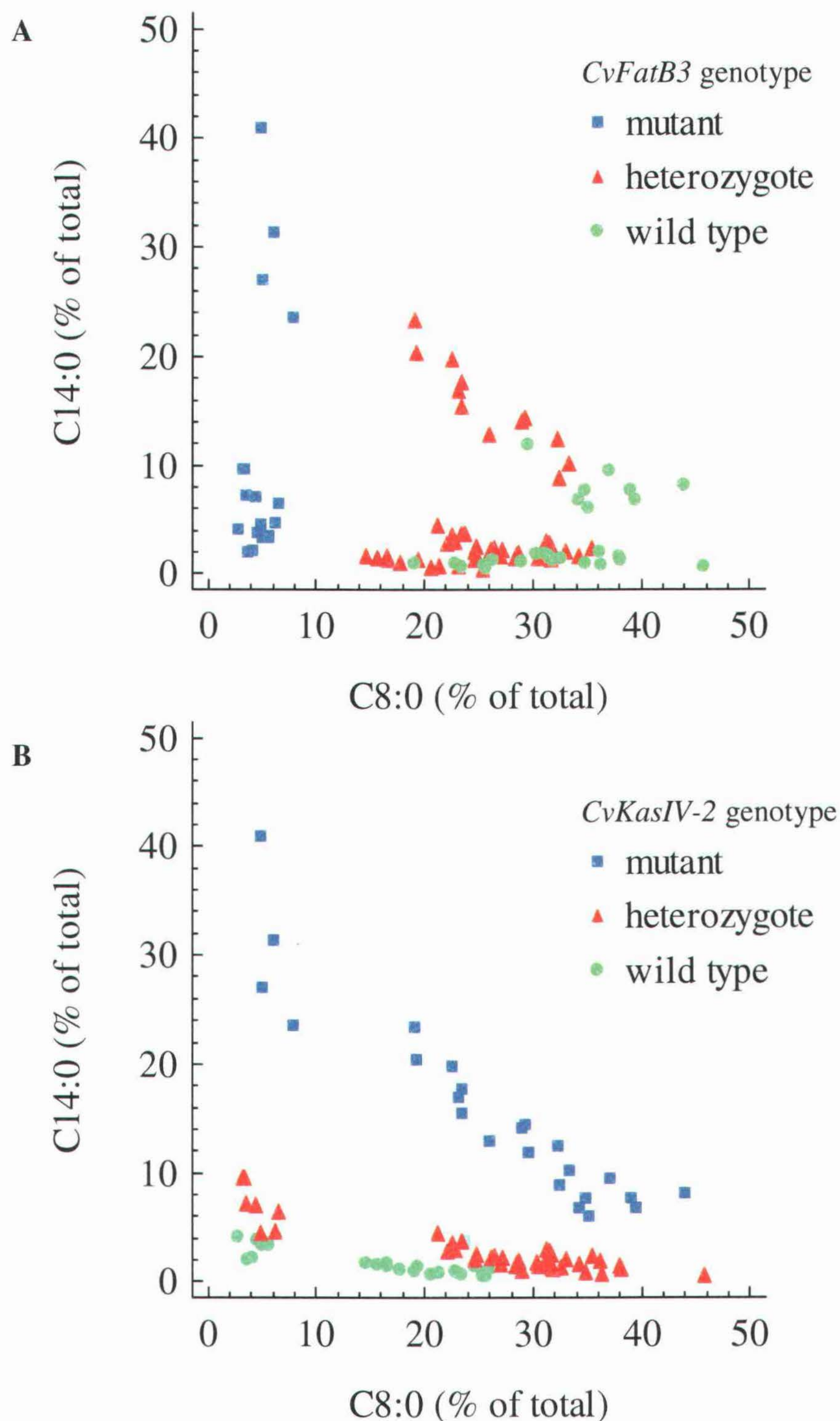


Figure 3-3. C8:0/C14:0 fatty acid phenotypes and *CvFatB3* (A) and *CvKasIV-2* (B) SNP genotypes among 92 F_2 progeny from VS342-6 x VL198. VS342-6 contributed the mutant allele of *CvFatB3* and VL198 contributed the mutant allele of *CvKasIV-2*. (A) *CvFatB3* SNP genotypes; WT=G/G, heterozygote=G/A, mutant =A/A. (B) *CvKasIV-2* SNP genotypes; WT=G/G, heterozygote=G/A, mutant =A/A.

Three genotypic classes were observed for each of the SNPs, which followed Mendelian inheritance patterns. The SNP in *CvFatB3* had 18 wild type individuals, 46 heterozygotes and 28 homozygous mutant individuals. The *CvKasIV-2* SNP had 25 wild type individuals, 43 heterozygotes, and 24 homozygous mutant individuals (Figures 3-4 and 3-5). The two mutations segregated independently. Four individuals that exhibited the VS335-like fatty acid profile were homozygous for both the *CvFatB3* and *CvKasIV-2* SNPs (Figure 3-3). Thus the *cei-2 cpr-1* double mutant phenotype co-segregated with the mutant SNP genotypes.

Analysis of additive and dominance effects of the *cei-2* and *cpr-1* candidate SNPs.

The putative *cei-2* genotype reduced C8:0 from 24% to 4.5%. In contrast, C10:0 increased from 68% to 82%, C14:0 from 1% to 3.2%, and C16:0 from 1.4% to 3.8% relative to wild type (Table 3-2, rows 1 and 3). The additive effect of *cei-2* on C8:0 was 13.3%, while the dominant effect was -6.10%, thus the degree of dominance of *cei-2* was 0.46 (Table 3-3). The *cei-2* additive effect on C10:0 was -4.6%, while the dominant effect was 3.0%, thus the degree of dominance of *cei-2* was 0.66 (Table 3-3).

Individuals homozygous for both mutations (putatively *cei-2/cei-2 cpr-1/cpr-1*) had reduced C8:0 from 24% to 6% and C10:0 from 68% to 36%. In contrast, C12:0 increased from 4.4% to 11%, C14:0 from 1% to 31%, and C16:0 from 1.4% to 11% relative to wild type (Table 3-2, rows 1 and 9). The *cpr-1* by *cei-2* additive by additive effect on C8:0 was -2.8%, which was the only significant interaction on the C8:0 levels (Table 3-3). In contrast, the C10:0 *cpr-1* by *cei-2* dominant by dominant effect was the only interaction that was not significant (Table 3-3). The additive by additive

effect was -4.5%, the additive by dominant effect was 3.5%, and the dominant by additive effect was 6.6% (Table 3-3).

Figure 3-4. The *CvFatB3* SNP (candidate for the *cei-2* mutation) assay results.

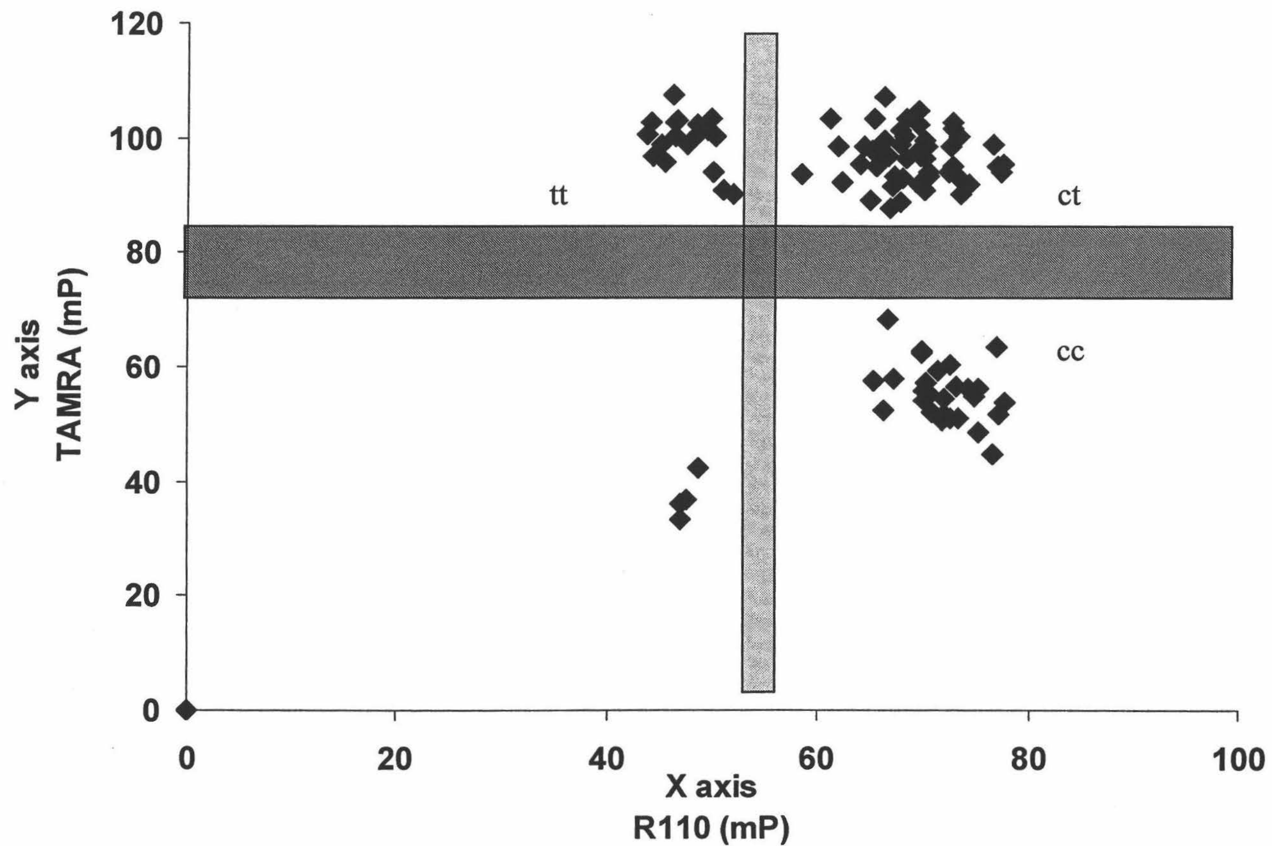


Figure 3-5. The *CvKasIV-2* SNP (candidate gene for the *cpr-1* mutation) assay results.

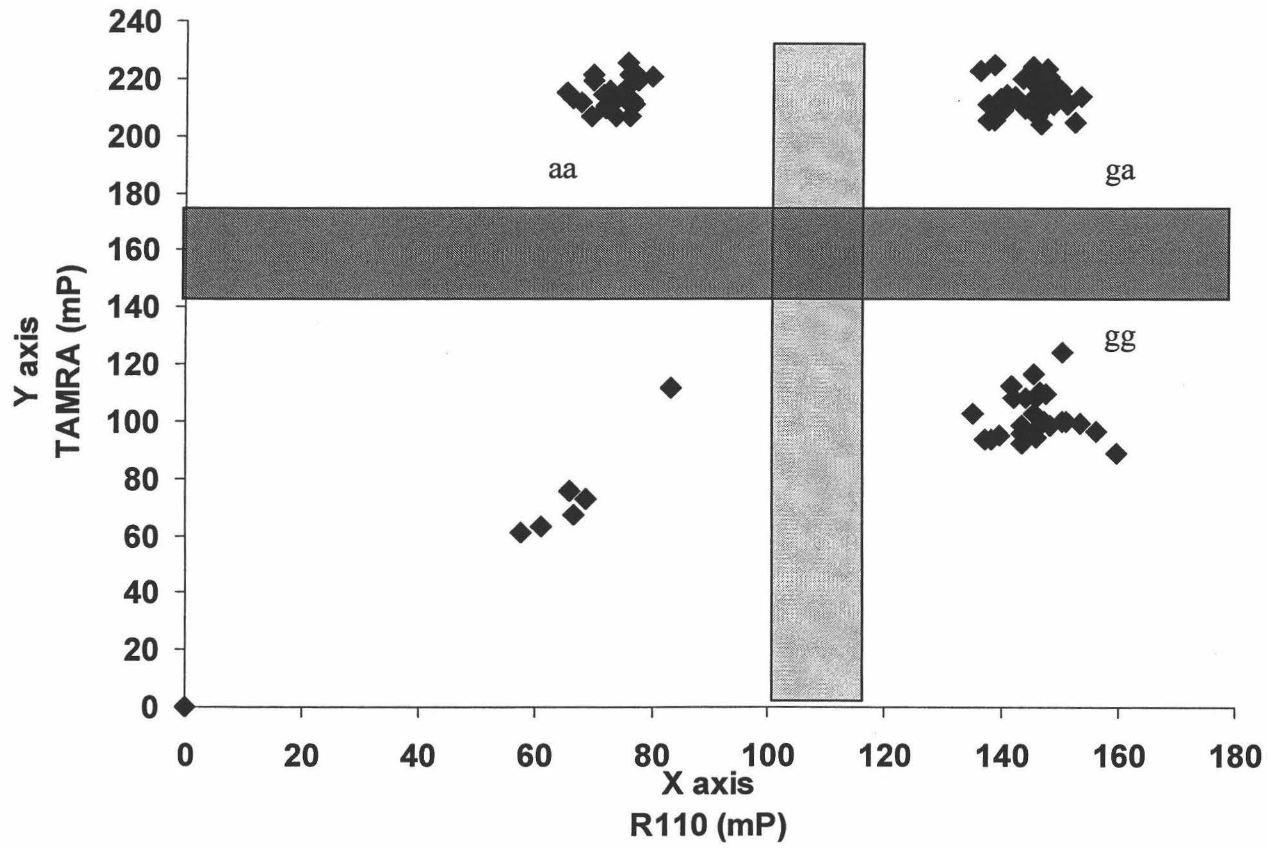


Table 3-2. The LSMeans for fatty acids of the nine genotypic classes of *cpr-1* and *cei-2*.

CPR	CEI	C6:0	C8:0	C10:0	C12:0	C14:0	C16:0
VS342-6	VL198	0	24.03	67.83	4.35	0.84	1.41
VS342-6	Het.	0.46	19.59	68.97	6.15	1.23	1.6
VS342-6	**VS342-6**	0.84	**4.5**	81.72	4.65	3.2	3.75
Het.	VL198	0.33	34.34	54.49	6.2	1.37	1.44
Het.	Het.	0.31	27.63	58.94	6.34	2.54	2.82
Het.	**VS342-6**	0.56	**4.6**	72.44	7.31	7.06	6.29
VL198	VL198	1.12	36.65	**40.35**	8.39	8.07	4.83
VL198	Het.	0.4	26.1	**39.46**	9.19	15.61	7.7
VL198	**VS342-6**	2.33	**5.95**	**36.07**	11.02	30.75	11.28

Note: Mutant genotypes for each SNP are in bold. Parental source of each genotype is designated. VS342-6 is the *cei-2* donor, VL198 is the *cpr-1* donor. Het.=heterozygous for the SNP.

Table 3-4 Additive and dominant effects of the candidate *cpr-1* and *cei-2* SNPs.

Fatty Acid	C6:0	P-value	C8:0	P-value	C10:0	P-value	C12:0	P-value	C14:0	P-value	C16:0	P-value
cpr-1												
Additive effect	-0.43	0.0036	-3.43	<0.0001	17.11	<0.0001	-2.24	<0.0001	-8.19	<0.0001	-2.84	<0.0001
Dominant Effect	0.45	0.036	-2.72	0.0029	-6.22	<0.0001	0.68	0.22	6.29	<0.0001	1.58	<0.0001
Degree of Dominance	1.05		0.79		0.36		0.30		0.77		0.56	
cei-2												
Additive effect	-0.38	0.011	13.33	<0.0001	-4.59	<0.0001	-0.67	0.072	-5.12	<0.0001	-2.27	<0.0001
Dominant Effect	0.47	0.024	-6.1	<0.0001	3.02	0.0013	-0.24	0.64	2.09	<0.0001	0.8	0.019
Degree of Dominance	1.24		0.46		0.66		0.36		0.41		0.35	
cpr-1* x *cei-2												
AA efffect	0.09	0.62	-2.79	0.0006	-4.54	<0.0001	0.58	0.23	5.08	<0.0001	1.03	0.0012
AD effect	-0.69	0.13	-0.26	0.82	3.53	0.0041	-1.09	0.12	-1.5	0.028	0.31	0.48
DA effect	-0.4	0.17	-2.32	0.57	6.57	<0.0001	-0.18	0.81	-3.42	<0.0001	0.23	0.63
DD effect	0.51	0.21	3.1	0.65	-2.24	0.21	-0.98	0.34	0.63	0.53	-0.37	0.57

Discussion

FatB and KAS enzymes are known to have an effect on MCFA profiles of *Cuphea* seeds (Davies et al. 1991, Dehesh et al. 1996a, b and 1998; Leonard et al. 1998). FatB thioesterases cleave the acyl group from the ACP, freeing the fatty acid and allowing it to be exported from the plastid to the cytosol for further processing and incorporation into triacylglycerol. The substrate range of FatB thioesterases varies in different species. *Umbellularia californica* (California bay) FatB thioesterase has a clear preference for C12:0-ACP, consistent with the accumulation of C12:0 as one of the major fatty acids the species utilizes for its seed reserves (Davies et al. 1991). Transformation of *Arabidopsis* and *Brassica napus* with the bay thioesterase re-directed a portion of seed fatty acid synthesis to C12:0 and C14:0 chain lengths (Voelker et al. 1992). In *C. palustris* two thioesterases, *Cp* FatB1 (Clade BIII) and *Cp* FatB2 (Clade BII), appear to be largely responsible for the unique oil profile of this species dominated by 64% and 20% of myristic and caprylic acids, respectively (Dehesh et al. 1996). Clade BII and Clade BIII Fat B cDNAs were also isolated from *C. wrightii*, a species rich in capric (29%) and lauric (69%) fatty acids. However, as assayed in transgenic *Arabidopsis*, *Cw*FatB1 and *Cw*FatB2 showed broader substrate specificity than expectedif they are the sole determinants of the capric and lauric acid rich seed oil of *C. wrightii* (Leonard et al. 1997). Leonard et al. (1997) suggest that other factors must contribute to the determination of the oil profile of *C. wrightii*.

One such factor may be the KAS enzymes, as they have also been found to have an effect on MCFA profiles (Dehesh et al. 1998, Leonard et al. 1998, Schutt et al.

2002). KAS enzymes build the acyl-ACP chain by adding two carbon units from acetyl-CoA. Because both KAS and FatB enzymes utilize the same acyl-ACP substrates to extend or terminate chain growth, respectively, the relative abundance and/or substrate specificities of these two enzyme classes is likely to have a profound impact on seed oil profiles. KASIV has been found to have impaired activity in the extension of C8:0 in the *cpr-1* mutant of *C. viscosissima* (Slabaugh et al. 1998). Comparison of wild type and VS-CPR-1 (*cpr-1*/*cpr-1*) in vitro fatty acid synthesis showed reduced extension to C10:0 and C12:0 by extracts prepared from *cpr-1* embryos (Slabaugh et al. 1998). After treatment with the KASI inhibitor cerulenin, wild type extracts exhibited drastically reduced extension beyond C10:0. Mutant extracts had reduced extension beyond C8:0, indicating that the *cpr-1* mutation plays a specific role in the extension of the C8:0-ACP (Slabaugh et al. 1998), implicating a mutation in a KAS gene. Comparison of wild type and *cpr-1* KAS proteins by Western blot analysis indicated a drastic reduction in a 46 kDa KAS band, denoted as KAS A1 in the report but subsequently renamed KASIV to indicate its putative role in MCFA biosynthsis. These results suggest that the *cpr-1* mutation affects the accumulation of KAS IV (Slabaugh et al. 1998). Sequencing of the *CvKASIV-2* gene from wild type and VS-CPR-1 revealed a single nucleotide difference that affected a splice site, such that a stop codon in the 11th intron terminates protein synthesis early, yielding a severely truncated protein (Slabaugh, unpublished data). This splicing defect and truncated protein supports the previous findings by Slabaugh et al. (1998) that the *cpr-1* mutant accumulates ~1/4 the amount of KASIV protein as wild type.

In the VS342-6 x VL198 F_2 population studied herein, the *CvKasIV-2* SNP genotype co-segregated with the low C10:0 phenotype (Table 3-2, Figure 3-3). The co-segregation of the phenotype and mutant genotype provide further support that the *CvKASIV-2* candidate gene is the gene affected by the *cpr-1* mutation.

Multiple sequence analysis indentifies three FatB clades among *Cuphea* species (Figure 3-1, Voelker 1996). In vitro and transgenic assays performed on several of these *Cuphea FatB* genes indicate that in general Clade BI acts on C16:0-ACP, Clade BII has substrate specificity toward C12:0- and C14:0-ACP, and Clade BIII has C8:0- and C10:0-ACP specificity. The *CvFatB* gene in which the *cei-2* associated SNP was identified is highly homologous to *ClFatB3*, from *C. lanceolata* (Figures 3-1 and 3-2, Voelker 1996). As compared to *C. viscosissima*, wild type *C. lanceolata* has very low C8:0 (~1%). The locus controlling caprylic acid levels in a *C. viscosissima* x *C. lanceolata* cross was designated *cei-1* and this locus was tightly linked to FatB3 in the interspecific map (Slabaugh, unpublished data). The SNP identified in *CvFatB3* in VS342-6, the *cei-2* donor parent, supports the hypothesis that a FatB is affected by the *cei-2* mutation. Thus *cei-1* of *C. lanceolata* and *cei-2* of *C. viscosissima* likely affect the same gene. A proposed three-dimensional structure for the FatB class of thioesterases has recently been published (Mayer and Shanklin 2005). The amino acid residue that is substituted due to the VS342-6 SNP is at a location identified as affecting substrate specificity in other FatB thioesterases (Mayer and Shanklin 2005). Substrate specificity determining residues are clustered in the N-terminal domain, while the catalytic residues are in the C-terminal domain (Mayer and Shanklin 2005). This modular domain structure may explain the great diversity of oil

profiles in *Cuphea* as this structure allows for rapid diversification of enzyme function (Patthy 2003).

The *CvFatB3* mutant genotype unambiguously co-segregated with the low C8:0 phenotype in the cross we studied (Figure 3-3). Nine distinct phenotypic classes were not observed in the scatter plots of fatty acid values shown in Figure 3-3 or in other analyses, e.g. C8:0 x C10:0, though certain classes were easily distinguished, including the homozygous double mutants (*cei-2/cei-2 cpr-1/cpr-1*, Figure 3-3). The quarter seed fatty acid analysis we performed is less precise than whole seed analysis, but necessary here as recovery of plants with the VS335-like fatty acid profile was desired and DNA for further analysis and correlation to the SNP assays was required. Previous phenotypic analysis of whole seeds from a population segregating for *cpr-1* and *cei-2* in an all *C. viscosissima* background (VS335 x VS55) showed nine distinct classes (Crane, unpublished data). Though VL198 is an interspecific line, it has primarily *C. viscosissima* background due to the backcrossing scheme used to develop it. Even so, the *C. lanceolata* background of VL198 may have contributed to the blurring of the nine phenotypic classes, as *C. lanceolata* also has low C8:0 levels relative to *C. viscosissima.* Brandt and Knapp (1993) found that *C. lanceolata* alleles were overabundant at fatty acid loci in an interspecific cross, indicating the *C. lanceolata* alleles may persist during a backcrossing scheme and thus affect the fatty acid profile in the VS342-6 x VL198 population. Analysis of the SNP and phenotype in a wild type by *cei-2* (VS342-6) population would remove this confounding factor and confirm (or refute) the correlation of the mutant SNP with the low C8:0 phenotype.

The phenotype of an individual cannot be simply predicted from only the *cpr-1* or *cei-2* effect because there are significant interactions between these two loci (Table 3-3). The genotype of the other locus must be considered for C8:0, C10:0, C14:0 and C16:0, as there is at least one significant interaction between the additive or dominant effects of *cpr-1* and *cei-2* (Table 3-3). As the fatty acid synthesis pathway is linear, it is not surprising that one mutation that affects a fatty acid pool will affect a mutation that affects a longer fatty acid. FatB and KAS enzymes are also presumed to be in competition for the same substrates (acyl-ACPs), therefore an alteration in the activity of one of these enzymes may alter the balance of this competition and cause a change in the resulting fatty acid profile. For example, as *cei-2* reduces the amount of C8:0-ACP that is cleaved, more C8:0-ACP is available to be extended by KAS, thus an increase of C10:0 is seen. However, when the *cpr-1* mutation is also present the C10:0 levels are suppressed and the increase is seen in C14:0 (Table 3-2). This result is curious as one would expect to see the increase in C12:0. A slight increase was observed, but the shift of fatty acids is primarily to C14:0 and C16:0 (Table 3-2). Why C12:0 seems to be skipped is a question of significance as lauric acid is the most economically important MCFA.

The *CvKASIV-2* gene is corroborated as the candidate gene affected by the *cpr-1* mutation based on biochemical evidence of reduced protein accumulation and altered activity (Slabaugh et al. 1998), identification of a SNP in the gene, and the co-segregation of the mutant SNP genotype with the low C10:0 phenotype (Figure 3-3). The *CvFatB3* gene is corroborated as the candidate gene affected by the *cei-2* mutation due to the identification of a SNP in the FatB thioesterase that corresponds to

the clade that is active on C8:0- and C10:0-ACP and the co-segregation of the mutant SNP with the low C8:0 phenotype. The combination of the *cpr-1* and *cei-2* mutations provides a unique fatty acid profile rich in myristic acid that will be useful in breeding domesticated *Cuphea*.

Works Cited

Brandt, T., and S.J. Knapp. 1993. Gemetic selection at fatty acid and allozyme marker loci and meiosis within *Cuphea viscosissima* x *Cuphea lanceolata* populations. Crop Science 33:1138-1143.

Chen, X., L. Levine, and P.-Y. Kwok. 1999. Fluorescence polarization in homogeneous nucleic acid analysis. Genome Res. 9:492-498.

Davies, H.M., L. Anderson, C. Fan, and D.J. Hawkins. 1991. Developmental induction, purification, and further characterization of 12:0-ACP thioesterase from immature cotyledons of *Umbellularia californica*. Archives of Biochemistry and Biophysics 290:37-45.

Dehesh, K., A. Jones, D.S. Knutzon, and T.A. Voelker. 1996a. Production of high levels of 8:0 and 10:0 fatty acids in transgenic canola by overexpression of Ch FatB2, a thioesterase cDNA from *Cuphea hookeriana*. Plant J 9:167-172.

Dehesh, K., P. Edwards, T. Hayes, A.M. Cranmer, and J. Fillatti. 1996b. Two novel thioesterases are key determinants of the bimodal distribution of acyl chain length of *Cuphea palustris* seed oil. Plant Physiol. 110:203-210.

Dehesh, K., P. Edwards, J. Fillatti, M. Slabaugh, and J. Byrne. 1998. KAS IV: a 3-ketoacyl-ACP synthase from *Cuphea* sp. is a medium chain specific condensing enzyme. The Plant Journal 15:383-390.

Don, R., P. Cox, B. Wainwright, K. Baker, and J. Mattick. 1991. 'Touchdown' PCR to circumvent spurious priming during gene amplification. Nucleic Acid Research 19:4008.

Gesch, R.W., F. Forcella, N.W. Barbour, W.B. Voorheees, and B. Phillips. 2003. Growth and yield response of *Cuphea* to row spacing. Field Crops Research 81:193-199.

Graham, S. 1988. Revision of *Cuphea* section *Heterodon* (Lythraceae). Syst. Bot. Mono. 20:1-168.

Knapp, S.J., and L. Tagliani. 1991. Two medium-chain fatty acid mutants of *C. viscosissima*. Plant Breeding 106:338-341.

Knapp, S.J. 1993. Modifying the seed storage lipids of *Cuphea*: a source of medium-chain tiglycerides, p. 142-154, In S. MacKenzie and D. Taylor, eds. Seed oils for the future. American Oil Chemists' Society, Champaign IL.

Leonard, J.M., M.B. Slabaugh, and S.J. Knapp. 1997. *Cuphea wrightii* thioesterases have unexpected broad specificities on saturated fatty acids. Plant Molecular Biology 34:669-679.

Leonard, J.M., S.J. Knapp, and M.B. Slabaugh. 1998. A *Cuphea* β-ketoacyl-ACP synthase shifts the synthesis of fatty acids towards shorter chains in *Arabidopsis* seeds expressing *Cuphea* FatB thioesterases. The Plant Journal 13:621-628.

Mayer, K.M., and J. Shanklin. 2005. A structural model of the plant acyl-acyl carrier protein thioesterase FatB comprises two helix/4-stranded sheet domains, the N-terminal domain containing residues that affect specificity and the C-terminal domain containing catalytic residues. J. Biol. Chem. 280:3621-3627.

Patthy, L.S. 2003. Modular assembly of genes and the evolution of new functions. Genetica 118:217-231.

Slabaugh, M.B., J.M. Leonard, and S.J. Knapp. 1998. Condensing enzymes from *Cuphea wrightii* associated with medium chain fatty acid biosynthesis. Plant J 13:611-620.

Thompson, A.E. 1984. *Cuphea*-a potential new crop. HortScience 19:352-354.

Topfer, R., N. Martini, and J. Schell. 1995. Modification of plant lipid wynthesis. Science 268:681-686.

Voelker, T.A., A.C. Worrell, L. Anderson, J. Bleibaum, C. Fan, D.J. Hawkins, S.E. Radke, and H.M. Davies. 1992. Fatty acid biosynthesis redirected to medium chains in transgenic oilseed plants. Science 257:72-74.

Voelker, T.A. 1996. Plant acyl-ACP thioesterases: chain-length determining enzymes in plant fatty acid biosynthesis, p. 111-133, In J. Setlow, ed. Genetic Engineering, Vol. 18. Plenum Press, New York.

Webb, D., S.J. Knapp, and L. Tagliani. 1992. An RFLP and allozyme linkage map of *Cuphea lanceolata*. Theoretical and Applied Genetics 83:528-532.

Xiao, M., A. Phong, K.L. Lum, R.A. Greene, P.R. Buzby, and P.-Y. Kwok. 2004. Role of excess inorganic pyrophosphate in primer-extension genotyping assays. Genome Research 14:1749-1755.

Breeding tools for *Cuphea*: An updated genetic linkage map and fatty acid mutant candidate gene analysis

CHAPTER 4

Conclusions

Medium chain fatty acids (MCFAs) are important industrial feedstocks for soaps, detergents, and surfactants, among other things (Topfer et al. 1995). Currently, coconut (*Cocus necifera* L.) and palm kernel (*Elaeis guinenes* Jacq.) oils are the utilized renewable supplies of MCFAs (Thompson et al. 1990). An interspecific line derived from the C10:0 rich species *Cuphea lanceolata* and *Cuphea viscosissima* is being domesticated as a temperate source of MCFAs. *Cuphea* will diversify the sources of MCFAs, stabilizing supplies and prices (Webb et al. 1992). *C. lanceolata* and *C. viscosissima* are good model organisms for studying the biosynthesis of MCFAs in seeds as they have small diploid genomes (Webb et al. 1992, Kopka et al. 1993). Significant strides in the domestication of *Cuphea* have been made, notably the identification of non-dormant and reduced seed shattering germplasm (Knapp 1993, Knapp and Crane 2000). As a new crop there are few molecular genetic tools available to speed the breeding and domestication efforts currently underway. The existing genetic linkage map of *C. lanceolata* consists of 32 RFLP and five allozyme markers (Webb et al. 1992). Increased marker density is desirable for marker assisted selection (MAS), quantitative trait loci (QTL) analysis and map-based cloning (Liu et al. 1996). A solid genetic linkage map acts as a basic tool in the modern-day plant breeder's toolbox.

SSRs and fatty acid synthesis loci were used here to increase the density of the *C. lanceolata* linkage map. SSR markers proved to be highly polymorphic and informative with heterozygosity ranging from 0.278 to 0.889 with a mean of 0.73. These markers amplified both *C. lanceolata* and *C. viscosissima* alleles, showing good cross taxa utility. The addition of SSR and fatty acid synthesis loci to the *C. lanceolata* map greatly increases the utility of the map, due to increased marker density and population of the map by PCR-based markers. The map is 340.8 cM in length, spread across eight linkage groups. Five of the linkage groups likely correspond to chromosomes, while the other three groups have only two or three markers. The lack of a large sixth group indicates the genome has not been fully covered. The fatty acid synthesis loci also will be useful in breeding, as they may be related to the oil profile. The map represents a significant step in the breeding of *Cuphea* as it provides the necessary framework with which to conduct QTL analysis and MAS.

Candidate gene analysis of fatty acid mutants sheds light on the relationship between enzymes in the fatty acid synthesis pathway. Based on the reduction of C8:0 in the *cei-2* mutant, a medium-chain specific FatB thioesterase was considered as the candidate for the gene affected by the *cei-2* mutation. The *cpr-1* mutation, which reduced C10:0, was known to co-segregate with a splice site mutation in the *CvKasIV-2* gene, reducing the extension of C10:0-ACP (Slabaugh, unpublished data). The identification of SNPs in both the *CvFatB3* and *CvKasIV-2* genes supported these genes as candidates for the *cei-2* and *cpr-1* mutations, respectively. Furthermore, in an F_2 population segregating for the two mutations, the SNP genotypes co-segregated

with the mutant phenotypes 100%. The combination of the *cei-2* and *cpr-1* mutations results in a novel fatty acid profile with a 30-fold increase in C14:0 and drastic reduction of C8:0 and C10:0.

The domestication and breeding of *Cuphea* is an ongoing process. The updated genetic linkage map of *C. lanceolata* is a tool which breeders can utilize for MAS and geneticists can track traits of interest through QTL analysis. Mutants are one method that novel oil profiles can be developed and exploited for breeding purposes. Candidate gene analysis of two such mutants, *cei-2* and *cpr-1*, allows for more effective utilization of these mutations in future breeding efforts as the root cause of the phenotypes is better understood and potential interactions recognized. The studies presented here represent molecular tools and analyses that will benefit further breeding and domestication efforts in *Cuphea*.

BIBLIOGRAPHY

Brandt, T., and S.J. Knapp. 1993. Gametic selection at fatty acid and allozyme marker loci and meiosis within *Cuphea viscosissima* x *Cuphea lanceolata* populations. Crop Science 33:1138-1143.

Chen, W., and W.W. Roath. 1995. Karyotype of *Cuphea lanceolata* Ait. and *Cuphea viscosissima* Jacq. Crop Science 35:246-250.

Chen, X., L. Levine, and P.-Y. Kwok. 1999. Fluorescence polarization in homogeneous nucleic acid analysis. Genome Res. 9:492-498.

Davies, H.M., L. Anderson, C. Fan, and D.J. Hawkins. 1991. Developmental induction, purification, and further characterization of 12:0-ACP thioesterase from immature cotyledons of *Umbellularia californica*. Archives of Biochemistry and Biophysics 290:37-45.

Dehesh, K. 2001. How can we genetically engineer oilseed crops to produce high levels of medium-chain fatty acids? European Journal of Lipid Science and Technology 103:688-697.

Dehesh, K., A. Jones, D.S. Knutzon, and T.A. Voelker. 1996a. Production of high levels of 8:0 and 10:0 fatty acids in transgenic canola by overexpression of Ch FatB2, a thioesterase cDNA from *Cuphea hookeriana*. Plant J 9:167-172.

Dehesh, K., P. Edwards, T. Hayes, A.M. Cranmer, and J. Fillatti. 1996b. Two novel thioesterases are key determinants of the bimodal distribution of acyl chain length of *Cuphea palustris* seed oil. Plant Physiol. 110:203-210.

Dehesh, K., P. Edwards, J. Fillatti, M. Slabaugh, and J. Byrne. 1998. KAS IV: a 3-ketoacyl-ACP synthase from *Cuphea* sp. is a medium chain specific condensing enzyme. The Plant Journal 15:383-390.

Don, R., P. Cox, B. Wainwright, K. Baker, and J. Mattick. 1991. 'Touchdown' PCR to circumvent spurious priming during gene amplification. Nucleic Acid Research 19:4008.

Dormann, P., F. Spener, and O.J. B. 1993. Characterization of two acyl-acyl carrier protein thioesterases from developing *Cuphea* seeds specific for medium-chain- and oleoyl-acyl carrier protein. Planta 189:425-432.

Earle, F.R., C.A. Glass, G.C. Geisinger, I.A. Wolff, and Q. Jones. 1960. Search for new industrial oils IV. Journal of American Oil Chemists Society 67:440-447.

Edwards, K., J. Barker, A. Daly, C. Jones, and A. Karp. 1996. Microsatellite libraries enriched for several microsatellite sequences in plants. BioTechniques 20:758-760.

FAOSTAT, d. 2005. [Online].

Fuhrmann, J., and K. Heise. 1993. Factors controlling medium-chain fatty acid synthesis in plastids from maturing *Cuphea* embryos. Z. Naturforsch 48c:616-622.

Gesch, R.W., F. Forcella, N.W. Barbour, W.B. Voorheees, and B. Phillips. 2003. Growth and yield response of *Cuphea* to row spacing. Field Crops Research 81:193-199.

Graham, S. 1988. Revision of *Cuphea* section *Heterodon* (Lythraceae). Syst. Bot. Mono. 20:1-168.

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

Jones, A., H.M. Davies, and T.A. Voelker. 1995. Palmitoyl-Acyl Carrier Protein (ACP) Thioesterase and the Evolutionary Origin of Plant Acyl-ACP Thioesterases. Plant Cell 7:359-371.

Karagyosov, L., I. Kalcheva, and V. Chapman. 1993. Construction of random small-insert genomic libraries highly enriched for simple sequence repeats. Nucleic Acid Research 21:3911-3912.

Karmarkar, V.M. 2004. Simple sequence repeat marker development and fingerprinting in *Cuphea lanceolata* Ait. and *C. viscosissima* Jacq. Masters, Oregon State University, Corvallis.

Knapp, S.J. 1993. Modifying the seed storage lipids of *Cuphea*: a source of medium-chain tiglycerides, p. 142-154, *In* S. MacKenzie and D. Taylor, eds. Seed oils for the future. American Oil Chemists' Society, Champaign IL.

Knapp, S.J., and L.A. Tagliani. 1989. Genetics of allozyme variation in *Cuphea lanceolata Ait. Genome 32:57-63.*

Knapp, S.J., and L. Tagliani. 1991. Two medium-chain fatty acid muntants of *C. viscosissima*. Plant Breeding 106:338-341.

Knapp, S.J., and J.M. Crane. 2000. Registration of reduced seed-shattering *Cuphea* germplasm PSR23. Crop Science 40:299-300.

Knapp, S.J., L. Tagliani, and B. Liu. 1991a. Outcrossing rates of experimental population of *Cuphea lanceolata*. Plant Breeding 106:334-337.

Knapp, S.J., L. Tagliani, and W.R. Roath. 1991b. Fatty acid and oil diversity of *Cuphea viscosissima*: a source of medium-chain fatty acids. Journal of American Oil Chemists Society 68:515-517.

Knapp, S.J., J. Crane, L. Tagliani, and M. Slabaugh. 1997. *Cuphea viscosissima* mutants with decreased capric acid. Crop Science 37:352-357.

Kopka, J., M. Robers, R. Schuch, and F. Spener. 1993. Acyl carrier proteins from developing seeds of *Cuphea lanceolata* Ait. Planta 191:102-111.

Lander, E., P. Green, J. Abrahamson, A. Barlow, M. Daly, S. Lincoln, and L. Newverg. 1987. MAPMAKER: an interactive computer package for constructing primary linkage maps of experimental and natural populations. Genomics 1:174-181.

Larson, T.R., T. Edgell, J. Byrne, K. Dehesh, and I.A. Graham. 2002. Acyl CoA profiles of transgenic plants that accumulate medium-chain fatty acids indicate inefficient storage lipid synthesis in developing oilseeds. Plant J 32:519-527.

Leonard, J.M., M.B. Slabaugh, and S.J. Knapp. 1997. *Cuphea wrightii* thioesterases have unexpected broad specificities on saturated fatty acids. Plant Molecular Biology 34:669-679.

Leonard, J.M., S.J. Knapp, and M.B. Slabaugh. 1998. A *Cuphea* β-ketoacyl-ACP synthase shigts the synthesis of fatty acids towards shorter chains in *Arabidopsis* seeds expressing *Cuphea* FatB thioesterases. The Plant Journal 13:621-628.

Liu, Z., R. Biyashev, and M. Saghai Maroof. 1996. Development of simple sequence repeat DNA markers and their integration into a barley linkage map. Theoretical and Applied Genetics 93:869-876.

Mayer, K.M., and J. Shanklin. 2005. A Structural Model of the Plant Acyl-Acyl Carrier Protein Thioesterase FatB Comprises Two Helix/4-Stranded Sheet Domains, the N-terminal Domain Containing Residues That Affect Specificity and the C-terminal Domain Containing Catalytic Residues. J. Biol. Chem. 280:3621-3627.

Milbourne, D., R.C. Meyer, A.J. Collins, L.D. Ramsay, C. Gebhardt, and R. Waugh. 1998. Isolation, characterisation and mapping of simple sequence repeat loci in potato. Molecular Genetics and Genomics 259:233-245.

Miller, R., F.R. Earle, I.A. Wolff, and Q. Jones. 1964. Search for new industrial oils IX *Cuphea*, a versatile source of fatty acids. Journal of American Oil Chemists Society 41:279-280.

Murphy, D.J. 2002. Biotechnology and the improvement of oils crops--genes, dreams and realities. Phytochemistry Reviews 1:67-77.

Patthy, L.s. 2003. Modular Assembly of Genes and the Evolution of New Functions. Genetica 118:217-231.

Porebski, S., L.G. Bailey, and B.R. Baum. 1997. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. Plant Molecular Biology Reporter 15:8-15.

Ramsay, L.D., M. Macaulay, S. degli Ivanissevich, K. MacLean, L. Cardle, J. Fuller, K. Edwards, S. Tuvesson, M. Morgante, A. Massari, E. Maestri, N. Marmiroli, T. Sjakste, M. Ganal, W. Powell, and R. Waugh. 2000. A simple sequence repeat-based linkage map of barley. Genetics 156:1997-2005.

Salas, J.J., and J.B. Ohlrogge. 2002. Characterization of substrate specificity of plant FatA and FatB acyl-ACP thioesterases. Archives of Biochemistry and Biophysics 403:25-34.

Sanguinetti, D., E. Neto, and A. Simpson. 1994. Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. BioTechniques 17:915-919.

Schuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology 18:233-234.

Schutt, B., A. Abbadi, B. Loddenkotter, M. Brummel, and F. Spener. 2002. β-Ketoacyl-acyl carrier protein synthase IV: a key enzyme for regulation of medium-chain fatty acid synthesis in *Cuphea lanceolata* seeds. Planta 215:847-854.

Slabaugh, M.B., J.M. Leonard, and S.J. Knapp. 1998. Condensing enzymes from *Cuphea wrightii* associated with medium chain fatty acid biosynthesis. Plant J 13:611-620.

Slabaugh, M.B., G.M. Huestis, J. Leonard, J.L. Holloway, C. Rosato, V. Hongtrakul, N. Martini, R. Toepfer, M. Voetz, J. Schell, and S.J. Knapp. 1997. Sequence-based genetic markers for genes and gene families: single-strand conformational polymorphisms for the fatty acid synthesis genes of *Cuphea.* Theoretical and Applied Genetics 94:400-408.

Stuber, C.W., M. Polacco, and M.L. Senior. 1999. Synergy of empirical breeding, marker-assisted selction, and genomics to increase crop yield potential. Crop Science 39:1571-1583.

Tang, S., J.-K. Yu, M.B. Slabaugh, D. Shintani, and S.J. Knapp. 2002. Simple sequence repeat map of the sunflower genome. Theoretical and Applied Genetics 105:1124-1136.

Temnykh, S., W.D. Park, N. Ayres, S. Cartinhour, N. Hauck, L. Lipovich, Y.G. Cho, T. Ishii, and S.R. McCouch. 2000. Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). Theoretical and Applied Genetics 100:697-712.

Thompson, A.E. 1984. *Cuphea*-a potential new crop. HortScience 19:352-354.

Thompson, A.E., D. Dierig, S.J. Knapp, and R. Kleiman. 1990. Variation in fatty acid content and seed weight in some lauric acid rich *Cuphea* species. Journal of American Oil Chemists Society 67:611-617.

Topfer, R., N. Martini, and J. Schell. 1995. Modification of Plant Lipid Synthesis. Science 268:681-686.

Voelker, T.A. 1996. Plant acyl-ACP thioesterases: chain-length determining enzymes in plant fatty acid biosynthesis, p. 111-133, *In* J. Setlow, ed. Genetic Engineering, Vol. 18. Plenum Press, New York.

Voelker, T.A., T.R. Hayes, A.M. Cranmer, J.C. Turner, and H.M. Davies. 1996. Genetic engineering of a quantitative trait: metabolic and genetic parameters influencing the accumulation of laurate in rapeseed. The Plant Journal 9:229-241.

Voelker, T.A., A.C. Worrell, L. Anderson, J. Bleibaum, C. Fan, D.J. Hawkins, S.E. Radke, and H.M. Davies. 1992. Fatty acid biosynthesis redirected to medium chains in transgenic oilseed plants. Science 257:72-74.

Webb, D., S.J. Knapp, and L. Tagliani. 1992. An RFLP and allozyme linkage map of *Cuphea lanceolata*. Theoretical and Applied Genetics 83:528-532.

Wu, K., and S. Tanksley. 1993. Abundance, polymorphism and genetic mapping of microsatellites in rice. Molecular Genetics and Genomics 245:187-194.

Xiao, M., A. Phong, K.L. Lum, R.A. Greene, P.R. Buzby, and P.-Y. Kwok. 2004. Role of excess inorganic pyrophosphate in primer-extension genotyping assays. Genome Research 14:1749-1755.

Yu, J.-K., S. Tang, M.B. Slabaugh, A. Heesacker, G. Cole, M. Herring, J. Soper, F. Han, W.-C. Chu, D.M. Webb, L. Thompson, K. Edwards, S. Berry, A.J. Leon, M. Grondona, C. Olungu, N. Maes, and S.J. Knapp. 2003. Towards a saturated molecular genetic linkage map for cultivated sunflower. Crop Science 43:367-387.

Ziegle, J.S., Y. Su, K.P. Corcoran, L. Nie, P.E. Mayrand, L.B. Hoff, L.J. McBride, M.N. Kronick, and S.R. Diehl. 1992. Application of automated DNA sizing technology for genotyping microsatellite loci. Genomics 14:1026-1031.

APPPENDIX

Assessment of amplification of SSR markers on a panel of six genotypes. Markers which amplified an allele are denoted Y, those that did not are denoted N and missing data is denoted by M.

	Genotype					
Marker	VS55	LN193	LN237	LN185	PSR108-2	PSR23
ORC12	N	N	N	N	N	N
ORC14	Y	Y	Y	Y	Y	Y
ORC16	Y	Y	Y	Y	Y	Y
ORC17	Y	Y	Y	Y	Y	Y
ORC18	Y	Y	Y	Y	Y	Y
ORC19	Y	Y	Y	Y	Y	Y
ORC20	Y	Y	Y	Y	Y	Y
ORC21	Y	Y	Y	Y	Y	Y
ORC22	Y	Y	Y	Y	Y	Y
ORC23	Y	Y	Y	Y	Y	Y
ORC26	N	N	N	N	N	N
ORC27	N	Y	Y	Y	Y	Y
ORC30	N	N	N	Y	N	N
ORC31	N	Y	Y	Y	Y	Y
ORC32	N	N	N	N	N	Y
ORC33	N	N	N	N	N	N
ORC34	Y	Y	Y	Y	Y	Y
ORC37	Y	Y	Y	Y	Y	Y
ORC38	Y	Y	Y	Y	Y	Y
ORC39	Y	Y	Y	Y	Y	N
ORC40	Y	Y	Y	Y	Y	Y
ORC41	Y	Y	Y	Y	Y	Y
ORC43	Y	Y	Y	Y	Y	Y
ORC46	N	N	N	N	N	N
ORC47	Y	Y	Y	Y	Y	Y
ORC48	N	N	N	N	N	N
ORC51	N	N	N	N	N	N
ORC54	N	Y	Y	Y	Y	Y
ORC55	Y	Y	Y	Y	M	Y
ORC56	N	N	N	N	N	N
ORC58	Y	Y	Y	Y	Y	Y
ORC60	Y	Y	Y	Y	Y	Y
ORC61	Y	Y	Y	Y	Y	Y
ORC62	Y	N	N	N	Y	Y
ORC64	Y	Y	Y	Y	Y	Y
ORC65	Y	Y	Y	Y	Y	Y
ORC68	N	N	N	N	N	N
ORC69	N	N	N	N	N	N
ORC70	Y	Y	Y	Y	Y	Y
ORC71	N	N	N	N	N	N
ORC73	Y	Y	Y	Y	Y	Y
ORC76	Y	Y	Y	Y	Y	Y
ORC77	N	N	N	N	N	N
ORC78	Y	Y	Y	Y	Y	Y

Continued

Marker	Genotype					
	VS55	LN193	LN237	LN185	PSR108-2	PSR23
ORC80	Y	Y	Y	Y	Y	Y
ORC82	Y	Y	Y	Y	Y	Y
ORC83	N	N	N	N	N	N
ORC86	N	N	N	N	N	N
ORC87	Y	Y	Y	Y	Y	Y
ORC89	Y	Y	Y	N	Y	Y
ORC90	Y	Y	Y	Y	Y	Y
ORC91	Y	Y	Y	Y	Y	Y
ORC92	Y	Y	Y	Y	Y	Y
ORC95	Y	Y	Y	N	Y	Y
ORC98	Y	Y	Y	N	Y	Y
ORC100	Y	Y	Y	Y	Y	Y
ORC101	N	N	N	N	N	N
ORC102	Y	Y	Y	Y	Y	Y
ORC103	Y	Y	Y	Y	Y	Y
ORC104	N	N	N	N	N	N
ORC107	N	N	N	N	N	N
ORC108	Y	Y	Y	N	Y	N
ORC110	Y	Y	Y	Y	Y	Y
ORC112	Y	Y	Y	Y	Y	Y
ORC113	N	N	N	N	N	N
ORC114	N	N	N	N	N	N
ORC117	N	N	N	N	N	N
ORC118	N	N	N	N	N	N
ORC121	N	N	N	N	N	N
ORC123	N	N	N	N	N	N
ORC124	Y	Y	Y	Y	Y	Y
ORC126	Y	Y	Y	Y	Y	Y
ORC127	Y	Y	Y	N	Y	Y
ORC128	N	N	N	N	N	N
ORC129	N	N	N	N	N	N
ORC130	N	N	N	N	N	N
ORC131	N	N	N	N	N	N
ORC132	N	N	N	N	N	N
ORC133	N	N	N	N	N	N
ORC134	N	N	N	N	N	N
ORC135	N	N	N	N	N	N
ORC137	Y	Y	M	N	M	N
ORC138	N	N	N	N	N	N
ORC142	N	N	N	N	N	N
ORC143	Y	Y	Y	Y	Y	Y
ORC145	Y	Y	Y	Y	Y	Y
ORC146	Y	Y	Y	Y	Y	Y
ORC147	Y	Y	Y	M	Y	Y
ORC149	N	N	N	N	Y	N
ORC151	N	N	N	N	N	N

Continued

Marker	Genotype					
	VS55	LN193	LN237	LN185	PSR108-2	PSR23
ORC153	N	Y	Y	Y	Y	Y
ORC154	N	N	N	N	N	N
ORC157	N	N	N	N	N	N
ORC159	Y	Y	Y	Y	Y	Y
ORC160	N	N	N	N	N	N
ORC161	N	N	N	N	N	N
ORC162	Y	N	Y	Y	Y	Y
ORC163	Y	N	Y	N	N	N
ORC164	Y	Y	Y	Y	Y	Y
ORC165	N	Y	Y	Y	Y	N
ORC166	Y	Y	Y	Y	Y	Y
ORC171	Y	Y	Y	Y	Y	Y
ORC172	Y	Y	Y	Y	Y	Y
ORC173	Y	Y	Y	Y	Y	Y
ORC174	N	N	N	N	N	N
ORC175	N	Y	Y	N	Y	Y
ORC176	N	Y	Y	Y	Y	Y
ORC181	Y	Y	Y	Y	Y	Y
ORC182	Y	Y	Y	Y	Y	Y
ORC183	Y	Y	Y	Y	Y	Y
ORC184	N	N	N	N	N	N
ORC186	N	N	N	N	N	N
ORC187	Y	Y	Y	Y	Y	Y
ORC189	N	N	Y	Y	Y	Y
ORC192	Y	Y	Y	N	Y	Y
ORC196	Y	Y	N	N	N	N
ORC198	N	N	N	N	N	N
ORC199	N	Y	Y	Y	Y	Y
ORC200	N	N	N	N	N	N
ORC201	N	N	N	N	N	N
ORC202	N	Y	M	N	Y	Y
ORC204	N	N	N	N	Y	Y
ORC205	M	Y	Y	Y	M	M
ORC206	N	N	N	N	N	N
ORC207	M	Y	M	M	M	M
ORC208	Y	Y	Y	Y	Y	Y
ORC209	N	N	N	N	N	N
ORC210	Y	Y	Y	Y	Y	Y
ORC211	N	N	N	N	N	N
ORC212	Y	Y	Y	Y	Y	Y
ORC213	Y	Y	Y	Y	Y	Y
ORC214	Y	Y	Y	Y	Y	Y
ORC215	Y	Y	Y	Y	Y	Y
ORC216	Y	N	Y	Y	Y	Y
ORC217	N	N	N	N	N	N
ORC218	Y	Y	Y	Y	Y	Y

Continued

Marker	Genotype					
	VS55	LN193	LN237	LN185	PSR108-2	PSR23
ORC219	N	Y	Y	Y	Y	Y
ORC220	N	N	N	N	N	N
ORC221	Y	Y	Y	Y	Y	Y
ORC222	Y	Y	Y	N	Y	Y
ORC223	Y	Y	Y	Y	Y	Y
ORC224	Y	Y	Y	Y	Y	Y
ORC225	Y	Y	Y	Y	Y	Y
ORC226	Y	Y	Y	N	Y	Y
ORC227	Y	N	Y	N	Y	Y
ORC228	Y	N	N	N	N	M
ORC229	N	Y	Y	N	Y	Y
ORC230	N	N	N	N	N	N
ORC231	N	Y	Y	Y	N	Y
ORC232	N	N	N	N	N	N
ORC234	N	N	N	N	N	N
ORC239	Y	Y	Y	Y	Y	Y
ORC241	Y	N	Y	Y	Y	Y
ORC242	Y	Y	Y	Y	Y	Y
ORC243	N	N	N	N	N	N
ORC244	N	Y	Y	N	M	M
ORC245	Y	Y	Y	Y	Y	Y
ORC246	Y	Y	Y	Y	Y	Y
ORC247	Y	Y	Y	Y	Y	Y
ORC251	Y	Y	Y	Y	Y	Y
ORC252	Y	Y	Y	Y	Y	Y
ORC253	Y	Y	Y	Y	Y	Y
ORC254	N	Y	N	Y	Y	Y
ORC255	Y	Y	Y	Y	Y	Y
ORC256	Y	Y	Y	Y	Y	Y
ORC257	N	Y	Y	Y	Y	Y
ORC258	Y	Y	Y	Y	Y	Y
ORC259	N	N	N	N	N	Y
ORC260	N	Y	N	N	Y	Y
ORC261	N	N	N	N	N	N
ORC262	N	N	N	N	N	N
ORC263	Y	Y	Y	Y	Y	Y
ORC264	Y	Y	Y	Y	Y	Y
ORC270	N	N	N	N	N	N
ORC271	Y	Y	Y	Y	Y	Y