AN ABSTRACT OF THE THESIS OF

Venkata Krishna Kishore for the degree of Doctor of Philosophy in Crop Science presented on March 21, 2002. Title: Mapping Quantitative Trait Loci Underlying Genome-Wide Recombination Rate and Mating System Differences in Meadowfoam

Redacted for Privacy

Abstract approved: _____

Steven J. Knapp

Meadowfoam (*Limnanthes alba* Bentham; Order: Brassicales; Family: Limnanthaceae) is a self-compatible, predominantly allogamous, insect pollinated species. Meadowfoam oil is a source of novel unsaturated very-long-chain (VLC) seed oils (C_{20} and C_{22}) with low concentrations of saturated fatty acids (typically less than 2%) and outstanding oxidative stability. Here we report the development of 389 SSR markers for meadowfoam. All the 389 SSRs were screened on 14 meadowfoam germplasm accessions to assess their utility and efficiency. Ninety-six percent of the SSR markers (373 out of 389) were polymorphic among the 14-germplasm accessions (from nine taxa) with a mean heterozygosity of 0.63.

We also report that the physical size of the meadowfoam genome was estimated to be 5.52 pg using flow cytometry; thus, the meadowfoam genome is ca. 16 times larger than the *Arabidopsis* genome. Karyotype analyses revealed that the meadowfoam genome is made up of two metacentric and three submetacentric chromosomes. Meadowfoam has two pairs of chromosomes with subterminal

nucleolar organizing regions (NOR's). A genetic map comprised of 84 SSR loci dispersed among five linkage groups with 11 to 22 SSR loci per linkage (6 SSR loci segregated independently) was constructed. The map was 988.7 cM long with a mean density of 11.8 cM and minimal clustering of loci.

A total of 20 quantitative trait loci (QTL) were identified for five mating system characters in meadowfoam, using the SSR linkage map of meadowfoam. Individual QTL for mating system traits [petal area (pa), seeds per plant (spp) and seeds per flower (spf)] account for up to 20% of the backcross phenotypic variance, with most traits showing QTL effects of 5-15%. The QTL for protandry and chiasma frequency were adjacent to the QTL for *spp* and *spf*. This study has provided evidence that the correlation between the chiasma frequency and the type of mating system is not a direct developmental relationship between these factors, but is due to a selective advantage of the combination of the characters found. The speculation that the genetic factors underlying chiasma frequency and autonomous seed set have co-evolved during evolution negates the self-fertilization as an "evolutionary dead end".

Mapping Quantitative Trait Loci Underlying Genome-Wide Recombination Rate and Mating System Differences in Meadowfoam

by

Venkata Krishna Kishore

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented March 21, 2002
Commencement June 2002

Doctor of Philosophy thesis of Venkata Krishna Kishore presented on March 21, 2002

APPROVED:

Redacted for Privacy

Major Professor, representing Crop Science

Redacted for Privacy

Chair of the Department of Crop and Soil Science

Redacted for Privacy

Dean of 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

Venkata Krishna Kishore, Author

ACKNOWLEDGMENTS

I am completely indebted to Dr. Steven J. Knapp, my research guide and major professor, for his uncountable practical suggestions during the conduct of research work, timely insights throughout the writing of this text, and critical review of each and every line of the text. I am also grateful to him for his encouragement and enthusiasm, which inspired me to work better. I also thank him for the financial support he provided throughout my doctoral program. I am thankful to members of my graduate committee, Dr. Patrick M. Hayes, Dr. James C Peterson, Dr. Tony Chen and Dr. Gregory Thompson for their time and guidance.

Special thanks and gratitude are extended to Jimmie Crane – "The walking and talking encyclopedia of meadowfoam" for sharing the information about cultivation of meadowfoam and also helping with the greenhouse work. I am thankful to Dr. Mary B. Slabaugh, who taught me all aspects of molecular techniques and was always generous with her time. Her encouragement, support, concern, and warm heart will be forever appreciated. Thanks are due to the individuals in the lab: Sonali, Shunxue, Pablo, Judy, Ju-Kyung, Felix, Cate and Adam for their warm smile, kindness and friendship throughout my study period. I am also grateful to Dr. Pablo Velasco-Pazos for his unreserved help, and technical suggestions especially during the preparation of the manuscript. The faculty and staff of the Department of Crop Science who directly or indirectly contributed to the success of this work are gratefully acknowledged.

I wish to thank Caprice Rosato who assisted in automated genotyping of SSR markers using GeneScan and Genotyper software. Help from Nanci Adair, and Joan

Rowe for sequencing DNA is appreciated. Appreciation is extended to Beth Parsons, Elias Sabry, Corwin Willard and Julie Oughton for their help with flow cytometry analysis. I am also thankful to Al Soeldner for his help with the electron microscopy analysis of the buds and flowers of meadowfoam. I owe a debt of gratitude to Dr. Oscar Riera-Lizarazu and Dr. Isabel Vales for allowing me to use their microscope facilities and help in getting started with the karyotype analysis.

The funding and support from USDA grant # 99-34407-7509 to Dr. S.J. Knapp is also acknowledged.

Though far from here, my work owes completely to my parents, sister, brother and my in-laws, for whom my success is their success. Special thanks are due to my brother Shreenivasarao, who took over and fulfilled my responsibilities towards our family during the course of my doctoral program.

Last but not the least, I thank my better half of the life "Sonali D. Gandhi", for her love and constant encouragement at home as well as at work. She has been an excellent team worker with me right from my bachelor's degree. I wouldn't have achieved whatever little I have so far without her patience, understanding and solid support.

CONTRIBUTION OF AUTHORS

Dr. Steven J. Knapp initiated, advised, and supervised all aspects of the projects. He substantially assisted in formulating hypotheses and revising manuscripts. Sonali D. Gandhi helped in acquiring data for karyotype and chaismata frequency analysis. She also helped in phenotyping the mating system traits. Jimmie Crane helped with all the technical aspects of growing meadowfoam in the greenhouse. He developed meadowfoam lines OMF64, OMF40-11 and OMF109-3 in association with Dr.Steven J. Knapp. Dr. Pablo Velasco helped in SSR screening for the SSR project. Dr. Mary B. Slabaugh contributed and helped in all aspects of molecular techniques and analyses.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

LIST OF FIGURES

Figure Page

LIST OF FIGURES (CONTINUED)

Figure Page

LIST OF TABLES

Mapping Quantitative Trait Loci Underlying Genome-Wide Recombination Rate and Mating System Differences in Meadowfoam

CHAPTER 1

INTRODUCTION

Cultivated meadowfoam (*Limnanthes alba* Benth.) is an annual oil seed crop plant native to Southern Oregon and Northern California (Mason 1952, Kalin 1971, Jain 1986). Meadowfoam belongs to Order Brassicales and Family Limnanthaceae (Wheeler et al. 2000). The genus *Limnanthes* is comprised of 17 diploid ($2n = 2x = 10$) species and subspecies belonging to two sections (Inflexae and Reflexae). *Limnanthes alba* Benth. belongs to section Inflexae. Meadowfoam seed oil contains unique unsaturated very long chain fatty acids (C_{20} and C_{22}) with outstanding oxidative stability (Isbell 1997). These novel characteristics of meadowfoam seed oil are of industrial interest and have fueled the development of meadowfoam as an oilseed crop (Knapp and Crane 1999).

Meadowfoam has been cultivated since 1974, when the first non-shattering cultivar, Foamore, was developed and released for commercial production (Calhoun and Crane 1975). Since then, four more cultivars have been released for cultivation in a span of 27 years. In order to improve our efficiency of developing new cultivars for increasing the productivity of meadowfoam it is essential to understand the genome of meadowfoam along with the genetics of economically

important traits. Our lab has focused on developing molecular breeding tools for meadowfoam for the past six years, as there were very few biochemical markers (Arroyo 1975; Brown and Jain 1979; Kesseli and Jain 1985) and virtually no DNA markers for meadowfoam until 2001. There is also no information published about the physical and cytological characteristics of the meadowfoam genome. With the advent of flow cytometry, one can easily quantify DNA in a large number of plants within a relatively short period of time compared to microdensitometry (Arumuganathan and Earle 1991a; 1991b). However the physical genome size of meadowfoam is not known yet. Propach (1934) and Resende (1937) studied the chromosome morphology of meiotic chromosomes of *Limnanthes douglasii* and *Limnanthes alba*. They disagreed on the number of satellite chromosomes in meadowfoam. Mason (1952) also studied the meiotic chromosome morphology and agreed with Resende (1937) that meadowfoam has only two pairs of satellite chromosomes. However, none of these investigators produced a detailed karyotype of meadowfoam chromosomes.

Genetic maps are important in plant breeding and are a powerful tool for localizing and isolating genes underlying both simple and complex traits. Katengam *et al.* (2002) produced the first genetic map of meadowfoam comprising of 103 amplified fragment length polymorphisms (AFLPs). AFLPs, however are dominant markers and difficult to use in regular breeding programs (Powell et al. 1996; Smith et al. 1997). Simple sequence repeat (SSR) markers have become

important genetic markers in many plant genomes because of the high level of polymorphism (Wang et al. 1994). In addition, they are polymerase chain reaction (PCR) based, facilitating easy screening compared to restriction fragment length polymorphisms (RFLPs). They can be easily scored and transferred among different laboratories. Genetic maps based on SSRs have been developed for wheat (Roder *et al.* 1998), rice (Temnykh *et al.* 2000), barley (Ramsay *et al.* 2000), sunflower (Tang *et al.* 2002) and many other crop plants.

The use of molecular markers and genetic maps has simplified the genetic analysis of quantitative traits. Quantitative trait loci (QTL) mapping has played an important role not only in breeding for better varieties but also in understanding various evolutionary processes. Bradshaw et al. (1995) and Lin and Ritland (1997) used QTL analysis to study the genetic factors underlying the reproductive isolation mechanisms in *Mimulus*. The mating systems of meadowfoam have been widely studied (Arroyo, 1975; Brown and Jain 1979; Brown et al. 1979; Kesseli and Jain 1985, Jain 1978; McNeill, 1983; McNeill and Jain 1985, Ritland, 1984). These studies have shown that the mating systems in meadowfoam vary from a predominantly allogamous (*L.alba*) to a completely autogamous (*L. floccosa*) system. Protandry and heterostyly are the major reasons for allogamy whereas cleistogamy is the major reason for autogamy. Commercial meadowfoam (*L.alba*) fields are primarily pollinated by domestic honeybees (*Apis millifera* L.). Poor seed yield have been routinely blamed on lack of pollination by the bees. Hence with

the long-range goals of developing self-pollinated meadowfoam, our lab undertook the study of genetics of self-pollination in meadowfoam.

The major goals for this study, which would be stepping stones towards the long range goals of the lab mentioned above are listed below:

Our goals (Chapter 2, 3, and 4) were

i) Development of public databases for sequence-based, highly polymorphic SSR markers using enriched small insert genomic libraries.

ii) Assessment of the efficiency of SSRs as genetic markers.

iii) Analyses of the genetic relationships among 14 meadowfoam accessions using the newly developed SSR markers.

iv) Describe the karyotype and physical size of the meadowfoam genome.

v) Develop a simple sequence repeat (SSR) map for meadowfoam.

vi) Elucidate the anatomical, developmental, genetical and evolutionary mechanisms underlying the shift in the mating system of meadowfoam from allogamy to autogamy.

CHAPTER 2

THE DEVELOPMENT AND UTILITY OF SIMPLE SEQUENCE REPEAT MARKERS FOR DNA FINGERPRINTING AND HIGH-THROUGHPUT GENOTYPING IN MEADOWFOAM

V.K.Kishore[1], P.Velasco[1], M.B. Slabaugh[1] and S.J. Knapp[1]

[1]Department of Crop and Soil Science, Oregon State University, Corvallis, Oregon, 97331, USA

Abstract

DNA markers, especially high-throughput sequence-based DNA markers, are powerful tools for intellectual property protection, genetic diversity analysis, genetic resource conservation management, cultivar identification, and molecular breeding in crop plants.. We isolated and sequenced 1,596 clones from genomic DNA libraries enriched for AG_n or AC_n. Six hundred and ninety-six clones harbored unique SSRs ranging in length from 6 to 150 bp; primers were designed and tested for 624 unique SSRs. Three hundred and eighty-nine primer pairs (62.3%) produced clean amplicons and yielded functional SSR markers. Ninety-six percent of the SSR markers (373 out of 389) were polymorphic among the 14-germplasm accessions (from nine taxa) and the heterozygosity (H) and power of discrimination (PD) scores ranged from 0.0 to 0.93 with a mean of 0.63 and 0.64 respectively. A strong positive correlation was observed between the Standard deviation of molecular weight (SDmw), H and maximum repeat count (MRC). Genetic distances ranged from 0.35 ± 0.022 to 0.69 ± 0.043 with a mean of 0.57 ± 0.030. Cluster and principal component analyses of the genetic distance matrix uncovered patterns of diversity concordant with species, subspecies, and breeding origin. The SSR markers described herein are powerful tools for efficiently and accurately discriminating between genotypes.

Keywords: SSR, repeat length, meadowfoam, diversity, endangered

Introduction

Meadowfoam (*Limnanthes alba* Benth.) is one of the diploid (x=5) species belonging to Brassicales order and Limnanthaceae family (Mason 1952; Ornduff and Carvello 1968). The genus *Limnanthes* is comprised of 17 species and subspecies belonging to two sections (Inflexae and Reflexae). *Limnanthes alba* Benth. belongs to section Inflexae. The *Limnanthes* are native to the vernal pools of California, southern Oregon and British Columbia. The seed oils of meadowfoam contain unique unsaturated very long chain fatty acids (C_{20}-C_{22}) with high oxidative stability (Smith *et al.* 1960, Bugby *et al.* 1961, Isbell 1997). Meadowfoam was first domesticated in 1973 (Jain 1986). It has been seriously cultivated for less than a decade, but has gained momentum due to progress in the marketing and development of specialty chemicals (Isbell 1997) from its novel seed oil. The cultivation of meadowfoam has also received a boost due to the release of improved varieties from this lab (Knapp and Crane 1999). In order to maintain and excel this tradition of meadowfoam cultivar improvement, it is necessary to have information about the amount of variation present in breeding material and exotic germplasm. The knowledge of the genetic relationship among breeding materials could help avoid the great risk of increasing uniformity in the elite germplasm and could ensure long term selection gains (Messmer *et al.* 1993). Further more, examining the genetic variability within a gene pool of exotic and elite breeding

material could make crop improvement more efficient by directed accumulation of favorable alleles.

Meadowfoam is a genus of annual, herbaceous wildflowers native to California, Southern Oregon, and British Columbia. Several rare and endangered species persist in the few remaining vernal pools once common in California's central valley, foothills and Southern Oregon. *Limnanthes flocossa* ssp. *californica*, *Limnanthes flocossa* ssp. *grandiflora, Limnanthes flocossa* ssp. *pumila, Limnanthes gracilis* ssp. *parishii*, *Limnanthes douglasii* ssp. *sulphurea* and *Limnanthes vinculans* are listed as endangered and *Limnanthes gracilis* ssp. *gracilis, Limnanthes bakeri* as rare (Eastman 1990; Dole and Sun, 1992). All have suffered from extensive agricultural and urban development. To maintain genetic diversity in conservation programs, it is essential that the genetic diversity be catalogued by using DNA markers.

Genetic diversity of germplasm collections can be established from pedigree records, morphological traits, isozyme and DNA markers (Smith *et al.* 1990; Mumm and Dudley 1994). Pedigree data are usually biased, as they don't account for mutation and random genetic drift. Moreover they are sometimes unavailable or erroneous due to the human error factor involved in maintaining the records. Isozyme markers have often failed in classification or identification of breeding genotypes because of poor genome coverage (Smith *et al.* 1990, Dudley 1993). Genetic variation in *Limnanthes* ssp. has been reported using 11 to 18

allozyme markers (Brown and Jain 1979, Mcneill and Jain 1983, Kessili and Jain 1985, Ritland and Jain 1984). Forty-one accessions of meadowfoam from section Inflexae were fingerprinted based on 176 Amplified Fragment Length Polymorphisms (AFLPs) by Katengam (2001). However AFLPs are shown to have less polymorphic information content than Restriction Fragment Length Polymorphisms (RFLPs) and Simple Sequence Repeats (SSRs). Moreover they are dominant markers and are technically difficult to use in regular breeding programs (Powell *et al.* 1996, Smith *et al.* 1997). So an ideal molecular genetic marker for plant genome analysis would be one that discloses multiple alleles (codominant), has an even distribution through out the genome and easily differentiate genetically similar individuals, be relatively easy to score and easily transferred among different laboratories. Microsatellites (Litt and Luty 1989) or SSRs (Weber and May 1989) possess all these characteristics.

Microsatellites are tandem arrays of short nucleotide repeats from 1 to 5 bases per unit. Simple sequence length polymorphisms (SSLPs) are based on the differences in the number of the DNA repeat units at a given locus and provide a valuable source of genetic markers. Microsatellites have been extensively exploited for genome mapping and for a wide range of population and evolutionary studies in human (Bowcock *et al.* 1994), mouse (Dietrich *et al.* 1996), *Drosophila* (Goldstein and Clark 1995, Schug *et al.* 1997), *Arabidopsis* (Innan *et al.* 1997) rice (Yang *et al.* 1994), sunflower (Yu *et al.* 2001, Tang *et al.* 2001) and other animal and plant

species. Additionally SSRs have been used for genotype identification and for purposes of intellectual property rights protection (Rongwen *et al.* 1995; Smith *et al.* 2000).

To our knowledge, no microsatellite marker has been yet reported in meadowfoam, likely due to the recent domestication of this crop and the large initial effort that these markers require. The aim of this research was to develop a large set of SSRs from enriched small insert genomic libraries, with a focus on the integration of advanced technology and improved analysis of the data. We have combined the informativeness of SSR markers with the precision of automated allele sizing of fluorescently labeled SSR alleles offering a rapid and standardized system for profiling DNA from fourteen germplasm accessions from several meadowfoam species and subspecies. Thus we report here a high throughput process of developing, screening and improved data analysis of 389 SSRs for meadowfoam.

Materials and Methods

Construction of small insert genomic libraries enriched for AG – or AC – repeats

Genomic DNA was isolated from the leaf tissue of OMF40-11 (*Limnanthes alba* ssp. *alba*) and OMF64 (*Limnanthes alba* ssp. *versicolor*) according to Lodhi *et al.* (1994) with minor modifications. The genomic DNA from both sources was pooled and sent to Genetic Identification Services (GIS). The SSR-enriched libraries were constructed by GIS according to the methods developed by Karagyozov *et al.* (1993) and Edwards *et al.* (1996). Several blunt-end restriction enzymes were used to partially digest genomic DNA to make sure that the enriched SSRs were evenly distributed in whole genome. The DNA fragments in the range 300 ~ 800 bp were collected for DNA library construction. These fragments were ligated to two DNA adapters with Hind III cloning sites. The ligated fragments were amplified using one adaptor sequence as a primer. The amplified fragments were captured by magnetic beads with targeted SSR sequences $(AG)_{15}$ or $(AC)_{15}$. The captured fragments were released and amplified by polymerase chain reaction (PCR) again, followed by a second round enrichment. The captured small inserts were released from magnetic beads by eluting into solution and ligated them into HindIII cloning sites of PUC19 plasmids. The plasmids were transformed into the *E. coli* strain DH5α, and the yield of recombinant cells was typically 5,000 ~ 10,000 recombinant cells/1.4mL library.

Colony PCR and DNA Sequencing

Fifty to one hundred microliters of the original stock solution of enriched library was spread on LB-agar media with 75μg/ml of ampicillin. The screening of recombinant clones was assessed through the IPTG and X-Gal procedure. White colonies were transferred to ampicillin–LB agar medium in 96-well format plates. The plates were incubated at 37°C overnight and the clone scrapes were used as templates for colony PCRs. The meadowfoam inserts were amplified using M13 universal forward (5'-CGCCAGGGTTTTCCCAGTCACGAC-3') and reverse (5'-TCACACAGGAAACAGCTATGAC-3') primers. The PCR amplifications were performed with Perkin Elmer 9600 or MJ PTC 200 thermocyclers. Each reaction was a total volume of 35 μL containing 1x PCR buffer, 2 mM Mg^{++}, 0.2 μM each of dNTPs, 1% Tween-20, 1.5 mM Cresol Red, 3% sucrose, 0.5 μM of each primer, 0.75 unit Taq polymerase (Qiagen, Valencia, CA, USA) and template DNA i.e. colony cell scrapes. The PCR plates were denatured at 95 °C for 3 minutes, followed by 35 cycles of 94 °C for 30 seconds, 57 °C for 30 seconds and 72 °C for one minute with a final extension for 10 minutes. After PCR, 5 μL PCR solution was loaded on to 1.5% agarose gel to check the insert number and size of each clone. Only the clones with a single insert ranging from 350 ~ 800 bp were selected as the inserts below 350bp were less likely to have enough sequences flanking a microsatellite for primer design, whereas inserts larger than 800 bp cannot be completely sequenced at one time. The clones that had more than one insert were discarded, as they cannot be sequenced accurately. The selected PCR products were purified using QIAquick PCR purification kit (Qiagen, Valencia, CA, USA). Half of the purifications were processed on a Qiagen BioRobot 3000 (Qiagen, Valencia,

CA, USA). About 25 ~ 50 ng purified PCR products were sequenced on ABI 377 or 3700 automated sequencer using a dye terminator fluorescent kit. Universal M13 forward primer (1.5 - 3.0 pmoles) was used for sequencing the PCR products. It is important to mention that the addition of 5% dimethyl sulfoxide (DMSO) as denaturant in the reaction mixture helped in improving the sequencing quality of the clones harboring SSRs with repeat sizes more than 40bp.

Sequence Analysis and Primer Design

The sequence analysis was carried using the SEQLAB module of the Genetics Computing Group (GCG) Software version.10.1 (Madison, WI). The sequences harboring SSRs (repeat number ≥6bp) were trimmed to remove the plasmid sequence and were checked for redundancy among the sequences using the PILEUP function. The unique sequences were divided into 3 categories: perfect, imperfect, and compound repeat sequences according to the standards of Weber (1990). Primers complementary to the flanking regions of the repeats were designed using PRIMER 3.0 software (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi). The primers were 20 ~ 27 bp long (optimum 22 bp) with G-C contents in the 35 ~ 60% (optimum 45%) range, T_m in the 59 ~ 63 °C (optimum 61 °C), and a primer pair Tm variance of 2.0 °C, maximum 3' stability 9.0, maximum self-complementarity 6.00, maximum 3' self-complementarity 2.0, and max ploy-X 4.0. Since one of our operational goals were to multiplex the PCR products before they were run on the gels, for multiplexing efficiency and high throughput genotyping, the primer pairs were designed systematically to vary the expected length of the PCR products from 108 bp to 554 bp in about 100 bp

increments. The 5' ends of all the forward primers were modified with one of three fluorescent tags, 6-FAM, HEX, or TET. Primers amplifying overlapping fragment sizes were tagged with different fluorescent dyes. MWG Biotech Inc. synthesized all the primers.

PCR profiling and Genotyping

A test array of 14 *Limnanthes* accessions (Table 2.1) was selected to represent the diversity in the entire *Limnanthes* genus. Five species and nine sub species (*Limnanthes alba* ssp. *alba*, *Limnanthes alba* ssp. *versicolor*, *Limnanthes floccosa* ssp. *floccosa*, *Limnanthes floccosa* ssp. *grandiflora*, *Limnanthes gracilis* ssp. *gracilis*, *Limnanthes gracilis* ssp. *parishii*, *Limnanthes montana* from the section inflexae and *Limnanthes douglasii* ssp. *douglasii*, *Limnanthes douglasii* ssp. *niveai*, *Limnanthes douglasii* ssp. *rosea* from section reflexae) from both sections of the genus were represented. For inbred accessions, ten plants were grown in the greenhouse and their leaf tissue was pooled for DNA isolation, however for outbreds a single plant was grown in the greenhouse for DNA isolation. Genomic DNA was isolated from the leaf tissue according to Lodhi *et al.* (1994) with minor modifications. Each PCR reaction was performed in 20 μL volume containing 1 x PCR buffer, 2.5 mM Mg^{++}, 0.2 μM each of dNTPs, 0.1% Tween-20, 5 - 7.5 pmoles of each primer, 0.75 unit Taq polymerase (Qiagen) and 15 - 20 ng of meadowfoam genomic DNA in a Perkin Elmer 9600 or MJ PTC 200 thermocyclers. A 'touchdown' PCR protocol (Don *et al.* 1991) was used to eliminate non-specific amplification. The PCR profile consisted of: (i) initial denaturation step at 95 °C for 3 minutes, followed by 1 cycle of 94 °C for 30 seconds, 68 °C for 30 seconds

and 72 °C for one minute. (ii) In each subsequent cycle, the annealing temperature was decreased 1°C till it reached 58 °C (the annealing temperature varied from 53 to 58 °C according to the Tm of different primer pairs). (iii) The amplification was continued for 30 cycles at 94 °C for 30 seconds, 58 °C for 30 seconds and 72 °C for one minute with a final extension for 20 minutes. Five microliters of the amplified PCR products were loaded onto 1.5% agarose gel to check the amplification efficiency of the primer pairs. Genotyping was done on a sequencing gel on an ABI 377 equipped with GeneScan Analysis version 2.1 (Applied Biosystems, Perkin Elmer, Foster City, CA). Three different amplified PCR products with 6-FAM, HEX, or TET fluorescent tags were multiplexed in one well, and diluted 20 times.

Table 2.1. Germplasm accessions of meadowfoam screened for simple sequence repeat (SSR) polymorphisms.

Accession	Section	Botanical name	Description
LDD	Reflexae	*Limnanthes douglasii spp. douglasii*	Allogamous and outbred (wild species)
LDN	Reflexae	*Limnanthes douglasii spp. nivea*	Allogamous and outbred (wild species)
LDR	Reflexae	*Limnanthes douglasii spp. rosea*	Allogamous and outbred (wild species)
LM	Inflexae	*Limnanthes montana*	Allogamous and outbred (wild species)
LGG	Inflexae	*Limnanthes gracilis spp. gracilis*	Allogamous and outbred (wild species)
LGP	Inflexae	*Limnanthes gracilis spp. parishii*	Autogamous (wild species)
LFF	Inflexae	*Limnanthes floccosa spp. floccosa*	Autogamous (wild species)
LFG	Inflexae	*Limnanthes floccosa spp. grandiflora*	Autogamous (wild species)
LE76	Inflexae	*Limnanthes alba spp. alba*	Allogamous and inbred
OMF40-11	Inflexae	*Limnanthes alba spp. alba*	Allogamous and inbred
OMF86	Inflexae	*Limnanthes alba spp. alba*	Allogamous and outbred
OMF156	Inflexae	*Limnanthes alba spp. alba*	Allogamous and outbred
OMF 64	Inflexae	*Limnanthes alba spp. versicolor*	Autogamous and inbred
OMF109-2	Inflexae	*Limnanthes alba spp. versicolor*	Autogamous and inbred (selected from Mermaid x OMF62 /OMF64)

Samples containing 0.5 μL diluted PCR products, 0.2 μL GeneScan 500 internal lane standard labeled with TAMRA and 50% formamide were heated to 92 °C for 5 minutes, chilled on ice for 5 minutes and then loaded on a denaturing sequencing gel with 48-well format on a ABI 377 automated sequencer with filter C. Amplified PCR products tagged with 6-FAM, TET and HEX showed up as blue, green and yellow colored products whereas the internal lane standard TAMRA 500 showed up as red product after excitement. GeneScan Analysis version 2.1 and Genotyper version 2.0 (Applied Biosystems, Perkin Elmer, Foster City, CA) were used for automated data collection and computation of allele sizes respectively.

Data Analysis

Exploratory data analysis was conducted with SAS for windows version 7. We used PROC FREQ, PROC GLM and PROC CORR as some of the procedures for finding out the significance of the data and also to unveil important associations among the various variables involved in the experiment. Heterozygosity (H) was estimated for each SSR marker using the formula:

$$H = 1 - \sum_{i=1}^{k} P_i^{\,2}$$

where P_i is the frequency of the ith allele and k is the number of alleles (Ott, 1991). The Power of Discrimination (PD) was calculated according to the formula mentioned above except where the allele frequency was replaced by the genotype frequency (Kloosterman *et al.* 1993). The different meaning of the two indexes is commented on in the discussion of results. The accessions showing only one fragment at a locus we considered homozygous for the fragment. As a

consequence, the Heterozygosity reported herein could be underestimated if null alleles occurred. Standard deviation of molecular weight (SDmw) expressed in base pairs was calculated for each SSR marker. SDmw is a derivative of the number and frequency of alleles on the one hand and the size range of the PCR fragments at each locus on the other, it provides a comprehensive measure of microsatellite genetic diversity (Cho *et al.* 2000). The Maximum Repeat Count (MRC) for each microsatellite locus was obtained using the formula:

MRC = [(Maximum allele Molecular Weight – reference allele Molecular Weight) / x] + reference repeat count (x = 2 or 3 for dinucleotide and trinucleotide repeats, respectively). The Mean Repeat Count (MERC) and the Most Common Repeat Count (MCRC) were also calculated according to the formula mentioned above by substituting Maximum allele molecular weight by Mean allele molecular weight for MERC and Most common allele molecular weight for MCRC (Goldstein and Clark 1995; Cho *et al.* 2000).

The genetic distance (GD) among genotypes was computed based on the microsatellite data with the program MICROSAT (Minch *et al.* 1997) that is available on the World Wide Web at http://hpgl.stanford.edu/projects/microsat/. This distance estimate is based on the proportion of shared alleles (ps) and is derived as: GD = 1-ps.

The proportion of shared alleles (ps) is defined as the mean of the minima of the relative frequencies of all alleles in the genotypes being compared i.e. ps = (sum over all alleles of MIN {P [A (i)], P [B (i)]})/n, where n is the total number of alleles for all loci (Bowcock *et al.* 1994). Bootstrap analysis was performed on the microsatellite data 10,000 times to get the standard errors of the GD estimates and

the mean GD estimates were plotted in a matrix. Cluster analysis of the GD estimates was performed using the Unweighted Paired Group Method using Averages (UPGMA). Cophenetic values based on the results of the UPGMA cluster analysis were calculated according to Sneath and Sokal (1973). Goodness of fit of the cluster analysis was tested using the cophenetic correlation values and the correspondence between the matrices was tested with the Mantel Z statistic (Mantel 1967). Significance of Z was determined by comparing the observed Z value with a critical Z value obtained from its permutational distribution. This distribution was derived by calculating Z values for one matrix with 1000 permuted variants of a second matrix. Associations among the 14-meadowfoam accessions were revealed by principal coordinate analyses (PCoA) based on the GD estimates. All these computations were performed with the appropriate procedures of the computer package NTSYS-pc version 2.02i (Rohlf, 1993).

Results

Preliminary screening of enriched libraries

Thirty-six clones were sequenced from across four libraries that were enriched for AC, AG, AAT and ATC repeat motifs, in order to identify those libraries most suitable for large scale SSR discovery. The level of SSR enrichment ranged from 0% (zero out of nine sequenced clones) in the library enriched for AAT to 55% (five out of nine sequenced clones) in the libraries enriched for AC or AG repeat motifs. The two libraries showing the highest enrichment for SSRs (AC or AG enriched) were selected for large scale SSR discovery in meadowfoam.

Large-scale SSR discovery and characterization of SSR loci

A total of 3264 recombinant clones were picked for colony PCR. After checking on agarose gel, 1668 clones containing more than one insert or inserts smaller than 350bp or larger than 800bp were discarded. One thousand five hundred and ninety-six colony PCR products (779 from AG and 817 from AC enriched library) were purified and sequenced. One thousand two hundred and thirty-seven sequences (617 from the AG enriched library and 620 from AC enriched library) harbored SSRs (repeat size $\geq$ 6bp). All of the 1237 sequences were aligned with the PILEUP function (GCG package version 10.1) to detect redundant sequences. A total of 541 sequences (202 from the AG enriched library and 339 from the AC enriched library) were found to be redundant and hence were not used further. The 696 unique sequences were used to design primers, however the sequences flanking 72 unique repeats (50 from the AG enriched library and 22

from the AC enriched library) were not long enough to design primers for amplifying repeat and flanking interstitial DNA sequences. Thus primers were designed and tested for 624 SSRs.

These 624 SSRs had repeat numbers ranging from 6 bp to 150 bp with a mean repeat number of 32 bp. The most abundant repeat numbers were 10 bp and 12 bp and each occurred 70 and 64 times respectively. Two hundred and forty-six of the 624 SSRs had repeat numbers ranging from 6 bp to 18 bp, whereas 220 SSRs had repeat numbers ranging from 20 bp to 40 bp and only 158 SSRs had repeat numbers higher than 40 bp. There were four different repeat motifs: 543 dinucleotides, 75 trinucleotides,
5 tetranuclotides and one heptanucleotide. We found 59 different motifs (Table 2.2) but the major motif types were AG (52.6%) and AC (21.8%). Categorization of the SSRs according to motif types, allowed a comparison of the numbers of perfect, imperfect and compound repeats. We found 150 perfect, 390 imperfect and 84 compound repeats.

Primer evaluation and SSR marker development

All the 624 primer pairs were first tested on the genomic DNA of OMF40-11 and OMF64, whose pooled genomic DNA was used for constructing the genomic DNA libraries. The amplicons were first checked on 1.5% agarose gels and then the accurate sizes of alleles were decided by the data recorded and analyzed by GeneScan Analysis 2.1 and Genotyper 2.0 (perkin Elmer Applied Biosystems). Three hundred and eighty-nine primers (279 from the AG enriched library and 110 from the AC enriched library) produced the expected amplicon on

at least one of the inbred lines. Two hundred and eleven primer pairs failed to produce an amplicon, whereas 24 primer pairs produced many bands along with the expected bands. The 389 primers (Table 2.3) were tested on the genomic DNA of an array of 12 meadowfoam accessions (excluding OMF40-11 and OMF64). There were 106 primer pairs which produced amplicons in all the meadowfoam accessions. Thus, there were 283 primer pairs which had null allels in at least one of the fourteen meadowfoam accessions. When only *Limnanthes alba* ssp. *alba* and *Limananthes alba* ssp. *versicolor* were taken into account there were only 93 null alleles. The number of alleles detected ranged from 1 (monomorphism) to 28 (Fig. 2.1) with a mean allele number of 6.25. The most common occuring number of alleles was 2 (56 markers out of 389 markers) followed by 3 alleles (52 out of 389 markers). There was a positive correlation between the number of the alleles detected by a primer pair and the MRC, MERC and MCRC (Fig. 2.2). However the number of alleles was highly significantly correlated ($r=0.47$, $p<0.0001$) with only MRC. The correlation between the number of alleles and MERC was 0.23 ($p < 0.0001$), whereas the correlation between MCRC and the alleles was 0.18 ($p = 0.002$). There was not a significant difference in the correlation of the perfect and imperfect repeats with the number of alleles. The MRC for the perfect repeats and imperfect repeats was positively correlated ($r=0.46$ & $r= 0.49$ respectively, $p<0.0001$) with the number of alleles, however the MRC for compund repeats showed no correlation ($r=0.25$, $p=0.17$) with the number of alleles.

Table 2.2. Frequency of the simple sequence repeats for which the primers were designed from 1596 genomic DNA clones of meadowfoam enriched for AC - and AG- repeats.

Motif	Frequency	Motif	Frequency	Motif	Frequency
aca-acc	1	cg-ca	1	ta-ca	4
at	10	ga	328	ta-ca-cg-ca	2
at-ca	2	gaa	2	ta-ga	1
at-ga	1	gaaa-agt	1		
ca	136	ga-ac	1		
caa	25	gaag	1		
caaa	1	gaa-ga-ga	1		
caa-aca	1	gaa-gta	1		
caa-caa	1	gaa-gtt	1		
caa-caa-cac	1	ga-at	1		
caa-cac	2	ga-ca	3		
caa-gaa-caa-caa	1	ga-caa	1		
ca-at	7	ga-cac	1		
cac	3	ga-cca	1		
ca-cg	5	ga-ct	1		
ca-ct	11	ga-gaa	1		
ca-ga	2	ga-gaa-gt	1		
ca-gc	2	ga-gagagac	1		
ca-gt-gc	1	ga-gga	1		
cat	2	ga-ggt	1		
ca-ta	17	ga-gt	4		
ca-ta-ca	1	ga-gt-ga	1		
ca-ta-cg	1	gat	4		
cca	14	gat-ggt	2		
ccaa	1	gat-gt	1		
cca-caa	1	gga	1		
cca-cct	1	ggga	1		
cca-gt	1	taa-ca	1		

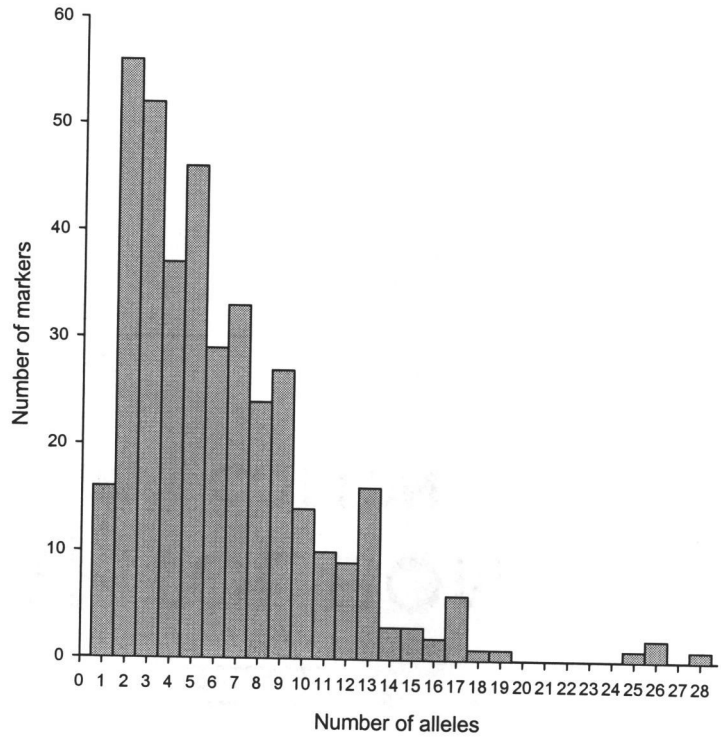
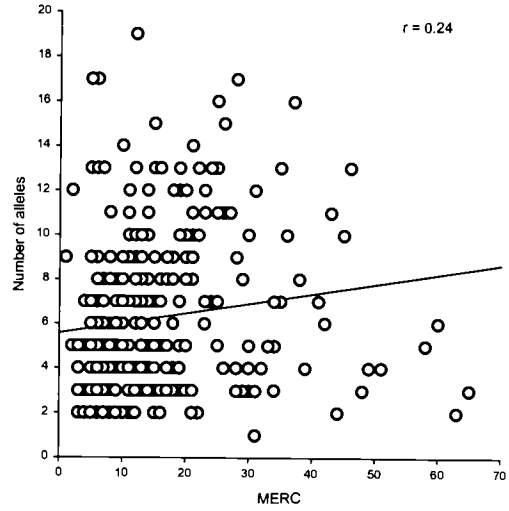
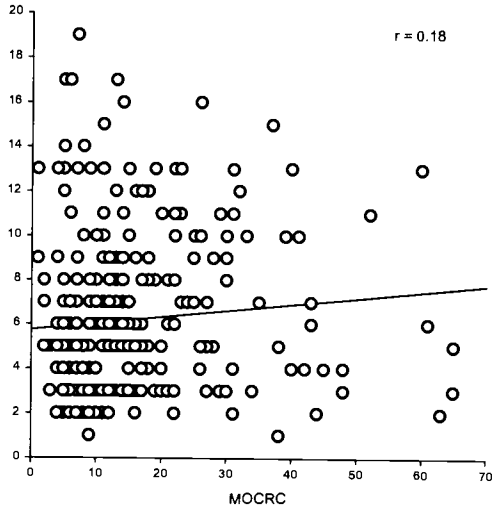


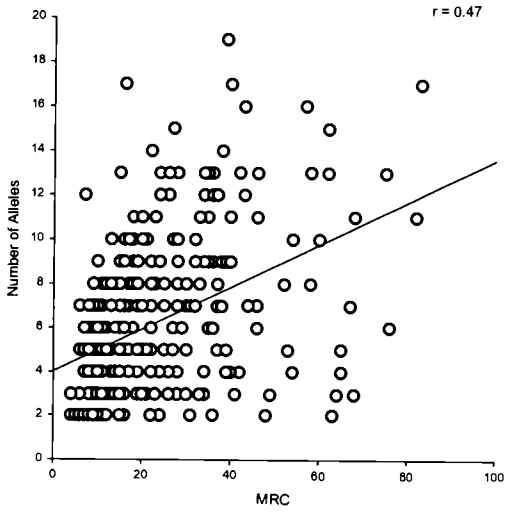
Figure 2.1. Frequency of the number of alleles occurring per SSR marker developed, when screened on fourteen meadowfoam accessions.



(A)



(B)



(C)

Figure 2.2. Correlation between the number of alleles per SSR marker and (A)- mean repeat count (MERC), (B)- most common repeat count (MOCRC) and (C) maximum repeat count (MRC).

Cross species amplification

From the total number of 624 primer pairs tested 389 amplified in the genomic DNA of OMF64 (*Limnanthes alba* ssp. *versicolor* - LAV). The number of amplified products ranged from 145 (*Limnanthes douglasii* ssp. *rosea* - LDR) to 389 (*Limnanthes alba* ssp. *versicolor* - LAV). The number of primers amplified in the *Limnanthes floccosa* ranged from 265 (*Limnanthes floccosa* ssp. *grandiflora* - LFG) to 282 (*Limnanthes floccosa* ssp. *floccosa* - LFF). In case of the *Limnanthes gracilis*, it ranged from 288 (*Limnanthes gracilis* ssp. *parishii* - LGP) to 303 (*Limnanthes gracilis* ssp. *gracilis* - LGG). The mean number of markers amplified in the section inflexae was 319.14, whereas in case of the reflxae it was 239. The degree of tranportability of the SSR markers developed across the genus *Limnanthes* ranged from 85.6% (333/389) (Fig. 2.3)for *Limnanthes montana* (LM) to 37.3% (145/389) for *Limnanthes douglasii* ssp.*rosea* (LDR).

Microsatellite polymorphism and discrimination power

From 389 total functional primer pairs, 373 primer pairs were polymorphic in the fourteen meadowfoam acessions. Heterozygosity for individual loci ranged from 0 to 0.93 (Fig. 2.4) with a mean heterozygosity of 0.63. Heterozygosity was higher on average for microsatellites produced from the AC enriched library (H= 0.66) in comparison with the microsatellites of the AG library (H=0.62). However this difference was not significant. The mean heterozygosities for di-, tri-, tetra- and heptanucleotides were 0.63, 0.63, 0.70 and 0.80 respectively, however these differences were not significant (p=0.87) as there were not sufficient observations in case of the tetra- and heptanucleotides.

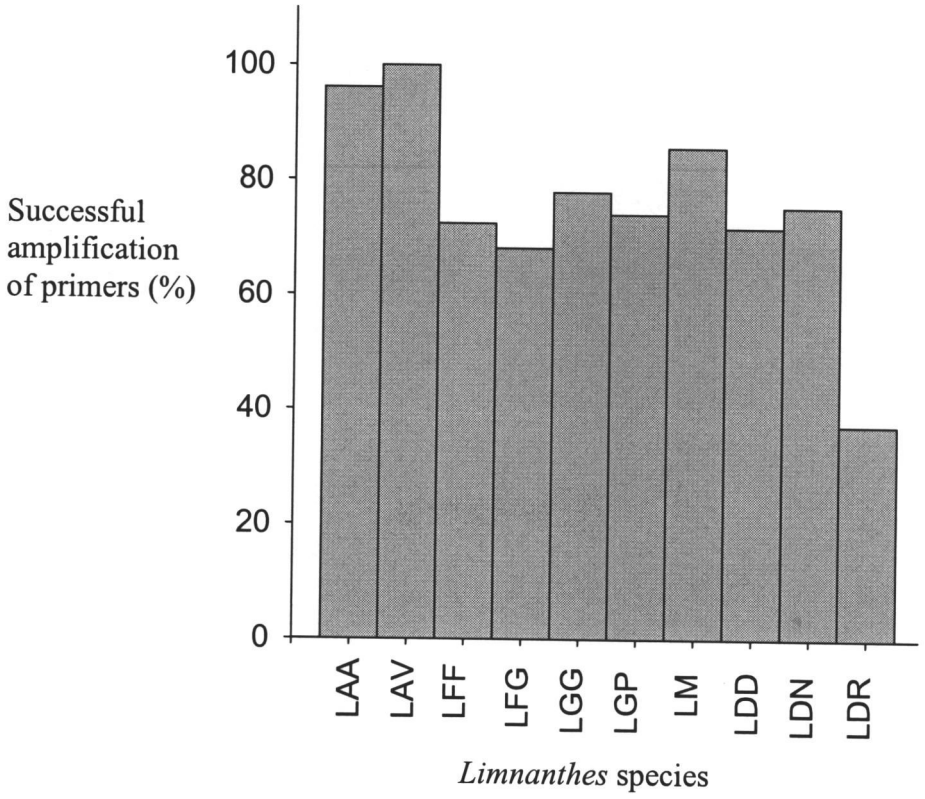


Figure 2.3. Successful amplification of primers in each species and sub-species (LAA- *Limnanthes alba* spp. *alba*, LAV- *Limnanthes alba* spp. *versicolor*, LFF- *Limnanthes floccosa* spp. *floccosa*, LFG- *Limnanthes floccosa* spp. *grandiflora*, LGG- *Limnanthes gracilis* spp. *gracilis*, LGP- *Limnanthes gracilis* spp. *parishii*, LM- *Limnanthes montana*, LDD- *Limnanthes douglasii* spp. *douglasii*, LDN- *Limnanthes douglasii* spp. *nivea*, LDR- *Limnanthes douglasii* spp. *rosea*) on our screening panel

Table 2.3 SSR markers developed along with their power of discrimination (PD) values.

Marker	Motif	PCR product (bp)	Ann.Tem. (°C)	PD	Dye	Forward primer	Reverse primer
MF002	ca-14	191	56	0.89	fam	ttcgaaggacttccgacaat	tcgtcagtaaaactcggtaaca
MF004	ca-11	210	56	0.85	fam	actggatttggcaattggtc	cccgtaatatttcgcttggt
MF006	ca-6	311	56	0.7	tet	ccacccatcaatctccctaa	tgccccttgtctctgaaaac
MF017	caa-5	199	56	0.5	fam	aagaggaagcccaaaccaac	ttcttcttgggtaggttgttca
MF019	caa-5,6	281	56	0.93	hex	tgcatttcagagggaagtga	ctttttgccaaggctaggtg
MF020	gt-23	199	56	0.91	fam	aatcaaccaagcctttcgtg	tcgacctccataaaagtttacc
MF022A	ga-5	201	56	0.58	fam	cacaaactctgtggctgctg	ctttgcctttgttgcctttc
MF022B	ga-7	197	56	0.13	fam	gaaaggcaacaaaggcaaag	tctttcacgggcttcttctg
MF024	ga-6,gagagac-14	318	56	0.47	tet	ccgtaagacgatggacgag	tcgtctctcgcctaacaactc
MF032	cttc-5	184	56	0.69	fam	ccctaaataccctacccatgt	tggccagtaaaaggaaggat
MF036	ga-34	302	56	0.46	hex	agccaaattggttttccatc	tgaggtattttgctcctctca
MF038	ct-10	363	56	0.85	tet	gcttcaacctcctccttcgt	ttacgccaagtttggtcctc
MF039	ga-13	170	56	0.92	fam	ggctaggttgcttgatggtc	gtttttccaccgagaagtgg
MF043	ca-12,15	288	56	0.88	hex	ggtggagctcaagtgaccat	tgtatcgggaaactggaagc
MF044	gt-12	245	56	0.77	hex	tcgtcagtaaaactcggtaaca	tgttgaagtcgaaggtgcag
MF045	ca-12,6,6	388	56	0.84	tet	agcgcgtgagattgctagtt	tcttcttatcatacccctctcg
MF046	gt-16	436	58	0.9	fam	ggcatgaacttgaccaaaca	ccgtactatggtgagccaca
MF047	ct-10	230	58	0.9	tet	aaatcctcctctccctgctc	agggtagtgtcgatcggtga

Table 2.3 (continued)

MF048	ga-22	299	58	0.92	hex	gagaggagccagaggaggat	ctggactgcgagtcttctcc
MF049	ga-7,6	322	58	0.41	hex	gaactggcggtacaagaggt	ggggctagattgcaaataagg
MF050	ct-9	192	58	0.87	hex	cgctcggtttttgatgagtc	taagagttgagcgagcgtga
MF052	ct-6	267	58	0.84	fam	gcagatccaaaacttgatgg	tgaagcagagcttgaactgg
MF053	cat-7	335	58	0.76	fam	tgagagaaaaatcgacacaaca	gtcccgctctaggaaaatcc
MF054	cat-7	389	58	0.62	tet	actttccccttcagtgacga	tctgaagatgaggctgaggaa
MF057	ca-25,at-5	325	58	0.81	fam	tgaacttcaaacccctctagttgt	taaatccggacccactttcc
MF063	ca-15	321	58	0.81	hex	gtcattgacccgtcctcgat	caaattgggatgccccatac
MF066	ca-16	377	58	0.58	hex	tgttcgtgtgcgagtgttttt	cttgggtagtagtttgagggatca
MF067	ca-27,at-6	217	58	0.9	tet	aacaacgagaatctgaggtgctt	gagtttgagaaagatgttcgagactg
MF070	ga-10	305	58	0.45	hex	gaggacaaagaggcgtcgag	tcgatccacaagcaacaatga
MF072	ga-11	335	58	0.57	tet	cttttgaacctcaccggttgt	aacgattcccttccaagacttt
MF073	ct-12	327	58	0.88	hex	ggaaatgtttgttggatgcag	gcggcaatatgcttgtttcat
MF074	ct-15	251	58	0.8	fam	tatctgcttggggagctcgta	cctgcagcaaattattatcacg
MF075	ct-5	266	58	0	tet	tacccattcgaccaccgtatc	tgttggatttacgaccactttcc
MF077	ct-9,cca-6	182	58	0	tet	cccaaatactcaaacgtcatcttc	gggttgctcattctccttgtt
MF078	ga-6	265	58	0.7	fam	aatgtgttggttcctcttgtgg	ggatgtctgtatccttctcctcca
MF079	ga-9,11	413	58	0.84	fam	tcaggttctctactgacattgaaca	cctttcgattccttggatcatt
MF081	ga-12,5	323	58	0.81	tet	accgccgcttatcttacacg	gtaatcacttcagggcgttcg
MF082	ga-16	238	58	0.73	tet	cttgatgatccaccgacgagt	acgtgatcagtggtcctaccc

Table 2.3 (continued)

MF083	ctt-19	391	58	0.13	fam	gaacacaaactcagacgtgcatt	aggaagaggagcaggaccaac
MF084	gaa-11	211	58	0.77	hex	tgccaaaggtcattctccagt	ttcagcctcgactatttcttcg
MF086	ga-11	294	58	0.89	fam	cccacgaccaagaggtcaat	tttcgttctccctagccttctt
MF088	ga-10	308	58	0.89	tet	tgcttcgtcatttgcgttaca	ttctgcttggctgttctattgtc
MF089	ct-12	300	58	0.89	hex	aagaaagattgtctggatcacacc	gttgttcatccctcctcatcg
MF090	ga-12	305	58	0.76	fam	ggaggacgtattggcaggtg	tgagacgtgagctgtgtttcc
MF091	ga-11	230	58	0.84	hex	aggcgtttctcttgacggtaa	tccgatgtctatttccactcaca
MF092	ga-6	190	58	0.13	tet	gaggcgagagaagaagggaga	gatccaccaccctccactaca
MF093	ga-10	247	58	0.81	tet	ggttcgctgagttcttgatcg	aacacccaaaaccggaagttt
MF095	ct-7	315	58	0.54	hex	tcctctgcctaaccctgtaacc	aacattgaggtaagtggttctgga
MF096	ga-20	317	58	0.91	hex	aaagaaaggaagcgtgcgagt	tgaaggaccactgatgatgatg
MF097	ct-9,caa-3	189	58	0.86	fam	ctacccgccaccatacgatt	ttcgagacgatgagaacgatg
MF098	ga-14	318	58	0.91	fam	gccatgactggagaacctgaa	tcttccgactctgtgtgtcca
MF099	ga-9	182	58	0.55	hex	tgcctgagttcgtgcatagtg	tcaaacaaacagcaccacttga
MF100	ga-16	398	58	0.89	tet	aaatcaagaccaatcgctacgag	atcatgccctgtctcacattg
MF101	ga-9	254	58	0.72	fam	cgagttttgtgatatgcaagga	accaccaccattgtccaaagt
MF102	ga-13	201	58	0.86	tet	ttgtttgagtccgggtaggtg	ctggagtagatccaactgtgactg
MF103	ga-16	245	58	0.78	fam	ttatatcggagttcgccgttg	tttcacccaatcacctatttcg
MF104	ct-23	340	58	0.71	hex	ccttgttctctttagcttctttgagag	gcagcagactcgacataggg
MF105	ga-14	221	58	0.8	hex	caggccctttaaaacgatgag	aacttgacccgagtccatga
MF106	ga-15	200	58	0.91	fam	cgcagtcgatgaggtagttgtc	gtgaatgaagacccagccaaa
MF107	ga-9	206	58	0.91	fam	cagtggggtatcaggggatag	gttacagcccccgttcaattt

Table 2.3 (continued)

MF108	ga-20	289	58	0.84	hex	ctagggttcatacggcatcca	tccagcccaacttctaatcca
MF109	ga-8	216	58	0.46	tet	tcaggggttggagaggatttt	ctcatctccctcacgttagcc
MF110	ct-13	230	58	0.73	fam	tcctgcttagggttgcgtact	gggctccctgtgtgtaatcct
MF111	ga-11	275	58	0.62	tet	tgtcgccggagttaaagtgat	cccatctcatcaaacatatccac
MF114	ga-21	196	58	0.86	hex	cactggtcgacggtgtttgta	ccttttgttacagccccattc
MF116	ga-13	264	58	0.8	hex	gagggatacgagacgcgaga	cttgcatatctccattgcatca
MF117	ga-5,7,6	301	58	0	tet	tgtgcagctttcgttaggtca	aactccgcgatgaggttctct
MF118	ct-15	377	58	0.83	fam	gcagatttccaactgtcattcc	atcattgtcggctgcacttc
MF121	ct-11	404	58	0.6	fam	ccttttgtgaaaggacacttga	gccttgctatgtcatccagaa
MF122	ct-6,6	399	58	0.86	hex	aagtgtgttgcaaatacggtca	cggagccatagctttgagaaa
MF124	at-5	201	58	0.24	hex	atggcaagagatgcaagatcg	ggcggttgtgtttgttctgac
MF125	gt-16,at-5,gt-3	386	58	0.68	fam	aaacccgtgaaagccagaatc	tccaagaaactcagcccaaga
MF126	ca-5	277	58	0.52	tet	cgctagatgtttccacccaac	atctcacgccttcgtcattgt
MF127	ggt-6	240	58	0.77	hex	ttgagtgtagagacagattgagtgttg	gcctccttaccacccttatttg
MF129	ca-11	290	58	0.83	fam	tggtcgaataaatcataccctcct	cgaccgatcataattttcaacc
MF131	cttt-5,tca-5	257	58	0.67	hex	cactcttgcccgtccactaga	tctcagaccctagcgatgttga
MF134	gt-5	381	58	0.83	hex	tcgtagtgtcccgttattatttcg	ttctactgctgcccaaactcc
MF136	ga-6	211	58	0.77	fam	ttgtggagaaagaatgatgttgtg	tcagaaaactcgcgggaagta
MF137	ca-12	256	58	0.88	fam	tcatcttcctcaagctgcaaa	catgtgagctactggttttacctg
MF139	ca-8	340	58	0.86	fam	tgaaagaggagaatagcaaaagagc	cattggtggctcaaggactgt
MF141	caa-28	195	58	0.64	tet	ccataatacatcagcccacaattag	ggtctggttggagttaggttatactg
MF142	ct-6	395	58	0.36	hex	tcgtttacggcatggactacc	gcgtgagtaagaagggcaatg

Table 2.3 (continued)

MF144	ga-5	454	58	0.13	tet	cgcaaagaattgtgatgagca	atacgctggagacaggaggtg
MF145	ca-14	293	58	0.65	tet	tcttcgggctgtctctttatca	ctagtggagcgagcggagtta
MF147	ga-8,gt-7,ga-8	272	58	0.9	hex	aggttggtcgaagcaccact	ggctctaaaaccctacttttcgtc
MF152	ga-6	246	58	0.53	hex	tacgcgatgtggtaggtgatg	cattgttcctaagcggcactc
MF154	gt-22	252	58	0.37	hex	cgcagaccatgcaatacaatg	aagtataacgacacttctcaca
MF155	ca-6	250	58	0.81	tet	aaatcaaactcgggtagcactattac	tgagaaacgaaggagaaagtgaga
MF156	ct-17	231	58	0.7	tet	gcaagccgttcagatatttcagt	caaacttcacttcattcagatgtgc
MF157	ga-6	267	58	0.56	tet	ggaggtgttggggaaggtatg	gcatgtttatcgagcgtcaga
MF158	ga-8	342	58	0.7	tet	agggtttgagtgagcagggta	attttgggaagggtcctctct
MF160	at-5	315	58	0.36	fam	gcgcagtaaattgttggacat	cacacacacacacacacacaca
MF161	ca-7,gc-5	300	58	0.46	fam	acctcctttctctcacaacacg	tcagcttgcttaatagccctca
MF162	gtt-5	255	58	0.47	hex	ccgttactggaaagacaagcaa	catagaaccacataccaaatccaca
MF163	ca-23,at-7	291	58	0.47	tet	cctagctggaggccgataaac	gacctttgatctttcagatgatgc
MF164	ga-14	256	58	0.92	tet	ctcgaccctggcacaatctac	tgaaggagatagcaaaggaacatc
MF165	ct-6, cac-11	345	58	0.79	fam	actcacgggtcctgcatgtt	tgaatgagcttggtgccttg
MF166	ct-12	246	58	0.83	fam	gcccactcactctcactttcc	tcgtacaacaacgccgattc
MF167	ct-10	233	58	0.84	hex	gcgcaaggatatgaaagtcc	cttgatttttggacagcatcg
MF168	ct-12	159	58	0.85	hex	accagtagtgacctgggcttg	gtttgaacgcagagatggtga
MF169	ct-8	294	58	0.88	hex	tgtctgatctttaacggcgagt	acagtcggtggtgtttgcac
MF170	ct-5	213	58	0.7	hex	cccgatacctctcgtcctcta	ccagggaaagattcagggaag
MF171	ga-17	196	58	0.87	fam	cacccaacacgggagatagac	tgcaagtgctagacatcgtca
MF172	ga-12	274	58	0.24	fam	tgctgacatagctccattgttg	tcagtttcttcaactcctctgacg

Table 2.3 (continued)

MF173	ga-18	310	58	0.88	tet	tgcacgtacctcagtcgattc	cccaaaaatgggtttatgctt
MF174	ga-8	331	58	0.67	tet	atggttcacccgagatcaatg	gtatgcgccggtaatggtaaa
MF178	ga-9	198	58	0.66	tet	gtagcccgattgcgatagctt	ctccaacaaaaaccagatcgt
MF179	ga-36	293	58	0.86	hex	cgccgagagaggaaattaaaaa	ggggaagaagaacccaacaac
MF184	cg-4,ca-27	343	58	0.82	hex	atcacgttattgacgcacacg	catcgatacgaattccatagtggt
MF186	ga-34	301	58	0.46	hex	agccggaactgaaggaagaag	gggggcacttcactaaacaaa
MF187	ctt-5, caa-19	271	58	0.85	tet	tcccattcccattcccataa	gagcgacgaggaagttagcag
MF189	ga-30	275	58	0.9	hex	gctggtgattatgaggcgttt	tgccggaactctttcagacat
MF190	ga-36	325	58	0.83	tet	atgcttcgttcgatggtgtct	cccctaaataaaacctcccaca
MF195	cca-14	188	58	0.5	hex	actaccggacgacgatcaca	ggtggtggtacttgtggaggt
MF196	ga-28	240	58	0.93	fam	aatggtcgcgcaattaatca	ccactgttctcaagatttttcaagg
MF212	ca-27,ta-8	466	58	0.55	fam	gggatgaagcggtcacataaa	tgttgagtggaattgagcccta
MF218	ct-22	259	58	0.86	fam	cgatgacccattaggcttgac	agtgatggtagccagaaagttga
MF226	ct-21	266	58	0.67	hex	ctatcgtgatgcaaccatcca	caagaaagccgatgagacgtt
MF231	ga-17,gt-5	254	58	0.37	tet	ttcgcagagaaaaaggaggag	ttctctctctaactaacctccatcacc
MF233	ga-33	279	58	0.88	tet	tcgccgagagaggaaattaaa	caccaacaactacctcccccta
MF234	ga-17	303	58	0.87	tet	accgttttttcaaccatggag	gtcccccaatttttcttccaa
MF237	ga-7	215	58	0.41	fam	cggatgaagaaagtggaaatagaga	tgatgtggtagttgatggattgg
MF238	ga-5	309	58	0.41	fam	cggactcatcacgctagaagtaa	tctccaccaacaatagcctctaca
MF239	ga-10	242	58	0.34	tet	tactcctcaatggtgctgtcct	ctgaaactcttctcaaccgcatc
MF240	ga-6	160	58	0.34	fam	tgaggttcaaaggattggctaaa	ctatccatgtcagtcgcatcatc

Table 2.3 (continued)

MF241	gat-26	262	58	0.55	fam	agttgggtgagaagaaactgagc	ttaattttgtacctacgcgagctt
MF242	gat-11, ggt-4	298	58	0.67	hex	tcctcagcccctatccatatcta	atgggcagcaaagatgaagtatg
MF243	ga-5	235	58	0.34	tet	acccatgaattttcgaggtcgt	ccatccctgaacaccttatcaac
MF244	ga-6	295	58	0.13	fam	agcgaatggagttttgtggagta	catcccatgataatagcctctcc
MF245	ga-5	390	58	0.26	tet	agtgaggaagaccttgtggatagg	catcaccagcattaccaactgag
MF246	ga-5	358	53	0.36	hex	tttgctaaaccgattccgagag	caccatcttctgtccgtacttga
MF247	ga-19	224	58	0.56	hex	tcgatggatgaggacgatgagt	tcataaaatggcgaagaaacagc
MF257	ga-4	312	58	0	fam	ggtccgagtacgctttcatgtta	agccattcgcagtttcacagt
MF258	ga-5	382	58	0.87	fam	tgcagagctttcagatttctcg	cacatctcatccatacctctcttctc
MF259	ga-5, 8	304	53	0.56	fam	agagagatacgagagcaaacatga	tctatttcaacataccttgtctcaca
MF260	ga-6	328	58	0.6	tet	aattctgaggagcattcaggact	ccttctcttgtttcccttgttca
MF261	ga-14	309	58	0.88	hex	tgatgtcctagagtgctgggtta	aatttccagggttcaagaccaac
MF262	ga-4	180	58	0.36	hex	aaatgagaatcgggtagggaatg	cacattcaatcaaccaccaattta
MF263	ga-6	221	58	0.36	hex	tgagtgggttgtatcgtggtt	accctaactaacctcagccaagt
MF264	ga-6, ac-4	483	58	0.5	tet	atcaaagaacggagccagatact	atacgcaggcaaccaaacagt
MF265	ga-5	348	58	0.24	fam	gctagagctggcaggtgaaataa	ctctcttccattggttcaacacg
MF266	ga-9	224	53	0.76	tet	gcaacgctaaaatcagaggtaaa	ctcatcactcaaagcaccaagat
MF267	ga-7	357	58	0.65	hex	ggctagaagtaaccagatgggttg	ccagaacctcctaaacgcaataa
MF268	ga-10	259	58	0.46	fam	ccacaatctcaatagtgcagtgg	catcaccctcttcgtcatcatct
MF269	ga-7	373	58	0.72	fam	cgatgcgtttccagtagtaaatg	gtgggactttcgttgtgggtat

Table 2.3 (continued)

MF270	ga-19	315	58	0.37	hex	agaaggtcgttgtcttgccatt	ttgaaacccgtaataaattggtca
MF271	ga-17	261	58	0.8	tet	aatgctactgactgttgcttcacc	ccctcattcccattaccctaaac
MF272	ga-9	276	58	0.54	fam	gcttcttcggcaatacgagctt	aagcaaagccacagaacctctct
MF273	ga-16	427	58	0.91	hex	aagctctgatccgcatagcatt	ttccaaaattgccccctttaagtt
MF274	ga-5, 5	406	58	0.5	tet	ttggaaatggagaaggttaatcg	caaggttaccataattccctcca
MF275	ga-11, 6	380	58	0.77	tet	tcgttgttctctcgatcctcttc	tgattctcttggctctcctcaac
MF276	ga-12	347	58	0.93	hex	ctcaaaaactcgaagaaccatcg	ttccaaagatacccccgatataaa
MF277	ga-7	318	58	0.71	fam	agctcatcgtccttgcagaatc	tgtcatctctaccacctgaggcta
MF278	ga-7	165	53	0.84	hex	attctgagtttgtctcttctgcttt	atcataccaccatagcctttcaa
MF279	ga-15	250	58	0.73	fam	acgtgtctttgatggagaaggaa	ctcattatggcgaaagtgtcatc
MF280	ga-11	195	58	0.62	tet	ggcgaatgatgaatgatgactg	agaatgtagcccatacacaccaa
MF281	ta-10, ga-11	249	58	0.85	hex	gccatgacagagggaaattagaa	acgtgacatcagacaatcagacg
MF282	ga-5	423	58	0.61	fam	gagaccaagaatgatgatgtgga	catgtttcgcatctctccattta
MF283	ga-24	188	53	0.86	fam	attaaatgaggagagagggttgc	agcggttctttaacaatcaatgc
MF284	ga-9	227	58	0.41	tet	acaaagaggaggaggacaggtt	ccggtgtccacattaaatacgtt
MF285	ga-7	278	58	0.81	hex	gaaataagagatagaaggagcgaacg	tttgaaccagcctccacactct
MF286	ga-27	321	58	0.36	tet	agtcgcgtcaagcatttgaga	tttctctccctaaacgcagcaac
MF287	ga-15	361	58	0	hex	ctttgccactttctgcatagcc	catggaagagtttgctgctgtg
MF288	ga-15	275	58	0.89	hex	aaacacgcaaaatctccagaaaa	tgctttgttgctactgatgatgaa
MF289	ga-20	272	58	0.37	fam	cacgcctgactctgtgattctt	tttatatgggctttagtcccctga

Table 2.3 (continued)

MF290	ga-5	333	53	0.87	fam	tttagctcttaggtccagaatcaga	atcgaaatctaaagccgtcgtc
MF291	ga-17	313	58	0.89	fam	acgagttctgtgtaatgcgagtc	tggtaagtaagtgaatgtgtgtgacc
MF292	ga-7, 4	272	58	0.64	hex	aaacccaccgttttacctttgac	gggaagattctctttacaggggta
MF293	ga-4	157	58	0.34	hex	atagatacgcgaaccaaggaagg	ttaagaagcggagtttcggttta
MF298	ga-14	397	58	0.8	tet	atctctcgatctttccctgtcgt	taattcactagctcctgggttcc
MF299	ga-10	346	58	0.85	tet	gggaaagagatgagagaggagaa	tatcatttgaccagccacatcag
MF301	ga-17	254	58	0.82	tet	attatgggtcgctgatctcatct	gtcggataaagttgggttgcata
MF302	ga-8	208	58	0.76	tet	atacggacaggagaccgagagc	aaacctattgtgaaatgaacccttc
MF304	ga-13	324	58	0.83	tet	gcatccaccagcaagaagattat	tggagattgtaaaggtcactctgc
MF305	ga-21	278	58	0.92	tet	aactggttcacaggtttgagtcg	tacagcccccattctatttttga
MF307	ga-11	253	58	0.62	fam	tagagttgaattggtgggtcgtc	catttaagacccgttacacaaacc
MF308	ga-6	222	58	0	hex	ctgaagagggagctggaaaaga	tttctgacttgccgatgctaaat
MF309	ga-6	192	58	0.26	tet	gtggtcgccaatttgataagagt	ctcccaaccaccacatttaagag
MF311	ga-5	425	58	0	fam	cggaaaagactcagggaagttg	cctctccctctcaaaacaatcct
MF312	ga-28	437	58	0.92	fam	gcacaagggctctcaggtaataa	tagcgcatatcgtcttcctctct
MF313	ga-26	370	58	0.61	tet	attttggtaaacggaaaaccaga	gaaaagctctatgacccgacaac
MF314	ga-9, gt-5	308	58	0.13	fam	agcccgtgatgattagggttagt	caccaaattattctctccatactcca
MF315	ga-10	280	58	0.83	tet	atatgatgtgtgcaacccaatcc	ttgaaacaggaccaggagtagttc
MF316	ga-11	500	58	0.34	fam	atcatctttggtggtgggaagag	aagcatttatggcaacctctcaa
MF317	ga-14	311	58	0.73	tet	gaatccaacgagaagagaacgaa	ccttaaacactgaaccaacgtga

Table 2.3 (continued)

MF318	ga-25	407	58	0.92	hex	aaagggtcaaacacaaactctctaca	gaatatggttcaaatgcagggta
MF319	ga-9	416	58	0.7	tet	tgatctctctggaccgtgatactc	ttgtgcagggctatgggatatta
MF320	ga-6, 6	431	58	0.47	tet	ttgtgtgattgaagatggctgat	aacccaaagagagctgtagcgta
MF321	ga-25	314	58	0.93	tet	ctacctctgtcgatcgatgtgaa	gcaaggagatctcgacgaactta
MF322	ga-22	442	58	0.9	hex	tgacaaatggatcatgcagtctt	gatgggctgtcgagaatctctt
MF323	ga-10	392	58	0.77	hex	gcgaatgattggagagaaggtt	ctttgggttgaaagctcaagaac
MF324	ga-12,gt-5	246	53	0.85	hex	ttttatgggtttatgtttagtttt	cactgtcattcatatttcgacta
MF325	ga-11	417	58	0.55	fam	ccttcgcattcaacgagttagag	ctaaaccgaaaccaccacaattt
MF326	ga-7	420	58	0.62	hex	tgctgaattaggaagcccattag	tggctctaaccctctaaactcttc
MF327	ga-10	459	58	0.87	hex	atgcagctaaacagatgaaatcg	cagtaaaggaagcaggaaaccaa
MF329	ga-3, 6	247	58	0.55	fam	gaaatagggcgctgagaaagac	tgtgttgtctagggaaaccttgag
MF330	ga-7	318	58	0.26	hex	catgcacactgagactagggatct	aaggtcctctcctaactctaacttca
MF331	ga-34	443	58	0.44	tet	gccaagaggagtgatttaacgag	accaccataaccaccgtcctaa
MF332	ga-8	427	58	0.55	tet	aggtttgatggatctgaacgaga	tgggacgtaggtaacaatggtg
MF333	ga-11	337	58	0.85	tet	ataaccatccaacacgaaccatc	ctttgccttggtcctctcttctt
MF334	gga-5	186	58	0	hex	atttagtgcgacggagctttg	ccaatacaaacaccaactcaagc
MF335	ga-37	251	58	0.13	fam	ccgttcaggcagaattaacaaac	ttaagtggatgtttacggcgtct
MF336	ga-10	370	53	0.77	fam	caacctgaaaccctagattatcg	ttttcacaaatctccttccaaaa
MF337	ga-16	185	58	0.56	fam	tatcctcggttaattgtgcatcc	tagatcgaagaagtccgccaaat
MF338	ga-11	199	58	0.26	tet	cagatattcccgtaaacgcatgt	gatcgtctcctcattcgctaaa

Table 2.3 (continued)

MF339	ga-17, ggt-5	323	58	0.69	tet	acgaatcagtggagaacgaagag	cggttcatatactgccttggttg
MF340	gat-6	431	58	0.83	fam	gcaaacaagatgacttcctccaa	cgagttctggttcatcatcactg
MF341	ga-23	424	58	0.6	tet	ggaggaccaatctagttcaagagc	aaggatgaatggatgagggactt
MF342	ga-26	321	58	0.9	fam	aagccgaacatcttatccttgct	ccctaaaatgggttaatgctcac
MF343	at-8, ag-19	358	58	0.77	fam	agtcaccaacatcgacatggata	ctgaagctacttggcattgacct
MF344	ga-5, 7	248	58	0.37	fam	cttcatcggctttacggaagttt	ccaccatcaccatctttctcttc
MF345	ga-20	312	58	0.13	hex	gagattcggacaagaagacgaaa	atgagattccctgtgatgagctt
MF346	ga-5	323	58	0.72	fam	tagtaagaccaaatcaccgttcg	tccgttccactatcaatcacctc
MF347	ga-5	277	58	0.26	tet	ttacgtttgaggagctggtctct	cccatctaggtcgttagttcgatt
MF349	ga-44	304	58	0.13	hex	acgtcgccagttgagtctaggta	cgatcagacaaacacaccagatt
MF350	ga-7	268	58	0.24	fam	tgcaagttcatagttgtgtggaga	ccaacatacccggaaatcataac
MF351	ga-28	253	58	0.13	hex	tgaacaacaatggtggtgttacc	caagaacatcaagaggacagagac c
MF352	ga-48	343	53	0.44	hex	tattatgtaatagccgagatgc	ctctcatatgtcttattttatctctc
MF353	ga-65	420	58	0.46	tet	tacccaacgcgagtaaggtcac	aaaacagattccgggagatgg
MF354	ga-36	304	58	0.93	tet	gaagcatctgaactgaatcgaaa	cccagcccgtaacaatataacc
MF355	ga-6	297	58	0.5	tet	gacaggaatagtagcagcattgg	tttaaggcaggctccacaactt
MF356	ga-16	364	53	0.93	hex	aaaagaagacctgtgatgatgct	ccgataacttcagagaaaacatgg
MF357	ga-8	377	58	0.82	hex	ggagtcgaaataatgaacgtgatg	catcgctctatccctattcttgg
MF359	ga-6, 9	388	58	0.46	hex	cggtgcctgtaactacgaatatacc	ggactcaagataattgattgcatttg

Table 2.3 (continued)

MF360	ga-7, ct-6	390	58	0.53	fam	tttcacaccattaggaagcctct	atcatcatcgtcgtcatcttcgt
MF361	ga-13	374	58	0.83	hex	gttggaggctgaaagaggctaat	gtgtattaccttgggtgcattcc
MF362	ct-5	260	58	0.89	fam	ttccaagataggtttatgcccttc	ccatgtctcctagttcccagaaa
MF363	ct-6	371	58	0.82	hex	cacctccacccaatcacataga	tcctggtcttgagttatcgaagg
MF366	ac-5	297	58	0.76	fam	cacacacaaacaacaaacgaacc	ccagatcgttccaccaagtaaaa
MF367	tc-4	297	58	0.56	hex	actgctgtgaaaggttcgtgttt	gcaactttacagcacacgaggta
MF368	ct-6	301	58	0.83	hex	tgtccttgttgttgggaagtttt	ccaaaagtgctgacccagtaaat
MF369	cac-9	311	58	0.37	fam	cgagatctacattgcaacagacg	caataaaccacgccttgtacagc
MF370	ct-5	431	58	0.8	hex	acagggacgacaaatggataaga	ggaaatatgtcagtgtgtgcagag
MF371	ct-12	394	58	0.88	fam	tcagtgactactggtgatctgacaa	gaaccaacatagggagtttgtgg
MF372	ct-7	435	58	0.88	hex	aacggcattcttgaaacaactaaa	aggagtctttgtgggctgtctcta
MF374	ct-7	290	58	0.84	tet	caaccacctcttcctcatcaact	ctctggactgagtttgtccgttt
MF375	ct-11	401	58	0.46	fam	ggttatcagaccaagcccaacta	ctttctctttgcttccgatttca
MF376	ct-18, ctt-8, ca-4	348	58	0.93	hex	ctgttgcttaacgccatctcaat	tttagtgaggcgtcggatttct
MF377	ct-6	381	58	0.49	hex	ttcagtggctatccctaatctgg	taggcagctttctggagagtttg
MF378	ga-14	325	58	0.93	hex	acgtagcagagagacttgtcgtgt	tgtataagagcatcaaactcccact
MF379	ct-8	257	58	0.55	tet	attcaacacagccgcctatttc	aaacctcaagcaataacacagtcg
MF381	ct-5	220	58	0.93	hex	ctgctcctcacctcacttcattt	tgttgtcgtaaagctctctcagg
MF382	ct-5	272	58	0	tet	tgaacttgaatcccttcgatgtc	accaaatcccgaatccgaaac
MF383	ct-12	156	58	0.44	tet	cgtctttcaacaacataattcattca	attcgaaatcggattgagattc

Table 2.3 (continued)

MF384	ct-14	281	58	0.91	fam	tttcctccctcttttctccctct	tcgcaatgaagaactaggttgaa
MF385	ct-9	131	58	0	fam	atcttcgcatctttatcgcaatc	aatccgttgatgcatctgagtgt
MF386	ct-18	285	58	0.55	hex	ttagctgaaccctcttttatctctttt	ggtgtttggtggcagtatttagg
MF387	ct-24	346	58	0.7	fam	gaatatgtcaccgacggggtat	tccctgtgaagagaatagctcca
MF388	ct-18	108	58	0.37	fam	gggttcagacaagaatccctaagc	tcgatgaagtctctggttgtgg
MF389	ct-10	130	58	0.36	hex	atgactctactgccctttcgttt	agagaaatgtttggaggagacca
MF390	ta-5	264	58	0.8	tet	tctttcgtatctcgtcttgtctcg	tgtatgatgaaatggggaaatga
MF391	ct-5	275	58	0.36	tet	tttggttctctgtggcaatca	tgggaacttactgaagcgaaact
MF392	ct-6	392	58	0.55	tet	cagtagcccaaacactaccctatc	tgggttcttggagagaggtcata
MF393	ct-28	319	58	0.78	tet	tctctggcttgggaagtagattg	ccttcgagagtgggatataacga
MF394	ct-4	155	58	0	hex	gccgatgctaaatccgtaatct	tttcgtgcatggagataagcac
MF395	ct-42	269	58	0.56	fam	taatcccactcgcactctttacc	gaggaagccaagaagaagagaca
MF396	ct-8	224	58	0.82	fam	tttcctgccctttcactctaaa	attggcgattgataaacatgaga
MF397	ct-6	255	58	0.41	fam	attaacgaccgatgttgaaggtt	caacgagcttctggattggatag
MF398	ct-4	311	58	0.7	hex	gggattatgatgttgactcgaatag	gaagtacgatctatcaccgggagt
MF401	cca-5, cct-3	310	58	0.85	fam	ccatccacaacatttttccctac	ggcaattcaagccctaaatctaa
MF402	ct-11	287	58	0.89	hex	cccctctcatcttcacattttca	cggcgattgtagtcatctttactt
MF403	ct-11	344	58	0.82	fam	cattgggtccaataaaaatcgaa	ctttcctaaaaatccgatggaga
MF404	ct-7	378	58	0.9	fam	agctccctcgaaagctcatgta	ctccgataacattccggtaagaa
MF405	ct-15	251	58	0.92	hex	cttggataccttctcgcaaacac	tttcctatactcaaccgctcgaa

Table 2.3 (continued)

MF406	ct-21	306	58	0.71	hex	cttccctccaaattccttcacat	cagatggtgattaatttgggatttt
MF407	ct-5, 6	288	53	0.7	tet	tcttctaaatccttcctccgttt	cgcactgaacaggaaagaaaatac
MF408	ct-26	256	58	0.7	hex	agactgttcctgctctacctaacctt	aataggcttggagttggagttgg
MF410	ct-21	187	53	0.63	tet	ccatccttatgtattccttaaat	gtttatgggtgattattctgttg
MF411	ct-15	229	58	0.89	hex	tgaaattcaaaacagcctcca	cagagaaaactaacaacgccaaa
MF414	ct-9	296	58	0.86	hex	cgacactctgctcaaattctcaa	tatgatctgtggttgctgttcgt
MF415	ct-7	449	58	0.83	hex	ggcaaaccaataaagtcctctca	caatccgcagacctctacttca
MF416	ct-19	381	58	0.82	tet	caaacccaccaatgattatacagaa	tgacggcaagtgataagagatga
MF418	ct-6, 5	338	58	0.26	tet	gtcatggcagttacctcaagcat	aagggaggaagataaagggttgg
MF419	ct-5, cct-4	433	58	0.8	fam	ccctccttcatttatccaaacc	taggcagctacagggatcttgtt
MF421	ct-7	358	58	0.65	hex	gacagagcatagcatcatagaggtt	agtcgtgtcaattaggaggatcg
MF422	ct-6	386	53	0.26	fam	gaatctcctctacaatcaaagtcaaa	aggttggtttccgtagcttct
MF426	ct-21	221	58	0.5	tet	tcataccttctttaatggcaca	atggctagggaggagagatttga
MF427	ct-8, 4, 7	282	58	0.69	tet	ttcccttaaccccaaccctaa	aacgatgtcgattgagaaacctc
MF428	ct-6	254	58	0.49	fam	tggcatccttctctaattccttc	agagctattgggtgagggttga
MF430	ct-7	290	58	0.36	fam	tgtatccaaaccaattcaaccaa	atgacgagcaggcagttatgagt
MF433	ct-21	300	58	0.8	fam	gtgggtttgtggatctttgtgag	ataccttcagatgtgcccgaga
MF434	ct-20	303	58	0	tet	gattctcgtagattggcagcttt	gcaccgagtcagatagagtttga
MF435	ct-23	209	58	0.93	hex	gatcatcaccacctcttccactt	agccaagaaattggatactcagg
MF436	ct-8, 8	302	58	0.85	tet	attcaacccccttcaatggtcagt	cccgaaagcaaacaggtaatcta

Table 2.3 (continued)

MF437	ct-5	330	58	0.77	hex	gttctgcaccttctgcgatattt	ttgtgggatctctgggttctagt
MF441	ct-23	208	58	0.61	fam	tcactgtaactggccttccaata	ggcgttgtgcgtatccttaatag
MF442	ct-22	396	58	0.71	hex	taaccatcacccctctgatcttg	tagtgaaggcagtagggaggctat
MF443	ct-48	475	58	0.57	fam	ccccaaatcttgtttacttagttgc	agggtccatcaggtaagtgtttg
MF444	ct-6	198	58	0.87	hex	aacttcttctctctccctgcaca	catggttgtggtcttgtgaaatc
MF446	ct-10	423	58	0.84	hex	attctgctggtgaagaggcatt	tctccaaagtgtaaatcgtcaaaga
MF451	ct-15	410	58	0.76	tet	aagagtctgtgttgcagtgttgg	ctgaatccattctccttcagctc
MF452	ga-62	284	58	0.24	fam	aagccaaatcctctcttaattcg	tctcgtacaacacttattgggtct
MF453	ct-15	347	58	0.71	fam	tggccaaacagctatctaagtatcc	cggtggtttcaaagatgaacatta
MF454	cca-5	262	58	0.66	hex	tcagacacctggctaactcttcct	gggttcaattcgatgaaagcact
MF456	ca-5	188	58	0.6	hex	gtccacacactttcccatctaca	atataagcataggtgtgcccgta
MF457	ca-9	277	58	0.93	fam	caacttatctgaccaggagcacac	aagcatttaacgggattgagctt
MF459	ca-27	479	53	0.13	hex	atattgaaaataacacacactaaaat	aaaagatgggtacaataatgaag
MF461	ca-10, ta-4	294	58	0.88	fam	gagaggtgtgtgaagagcatacga	catggtgtcaacaaagtgctcat
MF463	ct-4	300	58	0.37	hex	ttcaccaccatagcgttctgtct	ttcgcaacaaatggtatcaaagc
MF466	ca-13	297	58	0.91	tet	gtagccctaaaatgcagcaatcc	tttgggtgagttgttctcccttt
MF473	ca-15, ta-5	308	58	0.64	hex	cactcaggcggcacaaacttc	tggaaaatgatatgcttcctcacc
MF474	ca-9	223	58	0.53	tet	cactttcatttaaggggttggac	tggaagcatgttatcggaaagaa
MF476	ca-6	285	58	0.65	fam	cccatacttatcaccatcggcta	aagaccatgtccgaagagaagga
MF477	ca-13	288	58	0.8	fam	ataaatgggaacacgccagaact	aagattgatttgcgtgacaaacg

Table 2.3 (continued)

MF481	gc-3, ac-6	284	58	0.55	tet	tctgagcaaaggagtaataaagagca	tacgaagaagtgggtttggaaag
MF482	ac-49	361	58	0.53	tet	actcaaagttggctttctccaa	agcagtggtcaagctagtataggg
MF483	ca-21	250	58	0.92	hex	gtcgatggagtgctgactgattt	cccaattcgtttcttaccttgttc
MF484	ca-6	297	53	0.91	fam	cagaggttgaaagagtagcatcc	atcatcttgtacgtcctcggtct
MF485	ca-34, ta-12	295	58	0	hex	tggagtgtggcctaagagtaatatg	ttgttagtttcggtgggagagtt
MF487	ca-3, 3	279	58	0.77	fam	cacacacggacaatcgaaaata	ttgaagtgcctgagtgttgtagc
MF489	ca-6	333	58	0.79	hex	gaataacaaactgaacacccactagaa	ggctgacctgttctccttttga
MF494	ca-11	353	58	0.61	tet	cttgatccgattagaacaccaaa	cgatgaggcttttaaggaagaaa
MF496	ca-8	319	58	0.64	fam	agccttgacgcacacatcaaat	tgaatgtttatcgttgtggtgatagaa
MF497	ca-12, ta-5	206	58	0.89	hex	ttagagcctgacaaatcatggaa	gtcgccttggaaactgaagaaa
MF498	ca-55	347	58	0.34	tet	caacgacatacacagccacactac	tttggtgttattgtttcacagtgtttt
MF501	ca-5, ga-4	411	58	0.46	fam	ataaatacccacctacccgcatt	tatggacgacagattgctagacc
MF502	ca-51	286	58	0.36	fam	taggcatcatggttttggatttt	ttaaatggaaagcatatttaggaacc
MF503	gt-5, ga-4	272	58	0.83	fam	ttctattagcaaagggtggttca	cacaaactcgtccttagatccat
MF505	ca-24	287	53	0.88	tet	acaaactcacactcaagcacaaa	gacattattgacatccggctaca
MF506	ta-4	287	58	0.44	fam	catcatatacacctcgatcaccttt	tttctctagtcttgcatgttgtagttg
MF507	ca-46	382	53	0.53	hex	aaaaattaacatcgtcaaagaaa	ccatgtatatttagaatacagtcca
MF508	ct-12, ca-19	298	53	0.46	tet	ccctaacacttcaaatcctcttt	ttggttctatgtgtgtatttgagc
MF511	ca-19	157	58	0.49	tet	actctaccaccacataccatacgc	aagcggtctatgtatccgatttg
MF512	ca-16	356	58	0.81	tet	tagggacctaagacgtttccagt	gctcaaatgctagaaatacaagtga

Table 2.3 (continued)

MF515	ca-10	458	58	0.85	fam	caactcaaagaaaacacaaactaaaca	tgacctaattaatgcaggctagga
MF516	ca-9	357	58	0.82	fam	gggaacgaaagttctcacacata	agctttatcctggttgacacaat
MF517	ggt-12, gtt-16	197	58	0.81	tet	tgcccttgttagtgttctactttctc	ccaaccataccatcgagtctttc
MF520	gt-19	311	58	0.78	tet	cacaattctcactccgaagagttt	cccacttatcaatgctctctcca
MF523	gt-28	309	53	0.76	tet	atcattgggcattagaggagaa	catcatcatctttcaatcaatcacc
MF525	gt-4	307	58	0	fam	tactgttgatttgcaggtgttgg	accacttgagctgccctatttct
MF526	gt-56	454	58	0.89	tet	cgattaggttgctgcttgcttt	tcacgattcacacttacaaatatcaca
MF527	gt-26	305	58	0.93	fam	agattcttacgtggggattacctt	atttttgtagcgcttgatctgga
MF528	gt-41	296	58	0.64	tet	tgcgaaatttagtttcaatgagg	tggtcatcgatcagtaacttcagag
MF529	at-4	172	58	0.54	fam	cttggttcgactaaaggcatgtt	ttcaagatacacccattcattcttt
MF531	ggtt-5	281	58	0.67	hex	aaggacttctccatgttgagtgg	catctcgatccaacccttagaaa
MF534	ca-36	288	58	0.5	fam	aaacacacgagaacatacaatatacca	ttgtgggtaggtttgaataggttg
MF537	gt-15	299	58	0.13	tet	atgtcccactttggagaacactt	gaactggcctgagaaacagagaa
MF540	gt-14, ta-4	294	58	0.63	tet	ctcaaaaggtaactccccaaacc	aagataggtcgcgtttgaaattg
MF541	gt-6	312	53	0.71	tet	catgccataaccattattttcgt	atcaacatcaatcccgtcacc
MF544	gt-9, at-4	319	58	0.62	hex	aagtgccctggagaatttacttg	attgcccatttatgtcaagacca
MF545	gt-26	296	53	0.76	fam	ctaatatttcatgggttgtgtgc	ggctatggtttgaaagatacgaa
MF546	gt-62, at-23	465	58	0.56	tet	cctttagcgtggtggctagata	caaatcgactataggatatgtctttga
MF547	gt-21, at-9	328	58	0.7	fam	gcagtgtgcaaaagaagttcaaa	ccatcaacttctcgaaccaaaac
MF549	gt-12	301	58	0.8	hex	ttgttggttgagggtcagtaggt	tgtttggcttcctaatctgccta

Table 2.3 (continued)

MF550	gtt-6	291	58	0.24	fam	ggagtgttgctttattattggttga	atgatgctcactttcatgtcctg
MF551	gt-17	291	58	0.37	hex	tcactaaatggtagagccttctcg	ctgcccaaattaacccttcaaat
MF553	gt-35	314	58	0.46	fam	ttttgtaactgggcttgttgtga	tgcgctcctaacaataacaatca
MF556	gt-39	316	53	0.46	tet	gaatattcttagaaggtcaatggat	catatccgatccacatcagac
MF558	ga-19, gt-7	308	58	0	tet	agcccgtgatgattagggttagt	caccaaattattctctccatactcca
MF560	ct-4	354	58	0.36	fam	cgactcctataaatcggtcaagc	tcctctacaaggaatggatggaa
MF561	ta-4	260	58	0.62	hex	ttaacggtgttgagacgatcaaa	acaaaaccccccacaacctaaaag
MF562	gt-6, 3, at-3	300	58	0.24	tet	aaaggtttggtgggtaatttcgt	tcgaagtctgaaagatgtagatgga
MF563	gt-6	193	58	0.6	hex	gggtcttctttatcggcttcttg	ttacacctctacccacaaccaca
MF565	gt-6, at-7	259	58	0.46	hex	ggagcatcatattcagcctttatg	atcaaacctcacctagccaaaca
MF566	gtt-43	403	58	0.93	hex	gtgtttccaatctgtgttgatgc	cacacctgcaaatatcacaaacc
MF568	gt-36	383	53	0.77	fam	ttccaatatagaccatgaatatgaaaa	cctgtatggtaagtgatccatcg
MF569	at-13, gt-29	372	58	0.37	fam	cacaatcaagtgttccacgtttt	gatttgaggcattctcagcttct
MF570	gt-27	354	58	0.56	hex	ctttaacctgttgtgtgcctgct	caacatcaagtccaacaccaaga
MF571	gt-13, ga-9	233	53	0.86	fam	acctcctctggtgtttgattta	aagccctcctcttgttatctta
MF572	gt-21	375	58	0.92	fam	atattgcatcgcgacccactgct	cctccttttccacaccctaattcca
MF573	ca-4	147	58	0.61	tet	tgtgaatggaggtcagacaaatg	tcatctacaaccccaccacctta
MF574	gt-6, ga-6	233	58	0.76	hex	aagtggcgaagagcggtaataag	gagcagaaccctacaagttgacac
MF575	gat-11, ggt-4	337	58	0.85	fam	attagcacattcttgcctgttcc	ctagcaaacccaaggtgatatgg
MF576	gat-8	321	58	0.83	hex	ctggggatgatgaaagataatgc	caacacgggtcggacatactaat

Table 2.3 (continued)

MF578	ct-12	246	53	0	tet	tgaaggtctaaataacaaacaat	aaaatattttgagtcagtgagtat
MF579	ct-30	249	58	0.13	tet	acctctatcctctccttcctccat	gacggccattacgattgaaag
MF582	ga-28	315	53	0.78	fam	ccccatatagagatgtgatggag	aactgcaccaatcatttttagcc
MF583	ga-33	227	58	0.91	fam	tgaaactgactccagcagaaaca	aaacctcatcttctccaattctca
MF585	ga-26	281	58	0.73	fam	taacggagggtaatcggtgagt	tgacgacttaactacacggtcaaa
MF589	gt-37, at-8	365	58	0.86	tet	agcgtcaagtatgtcgttggag	tcgacctcataacaaggagaatatg
MF593	ct-12	295	58	0.63	fam	ccatgcagtttggcagtattgt	tcaagatgaagttcatctgagatgc
MF594	ct-34	284	58	0.82	fam	acattctgaggagttcagccttg	cgaggagaccgatcttctaatca
MF595	ag-5	350	58	0.69	hex	ttggtagccatgaaagagattgc	cttcttcttcttccccggtctc
MF597	ga-17	366	58	0.64	fam	ctagtaaagctgtttttccgcgta	gcatacaaatttcgctttgacgt
MF598	ct-6	335	53	0.84	tet	caccaacaactttccatgtacc	ttgatgtaactgaaggacttgatgtag
MF599	ca-49	316	58	0.47	fam	aacggtgagtccatcaggattt	cgttaggcttctcaattcattcg
MF603	ga-17	404	58	0.36	hex	ccgaagataacgagaggctgat	ggatcaaattaacaatggcgact
MF604	ga-10	344	58	0.81	hex	ccagacttttggggttacgatt	ctgctctatccactatggtttcg
MF605	ga-17	348	58	0.91	tet	tggtaatcgcagagaagcagac	tctccaccttaacattgaacttgat
MF608	ga-67	430	53	0.56	hex	tactagagagatgcgatacca	tccattatatcacgctttatctt
MF610	ca-25	264	58	0.85	fam	tttctcactcccattgtccaagt	agcttacatgcccaaagtcaaac
MF611	ca-24, ta-11	223	58	0.77	fam	tcaagaattcaccagttctaaaatca	tggtttgctactgtgagcatattgt
MF612	ga-18	372	53	0.61	hex	aagtaggagcagagggttttt	ccactactttaatttttattacaacac
MF613	ca-52	221	58	0.92	fam	aaaccatgatggagacatttatatcg	ttggtttgactcgttgatgtttg

Table 2.3 (continued)

MF614	ga-37	403	58	0.9	tet	ggtgatggctttcttttgatctc	ataccaatggtgcattcgaaaat
MF618	caa-32	300	58	0.44	fam	gactcatcgtgcgaattgtttg	tgttgagcttattgttgctgctc
MF621	gat-6	371	58	0.9	tet	taccgtgaaggatagggttgtg	ccatttccaggcatattctctagc
MF622	ca-38, cg-13	354	53	0.37	fam	tcgaataccaaaaatgttttctaaag	tgatgagtttgttttcagttgatga
MF624	ca-74	281	53	0.64	hex	atgccgtaataaataatcaatat	ctcttttcaaatctgataaact

The AC repeats had significantly (p= 0.03) higher mean heterozygosity (H=0.70) than the AG repeats (H= 0.62). The mean heterozygosities for the perfect repeats (H=0.71) was significantly higher than imperfect repeats (H=0.61, p=0.002) and compound repeats (H =0.62, p=0.05). The mean heterozygosities for repeat sizes more than 5 and less than 9 (H=0.54) was significantly lower than the heterozygosities for the repeat sizes more than 9 and less than 20 (H= 0.71; p<0.0001) and the repeat sizes more than 20 (H=0.69 ; p<0.0001). The mean values for the various measures of variation, H, PD, SDmw for the different classes of microsatellites are listed in Table 2.4. The PD followed the heterozygosity values ranging from 0 to 0.93 (Fig. 2.4) with a mean PD value of 0.64.

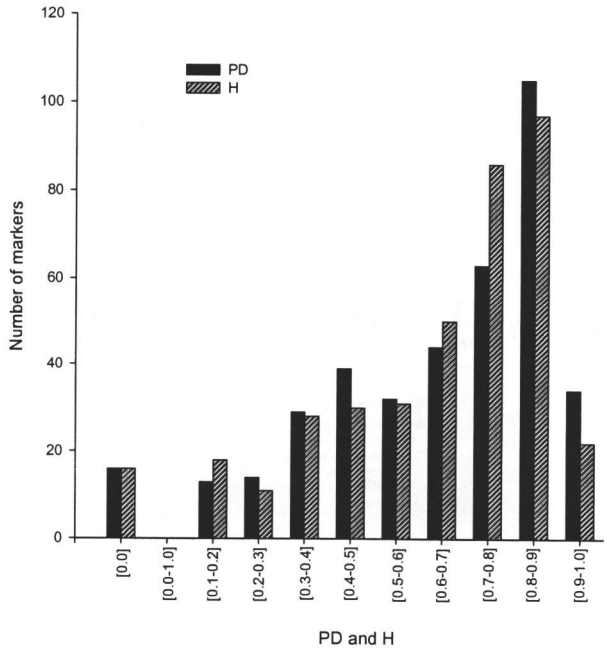


Figure 2.4. Distribution of the power of discrimination (PD) and heterozygosity (H) for the 389 SSR markers assayed on 14 meadowfoam accessions.

Table 2.4. The number and percentage of SSR markers developed and polymorphic SSR markers, heterozygosity (H), power of discrimination (PD), Standard deviation of molecular weight (SDmw) and the mean number of alleles per SSR marker (A) for 14 meadowofam accessions for different repeat classes. The percentage of functional markers was calculated by dividing the number of primers tested by the number of SSR markers developed. The percentage of polymorphic markers was calculated by dividing the number of polymorphic SSR markers by the number of SSR markers.

		SSR Markers developed		Polymorphic SSR Markers					
Class	Primer tested	Number	%	Number	%	H	PD	SDmw	A
Compound	84	37	44.0	34	91.9	0.62	0.61	12.1	6.5
Imperfect	390	272	69.7	263	96.7	0.61	0.62	9.2	5.9
Perfect	150	80	53.3	76	95.0	0.71	0.70	11.3	7.4
CA	136	64	47.1	63	98.4	0.70	0.69	12.1	6.5
GA	328	248	75.6	237	95.6	0.62	0.63	9.0	6.2
Dinucleotide	543	354	65.2	340	96.1	0.63	0.64	9.9	6.2
Trinucleotide	75	31	41.3	29	93.6	0.63	0.64	10.5	6.6
Tetranucleotide and above	6	4	66.6	4	100.0	0.73	0.63	8.0	4.3
$3 \leq N \leq 9$	246	161	65.4	150	93.2	0.54	0.56	6.6	4.6
$10 \leq N \leq 20$	220	139	63.2	135	97.1	0.71	0.71	10.3	7.1
$21 \leq N$	158	89	56.3	88	98.9	0.69	0.66	15.3	7.9
Total	624	389	62.3	373	95.9	0.63	0.64	9.9	6.3

There were a total of 233 markers which uniquely identified at least one meadowfoam accession from the panel of the fourteen meadowfoam accessions. Using a single marker we could uniqely identify from 1 to 11 accessions (Fig. 2.5) in the panel of 14 meadowfoam accessions, with a mean of 3.5 accessions per marker. Twenty five percent of the markers amplified in *Limnanthes floccosa* ssp. *floccosa* uniquely identified it from the rest of the accessions, which was followed by 24.0 % of the markers amplied in OMF40-11. *Limnanthes douglasii* ssp. *rosea* was the accession with the least amount of unique markers (14%) from the total number of markers amplified in the genomic DNA of *Limnanthes douglasii* ssp. *rosea*.

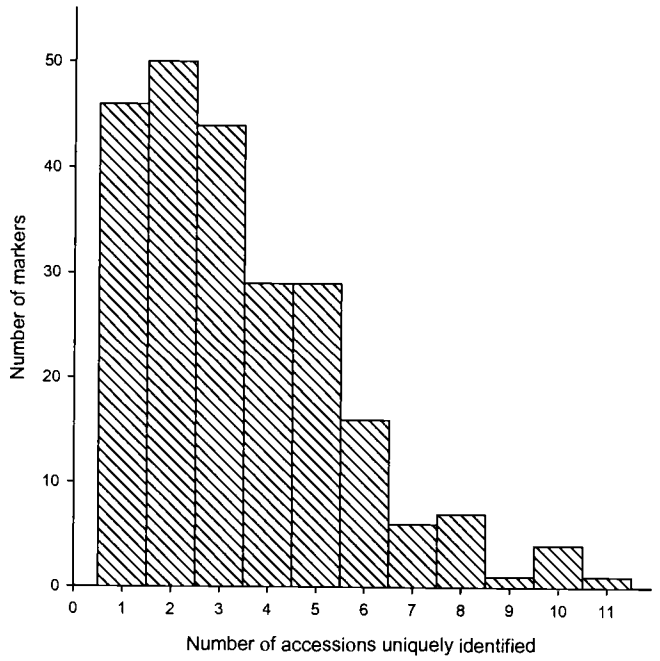


Figure 2.5. Distribution of the SSR markers that uniquely identify the meadowfoam accessions on our screening panel.

There were 50 markers which uniquely discriminated two meadowfoam accessions from the panel of 14 meadowfoam accessions, followed by 46 markers which unqiuely discriminated 1 meadowfoam accession from the panel of 14 meadwofoam accessions. There were 13 markers which uniquely discriminated more than half of the meadowfoam accessions on the panel.

Repeat unit length and variation

We examined the relationship between three measures of repeat unit length (MERC, MCRC and MRC) and two measures of variation (H and SDmw). MERC (310 polynucleotides) showed a low correlation ($r=0.23$, $p<0.0001$) with SDmw (Fig. 2.6) and a similar correlation with H ($r=0.24$, $p<0.0001$, Fig. 2.7). The correlation between MCRC (310 polynucleotides)and SDmw was lower ($r=0.21$, $p=0.0002$) and it also showed a similar correlation with H ($r=0.16$, $p=0.006$). However MRC (310 polynucleotides) showed higher correlations with SDmw ($r=0.60$, $p<0.0001$) and H ($r=0.44$, $p<0.0001$). Since only MRC showed higher correlations with the measures of variation, we further examined only the association of MRC with the SDmw and H in case of CA and GA repeats (Fig. 2. 8). It was found that for CA repeats (50 polynucleotides) MRC was strongly correlated ($r =0.64$, $p<0.0001$) with SDmw and H ($r =0.49$, $p=0.0003$). For GA repeats (198 polynucleotides) also the association between MRC, SDmw ($r = 0.59$, $p<0.0001$) and H ($r =0.45$, $p<0.0001$) was similar but of lesser magnitude than the CA repeats. In case of perfect repeats (65 polynucleotides) the asscoation between MRC, SDmw ($r =0.69$, $p<0.0001$) and H ($r =0.5$, $p<0.0001$) was very high, however in case of imperfect repeats (215 polynucleotides) the association between

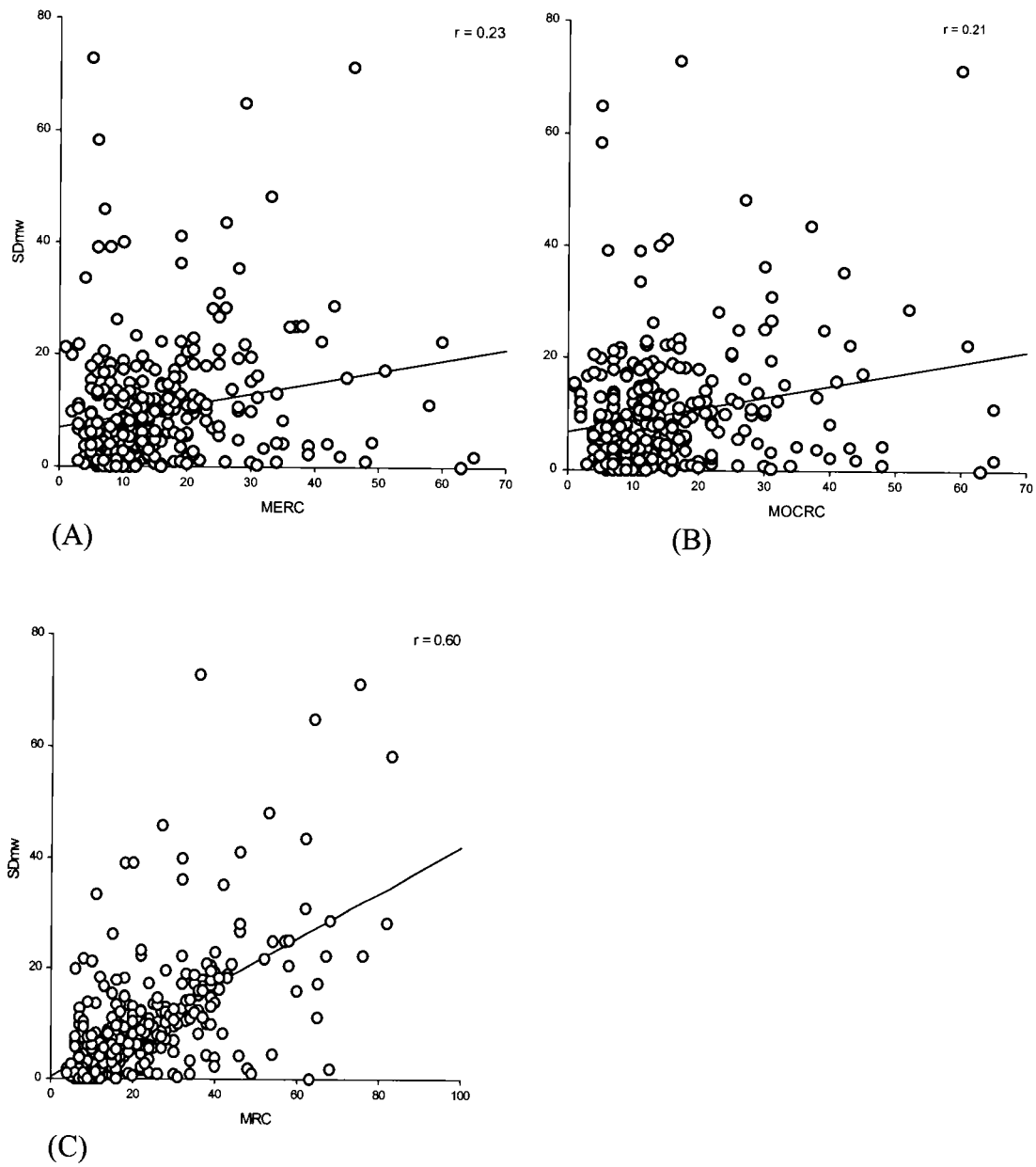
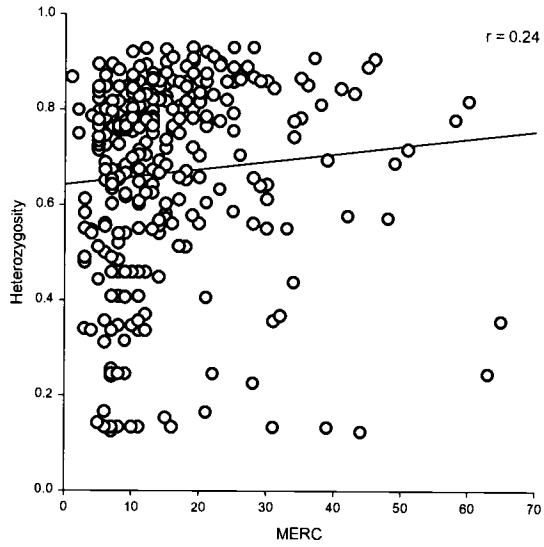
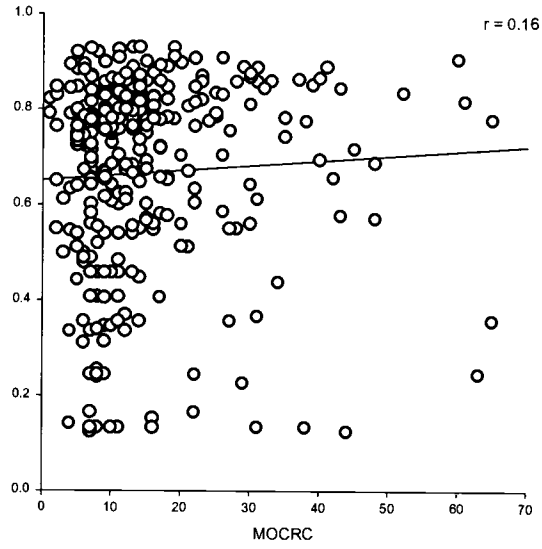


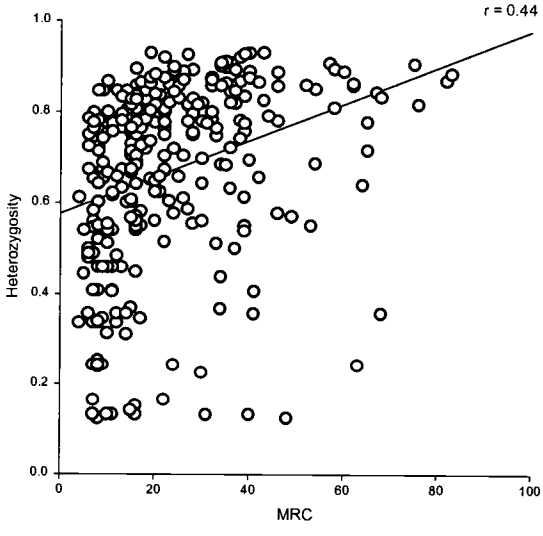
Figure 2.6. Correlation between the standard deviation of molecular weight (SDmw) for the SSR markers and (A)- mean repeat count (MERC), (B)- most common repeat count (MCRC) and (C) maximum repeat count (MRC).



(A)



(B)



(C)

Figure 2.7. Correlation between the heterozygosity for the SSR markers and (A)- mean repeat count (MERC), (B)- most common repeat count (MCRC) and (C) maximum repeat count (MRC).

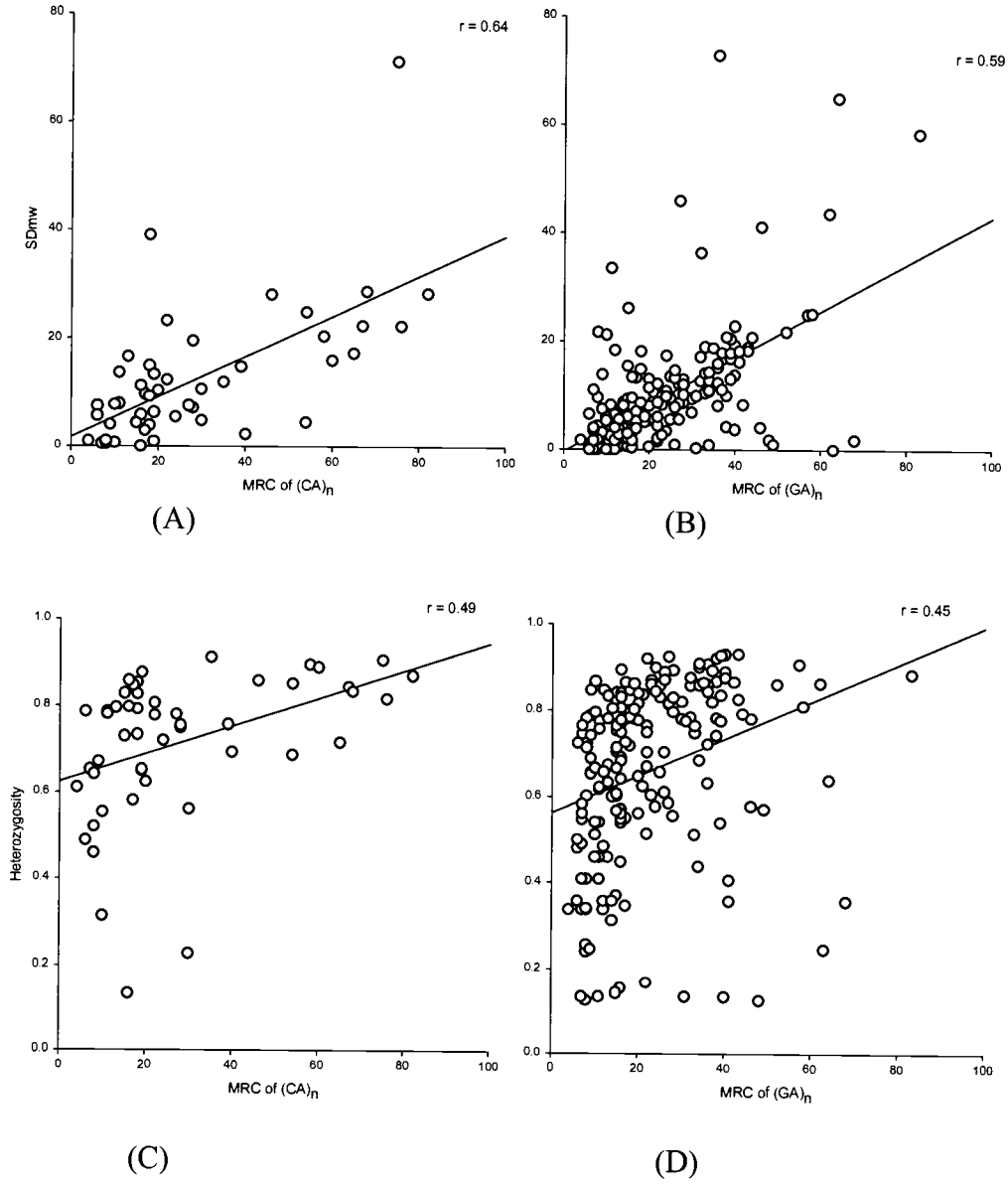


Figure 2.8. Correlation between the standard deviation of molecular weight (SDmw) and maximum repeat count for (A)- AC_n SSRs, (B)- AG_n SSRs; Heterozygosity and maximum repeat count for (C)- AC_n SSRs, (D)- AG_n SSRs.

MRC, SDmw ($r = 0.56$, $p<0.0001$) and H ($r = 0.42$, $p<0.0001$) was similar but of lesser magnitude.

Phenetic analysis based on SSR polymorphisms

Geneotypic data from 389 SSR markers was used to estimate the phentic relationships between 14 meadowfoam accessions. The genetic distance (GD) was based on the proportion of shared alleles. The GD ranged from 0.35 ± 0.022 (OMF64 and OMF109-2) to 0.69 ± 0.043 (LDR and LFG) with a mean of 0.57 ± 0.030. The distance estimated among inbred lines varied from 0.35 ± 0.022 (OMF64 and OMF109-2) to 0.55 ± 0.027 (LE76 and OMF64) with an average distance of 0.49 ± 0.025. In case of open pollinated cultivars and wild species the GD ranged from 0.44 ± 0.029 (LDD and LDN) to 0.69 ± 0.043 (LDR and LFG) with a mean GD of 0.58 ± 0.031. In case of section inflexae, the GD varied from 0.35 ± 0.022 (OMF64 and OMF109-2) to 0.63 ± 0.026 (OMF40-11 and LGG) with a mean GD of 0.55 ± 0.027, whereas for section reflaxe it varied from 0.44 ± 0.029 (LDD and LDN) to 0.50 ± 0.042 (LDR and LDD) with a mean GD of 0.46 ± 0.037. Cluster analysis using UPGMA on the bootstraped GD matrix revealed three major clusters (Fig. 2.9). The first cluster included the species and sub-species from the section reflexae and the remaining two clusters included species and sub-species from the section inflexae. The self pollinating wild species LFF , LFG and LGP were included in one cluster, whereas LAA (OMF 40-11, OMF 86, OMF 156), LAV (OMF64, OMF109-2), LM and LGG were included in one cluster. The goodness of fit for the cluster analysis was tested based on the cophenentic correlation value between the clustering tree matrix and the bootstrapped GD

matrix. The correlation was high ($r = 0.80$, $p=0.004$) indicating a good fit of the UPGMA cluster annalysis. The principal cooordinate ananlysis performed on the bootstrapaped GD matrix revealed the associations among the 14 meadowfoam accessions. The first three principal coordinates explained 43% of the genetic variance. The first two principal coordinates explained 18% and 14% of the genetic variances respectively. The first principal coordinate seperated the species and sub-species belonging to the section reflexae and section inflexae (Fig. 10). The second and the third principal coordinates separate the accessions in section inflexae into two groups where one group has a well understood breeding history and the other group has wild species.

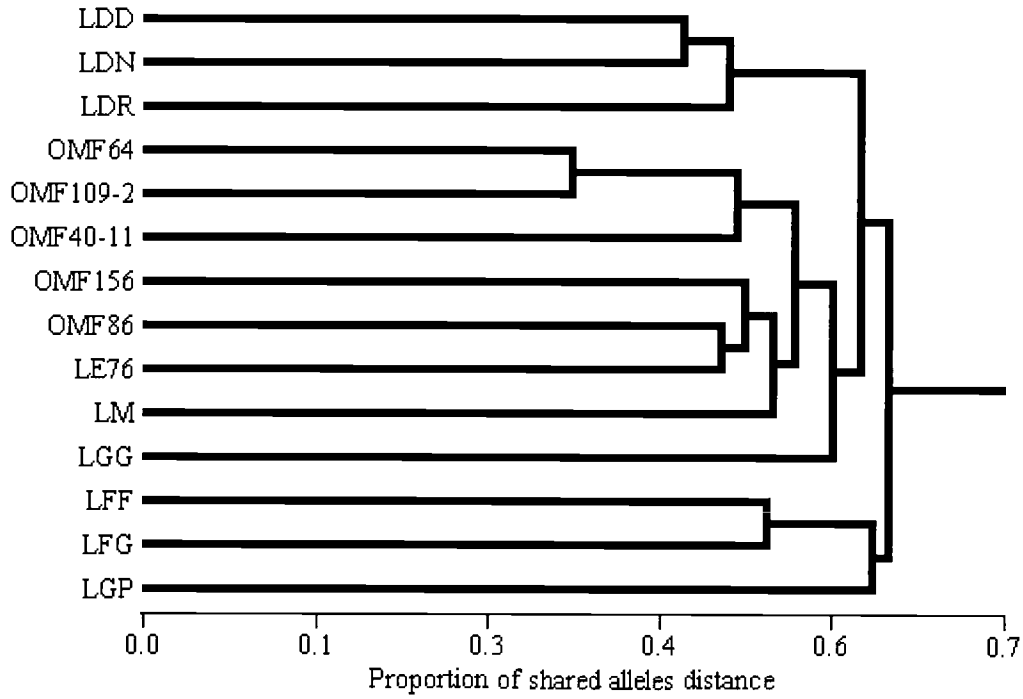


Figure 2. 9. A phenogram produced by UPGMA clustering of the proportion of shared alleles distance matrix estimated from 389 SSR fingerprints among 14 meadowfoam accessions.

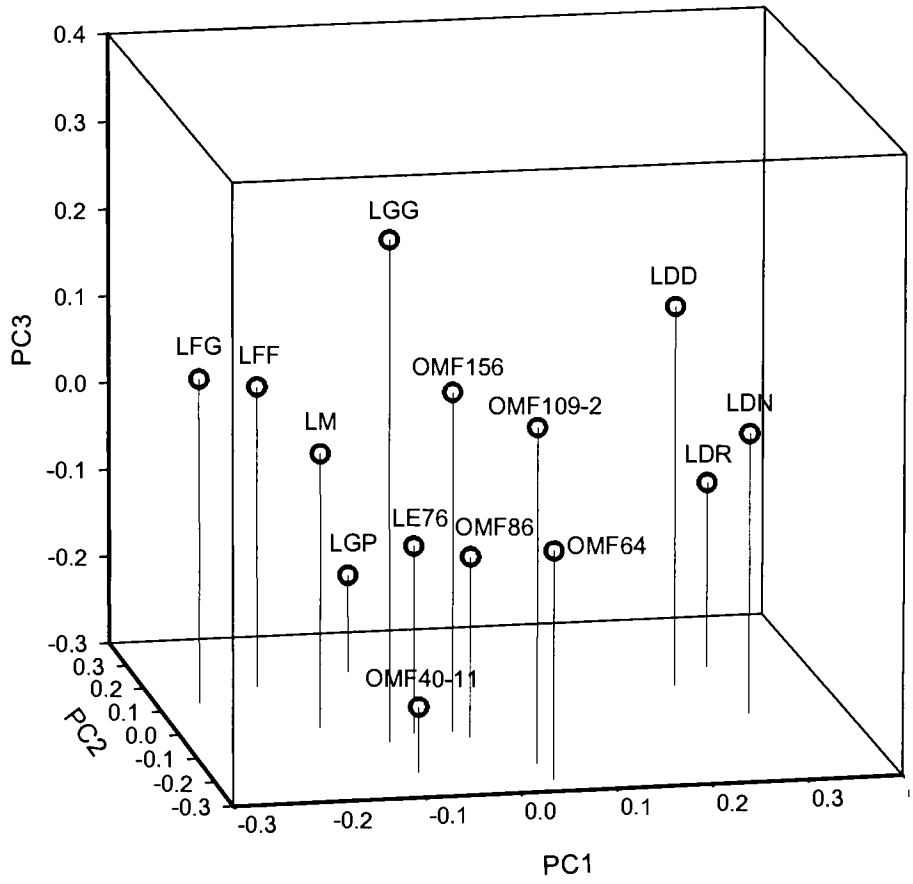


Figure 2.10. Associations among 14 meadowfoam accessions revealed by principal coordinate analysis performed on the genetic distance (GD) estimated from 389 SSR fingerprints.

Discussion

We describe herein the characterization of 389 SSR markers for meadowfoam. To characterize polymorphism at 389 microsatellite loci developed for meadowfoam we used a panel of 14 genotypes representing a diverse array of meadowfoam germplasm. This panel included 5 species and 9 sub-species from the sections, inflexae and reflexae. The enrichment procedure used in this study proved to be highly effective as it enabled the development of *Limnanthes alba* DNA libraries that were 77.5% enriched for SSR loci. This corresponds to a 260 to 775-fold enrichment when compared to the 0.1 to 0.3 % of the SSRs identified by screening non-enriched genomic libraries (Liu *et al.* 1995; Szewc-Mcfadden *et al.* 1996, Kubik *et al.* 1999, Saal and Wricke *et al.* 1999) of other crop plants. From the total clones sequenced, primer pairs could be designed for only 50.44% of the sequences confirming the huge attrition rates reported in SSR marker development process (Yu *et al.* 2001, Ramsay *et al.* 2000). One of the major factors involved in these huge attrition rates is the apparent duplication within the libraries presumably caused by selective amplification of certain sequences during the enrichment or possibly by the presence of SSRs within families of repeated sequences (Schmidt *et al.* 1991). Only sixty-two percent of the primers tested produced functional markers, suggesting that the sequences flanking the SSRs in meadowfoam are not conserved. However, 96% of the functional markers were polymorphic in the 14 genotypes screened. Forty-four percent (171 out of the 389) of the markers amplified a single locus whereas 56 % (218 out of 389) of the markers amplified 2 or more loci in an inbred. The high percentage of null alleles (24 % considering

only *Limnanthes alba* ssp. *alba* and *Limananthes alba* ssp. *versicolor*) along with the multilocus markers indicates the complexities of the meadowfoam genome.

One of the main attributes of microsatellite loci is their hypervariability, which makes them generally more informative than other molecular markers as RFLP (Wu & Tanksley 1993; Rongwen *et al.* 1995, Taramino and Tingey 1996). The average heterozygosity reported for microsatellites has generally been between 0.5 and 0.8 in mammals (Takezaki and Nei 1996) and between 0.5 and 0.9 in plants (Wu and Tanksley 1993, Rongwen *et al.* 1995, Taramino and Tingey 1996). Sixty six percent (255/ 389) of the markers developed had H values ranging from 0.7 to 0.93, whereas 63% (246/389) of the markers developed had PD values ranging from 0.7 to 0.93. The first index is an estimation of the probability that two alleles of any locus chosen at random from the populations of study are different from each other, whereas the discrimination power at a locus provides an estimate of the probability that two randomly sampled accessions in the study are differentiated by their allelic profile. High values for both H and PD suggest that meadowfoam is genetically diverse and the markers developed have higher discrimination power to identify the meadowfoam cultivars based on their allelic profile. Sixty percent of the markers developed uniquely identify atleast one meadowfoam accession from the panel of fourteen meadowfoam accessions. When a combination of the two markers, MF196 (uniquely identifies 11 acessions out of the 14) and MF356 (uniquely identifies 10 acessions out of 14), are used to fingerprint the whole 14 meadowfoam acessions all the 14 meadowfoam acessions are uniqely discriminated. This reinforces the conclusion that the markers developed have high discriminative power in identifying meadowfoam accessions.

SSR markers have been shown to be more polymorphic than RFLP markers in soybean (Akkaya *et al.* 1992 and Rongwen *et al.* 1995), wheat (Plaschke *et al.* 1995 and Roder *et al.* 1995), rice (Wu and Tanksley, 1993 and Olufowote *et al.* 1997), barley (Saghai Maroof *et al.* 1994) and rapeseed (Charters *et al.* 1996, Kresovich *et al.* 1995). Genetic variation in meadowfoam has so far been assessed only by using allozymes and AFLPs (Brown and Jain 1979, Mcneill and Jain 1985, Kessili and Jain 1985, Ritland and Jain 1984, Katengam, 2001). Both of these marker types are technically difficult to use in regular breeding programs (Powell *et al.* 1996, Smith *et al.* 1997). The average heterozygosity revealed by AFLPs was 0.31 whereas SSRs have revealed an average heterozygosity of 0.63. This difference is primarily due to the bi-allelic nature of AFLPs. The mean number of alleles estimated using SSRs was three times more than the mean number of alleles estimated in all of the studies using the allozyme markers.

It has been shown in early studies in human (Weber 1990) and confirmed in several organisms that the variability of microsatellite markers correlates well with the length of the tandem arrays (Goldstein & Clark 1995; Innan *et al.* 1997). The suggestion of stronger correlation between the MRC and the allele number than between the MERC, MCRC and allele number suggests that the mutation rate of the microsatellites does indeed increase with repeat count but not smoothly (Goldstein & Clark 1995). This is further confirmed by the stronger magnitudes of positive correlation between MRC and SDmw, which is three times higher than the correlation between MERC, MCRC and SDmw. Correlations between H and MRC, MERC and MCRC show similar patterns but are of lower magnitude, possibly reflecting the fact that H is a frequency measure and thus saturates rather quickly as

the number and frequency of alleles increase (Schug *et al.* 1998). The stronger magnitudes of positive correlation between MRC and SDmw for perfect repeats than the imperfect repeats confirms the previous results of Goldstein & Clark (1995), suggesting imperfections result in a lower variance than would be expected based on the size of the repeat string on the other side of the imperfection. The implication in that when imperfections split microsatellites, the mutation rate in the larger repeat string on one side of the imperfection is suppressed by the presence of the smaller string of repeats on the other side. Similar observations were made by Schug *et al.* (1998) in *Drosophila* except that much larger number of data points in the present study provides stronger validation of this trend. In fact the present study is the only study where such a large number of microsatellites have been used to test this hypothesis. Goldstein and Clark (1995) used 18 microsatellites; Schug *et al.* (1998) used 49 microsatellites whereas Cho *et al.* (2000) used a maximum of 120 microsatellites. It is important to understand that there is an underlying assumption based on which this analysis was carried out. The assumption being that the difference in the allele sizes for a marker in the fourteen-meadowfoam accessions screened was primarily due to change in the repeat numbers in the sequence but not due to any insertions or deletions in the regions flanking the repeats. The lack of sequence information on repeat length of loci assayed in all the fourteen meadowfoam genotypes prevents us from confirming the assumption in our study. However it has been shown in *Arabidopsis* that for most of the SSR loci, length polymorphism results only from variations in the number of repeats except for a few others where some variability was noted in the flanking sequences and for compound and interrupted loci containing two arrays of repeats, length variations

preferentially affect the longest one (Loridon *et al.* 1998). It is also important to remember that the number of units in a perfect array of tandem repeats rather than the absolute length of the array in the base pairs is the principal determinant of a given microsatellite's propensity to mutate (Cho *et al.* 2000).

Several factors apart from hypervariability have fueled the development of SSR markers in meadowfoam and other crop plants. High-throughput, semi-automated assay systems have been developed, whereby SSR markers can be multiplexed, by pooling separately produced amplicons. We performed SSR genotyping on a semi-automated, high-throughput system (the ABI 377) and designed SSR primers to facilitate multiplexing by dividing SSR markers into length and color bins. Reference allele amplicon lengths were uniformly sorted into 100 bp bins spanning ~100 to 550 bp and the three fluorophores (6 FAM, TET and HEX) were uniformly distributed among the length bins. Because the allele sizes for the screening panel genotypes were not known a priori, we multiplexed by color only in the primer and polymorphism screening process. The length and color bins we used for primer design routinely permit assays of 9 to 12 SSR markers per lane on ABI377 (Fig. 2.11).

Cultivated meadowfoam is based on *Limnanthes alba*, which belongs to section inflexae. This section comprises of 4 species namely *Limnanthes alba*, *Limnanthes floccosa*, *Limnanthes gracilis* and *Limnanthes montana*. The primary gene pool of meadowfoam is composed of *Limnanthes alba* ssp. *alba* and

Limnanthes alba ssp. *versicolor*, whereas *Limnanthes floccosa*, *Limnanthes gracilis* and *Limnanthes montana* are identified as a secondary gene pool of meadowfoam. Based on the fertility of inter-subspecific and inter-specific hybridizations with *Limnanthes alba* the species belonging to section reflexae including *Limnanthes bakeri*, *Limnanthes douglasii*, *Limnanthes macounii*, *Limnanthes striata* and *Limnanthes vinculans* are classified as tertiary gene pool of meadowfoam (Knapp and Crane 1999). The phenetic analysis of the 389 SSR fingerprints of fourteen meadowfoam accessions revealed the genus *Limnanthes* to be diverse. The section inflexae was more diverse than the section reflexae. This may be due to the fact that our screening panel included all the four species in section inflexae but only one species from section reflexae. The diversity among the inbreds was marginally lower than the diversity among the open pollinated cultivars and wild species, which was reflected in their mean genetic distances. The pattern of diversity as illustrated in the phenogram resulting from the cluster analysis was concordant with species, subspecies geography and breeding origin. All the sub-species belonging to *Limnanthes douglasii* in section reflexae were grouped into one cluster. The cluster analysis showed that there was only one cluster with a genetic distance below 0.40, which was the *Limnanthes alba* ssp. *versicolor* cluster, which included OMF 109-2 and OMF64. However the lower genetic distance results from the fact that OMF 64 is one of the parents of OMF 109-2. The two sub-species, *Limnanthes gracilis* ssp. *gracilis* and *Limnanthes gracilis* ssp. *parishii* were separated and were

far apart from each other. The two members of *Limnanthes gracilis* were found in different geographical area (Mason 1952). *Limnanthes gracilis* ssp. *gracilis* was found in Klamath mountain region of southwestern Oregon while *Limnanthes gracilis* ssp. *parishii* was found only in a few sites in San Diego County, California. *Limnanthes gracilis* ssp. *parishii* is highly autogamous and hence it was grouped with *Limnanthes floccosa* ssp. *floccosa* and *Limnanthes floccosa* ssp. *grandiflora* that are highly autogamous due to their cliestogamous nature of flowering. *Limnanthes montana* was found to be distributed intermediate in the range of two sub-species of *Limnanthes gracilis* from Mariposa County, Sierra Nevada southward to Tulare county California. The grouping of *Limnanthes gracilis* ssp. *gracilis* together with *Limnanthes alba* and *Limnanthes montana* agreed with the previous numerical taxonomic study using morphological traits by Ornduff and Crovello (1968) and an artificial hybridization study by Ornduff (1971). Hybrids between *Limnanthes alba* and *Limnanthes gracilis* ssp. *parishii* had relatively sterile pollen whereas those of *Limnanthes gracilis* ssp. *gracilis* and *Limnanthes alba* had highly viable pollen. Moreover, our results seem to support the hypothesis that was described by Mason (1952) that *Limnanthes montana* might be the remnant of these two populations of *Limnanthes gracilis*.

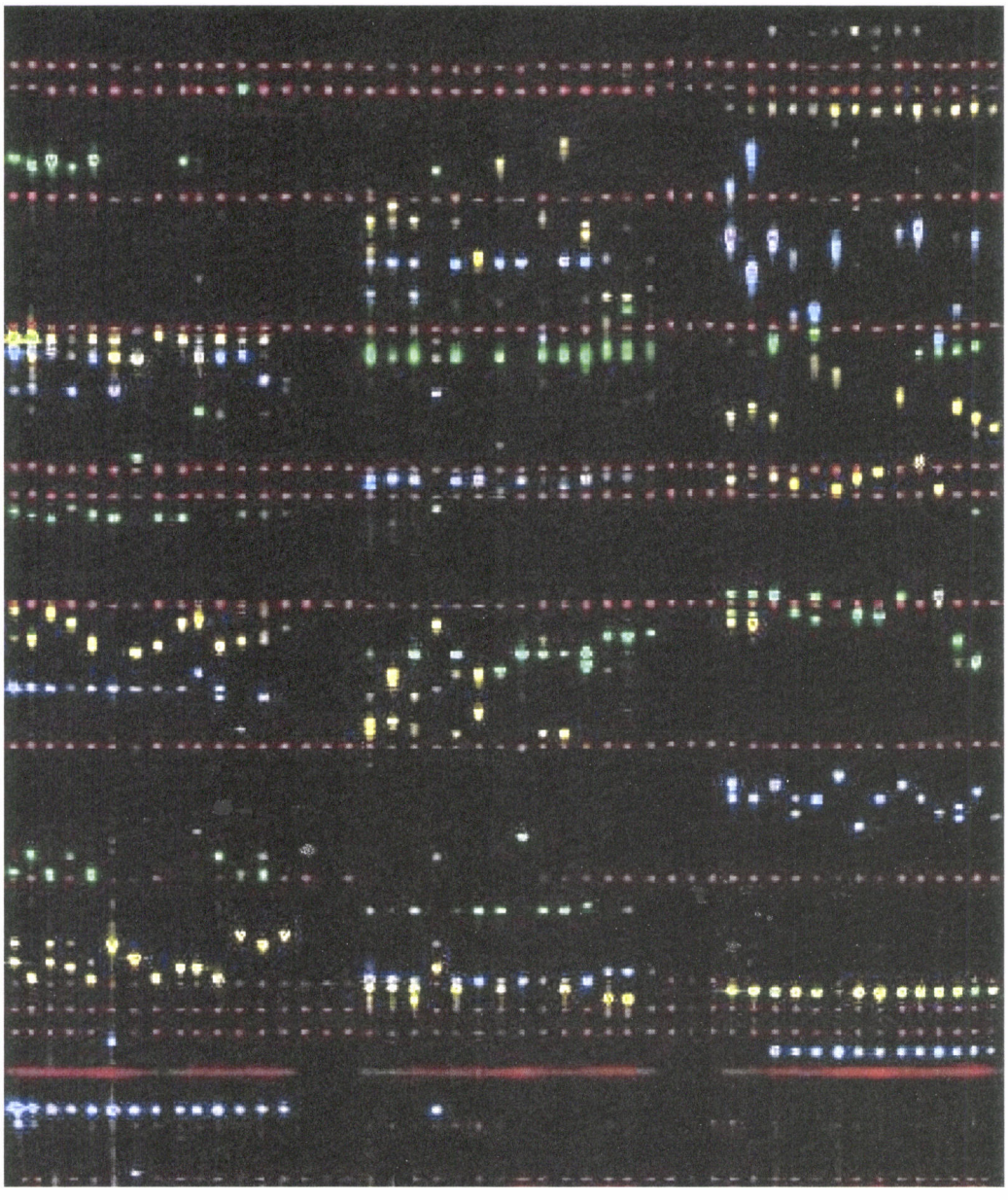


Figure 2.11. Gel image of 3 arrays of 14 meadowfoam genotypes screened with 3 sets of 9 multiplexed SSRs.

Once these two populations were continuously distributed, however climatic and geographical changes along with the extinction of many *Limnanthes* populations caused sub-division and subsequent isolation of these two subspecies.

Substantial yearly fluctuations in population sizes are seen in *Limnanthes*. Also these populations were considerably smaller (by approximately 90%) in 1989 and 1990 (Dole and Sun, 1992). Such demographic fluctuations could cause extinction of rare taxa, especially through their interaction with loss of genetic variation (Lande 1988). The genetic survey with the SSRs will allow us to estimate the extent of natural hybridization in sites where these species coexist. Genetic diversity within and between populations and gene flow between populations can also be estimated. Genetic data can be used to identify groups of populations that may contain high frequencies of locally adapted alleles. Marshall and Brown (1975) have suggested that locally common alleles should receive priority in sampling for conservation purposes. This is because wide spread common alleles will often be sampled without conscious efforts, whereas locally common alleles are much likely to be captured. These may be maintained at high local frequencies because they represent adaptations to local variation. All the SSR markers developed in the present study will be helpful in conservation of the endangered meadowfoam germplasm.

The principal coordinate analysis revealed similar patterns as that of the cluster analysis. The first three principal coordinates using SSR data explained 43% of the genetic variance that was higher than the 37% of the genetic variance explained by the first three principal coordinates using the AFLP data (Katengam 2000).

SSRs in this study showed high polymorphism among the fourteen meadowfoam accessions. Both the cluster analysis and principal coordinate analysis showed the evidence that SSRs are excellent tools for fingerprinting meadowfoam. This study demonstrates an automated system for unambiguous meadowfoam genotype identification. In addition, SSR technology presents the advantages of reliability, reproducibility and time and cost effectiveness over other marker systems. It is our intention to select a number of highly informative loci with high PD values which can be used to provide the basis for a meadowfoam DNA profile system. Such a system should be an excellent complement to the morphological markers that are currently used to obtain plant variety protection certificates for new meadowfoam cultivars. SSR markers reported here also provide an excellent tool for the construction of a genetic map which could be used for trait mapping and marker-assisted selection.

In conclusion, this paper reports the development and utility of a large set of highly polymorphic SSR markers for meadowfoam. Although direct costs involved in SSR development are relatively high, SSRs will almost certainly become the markers of the choice for the molecular breeding of meadowfoam.

References

Akkaya MS, Bhagwat AA, Cregan PB (1992). Length polymorphisms of simple sequence repeat DNA in soybean. Genetics, 132: 1131-1139.

Bagby MO, Smith CR, Miwa TK, Lohmar RL, Wolff LA (1961) A unique fatty acid from *Limnanthes douglasii* seed oil: The C_{22} diene. Journal of Organic Chemistry, 26: 1261-1265.

Bowcock AM, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd JR, Cavalli-Sforza LL (1994) High resolution of human evolutionary trees with polymorphic microsatellites. Nature, 368: 455-457.

Brown CR, Jain SK (1979) Reproductive system and pattern of genetic variation in two Limnanthes species. Theor. Appl. Genet., 54: 181-190.

Charters YM, Robertson A, Wilkinson MJ, Ramsay G (1996) PCR analysis of oilseed rape cultivars (*Brassica napus* L. ssp. *oleifera*) using 5'-anchored simple sequence repeat (SSR) primers. Theor. Appl. Genet., 92: 442-447.

Cho YG, Ishii T, Temnykh S, Chen X, Lipovich L, McCouch SR, Park, Ayres N, Cartinhour S (2000). Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). Theor. Appl. Genet., 100: 713-722.

Dietrich WF, Miller J, Steen R, Merchant MA, Damron- Boles D, Husain Z, Dredge R, Daly Mt, Ingalls KA, O'Connor JJ, Evans CA, DeAngelis MM, Levinson DM, Kruglyak L, Goodman N, Copelong NG, Jenkins NA, Hawkins TL, Stein L, Page DC, Lander ES (1996) A comprehensive genetic map of the mouse genome. Nature, 380: 149-152.

Dole JA, Sun M (1992) Field and genetic survey of the endangered Butte County meadowfoam-*Limnanthes floccossa* subsp. *californica* (Limnanthaceae). Conservation Biology, 6: 549-558.

Don RH, Cox PT, Wainwright BJ, Baker K, Mattick JS (1991) 'Touchdown' PCR to circumvent spurious priming during gene amplification. Nucleic Acid Res.,19: 4008.

Dudley, JW (1993) Molecular markers in plant improvement: manipulation of genes affecting quantitative traits. Crop Sci., 33: 660-668.

Eastman DC (1990) Rare and endangered plants of Oregon. Beautiful America publishing company, USA.

Edwards KJ, Barker JHA, Daly A, Jones C, Karp A (1996). Microsatellite libraries enriched for several microsatellite sequences in plants. BioTechniques, 20: 8-760.

Goldstein DB, Clark AG (1995) Microsatellite variation in North American populations of *Drosophila melanogaster*. Nucleic Acid Res., 23: 3882-3886.

Innan H, Terauchi R, Miyashita T (1997) Microsatellite polymorphism in natural populations of the wild plant *Arabidopsis thaliana*. Genetics, 146: 1441-1452.

Isbell TA (1997) Development of meadowfoam as industrial crop through novel fatty acid derivatives. Lipid Technology, 9: 140-144.

Jain SK (1986) Domestication of Limnanthes (Meadowfoam) as a new oil crop. In: Plant domestication induced mutation: Proceedings of an advisory group meeting on the possible use of mutation breeding for rapid domestication of new crop plants (ed. FAO/IAEA.), pp.121-134, Vienna, Austria.

Karagyozov L, Kalcheva ID, Chapman VM (1993). Construction of random small-insert genomic libraries highly enriched for simple sequence repeats. Nucleic Acid Res., 21:3911-3912.

Katengam S (2000) DNA finger printing and genome mapping in meadowfoam. PhD. Thesis, Oregon State University.

Kesseli RV, Jain SK (1985) Breeding systems and population structure in *Limnanthes*. Theor. Appl. Genet., 71: 292-299.

Kloosterman AD, Budowle B, Daselaar, P (1993) PCR-amplification and identification of the human D1S80 VNTR locus: Amplification conditions, population Genetics and application in forensic analysis. International Journal of Legal Medicine, 105: 257-264.

Knapp SJ, Crane JM (1999) The development of meadowfoam as an industrial oilseed: breeding challenges and germplasm resources. In: Proceedings of the Fourth national Symposium on New Crops (ed. J.Janick). Wiley, New York, USA.

Kresovich S, Szewc-McFadden AK, Blick SM (1995) Abundance and characterization of simple sequence repeats (SSRs) isolated from a size-fractionated genomic library of *Brassica napus* L. (rapeseed). Theor. Appl. Genet., 91: 206-211.

Kubik C, Meyer WA, Gaut BS (1999) Assessing the abundance and polymorphism of simple sequence repeats in perennial ryegrass. Crop Sci., 39: 1136-1141.

Lande R (1988) Genetics and demography in biological conservation. Science, 241:1455-1460.

Litt M, Luty JA (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. American Journal of Human Genetics, 44: 397-401.

Liu ZW, Jarret RL, Kresovich S, Duncan RR (1995) Characterization and analysis of simple sequence repeat (SSR) loci in seashore paspalum (*Paspalum vaginatum* Swartz). Theor. Appl. Genet., 91: 47-52.

Lodhi MA, Ye GN, Weeden NF, Resich BI (1994) A simple and efficient method for DNA extraction from grapevine cultivars and Vitis species. Plant. Mol. Biol. Rep. 12: 6-13.

Loridon K, Cournoyer B, Goubely C, Depeiges A, Picard G (1998) Length polymorphism and allele structure of trinucleotide microsatellites in

natural accessions of *Arabidopsis thaliana*. Theor. Appl. Genet., 97: 591-604.

Mantel NA (1967) The detection of disease clustering and a generalized regression approach. Can. Res., 27: 209-220.

Marshall DR, Brown AHD (1975) Optimum sampling strategies in genetic conservation. In: Crop genetic resources for today and tomorrow. (ed. OH Frankel and JG Hawkes) Cambridge University Press, Cambridge, England.

Mason CT (1952) A systematic study of the genus *Limnanthes*. Botanical Publications, 25: 455-512.

McNeill CI, Jain SK (1985) Genetic differentiation studies and phylogenetic inference in plant genus *Limnanthes* (section Inflexae). Theor. Appl. Genet., 66: 257-269.

Messmer MM, Melchinger AE, Herrmann RG, Boppenmaier J (1993) Relationship among early European maize inbreds: II. Comparison of pedigree and RFLP data. Crop Sci., 33: 944-950.

Minch E, Ruiz-Linares A, Goldstein D, Feldman M and Cavalli-Sforza LL (1997) MICROSAT (ver. 2.0d) A computer program for calculating various statistics on microsatellite allele data. Available at http://hpgl.stanford.edu/projects/microsat/.

Mumm RH, Dudley JW (1994) A classification of 148 U.S. maize inbreds: I. Cluster analysis based on RFLPs. Crop Sci., 34: 842-851.

Olufowote JO, Xu Y, Chen X, Park WD, Beachell HM, Dilday RH, Goto M, McCouch SR (1997) Comparative evaluation of within-cultivar variation of rice (*Oryza sativa* L.) using microsatellite and RFLP markers. Genome, 40: 370-378.

Ornduff R (1971) Systematic studies of Limnanthaceae. Madrona, 21: 103-111.

Ornduff R, Crovello TJ (1968) Numerical taxonomy of Limnanthaceae. American Journal of Botany, 55: 173-182.

Ott, J. 1991. Analysis of human genetic linkage. *John Hopkins University Press.* Baltimore, Maryland.

Plaschke J, Ganal MW, Roder MS (1995) Detection of the genetic diversity in closely related bread wheat using microsatellite markers. Theor. Appl. Genet., 91: 1001-1007.

Powell W, Morgante M, Andre C, Hanafey C, Vogel J, Tingey S, Rafalski A (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol. Breed., 2: 225-238.

Rae SJ, Aldam C, Dominguez I, Hoebrechts M, Barnes SR, Edwards KJ (2000) Development and incorporation of microsatellite markers into the linkage map of sugar beet (*Beta vulgaris* ssp.). Theor. Appl. Genet., 100: 1240-1248.

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

Ritland K, Jain SK (1984) A comparative study of floral and electrophoretic variation with life history variation in Limnanthes alba (Limnanthaceae). Theor. Appl. Genet., 63: 243-251.

Roder MS, Plaschke J, Konig SU, Borner A, Sorrels ME, Tanksley SD, Ganal SW (1995) Abundance, variability and chromosomal location of microsatellites in wheat. Molecular General Genetics, 246: 327-333.

Rohlf FJ (1993) NTSYS-pc numerical taxonomy and multivariate analysis system. Ver.2.02i Exster software, Setauket, New York.

Rongwen J, Akkaya MS, Bhagwat AA, Lavi V, Cregan PB (1995) the use of microsatellite DNA markers for soybean genotype identification. Theor. Appl. Genet., 90: 43-48.

Saal B, Wricke G (1999) Development of simple sequence repeat markers in rye (*Secale cereale* L.) Genome, 42: 964-972.

Saghai Maroof MA, Biyashev RM, Yang GP, Zhang Q, Allard RW (1994) Extraordinarily polymorphic microsatellite DNA in barley: species diversity, chromosomal locations, and population dynamics. Proc. Nat. Acad. Sci., 91: 5466-5470.

Schug MD, Hutter CM, Wetterstrand KA, Gaudette MS, Mackay TFC, Aquadro CF (1998) The mutation rates of di-, tri- and Tetranucleotide repeats in *Drosophila melanogaster*. Mol. Biol. Evol., 15: 1751-1760.

Schug MD, Mackay TFC, Aquordro CF (1997) Low mutation rates of microsatellite loci in *Drosophila melanogaster*. Nature Genetics, 15: 99-102.

Smith CR, Bagby MO, Miwa TK, Lohmar RL, Wolff IA (1960) Unique fatty acids from *Limnanthes douglasii* seed oil: the C_{20} – and C_{22} – monoenes. Journal of Organic Chemistry, 25: 1770-1774.

Smith JSC, Kresovich S, Hopkins MS, Mitchell SE, Dean RE, Woodman WL, Lee M, Porter K (2000) Genetic diversity among elite sorghum inbred lines assessed with simple sequence repeats. Crop Sci., 40: 226-232.

Smith MSC, Chin ECL, Shu H, Smith OH, Wall SJ, Senior ML, Mitchell SE, Kresovich S, Ziegle J (1997) An evaluation of utility of SSR loci as molecular markers in maize (*Zea mays* L.) comparisons with data from RFLPs and pedigree. Theor. Appl. Genet., 95: 163-173.

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

Sneath PHA Sokal RR (1973) Numeric Taxonomy. Freeman, San Francisco.

Szewc-McFadden AK, Kresovich S, Bliek SM, Mitchell SE, McFerson JR (1996) Identification of polymorphic, conserved simple sequence repeats (SSRs) in cultivated *Brassica* species. Theor. Appl. Genet., 93:534-538.

Takezaki N, Nei M (1996) Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. Genetics, 144: 389-399.

Tang S, Yu JK, Slabaugh MB, Shintani DK, Knapp SJ (2002) Simple sequence repeat map of the sunflower genome. Theor. Appl. Genet. (submitted).

Taramino G, Tingey S (1996). Simple sequence repeats for germplasm analysis and mapping in maize. Genome, 39: 227-287.

Weber JL (1990) Informativeness of Human $(dC\text{-}dA)_n \cdot (dG\text{-}dT)_n$ polymorphisms. Genomics,7:524-530.

Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Amer. J. of Human Genet., 44: 388-396.

Wu KS, Tanksley SD (1993) Abundance, polymorphism and genetic mapping of microsatellies in rice. Molecular General Genetics, 241: 225-235.

Yang GP, Saghai Maroof, MA Xu CG, Zhang Q, Biyashev RM (1994) Comparative analysis of microsatellite DNA polymorphisms in landraces and cultivars of rice. Molecular General Genetics, 245: 187-194.

Yu JK (2001) Simple sequence marker development and mapping in cultivated sunflower. PhD thesis, Oregon State University.

CHAPTER 3

THE MEADOWFOAM GENOME AND SIMPLE SEQUENCE REPEAT GENOME MAP

Venkata K. Kishore[1], Sonali D. Gandhi[1], Jimmie Crane[1] and Steven J.Knapp[1]

[1]Department of Crop and Soil Science, Oregon State University, Corvallis, Oregon, 97331, USA

Abstract

Meadowfoam (*Limnanthes alba* Benth.) is a predominantly allogamous, self-compatible, diploid with a small haploid chromosome number (x = 5), the same as *Arabidopsis*, a model species with a physically small genome (0.35 pg). The physical and cytological characteristics of the meadowfoam genome have not been described, and the present genetic map of meadowfoam is solely comprised of amplified fragment length polymorphisms. Our specific aims were to describe the karyotype and physical size of the meadowfoam genome and develop a simple sequence repeat (SSR) map for meadowfoam. The physical size of the meadowfoam genome was estimated to be 5.52 pg using flow cytometry; thus, the meadowfoam genome is ca. 16 times larger than the *Arabidopsis* genome. Karyotype analyses revealed that the meadowfoam genome is made up of two metacentric and three submetacentric chromosomes. Meadowfoam has two pairs of chromosomes with subterminal nucleolar organizing regions (NOR's). The genetic map was constructed by genotyping 96 (OMF40-11 x OMF64) x OMF64 BC_1 progeny with 90 SSR markers and was comprised of 84 SSR loci dispersed among five linkage groups with 11 to 22 SSR loci per linkage (6 SSR loci segregated independently). The five linkage groups presumably correspond to the five haploid chromosomes of meadowfoam. The map was 988.7 cM long with a mean density of 11.8 cM and minimal clustering of loci.

Key words: SSR, linkage map, karyotype, flow cytometry, meadowfoam

Introduction

Cultivated meadowfoam (*Limnanthes alba* Benth.) is an annual oil seed crop plant native to Southern Oregon and Northern California (Mason 1952, Kalin 1971, Jain 1986). Meadowfoam seed oil contains unique unsaturated very long chain fatty acids (C_{20} and C_{22}) with outstanding oxidative stability (Isbell 1997). These novel characteristics of meadowfoam seed oil are of industrial interest and have fueled the development of meadowfoam as an oilseed crop (Knapp and Crane 1999). Meadowfoam belongs to Order Brassicales and Family Limnanthaceae (Wheeler et al. 2000). The genus Limnanthes comprises of nine diploid (2n=2x=10) species including *Limnanthes alba*. These are the only non-Brassicaceae species in Brassicales order which have the same chromosome number as that of the model organism *Arabidopsis thaliana*.

Meadowfoam has been cultivated since 1974, when the first non-shattering cultivar, Foamore, was developed and released for commercial production (Calhoun and Crane 1975). Since then, four cultivars have been released for cultivation. In order to improve our efficiency of developing new cultivars for increasing the productivity of meadowfoam it is essential to understand the genome of meadowfoam along with the genetics of economically important traits. Our lab has focused on developing molecular breeding tools for meadowfoam for the past six years, as there were very few biochemical markers (Arroyo 1975; Brown and Jain 1979; Kesseli and Jain 1985) and virtually no DNA markers for meadowfoam

until 2001. There is also no information published about the physical and cytological characteristics of the meadowfoam genome. With the advent of flow cytometry, one can easily quantify DNA in a large number of plants within a relatively short period of time compared to microdensitometry (Arumuganathan and Earle 1991a; 1991b). However the physical genome size of meadowfoam is not known yet. Propach (1934) and Resende (1937) studied the chromosome morphology of meiotic chromosomes of *Limnanthes douglasii* and *Limnanthes alba*. They disagreed on the number of satellite chromosomes in meadowfoam. Mason (1952) also studied the meiotic chromosome morphology and agreed with Resende (1937) that meadowfoam has only two pairs of satellite chromosomes. However, none of these investigators produced a detailed karyotype of meadowfoam chromosomes.

Genetic maps are important in plant breeding and are a powerful tool for localizing and isolating genes underlying both simple and complex traits. Katengam *et al.* (2002) produced the first genetic map of meadowfoam using 103 amplified fragment length polymorphisms (AFLPs). AFLPs, however are dominant markers and difficult to use in regular breeding programs (Powell et al. 1996; Smith et al. 1997). Simple sequence repeat (SSR) markers have become important genetic markers in many plant genomes because of the high level of polymorphism (Wang et al. 1994). In addition, they are polymerase chain reaction (PCR) based, facilitating easy screening compared to restriction fragment length polymorphisms

(RFLPs). They can be easily scored and transferred among different laboratories. Genetic maps based on SSRs have been developed for wheat (Roder *et al.* 1998), rice (Temnykh *et al.* 2000), barley (Ramsay *et al.* 2000), sunflower (Tang *et al.* 2002) and many other crop plants. Kishore *et al.* (2002) developed 389 simple sequence repeat (SSR) markers for meadowfoam. Ninety six percent of the markers developed were polymorphic in a screening panel of fourteen meadowfoam accessions.

Our specific aims in this study were to describe the karyotype and physical size of the meadowfoam genome and develop a simple sequence repeat (SSR) map for meadowfoam.

Materials and Methods

Flow cytometry analysis

Nuclear genome sizes of meadowfoam, sunflower, barley, and chicken were measured using flow cytometry. Meadowfoam genome size was measured using two different methods and flow cytometers. In the first method we used 4 , 6-diamidino-2-phenylindole dihydrochloride (DAPI) for staining the suspension of intact nuclei and measured the DNA content in the nuclei using a Partec PA flow cytometer (Partec GMbH, Munster, Germany), with a mercury arc lamp for excitation in UV, blue and green light, which enabled detection of fluorescent signals from DAPI. Chicken red blood cells (CEN singlet, product # 1013 from Biosure® Controls, Grass Valley, CA 95945) were used as internal size standard, whereas sunflower (*Helianthus annuus* L. accession RHA 801) and barley (*Hordeum vulgare* cultivar Baronesse) leaf tissue were used as positive controls to verify the procedures followed and see if the genome size calculated by our method was similar to the previously reported genome sizes (Arumuganathan and Earle 1991b).

Fresh leaf, root, and stem tissues were collected from greenhouse grown plants of OMF40-11. OMF40-11 is a *Limnanthes alba* ssp. *alba* inbred line, developed by single seed descent (without artificial selection) from the open pollinated *L.alba* ssp. *alba* cultivar Mermaid. Tissues were placed on petri dishes

with 300µl of solution A added from nuclei isolation (High resolution DNA kit, type P: Partec GMbH, Munster, Germany). The tissue was chopped using a scalpel and remained in the solution for 5 min. Suspensions were filtered through a 30µm nylon mesh and transferred to a sample tube. Solution B (2ml) was then added for DAPI staining (High resolution DNA kit, type P: Partec GMbH, Munster, Germany) for 5 min. All the samples were analyzed with a Partec PA flow cytometer.

In the second method we stained suspensions of intact nuclei prepared by chopping pieces of young leaf tissues of meadowfoam in $MgSO_4$ buffer with propidium iodide and treating with DNAase-free RNAase (Arumuganathan and Earle, 1991a). We used an EPICS PROFILE flowcytometer (Coulter Electronics, Hialeah, FL, USA) with an argonion laser operating at a wavelength of 488 nm. Chicken red blood cells (CEN singlet, product # 1013 from Biosure® Controls, Grass Valley, CA 95945) were used as internal standards without positive controls.

For both the methods, we analyzed five replicates of each tissue. Thus in case of the first method we had 15 observations from three different tissues of meadowfoam and five observations each from sunflower and barley. However, in case of the second method we had only five observations from the leaf tissue of meadowfoam. In each sample, 10000 to 25000 nuclei were analyzed at a rate of one nucleus/second. The DNA content of 2C nuclei (those in G0/G1 phase of the cell cycle) was calculated by using nuclei from chicken red blood cells (CRBC) as

internal standards. For the absolute DNA content of CRBC nuclei, we used the value of 2.33pg/2C, as estimated chemically by Galbraith et al. (1983) after extraction of total DNA from a sample of cells. The 2C DNA content of the unknown sample was calculated using:

$$\text{2C DNA content} = \frac{\text{Unknown G1 peak mean} \times \text{Standard 2C DNA content}}{\text{Standard G1 peak mean}}$$

The Average packing ratio (APR) of meadowfoam DNA was calculated using:

$$\text{APR} = \frac{\text{Amount of 2C DNA present in a cell (pg)} \times 098 \times 10^{9} \times 3.4}{\text{Length of the total chromosomes at metaphase } (\mu\text{m}) \times 10^{4}}$$

The amount of DNA in pg was converted into base pairs by multiplying with 0.98 x 10^9, the distance between 2 bases is 3.4 A° and 1 μm = 10^4 A°, thus we can calculate the APR of DNA for any species.

Preparation of chromosomes

Mitotic chromosomes

Growing root tips of OMF40-11 were cut into 1 cm long pieces and treated with 0.05% (w/v) colchicine for 3 hrs in the dark. The treated samples were transferred into a fixative (3ethanol: 1acetic acid) plus staining solution (2% aceto–orcein). The root tips were left in this solution for 24 hrs at room temperature and then stored in the refrigerator (4°C) for later analysis. Squashes were prepared by placing in a petri dish with 45% acetic acid for 10 min. The samples were heated

for a few seconds. Each root tip was placed on a glass slide in a drop of 45% acetic acid. The root cap was removed using a sharp scalpel. A blunt instrument was used to apply pressure on the root tip in such a way that only the meristematic tissue would be released on the slide without breaking the whole root tip. The cells spread on the slide were covered with a coverslip and observed under a light microscope (Zeiss Axioscope 2).

Meiotic chromosomes

For observing meiotic events in pollen mother cells (PMCs), we collected floral buds of OMF40-11, OMF64, a self-pollinated inbred line, developed from wild *L.alba* ssp. *versicolor* population PI374801 (Knapp and Crane 1997; Crane and Knapp 2000) and the hybrid between the two inbred lines (OMF40-11 x OMF64). Floral buds varied in size from 2 to 4.5 mm with an increment of 0.25mm. They were fixed in an ethanol: acetic acid fixative (3:1). Buds ranging from 3 to 3.5 mm always had highest number of cells in metaphase (I) and anaphase (I). The buds were dissected and the anthers were placed on the slide with a drop of 2% acetocarmine and squashed. The cells were spread on the slide and a coverslip was placed. Then the slide was warmed up and turned upside down on a blotting paper and pressed behind the coverslip so as to flatten the cells. The slides thus prepared were observed under a light microscope (Zeiss axioscope 2). Photographs were taken using a CCD camera (CoolSNAP ^{TM}, Roper Scientific,

Tucson, AZ, USA) and processed with RSImage ™ software version 1.07 (Roper Scientific, Tucson, AZ, USA).

Analysis of chromosomes

For karyotype analysis, five cells from seven different plants with well-spread and similarly condensed chromosomes were identified. An image of each cell was captured by a CCD camera (CoolSNAP ™, Roper Scientific, Tucson, AZ, USA) and processed with RS Image ™ software version 1.07 (Roper Scientific, Tucson, AZ, USA). Chromosome measurements were made using digital vernier calipers on enlarged prints and converted to micrometers by relating measurements from enlarged prints with measurements made in a microscope with a micrometer. The chromosomes were identified on the basis of their total length, arm length ratio (long/short arm), and presence of satellites. Chromosomes in the karyotype were arranged in groups of two according to homology, following decreasing mean chromosome length.

To detect abnormalities in the meiosis of the hybrid (OMF40-11 x OMF 64), 1000 cells in anaphase I from three different hybrid plants were observed. The number of bridges observed in anaphase I was counted. Similarly the number of bridges observed in 1000 cells of anaphase I of OMF40-11 and OMF64 were also counted.

Pollen viability tests

A bulk pollen sample was collected from the parents (OMF40-11 and OMF64) and three hybrids (OMF40-11 x OMF64) at anthesis. Pollen grains were stained with I_2KI (Jensen 1989) and scored for stainability under a light microscope. The percentage of stainable pollen grains was estimated from 1,000 pollen grains.

Genetic mapping

Mapping population:

The genetic map was constructed using 96 (OMF40-11 x OMF64) xOMF64 backcross progeny, where OMF40-11 and OMF64 are S_5 lines. Emasculating and hand pollinating OMF40-11 and OMF40-11 x OMF64, we produced hybrid and backcross progeny, respectively. Seeds of the parents, hybrids and backcross progeny were produced in the greenhouse-grown plants. Whole seeds of the inbred lines, hybrid and 96 BC_1 progeny were germinated at 4°C in the dark on moistened blotter paper in 11x11x3 cm clear plastic boxes. Germinants were transplanted to potting soil (pumice: peat moss: sandy loam) in 7.5cm^2 plastic pots. The plants were grown at 15°C for 25 to 28 days in a growth chamber (Model CEL 37-14, Sherer-Gillett CO., Marshall, Mich.) with 8 h of fluorescent light per day. Four-week-old plants were transferred to a greenhouse and grown to maturity at 18°C (night) to 25°C (day) with 16 h of light per day. The bases of the pots and

protruding roots were continuously submerged in 1-3 cm of water in the growth chamber and greenhouse. Leaves from 50 to 55 day old plants were harvested, immediately frozen, and stored at –80°C.

Simple sequence repeat (SSR) marker development and screening:

Genomic DNA from frozen leaf tissue of OMF40-11, OMF64 and 96 backcross progeny was extracted using a protocol similar to that of Lodhi et al. (1994). One to two grams of leaf tissue were ground in liquid nitrogen and incubated with 2% cetyltrimetylammonium bromide (CTAB) extraction buffer for 1 hr at 65°C. The DNA was chloroform extracted once. The aqueous phase was mixed with a half volume of 5M NaCl, precipitated with two volumes of cold 95% ethanol, and refrigerated at 4°C overnight. The DNA pellets were dissolved in TE (10mM Tris HCL and 0.1 mM EDTA, pH 8.0) buffer. The dissolved DNA samples were treated with RNase (100 μg/ml) for 1 hour at 37°C.

We screened 389 SSR markers (Kishore *et al.* 2002) against genomic DNA of meadowfoam-inbred lines OMF40-11 and OMF64 to identify polymorphic SSRs. Polymorphic SSRs identified were genotyped on 96 backcross progeny, on an ABI Prism 377 DNA sequencer (Applied Biosystem, Perkin Elmer, Foster City, CA, USA) using polyacrylamide gels and fluorescently labelled amplicons as described by Kishore *et al.* (2002). Filter set C and the GeneScan 500 TAMRA internal standard were used for assays performed with 6FAM, TET and HEX

labelled amplicons. The amplicons for each SSR markers were separately produced, pooled post-PCR, and loaded into each lane; each amplicon in each pool lane was labelled with a different fluorophore (e.g. 6FAM, HEX and TET). Three-color multiplexes were used so that the DNA fragments produced by each SSR primer pair could be unambiguously identified. Allele lengths were recorded using GeneScan ver.2.1 and Genotyper ver.2.0 software (Applied Biosystem, Perkin Elmer, Foster City, CA, USA) and were manually checked. The individual loci amplified by multilocus SSR markers were labelled with consecutive letters (e.g. A, B, and C).

Statistical Analysis:

Genetic maps were constructed using MAPMAKER (Lander et al. 1987) and G-Mendel (Holloway and Knapp 1993). Log-likelihood ratio (G) tests for segregation distortion were performed for each locus. The observed ratio was significantly different from the expected ratio (1:1) when $G > \chi^2_{(1, 0.01)}$, where G is a log-likelihood ratio test statistic and $\chi^2_{(1, 0.01)}$ is a random variable from the χ^2 distribution with one degree of freedom. Loci were grouped using likelihood odds (LOD) threshold of 5.0 and a recombination frequency threshold of 0.25. Loci were ordered using the MAP function of MAPMAKER and the ORDER function of G-MENDEL. MAP estimates locus orders by comparing multipoint likelihoods, whereas ORDER estimates locus orders by comparing map lengths (sums of

adjacent recombination frequencies). Multipoint likelihood was used to select the final locus order estimate for each linkage group. If the likelihood for locus order produced by MAPMAKER was greater than the likelihood of the locus order produced by G-MENDEL, then the order produced by MAPMAKER was selected. Similarly, if the likelihood locus order produced by G-MENDEL was greater than the likelihood of the locus order produced by MAPMAKER, then the order produced by G-Mendel was selected.

Genome length and map coverage

The average marker spacing (s) was estimated by dividing the summed length of all linkage groups by the number of intervals (the number of markers minus the number of linkage groups). Genome length (L) was estimated by adding 2 s to the length of each linkage group, thereby accounting for chromosome ends beyond terminal markers (Fishman et al. 2001) and by multiplying the length of each linkage group by $(m+1)/(m-1)$ and adding the observed length of the linkage group, where m is the number of markers on each group (Chakravarti et al. 1991). The proportion of the genome within d cM of a marker, assuming a random distribution of markers, was estimated by $c=1-e^{-2dn/L}$, where L is the genome length estimate and n is the number of markers.

Results

Physical genome size

Using DAPI the 2C DNA content (pg) of meadowfoam leaves, stems, and roots were estimated to be 5.53±0.05, 5.49±0.04, and 5.33±0.00, respectively. The differences were not significant and the mean 2C content was 5.52 ± 0.08 pg. Using the fluorochrome PI, the mean 2C DNA content of meadowfoam leaves was estimated to be 5.43±0.07 pg. This estimate was not significantly different ($p = 0.61$) from the DAPI based estimate. The sunflower inbred line RHA801 and the barley cultivar Baronesse were used as positive controls and yielded 2C DNA contents (5.90 ± 0.07 and 11.34 ± 0.10 pg, respectively) in agreement with previously published estimates (Arumuganathan and Earle, 1991b). Using the standard ratio of 1 pg/980 Mbp (Bennett and Leitch, 2001), the physical length of meadowfoam genome was estimated to be 5,409.6 ± 78.4 Mbp.

Endopolyploidy was observed in the meadowfoam leaves and roots. DNA contents of leaf cellos ranged from diploid (2C) upto 16-ploid (16 C) (Fig. 1), whereas DNA contents of roots ranged from diploid to octaploid (8C) but were mostly diploid and tetraploid. Of the nuclei sampled in leaves, 23% had 4C DNA, 43% had 8C DNA, and 22% had 16C DNA contents, whereas only 12% had 2C DNA contents (Fig. 1). Stems and roots showed 2C, 4C and occasionally 8C nuclei,

however in the experiments performed it was seen that the nuclei from the roots and stem never exhibited nuclei with 16C DNA.

Meadowfoam chromosomes

We developed a detailed karyotype for five chromosomes in haploid meadowfoam genome (Fig. 2) using image analysis (Table 2 and Fig. 3). Chromosome arm length measurements were produced from replicate cells showing similar degrees of chromosome condensation. The chromosomes of meadowfoam were arranged in order of decreasing length, the longest being I (6.23 ± 0.14 μm) and the shortest being V (8.44 ± 0.014 μm) (Fig. 3). The chromosomes of meadowfoam were estimated to be 73.18 μm long; thus the average packing ratio for meadowfoam genome was estimated to be 24749. Chromosomes I and IV had nucleolar organizing regions (NORs). The satellite on the latter was shorter than the former. Chromosome I and III were metacentric (arm ratio 1.04 and 1.08, respectively), whereas chromosome II, IV and V were sub-metacentric (arm ratio 1.53, 1.48, and 1.50, respectively). Because of the small number of chromosomes ($2n+ 2x = 10$) and unique chromosome morphology and length combinations, the individual chromosomes could be easily distinguished. The two metacentric chromosomes (I and III) can be distinguished by the presence or absence of the satellite. Chromosome IV can be distinguished from the other two submetacentric

chromosomes (II and V) by the presence of a satellite, whereas chromosomes II and V can be distinguished by a 1.47 μm difference in length.

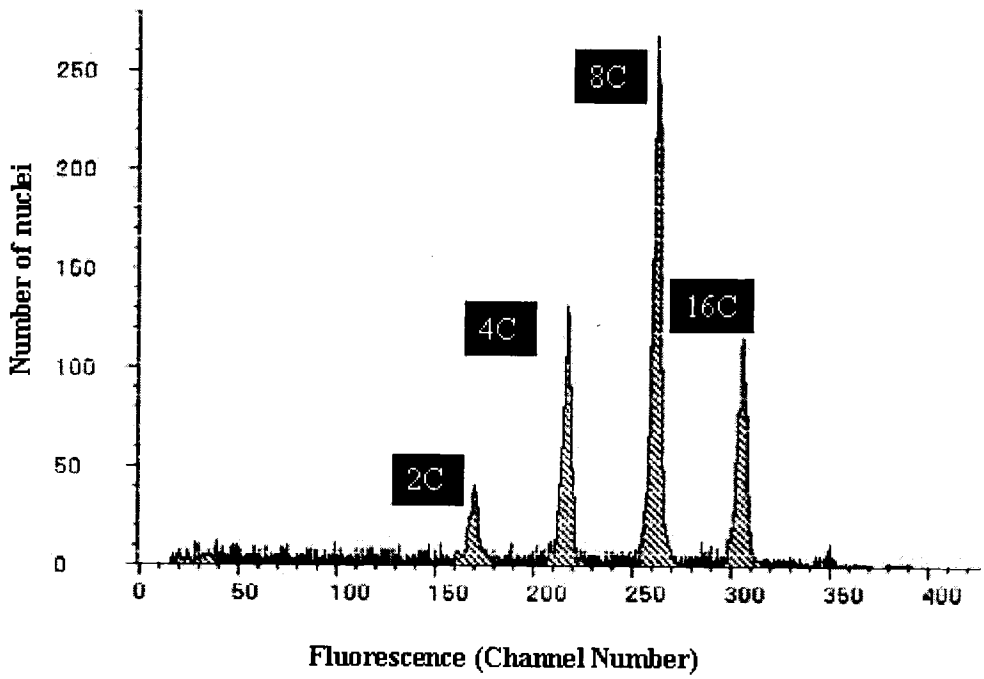


Figure 3.1. Histograms of intact nuclei of meadowfoam leaf tissue displaying endopolyploidy.

Table 3.1. Mean 2C DNA content of meadowfoam, sunflower, and barley.

Species	Stain	Flow Cytometer	Tissue	Number of Replicates	Mean 2C DNA Content (pg)	Standard error (pg)	Coefficient of variation (%)
Meadowfoam	DAPI	PA-Partec	Leaf	5	5.53	0.05	2.04
			Stem	5	5.49	0.04	1.93
			Roots	5	5.33	0.15	6.54
	PI	EPICS PROFILE Coulter	Leaf	5	5.43	0.07	3.02
Sunflower	DAPI	PA-Partec	Leaf	5	5.90	0.07	2.76
Barley	DAPI	PA-Partec	Leaf	5	11.34	0.1	2.12

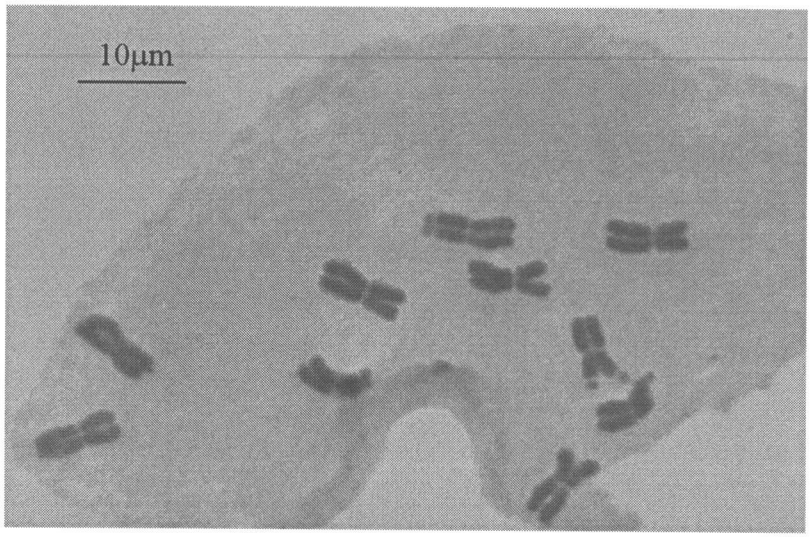


Figure 3.2. Orcein stained mitotic metaphase chromosomes of *Limnanthes alba* ssp. *alba*.

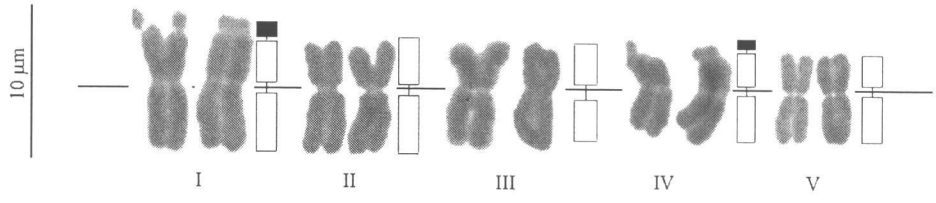


Figure 3.3. Karyotype and ideogram of *Limnanthes alba* ssp. *alba* based on the cell presented in Fig 3.2.

Table 3.2. The chromosomes of the *Limnanthes alba* spp. *alba*.

Chromosome number	Long arm mean ± SE	Short arm mean ± SE	Satellite size mean ± SE	Total Length mean ± SE	Arm ratio mean ± SE	Type
	μm					
I	4.29±0.09	2.89±0.07	1.26±0.09	8.44±0.14	1.04±0.02	metacentric
II	4.67±0.09	3.03±0.07		7.70±0.13	1.53±0.08	sub-metacentric
III	3.77±0.07	3.48±0.09		7.26±0.09	1.08±0.04	metacentric
IV	4.14±0.13	2.07±0.09	0.74±0.00	6.96±0.21	1.48±0.03	sub-metacentric
V	3.70±0.11	2.52±0.14		6.23±0.14	1.50±0.11	sub-metacentric

SSR map of meadowfoam

Of the 389 SSR markers developed for meadowfoam (Kishore et al. 2002), 218 were polymorphic between OMF64 and OMF40-11 and 67% amplified two or three loci. Most of the multiple loci were either monomorphic or were null in the donor parent (OMF40-11) and thus did not segregate in the backcross population. Three multilocus SSR marker loci (ORM116, ORM372 and ORM480) were mapped. Of the polymorphic SSR markers, 90 markers segregated in the backcross population.

Of the 90 segregating SSR marker loci, 84 coalesced into five linkage groups (Fig. 4). The other six SSR marker loci (ORM054, ORM179, ORM234, ORM268, ORM279 and ORM298) were unlinked. Seven dominant and 77 codominant SSR marker loci were mapped. The map was 988.7 cM long and had a mean density of 11.8 cM per SSR marker locus. The length of the linkage groups ranged from 128.6 cM (LG5) to 241.9 cM (LG4). The distribution of the loci across the genome was nearly uniform with minimal clustering on two linkage groups (LG1 and LG2). The number of loci per linkage group ranged from 11 on LG3 to 22 on LG1 with a mean of 16.8 markers per linkage group and the marker density ranged from 10.2 cM per locus on LG1 to 16.5 cM on LG3. The widest gap on the map and the only gap longer than 30 cM was between ORM048 and ORM439 (37.5 cM) on LG3. The longest gaps on the other four linkage groups ranged from 19.8 cM to 28.3 cM.

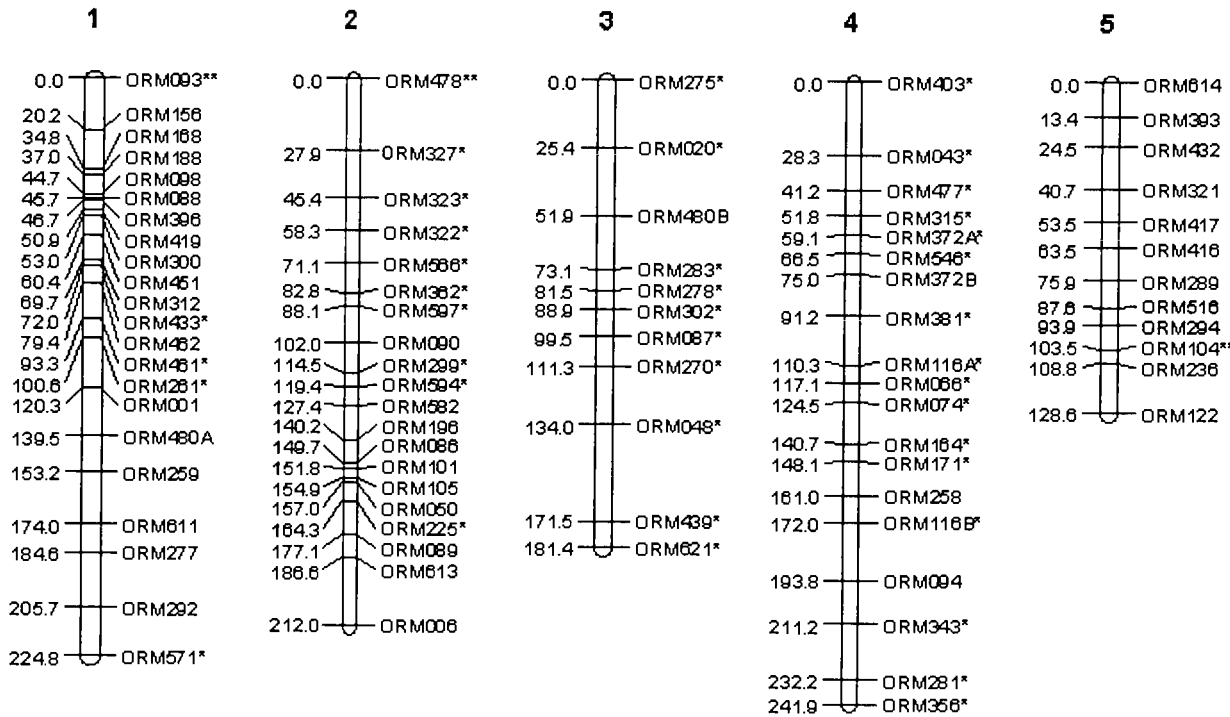


Figure 3.4. Linkage map of meadowfoam comprised of 84 simple sequence repeat loci.
* = Significantly ($p < 0.01$) distorted loci showing an excess of OMF40-11 alleles.
** = Significantly ($p < 0.01$) distorted loci showing excess of OMF64 alleles.

Genome length and coverage

While the SSR marker loci coalesced into the expected number of linkage groups based on chromosome number , six were unlinked and, theoretically, were either Type II errors or are located ~27.5 cM distal to terminal SSR marker loci on the genetic linkage map. Because the six unlinked SSR marker loci were singletons, we predicted that a minimum of ~165.0 cM were missing and the genetic linkage map of meadowfoam is 1,153.7 cM long or longer. Similarly, when 12.7 cM were added to each ends of each linkage group to account for chromosome ends, as proposed by Fishman et al. (2001), the length of the genetic linkage map was predicted to be 1,113.7 cM. Based on these genome length estimates, we estimated that the present map spans 85.7 to 88.8 % of the meadowfoam genome. Using an estimated genome length of 1,118.7 cM, we estimated that 78% of the genome is within 10cM and 95 % of the genome is within 20 cM of an SSR marker locus.

Transmission ratio distortion

The segregation ratios for 42 SSR loci, 47 % of the total were significantly distorted ($p \leq 0.01$). Three SSR marker loci had an excess of *L. alba* ssp. *versicolor* (OMF64) alleles, whereas 39 SSR marker loci had an excess of *L. alba* ssp. *alba* (OMF40-11) alleles. The former were terminal singletons on LG1, 2 and 5, whereas the latter were present in clusters on LG1, 2, 3 and 4 (Fig. 5).

Systematic patterns of segregation distortion along chromosomes signal gametic or other selection effects and are a norm in wide crosses (Zamir and Tadmor 1986, Paterson et al. 1991, Bernacchi and Tanksley 1997, Fishman et al. 2001). Because recombination frequency estimates produced from progeny with distorted segregation ratios are unbiased (Bailey, 1960), the distorted genetic markers should not have adversely affected locus grouping or ordering in the meadowfoam map. Eight SSR loci were polymorphic in the parents but monomorphic in the progeny. The genotypes of the parents and progeny were rechecked and replicated. We presume selection eliminated the *L. alba* ssp. *alba* genome as all these eight markers always showed the *L. alba* ssp. *versicolor* allele in the progeny.

The eight monomorphic SSR loci in the progeny prompted us to perform analyses of pollen and meiosis in pollen mother cells (PMCs) of the parents and the hybrid. The pollen viabilities of the parents (OMF40-11 and OMF64) were 90 and 93 % respectively, whereas the pollen viability of the hybrid was 70%. We observed anaphase I bridges in the hybrid but not the parents (Fig. 7). Bridges were present in 256 out of the 1,000 hybrid PMCs sampled; however, none showed more than two bridges per cell. Based on a transmission ration skewed towards OMF40-11, 30% pollen sterility in the hybrid and presence of anaphase bridges in 26% of the PMCs of the hybrid, we speculate that a biological mechanism underlies the systematic transmission of the alleles of *L. alba* ssp. *alba* favored over *L. alba* ssp. *versicolor* alleles.

Table 3.3. Summary of the meadowfoam simple sequence repeat map, estimated genome length and coverage.

Linkage group	Number of markers	Number of Loci	Distance between loci			Length of linkage group	Linkage group length[§]		Genome coverage[†]
			Min (cM)	Mean (cM)	Max (cM)	(cM)	Method1[Ψ]	Method2[ϕ]	
1	22	22	1.0	10.2	21.1	224.8	249.8	246.2	83
2	20	20	2.1	10.6	27.9	212.0	237.0	234.3	82
3	10	11	7.4	16.5	37.5	181.4	206.4	217.7	64
4	17	19	6.8	12.7	28.3	241.9	266.9	268.8	76
5	12	12	5.3	10.7	19.8	128.6	153.6	151.9	78
Total	81	84	1.0	11.8	37.5	988.7	1113.7	1118.9	78

§= An individual estimate of each chromosome length estimated by different methods
Method1[Ψ] = Adding 2*s* to the length of linkage group
Method2[ϕ] = Using method 4 of Chakravarti et al. (1991)
†= An individual estimate of genome covered (*c*) within 10 cM of a linked marker on each chromosome

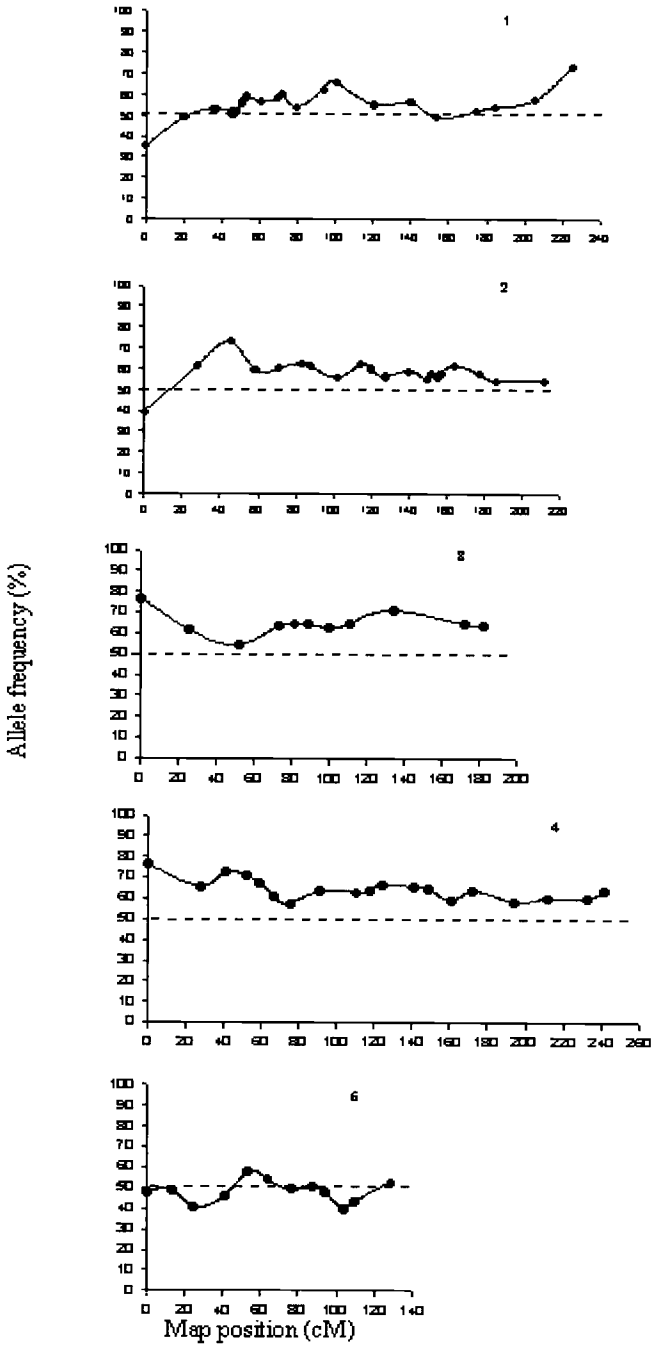


Figure 3.5. OMF40-11 (*Limnanthes alba* ssp. *alba*) simple sequence repeat loci allele percentages among (OMF40-11 x OMF64) x OMF64 BC1 progeny for ordered loci on the linkage map of meadowfoam.

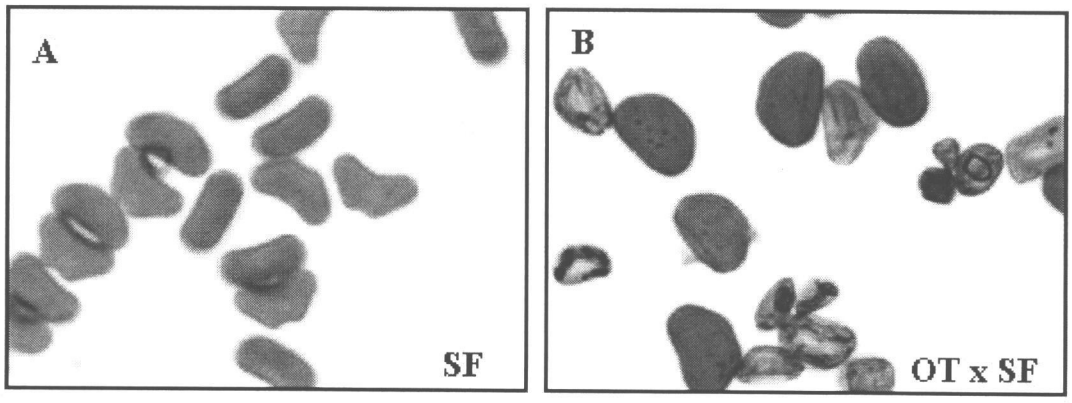


Figure 3.6. A - Pollen of OMF64 stained with I_2KI at a magnification of 100x, and B - Pollen of OMF40-11 x OMF64 stained with I_2KI at a magnification of 100x. SF= self-fertlizing parent, OtxSF= hybrid of outcrossing and self-fertilizing parents.

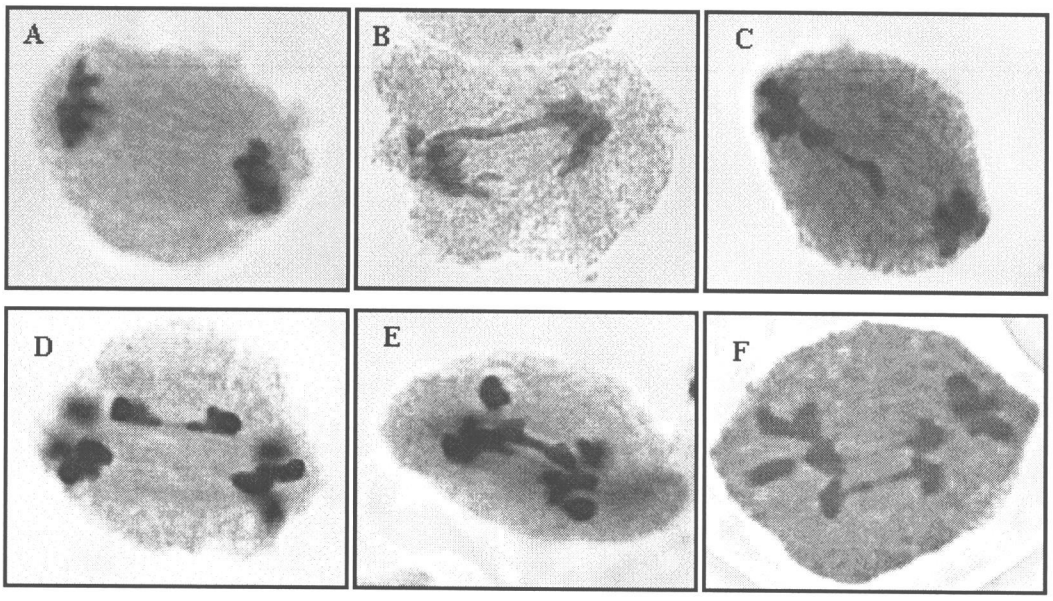


Figure 3.7. A- Anaphase I observed in the PMCs of OMF40-11 and OMF64, B, C, D – PMCs of OMF40-11 x OMF64 showing one bridge at anaphase I, E and F- PMCs of OMF40-11 x OMF64 showing two bridges at anaphase I

Discussion

Crop plant research is poised to make revolutionary strides in the era of genomics. The challenge is to develop a complete set of comprehensive and systematic resources to facilitate the research endeavors. A number of genomic resources have been developed for the model plant organism *Arabidopsis thaliana*, (Arabidopsis Genome Initiative 2000) and economically important crops like maize (Coe et al. 2002), soybean (Cregan et al.1999), rice (Harushima et al. 1998) and sunflower (Tang et al. 2002). However, it is not possible to develop such resources for every crop, especially new crops. The most efficient way to work with new crops would be to develop framework resources and then identify the closest relative for which sufficient genomic resources have been developed, and utilize comparative genomics to enhance the genetic knowledge about new crops. In this study we provide some of the basic information about the genome of meadowfoam. We also report the development of the first SSR marker based framework linkage map of meadowfoam. An attempt has also been made to identify the closest relative of meadowfoam (based on the information generated in this study and also from studies by other scientists) for whom extensive genomic resources have been developed.

Flow cytometry of nuclear DNA content has become an important technique for detection of genome sizes of various organisms (Arumuganathan and Earle 1991a). In many plant studies leaves (Arumuganathan and Earle 1991b) are

commonly used for flow cytometry. In this study we used three different tissues (leaves, stems and roots), two different fluorochromes (DAPI and PI) and two different flow cytometers to determine the 2C DNA content of meadowfoam. The two fluorochromes have different stain reactions. PI intercalates between base pairs of double stranded DNA with little or no base specificity (Properi et al. 1991), while DAPI is a nonintercalating stain that preferentially binds in a complex manner to A-T base regions (Godelle et al. 1993). As there were no significant differences observed in the 2C DNA content of meadowfoam estimated using DAPI and PI, it can concluded that the meadowfoam genome has equal amount of AT and GC base pairs unlike *Allium cepa* whose 2C DNA content estimated by DAPI was twice the 2C DNA content estimated by PI (Dolezel et al. 1992). The physical size (2C DNA) of the meadowfoam genome is 16 times larger than *Arabidopsis thaliana*, a model plant with the same basic chromosome number as meadowfoam.

Chromosomal studies on the genus *Limnanthes* have been limited to determination of chromosome number and the number of satellite chromosomes (Propach 1934; Resende 1937; Mason 1952). Detailed karyotype analysis presented in this study helped to identify the chromosomes of meadowfoam unambiguously. The two satellite pairs of chromosomes (I and IV) can be easily identified by the difference in the position of the centromere. Chromosome III is the only metacentric chromosome without any satellite and hence can be easily

distinguished from the other chromosomes. The difference in the average lengths of chromosome II and V is very large (1.5μm). Since meadowfoam chromosomes are morphologically distinct there is no need to use any special techniques like C-banding to identify them. The total diploid genome length of meadowfoam was five times larger than the reported diploid genome length of *Arabidopsis thaliana* at metaphase (Fransz et al. 1998). However the APR of meadowfoam DNA was only 3.23 fold greater than *Arabidopsis thaliana* (7656). Arabidopsis and meadowfoam both have two pairs of chromosomes with NOR's. The karyotype analysis of Arabidopsis by Fransz et al. (1998) mentions that the two longest chromosomes (1 and 5) are metacentric (m) and the two shortest chromosomes (2 and 4) are acrocentric (a) and carry NOR's subterminally in their short arms, while chromosome 3 is submetacentric (sm) and medium sized. The genome sequence of Arabidopsis has revealed that there is extensive duplication and reshuffling within and between the chromosomes of Arabidopsis (Blanc et al. 2000). Hence we speculate that Arabidopsis and meadowfoam came from a single ancestor with ten acrocentric chromosomes, and by the process of different cycles of Robertsonian fission and fusion, the karyotypes of Arabidopsis (2m +2a +1sm) and meadowfoam (2m+3sm) evolved. However both retained their subterminal NOR's on their chromosomes. The karyotype evolution of *Gibasis matudae* and *Gibasis pulchella* was explained by Jones (1998) in a similar way and Cerbah et al. (1998) provided a similar explanation for the evolution of *Hypochaeris maculata* and *H. glabra*.

Lagercrantz (1998) concludes that the rate of chromosome rearrangements and evolution in Brassicaceae is the highest reported so far among the higher plants. The key step in karyotype evolution is the fixation of newly arisen chromosomal rearrangements. Translocations and inversions are generally deleterious when heterozygous, but behave normally when homozygous (White 1973). The fixation of such rearrangements requires small, isolated populations and is aided greatly by self-fertilization (Lande 1979). Many wild Brassicaceae and Limnanthaceae species occupy marginal fragmented habitats such as maritime cliffs and are also found in vernal pools (Snogerup et al. 1990, Mithen et al. 1995, Mason 1952). This geographic distribution may support the higher rates of evolution of chromosomal rearrangements. However, there is little experimental support for these hypotheses. It is also not known if chromosome morphology provides a selective advantage within a particular lineage (De Villena 2001). However, with the advent of new molecular cytogenetic techniques, it would be valuable to perform a comparative karyotypic analysis on the chromosomes of meadowfoam and Arabidopsis using homologous probes to reveal more information about the evolution of their karyotypes.

Our results also demonstrated that in meadowfoam most of the somatic cells, especially in the leaf tissue, underwent several rounds of endoreduplication, resulting in cells with multiple ploidy levels. Systemic endopolyploidy has been reported in *A. thaliana* (Galbraith et al. 1991), cucumber (Gilissen et al. 1993),

tomato (Smulders et al. 1994), *Brassica rapa* and *Brassica oleracea* (Kudo and Kimura 2001). Since the 2C DNA values of meadowfoam are much larger than all of the plants mentioned above, the suggestion that endopolyploidy is prevalent in plants with small genomes (Nagl 1978, De Rocher et al. 1990) does not seem to be a universal generalization.

Endoreduplication has often been observed in association with cell growth and in plants a strong correlation between cell size and endoreduplication was found (Melaragno et al. 1993). These observations have led to the widely accepted view of that endoreduplication favors sustained cell elongation in the absence of mitosis. However, it could also be important for other reasons. Endoreduplication could present a means for organisms to increase the number of functional gene copies within each cell, thereby acting to mitigate any adverse effects of environmental influences on transcription of the genome. The tissue specific pattern of endopolyploidy suggests that endoreduplication cycles in plants constitute an essential part of the developmental program that are necessary for differentiation and the specialized function of given cells and tissues. Regulation of endopolyploidy may operate at multiple cellular levels. Transformation of the mitotic cycle by inhibition of the G_2/M transition is required, and the number of endoreduplication cycles is probably controlled by components of the cell cycle machinery. Grafi and Larkins (1995) demonstrated that endoreduplication in the development of maize endosperm proceeds as a result of both the inactivation of

M-phase related cyclin-dependent kinases (CDKs) with an inhibitor and the induction of S-phase related CDKs. Jacqumard et al. (1999) described a cell cycle gene, *CKS1At* that was associated with the endoreduplication cycle. Expression of *CKS1A* was present in endoreduplicating tissues in *A. thaliana*.

Meadowfoam belongs to Brassicales, the same taxonomic Order as *Arabidopsis thaliana* (Wheeler et al. 2000). Both have the same chromosome number and also have two pairs of chromosomes with NOR's. Endopolyploidy is a rule in *Limnanthes* and Arabidopsis rather than an exception. A phylogenetic tree constructed using the DNA sequences of the chloroplast *rbcL* gene and the 18S ribosomal RNA gene showed that *Limnanthes* is more closely related to *Arabidopsis* than members of other genera like *Gossypium*, *Bombax* and 12 other genera belonging to seven different orders (Rodman et al. 1998). Within Brassicales, *Limnanthes* was ranked as the eleventh closest member to Arabidopsis out of 18 species sampled.

Linkage maps provide a genetic framework for identifying quantitative trait loci (QTL) and analyzing genome structure. Here we present the first linkage map of meadowfoam based on microsatellites of a cross between *Limnanthes alba* ssp. *alba* and *Limnanthes alba* ssp. *versicolor*, a pair of closely related subspecies with widely divergent floral morphologies and mating systems. To generate a heterosubspecific linkage map suitable for QTL analysis and marker-assisted introgression of specific genomic regions, we took a thorough approach to

framework map construction. We used a stringent LOD score (5.0) to construct this framework meadowfoam genetic map that increased the efficiency of detecting true linkage and minimized false positives in assigning markers to linkage groups (Type I errors). The locus orders produced by MAPMAKER and G-MENDEL were nearly identical and only varied locally. The markers were fairly evenly distributed among the five linkage groups, in contrast to other SSR maps in several crops such as barley, wheat and sunflower (Ramsay et al.2000, Roder et al. 1998, Tang et al.2002, Yu 2001) where the markers tend to cluster in some regions, particularly around centromeric regions.

Several lines of evidence suggest that the framework linkage map provides thorough coverage of the *Limnanthes alba* ssp. *alba* x *Limnanthes alba* ssp. *versicolor* genome. The five linkage groups, which range in size from 128.6 to 241.9 cM and each contain at least 11 loci, presumably corresponding to the five chromosomes found in *Limnanthes alba*. The SSR linkage map (84 SSR loci) in this study was 290 cM longer than the AFLP linkage map described by Katengam et al. (2002). Estimates of the genome length based on the map length and distribution of the markers suggest that the map spans 88-89% of the genome and that 95% of the genome is within 20 cM of a linked marker.

Distorted segregation ratios are common in crosses between genetically divergent genomes (Zamir and Tadmor 1986, Paterson et al. 1991, Bernacchi and Tanksley 1997, Fishman et al. 2001). The distorted markers cover two complete

linkage groups (LG3 and LG4) and specific regions on LG1 and LG2. The distortion was unidirectional (93 % of the distorted loci exhibit an excess of *Limnanthes alba* ssp. *alba* alleles). This pattern suggests that biological mechanisms, rather than the chance or error underlie most of the observed transmission ratio distortion. Whatever the underlying process, this pattern may circumscribe the genetic composition of advanced generation hybrids in the lab or wild by favoring rapid fixation of *Limnanthes alba* ssp. *alba* alleles in some genomic regions and retarding introgression in others (Rieseberg et al. 1995, 2000).

Reasons for skewed segregation ratios include genetic factors operating in pre- and post- zygotic phases of reproduction (Zamir and Tadmor 1986), structural rearrangements (Stebbins 1950, Williams et al. 1990), or gametic selection (Zamir et al. 1982). The high level of segregation distortion obtained in this study indicates *L. alba* ssp. *alba* and *L. alba* ssp. *versicolor* have genomes that are genetically divergent and was expressed as a postmating mechanism. Post-mating mechanisms arise between species as a byproduct of evolutionary divergence, typically after the evolution of a pre-mating barrier (Levin 1978, Grant 1981). Cuckoo chromosomes have been identified in wheat whereby hetero- or hemizygous condition kill gametes lacking them (Mann 1975, Endo and Katayama 1978, Miller et al. 1982). This ensures that all gametes lacking the cuckoo genes are non-functional and that gametes containing genes are preferentially transmitted. Autosomal meiotic drive, in which a killer allele eliminates gametes carrying alternative alleles, can cause

severe segregation distortion. Such drive loci have been well characterized in Drosophila, mouse, and Neurospora (reviewed in Lyttle 1991). However, meiotic drive is an unlikely explanation for the widespread and generally moderate bias in genotype frequencies in our BC_1 progeny, because such a large number of independently segregating drive loci would render the heterozygous F_1 progeny partially or fully sterile.

Clearly, no single genetic mechanism can account for partial sterility (30% non-viable pollen), anaphase bridges in F_1 hybrid, and the patterns of gamete transmission absorbed in the BC_1 progeny. Our data on these post-mating barriers to hybridization and introgression not only provide new insight into genetic divergence in this rapidly evolving system, but also raise further questions about the underlying mechanisms. Extension of the results from this study to interpretation of an intersubspecific difference must be viewed with some caution. This report is based on a single cross, in one direction and under one environment. Different results may be obtained with additional crosses or under different conditions. The greenhouse environment in which the current cross was performed and plants raised is very moderate compared to field conditions. If selection for particular gametic combinations produced the skewed segregation ratios seen in this cross, then we can expect more severe selection under field conditions.

This study will provide a sound basis to carry out molecular breeding of this species. The SSR linkage map constructed will also be a useful framework work

map to add additional markers common to *Arabidopsis thaliana* and meadowfoam. This will help in understanding the comparative genome organization of both these species. The knowledge of linkage relationships between loci also provides a base line for selecting markers spanning the genome for use in population studies. Mapped loci at different levels of linkage are being used to develop new statistical approaches to assess the demographic history of a species. For example, the pattern of variation across unlinked microsatellite loci has been used to test whether the population size is constant or increasing (Goldstein et al. 1999). These methods would be particularly beneficial for some of the species of *Limnanthes* that have been listed as endangered (Dole and Sun 1992). The framework map of meadowfoam we present here will also be an invaluable guide for QTL analyses of phenotypic differences associated with mating system evolution.

References

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature, 408:796-812.

Arroyo MTK (1975) Electrophoretic studies of genetic variation in natural populations of allogamous *Limnanthes alba* and autogamous *Limnanthes floccosa* (Limnanthaceae). Heredity, 35:153-164.

Arumuganathan K, Earle ED (1991a) Estimation of nuclear DNA content of plants by flow cytometry. Plant Mol. Biol. Rep., 9:221-231.

Arumuganathan K, Earle ED (1991b) Nuclear DNA content of some important plant species. Plant Mol. Biol. Rep., 9:208-218.

Bailey NTJ (1961) Introduction to the mathematical theory of genetic linkage. Clarendon Press, Oxford.

Bennett MD, Leitch IJ (2001) Plant DNA C-values database (release 1.0, Sept. 2001) http://www.rbgkew.org.uk/cval/homepage.html

Bernacchi D, Tanksley SD (1997) An interspecific backcross of Lycopersicon esculentum x L. hirsutum: linkage analysis and a QTL study of sexual compatibility factors and floral traits. Genetics, 147:861-877.

Blanc G, Barakat A, Guyot R, Cooke R, Delseny M (2000) Extensive duplication and reshuffling in the Arabidopsis genome. Plant Cell, 12:1093-1101.

Brown CR, Jain SK (1979) Reproductive system and pattern of genetic variation in two Limnanthes species. Theor. Appl. Genet., 54:181-190.

Calhoun W, Crane JM (1975) Registration of "Foamore" meadowfoam. Oregon Agri. Exp. Sta. Corvallis, OR, USA.

Cerbah M, Coulaud J, Siljak-Yakovlev (1998) rDNA organization and evolutionary relationships in the genus Hypochaeris (Asteraceae). J. Heredity, 89:312-318.

Chakravarti A, Lasher LK, Reefer JE (1991) A maximum likelihood method for estimating genome length using genetic linkage data. Genetics, 128:183-193.

Coe E, Cone K, McMullen M, Chen S, Davis G, Gardiner J, Liscum E, Polacco M, Paterson A, Sanchez-Villeda H, Soderlund C, Wing R (2002) Plant Physiol., 128:9-12.

Crane JM, Knapp SJ (2000) Registration of the self-pollinated meadowfoam germplasm line OMF64. Crop Sci., 40:300-301.

Cregan PB, Jarvik T, Bush AL, Shoemaker RC, Lark KG, Kahler AL, Kaya N, VanToai TT, Lohnes DG, Chung J, Specht JE. (1999). An integrated genetic linkage map of the soybean genome. Crop Sci., 39:1464-1490.

De Rocher EJ, Harkins KR, Galbraith DW, Bohnert HJ (1990) Developmentally regulated systemic endopolyploidy in succulents with small genomes. Science, 250:99-101.

De Villena FP, Sapienza C (2001) Female meiosis drives karyotype evolution in mammals. Genetics, 159:1179-1189.

Dole JA, Sun M (1992) Field and genetic survey of the endangered Butte County meadowfoam-*Limnanthes floccossa* subsp. *californica* (Limnanthaceae). Conservation Biology, 6:549-558.

Dolezel J, Sgorbati S, Lucretti S (1992) Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. Physiological Plantarum 85:625-631.

Endo TR, Katayama Y 91978) Finding of a selectively retained chromosome of *Aegilops Caudata* L. in common wheat. Wheat Inf. Serv. 47-48:32-35

Fishman L, Kelly AJ, Morgan E, Willis JH (2001) A genetic map in the *Mimulus guttatus* species complex reveals transmission ratio distortion due to heterospecific interactions. Genetics, 159:1701-1716.

Fransz P, Armstrong S, Alonso-Blanco C, Fischer TC, Torres-Ruiz RA, Jones G (1998) Cytogenetics for the model system Arabidopsis thaliana. Plant J., 13:867-876.

Galbraith DW, Harkins KR, Knapp S (1991) Systemic endopolyploidy in *Arabidopsis thaliana*. Plant Physiol., 96:985-989.

Galbraith DW, Harkins KR, Maddox JM, Ayres NM, Sharma DP, Firoozabady E (1983) Rapid flow cytometric analysis of the cell cycle in intact plant tissues. Science., 220:1049-1051.

Gilissen LJW, Van Staveren MJ, Creemers-Molenaar J, Verhoeven HA (1993) Development of polysomaty in seedlings and plants of *Cucumis sativus* L. Plant Sci., 91:171-179.

Godelle B, Cartier D, Marie D, Brown SC, Siljak-Yakovlev (1993) Heterochromatin study demonstrating the non-linearity of fluorometry useful for calculating genomic base composition. Cytometry, 14:618-626.

Goldstein DB, Roemer GW, Smith DA, Reich DE, Bergman A, Wayne RK (1999) The use of microsatellite variation to infer population structure and demographic history in a natural model system. Genetics, 151:797-801.

Grafi G, Larkins BA (1995) Endoreduplication in maize endosperm: involvement of M phase-promoting factor inhibition and induction of S phase-related kinases. Science, 269:1262-1264.

Grant V (19810 Plant Speciation, Ed 2. Columbia University Press, New York.

Harushima Y, Yano M, Shomura A, Sato M, Shimano T, Kuboki Y, Yamamoto T, Lin SY, Antonio BA, Parco A, Kajiya H, Huang N, Yamamoto K, Nagamura Y, Kurata N, Khush GS, Sasaki T (1998) A high-density genetic linkage map with 2275 markers using a single F_2 population. Genetics, 148:479-494.

Holloway JL, and Knapp SJ (1993) Gmendel 3.0 Users Guide. Department of Crop and Soil Science, Oregon State University, OR, USA.

Isbell TA (1997) Development of meadowfoam as industrial crop through novel fatty acid derivatives. Lipid Technology, 9:140-144.

Jacqumard A, De Veylder L, Segers G, de Almedia Engler J, Bernier G, Van Montagu M, Inze D (1999) Expression of CKS1At in Arabidopsis thaliana indicates a role for the protein in both the mitotic and the endoreduplication cycle. Planta, 207:496-504.

Jain SK (1986) Domestication of Limnanthes (Meadowfoam) as a new oil crop. In: Plant domestication induced mutation: Proceedings of an advisory group meeting on the possible use of mutation breeding for rapid domestication of new crop plants (ed. FAO/IAEA.), pp.121-134, Vienna, Austria.

Jensen KB (1989) Cytology, fertility and origin of *Elymus abolinii* (Drob) Tzvelv and its F_1 hybrids with *Pseudoroegneria spicataq*, *E. lanceolatus E. dentatus E. ugamicus* and *E. drobovii* (Poaceae: Triticeae). Genome, 32:468-474.

Jones K (1998) Robertsonian fusion and centric fission in karyotype evolution of Higher plants. Botanical Review. 64:273-289.

Kalin MT (1971) The evolution of autogamy in *Limnanthes* section Inflexae. Ph.D. Dissertation, University of California, Berkeley.

Katengam S, Crane JM Knapp SJ (2002) The development of a genetic map for meadowfoam comprised of amplified fragment length polymorphisms. Theor. Appl. Genet., 104:92-96.

Kesseli RV, Jain SK (1985) Breeding systems and population structure in *Limnanthes*. Theor. Appl. Genet., 71:292-299.

Kishore VK, Velasco P, Slabaugh MB, Knapp SJ (2002) The development of simple sequence repeat markers for DNA finger printing and high throughput genotyping in meadowfoam (in preparation).

Knapp SJ, Crane JM (1997) The development of self-pollinated inbred lines of meadowfoam by direct selection in open pollinated populations. Crop Sci., 37:1770-1775.

Knapp SJ, Crane JM (1999) The development of meadowfoam as an industrial oilseed: breeding challenges and germplasm resources. In: Proceedings of the Fourth national Symposium on New Crops (ed. J.Janick). Wiley, New York, USA.

Kudo N, Kimura Y (2001) Flow cytometric evidence for endopolyploidy in seedlings of some Brassica species. Theor. Appl. Genet. 102:104-110.

Lagercrantz U (1998) Comparative mapping between Arabidopsis thaliana and Brassica nigra indicates that Brassica genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. Genetics, 150:1217-1228.

Lande R (1979) Effective deme size during long-term evolution estimated from rates of chromosomal rearrangements. Evolution, 33: 234-251.

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

Levin DA (1978) The origin of isolating mechanisms in flowering plants. Evol. Biol., 11:185-317.

Lodhi MA, Ye GN, Weeden NF, Resich BI (1994) A simple and efficient method for DNA extraction from grapevine cultivars and Vitis species. Plant. Mol. Biol. Rep., 12:6-13.

Lyttle TW (1991) Segregation distorters. Annu. Rev. Genet., 25:511-557.

Mann SS (1975) Exclusive preferential transmission of an alien chromosome in common wheat. Crop Sci., 16:580-583.

Mason CT (1952) A systematic study of the genus *Limnanthes*. Botanical Publications, 25:455-512.

Melaragno JE, Mehrotra B, Coleman Aw (1993) Relationship between endopolyploidy and cell size in epidermal tissues of Arabidopsis. Plant Cell, 5:1661-1668.

Miller TE, Hutchinson J, Chapman V (1982) Investigation of a preferentially transmitted *Aegilops sharonenesis* chromosome in wheat. Theor. Appl. Genet., 61:27-33.

Mithen R, Raybold AF and Giamoustaris (1995) Divergent selection for secondary metabolites between wild populations of Brassica Oleracea and its implications for plant-herbivore interactions. Heredity, 75:472-484.

Nagl W (1978) Endopolyploidy and polyteny in differentiation and evolution. Towards an understanding of quantitative and qualitative variation of nuclear DNA in ontogeny and phylogeny. Elsevier, North-Holland Publ, Amsterdam.

Paterson AH, Damon S, Hewitt JD, Zamir D (1991) Mendelian factors underlying quantitative traits in tomato: comparison across species, generations and environments. Genetics, 127:181-197.

Powell W, Morgante M, Andre C, Hanafey C, Vogel J, Tingey S, Rafalski A (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol. Breed., 2:225-238.

Propach H (1934) Cytologische untersuchungen an Limnanthes Douglasii. Zeitschr. Zellforsch, 21: 357-375.

Properi E, Giangare MC, Bottiroli G (1991) Nuclease induced DNA structural changes assessed by flow cytometry with the intercalating dye propidium iodide. Cytometry, 12:323-329.

Ramsay L, Macaulay M, Ivanissevich SD, Maclean K, Cardle L, Fuller J, Edwards KJ, Tuvesson S, Morgante M, Massari A, Maetri E, Marmiroli N, Sjakste T, Ganal M, Powell W, Waugh R (2000) A simple sequence repeat-based linkage map of barley. Genetics, 156:1997-2005.

Resende F (1937) Ueber die ubiquitat der sat-chromosomen bei den blutenpflanzen. Planta, 26:757-807.

Rieseberg LH, Baird SJE and Gardner KA (2000) Hybridization, introgression and linkage evolution. Plant Mol. Biol., 42:205-224

Rieseberg LH, Linder CR, Seiler GJ (1995) Chromosomal and genic barriers to introgression in Helianthus. Genomics, 141:1163-1171.

Roder MS Korzum V, Wendehake K, Plaschke J, Tixier MH, Leroy P, Ganal MW (1998) A microsatellite map of wheat. Genomics, 149:2007-2023.

Rodman JE, Soltis PS, Soltis DE, Sytsma KJ, Karol KG (1998) Parallel Evolution of Glucosinolate biosynthesis inferred from congruent nuclear and plastid gene phylogenies. Amer. J. Bot., 85: 997-1006.

Smith MSC, Chin ECL, Shu H, Smith OH, Wall SJ, Senior ML, Mitchell SE, Kresovich S, Ziegle J (1997) An evaluation of utility of SSR loci as molecular markers in maize (*Zea mays* L.) comparisons with data from RFLPs and pedigree. Theor. Appl. Genet., 95:163-173.

Smulders MJM, Rus-Kortekaas W, Gilissen LJM (1994) Development of polysomaty during differentiation in diploid and tetraploid tomato (*Lycopersicon esculentum*) plants. Plant Sci., 97:53-60.

Snogerup S, Gustafsson M, Von Bothmer R (1990) Brassica sect. Brassica (Brassicaceae).I. Taxonomy and variation. Willenowia, 19:271-365.

Stebbins GL (1950) Variation and Evolution in plants. Columbia University Press, New York.

Tang S, Yu JK, Slabaugh MB, Shintani DK, Knapp SJ (2002) Simple sequence repeat map of the sunflower genome. Theor. Appl. Genet., (submitted).

Temnykh S, Park WD, Ayres N, Cartinhour S, Hauck N, Lipovich L, Cho TG, Ishii T, McCouch SR (2000) Mapping and genome organization of microsatellite sequences in rice *Oryza sativa* L. Theor. Appl. Genet., 100:697-712.

Wang Z, Weber JL, Zhong G, Tanksley SD (19940 Survey of plant short tandem DNA repeats. Theor. Appl. Genet., 88:1-6.

Wheeler DL, Chappey C, Lash AE, Leipe DD, Madden TL, Schuler GD, Tatusova TA, Rapp BA (2000). Database resources of National Center for Biotechnology Information. Nucleic Acids Res., 28 (1): 4-10.

White M (1973) Animal Cytology and Evolution. William Clowes and Sons, London.

Williams JG, Kubelik AR, Livak KJ, Rafalski A and Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res., 18:6531-6535

Yu (2001) Simple sequence repeat marker development and mapping in cultivated sunflower, *Helianthus annuus* L. Ph.D. thesis. Oregon State University, Corvallis, OR, USA.

Zamir D, Tadmor Y (1986) Unequal segregation of nuclear genes in plants. Bot. Gaz., 147:355-358.

Zamir D, Tanksley SD, Jones RA (1982). Haploid selection for low temperature tolerance of tomato pollen. Genomics,101:129-137.

CHAPTER 4

MAPPING THE QUANTITATIVE TRAIT LOCI UNDERLYING THE SHIFTS IN THE MATING SYSTEM OF MEADOWFOAM FROM ALLOGAMY TO AUTOGAMY

Venkata K. Kishore[1], Sonali D. Gandhi[1], Jimmie Crane[1] and Steven J.Knapp[1]

[1]Department of Crop and Soil Science, Oregon State University, Corvallis, Oregon, 97331, USA

Abstract

Cultivated meadowfoam (*Limnanthes alba* Benth.), a novel very long-chain oilseed crop, is a predominantly allogamous, insect-pollinated species. The aim of this research was to elucidate the anatomical, developmental, genetical and evolutionary mechanisms underlying the shift from allogamy to autogamy in a meadowfoam population segregating for selfing rate and floral morphology. Electron microscopy analyses revealed that protandry is the major reasons for the differences in the autonomous seed set of OMF40-11 and OMF64. A total of 20 quantitative trait loci (QTL) were identified for five mating system characters in meadowfoam. Individual QTL for mating system traits (*pa*, *spp* and *spf*) account for up to 20% of the backcross phenotypic variance. Taken together, the QTL for *spp*, *spf* and *pa* accounted for 66, 54 and 28% of the backcross variation. The QTL for protandry and chiasma frequency were adjacent to the QTL for *spp* and *spf*. This study has provided evidence that the correlation between the chiasma frequency and the type of mating system is not a direct developmental relationship between these factors, but is due to a selective advantage of the combination of the characters found. The speculation that the genetic factors underlying chiasma frequency and autonomous seed set have co-evolved in the process of evolution negates the hypothesis that self-fertilization as an "evolutionary dead end."

Key words: SSR, evolution, meadowfoam, QTL, autogamy, recombination

Introduction

Self-fertilization is a common and widespread phenomena in angiosperms, especially among annual herbs. The shift from outcrossing to selfing is one of the most frequent evolutionary transitions in the plant kingdom (Stebbins, 1950; Grant, 1981). Autogamy has been proposed to have a selective advantage in several situations: the successful establishment of a taxon following chance long-distance dispersal (Baker 1955); as an alternative means of achieving a rapid build-up of a population with relatively uniform, well adapted genotypes from a few survivors or a single invader plant in ephemeral or weedy species (Stebbins, 1950; Mather, 1953); and as an alternative means of producing seed when pollen vectors are lacking (Arroyo, 1973). Despite providing the reproductive assurance, selfing may be a "evolutionary dead end" as inbred lineages continually go extinct and new lineages are founded from outcrossing progenitors (Stebbins, 1957, 1974; Grant, 1958; Wyatt, 1988). Chromosomal characteristics favoring high recombinational rates, specifically, large number of chromosomes and a high chiasma frequency should be selectively advantageous in selfing species (Stebbins, 1950). Inbreeders are expected to have a high recombination index (chiasma frequency per cell) thereby maximizing the benefits of occasional outcrosses by maximizing new gene combinations. The theory is attractive because predominantly selfing species can achieve a dynamic balance between genic homozygosity and heterozygosity. Cytological studies in selfing and outbreeding taxa of *Elymus*, *Sitanion,* and

Agropyron (Stebbins, Valencia and Valencia, 1946), *Gilia* (Grant, 1952), *Lolium* (Jones and Rees, 1966), *Limnanthes* (Arroyo, 1973) and *Senecio* (Gibbs et al. 1975) support this hypothesis. Evolutionists have not exploited the potential for gene mapping with molecular markers to identify and characterize loci underlying chiasma frequency differences.

The evolution of selfing has been studied in a variety of theoretical models, each with specific assumptions about its inheritance. A purely polygenic basis for selfing was invoked by Lande and Schemske (1985) in their models for the joint evolution of selfing and inbreeding depression. Fenster and Ritland (1994) estimated between five to 12 loci for the differences in several mating system characters between *M. guttatus* and two other inbreeding taxa *M. micranthus* and *M. laciniatus*, but these estimates had large standard errors. Quantitative trait loci (QTL) differentiating outbreeding *M. guttatus* from inbreeding *M. platycalyx* were reported by Lin and Ritland (1997). However, these investigators never reported QTL for autonomous seed set (seeds per flower in the absence of external pollinating agents), which would be one of the most important traits to distinguish outbreeding and inbreeding taxa. Autonomous seed set was not critically important for the study reported by Lin and Ritland (1997), as monkey flower is not a crop plant. In case of crop plants where the seeds are of economical importance if such a study were to be performed autonomous seed set would be the trait of most importance.

Meadowfoam (*Limnanthes alba* Bentham; Order: Brassicales; Family: Limnanthaceae) is a self-compatible, predominantly allogamous, insect pollinated species. The mating systems of meadowfoam have been widely studied (Arroyo, 1975; Brown and Jain 1979; Brown et al. 1979; Kesseli and Jain 1985, Jain 1978; McNeill, 1983; McNeill and Jain 1985, Ritland, 1984). These studies have shown that the mating systems in meadowfoam vary from a predominantly allogamous (*L.alba*) to a completely autogamous (*L. floccosa*) system. Protandry and heterostyly are the major reasons for allogamy whereas cleistogamy is the major reason for autogamy (Arroyo, 1975). Commercial meadowfoam (*L.alba*) fields are primarily pollinated by domestic honeybees (*Apis millifera* L.). Meadowfoam seed yields tend to be variable across years (Knapp and Crane 1999). Poor honeybee pollination is routinely blamed for low seed yield years. To address this problem, our lab undertook research on the genetics of self-pollination.

Partially autogamous variants have been discovered in wild populations of *L. alba* by Arroyo (1975), Brown (1977), Brown et al. (1979) and McNeill (1983). Strongly autogamous (non-cleistogamous) meadowfoam lines were developed by selecting for increased autogamy in wild populations of *L.alba* ssp. *versicolor* (Knapp and Crane, 1997; Crane and Knapp, 2000). Artificially selected autogamous accessions of meadowfoam can be crossed with allogamous meadowfoam to produce fertile hybrid progeny, thus allowing for the formal genetic analysis of the traits differentiating the two mating systems.

Kishore et al. (2002) developed 389 simple sequence repeat (SSR) markers for meadowfoam and developed a framework map using 84 SSRs. Thus enough groundwork has been laid to conduct a solid genetic analysis of the traits differentiating allogamy from autogamy in meadowfoam. The aim of this study was to elucidate the anatomical, developmental and genetical mechanisms underlying the shift in the mating system of meadowfoam from allogamy to autogamy.

Materials and Methods

Electron microscopy analysis

For studying the developmental patterns of the allogamous and autogamous bud and flower development, we collected buds and flowers of OMF40-11 and OMF64. It takes about 4 weeks from the day the bud is first seen on the plant till the flower matures and wilts. The buds and flowers were collected at four different time intervals (stage 1- seven days after the bud first appeared on the plant, stage 2 – seven days after stage 1, at a gap of one week each) and were fixed in a formalin fixative (90% of 70% ethanol, 5% acetic acid and 5% of formalin). All the samples were then moved through 50, 70 and 100% water/ethanol solutions, 20-30 minutes /change. The samples were later critical point dried in a Balzer CPD-020 (Balzer Union Ltd. Liechtenstein) dryer using carbon dioxide and following the method of Anderson (1951). They were later dissected and mounted on aluminum planchettes using Duco cement (Devcon Corp., Wood Dale, IL, USA). The specimens were later coated withy 20nm of 60/40-wt % Au/Pd alloy in a VARIAN VE-10 (VARIAN Inc, Palo Alto, CA, USA) vacuum evaporator at 1×10^{-5} Torr. The coated specimens were later observed under an AmRAY 3300 FE (AmRAY Inc. Bedford, MA, USA) scanning electron microscope operated in the electron microscope facility at the Department of Botany and Plant pathology at Oregon State University. Images were recorded on Polaroid Type 55 P/N positive /negative

4x5" format film. We expect this work to define stepwise morphological changes associated with floral development of both allogamous and autogamous meadowfoam.

Plant material and population structure

This study was performed using 160 families from backcross population ([OMF40-11x OMF64] x OMF64) described by Kishore et al. (2002). The donor parent (OMF40-11) is an inbred line developed from the *L. alba* ssp. *alba* cultivar Mermaid. The recurrent parent (OMF64) is an inbred line developed from the *L. alba ssp. versicolor* accession PI 374801 (Knapp and Crane, 1997). DNA was extracted from 160 BC_1 progeny as described by Kishore et al. (2002). These BC_1 progeny were selfed to produce BC_1S_1 seed. Ten seeds from each BC_1 family were randomly chosen and germinated at 4°C in the dark on moistened blotter paper in 11x11x3 cm clear plastic boxes. Germinants were transplanted to potting soil (pumice: peat moss: sandy loam) in 7.5cm^2 plastic pots. The plants were grown at 15°C for 25 to 28 days in a growth chamber (Model CEL 37-14, Sherer-Gillett CO., Marshall, Mich.) with 8 h of fluorescent light per day. Four-week-old plants were transferred to a green house and grown to maturity at 18°C (night) to 25°C (day) with 16 h of light per day. The bases of the pots and protruding roots were continuously submerged in 1-3 cm of water in the growth chamber and greenhouse. A total of 1209 plants (as some seeds didn't germinate and some plants couldn't

survive after germination) were phenotyped for the traits mentioned below. The number of BC_1S_1 progeny from one family varied from five to ten plants with an average of seven plants per family. The least square means of the families for the traits mentioned below were used for the analysis.

Phenotypic evaluation of the traits

Petal area (*pa*):

The corolla of meadowfoam flowers is made up of five petals. The petals of OMF40-11 (*Limnanthes alba* ssp. *alba*) are distinctively larger and different in appearance from the petals of OMF64 (*Limnanthes alba* ssp. *versicolor*). To measure the differences in the petal morphology of the mapping parents and progeny, we used a portable area meter (Model LI-3000, LI-COR, Inc., Lincoln, NE, USA), which scanned the each of the five petals and provided the total petal area in cm^2.

Seeds per plant (*spp*):

The plants were harvested at the end of their life cycle (four months in green house). The total number of seeds harvested from the plant was counted.

Seeds per flower (*spf*):

Total number of flowers on the plant was counted. Seeds per flower were derived by dividing the number of seeds per plant with the total number of flowers on the plant. As the plants were grown in a screened green house where there were no pollinators, the phenotypic values of seeds per plant and seeds per flower represent the autonomous seed set capacity of the plants.

Selective phenotyping

Ten remnant seed from 20 BC_1 families (ten from each end of the phenotypic distribution for seeds per flower) were germinated. A total of 183 plants (as some seeds didn't germinate and some plants couldn't survive after germination) were phenotyped for the traits mentioned below. The least square family means of the traits were used for the analysis.

Number of chiasmata per cell (*ch*):

For observing the number of chiasmata in meadowfoam pollen mother cells (PMC's), we collected floral buds of OMF40-11, OMF64, and the 183 BC_1S_1 plants. The floral buds varied in size from 2 to 4.5 mm with an increment of 0.25mm. They were fixed in an ethanol: acetic acid fixative (3:1). Buds of sizes ranging from 2.5 to 3.0 mm always had highest number of cells in diakinesis. The buds were dissected and the anthers were placed on the slide with a drop of 2% acetocarmine and squashed. The cells were spread on the slide and a coverslip was

placed. Then the slide was warmed up and turned upside down on a blotting paper and pressed behind the coverslip so as to flatten the cells. The slides thus prepared were observed under a light microscope (Zeiss axioscope 2). Photographs were taken by using a CCD camera (CoolSNAP ^{TM}, Roper Scientific, Tucson, AZ, USA) and processed with RSImage ^{TM} software version 1.07 (Roper Scientific, Tucson, AZ, USA). The number of chiasmata was counted from five cells in a plant. Thus making a total of approximately fifty observations for each BC_1 family.

Protandry (*pr*):

Protandry was measured as the time lag (hours) between anthesis and the receptivity of the stigma. The receptivity of stigma was tested using 1% benzidine in 60% ethanol, hydrogen peroxide (3%) and water, 4:11:22 by volume (Dafni, 1992). If the stigma is receptive (indicated by the presence of peroxidase), it breaks the hydrogen peroxide and oxidation of the benzidine gives a blue color. Thus the presence or absence of blue color would indicate the receptivity of the stigma. A single flower on each plant was checked at an interval of 12, 24, 26, 48, 60 and 72 hours from anthesis to record the receptivity of the stigma.

Length of the style (*ls*), filament (*lf*) and distance between the anther and stigma (*das*):

The length of the style, filament and the distance between the anther and stigma (millimeters) were measured for one randomly chosen flower for each plant

with digital calipers. The length of the style and filament was measured from the base of the calyx to the base of the stigma or base of the anther. The distance between the anther and stigma was measured as the of the distance between the center of the stigmatic lobes to the center of the farthest anther away from the stigma.

Linkage map construction

A linkage map covering almost 90% of the meadowfoam genome has been developed by Kishore et al. (2002) based on 96 BC_1 progeny. We used 57 well-spaced SSR markers dispersed across the genome to genotype all the 160 BC_1 progeny to create a framework map of meadowfoam. The SSRs identified were assayed on the 160 backcross progeny as described by Kishore *et al.* (2002), on an ABI Prism 377 DNA sequencer (Applied Biosystem, Perkin Elmer, Foster City, CA, USA) using polyacrylamide gels and fluorescently labelled amplicons. The linkage maps were constructed as explained in Kishore et al. (2002) using MAPMAKER (Lander et al. 1987) and G-Mendel (Holloway and Knapp 1993).

Statistical analysis

Trait means and correlations were calculated using SAS (SAS institute, Cary, N.C.). Simple linear regression analysis was performed for all the markers and traits using PROC REG in SAS. A significance level of $P \leq 0.05$ was used as

the threshold for QTL detection. Composite interval mapping was done using MQM mapping method in MapQTL version 4.0 (Van Ooijen and Maliepaard, 1996). Co-factors were selected by a backward elimination method at $P \leq 0.02$. The null hypothesis of no QTL was tested for positions throughout the genome by comparing LOD scores to an empirical genome-wise significance threshold calculated from 1000 permutations for $P = 0.05$ (Doerge and Churchill, 1996). Interactions between significant QTL for each trait were tested, as the interactions between the closet loci for two different QTL. This was performed using PROC ANOVA in SAS.

Results

Floral development in allogamous and autogamous meadowfoam

Style and filament length, the distance between the anther and the stigma and the time lag between anthesis and stigma receptivity (protandry) were significantly different in OMF40-11 and OMF64. Scanning electron microscopy studies of developing buds revealed marked differences in carpel and stamen morphology and development in the autogamous versus allogamous parents. of OMF64 and OMF40-11 (Fig. 4.1). We could clearly see that though both the buds were of the same developmental stage (Fig. 4.1A and 4.1B, seven days after the bud was first sighted on the plant), the development of the carpel and stamens of OMF64 was more synchronized than their development in OMF 40-11. Six days pre anthesis, the anthers were in much closer proximity to the developing stigmas in the selfer (Fig. 1D) than the outcrosser (Fig. 1B). The stylar separations were greater in the selfer (Fig. 1E) than the outcrosser (Fig. 1G) at anthesis. Through-out the floral development the time lag between the maturation of the stamen and the carpel in case of OMF40-11, whereas in case of OMF64 this time lag is small. Besides these visual observations, several traits were measured for both OMF40-11 and OMF64 and their backcross progeny. The floral parts of the autogamous parent were significantly reduced compared to the allogamous parent (Table 4.1). OMF64 individuals had smaller petal areas (Fig. 4.4), shorter styles and filaments,

and slightly shorter distances between the anthers and stigmas at anthesis. Moreover, OMF64 had a significantly shorter protandry period and produced significantly more chiasmata per cell than OMF40-11. The seeds per flower distribution was exponential and left-skewed towards the phenotypic mean for the outcrossing parent (OMF40-11) (Fig. 4.2). Transgressive segregation was observed among BC_1S_1 progeny for increased autonomous seed set. Of the 160 BC_1 families phenotyped 12 families produced more seeds per flower than the selfing parent (0.72 seeds per flower) (Fig. 4.2).

Linkage map

We constructed a genetic linkage map using previously mapped SSR markers (Kishore et al. 2002) and DNA from 160 BC_1 progeny. The map was 934 cM long (Fig. 4.3), and a mean density of 16 cM per locus. The number of loci in each linkage group varied from 9 (LG 3, 4, and 5) to 16 (LG2). The map was predicted to cover approximately 90% of the meadowfoam genome (Kishore et al. 2002).

Summary of QTL detected

Petal area (*pa*): Three QTL were detected for petal area on a different linkage groups 2,4, and 5 (Fig. 4.6) . The QTL on linkage group 2, *pa2.1,* was the most significant QTL as it had a LOD score of 4.22 and explained 12% of the

phenotypic variance. *pa 5.1*, on linkage group 5 which explained 11% of the phenotypic variance, followed it. The QTL on linkage group 3, *pa 3.1*, was marginally significant as it had a LOD score of 1.70 and it explained only 5% of the phenotypic variance. Thus a total of only 28% of the total phenotypic variance was explained by all the three QTLs. OMF40-11 alleles (*L. alba* ssp. *alba*) for two QTL (*pa 2.1* and *pa 3.1*) were associated with increased petal area , whereas the allele from OMF64 (*L.alba* ssp. *versicolor*) was associated with *pa 5.1*.

Seeds per flower (*spf*): Five QTL for seeds per flower were found on four linkage groups (1, 2, 4, and 5) and accounted for 54% of the phenotypic variance (Fig 4.6) (Table 4.2). *spf 5.1,* the QTL on linkage group 5 was the most significant QTL for seeds per flower, however it accounted for only 15% of the total phenotypic variance, which is lower the variance explained by *spf 1.1* (20 % at a LOD of 2.78), the QTL on linkage group 1. The second QTL on linkage group 1, *spf 1.2,* was marginally significant (LOD 1.68) and accounted for only 7% of the phenotypic variance. The QTL on linkage groups 2 and 4, *spf 2.1* and *spf 4.1,* accounted for 7 and 5 % of the variance explained. Surprisingly, OMF40-11 alleles for two QTL (*spf 1.2* and *spf 4.1*) were associated with an increase in seeds per flower. This result was not expected because the autonomous seed set of OMF40-11was almost zero (Table 4.1). For the remaining QTL, the alleles from OMF64 were associated with the increase in seeds per flower, as expected.

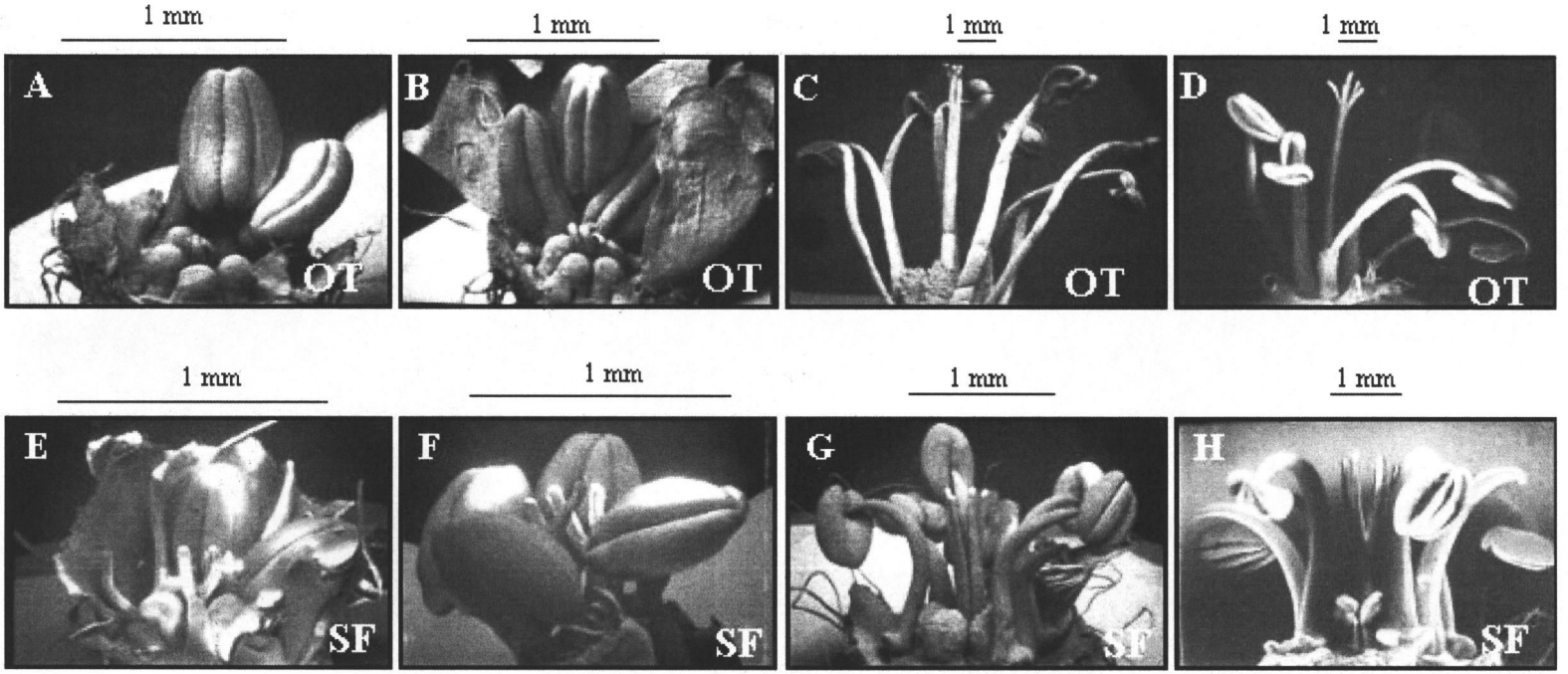


Figure 4.1. Electron micrographs showing the different floral developmental stages of allogamous and autogamous meadowfoam. A, C, E and G are the floral developmental stages for OMF40-11, whereas B, D, F and H are the floral developmental stages for OMF64. A and B represent the stage of the bud seen seven days after the bud is first sighted on the plant. C and D represent the stage of the bud seen 14 days after the bud is first sighted on the plant. E represents the stage of anthesis, whereas F represents the stage when both the anthers and the stigma are about to mature in an open flower. G and H represent when the stigma is receptive in an open flower. Note the differences in scales for each electron micrograph

Table 4.1. Phenotypic means, standard deviations (SD) and ranges for OMF40-11, OMF64 and 160 BC_1 [(OMF40-11 x OMF64) x OMF64] families

.

Trait	Mean±SD			Number of BC_1 families	BC_1 families range	
	OMF40-11	OMF64	BC_1 families		Minimum	Maximum
Petal area (cm^2)	4.49±0.54	2.47±0.45	3.20±0.55	160	1.69	4.86
Seeds per plant	1.14±1.07	9.38±4.31	2.74±2.58	160	0.0	15.0
Seeds per flower	0.09±0.09	0.72±0.33	0.25±0.28	160	0.0	1.91
Style length (mm)	5.05±0.30	3.6±0.35	3.90±0.61	20	1.98	5.39
Filament length (mm)	5.70±0.81	3.6±0.43	4.20±0.65	20	2.57	5.83
Distance between anther and stigma (mm)	3.24±0.57	2.80±0.54	2.76±0.72	20	1.23	4.90
Protandry (hrs)	61.5±7.69	39.6±5.79	38.57±10.32	20	20.0	50.4
Number of chiasmata	6.75±0.71	9.80±1.32	7.57±1.35	20	5.6	9.2

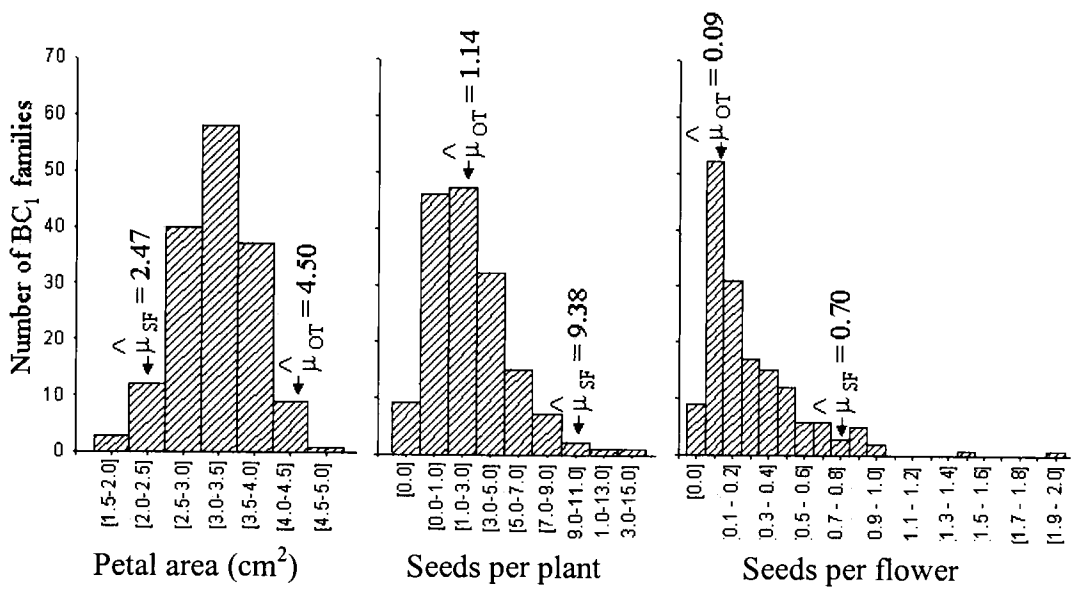


Figure 4.2. Histograms showing the phenotypic distributions of mating system characters in the backcross progeny.

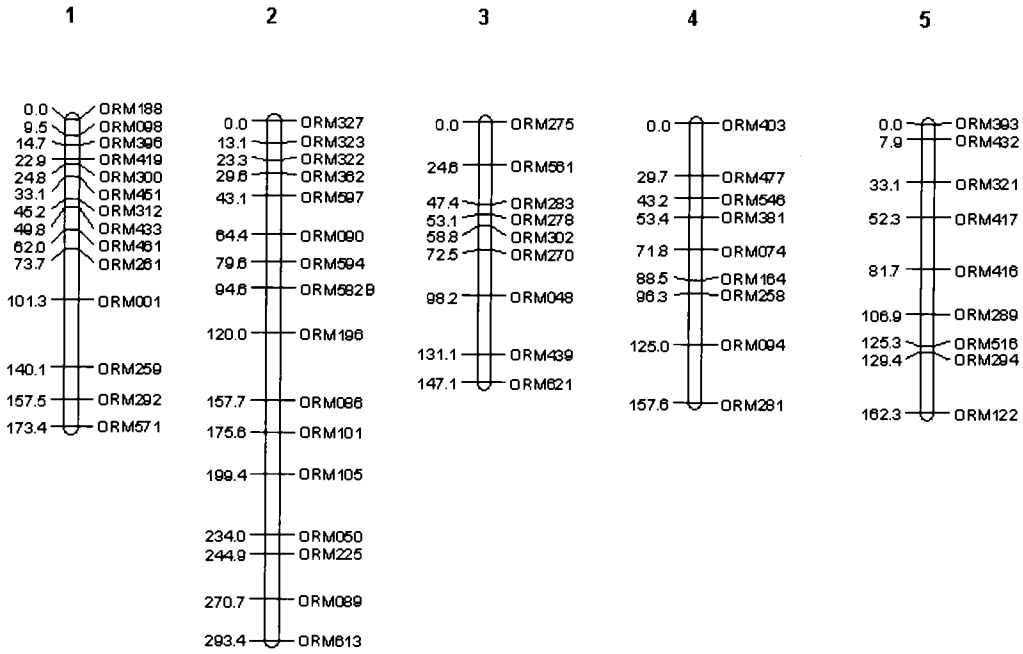


Figure 4.3. Linkage map of the meadowfoam developed screening 57 simple sequence repeat markers on 160 backcross progeny.

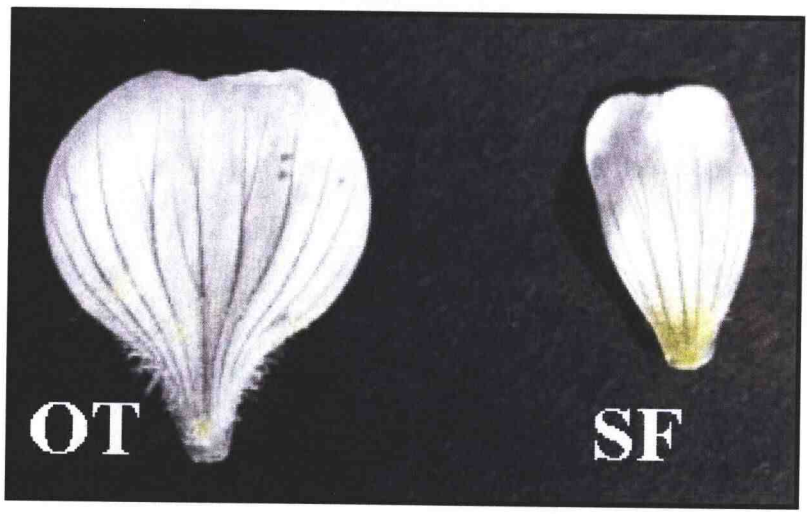


Figure 4.4. Differences in the petal dimensions of the outcrossring (OT) and self-fertlizing (SF) meadowfoam lines.

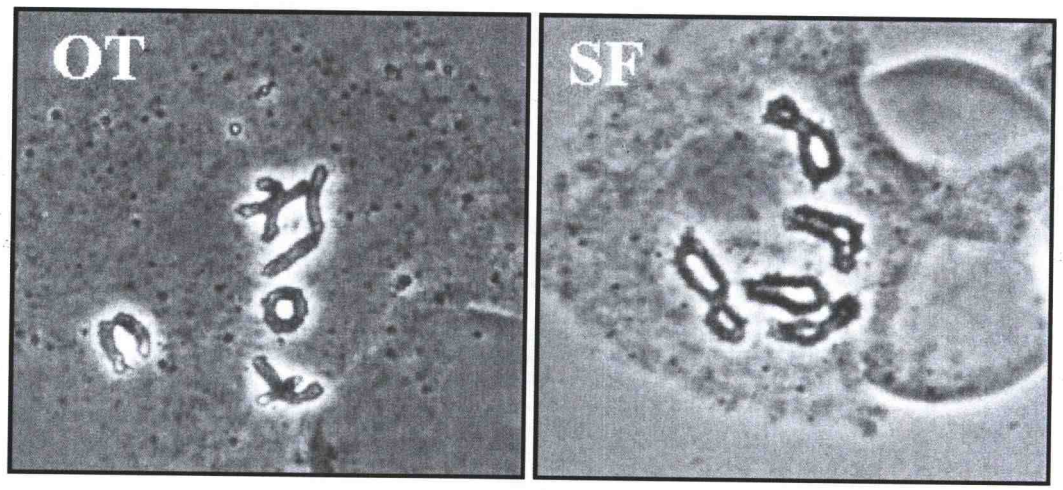


Figure 4.5 Differences in the chiasmata frequency in the PMCs of the outcrossing (OT) and self-fertilizing (SF) meadowfoam lines.

Seeds per plant (*spp*): Eight QTL were detected for seeds per plant. Five QTL (*spp 1.1*, *spp1.2*, *spp2.2*, *spp 4.1* and *spp 5.1*) overlapped with the QTL for seeds per flower. One QTL partially (*spp2.4*) overlapped with a QTL for petal area (*pa 2.1*). *spp 2.1* and *spp 2.3* were the two QTL which were unique for seeds per plant. All the eight QTL explained a total of 66% of the phenotypic variance for seeds per plant. *spp 1.2* and *spp 5.1* each explained 12 % of the phenotypic variance. The QTL on linkage group 2, *spp 2.1*, *spp 2.2*, *spp 2.3* and *spp 2.4* explained 7, 6, 7 and 9 % of the phenotypic variance respectively. Interestingly for half of the *spp* QTL , the alleles from *L. alba* ssp. *alba* were associated with increase in the seeds per plant.

Protandry (*pr*): A single QTL, *pr 5.1* was located on linkage group 5 (Fig. 4.7). It was marginally significant (LOD, 1.94) however it explained 30 % of the total phenotypic variance. *pr 5.1* coincided with the QTL *pa 5.1*. As expected, for *pr 5.1*, the allele from *Limnanthes alba* ssp. *alba* was associated with the increase in time lag between the time of anthesis and stigma receptivity.

Number of chiasmata per cell (*ch*): Three QTL located on two linkage groups (3 and 5) were identified for the number of chiasmata per cell (Fig. 4.7). A total of 63 % of the phenotypic variance was explained by these three QTL. *ch 5.2* was the most significant (LOD, 8.79) QTL identified and explained 30% of the

phenotypic variance. The other two QTL *ch 3.1* and *ch 5.1* explained 15 and 18 % of the phenotypic variance. For all the three QTL the alleles from *Limnanthes alba* ssp. *versicolor* had a positive effect in increasing the number of chiasmata per cell.

Correlation between traits

Surprisingly, there was no significant correlation between the seeds per plant and any of the traits related with spatial isolation of the male and female reproductive organs, i.e. *ls*, *lf* and *das*. Seeds per flower was negatively correlated with *lf*, however it was not significantly correlated with *ls* and *das*. As no QTL could also be detected for *ls*, *lf* and *das* it seems plausible that the data obtained for these traits from 20 BC_1 families was not sufficient to detect any relationships with seeds per plant. The correlation between seeds per flower and seeds per plant was strong (0.90) and highly significant. Despite the strong correlation three *spp* QTL on linkage group 2, had no corresponding *spf* QTL. Petal area was also moderately correlated with *spp* and *spf*. There was a partial overlap between *pa 2.1* and *spp 2.4*, suggesting either pleiotropy or tight linkage between QTL. Protandry was negatively correlated (-0.96, $p < 0.0001$) with seeds per flower (Fig 4.8A). The single QTL identified for protandry *pr 5.1* was adjacent to the *spp 5.1* and *spf 5.1*. However there was no over lap between *pr 5.1* and *spp 5.1* or *spf 5.1*. Surprisingly, though the correlation between protandry and petal area was not significant, *pr 5.1* coincided with *pa 5.1*. The number of chiasmata per cell was significantly

correlated with *spp*, *spf*, *pr* and *ls*. Though it was strongly correlated with *spp* (0.75), *spf* (0.85) (Fig 4.8B) and *pr* (-0.85) none of the QTL identified for *ch* coincided with the QTL for *spp*, *spf* or *pr*. The two QTL *ch 5.1* and *ch 5.2* were adjacent to *spp5.1* and *spf 5.1*. In general, from the correlations and the QTL positions for the mating system related characters, it appears that the tight linkage between the loci controlling these characters rather than pleiotropy plays a significant role in the evolution of mating systems.

Table 4.2. Summary of the QTLs detected for each of the mating system traits.

Trait	QTL	Linkage group	NML*	LOD	PVE (%)	Trait means at QTL		Effect of substitution of the QTL
						AA	Aa	
Petal area	*pa 2.1*	2	ORM086	4.22	12	3.00	3.38	-0.38
	pa 3.1	3	ORM275	1.70	05	3.01	3.28	-0.27
	pa 5.1	5	ORM432	2.78	11	3.42	3.05	0.37
Seeds per flower	*spf 1.1*	1	ORM098	2.78	20	0.58	-0.11	0.69
	spf 1.2	1	ORM419	1.68	07	0.05	0.45	-0.40
	spf 2.1	2	ORM090	2.42	07	0.36	0.14	0.22
	spf 4.1	4	ORM403	1.62	05	0.15	0.30	-0.14
	spf 5.1	5	ORM417	4.73	15	0.43	0.18	0.25
Seeds per plant	*spp 1.1*	1	ORM098	2.06	08	4.70	0.55	4.15
	spp 1.2	1	ORM419	2.56	12	-0.46	5.73	-6.19
	spp 2.1	2	ORM362	2.88	07	1.50	4.28	-2.78
	spp 2.2	2	ORM090	2.02	06	4.01	1.76	2.25
	spp 2.3	2	ORM582	2.72	07	3.96	1.82	2.14
	spp 2.4	2	ORM086	1.67	09	2.05	3.94	-1.90
	spp 4.1	4	ORM403	1.77	05	1.79	3.10	-1.31
	spp 5.1	5	ORM417	3.56	12	4.18	2.20	1.97
Protandry	*pr 5.1*	5	ORM432	1.94	32	30.0	50.0	-20.0

Table 4.2 (continued)

Chaismata	*ch 3.1*	3	ORM048	4.65	15	8.86	6.29	2.57
	ch 5.1	5	ORM289	6.28	18	8.94	6.16	2.78
	ch 5.2	5	ORM294	8.79	30	8.95	6.17	2.78

*NML= nearest marker locus to the QTL
AA = Trait means for homozygous OMF64
Aa = Trait means coded for heterozygous (OMF64 and OMF40-11)
PVE (%) = is the percentage of phenotypic variation explained

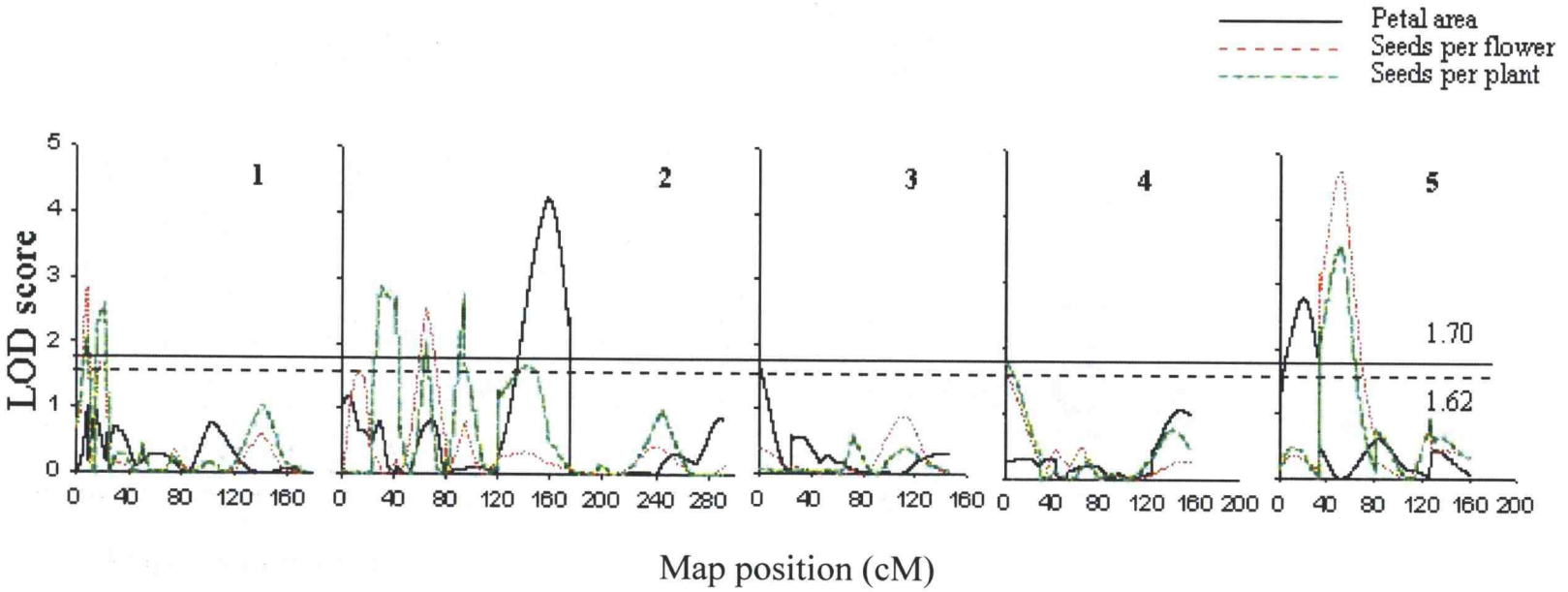


Figure 4.6. LOD plots of the QTL identified for petal area, seeds per flower and seeds per plant. The horizontal lines are the threshold lines at 0.05 significance. LOD 1.70 is the threshold for petal area, whereas LOD of 1.62 is the threshold for seeds per flower and seeds per plant. The numbers 1,2,3,4, and 5 represent the linkage groups.

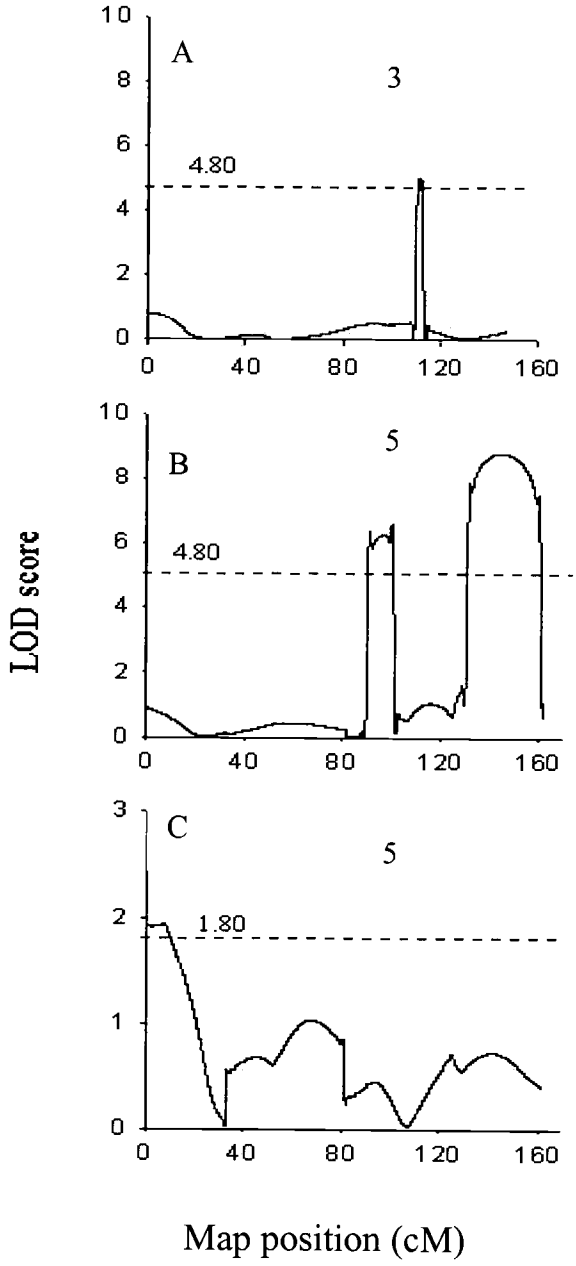


Figure 4.7. LOD plots for chaisma frequency (A and B) and protandry (C). The numbers 3 and 5 represent the linkage groups. The horizontal lines represent the threshold lines at 0.05 significance.

Table 4.3. Phenotypic correlation coefficients between mating system characters in a backcross population of *L. alba* ssp. *alba* x *L. alba* ssp. *versicolor*

	pa	*spp*	*spf*	*ls*	*lf*	*das*	*pro*	*ch*
pa	1.0	0.32**	0.32**	-0.34	-0.43*	-0.25	-0.41	0.45
spp		1.0	0.90**	0.01	-0.41	-0.28	-0.92**	0.75**
spf			1.0	0.03	-0.43*	-0.26	-0.96**	0.85**
ls				1.0	0.65**	0.15	-0.01	-0.01
lf					1.0	0.28	0.48*	-0.51*
das						1.0	0.22	-0.20
pro							1.0	-0.85**
ch								1.0

* = $P \leq 0.05$

** = $P \leq 0.0001$

Correlations for the traits underlined are based on the values of 20 BC_1 families.

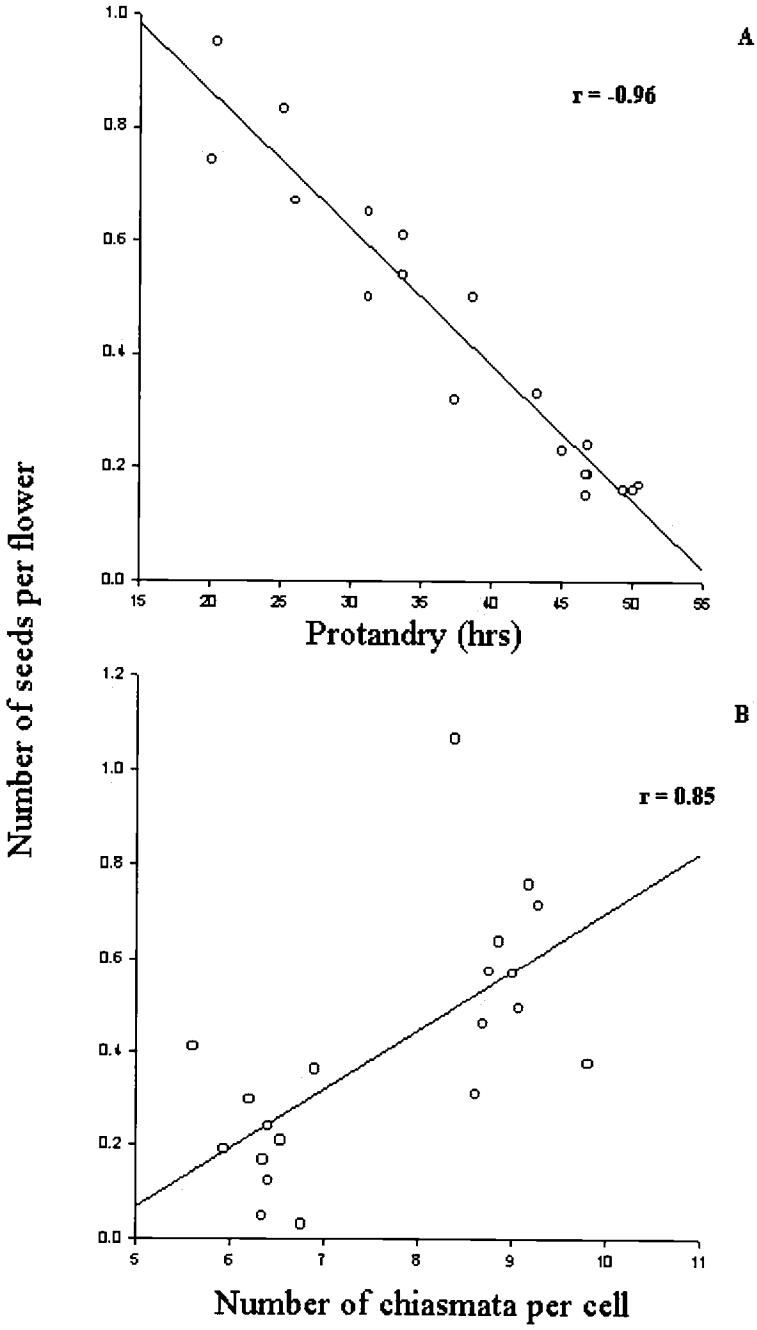


Figure 4.8. Scatter plots showing the relation ship between protandry and seeds per flower (A); Number of chaismata per cell and seeds per flower (B).

Discussion

The selective reasons for the change in mating systems from allogamy to autogamy have been a subject of debate since the publication of Charles Darwin's (1877) "The effects of Cross- and self-fertilization in vegetable kingdom". Early workers including Darwin(1877) and Muller (1883) were impressed with the abundance of autogamous taxa in the high alpine floras of Europe and concluded that autogamy evolved to ensure pollination and seed-set where weather conditions were unfavorable for insect pollination. For Darwin and Muller, autogamy was "fertility insurance", but it was only purchased at the price of decreasing benefits of cross-pollination. Stebbins (1950) proposed that chromosomal characteristics favoring high recombinational rates would be selectively advantageous for autogamous species. Chiasma frequency has been shown to reflect the amount of genetic recombination in chromosomes at meiosis in diverse organisms (Lewis and John, 1963; Bodmer and Parsons, 1962). Chiasma frequency like any other genetic factors affecting the genetic variability is subject to selection, and may increase or decrease the amount of genetic variability in populations due to recombination alone (Lewis and John, 1963). This study is the first study to use genetic markers to document QTL differences between taxa that differ in chiasma frequencies. Besides, this is the only study where QTL for autonomous seed set (seeds per flower) have been identified and are perceived as the major focus of mating system

differences rather than related characters like flower length, pistil length etc (Lin and Ritland, 1997; Fenster and Ritland, 1994).

Morphological observations along with the electron microscopy studies revealed that protandry and the spatial isolation of the male and the female reproductive organs are the major reasons for the differences in the capability of autonomous seed set of OMF40-11 and OMF64. However only protandry showed high correlation with *spp* and *spf*. The traits related to the spatial isolation of the male and the female reproductive organs (*ls*, *lf* and *das*) showed no significant correlation with *spp* or very weak correlation with *spf*. Hence our assumption that sampling the progeny from the extreme ends of the phenotypic distribution for *spf* would lead to the sampling of the progeny from the extreme ends of the phenotypic distribution for *ls*, *lf* and *das* failed. Hence no QTL could be identified for *ls*, *lf* and *das* among selectively genotyped progeny. However it doesn't mean that the spatial isolation of the male and the female reproductive organs doesn't play a role in autonomous seed set. In one of the selection experiments for developing highly autogamous meadowfoam in our lab (data not shown), transgressive progeny for autonomous seed set from a cross between autogamous and allogamous meadowfoam showed both decreased protandry and a novel combination of the *ls*, *lf* and *das* which made it more strongly autogamous than its autogamous parent. OMF109-3 illustrated in the fig. 4.9 is one of the transgressive progeny, which has the length of the style of the allogamous parent and the spread of the stigma from

the autogamous parent. In this study also it can be seen that two out of the five QTL identified for *spf* and four out of the eight QTL identified for *spp* had favorable alleles from OMF40-11, the allogamous parent. Hence we speculate that these QTL may be responsible for the length of the style and the length of the stylar split and consequent physical changes in stigma drooping , which would play an important role in a novel recombinant like OMF109-3.

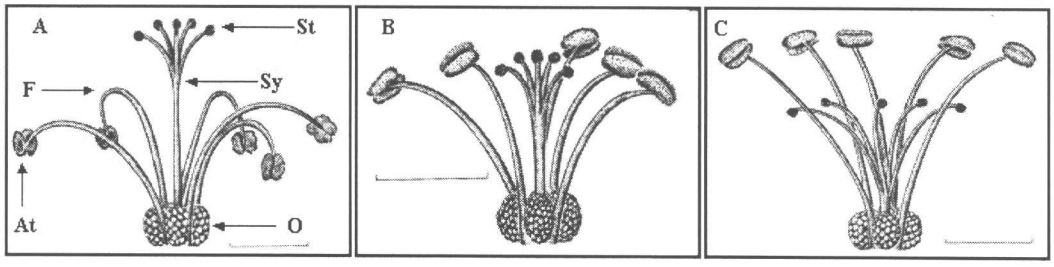


Figure 4.9. An artist's representation of reproductive organs of (A) allogamous parent of OMF109-3 (B) autogamous parent of OMF109-3 (C) OMF109-3. This is the stage of floral development when the stigma is receptive to pollen for all the three types of meadowfoam. St- stigma, Sy- style, F- filament, At – anther and O – Ovule. The bars represent 1mm in scale.

In an evolutionary comparison of two taxa, one cannot infer the direction of evolution of phenotypic traits and their underlying QTL as the true progenitor condition is unknown. However, most of the models and theories of evolution of mating systems (Jain 1976) suggest that allogamy is the ancestral mating system. Arroyo (1973) concluded that completely autogamous *L. floccosa* evolved from

the predominantly allogamous *L. alba*. In a study of evolution of autogamy in *L. floccosa*, Arroyo (1973) concluded that autogamy evolved in *Limnanthes* in relation to its pre-fertilization advantages of securing seed production in the absence of insect-pollinators and the associated tendency of seed production to occur rapidly, in marginal and uncertain habitats. OMF64 was selected for increased autogamy from a population of *L. alba* ssp. *versicolor* (PI374801) found at Shasta, California (Knapp and Crane, 2000). *L. floccosa* is also found around this region (Arroyo, 1973). As OMF64 is not cliestogamous as *L. floccosa*, we speculate that OMF64 is a transient form in the process of evolution of completely autogamous *L. floccosa*. Hence we assume that the QTL identified in this study for the different mating system characters are the QTL underlying the shift in the mating system of meadowfoam from allogamy to autogamy.

We found that between *L. alba* ssp. *alba* and *L. alba* ssp. *versicolor*, individual QTL for mating system traits (*pa*, *spp* and *spf*) account for up to 20% of the backcross phenotypic variance, with most QTL explaining 5 to15% of the phenotypic variance. Collectively, the QTL for spp, spf and pa accounted for 66, 54 and 28% of the backcross variation. Given that we used only 160 progeny in this study, only loci with relatively large effects can be detected (Beavis, 1994). The unaccounted variation is likely due to loci with smaller effects detectable only with a larger mapping population. These results are in consistent with the hypothesis that quantitative variation in mating systems is primarily controlled by loci with small

effects (Fenster and Ritland, 1994; Lin and Ritland, 1997). However, Bradshaw et al. (1995) detected at least one QTL accounting for more than 25% of the phenotypic variance for each of the eight floral traits distinguishing the bird-pollinated *M. cardinalis* from bee-pollinated *M. lewisii* and a single QTL explaining > 40% of the phenotypic variance was very common.

The QTL identified for protandry *pr 5.1* explained 32 % of the phenotypic variance, whereas all the three QTL identified for the number of chiasmata per cell explained a total of 63 % of the total phenotypic variance. However, these estimates may be biased as we used selective genotyping (reviewed in Lynch and Walsh 1998). Selective genotyping can result in a large increase in power to detect the QTL, for the simple reason that much of the linkage information resides in individuals with extreme phenotypes (Lebowitz et al. 1987; Lander and Botstein, 1989; Carey and Williamson, 1991; Darvasi and Soller, 1992). It has been estimated that a single monogenic trait can easily be mapped with 40 informative meioses (equivalent to 20 F_2 progeny) using markers at 20cM intervals (Lander and Schork, 1994; Bonyadi et al. 1997). Ohno et al. (2000) used a total of 18 progeny from the extreme ends of the distribution of 78 BC_1 progeny of rats segregating for hypertension to map the QTL underlying hypertension in rats. Hence we suggest though the estimates of the QTL detected for protandry and the number of chaismata per cell may be biased their locations are precise.

Out of 20 QTL identified for the various mating system characters, 14 QTL overlapped. Ten of the 14 QTL were for *spp* and *spf*. Since *spf* is a derived character from *spp*, we assume that the genes underlying the five QTL for *spf* are the same as the genes underlying the five overlapping QTL of *spp*. More closely spaced markers and greater numbers of advanced generation progeny like QTL-NILs are needed to exclude the possibility of a single QTL not being a composite of different loci. In this study it was seen that though protandry and number of chaismata per cell were highly correlated with *spp* and *spf*, the position of their QTLs didn't overlap but they were adjacent to each other. Hence we speculate that these characters co-evolved along with higher autonomous seed set.

The correlation between the chiasma frequency and the type of mating system is not a direct developmental relationship between these factors, but is due to a selective advantage of the combination of the characters found. In allogamous meadowfoam, an individual's variability and heterozygosity are secured in every generation by means of cross-pollination. Consequently the additional amount of gene recombination, which would be caused by large amount of crossing over, does not have a high selective value. In autogamous meadowfoam, however predominant self-fertilization has led to a variation pattern that consists of large number of homozygous or nearly homozygous races, with heterozygosity obtained periodically through occasional crossing between races. Then the ability of such species to adapt themselves to new conditions would depend largely on the ability

of the heterozygous interracial hybrids to produce many new gene combinations, some of which would have selective value, and thus form the beginning of a new race. Under these conditions, any factor increasing the amount of gene recombination would have an enhanced selective value, and this would hold for reduction of the amount of linkage by a higher chiasma frequency and frequency of crossing over. Kannanberg and Allard (1967) pointed out that very few autogamous species are obligately self-pollinating, while the majority shows small to medium amounts of cross-pollination. Hence we speculate that most of the autogamous species use the rare cross-pollination events to generate novel gene combinations by increasing the number of crossovers. Thus we suggest that the hypothesis that self-fertilization being an "evolutionary dead-end" is not true.

One of the important points to note is that, it is possible that the different chiasmata frequencies exhibited by allogamous and autogamous meadowfoam (Fig. 4.5) could be the result of differences in terminalization rates. If rates of terminalization were slower in autogamous meadowfoam, the chiasma frequencies would appear relatively higher than in out crossing species. This possibility was discarded by Arroyo (1973) by comparing the rates of terminalization against chiasma frequencies of four different species of meadowfoam. She concluded that there was no relationship between these two variables for *Limnanthes*. In this study we also tested the relationship between terminalization and chaisma frequency for OMF40-11 and OMF64 (data not shown), and couldn't find any relationship

between these two variables. It is therefore concluded that observed differences in chiasma frequencies obtained at diakinesis reflect differences in the number of chiasmata formed early in meiotic prophase.

Despite providing excellent information regarding the genetic factors underlying the mating system differences in meadowfoam and the co-evolution of these factors with other factors like chaisma frequency, there are some limitations to this study.

The modest size of the BC_1 mapping population (n = 160) and the non-normal distributions of the two important phenotypic traits(*spp* and *spf*) may have limited our ability to identify QTLs and accurately estimate the magnitude of their effects (reviewed in Lynch and Walsh, 1998). Part of the difficulty of mapping genes affecting autonomous seed set is that multiple physiological and development mechanisms underlie autogamy and the underlying phenotypes are threshold traits. Suppose multiple loci affect protandry, heterostyly, and other components of autonomous seed set, then progeny with no autonomous seed set are an amalgam of genotypes without the critical alleles or allele combinations required for *any* degree of autonomous seed set. While these progeny represent an array of genotypes, they are virtually uninformative for mapping because there are no phenotypic differences among them (the phenotype is 0.0 seeds per flower regardless of the genotype for progeny below the critical threshold). These progeny reduce power for mapping QTL underlying autonomous seed set.

One downside to using BC progeny is that gene action cannot be estimated because the contrast is between homozygotes and heterozygotes (e.g., y_{AA} - y_{Aa}). However, we used backcross progeny because RILs or other fixed progeny have not been produced for meadowfoam and might be too depressed by inbreeding to employ for phenotypic analysis.Second, we used backcross progeny because we speculated that the threshold trait effect would be most severe in F_2 and F_3 progeny and least severe in backcross progeny produced using the autogamous inbred as the recurrent parent.Using the latter increases the frequency of alleles for autonomous seed set and should reduce the threshold trait effect. The shortcomings of backcross and other balanced segregating population approaches for threshold traits might be decreased by using advanced backcross approaches, e.g., by producing and comparing QTL-NILs.

When interactions between significant QTL were tested it was found that none of interactions were significant. Thus suggesting that epistasis was not involved in the QTL identified in this study. However we need to recognize that interactions may be present between the QTL and the loci without significant main effects (Knapp, personal communication), which was not tested.

The use of artificially selected forms of *L.alba* ssp. *alba* and *L.alba* ssp. *versicolor* from the natural populations (Knapp and Crane, 2000) in the backcross mating design is one of the major limitations of the study. Although this facilitated linkage and QTL mapping, we have to be very cautious in extrapolating these

results to the natural populations. Hence, one has to realize that the QTL described in this study differentiating the two different mating systems in meadowfoam are not the only QTL differing in the wild populations of both the mating systems. However this doesn't detract us from concluding that the QTL described in this study are some of the QTL responsible for the mating system differences in meadowfoam.

In conclusion, this study has elucidated the developmental differences in the floral morphology of autogamous and allogamous meadowfoam along with the mapping of the genetic factors underlying these differences. It has provided evidence that the correlation between the chiasma frequency and the type of mating system is not a direct developmental relationship between these factors, but is due to a selective advantage of the combination of the characters found. The speculation that the genetic factors underlying chiasma frequency and autonomous seed set have co-evolved in the process of evolution negates the self-fertilization as an "evolutionary dead end" hypothesis.

References

Arroyo MTK (1973) Chiasma frequency evidence on the evolution of autogamy in Limnanthes floccosa (Limnanthaceae) Evolution, 27:679-688.

Arroyo MTK (1975) Electrophoretic studies of genetic variation in natural populations of allogamous *Limnanthes alba* and autogamous *Limnanthes floccosa*. Heredity, 35:153-164.

Baker HG (1955) Self-incompatibility and establishment after long distance dispersal. Evolution, 9:347-349.

Beavis WD (1994) The power and deceit of QTL experiments: lessons from comparative QTL studies. Proceedings of the 49th Annual Corn and Sorghum industry research conference, 250-266, Chicago, IL, USA.

Bodmer WF, Parsons PA (1962) Linkage and recombination in evolution. Advances in Genet., 11:1-100.

Bonyadi M, Rusholme SA, Cousins FM, Su HC, Biron CA, Farrall M, Akhurst RJ (1997) Mapping of a major genetic modifier of embryonic lethality in TGF beta 1 knockout mice. Nature Genet., 15:207-211.

Bradshaw HD, Wilbert SM, Otto KG and Schemske DW. 1995. Genetic mapping of floral traits associated with reproductive isolation in monkey flowers (Mimulus). Nature, 376:762-765.

Brown CR (1977) A comparison of patterns of variability in *Limnanthes alba* Benth. and *L. floccosa* Howel. Ph.D. Diss., Univ. Calif., Davis.

Brown CR, Hauptli H, Jain SK (1979) Variation in *Limnanthes alba*: A biosystematic survey of germplasm resources. Econ. Bot., 33:267-274.

Brown CR, Jain SK (1979) Reproductive system and pattern of genetic variation in two *Limnanthes* species. Theor. Appl. Genet., 54:181-190.

Carey G, Williamson J (1991) Linkage analysis of quantitative traits: increased power by using selected samples. Amer. J. Human Genet., 49:786-796.

Crane JM and Knapp SJ (2000) Registration of the self-pollinated meadowfoam germplasm line OMF64. Crop Sci., 40:300-301.

Dafni A (1992) Pollination ecology: a practical approach. IRL press at Oxford University press, UK.

Darvasi A, Soller M (1992) Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. Theor. Appl. Genet., 85:353-359.

Darwin C (1877) The effects of cross and self-fertilization in the vegetable kingdom. John Murray, London, UK.

Doerge RW, Churchill GA. 1996. Permutation tests for multiple loci affecting a quantitative character. Genetics, 142: 285-294.

Fenster CB, Ritland K. 1994.Quantitative genetics of mating system divergence in the yellow monkey flower species complex. Heredity, 73:422-435

Gibbs PE, Milne C, Varga Carrillo, M (1975) Correlation between the breeding system and recombination index in five species of *Senecio*. New Phytol., 75:619-626.

Grant V (1952) Cytogenetics of the hybrid *Gilia millaefolia* x *G. achilleaefolia* I. Variation in meiosis and polyploidy rate as affected by nutritional and genetic conditions. Chromosoma, 5:372-390.

Grant V (1958) The regulation of recombination in plants. Cold springs harbor Symposia on Quantitative Biology, 23:337-363.

Grant V (1981) Plant Speciation. Columbia University Press, New York, USA.

Holloway JL, and Knapp SJ (1993) Gmendel 3.0 Users Guide. Department of Crop and Soil Science, Oregon State University, OR, USA.

Isbell TA.1997. Development of meadowfoam as industrial crop through novel fatty acid derivatives. Lipid Technology, 9:140-144.

Jain SK (1978) Breeding systems in *Limnanthes alba*: Several alternative measures. Amer. J. Bot., 65:272-275.

Jain SK (1976) The evolution of inbreeding in plants. Ann. Rev. Ecol. Syst. 7: 469-495.

Jones N, Rees H (1966) Chiasma frequencies and potential genetic variability of *Lolium* populations. Nature, 211:432-433.

Kannanberg LW, Allard RW (1967) Population studies in predominantly self-pollinated species. VIII. Genetic variability in the *Festuca microstachys* complex. Evolution, 21:227-240.

Kesseli RV, Jain SK (1985) Breeding systems and population structure in Limnanthes. Theor. Appl. Genet., 71:292-299.

Kishore VK, Gandhi SD, Crane JM, Knapp SJ (2002) The meadowfoam genome and simple sequence repeat genome map (in preparation).

Knapp SJ, Crane JM (1997) The development of self-pollinated inbred lines of meadowfoam by direct selection in open-pollinated populations. Crop Sci., 37:1770-1775.

Lande R, Schemske DW (1985). The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. Evolution, 39: 24-40.

Lander ES and Schork NJ (1994) Genetic dissection of complex traits. Science, 265:2307-2048.

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics, 121:185-199.

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

Lebowitz RJ, Soller M, Beckman JS (1987) Trait-based analyses for the detection of linkage between marker loci and quantitative trait loci in crosses between inbred lines. Theor. Appl. Genet., 73:556-562.

Lewis KR, John B (1963) Chromosome marker. J. and A. Churchill, London UK.

Lin JZ, Ritland K (1997) Quantitative trait loci differentiating the outbreeding *Mimulus guttatus* from the inbreeding *M. platycalyx*. Genetics, 146:1115-1121.

Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer Associates Inc. Publishers, MA, USA.

Mather K (1953) The genetical structure of populations. Symp. Soc. Exp. Biol. 7:66-95.

McNeill CI (1983) genetic differentiation, species relationships and the evolution of mating systems in plant genus Limnanthes, section Inflexae. Ph.D. Diss. Univ. Calif., Davis.

McNeill CI, Jain SK (1985). Genetic differentiation studies and phylogenetic inference in plant genus *Limnanthes* (section Inflexae). Theor. Appl. Genet., 66:257-269

Muller H (1883) The fertilization of flowers. Macmillan, London, UK.

Ohno Y, Tanase H, Nabaki T, Otsuka K, Sasaki T, Suzawa T, Morii T, Yamori Y, Saruta T (2000) Selective Genotyping with epistasis can be utilized for a major quantitative trait locus mapping in hypertension in rats. Genetics, 155:785-792.

Ritland K (1984) A comparative study of floral and electrophoretic variation with life history variation in Limnanthes alba (Limnanthaceae). Oecologia, 63:243-251.

Stebbins GL (1950) Variation and evolution in plants. Columbia, New York, USA.

Stebbins GL (1957) Self-fertilization and population variability in plants. American Naturalist, 91:337-354.

Stebbins GL (1974) Flowering plants. Evolution above the species level. Belknap Press of Harvard University press, Cambridge Massachusetts, USA.

Stebbins GL, Valencia J, Valencia RM (1946) Artificial and natural hybrids in the Gramineae, Tribe Hordeae I. *Elmus*, *Sitanion*, *Agropyron*. Amer. J. Bot., 33:338-351.

Van Ooijen JW, Maliepaard C 1996. MapQTL (tm) version 4.0: Software for the calculation of QTL positions on genetic maps. Plant Genome IV Abstracts. World Wide Web: http://probe.nal.usda.gov/.

Wyatt, R (1988) Phylogenetic aspects of the evolution of self-pollination. In L.D. Gottlieb and S.K. Jain 9eds.), Plant evolutionary biology, 109-131. Chapman and Hall, London, UK.

CHAPTER 5

CONCLUSIONS

Meadowfoam (*Limnanthes alba* Benth.) is an annual oil seed crop plant native to Southern Oregon and Northern California (Mason 1952, Kalin 1971, Jain 1986). Meadowfoam seed oil contains unique unsaturated very long chain fatty acids (C_{20} and C_{22}) with outstanding oxidative stability (Isbell 1997). These novel characteristics of meadowfoam seed oil are of industrial interest and have fueled the development of meadowfoam as an oilseed crop (Knapp and Crane 1999). It has been cultivated since 1974, when the first non-shattering cultivar, Foamore, was developed and released for commercial production (Calhoun and Crane 1975). Since then, four more cultivars have been released for cultivation in a span of 27 years. In order to improve our efficiency of developing new cultivars for increasing the productivity of meadowfoam it is essential to understand the genome of meadowfoam along with the genetics of economically important traits. There were very few biochemical markers (Arroyo 1975; Brown and Jain 1979; Kesseli and Jain 1985) and virtually no DNA markers for meadowfoam until 2001. Katengam et al. (2002) developed amplified fragment length polymorphisms (AFLPs) for meadowfoam. AFLPs, however are dominant markers and difficult to use in regular breeding programs (Powell et al. 1996; Smith et al. 1997). Hence we decided to develop simple sequence repeat (SSR) markers for meadowfoam.

In our first project, we isolated and sequenced 1,596 clones from genomic DNA libraries enriched for AG_n or AC_n. Six hundred and ninety-six clones harbored unique SSRs ranging in length from 6 to 150 bp; primers were designed and tested for 624 unique SSRs. Three hundred and eighty-nine primer pairs (62.3%) produced clean amplicons and yielded functional SSR markers. Ninety-six percent of the SSR markers (373 out of 389) were polymorphic among the 14-germplasm accessions (from nine taxa) and the heterozygosity (H) and power of discrimination (PD) scores ranged from 0.0 to 0.93 with a mean of 0.63 and 0.64 respectively. A strong positive correlation was observed between the Standard deviation of molecular weight (SDmw), H and maximum repeat count (MRC). Genetic distances for the 14 meadowfoam accessions ranged from 0.35 ± 0.022 to 0.69 ± 0.043 with a mean of 0.57 ± 0.030. Cluster and principal component analyses of the genetic distance matrix uncovered patterns of diversity concordant with species, subspecies, and breeding origin. The SSR markers developed are excellent resources for molecular breeding in meadowfoam.

Genetic maps are important in plant breeding and are a powerful tool for localizing and isolating genes underlying both simple and complex traits. In our second project, our specific aims were to describe the karyotype and physical size of the meadowfoam genome and develop a simple sequence repeat (SSR) map for meadowfoam. The physical size of the meadowfoam genome was estimated to be 5.52 pg using flow cytometry; thus, the meadowfoam genome is ca. 16 times larger

than the *Arabidopsis* genome (0.35pg). Karyotype analyses revealed that the meadowfoam genome is made up of two metacentric and three submetacentric chromosomes. Meadowfoam has two pairs of chromosomes with subterminal nucleolar organizing regions (NOR's). The genetic map was constructed by genotyping 96 (OMF40-11 x OMF64) x OMF64 BC_1 progeny with 90 SSR markers and was comprised of 84 SSR loci dispersed among five linkage groups with 11 to 22 SSR loci per linkage (6 SSR loci segregated independently). The five linkage groups presumably correspond to the five haploid chromosomes of meadowfoam. The map was 988.7 cM long with a mean density of 11.8 cM and minimal clustering of loci. Fifty percent of the mapped loci were distorted towards OMF40-11. Thus based on the transmission ratio distortion favoring one of the mapping parents (OMF40-11), 30% pollen inviability in the hybrid between the mapping parents and presence of anaphase bridges in 26% of the PMCs of the hybrid, lead us to speculate that there is a biological mechanism underlying the systematic transmission of the alleles of *Limnanthes alba* ssp. *alba* alleles favored over *Limnanthes alba* ssp. *versicolor* alleles.

The mating systems of meadowfoam have been widely studied (Arroyo, 1975; Brown and Jain 1979; Brown et al. 1979; Kesseli and Jain 1985, Jain 1978; McNeill, 1983; McNeill and Jain 1985, Ritland, 1984). These studies have shown that the mating systems in meadowfoam vary from a predominantly allogamous (*L.alba*) to a completely autogamous (*L. floccosa*) system. Protandry and

heterostyly are the major reasons for allogamy whereas cleistogamy is the major reason for autogamy. With the newly developed molecular tools (chapter 2 and chapter 3) we can identify the genetic factors underlying the evolution of mating systems in meadowfoam. The aim of the third project was to elucidate the anatomical, developmental, genetical and evolutionary mechanisms underlying the shift from allogamy to autogamy in a meadowfoam population segregating for selfing rate and floral morphology. Electron microscopy analyses revealed that protandry is the major reasons for the differences in the autonomous seed set of OMF40-11 and OMF64. A total of 20 quantitative trait loci (QTL) were identified for five mating system characters in meadowfoam. Individual QTL for mating system traits [petal area (*pa*), seeds per plant (*spp*) and seeds per flower (*spf*)] account for up to 20% of the backcross phenotypic variance, with most traits showing QTL effects of 5-15%. Taken together, the QTL for *spp*, *spf* and *pa* accounted for 66, 54 and 28% of the backcross variation. The QTL for protandry and chiasma frequency were adjacent to the QTL for *spp* and *spf*. This study has provided evidence that the correlation between the chiasma frequency and the type of mating system is not a direct developmental relationship between these factors, but is due to a selective advantage of the combination of the characters found. The speculation that the genetic factors underlying chiasma frequency and autonomous seed set have co-evolved in the process of evolution negates the self-fertilization as an "evolutionary dead end"

This study presented,

(i) Development of SSR markers for meadowfoam and their utility

(ii) Characterized the genome size and karyotype of meadowfoam

(iii) The first linkage map of meadowfoam based on SSR markers.

(iv) Elucidated the anatomical, developmental, genetical and evolutionary mechanisms underlying the shift from allogamy to autogamy in a meadowfoam.

BIBLIOGRAPHY

Akkaya MS, Bhagwat AA, Cregan PB (1992). Length polymorphisms of simple sequence repeat DNA in soybean. Genetics, 132:1131-1139.

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature, 408:796-812.

Arroyo MTK (1973) Chiasma frequency evidence on the evolution of autogamy in *Limnanthes floccosa* (Limnanthaceae). Evolution, 27:679-688.

Arroyo MTK (1975) Electrophoretic studies of genetic variation in natural populations of allogamous *Limnanthes alba* and autogamous *Limnanthes floccosa* (Limnanthaceae). Heredity, 35:153-164.

Arumuganathan K, Earle ED (1991a) Estimation of nuclear DNA content of plants by flow cytometry. Plant Mol. Biol. Reporter, 9:221-231.

Arumuganathan K, Earle ED (1991b) Nuclear DNA content of some important plant species. Plant Mol. Biol. Reporter, 9:208-218.

Bagby MO, Smith CR, Miwa TK, Lohmar RL, Wolff LA (1961) A unique fatty acid from *Limnanthes douglasii* seed oil: The C_{22} diene. Journal of Organic Chemistry, 26:1261-1265.

Bailey NTJ (1961) Introduction to the mathematical theory of genetic linkage. Clarendon Press, Oxford.

Baker HG (1955) Self-incompatibility and establishment after long distance dispersal. Evolution, 9:347-349.

Beavis WD (1994) The power and deceit of QTL experiments: lessons from comparative QTL studies. Proceedings of the 49th Annual Corn and Sorghum industry research conference, 250-266, Chicago, IL, USA.

Bennett MD, Leitch IJ (2001) Plant DNA C-values database (release 1.0, Sept. 2001) http://www.rbgkew.org.uk/cval/homepage.html

Bernacchi D, Tanksley SD (1997) An interspecific backcross of Lycopersicon esculentum x L. hirsutum: linkage analysis and a QTL study of sexual compatibility factors and floral traits. Genetics, 147:861-877.

Blanc G, Barakat A, Guyot R, Cooke R, Delseny M (2000) Extensive duplication and reshuffling in the Arabidopsis genome. Plant Cell, 12:1093-1101.

Bodmer WF, Parsons PA (1962) Linkage and recombination in evolution. Advances in Genet., 11:1-100.

Bonyadi M, Rusholme SA, Cousins FM, Su HC, Biron CA, Farrall M, Akhurst RJ (1997) Mapping of a major genetic modifier of embryonic lethality in TGF beta 1 knockout mice. Nature Genet., 15:207-211.

Bowcock AM, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd JR, Cavalli-Sforza LL (1994) High resolution of human evolutionary trees with polymorphic microsatellites. Nature, 368:455-457.

Bradshaw HD, Wilbert SM, Otto KG and Schemske DW. (1995). Genetic mapping of floral traits associated with reproductive isolation in monkey flowers (Mimulus). Nature, 376:762-765.

Brown CR (1977) A comparison of patterns of variability in *Limnanthes alba* Benth. and *L. floccosa* Howel. Ph.D. Diss., Univ. Calif., Davis.

Brown CR, Hauptli H, Jain SK (1979) Variation in *Limnanthes alba*: A biosystematic survey of germplasm resources. Econ. Bot., 33:267-274.

Brown CR, Jain SK (1979) Reproductive system and pattern of genetic variation in two Limnanthes species. Theor. Appl. Genet., 54:181-190.

Calhoun W, Crane JM (1975) Registration of "Foamore" meadowfoam. Oregon Agri. Exp. Sta. Corvallis, OR, USA.

Carey G, Williamson J (1991) Linkage analysis of quantitative traits: increased power by using selected samples. Amer. J. Human Genet., 49:786-796.

Cerbah M, Coulaud J, Siljak-Yakovlev (1998) rDNA organization and evolutionary relationships in the genus Hypochaeris (Asteraceae). J. Heredity, 89:312-318.

Chakravarti A, Lasher LK, Reefer JE (1991) A maximum likelihood method for estimating genome length using genetic linkage data. Genetics, 128:183-193.

Charters YM, Robertson A, Wilkinson MJ, Ramsay G (1996) PCR analysis of oilseed rape cultivars (*Brassica napus* L. ssp. *oleifera*) using 5'-anchored simple sequence repeat (SSR) primers. Theor. Appl. Genet., 92: 442-447.

Cho YG, Ishii T, Temnykh S, Chen X, Lipovich L, McCouch SR, Park, Ayres N, Cartinhour S (2000). Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). Theor. Appl. Genet., 100:713-722.

Coe E, Cone K, McMullen M, Chen S, Davis G, Gardiner J, Liscum E, Polacco M, Paterson A, Sanchez-Villeda H, Soderlund C, Wing R. (2002). Access to the maize genome: An integrated physical and genetic map. Plant Physiol., 128:9-12.

Crane JM, Knapp SJ (2000) registration of the self-pollinated meadowfoam germplasm line OMF64. Crop Sci., 40:300-301.

Cregan PB, Jarvik T, Bush AL, Shoemaker RC, Lark KG, Kahler AL, Kaya N, VanToai TT, Lohnes DG, Chung J, Specht JE. (1999). An integrated genetic linkage map of the soybean genome. Crop Sci., 39:1464-1490.

Dafni A (1992) Pollination ecology: a practical approach. IRL press at Oxford University press, UK.

Darvasi A, Soller M (1992) Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. Theor. Appl. Genet., 85:353-359.

Darwin C (1877) The effects of cross and self-fertilization in the vegetable kingdom. John Murray, London, UK.

De Rocher EJ, Harkins KR, Galbraith DW, Bohnert HJ (1990) Developmentally regulated systemic endopolyploidy in succulents with small genomes. Science, 250:99-101.

De Villena FP, Sapienza C (2001) Female meiosis drives karyotype evolution in mammals. Genetics, 159:1179-1189.

Dietrich WF, Miller J, Steen R, Merchant MA, Damron- Boles D, Husain Z, Dredge R, Daly Mt, Ingalls KA, O'Connor JJ, Evans CA, DeAngelis MM, Levinson DM, Kruglyak L, Goodman N, Copelong NG, Jenkins NA, Hawkins TL, Stein L, Page DC, Lander ES (1996) A comprehensive genetic map of the mouse genome. Nature, 380:149-152.

Doerge RW, Churchill GA. 1996. Permutation tests for multiple loci affecting a quantitative character. Genetics, 142:285-294.

Dole JA, Sun M (1992) Field and genetic survey of the endangered Butte County meadowfoam-*Limnanthes floccossa* subsp. *californica* (Limnanthaceae). Conservation Biology, 6:549-558.

Dolezel J, Sgorbati S, Lucretti S (1992) Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. Physiological Plantarum, 85:625-631.

Don RH, Cox PT, Wainwright BJ, Baker K, Mattick JS (1991) 'Touchdown' PCR to circumvent spurious priming during gene amplification. Nucleic Acids Res.,19:4008.

Dudley JW (1993) Molecular markers in plant improvement: manipulation of genes affecting quantitative traits. Crop Sci., 33:660-668.

Eastman DC (1990) Rare and endangered plants of Oregon. Beautiful America publishing company, USA.

Edwards KJ, Barker JHA, Daly A, Jones C, Karp A (1996). Microsatellite libraries enriched for several microsatellite sequences in plants. BioTechniques, 20:758-760.

Endo TR, Katayama Y 91978) Finding of a selectively retained chromosome of *Aegilops Caudata* L. in common wheat. Wheat Inf. Serv., 47-48:32-35

Fenster CB, Ritland K. 1994. Quantitative genetics of mating system divergence in the yellow monkey flower species complex. Heredity, 73:422-435

Fishman L, Kelly AJ, Morgan E, Willis JH (2001) A genetic map in the *Mimulus guttatus* species complex reveals transmission ratio distortion due to heterospecific interactions. Genetics, 159:1701-1716.

Fransz P, Armstrong S, Alonso-Blanco C, Fischer TC, Torres-Ruiz RA, Jones G (1998) Cytogenetics for the model system Arabidopsis thaliana. Plant J., 13:867-876.

Galbraith DW, Harkins KR, Knapp S (1991) Systemic endopolyploidy in *Arabidopsis thaliana*. Plant Physiol., 96:985-989.

Galbraith DW, Harkins KR, Maddox JM, Ayres NM, Sharma DP, Firoozabady E (1983) Rapid flow cytometric analysis of the cell cycle in intact plant tissues. Science, 220:1049-1051.

Gibbs PE, Milne C, Varga Carrillo, M (1975) Correlation between the breeding system and recombination index in five species of *Senecio*. New Phytol, 75:619-626.

Gilissen LJW, Van Staveren MJ, Creemers-Molenaar J, Verhoeven HA (1993) Development of polysomaty in seedlings and plants of *Cucumis sativus* L. Plant Sci, 91:171-179.

Godelle B, Cartier D, Marie D, Brown SC, Siljak-Yakovlev (1993) Heterochromatin study demonstrating the non-linearity of fluorometry useful for calculating genomic base composition. Cytometry, 14:618-626.

Goldstein DB, Clark AG (1995) Microsatellite variation in North American populations of *Drosophila melanogaster*. Nucleic acids Res., 23:3882-3886.

Goldstein DB, Roemer GW, Smith DA, Reich DE, Bergman A, Wayne RK (1999) The use of microsatellite variation to infer population structure and demographic history in a natural model system. Genetics, 151:797-801.

Grafi G, Larkins BA (1995) Endoreduplication in maize endosperm: involvement of M phase-promoting factor inhibition and induction of S phase-related kinases. Science, 269:1262-1264.

Grant V (1952) Cytogenetics of the hybrid *Gilia millaefolia* x *G. achilleaefolia* I. Variation in meiosis and polyploidy rate as affected by nutritional and genetic conditions. Chromosoma, 5:372-390.

Grant V (1958) The regulation of recombination in plants. Cold springs harbor Symposia on Quantitative Biology, 23:337-363.

Grant V (1981) Plant Speciation. Columbia University Press, New York, USA.

Harushima Y, Yano M, Shomura A, Sato M, Shimano T, Kuboki Y, Yamamoto T, Lin SY, Antonio BA, Parco A, Kajiya H, Huang N, Yamamoto K, Nagamura Y, Kurata N, Khush GS, Sasaki T (1998) A high-density genetic linkage map with 2275 markers using a single F_2 population. Genetics, 148:479-494.

Holloway JL, Knapp SJ (1993) Gmendel 3.0 Users Guide. Department of Crop and Soil Science, Oregon State University, OR, USA.

Innan H, Terauchi R, Miyashita T (1997) Microsatellite polymorphism in natural populations of the wild plant *Arabidopsis thaliana*. Genetics, 146:1441-1452.

Isbell TA (1997) Development of meadowfoam as industrial crop through novel fatty acid derivatives. Lipid Technology, 9:140-144.

Jacqumard A, De Veylder L, Segers G, de Almedia Engler J, Bernier G, Van Montagu M, Inze D (1999) Expression of CKS1At in Arabidopsis thaliana indicates a role for the protein in both the mitotic and the endoreduplication cycle. Planta, 207:496-504.

Jain SK (1976) The evolution of inbreeding in plants. Ann. Rev. Ecol. Syst., 7: 469-495.

Jain SK (1978) Breeding systems in *Limnanthes alba*: Several alternative measures. Amer. J. Bot., 65:272-275.

Jain SK (1986) Domestication of Limnanthes (Meadowfoam) as a new oil crop. In: Plant domestication induced mutation: Proceedings of an advisory group meeting on the possible use of mutation breeding for rapid domestication of new crop plants (ed. FAO/IAEA.), pp.121-134, Vienna, Austria.

Jensen KB (1989) Cytology, fertility and origin of *Elymus abolinii* (Drob) Tzvelv and its F_1 hybrids with *Pseudoroegneria spicataq*, *E. lanceolatus E. dentatus E. ugamicus* and *E. drobovii* (Poaceae: Triticeae). Genome, 32:468-474.

Jones K (1998) Robertsonian fusion and centric fission in karyotype evolution of Higher plants. Botanical Review, 64:273-289.

Jones N, Rees H (1966) Chiasma frequencies and potential genetic variability of *Lolium* populations. Nature, 211:432-433.

Kalin MT (1971) The evolution of autogamy in *Limnanthes* section Inflexae. Ph.D. Dissertation, University of California, Berkeley.

Kannanberg LW, Allard RW (1967) Population studies in predominantly self-pollinated species. VIII. Genetic variability in the *Festuca microstachys* complex. Evolution, 21:227-240.

Karagyozov L, Kalcheva ID, Chapman VM (1993). Construction of random small-insert genomic libraries highly enriched for simple sequence repeats. Nucleic Acids Res., 21:3911-3912.

Katengam S (2000) DNA finger printing and genome mapping in meadowfoam. PhD. Thesis, Oregon State University.

Katengam S, Crane JM Knapp SJ (2002) The development of a genetic map for meadowfoam comprised of amplified fragment length polymorphisms. Theor. Appl. Genet., 104:92-96.

Kesseli RV, Jain SK (1985) Breeding systems and population structure in *Limnanthes*. Theor. Appl. Genet., 71:292-299.

Kishore VK, Gandhi SD, Crane JM, Knapp SJ (2002) The meadowfoam genome and simple sequence repeat genome map (in preparation).

Kishore VK, Velasco P, Slabaugh MB, Knapp SJ (2002) The development of simple sequence repeat markers for DNA fingerprinting and high throughput genotyping in meadowfoam (in preparation).

Kloosterman AD, Budowle B, Daselaar, P (1993) PCR-amplification and identification of the human D1S80 VNTR locus: Amplification conditions, population genetics and application in forensic analysis. International Journal of Legal Medicine, 105:257-264.

Knapp SJ and Crane JM (1997) The development of self-pollinated inbred lines of meadowfoam by direct selection in open-pollinated populations. Crop Sci., 37:1770-1775.

Knapp SJ, Crane JM (1999) The development of meadowfoam as an industrial oilseed: breeding challenges and germplasm resources. In: *Proceedings of the Fourth national Symposium on New Crops* (ed. J.Janick). Wiley, New York, USA.

Kresovich S, Szewc-McFadden AK, Blick SM (1995) Abundance and characterization of simple sequence repeats (SSRs) isolated from a size-fractionated genomic library of *Brassica napus* L. (rapeseed). Theor. Appl. Genet., 91:206-211.

Kubik C, Meyer WA, Gaut BS (1999) Assessing the abundance and polymorphism of simple sequence repeats in perennial ryegrass. Crop Sci., 39:1136-1141.

Kudo N, Kimura Y (2001) Flow cytometric evidence for endopolyploidy in seedlings of some Brassica species. Theor. Appl. Genet., 102:104-110.

Lagercrantz U (1998) Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that *Brassica* genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. Genetics, 150:1217-1228.

Lande R (1979) Effective deme size during long-term evolution estimated from rates of chromosomal rearrangements. Evolution. 33: 234-251.

Lande R (1988) Genetics and demography in biological conservation. Science, 241:1455-1460.

Lande R, Schemske DW 1985. The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. Evolution, 39:24-40.

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics, 121:185-199.

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

Lander ES, Schork NJ (1994) Genetic dissection of complex traits. Science, 265:2307-2048.

Lebowitz RJ, Soller M, Beckman JS (1987) trait-based analyses for the detection of linkage between marker loci and quantitative trait loci in crosses between inbred lines. Theor. Appl. Genet., 73:556-562.

Levin DA (1978) The origin of isolating mechanisms in flowering plants. Evol. Biol., 11:185-317.

Lewis KR, John B (1963) Chromosome marker. J. and A. Churchill, London UK.

Lin JZ, Ritland K (1997) Quantitative trait loci differentiating the outbreeding *Mimulus guttatus* from the inbreeding *M. platycalyx*. Genetics, 146:1115-1121.

Litt M, Luty JA (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Amer. J. Human Genet., 44: 397-401.

Liu ZW, Jarret RL, Kresovich S, Duncan RR (1995) Characterization and analysis of simple sequence repeat (SSR) loci in seashore paspalum (*Paspalum vaginatum* Swartz). Theor. Appl. Genet., 91:47-52.

Lodhi MA, Ye GN, Weeden NF, Resich BI (1994) A simple and efficient method for DNA extraction from grapevine cultivars and Vitis species. Plant. Mol. Biol. Rep., 12:6-13.

Loridon K, Cournoyer B, Goubely C, Depeiges A, Picard G (1998) Length polymorphism and allele structure of trinucleotide microsatellites in natural accessions of *Arabidopsis thaliana*. Theor.Appl. Genet., 97:591-604.

Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer Associates Inc. Publishers, MA, USA.

Lyttle TW (1991) Segregation distorters. Annu. Rev. Genet., 25:511-557.

Mann SS (1975) Exclusive preferential transmission of an alien chromosome in common wheat. Crop Sci., 16:580-583.

Mantel NA (1967) The detection of disease clustering and a generalized regression approach. Canadian Research, 27:209-220.

Marshall DR, Brown AHD (1975) Optimum sampling strategies in genetic conservation. In: *Crop genetic resources for today and tomorrow*. (ed. OH Frankel and JG Hawkes) Cambridge University Press, Cambridge, England.

Mason CT (1952) A systematic study of the genus *Limnanthes*. Botanical Publications, 25:455-512.

Mather K (1953) The genetical structure of populations. Symp. Soc. Exp. Biol. 7:66-95.

McNeill CI (1983) Genetic differentiation, species relationships and the evolution of mating systems in plant genus Limnanthes, section Inflexae. Ph.D. Diss. Univ. Calif., Davis.

McNeill CI, Jain SK (1985). Genetic differentiation studies and phylogenetic inference in plant genus *Limnanthes* (section Inflexae). Theor. Appl. Genet. 66:257-269

Melaragno JE, Mehrotra B, Coleman Aw (1993) Relationship between endopolyploidy and cell size in epidermal tissues of Arabidopsis. Plant Cell 5:1661-1668.

Messmer MM, Melchinger AE, Herrmann RG, Boppenmaier J (1993) Relationship among early European maize inbreds: II. Comparison of pedigree and RFLP data. Crop Sci., 33:944-950.

Miller TE, Hutchinson J, Chapman V (1982) Investigation of a preferentially transmitted *Aegilops sharonenesis* chromosome in wheat. Theor. Appl. Genet., 61:27-33.

Minch E, Ruiz-Linares A, Goldstein D, Feldman M and Cavalli-Sforza LL (1997) MICROSAT (ver. 2.0d) A computer program for calculating various statistics on microsatellite allele data. Available at http://hpgl.stanford.edu/projects/microsat/.

Mithen R, Raybold AF and Giamoustaris (1995) Divergent selection for secondary metabolites between wild populations of *Brassica Oleracea* and its implications for plant-herbivore interactions. Heredity, 75:472-484.

Muller H (1883) The fertilization of flowers. Macmillan, London, UK.

Mumm RH, Dudley JW (1994) A classification of 148 U.S. maize inbreds: I. Cluster analysis based on RFLPs. Crop Sci., 34:842-851.

Nagl W (1978) Endopolyploidy and polyteny in differentiation and evolution. Towards an understanding of quantitative and qualitative variation of nuclear DNA in ontogeny and phylogeny. Elsevier, North-Holland Publ, Amsterdam.

Ohno Y, Tanase H, Nabaki T, Otsuka K, Sasaki T, Suzawa T, Morii T, Yamori Y, Saruta T (2000) Selective Genotyping with epistasis can be utilized for a major quantitative trait locus mapping in hypertension in rats. Genetics, 155:785-792.

Olufowote JO, Xu Y, Chen X, Park WD, Beachell HM, Dilday RH, Goto M, McCouch SR (1997) Comparative evaluation of within-cultivar variation of rice (*Oryza sativa* L.) using microsatellite and RFLP markers. Genome, 40:370-378.

Ornduff R (1971) Systematic studies of Limnanthaceae. Madrona, 21: 103-111.

Ornduff R, Crovello TJ (1968) Numerical taxonomy of Limnanthaceae. Amer. J. Bot., 55:173-182.

Ott, J. 1991. Analysis of human genetic linkage. *John Hopkins University Press*. Baltimore, Maryland.

Paterson AH, Damon S, Hewitt JD, Zamir D (1991) Mendelian factors underlying quantitative traits in tomato: comparison across species, generations and environments. Genetics, 127:181-197.

Plaschke J, Ganal MW, Roder MS (1995) Detection of the genetic diversity in closely related bread wheat using microsatellite markers. Theor. Appl. Genet., 91:1001-1007.

Powell W, Morgante M, Andre C, Hanafey C, Vogel J, Tingey S, Rafalski A (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol. Breed., 2:225-238.

Propach H (1934) Cytologische untersuchungen an Limnanthes Douglasii. Zeitschr. Zellforsch, 21:357-375.

Properi E, Giangare MC, Bottiroli G (1991) Nuclease induced DNA structural changes assessed by flow cytometry with the intercalating dye propidium iodide. Cytometry, 12:323-329.

Rae SJ, Aldam C, Dominguez I, Hoebrechts M, Barnes SR, Edwards KJ (2000) Development and incorporation of microsatellite markers into the linkage map of sugar beet (*Beta vulgaris* ssp.). Theor. Appl. Genet., 100:1240-1248.

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

Resende F (1937) Ueber die ubiquitat der sat-chromosomen bei den blutenpflanzen. Planta., 26:757-807.

Rieseberg LH, Baird SJE and Gardner KA (2000) Hybridization, introgression and linkage evolution. Plant Mol. Biol., 42:205-224

Rieseberg LH, Linder CR, Seiler GJ (1995) Chromosomal and genic barriers to introgression in Helianthus. Genetics, 141:1163-1171.

Ritland K (1984) A comparative study of floral and electrophoretic variation with life history variation in *Limnanthes alba* (Limnanthaceae). Oecologia, 63:243-251.

Ritland K, Jain SK (1984) A comparative study of floral and electrophoretic variation with life history variation in *Limnanthes alba* (Limnanthaceae). Theor. Appl. Genet., 63: 243-251.

Roder MS Korzum V, Wendehake K, Plaschke J, Tixier MH, Leroy P, Ganal MW (1998) A microsatellite map of wheat. Genetics, 149:2007-2023.

Roder MS, Plaschke J, Konig SU, Borner A, Sorrels ME, Tanksley SD, Ganal SW (1995) Abundance, variability and chromosomal location of microsatellites in wheat. Mol. Gen. Genet., 246:327-333.

Rodman JE, Soltis PS, Soltis DE, Sytsma KJ, Karol KG (1998) Parallel Evolution of Glucosinolate biosynthesis inferred from congruent nuclear and plastid gene phylogenies. Amer. J. Bot., 85:997-1006.

Rohlf, FJ (1993) *NTSYS-pc numerical taxonomy and multivariate analysis system.* Ver.2.02i Exster software, Setauket, New York.

Rongwen J, Akkaya MS, Bhagwat AA, Lavi V, Cregan PB (1995) the use of microsatellite DNA markers for soybean genotype identification. Theor. Appl. Genet., 90:43-48.

Saal B, Wricke G (1999) Development of simple sequence repeat markers in rye (*Secale cereale* L.). Genome, 42:964-972.

Saghai Maroof MA, Biyashev RM, Yang GP, Zhang Q, Allard RW (1994) Extraordinarily polymorphic microsatellite DNA in barley: species diversity, chromosomal locations, and population dynamics. Proc. Nat. Acad. Sci., 91:5466-5470.

Schug MD, Hutter CM, Wetterstrand KA, Gaudette MS, Mackay TFC, Aquadro CF (1998) The mutation rates of di-, tri- and Tetranucleotide repeats in *Drosophila melanogaster*. Mol. Biol. Evol., 15:1751-1760.

Schug MD, Mackay TFC, Aquordro CF (1997) Low mutation rates of microsatellite loci in *Drosophila melanogaster*. Nature Genetics, 15:99-102.

Smith CR, Bagby MO, Miwa TK, Lohmar RL, Wolff IA (1960) Unique fatty acids from *Limnanthes douglasii* seed oil: the C_{20} – and C_{22} – monoenes. Journal of Organic Chemistry, 25:1770-1774.

Smith JSC, Kresovich S, Hopkins MS, Mitchell SE, Dean RE, Woodman WL, Lee M, Porter K (2000) Genetic diversity among elite sorghum inbred lines assessed with simple sequence repeats. Crop Sci., 40: 226-232.

Smith MSC, Chin ECL, Shu H, Smith OH, Wall SJ, Senior ML, Mitchell SE, Kresovich S, Ziegle J (1997) An evaluation of utility of SSR loci as molecular markers in maize (*Zea mays* L.) comparisons with data from RFLPs and pedigree. Theor. Appl. Genet., 95:163-173.

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

Smulders MJM, Rus-Kortekaas W, Gilissen LJM (1994) Development of polysomaty during differentiation in diploid and tetraploid tomato (*Lycopersicon esculentum*) plants. Plant Sci., 97:53-60.

Sneath PHA Sokal RR (1973) *Numeric Taxonomy*. Freeman, San Francisco.

Snogerup S, Gustafsson M, Von Bothmer R (1990) Brassica sect. Brassica (Brassicaceae).I. Taxonomy and variation. Willenowia, 19:271-365.

Stebbins GL (1950) Variation and evolution in plants.Columbia, New York, USA.

Stebbins GL (1957) Self-fertilization and population variability in plants. American Naturalist, 91:337-354.

Stebbins GL (1974) Flowering plants. Evolution above the species level. Belknap Press of Harvard University press, Cambridge Massachusetts, USA.

Stebbins GL, Valencia J, Valencia RM (1946) Artificial and natural hybrids in the Gramineae, Tribe Hordeae I. *Elmus*, *Sitanion*, *Agropyron*. Amer. J. Bot., 33:338-351.

Szewc-McFadden AK, Kresovich S, Bliek SM, Mitchell SE, McFerson JR (1996) Identification of polymorphic, conserved simple sequence repeats (SSRs) in cultivated *Brassica* species. Theor. Appl. Genet., 93:534-538.

Takezaki N, Nei M (1996) Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. Genetics, 144:389-399.

Tang S, Yu JK, Slabaugh MB, Shintani DK, Knapp SJ (2002) Simple sequence repeat map of the sunflower genome. Theor. Appl. Genet. (submitted).

Taramino G, Tingey S (1996). Simple sequence repeats for germplasm analysis and mapping in maize. Genome, 39:227-287.

Temnykh S, Park WD, Ayres N, Cartinhour S, Hauck N, Lipovich L, Cho TG, Ishii T, McCouch SR (2000) Mapping and genome organization of microsatellite sequences in rice *Oryza sativa* L. Theo. Appl. Genet., 100:697-712.

Van OoijenJW, Maliepaard C 1996. MapQTL (tm) version 4.0: Software for the calculation of QTL positions on genetic maps. Plant Genome IV Abstracts. World Wide Web: http://probe.nal.usda.gov/.

Wang Z, Weber JL, Zhong G, Tanksley SD (19940 Survey of plant short tandem DNA repeats. Theor. Appl. Genet., 88:1-6.

Weber JL (1990) Informativeness of Human $(dC\text{-}dA)_n \cdot (dG\text{-}dT)_n$ polymorphisms. Genomics, 7:524-530.

Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Amer. J. Human Genet., 44:388-396.

Wheeler DL, Chappey C, Lash AE, Leipe DD, Madden TL, Schuler GD, Tatusova TA, Rapp BA (2000). Database resources of National Center for Biotechnology Information. Nucleic Acids Res., 28:4-10.

White M (1973) Animal Cytology and Evolution. William Clowes and Sons, London.

Williams JG, Kubelik AR, Livak KJ, Rafalski A and Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res., 18:6531-6535

Wu KS, Tanksley SD (1993) Abundance, polymorphism and genetic mapping of microsatellies in rice. Mol. Gen. Genet., 241:225-235.

Wyatt, R (1988) Phylogenetic aspects of the evolution of self-pollination. In L.D. Gottlieb and S.K. Jain 9eds.), Plant evolutionary biology, 109-131. Chapman and Hall, London, UK.

Yang GP, Saghai Maroof, MA Xu CG, Zhang Q, Biyashev RM (1994) Comparative analysis of microsatellite DNA polymorphisms in landraces and cultivars of rice. Mol. Gen. Genet., 245:187-194.

Yu (2001) Simple sequence repeat marker development and mapping in cultivated sunflower, *Helianthus annuus* L. Ph.D. thesis. Oregon State University, Corvallis, OR, USA.

Zamir D, Tadmor Y (1986) Unequal segregation of nuclear genes in plants. Bot. Gaz., 147:355-358.

Zamir D, Tanksley SD, Jones RA (1982). Haploid selection for low temperature tolerance of tomato pollen. Genetics,101:129-137.