ABSTRACT

LAWTON-RAUH, AMY LOUISE.
Molecular Evolution and Population Genetics of Duplicated Floral Regulatory Genes.  
(Under the direction of Michael D. Purugganan)

Duplicated genes are the building blocks for the origin and diversification of gene families and may play a pivotal role in organismal diversity. Many regulatory genes involved in key developmental pathways are derived from gene duplications and divergence processes underlying gene family expansion. Thus, studies of duplicated genes shed light on genome evolution, the origin and diversification of gene families, and the potential adaptive significance of gene duplications. This dissertation research examined the molecular evolution and population genetics of duplicated floral regulatory genes in two plant systems: Arabidopsis and the allopolyploid Hawaiian silversword alliance adaptive radiation.

In the Arabidopsis system, the molecular evolution of the ancestral APETALA3/PISTILLATA gene duplication and the more recent (within Brassicaceae) APETALA1/CAULIFLOWER duplication was analyzed among four taxonomic levels: between sibling Arabidopsis species (A. thaliana and A. lyrata), among several Brassicaceae species, and among distantly related species from two different plant families (Antirrhinum majus from Scrophulariaceae and Silene latifolia from Caryophyllaceae). Examination of all four loci uncovered contrasting dynamics of molecular evolution between these two sets of paralogous regulatory genes. The CAULIFLOWER (CAL) locus is evolving at a significantly faster rate than its paralog APETALA1 (AP1) and significant differences in substitional distributions exist along these genes, reflecting the possible adaptive differentiation of CAL from AP1. The APETALA3/PISTILLATA paralog pair does not have detectable differences in patterns of substitution. Taken together, these results suggest that altered rates and patterns of sequence evolution may lead to divergence in developmental functions of some paralogous regulatory genes.
In the Hawaiian silversword alliance, the molecular population genetics of duplicated \textit{APETALA1} and \textit{APETALA3/TM6} orthologs was investigated in three species from different Hawaiian sublineages (\textit{Argyroxyphium sandwicense}, \textit{Dubautia ciliolata}, and \textit{Dubautia arborea}). Comparisons of these duplicated orthologs (\textit{ASAPETALA1-A}, \textit{ASAPETALA1-B} and \textit{ASAPETALA3/TM6-A}, \textit{ASAPETALA3/TM6-B}) in \textit{A. sandwicense} and \textit{D. ciliolata} indicate that two regulatory gene homoeologs (\textit{ASAPI-1-A} and \textit{ASAPI-1-B}) appear to be evolving in a similar fashion while the other two homoeologs (\textit{ASAP3/TM6-A} and \textit{ASAP3/TM6-B}) have patterns of nucleotide diversity consistent with divergent evolutionary trajectories. This divergent evolution between two homoeologous regulatory genes suggests that selection may be partitioning the functional trajectories of these two gene copies.

A multilocus study of six genes (\textit{ASAPI-1-A}, \textit{ASAPI-1-B}, \textit{ASAP3/TM6-A}, \textit{ASAP3/TM6-B}, \textit{ASCAB9} and \textit{ASNAD1}) among all three Hawaiian silversword alliance species was also conducted. This study compared morphological divergence between two recently derived sibling species (\textit{Dubautia ciliolata} and \textit{D. arborea}) to the genetic divergence across six genes. These two species differ significantly in morphology; yet have a genetic distance similar to levels expected between intraspecific populations. The extent of genetic divergence and historical demographics of diversification was also examined among all three species. Haplotype divergence, likelihood ratio tests of exponential growth rate, and Bayesian coalescent simulations of migration and divergence across these genes suggest demographic patterns associated with the evolution of these adaptively-radiating species.
MOLECULAR EVOLUTION AND POPULATION GENETICS OF Duplicated FLORAL REGULATORY GENES

by

AMY LOUISE LAWTON RAUH

A dissertation submitted to the Graduate Faculty of North Carolina State University in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

DEPARTMENT OF GENETICS

Raleigh, NC

2003

APPROVED BY:

[Signatures]

Michael D. Purugganan, Thesis Advisor & Co-chair, Genetics

Trudy F. C. Mackay, Committee Co-chair, Genetics

Spencer V. Muse, Committee Member, Statistics

Dominique Robertson, Committee Member, Botany
DEDICATION

This dissertation is dedicated to my favorite fans, my nieces and my nephew: Joseph and Katherine (Fairman) and Jessica, Jillian, and Courtney (Timmer). Remember to believe in your inner voice, stay true to yourself, keep balance and creativity in your life, maintain a compassionate view of the world, and never be afraid to challenge yourself physically and mentally.
BIOGRAPHY

I was born in Hermann, Missouri (a small German settlement town on the bluffs of the Missouri River) in 1973 to two creative, supportive parents that have taught me much about the importance of seeking art in life, listening to the history of place, and the importance of appreciating a balanced life in the present. In 1980, we moved to St. Peters, Missouri where I attended Progress South Elementary School, South Middle School, and Fort Zumwalt South High School, all in the Fort Zumwalt School District in St. Charles County.

When I was about 6 years old, I discovered my inner scientist. I was over at my parents’ friend’s house during an airing of the PBS show “3-2-1 Contact”, a kids show that presented many different aspects of science. I think my eyes were fixated on the TV for that entire show! The difficult part for me was deciding what type of science that I wanted to pursue. When I was 10 years old, I wanted to be a part time biologist, part time artist (I was also fond of drawing and painting). Then, combining my love of adventure and passion for science, I became interested in astrophysics and aerospace engineering (designing and flying spaceships) and longed to be a part of the space program. I saved my allowance for several years to pay for my trip to the U.S. Space Academy at age 14. It was a wonderful experience and it helped me realize that I am motivated by basic research and that astronauts are true heroes, bravely accepting enormous risks with the hope of serving humanity through science.

After a year of engineering coursework at the University of Missouri, Columbia (affectionately known as ‘Mizzou’), I decided to concentrate on a curriculum combining Biology and minor concentrations in Chemistry and Anthropology. Spurred by an intriguing anthropology lecture on ‘Mitochondrial Eve’, I spent 3 summers and 2 academic years researching mitochondrial defects as a Howard Hughes and NSF-REU research intern in Dr. Kathy Newton’s maize mitochondrial genetics lab. Dr. Newton’s infectious enthusiasm and the hospitality of the maize genetics community taught me much about research life,
including fieldwork, benchwork, and communication skills. I also spent a year abroad at the University of East Anglia in Norwich, England through an E.C. Science exchange student program. While at U.E.A., I took an excellent course in plant biology taught jointly by U.E.A. and John Innes Institute professors that piqued my interest in plant adaptation and evolution. Needless to say, this interest has since grown into an obsession of sorts (a healthy one, I hope!).

In August 1997 I moved to Raleigh, North Carolina and entered the N.C. State University department of Genetics doctoral program under the direction of Dr. Michael Purugganan. As a graduate student, I learned much about professionalism and how to develop a research program while watching Michael’s lab triple in size and scope from a handful of people working on the molecular evolution of developmental regulatory genes to the incorporation of quantitative genetics, molecular population genetics, and genome-wide aspects of *Arabidopsis* evolution. In addition to exploring avenues of research life, I sought out opportunities to grow as a future educator by participating in workshops and programs involved with teaching at the university level, as I believe that public education is an essential contribution and is one of my ultimate goals. I have accepted a postdoctoral research scientist position in the Department of Genetics and Evolution at the Max Planck Institute of Chemical Ecology in Jena, Germany to work with Dr. Thomas Mitchell-Olds on aspects of population structure effects on molecular population genetic signatures.
EDUCATION

Honors B.S. (Biological Sciences), minors in Chemistry and Anthropology
University of Missouri-Columbia (May 1997)

RESEARCH EXPERIENCE

North Carolina State University, Graduate Research Assistant
Advisor: Dr. Michael D. Purugganan, Department of Genetics (1997-2003)
Title of thesis: Molecular Evolution and Population Genetics of Duplicated Floral Regulatory Genes

University of Missouri-Columbia, Undergraduate Research Intern
Advisor: Dr. Kathleen J. Newton, Dept. of Biology (Summer 1994, 1995-1997)
Maize Mitochondrial Genetics
1) NSF/REU Summer Science Research Internship, May – August 1994
   Investigation of malate dehydrogenase mutations in maize
2) Howard Hughes Undergraduate Research Fellowship/Internship, July 1995 – August 1996;
   Undergraduate Research Assistant, August 1996 – July 1997
   P2: an unstable maize mitochondrial genome induced by nuclear factor(s)

PUBLICATIONS


TEACHING EXPERIENCE

Graduate Teaching Assistant/Lecturer (advisor: Dr. Betty Gardner)
   General Genetics Laboratory (GN412) (1 credit)-1 class, 24 students (Fall 2000)
   Developed and implemented a new lab component: ‘Genetics of Flower Development’

Preparing the Professoriate (mentor: Dr. Betty Gardner)
   One of ten campus-wide awardees to participate in year-long teaching workshop series (2000-01)

Graduate Teaching Assistant (mentor: Dr. Ted Emigh)
   General Genetics Recitation (GN411P) (3 credits)-2 classes, 48 students (Spring 2000)

Undergraduate Teaching Assistant (mentor: Dr. Robin Kennedy)
   Plant Taxonomy (4 credits) 2-2 hour plant systematics labs weekly (Spring 1997)
AWARDS

A. Research
   Sigma Xi Grant-in-Aid of Research (May 2001)
   Professor Stanley Zimmering Prize for Outstanding Undergraduate Research,
   Biological Sciences (May 1997)
   Howard Hughes Medical Institute Undergraduate Research Fellowship (1995-96)
   NSF/REU Undergraduate Research Fellowship (Summer 1994)
   National Science Scholar, Missouri (1992-97)

B. Academic and professional development
   Graduate Professional Skills Development Workshop, selected participant (Oct. 2001)
   Gamma Sigma Delta Honor Society (1999 - present)
   Biological Sciences Divisional Honors Certificate (May 1997)
   University Honors College Scholar, Certificate (1992-97, May 1997)
   Golden Key National Honor Society (1995 - present)
   E.C. Science Study Abroad Fellowship
      (exchange student, University of East Anglia-Norwich, England (Sept. 1994 - June 1995))
   University of Missouri Curators Scholar (1992-97)

C. Teaching
   Outstanding Teaching Award (May 2001)
   Preparing the Professoriate Award (2000-01)

AFFILIATIONS:

   Society for the Study of Evolution (March 2001 - present)
   Genetics Graduate Student Association-President (May 1999 - May 2001)
   University Graduate Student Grievance Panel-Student Representative (Fall 1999)
   University Graduate Student Association-Genetics Representative (1998-99)
   Student Pugwash (2000 – present)

SELECTED PRESENTATIONS:

   Lawton-Rauh, A.*, Robichaux, R.H. and Purugganan, M.D.  Duplicated regulatory genes in the Hawaiian

   Lawton-Rauh, A.*, Robichaux, R.H. and Purugganan, M.D.  What happens to duplicated regulatory genes
   in an adaptive radiation? Oral presentation, Annual joint meeting of the Society for the Study of Evolution,

   Lawton-Rauh, A.*, Robichaux, R.H. and Purugganan, M.D.  Molecular evolution and population genetics
   of floral homeotic genes in the Hawaiian silversword alliance. Oral presentation, Annual joint meeting of
   the Society for the Study of Evolution, Society of Systematic Biologists, and the American Society of

   Lawton-Rauh, A.L.*, Robichaux, R.H. and Purugganan, M.D.  Molecular population genetics of duplicate
   *ASAPETALAI* loci among two closely-related Hawaiian silversword alliance species.

   Lawton-Rauh, A.L.*, Robichaux, R.H. and Purugganan, M.D.  Molecular population genetics of duplicate
   *ASAPETALAI* loci among two closely-related Hawaiian silversword alliance species.
ACKNOWLEDGMENTS

I am sincerely thankful for the experiences that have granted me amazing opportunities to explore ideas and pursue my dreams. I have been fortunate to have support and thoughtful contributions from many people throughout my life and would like to extend special thanks to the following people and organizations for making my educational endeavors possible:

My major thesis advisor, Michael Purugganan, for his encouragement and generous support to pursue my graduate degree, explore my wide-ranging interests, and for sharing valuable research and professional insight.

Other members of my advisory committee, Trudy Mackay, Niki Robertson and Spencer Muse, for their patience, encouragement, and thoughtful advice during the preparation of this dissertation.

Members of the Purugganan lab, especially Jane Suddith (the ‘only child’), Marianne Barrier, Jeni Cork, Sola Halldorsdottir (3 weeks of fieldwork in Hawaii), Ken M. Olsen (Glutinous Knock Me Over), Dave Remington, Kristin Shepard, and Mark Ungerer for encouragement, advice, and patience.

Rob Robichaux, the Hawaiian silversword alliance ‘guru’, for encouragement, inspiration, and patience during manuscript preparations, tireless fieldwork efforts, and for being a wonderful role model.

Elizabeth Friar, for input regarding my Hawaiian silversword alliance project as well as for Hawaii fieldwork memories involving many tiny measurements recorded at ‘Uncle Billy’s Hilo Bay Hotel’, bunkbeds at the Nature Conservancy field station and the Kapalaoa cabin both in Haleakala Crater State Park in Maui, and for surviving man-eating ferns to take a peek at some cool plants.

Co-conspiring fellow graduate students from the Department of Genetics, with extra special thanks to Michael Buck, Andrea Johnson, Arnar Palsson, Gretchen Geiger Thornsberry and Jennifer Ingram for friendship, support, and fond memories.

Jeff Thornsberry, Carlyn Buckler and Ed Buckler for friendship, encouragement, and priceless advice.

Betty Gardner, for being an excellent teaching mentor, sharing inspiring conversations, and for teaching me the ‘ins and outs’ of managing a lab-based genetics course.

The supportive Pholks at Phinished and my beloved 12 year old flame point Siamese kitty ‘Azure’.
My educational endeavors obviously did not begin with my graduate education, and I would like to pay a special thanks to the following people and programs that have contributed significantly to my academic and professional education prior to my university education:

Mrs. Nelson, my 6\textsuperscript{th} grade social studies teacher that taught me that ‘ethnocentrism’ is the fundamental reason for barriers between cultures and source of miscommunication ...an anthropological perspective which I strive to incorporate in my own communications as a scientist and as a member of society.

The American Youth Foundation/Danforth Foundation for selection and sponsorship to participate in the international ‘I Dare You!’ leadership conference in 1991.

The Fort Zumwalt South High Foreign Language club scholarship for participation in a high school youth exchange program that sent me to a Bavarian hog farm for a month in 1992 (I don’t have enough room to explain this here!).

During my undergraduate years, I was fortunate to receive a National Science Scholar fellowship (representing one of 18 recipients from the state of Missouri), an MU Curator’s Scholarship, an E.C. Science Study Abroad fellowship, several Thomas Simons essay scholarships (1992-1996), NSF-REU and Howard Hughes Undergraduate Research program fellowships, and the Prof. Stanley Zimmering Prize for Undergraduate Research. I am sincerely thankful to these programs, which contributed significantly to my undergraduate education.

Drs. Kathy Newton, Candace Galen, and Karen Cone for encouragement and advice while I was an undergraduate at the University of Missouri.

My forever friends that know me better than most people: Jon Wilkinson (my high school chemistry lab partner and ‘big brother’) and Jill Groboski (my college roommate, ‘maid of honor’ and very close friend) for their ‘celebrate at will’ approach to life and for their loving and honest friendship.

I am truly thankful for the endless encouragement and love from my family: my mom and dad (William and Cecelia Lawton), my sisters Jennifer (Fairman) and Michelle (Timmer), my brothers-in-law Arland (Fairman) and Michael (Timmer), and my special cheerleaders Joseph and Katie (Fairman), Jessica, Jillian, and Courtney (Timmer). You mean more to me than I could ever possibly express in words.

I am especially grateful to my husband Brad Rauh...my present, past and future, and my Best Friend. Brad, it is the depth of your unconditional love, immeasurable support, patience, and creative spirit that gives me a special strength and balance. I am still amazed that we found each other and am looking forward to the many adventures we have yet to share!
# TABLE OF CONTENTS

List of Tables ................................................................................................................... xiv  
List of Figures .................................................................................................................. xvi

PREFACE ................................................................................................................................. 1

CHAPTER 1 ............................................................................................................................... 4

**EVOLUTIONARY DYNAMICS OF DUPLICATED GENES IN PLANTS**  
Abstract ................................................................................................................................. 5  
Introduction ............................................................................................................................ 6  
Mechanisms and Occurrence of Gene Duplication ............................................................... 8  
Evolutionary Dynamics of Duplicate Genes Occur in Different Temporal Windows .......... 13  
Potential for Adaptive Evolution .......................................................................................... 21  
Research Implications of Duplicate Gene Copies ............................................................... 26  
Acknowledgments .................................................................................................................. 30  
Literature Cited ....................................................................................................................... 31  
Figures .................................................................................................................................. 40

CHAPTER 2 ............................................................................................................................... 43

**MOLECULAR EVOLUTION OF FLOWER DEVELOPMENT**  
Abstract ................................................................................................................................. 44  
Introduction ............................................................................................................................ 45  
Evolution of MADS-box Genes ............................................................................................ 46  
Timing of floral homeotic gene origins .................................................................................. 50
Evolution of floral gene expression ......................................................... 52
The microevolution of flower development ........................................... 55
Prospects ................................................................................................. 57
Acknowledgements ................................................................................ 58
Literature cited ........................................................................................ 59
Figures ....................................................................................................... 63

CHAPTER 3 .......................................................................................................... 65
ADAPTIVE RADIATION AND THE HAWAIIAN SILVERSWORD ALLIANCE

Introduction ................................................................................................... 66
  Interspecific hybrid ancestry
  Diversity and divergence along the Hawaiian archipelago
  Incomplete genetic divergence despite accelerated rate of evolution

The Hawaiian Silversword Alliance ................................................................ 69

Literature cited ............................................................................................ 74

Figure .............................................................................................................. 79

CHAPTER 4 .......................................................................................................... 80
PATTERNS OF MOLECULAR EVOLUTION AMONG PARALOGOUS
FLORAL HOMEOTIC GENES

Abstract ....................................................................................................... 81

Introduction .................................................................................................. 82

Materials and Methods ............................................................................. 86
  Isolation and Sequencing of Floral Homeotic Genes
  Data Analysis
CHAPTER 6 ................................................................. 150

SPECIATION IN AND ISLAND ADAPTIVE RADIATION: A MULTILOCUS STUDY OF THE HAWAIIAN SILVERSWORD ALLIANCE

Abstract ................................................................. 151

Introduction ............................................................ 152

Materials and Methods ................................................. 157
  Morphological data: measurements and analysis
  Nucleotide sequence data: sample collections and DNA extraction
  PCR amplifications and gene sequencing
  Sequence data analyses

Results ........................................................................ 163
  Dubautia ciliolata and D. arborea: morphological data analyses
  Intraspecific genetic variation
  Interspecific genetic divergence: A. sandwicense vs. each Dubautia species
  Interspecific genetic divergence: between sibling species
  Haplotype networks: A. sandwicense vs. each Dubautia species
  Haplotype networks: nuclear genes between sibling species
  Assessment of genome-wide signatures
  Estimates of population growth rate
  Estimates of gene divergence times and species divergence times
  Estimates of migration rates

Discussion ..................................................................... 172
  Divergence between D. arborea and D. ciliolata
  Signatures of population expansion in A. sandwicense, D. ciliolata and D. arborea
  High migration rates between D. arborea and D. ciliolata

Acknowledgements ....................................................... 179

Literature Cited ........................................................... 180

Tables ......................................................................... 183
LIST OF TABLES

CHAPTER 4: PATTERNS OF MOLECULAR EVOLUTION AMONG PARALOGOUS FLORAL HOMEOTIC GENES

1. Floral homeotic genes used in this study ................................................................. 103
2. Sequence distances between Arabidopsis thaliana and other Brassicaceae floral homeotic gene orthologs ................................................................. 104
3. Sequence distances between coding regions of Arabidopsis thaliana and distant floral homeotic gene orthologs ................................................................. 105
4. Distribution of coding region differences between functional domains .............. 106
5. Comparison of nonsynonymous substitutions between core and noncore domains . 107

CHAPTER 5: PATTERNS OF NUCLEOTIDE VARIATION IN HOMOELOGOUS REGULATORY GENES IN THE ALLOTETRAPLOID HAWAIIAN SILVERSWORD ALLIANCE (ASTERACEAE)

1. Molecular variation in ASAPETALA1 and ASAPETALA3/TM6 homoeologs (A and B) from Dubautia ciliolata subsp. glutinosa and Argyroxiphium sandwicense subsp. macrocephalum ................................................................. 141
2. Hudson-Kreitman-Aguade contingency test values for ASAPETALA1 and ASAPETALA3/TM6 homoeologs (A and B) ................................................................. 142

CHAPTER 6: SPECIATION IN AN ISLAND ADAPTIVE RADIATION: A MULTILOCUS STUDY OF THE HAWAIIAN SILVERSWORD ALLIANCE

1. Morphological trait means in D. ciliolata and D. arborea .................................. 183
2. ANOVA analyses of morphological traits in D. ciliolata and D. arborea ............ 184
3. Nucleotide variation for loci examined in this study ........................................... 186
4. AMOVA and fixation indices for loci examined in this study: *A. sandwicense* vs. 
   *D. arborea* and *D. ciliolata* .......................................................... 187

5. AMOVA and fixation indices for loci examined in this study: *D. arborea* vs. 
   *D. ciliolata* .................................................................................. 188

6. Parameter estimates and test statistics examined for demographic trends .......... 189

7. MDIV joint estimates of gene and species divergence times, migration rates, and 
   population-mutation parameters ......................................................... 191

8. Nucleotide sequences of *ASAPETALA1-A* haplotypes ................................. 192

9. Nucleotide sequences of *ASAPETALA1-B* haplotypes .................................. 193

10. Nucleotide sequences of *ASAPETALA3/TM6-A* haplotypes ......................... 194

11. Nucleotide sequences of *ASAPETALA3/TM6-B* haplotypes ......................... 195

12. Nucleotide sequences of *ASCAB9* haplotypes ........................................... 196
LIST OF FIGURES

CHAPTER 1: EVOLUTIONARY DYNAMICS OF DUPLICATED GENES IN PLANTS

1. Spectrum of historical relatedness between progenitor genomes suggest three polyploidization categories ................................................................. 40

2. Relationships among homoeologous genes derived from gene and genome-wide duplications ...................................................................................... 41

3. Possible functional fates of duplicate gene copies follow a hierarchy of trajectory ..... 42

CHAPTER 2: MOLECULAR EVOLUTION OF FLOWER DEVELOPMENT

1. Major relationships within the plant MADS-box regulatory gene family ............... 63

2. Box: The genetics of flower development ................................................................ 64

CHAPTER 3: ADAPTIVE RADIATION AND THE HAWAIIAN SILVERSWORD ALLIANCE

1. Time-calibrated phylogeny including the Hawaiian silversword alliance ............... 79

CHAPTER 4: PATTERNS OF MOLECULAR EVOLUTION AMONG PARALOGOUS FLORAL HOMEOTIC GENES

1. Structure of plant MADS-box proteins .................................................................... 108

2. Schematic diagram of the Arabidopsis thaliana flower developmental pathway and portions of the plant MADS-box gene family phylogeny ......................... 109

3. Phylogenies of APETALA1 and CAULIFLOWER, APETALA3, and PISTILLATA genes .............................................................................................................. 110

4. Amino acid site diversity profiles of paralogous floral homeotic proteins ............. 111
CHAPTER 5: PATTERNS OF NUCLEOTIDE VARIATION IN HOMOEEOLOGOUS REGULATORY GENES IN THE ALLOTETRAPLOID HAWAIIAN SILVERSWORD ALLIANCE (ASTERACEAE)

1. Gene maps and summary figures of nucleotide variation in *ASAPETALA1* homoeologs ................................................................. 144

2. Gene maps and summary figures of nucleotide variation in *ASAPETALA3* homoeologs ................................................................. 145

3. Statistical parsimony haplotype networks of *ASAPETALA1* homoeologs .......... 147

4. Statistical parsimony haplotype networks of *ASAPETALA3* homoeologs .......... 148

CHAPTER 6: SPECIATION IN AN ISLAND ADAPTIVE RADIATION: A MULTILOCUS STUDY OF THE HAWAIIAN SILVERSWORD ALLIANCE

1. Populations of endemic *Dubautia ciliolata* subsp. *glutinosa* and *D. arborea* sampled from Mauna Kea, Hawaii ................................................................. 197

2. Populations of endemic *Argyrixiphium sandwicense* subsp. *macrocephalum* sampled in Haleakala National Park, Maui ......................................................... 198

3. Statistical parsimony haplotype network of *ASAPETALA1-A* ................................ 200

4. Statistical parsimony haplotype network of *ASAPETALA1-B* ................................ 201

5. Statistical parsimony haplotype network of *ASAPETALA3/TM6-A* ...................... 202

6. Statistical parsimony haplotype network of *ASAPETALA3/TM6-B* ...................... 203

7. Statistical parsimony haplotype network of *ASCAB9* ........................................ 204

APPENDIX: SUPPLEMENTAL FIGURES FOR CHAPTER 6

A1. Means comparisons of the number of capitula per capitulescense ......................... 212

A2. Means comparisons of the number of florets per capitulum ................................. 213
A3. Means comparisons of receptacular bract length .................................................... 214
A4. Means comparisons of sepal length ................................................................. 215
A5. Means comparisons of corolla length ............................................................... 216
A6. Means comparisons of ovary length ............................................................... 217
A7. Means comparisons of leaf length ................................................................. 218
A8. Means comparisons of maximum leaf width .................................................. 219
A9. Statistical parsimony haplotype network and polymorphism table of
    *ASNAD1* alleles examined in this study ........................................................... 220
PREFACE

This dissertation research is an integration of several elements of evolutionary genetics: molecular population genetics, regulatory gene evolution, evolutionary dynamics of duplicated genes, and speciation in an island adaptive radiation. To address this array of topics, the following objectives are addressed in the research discussed in the proceeding chapters:

(1) Investigate the molecular evolution of duplicated floral regulatory genes at different temporal levels, including paralogous genes in Brassicaceae species (Brassica oleracea and the sibling Arabidopsis species A. thaliana and A. lyrata) and their orthologs in other plant species from different plant families (Antirrhinum majus in Scrophulariaceae and Silene latifolia in Caryophyllaceae).

(2) Study the molecular population genetic dynamics among duplicate floral regulatory genes in the allotetraploid Hawaiian silversword alliance adaptive radiation by examining duplicate copies (homoeologs) of ASPETALA1 and ASPETALA3/TM6 in Argyroxiphium sandwicense subsp. macrocephalum (A. sandwicense) and Dubautia ciliolata subsp. glutinosa (D. ciliolata).

(3) Employ a multilocus approach using duplicated genes to study the population genetics of speciation of an island adaptive radiation by examining six genes among A. sandwicense and several populations of Dubautia arborea and D. ciliolata.

As an orientation to the motivations behind these objectives, this dissertation is organized as a series of literature review chapters followed by a series of research chapters. Chapters one through three are literature reviews and chapters four through six are primary research chapters.
*Literature reviews*

The literature review in this dissertation has been divided into separate chapters because chapters one and two represent publications or submitted manuscripts. Chapter one is a review of the evolutionary dynamics of duplicated genes in plants that was recently submitted as an invited review for a special issue on ‘plant molecular evolution’ for the journal *Molecular Phylogenetics and Evolution* [Lawton-Rauh AL. (2003). Evolutionary dynamics of duplicated genes in plants. *Molecular Phylogenetics and Evolution*, submitted]. This chapter serves as the background for the study of duplicated gene evolution in chapters four and five.

Chapter two is a review of the molecular evolution of floral regulatory genes that was published in *Trends in Ecology and Evolution* [Lawton-Rauh AL, Alvarez-Buylla E, Purugganan MD. (2000). The molecular evolution of flower development. *Trends in Ecology and Evolution*. 15, 4, 144-149]. This chapter discusses the role that several floral homeotic genes play in plant development, including *APETALA1, CAULIFLOWER, APETALA3,* and *PISTILLATA* which are the focus of the first research objective discussed in chapter four. The Asteraceae orthologs of two of these genes (*ASAPETALA1* and *ASAPETALA3/TM6*) are the focus of objective two and chapter five, were utilized to address objective three, and were included in the analyses of chapter six.

Chapter three is an overview of the Hawaiian silversword alliance. This chapter discusses the ancestry, trait variation, and rapid evolution of this insular plant group. The role of selection and rapid speciation are also discussed and several questions are posed regarding the speciation dynamics of the Hawaiian silversword alliance. This discussion serves as the foundation for objective three, the organismal background for chapter five, and as the motivation for chapter six.
**Research chapters**

Chapters four through six are research summaries presented as manuscripts for publication. Chapter four addresses the first objective and consists of a manuscript published in the journal *Molecular Biology and Evolution* [Lawton-Rauh AL, Buckler ES, Purugganan MD. (1999). Patterns of molecular evolution among paralogous floral homeotic genes. *Molecular Biology and Evolution*. 16, 8, 1037-1045]. Chapter five focuses on objective two and is a manuscript that was recently accepted for publication in a special issue of ‘genes in ecology’ for the journal *Molecular Ecology* [Lawton-Rauh AL, Robichaux RH, Purugganan MD. (2003). Patterns of nucleotide variation in homoeologous regulatory genes in the allotetraploid Hawaiian silversword alliance (Asteraceae). *Molecular Ecology*. In press]. Chapter six addresses the third objective and is a manuscript that is being prepared for submission to the journal *Evolution*. 
CHAPTER 1

EVOLUTIONARY DYNAMICS OF DUPLICATED GENES IN PLANTS

ABSTRACT

Gene duplication, arising from region-specific duplication or genome-wide polyploidization, is a prominent feature in plant genome evolution. While most studies of plant genome evolution have focused on the distribution relationship among genes (namely relative chromosomal location, quantitative trait locus, and quantitative trait nucleotide mapping), the genomics explosion permits deeper analyses of the occurrence and mechanisms associated with gene and genome duplication. Understanding the mechanisms generating duplicate gene copies and the subsequent dynamics among gene duplicates is vital to evolutionary genetics. These investigations shed light on regional and genome-wide aspects of evolutionary forces shaping intra- and interspecific genome contents, evolutionary relationships, and interactions. This review discusses recent gene duplication analyses, focusing on the molecular and evolutionary dynamics occurring at different timescales following duplication and the potential adaptive significance of these events in plant evolution.
INTRODUCTION

Gene duplication is potentially the most prominent genomic precursor of evolutionary diversification. Detection and analysis of lineage-specific gene family expansion among yeast (*Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*), nematode (*Caenorhabditis elegans*), fruit fly (*Drosophila melanogaster*) and plant (*Arabidopsis thaliana*) genome sequences suggests that gene duplication has been rampant following the divergence of the major lineages of the eukaryotic crown group (Lespinet *et al.* 2002). All currently characterized eukaryotic genomes also exhibit evidence of very recent gene duplications; thus gene duplication continues to shape organismal diversity (Lynch 2002). Furthermore, gene duplication among all eukaryotic lineages has been implicated as a possible source for adaptive variation, including the diversification of *HOX* clusters in early vertebrates (Malago-Trillo 2001), the accelerated rate of regulatory gene evolution in catostomid fish species (Ferris & Whitt 1979) and the addition to multigene families such as the MADS-box gene family, which has played a major role in angiosperm floral and reproductive diversity (Purugganan *et al.* 1995; Kramer *et al.* 1998; Alvarez-Buylla *et al.* 2000; Albert *et al.* 2002).

The origins of most recently arisen genes are through duplication, chromosomal rearrangement and the subsequent divergence of pre-existing genes (Bennetzen 2002). Most genes are members of multigene families, indicative of the prevalence of gene duplication in the origin and formation of diverse gene functions (Wendel 2000). Studies in *Arabidopsis thaliana* (*A. thaliana*) indicate that 65% of all genes belong to gene families, 37.4% of which are members of families composed of at least five genes (Arabidopsis Genome Initiative 2000) (Yuan *et al.* 2002). Gene family members are drawn from a range of homologous genes including tandem duplicates, dispersed duplications, and genome-wide duplications (Yuan *et al.* 2002). While gene duplicates have been discovered across eukaryotic lineages, gene duplication appears to be most common in plants.
Although gene duplication can occur through several regional and genome-wide mechanisms (see below), most studies of duplicated gene evolution in plants have centered on theoretical aspects of duplicate gene fates, and the phenotypic effects, as well as genomic alterations, associated with polyploidization. The recent surge of sequence-based comparative analyses has contributed a significant amount of empirical knowledge regarding the functional trajectories of duplicate gene copies and possible molecular and evolutionary mechanisms that confer genetic differentiation among duplicate polyploid loci. These data also provide tools to investigate similar aspects of region-specific duplication mechanisms.

Seminal publications by several researchers including Ohno’s ‘Evolution by Gene Duplication’ (Ohno 1970) and Stebbin’s ‘Variation and evolution in plants’ (Stebbins 1950) established much of the current framework of discussion regarding the potential contribution of gene duplication in plant evolution. Several recent studies investigating the molecular mechanisms and evolutionary dynamics of duplicate genes have begun to address hypotheses espoused in these classic publications. These studies incorporate different aspects of gene duplication, including chromosomal and gene specific modifications, phenotypic shifts, the population dynamics of inter-locus interactions, long-term alternative functional fates among duplicated genes, and the potential adaptive significance.

Following a synopsis of the mechanisms and occurrence of gene duplication, this review examines recent studies of plant evolution by gene duplication at three different relative timescales: immediate features regarding the establishment and persistence of cytotypes (representing different gene copy numbers), interactions among cytotypes housing different gene copy numbers, and possible functional differentiation among gene duplicates. The potential adaptive significance and research implications of gene duplication are also discussed.
MECHANISMS AND OCCURRENCE OF GENE DUPLICATION

Duplicated genes arise via two primary mechanisms: (1) regional genomic events and (2) genome wide events (polyploidization). In plants, most duplicate gene copies are considered to be derivatives of polyploidization events, yet recent results indicate that regional events may be responsible for a significant number of duplicated genes. Therefore, this review includes the contributions of both regional and genome-wide gene duplication mechanisms to genomic and organismal diversity.

Regional duplication

Regional events include local duplications as well as dispersive mechanisms (Schmidt 2002). Local duplications include tandem duplications that occur through transposable element activity and replication slippage (Bennetzen 2002). Other regional duplication events are dispersive processes resulting in the transfer of genes or chromosomal segments to more distant chromosomal locations. Dispersive processes include single gene duplications from active transposable elements and duplications of chromosomal regions of various sizes through aberrant crossing over events (inversion, translocation, unequal recombination, non-reciprocal translocation) (Bennetzen 2002).

The frequency of regional gene duplication is difficult to assess, as occurrence of region-specific events requires extensive genome-wide information. Fortunately, the availability of the entire genomic sequence of *A. thaliana* has permitted a direct assessment of relatedness among genes and chromosomal blocks in this species. Analysis of composite open reading frames (cORFs) revealed 103 duplicated blocks ranging in size from 12.5 – 149.5 cORFs (Vision et al. 2000). Identification of collinear gene clusters using TBLASTX suggest that large duplicated segments account for 60% of the *A. thaliana* genome (Arabidopsis Genome Initiative 2000; Mitchell-Olds & Clauss 2002). A significant
proportion of *A. thaliana* genes (17%) are arranged as tandemly repeated segments (4140 genes in 1528 arrays), indicating that regional duplication and persistence occurs at a substantial rate in *A. thaliana*, and perhaps in most plant species genome (Arabidopsis Genome Initiative 2000; Mitchell-Olds & Clauss 2002).

Genome studies in rice (*Oryza sativa* spp. *japonica* cv. Nipponbare) have also revealed the prominence of duplicated genes. The long arm of chromosome 10L in rice contains a gene-rich region with a high frequency of locally duplicated genes, including 24 copies of a glutathione S-transferase gene (Yuan *et al.* 2002). Continued sequencing of additional plant species will undoubtedly reveal a great extent of regionally duplicated genes that occur through both dispersive and local events. It will be very interesting to learn the relative occurrence rates and distributions of regional duplication events among taxonomic lineages.

**Polyploidization**

A better-characterized mechanism of gene duplication in plants is genome-wide duplication (polyploidization). In a broad sense, polyploidization is the duplication of a genome occurring either within a species due to incomplete segregation of chromosomes during meiosis (autopolyploidy) or as a result of unreduced gametes during hybridization between species (allopolyploidy) (Figure 1).

Polyploidization is considered to be the primary source of duplicate loci in plants, with the highest ploidy levels up to 2n = 640 (approx. 80x) in angiosperms [the stonecrop *Sedum suaveolens* (Uhl 1978; Leitch & Bennett 1997; Otto & Whitton 2000)], and in land plants up to 2n = 1260 (approx. 84x) [the fern *Ophioglossum pycnostichum* (Love *et al.* 1977; Otto & Whitton 2000)]. Polyploidization is very common in plants and occurs at different frequencies among angiosperm families but is not a common feature of
gymnosperm genomes (Leitch & Bennett 1997). Estimates suggest that 40 – 70 % of land plants have histories of at least one relatively recent episode, with estimates up to 95 % among pteridophytes (Grant 1981; Masterson 1994; Wendel 2000; Eckardt 2001).

The most probable mechanism of polyploid genome formation is the fusion of unreduced gametes that contain diploid, rather than haploid, chromosomal complements and subsequent crossing with other such individuals (Pikaard 2001). Various types of polyploidization may be described depending on the relative historical relatedness between the progenitor genomes. For example, intraspecific polyploidization merges genomes of high historical relatedness and interspecific polyploidization in a wide taxonomic cross arises from genomes of low historical relatedness. Between these two extremes lie the various degrees of initial divergence between the merging genomes. The initial divergence between progenitor genomes is influenced by distribution and allelic diversity factors such as population differentiation and cross-ability between taxa of increasing phylogenetic distance.

The midpoint in the spectrum of historical relatedness between merging genomes is segmental allopolyploidization (Figure 1). Segmental allopolyploidization refers to the merging of genomes from two species that differ with respect to chromosomal segment arrangement. The progenitor genomes may have highly similar gene content along large chromosomal segments, but have differences that may be due to inversions and translocations (Stebbins 1971). The hybrids between these two species end up with reduced chromosomal pairing and irregular meiosis (Stebbins 1950, 1971).

Using coalescent-based sequence analysis of 14 pairs of duplicated loci to detect approximate divergence times, Gaut and Doebley have shown that the *Zea mays ssp. mays* (maize) genome is the result of two genome-wide divergence events occurring approximately 20.5 and 11.4 million years ago (Mya) (Gaut & Doebley 1997). These two dates correspond to a divergence time of approximately 20.5 Mya between the two diploid progenitor species and an allotetraploidization event approximately 11.4 Mya (Gaut & Le Thierry d'Ennequin
Inferring a segmental allopolyploidization event versus a genomic allopolyploidization event was due to coalescent model-based indications of 2 divergence dates, suggesting that the homologous loci from the two ancestral species were not distinct prior allopolyploidization.

Genome-wide duplication occurred frequently in angiosperm evolutionary history and continues to occur (Grant 1981; Masterson 1994) (Soltis & Soltis 1999). Several studies indicate that the majority of flowering plant genomes harbor remnants of multiple polyploidization events of various ages (Vision et al. 2000; Zhang et al. 2001; Mitchell-Olds & Clauss 2002). While under debate, several analyses in *A. thaliana* suggest that polyploidization may have taken place several times in the *Arabidopsis* lineage, including an early genome duplication event that occurred 100 – 200 Mya and from one to three subsequent polyploidization events (Ku et al. 2000; Vision et al. 2000; Mitchell-Olds & Clauss 2002). This early date suggests an ancient duplication predating the radiation of dicotyledonous plants (Ku et al. 2000) (Vision et al. 2000). Other analyses suggest only a single polyploidization event occurring in the *Arabidopsis* lineage. Further study is required to verify the approximate number of polyploidization events molding the *Arabidopsis* genome (Arabidopsis Genome Initiative 2000; Mitchell-Olds & Clauss 2002).

Relationships among duplicated genes are complex

Duplicate genes are characterized by highly similar (homologous) nucleotide and/or protein sequences, with some shared terminology between regional and polyploid duplicates. Several definitions are used to generalize the relationships among genes with homologous nucleotide and protein sequences (Figure 2). Paralogs are genes with homologous sequences that result from a gene duplication within specific clades, such as the duplication of *APETALAI* resulting in the *CAULIFLOWER* paralog restricted to the Brassicaceae (Purugganan 1997). Orthologous genes are the same gene copy found in different species,
such as the *Arabidopsis* gene *CAULIFLOWER* and its *Brassica* ortholog *BoCAULIFLOWER* (Kempin et al. 1995). Autopolyploids contain paralogs and orthologs. In addition to paralogs and orthologs, allopolyploid genomes contain homoeologs, which are duplicate loci derived from different parental genomes. The relationship among duplicate gene copies in polyploid species occurs as a continuous distribution of relatedness resulting from the time since duplication. Due to the spectrum of historical relatedness between hybridizing progenitor genomes in allopolyploids, it can be difficult to distinguish among orthologs, paralogs, and homoeologs (Barnes 2002).

Further complicating the historical relatedness between gene copies is the dynamic nature of genomes which can distort the footprints of past gene and genome duplication events. Plant genomes are dynamic, with complicated fluctuations between genome expansion and contraction (Bennetzen 2002; Wendel et al. 2002). Additionally, evidence suggests that genomic restructuring increases dramatically following polyploidization (Levin 1983; Ozkan et al. 2001; Shaked et al. 2001). These ‘coping’ mechanisms may occur after the ‘genomic shock’ of polyploidization and are possibly related to diploidization and nuclear maintenance (McClintock 1967; Levin 1983; Shaked et al. 2001). This increased genomic activity leads to the difficulty in discerning ancient duplication events due to the noise that arises from these genomic ‘coping’ mechanisms and the increased diversification among duplicate gene copies. This noise can make it impossible to determine the extent of initial similarities between some gene duplicates (particularly more ancient duplications).
EVOLUTIONARY DYNAMICS OF DUPLICATE GENES OCCUR IN DIFFERENT TEMPORAL WINDOWS

While gene duplication has often been discussed in terms of potential adaptive significance, direct evidence for a causal link between a significant signal of natural selection correlated with an adaptive phenotype and a specific gene duplication event is lacking (Wendel 2000). The primary reasons for this concern the required timeline for demonstrating the trajectory of a progenitor gene (or genome), a specific duplication event, and direct evidence for adaptive value and selective signature in the offspring species. This is largely due to the fact that these required elements occur in different temporal windows, thus making causal links among these factors challenging to connect. Recent studies have made great contributions to further our understanding of mechanisms occurring at these different temporal windows, but we have yet to directly link the entire timeline.

There are three different temporal windows to consider in terms of the dynamics of gene duplication: the establishment and persistence of cytotypes (short-term), the interactions among duplicate gene copies following cytotype establishment (intermediate-term), and the possible functional fates of duplicated genes (long-term).

Establishment and persistence of cytotypes

The vast majority of investigations into the potential mechanisms of plant establishment and persistence following gene duplication have focused on polyploidization. This reflects the difficulty of detecting the onset of regional (especially gene-specific) duplications within populations. With increased availability of genomic sequences in several plant species, this avenue of research into single-gene and regional duplication will be an important contribution to evolutionary genetics.
Unless newly arising polyploids are self-compatible, the establishment of polyploid plants within diploid progenitor populations is a frequency-dependent process (Rodriguez 1996). The greater the frequency of polyploid individuals, the greater the probability of successful fertilization between compatible cytotypes (diploid versus polyploid cytotypes). Intercytotypic crosses are inviable, thus the establishment of a new cytotype (the polyploid cytotype) in outcrossing species is highly dependent upon the frequency of polyploid individuals in the population for successful cytotype-appropriate crossing. The frequency-dependent nature of polyploid establishment rapidly decreases with increased rates of selfing. This decreased reliance on density-dependence suggests that self-compatible polyploid offspring can potentially invade new niches relative to their diploid progenitor(s), thus polyploids are often weedier and more invasive than their diploid relatives. Furthermore, it has been suggested that the two cytotypes must actually have striking ecological differences for coexistence of both cytotypes (Levin 1983). Thus, polyploidy is one possible mechanism promoting adaptive evolution.

As Otto and Whitton suggest, one assessment to support the hypothesis of promoted adaptive evolution by polyploidy would be a phylogenetic-based examination of multiple independent transitions to polyploidy to test for increased rates of speciation or morphological evolution (Otto & Whitton 2000). This would not provide definitive proof, however, but would serve as a scaffold to test the hypothesis of increased adaptive potential in polyploids more succinctly and directly.

*Interactions among duplicate gene copies*

The establishment of duplicate-gene cytotypes can result in operational heterozygosity and increased genetic variation, particularly if duplicate gene copies confer similar function (Pikaard 2001). High levels of segregating genetic variation can be maintained in polyploid species due to the merging of diversity from diploid progenitor species and subsequent segregation and recombination following polyploidization (Soltis &
Unstable habitats may provide polyploid plants opportunities to exploit their increased genetic variation (Stebbins 1971). Stable environments without new ecological niches are actually less adaptive for polyploid plants than for their established diploid counterparts. The increased variation in polyploids leads to competition favoring the established diploid plants that have persisted in the niches preceding the appearance of polyploid individuals. The scenario is often different in dynamic environments. When new ecological niches are exposed in rapidly changing environments, polyploids may contain new combinations of alleles that enable establishment in areas that other plants cannot grow. The novel combinations of alleles found in polyploids may therefore lead to increased tolerance to wider ranges of environmental conditions (Stebbins 1971).

Allopolyploidy potentially leads to greater alterations in gene regulation than autopolyploidy due to increased incompatibilities between merging genomes of lower historical relatedness (Comai 2000). Gene expression as well as epistatic and pleiotropic interactions within the genomes of each progenitor species confer phenotypes and aspects of life history that may differ substantially. The result is that allopolyploids have the capacity for expressing unique gene product combinations that confer novel phenotypes. These unique gene product combinations may result from the ‘permanent heterozygosity’ due to the inheritance of one allele set from both progenitor lineages (Pikaard 2001). This increased number of alleles potentially decreases the deleterious effects of inbreeding depression, permitting inbreeding and/or self-fertilization.

On the genome-wide level, early models of polyploidization suggested that the increased DNA found in polyploids results in a ‘buffering’ such that single alleles are less likely to be hit by mutation and recombination and that genetic redundancy increases this ‘buffer zone’ (Levin 1983). It was therefore thought that mutation and recombination are
less effective at constructing novel adaptive complexes in polyploids than in diploids. However, several recent studies demonstrate that some polyploid genomes are dynamic, with rather substantial intra- and intergenomic changes occurring rapidly following polyploidization (Song et al. 1995; Ozkan et al. 2001; Shaked et al. 2001).

Studies in synthetic polyploids of *Brassica* (Song et al. 1995) and allotetraploid wheat (Ozkan et al. 2001; Shaked et al. 2001) demonstrate that extensive genomic changes can occur within a few generations. These genomic alterations can result in reproducible progenitor genome-specific sequence elimination can begin as early as the first generation following allopolyploidization and be mostly complete by the second or third generation (Ozkan et al. 2001; Shaked et al. 2001). Many of these rapid widespread genomic alterations may arise from changes in DNA methylation patterns and may occur through progenitor genome-specific targeting mechanisms (Ozkan et al. 2001; Shaked et al. 2001). These rapid genomic changes may lead to restored fertility in some allopolyploids (Ozkan et al. 2001; Rieseberg 2001; Shaked et al. 2001). Other studies suggest that polyploidization is not necessarily always followed by immediate genomic changes, as discovered in *Gossypium* [Malvaceae, (Liu et al. 2001; Pikaard 2001)] and *Spartina anglica* [Poaceae, (Baumel et al. 2002)]. Taken together, these studies suggest that aspects of genome maintenance can vary among plant lineages, are not necessarily deleterious and may lead to increased genetic diversity.

Comparative chromosomal analyses using genetic mapping and fluorescence *in situ* hybridization (FISH) between natural allopolyploids and modern descendents of putative progenitors indicate that several types of chromosomal changes can occur following polyploidization (Comai 2000; Pikaard 2001). These changes include DNA sequence elimination, reciprocal inversions and translocations, and heterochromatin expansion. The genomic lability of polyploids suggests that substantial genome-wide changes can occur that may affect relative positions and linkage relationships among genes as well as the orchestration of epistatic and pleiotropic expression patterns of duplicated genes.
Indeed, a recent study by Lee and Chen indicates that progenitor genome-specific expression levels can be modified in polyploids. Using an amplified fragment length polymorphism-cDNA display approach (AFLP-cDNA), diploid progenitor plants (Arabidopsis thaliana and Cardaminopsis arenosa) and offspring allotetraploid Arabidopsis suecica plants were screened for differentially expressed progenitor genome-specific genes (Lee & Chen 2001). Ten genes were differentially expressed in A. suecica compared to A. thaliana and C. arenosa, with no observations of allopolyploid-specific expression. These results indicate that genes of diverse functions can be differentially expressed in allopolyploids and that gene silencing is not restricted to transposable elements thus can directly affect coding regions (Lee & Chen 2001). Furthermore, this study suggests that differential evolution between duplicated genes is a possible outcome of gene duplication.

Possible functional fates of duplicate genes

A large fraction of duplicated genes appear to persist for a longer duration than previously predicted (Otto & Whitton 2000). Estimates in maize indicate that approximately 72% of duplicate gene copies have escaped sequence elimination leading to gene loss (Ahn & Tanksley 1993; Gaut & Doebley 1997; Otto & Whitton 2000). Assuming that most duplicated genes are the result of polyploidization, this suggests that most duplicate genes can escape immediate loss during early reorganization events following genome-wide duplication. If both gene copies persist for a considerable amount of time following duplication, what happens to the functional relationship between duplicate gene copies over time?

At the gene level, there are several potential functional fates of duplicated genes: functional redundancy, subfunctionalization, neofunctionalization or pseudogene formation [Figure 3, (Ohno 1970; Walsh 1995; Force et al. 1999)]. If both gene copies retain the original function and are equally maintained such that one gene copy can substitute for
another, then both copies are functionally redundant. In a second scenario, both gene copies may accumulate compensatory mutations such that the combined activities of both genes are required for the original function, leading to subfunctionalization. In a third scenario, one gene copy may retain the original function and the second gene copy may accumulate mutations such that the second copy is no longer involved in the original function. Regarding the fate of the second gene copy, there are two possible outcomes of this third scenario: neofunctionalization or pseudogene formation. Mutational accumulation in the second gene copy may effectively alter gene function, leading either to incorporation of the gene into a different pathway or to a gain of novel function, leading to neofunctionalization. Alternatively, changes via mutational accumulation or epigenetic gene silencing may render the second gene copy nonfunctional, resulting in pseudogene formation.

A study in *Gossypium* indicates the potential for functional redundancy in plants. Analysis of ribosomal DNA tandem arrays (rDNA) in five allopolyploid species of *Gossypium* and their diploid progenitor species indicates that duplicate genes can be equivalently conserved through concerted evolution following allopolyploidization (Wendel *et al.* 1995). Investigation of the internal transcribed spacer regions (ITS1 and ITS2) and the 5.8S gene show that the rDNA arrays are very similar among all of the species investigated (both diploid and allopolyploid). These rDNA arrays are found in four separate chromosomal positions, with two loci from each progenitor genome. The homogenization of the four different arrays therefore suggests that these repeats were maintained by interlocus concerted evolution, possibly occurring through unequal crossing over events among these loci (Wendel *et al.* 1995; Leitch & Bennett 1997).

An example of subfunctionalization is the non-overlapping expression of two floral homeotic genes in maize (Mena *et al.* 1996; Lawton-Rauh *et al.* 2000). Based on initial mutational analyses of flower development, there are three central mutation classes involved with homeotic development described as A-class, B-class and C-class. C-class mutations, in particular, lead to loss of both stamen and carpel identity in the inner floral whorls of *A.*
In the developing maize flower, C-class floral homeotic function results from non-overlapping expression of two duplicate maize genes (zag1 and zmm2), indicating that C-class floral homeotic function in maize has been partitioned between these two gene duplicates (subfunctionalization). These two genes are differentially expressed in the developing inner whorls of maize florets. These results suggest that zag1 and zmm2 have subfunctionalized the floral homeotic C-class function. Furthermore, phylogenetic analysis suggests that this subfunctionalization occurred subsequent to the divergence of these two genes within Poaceae approximately 60 Mya, which is prior to the divergence of most gene pairs in maize (11-20 My) (Mena et al. 1996).

The signature of divergence between gene duplicates indicating deviating trajectories among duplicated genes is differential evolution. Molecular population genetic evidence for differential evolution between duplicate gene copies in plants has been primarily investigated in Gossypium (Cronn et al. 1999; Small et al. 1999; Small & Wendel 2000, 2002). Analysis of the relative rates of evolution at both nonsynonymous and silent sites among several members of the Gossypium Adh gene family, indicate that AdhC is evolving at a faster rate than AdhA (Small & Wendel 2000). In a recent study, Small and Wendel take this analysis to the molecular population genetic level by comparing the levels and patterns of nucleotide diversity between paralogous Adh loci (AdhA and AdhC) in allopolyploid Gossypium (Small & Wendel 2002). As expected, the nucleotide diversity levels of these loci differ, with AdhC diversity being greater than AdhA diversity, indicating that differential evolution between these two gene copies. Further analysis comparing loci two sets of homoeologous Adh loci suggests that this differential evolution between duplicate genes may be the result of differentiation between the progenitor genomes after merging during allotetraploidization (Small & Wendel 2002). This progenitor genome-specific divergence may not occur in all plant species, as shown in a recent study of homoeologous regulatory genes in members of the Hawaiian silversword alliance (see below) (Lawton-Rauh et al. 2003).
The alternative functional fates of duplicated genes may follow a hierarchy of states that depend upon the historical relatedness and initial similarity of the duplicate gene copies (Figure 3). The greater the initial structural and sequence similarity between the duplicate gene copies, the greater the probability of functional redundancy between both copies. Therefore, assuming that intraspecific diversity is lower than interspecific diversity, duplicate copies derived from intraspecific duplication (such as regional events or autopolyploidization) have a higher chance of conferring similar function than duplicates derived from interspecific duplication (allopolyploidization).

As suggested in Figure 3a, these alternative functional fates follow irreversible trajectories, such that functional similarity may not be regained after decreased functional co-dependence between gene duplicates. In this hierarchy, functional redundancy can be followed by subfunctionalization, neofunctionalization and pseudogene formation. A gene copy at the latter three states does not revert back to functional redundancy, unless rare back mutations occur. If duplicate gene copies are expressed as sub-components for the same function (subfunctionalization), then these gene copies can either further differentiate such that one copy is incorporated into an entirely new pathway (neofunctionalization) or is lost and becomes a pseudogene. After being incorporated into entirely different functions, a duplicate gene copy loses co-function with the other gene copy and continues along a separate path, possibly leading to pseudogene status. At the other end of this hierarchy, after a gene copy attains pseudogene status, it will no longer be involved with the original function and will be lost.

This hierarchical scheme is a general guideline regarding the potential evolutionary trajectories differentiating duplicate gene copies. At any point in this pathway, a duplicate copy may be lost due to genomic and chromosomal-mediated sequence elimination mechanisms. Additionally, the duration of each status can vary significantly and be influenced by many factors affecting fitness such as position in critical pathways and the potential adaptive significance for differential evolution between duplicate gene copies.
POTENTIAL FOR ADAPTIVE EVOLUTION

The increased rate of gene duplication compared with nucleotide sequence changes per gene involved with reproductive isolation suggests that regional gene duplication is similar in level to the effects of intragenic adaptive nucleotide changes as a mechanism for incompatibilities between species (Lynch 2002). Additionally, while no estimates have been reported regarding the influence of regional duplication events on plant divergence, polyploidization alone is estimated to account for ~ 2–4% of angiosperm and 7% of fern speciation events (Otto & Whitton 2000). Hence, polyploidy is potentially a principal mechanism of sympatric speciation in plants (Otto & Whitton 2000). Although no direct evidence has linked altered rates and patterns of diversity within genes conferring phenotypes associated with specific gene duplication events, studies suggest that genetic system and phenotypic alterations following gene duplication (particularly polyploidization) may be associated with increased diversification relative to single gene-copy progenitors (Comai et al. 2000) (Otto & Whitton 2000).

Potential adaptive divergence of regionally duplicated genes

One example of potential adaptive divergence among regionally duplicated genes is the recently described CBF/DREB1 tandem array in *A. thaliana* (Gilmour et al. 1998; Medina et al., 1999; Haake et al. 2002). Cis-acting and trans-acting elements, known collectively as CBF/DREB1, exist as three tandemly arrayed genes (CBF1, 2 and 3; the CBF1-3 tandem array) occurring within an 8.7 kb region on chromosome IV (Gilmour et al. 1998) (Medina et al. 1999) (Haake et al. 2002). The expression of these duplicated transcriptional activation genes is induced by low temperature and these genes subsequently activate many other low temperature-responsive genes (Seki et al. 2001). A fourth copy of the CBF/DREB1 genes (CBF4) has been reported that confers an alternative function, indicating functional divergence following duplication (Haake et al. 2002). While gene
expression of the three previously uncovered copies of CBF/DREB1 (the CBF1-3 tandem array) is activated under low temperatures, CBF4 exhibits increased expression under drought stress, but not cold stress.

Further analysis of CBF4 indicates that the divergence of this gene from the CBF1-3 tandem array is not due to selection of entirely new functions. Rather this divergence is due to diversification of promoter regulatory elements responding to different, but closely-related, environmental signals (Haake et al. 2002). CBF4 is located on a different chromosome (chromosome V) and phylogenetic analysis revealed that an ancestral gene duplication preceded the split between the CBF1-3 tandem array and CBF4. This gene duplication splitting the CBF1-3 tandem array and CBF4 gene lineages was probably mediated through segmental duplication, with subsequent tandem duplication generating the CBF1, 2 and 3 array on chromosome IV (Haake et al. 2002). The divergence between the CBF1-3 array and CBF4 is coincident with divergent functional trajectories between these two clades. Other regional gene duplications in A. thaliana that may be associated with increased adaptive potential include the diversification of 2-oxoglutarate-dependent dioxygenases, which are involved in conversions of several glucosinolate secondary metabolites (AOP1, AOP2, and AOP3 on chromosome IV (Kliebenstein et al. 2001)), and addition to multigene families such as the rbcS family (Sasanuma 2001).

Potential adaptive significance of polyploid gene duplicates

Polyploidization has also been shown to co-occur with phenotypic alterations. Due to the high incidence of polyploidy in crop plants, investigations of phenotypic shifts in polyploids has been intensely investigated in cultivated species and center on phenotypes with direct effects on yield and persistence (Hilu 1993). A number of reproductive traits have been found to differ sharply between some polyploid species and their diploid progenitors, such as the initiation and duration of flowering time, fertility, self-compatibility,
apomixis, germination, and organ size (Jackson 1976; Lewis 1980; Lumaret 1988; Segraves & Thompson 1999; Schranz & Osborn 2000; Cook & Soltis 2000). Several other studies revealed alterations in vegetative traits including drought tolerance, growth rate, and susceptibilities to arthropod infestation and fungal diseases (Jackson 1976; Lewis 1980; Lumaret 1988).

Shifts in flowering time variation have been documented in resynthesized *Brassica napus* polyploids (Schranz & Osborn 2000). In this study, *Brassica napus* polyploids were synthesized from reciprocal interspecific hybridization between *B. rapa* and *B. oleracea*. A single amphidiploid plant from each reciprocal cross was colchicine-treated and self-pollinated. The resultant progeny were then maintained as separate lineages through careful self-pollinations. Several subsequent generations were then assessed for number of days to flowering and leaf number. Analyses of these traits indicate that the sixth generation had a wide range of variation in days to flowering that differed significantly between the separately maintained lineages. Additional analyses of further generations and testcrosses indicate that this phenotypic variation is heritable. These results indicate that heritable alterations in temporal developmental variation can occur rapidly following polyploidization (Schranz & Osborn 2000).

Several phenotypic shifts that occur in polyploid species may contribute to niche separation and further reproductive isolation between polyploids and their diploid progenitors. With shifts in traits such as flowering time (hence fruiting time) polyploids may have shifted ecological connections to pollinators, seed dispersers, herbivores, predators, and pathogens by decreasing time of overlap between temporal schedules. Studies of polyploidy indicate no distinct polarities of change in response or performance relative to their diploid progenitors (Stebbins 1971), yet these studies illustrate a tremendous opportunity to investigate the mechanisms of genetic information storage, the nature of gene expression, and the affect of progenitor genomic components on offspring adaptability and persistence.
One naturally distributed plant system illustrating the potential adaptive lability conferred by polyploidization is the allopolyploid Hawaiian silversword alliance (Asteraceae, Heliantheae, Madiinae). A premiere example of plant adaptive radiation, the Hawaiian silversword alliance is a group of 30 species distributed on six of the eight main islands of the Hawaiian archipelago (Kauai, Oahu, Molokai, Lanai, Maui and Hawaii) (Carr 1985; Robichaux et al. 1990; Baldwin & Robichaux 1995; Baldwin 1997). All but five of the Hawaiian silversword alliance members exist as single-island endemics and exhibit a remarkable range of habitat preferences and morphological growth forms along the steep range of available habitats afforded by the Hawaiian island geological and ecological diversity (Carr 1985). The current distribution and morphological range of the Hawaiian silverswords is much more diverse than conditions inferred for the progenitor diploid North American tarweed lineages (Baldwin & Robichaux 1995), a predicted shift in polyploids relative to diploid progenitors (Stebbins 1971).

The Hawaiian silversword alliance species are allopolyploids derived from an interspecific hybridization event between two lineages of the North American tarweeds more than 5.2 ± 0.8 million years ago (Robichaux et al. 1990; Baldwin et al. 1991; Baldwin and Robichaux 1995; Baldwin 1997; Barrier et al. 1999). Thus, the Hawaiian silversword alliance species harbor two distinct gene copies (homoeologs) derived from separate North American tarweed lineages.

Molecular evolutionary analysis of several homoeologous floral regulatory genes (ASAPETALA1-A, ASAPETALA1-B; and ASAPETALA3-A, ASAPETALA3-B) in the Hawaiian silversword alliance indicate a significant increase in the rate of protein evolution in the Hawaiian silversword alliance species compared to extant North American tarweed species (Barrier et al. 2001). Whether this increased rate in evolution was propelled by genomic events related to allopolyploidization is not clear and will require further investigation of genome-wide signatures confirming possible chromosomal and gene-specific alterations subsequent to polyploidization. Unfortunately, this origin of this adaptive radiation is
estimated to be at least 5.2 ± 0.8 Mya (Baldwin & Sanderson 1998), thus too much time has passed to make justifiable comparisons between putative diploid progenitor lineages and the allopolyploid Hawaiian silversword alliance.

While the timeframe of the Hawaiian silversword alliance adaptive radiation makes diploid versus allopolyploid comparisons impossible, it does permit appropriate studies of longer-term molecular population genetics associated with duplicate gene copies. A recent study investigated the molecular population genetics of two sets of homoeologous floral regulatory genes (ASAPETALA1-A, ASAPETALA1-B; ASAPETALA3-A, ASAPETALA3-B) among three species representing two major lineages of the Hawaiian silversword alliance (Lawton-Rauh et al. 2003). Nucleotide diversity estimates and haplotype network analyses indicate differential evolution between two homoeologs (ASAP3-A and ASAP3-B), but not between the other two homoeologs (ASAP1-A and ASAP1-B) in recently derived species. These results suggest that differential evolution between homoeologous gene copies can occur in naturally distributed plants undergoing adaptive radiation.
RESEARCH IMPLICATIONS OF DUPLICATE GENE COPIES

Detection, identification and surveys of gene duplication have become a recent possibility due to the increased implementation of genomics tools, including refined comparative mapping and genome sequencing projects (Wilson et al. 1983; Reinisch et al. 1994; Moore et al. 1995; Lagercrantz & Lydiate 1996; Shoemaker et al. 1996; Bennetzen & Freeling 1997; Gaut & Doebley 1997; Gomez et al. 1998; Kellogg 1998; Lagercrantz 1998; Muravenko et al. 1998; Sossey-Alaouni et al. 1998; Brubaker et al. 1999; Paterson et al. 2000; Wendel et al. 2002). Fortunately, the consistent influx of molecular advances and practical applications will continue to increase the feasibility and precision of investigating the evolutionary impact of duplicated genes and genomes. Increased understanding of the dynamics of regional and genome-wide duplication will greatly improve the capacity of probing the genomic dynamics and relationships both among and between species.

Phylogenetic framework

A prerequisite for continued progress in expanding studies of the evolutionary dynamics of gene duplication to more species is a clear phylogenetic framework among clades under comparison. A well-understood phylogeny is required for inferring the directionality of genome size changes and correct gene relationships as well as to clearly assess genome-wide dynamics occurring within and between taxa over evolutionary time (Bennetzen & Kellogg 1997; Wendel et al. 2002). The establishment of phylogenetic relationships among taxa and genes as well as genomic elements from progenitor lineages permits stronger connections between the causality, occurrence, and effects of genome-wide events.
Identification of homologous relationships among duplicated genes

Following a well-established phylogenetic framework, appropriate identification of relationships among duplicate genes is a fundamental exercise, as analyses among mis-identified homologous genes may lead to erroneous interpretations. Because of the high occurrence of gene and genome duplication, this is a particularly important consideration for plant molecular evolutionary studies. Identification of homologous relationships is tricky, particularly in plant systems that are not genetically well characterized, due to the complexity of gene and genome duplication mechanisms and dynamics.

In addition to the possibilities of rapid genomic alterations mentioned previously, examination of duplicate genes requires special attention to several other aspects regarding the occurrence and dynamics of gene and genome duplication. Discerning relationships among duplicated genes can be further complicated by evolutionary histories comprised of multiple, nested duplication events and the possibility of several independent gene duplication events among closely related species.

Nested rounds of gene and genome duplication have occurred among several investigated species, even in presumably diploid species as mentioned above in *A. thaliana* (Ku *et al.* 2000; Vision *et al.* 2000; Mitchell-Olds & Clauss 2002). Genomic sequencing and mapping studies of taxa that were previously considered diploid species have revealed ancient rounds of polyploidization (Eckardt 2001) in soybean [*Glycine* (Shoemaker *et al.* 1996; Lee *et al.* 2001)], cotton [*Gossypium* (Muravenko *et al.* 1998)], and sorghum (Gomez *et al.* 1998). Further studies indicate that many angiosperm genomes have signatures consistent with multiple episodes of polyploidization with subsequent cycles of genomic rearrangement and gene loss at several points in time (Wendel 2000).

Another consideration in studies of duplicate gene copies among closely related taxa is the possibility of multiple independent origins. Investigation of *Tragopogon* (Asteraceae,
Compositae) using RAPD markers revealed recent, independent origins of allopolyploid derivatives that occurred within a small geographic distribution at high frequency over a very short timespan (the last 60 – 70 years). Two goatsbeard species (Tragopogon mirus and T. miscellus) formed independently 12 and 20 times, respectively, with some co-occurrences within single small towns (Soltis & Soltis 1991; Soltis et al. 1995; Cook et al. 1998; Soltis & Soltis 1999). Less exaggerated examples of multiple independent polyploid events include the five origins of Heuchera grossulariifolia (Saxifragaceae, gooseberry leaf alumroot) (Segraves et al. 1999; Soltis & Soltis 1999) and the 13 origins of Draba norvegica (Brassicaceae, Norwegian draba) (Brochmann & Elven 1992; Soltis & Soltis 1995; Soltis & Soltis 1999).

Prospects

The genomics explosion has made available a variety of tools and techniques that will continue to greatly influence plant evolutionary genetics, and comparative evolutionary genomics in particular. With the increasing number of entirely and partially sequenced plant genomes will come an amazing potential to address many genome-wide and gene-specific questions across species. This will result in a deeper understanding of genome-wide events and how such dynamics influence the origin and diversification of specific genes, potentially of confirmed functional interactions with other genes and phenotypic manifestations.

Several important questions regarding the molecular evolutionary dynamics of gene duplication remain particularly elusive, including the role that gene and genome duplication play in the evolutionary histories and distributions of naturally-distributed plant taxa, the molecular population dynamics associated with the maintenance and eventual fate of duplicated genes, and whether different gene classes (such as transcriptional regulators versus structural housekeeping genes) have distinctly different evolutionary rates or divergence patterns between gene duplicates.
Extending these questions to comparisons between various mechanisms of gene duplication will be very exciting. Several potential comparisons between regional and genome-wide gene duplication dynamics include the rates and probabilities of establishment, persistence and alternative functional fates between duplicated genes. Some polyploid-specific questions include the extent of progenitor-genome specific patterns and the distribution of the effects of relative historical relatedness among progenitor genomes on the probabilities of alternative functional diversification between duplicates. Another consideration is whether genomic stability in polyploids is affected by nuclear-cytoplasmic interactions between polyploid nuclear genomes and chloroplast and mitochondrial genomes (Gill 1991; Song et al. 1995; Leitch & Bennett 1997; Soltis & Soltis 1999; Wendel 2000).
ACKNOWLEDGMENTS

I am grateful to my thesis advisor, Michael Purugganan, for his guidance, support and helpful suggestions for this manuscript. I thank Rich Moore and Brad Rauh for insightful discussions. This research was supported by an NSF grant to Michael Purugganan and Robert Robichaux and a Sigma Xi Grant-in-aid of Research to ALR.
LITERATURE CITED


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 recombinations. Genetics 150, 1217-1228.


Figure 1. Spectrum of historical relatedness between progenitor genomes suggest three polyploidization categories
Figure 2. Relationships among homologous genes derived from gene and genome-wide duplications

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>B vs C and B’ vs C’</td>
<td>orthologs</td>
</tr>
<tr>
<td>B vs B’ and C vs. C’</td>
<td>paralogs</td>
</tr>
<tr>
<td>B B’ vs. C C’</td>
<td>homoeologs (in allopolyploids)</td>
</tr>
<tr>
<td>A vs B vs B’ vs C vs C’</td>
<td>homologous chromosomes, regions, segments, or genes</td>
</tr>
</tbody>
</table>
Figure 3. Possible functional fates of duplicate gene copies follow a hierarchy of trajectory.
CHAPTER 2

Molecular Evolution of Flower Development

Amy L. Lawton-Rauh, Elena R. Alvarez-Buylla† and Michael D. Purugganan

† Instituto de Ecologia, Laboratorio Genetica Molecular y Evolucion, Universidad Nacional Autonoma de Mexico, AP Postal 70-275, Mexico D.F. 04510, Mexico

This chapter consists of a paper published as:
ABSTRACT

Flowers, as reproductive structures of the most successful group of land plants, have been a central focus of study for both evolutionists and ecologists. Recent advances in unravelling the genetics of flower development have provided insight into the evolution of floral structures among angiosperms. The study of the evolution of genes that control floral morphogenesis permits us to draw inferences on the diversification of developmental systems, the origin of floral organs and the selective forces that drive evolutionary change among these plant reproductive structures.
INTRODUCTION

Flowers are among the most visible and spectacular products of evolution. Floral structures first appeared in the fossil record among the seed plants, as determinate, sporophyll-bearing shoots, sometime during the early Cretaceous ~130 million years ago (Mya) (Sun et al. 1998, Swisher et al. 1999); since then they have undergone considerable diversification in both form and function. These specialized reproductive structures display a stunning array of morphologies in extant flowering plant species, driven, in part, by selection associated with varied pollination and dispersion strategies. In recent years, studies of floral morphology in an evolutionary context have been buoyed by genetic analysis of developmental pathways underlying inflorescence and flower morphogenesis (Yanofsky 1995; Weigel 1995). Isolation and characterization of loci known to participate in the genetic control of floral development have led to the formulation of models describing how floral regulatory genes interact with each other to pattern development in the angiosperm flower, including the classic ABC genetic model of floral organ identity (Box 1) (Weigel 1995). The genetic control of flower development is controlled by relatively few regulatory genes, which has provided evolutionary biologists with new opportunities to dissect the molecular basis of evolutionary change in plant reproductive morphology (Baum 1998; Purugganan 1998).
EVOLUTION OF MADS-BOX GENES

Most floral regulatory genes that have been identified encode various sequence-specific DNA-binding transcriptional activators, including homeodomain (BEL1) (Reiser et al. 1997), zinc-finger (SUPERMAN) (Sakai et al. 1995) and novel regulatory proteins (APETALA2 and LEAFY) (Jofuku et al. 1994, Weigel et al. 1992). However, many of these loci belong to the eukaryotic MADS-box gene family whose products are characterized by the presence of the highly conserved 57-amino acid DNA-binding MADS domain (Riechmann & Meyerowitz 1997). In Arabidopsis thaliana (thale cress), at least 47 MADS-box sequences are known (E.R. Alvarez-Buylla et al., unpublished), including floral homeotic genes such as AGAMOUS (AG), APETALA3 (AP3), PISTILLATA (PI) and several AGAMOUS-LIKE genes (Ma et al. 1991; Rounsley et al. 1995). Given the large number of MADS-box genes involved in floral development, most studies on the molecular basis of floral developmental evolution have focused on these regulatory loci.

The first plant MADS-box genes identified encoded proteins that shared a common structure consisting of four separate domains. These included: the highly conserved DNA-binding MADS domain at the amino-terminus; a moderately conserved K domain, which probably forms a coiled-coil structure and participates in protein–protein interactions; a weakly conserved intervening region linking the MADS and K domains; and a poorly conserved carboxy terminal domain that might function as a transactivation domain. The K domain is absent in all nonplant MADS-box genes and was thought to be a specific feature of these proteins only in plants. However, a recent study suggests that several plant MADS-box genes found at the base of the gene phylogeny do not share this stereotypical structure, and that the K domain appears to have evolved at or around the time of the major diversification of the gene family (E.R. Alvarez-Buylla et al., unpublished).

Early phylogenetic studies of MADS-box genes isolated from more than 19 plant species provided preliminary glimpses into patterns of diversification and evolution of
developmental function in this important regulatory gene family (Purugganan et al. 1995; Purugganan 1997; Theissen et al. 1996). Molecular evolutionary analyses indicate that major duplication events within the plant MADS-box gene family resulted in the establishment of at least four monophyletic floral homeotic gene groups: AGAMOUS (AG), APETALA3 (AP3), PISTILLATA (PI) and APETALA1/AGAMOUS-LIKE9 (AP1/AGL9) (Purugganan et al. 1995; Purugganan 1997; Theissen et al. 1996) (Fig. 1). Genetic and expression analyses indicate that members of a floral homeotic gene group tend to share similar developmental functions in flower and inflorescence morphogenesis (Purugganan et al. 1995; Purugganan 1997; Theissen et al. 1996), thus reflecting high conservation among evolutionarily related regulatory genes.

Members of the AGAMOUS (AG) group include the Arabidopsis AGAMOUS (AG), Antirrhinum (snapdragon) PLENA, and Zea mays (maize) zag1 and zmm2 loci; all are C-function genes involved in stamen and carpel development (Box 1). The Arabidopsis AG gene and its orthologs in other angiosperm species are expressed specifically in the floral reproductive organs (Purugganan et al. 1995), while gymnosperm orthologs, such as the Gnetum GMM3 locus, are expressed in the pollen- and ovule-producing structures of the strobili (Winter et al. 1999). The expression patterns of AG group members at different taxonomic levels suggest that these MADS-box genes diversified from their ancestral function to control reproductive organ differentiation. This diversification led to present-day C-function genes, which compartmentalize reproductive structures from sterile, nonreproductive tissues in the developing sporophyll. Interestingly, several genes in the AG clade (including AGL1 and AGL5) are expressed in derived angiosperm-specific structures, such as carpels and fruits.

Conservation of function among evolutionarily related MADS-box genes is observed in the AP3 and PI group genes, which share B-class floral homeotic functions in petal and stamen differentiation (Jack et al. 1992). These two gene groups appear to be sister to one another and arose via duplication from a single ancestral gene (Purugganan et al. 1995;
In extant higher eudicots, such as Antirrhinum and Arabidopsis, petal and stamen differentiation requires loci from both the AP3 and PI gene groups (Jack et al. 1992). The ancestral seed plant B-function locus might have served as a sex differentiation gene distinguishing male versus female reproductive organs in sporophylls (Purugganan 1998). For example, angiosperm B-function genes specify stamen, but not carpel, identity in flowering plants (Jack et al. 1992), and gymnosperm loci related to AP3 and PI [Gnetum GMM2 (Winter et al. 1999) and Picea abies (Norway spruce) DAL11–DAL13 (Sundstrom et al. 1999)] also are expressed in male, but not female, reproductive organs.

Although members of the AP3, PI and AG gene groups exhibit general conservation of developmental function, the AP1/AGL9 group is more diverse and contains genes expressed in a wider range of plant tissues. The AP1/AGL9 clade contains more members than the AP3, PI and AG floral homeotic gene groups and the basal (earliest diverging) loci in the AP1/AGL9 clade are expressed in both leaves and reproductive structures (Ma et al. 1991, Rounsley et al. 1995). In addition to floral meristem identity, APETALA1 is an A-class floral homeotic gene, which specifies sepal and petal development. However, the expression of other genes in the AP1/AGL9 group is not restricted to reproductive structures; AGL3 is expressed in vegetative structures of the inflorescence shoot (Ma et al. 1991), and AGL8 (FRUITFULL) coordinates tissue growth during fruit morphogenesis and is expressed in leaves (Gu et al. 1998).

Other plant MADS-box loci with diverse expression patterns do not belong to these floral homeotic gene groups (Rounsley et al. 1995) (Fig. 1). Phylogenetic analyses indicate that these non-floral homeotic group loci represent the most basal members of the plant MADS-box gene family (Theissen et al. 1996). Several of these loci are expressed in inflorescences, although others are expressed in vegetative and/or embryonic structures of plants. For example, the Arabidopsis AGL15 gene is expressed in embryos (Rounsley et al. 1995) and members of two other distinct, monophyletic MADS-box gene groups are
expressed specifically in roots (E.R. Alvarez-Buylla et al., unpublished). The basal position of many of these loci suggests that plant MADS-box genes might originally have served to regulate vegetative and/or embryonic development, and subsequent duplications resulted in the derivation of new genes that control reproductive development in land plants (Purugganan 1998).

Duplications within the MADS-box regulatory gene family, which gave rise to the major floral homeotic gene groups, appear to have occurred ~285 Mya; however, floral homeotic functions continue to diversify at more recent timescales. For example, the ABC model of flower development predicts that C-class mutations lead to loss of both stamen and carpel identity in the inner floral whorls. Studies of the duplicate maize loci zag1 and zmm2 suggest that C-class floral homeotic function in this grass species results from nonoverlapping expression of these two loci in the maize flower (Mena et al. 1998). zag1 and zmm2 are expressed specifically in carpels and stamens, respectively, suggesting that these two genes have partitioned the floral homeotic C function since they last shared a common ancestor within the grass family ~60 Mya. The recent partitioning of C-class floral homeotic functions within some grass groups illustrates the evolutionary lability of the genetic mechanisms governing fundamental developmental programming among flowering plant taxa.
TIMING OF FLORAL HOMEOTIC GENE ORIGINS

The timing of duplication events that led to the floral homeotic genes is of great evolutionary interest: are these floral developmental genes specific to flowering plants or did they pre-date the origin of flowers? Many MADS-box genes isolated from the gymnosperms *P. abies* (Norway spruce) (Sundstrom *et al.* 1999, Tandre *et al.* 1995), *P. mariana* (black spruce) (Rutledge *et al.* 1999) and *Gnetum gnemon* (melindjo) (Winter *et al.* 1999) are orthologous to known angiosperm floral homeotic loci and are members of previously identified floral homeotic groups (Purugganan 1997; Winter *et al.* 1999; Rutledge *et al.* 1999). Phylogenetic analyses demonstrate that floral homeotic gene groups observed in angiosperms pre-date the divergence of flowering plants and gymnosperms ~285 Mya (Purugganan *et al.* 1995; Purugganan 1997) (Fig. 1). Indeed, molecular clock studies suggest that the divergence of the four major floral homeotic gene groups might have occurred as early as 486 +/- 45 Mya, during the Ordovician, and that the establishment of these genes coincided with the rise of land plants (Purugganan 1997). However, these molecular clock estimates assume evolutionary rate homogeneity among MADS-box loci; any significant, undetected temporal differences in substitution rates could distort estimates. If correct, however, these early estimates of MADS-box gene group divergences suggest that the major diversification of floral regulatory loci might be associated with the evolution of more elaborate and specialized reproductive morphologies during the early evolution of land plants (Theissen *et al.* 1996; Purugganan 1997).

More recent molecular analyses, which include MADS-box genes from two pteridophyte (fern) genera (*Ceratopteris* and *Ophioglossum*) (Munster *et al.* 1997; Hasebe *et al.* 1998), suggest that the establishment of the floral homeotic gene groups occurred more recently than the dates derived from molecular clock estimates. These studies also indicate that although floral homeotic gene groups predate seed plants, the gene groups arose after the separation of seed plants and ferns. However, the isolated pteridophyte loci might simply represent orthologs to more basal, nonfloral homeotic gene group lineages and indicate that
pteridophyte orthologs to angiosperm floral genes have yet to be isolated. Alternatively, orthologs to the angiosperm floral homeotic loci might have been lost in pteridophytes after the divergence of ferns and seed plants. However, to date, all studies agree that distinct floral regulatory gene groups were present in ancestral seed plant taxa before the evolution of flowers, probably to control reproductive organ differentiation, and that these developmental loci were co-opted to control floral morphogenesis when flowering plants evolved. Detailed genetic studies in other basal land plant taxa, including bryophytes (mosses, liverworts and hornworts) and lycophytes, could shed light on the functions of these regulatory loci in species whose reproductive structures are distantly related to angiosperm flowers.
EVOLUTION OF FLORAL GENE EXPRESSION

Comparative gene expression studies between the families Brassicaceae (Arabidopsis), Scrophulariaceae (Antirrhinum) and Solanaceae (Petunia) indicate strong conservation of floral developmental gene functions across broad taxonomic levels. This functional conservation can be used to determine reproductive morphological homologies between distantly related angiosperm taxa, including organs representing unusual floral innovations. If we assume that the expression of a particular regulatory gene is both necessary and sufficient to control development of a particular organ in one or more model plant species, then expression patterns of the gene might identify homologous structures in diverse taxa. Thus, expression of orthologous MADS-box genes between distant species might allow us to identify structural homologies among derived floral organs and to examine the extent of developmental gene function conservation across angiosperm taxa. However, the use of such molecular expression markers should be used with caution because evolution (including co-option) of gene functions across distant lineages might lead to erroneous conclusions in assigning morphological homologies (Wray & Abouheif 1998). Nevertheless, initial use of floral homeotic genes to examine homologies indicates that this approach might offer additional evidence supporting the identification of morphological homologies among derived flower organs.

One example comes from Asteraceae species, where flowers display an array of bristles (pappus), which surround the corolla and serve as a seed dispersal aid. A longstanding debate is whether the pappus, which is positionally homologous to the calyx, is indeed a true calyx. A recent study of floral homeotic gene orthologs from Gerbera hybrida supports the argument that a pappus is, in fact, a true calyx (Yu et al. 1999). Transgenic studies with Gerbera indicate that reduction of expression of either the AG ortholog gaga2 or the AP3 ortholog gglo1 is accompanied by the transformation of carpels and petals, respectively, to pappus-like structures. Additionally, ectopic expression of gglo1 leads to the
replacement of a pappus with petaloid structures in the first whorl. These results are consistent with the hypothesis that pappus bristles are modified sepal structures.

Expression and genetic analyses of MADS-box genes also have been used to address questions regarding the evolution of floral organs in monocots (Schmidt & Ambrose 1998). For instance, the lodicules are enigmatic floral structures in grasses that have been described alternatively as modified petals or reduced staminodes. Another persistent question is whether the palea and lemma, inner bract-like organs in grass flowers, are homologous to eudicot sepals. In both cases, studies in maize (Schmidt & Ambrose 1998) and *Oryza sativa* (rice) (Kang et al. 1998) have begun to provide molecular genetic clues to morphological homologies. Mutants in the maize *AP3* ortholog (*silky1*) (Schmidt & Ambrose 1998) and in the rice *PI* ortholog (*OSMADS4*) (Kang et al. 1998) exhibit phenotypes remarkably similar to B-class homeotic mutations in eudicots, thus resulting in the development of carpelloid structures in whorl 3 and the replacement of lodicules with palea-like organs. Together, these investigations provide strong evidence that lodicules are homologous to eudicot petals and that paleae are homologous to eudicot sepals.

Recent molecular studies also have provided evidence supporting previous morphological investigations indicating that petals evolved independently several times during angiosperm evolution (Kramer et al. 1998; Kramer & Irish 1999). Morphological studies of flowers suggest that angiosperm petals are derived from either stamens (andropetals) or sepals and other sterile subtending organs (bracteopetals). Bracteopetals are distributed within the Magnoliid dicot orders Magnoliales, Piperales and Aristolochiales, but andropetals have evolved many times within lower eudicots, and at least once at the base of higher eudicots and monocots (Takhtajan 1991). The independent origins of petals are supported by studies of *AP3* and *PI* homologs from several lower eudicot species whose petals are derived from bracts and, therefore, are not homologous to andropetals in *Arabidopsis* and *Antirrhinum*. Among *Arabidopsis*, *Antirrhinum* and other higher eudicots, the B-function genes *AP3* and *PI* are expressed in petals throughout organ
development. By contrast, orthologs to these loci in Ranunculids (lower eudicots) are expressed only weakly in petal primordia and progressively decrease in expression upon further differentiation of petals (Kramer & Irish 1999). The dramatic differences in expression patterns of these loci suggest that petal identity in higher and lower eudicots is controlled by different loci, and provides strong molecular evidence supporting multiple independent origins of perianth organs among angiosperms.
THE MICROEVOLUTION OF FLOWER DEVELOPMENT

The diversification of floral developmental patterns at the macroevolutionary level must originate from molecular variation present within species. A comprehensive understanding of the evolutionary dynamics of flower development thus requires investigation of the evolution of developmental loci at the population level or between closely related species. One approach has involved examining domesticated plant species, such as maize (Wang et al. 1999) and Brassica oleracea (the cole crops) (Kempin et al. 1995), which both display clear divergence in within-species floral developmental programs arising from artificial selection by early farmers. For example, the evolution of the domesticated cauliflower (B. oleracea spp. botrytis) appears to be associated with mutations in the MADS-box floral meristem identity genes CAULIFLOWER and APETALA1 (Kempin et al. 1995, Lowman & Purugganan 1999). The dramatic inflorescence architectural differences between domesticated maize and its teosinte relatives also appear to arise, in part, from selection on the teosinte-branched1 (tb1) gene, a maize locus that belongs to the same regulatory gene family (the TCP gene family) as the Antirrhinum floral symmetry gene CYCLOIDEA (Wang et al. 1999). Molecular population genetic analysis of tb1 demonstrates that the evolution of the unique inflorescence architecture found in maize is associated with positive selection and adaptive divergence on the promoter of this developmental regulatory locus.

Although studies using domesticated species are instructive, the availability of sufficient genetic variation in developmental regulatory loci for adaptive diversification within natural populations remains unclear. At least one species, Clarkia concinna (pink ribbons), has been shown to exhibit a high frequency of natural floral homeotic conversion of petals to sepals in a wild population, as a result of the segregation of a single floral homeotic gene, BICALYX (Ford & Gottlieb 1992). Low-frequency natural variation in floral symmetry within Linaria vulgaris (common toadflax) also has been documented since the time of Linnaeus. In this case, the occurrence of radially symmetric flowers among the bilaterally...
symmetric (zygomorphic) flowers of this species has been attributed to epigenetic methylation changes in the *CYCLOIDEA* locus (Cubas et al. 1999).

Population genetic studies of floral homeotic genes in the wild weed *A. thaliana* indicate that the *CAL* (Purugganan & Suddith 1998), *AP3* and *PI* (Purugganan & Suddith 1999) genes harbor considerable within-species diversity at the molecular level. These three loci display elevated levels of intra-specific amino acid diversity in protein sequence, as well as evidence of non-neutral evolution. However, the levels and distribution of allelic variation at these loci do not appear to be controlled by recent episodes of adaptive selection, but are shaped largely by demographic forces operating on this selfing species (Purugganan & Suddith 1999). Nevertheless, there is evidence that molecular variation at some of these loci can result in differences in floral developmental functions. For example, naturally occurring alleles at the *CAL* locus can be distinguished by their differential capacities to direct floral meristem development (Kempin et al. 1995; Purugganan & Suddith 1998). The significant within-species molecular diversity in floral developmental genes, some of which might result in functionally distinguishable wild alleles, provides genetic material for selective forces to operate and potentially leads to macroevolutionary diversity in floral developmental programs.
PROSPECTS

The recent wealth of data on the developmental genetics of flower and inflorescence morphogenesis have provided the impetus for new studies on the underlying molecular mechanisms of floral evolution. However, major gaps in our understanding still remain; for example, we still have little information on the function and evolution of orthologs of the eudicot floral homeotic genes in many basal land plant taxa, including the bryophytes, other pteridophyte and gymnosperm groups, and basal angiosperms. Data from a wider sampling of land plant groups can help address numerous issues, including the history of the diversification of genes involved in reproductive development, the role that duplication and diversification of genes play in plant diversity, and the potential identification of homologous structures among distantly related taxa. Moreover, studies on the evolution and function of MADS-box genes that function in other, nonfloral aspects of plant development should yield general insights into the mechanisms behind functional diversification of developmental gene families. Previous efforts to dissect the evolution of flower development also largely have ignored floral regulatory genes outside of the MADS-box loci. To fully understand the genetic interactions that result in floral innovations, a parallel effort must explore the evolution of non-MADS-box genes among both flowering and nonflowering plants. Finally, we need to understand the microevolutionary forces that shape the diversification of floral regulatory loci, which requires an analysis of the evolution of floral developmental genes in the context of both population genetics and ecology.

One of the more exciting aspects of current research in this area is the close collaboration of evolutionary, developmental and molecular geneticists. The result is a continued forging of links between classical plant evolutionary biology and molecular genetics. In the next few years, progress in this area will depend largely on the continuing interaction of molecular and organismal evolutionists. As is evident from the nascent work in this field, bridging the gap between evolutionary and developmental geneticists provides a fruitful source of new ideas, and a synthetic approach to general questions on the evolution of floral and plant development.
ACKNOWLEDGEMENTS

The authors would like to thank V.F. Irish, G. Theissen, M.F. Yanofsky, R.J. Schmidt, B. Ambrose, D. Baum, J. Sundstrom and P. Engstrom for providing preprints of papers and for helpful discussions. We also thank Patrick Reeves, Richard Olmstead and three other anonymous reviewers for helpful suggestions. This work was supported in part by grants from UNAM, CONACYT and the Human Frontiers Program to E.A.B., and the National Science Foundation, US Dept of Agriculture and the Alfred P. Sloan Foundation to M.D.P.


Munster T, Pahnke J, DiRosa A et al. (1997) Floral homeotic genes were recruited from homologous MADS-box genes preexisting in the common ancestor of ferns and seed plants. Proceedings of the National Academy of Sciences of the USA 94, 6, 2415-2420.


Purugganan MD and Suddith JI (1999) Molecular population genetics of floral homeotic loci: departures from the equilibrium neutral model at the AP3 and PI genes of Arabidopsis thaliana. Genetics 151, 839-848.


Sundstrom J, Carlsbecker A, Svensson ME et al. (1999) MADS-box genes active in developing pollen cones of Norway spruce (Picea abies) are homologous to the B-class floral homeotic genes in angiosperms. Developmental Genetics 25, 3, 253-266.


Winter KU, Becker A, Munster T et al. (1999) MADS-box genes reveal that gnetophytes are more closely related to conifers than to flowering plants. Proceedings of the National Academy of Sciences of the USA 96, 13, 7342-7347.


Fig. 1. Major relationships within the plant MADS-box regulatory gene family. This composite supergene is derived from several different analyses. Not all major genes are shown, and the phylogeny from Theissen et al.13 provides the backbone of the composite tree. The bootstrapping support at nodes is taken from the individual studies, and only relationships consistent across the published phylogenies are shown. The major floral homeotic gene groups are indicated, as well as the fern MADS-box genes. Phylogenetic analyses also indicate a major split among angiosperm MADS-box genes before the divergence of the plant and animal lineages, resulting in a separate plant SRM-like MADS-box gene group (J.R. Alvarez-Eyzaguirre et al. unpublished). The species designations listed below are shown in brackets in the figure. Eudicot: At, Arabidopsis thaliana; At, Nicotiana tabacum; Le, Lycopersicon esculentum; Ph, Petunia hybrida; Am, Antirrhinum majus; St, Solanum tuberosum; Nonocot: Zm, Zea mays; Ad, Arabidopsis x Debenth orchid hybrid. Gymnosperms: Gg, Gnetum gnetum; Fa, Picea abies. Ferns: Cr, Ceratopteris richardii.
Box 1. The genetics of flower development

Development in plants relies on groups of undifferentiated embryonic cells called meristems, whose activities determine plant architecture and morphology. In Arabidopsis thaliana (Brassicaceae, thale cress) the reproductive phase begins when the shoot apical meristem transforms from being a vegetative meristem to an inflorescence meristem, in response to internal signals or external environmental cues. This reproductive, inflorescence meristem is indeterminate in Arabidopsis and primula at the flanks of the growing meristem can either form secondary inflorescence shoots or become floral meristems that will develop into flowers.

Based on extensive molecular expression and genetic data, floral homoeotic genes can be divided into two categories according to developmental fate: floral meristem identity genes, which regulate the identity of reproductive meristems; and organ identity genes, which control the identities of floral organs that form in developing flowers. Expression of floral meristem identity genes, such as LEAFY (LFY), APETALA1 (API) and CAULIFLOWER (CAF), leads to the formation of floral meristems in regions flanking the inflorescence meristem. Conversely, the expression of TERMINAL FLOWER (TFL), which maintains inflorescence meristem identity, downregulates the expression of the floral meristem gene LFY.

The floral meristem is partitioned into three overlapping fields of gene activity leading to the definition of four concentric rings or whorls of floral organs. There are four floral organ whorls in Arabidopsis: whorl one (four sepals), whorl two (four petals), whorl three (six stamens) and whorl four (two fused carpels). The figure shows the ABC model of floral organ identity. The A-, B- and C-class floral organ identity genes are expressed in two adjacent whorls of the flower. At least one of the A-class genes (APETALAS) negatively regulates expression of the C-class AGAMOUS gene in the first two whorls. Conversely, AG expression in whorls 3 and 4 negatively regulates both AP2 and API expression. Analysis of similar floral homoeotic mutants between Arabidopsis and Antirrhinum majus (Scrophulariaceae) indicate that floral organ identity loci fall into one of three general classes: A-, B- and C-class homoeotic genes. Genes of each class regulate floral organ development in two adjacent whorls and can interact in combination to determine the fate of organ primordia (Fig. 1). In general, A-class genes (including APETALAS and APETALAL2) affect development in whorls 1 and 2 (sepals and petals), B-class genes (including APETALAS and PISTILLATA) affect development in whorls 2 and 3 (petals and stamens) and C-class genes (AGAMOUS) affect development in whorls 3 and 4 (stamens and carpels). The ABC model further suggests that some A- and C-class genes are mutually antagonistic, such that mutations in the A-class gene AP2 lead to the expansion of C-class gene expression to all four floral whorls. Conversely, mutations in the C-class AG locus lead to expression of both API and AP2 in the third and fourth whorls.
CHAPTER 3

ADAPTIVE RADIATION AND
THE HAWAIIAN SILVERSWORD ALLIANCE
INTRODUCTION

The genetic basis for adaptation is of tremendous interest to evolutionary biologists, as it is the heritable component of favorable traits that ultimately leads to the capacity for a naturally-distributed species to dominate a landscape, optimize resources for persistence, or escape extinction. Thus, the neo-Darwinian synthesis that incorporated genetics into evolutionary studies prompted significant discussion regarding the role of genetics in adaptive evolution and the relative importance among selective processes and “dispersion forces” (Wade & Goodnight 1991,1998) operating on genetic mechanisms such as drift and mutation. Despite these discussions over the last century, we still know very little about the genetic basis of adaptation (Orr 1998, 2001).

Many investigations of the genetic basis of adaptation have focused on phenotypic, genome-wide, and gene specific searches for fixed or highly divergent factors between recently-derived species. Prior to recent quantitative genetic applications, these studies relied upon theoretical manipulations of population genetic parameters and simulation studies. Orr (1998) lists two main reasons for the lack of studies dealing directly with the population genetics of adaptation: (1) high attention to deleterious mutations (it is more tractable to detect the effects of deleterious genetic changes in short timescales), and (2) reliance upon Fisher’s mathematical proof that factors of large phenotypic effect are insignificant in adaptive processes (Fisher 1930; Orr & Coyne 1992; Orr 1998).

One valuable resource for investigating adaptive evolution is to study trends in rapidly-evolving species groups that have differentiated into multiple lineages through adaptation to diverse conditions. Indeed, such ‘adaptive radiations’ have been implicated as the source of several explosions of diversity including the diversification of early vertebrates during the ‘Cambrian explosion’ (Stanley 1973). Some scientists even suggest that adaptive radiation is responsible for the origin and diversification of all organismal lineages (Simpson 1953, discussed in Schluter 2000). While this suggestion remains unsubstantiated, regardless
of frequency, adaptive radiations undeniably provide the opportunity to understand accelerated evolution and the underlying basis for adaptation.

‘Adaptive radiation’ refers to the differentiation of an ancestral species into an array of rapidly-multiplying derived lineages that inhabit diverse environments and exhibit different character traits that exploit these environments (Lack 1947; Simpson 1953; Mayr 1963; Carlquist 1974; Grant 1986; Futuyma 1986; Skelton 1993; Schulter 2000). A key component of this definition is that the differentiation of the ancestral species coincides with the origin of new species through the evolution of ecological traits partitioning the adaptively radiating lineages. The most prevalent theory of adaptive radiation is that this differentiation occurs through divergent natural selection arising from adaptations to environments, resources, and resource competition. This adaptive evolution may occur through three processes (Schluter 2000): (1) phenotypic differentiation between populations and species due to environmental and resource differences, (2) phenotypic divergence caused by resource competition among lineages, and (3) divergent natural selection leading to increased ecological and phenotypic divergence between lineages that results in speciation.

The relative contributions of each of these three processes remain controversial. Two aspects that remain particularly unresolved include the separation of taxa through adaptation to different environments by crossing the fitness valleys espoused in Wright’s Shifting Balance theory (Wright 1980) and the connection between ecological opportunity and key organismal innovations (Schluter 2000). The foundation for these two unresolved issues is determining if divergent natural selection is the key element of rapid speciation during adaptive radiations. The emphasis upon divergent natural selective forces differentiating species during adaptive radiation contrasts with the competing theory that this divergence occurs through accelerated neutral-equilibrium processes that do not invoke natural selection (Schluter 2000).
According to Schluter, adaptive radiations follow four criteria: common ancestry among species in the radiation, phenotype-environment correlation, trait utility, and rapid speciation (Schluter 2000). Ability to detect all four of these features continues to be a significant challenge, and becomes more difficult as we move down this list. The most straightforward criterion to detect is common ancestry. Methods of phylogenetic analyses have been greatly improved in the last decade, making it much more feasible to determine if all members of a putative radiation descend from a common ancestral lineage. The next criterion is a significant association between environmental resources and the morphological and physiological traits that utilize these different resources. Unmasking such a significant correlation implies exploring fitness in an ecological context and requires verification that the morphologies and/or physiologies have a heritable, genetic basis. The third criterion is evidence that these correlated phenotypes are employed and actually improve fitness in the particular environment. The establishment of this connection between a specific phenotype, environment, and fitness establishes a direct correlation between heritable variation and adaptation that may ultimately lead to lineage divergence within a taxon. The increased differentiation among lineages with alternative adaptive trait values is the fourth criterion--rapid speciation. Verification of ‘rapid’ speciation across a species group with confirmed heritable phenotypic divergence associated with trait utility in alternative environments remains the ultimate goal.

Several examples of adaptive radiation have been investigated for these correlations, including some of the most dramatic examples of adaptive radiation such as Darwin’s finches on the Galapagos Islands (Lack 1947), West Indian Anolis lizards on the Caribbean Islands (Pianka 1969, 1986; Losos 1990a, 1990b, 1992; Garland 1994; Losos 1996; Irschick 1997), Columbines in North America (Chase & Raven 1975; Miller 1981; Hodges 1997; Fulton & Hodges 1999), and the ‘greatest living example’ of plant adaptive radiation--the Hawaiian silversword alliance (Robichaux 1984; Robichaux & Canfield 1985; Carr et al. 1989; Schluter 2000).
THE HAWAIIAN SILVERSWORD ALLIANCE ADAPTIVE RADIATION

The adaptive radiation that resulted in the Hawaiian silversword alliance ranks ‘with the dinosaur extinctions and the origin of our own species’ as one of the most celebrated events in the history of life on Earth (Schluter 2000). While the direct connections between heritable variation conferring specific phenotypes that result in environmental utility have yet to be ascertained, studies do indicate that this insular plant group may conform to Schluter’s four criteria for adaptive radiation. The Hawaiian silversword alliance species share a recent common ancestry, exhibit several physiological attributes that may correlate with adaptation to different environments, and have an excessive number of descendant species (Baldwin & Sanderson 1998; Schluter 2000).

Interspecific hybrid ancestry

The Hawaiian silversword alliance descended from the North American Madia lineage of the Madiinae (Heliantheae), which is known as the tarweeds (Baldwin 1995; Baldwin 1997; Robichaux 1990; Baldwin 1991; Baldwin ref). Based on a calibrated, rate-constant ITS tree, the onset of diversification in this monophyletic plant group appears to have occurred within the last 5.2 ± 0.8 million years (Figure 1, Baldwin & Sanderson 1998). This date of diversification coincides with the emergence of the island of Kauai. Further analyses of diversification dates suggest that the Hawaiian silverswords radiated along the age-progressive Hawaiian Islands as they formed from a ‘hot spot’ along the Hawaiian-Emperor chain that gradually drifted northwest due to sea floor spreading (Clague & Dalrymple 1987). Indeed, the periodic availability of new habitats generated by island dynamics and extensive chromosomal evolution are believed to have contributed to the adaptive radiation of the Hawaiian silversword alliance (Robichaux 1990, Wagner & Funk 1995).
Analysis of several homoeologous regulatory genes and a structural gene among North American tarweed and Hawaiian silversword alliance species revealed that the Hawaiian silversword alliance species are allopolyploids that appear to be derived from an interspecific hybridization event between two North American tarweed lineages (Barrier 1999). Phylogenetic analyses of \textit{ASAPETALA1} and \textit{ASAPETALA3/TM6} indicate that both of these genes are present as two distinct copies (homoeologs) derived from separate North American tarweed lineages. The A homoeologs derive from the ancestral lineage of \textit{Anisocarpus scabridus} and the B homoeologs derive from an ancestral lineage including \textit{Carlquistia muirii} (Barrier 1999).

The chromosomal contributions of both North American tarweed progenitor taxa (n = 7) remained unreduced; thus the entire Hawaiian silversword alliance is allopolyploid (n = 13 to n = 14) (Barrier 1999). Subsequent to initial colonization, the Hawaiian silversword alliance taxa adaptively radiated and speciated along the Hawaiian archipelago during the formation and weathering of these islands. The dynamics of island formation and ageing made available a vast array of dynamic ecological habitats, many of which have been exploited by this insular plant group. These include exposed lava, dry scrub, dry coast, mesic forest, wet forest and bogs. These habitats are the result of diverse ages of lava formations, rapid changes in elevation (Hawaiian silversword alliance species occupy habitats ranging in elevation from 75 m to 3750 m), and average annual rainfall levels ranging from less than 400 mm in dry scrub habitats to over 12,300 mm in wet forest and bog habitats (Smathers & Mueller-Dombois 1974; Carlquist 1980; Carr & Kyhos 1981; Robichaux 1985; Robichaux & Canfield 1985).

The current distribution and morphological range of the Hawaiian silverswords is much more diverse than conditions inferred for the progenitor North American tarweed lineages (Baldwin 1995). The North American tarweeds display nearly uniform biogeography generally restricted to the summer-dry Californian Floristic Province while the Hawaiian silversword alliance species are endemic to the Hawaiian Islands and occupy the
extremely diverse combinations of environmental variables mentioned above (Baldwin 1998). Additionally, the North American tarweeds are generally ephemeral herbs with similar reproductive morphologies while the Hawaiian silversword species exhibit a wide range of physiological variation associated with water retention (Robichaux & Canfield 1985), vegetative growth forms, and reproductive architectures (Carr 1985).

Diversity and divergence along the Hawaiian archipelago

The Hawaiian silversword alliance is a spectacular textbook example of plant adaptive radiation (Baldwin 1995, Baldwin 1997, Robichaux 1990, Schluter 2000). This insular plant group includes 30 perennial species composing three endemic genera: *Argyroxyphium*, *Dubautia* and *Wilkesia* (Baldwin 1995, Baldwin 1997, Robichaux 1990, Carr 1985). These species are distributed on six of the eight main islands of the Hawaiian archipelago (Kauai, Oahu, Molokai, Lanai, Maui, and Hawaii), with all but five species being single-island endemics.

The Hawaiian silversword alliance species display a remarkable array of morphological growth forms, including unbranched monocarpic rosette plants, branched polycarpic rosette plants, large shrubs, mat-forming subshrubs, trees, cushion plants and a liana. These species also exhibit significant diversity in flowering stalk, inflorescence and floral architectures (Carr 1985). For example, capitula (a composite flowering head) range in capacity from two florets per capitulum in *Dubautia pauciflorula* to as many as 650 florets per capitulum in *Argyroxyphium sandwicense* subsp. *macrocephalum* (Carr 1985). Capitulescence (inflorescence) types also vary across this insular plant group, from racemiformous in *A. caliginis* and *W. hobdyi*, paniculiformous in *A. sandwicense* and *W. gymnoxiphium*, solitary form in *D. ciliolata* subsp. *glutinosa*, corymbiformous in *D. scabra*, elongated paniculiformous in *D. plantaginea*, *D. latifolia*, and *D. pauciflorula*, and racemiformous in *D. menziesii*, *D. arborea*, and *D. platyphylla* (Carr 1985). In addition to
reproductive features, Hawaiian silversword alliance species also vary in vegetative traits, including leaf morphologies ranging from small xeromorphic leaves less than 1 cm in length in *D. ciliolata* to expansive mesomorphic leaves greater than 25 cm in length in *D. raillardioides* and leaf shapes such as linguate, elliptic, lanceolate, linear, ovate, and oblong forms. Other morphologies that vary in the Hawaiian silversword alliance include vascular bundle anatomy, leaf arrangement, stigmatic branching, pappus types and trichome types (Carr 1985).

The vast majority of the 30 Hawaiian silversword alliance species are in the genus *Dubautia*, which is composed of 21 endemic species. All *Argyrotophiun* and *Wilkesia* species and 12 *Dubautia* species contain 14 chromosome pairs, while nine *Dubautia* species have 13 chromosomal pairs, suggesting an aneuploid reduction during the diversification of the *Dubautia* lineage (Carr 1978; Carr & Kyhos 1981; Robichaux & Canfield 1985). This reduction in chromosome number in the *Dubautia* lineage may have coincided with significant changes at the physiological level (Robichaux & Canfield 1985). Examination of turgor maintenance capacity among several 13-paired and 14-paired *Dubautia* species indicates that the derived 13-paired *Dubautia* species may have increased tolerance for low moisture availability than the 14-paired *Dubautia* species. This increased tolerance for arid conditions may have led to the successful colonization of the 13-paired species in habitats that are significantly drier than those inhabited by the ancestral 14-paired *Dubautia* species (Robichaux & Canfield 1985; Robichaux et al. 1990).

*Incomplete genetic divergence despite accelerated rate of evolution*

While the Hawaiian silversword alliance species display an amazing array of divergent morphologies and physiologies, these species are not completely reproductively isolated. Artificial intergeneric and infraspecific hybrids can be created between many species of this plant group (Carr & Kyhos 1981, 1986; Carr et al. 1989, 1996; Baldwin et al. 1997).
1991, 1998; Baldwin 1997). Additionally, several naturally-occurring intergeneric hybrids have been documented, including (among others) hybridizations between *A. sandwicense* subsp. *sandwicense* and *D. arborea*, and hybridizations between *A. caliginis* and *D. scabra* subsp. *scabra* (Carr & Kyhos 1981, 1986; Carr 1995; Carr *et al.* 1996).

Analysis of ten allozyme loci further indicates low genetic differentiation among Hawaiian silversword alliance species (Witter & Carr 1988). In this study, even species that diverged approximately 500,000 to 1.5 million years ago have highly similar genetic identities, including many Nei distance coefficients in the range of values typically expected in comparisons between populations within species (*I* > 0.90). The low genetic differentiation among species in the Hawaiian silversword alliance contrasts with recent analyses that indicate an accelerated rate of protein evolution across this insular plant group (Barrier *et al.* 2001). Accelerated protein evolution coupled with adaptive divergence in these species might suggest decreased genetic divergence among these species if differentiation is the result of accelerated lineage sorting processes affecting the entire genome. This contrast of low genetic divergence and accelerated evolution in these species suggests that processes associated with the natural selection of few loci with large phenotypic effects may be operating.


Chase VC and Raven PH (1975) Evolutionary and ecological relationships between *Aquilegia formosa* and *A. pubescens* (Ranunculeae), two perennial plants. *Evolution* 29, 474-486.


Figure 1. Time-calibrated phylogeny of one of four most-parsimonious rDNA ITS trees of the Hawaiian silversword alliance (Argyroxyphium, Dubautia, and Wilkesia) and closest continental perennial relatives in Madia and Raillardiopsis (Baldwin 1992; Baldwin & Robichaux 1995). Outgroup tarweed taxa, Adenothamnus validus and Raillardella pringlei, are not shown. Bootstrap values are shown along branches. Consistency index = 0.81, Retention index = 0.85. The tree shown is a clade that nests within the much larger tarweed radiation (ca. 114 species in 17 genera) (Baldwin 1996). This figure is from Baldwin & Sanderson 1998.
CHAPTER 4

PATTERNS OF MOLECULAR EVOLUTION AMONG PARALOGOUS FLORAL HOMEOTIC GENES

Amy L. Lawton-Rauh, Edward S. Buckler IV, and Michael D. Purugganan

This chapter consists of a paper published as:
The plant MADS-box regulatory gene family includes several loci that control different aspects of inflorescence and floral development. Orthologs to the *Arabidopsis thaliana* MADS-box floral meristem genes *APETALA1* and *CAULIFLOWER* and the floral organ identity genes *APETALA3* and *PISTILLATA* were isolated from the congeneric species *Arabidopsis lyrata*. Analysis of these loci between these two *Arabidopsis* species, as well as three other more distantly related taxa, reveal contrasting dynamics of molecular evolution between these paralogous floral regulatory genes. Among the four loci, the *CAL* locus evolves at a significantly faster rate, which may be associated with the evolution of genetic redundancy between *CAL* and *AP1*. Moreover, there are significant differences in the distribution of replacement and synonymous substitutions between the functional gene domains of different floral homeotic loci. These results indicate that divergence in developmental function among paralogous members of regulatory gene families is accompanied by changes in rate and pattern of sequence evolution among loci.
INTRODUCTION

Developmental processes are largely controlled by regulatory loci that modulate gene expression patterns. Molecular genetic studies have identified a number of loci that regulate developmental processes, and many of these genes have been shown to be members of regulatory gene families. Developmental genetic investigations have demonstrated that these regulatory gene families evolve primarily by gene duplication and divergence (Scott & Weiner 1984; Purugganan 1998), leading to distinct paralogous loci regulating different aspects of organismal morphogenesis. The evolutionary divergence of gene function among regulatory gene family members provides a mechanism for the elaboration of developmental genetic networks and the increasing complexity of morphological structures. There has thus been continued interest in studying the evolutionary histories and dynamics of developmental regulatory genes, as well as in increasing efforts to investigate the evolution of homeodomain (Zhang & Nei 1996; Bailey et al. 1997), bHLH (Atchley & Fitch 1995), myb-class (Rosinski & Atchley 1998), and other regulatory gene families. It remains unclear, however, to what extent functional diversification among regulatory loci is mirrored by differences in the extent and patterning of sequence evolution between homologous developmental genes.

The plant MADS-box regulatory gene family comprises a group of developmental regulatory loci that encode sequence-specific DNA-binding transcriptional activators (Riechmann & Meyerowitz 1997). Plant MADS-box proteins are about 240–260 amino acids (aa) in length and are characterized by the presence of a highly conserved 57-aa MADS-box (see fig. 1). This MADS-box protein region is widely distributed among eukaryotic genomes within humans (SRF), Drosophila (MEF2C), and yeast (MCM1) transcriptional activators (Pollock & Treisman 1991). Plant MADS-box proteins also possess a moderately conserved 70-aa domain called the K-box, which may form coiled-coil structures that participate in protein–protein dimerization interactions (Ma et al. 1991; Riechmann & Meyerowitz 1997). Plant MADS-box proteins also include the I- and C-regions that are poorly conserved at the sequence level (Purugganan et al. 1995).
Molecular studies indicate that homo- or heterodimerization of MADS-box proteins is necessary for sequence-specific DNA-binding activity of these transcriptional activators (Riechmann & Meyerowitz 1997). Domain analyses of Arabidopsis MADS-box proteins indicate that the MADS-box/I-region/K-box sequence (the MIK region) constitutes the functional core domain necessary for both dimerization and DNA binding (see fig. 1). Previous molecular studies have demonstrated that at least one Arabidopsis MADS-box transcriptional activator binds its DNA target as a homodimer (Krizek & Meyerowitz 1996; Riechmann et al. 1996; Riechmann et al. 1996). Moreover, biochemical studies indicate that the Arabidopsis APETALA3 and PISTILLATA proteins bind to target promoter sequences as a heterodimer (Riechmann & Meyerowitz, 1997). The dimerization domains differ between various MADS-box proteins; in vitro experiments suggest that the MADS-box and the I-region are both necessary for dimerization and that the K-box serves to stabilize protein–protein interactions (Riechmann & Meyerowitz 1997). Domain deletion studies of the related MADS-box transcriptional activator AGAMOUS indicates that a protein containing only the MADS-box, the I-region, and the K-box (the MIK region) can form efficient dimers (Mizukami et al. 1996).

Mutations at several Arabidopsis MADS-box genes result in floral phenotypes characterized by alterations in floral organ development (Bowman et al. 1991). These loci are referred to as floral homeotic genes, and genetic and molecular studies indicate that these loci fall into two broad classes: floral meristem identity genes, specifying inflorescence and floral meristem identity, and floral organ identity genes, defining identities of organs at specific locations in the developing flower (Yanofsky 1995) (see fig. 2A). APETALAI and CAULIFLOWER are two MADS-box meristem identity loci. API and CAL perform partially redundant developmental functions, specifying the identity of the floral meristems flanking inflorescence shoots (Mandel et al. 1992; Bowman et al. 1993; Kempin et al. 1995). API appears to possess additional organ identity functions involving sepal and petal differentiation. Phylogenetic studies indicate that API and CAL arose from a recent gene
duplication event, as the duplicate CAL locus is found only within the Brassicaceae family (see fig. 2B) (Purugganan, 1997; Lowman & Purugganan, 1999). The Arabidopsis loci APETALA3 and PISTILLATA are floral organ identity genes required for petal and stamen development (see fig. 2A) (Jack et al. 1992; Goto & Meyerowitz 1994). Like AP1 and CAL, AP3 and PI are closely related to one another and arose via a gene duplication event. However, the AP3-PI duplication event occurred before the major angiosperm diversification, considerably earlier than the AP1-CAL duplication (see fig. 2B) (Purugganan 1997; Kramer et al. 1998). Additionally, phylogenetic studies indicate that the last common ancestor of the AP3 and PI group of genes and the AP1/CAL gene subgroup predates the angiosperm/gymnosperm separation 285 MYA (Purugganan 1997). Orthologs to the Arabidopsis floral homeotic genes AP3, PI, and AP1/CAL have been identified in several other angiosperm species (see table 1), including Antirrhinum majus (Scrophulariaceae) (Coen 1991; Saedler & Huijser 1993) and Silene latifolia (Caryophyllaceae) (Hardenack et al. 1994). In Antirrhinum, the DEFICIENS, GLOBOSA, and SQUAMOSA genes are orthologs to the Arabidopsis AP3, PI, and AP1 loci, respectively (Sommer et al. 1990; Huijser et al. 1992; Trobner et al. 1992). Experiments with Antirrhinum indicate that DEF and GLO are also involved in petal and stamen differentiation in snapdragon flowers, while mutations in SQUA result in the formation of bract-forming shoots in place of flowers. In Silene, a number of MADS-box genes (SLM2–SLM4) have been shown to be orthologs to different Arabidopsis floral homeotic loci (Hardenack et al. 1994; Purugganan 1997) (see table 1). Genetic studies among distantly related flowering plant species with distinct floral morphologies suggest evolutionary conservation of basic developmental function between these orthologous floral homeotic loci (Coen 1991).

An analysis of the molecular evolution of the MADS-box floral homeotic genes constituting the flower developmental pathway may provide insights into evolutionary patterns that accompany diversification of genes within a regulatory network. This analysis would benefit from the significant amount of structural and functional information on these
floral homeotic loci that has been obtained with both genetic and molecular approaches (Riechmann & Meyerowitz 1997). Moreover, the use of paralogous members of a single regulatory gene family provides a common structural and functional context to compare evolutionary patterns across developmentally distinct loci. In this paper, we describe patterns of sequence evolution among the genes *APETALA3, PISTILLATA, APETALA1,* and *CAULIFLOWER.* By utilizing genes from species that have diverged at various times, one can analyze changes in the rates of molecular evolution across different temporal scales. Our analysis includes comparisons of loci within the Brassicaceae (5–40 Myr divergence times) and among genes found between Brassicaceae, Scrophulariaceae, and Caryophyllaceae species (~ 60 Myr divergence) (Crane *et al.* 1995).

We find that despite similarities in overall structure and basic functions of these different floral homeotic MADS-box genes, there appear to be significant differences in the rates of nucleotide substitution between functionally distinct loci. Genes that control differing aspects of floral development also show contrasting distributions in sequence changes across structural and functional domains. Altogether, this analysis provides a context for looking at sequence divergence among functionally distinct loci within a developmental regulatory gene family.
MATERIALS AND METHODS

Isolation and Sequencing of Floral Homeotic Genes

Arabidopsis lyrata seed was provided by C. H. Langley. Tissue for the Arabidopsis thaliana Landsberg erecta ecotype was obtained from single-seed propagated leaf material provided by the Arabidopsis Biological Resource Center.

The Arabidopsis lyrata AP3, PI, CAL, and AP1 genes were isolated by PCR amplification using primers designed from A. thaliana sequences. Miniprep DNA was isolated from young leaves as previously described (Ausubel 1992). PCR was performed, with 40 cycles of 1 min at 95°C, 1 min at 52°C, and 3 min at 72°C, followed by 15 min at 72°C. The error-correcting recombinant Tth polymerase XL formulation (Perkin Elmer) was used to minimize nucleotide misincorporation. The error rate for this polymerase formulation, based on multiple amplification and resequencing of known genes, is less than 1 in 7,000 bp (unpublished data).

The isolation of the A. lyrata CAL gene is reported elsewhere (Purugganan & Suddith 1998). PCR primers were designed based on genomic sequences provided by M. F. Yanofsky (AP1 and CAL) and E. M. Meyerowitz (AP3 and PI). The AP3-specific primers AP3F (for exon 1 forward) (5’ –GAATATGGCGAGAGGGAAGATCC- 3’) and AP3R (for exon 7 reverse) (5’-GCCTTTAATTATTCAAGAAGATGG-3’) and the PI-specific primers PI-1F (for exon 1 forward) (5’-GAGAAAAGATGGGTAGAGGAAG-3’) and PI-1R (for exon 6 reverse) (5’-ATCTCGATGATCAATCGATGACC-3’) were used in PCR reactions to amplify the A. thaliana and A. lyrata AP3 and PI genes, respectively. The A. lyrata and A. thaliana AP1 genes were isolated as two overlapping fragments. This reaction utilized primers AP1FPCR (5’-ATGGGAAGGGGTAGGGTTCA-3’) and AP1X2R (5’-ATTAATTCTGCTCCACCGATCC-3’) for the 5’ fragment and primers AP1X2F (5’-
GTAAAAGGTACTATTGAGG-3’ and (5’-AAGGTTGCAGTT-GTAAACGGG-3’) for the 39 AP1 fragment. Amplified DNA was cloned into pCR2.1 using the TA cloning kit (Invitrogen). DNA sequencing for both genes was conducted with the ABI377 automated sequencer using a series of nine nested internal sense and antisense primers. Cloned genes were sequenced at least twice, and ambiguous sites were visually rechecked from chromatograms. The DNA sequences are available from GenBank (accession numbers AF143379–AF143382).

Data Analysis

Sequences for floral homeotic genes from A. thaliana, Brassica oleraceae, Antirrhinum majus, and Silene latifolia were obtained from GenBank. Sequences between A. thaliana and A. lyrata were visually aligned; more distantly related sequences were aligned using published alignment frameworks for plant MADS-box genes (Purugganan et al. 1995).

The nucleotide substitution distances between sequences were estimated using the Tajima-Nei model (Tajima & Nei 1984). For coding region sequences, synonymous (KS) and nonsynonymous (KA) substitution distances were estimated with Jukes-Cantor corrections (Nei & Gojobori 1986). The statistical significance of nucleotide substitution distance differences was evaluated using a t-test. Synonymous codon usage for each gene was determined using the MEGA program package (Kumar et al. 1994), and the effective number of codons was estimated (Wright 1990).

Phylogenies of the different floral homeotic genes were estimated using both maximum-parsimony (MP) (Swoford 1993) and neighbor-joining (NJ) (Saitou & Nei 1987) techniques. For the MP technique, the heuristic search algorithm with the tree bisection-reconnection procedure of the PAUP program was used (Swoford 1993) with random addition of genes and with the MULPARS and collapse options in effect. Node confidence
was assessed with 500 bootstrap replicates of the data. For the NJ analysis, genetic distances were estimated using either the Tajima-Nei distances or nonsynonymous substitution distances. Confidence estimates were assessed with 500 bootstrap replicates of the data. The NJ analysis was undertaken using the MEGA program package.

Nucleotide substitutions were mapped onto the gene phylogenies using MacClade to determine the proportion of molecular change across various gene domains (Maddison & Maddison 1992). The number of inferred replacement changes along phylogenetic branches was plotted in a sliding-window analysis with a sequential overlapping 25-aa partition. Significance of differences in the distribution of replacement changes at different structural regions between paralog pairs was determined by calculating \( D_i \): 
\[
D_i = \left| \text{proportion of protein 1 - protein 2 replacement differences within window } i \right| 
\]
for each sliding window \( i \), and 
\[
D = \max(D_i) \text{ over all } i \text{ sliding window positions.} 
\]
The distribution of \( D \) for any pair of proteins was estimated by calculating the maximum sliding window difference for 1,000 permutations of the protein residues.

Contingency tests for independence of coding region substitution categories were conducted using Fisher’s exact test to evaluate significance. The coding-region variation was partitioned into functional (core and noncore) domains (Riechmann & Meyerowitz 1997) for separate contingency analyses (Templeton 1996). The relative-ratio test (Muse & Gaut 1997) was used to examine locus-by-lineage effects in gene evolution. The test was conducted between gene pairs; for those pairs involving \( PI \), we excluded the \( B. \) oleracea sequences. This test was implemented using the program PROPML (Proportional Maximum Likelihood) provided by S. V. Muse (North Carolina State University).
RESULTS AND DISCUSSION

Comparison of *A. thaliana* and *A. lyrata* Floral Homeotic Genes: Evolution at Short Temporal Scales

This study of the molecular evolution of developmental regulatory genes revolves around two questions: (1) Are there differing patterns of molecular evolution among orthologous developmental genes from species with diverse morphologies and (2) is the diversification of paralogous genes with specialized developmental functions accompanied by divergent patterns of molecular evolution?

Since the structural similarities among MADS-box floral homeotic genes reflect the biochemical similarity in their roles as transcriptional activators (Riechmann & Meyerowitz 1997), these four regulatory loci should display similar rates and patterns of molecular evolution. Alternatively, divergence in developmental function among these four regulatory loci in various species and different gene lineages may result in species- or lineage-specific variation in rates and patterns of molecular evolution between these developmental control genes (Purugganan 1998). In order to address this issue, the molecular evolution of four paralogous floral regulatory genes was analyzed. We isolated orthologs to the *A. thaliana* *AP1*, *CAL*, *AP3*, and *PI* loci from the congeneric Brassicaceae species *A. lyrata*. Orthologs to these *Arabidopsis* genes have also been identified in another Brassicaceae species (*B. oleracea*) as well as in *A. majus* (Scrophulariaceae) and *S. latifolia* (Caryophyllaceae) (table 1).

Neighbor-joining phylogenies of the different loci (see fig. 3) are congruent with previously established phylogenies (Purugganan *et al.* 1995; Purugganan 1997). The phylogenetic trees also show that the *AP1* and *CAL* genes duplicated before the split of *Arabidopsis* and *Brassica*, but after the separation of these Brassicaceae species from Scrophulariaceae and Caryophyllaceae (see fig. 3). Comparison between *A. thaliana* and *A.
*lyrata* genes permits an analysis of regulatory gene diversification at relatively short evolutionary timescales (<5 MYA). The close relationship between *A. thaliana* and *A. lyrata* is reflected in the high sequence similarity between genes from these two species, with genomic sequences from the two taxa differing by approximately 3% at the nucleotide level. The intron/exon structures of the floral homeotic genes are also similar between the two species. There are several insertion/deletion (indel) differences within the intron regions of the various loci contributing to differences in genomic sequence lengths between the two *Arabidopsis* species. The *AP1* and *CAL* orthologs have 90 and 70 indels, respectively, ranging in size from 1 to 127 bp. There are fewer indels in *AP3* and *PI* genes between *A. thaliana* and *A. lyrata*. These two floral homeotic genes have 25 and 28 indels, respectively, from 1 to 50 bp in length.

The genes in these two species show similar patterns of moderate codon bias, with the effective number of codons (ENC) (Wright 1990) being approximately 56.7 for *CAL* and 55.9 for *AP3* and *PI*. The *AP1* locus in both species displays a slightly higher bias (ENC = 47.3) and is higher than the average for *Arabidopsis* nuclear genes (Miyashita *et al.* 1998). Since codon bias appears to be correlated with gene expression levels, this suggests that *AP1* may be expressed at higher levels in these species than are the three other floral regulatory genes (Sharp & Li 1986). Indeed, the multiple roles of *AP1* in floral meristem and organ identity (and the more limited developmental role of the three other loci) provide support for the hypothesis that the higher codon bias for this gene may be associated with selection for translational efficiency.

The rates of sequence substitution at these floral homeotic loci can be directly compared between the two species. The four floral homeotic genes evolve at different rates between the two species (see table 2). The Tajima-Nei sequence distance values (*K*) for three genes (*AP1*, *AP3* and *PI*) are comparable, ranging from 0.079 to 0.089 nucleotide substitutions per site across the entire gene. In contrast, the *CAL* locus evolves at a faster rate than the other regulatory loci (*K* = 0.120, *P* < 0.05). The high rate in *CAL* is the product of a
high nonsynonymous substitution rate compared with those of the other loci ($P < 0.05$ for $AP1$ and $AP3$).

The faster rate of protein evolution for $CAL$ is also reflected in the ratio of nonsynonymous to synonymous substitutions ($Ka/Ks$). The $CAL$ gene has a $Ka/Ks$ of 0.28, while the other three floral homeotic loci have a $Ka/Ks$ ratio of 0.13 to 0.16 (see table 1). These estimates indicate that the rate of protein evolution is greater for $CAL$ than for the other three MADS-box floral homeotic genes between these two closely related Arabidopsis species. Similar results are seen when $A. thaliana$ and $A. lyrata$ $CAL$ genes are compared with the $B. oleracea$ ortholog (table 2). The accelerated evolution of $CAL$ within members of the Brassicaceae family may reflect its more recent origin within this group. The duplication of $CAL$ and $AP1$ occurred sometime during the evolution of the Brassicaceae, possibly early in the history of this eudicot family (Purugganan 1997). Genetic studies indicate a degree of genetic redundancy between $AP1$ and $CAL$ in floral meristem identity function (Kempin et al. 1995). The redundancy of $CAL$ to $AP1$ may be reflected in the higher rate of evolution for this locus than for the other floral homeotic genes that are presumably under stronger stabilizing selection.

The Floral Homeotic Genes Do Not Display Significant Locus-by-Lineage Effects

Plant genes may evolve at different rates along different species lineages (Gaut et al. 1997). Several mechanisms, such as generation time, life history, and global mutation rates, will affect all loci within a genome to a similar extent. Alternatively, selective forces affecting individual loci that may vary between different species will result in uncorrelated evolutionary rates along specific taxonomic lineages (Gillespie 1991). The relative ratio test evaluates whether the rates of molecular evolution are correlated among loci in various species lineages (Muse & Gaut 1997). Locus-by-lineage effects are characterized by uncorrelated variation among loci in the rates of evolution in different species lineages and may reflect variation in selective forces among different genes in different species.
The relative-ratio test does not reject the null hypothesis of correlated rates of molecular evolution among all four floral homeotic loci in different species lineages. The relative ratios of nonsynonymous distances between floral homeotic genes among different species lineages are remarkably similar, implying similar patterns of evolutionary rate among the genes between taxa. This correlation in rates is present even though the different genera utilized in the study (*Arabidopsis, Brassica, Antirrhinum,* and *Silene*) display considerable variation in inflorescence form, symmetry and size of flowers, and number, size, and shape of floral organs. This suggests that selection for floral morphological diversification between taxa is not associated with large, statistically significant differences in rates of molecular evolution among these specific taxonomic lineages.

**Patterns of Divergence Between Functional Domains**

Molecular genetic studies have delineated the functional core sequences of several MADS-box floral homeotic loci. This core region, which encompasses the MADS-box, the I-region, and the first 16 aa of the K domain, has been shown to be necessary for dimerization and DNA-binding activities of these proteins (see fig. 1) (Riechmann & Meyerowitz 1997). The noncore region, which includes the 3’ half of the K-domain and all of the C-terminal region, does not appear to be important for DNA binding. The noncore region, however, includes sequences that may serve as the transcriptional activation domain (Riechmann & Meyerowitz 1997).

Previous work indicates that the levels of nonsynonymous substitutions are generally higher in the sequence encoding the C-terminal regions of plant MADS-box genes. Calibrated rates of nonsynonymous substitutions in the C-terminal region are found to be $79 \times 10^{10}$ nonsynonymous substitutions per site per year, compared with $3 \times 10^{10}$ nonsynonymous substitutions per site per year for the highly conserved MADS-box region,

92
suggesting that the noncore region contains sequence elements that consistently display a greater degree of sequence divergence (Purugganan et al. 1995). However, mutations in the C-terminal region of several MADS-box genes are known to produce floral homeotic phenotypes (Kempin et al. 1995), indicating that this domain is required for protein function despite its rapid evolutionary rate.

The distributions of nucleotide substitutions between these sequence regions provide information on the differences in selective forces that operate between these two functional domains. Similar selective constraints on all four floral homeotic genes should result in a similar distribution of replacement substitutions between core and noncore domains (Templeton 1996). A sliding-window analysis plotting the number of amino acid replacements detected within a gene from all five study species indicates that different loci have different patterns of conservation and divergence (see fig. 4). Both visual inspection of these graphs and permutation testing suggest that the closely related paralogs CAL and AP1 exhibit nearly identical patterns. This result may reflect both the relatively recent duplication of these loci and the use of the same Silene and Antirrhinum genes as outgroup sequences in the analysis of both CAL and AP1 variation. The paralogs AP3 and PI do show substantial differences from one another. A permutation test indicates that variation between portions of the AP3 and PI I- and C-terminal regions are significant ($P < 0.05$). These differences in amino acid replacement patterns suggest that each locus evolves differently despite similarities in overall structural organization between the paralogs.

A contingency analysis was undertaken to test whether relative levels of replacement and synonymous substitutions were similar between functional regions in all floral homeotic genes (Templeton 1996). These contingency tests indicate that the distribution patterns of nucleotide substitutions differed between floral homeotic loci when A. thaliana and A. lyrata were compared. The AP3 gene appears to show substantial constraint in both the core and the noncore domains (see table 4). Only 2 of the 6 substitutions in the core domain and 2 of the 10 substitutions in the noncore domain were replacement changes; the differences
between core and noncore regions are not statistically significant (Fisher’s exact test, \( P = 0.604 \)). In \texttt{AP1}, \texttt{CAL}, and \texttt{PI}, however, the noncore domain exhibited as many or more nonsynonymous substitutions relative to synonymous substitutions than did the core domains (table 4). The differences between core and noncore domains, however, are significant only for \texttt{AP1}, for which 10 out of 17 differences in the noncore domain are nonsynonymous, while no substitutions in the core region result in amino acid replacements (Fisher’s exact test, \( P < 0.019 \)).

The patterns of nonsynonymous evolution differ between different evolutionary timescales (see table 5). In general, the rate of nonsynonymous evolution for the noncore region is approximately two to three times the rate for the core domain when very distantly related species (\texttt{A. thaliana} vs. \texttt{A. majus} or \texttt{S. latifolia}) are examined. This trend does not appear to hold for \texttt{AP3}; the noncore region has half as many nonsynonymous substitutions as the core region in comparisons between \texttt{A. thaliana} and \texttt{B. oleracea} (0.034 and 0.017 nonsynonymous substitutions per site for the \texttt{AP3} core and noncore domains, respectively), and over longer evolutionary distances, the situation is reversed in \texttt{AP3}.

\textit{Evolutionary Dynamics of a Regulatory Gene Pathway}

It has been shown in several cases that gene families controlling morphogenesis typically contain both paralogous and orthologous members with diverged developmental functions (Ruddle \textit{et al.} 1994). This variation in regulatory function, however, proceeds in the context of remarkable conservation of structure in key domains within such control loci (Gerhart & Kirschner 1997). This apparent structural conservation underlies the similarity in biochemical function (e.g., DNA-binding transcriptional activation) among members of a regulatory gene family. These observations suggest that changes in regulatory gene sequences play a secondary role during the evolution of eukaryotic developmental systems.
and that changes in regulatory gene expression patterns provide the major mechanism by which control genes diverge (Gerhart & Kirschner 1997).

If differential expression patterns and not the actual sequences of regulatory loci are crucial to the divergent functions, then the evolutionary dynamics of regulatory protein structure should be similar across homologous loci. Our analyses of floral homeotic genes of different taxa indeed suggest that there are no apparent significant differences in the rates of evolution of orthologous plant MADS-box loci. Relative-rate tests do not show any appreciable rate differences between orthologous loci (Purugganan 1997). Moreover, relative-ratio tests do not detect significant floral homeotic locus-by-lineage effects. This indicates that the rates of molecular evolution of different floral homeotic genes are similar across the flowering plant taxa analyzed.

Such similarity in molecular evolutionary rates across different taxonomic lineages is in contrast to the significant variation in floral and inflorescence morphologies displayed by some of the species in this study. This suggests that selective differences between orthologous loci in these different taxa are not manifested in dramatic differences in the rates of nucleotide substitution. This would lend support to the assertion that differential regulation, and not differential structure, may be the major component of regulatory gene evolution (Gerhart & Kirschner 1997). However, it is also probable that orthologous regulatory genes may be less likely to display significantly different rates of molecular evolution as a result of species-specific variation in selection pressures. For example, if evolutionary changes in the sequences of these regulatory proteins are crucial to interspecific diversification of function, then these changes may be confined to only a few amino acid positions and would not significantly impact the overall between-species nucleotide substitution rates. There is evidence, moreover, for accelerated protein evolution rates in orthologous floral homeotic genes at short timescales in plant adaptive radiations (unpublished data).
While variation in the patterns of molecular substitutions is not evident between orthologous floral homeotic genes among our taxa, there appear to be significant differences in the evolutionary dynamics of paralogous regulatory genes performing distinct developmental functions. The divergence in floral developmental function of the four homeotic genes in this study appears to be accompanied by variation in the patterns of molecular evolution among these loci. Comparisons between the floral homeotic genes found in Brassicaceae species, for example, suggest that at least one locus (CAL) evolves at a significantly higher rate than other paralogous genes. This increased rate of CAL evolution is also observed at nonsynonymous sites, but not at synonymous nucleotide positions, suggesting that variation in the rate of protein evolution that exists among paralogous loci is not merely due to differences in the neutral mutation rates.

Our results also indicate that paralogous developmental regulatory genes may display differences in the patterns of nucleotide substitutions between functional domains, most likely as a result of differences in the selective pressures experienced by each gene. Similarity in selective constraints between loci would be expected to manifest itself in the conservation of nucleotide substitution patterns between functional domains among the four floral homeotic genes. Analyses of the distribution of nucleotide substitutions indicate that molecular changes are not partitioned equally among the functional domains. In general, the AP1, CAL, and PI genes display greater degrees of sequence constraint in the core domain, with the noncore regions showing higher levels of nonsynonymous substitutions. AP3, however, has consistently high levels of sequence constraint across both domains. There are differences in the patterns of substitution despite similarities in developmental function among the loci. The AP3 and PI genes, for example, are both required for petal and stamen development, but the distribution of replacement substitutions between functional domains differs significantly between these two floral homeotic genes.

The emerging picture from this analysis is a contrasting portrait of regulatory gene evolution between paralogous members of a developmental gene family and orthologs of
specific genes found in morphologically distinct taxa. Interspecific evolution of these floral regulatory loci does not appear to be associated with striking variation in evolutionary rates between orthologs. Divergence of developmental function between paralogous regulatory loci, on the other hand, appears to be associated with significant differences in the rates and patterns of molecular evolution. Our results suggest that diversification of regulatory gene families by duplication results in the formation of loci with distinct and contrasting evolutionary dynamics.
ACKNOWLEDGMENTS

We would like to thank S. V. Muse and J. Thorne for stimulating discussions. We would also like to thank M. F. Yanofsky and E. M. Meyerowitz for providing us with sequence information. This work was supported in part by an NIH postdoctoral fellowship to E.S.B. and grants from the USDA NRICGP and the National Science Foundation to M.D.P.
LITERATURE CITED


Table 1. Floral Homeotic Genes Used in this Study.

<table>
<thead>
<tr>
<th>Arabidopsis thaliana</th>
<th>Arabidopsis lyrata</th>
<th>Brassica oleracea</th>
<th>Antirrhinum majus</th>
<th>Silene latifolia</th>
</tr>
</thead>
<tbody>
<tr>
<td>APETALAI</td>
<td>AIAPETALAI</td>
<td>BoAPETALAI</td>
<td>SQUAMOSA</td>
<td>SLM4</td>
</tr>
<tr>
<td>CAULIFLOWER</td>
<td>AICAULIFLOWER</td>
<td>BoCAULIFLOWER</td>
<td>SQUAMOSA*</td>
<td>SLM4*</td>
</tr>
<tr>
<td>APETALAI3</td>
<td>AIAPETALAI3</td>
<td>BoAPETALAI3</td>
<td>DEFICIENS</td>
<td>SLM3</td>
</tr>
<tr>
<td>PISTILLATA</td>
<td>AIPISTILLATA</td>
<td>n.a.</td>
<td>GLOBOSA</td>
<td>SLM2</td>
</tr>
</tbody>
</table>

*Since the *AP1* and *CAL* genes are duplicate within the Brassicaceae, these loci are orthologous to both Brassicaceae *AP1* and *CAL.*
Table 2. Sequence Distances Between *Arabidopsis thaliana* and Other Brassicaceae Floral Homeotic Gene Orthologs

<table>
<thead>
<tr>
<th>Gene</th>
<th>versus <em>Arabidopsis lyrata</em></th>
<th>versus <em>Brassica oleracea</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K(genomic)</td>
<td>Ka</td>
</tr>
<tr>
<td>APETALA1</td>
<td>0.079 ± 0.036</td>
<td>0.018 ± 0.006</td>
</tr>
<tr>
<td>CAULIFLOWE</td>
<td>0.120 ± 0.007</td>
<td>0.039 ± 0.008</td>
</tr>
<tr>
<td>APETAL3</td>
<td>0.089 ± 0.008</td>
<td>0.013 ± 0.005</td>
</tr>
<tr>
<td>PISTILLATA</td>
<td>0.087 ± 0.007</td>
<td>0.022 ± 0.007</td>
</tr>
</tbody>
</table>

<sup>a</sup> Genomic sequence distance not available due to lack of intron information for *Brassica oleracea* genes.
Table 3. Sequence Distances Between Coding Regions of *Arabidopsis thaliana* and Distant Floral Homeotic Gene Orthologs

<table>
<thead>
<tr>
<th></th>
<th>Versus <em>Brassica oleracea</em></th>
<th>Versus <em>Antirrhinum majus</em></th>
<th>Versus <em>Silene latifolia</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>APETALA1</td>
<td>0.070 ± 0.010</td>
<td>0.365 ± 0.028</td>
<td>0.449 ± 0.034</td>
</tr>
<tr>
<td>CAULIFLOWER</td>
<td>0.140 ± 0.015</td>
<td>0.410 ± 0.031(^a)</td>
<td>0.458 ± 0.034(^a)</td>
</tr>
<tr>
<td>APETALA3</td>
<td>0.083 ± 0.012</td>
<td>0.437 ± 0.033</td>
<td>0.479 ± 0.036</td>
</tr>
<tr>
<td>PISTILLATA</td>
<td>NA(^b)</td>
<td>0.489 ± 0.038</td>
<td>0.538 ± 0.041</td>
</tr>
</tbody>
</table>

\(^a\) Estimates of *CAL* and *AP1* divergence between Arabidopsis thaliana and non-Brassicaceae species are not independent due to recent *AP1/CAL* duplication.

\(^b\) *Brassica oleracea PI* ortholog not available.
Table 4. Distribution of coding region differences between functional domains.

<table>
<thead>
<tr>
<th></th>
<th>APETALA1</th>
<th>CAULIFLOWER</th>
<th>APETALA3</th>
<th>PISTILLATA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Core</td>
<td>Noncore</td>
<td>Core</td>
<td>Noncore</td>
</tr>
<tr>
<td>Replacement</td>
<td>0</td>
<td>10</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>Silent</td>
<td>6</td>
<td>7</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ P = 0.019^* \] \[ P = 0.315 \] \[ P = 0.604 \] \[ P = 0.091 \]

* Significant at the 5% level.
Table 5. Comparison of Nonsynonymous Substitutions Between Core and Noncore Domains.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Domain</th>
<th>vs. Brassica oleracea</th>
<th>vs. Antirrhinum majus</th>
<th>vs. Silene latifolia</th>
</tr>
</thead>
<tbody>
<tr>
<td>APETALAI</td>
<td>core</td>
<td>0.004 ± 0.004</td>
<td>0.138 ± 0.025</td>
<td>0.184 ± 0.030</td>
</tr>
<tr>
<td></td>
<td>noncore</td>
<td>0.039 ± 0.011</td>
<td>0.263 ± 0.033</td>
<td>0.350 ± 0.040</td>
</tr>
<tr>
<td>APETALAI3</td>
<td>core</td>
<td>0.034 ± 0.012</td>
<td>0.178 ± 0.030</td>
<td>0.222 ± 0.034</td>
</tr>
<tr>
<td></td>
<td>noncore</td>
<td>0.017 ± 0.008</td>
<td>0.376 ± 0.040</td>
<td>0.395 ± 0.050</td>
</tr>
<tr>
<td>PISTILLATA</td>
<td>core</td>
<td>n.a.</td>
<td>0.178 ± 0.030</td>
<td>0.176 ± 0.030</td>
</tr>
<tr>
<td></td>
<td>noncore</td>
<td>n.a.</td>
<td>0.560 ± 0.070</td>
<td>0.570 ± 0.070</td>
</tr>
</tbody>
</table>
Figure 1. Structure of the plant MADS-box proteins. The functional core and noncore domains are indicated in brackets.
Figure 2.  

A. Schematic diagram of the Arabidopsis thaliana flower developmental pathway. The genes utilized in this study are indicated in boldface.  

B. Portions of the plant MADS-box gene family phylogeny, showing the relationships of the genes used in this study. Molecular-clock estimates of divergence times are indicated (Purugganan 1997).
Figure 3. Phylogenies of the (A) APETALA1 and CAULIFLOWER, (B) APETALA3, and (C) PISTIL-LATA genes. The branch lengths, estimated from neighbor-joining analysis, are indicated along each branch.
Figure 4. Amino acid site diversity profiles of paralogous floral homeotic proteins. The number of replacement substitutions was counted over a maximum-parsimony tree of the genes under study. The relative positions of the various protein structural domains are shown at the top of each profile. Regions of significant differences between the AP3 and PI paralog pair are shown in brackets ($P < 0.05$). There are no regions of significant difference between AP1 and CAL proteins.
CHAPTER 5

PATTERNS OF NUCLEOTIDE VARIATION IN HOMOELOGOUS REGULATORY GENES IN THE ALLOTETRAPLOID HAWAIIAN SILVERSWORD ALLIANCE (ASTERACEAE)

Amy Lawton-Rauh, Robert H. Robichaux*, and Michael D. Purugganan

*Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721

Genome-wide duplication (polyploidization) is prevalent in a large number of eukaryotic organisms and is particularly widespread in flowering plants. Polyploid species appear to vary from their diploid progenitors in a variety of ecologically important traits, suggesting that genome duplications provide a mechanism for ecological diversification. Studies of nucleotide variation at duplicate genes that arise via polyploidization allow us to infer the evolutionary forces that act on these polyploid loci. In an effort to examine the evolutionary dynamics of homoeologous loci, molecular population genetic analyses were undertaken for duplicate regulatory genes in the allopolyploid Hawaiian silversword alliance, a premier example of adaptive radiation. The levels and patterns of nucleotide variation for the floral homeotic genes ASAPETALAI (ASAP1) and ASAPETALAI3/TM6 (ASAP3/TM6) were studied in two species representing different major sub-lineages of the Hawaiian silversword alliance: Argyroxiphium sandwicense ssp. macrocephalum and Dubautia ciliolata ssp. glutinosa. While there is no evidence of a general pattern of positive selection at these developmental loci was detected in both species, homoeologous copies of ASAP1 and ASAP3/TM6 show differing levels and patterns of nucleotide polymorphism. Duplicate ASAP1 copies have similar levels of nucleotide diversity and haplotype divergence in both species; by contrast, duplicate ASAP3/TM6 genes display different levels and patterns of variation in Dubautia ciliolata ssp. glutinosa. Additionally, D. ciliolata ssp. glutinosa appears to be segregating for a moderate frequency null allele in one ASAP3/TM6 homoeolog. These results suggest that differing evolutionary forces can affect duplicate loci arising from allopolyploidization.
INTRODUCTION

Gene duplication may arise from tandem repetition, retrotransposon activity, segmental duplication, or genome-wide duplication (Paterson et al. 2000; Bennetzen 2002; Schmidt 2002). The evolutionary pathways of duplicate gene copies often result in the origin and diversification of gene families and appear to play an important role in organismal complexity and evolution (Ohno 1970). Genome duplication (polyploidization) can occur either within a species (autopolyploidization) or as a result of an unreduced interspecific hybridization between two species (allopolyploidization). Polyploidization has been shown to occur across several eukaryotic groups, including yeast (Coissac et al. 1997; Mewes et al. 1997; Wolfe & Shields 1997), invertebrates (Foighil & Thiriot-Quievreux 1999) and vertebrates (Spring 1997; Pebusque et al. 1998; McLysaght et al. 2002).

Polyploidization is particularly common in plants, with estimates suggesting that 40-70% of land plants have evolutionary histories that include at least one episode of genome doubling (Stebbins 1971; Masterson 1994; Leitch & Bennett 1997). Additional studies indicate that genome duplication continues to occur in plants (Grant 1981; Masterson 1994). Furthermore, calculations suggest that ~2-4% of angiosperm speciation events involve polyploidization, suggesting that genome duplication may be associated with species diversification (Otto & Whitton 2000). Several studies reveal that multiple, nested episodes of polyploidy occurred across modern angiosperms, thus the majority of flowering plants harbor remnants of genome-wide duplications of varying ages (Schmidt et al. 2001; Zhang et al. 2001; Mitchell-Olds & Clauss 2002). In the model plant Arabidopsis thaliana, for instance, there appear to have been at least four separate large-scale genome-wide duplications (Vision et al. 2000; Mitchell-Olds & Clauss 2002). Several of these duplication events occurred 100 – 200 million years ago, coincident with the Mesozoic era and possibly near the origin and expansion of early angiosperms (Vision et al. 2000).
Several important phenotypic traits differ markedly between polyploid plant taxa and their diploid progenitors. Examples of reproductive traits that may vary due to changes in ploidy level include the duration and initiation of flowering time, fertility, self-compatibility, apomixis, germination, and organ size (Jackson 1976; Lewis 1980; Lumaret 1988; Segraves & Thompson 1999; Schranz & Osborn 2000; Cook & Soltis 2000). Changes in vegetative traits such as growth rate and drought tolerance, as well as susceptibilities to arthropod infestation and fungal diseases, have also been documented (Jackson 1976; Lewis 1980; Lumaret 1988).

Recent studies regarding the immediate effects of genome duplication have shed light on the structural events and expression changes that take place following polyploidization. Significant genome restructuring and gene silencing can occur quite rapidly following polyploidization, as found in Arabidopsis thaliana allotetraploids (Comai et al. 2000), cotton (Zhao et al. 1998) and synthetic polyploids in Brassica (Song et al. 1995; Leitch & Bennett 1997) and in wheat (Feuillet et al. 2001; Ozkan et al. 2001). Evolutionary genetic models have explored alternate fates of duplicate loci, including neo- and sub-functionalization, redundancy and pseudogene formation (Ohno 1970; Walsh 1995; Force et al. 1999; Lynch & Conery 2000). Moreover, a recent study in allopolyploid cotton indicates that duplicate gene copies can accumulate different levels of nucleotide variation, and that differential selection with respect to genomic origin of each gene duplicate is possible (Small & Wendel 2002). The above examples suggest significant short-term and long-term genetic consequences following gene duplication. The dramatic effects of gene duplication by polyploidization on phenotypic and molecular evolution emphasize the need for further studies in different organismal and ecological contexts.
Floral regulatory genes

The molecular genetics of floral and inflorescence development has been the subject of intense study (Weigel 1995; Yanofsky 1995; Liljegren & Yanofsky 1996). Much of this work has focused on *Arabidopsis thaliana* (Brassicaceae), where several genes controlling various aspects of flower development have been identified. Mutational lesions in these genes result in the formation of aberrant floral organ types, thus these loci are referred to as floral homeotic genes (Bowman *et al.* 1991). Among the best characterized floral homeotic genes are the floral meristem identity gene *APETALA1* (*AP1*) (Mandel *et al.* 1992; Bowman *et al.* 1993; Gustafson-Brown *et al.* 1994; Irish 1998) and the floral organ identity gene *APETALA3* (*AP3*) (Bowman *et al.* 1991; Jack *et al.* 1992; Irish & Yamamoto 1995). Both *AP1* and *AP3* are transcription factors and are members of the extensive MADS-box gene family (Riechmann & Meyerowitz 1997; Alvarez-Buylla *et al.* 2000).

The *AP1* gene in *Arabidopsis thaliana* controls floral meristem identity as well as sepal and petal organogenesis (Bowman *et al.* 1993; Irish 1998). Mutants at this locus display a loss of sepal and petal formation and a conversion of floral meristems into partial inflorescence-like shoots (Bowman *et al.* 1993; Irish 1998). The *Arabidopsis* *AP3* gene is a floral organ identity gene that regulates petal and stamen development. *AP3* mutants display a homeotic transformation of these organs to sepaloid and carpelloid structures (Bowman *et al.* 1991). Homologs of *AP1* and *AP3* have been discovered in other plant taxa, indicating that these genes have relatively conserved functions among various angiosperm lineages. The function of the Solanaceae *AP3*-like paralog *Lycopersicon TM6* gene is unknown, although its expression pattern is consistent with roles in petal and stamen development (Pnueli *et al.* 1991; Kramer *et al.* 1998). Asteraceae orthologues of *AP1* and *AP3* have been isolated in *Gerbera hybrida* and appear to perform functions that are similar to their *Arabidopsis* counterparts (Yu *et al.* 1999).
The Hawaiian silversword alliance

The Hawaiian silversword alliance (Asteraceae, Heliantheae, Madiinae) is a premier example of plant adaptive radiation (Robichaux et al. 1990; Baldwin & Robichaux 1995; Baldwin 1997). The alliance includes 30 perennial species in three endemic genera: *Argyroseriphium*, *Dubautia* and *Wilkesia* (Carr 1985; Robichaux et al. 1990; Baldwin & Robichaux 1995; Baldwin 1997). These species are distributed on six of the eight main islands of the Hawaiian archipelago (Kaua‘i, Oahu, Moloka‘i, Lana‘i, Maui and Hawai‘i), with all but five species occurring as single-island endemics. Members of the alliance also display a wide array of habitat preferences and morphological growth forms, including a striking diversity in inflorescence and floral architectures (Carr 1985). The periodic availability of new habitats (due to close proximity, formation, and weathering of the Hawaiian islands) and extensive chromosomal evolution are believed to have contributed to the adaptive radiation of these taxa (Robichaux et al. 1990; Wagner & Funk 1995).

The Hawaiian silversword alliance descended from the *Madia* lineage of the North American tarweeds (Asteraceae, Heliantheae, Madiinae), which is known as the tarweeds (Robichaux et al. 1990; Baldwin et al. 1991; Baldwin & Robichaux 1995; Baldwin 1997). The onset of diversification in this adaptive radiation appears to have occurred within the last 5.2 ± 0.8 million years, based on a calibrated, rate-constant ITS tree (Baldwin & Sanderson 1998). Evolutionary analysis of the Hawaiian silversword alliance indicates that the species are allopolyploids that apparently originated from an interspecific hybridization event between two North American tarweed species (Barrier et al. 1999). Orthologues of several *Arabidopsis thaliana* floral developmental loci have been isolated in the Hawaiian silversword alliance, including orthologs of the *Arabidopsis thaliana* genes *APETALA1* (*ASAPI*) and *APETALA3* (*ASAP3/TM6*) genes (Barrier et al. 1999). Phylogenetic analyses of *ASAPI* and *ASAP3/TM6* indicate that both
genes are present as two copies (homoeologs) derived from separate North American tarweed lineages. The A homoeolog originated from the ancestral lineage of *Anisocarpus scabridus* and the B homoeolog appears to derive from an ancestral lineage including *Carlquistia muirii* (Barrier et al. 1999). Both homoeologs of each isolated floral regulatory gene (*ASAP1-A, ASAP1-B, ASAP3/TM6-A, and ASAP3/TM6-B*) are expressed in the Hawaiian species (Barrier et al. 2001). Additionally, molecular evolutionary analysis of these homoeologous regulatory genes suggests an increase in the rate of protein evolution in the Hawaiian silversword alliance species compared to the North American tarweed species (Barrier et al. 2001).

In this study, we examined the effects of polyploidy on regulatory gene evolution in a plant adaptive radiation by investigating the molecular population genetics of homoeologous floral regulatory genes in two species of the Hawaiian silversword alliance. This study addresses the following questions: (i) Do homoeologous regulatory genes harbor different levels and patterns of nucleotide variation? (ii) Is an excess of replacement substitutions in regulatory genes across the Hawaiian silversword alliance reflected in a non-neutral pattern of within- and between-species protein evolution? (iii) Is there a progenitor genome-specific pattern of sequence variation in homoeologous regulatory genes?

The levels and patterns of nucleotide variation were analyzed for the *ASAP1* and *ASAP3/TM6* regulatory genes in *Argyroserinum sandwicense* ssp. *macrocephalum* and *Dubautia ciliolata* ssp. *glutinosa*. Our results suggest that evolutionary forces are similar among the *ASAP1* homoeologs in both species. In contrast, the *ASAP3/TM6* homoeologs in *D. ciliolata* ssp. *glutinosa* appear to have different levels of nucleotide diversity and haplotype structure, suggesting that evolutionary forces have acted differently between the duplicate *ASAP3/TM6* copies in this allopolyploid species. These results demonstrate that differential evolution between homoeologous copies may accompany duplicate gene evolution in plant species undergoing adaptive radiation.
MATERIALS AND METHODS

Sample collections and DNA extraction

_Dubautia ciliolata_ ssp. _glutinosa_ leaf tissue samples were collected from 10-24 randomly selected individuals from three localities on Mauna Kea, Hawaii: Puu Kanakaleonui, Puu Kawiiwi and adjacent to Waipahoehe gulch as indicated in Table 1. Leaf tissue samples were also collected from 15-16 individuals of _Argyroxyphium sandwicense_ ssp. _macrocephalum_ obtained from four localities on Haleakala, Maui: Silversword Loop, Puu o Pele, Ka moa o Pele, and Puu Naue (Table 1).

The lineages containing these two taxa represent different major sub-lineages and apparently diverged relatively early during the radiation of the Hawaiian silversword alliance (Baldwin & Wessa 2000). _Argyroxyphium sandwicense_ ssp. _macrocephalum_ and _D. ciliolata_ ssp. _glutinosa_ have strongly contrasting growth forms and reproductive characteristics. Both taxa grow in dry, high-elevation habitats and are endemic to two different younger islands of the Hawaiian archipelago ( _Argyroxyphium sandwicense_ ssp. _macrocephalum_ on Maui; _Dubautia ciliolata_ ssp. _glutinosa_ on Hawaii) [Baldwin & Robichaux 1995; Carr 1985]. These species are referred to as _D. ciliolata_ and _A. sandwicense_ for the remainder of this paper.

Genomic DNA was extracted from leaf tissue using a general rapid plant CTAB protocol. These genomic DNA extractions were purified with an EluQuick glass bead purification procedure (Schleichter & Schuell) to reduce the amount of pectin and secondary-product contamination.
PCR amplifications and sequencing

DNA fragments for both copies (homoeologs) of ASAP1 and ASAP3/TM6 were obtained via PCR amplification with the error-correcting Pwo polymerase (Roche, Indianapolis). The error-correcting Pwo polymerase was utilized to minimize PCR errors due to nucleotide misincorporation. Data from multiple, independent amplifications and re-sequencing of several genes indicates an error rate of less than one error in 7–10 kb (unpublished observations).

Gene-specific primers for ASAP1 and ASAP3/TM6 were designed based on cDNA sequences as described in Barrier et al (1999). Homoeologous copies of ASAP1 (ASAP1-A and ASAP1-B) spanning exons 3 to 8 were amplified using primers AP1-3X (5’-CTGGACCATGGAGTACAACAAAC-3’) and AP1-8XR2 (5’-ATCGGCTGCAGAC-TCAGGTC-3’) [Fig. 1]. The PCR conditions were generally as follows: 94°C for a 5 min. hot-start, followed by 10 cycles of 94°C for 30 sec., 52°C for 30 sec., 68°C for 2 min., then 20 cycles of 94°C for 30 sec., 52°C for 30 sec., and 68°C for 2 min. with a 20 sec. per cycle auto-incrementation, and a final 7 min. extension step at 68°C. The ~1.2 kb PCR products were purified using either Spin-X and Ultrafree-MC filter purification (Millipore) or a QiaQuick extraction from agarose gels (Qiagen). Amplified product bands were cloned using either TA cloning or Zero Blunt TOPO TA cloning kits (Invitrogen, San Diego). Restriction enzyme digests on cloned ASAP1 gene fragments were utilized to screen for each homoeologous copy (ASAP1-A versus ASAP1-B) from multiple clones per individual.

For the ASAP3/TM6-A and -B copies, primers APETALA3-3 (5’-TACAAACAGGC-AGGTGACATCTC-3’) and ASAP3-2R4 (5’-CTGCTGCTCGAGAATGGTTAGATC-3’) were used to simultaneously amplify a region spanning exons 1–4 from both homoeologs (Fig. 2). The following general PCR amplification conditions were utilized: 94°C for a 5 min. hot-start followed by 10 cycles of 94°C for 15 sec., 55°C for 30 sec., 72°C for 1 min., then 20 cycles of 94°C for 15
sec., 55°C for 30 sec., and 72°C for 1 min. with a 5 sec. per cycle auto-incrementation, followed by a final 7 min. extension step at 72°C. Amplified PCR products were isolated and cloned using the same protocol reported above for the ASAP1 homoeologs. The homoeologous ASAP3/TM6 copies are significantly different in size (1.1 kb vs. 1.4 kb for ASAP3/TM6-A and -B, respectively), and thus can be readily differentiated by fractionation in a 1.5% agarose gel.

All genes were sequenced using automated DNA sequencers at various facilities (NCSU DNA Sequencing Facility, Iowa State University Sequencing Facility, NCSU Genome Research Laboratory) using a nested primer series designed in both directions. Sequences were aligned and visually refined. Polymorphic sites were confirmed by visual inspection of chromatograms. The DNA sequences are available from GenBank (accession numbers in progress).

Sequence data analyses

Sequence variation was estimated at silent sites using nucleotide diversity ($\pi_{\text{silent}}$) (Nei 1987) and at all sites using the population mutation parameter ($\theta_W$) (Watterson 1975) in DnaSP version 3.53 (Rozas & Rozas 1999). Coalescent simulations with 2000 replications were conducted to calculate the confidence interval of $\theta_W$. Haplotype networks were constructed using the 95% statistical parsimony support criterion for inferred estimated gene genealogies (Templeton et al. 1992) as implemented in the TCS program (Clement et al. 2000). The Hudson-Kreitman-Aquade (HKA) test for selection was performed using the number of segregating sites and average number of differences for each locus (Hudson et al. 1987).

The neutral-equilibrium model predicts the expected frequency of polymorphisms in an allelic sample and Tajima’s test of selection examines deviations from this neutral expectation (Tajima 1989). This test of selection calculates the statistic $D$; loci evolving
neutrally have D values equal to zero, while positive D values suggest an excess of intermediate frequency polymorphism. Negative values of Tajima’s D indicate an excess of low-frequency polymorphisms, which may arise from a recent selective sweep or demographic factors such as population expansion. Tajima's test (Tajima 1989) for selection was performed, and significance assessed based on coalescent simulations of 10,000 runs using the number of segregating sites and the estimated recombination parameter. The population recombination parameter was calculated from the data using SITES (Hey & Wakeley 1997). The McDonald-Kreitman test (McDonald & Kreitman 1991) was performed to test for neutral evolution in protein-coding regions using intraspecific data from *D. ciliolata* and designation of *A. sandwicense* as outgroup for interspecific divergence comparison.
RESULTS

Nucleotide variation at homoeologous ASAPETALAI loci

The levels and patterns of nucleotide variation provide information on the pattern and type of evolutionary forces that have acted on genes. Alleles for homoeologs of ASAP1 and ASAP3/TM6 were isolated and sequenced from D. ciliolata individuals sampled among three Mauna Kea, Hawaii localities and from A. sandwicense individuals sampled among four Haleakala, Maui localities (see nlocality, Table 1).

Molecular variation was estimated for a region spanning exons 3 to 8 (and the intervening five introns) of each ASAP1 homoeolog (Fig. 1). The amplified portions of both genes include the coding sequences for the K-box and C-terminal domains of the encoded MADS-box transcription activator. Sequence analysis indicates that the D. ciliolata ASAP1-A alleles from 20 individuals form 11 haplotypes with 27 segregating sites and one 1-bp insertion/deletion (indel) polymorphism (Table 1 and Fig. 1). Among the polymorphic sites in exons, six are replacement mutations and none is a synonymous mutation. Silent-site nucleotide diversity \((\pi_{\text{silent}})\) at this locus is 0.0035, which is comparable to the nucleotide diversity levels reported for the orthologous AP1 gene in Arabidopsis thaliana \([\pi_{\text{silent}} = 0.004]\) (Olsen et al. 2002) and similar to the estimated population mutation parameter \((\theta_W = 0.0046)\). The A. sandwicense ASAP1-A alleles from 15 individuals fall into nine haplotypes and have 21 segregating sites (Table 1 and Fig. 1). There is one replacement and one synonymous polymorphism in the coding region of this gene. Silent site nucleotide diversity at this locus is 0.0044, comparable to the estimate for this gene in D. ciliolata as well as the population mutation parameter (0.0038).

The levels of nucleotide variation appear to be lower for the B homoeolog of the ASAP1 gene than for the A homoeolog. Sequence analysis of ASAP1-B alleles from the
24 *D. ciliolata* individuals indicates that this locus has nine haplotypes and 13 segregating sites (Table 1 and Fig. 1). The protein coding region of *ASAP1-B* has one replacement and no synonymous polymorphisms. There are no insertion/deletion polymorphisms. Silent site nucleotide diversity (πsilent) at this locus is 0.0019, which is similar to the population mutation parameter (θW = 0.0020) and is apparently about half the value of πsilent for *ASAP1-A* in this species. The *A. sandwicense* *ASAP1-B* allele dataset from 15 sampled individuals includes 11 haplotypes and 18 segregating sites (Table 1 and Fig. 1). Among protein coding region polymorphic sites, two are replacement changes and none is a synonymous change. Silent site nucleotide diversity (πsilent) at this locus is 0.0029, which is similar to the population mutation parameter (θW = 0.0032) and apparently two-thirds the value of πsilent for *ASAP1-A* in this species.

*Nucleotide variation at homoeologous ASAPETALA3/TM6 loci*

Nucleotide variation was examined in a region spanning exons 1 to 4 of each ASAP3/TM6 homoeolog. The amplified portion of both genes includes the protein coding region for the DNA-binding MADS-box domain of this transcriptional activator and a portion of the K-box coding sequence (Figure 2). ASAP3/TM6-A alleles were sequenced from 21 individuals of *D. ciliolata* and from 16 individuals of *A. sandwicense* (Table 1). Nucleotide sequence analysis of the *D. ciliolata* ASAP3-A alleles reveals 13 haplotypes and 22 segregating sites, with three 1-bp indels (Table 1 and Fig. 2). There are one replacement and two synonymous polymorphisms in the protein coding region of this gene. Silent site nucleotide diversity (πsilent) at this locus is 0.0052, which is similar in magnitude to the estimated population mutation parameter for this gene (θW = 0.0064) and lower than the value reported for *AP3* in *Arabidopsis thaliana* [πsilent = 0.0076] (Olsen *et al.* 2002). The *A. sandwicense* ASAP/TM6-A allele sequences are distributed among 8 haplotypes and have 12 segregating sites (Table 1 and Fig. 2). There is one replacement and no synonymous mutations in the exons of this gene. Silent site
nucleotide diversity ($\pi_{\text{silent}}$) at this locus is 0.0031, which is similar to the population mutation parameter value ($\theta_w = 0.0037$).

Alleles from the homoeologous ASAP3/TM6-B gene were sequenced from ten individuals of *D. ciliolata* and from 16 individuals of *A. sandwicense*. We were unable to amplify the B homoeolog of ASAP3/TM6 from several additional *D. ciliolata* individuals. In an effort to determine if this copy was present, we utilized a saturated PCR approach, using 41 primer combinations throughout the gene. This PCR-based approach to determine the presence of this copy was necessitated by the difficulties in conducting Southern blot analyses within Madiinae species, which apparently have large genomes and high DNA methylation levels (Barrier et al. 1999). The approach utilized in this study was used to successfully co-amplify both ASAP3/TM6 homoeologs (*ASAP3/TM6-A* and *–B*) in many other species across the Hawaiian silversword alliance (Barrier et al. 1999). This co-amplification strategy was not consistently successful, however, for amplification of both ASAP3/TM6 homoeologs in *D. ciliolata*. Whereas *ASAP3/TM6-A* was successfully amplified using most of the 41 primer combinations in several samples, amplification of *ASAP3/TM6-B* from 14 out of 24 individuals (42% success rate) in *D. ciliolata* was not successful. Taken together, this suggests that *ASAP3/TM6-B* is deleted in the plants from which we were unable to amplify both ASAP3/TM6 homoeologs.

The *D. ciliolata* ASAP3/TM6-B allele sequences we were able to obtain have the lowest diversity of the genes in this study, with only three haplotypes and two segregating sites, of which one segregating site is a singleton (Table 1 and Fig. 2). There are no polymorphisms in the protein-coding region of this gene, and polymorphisms in *D. ciliolata* ASAP3/TM6-B are in introns 2 and 3 (Figure 2). Silent site nucleotide diversity ($\pi_{\text{silent}}$) at this locus is 0.0005, which is apparently 4-10 fold lower than calculations in other currently sequenced genes in this species. This pattern in $\pi_{\text{silent}}$ is further reflected in calculated $\theta_w$ values, where the 95% confidence intervals do not overlap between ASAP3/TM6-A and ASAP3/TM6-B values for *D. ciliolata*. 
Silent site nucleotide diversity at *A. sandwicense* ASAP3/TM6-B is also low, but not nearly as low as the estimate for this gene in *D. ciliolata*. In the *A. sandwicense* ASAP3/TM6-B allele sample, there are five haplotypes and five segregating sites (Table 1 and Fig. 2). Among protein coding region polymorphic sites, there are two replacement polymorphisms in exon 2, with no observed synonymous variation. Silent site nucleotide diversity ($\pi_{\text{silent}}$) at this locus is 0.0014, which is comparable to the calculated Watterson’s population mutation parameter ($\theta_W$) [Table 1].

**Haplotype networks of floral regulatory genes**

The statistical parsimony-based haplotype networks of *D. ciliolata* and *A. sandwicense* for each homoeologous gene pair are displayed in Figures 3 and 4. The ASAP1-A and ASAP1-B haplotype networks reveal separate clades for alleles in *D. ciliolata* and *A. sandwicense*, indicating several fixed differences between species haplotypes and the lack of shared polymorphisms (Figure 3). The inferred haplotype networks of ASAP3/TM6 homoeologs also show reciprocal monophyly between species, with no shared haplotypes between the two taxa (Figure 4). The ASAP3/TM6-A haplotype network shows non-linear connections among several *D. ciliolata* haplotypes, reflecting the moderate level of recombination at this locus that was also detected in our analyses of recombination (results not shown).

**Reduced variation in the ASAPETALA3/TM6-B gene**

The Hudson-Kreitman-Aguade (HKA) test compares levels of intraspecific polymorphism and interspecific divergence among loci to determine these have significantly different levels of nucleotide variation (Hudson et al. 1987). This test can be applied to homoeologous gene copies to examine whether loci that arise from
allopolyploidization are subject to contrasting evolutionary forces. The HKA test results in this study are displayed in Table 2.

The level of silent site diversity ($\pi_{\text{silent}}$) is almost two-fold higher for $\text{ASAP1-A}$ compared to $\text{ASAP1-B}$ in $D. \text{ciliolata}$ (Table 1). Based on HKA test results, however, this difference in diversity is not significant for this homoeologous gene pair ($P > 0.10$, Table 2). A similar trend is observed between the homoeologous $\text{ASAP1}$ gene copies in $A. \text{sandwicense}$. Silent site nucleotide diversity is approximately 1.5 times higher in $\text{ASAP1-A}$ compared to $\text{ASAP1-B}$ (Table 1) but this difference is not significant in an HKA test ($P > 0.10$, data not shown).

The results for $\text{ASAP1}$ in both species are in contrast to the results for $\text{ASAP3/TM6}$ loci. Sequence analysis reveals a difference in silent site nucleotide diversity levels between the $\text{ASAP3/TM6}$ homoeologs obtained from $D. \text{ciliolata}$, with the $\text{ASAP3/TM6-B}$ homoeolog having an apparent 10-fold lower nucleotide diversity compared to the $\text{ASAP3/TM6-A}$ homoeolog (Table 1). This pattern is further reflected in $\theta_W$ values of $\text{ASAP3/TM6-A}$ and $\text{ASAP3/TM6-B}$, which differ by an order of magnitude and have non-overlapping 95% confidence intervals (Table 1). The HKA test between $\text{ASAP3/TM6-A}$ and -$B$ genes in $D. \text{ciliolata}$ indicates a significant deviation in sequence variation between these two loci ($P < 0.05$, Table 2). This significant difference appears to result in part from the observed number of polymorphisms being 144% of the expected value in $\text{ASAP3/TM6-A}$ and 37% of the expected value in $\text{ASAP3/TM6-B}$ (Table 2).

As in $D. \text{ciliolata}$, the $\text{ASAP3/TM6}$ gene pair also shows a difference in levels of silent site nucleotide diversity in $A. \text{sandwicense}$. The $\text{ASAP3/TM6-B}$ homoeolog is approximately 2.2-fold lower in diversity compared to the $\text{ASAP3/TM6-A}$ homoeolog (Table 1). Unlike in $D. \text{ciliolata}$, however, the HKA test for these two loci using $A. \text{sandwicense}$ intraspecific polymorphism is not significant ($P > 0.10$, data not shown).

*Excess low-frequency polymorphisms at floral regulatory loci*
Almost all genes examined in this study, with the exception of *A. sandwicense ASAP3/TM6-B*, have negative Tajima's D values. Observed Tajima’s D values range from –0.777 to –0.110 for the *ASAP1* homoeologs (Table 1). Negative values are also observed for *ASAP3/TM6-A* in *D. ciliolata* (-1.257) and *A. sandwicense* (-0.578), and for *ASAP3/TM6-B* in *D. ciliolata* (-0.691) [Table 1]. The only observed positive Tajima’s D value among all loci examined in both of these species is the one estimated for the *A. sandwicense ASAP3/TM6-B* homoeolog. Of all genes in this study, however, only the *ASAP3/TM6-A* gene in *D. ciliolata* has an excess of low-frequency variants that is significantly different from the neutral-equilibrium model expectations (*P* < 0.05). Although not significantly different from neutral expectations, negative Tajima's D values in all but one of the floral regulatory genes in these two species suggests a trend of excess low-frequency polymorphisms at most loci. It should be noted, however, that the small number of polymorphisms in some samples (e.g., *D. ciliolata ASAP3/TM6-B*) reduces the power of this test to detect signatures of selection.

*Evolution of protein coding regions*

Barrier *et al* (2001) found higher ratios of nonsynonymous (Ka) to synonymous (Ks) mutations in the coding regions of the floral regulatory genes in the Hawaiian silversword alliance than in the North American tarweeds. Among the Hawaiian species, many of the pairwise interspecific Ka/Ks values for the *ASAP1* and *ASAP3/TM6* loci were greater than 1, which suggests that selection and adaptive divergence may have operated to shape the structure of these loci. Selection in protein-coding regions can also be inferred with population-level sequence data using the McDonald-Kreitman test, which examines the relative levels of within- and between-species replacement to synonymous changes (McDonald & Kreitman 1991). None of the McDonald-Kreitman tests, however, was significant (G-tests, *P* = 0.16 to 1.00).
DISCUSSION

It has been proposed that the increased number of possible gene products encoded by duplicate gene copies may lead to potential increases in genetic adaptability in polyploid species, particularly allopolyploid taxa, relative to their diploid progenitors (Leitch & Bennett 1997). Polyploidization has been postulated as playing a significant role in diversification, including the evolution of HOX clusters in early vertebrates (Malago-Trillo 2001), the addition to multigene families such as the rbcS family (Sasanuma 2001) and the radiation of early vertebrates during the Cambrian explosion (Spring 1997; Valentine 1999; Miyata 2001). Furthermore, regulatory and structural gene evolution may be quickly uncoupled in polyploid taxa due to an accelerated rate of mutational accumulation at duplicate regulatory genes relative to duplicate structural genes (Ferris & Whitt 1979). Taken together, the increased availability of unique gene product combinations and an accelerated rate of evolution in duplicated regulatory genes may play significant roles in the adaptive diversification of naturally occurring polyploids.

We examined patterns of variation in homoeologous copies of the floral regulatory genes ASAP1 and ASAP3/TM6 in two species of the allopolyploid Hawaiian silversword alliance, D. ciliolata and A. sandwicense. Orthologues to these genes in other dicot species suggest that these loci control inflorescence and floral development (Ng & Yanofsky 2001). Species in the Hawaiian silversword alliance differ greatly in inflorescence and floral morphologies (Carr 1985), suggesting that floral regulatory gene evolution may underlie the diversification of reproductive morphologies in this group. Indeed, a previous study indicates accelerated protein evolution at these regulatory loci within the Hawaiian silversword alliance (Barrier et al. 2001). The range of morphological and physiological traits found in the Hawaiian silversword alliance is remarkable given the relatively low genetic distances among these taxa. Population studies, for example, indicate very little genetic differentiation in 11 allozyme loci between morphologically distinct species in the Hawaiian silversword alliance (Witter &
Carr 1988). This suggests that relatively few genetic changes of large phenotypic effect, possibly at regulatory loci, may be responsible for the wide phenotypic range in these allopolyploid taxa. Such a pattern would be consistent with an extensive analysis of multilocus isozymes in several polyploid catostomid fish species which indicates that polyploidization may be followed by an accelerated rate of mutational accumulation in regulatory genes (Ferris & Whitt, 1979).

Our molecular population genetic data, however, do not reveal any evidence of consistent positive selection accompanying the evolution of these loci within- and between two species representing different major sub-lineages of the Hawaiian silversword alliance. Except for a few exceptions, neither the Tajima's nor the HKA tests for selection detected any significant deviation from predictions of the neutral-equilibrium model of molecular evolution. Moreover, while there appears to be an excess of replacement substitutions for these regulatory genes across the entire Hawaiian silversword alliance (Barrier et al. 2001), this is not reflected in the patterns of protein evolution within- and between-species. None of the McDonald-Kreitman tests for protein evolution indicates significant deviations from neutral expectations (unpublished results).

While not significant, the negative Tajima’s D values calculated for most loci studied suggest a trend of excess low-frequency polymorphisms. This general excess of low-frequency polymorphisms, however, suggests that it is a genome-wide signature that is consistent with the effects of demographic changes such as rapid population expansion in these species. Rapid population expansion, possibly occurring as a result of colonization of new island habitats, may thus have contributed to the patterns of allelic diversity observed in *A. sandwicense* and *D. ciliolata*.

There do appear, however, to be differences in the patterns of evolution between some homoeologous floral regulatory gene copies. There is a marked difference in levels of nucleotide diversity as well as the number and distribution of haplotypes between
ASAP3/TM6 homoeologs within D. ciliolata, suggesting that the duplicate copies of this floral regulatory gene have been subject to different evolutionary forces (Table 2, Figure 4B). There is evidence that a deletion allele is segregating at moderate frequency in D. ciliolata, leading to only one ASAP3/TM6 copy in several individuals of this species. This segregation pattern does not appear to be confined to specific populations; the ASAP3/TM6-B alleles that were successfully amplified from D. ciliolata ssp. glutinosa were found in all three sampling localities (see n_{locality}, Table 1). Additionally, microsatellite-based studies, including samples from the localities used in this study, reveal no evidence of population structure D. ciliolata (E. Friar, personal communication).

In the extant ASAP3/TM6-B alleles of D. ciliolata, however, there is an apparent 10-fold reduction in the levels of silent site nucleotide variation for this copy compared to ASAP3/TM6-A (\(\pi_{\text{silent}} = 0.0005\) vs. 0.0052, respectively). Although the sample size of D. ciliolata ASAP3/TM6-B alleles is reduced due to the putative deletion allele, this should not significantly affect estimates of nucleotide variation levels since of \(\pi_{\text{silent}}\) and \(\theta_W\) estimates are corrected for sample size variation. The HKA test of selection comparing both copies indicates that the observed reduction in nucleotide polymorphism levels for ASAP3/TM6-B in D. ciliolata is significant compared to the A homoeolog of the same locus (Table 2). The difference in levels and patterns of nucleotide diversity between these two genes is also reflected in the haplotype networks of these two loci, which display a reduced number and decreased diversity of ASAP3/TM6-B haplotypes relative to ASAP3/TM6-A haplotypes (see Fig. 4).

This difference between ASAP3/TM6-B homeologs is found only in D. ciliolata. In A. sandwicense, there is a 2.5-fold reduction in silent site diversity between the ASAP3/TM6-A and -B copies (\(\pi_{\text{silent}} = 0.0036\) vs. 0.0014, respectively). Unlike in D. ciliolata, however, this difference is not significant between the two gene copies based on an HKA test (data not shown). Moreover, this pattern of differential variation in ASAP3/TM6 homoeologs in D. ciliolata is in contrast to the ASAP1 loci in both D.
*ciliolata* and *A. sandwicense*. In both of these species, there are slight reductions in the levels of silent site nucleotide diversity ($\pi_{\text{silent}}$) for *ASAP1-B* compared to *ASAP1-A*, but these differences are not significant in an HKA test [Table 2].

The reduced levels of variation observed for the *ASAP3-B* homoeolog alleles segregating in *D. ciliolata* may arise from positive selection on these alleles (Fay & Wu 2000; Charlesworth *et al.* 2001). If extant *ASAP3/TM6-B* alleles are evolving neutrally, then these alleles would have variation levels comparable to other neutral loci in this species. It may be that the deletion allele for *ASAP3/TM6-B* is being selected against, and that the extant alleles are in the process of sweeping through *D. ciliolata*, resulting in decreased variation at *ASAP3/TM6-B* in this species. Physical mapping of the putative deletion and an investigation of the levels and patterns of variation in flanking sequences may help clarify the evolutionary dynamics of this locus in *D. ciliolata*.

The patterns of molecular evolution within these floral regulatory loci in allopolyploid Hawaiian silversword alliance species can be compared to patterns observed between homoeologous genes in allopolyploid cotton taxa. Molecular evolutionary studies of 16 homoeologous gene pairs in polyploid and diploid *Gossypium* species indicate that there are no significant differences in rates of evolution between homoeologous gene copies following allopolyploidization (Cronn *et al.* 1999). Molecular population genetic studies in these species, however, indicate that genome-specific differential evolution can occur between homoeologous genes (Small & Wendel 2002). Within-species pairwise comparisons of *AdhA* (from the A progenitor genome) and *AdhC* (from the D progenitor genome) from two independently-derived allopolyploid species show a consistent difference correlated with genome origin. Relative rate tests also suggest that *AdhC* is evolving more quickly than *AdhA* (Small & Wendel 2002). Additionally, a higher nonsynonymous diversity in *AdhC* implies relaxed selection on the D-subgenome and purifying selection on the A-subgenome in *Gossypium hirsutum* and another allotetraploid *G. barbadense*. 

132
In contrast to results in allopolyploid *Gossypium*, our study indicates that
differential evolution between homoeologs is not necessarily associated with the genome
origins of the duplicate loci. We found contrasting patterns of evolution between two
sets of duplicate floral regulatory genes in two species of the Hawaiian silversword
alliance. The homoeologous copies of *ASAP1* within both *A. sandwicense* and *D.
ciliolata* do not have significantly different levels and patterns of nucleotide variation,
suggesting that similar evolutionary forces have acted on both gene copies. In contrast,
*D. ciliolata ASAP3/TM6* homoeologs show evidence of differential evolution between
gene copies. This pattern, however, is not observed for the same two homoeologous loci
in *A. sandwicense*, suggesting a species-specific effect associated with differential
evolution.

It remains unclear to what extent genome restructuring as a result of
polyploidization continues to occur in these species, given that the allopolyploidization
event likely occurred more than 5.2 ± 0.8 mya (Baldwin & Sanderson 1998; Barrier *et al*
1999). Cytogenetic studies suggest continuing chromosomal repatterning, with eight
genomic rearrangements distinguished by reciprocal translocations and an aneuploid
reduction found between several species (Carr & Kyhos 1981, 1986). More studies will
be needed to unravel the precise relationship between polyploidization, the diversification
of homoeologous regulatory gene copies and the adaptive radiation of the Hawaiian
silversword alliance.
ACKNOWLEDGEMENTS

The authors thank E. Friar, D. Remington and members of the Purugganan lab for insightful discussions and fieldwork assistance. *Argyroxiphium sandwicense* ssp. *macrocephalum* leaf tissue samples from Puu o Pele, Ka moa o Pele, and Puu Naue localities were provided by E. Friar. This work is supported in part by a grant from the National Science Foundation (to MDP and RHR), an Alfred P. Sloan Foundation Young Investigator Award (to MDP) and a Sigma Xi Grant-in-Aid of Research Award (to ALR).
LITERATURE CITED


Feuillet C, Penger A, Gellner K et al. (2001) Molecular evolution of receptor-like kinase genes in hexaploid wheat. Independent evolution of orthologs after polyploidization and


Table 1. Molecular variation in *ASAPETALAI* and *ASAPETALAI3/TM6* homoeologs (*A* and *B*) from *Dubautia ciliolate* ssp. *glutinosa* and *Argyroxyphium sandwicense* ssp. *macrocephalum*.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Species</th>
<th>Length (bp)</th>
<th>n</th>
<th>nlocality</th>
<th>Nhap</th>
<th>S</th>
<th>π silent</th>
<th>θw (95% C.I)</th>
<th>Tajima’s D</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASAP1-A</td>
<td><em>D. ciliolata</em></td>
<td>1661</td>
<td>20</td>
<td>8/6/6</td>
<td>11</td>
<td>27</td>
<td>0.0035</td>
<td>0.0046 (0.0019, 0.0085)</td>
<td>-0.649 (P &gt; 0.10)</td>
</tr>
<tr>
<td></td>
<td><em>A. sandwicense</em></td>
<td>1696</td>
<td>15</td>
<td>8/3/1/3</td>
<td>9</td>
<td>21</td>
<td>0.0044</td>
<td>0.0038 (0.0015, 0.0073)</td>
<td>-0.110 (P &gt; 0.10)</td>
</tr>
<tr>
<td>ASAP1-B</td>
<td><em>D. ciliolata</em></td>
<td>1736</td>
<td>24</td>
<td>7/12/5</td>
<td>9</td>
<td>13</td>
<td>0.0019</td>
<td>0.0020 (0.0006, 0.0041)</td>
<td>-0.617 (P &gt; 0.10)</td>
</tr>
<tr>
<td></td>
<td><em>A. sandwicense</em></td>
<td>1709</td>
<td>15</td>
<td>8/2/2/3</td>
<td>11</td>
<td>18</td>
<td>0.0029</td>
<td>0.0032 (0.0011, 0.0068)</td>
<td>-0.777 (P &gt; 0.10)</td>
</tr>
<tr>
<td>ASAP3/</td>
<td><em>D. ciliolata</em></td>
<td>949</td>
<td>21</td>
<td>8/8/5</td>
<td>13</td>
<td>22</td>
<td>0.0052</td>
<td>0.0064 (0.0032, 0.0103)</td>
<td>-1.257 (P &lt; 0.05)</td>
</tr>
<tr>
<td>TM6-A</td>
<td><em>A. sandwicense</em></td>
<td>1012</td>
<td>16</td>
<td>8/4/2/2</td>
<td>8</td>
<td>12</td>
<td>0.0031</td>
<td>0.0037 (0.0015, 0.0063)</td>
<td>-0.578 (P &gt; 0.10)</td>
</tr>
<tr>
<td>ASAP3/</td>
<td><em>D. ciliolata</em></td>
<td>1247</td>
<td>10</td>
<td>4/4/2</td>
<td>3</td>
<td>2</td>
<td>0.0005</td>
<td>0.0006 (0.0000, 0.0017)</td>
<td>-0.691 (P &gt; 0.10)</td>
</tr>
<tr>
<td>TM6-B</td>
<td><em>A. sandwicense</em></td>
<td>1230</td>
<td>16</td>
<td>8/3/2/3</td>
<td>5</td>
<td>5</td>
<td>0.0014</td>
<td>0.0012 (0.0002, 0.0029)</td>
<td>0.823 (P &gt; 0.10)</td>
</tr>
</tbody>
</table>

n is the number of individuals/alleles sampled.
nlocality is the distribution of individual/allele sampling by locality
   for *D. ciliolata* samples: Puu Kanakaleonui / Puu Kwiiwi / Waipahoeohoe gulch
   for *A. sandwicense* samples: Silversword Loop / Puu o Pele / Ka moa o Pele / Puu Naue
Nhap is the number of observed haplotypes.
S is the number of observed segregating sites.
πsilent is nucleotide diversity at silent sites.
θw (C.I.) is Watterson’s estimate of θ with 95% confidence interval (95% C.I.).
Table 2. Hudson-Kreitman-Aguade contingency test values for *ASAPETALAI* and *ASAPETALA3/TM6* homoeologs (*A* and *B*) using intraspecific data from *Dubautia ciliolata* ssp. *glutinosa* and interspecific divergence between *Dubautia ciliolata* ssp. *glutinosa* and *Argyroxyphium sandwicence* ssp. *macrocephalum*. * Significant at the $\alpha = 0.05$ level.

<table>
<thead>
<tr>
<th></th>
<th>Intraspecific Polymorphism</th>
<th>Interspecific Divergence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed</td>
<td>Expected</td>
</tr>
<tr>
<td><em>ASAP1</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>A</em></td>
<td>20</td>
<td>16.67</td>
</tr>
<tr>
<td><em>B</em></td>
<td>9</td>
<td>12.33</td>
</tr>
<tr>
<td><em>ASAP3/TM6</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>A</em></td>
<td>22.68</td>
<td>26.01</td>
</tr>
<tr>
<td><em>B</em></td>
<td>22.57</td>
<td>19.24</td>
</tr>
<tr>
<td><em>A</em></td>
<td>1380.54</td>
<td>1427.86</td>
</tr>
<tr>
<td><em>P</em>-value</td>
<td>0.311</td>
<td>0.034*</td>
</tr>
</tbody>
</table>

$\chi^2$ values

Interspecific Divergence

# silent sites

142
Figures 1 and 2. Gene maps and summary figures of nucleotide variation in *ASAPETAL1* homoeologs (Fig. 1) and *ASAPETAL3/TM6* homoeologs (Fig. 2) from *Argyroxyphium sandwicense* ssp. *macrocephalum* and *Dubautia ciliolata* ssp. *glutinosa*. In both gene maps, exons are shown as numbered boxes. Bent arrows indicate positions of PCR primers used to amplify loci. Gray triangles above each gene map represent major deletions and insertions of *B* homoeologs relative to *A* homoeologs, with italicized numbers above each gray triangle indicating indel size. The MADS box and K-box regions are shown as darkened sections of exons as labeled. A scale bar represents approximately 100 bp. Dotted lines between exons indicate an extended length beyond the scale bar. In every summary figure, vertical bars above the line represent non-singleton polymorphic sites and vertical bars below the line represent singleton sites. Each filled circle represents a replacement polymorphism and each triangle represents an indel polymorphism. All polymorphic sites indicated in the summary figures are general locations based on joint alignments including both homoeologs of each floral regulatory gene.
Figure 1. **ASAPETALAI**
Figure 2. **ASAPETALA3/TM6**

**ASAPETALA3/TM6-A**

Argyroxyphium sandwicense

Dubautia ciliolata

**ASAPETALA3/TM6-B**

Argyroxyphium sandwicense

Dubautia ciliolata
Figure 3. Statistical parsimony haplotype networks of *ASAPETALAI* homoeologs in *Dubautia ciliolata* ssp. *glutinosa* and *Argyroxiphium sandwicense* ssp. *macrocephalum*. Figure A is the *ASAP1-A* homoeolog haplotype network and Figure B is the *ASAP1-B* homoeolog haplotype network. Circles represent *D. ciliolata* haplotypes and squares represent *A. sandwicense* haplotypes. Numbers within and relative sizes of haplotype shapes indicate the number of sampled alleles of specific haplotypes. Small circles represent missing haplotypes with > 95% statistical parsimony support.
Figure 4. Statistical parsimony haplotype networks of ASAPETALA3/TM6 homoeologs in Dubautia ciliolata ssp. glutinosa and Argyroxiphium sandwicense ssp. macrocephalum. Figure A is the ASAP3/TM6-A homoeolog haplotype network and Figure B is the ASAP3/TM6-B homoeolog haplotype network. Circles represent D. ciliolata haplotypes and squares represent A. sandwicense haplotypes. Numbers within and relative sizes of haplotype shapes indicate the number of sampled alleles of specific haplotypes. Small circles represent missing haplotypes with > 95% statistical parsimony support.
CHAPTER 6

SPECIATION IN AN ISLAND ADAPTIVE RADIATION:
A MULTILOCUS STUDY OF
THE HAWAIIAN SILVERSWORD ALLIANCE

This chapter consists of a manuscript in preparation for submission to the journal *Evolution*. This paper will be co-authored with Robert H. Robichaux (Department of Ecology and Evolutionary Biology, University of Arizona, Tucson).
ABSTRACT

The evolutionary mechanisms underlying adaptive radiations remain elusive, yet are critical to understanding organismal diversification. In this investigation, we studied the morphological and molecular divergence in three species of the Hawaiian silversword alliance, a premier example of plant adaptive radiation. Examination of the extent of molecular divergence and morphological differentiation between two sibling species (*Dubautia ciliolata* and *D. arborea*) across eight morphological characters indicates that these recently derived species differ significantly for each trait, yet have a genetic distance similar to levels expected between intraspecific populations. A multilocus molecular population genetic approach was also used to determine the extent of genetic divergence and historical demographics of diversification among three species: the recently derived sibling species (*Dubautia ciliolata* and *Dubautia arborea*) and another species from a separate clade of the Hawaiian silversword alliance (*Argyroserjium sandwicense*). Haplotype divergence, likelihood ratio tests of exponential growth rate, and Bayesian coalescent simulations of migration and divergence across these genes reveal a pattern consistent with significant demographic influences on genetic diversification in these species. Taken together, these results indicate that species within the Hawaiian silversword alliance exhibit rapid morphological divergence and high migration rates between sibling species that may be sufficient to supercede genetic divergence caused by drift. Moreover, these adaptively radiating species have histories including population expansion. These results are discussed in terms of the evolutionary mechanisms of adaptive radiations.
INTRODUCTION

Understanding the mechanisms underlying adaptive evolution and organismal divergence is a major goal of evolutionary biology. These two processes act in concert to yield the most dramatic examples of organic diversification on this planet--adaptive radiations. Adaptive radiations are defined as rapidly evolving species groups that have differentiated into numerous lineages through resource utilization and environmental heterogeneity (ecological adaptation, Schluter 2000). This definition distinguishes these species groups from non-adaptive radiations, which occur with little to no relationship to ecological adaptation (Brooks et al. 1985; Gittenberger 1991; Schluter 2000).

Adaptive radiations are associated with accelerated speciation correlated with elevated morphological, physiological, and ecological divergence among lineages. This differentiation may occur through three processes (Schluter 2000): (1) phenotypic differentiation among populations and species due to alternative environments and resources, (2) phenotypic divergence due to competition for resources among lineages, and (3) divergent natural selection leading to increased ecological and phenotypic divergence among lineages, resulting in speciation. These processes result in four main group characteristics, including common ancestry, phenotype-environment correlation, trait utility, and rapid speciation (Schluter 2000). While these features are difficult to empirically detect, progress has been made in the investigation of these characteristics in several of the most dramatic examples of adaptive radiation such as Darwin’s finches on the Galapagos Islands (Lack 1947; Grant 1984; Grant & Grant 1994), West Indian Anolis lizards on the Caribbean Islands (Pianka 1969, 1986; Losos 1990a, 1990b, 1992; Garland 1994; Losos 1996; Irschick 1997), Columbines in North America (Aquilegia) [Chase & Raven 1975; Miller 1981; Hodges 1997; Fulton & Hodges 1999], and the ‘greatest living example’ of plant adaptive radiation--the Hawaiian silversword alliance (Robichaux 1984; Robichaux & Canfield 1985; Carr et al. 1989; Baldwin & Sanderson 1998; Barrier et al. 1999; Barrier et al. 2000; Schluter 2000).
Most studies of adaptive radiations have focused on the macroevolutionary effects of rapid evolution on phenotypes and their correlations with trait utility and divergence between species. In contrast, very little is known about the contribution of demographic forces to the differentiation of populations into divergent taxa in adaptive radiations. In this study, a multilocus approach is utilized to examine the potential demographic effects of lineage sorting, migration, and population growth rate on nucleotide diversity in the Hawaiian silversword alliance adaptive radiation.

The Hawaiian Silversword Alliance Adaptive Radiation

The Hawaiian silversword alliance (Asteraceae, Heliantheae, Madiinae) is one of the most remarkable extant adaptive radiations and ‘ranks with the dinosaur extinctions and the origin of our own species among the most celebrated events in the history of life’ (Schluter 2000). This insular radiation is a morphologically divergent group descended from the self-incompatible Madia lineage of the North American tarweeds (Californian Madiinae, Heliantheae) [Robichaux et al. 1990; Baldwin et al. 1991; Baldwin & Robichaux 1995; Baldwin 1997]. This adaptive radiation is also characterized by considerable chromosomal evolution, with an inferred eight genomic rearrangements based on chromosomal pairing in natural and artificial hybrids (Carr and Kyhos 1981, 1986).

A calibrated, rate-constant ITS tree indicates that the Hawaiian silversword alliance diversified within the last $5.2 \pm 0.8$ million years (Baldwin & Sanderson 1998). Further phylogenetic analyses show that these species are allopolyploids presumably derived from an interspecific hybridization event between two North American tarweed species (Barrier et al. 1999). Examination of duplicated regulatory genes and a structural gene indicates a genome-wide acceleration of protein evolution in the Hawaiian silversword species compared to the extant North American tarweed species, which appears to be slightly elevated in regulatory versus structural genes (Barrier et al. 2001).
The Hawaiian silversword alliance is composed of three main taxonomic lineages that diverged from each other immediately after the origin of this adaptive radiation 5.2 ± 0.8 million years ago (Baldwin & Sanderson 1998). The ancestral species colonized the Hawaiian archipelago and diversified through adaptive radiation during the formation and weathering of these islands through at least thirteen founder events associated with inter-island dispersal (Baldwin and Robichaux 1995). This adaptive radiation resulted in three endemic genera comprising the Hawaiian silversword alliance (Argyrospirum DC, Dubautia Gaudich, and Wilkesia A. Gray) [Carr 1985; Robichaux et al. 1990; Baldwin et al. 1991; Baldwin & Robichaux 1995; Baldwin 1997; Baldwin & Sanderson 1998; Barrier et al. 1999; Baldwin & Wessa 2000].

This study includes three species: Argyroxiphium sandwicense subsp. macrocephalum, Dubautia ciliolata subsp. glutinosa, and Dubautia arborea. These species represent two different major sublineages of the Hawaiian silversword alliance that diverged in the earliest phase of this adaptive radiation approximately 5 million years ago: the Argyroxiphium genus and the Raillardia clade (Baldwin & Sanderson 1998). Sampling from two different major sublineages of this adaptive radiation permits examination of the mechanisms of divergence since the early ancestral differentiation of members of this species alliance.

The extraordinary diversity observed across this adaptive radiation is the result of divergence at the population level, thus diversity and divergence between two sibling species in this insular plant group was examined. Two sibling species selected for this study (Dubautia ciliolata subsp. glutinosa and Dubautia arborea) appear to be the most recently derived species of the Hawaiian silversword alliance. These two species are endemic to the island of Hawaii, which emerged approximately 0.5 million years ago and are both members of the Raillardia clade. The recent divergence of these sibling species provides the opportunity to examine the population-level mechanisms of adaptive radiation, as the
diversity evident across higher taxonomic levels originates from population genetic dynamics differentiating closely related lineages (Simpson 1953; Schluter 2000).

The morphological diversity in these species is evident in comparisons of several traits within these two major clades. *Argyroxiphium* and *Dubautia* species have dramatically different growth habits and reproductive morphologies. The *Argyroxiphium* genus includes 5 species with habits ranging from erect, unbranched, monocarpic rosette types (*A. sandwicense*) to branched, spreading, polycarpic species (*A. caliginis*) (Carr 1985). The *Dubautia* genus contains 23 perennial species that exhibit a greater range of growth forms than *Argyroxiphium* species and includes mat-forming subshrubs, cushion plants, larger subshrubs, trees, and a liana (Carr 1985). *Argyroxiphium sandwicense* subsp. *macrocephalum* have radiate capitula containing upwards of 650 florets while *Dubautia ciliolata* subsp. *glutinosa* and *Dubautia arborea* have discoid capitula containing an average of 8 and 22 florets, respectively (Carr 1985).

Analysis of ten allozyme loci indicates low genetic differentiation among Hawaiian silversword alliance species (Witter & Carr 1988). Furthermore, even species that diverged approximately 500,000 to 1.5 million years ago have highly similar genetic identities with many Nei genetic identity coefficients in the range of values typically expected in comparisons between populations within species (*I* > 0.90; Nei 1972; Witter & Carr 1985). In this study, the questions addressed by Witter and Carr (1988) were revisited and expanded by examining nucleotide variation for evidence of allelic divergence and then investigating patterns of haplotype and nucleotide differentiation to infer the demographic forces shaping genetic variation in this adaptive radiation. Multiple nuclear genes [homoeologous copies of two floral regulatory genes: *ASAPETALA1* (A and B copy), homoeologous copies of *ASAPETALA3* (A and B copy) and the housekeeping gene chlorophyll a/b binding protein (*ASCAB9*)] and one mitochondrial gene (*ASNAD1*) were used to assess genome-wide patterns of nucleotide variation that reflect demographic processes acting on all genes simultaneously in a genome. NADH dehydrogenase subunit 1 (NAD1) was selected because
it was previously shown to be phylogenetically informative among *Cucurbita* subspecies (Demseure *et al.* 1995; Sanjur *et al.* 2002)

More specifically, this study investigated divergence and speciation in the Hawaiian silversword alliance adaptive radiation and focused on three main objectives: (1) to evaluate morphological divergence between the sibling species *D. arborea* and *D. ciliolata* subsp. *glutinosa*, (2) to examine variation across loci and determine if the observed pattern reflects selection or demographic effects such as a population bottleneck or population expansion among all three species in this study, and (3) to address the effects of migration/selection versus lineage sorting on haplotype diversity in this adaptive radiation by examining joint estimates of migration rate, time to most recent common ancestor (TMRCA) of all alleles, and TMRCA between species.
MATERIALS AND METHODS

Morphological data: measurements and analysis

Six reproductive traits (number of capitula per capitulescence, number of florets per capitulum, receptacular bract length, sepal length, petal length, and ovary length) and two vegetative traits (leaf length, maximum leaf width) were measured to assess the morphological divergence between the sibling species *Dubautia arborea* and *Dubautia ciliolata* subsp. *glutinosa* on Mauna Kea, Hawaii. Measurements were recorded from 50 random individuals per population for each species, with 3 replicate measurements within each individual. The three populations sampled from *Dubautia arborea* include individuals from Puu Laau, Puu Mali and a woodland near Waipahoeohoe gulch (Figure 1). The three populations sampled from *Dubautia ciliolata* subsp. *glutinosa* include individuals from Puu Kanakaleonui, Puu Kawiwi and a shrubland near Waipahoeohoe gulch (Figure 1). An additional 100 *D. arborea x D. ciliolata* ssp. *glutinosa* hybrid individuals from Waipahoeohoe gulch were also measured. These species are referred to as *D. ciliolata* and *D. arborea* for the remainder of this paper.

Analysis of variance (ANOVA) fitting a standard least squares means model was used to partition the variance in each trait attributable to the following effects: among species, populations within species, and individuals within populations per species, including the variance from replicate measurements within individuals as the error term. All morphological data analyses were conducted using JMP version 4.0.1 (SAS Institute, Cary, NC).
Nucleotide sequence data: sample collections and DNA extraction

Leaf tissue samples were haphazardly collected from approximately 5 – 10 individuals from the same 3 populations listed above from each of the two Dubautia species included in this study (Dubautia arborea and Dubautia ciliolata subsp. glutinosa) [Figure 1]. Leaf tissue samples from 16 haphazardly-selected plants from Argyroxiphium sandwicense subsp. macrocephalum were obtained from 4 localities in Haleakala Crater Park, Maui: Silversword Loop, Puu o Pele, Ka Moa o Pele, and Puu Naue (Figure 2). This species is referred to as A. sandwicense for the remainder of this paper.

Genomic DNA isolations were extracted from young leaf tissues using a rapid plant CTAB (hexadecyltrimethylammonium bromide) protocol (Saghai-Maroof et al. 1984). EluQuick (Schleicher and Schuell, Keene, NH) glass bead purification was then used on all genomic DNA extractions to significantly reduce the amount of pectin and secondary-product contamination.

PCR amplifications and gene sequencing

To limit misincorporation of incorrect nucleotides, the error-correcting Pwo polymerase (Roche, Indianapolis) was used in all PCR amplifications. An error rate calculation based on multiple amplifications and re-sequencing of several genes indicates an error rate of less that one error in 7 – 10 kb (unpublished observations). Generation of gene-specific primers for ASAPETALA1 and ASAPETALA3/TM6 based on cDNA sequences was described in Barrier et al. 1999. Gene-specific primers for ASCAB9 based on cDNA sequences were designed as described in Barrier et al. 2001. Due to similarity in size, homoeologous copies of ASAPETALA1 (ASAP1-A and ASAP1-B) and ASAPETALA3/TM6 (ASAP3/TM6-A and ASAP3/TM6-B) were co-amplified, band-extracted and clone identity was then assessed using a restriction digest-based screen as described in Lawton-Rauh et al. 2003. This approach was used to amplify a region spanning exons 3 – 8 of the ASAP1-A and
ASAP1-B homoeologs (both fragments are ~ 1.8 kb) and a region spanning exons 1 – 4 of the ASAPETALA3/TM6 homoeologs (the obtained ASAP3-A fragments are ~1.1 kB, the ASAP3-B fragments are ~1.4 kB).

ASCAB9 was amplified using primers CAB9-e1f (5’- TTCGGACT-TGGTAAGAAAC-3’) and ASCAB9lr1 (5’-AAGCCAATTGATTCTGTA-3’) generally using the conditions: 94°C for a 5 min. hot-start followed by 10 cycles of 94°C for 15 sec., 58°C for 30 sec., 72°C for 1 min. then 20 cycles of 94°C for 15 sec., 58°C for 30 sec., and 72°C for 1 min. with a +5 sec. per cycle autoincrementation followed by a 7 min. extension at 72°C. One ASCAB9 homoeolog was extracted using the QiaQuick band extraction kit (Qiagen) and then cloned into the pCR-BluntII-TOPO vector using the Zero Blunt TOPO TA cloning kit (Invitrogen). A region spanning exons B and C of the mitochondrial gene ASNAD1 was amplified with universal primers used in Sanjur et al. 2002 and designed in Demesure et al. 1995. The ASNAD1 PCR products were directly sequenced due to the presence of single copies and single alleles.

All genes were sequenced by primer walking using automated DNA sequencing machines (NCSU DNA Sequencing Facility, Iowa State University Sequencing Facility, and NCSU Genome Research Laboratory). Each sequence was then aligned and edited using the quality-based scoring algorithms of the BioLign alignment and editing suite (Tom Hall, 2000-2001) and the CodonCode phred/phrap quality index program. All polymorphic sites were confirmed by visual inspection of chromatograms. The DNA sequences will be available through GenBank.

Sequence data analyses

Nucleotide diversity levels were estimated as mean pairwise differences (π) at silent sites, intron sites, synonymous sites, and nonsynonymous sites. Estimates of haplotype
diversity and nucleotide diversity ($\theta_s$) at all sites and several tests for selection were conducted using the DnaSP program version 3.53 (Rozas & Rozas 1999).

The Tajima’s $D$ test statistic was used to evaluate deviations from the expectation of the neutral-equilibrium model (Tajima 1989). This test is generally used to detect gene-specific signatures consistent with deviations from neutrality but can be extended to evaluate the role of genome-wide demographic processes. A multi-locus trend of positive Tajima’s $D$ test statistic values is expected following population subdivision, while negative Tajima’s $D$ statistic values are expected following population expansion (Tajima 1989). The Tajima’s $D$ test was performed in DnaSP version 3.53 (Rozas & Rozas 1999). The significance of Tajima’s $D$ test statistics were evaluated via coalescent simulations in DnaSP using 10,000 runs, setting the number of segregating sites to values directly calculated using the data, and the population recombination parameter estimated using SITES (Hey & Wakeley 1997).

Analyses of molecular variance (AMOVAs, Excoffier et al. 1992) were conducted to partition molecular variance among species and populations using Kimura 2- parameter distances among haplotypes in Arlequin version 2.000 (Excoffier et al. 1992). These analyses were employed in pairwise comparisons between all three species at all loci.

Haplotype networks for each locus were calculated using the 95% statistical parsimony support criterion for inferred estimated gene genealogies (Templeton et al. 1992) as implemented in the TCS version 1.13 program (Clement et al. 2000). The 95% connection limit was set to the recommended value of 20 inferred missing haplotypes for every network except for ASAP1-A, and ASAP1-B, where the value was set to 23 in order to connect $A.\ sandwicense$ haplotypes to the same network as $D.\ ciliolata$ and $D.\ arborea$.

The likelihood surfaces for the per gene population-mutation parameter ($\theta$, denoted as $\theta_F$ for the value estimated using the Fluctuate program) and the population growth rate parameter ($g$) were jointly estimated and optimized from silent site data using the
Metropolis-Hastings Markov Chain Monte Carlo genealogical approach implemented in the program Fluctuate (Kuhner et al. 1998). Analyses were repeated three times to ensure estimation stability using the following settings: population-mutation parameter estimated from the data (Watterson 1975), search among unconstrained possible exponential growth parameter estimates, randomly-inferred genealogy, a starting growth rate of \( g = 100 \) to \( 200 \) for each initial search, and the recommended search strategy chain lengths. The likelihood ratio test was used to evaluate support for population growth \( (g > 0) \). The relationship among the parameters of effective population size \( (N_e) \), diversity \( (\theta_F) \), mutation rate per generation \( (\mu) \), growth rate \( (g) \), and time in generations \( (t) \) is described by the equation: 

\[
N_e = \left( \frac{\theta_F}{2\mu} \right) e^{\frac{-g\mu t}{2}}
\]  

(Kuhner et al. 1998).

To examine the relative effects of lineage sorting and migration on haplotype diversity between species, the Bayesian inference-based Markov chain Monte Carlo framework of the MDIV program was employed (Nielsen & Wakeley 2001). This algorithm estimates the joint non-equilibrium values of \( \theta = 4N_e \mu \), migration rate \( M = 2N_eM \) and the time to most recent common ancestor \( [\text{TMRCA} = t / (2N_e)] \) between taxa and among all alleles. This procedure was developed to choose between two alternative hypotheses regarding the divergence between populations or between unstructured species: short divergence times with little migration between taxa versus long divergence times with strong migration. Although the credibility intervals of this nonequilibrium-based estimation of migration and species divergence times are broad, these estimates are a significant improvement from other available estimation algorithms that are based on equilibrium models. Other methods assume strict isolation between taxa and ignore incomplete lineage sorting hence are more biased towards shared ancestral polymorphisms (incomplete lineage sorting) (Nielsen & Wakeley 2001; Griswold & Baker 2002). By utilizing the coalescent, the Mdiv algorithm relaxes these assumptions by incorporating both the species coalescent and the gene coalescent (Nielsen & Wakeley 2001).
This approach was used to determine the relative contributions of migration and incomplete lineage sorting or lack of genetic divergence across species. Three to eight replicate runs of each pairwise comparison were conducted using different random seeds while optimizing prior maximum value inputs for species divergence time and migration rate to assess both the convergence of estimations and to establish appropriate 95% credibility intervals. These simulations were run using the recommended settings: HKY mutation model and 2,000,000 parameter space exploration chains following a burn-in of 50,000 chains. The starting prior maximum parameter values differed in each run and the high similarity of posterior distributions indicates that estimates converged to the ergodic average.
RESULTS

Dubautia ciliolata and D. arborea: Morphological Data Analyses

The extent of morphological divergence between two closely related species can be used to support the taxonomic division of the two species and, as in this study, it can illustrate the magnitude of phenotypic differentiation among sibling species. Eight morphological traits (six reproductive and two vegetative) were analyzed between the sibling species D. ciliolata and D. arborea to determine the extent of phenotypic divergence between these two species using ANOVA to partition the variance into sources attributable to species, population within species, individuals within populations.

The mean measurements of all phenotypes examined in this study are shown in Table 1. For every trait in this study, population-level and individual-level variances were significant by ANOVA, \( P < 0.001 \), Table 2). For all traits except receptacular bract length, sepal length, and ovary length (each with \( P > 0.10 \)), species-level variance is significant \( (P < 0.01) \). The additivity of sums of squares permits comparisons among sources of variation. These percentages of variation by source out of the total variation were calculated for each trait and are shown in the last column of Table 2. Examination of sums of squares for each source of variation indicates that, in several traits, a substantial percentage of the variation is due to among-species differentiation. In two out of the six reproductive traits examined (number of capitula per capitulescence and number of florets per capitulum) and in both vegetative traits (leaf length and maximum leaf width), the majority of the variation is due to between-species variation \( (SS_{\text{source}}/SS_{\text{total}} = 74\%, 63.6\%, 83.1\%, \text{and} 83.9\%, \text{respectively}) \). Comparing among sums of squares for receptacular bract length indicates that most variation in this trait occurs among individuals within populations. Sepal length, corolla length, and ovary length have high error terms, accounting for 19 - 29 % of the total sums of squares (Table 2).
**Intraspecific Genetic Variation**

Evolutionary history can be inferred from the levels and patterns of molecular variation within and among genes. In this study, we investigated six genes to examine the population genetics of the Hawaiian silversword alliance adaptive radiation.

The alleles for *ASAP1-A, ASAP1-B, ASAP3/TM6-A, ASAP3/TM6-B, ASCAB9, and ASNAD1* were isolated and sequenced from *A. sandwicense, D. ciliolata, and D. arborea*. The aligned sequence length and population sampling distributions are indicated in Table 3 and Figures 1 and 2. In general, sampling within *D. ciliolata* and *D. arborea* is evenly distributed among study populations, with the exception of *ASAP3/TM6-B* in *D. ciliolata* and *D. arborea*. As described in Lawton-Rauh et al. 2003, we were unable to amplify this gene from 14 out of 24 individuals in *D. ciliolata* and from 23 out of 24 individuals in *D. arborea*, suggesting that a deletion allele for this gene is present in these two species.

Nucleotide diversity levels (π) are very low at all loci for all three species (Table 3). The mitochondrial *ASNAD1* gene is monomorphic in all species, except for a single polymorphism in one *D. ciliolata* individual (Appendix Figure A9). Estimates of nucleotide diversity at all nuclear loci in all three species are low, ranging from 0.0005 to 0.0075 (Table 3). A previous study tested for evidence of selection and found that *ASAP3/TM6-B* in *D. ciliolata* showed evidence of a possible selective sweep and divergent evolution from *ASAP3/TM6-A* (Lawton-Rauh 2003). In this current study, these same tests were employed for every gene among all three species to test for deviations from neutral-equilibrium model expectations: the Hudson Kreitman Aguade (HKA) test of intraspecific variation and intraspecific divergence between gene copies (results not shown; Hudson et al. 1987), the McDonald-Kreitman test of evolutionary rates between gene copies (results not shown; McDonald & Kreitman 1991), and the Tajima’s D test statistic (Table 6). For all but the
previously published result of a possible selective sweep in ASAP3/TM6-B in D. ciliolata, these analyses suggest that we cannot reject the neutral-equilibrium model.

Sequence analysis indicates that ASAP1-A alleles from 15 A. sandwicense individuals consist of 9 haplotypes and ASAP1-B alleles from 15 individuals form 11 haplotypes (Table 3, Figs. 1 and 2). ASAP3/TM6-A alleles from 15 individuals of this same species consist of 8 haplotypes and ASAP3/TM6-B alleles from 16 individuals form 5 haplotypes (Table 3, Figs. 3 and 4). ASCAB9 alleles in A. sandwicense consist of 3 haplotypes (Table 3 and Fig. 5). No sequence variation was found among sequenced ASNAD1 alleles in A. sandwicense (Appendix Figure A9).

Analyses of D. ciliolata indicate that ASAP1-A alleles from 20 individuals form 11 haplotypes and ASAP1-B alleles from 24 individuals form 9 haplotypes (Table 3, Figs. 1 and 2). ASAP3/TM6-A alleles in this same species from 21 individuals form 13 haplotypes and ASAP3/TM6-B alleles from 10 individuals form 3 haplotypes (Table 3, Figs. 3 and 4). ASCAB9 alleles in D. ciliolata from 25 individuals form 15 haplotypes and ASNAD1 alleles from 20 individuals form 2 haplotypes (Table 3, Fig. 5, Appendix Figure A9).

Dubautia arborea ASAP1-A alleles from 20 individuals reveals 7 haplotypes and ASAP1-A alleles from 19 individuals consist of 9 haplotypes (Table 3, Figs. 1 and 2). ASAP3/TM6-A alleles in D. arborea from 18 individuals form 10 haplotypes and ASCAB9 alleles from 16 individuals form 10 haplotypes (Table 3, Figs. 3 and 5). As mentioned above, ASAP3/TM6-B was only obtained from one D. arborea individual and ASNAD1 has no sequence variation (Table 3, Fig. 4, Appendix Figure A9).

Interspecific Genetic Divergence: A. sandwicense vs. each Dubautia species
Analysis of molecular variance (AMOVA) was used to partition haplotypic diversity covariation into three sources: among-species, among-populations, and within-populations (Excoffier et al. 1992; Tables 4 and 5). AMOVA comparisons between *A. sandwicense* and both *D. ciliolata* and *D. arborea* indicate that most molecular variation is due to among-species divergence in all genes (Table 4). The only exception is in *ASAP3/TM6-A* between *A. sandwicense* and *D. ciliolata*, where most variation occurs at the intrapopulation level such that within-population variation is 57.6 % and among species variation accounts for 43.5% of total variation.

The fixation indices (Φ) estimated from AMOVA are defined as the correlation of random haplotypes within a subgroup drawn from the entire group. The correlation of interspecific haplotypes drawn from both species is Φ<sub>CT</sub>, interpopulation haplotypes drawn within a species is Φ<sub>SC</sub>, and inter-population haplotypes drawn from both species is Φ<sub>ST</sub> (Excoffier et al. 1992; Table 4). The interspecific-total (Φ<sub>CT</sub>) and interpopulation-total (Φ<sub>ST</sub>) fixation indices are similar for all genes, with the highest values occurring between *A. sandwicense* and *D. ciliolata* in *ASAP3/TM6-B* where these fixation indices are 0.854 and 0.888, respectively (Table 4).

Interspecific Genetic Divergence: Between sibling species

For the comparisons between *D. ciliolata* and *D. arborea*, most molecular covariation is found within species, with approximately 80% of the molecular covariation occurring within populations. Overall, the interpopulation-total fixation indices (Φ<sub>ST</sub>) are higher than the interspecific-total indices (Φ<sub>CT</sub>) and the interpopulation-species indices (Φ<sub>SC</sub>) (Table 5). The interpopulation-total fixation index (Φ<sub>ST</sub>) is notably high in *ASAP1-A* and *ASAP1-B* at
0.199 and 0.230. The interpopulation-species fixation index ($\Phi_{SC}$), which estimates between-population differentiation within species (as opposed to $\Phi_{ST}$, which estimates differentiation across all populations regardless of species) is very low in all four genes, ranging from 0.016 at ASAP1-A to 0.075 in ASAP1-B.

AMOVA analyses between $D. ciliolata$ and $D. arborea$ reveal moderate to high levels of interspecific-total covariation in ASAP1-A, ASAP1-B, and ASAP3/TM6-A, with $\Phi_{CT}$ estimates at least four-fold higher than the value estimated in ASCAB9. Interestingly, analysis of 166 AFLP markers using individuals from the same populations indicates a genome-wide fixation index estimate of 0.086 between these two species, which is roughly half the estimates for ASAP1-A, ASAP1-B, and ASAP3/TM6-A (D. Remington, personal communication).

Haplotype networks: A. sandwicense vs. each Dubautia species

The inferred 95% statistical parsimony haplotype networks for ASAP1-A, ASAP1-B, ASAP3/TM6-A, ASAP3/TM6-B, and ASCAB9 are shown in figures 3 through 7 and the network for the mitochondrial gene ASNAD1 is shown in Appendix Figure A9. The haplotype sequence tables are shown in Tables 8 - 12. All inferred nuclear gene networks indicate separate clades with no shared haplotypes for A. sandwicense alleles compared to alleles from both Dubautia species. While there are no shared haplotypes between the genera, the number of inferred missing haplotypes between A. sandwicense and both Dubautia species is notably fewer in ASAP3/TM6-A compared to the other nuclear genes in this study. The haplotype network for the mitochondrial gene ASNAD1 illustrates the complete monomorphism of this gene among A. sandwicense, D. arborea alleles and all but one D. ciliolata allele (Figure A9).
None of the nuclear genes investigated in this study shows reciprocal monophyly between alleles from the sibling species *D. ciliolata* and *D. arborea* (figs. 3 through 7). Several haplotypes in most genes are shared between these two species (*ASAP1-B, ASAP3/TM6-A, ASAP3/TM6-B, ASCAB9*) and no haplotypes are shared among *ASAP1-A* alleles. Several haplotypes appear to derive from ancestral haplotypes in all but *ASAP3/TM6-B*. This includes haplotypes L and P in *ASAP1-A, D. ciliolata* haplotype G in *ASAP1-B*, haplotype T in *ASAP3/TM6-B*, and haplotypes S and T in *ASCAB9*. Additionally, the inferred haplotype networks for *ASAP3/TM6-A* and *ASCAB9* indicate evidence of homoplasy between two haplotypes as shown by marked connecting lines (figures 5 and 7). This homoplasy is probably due to recombination and is consistent with the moderate level of estimated recombination in these two genes (results not shown).

Trends in genome-wide diversity signatures were assessed using Tajima’s D test statistic at silent sites as well as haplotype diversity and nucleotide diversity at all sites (Table 6). Tajima’s D statistic examines the distribution of low frequency polymorphisms and similarities in D values at multiple loci may arise from demographic forces acting on species and leaves a genome-wide signature. Except for *ASAP3/TM6-B* in *A. sandwicense*, all Tajima’s D test statistic values are negative. Tajima’s D for *ASAP3/TM6-A* in *D. ciliolata* is the only significant value (Table 6). Consistently negative values for Tajima’s D across a
genome indicate an excess of low frequency polymorphisms, and is expected following expansion within an unstructured population or species (Tajima 1989).

The relative levels nucleotide and haplotype diversity can be compared to determine if observed high levels of variation is due to population expansion (Grant & Bowen 1998). Haplotype diversity estimates indicate the frequency and number of haplotypes among individuals and vary between 0 and 1.0. Nucleotide diversity estimates reflect the average number of pairwise differences per site within a sample and vary from 0.0000 under no variation to over 0.1000 under very deep divergences among sampled alleles (Grant & Bowen 1998). The haplotype diversity and average nucleotide diversity values at all sites are shown in Table 6. For all genes within each species, haplotype diversity estimates are high and all nucleotide diversity estimates are low. Haplotype diversity estimates in A. sandwicense range from 0.448 in ASCAB9 to 0.952 in ASAP1-B and nucleotide diversity estimates range from 0.0006 in ASCAB9 to 0.0037 in ASAP1-A. In D. ciliolata, haplotype diversity estimates range from 0.511 in ASAP3/TM6-B to 0.950 in ASCAB9 and nucleotide diversity estimates range from 0.0005 in ASAP3/TM6-B to 0.0043 in ASAP3/TM6-A. Haplotype diversity estimates in D. arborea range from 0.732 in ASAP1-A to 0.933 in ASCAB9 and nucleotide diversity estimates range from 0.0018 in ASAP1-A to 0.0057 in ASCAB9. This observed combination of high haplotype diversity and low nucleotide diversity is what is expected following population expansion (Grand & Bowen 1998).

Estimates of Population Growth Rate

The joint maximum likelihood estimates of nucleotide diversity at silent sites ($\theta_F$) and exponential growth rate (g) were conducted to determine if these Hawaiian silversword alliance species show evidence of population expansion. These parameter estimates and the conservative $P < 0.01$ log likelihood ratio test results for the estimated growth rate parameter versus the model of no growth rate are shown in the last three columns of Table 6. The
exponential growth rate parameter estimate \( (g) \) in \( A. \) sandwicense \( ASAP3/TM6-B \) was not recovered, most likely due to a flat likelihood surface preventing a successful search for appropriate likelihood estimates. Among recovered estimates, half of the \( A. \) sandwicense genes show significant deviation from a static population model \( (g = 0) \) with exponential growth rate parameters estimated at \( \sim 10.0 \times 10^3 \) using \( ASCAB9 \) and \( \sim 1.5 \times 10^3 \) using \( ASAP1-B \). Joint estimates of nucleotide diversity and the exponential growth parameter failed to converge in \( ASAP3/TM6-B \) in \( A. \) sandwicense. In \( D. \) ciliolata, four out of five genes examined indicate population expansion, with values of \( \sim 2.2 \times 10^3 \) using \( ASAP1-B \), \( \sim 0.4 \times 10^3 \) using \( ASAP3/TM6-A \), \( \sim 10.0 \times 10^3 \) using \( ASAP3/TM6-B \) and \( \sim 1.3 \times 10^3 \) using \( ASCAB9 \). Half of the \( D. \) arborea genes have joint estimates that are significantly different from expectations without population growth, with significant growth rate parameters of \( \sim 1.2 \times 10^3 \) using \( ASAP1-B \) and \( \sim 0.4 \times 10^3 \) using \( ASCAB9 \). These positive exponential growth rate parameter estimates among all three species indicates that these species have histories of exponential growth rate \( (g>0) \).

**Estimates of Gene Divergence Times and Species Divergence Times**

The pairwise estimates of the time to most recent common ancestor (TMRCA) of alleles (gene divergence) and species (species divergence) are presented in Table 6. Across loci and among pairwise comparisons, estimates of time to most recent common ancestor of alleles is greater than the estimate of time since species divergence except in \( ASAP3/TM6-B \) between \( A. \) sandwicense and \( D. \) ciliolata where species divergence is greater than the estimate of allele coalescence (Table 7). There is a notable difference between gene and species coalescence in the comparisons between \( D. \) ciliolata and \( D. \) arborea, where species divergence is fairly low (ranging from 0.18 to 0.89) and gene divergence is high and similar to gene divergence estimates between both \( Dubautia \) species and \( A. \) sandwicense.

**Estimates of Migration Rates**
Estimates of migration rate (M) and the population-mutation parameter per gene (θ) were obtained concurrently with the above divergence times using silent sites and are also presented in Table 7. Migration rate estimates are low in all comparisons between A. sandwicense and both D. ciliolata and D. arborea, with migration rate estimates ranging from 0.01 in most comparisons to 0.15 at ASCAB9. It should be noted that these estimates are slightly biased upwards because they are at the very edge of parameter space (Nielsen & Wakeley 2001). Nucleotide divergence estimates between A. sandwicense and D. ciliolata range from 1.12 substitutions per gene in ASCAB9 to 5.65 in ASAP1-B. Between A. sandwicense and D. arborea, nucleotide divergence estimates range from 1.40 in ASAP3/TM6-A to 7.17 in ASAP1-B.

The migration rates in D. ciliolata and D. arborea are high, with values ranging from 0.52 at ASAP1-A to 18.8 at ASAP3/TM6-A. This high migration rate suggests that migration is at a level sufficient to overcome genetic divergence between these two species caused by drift (Griswold & Baker 2002). Nucleotide divergence between these two species, while lower than the values calculated in A. sandwicense comparisons, they are high with estimates ranging from 1.30 at ASCAB9 to 3.44 at ASAP1-B.
DISCUSSION

The Hawaiian silversword alliance exhibits a remarkable array of morphologies and physiologies despite overall low levels of genetic divergence. This rapid morphological divergence is clearly illustrated in the examination of six reproductive and two vegetative traits presented here. Analysis of variance (ANOVA) for each trait between the sibling species *D. ciliolata* and *D. arborea* indicates significant divergence of these traits, most notably number of capitula per capitulescence, number of florets per capitulum, leaf length and maximum leaf width (Table 2). This morphological divergence exists even though these two species have a high average Nei genetic identity among 10 allozymes (*I* _ave_ = 0.944; Nei 1972; Witter & Carr 1985). The high level of morphological divergence between sibling species is the hallmark of this plant group and illustrates the rapid phenotypic diversification associated with the Hawaiian silversword alliance adaptive radiation.

In a previous study, four genes were initially chosen as candidate genes to investigate the molecular population genetics of duplicated floral regulatory genes in *A. sandwicense* and *D. ciliolata* (*ASAP1-A, ASAP1-B, ASAP3/TM6-A, and ASAP3/TM6-B*). This study indicated that one homoeologous pair of these candidate genes shows evidence of divergent evolution in *D. ciliolata* (*ASAP3/TM6-A and ASAP3/TM6-B*, Lawton-Rauh _et al._ 2003). These same tests for selection were also conducted for the additional genes included in this study. Except for *ASAP3/TM6-B* in *D. ciliolata*, all loci included in this present study have patterns consistent with the neutral-equilibrium model in HKA and McDonald-Kreitman tests.

This previous study also uncovered low nucleotide diversity estimates and negative Tajima’s D test statistic values across loci in both species. The combination of these two results suggested that demographic forces affecting all genes simultaneously might be operating to affect patterns of nucleotide variation in this adaptive radiation. To investigate possible demographic forces shaping nucleotide variation and allelic differentiation, this study includes two additional genes (*ASCAB9* and *ASNAD1*) and an additional species of the Hawaiian silversword alliance that is closely related to *D. ciliolata* (*D. arborea*). By
examining closely related species, specific questions regarding the population genetic mechanisms underlying the adaptive radiation of this dramatically divergent insular group were more directly addressed.

*Divergence between D. arborea and D. ciliolata*

*Dubautia ciliolata* and *D. arborea* are sibling species that appear to be the most recently derived species of the Hawaiian silversword alliance. Both of these species are endemic to Hawaii, the youngest island of the Hawaiian archipelago (Baldwin & Robichaux 1995). This island emerged approximately 500 thousand years ago, suggesting a very recent origin of these species (MacDonald et al. 1983; Carr 1985; Clague & Dalrymple 1987; Witter & Carr 1988; Baldwin & Sanderson 1998). *Dubautia ciliolata* and *D. arborea* are endemic to the island mountain Mauna Kea and grow at different elevations. There is at least one hybrid zone containing what appear to be members of both species and early generation hybrids between these species (D. Remington, personal communication). The relatively recent speciation between *D. ciliolata* and *D. arborea*, however, is associated with significant morphological divergence in reproductive and vegetative morphologies, as shown by the analyses of variance within traits presented above (Tables 1 and 2).

The morphological divergence between these two sibling species is in contrast to the lack of complete genetic differentiation between them. Previous analyses including allozymes (Nei’s genetic distance = 0.058, Nei 1972; Witter & Carr 1988) and AFLP markers (Wright’s $F_{ST} = 0.086$, D. Remington, personal communication; Wright 1951, 1965) indicate that *D. ciliolata* and *D. arborea* are not strongly differentiated. The lack of fixed differences observed in this present study between these two species is consistent with a previous allozyme-based approach which indicated that the more recently-derived *Dubautia* species ($n = 13$) are genetically similar, with high average genetic identity ($I = 0.95$, Witter & Carr 1988).
The analyses of molecular variance (AMOVA) included in the current study indicate that species divergence ($\Phi_{CT}$) between these two sibling species is low in the housekeeping gene $ASCAB9$ and is moderate to high in the three floral regulatory genes ($ASAP1-A$, $ASAP1-B$, and $ASAP3/TM6-A$) [Table 5]. These moderate to high levels of interspecific distances occur despite the lack of fixed differences between these two species in all genes, resulting in a lack of reciprocal monophyly. It is not clear why estimates of differentiation are approximately two-fold different between sequence-based estimates and AFLP marker-based estimates. This may arise from the specific types of polymorphisms that are scored in these markers. Alternatively, it may suggest that regulatory genes show greater differentiation between species than either structural genes or random genome-wide markers. Although the latter is intriguing, the lack of evidence of allele fixation between the sibling species suggests that these differences in fixation indices may be due to strong selection on regulatory genes associated with the clear morphological differences between the species.

While there are no fixed molecular differences between $D. ciliolata$ and $D. arborea$, haplotype networks including alleles from both species indicate several species-specific haplotypes (Figs. 3 - 7). Moreover, there are differences in the frequencies of shared haplotypes within each species which also contributes to the observed levels of nucleotide differentiation between $D. ciliolata$ and $D. arborea$. Thus, although overall genetic distances between these two species is low, haplotype distributions suggest that some degree of differentiation is occurring between these sibling species among all loci in this study.

*Signatures of population expansion in A. sandwicense, D. ciliolata and D. arborea*

The high haplotype diversity and low nucleotide diversity estimates among all genes, together with negative Tajima’s D test statistic values, suggests that all three species have histories that include population expansion (Table 6). Tests for evidence of exponential growth across genes are generally consistent with this trend, with two out of four genes
significant in *A. sandwicense*, four out of six genes significant in *D. ciliolata* and two out of four significant in *D. arborea* (*P* < 0.01, Table 6). While *A. sandwicense* has experienced recent human-mediated population decline within the last 200 years, the long generation time of this species (approx. 50 years) suggests that the nucleotide diversities observed in this study represent the effects of events that occurred prior to this decline. Thus, the estimates of population expansion in this species reflect population growth prior to more recent human influences on distribution of this species in Haleakala Crater, Maui.

The evidence of exponential population growth in *D. ciliolata* and, to some extent, in *D. arborea* suggests that the nucleotide diversity patterns are also consistent with patterns expected following population expansion (Table 6). At the genome-wide level, this further suggests that nucleotide diversity across all genes in these species have been influenced by population expansion and that patterns that appear to be consistent with non-neutral evolution must consider these demographic factors. In light of these results, it is possible that the inferred differentiation between *ASAP3/TM6-A* and *ASAP3/TM6-B* in *D. ciliolata* and *D. arborea* may be due to concurrent demographic and selective sweep events. The concurrence of a selective sweep during a population expansion suggests that the radiation of these two species may be the result of adaptive evolution occurring through natural selection coincident with or possibly resulting in exponential population growth.

*High migration rates between D. arborea and D. ciliolata*

The results of these analyses suggest an intriguing hypothesis regarding the role of migration in adaptive radiation. The results presented here suggest that migration rates may be sufficiently high among diverging Hawaiian silversword alliance species to reduce the influence of equilibrium-neutral lineage sorting processes. It is clear that opportunities for gene flow have occurred, and continue to occur. *Dubautia ciliolata* and *D. arborea* are interfertile with each other and their species ranges overlap on the island mountain Mauna
Kea. Additionally, at least one hybrid zone between these two species is known (Waipahoeohoe gulch). Taken together, this indicates that hybridization between the two species can occur in nature and may serve as the basis for historical and current gene flow. This also suggests that speciation in this adaptive radiation may occur through selection of a limited number of adaptive alleles that remain fixed in the face of extensive migration, rather than through neutral-equilibrium lineage sorting.

This suggestion of divergent selection during adaptive differentiation accompanied by high rates of migration among sympatric taxa has also been documented in a recent study of sympatric smelt (Saint-Laurent et al. 2003). In this study, sympatric normal and dwarf ecotypes of rainbow smelt exhibit significant phenotypic divergence in several optical, jaw, and fin characteristics yet have low levels of genetic differentiation and high rates of gene flow. The results of this study suggest that gene flow in this system has not prevented a directional effect of selection on phenotypic divergence between these two ecotypes (Saint-Laurent et al. 2003). In the context of the differentiation between *D. ciliolata* and *D. arborea*, gene flow may moderate the extent of adaptive divergence between diverging species and divergent selection appears to have been strong enough to maintain phenotypic differentiation despite high rates of migration between these two species.
Adaptive radiation, especially among insular species, occurs through a dynamic series of ecological interactions across stochastic environmental gradients. Examining hypothetical population demographic models of adaptive radiations can therefore be complex. For this reason, the evolutionary forces shaping population divergence in adaptive radiations has primarily been assumed rather than empirically investigated (Schluter 2000). To further dissect the influences of factors such as genetic differentiation, population expansion, and migration in the Hawaiian silversword alliance adaptive radiation, future analyses could include samples among more closely related species pairs across this insular group to determine the extent of migration in differentiating species. These analyses would shed light on the relative contributions of isolation and migration across the Hawaiian silversword alliance and test hypotheses regarding the population demographics of adaptive radiations.
ACKNOWLEDGMENTS

The authors thank Elizabeth F. Friar, Jody Hey, Andrea M. Johnson, Trudy F.C. Mackay, Ken M. Olsen, Arnar Palsson, David L. Remington, and Mark C. Ungerer for insightful discussions. Ed S. Buckler generously provided computer usage for the time and memory intensive Mdiv program analyses. Leaf tissue samples from *Argyroxciphium sandwicense* subsp. *macrocephalum* from Puu o Pele, Ka moa o Pele, and Puu Naue localities were generously provided by Elizabeth F. Friar. Elizabeth Friar, Marianne Barrier, Sola Halldorsdottir and Dave Remington assisted with field measurements of *Dubautia arborea* and *D. ciliolata* morphologies. Carlyn Keith Buckler performed several of the *D. arborea* and *D. ciliolata* DNA extractions used in this study. This work is supported by a grant from the National Science Foundation (to MDP and RHR), an Alfred P. Sloan Foundation Young Investigator Award (to MDP) and a Sigma Xi Grant in Aid of Research Award (to ALR).
LITERATURE CITED


Griswold CK and Baker AJ (2002) Time to the most recent common ancestor and divergence times of populations of common chaffinches (Fringilla coelebs) in Europe and North Africa: insights into pleistocene refugia and current levels of migration. Evolution 56, 143-153.


Sanjur OI, Piperno DR, Andres TC, and Wessel-Beaver L (2002) Phylogenetic relationships among domesticated and wild species of Cucurbita (Cucurbitaceae) inferred from a


Table 1. Morphological trait means in *Dubautia ciliolata* and *D. arborea* with confidence intervals (C.I.) and standard errors (SE). All measurements are in millimeters.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Species</th>
<th>Trait Mean (C.I.)</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td># capitula/capitulescence</td>
<td><em>D. ciliolata</em></td>
<td>3.120 (2.73, 3.51)</td>
<td>0.174</td>
</tr>
<tr>
<td></td>
<td><em>D. arborea</em></td>
<td>15.682 (15.29, 16.07)</td>
<td>0.247</td>
</tr>
<tr>
<td># florets / capitulum</td>
<td><em>D. ciliolata</em></td>
<td>8.646 (8.08, 9.21)</td>
<td>0.252</td>
</tr>
<tr>
<td></td>
<td><em>D. arborea</em></td>
<td>22.704 (22.14, 23.27)</td>
<td>0.252</td>
</tr>
<tr>
<td>receptacular bract length</td>
<td><em>D. ciliolata</em></td>
<td>8.039 (7.90, 8.18)</td>
<td>0.062</td>
</tr>
<tr>
<td></td>
<td><em>D. arborea</em></td>
<td>8.860 (8.72, 9.00)</td>
<td>0.062</td>
</tr>
<tr>
<td>sepal length</td>
<td><em>D. ciliolata</em></td>
<td>5.147 (5.07, 5.23)</td>
<td>0.036</td>
</tr>
<tr>
<td></td>
<td><em>D. arborea</em></td>
<td>4.389 (4.31, 4.47)</td>
<td>0.037</td>
</tr>
<tr>
<td>corolla length</td>
<td><em>D. ciliolata</em></td>
<td>5.412 (5.33, 5.49)</td>
<td>0.035</td>
</tr>
<tr>
<td></td>
<td><em>D. arborea</em></td>
<td>6.532 (6.45, 6.61)</td>
<td>0.035</td>
</tr>
<tr>
<td>ovary length</td>
<td><em>D. ciliolata</em></td>
<td>4.407 (4.33, 4.48)</td>
<td>0.033</td>
</tr>
<tr>
<td></td>
<td><em>D. arborea</em></td>
<td>4.836 (4.76, 4.91)</td>
<td>0.033</td>
</tr>
<tr>
<td>leaf length</td>
<td><em>D. ciliolata</em></td>
<td>14.190 (13.4, 15.0)</td>
<td>0.357</td>
</tr>
<tr>
<td></td>
<td><em>D. arborea</em></td>
<td>47.750 (46.9, 48.6)</td>
<td>0.357</td>
</tr>
<tr>
<td>maximum leaf width</td>
<td><em>D. ciliolata</em></td>
<td>2.940 (2.69, 3.19)</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td><em>D. arborea</em></td>
<td>13.600 (13.35, 13.85)</td>
<td>0.11</td>
</tr>
</tbody>
</table>
Table 2. ANOVA analyses of morphological traits in *Dubautia ciliolata* and *D. arborea*.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Source of Variation</th>
<th>DF</th>
<th>MS</th>
<th>F ratio</th>
<th>Prob &gt; F</th>
<th>SS_{source}/SS_{total}</th>
</tr>
</thead>
<tbody>
<tr>
<td># capitula per capitulescence</td>
<td>Species</td>
<td>1</td>
<td>35170.08</td>
<td>192.23</td>
<td>&lt;0.001</td>
<td>0.7404</td>
</tr>
<tr>
<td></td>
<td>Populations within species</td>
<td>4</td>
<td>182.96</td>
<td>6.80</td>
<td>&lt;0.001</td>
<td>0.0154</td>
</tr>
<tr>
<td></td>
<td>Individuals within population per species</td>
<td>294</td>
<td>26.91</td>
<td>4.52</td>
<td>&lt;0.001</td>
<td>0.1666</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>595</td>
<td>5.96</td>
<td></td>
<td></td>
<td>0.0747</td>
</tr>
<tr>
<td># florets per capitulum</td>
<td>Species</td>
<td>1</td>
<td>44342.99</td>
<td>23.07</td>
<td>&lt;0.01</td>
<td>0.6360</td>
</tr>
<tr>
<td></td>
<td>Populations within species</td>
<td>4</td>
<td>1921.92</td>
<td>47.57</td>
<td>&lt;0.001</td>
<td>0.1103</td>
</tr>
<tr>
<td></td>
<td>Individuals within population per species</td>
<td>294</td>
<td>40.40</td>
<td>4.10</td>
<td>&lt;0.001</td>
<td>0.1704</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>595</td>
<td>9.85</td>
<td></td>
<td></td>
<td>0.0841</td>
</tr>
<tr>
<td>Receptacular bract length</td>
<td>Species</td>
<td>1</td>
<td>151.10</td>
<td>4.31</td>
<td>&gt;0.10</td>
<td>0.0893</td>
</tr>
<tr>
<td></td>
<td>Populations within species</td>
<td>4</td>
<td>35.07</td>
<td>9.66</td>
<td>&lt;0.001</td>
<td>0.0829</td>
</tr>
<tr>
<td></td>
<td>Individuals within population per species</td>
<td>294</td>
<td>3.63</td>
<td>6.48</td>
<td>&lt;0.001</td>
<td>0.6315</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>595</td>
<td>0.56</td>
<td></td>
<td></td>
<td>0.1969</td>
</tr>
<tr>
<td>Sepal length</td>
<td>Species</td>
<td>1</td>
<td>127.46</td>
<td>4.20</td>
<td>&gt;0.10</td>
<td>0.1933</td>
</tr>
<tr>
<td></td>
<td>Populations within species</td>
<td>4</td>
<td>30.32</td>
<td>32.26</td>
<td>&lt;0.001</td>
<td>0.1839</td>
</tr>
<tr>
<td></td>
<td>Individuals within population per species</td>
<td>294</td>
<td>0.94</td>
<td>4.27</td>
<td>&lt;0.001</td>
<td>0.4178</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>595</td>
<td>0.22</td>
<td></td>
<td></td>
<td>0.2018</td>
</tr>
<tr>
<td>Corolla length</td>
<td>Species</td>
<td>1</td>
<td>280.30</td>
<td>21.33</td>
<td>&lt;0.01</td>
<td>0.3650</td>
</tr>
<tr>
<td></td>
<td>Populations within species</td>
<td>4</td>
<td>13.14</td>
<td>13.41</td>
<td>&lt;0.001</td>
<td>0.0679</td>
</tr>
<tr>
<td></td>
<td>Individuals within population per species</td>
<td>294</td>
<td>0.98</td>
<td>3.92</td>
<td>&lt;0.001</td>
<td>0.3753</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>595</td>
<td>0.25</td>
<td></td>
<td></td>
<td>0.1875</td>
</tr>
<tr>
<td>Ovary length</td>
<td>Species</td>
<td>1</td>
<td>41.75</td>
<td>6.65</td>
<td>&gt;0.10</td>
<td>0.0861</td>
</tr>
<tr>
<td></td>
<td>Populations within species</td>
<td>4</td>
<td>6.28</td>
<td>6.61</td>
<td>&lt;0.001</td>
<td>0.0518</td>
</tr>
<tr>
<td></td>
<td>Individuals within population per species</td>
<td>294</td>
<td>0.95</td>
<td>3.96</td>
<td>&lt;0.001</td>
<td>0.5750</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>595</td>
<td>0.24</td>
<td></td>
<td></td>
<td>0.2883</td>
</tr>
<tr>
<td>Trait</td>
<td>Source of Variation</td>
<td>DF</td>
<td>MS</td>
<td>F ratio</td>
<td>Prob &gt; F</td>
<td>$\frac{SS_{source}}{SS_{total}}$</td>
</tr>
<tr>
<td>----------------------</td>
<td>-----------------------------------</td>
<td>-----</td>
<td>-----</td>
<td>---------</td>
<td>----------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>Leaf length</td>
<td>Species</td>
<td>1</td>
<td>253505.54</td>
<td>130.46</td>
<td>&lt;0.001</td>
<td>0.8310</td>
</tr>
<tr>
<td></td>
<td>Populations within species</td>
<td>4</td>
<td>1943.24</td>
<td>16.53</td>
<td>&lt;0.001</td>
<td>0.0255</td>
</tr>
<tr>
<td></td>
<td>Individuals within population per species</td>
<td>294</td>
<td>117.56</td>
<td>7.64</td>
<td>&lt;0.001</td>
<td>0.1133</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>600</td>
<td>15.39</td>
<td>0.0303</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum leaf width</td>
<td>Species</td>
<td>1</td>
<td>25569.08</td>
<td>182.43</td>
<td>&lt;0.001</td>
<td>0.8385</td>
</tr>
<tr>
<td></td>
<td>Populations within species</td>
<td>4</td>
<td>140.16</td>
<td>12.13</td>
<td>&lt;0.001</td>
<td>0.0184</td>
</tr>
<tr>
<td></td>
<td>Individuals within population per species</td>
<td>294</td>
<td>11.56</td>
<td>7.18</td>
<td>&lt;0.001</td>
<td>0.1114</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>600</td>
<td>1.61</td>
<td>0.0317</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3. Nucleotide variation for loci examined in this study.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Species</th>
<th>Length (bp)</th>
<th>n</th>
<th>n&lt;sub&gt;locality&lt;/sub&gt;</th>
<th>N&lt;sub&gt;hap&lt;/sub&gt;</th>
<th>S</th>
<th>π&lt;sub&gt;silent&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASAP1-A</td>
<td>A. sandwicense</td>
<td>1683</td>
<td>15</td>
<td>8 / 3 / 1 / 3</td>
<td>9</td>
<td>21</td>
<td>0.0044</td>
</tr>
<tr>
<td></td>
<td>D. ciliolata</td>
<td>1661</td>
<td>20</td>
<td>8 / 6 / 6</td>
<td>11</td>
<td>27</td>
<td>0.0035</td>
</tr>
<tr>
<td></td>
<td>D. arborea</td>
<td>1680</td>
<td>20</td>
<td>7 / 6 / 7</td>
<td>7</td>
<td>8</td>
<td>0.0016</td>
</tr>
<tr>
<td>ASAP1-B</td>
<td>A. sandwicense</td>
<td>1709</td>
<td>15</td>
<td>8 / 2 / 2 / 3</td>
<td>11</td>
<td>18</td>
<td>0.0029</td>
</tr>
<tr>
<td></td>
<td>D. ciliolata</td>
<td>1736</td>
<td>24</td>
<td>7 / 12 / 5</td>
<td>9</td>
<td>13</td>
<td>0.0019</td>
</tr>
<tr>
<td></td>
<td>D. arborea</td>
<td>1735</td>
<td>19</td>
<td>7 / 5 / 7</td>
<td>9</td>
<td>18</td>
<td>0.0021</td>
</tr>
<tr>
<td>ASAP3/</td>
<td>A. sandwicense</td>
<td>1012</td>
<td>15</td>
<td>8 / 3 / 2 / 2</td>
<td>8</td>
<td>12</td>
<td>0.0036</td>
</tr>
<tr>
<td>TM6-A</td>
<td>D. ciliolata</td>
<td>949</td>
<td>21</td>
<td>8 / 8 / 5</td>
<td>13</td>
<td>22</td>
<td>0.0052</td>
</tr>
<tr>
<td></td>
<td>D. arborea</td>
<td>949</td>
<td>18</td>
<td>6 / 5 / 7</td>
<td>10</td>
<td>13</td>
<td>0.0030</td>
</tr>
<tr>
<td>ASAP3/</td>
<td>A. sandwicense</td>
<td>1230</td>
<td>16</td>
<td>8 / 3 / 2 / 3</td>
<td>5</td>
<td>5</td>
<td>0.0014</td>
</tr>
<tr>
<td>TM6-B</td>
<td>D. ciliolata</td>
<td>1247</td>
<td>10</td>
<td>4 / 4 / 2</td>
<td>3</td>
<td>2</td>
<td>0.0005</td>
</tr>
<tr>
<td></td>
<td>D. arborea</td>
<td>1247</td>
<td>1</td>
<td>0 / 0 / 1</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>ASCAB9</td>
<td>A. sandwicense</td>
<td>952</td>
<td>15</td>
<td>7 / 3 / 2 / 3</td>
<td>3</td>
<td>3</td>
<td>0.0011</td>
</tr>
<tr>
<td></td>
<td>D. ciliolata</td>
<td>949</td>
<td>25</td>
<td>8 / 12 / 5</td>
<td>15</td>
<td>17</td>
<td>0.0044</td>
</tr>
<tr>
<td></td>
<td>D. arborea</td>
<td>952</td>
<td>16</td>
<td>7 / 5 / 4</td>
<td>10</td>
<td>19</td>
<td>0.0075</td>
</tr>
<tr>
<td>ASNAD1</td>
<td>A. sandwicense</td>
<td>767</td>
<td>16</td>
<td>8 / 3 / 2 / 3</td>
<td>1</td>
<td>0</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>D. ciliolata</td>
<td>760</td>
<td>20</td>
<td>8 / 7 / 5</td>
<td>2</td>
<td>1</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>D. arborea</td>
<td>759</td>
<td>19</td>
<td>7 / 6 / 6</td>
<td>1</td>
<td>0</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

n is the number of individuals/alleles sampled.

n<sub>locality</sub> is the distribution of individual/allele sampling by locality

for A. sandwicense samples: Silversword Loop / Puu o Pele / Ka moa o Pele / Puu Naue

for D. ciliolata samples: Puu Kanakaleonui / Puu Kawaiwi / Waipahoeohoe gulch

for D. arborea samples: Puu Laau / Puu Mali / Waipahoeohoe gulch

S is the number of observed segregating sites.

N<sub>hap</sub> is the number of haplotypes observed in the given dataset, all sites considered.

π<sub>silent</sub> is nucleotide diversity at silent sites.

na (not amplified) refers to lack of ASAP3/TM6-B amplification from all but one D. arborea individual examined.
Table 4. AMOVA and fixation indices for loci examined in this study: *A. sandwicense* vs. *D. arborea* (*D.arb*) and *D. ciliolata* (*D.cilio*).

<table>
<thead>
<tr>
<th>Source of Variation (%)</th>
<th>ASAP1-A</th>
<th>ASAP1-B</th>
<th>ASAP3/TM6-A</th>
<th>ASAP3/TM6-B</th>
<th>ASCAB9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>D.arb</em></td>
<td><em>D.cilio</em></td>
<td><em>D.arb</em></td>
<td><em>D.cilio</em></td>
<td></td>
</tr>
<tr>
<td>Among species</td>
<td>83.24</td>
<td>74.05</td>
<td>78.03</td>
<td>79.52</td>
<td></td>
</tr>
<tr>
<td>Among populations within species$^a$</td>
<td>0.71</td>
<td>-0.40</td>
<td>-0.97</td>
<td>-1.29</td>
<td></td>
</tr>
<tr>
<td>Within populations</td>
<td>16.05</td>
<td>26.36</td>
<td>22.94</td>
<td>21.77</td>
<td></td>
</tr>
</tbody>
</table>

Fixation indices

<table>
<thead>
<tr>
<th></th>
<th>$\Phi_{CT}$ (species / total)</th>
<th>$\Phi_{SC}$ (population / species)$^a$</th>
<th>$\Phi_{ST}$ (population / total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Phi_{CT}$</td>
<td>0.832</td>
<td>0.042</td>
<td>0.839</td>
</tr>
<tr>
<td>$\Phi_{SC}$</td>
<td>0.740</td>
<td>-0.016</td>
<td>0.736</td>
</tr>
<tr>
<td>$\Phi_{ST}$</td>
<td>0.780</td>
<td>-0.044</td>
<td>0.771</td>
</tr>
<tr>
<td></td>
<td>0.795</td>
<td>-0.063</td>
<td>0.782</td>
</tr>
<tr>
<td></td>
<td>0.618</td>
<td>0.062</td>
<td>0.641</td>
</tr>
<tr>
<td></td>
<td>0.434</td>
<td>-0.019</td>
<td>0.424</td>
</tr>
<tr>
<td></td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td></td>
<td>0.854</td>
<td>0.230</td>
<td>0.888</td>
</tr>
<tr>
<td></td>
<td>0.562</td>
<td>0.069</td>
<td>0.592</td>
</tr>
<tr>
<td></td>
<td>0.553</td>
<td>0.229</td>
<td>0.655</td>
</tr>
</tbody>
</table>

na (not amplified) refers to lack of *ASAP3/TM6-B* amplification from all but one *D. arborea* individual.

$^a$Negative variance measures and fixation indices occur because these values are neither covariances nor correlation coefficients, respectively, and indicate lack of within-species genetic structure.
Table 5. AMOVA and fixation indices for loci examined in this study: *D. arborea* vs. *D. ciliolata*

<table>
<thead>
<tr>
<th>Source of Variation (%)</th>
<th>ASAP1-A</th>
<th>ASAP1-B</th>
<th>ASAP3/TM6-A</th>
<th>ASCAB9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among species</td>
<td>18.61</td>
<td>16.74</td>
<td>16.66</td>
<td>4.11</td>
</tr>
<tr>
<td>Among populations within species</td>
<td>1.27</td>
<td>6.26</td>
<td>-5.03</td>
<td>6.45</td>
</tr>
<tr>
<td>Within populations</td>
<td>80.13</td>
<td>77.00</td>
<td>88.37</td>
<td>89.44</td>
</tr>
</tbody>
</table>

Fixation indices

<table>
<thead>
<tr>
<th></th>
<th>Φ_{CT} (species / total)</th>
<th>Φ_{SC} (population / species)</th>
<th>Φ_{ST} (population / total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Φ_{CT} (species / total)</td>
<td>0.186</td>
<td>0.016</td>
<td>0.199</td>
</tr>
<tr>
<td>Φ_{SC} (population / species)</td>
<td>0.167</td>
<td>0.075</td>
<td>0.230</td>
</tr>
<tr>
<td>Φ_{ST} (population / total)</td>
<td>0.167</td>
<td>-0.060</td>
<td>0.116</td>
</tr>
</tbody>
</table>

Negative variance measures and fixation indices occur because they are neither covariances nor correlation coefficients and indicate lack of within-species genetic structure at *ASAP3/TM6-A*. 
Table 6. Summary of parameter estimates and test statistics examined for demographic trends. The last three columns ($\theta_F$ and g) are joint parameters estimated in FLUCTUATE. Numbers in parentheses are standard deviations of estimates.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Species</th>
<th>Tajima’s D</th>
<th>Haplotype Diversity</th>
<th>$\theta_e$</th>
<th>$\theta_F$</th>
<th>g</th>
<th>$\Delta \ln L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ASA P1-A$</td>
<td>$A. sandwicense$</td>
<td>-0.1100</td>
<td>0.886 (0.069)</td>
<td>0.0037 (0.0003)</td>
<td>0.0076 (0.0019)</td>
<td>461 (214)</td>
<td>1.071</td>
</tr>
<tr>
<td></td>
<td>$D. ciliolata$</td>
<td>-0.6487</td>
<td>0.900 (0.052)</td>
<td>0.0038 (0.0009)</td>
<td>0.0048 (0.0011)</td>
<td>148 (149)</td>
<td>0.183</td>
</tr>
<tr>
<td></td>
<td>$D. arborea$</td>
<td>-1.1372</td>
<td>0.732 (0.092)</td>
<td>0.0018 (0.0002)</td>
<td>0.0010 (0.0003)</td>
<td>38 (601)</td>
<td>0.001</td>
</tr>
<tr>
<td>$ASA P1-B$</td>
<td>$A. sandwicense$</td>
<td>-0.7772</td>
<td>0.952 (0.040)</td>
<td>0.0026 (0.0003)</td>
<td>0.0153 (0.0050)</td>
<td>1452 (371)</td>
<td>6.489**</td>
</tr>
<tr>
<td></td>
<td>$D. ciliolata$</td>
<td>-0.6168</td>
<td>0.880 (0.035)</td>
<td>0.0017 (0.0002)</td>
<td>0.0047 (0.0010)</td>
<td>2156 (465)</td>
<td>4.454**</td>
</tr>
<tr>
<td></td>
<td>$D. arborea$</td>
<td>-1.2315</td>
<td>0.871 (0.051)</td>
<td>0.0020 (0.0004)</td>
<td>0.0083 (0.0024)</td>
<td>1154 (393)</td>
<td>3.400**</td>
</tr>
<tr>
<td>$ASA P3/TM6-A$</td>
<td>$A. sandwicense$</td>
<td>-0.5782</td>
<td>0.905 (0.046)</td>
<td>0.0031 (0.0007)</td>
<td>0.0061 (0.0019)</td>
<td>328 (221)</td>
<td>0.555</td>
</tr>
<tr>
<td></td>
<td>$D. ciliolata$</td>
<td>-1.2572*</td>
<td>0.905 (0.047)</td>
<td>0.0043 (0.0007)</td>
<td>0.0187 (0.0038)</td>
<td>389 (97)</td>
<td>3.081**</td>
</tr>
<tr>
<td></td>
<td>$D. arborea$</td>
<td>-0.6855</td>
<td>0.863 (0.064)</td>
<td>0.0033 (0.0007)</td>
<td>0.0082 (0.0028)</td>
<td>563 (259)</td>
<td>1.105</td>
</tr>
<tr>
<td>$ASA P3/TM6-B$</td>
<td>$A. sandwicense$</td>
<td>0.8234</td>
<td>0.683 (0.091)</td>
<td>0.0015 (0.0002)</td>
<td>0.0003 (0.0001)</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td></td>
<td>$D. ciliolata$</td>
<td>-0.6910</td>
<td>0.511 (0.164)</td>
<td>0.0005 (0.0002)</td>
<td>0.0016 (0.0010)</td>
<td>10000 (4413)</td>
<td>4.139**</td>
</tr>
<tr>
<td></td>
<td>$D. arborea$</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>$ASCAB9$</td>
<td>$A. sandwicense$</td>
<td>-1.0095</td>
<td>0.448 (0.134)</td>
<td>0.0006 (0.0003)</td>
<td>0.0022 (0.0015)</td>
<td>10000 (8707)</td>
<td>5.61**</td>
</tr>
<tr>
<td></td>
<td>$D. ciliolata$</td>
<td>-1.0541</td>
<td>0.950 (0.024)</td>
<td>0.0033 (0.0004)</td>
<td>0.0581 (0.0072)</td>
<td>1276 (77)</td>
<td>21.929**</td>
</tr>
<tr>
<td></td>
<td>$D. arborea$</td>
<td>-0.1989</td>
<td>0.933 (0.040)</td>
<td>0.0057 (0.0005)</td>
<td>0.0198 (0.0066)</td>
<td>446 (149)</td>
<td>2.969**</td>
</tr>
</tbody>
</table>

** $P < 0.01$

$\theta_e$ (nucleotide diversity) is the average number of nucleotide differences per site between two sequences, all sites considered.

$\theta_F$ is the maximum likelihood estimate of the population-mutation parameter at silent sites jointly estimated with the exponential growth parameter (g).

na (not amplified) refers to lack of $ASA P3/TM6-B$ amplification from all but one $D. arborea$ individual.

NC (not computable) refers to non-convergence of the exponential population growth parameter at $ASA P3/TM6-B$ in $A. sandwicense$. 

Table 7. MDIV joint estimates of gene and species divergence times (TMRCA = t / 2N_e), migration rate (M = 2N_e m), and the population mutation parameter (θ = 4N_e µ). 95% credibility intervals are indicated in parentheses.
<table>
<thead>
<tr>
<th></th>
<th>Gene divergence</th>
<th>Species divergence</th>
<th>$\mathbf{M}$</th>
<th>$\theta$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. sandwicense</strong> vs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASAP1-A</td>
<td>3.76</td>
<td>2.81 (1.15, 6.92)</td>
<td>0.01</td>
<td>4.77</td>
</tr>
<tr>
<td><strong>D. ciliolata</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASAP1-B</td>
<td>4.21</td>
<td>2.66 (1.30, 9.92)</td>
<td>0.02</td>
<td>5.65</td>
</tr>
<tr>
<td>ASAP3/TM6-A</td>
<td>3.00</td>
<td>1.56 (0.68, 9.52)</td>
<td>0.04</td>
<td>2.72</td>
</tr>
<tr>
<td>ASAP3/TM6-B</td>
<td>9.36</td>
<td>10.53 (2.91, 15.00)</td>
<td>0.01</td>
<td>1.12</td>
</tr>
<tr>
<td>ASCAB9</td>
<td>3.24</td>
<td>1.96 (0.56, 9.54)</td>
<td>0.15</td>
<td>1.12</td>
</tr>
<tr>
<td><strong>A. sandwicense</strong> vs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASAP1-A</td>
<td>5.83</td>
<td>4.42 (1.64, 9.86)</td>
<td>0.01</td>
<td>2.88</td>
</tr>
<tr>
<td><strong>D. arborea</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASAP1-B</td>
<td>3.40</td>
<td>2.25 (0.91, 4.88)</td>
<td>0.01</td>
<td>7.17</td>
</tr>
<tr>
<td>ASAP3/TM6-A</td>
<td>4.20</td>
<td>2.94 (1.14, 9.68)</td>
<td>0.01</td>
<td>1.40</td>
</tr>
<tr>
<td>ASCAB9</td>
<td>3.30</td>
<td>2.14 (0.56, 9.52)</td>
<td>0.02</td>
<td>1.56</td>
</tr>
<tr>
<td><strong>D. ciliolata</strong> vs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASAP1-A</td>
<td>3.21</td>
<td>0.42 (0.18, 9.50)</td>
<td>0.52</td>
<td>1.72</td>
</tr>
<tr>
<td><strong>D. arborea</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASAP1-B</td>
<td>1.79</td>
<td>0.89 (0.18, 4.81)</td>
<td>0.87</td>
<td>3.44</td>
</tr>
<tr>
<td>ASAP3/TM6-A</td>
<td>3.69</td>
<td>0.18 (0.04, 19.00)</td>
<td>18.8</td>
<td>2.03</td>
</tr>
<tr>
<td>ASCAB9</td>
<td>3.40</td>
<td>0.65 (0.24, 11.59)</td>
<td>8.13</td>
<td>1.30</td>
</tr>
</tbody>
</table>
Table 8. Nucleotide sequences of *ASAPETAL1-A* haplotypes. Position in alignment is indicated by top numbers. Labels are cross-referenced with specific haplotypes in figure 3. Haplotypes A - R refer to *Dubautia ciliolata* and *D. arborea* haplotypes and haplotypes i - ix refer to *Argyroxiphium sandwicense* haplotypes.

```
A   TAGAGCCGCACCTCTCCGCATCCCGTTCTCTAAAAGTGCCGCCCTCAATATTATGCGGTAGATCGAGAAGAG
B  ...................................A...................................
C     ...G...............................A............................G......
D ...............T................................................G......
E ...............T..........T.....................................G......
F ...............T....................T...........................G......
G    ...............T....................T.....................G.....G......
H ................................................G....A.......C..........
I .............C..................................G....A.......C..........
J .............C..A...............................G....A.......C..........
K   .............C...A..............................G....A.......C..........
L      .............C..................................G....A..A....C..........
M .......................................C........G....A.......C..........
N .......................................C...............................
O      ....................T..................C...............................
P      .......................................C.................G.............
Q    ..C.......AC.........T..CC...-...C.....-.T.......G.........A.........GA
R ..C....T..AC......G..T..CC...-...C.....-.T.......G.........A.........GA
i   C.....G.AG....T....CC..CC.....TTC.....TG..T..GGC...CC.T.....G....A.G.G.
ii   C.....G.AG....T....CC..CC....ATTC.....TG..T..GGC...CC.T.....G.T..A.G.G.
iii CT....G.AG....T....CC..CC.....TTC.A...TG..T..GGC...CC.TA....G....A.G.G.
iv CT....G.AG....T....CC..CC.....TTC.A...TGT.T..GGC...CC.TA....G....A.G.G.
v CT...AG..G....T....CC..CC..G..TTC....TTG..T..GGC...CC.T.....G..C.A.GAG.
viii CT....G..G....T....CC.ACC..GG.TTC.....TG..T.TGGC...CC.T.....G....A.G.G.
nix CT....G..G....T....CC.ACC..GG.TTC.....TG..T.TGGC...CC.T.....G....A.G.G.
```
Table 9. Nucleotide sequences of *ASAPETALA1-B* haplotypes. Position in alignment is indicated by top numbers. Labels are cross-referenced with specific haplotypes in figure 4. Haplotypes A - P refer to *Dubautia ciliolata* and *D. arborea* haplotypes. Haplotypes i - xi refer to *Argyroxyphium sandwicense* haplotypes.
Table 10. Nucleotide sequences of ASAPETALA3-A haplotypes. Position in alignment is indicated by top numbers. Labels are cross-referenced with specific haplotypes in figure 5. Haplotypes A - U refer to *Dubautia ciliolata* and *D. arborea* haplotypes and haplotypes i - vii refer to *Argyroxiphium sandwicense* haplotypes.
Table 11. Nucleotide sequences of ASAPETAL3-B haplotypes. Position in alignment is indicated by top numbers. Labels are cross-referenced with specific haplotypes in figure 6. Haplotypes A - C refer to *Dubautia ciliolata* and *D. arborea* haplotypes and haplotypes i - v refer to *Argyroserium sandwicense* haplotypes.

- **A**: CGAAACGCAGCTCCTCG
- **B**: ................T
- **C**: ..........A......
- **i**: ..CC.AATTA..AG.T.
- **ii**: ..CCTAATTA..AG.T.
- **iii**: .ACC.AATTA..AG.T.
- **iv**: TACC.AATTA..AG.T.
- **v**: ..CC.AATTA.CAGGT.
Table 12. Nucleotide sequences of ASCAB9 haplotypes. Position in alignment is indicated by top numbers. Labels are cross-referenced with specific haplotypes in figure 7. Haplotypes A - T refer to *Dubautia ciliolata* and *D. arborea* haplotypes and haplotypes i - iii refer to *Argyroxyphium sandwicense* haplotypes.
Figure 1. Populations of endemic *Dubautia ciliolata* subsp. *glutinosa* and *Dubautia arborea* sampled from Mauna Kea, Hawaii. *Dubautia ciliolata* populations are indicated by circles with the following labels: (1) Puu Kanakaleonui, (2) Puu Kawiiwi, and (3) Waipahoeohoe gulch. *Dubautia arborea* populations are indicated by triangles with the following labels: (1) Puu Laau, (2) Puu Mali, and (3) Waipahoeohoe gulch.
Figure 2. Populations of endemic *Argyroxyphium sandwicense* subsp. *macrocephalum* sampled in Haleakala National Park, Maui. These populations are labeled as follows:

(1) Silversword Loop, (2) Puu o Pele, (3) Ka Moa o Pele, (4) Puu Naue.
Figures 3 through 7. Statistical parsimony haplotype networks of all genes examined in this study from *A. sandwicense*, *D. arborea*, and *D. ciliolata*. Squares indicate *A. sandwicense* haplotypes (labeled in lowercase roman numerals) with the number of sampled individuals of each haplotype denoted by subscripted numbers. Circles indicate *D. ciliolata* and *D. arborea* haplotypes (labeled in uppercase letters) with *D. ciliolata* haplotypes depicted as white circles and *D. arborea* haplotypes depicted as black circles. Shared haplotypes between *D. ciliolata* and *D. arborea* are denoted as pie charts. The first subscript number below *Dubautia* haplotypes denotes the number of *D. ciliolata* sampled individuals with the designated haplotype and the second subscript number denotes the number of *D. arborea* individuals with the same haplotype. Haplotype labels are cross-referenced with haplotype sequences shown in Tables 8 through 12.
Figure 3. Statistical parsimony haplotype network of *ASAPETALAI-A*. 
Figure 4. Statistical parsimony haplotype network of *ASAPETALA1-B*.
Figure 5. Statistical parsimony haplotype network of ASAPETALA3/TM6-A. * indicates homoplasy between haplotypes G and L.
Figure 6. Statistical parsimony haplotype network of \textit{ASAPETALA3/TM6-B}.
Figure 7. Statistical parsimony haplotype network of \textit{ASCAB9}. * indicates homoplasy between haplotypes A and P.
CHAPTER 7

SUMMARY AND PROSPECTS
Gene and genome duplication are prominent feature in plant evolutionary history and appears to play a major role in organismal diversification. Understanding the short and long term dynamics among gene duplicates is an important problem in evolutionary genetics, as these events strongly influence intra- and interspecific genome diversity and may be involved with adaptive differentiation. Furthermore, duplicated genes are the building blocks for the origin and diversification of gene families, particularly regulatory genes that strongly influence many aspects of development and disease resistance. In the preceding chapters, the molecular evolution and population genetics of duplicated floral regulatory genes are discussed with respect to two systems: paralogous and orthologous genes in *Arabidopsis* and homoeologous genes in the allopolyploid Hawaiian silversword alliance adaptive radiation.

In the first research chapter (chapter 4), molecular evolutionary patterns of the ancestral *APETALA3/PISTILLATA* (*AP3/PI*) gene duplication and the more recent *APETALA1/CAULIFLOWER* (*AP1/CAL*) duplication (within Brassicaceae) were compared among Brassicaceae paralogs and their orthologs in species from two distantly-related plant families (Scrophulariaceae and Caryophyllaceae). Examination of all four loci uncovered contrasting dynamics of molecular evolution between these two sets of paralogous regulatory genes. These results indicate that changes in rates and patterns of sequence evolution may result in divergent developmental functions of some paralogous regulatory genes.

The dramatic influence of floral regulatory genes on plant development suggests that duplications of these genes may diverge in lineages with striking floral trait differences. To investigate the evolutionary dynamics of duplicated floral regulatory genes in an adaptive radiation exhibiting diverse reproductive morphologies, chapters 5 and 6 examined the extent of nucleotide divergence of duplicated floral regulatory genes among three species of the Hawaiian silversword alliance from two major lineages of this radiation: *Argyroxyphium sandwicense* subsp. *macrocephalum* (*A. sandwicense*), *Dubautia ciliolata* subsp. *glutinosa* (*D. ciliolata*), and *Dubautia arborea*. 

206
Duplicated orthologs of the Arabidopsis *APETALA1* (*ASAP1*) and *APETALA3* (*ASAP3/TM6*) were isolated and sequenced from *A. sandwicense* and *D. ciliolata*. The duplicate copies of *ASAP1* and *ASAP3/TM6* show contrasting patterns. Homoeologous *ASAP1* genes show similar levels of nucleotide diversity and haplotype divergence in both species, while homoeologous *ASAP3/TM6* genes have different levels and patterns of variation in *Dubautia ciliolata*. Additionally, *D. ciliolata* may be segregating for a moderate frequency null allele in one of the *ASAP3/TM6* homoeologs. These results suggest that evolutionary forces can lead to divergent evolutionary trajectories between duplicated genes. Furthermore, similar patterns of low frequency polymorphisms in both species were observed across most loci. This observed low level of nucleotide diversity and excess number of singletons among all four regulatory genes suggest that processes affecting the entire genome may be significant contributors to genetic variation in *A. sandwicense* and *D. ciliolata*.

To further explore the extent of genome-wide forces shaping genetic variation, a multilocus study of gene evolution in *A. sandwicense*, *D. ciliolata* and the sibling species of *D. ciliolata* (*D. arborea*) was conducted, as presented in chapter six. Multiple genes were examined for nucleotide and haplotype diversity, including homoeologous floral regulatory genes that show evidence of divergence in *D. ciliolata* (*ASAP3/TM6-A* and *ASAP3/TM6-B*), and other genes that do not deviate from equilibrium-neutral model expectations (homoeologous floral regulatory genes *ASAP1-A* and *ASAP1-B*, the housekeeping gene *ASCAB9*, and the mitochondrial gene *ASNAD1*). Using this multilocus approach, we uncovered low nucleotide diversity and patterns consistent with population expansion in all three species. Investigation of the recently-derived sibling species (*D. ciliolata* and *D. arborea*) revealed high morphological divergence between these two species at eight traits, low overall genetic divergence, low levels of shared haplotypes, lack of population substructure within either sibling species, and high rates of migration between these two species that may perturb genetic divergence between sibling species through drift.
The differential evolution uncovered between homoeologous copies of \textit{ASAPETALA3/TM6} in \textit{Dubautia ciliolata} (and possibly \textit{D. arborea}) is intriguing. The \textit{Arabidopsis} ortholog of \textit{ASAPETALA3/TM6} (\textit{APETALA3}) has been shown to be involved with floral organ identity by regulating stamen and petal formation, and the size of these organs have been mapped to QTL in \textit{A. thaliana} that include \textit{APETALA3} (Juenger et al. 2000). Although \textit{ASAPETALA3/TM6} function has not been characterized in the Hawaiian silverswords, our analyses of \textit{ASAPETALA3/TM6} and morphological data suggest a plausible link between the evolutionary divergence of homoeologous gene function and morphological differentiation between \textit{Dubautia arborea} and \textit{Dubautia ciliolata} ssp. glutinosa.

The apparent selective sweep differentiating \textit{ASAPETALA3/TM6-B} from \textit{ASAPETALA3/TM6-A} suggests functional divergence between these two gene copies. Because these are MADS-box transcription factors, such divergence could possibly result in altered regulatory activity. An experiment to determine gene expression patterns of the homoeologous copies of \textit{ASAPETALA3/TM6} could be performed to probe for a correlation between the divergence of these homoeologs and differential gene expression. The relative expression levels of both gene copies could be quantified in individuals with differing quantitative measures of floral morphologies in both species and in hybrids. To quantify the relative expression levels of these homoeologous genes quantitative real-time RT-PCR could be used to compare homoeologous copies within species as well as to compare gene expression between taxa.

Adaptive radiation, especially among insular species, occurs through a dynamic series of ecological interactions across stochastic environmental gradients. This complicates the examination of evolutionary models of speciation and differentiation. In our analysis, we find three factors associated with the establishment of new species in this radiation: (1) morphological differentiation proceeds more rapidly than genetic differentiation, (2) population expansion is prevalent, and (3) migration rates can be high in morphologically divergent sibling species. To further dissect the patterns of genetic differentiation,
population expansion, and migration in the Hawaiian silversword alliance adaptive radiation, future analyses could include samples among additional closely related species pairs across this insular group. One possibility would be to compare our results sampled from the recently derived *Dubautia* species with 13 chromosome pairs (n = 13) to another group of *Dubautia* species with 14 chromosome pairs (n = 14), such as *Dubautia scabra* subsp. *leiophylla* and *Dubautia plantaginea* (Koolau). These two species exhibit dramatically different morphologies and have a low Nei’s genetic-distance coefficient (*I* = 0.097) [Nei 1972; Witter & Carr 1988]. The n = 14 clade has been shown to have increased genetic divergence among morphologically distinct closely related species (Witter & Carr 1988). It would also be interesting to examine other members of the same *Dubautia* clade examined in this current study (n = 13) to assess the distribution of ancestral alleles among and within species and to investigate the extent of divergence between *ASAP3/TM6-A* and *ASAP3/TM6-B* across this plant group. Such analyses would shed light on the relative contributions of isolation and migration across the Hawaiian silversword alliance, further probe the dynamics of duplicated regulatory genes (hence contribute to our understanding of the origin and diversification of gene family members) and present testable hypotheses that can be further examined in other examples of adaptive radiation and polyploid groups.
LITERATURE CITED


APPENDIX

SUPPLEMENTAL FIGURES FOR CHAPTER
Figure A1. Means comparisons of the number of capitula per capitulescense (HC) by population.

<table>
<thead>
<tr>
<th>POPULATION</th>
<th>MC 0.05</th>
<th>All Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>D. arborea_Laau</td>
<td>0.391</td>
<td>0.391</td>
</tr>
<tr>
<td>D. arborea_Mali</td>
<td>0.391</td>
<td>0.391</td>
</tr>
<tr>
<td>D. ciliolata_Gulch</td>
<td>212.212</td>
<td>212.212</td>
</tr>
<tr>
<td>D. ciliolata_Kawiiwi</td>
<td>0.391</td>
<td>0.391</td>
</tr>
<tr>
<td>D. ciliolata_Kanakaleonui</td>
<td>0.391</td>
<td>0.391</td>
</tr>
</tbody>
</table>

Comparisons for all population pairs using Tukey-Kramer HSD, $q^* = 2.95415$; positive values show pairs of means that are significantly different.

### Dubautia arborea populations

<table>
<thead>
<tr>
<th>Abs(Dif)-LSD</th>
<th>Puu Mali</th>
<th>Puu Laau</th>
<th>D. arborea_gulch</th>
<th>Hybrid zone</th>
<th>D. ciliolata_Gulch</th>
<th>Puu Kawiiwi</th>
<th>Puu Kanakaleonui</th>
</tr>
</thead>
<tbody>
<tr>
<td>D. arborea_Mali</td>
<td>-1.324</td>
<td>0.391</td>
<td>1.172</td>
<td>6.254</td>
<td>11.629</td>
<td>12.879</td>
<td>13.419</td>
</tr>
<tr>
<td>D. arborea_Lauu</td>
<td>0.391</td>
<td>-1.333</td>
<td>-0.552</td>
<td>4.530</td>
<td>9.904</td>
<td>11.154</td>
<td>11.694</td>
</tr>
<tr>
<td>D. arborea_gulch</td>
<td>1.172</td>
<td>-0.552</td>
<td>-1.320</td>
<td>3.762</td>
<td>9.137</td>
<td>10.387</td>
<td>10.927</td>
</tr>
<tr>
<td>Hybrid_zone</td>
<td>6.254</td>
<td>4.530</td>
<td>3.762</td>
<td>-1.329</td>
<td>4.046</td>
<td>5.296</td>
<td>5.836</td>
</tr>
<tr>
<td>D. ciliolata_Gulch</td>
<td>11.629</td>
<td>9.904</td>
<td>9.137</td>
<td>4.046</td>
<td>-1.324</td>
<td>-0.074</td>
<td>0.466</td>
</tr>
<tr>
<td>D. ciliolata_Kawiiwi</td>
<td>12.879</td>
<td>11.154</td>
<td>10.387</td>
<td>5.296</td>
<td>-0.074</td>
<td>-1.320</td>
<td>-0.780</td>
</tr>
<tr>
<td>D. ciliolata_Kanakaleonui</td>
<td>13.419</td>
<td>11.694</td>
<td>10.927</td>
<td>5.836</td>
<td>0.466</td>
<td>-0.780</td>
<td>-1.320</td>
</tr>
</tbody>
</table>

### Dubautia ciliolata populations

<table>
<thead>
<tr>
<th>Abs(Dif)-LSD</th>
<th>Puu Mali</th>
<th>Puu Laau</th>
<th>D. ciliolata_gulch</th>
<th>Hybrid zone</th>
<th>D. ciliolata_Kawiiwi</th>
<th>Puu Kawiiwi</th>
<th>Puu Kanakaleonui</th>
</tr>
</thead>
<tbody>
<tr>
<td>D. ciliolata_Gulch</td>
<td>11.629</td>
<td>9.904</td>
<td>9.137</td>
<td>4.046</td>
<td>-1.324</td>
<td>-0.074</td>
<td>0.466</td>
</tr>
<tr>
<td>D. ciliolata_Kawiiwi</td>
<td>12.879</td>
<td>11.154</td>
<td>10.387</td>
<td>5.296</td>
<td>-0.074</td>
<td>-1.320</td>
<td>-0.780</td>
</tr>
<tr>
<td>D. ciliolata_Kanakaleonui</td>
<td>13.419</td>
<td>11.694</td>
<td>10.927</td>
<td>5.836</td>
<td>0.466</td>
<td>-0.780</td>
<td>-1.320</td>
</tr>
</tbody>
</table>
Figure A2. Means comparisons of the number of florets per capitulum (FH) by population.

Comparisons for all population pairs using Tukey-Kramer HSD, $q^* = 2.95415$; positive values show pairs of means that are significantly different.
Figure A4. Means comparisons of the receptacular bract length (LRB, in millimeters) by population.

<table>
<thead>
<tr>
<th></th>
<th>Dubautia arborea populations</th>
<th>Dubautia ciliolata populations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Abs(Dif)-LSD</td>
<td>Puu Mali</td>
</tr>
<tr>
<td>D.arborea_Mali</td>
<td>-0.4408</td>
<td>0.0123</td>
</tr>
<tr>
<td>D.arborea_Laau</td>
<td>0.0123</td>
<td>-0.4453</td>
</tr>
<tr>
<td>D.arborea_gulch</td>
<td>0.2885</td>
<td>-0.1691</td>
</tr>
<tr>
<td>Hybrid_zone</td>
<td>0.3739</td>
<td>-0.0837</td>
</tr>
<tr>
<td>D.ciliolata_Gulch</td>
<td>0.5910</td>
<td>0.1334</td>
</tr>
<tr>
<td>D.ciliolata_Kawiiwi</td>
<td>1.0252</td>
<td>0.5676</td>
</tr>
<tr>
<td>D.ciliolata_Kanakaleomui</td>
<td>1.2279</td>
<td>0.7703</td>
</tr>
</tbody>
</table>

Comparisons for all population pairs using Tukey-Kramer HSD, $q^* = 2.95415$; positive values show pairs of means that are significantly different.
Figure A4. Means comparisons of the sepal length (LS, in millimeters) by population.

Comparisons for all population pairs using Tukey-Kramer HSD, \( q^* = 2.95415 \); positive values show pairs of means that are significantly different.

### Dubautia arborea populations

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Puu Mali</th>
<th>Puu Laau</th>
<th>D. arborea gulch</th>
<th>Hybrid zone</th>
</tr>
</thead>
<tbody>
<tr>
<td>D. arborea Mali</td>
<td>-0.2463</td>
<td>0.1115</td>
<td>0.1812</td>
<td>0.2277</td>
</tr>
<tr>
<td>D. arborea Laau</td>
<td>0.1115</td>
<td>-0.2471</td>
<td>-0.1774</td>
<td>-0.1309</td>
</tr>
<tr>
<td>D. arborea gulch</td>
<td>0.1812</td>
<td>-0.2477</td>
<td>-0.2488</td>
<td>-0.2023</td>
</tr>
<tr>
<td>Hybrid zone</td>
<td>0.2277</td>
<td>-0.1309</td>
<td>-0.2023</td>
<td>-0.2463</td>
</tr>
</tbody>
</table>

### Dubautia ciliolata populations

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Puu Kawiwi</th>
<th>Puu Kanakaleomui</th>
</tr>
</thead>
<tbody>
<tr>
<td>D. ciliolata gulch</td>
<td>0.4908</td>
<td>0.8048</td>
</tr>
<tr>
<td>D. ciliolata Kawiwi</td>
<td>0.8048</td>
<td>1.3657</td>
</tr>
<tr>
<td>D. ciliolata_Kanakaleomui</td>
<td>1.3657</td>
<td>0.3138</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Puu Kawiwi</th>
<th>Puu Kanakaleomui</th>
</tr>
</thead>
<tbody>
<tr>
<td>D. ciliolata gulch</td>
<td>0.4908</td>
<td>0.8048</td>
</tr>
<tr>
<td>D. ciliolata Kawiwi</td>
<td>0.8048</td>
<td>1.3657</td>
</tr>
<tr>
<td>D. ciliolata_Kanakaleomui</td>
<td>1.3657</td>
<td>0.3138</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Puu Kawiwi</th>
<th>Puu Kanakaleomui</th>
</tr>
</thead>
<tbody>
<tr>
<td>D. ciliolata gulch</td>
<td>0.4908</td>
<td>0.8048</td>
</tr>
<tr>
<td>D. ciliolata Kawiwi</td>
<td>0.8048</td>
<td>1.3657</td>
</tr>
<tr>
<td>D. ciliolata_Kanakaleomui</td>
<td>1.3657</td>
<td>0.3138</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Puu Kawiwi</th>
<th>Puu Kanakaleomui</th>
</tr>
</thead>
<tbody>
<tr>
<td>D. ciliolata gulch</td>
<td>0.4908</td>
<td>0.8048</td>
</tr>
<tr>
<td>D. ciliolata Kawiwi</td>
<td>0.8048</td>
<td>1.3657</td>
</tr>
<tr>
<td>D. ciliolata_Kanakaleomui</td>
<td>1.3657</td>
<td>0.3138</td>
</tr>
</tbody>
</table>
Figure A5. Means comparisons of the corolla length (LC, in millimeters) by population.

Comparisons for all population pairs using Tukey-Kramer HSD, $q^* = 2.95415$; positive values show pairs of means that are significantly different.

**Dubautia arborea populations**

<table>
<thead>
<tr>
<th>Abs(Dif)-LSD</th>
<th>Puu Mali</th>
<th>Puu Laau</th>
<th>D.arborea_gulch</th>
<th>Hybrid_zone</th>
<th>D.ciliolata_gulch</th>
<th>Puu Kawiwi</th>
<th>Puu Kanakaleonui</th>
</tr>
</thead>
<tbody>
<tr>
<td>D.arborea_Mali</td>
<td>-0.2552</td>
<td>0.1234</td>
<td>0.4358</td>
<td>0.5708</td>
<td>0.9568</td>
<td>1.3203</td>
<td>1.3868</td>
</tr>
<tr>
<td>D.arborea_Laau</td>
<td>0.1234</td>
<td>-0.2578</td>
<td>0.0546</td>
<td>-0.1210</td>
<td>0.5757</td>
<td>0.9392</td>
<td>1.0057</td>
</tr>
<tr>
<td>D.arborea_gulch</td>
<td>0.4358</td>
<td>0.0546</td>
<td>-0.2560</td>
<td>-0.1210</td>
<td>0.2651</td>
<td>0.6285</td>
<td>0.6951</td>
</tr>
<tr>
<td>Hybrid_zone</td>
<td>0.5708</td>
<td>0.1897</td>
<td>-0.1210</td>
<td>-0.2586</td>
<td>0.1274</td>
<td>0.4909</td>
<td>0.5574</td>
</tr>
</tbody>
</table>

**Dubautia ciliolata populations**

<table>
<thead>
<tr>
<th>Abs(Dif)-LSD</th>
<th>Puu Mali</th>
<th>Puu Laau</th>
<th>D.arborea_gulch</th>
<th>Hybrid_zone</th>
<th>D.ciliolata_gulch</th>
<th>Puu Kawiwi</th>
<th>Puu Kanakaleonui</th>
</tr>
</thead>
<tbody>
<tr>
<td>D.ciliolata_Gulch</td>
<td>0.9568</td>
<td>0.5757</td>
<td>0.2651</td>
<td>0.1274</td>
<td>-0.2552</td>
<td>0.1083</td>
<td>0.1748</td>
</tr>
<tr>
<td>D.ciliolata_Kawiwi</td>
<td>1.3203</td>
<td>0.9392</td>
<td>0.6285</td>
<td>0.4909</td>
<td>0.1083</td>
<td>-0.2560</td>
<td>-0.1895</td>
</tr>
<tr>
<td>D.ciliolata_Kanakaleonui</td>
<td>1.3868</td>
<td>1.0057</td>
<td>0.6951</td>
<td>0.5574</td>
<td>0.1748</td>
<td>-0.1895</td>
<td>-0.2552</td>
</tr>
</tbody>
</table>
Figure A6. Means comparisons of the ovary length (LO, in millimeters) by population.

Comparisons for all population pairs using Tukey-Kramer HSD, $q^* = 2.95415$; positive values show pairs of means that are significantly different.

<table>
<thead>
<tr>
<th>Dubautia arborea populations</th>
<th>Dubautia ciliolata populations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abs(Dif)-LSD</td>
<td>Puu Mali</td>
</tr>
<tr>
<td>D.arborea_Mali</td>
<td>-0.24511</td>
</tr>
<tr>
<td>D.arborea_Lauu</td>
<td>-0.03199</td>
</tr>
<tr>
<td>D.arborea_gulch</td>
<td>0.01672</td>
</tr>
<tr>
<td>Hybrid_zone</td>
<td>0.05601</td>
</tr>
<tr>
<td>D.ciliolata_Gulch</td>
<td>0.27734</td>
</tr>
<tr>
<td>D.ciliolata_Kawiwi</td>
<td>0.43819</td>
</tr>
<tr>
<td>D.ciliolata_Kanakaleonui</td>
<td>0.53735</td>
</tr>
</tbody>
</table>
Figure A7. Means comparisons of the leaf length (LL, in millimeters) by population.

Comparisons for all population pairs using Tukey-Kramer HSD, $q^* = 2.95415$; positive values show pairs of means that are significantly different.

<table>
<thead>
<tr>
<th></th>
<th>Dubautia arborea populations</th>
<th>Dubautia ciliolata populations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Abs(Dif)-LSD</td>
<td></td>
</tr>
<tr>
<td>D.arborea_Mali</td>
<td>PUU Mali 1.414</td>
<td>PUU Kawiwi 9.059</td>
</tr>
<tr>
<td>D.arborea_Lauu</td>
<td>PUU Laau 1.599</td>
<td>PUU Kanakaleonui 9.158</td>
</tr>
<tr>
<td>D.arborea_gulch</td>
<td>1.599 1.414</td>
<td>11.257</td>
</tr>
<tr>
<td>Hybrid_zone</td>
<td>6.849 4.599</td>
<td>11.592</td>
</tr>
<tr>
<td>D.ciliolata_Gulch</td>
<td>11.257 9.008</td>
<td></td>
</tr>
<tr>
<td>D.ciliolata_Kawiwi</td>
<td>11.309 9.059</td>
<td></td>
</tr>
<tr>
<td>D.ciliolata_Kanakaleonui</td>
<td>11.592 9.158</td>
<td></td>
</tr>
</tbody>
</table>
Figure A8. Means comparisons of the maximum leaf width (MWL, in millimeters) by population.

Comparisons for all population pairs using Tukey-Kramer HSD, $q^* = 2.95415$; positive values show pairs of means that are significantly different.

<table>
<thead>
<tr>
<th></th>
<th>Dubautia arborea populations</th>
<th>Dubautia ciliolata populations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Abs(Dif)-LSD</td>
<td>Puu Mali</td>
</tr>
<tr>
<td>D.arborea_Mali</td>
<td>-0.835</td>
<td>1.414</td>
</tr>
<tr>
<td>D.arborea_Laau</td>
<td>1.414</td>
<td>-0.835</td>
</tr>
<tr>
<td>D.arborea_gulch</td>
<td>1.599</td>
<td>-0.651</td>
</tr>
<tr>
<td>Hybrid_zone</td>
<td>6.849</td>
<td>4.599</td>
</tr>
<tr>
<td>D.ciliolata_Gulch</td>
<td>11.257</td>
<td>9.008</td>
</tr>
<tr>
<td>D.ciliolata_Kawiwi</td>
<td>11.309</td>
<td>9.059</td>
</tr>
<tr>
<td>D.ciliolata_Kanakaleomui</td>
<td>11.592</td>
<td>9.343</td>
</tr>
</tbody>
</table>
Figure A9. Statistical parsimony haplotype network and polymorphism table of *ASNAD1* alleles examined in this study.