MITOCHONDRIAL HETEROPLASMY IN MIMULUS GUTTATUS

A thesis submitted
to Kent State University in partial
fulfillment of the requirement of the
degree of Master of Science

by

Eric R. Floro

May 2011
# TABLE OF CONTENTS

LIST OF FIGURES .........................................................................................................................iv  
LIST OF TABLES ..............................................................................................................................v  
ACKNOWLEDGEMENTS ...................................................................................................................vi  
CHAPTER 1: ......................................................................................................................................1-12  
CHAPTER 2: ......................................................................................................................................13-47  
  INTRODUCTION .............................................................................................................................13-14  
  METHODS .......................................................................................................................................15-30  
  RESULTS ........................................................................................................................................31-41  
  DISCUSSION ..................................................................................................................................42-47  
CHAPTER 3: ......................................................................................................................................48-70  
  INTRODUCTION .............................................................................................................................48-50  
  METHODS .......................................................................................................................................50-56  
  RESULTS ........................................................................................................................................56-65  
  DISCUSSION ..................................................................................................................................66-70  
CHAPTER 4: ......................................................................................................................................71-77  
REFERENCES ....................................................................................................................................78-83
LIST OF FIGURES

FIGURE 2.1...........................................................................................................17
FIGURE 2.2...........................................................................................................19
FIGURE 2.3...........................................................................................................22
FIGURE 2.4...........................................................................................................33
FIGURE 2.5...........................................................................................................34
FIGURE 2.6...........................................................................................................37
FIGURE 2.7...........................................................................................................40-41
FIGURE 3.1...........................................................................................................59-60
FIGURE 3.2...........................................................................................................64-65
ACKNOWLEDGEMENTS

The help and assistance of many people have made this thesis possible. Most importantly I am extremely grateful for my Advisor Dr. Andrea Case whose encouragement, enthusiasm, and guidance have made this thesis possible and successful. I thank Dr. Chris Blackwood for being part of my committee and for his assistance with all things qPCR. I would also like to thank Dr. Dave McCauley and Dr. Randy Hoeh for being on my guidance committee and for their helpful insights. I thank Chris Rizzo for his assistance with plant and greenhouse care, and Young Wha Lee, and Julius Mojica for their assistance with the location of study populations. I would also like to thank my fellow graduate students for their assistance with lab work, and advice. I am particularly grateful to Larry Feinstein, Stephanie Hovatter, Hannah Madson, Mike Monfredi, Julie Proell, Justin Reeves, and Daniel Sprockett. Additionally I would like to thank my family for their love and encouragement throughout my life and academic career. Last, but definitely not least, I would like to thank Vanessa Potter for her love, encouragement, and support over the last 5 years. Without all of you this thesis would not have been possible, thank you!
CHAPTER 1:

GENERAL INTRODUCTION

Eukaryotic organisms have two distinct genomes—the nuclear genome and the cytoplasmic genome. Aside from their locations within the cell, they differ in size, physical structure, function, and in the way they are transmitted to offspring. The nuclear genome of any individual is made up of DNA from both of its parents; each parent contributes half of the genome to an offspring. The organellar genome in eukaryotes is usually passed uniparentally, and usually from the maternal parent (Birky 2001) to offspring. This form of uniparental transmission should result in individuals having only one genetically unique organellar genome, referred to as homoplasmy (Birky 2001).

Organellar genomes and eukaryotes are thought to have arisen through a process called endosymbiosis. Their evolutionary origins may have resulted in within-individual genetic conflict that directly contributes to uniparental inheritance (Burt & Trivers 2006). The theory of endosymbiosis is that organelles containing DNA within eukaryotic cells evolved from free-living prokaryotic ancestors (Kutschera & Niklas 2004). During their early evolution, there would have been competition among the precursors to organelles for survival within their host cells. In this case, natural selection would favor the organellar precursors with the
smallest genomes, as they would be able to replicate the most quickly. Within-cell competition could have been harmful to the host cell if the host cell relied on the organelles to carry out metabolic functions, because smaller genome size could equate to missing portions of essential organellar DNA. Selection should therefore favor a situation where there is limited within-individual competition, reducing the potentially negative effects of that competition to the eukaryotic host. The easiest way for this to happen is for the organellar genome to be uniparentally transmitted to offspring; if only one unique genome is present, within-individual selection is not possible.

In addition to limiting within-individual competition, enforced uniparental inheritance can maintain intergenomic compatibility. Nuclear genes regulate most of the transcription and translation in organelles, thus intergenomic compatibility and co-evolution are of critical importance (Cosmides & Tooby 1981, Rand et al. 2004). This extends another layer of control over the organelles, and because the organelles cannot function without an appropriately matched nuclear genome this is a symbiotic relationship. Recent evidence not only suggests that there is occasionally more than one unique organellar genome present within individuals, known as heteroplasmy, but that heteroplasmy could potentially disrupt intergenomic compatibility. Such wide-ranging evolutionary implications call for new studies of the frequency, cause, and consequence of organellar heteroplasmy in eukaryotes.
Goal of this thesis

Through this thesis, I explore the causes and consequences of non-maternal inheritance and organellar heteroplasmy in *Mimulus guttatus* (Phrymaceae), presented in two data chapters. In Chapter 2, my goal was to investigate how mitochondrial ratios differ both within and among populations and within individual plants. I did this to understand what processes may contribute to the origin and maintenance of heteroplasmy in lineages of *M. guttatus*. I screened multiple *M. guttatus* individuals to record the presence of heteroplasmy, first in natural populations, and then in their open-pollinated offspring. I also screened multiple tissue types from single individuals to assess the potential for differing mitotype ratios within individual plants. In Chapter 3, I attempted to ascertain whether heteroplasmy might affect the fitness of individual plants. I also performed a number of controlled crosses to assess the potential for the generation of heteroplasmy.

What is heteroplasmy?

Heteroplasmy is defined as the presence of more than one unique organellar genome, either mitochondria or plastid, present within an individual (Kmiec et al. 2006). Depending on how heteroplasmy arises, it may be inconsistent with strict uniparental inheritance. There are at least three ways through which heteroplasmy can arise. The first two (recombination and mutation) are completely consistent
with maternal inheritance. The third involves bi-parental inheritance, sometimes referred to as “paternal leakage,” that occurs during fertilization.

Molecular recombination between two organellar genomes within a homoplasmic individual could generate unique organellar genomes that differ from their progenitors in gene order. If both the original and the recombined genomes are maintained within individuals, this results in heteroplasm (Kmiec et al. 2006). Evidence of organellar recombination has been recorded in multiple lineages: humans (Kraytsberg 2004), other animals (Ladoukakis & Zouros 2001, Hoarau 2002, Zhao et al. 2004, Gantenbein et al. 2005, Armstrong 2007), plants (Bergthorsson 2005), and yeast (MacAlpine 1998). While most recombination events will alter gene order, and not likely have much of an effect on organellar function, recombination can create novel “chimeric” genes, where sections of different genes are linked together into a single reading frame (Budar et al. 2003, McCauley & Olson 2008). This is commonly observed in the mitochondrial genomes of plants (Chase 2007), and can produce novel (sometimes detrimental) phenotypes (Zeh & Zeh 2005, Delph et al. 2007).

Similarly, point mutations within organellar genomes could also result in heteroplasm by altering gene sequences among individual genomes already present within cells. If the mutation isn’t lethal, then the new genome could be replicated and maintained within the organism resulting in heteroplasmy. Mutations in mitochondrial genomes resulting in non-lethal disease have been
reported in humans in several cases (Anitori et al. 2005, Alston et al. 2010, and reviewed by Mazunin et al. 2010).

The third way in which heteroplasmy could be generated is not consistent with maternal inheritance. Paternal leakage, the occasional transmission of organellar genomes from the paternal parent to offspring in combination with the normal maternal inheritance, could also result in heteroplasmy (White et al. 2008). Paternal leakage resulting in heteroplasmy has been identified in multiple lineages (Kvist et al. 2003, Breton et al. 2007, Ellis et al. 2008, Bentley et al. 2010). One way in which heteroplasmy seems to be generated in nature is during hybridization events between closely related species (Coyer et al. 2004, Kijewski et al. 2006, Hansen et al. 2007). These examples suggest that the mechanisms regulating uniparental inheritance may be different among species, and consequently break down during hybridization.

**Why study heteroplasmy?**

At present, the frequency and distribution of non-uniparental inheritance in eukaryotes is not known, and as discussed previously, heteroplasmy has been documented in a wide range of organisms both in the lab and in natural populations. Despite its overall rarity, its occurrence in multiple unrelated lineages indicates that it may naturally occur much more widely than previously appreciated. There may also be similarities in mechanisms across taxa, either between how heteroplasmy arises or in how it is maintained.
In addition to basic questions about the potential causes and consequences of heteroplasmy within individuals, a more practical consideration lies in the fact that evolutionary inferences made from organellar genetic markers may not be robust if non-maternal inheritance is appreciable. Because mitochondrial DNA (mtDNA) in animals and chloroplast DNA (cpDNA) in plants are less conserved, i.e., have higher point mutation rates, compared to nuclear DNA, they are often used in conjunction with nuclear DNA to infer evolutionary relationships (Wolfe & Randle 2004). The non-coding regions of the organellar genome are useful for inferring short evolutionary histories, and coding regions are more useful for longer-term histories (Wolfe & Randle 2004). Because the organellar genome is usually maternally inherited, it is also used to infer genealogy, as in the case of mitochondrial Eve in humans (Cann et al. 1987). Heteroplasmy, whether consistent with maternal inheritance or not, could skew the evolutionary relationships that are inferred from them (Kvist et al. 2003, Wolfe & Randle 2004, Hansen et al. 2007, White et al. 2008).

Heteroplasmy also provides an opportunity to study, indirectly, the mechanisms that control organelle inheritance, both mitotically (through cell divisions within an individual) and meiotically (in the production of gametes). This can be done by examining the differences between parental and offspring haplotype ratios, and also by identifying the potential differences in haplotype ratios between different tissues within an individual (referred to as vegetative sorting). Vegetative sorting occurs when stochastic processes act on the mitotype ratios resulting in different mitotype ratios in different tissues of a single individual (Mogensen 1996).
The presence of more than one unique organellar genome may have effects on individuals if there are any incompatibilities between the nuclear and organellar genomes. Thus, it would be useful to examine any potential consequences of heteroplasmy, whether positive or negative, to individuals. If there is within individual competition between haplotypes for expression, there may be fitness-related effects. It is also possible that heteroplasmy and recombination could act to relieve mutational load within organelles (Barr et al. 2005). Studying heteroplasmy when and where it occurs is important to document its occurrence and its potential effects, to both individual fitness and our theories about how the intergenomic relationships, and eukaryotic evolution occurred.

**Why study mitochondrial heteroplasmy in plants?**

Plants present some unique opportunities when it comes to studying heteroplasmy. First, mitochondrial heteroplasmy has been discovered in multiple lineages of plants: *Silene vulgaris* (McCauley et al. 2005), *Pelargonium zonale* & *P. inquinans* hybrids (Weihe et al. 2009), and *Phaseolus vulgaris* (Woloszynska & Trojanowski 2009). All of the previous examples have been attributed to paternal leakage, which is significant. If, for example, heteroplasmy had been found in only one taxonomic group of plants and not anywhere else, it could indicate that there was a lineage-specific change at some point in evolutionary history that resulted in heteroplasmy in only that lineage. It shows that mitochondrial heteroplasmy is not the result of a single lineage of plants having a different or less specific form of
organellar inheritance, but rather is an uncommon occurrence found in many lineages.

Second, the large mitochondrial genomes of plants often contain large repeats and heteroplasmy can lead to recombination between mitotypes (Woloszynska 2010). If recombination interferes with cytonuclear compatibility, it could result in negative effects to the individual. One excellent example of this is cytoplasmic male sterility (CMS). CMS is a cytonuclear incompatibility caused by a disruption of mitochondrial function that prevents viable pollen formation (McCauley & Olson 2008); recombination is one mechanism that is thought lead to the formation of novel CMS genes (Chase 2007). When these genes are expressed, they produce novel proteins (Delph et al. 2007) that can disrupt pollen production.

**Why document variation in heteroplasmy?**

Documenting heteroplasmy in natural populations, and in multiple maternal families, is an important step as it enables researchers to determine that this occurs naturally, rather than as only occurring as a result of controlled crosses or crosses between species. Additionally, documenting heteroplasmy in multiple lineages allows us to further ascertain that this phenomenon is not lineage specific, i.e. it is found in plants, animals, and fungi, which is important to know because it may lend further evidence indicating whether or not strict maternal inheritance is the norm in eukaryotes.
It is also important to consider within-plant variability when studying heteroplasmy, primarily vegetative sorting of unique mitotypes within individual plants. If vegetative sorting were to occur, then one tissue sampled might indicate that an individual is homoplasmic while another tissue from the same plant might be heteroplasmic. If the two tissues in question are both floral buds then both heteroplasmy and homoplasmzy could be passed from the maternal plant potentially resulting in offspring with differing mitotype ratios.

**Why quantify mitotype ratios rather than simply homoplasmic/heteroplasmic?**

When considering heteroplasmy, additional consideration is required when using quantitative methods because the ratios of mitotypes can vary on an almost continuous scale. As a result, determining a level at which a plant is heteroplasmic is a difficult choice. While it is easy to say a plant with a 50/50 mix of two divergent mitotypes is heteroplasmic, determining the level at which heteroplasmy becomes biologically relevant is more difficult. Pearl et al. (2009) make the case that when a haplotype comprises less than 0.05% of the total copies, it may be biologically irrelevant. They consider any occurrence below that level to be homoplasmic and any individual having a value above this cutoff would be considered heteroplasmic. However, there is no direct evidence that this level of occurrence has any biological significance.

Using quantitative PCR (qPCR) to determine the ratios of mitotypes present in a single individual has some distinct advantages over an approach that merely
identifies the presence/absence of heteroplasmy. First, it enables me to quantify the difference between parent and offspring heteroplasmy levels (ratio of one mitotype to another), and allows me to determine if heteroplasmy levels change between generations rather than simply saying that heteroplasmy is or isn’t present in both generations. This is an important step if we hope to understand heteroplasmy at a detailed level, particularly if we want to determine the potential effects of heteroplasmy. While there is currently little direct evidence that heteroplasmy has any effect on individuals, there may be a direct relationship between the occurrence of a mitotype and the expression of a phenotypic effect if both mitotypes are expressed in an individual.

Using quantitative methods also allow me to determine potentially differing levels of heteroplasmy in different tissues (vegetative sorting). Sorting could influence how the different mitotypes are passed to subsequent generations. For example, stochastic processes within a developing plant could by chance result in only one mitotype being present in reproductive tissues. If the other mitotype were not found in the reproductive tissues, this would prevent it (and potentially prevent heteroplasmy) from being passed to the next generation.

Why *Mimulus guttatus* is an ideal study organism

*Mimulus guttatus* (Phrymaceae) is an excellent study organism for many types of ecological and evolutionary studies. There has been extensive research done in the past on this and closely related species so there is ample data available
on a large variety of topics for reference (as reviewed by Wu et al. 2008).

Additionally, there are extensive molecular resources and data available for this plant, and the whole genome has been sequenced and annotated recently (J.H. Willis, unpub. data). There is a long history of research on populations of this plant in the Oregon Cascades; this is the area where heteroplasmy was first identified (Case & Willis 2008), and extensive seed collections of plants from the entire species range are available. Studies by Fishman and Willis (2006) and Case and Willis (2008) identified a cytonuclear incompatibility in crosses between *M. guttatus* and a closely related species *Mimulus nasutus*. This cytonuclear incompatibility was expressed as complete male (pollen) sterility. One of these studies identified two unique mitochondrial genomes (mtF and mtS); the mtS identified genome is associated with the CMS gene expressed in the crosses between those two plants. Additionally, this plant is relatively easy to grow in controlled conditions, and has a relatively short life span making it ideal for crossing studies. These resources make this an excellent plant model system in which to answer many types of questions ecological and evolutionary questions, including the ones asked in the following studies.

**What contributions will this thesis make to science?**

The findings in this thesis will contribute data to both the *Mimulus* research community and to the literature on heteroplasmy. This thesis contributes the first quantification of mitochondrial heteroplasmy in *M. guttatus*. This is an important
addition to the body of research of cytonuclear interactions in *M. guttatus*, including a known cytonuclear incompatibility resulting in CMS in crosses between *M. guttatus* and another closely related species *Mimulus nasutus* (Fishman & Willis 2006, Case & Willis 2008). This thesis will add to the already extensive body of work concerning *Mimulus guttatus*, and further its usefulness as a model study organism. Additionally this thesis has broader ranging consequences. It stands as one of the first series of studies to investigate the variation of mitotypes within both plants and populations with the intent of linking the two. It is also the first study to try to link measurable phenotypic traits (a proxy for fitness) to the incidence of heteroplasmy.
CHAPTER 2:

PATTERNS OF HETEROPLASMY IN *M. guttatus*

Introduction

Mitochondrial heteroplasmy is much more common than previously thought. Several studies have documented the occurrence of heteroplasmy in natural populations (Coyer et al. 2004, Welch et al. 2006) and compared mitotype ratios in parents and offspring (McCauley et al. 2005, Hoarau et al. 2009). But so far, no study has explicitly documented the variation of mitotype ratios within plants. In order to thoroughly study heteroplasmy as it occurs naturally, it is first necessary to document its variation within and between populations, within plants, and between parents and offspring. Documenting variation should ideally be done using natural populations and open-pollinated offspring. This chapter will explore the issues discussed above, and document naturally occurring heteroplasmy in *Mimulus guttatus*. The data obtained by these studies will allow for more comprehensive testing of how heteroplasmy arises, how it is lost, and its potential effects on individual fitness.

Within plant variation is of particular interest, as two studies have hypothesized that vegetative sorting of mitotypes contributes to the loss of heteroplasmy between generations (Andersson 1999, Pearl et al. 2009, Bentley et al. 2005).
Vegetative sorting is due to stochastic processes that occur within a plant during development. As cells divide during development, the nuclear genome is divided equally into the daughter cells, but the organellar genome is split into the daughter cells at random (Sheahan et al. 2007). If more than one mitotype is present in a cell they may be divided into the daughter cells in uneven ratios. This could result in vegetative sorting of mitotypes in different tissues throughout the plant. This stochastic process may result in a loss of heteroplasmy between generations if the heteroplasmic cells do not end up in reproductive tissue. However this may also be a means by which rare mitotypes, or more than one mitotype (heteroplasmy), could be transmitted maternally to offspring if by chance they end up in reproductive tissue and are passed to the next generation.

For the studies in this chapter, I used quantitative PCR (qPCR) to document the presence and variation of mitochondrial heteroplasmy in seven populations of *M. guttatus* from the Oregon Cascades. The primers I used for this study amplify unique genetic sequences found in two divergent mitochondrial types found in natural populations of *M. guttatus*. I also performed a sensitivity analysis to determine how precise qPCR is when being used to determining mitotype ratios in *M. guttatus*. I then used the same qPCR methods to examine the open-pollinated offspring from two populations for evidence of heteroplasmy to determine the rate of loss or gain of heteroplasmy. Finally, I compared qPCR data from three different tissue types from these plants to look for evidence of compartmentalization (vegetative sorting) of divergent mitotypes.
Materials and Methods

Plants used in study and study populations

The common yellow monkeyflower, *Mimulus guttatus* is a relatively short-lived, highly self-compatible wildflower native to the western coast of North America (Vickery 1978). Although I have collected samples from 17 total populations (Table 2.1), these studies focus on eight populations of *M. guttatus* located in the Oregon Cascades (Figure 2.1). This was the area in which the mtS mitotype was first reported and to which it seems to be restricted (Case & Willis 2008). The eight populations used for this study were populations that were previously identified by researchers from Duke University and elsewhere (complete list available at http://openwetware.org/wiki/Mimus_Community#Seed_Collections). The other nine populations were previously unreported populations that I found in spring 2009.

In addition to plants from natural populations, I used two homoplasmic inbred lines that were fixed for the two opposing mitotypes as controls in all of my studies. The IM62 inbred lined is derived from a population of *M. guttatus* found at Iron Mountain (IM), Oregon, and fixed for the S mitotype (hereafter mtS). The second is derived from a closely related species, *Mimulus nasutus*, from Shearer's Falls, Oregon (SF), an obligate self-pollinating species in which the F mitotype was originally characterized. This (mtF) is also the mitotype found in most other *M.*
*guttatus* plants (see Case & Willis 2008). The markers used in this thesis distinguish between the two unique mitotypes (mtS and mtF), therefore tissue from both inbred lines were included in all qPCR runs, as positive controls for their own mitotype and negative controls for the alternate mitotype. Any qPCR run showing amplification in the negative controls was repeated to ensure samples were not contaminated. DNA from these inbred lines was also used in a serial dilution experiment to determine the sensitivity of the qPCR methods used to estimate mitotype ratios.

**Primer Design, PCR Screens, and Quantification of Heteroplasmy**

I used both standard PCR and qPCR to first screen for, and then quantify, mitochondrial heteroplasmy. Heteroplasmy in *M. guttatus* was first identified with standard PCR and the primers (commonR and mtS & mtF standard PCR) used by Case and Willis (2008). The primers used for the qPCR analysis were designed using Primer3 (Rozen & Skaletsky 2000) and were based on sequences obtained by Case and Willis (2008) to amplify sequences 255 bp (mtS) and 237 bp (mtF) in length. This marker (with both types of primers) spans a physical rearrangement point between the two mitotypes, allowing my primer pairs to share a reverse primer (Figure 2.2 and Table 2.1). This means that I was unable to assay relative mitotype abundances in duplexed reactions with labeled primers, as in some other studies (Welch et al. 2006, Pearl et al. 2009). For all tissue collected I extracted total genomic DNA using a modified CTAB extraction method (http://openwetware.org/wiki/Mimulus:DNA_Extraction_Protocol).
Figure 2.1: Location of *Mimulus guttatus* populations in Oregon Cascades (Oregon, USA) screened for evidence of heteroplasmy

FV-Fernview; IM-Iron Mountain; CP-Cone Peak; HAC-Hackleman; BR-Browder Ridge; BLU-Blue River Reservoir; CGR-Cougar Reservoir; CGRE-Cougar Reservoir East
Samples were eluted in 100 μL of water, and two separate 10 μL aliquots were taken and qPCR was used, as described below, to analyze tissue samples for evidence of mitochondrial heteroplasmy. The standard PCR primers used amplify a 916 bp sequence (mtS) and a 626 bp sequence (mtF). Each 10μL reaction contained 1μL DNA template, 1μL of each 1μM primer, 0.8μL of 2.5μM dNTPs, 0.1μL Taq Polymerase, 2μL 5x Taq buffer, 0.8μL of 25μM MgCl₂, and 3.3μL of purified water. I amplified DNA fragments using BioRad iCycler thermal cyclers with denaturation at 94° C for 1 minute; followed by 5 cycles with denaturation at 94°C for 15 seconds denaturation, annealing at 72° C (decreasing by 1° C every cycle) for 30 seconds and elongation at 72° C for 2 minutes 30 seconds, this was followed by 24 cycles of denaturation at 94° C for 15 seconds, annealing at 67° for 30 seconds, and elongation at 67° C for 2 minutes 30 seconds, and lastly elongation at 72° C for 7 minutes.

Plants from natural populations were first screened with standard PCR to look for evidence of heteroplasmy. Following PCR amplification, I ran the samples on 1% agarose gels and examined the results visually with a UV transilluminator to look for amplification of the two mitotypes. I compared the size of the amplicons to a mid-range DNA ladder (Fermentas FastRuler™ Middle Range DNA Ladder, Glen Burnie, MD), to ensure that they were the expected sizes.
Figure 2.2: Location of primers amplifying alternate *M. guttatus* mitochondrial genomes

This figure shows the location of primers upstream from mitochondrial NADH-dehydrogenase subunit-6 (*nad6*) gene. A physical rearrangement (yellow triangle) differentiates two mitotypes (mtS & mtF), permitting amplification of mt-specific fragments as markers. Shown are locations for both standard primers and qPCR primers used in this study, and an additional control primer (ctrl) that amplifies a region of shared sequence in both mitotypes. Modified from Case & Willis 2008.
Table 2.1: Primers sequences used in both the standard and qPCR to identify alternate mitotypes in *M. guttatus*

The common reverse primer for both sets of primers is anchored in an area of shared sequence in both mtF and mtS. The primers amplify sequences that identify unique rearrangements upstream from the conserved mitochondrial gene, nad6 (Case & Willis 2008). The mtF primers were designed based on *M. nasutus* and the mtS primers were designed based on the IM62 inbred line of *M. guttatus.*

<table>
<thead>
<tr>
<th>Primer Sequence</th>
<th>Mitotype Amplified</th>
<th>PCR type</th>
<th>qPCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’ TCATGAATACAGATTCCCTCCCTTCTCG 3’</td>
<td>mtS</td>
<td>FWD</td>
<td>X</td>
</tr>
<tr>
<td>5’ CAGAGGAGATCAGATACCGGAAACGAA 3’</td>
<td>mtF</td>
<td>FWD</td>
<td>X</td>
</tr>
<tr>
<td>5’ TATACTCGGCTGAGGATAGCCTTACG 3’</td>
<td>mtS</td>
<td>REV</td>
<td>X</td>
</tr>
<tr>
<td>5’ GAGGAGCTGCTGAATTTTCTCGATTT 3’</td>
<td>mtF</td>
<td>FWD</td>
<td>X</td>
</tr>
<tr>
<td>5’ CACGAAATAAGGAAAGGAAAGGCAAC 3’</td>
<td>mtF</td>
<td>REV</td>
<td>X</td>
</tr>
</tbody>
</table>
When I saw amplification of both primer pairs in the same individual, they were judged to be heteroplasmic, and while this is not the most sensitive method available, it does allow me to efficiently identify individuals with relatively high degrees of heteroplasmy in natural populations (Figure 2.3).

While standard PCR is useful for screening large numbers of individuals for heteroplasmy, it doesn’t allow me to determine the exact ratios of mitotypes present. I used qPCR to quantify the level of heteroplasmy of individuals, allowing me to accurately track the change in mitotype ratios between parents and offspring and even within individual plants.

To ensure that the standard PCR and qPCR primers worked equally well to detect the two mitotypes, I first used standard PCR to amplify total genomic DNA from five separate individuals. Three of the plants had varying levels of heteroplasmy and two inbred lines were homoplasmic for the different mitotypes. I then used qPCR to assess the actual proportion mtS present in the individuals (Figure 2.3). This test showed that the qPCR primers function as expected, and that the standard PCR primers are capable of detecting the presence or absence of heteroplasmy. Therefore standard PCR represents an effective way to screen populations for heteroplasmy in an efficient manner. The results of qPCR match the findings of standard PCR, but take it one step further allowing me to calculate the ratio of the two mitotypes for quantitative analysis.
Figure 2.3: Sensitivity of standard PCR for detecting mitochondrial heteroplasmy as compared to qPCR methods

For each of 5 plants, I compared visible banding patterns from standard PCR to the proportion mtS based on quantitative PCR of the *nad6* region of the mitochondrial genome (see Fig. 2.2). I chose three heteroplasmic plants from natural populations that varied in their degree of heteroplasmy, and two inbred lines (IM62 and SF) each fixed for alternate mitotypes (mtS and mtF, respectively). While standard PCR works well to pick up a bias toward one mitotype or the other, qPCR is a much more precise, and accurate way to measure mitotype ratios.
For all studies in this thesis where a ratio of two mitotypes is reported, I ran separate, duplicated reactions for each primer pair from the two separate 10 μL aliquots takes from the original 100μL template. This was done to ensure the accuracy of my measurements as a representative of the original sample. This means that I ran a total of four qPCR reactions per DNA sample (two with each primer pair). Duplicating each qPCR run and using the average value from all runs helps to account for variation between subsamples due to pipetting error. Each 13μL qPCR reaction contained 2.5μL of DNA template, 1μL of each 1μM non-labeled primer, 2.25μL purified water, and 6.25μL of master mix (PerfeCTa™ SYBR® Green SuperMix, Low ROX, Quanta Biosciences, Gaithersburg, MD). The DNA template was amplified on a MX3005P® QPCR System (Stratagene, La Jolla, CA) with denaturation at 94°C for one minute, followed by 40 cycles of denaturation at 94°C for 15 seconds, annealing at 67°C for 30 seconds, elongation at 72°C for 2 minutes 30 seconds, and elongation at 77°C for 30 seconds; then denaturation for 1 minute at 94°C; followed by 45 cycles of 67°C for 33 seconds with the temperature rising 0.5°C every cycle to form a dissociation curve.

The output from each run allows me to quantify the total copies, and the mitotype ratios present in the samples. The number of copies present in the sample is calculated by comparing them to a serial dilution. The serial dilution was created with standard PCR product amplified from two homoplasmic, inbred lines (IM62 & SF). I purified the dsPCR product from standard PCR reactions, and used a Pico Green Assay (Invitrogen, Carlsbad, CA) to quantify each sample on a Synergy 2
microplate reader (Biotek, Winooski, VT) as per the manufacturers instructions. I then created a serial dilution with the dsPCR product (10^{-1} to 10^{-12}); this dilution was included in duplicate with each qPCR run. I also included positive and negative controls containing DNA from the two inbred lines in each run, and one non-template negative control (NTC) per 7 samples analyzed, one in each 8-tube strip. These extensive controls allow me an additional way to detect contamination and reduce false-positive results. Any qPCR run showing indication of contamination was re-run.

Calculation of mitotype copy number and mitotype ratios

To quantify the amount of each mitotype present in the unknown samples, I compared the relative concentration of each sample to a serial dilution between 2.59 \times 10^8 and 2.59 \times 10^{-4} \text{ copies/µl (mtS)}, and 2.52 \times 10^8 and 2.52 \times 10^{-4} \text{ copies/µL (mtF)}. I calculated the actual copy number/µL present in each sample using standard-curve (absolute) quantification. The copy number is calculated by comparing the number of amplification cycles between the start of the qPCR run and when an unknown sample crosses the threshold of detection (Ct) to where the known dilutions cross that threshold. I then use the molecular weight of the respective amplicons, and Avagadro’s number to calculate the number of mitochondrial copies present in the original sample. I also calculated the efficiency of each reaction by calculating how quickly the DNA was replicated for the first four sensor readings after crossing the threshold of detection (both methods are
described in Kontanis & Reed (2006). For ease of interpretation, I report the results as proportion mtS, based on the ratio of mtS copies to total (mtS + mtF) copies/µl of each mitotype. A value of 1 represents homoplasmy for mtS, and 0 homoplasmy for mtF. With each qPCR run, I compared the mitotype ratios between replicates to ensure that the sub-samples were similar to each other between runs (linear regression for all sub-samples from screening of maternal parents, \( R^2 = 0.95, P < 0.0001 \)) by calculating the pairwise difference in mtS ratio between the subsamples. Proportion mtS values between duplicate runs for all templates (n=963) differed by an average of 0.3 percentage points; most pairs of subsamples differed by less than 0.1 percentage points, but 72 samples had mtS ratios that differed by 1-9 percentage points between runs. Subsamples that differed from each other by >10 percentage points between runs, or if only one of the two subsamples crossed the threshold of being classified as heteroplasmic, the original template was run a third time (n=46). In all analyses, I use the average of all subsamples as the proportion of mitotypes present in each sample to determine if a sample is heteroplasmic.

Quantitative PCR Sensitivity Analysis

I used the 0.005 minority mitotype proportion (i.e., 0.5% or more of the total copies detected are different mitotype) as a threshold between functionally homoplasmic and heteroplasmic states (see Chapter 1 and Pearl et al. 2009). Therefore, I needed to confirm that my qPCR methods could distinguish samples above and below that threshold. As an explicit sensitivity analysis, I isolated DNA
from bud tissue from homoplasmic IM62 and SF plants and quantified the total genomic DNA present (in ng/µL) in each of the original samples using a PicoGreen Assay. I created a series of mixtures of total genomic DNAs in the following ratios (ng IM62 DNA: ng SF DNA): 0:1, 0.05:99.5, 1:99, 1:9, 1:3, 1:1, 1:3 1:9, 1:19, 1:1, 3:1, 9:1, 19:1, 99:1, 99.5: 0.05, 1:0. I assumed the amount of mitochondrial DNA present in the samples would scale with total genomic DNA, therefore allowing me to make mixtures of the two mitotypes based purely on the amount of total genomic DNA. Using each of these mixtures as a DNA template for qPCR, I expected the proportion of mtS to scale perfectly with the proportion of IM62 DNA in each mixture, indicating that my qPCR approach was sensitive enough to quantify heteroplasmy in unknown samples at a threshold of 0.005.

**Heteroplasmy in Natural Populations**

To determine levels of heteroplasmy present in natural populations, I collected leaf tissue samples and open-pollinated seeds from mature plants in the field. I started with samples collected by a colleague at Duke University in June 2005 from three populations in the Oregon Cascades (IM, BR, CP; Figure 2.1). In June 2008, I collected and screened leaf tissue samples from the same three populations plus five additional populations. Seeds were placed in petri dishes between two sheets of dampened filter paper and cold stratified at 4°C for 7 days to break seed dormancy. The seeds were left in the petri dishes on a 16/8h dark/light cycle at 24°C until germination, and were then transplanted into plug trays (Hummert
International, Earth City, MO), filled with Fafard 4P potting soil (Fafard Soils, BFG Supply, Burton, OH). Once plants had 4 true leaves, they were transplanted into 3” square pots in Fafard 4P potting soil and placed in the greenhouse with the same 16/8-hour light cycles. After the plants had produced one flower I collected leaf tissue from each plant, and the samples were frozen at -80°C until DNA could be extracted and analyzed with qPCR.

**Vegetative Sorting**

To look for evidence of vegetative sorting, I collected two flower buds and one leaf from 85 open-pollinated individuals from the two populations (BR & IM) known to contain both haplotypes. The seeds from 2005 were sown in February 2008 in the greenhouse as described above. I collected tissue, extracted DNA and performed qPCR as described above and then compared the heteroplasmy levels of the three tissue types.

**Heteroplasmy in seed offspring of open-pollinated plants**

To examine the relationships between offspring and parents, I sowed (February 2008) open-pollinated seeds from 46 maternal lineages and two populations (IM & BR) in which heteroplasmy had been previously identified. I collected tissue from the offspring (n=157), extracted DNA and performed qPCR as described above and then compared the heteroplasmy levels of the maternal parents and open-pollinated offspring. Loss of heteroplasmy would indicate that
some mechanism or stochastic process is at work reducing the transmission of one mitotype over another. However, if heteroplasmy is maintained between generations (passed to offspring), it may mean that mechanisms enforcing strict maternal inheritance cannot recognize and prevent heteroplasmy as long as both mitotypes are inherited maternally. For example, paternal mitochondria may be actively targeted for destruction as in ubiquitin tagging in mammals (Sutovsky et al. 2004) or otherwise excluded from offspring. Evidence of heteroplasmy in offspring from mothers determined to be homoplasmic by screening leaf tissue could be a result of two separate processes, either paternal leakage of a different mitotype during pollination, or stochastic processes in the mother.

**Statistical Analysis**

*Sensitivity of qPCR for detecting heteroplasmy.* To test the effectiveness of the qPCR methods in the sensitivity analysis, I used linear regression to compare the proportion mtS from qPCR (mitochondrial DNA) using the mixed dilution series to ratios of total genomic DNA. All tests were carried out in JMP version 8.0.1 (SAS Institute Inc, 2009).

*Among-population variation in heteroplasmy.* To test whether natural populations differed in the level of heteroplasmy of individuals, I used Levene’s Test (SAS Institute Inc, 2009) to compare variances in proportion mtS within populations. I then used Welch’s Test (Zar 1999) to examine differences in mean proportion mtS among populations.
Within-plant variation in heteroplasmy. To check for evidence of vegetative sorting of mitotypes within plants, I calculated the average proportion mtS present in leaf and bud tissue samples, and split the plants into three categories (homoplastic mtS, homoplastic mtF, and heteroplastic) based on the mitotype present in the leaf tissue samples.

I calculated the pairwise differences between the leaf proportion mtS and each bud proportion mtS value, then used the absolute values to calculate the average of these two differences for each plant as an index of within-plant variation. I used ANOVA to test whether this index of within-plant variation differed among plants with homoplastic mtS, heteroplastic, and homoplastic mtF leaves.

Calculating prediction intervals. I used the amount of within-plant mitotype variation based on the leaf and bud samples from individual plants to calculate the expected variation between mothers and offspring in the open-pollinated maternal families that would be consistent with the effects of vegetative sorting in the mothers. Values falling outside these calculated prediction intervals would be attributed to other processes, such as paternal leakage.

For each of the three leaf-mitotype classes (homoplastic mtF, heteroplastic, homoplastic mtS), I calculated 99% confidence intervals around mean of the within-plant variation indices (above) for each plant. Prediction intervals are based on these 99% confidence intervals, and were calculated separately for each heteroplasmy class. Because I used absolute values to calculate the means of the within-plant variation data, I applied one-sided upper prediction intervals to the
absolute differences between mothers and offspring in the open-pollinated families. This facilitates evaluation of whether the deviation between mothers and offspring exceeds expected in magnitude, regardless of the direction of the difference.

Prediction intervals define the range of expected values for a given number of future samples ($m$) based the amount of variation in a number of observed samples ($n$; SAS Institute, 2009). The upper limit of the one-sided prediction interval ($Y_u$) is calculated from the mean of the observed data ($X$), and the confidence interval around that mean ($t$), taking into account the number of both future ($m$) and observed samples ($n$), and the estimated standard deviation ($s$) of the observed samples:

$$[Y_L, Y_u] = \overline{X} \pm t(1-\alpha; n-1) \times \sqrt{\frac{1}{m} + \frac{1}{n} \times s}$$

Because I used $1-\alpha$ of 0.99, I was 99% confident that each absolute difference in proportion mtS between mothers and offspring in the open-pollinated families would be less than the upper-limit of the calculated prediction interval. Prediction intervals were calculated in JMP for each of the leaf-mitotype classes, specifying $m$ based on the total number of offspring produced by all of the mothers in each leaf-mitotype class.
Results

Sensitivity Analysis

The sensitivity analysis resulted in a significant linear relationship (Figure 2.4) between the amount of mitochondrial DNA detected in qPCR and mixtures of total genomic DNA. The slope of the best-fit line was not significantly different from unity (slope=0.97±0.07), but the y-intercept was positive (b =0.05; P>0.10). The observed proportions mtS from qPCR were greater than expected in samples with more even mixtures of genomic DNA, particularly in samples where the minority genotype comprised at least 10% of the total genomic DNA (midrange values in Figure 2.4). As the DNA mixtures approached pure samples of either genotype, the observed proportions mtS closely matched the expected values. This result indicates that my qPCR methods are capable of distinguishing variable mitotype ratios, and are particularly accurate near the 0.005 threshold for detecting heteroplasmy.

Evidence of heteroplasmy in natural populations

The frequency and occurrence of heteroplasmic individuals varied among populations. Heteroplasmic plants were found in five of the eight study populations (Figure 2.5), and all populations also contained homoplasmic individuals. Variance in individual mitotype ratios differed significantly among populations (Levene’s test F-Ratio=6.28, DF=7, P<0.0001), as did the mean proportion mtS (Welch’s test F-
Ratio=4741 DF=7, P<0.0001). One population (HAC) contained plants homoplasmic for both mitotypes, but no heteroplasmic individuals. Two populations contained only individuals homoplasmic for mtF (CGR & CP), and in all but two populations (IM & BR), heteroplasmy was very rare. The two sites where heteroplasmy was most common were dominated by alternate mitotypes; IM contained only homoplasmic mtS and mtS-biased individuals, and BR contained only homoplasmic mtF and mtF-biased individuals. There did not appear to be any geographic structure to the mitotypes present. In fact the two closest populations (IM & CP) had completely different dominant mitotypes. Overall, mtF appears to be the most common mitotype in these populations of *M. guttatus* (Table 2.2).

**Vegetative Sorting**

Of the 82 individuals sampled, 19 plants showed evidence of vegetative sorting (Figure 2.6). Only two plants (from the same maternal family) had all three tissues that were heteroplasmic. Of the remaining 17 plants showing evidence of sorting, 14 had heteroplasmic floral buds, and homoplasmic leaves. Two of the remaining three plants had heteroplasmic leaves and homoplasmic flowers, and the third had a heteroplasmic flower a heteroplasmic leaf and a homoplasmic flower. I found that within-plant mitotype variation differed significantly by leaf-mitotype category (ANOVA F-ratio=18.9, DF=2, P<0.0001; see Table 2.3). Plants with homoplasmic leaves are more uniform in their distribution of mitotypes than plants with heteroplasmic leaves.
Figure 2.4: Sensitivity of qPCR for detecting quantitative levels of mitochondrial heteroplasmy.

Total genomic DNA from two homoplasmic inbred lines (M. guttatus IM62 and M. nasutus SF) were mixed in symmetrical proportions (x-axis): 0, 0.001, 0.01, 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95, 0.99, 0.999, 1. These mixtures were used as templates for qPCR (y-axis) to assess the accuracy of our approach. The gray line represents a 1:1 line, and the red line is a best-fit line ($R^2=0.971$; slope=0.978).
Figure 2.5: Comparison of the occurrence of mitochondrial heteroplasmy in natural populations of *M. guttatus*

Heteroplasmy in natural populations based on qPCR analyses of field collected tissue samples, and open-pollinated seed samples. The circles represent heteroplasmic individuals (proportion of their minority mitotype ≥0.005), and the short horizontal markers are homoplasmic individuals. Populations are arranged geographically from northeast to southwest.
Table 2.2: Populations of *M. guttatus* sampled

Populations are ordered by mitotypes present, tissue type sampled and the date tissue was collected. The columns labeled PCR/qPCR indicate which method was used to determine the mitotype present. The coordinates are latitude and longitude for population locations in the Oregon Cascades.
<table>
<thead>
<tr>
<th>Population</th>
<th>mtF bias</th>
<th>mtS bias</th>
<th>Tissue/Date collected</th>
<th>PCR</th>
<th>qPCR</th>
<th>Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iron Mountain</td>
<td>X</td>
<td>X</td>
<td>Leaf tissue and Seeds 2005 &amp; 2008</td>
<td>X</td>
<td>X</td>
<td>122°8’34.80&quot; 44°24’15.12&quot;</td>
</tr>
<tr>
<td>Hacklemans</td>
<td>X</td>
<td>X</td>
<td>Seeds 2008</td>
<td>X</td>
<td>X</td>
<td>122°7’5.64&quot; 44°23’42.66&quot;</td>
</tr>
<tr>
<td>Cougar Reservoir East</td>
<td>X</td>
<td>X</td>
<td>Seeds 2008</td>
<td>X</td>
<td>X</td>
<td>122°13’21.36&quot; 44°5’32.88&quot;</td>
</tr>
<tr>
<td>Blue River</td>
<td>X</td>
<td>X</td>
<td>Seeds 2008</td>
<td>X</td>
<td>X</td>
<td>122°16’26.10&quot; 44°10’48.66&quot;</td>
</tr>
<tr>
<td>Browder Ridge</td>
<td>X</td>
<td>X</td>
<td>Leaf tissue and Seeds 2005 &amp; 2008</td>
<td>X</td>
<td>X</td>
<td>122°7’49.56&quot; 44°22’24.66&quot;</td>
</tr>
<tr>
<td>Fernview</td>
<td>X</td>
<td>X</td>
<td>Seeds 2008</td>
<td>X</td>
<td>X</td>
<td>122°18’8.58&quot; 44°24’14.46&quot;</td>
</tr>
<tr>
<td>Cone Peak</td>
<td>X</td>
<td></td>
<td>Leaf tissue and Seeds 2005 &amp; 2009</td>
<td>X</td>
<td>X</td>
<td>122°8’3.06&quot; 44°24’19.20&quot;</td>
</tr>
<tr>
<td>Cougar Reservoir</td>
<td>X</td>
<td></td>
<td>Seeds 2008</td>
<td>X</td>
<td>X</td>
<td>122°14’31.62&quot; 44°6’51.90&quot;</td>
</tr>
<tr>
<td>Browder Ridge Upper</td>
<td>X</td>
<td></td>
<td>Tissue 2009</td>
<td>X</td>
<td></td>
<td>122°7’56.04&quot; 44°22’51.00&quot;</td>
</tr>
<tr>
<td>Browder Ridge 3</td>
<td>X</td>
<td></td>
<td>Tissue 2009</td>
<td>X</td>
<td></td>
<td>122°7’41.58&quot; 44°23’3.06&quot;</td>
</tr>
<tr>
<td>Gate Creek Trail 1</td>
<td>X</td>
<td></td>
<td>Tissue 2009</td>
<td>X</td>
<td></td>
<td>122°5’9.00&quot; 44°21’56.94&quot;</td>
</tr>
<tr>
<td>Gate Creek Trail 2</td>
<td>X</td>
<td></td>
<td>Tissue 2009</td>
<td>X</td>
<td></td>
<td>122°6’54.24&quot; 44°22’41.16&quot;</td>
</tr>
<tr>
<td>Cone Peak 1</td>
<td>X</td>
<td></td>
<td>Tissue 2009</td>
<td>X</td>
<td></td>
<td>122°7’58.56&quot; 44°24’7.14&quot;</td>
</tr>
<tr>
<td>1598-1</td>
<td>X</td>
<td></td>
<td>Tissue 2009</td>
<td>X</td>
<td></td>
<td>122°4’19.56&quot; 44°21’15.54&quot;</td>
</tr>
<tr>
<td>1598-2</td>
<td>X</td>
<td></td>
<td>Tissue 2009</td>
<td>X</td>
<td></td>
<td>122°5’9.48&quot; 44°21’16.86&quot;</td>
</tr>
<tr>
<td>Trailbridge</td>
<td>X</td>
<td></td>
<td>Tissue 2009</td>
<td>X</td>
<td></td>
<td>122°3’3.66&quot; 44°16’53.04&quot;</td>
</tr>
</tbody>
</table>
Figure 2.6: Evidence of vegetative sorting of alternate mitotypes in *Mimulus guttatus*

The data represent 19 individuals (out of a total of 82) that showed evidence of vegetative sorting among three separate tissue samples. Individuals are ranked by mean proportion mtS among all three tissues (x-axis) based on qPCR assays (see Text for details). Tissue samples from each individual plant are stacked along the y-axis, connected by gray vertical dashed lines. The black symbols are floral bud samples and green symbols represent leaf tissue samples. The short horizontal lines indicate homoplasmic tissues and the open circles represent heteroplasmic tissues.
Table 2.3: Variation in heteroplasmy among leaf and bud tissues of individual *Mimulus guttatus* plants

The mitotype category represents the majority mitotype present in leaf tissue, and the mean pairwise differences are calculated for each plant as the absolute average difference between its leaf proportion mtS and the proportion mtS in each of two bud samples. The 99% confidence intervals are calculated based on the mean pairwise differences, and the far-right column indicates how many plants of each class were analyzed. Within-plant mitotype variation differs among mitotype categories. Plants with heteroplasmic leaves have much higher average within-plant mitotype variation than plants with homoplasmic leaves of either mitotype.

<table>
<thead>
<tr>
<th>Mitotype category based on leaf tissue proportion mtS</th>
<th>Mean pairwise difference in proportion mtS between leaves and each flower bud</th>
<th>99% Confidence Intervals</th>
<th>Plants sampled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homoplasmic mtF</td>
<td>0.002</td>
<td>-0.001</td>
<td>0.005</td>
</tr>
<tr>
<td>Heteroplasmic</td>
<td>0.107</td>
<td>-0.079</td>
<td>0.293</td>
</tr>
<tr>
<td>Homoplasmic mtS</td>
<td>0.008</td>
<td>0.002</td>
<td>0.014</td>
</tr>
</tbody>
</table>
Maternal vs. offspring heteroplasmy values

*Categorical analysis.* Of the 46 open-pollinated maternal families planted (proportion mtS determined by bud tissue), 20 mothers were homoplasmic for mtS, 4 mothers were homoplasmic for mtF, and 22 mothers were heteroplasmic. The number of offspring screened per maternal family ranged from 1 to 12 with a mean of 3.3 offspring per maternal family (total=154). The majority (92%) of all offspring from all groups were homoplasmic however, there were 3 heteroplasmic offspring produced by homoplasmic mothers (Figure 2.7). Heteroplasmic mothers produced eight heteroplasmic offspring; but the remainder of the offspring produced by these mothers were homoplasmic, containing only the mitotype that was more common in the mother.

*Quantitative analysis.* Based on the prediction intervals calculated for each class, none of the offspring from the 46 maternal families were different enough from their mother’s mtS ratio to invoke paternal leakage (Figure 2.7). In other words, regardless of the mother’s heteroplasmy class, the changes between the mother and offspring values could be attributed completely to within-plant vegetative sorting processes. Although offspring of heteroplasmic mothers were the most variable (Figure 2.7), the prediction interval for mothers with heteroplasmic leaves was broad enough to include all possible offspring values, meaning that we were not able to discriminate any evidence of paternal leakage when mothers were heteroplasmic.
**Figure 2.7: Comparison of maternal proportion mtS and open-pollinated offspring proportion mtS in *Mimulus guttatus* plants from the Oregon Cascades**

Mothers (total n=46) were classified as homoplasmic mtS, homoplasmic mtF, or heteroplasmic based on the 0.005 proportional threshold value for minority mitotypes. The maternal parents are ordered by their proportion mtS. Short horizontal bars represent homoplasmic offspring, and open circles represent heteroplasmic offspring (total offspring=154). The x-axis is not to scale, and is spread out to show each maternal family individually. Offspring displayed in black were produced by homoplasmic maternal parents, blue points are from mtF-biased heteroplasmic moms (0.005 < proportion mtS < 0.5), red points are from mtS-biased moms (0.5 < proportion mtS < 0.995). The dashed vertical lines delineate maternal values of interest. Some mothers had more even proportions mtS (0.1-0.9), however, with only one exception they produced either all homoplasmic offspring or that were offspring less heteroplasmic.
Discussion

The studies in this chapter have shown that heteroplasmy occurs in multiple populations of *M. guttatus*, the degree of heteroplasmy can vary drastically among individuals from the same population, and that heteroplasmy is sometimes passed to offspring but is more often lost between generations. It has also shown that mitotypes are found in different ratios within individual plants through a stochastic process known as vegetative sorting.

**Natural population patterns and potential mechanisms**

Recently there have been a number of studies focusing on heteroplasmy in flowering plants, particularly by those working with *Silene vulgaris* (Welch et al. 2006, Pearl et al. 2009, Bentley et al. 2010). These studies found mitochondrial heteroplasmy in multiple populations, however they also found that some populations showed no evidence of heteroplasmy while others had rather high rates of occurrence. In the evidence presented here, I show a similar pattern, where heteroplasmy is only found in some populations in the Oregon Cascades and that there are populations homoplasmic for both mitotypes.

An interesting pattern is evident in the natural population survey data: none of the populations that contained heteroplasmic plants contains homoplasmic plants of both mitotypes. Instead, the homoplasmic plants in these populations contained only one or the other mitotype; this is true even in populations where heteroplasmy
was quite common (IM population) and even biased in both directions (BR population). Stochastic processes may help to explain this phenomenon. When one mitotype is extremely common, stochastic processes are more likely to act on it by chance and keep it common within a plant and population, simply because it is more common. While these processes could also act to make the less common mitotype become the only one in a plant, the chances are small due to its rarity within the population in question. However, an occasional input of a differing mitotype via pollen flow may be enough to result in heteroplasmy that can be passed to subsequent generations. Occasional pollen flow could be enough to counteract stochastic processes that result in the loss of heteroplasmy, and result in detectable heteroplasmy during population screens. This does not explain why no populations were found with both heteroplasmic and the two opposing homoplasmic mitotypes. One possible explanation is that some populations may contain both types of homoplasmic individuals and that the number of plants I screened was too small to locate those plants. Because populations often contain many thousands of plants, I may be sampling what equates to less than 1% of the overall population so it is possible that my sampling methods are missing much of the variability that exists.

I also found one population (HAC) that was polymorphic and contained individuals homoplasmic for both mitotypes but no heteroplasmic individuals. This indicates that either there is a mechanism that enforces maternal inheritance in this population, or that there is assortative mating taking place, limiting the chance for detectable paternal leakage to occur. Assortative mating is very unlikely in this case
due to the proximity of plants to one another, and the sharing of pollinators, so the more plausible explanation is that there is a mechanism at work that enforces maternal inheritance, and therefore preventing heteroplasmy from occurring in this particular population. However, it is possible that this mechanism is equally as stringent in the other populations, and that leakage events occurred long ago and heteroplasmy has been maintained via maternal inheritance and vegetative sorting.

The results of the study examining open-pollinated families also highlight several important points. First, all of these plants exhibited the effects of vegetative sorting on the transmission of mitotypes to offspring. Stochastic within-plant variation may affect studies documenting patterns of heteroplasmy, and may form a large part of the processes that contribute to and maintain heteroplasmy in nature. More detailed information on the cellular mechanisms of vegetative sorting will be important for understanding its short-term and long-term consequences. Second, offspring of maternal seed families were consistently less heteroplasmic than their mothers, and less heteroplasmic than the field tissue samples collected from their home populations. This is a puzzling discrepancy, and it’s unclear whether the low levels of heteroplasmy in open-pollinated offspring has something to do with being grown in benign conditions or some other phenomenon that tends to increase the occurrence of heteroplasmy in field-grown adult plants. Nevertheless, heteroplasmy seems to be much more common in natural populations than in the seeds collected from those populations, a pattern that warrants further investigation.
Some studies have shown experimental evidence for paternal leakage and vegetative sorting. For example Pearl et al. (2009) found limited evidence for both paternal leakage and vegetative sorting, where paternal leakage is a way by which new heteroplasmy arises and vegetative sorting as the likely way by which heteroplasmy is lost between generations in *Silene vulgaris*. Hansen et al. (2007) found evidence for paternal and biparental transmission of the chloroplast genome in crosses between different *Passiflora spp*. Evidence of paternal leakage has also been found along hybrid zones in natural populations of *Fucus* a brown algae (Hoarau et al. 2009). It seems that paternal leakage in hybrids is more likely to occur regularly given the experimental evidence; this is probably due to incompatibilities in the mechanisms that prevent leakage in these matings. That paternal leakage occurs within species is particularly important because of the assumptions made in many types of research. Research into heteroplasmy also assumes that the values generated by qPCR methods are real, however there are several forms of error that must be accounted for.

**Identifying and accounting for error**

Possible sources of error in studies of heteroplasmy include the qPCR methods used to quantify mitotypes, and variability of mitotypes within plants (vegetative sorting). Each type presents unique challenges; the qPCR methods used in my study introduce potential error in several ways. Pipetting error is the most obvious, and in my case easiest to control. Pipetting error would be due to pipetting
by chance more or less of a given mitotype from a sample. I control for this by analyzing and comparing two subsamples for consistency in mitotype ratio. To keep the difference between runs from affecting my overall results I allowed no more than a 10 percentage-point difference between runs however most were far smaller than that. Using this cut off allowed me to ensure that my results were consistent and as accurate as possible. However, this is not the only type of potential error present in my studies.

Within-plant mitotype variation due to vegetative sorting can also be a form of error in studies of heteroplasmy that, depending on sampling strategies, could result in misclassified samples. While my research has shown no conclusive evidence for paternal leakage, I have found substantial evidence for vegetative sorting sometimes resulting in drastic differences in the ratio of mitotypes present in different tissue types and this must be taken into account when analyzing any data investigating heteroplasmy in *M. guttatus*. This might occur when the less common mitotype ends up in non-reproductive tissues, such as leaves. However, this may also be a mechanism by which heteroplasmy is passed to the next generation, even an extremely rare mitotype may be stochastically shunted into a floral bud and subsequently passed to that plant’s offspring.

**Accounting for error in controlled crosses**

While there are some ways to account for sorting when examining plants, such as extracting DNA from an entire plant, in *M. guttatus* distinguishing between
paternal leakage and vegetative sorting can be a difficult task. Sampling tissue for DNA extraction is by definition destructive, and this prevents me from determining an individual floral bud’s mitotype prior to using it for a controlled cross. One way to get around this problem would be to take several tissue samples from each individual used for crosses and check all of them for consistent heteroplasmy/homoplasmy values. Additionally, crosses could be made using multiple flowers from one plant as the mother and then using one flower from another plant as the father for all offspring. The offspring from all of these crosses could be screened (using multiple tissues) for heteroplasmy to determine whether they have similar mitotype values. Alternatively, the average rate of vegetative sorting for a parent (by mitotype) can be used as a proxy to determine how much of an mt ratio difference between the parent and offspring proportion mtS can be expected due to sorting. Any variation above those levels could indicate that paternal leakage has taken place.

Regardless of how heteroplasmy arises in *M. guttatus* I have shown that it does occur, and as a result may have potential consequences to heteroplasmic individuals. There are still several questions regarding heteroplasm remaining to be answered in *M. guttatus*. In Chapter 3, I test if heteroplasmy is generated via paternal leakage, and is there any negative effect to individual plants due to heteroplasmy.
Chapter 3:

CAUSES AND CONSEQUENCES OF HETEROPLASMY

Introduction

Not much is currently known about the consequences of heteroplasmy and the processes leading to, and maintaining, it between generations. Both are lines of study that should be undertaken in organisms found to contain heteroplasmy in order to understand how heteroplasmy occurs and its effects on individuals.

To my knowledge, no published experimental study has attempted to link fitness effects, negative or positive, to the occurrence of heteroplasmy. There are several reasons that plants could be negatively affected by having two separate mitochondrial genomes. First, there is extensive interaction between the mitochondrial and nuclear genomes. The nuclear genome regulates most of the transcription and translation of the organellar genomes, thus intergenomic compatibility and co-evolution are of utmost importance (Cosmides & Tooby 1981, Rand et al. 2004). If there were to be any incompatibility between the two genomes, it is very possible that there could be a substantial effect on mitochondrial function, and, in turn, plant fitness. Another possibility is that heteroplasmy could lead to recombination between two mitochondrial genomes. Recombination could disrupt mitochondrial function by creating new novel reading frames, and as a result, affect
plant development and fitness. Mitochondrial recombination, once thought to be unlikely, has been found to occur at least occasionally in several plant species (reviewed by White et al. 2008), leaving open the possibility that it could have effects on individual plant fitness. Whether or not there are measurable fitness effects, exploring heteroplasmy should also extend into how it is passed between generations, and how it arises to determine if non-maternal inheritance is the driving force behind heteroplasmy.

To date, only one study has directly tested the hypothesis that paternal leakage of mitochondria leads to heteroplasmy (Bentley et al. 2010). While this study found that paternal leakage does occur in *Silene vulgaris* resulting in heteroplasmy, it was unable to account for the additional effects of within-plant variability in the transmission of heteroplasmy. Stochastic processes leading to differing mitotype ratios in different tissues within a plant (vegetative sorting) must not be overlooked when considering how heteroplasmy arises and is transmitted. Though no direct evidence of vegetative sorting was shown, a study by Andersson (1999) suggested that the mitotypes present in different floral buds (potentially due to vegetative sorting) resulted in different sex ratios in the offspring of gynodioecious *Silene vulgaris*. Paternal leakage and stochastic processes may both be important to the generation and transmission of heteroplasmy in flowering plants and should both be considered when performing controlled crosses and considering the outcome of those crosses. The advantage of controlled crosses when compared to open-pollinated families is that the mitotype of both parents is known.
In the studies in this chapter, I used the primers and qPCR methodology discussed in Chapter 2 to test for potential negative effects on fitness due to heteroplasmy. I measured a number of potentially fitness-related phenotypic traits to compare to the mitotype ratios in the plants. I also performed a number of controlled crosses to look for evidence of paternal leakage resulting in heteroplasmy while attempting to take into account any potential stochastic processes taking place within individual plants.

**Material and Methods**

**Study species and populations**

The common yellow monkeyflower, *Mimulus guttatus* is a wildflower often found on hillside seeps, and road cuts on the western coast of North America. The studies in this chapter focus on eight populations of *M. guttatus* from the Oregon Cascades (Figure 2.1), all populations studied were previously identified by other researchers and listed on the *Mimulus* wiki website ([http://openwetware.org/wiki/Mimulus_Community#Seed_Collections](http://openwetware.org/wiki/Mimulus_Community#Seed_Collections)). The open-pollinated seeds used for these studies were either obtained from Duke University (collected in 2005) or myself (collected spring 2009). Tissue from these plants was analyzed with quantitative PCR (qPCR) to determine the proportion of mtS present. The methods of DNA extraction, qPCR analysis and the primers used to identify the two mitotypes (mtS & mtF) in this chapter are the same as described in detail
previously (see Chapter 2). However in this chapter I further identify the specific tissues used in each data set as a way to account for vegetative sorting.

**Plant fitness and heteroplasmy**

In order to determine any potential phenotypic effects of heteroplasmy, I compared the proportion mtS to a variety of vegetative and floral fitness-related traits (Hall and Willis 2006). I sowed (as previously described) up to 10 open-pollinated seeds from each of 147 maternal families from eight populations of *M. guttatus*, and recorded the time to germination, time to first true leaf, and time to first flower of the resulting 697 offspring (not all seeds germinated). Once the plants had produced their first flower, I measured and recorded the size of the largest cauline leaf (width and length), and overall plant height (soil level to shoot apex). I also measured 4 floral dimensions (see Figure 1 in Fishman et al. 2002) of the first two flowers produced (Table 3.2), and examined anther morphology to determine if anthers were fully developed (Fishman & Willis 2006, Case & Willis 2008). Malformed anthers containing no pollen have been associated with cytoplasmic male sterility (CMS) in this plant (Case & Willis 2008). The leaf and floral dimensions were measured with a digital caliper (World Precision Instrument Inc. Sarasota, FL), and the plant height was measured with a stick ruler. To screen plants for both pollen fertility and pollen quantity, I collected two pollen samples from each of the first two flowers produced for a total of four pollen samples per plant. I stained the pollen samples by placing the anthers from each flower into a separate
0.5 mL microcentrifuge tube with 60 mL of lactophenol-aniline blue pollen stain (Kearns & Inouye, 1993). I placed all samples in a Branson 1510 ultrasonic cleaner (Bransonic Danbury, CT) for 5 minutes and then vortexed each tube for 30 seconds to ensure dehiscence of the anthers and full exposure of the pollen to stain. I stored the samples at room temperature until they were counted.

To count the pollen grains, I placed 1 mL samples from each tube into the counting chambers of an Improved Neubaeur hemocytometer (Levy Ultra Plane, C.A. Hausser & Son Philadelphia, PA), counted all pollen grains present in the subsamples and scored them for fertility using a compound light microscope. Aniline blue stains the starches in the pollen grain (Kearns & Inouye 1993), so an individual pollen grain was classified as fertile if it had dark, even staining; pollen grains with either no stain uptake or abnormal staining patterns were scored as infertile. Two offspring from each of 4 to 15 maternal lineages per population (n=140 plants) were assessed for pollen quantity, pollen fertility and screened for heteroplasmy. In some cases (e.g., the CP population) only one offspring was available from each maternal lineage. To screen for heteroplasmy in these plants, I collected one leaf per plant on dry ice, and stored it at -80°C until DNA could be extracted. Using the previously described methodology (see Chapter 2), I extracted DNA from the leaf samples, and used qPCR to determine the proportion mtS present in the samples. The proportion mtS in leaves was then compared to the morphological measurements; if there was an effect of heteroplasmy on plants, I would expect the heteroplasmic and homoplasmic plants to differ in the phenotypic
variables. For example, heteroplasmic plants might be smaller, flower later, or have either less pollen or lower pollen fertility.

**Paternal leakage as assessed by controlled crosses**

To determine the potential for paternal leakage in *M. guttatus*, I selected plants with known mitotypes from the fitness study (described above) with which to make controlled crosses (Table 3.1). I first emasculated the flowers used for crosses prior to anthesis. Using forceps, I removed all anthers from each sire crushed them on a glass microscope slide to extract the pollen, then placed it onto the stigmatic surface of the maternal plant using toothpicks. *Mimulus guttatus* plants have sensitive stigmas that close when touched, and remain closed when saturated with pollen. After the crosses were made, the plants were left on the counter and re-examined after 2 hours to ensure pollen saturation.

The first set of crosses was made between plants from multiple populations known to be homoplasmic for the two opposing mitotypes. As a means of detecting potential cryptic heteroplasmy, I made a second set of crosses between individuals homoplasmic for the same mitotypes. The third set of crosses was between plants homoplasmic for one mitotype and individuals that were heteroplasmic with a strong bias towards the other mitotype. I collected mature fruits from the plants into labeled paper coin envelopes, allowed them to dry thoroughly and then stored the seeds in a refrigerator at 10°C.
Table 3.1: Crosses to assess evidence of paternal leakage and loss of heteroplasmacy in *Mimulus guttatus*

The table below shows the possible results of crosses between parents based on their mitochondrial haplotypes; n indicates the number of crosses of each type made for this study. Crosses with results in black text illustrate expected results of crosses to detect cryptic heteroplasmacy. **Red** text indicates results of strict maternal transmission, **blue** shows expected results of strict paternal transmission, and **green** is expected from bi-parental transmission of mitochondrial haplotypes. The unequal sample sizes for the each category of cross type are due to mortality following heat stress in the greenhouse prior to making the crosses.

<table>
<thead>
<tr>
<th>Mom Mitotype</th>
<th>Homoplasmic mtS</th>
<th>Dad Mitotype</th>
<th>Heteroplasmic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homoplasmic mtS</td>
<td>n=5 Homoplasmic mtS</td>
<td>n=13 Homoplasmic mtS</td>
<td>n=1 Homoplasmic mtS</td>
</tr>
<tr>
<td>Homoplasmic mtF</td>
<td>n=16 Homoplasmic mtF Homoplasmic mtS Heteroplasmic (F/S)</td>
<td>n=7 Homoplasmic mtF</td>
<td>n=2 Homoplasmic mtF</td>
</tr>
<tr>
<td>Heteroplasmic</td>
<td>n=2 Heteroplasmic Homoplasmic mtS Heteroplasmic (S&gt;F)</td>
<td>n=2 Heteroplasmic Homoplasmic mtF Heteroplasmic (S&lt;F)</td>
<td>n=0 Heteroplasmic S/F</td>
</tr>
</tbody>
</table>
I cold-treated, then sowed 15 seeds (as previously described), and collected tissue from all plants in each cross to examine the mitotypes of the offspring (sample sizes for each cross are given in Table 3.1).

Although parental mitotype ratios were assessed from leaf tissue, offspring ratios were determined by bulking multiple tissues into a single tube for analysis. This was done to account for any potential vegetative sorting in the offspring. I collected and bulked three tissue samples (one leaf sample and two buds) from each offspring into one tube. In cases where offspring did not flower (n=23), I collected three separate leaf samples. There was no obvious pattern in the occurrence of vegetative plants. Tissue samples were collected, DNA extracted, and qPCR performed as described previously (Chapter 2). I then compared the parental proportion mtS to the offspring proportion mtS to determine how the mitochondrial genome had been transmitted to the offspring in the different crosses.

**Statistical analysis**

*Fitness effects.* To test for possible fitness effects due to heteroplasmy, I first used principal component analysis to summarize the plant characteristics using sets of measurements: all plant size measures and all flower size measures. In both cases, the eigenvectors loaded positively and the first PC explained at least 80% of the variation in the data. I calculated the average pollen fertility, and pollen quantity for each plant separately from the PCA. I then used a repeated-measures multivariate analysis of variance (MANOVA) to test for an effect of heteroplasmy on
the plants. I did this by classifying plants as either homoplasmic or heteroplasmic. The repeated measures in this test were the four plant fitness traits and the independent variables were the plant’s heteroplasmic state and the population of origin. Previous studies have shown variation in these traits at the population level (Fishman & Willis 2002, Hall et al. 2006).

Evidence for paternal leakage in controlled crosses. To assess the evidence of paternal leakage in the controlled crosses, I first calculated the proportion mtS for each parent and all offspring. I then calculated the absolute pairwise difference in proportion mtS between mothers and offspring. I recalculated prediction intervals as described in Chapter 2, but used the total number of offspring in each maternal leaf-mitotype class as the number of future samples \((m)\). This prediction interval determines how many mother-offspring pairwise differences could be explained by vegetative sorting. Pairwise mother-offspring differences that exceeded the prediction intervals for each class (homoplasmic mtS, homoplasmic mtF, and heteroplasmic) were assumed to have resulted from paternal leakage (in crosses with alternate mitotypes) or cryptic heteroplasmy (in crosses with shared mitotypes).

Results

Fitness effects

The statistical results for fitness effects only take into account the sub-sample of plants analyzed with qPCR, but represent all eight populations used for
this experiment. I used qPCR to analyze 139 plants from the eight populations; the number of plants sampled from each population varied from 4 - 32, (mean=17). I found that there was a relationship between the population of origin and the principal component values I calculated, and pollen quantity and fertility (n=132, $F=10.3$, $DF=7$, $p<0.0001$). This indicates that there are significant phenotypic differences among populations. Further evidence of this is found in a significant interaction between fitness traits and population (n=132, $F=5.5$, $DF=21$, $P<0.0001$ and Table 3.2). However, I found no significant effect of mitotype and the fitness traits (n=132, $F=0.95$, $DF =1$, $p=0.3308$). This shows that there does not seem to be a fitness effect of heteroplasmy to plants in a controlled lab setting. However, it should be pointed out that this data set contains only eight heteroplasmic individuals compared to 131 homoplasmic individuals. Regardless, the fitness traits of heteroplasmic individuals did not differ from the homoplasmic values (Figure 3.1).
Table 3.2: Vegetative and floral size measurements of *M. guttatus* plants measured to look for possible fitness effects of heteroplasmy

The table below shows the mean and standard error of each trait measured by population. Sample sizes for each population are displayed below the population name. The results from the second flower measured closely matched the first, and so only the measurements from the first flower are displayed. Pollen results are the average of two counts from a single flower.

<table>
<thead>
<tr>
<th>Traits</th>
<th>Populations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BLU n=119</td>
</tr>
<tr>
<td>Plant height (mm)</td>
<td>64±1.7</td>
</tr>
<tr>
<td>Leaf width (mm)</td>
<td>17±0.4</td>
</tr>
<tr>
<td>Leaf length (mm)</td>
<td>22±0.5</td>
</tr>
<tr>
<td>Throat width (mm)</td>
<td>6±0.1</td>
</tr>
<tr>
<td>Corolla length (mm)</td>
<td>15±0.2</td>
</tr>
<tr>
<td>Upper petal width (mm)</td>
<td>7±0.1</td>
</tr>
<tr>
<td>Corolla width (mm)</td>
<td>12±0.2</td>
</tr>
<tr>
<td>Pollen quantity</td>
<td>183±18</td>
</tr>
<tr>
<td>Pollen fertility</td>
<td>77±6.7</td>
</tr>
</tbody>
</table>
Figure 3.1: Comparison of vegetative and floral measurements between heteroplasmic and homoplasmic *Mimulus guttatus* plants from 8 populations in the Oregon Cascades

The four panels below display a comparison of vegetative and floral characteristics of both heteroplasmic and homoplasmic *M. guttatus* plants. Homoplasmic plants were grouped regardless of mitotype. The y-axes in panel A (floral measurements) and panel B (vegetative measurements) show the values for the first principal component (PC1). In both cases, PC1 explained at least 80% of the variation in the data and all eigenvectors loaded positively. The y-axis in panel C is the average pollen quantity of two subsamples from each of two flowers per plant. The y-axis in Panel D displays the logit transformed average pollen fertility of the two flowers measured. As a result of this transformation a value of 4.59 is equal to 99% pollen fertility, and a value of -4.59 is equal to 1% pollen fertility. None of the traits means differed significantly between heteroplasmic and homoplasmic groups.
A

B

C

D

Floral PC1

Vegetative PC1

Pollen Quantity

Pollen Fertility
**Controlled crosses**

The unequal sample sizes among crosses (Table 3.1) reflect high mortality in the greenhouse during the parental generation. Just prior to conducting these crosses, the temperature in the greenhouse rose to a level that most plants did not survive. Although I saw no obvious trends among the plants that survived, this certainly affected the sample sizes for the crosses. I made all possible combinations of crosses with the plants that remained, resulting in the unequal sample sizes seen below.

*Crosses with homoplasmic mtF mothers.* I made a total of 22 crosses with maternal plants homoplasmic for mtF (Figure 3.2A). Seven of these crosses were to dads also homoplasmic for mtF; these crosses act as a control for cryptic heteroplasmy in this crossing study, and should produce only offspring homoplasmic for mtF. In one of these crosses (FV 5-1 x HAC 9-8), five of the eight offspring were heteroplasmic with a bias towards mtF. In two of these plants, the mom-offspring pairwise differences were greater than expected by the calculated prediction interval for homoplasmic mtF mothers (mom-offspring differences in proportion mtS = 0.07 and 0.08 vs. 0.04 prediction interval). Because both parents were determined to be homoplasmic mtF, these offspring should represent the error due to cryptic heteroplasmy in the study. However, these crosses were unusual in that the total copy number estimated for all of the offspring and for the maternal FV 5-1 parent were extremely low (n=8, mean=157 total copies) compared to plants in all other qPCR datasets (mean=2.6x10^7 total copies).
Three crosses were made between mtF moms and heteroplasmic mtS-biased dads; all 23 resulting offspring were homoplasmic for mtF. Finally, I made 12 crosses between mtF moms and mtS dads resulting in 81 total offspring. Sixty-six of these offspring were homoplasmic for mtF. The remaining 15 offspring (from five crosses) showed evidence of heteroplasmy; 11 of these offspring had a mom/offspring pairwise difference greater than the prediction interval, consistent with paternal leakage. However, seven of these plants were produced by the same mom (FV5-1) as the cross above having extremely low mitotype copy numbers.

Crosses with heteroplasmic mothers. Five crosses (Figure 3.2B) were made with heteroplasmic moms. One cross, made between a heteroplasmic mtF-biased mom and a heteroplasmic mtS-biased dad produced two homoplasmic mtF offspring. I also made four crosses between heteroplasmic mtS-biased moms, and homoplasmic-mtF dads. Of 25 offspring, 20 were homoplasmic for mtS, and five were heteroplasmic with mtS biases.

The prediction interval for heteroplasmic mothers was broad enough to include all possible values for offspring proportion mtS, meaning that I could not detect whether paternal leakage occurred in the offspring of any heteroplasmic mothers. In this case, I assumed that all mother-offspring differences could be explained by vegetative sorting.

Crosses with homoplasmic mtS mothers. I made 17 crosses using plants homoplasmic for mtS (Figure 3.2C) for a total of 119 offspring. Five of these crosses were to plants with the same mitotype (mtS), resulting in 33/36 offspring that were
homoplasmic for mtS. The remaining three offspring (from two separate crosses) were heteroplasmic with an mtS bias, however the mom/offspring pairwise ratios of these plants were within the calculated prediction interval (the maximum difference between mothers and offspring in proportion mtS should be 0.06). This suggests that these values are consistent with within-plant stochastic processes. I also made 11 crosses to plants that were homoplasmic for mtF; of 74 total offspring, only three (from three different crosses) were heteroplasmic. None of the three crosses had mom/offspring pairwise differences exceeding the prediction interval. Finally, I made one cross between an mtS mom and a heteroplasmic mtF-biased dad; all nine offspring resulting from this cross were homoplasmic for mtS.
Figure 3.2: Results of controlled crosses to detect paternal leakage in *Mimulus guttatus*

The three panels below display the results of controlled crosses to detect paternal leakage. Crosses were made with three categories of maternal plants: (A) homoplasmic mtF moms, (B) heteroplasmic moms, and (C) homoplasmic mtS moms. In each panel, the crosses are ordered by the paternal proportion mtS. The maternal parent in panel B is represented by an (X); in all three panels, the paternal parent is represented by (Y), and offspring are either a short horizontal line (if homoplasmic) or an open circle (if heteroplasmic). **Red** open circles are heteroplasmic offspring whose mom/offspring differences are greater than could be accounted for by stochastic processes (e.g., vegetative sorting) affecting mitotype ratios.
Discussion

The results in this chapter have shown that heteroplasmy had no effect on the fitness traits measured here, but that there is a difference in plant size and pollen quantity and quality among populations. The controlled crosses showed that paternal leakage may occasionally occur in *M. guttatus*, but that vegetative sorting regulates the transmission or loss of mitochondrial heteroplasmy between generations more often than paternal leakage.

**Heteroplasm does not affect fitness**

The data from the fitness study showed no obvious links between the fitness parameters I measured and a heteroplasmic state in *M. guttatus*. Although the sample size of heteroplasmic plants was extremely small, there were no apparent trends in the data; values for heteroplasmic plants were within 2 standard deviations of the mean measurements for their respective populations. It is however possible that an increased sample of heteroplasmic plants would show a larger amount of variance in the measurements of heteroplasmic plants.

A lack of fitness effects of heteroplasmy suggests that incompatibility between genomes is unlikely in *M. guttatus*. Additionally, both mtS and mtF are present in natural populations (see Chapter 2 and Case & Willis 2008. At least one population (HAC; see Chapter 2) has both mitotypes present in homoplasmic individuals with no noticeable differences between the plants. Also, the populations
considered here are somewhat close together, so frequent gene flow between populations could have exposed and selected out any possible incompatibilities between the two genomes.

Heteroplasmic plants are also found in several natural populations, and data collected in the fitness study suggests that would be expected to function normally. Thus, it may have been unlikely to find an effect of heteroplasmy on the plant traits measured, unless there were to be a recombination event that disrupted mitochondrial function. Any recombination event would be unlikely to occur, and would probably not cause significant changes to the results of a study of this nature, and future studies should explore this possibility. The significant effect of population on plant traits found in the fitness study is not unexpected. Several other studies in *M. guttatus* have measured phenotypic traits in different populations and found similar differences between those populations (Fishman & Willis 2002, Hall et al. 2006).

**Paternal leakage and stochastic processes contribute to heteroplasmy**

The controlled crosses performed in this chapter have shown that it is likely that both paternal leakage and stochastic processes are acting to maintain heteroplasmy in *M. guttatus*. In the previous chapter, I showed evidence for vegetative sorting. In the crossing study, I tried to take sorting into account in two ways. First, I collected three tissue samples from each offspring of the controlled crosses to try to account for possible sorting in those plants. Next, because the
parental proportion mtS was determined by leaf tissue, I calculated prediction intervals to determine how great of a change in potential offspring values could be explained by vegetative sorting.

There are several crosses to be noted when discussing stochastic processes. First, most of the heteroplasmic offspring (Figure 3.2) from the mtS x mtF and mtF x mtS crosses were similar enough to their mothers in mitotype frequency to have resulted from stochastic processes alone. Second, in the crosses between heteroplasmic moms and homoplasmic dads, five of the offspring from these crosses were heteroplasmic and the rest were homoplasmic. Because of the wide variation that can result from vegetative sorting in heteroplasmic parents, all of these offspring were also consistent with stochastic processes. Greater mother-offspring differences were required in order for me to attribute inheritance to paternal leakage events rather than within-plant stochastic processes.

Previously, only one other study (Bentley et al. 2010) has shown experimental evidence for paternal leakage of mitochondria in controlled crosses. Ten plants from the controlled crosses in this study showed evidence consistent with paternal leakage. In these crosses, proportion mtS in the offspring varied more than would have been expected after stochastic processes were taken into account. In one case (cross 13, Figure 3.2A), the extreme change in proportion mtS seen in the offspring is best explained by paternal leakage. This suggests that paternal leakage may be the process that leads to heteroplasmy in *M. guttatus.*
However, the extreme changes seen in one other cross (cross 11, Figure 3.2A) may be due to something else. The mother for this cross was the same as for cross 5 in the same panel (FV 5-1); note that the offspring from both of these crosses all have very low total copy numbers. The copy numbers of several of the plants collected from the original population have very low total copy numbers as well. When all of this information is considered, it suggests that there may be a third mitotype present in the plants from this population that represents a “null allele” with respect to the marker we are using. Interestingly, there were also two genotypes screened by Case and Willis (2008) where no amplification of either known mitotype was found; this also suggests the presence a third (null) mitotype in \textit{M. guttatus}.

\textbf{What is the next step?}

The studies in this chapter have shown that there appears to be no fitness effect of heteroplasmy in a lab setting, and that both paternal leakage and stochastic processes appear to be involved in the maintenance of heteroplasmy in \textit{M. guttatus}. However, they do raise some additional and interesting questions. First, while no obvious fitness effects were seen in the lab, it is possible that the selective pressures that act on heteroplasmic plants are mostly eliminated in the carefully controlled conditions present in the environmental chamber. Testing this would require a reciprocal transplant study where plants of each mitotype from each population would be placed at all other populations. I would then record plant measurements
as I did in the transplant study to test for effects of heteroplasmy. I would also collect and count seeds produced by the plants to look for differences in seed set due to heteroplasmy or population of origin. Second, are there more than two mitochondrial haplotypes in *M. guttatus*, and if so how widespread are they? Both my thesis and a study by Case and Willis (2008) have shown circumstantial evidence for possible additional mitotypes. The best way to screen for additional mitotypes would be to sequence whole mt genomes in plants containing known and unknown mitotypes. The genome sequences could then be used to screen for and identify mitotype specific markers. These markers could be used to screen plants in natural populations to confirm the presence of multiple mitotypes.
CHAPTER 4:

CONCLUSIONS

The findings in this thesis contribute new data to the *Mimulus* research community and to the literature investigating heteroplasmy. These contributions are discussed below. However, there are still many questions that remain to be answered, both in *Mimulus guttatus* and about heteroplasmy in general.

**What do we now know about heteroplasmy in *Mimulus guttatus***?

Mitochondrial heteroplasmy occurs in most, but not all, populations of *M. guttatus* in the Oregon Cascades. The occurrence of heteroplasmy in the populations sampled in this thesis show no evidence of geographic or spatial structure. I also found that even ratios of the two mitotypes does not often occur, that most plants that are heteroplasmic have values very close to the threshold between homoplasmy and heteroplasmy, and that mitotype ratios can vary substantially within plants.

I found evidence of vegetative sorting in multiple plants from the two populations I sampled (IM & BR). However, the majority of the plants (17 of 19) that showed evidence of sorting were from the IM population. This seems to suggest that
sorting may be more common in some populations than others, but this trend may be simply due to heteroplasm being more common in the IM population. Regardless, vegetative sorting has important implications for the study of heteroplasm in *M. guttatus* as it can make determining an actual mitotype ratio difficult. Though informative, the studies in this thesis leave several questions unanswered when considering heteroplasm in *M. guttatus*.

**What do we know about heteroplasm?**

One of the first points discussed in this thesis is whether heteroplasm is consistent with strict maternal inheritance. Although I found limited evidence of paternal leakage, I think that uniparental inheritance has been and will continue to be the most common form of organelle inheritance found in eukaryotic organisms. Several recent studies have looked for evidence of paternal leakage and heteroplasm in a variety of organisms (Shepherd et al. 2008, Wolff & Gemmell 2008, White & Martin 2009, Okada et al. 2010). None of these studies found evidence that paternal leakage resulting in heteroplasm had occurred. Additionally most of the cases of heteroplasm I found in controlled crosses could be explained by within-plant processes, indicating that once present, heteroplasm can be passed to offspring via maternal inheritance. This lends further evidence that maternal inheritance is the common form of inheritance.

However, there have been several studies that have shown evidence of paternal leakage resulting in heteroplasm in hybridization events between closely
related species (reviewed by Barr et al. 2005, Hansen et al. 2007, Hoarau et al. 2009). This seems to indicate that the mechanisms that prevent leakage may differ between species more than between populations of the same species. Additionally, neither recombination nor mutation is the likely explanation for how heteroplasmy arises; both of these mechanisms would result in one single mitochondrial genome being different than the rest in a given individual. While paternal leakage results in very few unique mitochondrial genomes being introduced, there is still a distinct numerical advantage. There are very few mitochondria present in a newly formed zygote compared to the number present in a grown individual. This single advantage leaves paternal leakage as the most likely explanation for the occurrence of heteroplasmy.

Regardless of where it occurs, or how often, heteroplasmy may have important consequences in both research and evolution. Whether or not heteroplasmy has consequences depends on the context in which it is investigated.

There are also several evolutionary implications related to the occurrence of heteroplasmy whether or not it occurs via non-maternal inheritance. Recombination can have either positive or negative effects on individuals depending on how it occurs. Recombination can occur between either two similar or two dissimilar organellar genomes, but either situation could have similar effects on individuals. Recombination has been suggested (Barr et al. 2005) to be a way by which organellar genome mutations could be purged. If this is the case, then selection could favor occasional paternal leakage if it benefits organisms. However,
recombination could have negative effects on individual fitness if it disrupts the tightly co-evolved relationship between the nuclear and cytoplasmic genomes.

On the other hand, if recombination doesn’t occur, and the less common mitotype is not expressed, it may have no measurable effect on individuals. However, studies using cytoplasmic markers, whether as a way to infer phylogeny, or in studies investigating relationships among populations, may see measurable differences in the results of the study if heteroplasmy is present. If more than one unique genome is present and multiple markers are used, those markers may either differ between the cytotypes or even be contained in different genomes. The presence or lack, of those markers in other individuals could be viewed as loss or gain of unique cytotypes in the respective population. However, one question unanswered is why heteroplasmy only occurs in some populations and not others. The natural population screens I conducted, and other studies in *Silene vulgaris* (Welch et al. 2006, Pearl et al. 2009), have found that heteroplasmy only occurs in some populations. This is even the case in *M. guttatus* when two mitotypes co-occur in a population (as in HAC) but no evidence of heteroplasmy was found.

**What do we need to do now?**

Much of the work remaining to be done in *M. guttatus* focuses on the effects of heteroplasmy. Specifically finding at what ratio of mitotypes heteroplasmy becomes biologically relevant. Additionally it would be interesting to investigate why heteroplasmy only occurs in some populations.
The work in this thesis has shown no effect of heteroplasmy on the phenotype parameters measured. However it is possible that the effect of heteroplasmy was not expressed in the controlled conditions of the lab. It would be a useful contribution to conduct a reciprocal transplant study in natural populations to assess the effects of mitochondrial heteroplasmy on fitness. This experiment would need to account for potential population effects. A total of 60 plants would be placed at each of four population sites (IM, BR, HAC, CGR) 30 of the plants would be homoplasmic (15 for each mitotype). The remaining 30 plants would be heteroplasmic with 15 plants biased towards each mitotype. To find enough of these plants I would screen multiple natural collected maternal families for evidence of heteroplasmy. I would germinate seedlings in controlled conditions in the lab to prevent confusing the treatment seedlings with plant from the seed bank. As seeds germinated I would transplant them to one area in each population and measure the same phenotypic traits I measured in the lab-based study. I would also allow plants to be pollinated and collect all fruits produced to look for differences in seed set. If there really is no difference between heteroplasmic and homoplasmic plants I would expect to find the same general results as in the lab study, where I found no effect of mitotype on phenotypic traits. If there is an effect of mitotype on phenotypic traits I would expect to find that heteroplasmic plants are either smaller or set fewer seeds due to conflicts between the nuclear and cytoplasmic genomes.

It would also be useful to determine if both mitotypes are expressed in heteroplasmic individuals. Northern blots could be used to determine expression
levels of heteroplasmic individuals. To determine whether or not both mitochondrial haplotypes are expressed in heteroplasmic plants I would first germinate seeds from multiple maternal lineages. I would include seeds from both inbred lines (IM & SF) and multiple maternal lineages known to produce heteroplasmic offspring. I would screen the offspring with qPCR and the primers used in this thesis to determine which plants were heteroplasmic and homoplasmic. Following this I would use the northern blot protocol utilized by Case and Willis (2008) to compare expression levels in plants with different mitotypes.

Another useful and informative avenue of study in M. guttatus would be to determine the potential for CMS expression in heteroplasmic individuals. This study would determine at which ratio of the two mitotypes CMS is expressed. Two studies (Fishman & Willis 2006, Case & Willis 2008) showed evidence of a cytonuclear hybrid incompatibility in crosses between M. guttatus and Mimulus nasutus. This incompatibility is expressed as complete male (pollen) sterility in reciprocal backcrosses to an M. nasutus plant. To determine the potential for CMS expression in heteroplasmic individuals, I would repeat the crossing designs of Fishman and Willis (2006). However, I would use heteroplasmic plants as well as homoplasmic plants in crosses to the inbred lines. Doing crosses in this way would create plants in the F2 generation with no nuclear restorers and varying amounts of the two mitotypes. I would visually inspect anthers for signs of sterility, and stain and count pollen to determine the amount of fertile pollen present. Finally I would use qPCR to assess the ratio of mitotypes in each of the plants in this generation and compare
that ratio to the incidence of sterility expression to determine whether or not heteroplasm effects sterility expression. I would expect that, in the lack of restorers, that pollen fertility would decrease with an increase in the incidence of the ‘S’ mitotype. If sterility were not expressed in heteroplasmic individuals then it may mean that only one mitotype is being expressed in the plants.

Finally, the studies in this thesis, and a study performed by Case and Willis (2008) have shown indirect evidence suggesting the presence of other mitotypes in *M. guttatus* plants. Case and Willis found plants from two populations in which neither known mitotype amplified, and I found plants from one population in which mitochondrial copy numbers were extremely low. Both of these suggest that there may be additional unique mitochondrial genomes in some plants. In order to detect other possible mitotypes, and further confirm heteroplasmy in *Mimulus*, it is necessary to develop additional markers within the mitochondrial genome. In fact one study (Mower et al. in prep) has described the entire mitochondrial genome in *M. guttatus*, and more markers could easily be identified from that data. This could potentially identify differences in length or sequence and used to screen the populations and plants suspected to have additional mitotypes. All the molecular tools that have been recently created for use in *M. guttatus* are based on genotypes from the IM population. This is the population where mtS was first discovered and seems to be centralized (Case & Willis 2008). Therefore, whole mitochondrial genomes could also be sequenced to more easily identify differences between mitotypes, and in fact this might be a better solution.


