2019

CONSIDERING CYTONUCLEAR INTERACTIONS IN THE FACE OF HETEROPLASMY: EVIDENCE FROM DAUCUS CAROTA (APIACEAE), A GYNODIOECIOUS PLANT SPECIES

Adam Joseph Ramsey

CONSIDERING CYTONUCLEAR INTERACTIONS IN THE FACE OF HETEROPLASMY: EVIDENCE FROM *DAUCUS CAROTA* (APIACEAE), A GYNODIOECIOUS PLANT SPECIES

by

Adam Joseph Ramsey

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Biology

The University of Memphis

December 2019

I dedicate this dissertation to the children we have had the pleasure to foster during my time at the University of Memphis: Dasani, Zoey, Isaiah, Caleb, Travaris, and Bodie.

And to our sons, Wesley and Zeke.

Thank you for making life fun.

ACKNOWLEDGEMENTS

I have so many people to thank for the completion of this dissertation. First, I thank all my colleagues and collaborators I have had the opportunity to work with during these last five years: Dave McCauley, Vinhthuy Phan, Bernie Daigle, Caroline Melton, Diem-Trang Pham, Phil Simon, Massimo Iorizzo, and particularly Mike Ballou who helped me with field research on Nantucket Island. In addition, many undergraduates and graduates have assisted in my research. Thank you to Robby Schranze, Kendall Major, Blair Combest, Matthew Tom, Simerjot Virdi, Jakyra Scaife, Matthew DeVazier, Amanda Gelsinger, Nick Bailey, and Nathanael Prather. Several of these have gone on to further their education with advanced degrees of their own. A special thank you to Rob Coleman who assisted me early in my research, and from whom I learned many techniques and had a great many laughs.

Several people with whom I have had numerous conversations in person and via email, phone, and Skype have helped my understanding of mitochondria, paternal leakage, heteroplasmy, and cytonuclear interactions: Patrick Abbot, Jacob Holley, Dviti Mody, James Moore, Damian Dowling, Joel Sharbrough, Hansong Ma, and Justin Havird, among many others who commented on my research at various conferences. I have sometimes doubted the veracity of some of my hypotheses, yet interactions with these colleagues have reinforced my ideas and helped refine hypotheses I wish to address in my future work.

My lab mates have made life and lab work consistently exciting and engaging, all the while suffering my choice of papers for Friday lab meetings. A very heartfelt thank you to Ram Thapa, Carolina Siniscalchi, Paige Murin, Erika Moore, and again, Mike

Ballou, in addition to many others who visited our lab over the last five years. Several other graduate students helped my transition to the University of Memphis and made my time here memorable. Thank you to all former and current graduate students, but in particular, thank you, Denita Weeks, for your friendship and support as we made our way towards graduation.

It is of incredible importance for me to say that my family has been instrumental in the completion of this dissertation. From my extended family to my immediate family – so many to list out here – all have been supportive and believed in my ability to succeed. Thank you, especially, to Zach and Wesley. Zach has been my advocate for years and has sacrificed much to see me succeed. Wesley helped in collecting data for one of my chapters and has been so interested in what I do that it has inspired me continue to the best of my abilities. And baby Zeke. Thank you for being such fun. Seeing you grow and develop is such joy, and it has been a nice distraction from my work when I needed it most.

I would be remiss if I did not thank my wonderful committee. Duane McKenna, you have kept me on my toes with your pointed questions of the deeper impacts of my research. I have so enjoyed your curiosity, Judy Cole. I have always appreciated your desire to learn about and understand my research. Anna Bess Sorin, thank you, for your directed questions, forthright manner, and attention to detail. I will strive to stop writing "This…". In addition to helping keep my botanical knowledge fresh in the world of mitochondria in which I have found myself, you, Randy Bayer, have become a friend and mentor to me and my family. And Jennifer Mandel. As you have been my mentor, colleague, proponent, and friend, thank you. There is no doubt: this dissertation would

not have been completed without your persistence and belief in me. You have pushed me beyond what is essential and necessary for a successful graduate student. Because of your mentorship and guidance over *countless* hours of Monday meetings, I am such a better researcher, thinker, and writer than I would otherwise be if I had mentor of only average commitment to her students. The time you invested in me is greatly cherished. To my committee as a whole, the high confidence I have received from each of you has been much appreciated, and often needed.

I have been fortunate to have many sources of funding during my research. Not only have I received travel support to domestic conferences from the Association of Southeastern Biologists, The University of Memphis College of Arts & Sciences, the Southern Appalachian Botanical Society, the Botanical Society of America (and separately from the Genetics Section thereof), and the Society for Integrative and Comparative Biology, I also received funding from the Botanical Society of America and International Botanical Congress to attend the International Botanical Congress in Shenzhen, China and from the Society for Molecular Biology and Evolution to attend the inaugural Mitochondrial Genomics and Evolution conference in Ein Gedi, Israel. Much of my research was also externally funded. In addition to two Small Research Grants from the Nantucket Biodiversity Initiative, I received a Graduate Student Research Award from both the New England Botanical Club and Botanical Society of America, the Earl Core Student Research Award from the Southern Appalachian Botanical Society, and from the University of Memphis Department of Biological Sciences, the Bill A. Simco Student Research Award.

Finally, I say “Thank you” to the many quadrillions of mitochondria who coordinated with, and persevered under, the regime of my nuclear genome to successfully produce an uncountable number of ATP needed to fuel my brain and body during my efforts to understand your kin in the plant kingdom. I hope I did you well.

PREFACE

My dissertation research at the University of Memphis investigated the effects of organellar heteroplasmy at two organizational levels: 1) individual fitness; and 2) population genetic. I used the species *Daucus carota* ssp. *carota* (wild carrot) as it is known to exhibit organellar heteroplasmy, paternal leakage and gynodioecy, a rare breeding system in which these characteristics, among others, have been documented. Here, I present one chapter as a comprehensive literature review of organellar heteroplasmy in plants, two chapters directly addressing my dissertation research goals, and one chapter directly related to my study species but indirectly related to my dissertation research goals.

Chapter 2, titled "When One Genome Is Not Enough: Organellar Heteroplasmy in Plants", has been published in the journal *Annual Plant Reviews Online* and is formatted as such (Ramsey and Mandel, 2019). Chapter 3, titled "Mitochondrial and Plastid Heteroplasmy, Phenotypic Differentiation, and Fitness in Wild Carrot, *Daucus carota*" is formatted for, and planned for submission to, *American Journal of Botany*. Chapter 4, titled "Heteroplasmy and Patterns of Cytonuclear Linkage Disequilibrium in Wild Carrot", has been published in the journal *Integrative and Comparative Biology* and is such formatted (Ramsey, McCauley, et al., 2019). Chapter 5, titled "Altered pollination ecology of a native plant species in the presence of an attractive non-native plant species" is formatted in the style of the journal *Rhodora* in which it has been published (Ramsey, Ballou, Jr., et al., 2019). This Preface, Chapter 1: Introduction, and Chapter 6: Conclusion are formatted in the style of *American Journal of Botany*. Supplementary material for Chapter 4 is available in Appendix 1 and on the website of *Integrative and*

Comparative Biology, and the same for Chapter 5 is available in Appendix 2 and on FigShare (https://doi.org/10.6084/m9.figshare.7542191.v2).

REFERENCES

Ramsey, A. J., S. M. Ballou, Jr., and J. R. Mandel. 2019. Altered pollination ecology of a native plant species in the presence of an attractive non-native plant species. *Rhodora* 121: 159–173.

Ramsey, A. J., and J. R. Mandel. 2019. When One Genome Is Not Enough: Organellar Heteroplasmy in Plants. *Annual Plant Reviews Online* 2: 1–40.

Ramsey, A. J., D. E. McCauley, and J. R. Mandel. 2019. Heteroplasmy and Patterns of Cytonuclear Linkage Disequilibrium in Wild Carrot. *Integrative and Comparative Biology* icz102: 1–11.

ABSTRACT

Ramsey, Adam Joseph. Ph.D. The University of Memphis. December 2019. Considering Cytonuclear Interactions in the Face of Heteroplasmy: Evidence from *Daucus carota* (Apiaceae), a Gynodioecious Plant Species. Major Professor: Jennifer R. Mandel.

Cytonuclear interactions in plants are the complex result of coevolution between the nuclear, mitochondrial, and plastid genomes. After several billion years of coevolution, metabolic activities are now tightly coordinated. Organellar genomes were believed to be under regulation by the nuclear genome, resulting in uniparental inheritance of small, non-recombining, and haploid organellar genomes. Apparent deviations from this coordination have often been perceived as deleterious to organismal fitness and not tolerated by natural selection. Recently, this assumption has been challenged. Although cytonuclear interactions are indeed highly coordinated, deviations are common. Across species, organellar genomes are cytoploid. They vary in size, recombine, and are neither strictly-uniparentally inherited nor haploid. A common state of organellar genomes is heteroplasmy – intraindividual organellar genetic variation. Mitochondrial heteroplasmy is common throughout eukaryotes, as is plastid heteroplasmy in plants. Interest in cytonuclear interactions has grown, and the presence of two autonomous organellar genomes adds a level of complexity. Even though heteroplasmy is becoming recognized as a compounding factor in cytonuclear interactions, little research in plants has focused on the interface between these phenomena. Here, I present data on that interface using *Daucus carota*. I focused on how heteroplasmy may impact cytonuclear interactions at the individual level, the population level, and the level of genome evolution. I performed a long-term greenhouse study of

individuals from across the range of *D. carota* to characterize phenotypic differences between heteroplasmic and non-heteroplasmic (homoplasmic) individuals. I then revealed patterns of linkage disequilibrium between and within organellar and nuclear genomes to determine if these patterns were consistent between heteroplasmic and homoplasmic individuals. While I do not suggest causation, this research demonstrates that heteroplasmy may impact aspects of cytonuclear interactions by affecting phenotypes and altering levels of nuclear and organellar linkage. Many new questions have arisen, as these results only address a small number of questions relating to heteroplasmy and cytonuclear interactions and offer a limited insight into this intriguing interplay of phenomena.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

After the separate endosymbiotic events that likely gave rise to mitochondria (mt) and plastids (pt), evolution began acting to ensure tight coordination between the endosymbiont and host genomes. As metabolic activities between the endosymbiont and host became interconnected, neither could survive independently. The separate genomes coevolved together, and the interactions between them became ever more tightly coordinated. Over evolutionary time, most endosymbiotic genes were transferred to the host's genome, yet across eukaryotes as a whole, the conservation of the genes remaining in the mt is striking, and in plants, the same is true for the pt (Blanchard and Lynch, 2000; Alberts et al., 2002).

After some 2.5 billion years of coevolution, most eukaryotes still retain their mt. In only one parasitic genus of protozoa, *Monocercomonoides*, has neither mitochondrial genome (mtDNA) nor mitochondrial-like organelle been discovered (Karnkowska et al., 2016). In plants, only parasitic lineages, particularly in the families Orobanchaceae and Rafflesiaceae, seem to have degenerate plastids. In only one plant species has no complete plastid genome (ptDNA) been documented, the parasitic *Rafflesia lagascae* Blanco (Molina et al., 2014). The loss of these organelles and their genomes in parasitic lineages is notable. Parasitic species rely on their hosts for nutrients and survival. Presumably, these species no longer rely on their organelles. As such, relaxed selection on the tight coordination of nuclear and organellar genes has permitted organellar degeneration and loss.

However, in the overwhelming majority of eukaryotic species, selection has acted strongly to maintain effective coordination between nuclear and organellar genomes for

metabolic activities. Mitochondria produce the ATP needed to maintain the high energy demands required by eukaryotes, and pt have evolved several specialized metabolic functions required in plants (e.g., sugar production, chloroplasts; pigment storage, chromoplasts; starch storage, amyloplasts). Even though mt and pt still contain their own genomes, thousands of genes are currently found in the nuclear genome whose products function in either the mt, pt, or both (Peeters and Small, 2001). Sets of organellar and nuclear genes may interact within protein complexes at the molecular level, and variants of one or both may disrupt the function or efficiency of protein complexes, thereby negatively affecting metabolic functions. The results of protein mismatches may affect only a few cells or, more drastically, an individual's overall fitness.

Within the last several decades the importance of interactions between the nuclear and organellar genomes (i.e., cytonuclear interactions) has become clearer. Not only do cytonuclear interactions play an integral role in basic metabolism, they have been shown to affect an organism's overall phenotype, i.e., the phenome. For instance, in *Drosophila melanogaster*, mutations in cytochrome oxidase, found in the mtDNA, cause flies to be sensitive to high temperatures (Ma et al., 2014), and when different fly lines are crossed to produce various mtDNA and nuclear combinations, fitness effects are sex-dependent (Dowling et al., 2007). In plants cytonuclear interactions between mtDNA mutations and nuclear genes can dictate an individual's male fertility (McCauley, 1998) and affect phenotypes relating to phenology and overall reproduction (Roux et al., 2016). Even more striking, a recent hypothesis has been put forth that cytonuclear interactions drive coloration in birds, a phenomenon that contributes to speciation in taxa with strong sexual selection (Hill, 2017).

Much of this work has focused on interactions between the mtDNA and nuclear genome, but few studies have considered organellar genomes as anything more than haploid. However, the traditional view of mtDNA (and to a lesser degree, ptDNA) is being challenged. In contrast to mtDNA being small, circular, and non-recombining molecules which are haploid and uniparentally-inherited, there is a growing body of evidence demonstrating that mtDNA should otherwise be characterized (Burger et al., 2003). Although the putative α-proteobacterial ancestor of the mt presumably had some of the characteristics currently applied to the mt (López-García and Moreira, 2015), several billion years of separate evolutionary histories have changed the structure and function of the mtDNA into something more dynamic and complex. The current view is suggested to have come from the biased sampling of bilaterian animals (Gissi et al., 2010; Lavrov and Pett, 2016), but a large number of taxa in both plants and animals show strikingly different mtDNA characteristics (Lavrov and Pett, 2016; Gualberto and Newton, 2017).

Most notably and of interest to the study of cytonuclear interactions is non-uniparental inheritance and the presence of organellar heteroplasmy, or within individual organellar genetic variation. Both these characteristics challenge the accepted notion of organelles as haploid and being inherited in a non-Mendelian (i.e., uniparental) fashion. Uniparental inheritance was independently described in plants by Correns and Baur in 1909 (Baur, 1909; Correns, 1909; reviewed in Hagemann, 2000). However, both noted and described incidents of plastid variegation (i.e., heteroplasmy), explained by Baur to be the result of biparental inheritance of pt (Baur, 1909). As early as the 1930s mt were also characterized as following a non-Mendelian inheritance pattern (Anderson, 1936).

Even though this early work demonstrated exceptions to non-Mendelian inheritance, it has been overlooked as a transient phenomenon and of little importance, and thus has been little studied (Reboud and Zeyl, 1994).

Until recently, surveys seeking biparental inheritance or heteroplasmy in various species have found no evidence for their presence, or when discovered, they seemed to be exceptional events. However, as more refined methods for detecting these phenomena developed, they seemed to become less exceptional and rather more common. In populations containing no organellar haplotype variants, biparental inheritance would be "cryptic", as parental haplotypes would be indistinguishable. Yet given different organellar haplotypes in populations, instances of biparental inheritance during fertilization (so called maternal or paternal "leakage") would result in heteroplasmy. Heteroplasmy originating in such a manner can be persistent. It may be maintained within an individual and inherited across generations, thus increasing in frequency within populations, as has been shown in cucumber (Shen et al., 2019).

A type of breeding system often associated with leakage of organelles and heteroplasmy is gynodioecy, in which some individuals of a species are male-sterile while others are hermaphroditic (McCauley and Bailey, 2009). There are a variety of mt mutations within and among species thought to cause cytoplasmic male sterility (CMS; Jacobs and Wade, 2003). Interestingly, the nuclear genome often contains "restorer" genes which counteract the mt mutation. A variety of male sterile and nuclear restorer genes can be found within a single species. For instance, in *Daucus carota* and *Plantago coronopus*, there are least four separate CMS-restorer systems, with each mt gene possibly associated with more than one nuclear locus and are thus in a complex state of

cytonuclear interactions. Balancing selection is often cited as a mechanism to maintain such CMS and nuclear restorer pairs (Adhikari et al., 2019). Yet in crosses of lines between different sets of CMS-nuclear restorers, particularly between divergent populations (McCauley, 2013; Breton and Stewart, 2015) or even species (Burt and Trivers, 2006), F_1 progeny often show an increase in the frequency of CMS, due to a disruption of cytonuclear compatibility (Nothnagel et al., 2000; van Damme et al., 2004). Theoretical (Burt and Trivers, 2006; McCauley, 2013) and experimental (Budar et al., 2003; de Cauwer et al., 2011) work has demonstrated fitness differences between male-sterile and hermaphroditic individuals. de Cauwer et al. (2011) went further and showed that hermaphrodites with restored male fertility (due to the presence of a nuclear restorer gene) nonetheless suffer a fitness cost through a reduction in pollen production and fruit set. The authors suggested this is the result of complex cytonuclear interactions involved in male fertility restoration (de Cauwer et al., 2011).

Although cytonuclear interactions are becoming more understood, studies considering the additional phenomenon of heteroplasmy are lacking. As it is often suspected of being a transient occurrence, little emphasis has been placed on the interface of heteroplasmy and cytonuclear interactions. I am particularly interested in the effect(s) of heteroplasmy at two biological organizational levels: the individual and population levels. This project was designed to address the following questions: What are the effects of heteroplasmy on an individual's phenome, and how does it affect linkage with the nuclear genome within and among populations? My goals in addressing these questions were 1) to compare the phenology and fitness-related characteristics of heteroplasmic and non-heteroplasmic (homoplasmic) individuals from a broad distribution of a species'

native and introduced ranges, and 2) assess levels of cytonuclear and organellar linkage disequilibrium between heteroplasmic and homoplasmic individuals within populations in two separate geographic regions and among those regions. I used the gynodioecious species *Daucus carota* ssp. *carota*, as it is known to contain organellar polymorphisms within populations (Mandel et al., 2012), populations with heteroplasmic individuals (Mandel and McCauley, 2015), and the molecular resources (all three sequenced genomes) are readily at disposal (Ruhlman et al., 2006; Iorizzo et al., 2012, 2016).

REFERENCES

Adhikari, B., C. M. Caruso, and A. L. Case. 2019. Beyond balancing selection: frequent mitochondrial recombination contributes to high female frequencies in gynodioecious *Lobelia siphiliticia* L. (Campanulaceae). *New Phytologist*: nph.16136.

Alberts, B., A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. 2002. The Genetic Systems of Mitochondria and Plastids.

Anderson, L. E. 1936. Mitochondria in the Life Cycles of Certain Higher Plants. *American Journal of Botany* 23: 490.

Baur, E. 1909. Das Wesen und die Erblichkeitsverhältnisse der ‘Varietates albomarginatae hort.’ von *Pelargonium zonale*. *Ztschr f ind Abstamm u Vererbungsl* 1: 330–351.

Blanchard, J. L., and M. Lynch. 2000. Organellar genes. Why do they end up in the nucleus? *Trends in Genetics* 16: 315–320.

Breton, S., and D. T. Stewart. 2015. Atypical mitochondrial inheritance patterns in eukaryotes: a review. *Genome* 58: 423–431.

Budar, F., P. Touzet, R. De Paepe, R. De Paepe, R. De Paepe, and R. De Paepe. 2003. The nucleo-mitochondrial conflict in cytoplasmic male sterilities revisited. *Genetica* 117: 3–16.

Burger, G., M. W. Gray, and B. F. Lang. 2003. Mitochondrial genomes: Anything goes. *Trends in Genetics* 19: 709–716.

Burt, A., and R. Trivers. 2006. Genes in Conflict: The Biology of Selfish Genetic Elements. Harvard University Press.

de Cauwer, I., J. F. Arnaud, A. Courseaux, and M. Dufay. 2011. Sex-specific fitness variation in gynodioecious *Beta vulgaris* ssp. *maritima*: Do empirical observations fit theoretical predictions? *Journal of Evolutionary Biology* 24: 2456–2472.

Correns, C. 1909. Zur Kenntniss der Rolle von Kern und Plasma bei der Vererbung. *Ztschr f ind Abstamm u Vererbungsl* 2: 331–340.

van Damme, J. M. M., M. P. J. Hundscheid, S. Ivanovic, and H. P. Koelewijn. 2004. Multiple CMS-restorer gene polymorphism in gynodioecious *Plantago coronopus*. *Heredity* 93: 175–81.

Dowling, D. K., U. Friberg, F. Hailer, and G. Arnqvist. 2007. Intergenomic epistasis for fitness: Within-population interactions between cytoplasmic and nuclear genes in *Drosophila melanogaster*. *Genetics* 175: 235–244.

Gissi, C., G. Pesole, F. Mastrototaro, F. Iannelli, V. Guida, and F. Griggio. 2010. Hypervariability of Ascidian Mitochondrial Gene Order: Exposing the Myth of Deuterostome Organelle Genome Stability. *Molecular Biology and Evolution* 27: 211–215.

Gualberto, J. M., and K. J. Newton. 2017. Plant Mitochondrial Genomes: Dynamics and Mechanisms of Mutation. *Annual Review of Plant Biology* 68: annurev-arplant-043015-112232.

Hagemann, R. 2000. Erwin Baur or Carl Correns: who really created the theory of plastid inheritance? *The Journal of Heredity* 91: 435–440.

Hill, G. E. 2017. The mitonuclear compatibility species concept. *The Auk* 134: 393–409.

Iorizzo, M., S. Ellison, D. Senalik, P. Zeng, P. Satapoomin, J. Huang, M. Bowman, et al. 2016. A high-quality carrot genome assembly provides new insights into carotenoid accumulation and asterid genome evolution. *Nature Genetics*.

Iorizzo, M., D. Senalik, M. Szklarczyk, D. Grzebelus, D. Spooner, and P. Simon. 2012. De novo assembly of the carrot mitochondrial genome using next generation sequencing of whole genomic DNA provides first evidence of DNA transfer into an angiosperm plastid genome. *BMC Plant Biology* 12: 1–17.

Jacobs, M. S., and M. J. Wade. 2003. A synthetic review of the theory of gynodioecy. *The American naturalist* 161: 837–851.

Karnkowska, A., V. Vacek, Z. Zubáčová, S. C. C. Treitli, R. Petrželková, L. Eme, L. Novák, et al. 2016. A eukaryote without a mitochondrial organelle. *Current Biology*

26: 1274–1284.

Lavrov, D. V., and W. Pett. 2016. Animal mitochondrial DNA as we do not know it: Mt-Genome Organization and Evolution in Nonbilaterian Lineages. *Genome Biology and Evolution* 8: 2896–2913.

López-García, P., and D. Moreira. 2015. Open Questions on the Origin of Eukaryotes. *Trends in Ecology and Evolution* 30: 697–708.

Ma, H., H. Xu, and P. H. O'Farrell. 2014. Transmission of mitochondrial mutations and action of purifying selection in *Drosophila melanogaster*. *Nature Genetics* 46: 393–397.

Mandel, J. R., E. V. McAssey, K. M. Roland, and D. E. McCauley. 2012. Mitochondrial gene diversity associated with the atp9 stop codon in natural populations of wild carrot (*Daucus carota* ssp. *carota*). *Journal of Heredity* 103: 418–425.

Mandel, J. R., and D. E. McCauley. 2015. Pervasive Mitochondrial Sequence Heteroplasmy in Natural Populations of Wild Carrot, *Daucus carota* spp. *carota* L. *Plos One* 10: 1–12.

McCauley, D. E. 2013. Paternal leakage, heteroplasmy, and the evolution of plant mitochondrial genomes. *New Phytologist* 200: 966–977.

McCauley, D. E. 1998. The Genetic Structure of a Gynodioecious Plant: Nuclear and Cytoplasmic Genes. *Evolution* 52: 255.

McCauley, D. E., and M. F. Bailey. 2009. Recent advances in the study of gynodioecy:

The interface of theory and empiricism. *Annals of Botany* 104: 611–620.

Molina, J., K. M. Hazzouri, D. Nickrent, M. Geisler, R. S. Meyer, M. M. Pentony, J. M. Flowers, et al. 2014. Possible loss of the chloroplast genome in the parasitic flowering plant *Rafflesia lagascae* (Rafflesiaceae). *Molecular Biology and Evolution* 31: 793–803.

Nothnagel, T., P. Straka, and B. Linke. 2000. Male sterility in populations of *Daucus* and the development of alloplasmic male-sterile lines of carrot. *Plant Breeding* 119: 145–152.

Peeters, N., and I. Small. 2001. Dual targeting to mitochondria and chloroplasts. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1541: 54–63.

Reboud, X., and C. Zeyl. 1994. Organelle inheritance in plants. *Heredity* 72: 132–140.

Roux, F., T. Mary-Huard, E. Barillot, E. Wenes, L. Botran, S. Durand, R. Villoutreix, et al. 2016. Cytonuclear interactions affect adaptive traits of the annual plant *Arabidopsis thaliana* in the field. *Proceedings of the National Academy of Sciences* 113: 3687–3692.

Ruhlman, T. A., S. Lee, R. K. Jansen, J. B. Hostetler, L. J. Tallon, C. D. Town, and H. Daniell. 2006. Complete plastid genome sequence of *Daucus carota*: implications for biotechnology and phylogeny of angiosperms. *BMC genomics* 7: 222.

Shen, J., W. Shou, Y. Zhang, G. Yuan, Y. Zhao, J. Chen, and M. J. Havey. 2019. Rare maternal and biparental transmission of the cucumber mitochondrial DNA reveals sorting of polymorphisms among progenies. *Theoretical and Applied Genetics* 132:

1223–1233.

CHAPTER 2: WHEN ONE GENOME IS NOT ENOUGH: ORGANELLAR HETEROPLASMY IN PLANTS

1 Introduction

Heteroplasmy was discovered over one hundred years ago, yet it is often an overlooked phenomenon. Heteroplasmy is a mixture of detectably different organellar genomes (either mitochondrial or plastid) within a cell or individual (see Glossary). Heteroplasmy is taxonomically widespread across eukaryotes, and the presence of within-individual organellar genomic variation is commonly overlooked (Barr, Neiman, & Taylor, 2005; Basse, 2010; Kmiec, Woloszynska, & Janska, 2006). A single locus can be described as heteroplasmic, but this designation can extend to the genome-wide level. The definition with respect to plastids has sometimes been expanded to include morphological, but non-genetic, differences between functioning and nonfunctioning chloroplasts (Sakamoto, 2003). Mosaics of leaf color, or variegation, was first described by Erwin Baur and Carl Correns in 1909 (reviewed in Hagemann, 2000). Variegation can be caused by developmental changes in specific tissues, but Baur (1909) described it as the somatic segregation of two types of plastids, green and white, which we now call heteroplasmy. Since Baur's work with visual polymorphisms in 1909 and thereafter, other types of plastid heteroplasmy in plants and mitochondrial heteroplasmy in plants, animals, and fungi have been discovered, including structural rearrangements (gene rearrangements and gene chimeras caused by recombination) and sequence changes (insertions or deletions and point mutations).

Glossary:

- **Biparental Inheritance** – when organellar genomes are commonly inherited from both parents; a quantitative trait that exists on a continuum from strict maternal to strict paternal inheritance.
- **Cytonuclear Interactions** – the coordinated network of genes from organellar and nuclear genomes interacting through anterograde and retrograde signaling and proper protein-protein functioning.
- **Cytoploid** – the nature of organellar genomes as having complex inheritance patterns and a multipartite, flexible structure, existing in multiple copies, within which recombination occurs.
- **Gynodioecy (gynodioecious)** – the breeding system in which hermaphrodites and male-sterile individuals occur together in populations, often caused by cytoplasmic male sterility.
- **Heteroplasmy** – the occurrence of multiple, but distinguishable plastid or mitochondrial genomes within a cell or individual.
- **Paternal/maternal Leakage** – the exceptional occurrence of varying levels of biparental inheritance of one or both organellar genomes in species which otherwise have strict uniparental inheritance.
- **Substoichiometric Shifting** – the change in frequencies of cytotypes within an individual through cell divisions, due to the cytoploid nature of organellar genomes.
- **Variegation** – the white and green patterning found on leaves caused by vegetative segregation of heteroplasmy, nuclear control, or a disruption in cytonuclear interactions.
- **Vegetative Segregation** – the loss of heteroplasmy or the shift in substoichiometric levels due to repeated sampling events during cell division, resulting in parts of plants with different cytotypes than one another.

Much work has been carried out in animal systems (especially humans) to study mitochondrial heteroplasmy due to the strong link between its occurrence and mitochondrial diseases (Chinnery & Turnbull, 1999). Mitochondrial heteroplasmy has been detected in a wide variety of animal taxa (Barr et al., 2005; Konrad et al., 2017; Ma, Xu, & O'Farrell, 2014; Robison, Balvin, Schal, Vargo, & Booth, 2015; Wolff et al., 2014; Xiong, Barker, Burger, Raoult, & Shao, 2013). For example, in natural populations of the fruit fly *Drosophila melanogaster*, 14% of individuals were found to be heteroplasmic for the COI (*cox1* in plants) mitochondrial locus (Nunes, Dolezal, & Schlötterer, 2013). A remarkable finding in mice was that while the offspring of a heteroplasmic mother can vary substantially in their levels of heteroplasmy (possibly due to vegetative segregation), the mean offspring level of heteroplasmy is roughly equal to that found in the mother (Jenuth, Peterson, & Shoubridge, 1997). Likewise, in *Caenorhabditis elegans*, heteroplasmy was stable for over 100 generations (Tsang &

Lemire, 2002). In fungi, heteroplasmy of wild type and mutant mitochondrial genomes has been reported in *Neurospora* and *Podospora* (Bertrand, Chan, & Griffiths, 1985; Bertrand, Collins, Stohl, Goewert, & Lambowitz, 1980; Griffiths, 1992). Truly, one only needs to perform a Google Scholar search with heteroplasmy and "organism of choice" to start down the rabbit hole that is the phenomenon of heteroplasmy (heteroplasmy alone yields 25,100 hits as of June 2018).

Heteroplasmy in plants is similarly widespread (Table 1), and indeed, as described by the title of Kmiec et al. (2006), mitochondrial heteroplasmy is often a "common state". Likewise, plastid heteroplasmy has been characterised in numerous plant species (e.g., Ellis, Bentley and McCauley, 2008; Frey, Frey and Forcioli, 2005; Johnson and Palmer, 1989; Mandel *et al.*, 2016; Wheeler *et al.*, 2014). For example, in the green alga, *Chlamydomonas reinhardtii*, persistent chloroplast heteroplasmy was reported in multiple studies (Bolen, Gillham, & Boynton, 1980; Spreitzer & Chastain, 1987). In our own work with wild carrot, *Daucus carota* ssp. *carota*, more than 60% of the populations surveyed harbored heteroplasmic individuals for either the mitochondrial *atp9* or *cox1* locus (Mandel & McCauley, 2015). Studies such as ours, demonstrate that heteroplasmy is widespread (found in many individuals) and abundant (found in observable proportions within individuals) in *D. carota*. These findings are also noteworthy given that heteroplasmy is typically assayed at just a few genes, or loci, and not at the genome-wide level. This suggests that heteroplasmy could be more widespread, since if heteroplasmic genomes do not differ at the specific regions surveyed, it will be missed, likely underestimating the true prevalence of heteroplasmy in natural populations.

Table 1. Some examples of mitochondrial and plastid heteroplasmy and their detection method in green plants and algae.

Order	Family	Species	Common Name	Detection	Genome	References
Alismatales	Potamogetonaceae	*Potamogeton anguillanus*	water plantain	Sanger	pt	Iida et al., 2007
Apiales	Apiaceae	*Daucus carota*	carrot	length variation	pt	Mandel et al., 2016
				Sanger	mt	Szklarczyk et al., 2014
Arecales	Arecaceae	*Phoenix dactylifera*	date palm	NGS	pt	Sabir et al., 2014
				NGS	mt	Sabir et al., 2014
Asterales	Asteraceae	*Helianthus verticillatus*	whorled sunflower	length variation	pt	Ellis et al., 2008
Caryophyllales	Caryophyllaceae	*Silene vulgaris*	bladder campion	length variation	mt	McCauley et al., 2005
Chlamydomonadales	Chlamydomonadaceae	*Chlamydomonas reinhardtii*	green algae	length variation	pt	Nishimura and Stern, 2010
Cucurbitales	Cucurbitaceae	*Cucumis*	cucumber	length variation	mt	Lilly et al., 2001
Ericales	Actinidiaceae	*Actinidia deliciosa*	kiwifruit	length variation	pt	Chat et al., 2002
Fabales	Fabaceae	*Astragalus membranaceus*	milkvetch	NGS	pt	Lei et al., 2016
Fabales	Fabaceae	*Cicer arietinum*	chickpea	length variation	pt	Kumari et al., 2011
Fabales	Fabaceae	*Medicago sativa*	alfalfa	length variation	pt	Lee et al., 1988
				length variation	mt	Pupilli et al., 2001
Geraniales	Geraniaceae	*Pelargonium*	geranium	Sanger	pt	Weihe et al., 2009
				Sanger	mt	Weihe et al., 2009
Lamiales	Phrymaceae	*Mimulus guttatus*	monkeyflower	qPCR	mt	Floro, 2011
Lamiales	Oleaceae	*Olea europaea*	olive	Sanger	mt	García-Díaz et al., 2003

Table 1 (continued).

Order	Family	Species	Common Name	Detection	Genome	References
Lamiales	Plantaginaceae	*Plantago lanceolata*	ribwort plantain	Sanger	mt	Levsen et al., 2016
Malpighiales	Passifloraceae	*Passiflora*	passion flower	Sanger	pt	Hansen et al., 2006
Myrtales	Onagraceae	*Oenothera*	primrose	length variation	pt	Stoike and Sears, 1998
Pinales	Pinaceae	*Pinus*	pine	length variation	pt	Govindaraju et al., 1988
Poales	Poaceae	*Oryza sativa*	rice	length variation	pt	Moon et al., 1987
Poales	Poaceae	*Triticum*	durum wheat	NGS	mt	Noyszewski et al., 2014
Poales	Poaceae	*Zea mays*	maize	length variation	mt	Yamato and Newton, 1999
Saxifragales	Cynomoriaceae	*Cynomorium coccineum*	suo yang	Sanger	pt	García et al., 2004

pt = plastid genome, mt = mitochondrial genome

In this review, we cover both plant organellar genomes, discuss the discovery of plastid and mitochondrial heteroplasmy in plants, and describe broad taxonomic patterns for plants when they occur. Next, we introduce and discuss the "cytoploid" nature and inheritance patterns of organellar genomes, their variability in size and content, as well as, their structure and evidence for recombination. We also provide details about molecular methods for detecting heteroplasmy including more recent techniques employing next-generation sequencing. Next, we investigate the potential link between heteroplasmy and the gynodioecious breeding system (having mixtures of females and hermaphrodites within a population). We then discuss some consequences of heteroplasmy and topics in need of further study including the effects on fitness from heteroplasmy. Finally, we conclude with a discussion of future research directions within the field of heteroplasmy including gaps in our understanding of the ecological and evolutionary consequences of this within-individual organellar variation.

2 Discovery of heteroplasmy

The discovery of heteroplasmy as a phenomenon (sometimes called intra-individual or within-individual genetic variation) was virtually simultaneous with the development of the theory of plastid inheritance (i.e., uniparental inheritance) in studies of angiosperms. As Mendelian inheritance of nuclear genes was discovered using visual polymorphisms in pea (Mendel, 1866), so was the non-Mendelian inheritance of organellar genomes. In 1909, Erwin Baur and Carl Correns independently described the inheritance of plastid variegation in plants, though it was not understood that they carried their own genetic material until Ruth Sager and others' work in the 1960s (Chun,

Vaughan, & Rich, 1963; Sager & Ishidat, 1963). Baur (1909) studied *Pelargonium* while Correns (1909) used *Mirabilis, Urtica,* and *Lunaria* to track variegation patterns within a single plant and in controlled crosses. Correns posited that the cytoplasm was the cause of healthy and ill plastids, yet it was Baur who was correct in concluding that it was the plastids themselves that carried the phenotype and thus the genes. Baur established that progenitor cells giving rise to variegated leaves must contain a mixture of white and green plastids (heteroplasmy) (reviewed in Birky, 1995; Hagemann, 2000). Indeed, leaf variegation often results from a mixture of mutated plastid genomes, (leading to defective plastids that produce white sections) with normal plastid genomes (leading to functional plastids that produce green sections) (Chiu, Stubbe, & Sears, 1988).

The term "heteroplasmy", however, was not widely used in the literature to refer to a mixture of organellar genomes until at least the 1970s when Birky and colleagues (Birky, 1978, 1983; Birky, Maruyama, & Fuerst, 1983) began thoroughly detailing the theoretical implications and experimental evidence of organellar inheritance, evolution, and vegetative sorting (within individual drift of organellar variants). The terms "heteroplasmon" and "heteroplasmonic" were also used in the literature around this time to describe organellar variants of fungi, for example, in Lewontin (Lewontin, 1970). Birky (1976) defined a heteroplasmic cell as one “containing mitochondria or chloroplasts of two different genotypes.” Much of the work describing within individual organellar variation was coupled with understanding the inheritance of mitochondrial and plastid genomes in the genetic model systems of paramecia, algae, and yeast during the early 1970s (Adoutte & Beisson, 1972; Beale, Knowles, & Tait, 1972; Birky, 1973). Ruth Sager, well-known for her work in cancer genetics, also worked extensively with

Chlamydomonas to understand plastid inheritance and its evolutionary significance. In fact, Sager (1977) was likely observing heteroplasmic samples when she reported "persistent heterozygotes" in the offspring of *Chlamydomonas* crosses. A number of prominent books about organellar genomics were published at this time, including by Sager: "Cytoplasmic Genes and Organelles" in 1972, and Birky and colleagues: "Genetics and Biogenesis of Mitochondria and Chloroplasts" in 1975. Since this work in the 1970s, advances in our understanding of heteroplasmy have continued to be linked with the genetic composition and inheritance of organellar genomes. Though outside the scope of this plant review, heteroplasmy is an important field across all eukaryotes with major implications for human health and disease (Stewart & Chinnery, 2015), therefore, this field of study is rapidly moving forward since its plant beginnings in *Pelargonium* and *Mirabilis*.

3 The cytoploid nature of organellar genomes: consistently not asexual

In disagreement with various studies showing the flexibility of organellar genomes, the consensus view of organellar genomes as asexual, static molecules remains stubbornly intact. Even within animals there is disagreement about the existence of a typical mitochondrial genome (described as a “myth” in Lavrov and Pett, 2016). Inheritance patterns are increasingly being shown to be flexible (and quantitative rather than categorical; Figure 1). Organellar genomes are fragmented, yielding sublimons and subgenomic fragments, and the circular structure of these genomes is highly contested (a few circular plastid genomes are notable exceptions). As such, the genome content within a single organelle varies from none, to an incomplete genome, and to multiple copies of

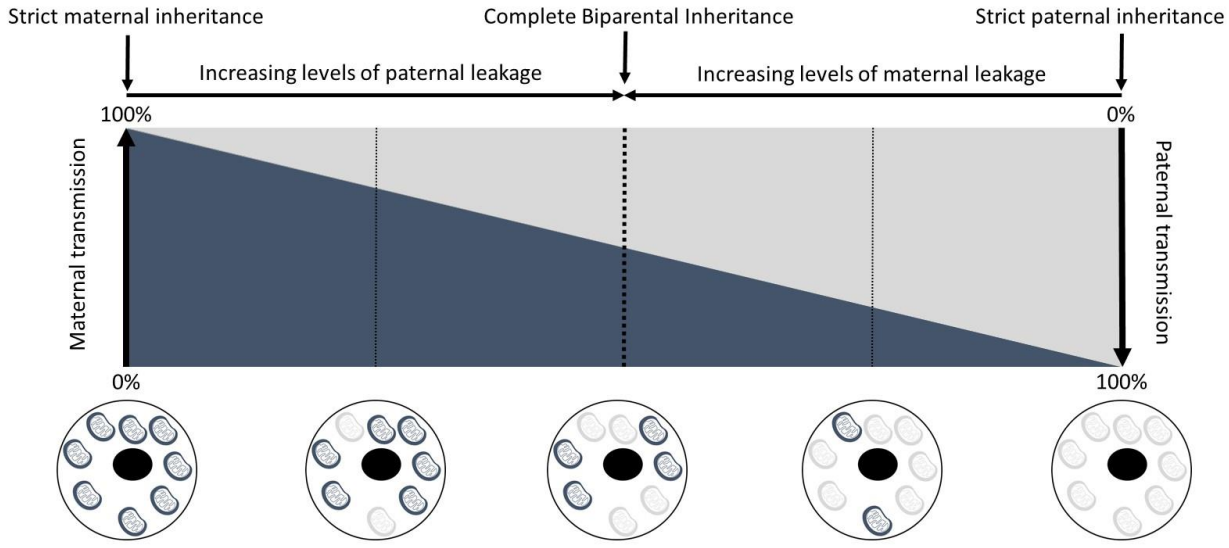


Figure 1. The quantitative nature of organellar inheritance. Strict maternal inheritance is demonstrated on far left (100% maternal transmission), and strict paternal inheritance on far right (100% paternal transmission). Area between represents differing levels of maternal or paternal contribution to the offspring with complete biparental inheritance in the middle (50:50 maternal to paternal transmission). When maternal and paternal organellar genomes differ, departure from either strict inheritance results in heteroplasmy. Circles below the graph represent the homoplasmic offspring of strict maternal and paternal inheritance (assuming neither parent is heteroplasmic) on the left and right, respectively. Circles in the middle represent heteroplasmy generated from varying levels of biparental inheritance 75:25, 50:50, and 25:75, left to right.

the entire genome, which can change through a cell's life cycle (Cupp & Nielsen, 2014; Morley & Nielsen, 2016; Delene J. Oldenburg & Bendich, 2015). The size of the plastid genome is variable, yet fairly constant across plant taxa (~150 kb). But the size range of the mitochondrial genome is regularly being expanded, the current smallest being 222 kb (Handa, 2003). The largest has been increasing over time from 337 kb (Unseld, Marienfeld, Brandt, & Brennicke, 1997) to 2.7 Mb (Rodríguez-Moreno et al., 2011). Currently, the largest mitochondrial genome is found in *Silene conica* at 11.3 Mb (Sloan, Alverson, Chuckalovcak, et al., 2012). These highly variable characteristics are not typical of an asexual (haploid) genome, nor is it accurate to call them fully sexual (diploid +) genomes. We, therefore, define and refer to the nature of plant organellar

genomes as “cytoploid”. In this section we discuss the cytoploid characteristics of plant organellar genomes.

3.1 Inheritance patterns

Organellar genomes are unlike nuclear genomes. One of the most striking differences is the mode of inheritance—both from parent to offspring during reproduction and from cell to cell within individuals during cell division. Nuclear genomes of plants are typically biparentally-inherited, with the maternal and paternal parents contributing equally to their offspring. In contrast, the genomes of mitochondria and plastids are typically described as uniparentally inherited (usually said to be maternally-inherited in angiosperms, but paternally-inherited for the plastids of gymnosperms with no maternal contribution to the offspring, Birky, 1995; Reboud and Zeyl, 1994). However, these qualitative statements about organellar inheritance have been challenged (Birky, 1994; McCauley, 2013; McCauley & Olson, 2008; Wu, Cuthbert, Taylor, & Sloan, 2015). With these reconsiderations, the additional findings of Baur and Correns in 1909 (i.e., rare biparental inheritance of plastids) have been revisited (reviewed in Hagemann, 2000).

In many cases where a departure from strict uniparental inheritance is documented, it occurs along with, but not instead of, maternal inheritance. This type of untraditional biparental inheritance is commonly termed "paternal leakage" (Burt & Trivers, 2006; McCauley, 2013). As organelles are considered maternally-inherited, instances of leakage are defined as paternal, yet “maternal leakage” is also a possibility in species with paternally-inherited organelles (e.g., conifers). Such is the case in Kiwifruit; plastids are inherited paternally and mitochondria maternally (Chat, Chalak, & Petit,

1999), yet Chat et al. (2002) found one instance of plastid heteroplasmy which the authors described as being the result of biparental inheritance which we suggest is maternal leakage. These findings suggest organellar inheritance can be described as a quantitative trait (*sensu* Birky, 1995, 2008; Figure 1). In fact, should leakage occur such that the typical minor parental haplotype outnumbers the typical major parental haplotype, a full "shift" in the inheritance pattern could occur, particularly in crosses of divergent populations or species due to a breakdown of mechanisms preventing biparental inheritance or paternal leakage (Barnard-Kubow, So, & Galloway, 2016; Nagata, 2010). Biparental inheritance and paternal or maternal leakage are often synonymous in the literature; however, we suggest that biparental inheritance be used to describe the inheritance pattern, whereas paternal or maternal leakage be used to describe the shift in inheritance pattern.

Organellar genomes are reported to be biparentally-inherited (19% of plant species exhibit potential biparental inheritance (PBPI), Zhang and Sodmergen, 2010) in many taxonomically diverse plants (Barr et al., 2005; Corriveau & Coleman, 1988; Hu, Zhang, Rao, & Sodmergen, 2008; McCauley, 2013; White, Wolff, Pierson, & Gemmell, 2008; Zhang, Liu, & Sodmergen, 2003). However, biparental transmission does not necessarily result in 50:50 ratios between parents nor does it occur during every fertilization event, even in those species in which it has been detected (Birky, 1995). Inheritance of organelles can be independent of one another. That is, paternal leakage may involve the transmission of mitochondria or plastids, or both, into the fertilised egg. When maternal and paternal organellar genomes are different from one another in size, content, or structure, the offspring is said to be heteroplasmic. A variety of mechanisms

appear to have independently evolved to prevent biparental inheritance of organellar genomes and maintain genetic homogeneity, or homoplasmy, across a large number of taxa (Birky, 2001; Nagata, 2010). Organelles, or their DNA, can be physically excluded from the egg during fertilization, or the organellar genomes within the sperm cell can be degraded before or after fertilization through an assortment of digestive pathways (Birky, 1995, 2001; Clay Montier, Deng, & Bai, 2009; Nagata, 2010). A broad hypothesis explaining paternal leakage is that the taxon-specific mechanism(s) preventing paternal inheritance break down for some, usually, unknown reason (Breton & Stewart, 2015). In some species, it tends to occur in specific populations or particular intra- and interspecific crosses more frequently than in others, potentially due to disrupted cytonuclear co-adaptations (Bentley, Mandel, & McCauley, 2010; Burt & Trivers, 2006). Another possibility is that if biparental inheritance does occur, but the two parental genomes are not divergent (at least in the particular regions assayed), the inheritance pattern will not be detected (discussed in Section 4).

A second major difference in organellar genomes from their nuclear counterpart, occurs during mitotic and meiotic cell division, where each daughter cell receives roughly half of the organelles from the mother cell in a randomly sorted fashion—a process termed vegetative segregation akin to genetic drift (Birky, 1978) and as mentioned above can be described as a quantitative trait (Birky, 1995). A consequence of strict uniparental inheritance and vegetative segregation would be homoplasmy, or within-individual genetic homogeneity, of organellar genomes. This is both because uniparental inheritance precludes the mixing of parental organellar genomes at fertilization, and homoplasmy is enforced by repeated sampling events (within-individual

drift as above) associated with the transmission of a finite number of organelles from mother to daughter cells during division (Birky, 2001). Given this, any heteroplasmy could be lost over cell divisions and consequently should be rare. This drift, or vegetative sorting, can occur within a single individual and across generations as oocytes also obtain a random sampling of organelles from the maternal parent. It would be very intriguing to assay large clonal species (e.g., Quaking Aspen or Creosote Bush) for heteroplasmy in order to better understand vegetative segregation.

Some research in animal systems has shown that segregation to homoplasmy is not the case (Hagström, Freyer, Battersby, Stewart, & Larsson, 2014; Ma & O'Farrell, 2015; Ma et al., 2014). Even in the mitochondrial genome of animals, heteroplasmy can be maintained for up to 50 generations in mice (Hagström et al., 2014), and it is possible that intracellular selection occurs in developing egg cells such that heteroplasmy is selected for and maintained across generations (Ma et al., 2014). When this is the case (and barring extraneous paternal leakage), offspring would be expected to exhibit low variance around the mean level of heteroplasmy (which should be close to the maternal level of heteroplasmy), and few offspring would be expected to revert to a homoplasmic state. An additional example is found in *Caenorhabditis elegans* where one copy of the mitochondrial genome contains a deletion of 11 genes, which on its own, would be lethal, yet the mutant genome appears to have a replicative advantage resulting in an animal strain in which heteroplasmy (mutant/wild type) has been maintained for over 100 generations (Tsang & Lemire, 2002).

3.2 Size and content

Earlier studies of land plants reported that mitochondrial genomes showed the lowest synonymous substitution rates among the three genomes and that the plastid genome was the next lowest at about half the rate of the nuclear genome (Wolfe, Li, & Sharp, 1987) and when compared to animal mitochondrial genomes, both organellar genomes were lower. Overall, this has remained the case with a few notable exceptions of angiosperms with accelerated rates of mitochondrial genome evolution including bugleweed, *Ajuga reptans* (Zhu, Guo, Jain, & Mower, 2014) and certain campion (*Silene*) species (Sloan, Alverson, Wu, Palmer, & Taylor, 2012). These lower levels of substitution rates and thus low sequence diversity (especially within species) potentially make detecting heteroplasmy more difficult in plants than animals since divergent genomes may not differ at loci used for assaying heteroplasmy. Typically, non-coding regions of genomes accumulate mutations faster, and are hence more variable, than coding regions making these genic regions less informative for studies of heteroplasmy. However, the coding regions of mitochondrial genomes in gynodioecious species are often highly variable and sometimes as variable as the non-coding regions (see section 5 below). This leads to unexpectedly high intraspecies diversity assisting in the detection of heteroplasmy (see section 4 below).

In terms of DNA content, the numbers of organelles and copies of their genome are not necessarily associated with one another and are highly variable across taxa ranging from hundreds to thousands of copies per cell (Coleman & Nerozzi, 1999; Preuten et al., 2010; Zoschke, Liere, & Börner, 2007). Replication of DNA within organelles is disjunct from nuclear DNA replication. That is, replication of these genomes

does not parallel that of the nuclear genome nor, in general, the cell cycle (Clay Montier et al., 2009). In fact, the number of organelles found within a cell varies greatly, from as low as 10 to hundreds of thousands for mitochondria (Galtier, 2011) and 20 to hundreds for plastids in angiosperms (Sakamoto, Miyagishima, & Jarvis, 2008). The number of genomes found within a single organelle can also vary. A single mitochondrion may contain only a small part of the entire mitochondrial genome or even no DNA whatsoever, while others may contain several copies of the whole genome (Preuten et al., 2010). Indeed, the notion that mitochondria are discrete entities within the cell may not be entirely correct. Studies of mitochondrial dynamics have revealed that mitochondria are fluid structures with indefinite shapes and sizes that fuse and divide somewhat randomly within the cell (Logan, 2006; Logan & Leaver, 2000; Wang et al., 2015). Furthermore, mitochondria form tubules that connect mitochondria to one another, forming a large mitochondrial network across the cell, and since the mitochondrial tubules contain mitochondrial matrix, movement of protein products and DNA is possible across the cell's volume (Wang et al., 2015). This fluidity allows the mitochondrial genome to be scattered around a cell's population of mitochondria in what is known as a chrondriome (Logan, 2006). For a single chloroplast, the number of genomes can vary from at least 10 to hundreds and varies depending upon developmental stage (Zoschke et al., 2007). Even the plastid genome may find itself in a linear, concatemeric, or even branched form (reviewed in (Maréchal and Brisson, 2010). Plastids also have stromules that connect one to another under stress situations, though this is less studied (Brunkard, Runkel, & Zambryski, 2015; Schattat, Klösgen, & Mathur, 2012). As such, even though

the number of genes can vary, the total mitochondrial or plastid genome is not haploid, but polyploid at the cell level (Preuten et al., 2010).

3.3 Structure, recombination, and substoichiometric shifting within organellar genomes

The structure of organellar genomes is highly variable and is consistently not detected as a static circular genome (Delene J. Oldenburg & Bendich, 2015), although a "master circle" is a convenient mapped arrangement for many species (Mower, Case, Floro, & Willis, 2012; Delene J. Oldenburg & Bendich, 2015; Sloan, 2013). In fact, many different structures are found within organellar genomes *in vivo*. Only some plastid genomes have been observed as circular, e.g. *Arabidopsis*, *Nicotiana*, and *Pisum* (Lilly, Havey, Jackson, & Jiang, 2001). A convincing visual example is found in tobacco. Lilly et al. (2001) found circular plastid genomes using fiber-FISH *in vivo*. A circular molecule was potentially found in the mitochondrial genome of the moss *Physcomitrella* as a fragment length variant (~300 kb; Terasawa et al., 2007), although the authors also suggest it could be a composite of linear multimers as is found in *Marchantia* (D J Oldenburg & Bendich, 1998; Delene J. Oldenburg & Bendich, 2015). Tobacco mitochondrial DNA molecules have been found as branched linear fragments, as well. This is posited to be the result of recombination-dependent DNA replication (D. Oldenburg & Bendich, 1996). In contrast to most of the sampled mitochondrial genomes in animals, plant organellar genomes contain introns. For example, *Chlamydomonas* has introns in both plastid (*psaB* and *psbC*) and mitochondrial (*rrnL*) genes (Korpelainen, 2004), and the tobacco plastid genome has 15 genes that contain introns (Shinozaki et al.,

1986). Palmer et al. (2000) determined that there has been an extraordinarily numerous acquisition of homing intron I group in 48 of 281 angiosperm species, notably within the mitochondrial *cox1* gene.

Recombination occurs in the plastid genome but usually not as frequently as in the mitochondrial genome (reviewed in Maréchal and Brisson, 2010; Wolfe and Randle, 2004). A mitochondrial example was provided in two plant studies where the four-gamete rule was demonstrated for *atpA* and *cob* in *Silene* (McCauley & Ellis, 2008) and *atp9* and *atp1* in wild carrot, *D. carota* (Mandel, McAssey, Roland, & McCauley, 2012). Recombination within organellar genomes is facilitated by the frequent fission and fusion of organelles (mitochondria, Logan, 2010); less likely are the plastids, Medgyesy, Fejes and Maliga, 1985); (Hanson & Sattarzadeh, 2011). These and other studies have pointed out that, should recombination occur between genomes that are not sufficiently variable from one another, no new allelic combinations would be found. Indeed, the genus *Silene* is highly variable in regard to the mitochondrial genome structure: within the genus the size of the genome ranges from 250 kb to 11.3 Mb in size. Incredibly, this is nearly the full range of the mitochondrial genome found within all angiosperms (Sloan et al., 2012).

Recombination within repeat regions creates an interesting feature of these genomes: subgenomic fragments (sublimons, chromosomes, subgenomic circles/molecules). The subgenomic nature of the mitochondrial genome reaches its fullest extreme in *Silene*. With such large mitochondrial genomes, some species of *Silene* are described as being comprised of chromosomes. Chromosomes within the mitochondrial genome are also variable in structure and number. That is, they are not "finished" chromosomes, as described in Bendich (2007), especially when compared to

nuclear chromosomes. For example, even though the mitochondrial genome contains only a few dozen genes and would not be expected to have more chromosomes than genes, *S. noctiflora* has a surprising 50 mitochondrial chromosomes, totaling some 7 Mb in size, yet even within the species, two populations were found to vary by 19 entire chromosomes, even though high similarity was found in the coding regions, with only 9 SNP sites found in 25 kb of protein coding sequences (Wu et al., 2015).

The multipartite structure (being composed of subgenomic fragments which have flexible structures) of organellar genomes allows for a phenomenon called substoichiometric shifting (SSS). In *Arabidopsis* the single stranded DNA binding protein OSB1 is required for mitochondrial DNA stability and the correct stoichiometric transmission of mitochondrial DNA (Zaegel et al., 2006). OSB1 is expressed in young/gametophytic tissues (i.e., bud shoots/megagametophytes). When it is knocked-out through a T-DNA insertion, mitochondrial recombination products accumulate, and phenotypes of leaf distortion and variegation develop. Through DNA gel blot hybridization, it was shown that mutant DNA patterns differed from wild-type around several genes, and the substoichiometry was altered as additional DNA fragments were found whereas others were not visible in the mutants (Zaegel et al., 2006). An interesting study was carried out by Preuten et al. (2010). Here the authors investigated the copy numbers of individual mitochondrial genes in different *Arabidopsis* organs and tobacco leaves. They determined that *atp1*, *rps4*, *nad6* and *cox1* differ in copy numbers from one another and between organs and during development from the cotyledon to leaf stage in *Arabidopsis*, although there was an increase in the number of organelles and organellar genomes with an increasing size (Preuten et al., 2010).

4 Molecular methods of detection

The molecular detection of heteroplasmy is contingent upon the degree of divergence among genomic variants, the ratio of genomic variants within the individual, and the sensitivity of the detection method. For example, in methods that rely upon PCR for the detection of heteroplasmy, the selective amplification of the major variant in a sample may hinder identifying rare variants. Methods of detecting heteroplasmy have progressed as molecular techniques have advanced, especially sequencing technologies. Numerous studies have detected heteroplasmy from controlled cross experiments where departures from strict maternal inheritance (i.e., paternal leakage) are found (e.g., *Aegilops* and *Triticum*, Aksyonova et al., 2005; *Helianthus*, Ellis, Bentley and McCauley, 2008; *Passiflora*, Hansen et al., 2007; *Medicago*, Masoud, Johnson and Sorensen, 1990; *Silene*, McCauley et al., 2005). Evidence for such departures are often noted in crosses between divergent parents, often different species, leading to the concept that perhaps paternal leakage, and the resulting heteroplasmy, are more common in wide crosses (Barr, Neiman and Taylor, 2005; Rokas, Ladoukakis and Zouros, 2003; Xu, 2005). However, it is difficult to disentangle whether increased evidence for paternal leakage and heteroplasmy result from a breakdown of the mechanisms that maintain maternal inheritance or, since divergent parents are likely more different at the sequence level, a detection bias exists. For example, some gynodioecious species (see Section 5) also have high levels of paternal leakage (McCauley, 2013; McCauley & Olson, 2008; Pearl, Welch, & McCauley, 2009), and given that paternal leakage occurs, the differing

haplotypes create sufficient intra-individual variation that heteroplasmy can be detected (McCauley, 2013).

In the following section, we review the different kinds of molecular methods for heteroplasmy detection and key studies that have utilised a given method including, when possible, information about the experimental setting for the finding. We also note when structural (large-scale rearrangements) versus sequence (single nucleotide polymorphisms, SNPs or insertion-deletion polymorphisms, INDELS) heteroplasmy has been identified.

4.1 As fragment length variation

Restriction Fragment Length Polymorphisms (RFLPs) are variants that show length variation of DNA fragments (as visualised using gel electrophoresis) following digestion by restriction enzymes. Both structural (large-scale deletions, duplications) and sequence variants of the organelle genome can be detected via Southern blot type analyses. In the case of detecting heteroplasmy, typically plastid DNA is isolated from nuclear and mitochondrial DNA in tissue samples. In *Medicago*, Johnson and Palmer (1989) detected plastid heteroplasmy using restriction enzyme analysis (*Xba*I), filter hybridization, and autoradiography techniques using plastid probes taking advantage of RFLPs among accessions of *M. sativa* including several subspecies and other *Medicago* species. The authors identified heteroplasmic plants in the cultivar line *M. sativa* cv Regen S, as well as, one *M. scutellata* plant and interpreted its presence as resulting from biparental inheritance of plastid DNA in *Medicago*. Prior to their work, biparental transmission of plastids was reported to occur at a high frequency in *Medicago* (Smith,

Bingham, & Fulton, 1986). Hansen et al. (2007) employed a similar approach in *Passiflora* experimental crosses but digested total extracted DNA and used radiolabeled *Nicotiana tabacum* L. plastid DNA probes. This study detected heteroplasmy and extensive paternal leakage in both intra- and interspecific crosses.

Instead of first digesting whole organellar DNA, a PCR of specific genes or regions of the organellar genome can be amplified first and then incubated with restriction enzymes that target particular restriction sites in the gene (RFLP-PCR). In Kiwifruit, *Actinidia deliciosa,* Chat et al. (2002) used fragment length variation in the *psbC* and *trnS* region (along with pGEM-T Vector cloning of the PCR product and Sanger sequencing) to report a mixed population of plastid genomes within an individual. Heteroplasmy in the gynodioecious *Silene vulgaris* was detected in a small number of individuals using RFLPs obtained by digesting (using *Alu*I, *Msp*I, or *Dde*I) the PCR products from the amplification of the mitochondrial genes *atp1* and *cox1* and visualizing the products using agarose gel electrophoresis (McCauley et al., 2005). Given that heteroplasmy may be in uneven ratios, (e.g., 95% Type 1 to 5% Type 2), and thus the major variant may outcompete the minor during PCR or be too little to detect via agarose gel electrophoresis, McCauley et al. (2005) devised a clever variant on the RFLP-PCR method to "knockdown" the major variant in the sample. Taking advantage of a *Sma*I restriction site in a subset of the haplotypes discovered in their population sampling, the authors digested whole genomic DNA with *SmaI* prior to PCR amplification which should inhibit the amplification of the haplotypes containing the restriction site. This approach revealed cryptic heteroplasmy in 24 out of 119 sampled individuals.

The above approaches usually target SNP variation by taking advantage of the presence or absence of a restriction site within a sample of DNA or a target sequence. Microsatellite repeats, or Simple Sequence Repeats (SSRs), present within an organellar genome, usually the plastid genome, have also been employed to detect within individual variation. Typically, following amplification of the SSR locus, fragments are visualised using fluorescence-based capillary electrophoresis (such as the Fragment Analyzer by Advanced Analytical Technologies or an Applied Biosystems DNA Analyzer). Employing this approach, Ellis et al. (2008) studied the inheritance of plastid SSRs in 45 controlled intraspecific crosses of *Helianthus verticillatus*. Their study identified 6/323 offspring which were heteroplasmic comprising the variant of the maternal parent, as well as, the variant of the paternal parent. In each of these cases, a primary or major haplotype was detected, and the authors only considered cases in which the minor variant was at least 10% of the primary fragment peak fluorescence level. Overall, their study reported heteroplasmy in 1.86% of the sampled individuals, yet they noted that because the detection sensitivity of their approach required a minimum of 10% fragment peak fluorescence, some heteroplasmy was likely undetected and remained cryptic.

Length variants resulting from insertion-deletion polymorphisms (INDELS) within the plastid genome also serve as a marker to detect heteroplasmy. In natural populations of wild *D. carota* and several cultivar *D. carota* lines, Mandel et al. (2016) fluorescently labeled PCR product of the *trnS* to *trnG* spacer region using the ‘S’ and ‘G’ primers from Hamilton (1999). Labeled fragments were visualised via fluorescence-based capillary electrophoresis and individuals showing more than one allele signal were

characterised as heteroplasmic. Their study detected heteroplasmy in 36 wild *D. carota* plants and in five cultivar *D. carota* lines.

4.2 By Quantitative (Real-time) PCR

While the above methods provide qualitative evidence of heteroplasmy, approaches using Quantitative, or Real-time, PCR (Q-PCR) allow a quantitative assessment of heteroplasmy. Investigating previously documented (via PCR-RFLP) R- and S-haplotype heteroplasmy in *Senecio vulgaris*, Frey et al. (2005) made use of allele-specific primers for the R-haplotype for a plastid triazine-resistance determining site in *psbA*. This approach used SYBR Green as a fluorescent reporter in the Q-PCR and included mixtures of the R- and S-haplotypes as standards comprising 100, 25, 5, and 1%. Unknowns were run and quantified using mixtures as data standards to quantify heteroplasmy. Similar studies have been carried out in the gynodioecious *Silene vulgaris* using Q-PCR using of TaqMan probes (instead of SYBR Green) that distinguish different mitochondrial gene haplotypes (Bentley et al., 2010; Feng, Kaur, Mackenzie, & Dweikat, 2009; Welch, Darnell, & McCauley, 2006). In these studies, Q-PCR was carried out employing two haplotype-specific TaqMan probes for allelic discrimination. The probes each included a 5’ reporter dye either 6-FAM or VIC. As in the *Senecio* study (and *Plantago* described below), artificial mixtures of the two haplotypes (alleles) were created to use in linear regression for quantifying the level of heteroplasmy. These Q-PCR *Silene* studies also used the "knockdown" method of McCauley et al. (2005) to allow for the detection of rare heteroplasmic variants. From this work, McCauley and colleagues demonstrated that mitochondrial heteroplasmy of the *cox1* and *atp1* genes was

pervasive and often resulted from occasional paternal leakage in the species. Following similar approaches targeting SNP variants and Q-PCR, work from our own lab (Mandel & McCauley, 2015; Mandel et al., 2016) identified extensive heteroplasmy in the *cox1* and *atp9* genes in *D. carota*, another gynodioecious species.

4.3 By chain-termination DNA sequencing

The detection of heteroplasmy via chain-termination sequencing and capillary electrophoresis (Sanger Sequencing) comprises amplifying a region of the mitochondrial or plastid genome via PCR and then directly sequencing the product or via cloning (e.g., pGEM-T Vector) the PCR product and then sequencing. Heteroplasmy is usually detected as an overlay of more than one SNP signal (peaks) in the electropherogram (Roy & Schreiber, 2014). In a study of mitochondrial variants of the gynodioecious *Plantago lanceolata,* Levsen et al. (2016) used PCR amplification of two genes, *atp6* and *rps12,* along with cloning and sequencing of the resulting PCR product. In order to estimate the heteroplasmy ratios for each gene, the authors generated artificial mixtures of the two SNPs found within their samples (100% T, 5:95, 10:90, 20:80, 30:70, 40:60, 60:40, 70:30, 80:20, 90:10, 95:5 to 100% C). The authors then used linear regression of the mixture ratio against the mean peak height from the sequencing output chromatograms. Using this approach, 15/179 heteroplasmic variants were detected from 6/10 European populations. Another study using Sanger sequencing for heteroplasmy detection was reported in a follow-up to the *Medicago* work above (Johnson & Palmer, 1989) describing heteroplasmy in the *M. sativa* cv Regen S line (Fitter et al., 1996). In this

subsequent work, Fitter et al. (1996) confirmed Johnson and Palmer's finding using DNA sequencing by the chain-termination method.

4.4 By Next-Generation Sequencing

The detection of heteroplasmy via massively parallel, or Next-Generation Sequencing (NGS), is becoming increasingly common. Whereas most of the work described above has focused on a few genes or regions of an organellar genome, NGS allows for the detection of heteroplasmy across entire genomes. A typical workflow consists of extracting organellar reads by mapping/aligning (e.g., using Bowtie2, Langmead and Salzberg 2012 or BWA, http://bio-bwa.sourceforge.net/) to a reference mitochondrial or plastid genome. Following this, various SNP calling tools are used including GATK (McKenna et al., 2010), SAMTOOLS (Li & Durbin, 2009), or VARSCAN (Koboldt et al., 2012). Scarcelli et al. (2016) provide a thorough comparison on these SNP calling methods (including using them to identify heteroplasmy) and compare findings to Sanger sequencing results in the palm genus *Podococcus*. These authors concluded from an evaluation of nine bioinformatics pipelines, VARSCAN was most effective for calling SNPs in the plastid genome for their species, and they provided recommendations for handling plastid polymorphic SNPs.

Since other processes may generate false positives of heteroplasmy, including sequencing/genotyping error, and transfers among the three plant genomes, taking a conservative approach toward identifying heteroplasmy from NGS data is warranted. While these processes that can generate false heteroplasmy are issues for all detection methods of heteroplasmy, they may be especially problematic when sequencing total

genomic DNA (nuclear, mitochondrial, plastid) using methods that have somewhat high error rates (e.g., particular types of NGS approaches). Therefore, it may be advisable to take a conservative approach when identifying heteroplasmy from NGS data. To this end, Sabir et al. (2014) excluded any region which mapped to both the mitochondrial or plastid genome when detecting heteroplasmy in Date Palm (*Phoenix dactylifera*). When they did this, they eliminated most of the heteroplasmic sites of the mitochondrial genome, but at least 30 sites in the plastid genome remained. One step further, when possible, would be to map to the nuclear genome and eliminate any regions that also map to this genome to avoid potential false positives resulting from DNA transfer--an approach we are taking in our own work (unpublished data).

5 Gynodioecy, paternal leakage, heteroplasmy

Among flowering plants, several different types of breeding systems are found (Glick, Sabath, Ashman, Goldberg, & Mayrose, 2016; McCauley & Bailey, 2009; Neal & Anderson, 2005; Pannell, 2002). The majority of angiosperms are hermaphroditic (>72 %; Yampolsky and Yampolsky, 1922), but gynodioecy, in which females coexist with hermaphrodites (Glick et al., 2016), is the second most common breeding system (as defined by Neal and Anderson, 2005). Still, the number of species exhibiting gynodioecy is relatively small: out of the 250,000+ angiosperm species, 1,325 of them (<<1 %) and only 275 out of the 14,559 (1.8%) angiosperm genera have been described as being gynodioecious (Caruso, Eisen, & Case, 2016; Charlesworth, 2002; Renner, 2014). Although the number of gynodioecious species is relatively small, it is found in many diverse angiosperm families (at least 50, see Table 1 for a subset of families; Jacobs and

Wade, 2003) and has been associated with temperate climates and herbaceous growth patterns (Caruso et al., 2016). Notable genera containing gynodioecious species include *Bidens* (Schultz, 2009), *Silene* (Garraud, 2011), *Brassica* (Montgomery, Bailey, Brown, & Delph, 2014), *Daucus* (Mandel et al., 2016), *Fragaria* (Ashman et al., 2015), *Plantago* (Levsen et al., 2016), *Oryza* (Yamagata, Doi, Yasui, & Yoshimura, 2007), and *Zea* (Forde, Oliver, & Leaver, 1978) from the families Asteraceae, Caryophyllaceae, Brassicaceae, Apiaceae, Rosaceae, Plantaginaceae, and Poaceae, respectively.

Unique features underlie the gynodioecious breeding system. Gynodioecious species contain male-sterile (female) and hermaphroditic individuals. Male-sterility can be a nuclear-encoded trait (Burgess et al., 2002; Junfang Chen, Hu, Vick, & Jan, 2006; De Block, Debrouwer, & Moens, 1997; Mariani, Beuckeleer, Truettner, Leemans, & Goldberg, 1990), but in many instances, the cause of male-sterility is cytoplasmic (Jacobs & Wade, 2003; McCauley, 2013; Wade & McCauley, 2005). There are several known mechanisms of cytoplasmic male sterility (CMS), and the genes involved have been identified in only a few species, such as maize (*Zea mays*), beans (*Phaseolus vulgaris*) and in species of *Brassica* (Mackenzie & Chase, 1990; Pelletier & Budar, 2007). In some types of CMS, in otherwise typically wild-type individuals, a ‘brown anther’ phenotype develops, where a mitochondrial mutation causes apoptosis in developing anthers and no viable pollen. Alternatively, there are some homeotic mutations arising from the mitochondrial genome. These genes code for a retroactive signal to the nucleus where it alters transcription factor activity of MADS-box genes, i.e. those involved in floral development (Linke, Nothnagel, & Börner, 2003). Consequently, various homeotic mutants have been characterised. In some individuals, would-be anthers develop into

petals, sepal-like structures, or even carpels (Kitagawa, Posluszny, Gerrath, & Wolyn, 1994). An example is found in *D. carota* in which all three types of CMS have been observed (Linke et al., 2003; Scheike, Gerold, Brennicke, Mehring-Lemper, & Wricke, 1992). Restorers to such CMS alleles are located in the nuclear genome (Bailey & Delph, 2007; Barr & Fishman, 2010; Wade & McCauley, 2005), and many CMS-nuclear restorers can exist within a single species (Touzet & H. Meyer, 2014; Wade & McCauley, 2005), as in *Plantago coronopus* (van Damme, Hundscheid, Ivanovic, & Koelewijn, 2004). Disruptions between these CMS and restorer alleles can cause an increase in the CMS phenotype, particularly in hybrid species, such as that found in *Mimulus* hybrids (Fishman & Willis, 2006).

There is clear evidence of paternal leakage and heteroplasmy occurring often in gynodioecious species, potentially more than would be predicted. Interestingly, another characteristic of gynodioecious species is that the coding regions of the mitochondrial genomes are highly variable, much more so than non-gynodioecious species. So, should paternal leakage occur in these species, there is a likely probability that heteroplasmy will be created in the resulting offspring. The high amount of mitochondrial genome diversity found within these species allows for the detection of paternal leakage and heteroplasmy with relative ease using a variety of molecular techniques (see section 4). Evidence continues to accumulate from two well-studied genera, *Silene* and *Daucus*. In *Silene vulgaris*, for example, Welch, Darnell and McCauley (2006) report that the *atpA* gene has heteroplasmy levels up to 29% in some wild populations, and that patterns of inheritance suggest heteroplasmy of both the *atpA* and *cox1* genes. Additionally, patterns of non-neutral evolution were significant for *atp1* and *atp9*, indicative of a balanced and

maintained polymorphism, and recombination between mitochondrial genes and within the mitochondrial gene *atp1* implies that there has been at least transient mitochondrial heteroplasmy in ancestral lineages (Houliston & Olson, 2006). In 318 *S. vulgaris* controlled crosses, only 96% showed strict maternal inheritance, with the remaining receiving some level of paternal leakage of mitochondrial (McCauley et al., 2005). Likewise, in *D. carota*, there is evidence of paternal leakage and heteroplasmy. All four possible two-locus genotypes were observed between the *atp1* and *atp9* genes in wild populations of *D. carota* in two separate studies (Mandel et al., 2012; Mandel & McCauley, 2015). This suggests that previous biparental inheritance of the mitochondrial genome occurred which allowed for recombination between the genes. In other wild populations of *D. carota*, heteroplasmy was found at levels up to 30% for the *atp9* and 50% for *cox1* (Mandel et al., 2016).

The maintenance of heteroplasmy (even low levels) within individuals and across generations is remarkable. Within an individual, bottlenecks occur with each cell division event, and the germline bottleneck was believed to limit organellar variation from one generation to the next (modeled in Bergstrom and Pritchard, 1998; Clark, 1988; Roze, Rousse and Michalakis, 2005). However, this is not necessarily the case. With an increase in paternal leakage, intra- and interindividual variation increases (Figure 1). Clark (1988) modeled sampling, mutation, and paternal leakage and found that a balance between sampling effects allows heteroplasmy to be maintained within a population. Dornier and Dufay (2013) modeled gynodioecy and sex ratios with the effect of selfing rate, inbreeding depression, and pollen limitation. The authors did not investigate paternal leakage or heteroplasmy, but they determined that polymorphisms governing the

gynodioecious breeding system could be maintained depending on whether the nuclear restorer cost affects males or females (Dornier & Dufay, 2013). Going further, Wade and McCauley (2005) determined that low levels of paternal leakage were sufficient to maintain stable and persistent heteroplasmy in populations. It has long been suggested that female fitness increases relative to hermaphrodites (Bailey & Delph, 2007; Birky, 2006; Delph & Kelly, 2014; Touzet, 2012) but only at low female:hermaphroditic ratios. In such a situation, the CMS cytoplasm would be expected to increase over time, yet at some point the population would reach pollen-limitation due to the high levels of females with local extinction likely. As there are typically several CMS-nuclear restorers in a species, paternal leakage of wild-type alleles could reduce the possibility of allelic fixation (Wade & McCauley, 2005). Thus, balancing selection potentially plays a role in maintaining CMS polymorphisms due to selection for paternal leakage of organelles which likely generates (or supports levels of) heteroplasmy in the offspring.

6 Consequences of heteroplasmy and topics in need of further study

6.1 The potential to generate genotypic novelty and contribute to organellar evolution

The long-held view that organellar genomes evolve only as asexual entities with little opportunity for the evolutionary processes that act in sexual lineages is increasingly being challenged (reviewed in McCauley, 2013). Recombination that generates genotypic novelty, a key feature of sexual systems, is a potential outcome when heteroplasmy is present (through recombination between differing genomes) and suggests that organellar

genomes need not evolve in an entirely asexual fashion. Instead, organellar genomes may be placed on an asexual-to-sexual continuum (as described in McCauley, 2013) and genetic/genomic information can be used to quantify the degree of sexuality. In particular, heteroplasmy provides opportunity for evolutionary processes such as selection, mutation, and drift to act within individuals (Birky, 1973; Eberhard, 1980; Galtier, 2011).

Recent mathematical modeling has shown that given any non-neutral mutation (advantageous or deleterious), selection against heteroplasmy can lead to the evolution of uniparental inheritance, as is the norm for eukaryotes (Christie & Beekman, 2017b, 2017a; Christie, Schaerf, & Beekman, 2015). Based on these observational and theoretical findings, it would follow that there could be strong selection against the state of heteroplasmy in organisms (Rand, 2001). However, others have postulated that heteroplasmy could be maintained by balancing selection (e.g., Doublet et al., 2008; see Section 5). Some authors have suggested that within-individual selection on organellar genomes provides a potential mechanism for the individual to rapidly adapt (e.g., Christie and Beekman, 2017a, 2017b; Korpelainen, 2004; Woloszynska, 2010)). In addition, the potential for genetic or genomic conflict with the nuclear genome is of much interest (Burt & Trivers, 2006; McCauley, 2013). Thus, the role of adaptive selection in heteroplasmy, cytonuclear interactions, and the evolution of organellar genomes remain open questions.

In our own work, we have documented heteroplasmy in the mitochondrial and plastid genomes of wild carrot, *D. carota* ssp. *carota* (Mandel et al., 2012, 2016; Mandel & McCauley, 2015). Wild carrot is an adventive plant species that is native to the Middle

East and is the progenitor of the cultivated carrot (ssp. *sativus*). In natural populations of wild carrot within the United States, we have documented both indirect and direct evidence for substantial levels of heteroplasmy at two mitochondrial loci and one plastid locus. An intriguing finding of this work is that heteroplasmy can provide the raw material for recombination to generate genotypic novelty (Mandel et al., 2012; see also McCauley and Ellis, 2008). We postulate that rare paternal inheritance (i.e., leakage) may be one source of heteroplasmy in carrot. To this end, we have also investigated whether the assumption of strict maternal inheritance is valid using controlled crosses as a part of a USDA-NIFA funded project to understand pollen flow between crop and wild carrot aimed at assessing the risk of paternal leakage in this system. Our results demonstrate that some degree of paternal inheritance can occur in this species and similar findings have previously been reported in the genus *Daucus* (Boblenz, Nothnagel, & Metzlaff, 1990).

Low levels of heteroplasmy (either within or among individuals) may not significantly contribute to evolutionary processes and are effectively neutral sources of variation. Substantial levels (and these levels may vary in different systems) of heteroplasmy may in turn provide genetic variation for selection to act upon, may allow for the generation of genotypic novelty via recombination, and may alter cytonuclear interactions. The argument that heteroplasmy facilitates co-adaptation between the genomes by maintaining associations between the nuclear and organellar genomes has been put forth as an explanation of heteroplasmy in animal systems (Havird, Whitehill, Snow, & Sloan, 2015; Sharpley et al., 2012). Given the complex interactions and coordination between the nuclear and organellar genomes, Wolff et al. (2014) suggest

that heteroplasmy may allow for 'matching' of nuclear and mitochondrial encoded gene products (which interact closely) in the face of nuclear heterozygosity and/or gene duplicates. This suggestion provides the testable hypothesis of whether interacting organellar-nuclear genes show correspondingly higher levels of heteroplasmy-heterozygosity. As an extension of our heteroplasmy work in carrot, we have also investigated whether patterns of cytonuclear linkage disequilibrium are different in heteroplasmic and homoplasmic individuals. Intriguingly, for both the mitochondrial and plastid genomes, heteroplasmic individuals demonstrate greater levels of cytonuclear linkage disequilibrium when compared to homoplasmic individuals (Ramsey & Mandel, unpublished data). This is an area of inquiry in plants where more work is needed to explore the potential for cytonuclear interaction and coadaptation.

Interesting discoveries have recently been made in the levels of linkage disequilibrium (LD) in *Silene* and *Daucus*. In a static, circular genome, genes are physically linked to one another, and thus, mitochondrial genomes are expected to have high levels of LD. However, genes found on the multipartite mitochondrial genome of plants can be physically unlinked from one another, analogous to genes found on separate chromosomes within the nuclear genome. As such, in the heteroplasmic state, independent assortment of mitochondrial genes occurs. This should lower the levels of LD on these genomes, and in fact, overall levels of LD have been found to be much lower than expected for the mitochondrial genome (Fields et al., 2014; Mandel et al., 2012; Mandel and McCauley, 2015; Ramsey & Mandel, unpublished data). Modest levels of LD (~0.5) were recently found in *D. carota* (Mandel et al., 2012; Mandel & McCauley, 2015) and *S. vulgaris* (Bentley et al., 2010; McCauley & Ellis, 2008). In

McCauley and Ellis (2008) LD varied between pairs of loci, with a range of 0.17 – 0.78, indicating modest LD and the presence of an intermediate level of recombination in *S. vulgaris* as LD did not equal 0 nor 1. Additionally, LD values have been shown to vary over time. Four of 14 cytonuclear associations showed a significant shift over ~7 generations in another *Silene*, *S. latifolia* (Fields et al., 2014). As yet, no published study of cytonuclear LD has investigated how heteroplasmy may influence levels of LD within the mitochondrial genome or LD between the mitochondrial and nuclear genomes. However, there is evidence that LD varies between heteroplasmic and homoplasmic individuals of *D. carota* (Ramsey & Mandel, unpublished data).

6.2 Gene flow and measures of population structure

Since the composition of organellar genomes is usually considered to be homogeneous (homoplasmic) and they are typically uniparentally inherited, these genomes are often described as haploid (i.e., even though there may be many copies, they are identical). Given this, the effective population size should be reduced (roughly four-fold) when compared to that of a diploid, biparentally-inherited nuclear genome, and as such organellar genomes should experience more effective genetic drift (Birky, 1983; Takahata & Maruyama, 1981). This difference, coupled with the generally lower dispersal of seed vs. pollen, is given as the explanation as to why measures of genetic population structure are higher for organellar genomes when compared to the nuclear genome. For example, Petit et al. (2005) reviewed measures of G_{ST} (an analog of F_{ST}) in the literature and reported a mean of 0.637 across maternally-inherited angiosperm organellar genomes compared with 0.184 for the biparentally-inherited nuclear genome.

If, however, heteroplasmy results from paternal leakage of plastid or mitochondrial genomes in pollen, then measures of population structure could be lower (i.e., more gene flow). This was proposed as an explanation for the lower than expected organellar FST measures in a study of mitochondrial and plastid genes in wild carrot (Mandel et al., 2012, 2016). Inspection of the raw data from Petit et al. (2005) reveals reported measures of population structure of maternally-inherited genomes as low at 0.05 (18 studies finding measures of G_{ST} 0.30 and below). Certainly, the focal taxa of these studies (in families: Apiaceae, Asteraceae, Betulaceae, Brassicaceae, Caryophyllaceae, Lamiaceae, Myrtaceae, Orchidaceae, Rosaceae, Salicaceae, Vitaceae) would interesting groups to address paternal leakage and the potential for heteroplasmy.

6.3 The potential to influence phenotype, especially fitness

Phenotypic consequences of heteroplasmy, non-uniparental inheritance, recombination, and differential modes of selection have been explored in animal mitochondrial genomes (reviewed in Barr et al. 2005), and mitochondrial heteroplasmy has long been suggested as a contributor to disease and aging in mammals (Lightowlers, Chinnery, Turnbull, & Howell, 1997). In plants, these processes have been documented but without such a link to the consequences for the fitness of the organism or a deeper understanding of their influence on variation and evolution of organellar genomes. Therefore, an unanswered question in plants is whether phenotypic differences exist between heteroplasmic and homoplasmic individuals, and if so, which genes/regions when found in the heteroplasmic state affect phenotypes? Few studies have explicitly attempted to link heteroplasmy and phenotypic effects in plants (Albert et al., 2003;

dissertation of Floro, 2011). Addressing this question would require determining if there is a threshold effect. It is known that stable heteroplasmy levels within an individual can be quite low, so it seems likely that only at some threshold level would fitness effects be detectable.

Heteroplasmy of normal and defective mitochondrial mutant genomes can lead to phenotypic abnormalities such as CMS (see section 5), where male reproductive function is partly or completely disrupted (Schnable, P. S. & Wise, 1998). In an interesting case of SSS in *Brassica napus*, Chen et al. (2011) suggested that differences in copy numbers of the pol-CMS and nap mitotypes could explain male sterile and fertile phenotypes. In this example, a heteroplasmy-threshold effect may confer a fitness advantage to females. Paternal leakage and heteroplasmy are often posited to disrupt cytonuclear interactions, but Barnard-Kubow, McCoy and Galloway (2016) found biparental inheritance commonplace in *Campanulastrum americanum*, which restored cytonuclear incompatibility which would otherwise occur through hybridization between divergent populations in the eastern United States. Though beyond these examples, very little work in plants has been carried out (McCauley, 2013). In tobacco, a simple rearrangement which resulted from a recombination event was associated with a greater stem height and later flowering, traits suggested to be advantageous for the plant (Albert et al., 2003). Findings such as these suggest that heteroplasmy can provide the genetic material to generate genotypic novelty and have phenotypic consequences (for example, implications for crop improvement) within an individual in the span of just one generation.

The state of being heteroplasmic (regardless of the amount or degree) for mitochondrial genomes is often deleterious to organismal phenotypes or health (and very

little is known about plastid heteroplasmy and its link to phenotypes). In humans, mitochondrial heteroplasmy is often associated with mitochondrial disease, aging, and cancer (reviewed and discussed in (Wallace, 2005; Wallace & Chalkia, 2013). In fact, the study of heteroplasmy has major implications for understanding human health where mitochondrial heteroplasmy is associated with numerous disorders (Stewart & Chinnery, 2015). As such, much work has been carried out to understand the causes and consequences of mitochondrial heteroplasmy in animal systems, yet studies investigating fitness gains/losses in plants are lacking.

6.4 Consequences for phylogenetic studies

Heteroplasmy is important to recognise in phylogenetic studies both in generating and analyzing data and in the interpretation of findings (reviewed in Wheeler et al., 2014 and Wolfe and Randle, 2004). Heteroplasmy has been noted as minor peaks in the chromatograms of Sanger Sequencing reads (see Section 4 above). Often, these minor variants are not considered while analyzing data for phylogenetic studies, but if variants are relatively common in the sample, haplotypes may be scored incorrectly (Wolfe & Randle, 2004). In addition, the presence of heteroplasmy in NGS studies has been noted to cause difficulty in *de novo* assembly of heteroplasmic regions (Lei et al., 2016). In our own work reconstructing Asteraceae phylogenies, we suspect heteroplasmy has caused a similar problem when attempting to assemble *de novo* plastid loci (Mandel, unpublished). Also, within individual variation of plastid loci may be one cause of the conflicting signals detected by Zeng et al. (2014) from plastid genome datasets when reconstructing major angiosperm lineages.

Widespread heteroplasmy in phylogenetic studies could result in the same types of difficulties in identifying orthologs from paralogs of nuclear genes if the different organellar genomes have experienced different phylogenetic histories. Heteroplasmy may also lead to discordance among mitochondrial, plastid, and nuclear genomes. For example, Hansen et al. (2007) report paternal leakage and heteroplasmy in the genus *Passiflora* and note that a plastid phylogeny may not reflect accurate ancestor-descendant relationships. Their work directly demonstrated this when a heteroplasmic individual harboring two divergent plastid genomes was reconstructing in two different clades (Hansen et al., 2006). Interestingly, we find evidence for heteroplasmy and paternal leakage in crop and wild carrot (Mandel & McCauley, 2015; Mandel et al., 2016) and previous authors have noted discordance between nuclear and organellar carrot phylogenies (e.g., (Bradeen & Bach, 2002) that could be linked to the difficulties in assigning orthology when genomes are heteroplasmic. Given these concerns, recent authors have called for caution to be exercised suggesting that the presence of heteroplasmy be assayed when using organellar data for phylogenetic studies (Wheeler et al., 2014).

7 Conclusion

The occurrence of heteroplasmy has been described and studied in plants for over 100 years; however, recent studies have revealed that heteroplasmy is more common and widespread than once thought (Barr et al., 2005; Kmiec et al., 2006; McCauley, 2013). The view of organellar genomes being static, stable molecules is regularly shown to be false. The results of the cytoploid nature of organellar genomes commonly allows for

persistent heteroplasmy, often at high levels, which provides the opportunity for evolutionary processes such as selection, mutation, and drift to act within individuals. Recombination that generates genotypic novelty, a key feature of sexual systems, is a potential outcome when heteroplasmy is present (through recombination between differing genomes) and suggests that organellar genomes need not evolve in an entirely asexual fashion. In fact, the genomes may be placed on an asexual-to-sexual continuum (Fields et al., 2014)) and genetic/genomic information could be used to quantify the degree of sexuality. Advances in sequencing technologies now permit the study of entire organellar genomes within individuals as compared to only one or a few genes, so heteroplasmy may likely be common and could thus be found more often than once thought.

The occurrence of within-individual genetic variation raises profound questions relating to evolutionary processes: How sexual are organellar genomes? How often does within individual selection on organellar genomes occur? Does heteroplasmy influence cytonuclear interactions? Is phenotypic variation associated with heteroplasmy? There is a growing body of literature about the prevalence and consequences of heteroplasmy in natural populations of plants. However, many questions still remain, and additional studies including genome-wide analysis of heteroplasmy coupled with phenotypic assays are needed to address questions about the cytoploid nature of organellar genomes and the fitness effects and evolutionary significance of heteroplasmy.

Acknowledgements

We dedicate this review to David E. McCauley whose work in populations genetics, heteroplasmy, and paternal leakage inspired many of the discussion points in this review. We would also like to thank Michael Wade, C. W. Birky, Patrick Abbot, Jacob Holley, Dviti Mody, Victoria Scott, and members of the Mandel Lab for helpful discussions about the subject of this review. Financial support for this research was provided by the United States Department of Agriculture—National Institute of Food and Agriculture Award 2014-33522-21826 to JRM.

References

Adoutte, A., & Beisson, J. (1972). Evolution of Mixed Populations of Genetically Different Mitochondria in *Paramecium aurelia*. *Nature*, *235*, 393. Retrieved from http://dx.doi.org/10.1038/235393b0

Aksyonova, E., Sinyavskaya, M., Danilenko, N., Pershina, L., Nakamura, C., & Davydenko, O. (2005). Heteroplasmy and paternally oriented shift of the organellar DNA composition in barley–wheat hybrids during backcrosses with wheat parents. *Genome*, *48*(5), 761–769. https://doi.org/10.1139/g05-049

Albert, B., Lelandais, C., Pla, M., Leuret, C., Vitart, V., Mathieu, C., … De Paepe, R. (2003). Amplification of *Nicotiana sylvestris* mitochondrial subgenomes is under nuclear control and is associated with phenotypic changes. *Genetica*, *117*(1), 17–25. https://doi.org/10.1023/A:1022356330794

Asai, S., Shimoda, C., Nishiumi, I., Eguchi, K., & Yamagishi, S. (1999). Isolation of microsatellite loci for paternity testing in the rufous vanga *Scheba rufa*. *Molecular*

Ecology, *8*, 513–525. https://doi.org/10.1111/j.1365-294x.1998.00368.x

Ashman, T.-L., Tennessen, J. A., Dalton, R. M., Govindarajulu, R., Koski, M. H., & Liston, A. (2015). Multilocus Sex Determination Revealed in Two Populations of Gynodioecious Wild Strawberry, *Fragaria vesca* subsp. bracteata. *G3*, *5*(December), 2759–2773. https://doi.org/10.1534/g3.115.023358

Bailey, M. F., & Delph, L. F. (2007). Sex-ratio evolution in nuclear-cytoplasmic gynodioecy when restoration is a threshold trait. *Genetics*, *176*(4), 2465–2476. https://doi.org/10.1534/genetics.107.076554

Barnard-Kubow, K. B., McCoy, M. A., & Galloway, L. F. (2016). Biparental chloroplast inheritance leads to rescue from cytonuclear incompatibility. *New Phytologist*, *213*(3), 1466–1476. https://doi.org/10.1111/nph.14222

Barnard-Kubow, K. B., So, N., & Galloway, L. F. (2016). Cytonuclear incompatibility contributes to the early stages of speciation. *Evolution*, *70*(12), 2752–2766. https://doi.org/10.1111/evo.13075

Barr, C. M., & Fishman, L. (2010). The nuclear component of a cytonuclear hybrid incompatibility in *Mimulus* maps to a cluster of pentatricopeptide repeat genes. *Genetics*, *184*(2), 455–465. https://doi.org/10.1534/genetics.109.108175

Barr, C. M., Neiman, M., & Taylor, D. R. (2005). Inheritance and recombination of mitochondrial genomes in plants, fungi and animals. *The New Phytologist*, *168*(1), 39–50. https://doi.org/10.1111/j.1469-8137.2005.01492.x

Basse, C. W. (2010). Mitochondrial inheritance in fungi. *Current Opinion in Microbiology*, *13*, 712–719. https://doi.org/10.1016/j.mib.2010.09.003

Baur, E. (1909). Das Wesen und die Erblichkeitsverhältnisse der “Varietates

albomarginatae hort." von *Pelargonium zonale*. *Ztschr f Ind Abstamm u Vererbungsl*, *1*, 330–351.

Beale, G., Knowles, J., & Tait, A. (1972). Mitochondrial Genetics in *Paramecium*. *Nature*, *235*, 396. Retrieved from http://dx.doi.org/10.1038/235396a0

Bendich, A. J. (2007). The size and form of chromosomes are constant in the nucleus, but highly variable in bacteria, mitochondria and chloroplasts. *BioEssays*, *29*(5), 474–483. https://doi.org/10.1002/bies.20576

Bentley, K. E., Mandel, J. R., & McCauley, D. E. (2010). Paternal Leakage and Heteroplasmy of Mitochondrial Genomes in *Silene vulgaris*: Evidence From Experimental Crosses. *Genetics*, *185*(3), 961–968. https://doi.org/10.1534/genetics.110.115360

Bergstrom, C. T., & Pritchard, J. (1998). Germline bottlenecks and the evolutionary maintenance of mitochondrial genomes. *Genetics*, *149*(4), 2135–2146. https://doi.org/10.1146/annurev.ge.25.120191.000401

Bertrand, H., Chan, B. S.-S., & Griffiths, A. J. F. (1985). Intertion of a foreign nucleotide sequence into mitochondrial DNA causes senescence in *Neurospora intermedia*. *Cell*, *41*(3), 877–884.

Bertrand, H., Collins, R., Stohl, L. L., Goewert, R. R., & Lambowitz, a M. (1980). Deletion mutants of *Neurospora crassa* mitochondrial DNA and their relationship to the "stop-start" growth phenotype. *Proceedings of the National Academy of Sciences*, *77*(10), 6032–6036. https://doi.org/10.1073/pnas.77.10.6032

Birky, C. W. (1973). On the origin of mitochondrial mutants: evidence for intracellular selection of mitochondria in the origin of antibiotic resistant cells in yeast. *Genetics*,

74(3), 421–432.

Birky, C. W. (1976). The Inheritance of Genes in Mitochondria and Chloroplasts. *BioScience*, *26*(1), 26–33. https://doi.org/10.1146/annurev.genet.35.102401.090231

Birky, C. W. (1978). Transmission Genetics of Mitochondria and Chloroplasts. *Ann. Rev. Genet*, *12*, 471–512. https://doi.org/doi:10.1146/annurev.ge.12.120178.002351

Birky, C. W. (1983). Relaxed cellular controls and organelle heredity. *Science (New York, N.Y.)*, *222*, 468–475. https://doi.org/10.1126/science.6353578

Birky, C. W. (1994). Relaxed and Stringent Genomes: Why Cytoplasmic Genes Don't Obey Mendel's Laws. *J Hered*, *85*, 355–365.

Birky, C. W. (1995). Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(December), 11331–11338. https://doi.org/10.1073/pnas.92.25.11331

Birky, C. W. (2001). The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. *Annual Review of Genetics*, *35*, 125–148. https://doi.org/10.1146/annurev.genet.35.102401.090231

Birky, C. W. (2006). Mitochondrial Non-Mendelian Inheritance : Evolutionary Origin and Consequences. In *Major Reference Works*. *eLS*. https://doi.org/doi:10.1002/9780470015902.a0006129

Birky, C. W. (2008). Uniparental inheritance of organelle genes. *Current Biology*, *18*(16), 692–695. https://doi.org/10.1016/j.cub.2008.06.049

Birky, C. W., Maruyama, T., & Fuerst, P. (1983). An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some

results. *Genetics*, *103*, 513–527.

Boblenz, K., Nothnagel, T., & Metzlaff, M. (1990). Paternal inheritance of plastids in the genus *Daucus*. *MGG Molecular & General Genetics*, *220*(3), 489–491. https://doi.org/10.1007/BF00391760

Bolen, P. L., Gillham, N. W., & Boynton, J. E. (1980). Evidence for persistence of chloroplast markers in the heteroplasmic state in *Chlamydomonas reinhardtii*. *Current Genetics*, *2*(2), 159–167. https://doi.org/10.1007/BF00420628

Bradeen, J., & Bach, I. (2002). Molecular diversity analysis of cultivated carrot (*Daucus carota* L.) and wild *Daucus* populations reveals a genetically nonstructured composition. *Journal of the American Horticultural Society*, *127*(3), 383–391. Retrieved from http://journal.ashspublications.org/content/127/3/383.short

Breton, S., & Stewart, D. T. (2015). Atypical mitochondrial inheritance patterns in eukaryotes: a review. *Genome*, *58*(10), 423–431. https://doi.org/10.1139/gen-2015-0090

Brunkard, J. O., Runkel, A. M., & Zambryski, P. C. (2015). Chloroplasts extend stromules independently and in response to internal redox signals. *Proceedings of the National Academy of Sciences*, *112*(32), 10044–10049. https://doi.org/10.1073/pnas.1511570112

Burgess, D. G., Ramsey, M., Hanson, W. G., Heckert, M., Ho, M., Jenq, T., … Gutterson, N. (2002). A novel, two-component system for cell lethality and its use in engineering nuclear male-sterility in plants. *Plant Journal*, *31*(1), 113–125. https://doi.org/10.1046/j.1365-313X.2002.01330.x

Burt, A., & Trivers, R. (2006). *Genes in Conflict: The Biology of Selfish Genetic*

Elements. Harvard University Press.

Caruso, C. M., Eisen, K. E., & Case, A. L. (2016). An angiosperm-wide analysis of the correlates of gynodioecy. *International Journal of Plant Sciences*, *177*(2), 115–121. https://doi.org/10.1086/684260

Charlesworth, D. (2002). What maintains male-sterility factors in plant populations? *Heredity*, *89*(6), 408–409. https://doi.org/10.1038/sj.hdy.6800193

Chat, J., Chalak, L., & Petit, R. J. (1999). Strict paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in intraspecific crosses of kiwifruit. *Theoretical and Applied Genetics*, *99*(1–2), 314–322. https://doi.org/10.1007/s001220051238

Chat, J., Decroocq, S., Decroocq, V., & Petit, R. (2002). A case of chloroplast heteroplasmy in kiwifruit (*Actinidia deliciosa*) that is not transmitted during sexual reproduction. *Journal of Heredity*, *93*(4), 293–300. https://doi.org/10.1093/jhered/93.4.293

Chen, Jianmei, Guan, R., Chang, S., Du, T., Zhang, H., & Xing, H. (2011). Substoichiometrically Different Mitotypes Coexist in Mitochondrial Genomes of *Brassica napus* L. *PLoS ONE*, *6*(3), e17662. https://doi.org/10.1371/journal.pone.0017662

Chen, Junfang, Hu, J., Vick, B. A., & Jan, C. C. (2006). Molecular mapping of a nuclear male-sterility gene in sunflower (*Helianthus annuus* L.) using TRAP and SSR markers. *Theoretical and Applied Genetics*, *113*(1), 122–127. https://doi.org/10.1007/s00122-006-0278-2

Chinnery, P. F., & Turnbull, D. (1999). Mitochondrial DNA and disease. *The Lancet*,

354, S17–S21.

Chiu, W. L., Stubbe, W., & Sears, B. B. (1988). Plastid inheritance in *Oenothera*: organelle genome modifies the extent of biparental plastid transmission. *Current Genetics*, *13*(2), 181–189. https://doi.org/10.1007/BF00365653

Christie, J. R., & Beekman, M. (2017a). Selective sweeps of mitochondrial DNA can drive the evolution of uniparental inheritance. *Evolution*, *71*(8), 2090–2099. https://doi.org/10.1111/evo.13291

Christie, J. R., & Beekman, M. (2017b). Uniparental inheritance promotes adaptive evolution in cytoplasmic genomes. *Molecular Biology and Evolution*, *34*(3), 677–691. https://doi.org/10.1093/molbev/msw266

Christie, J. R., Schaerf, T. M., & Beekman, M. (2015). Selection against Heteroplasmy Explains the Evolution of Uniparental Inheritance of Mitochondria. *PLoS Genetics*, *11*(4), e1005112. https://doi.org/10.1371/journal.pgen.1005112

Chun, E. H. L., Vaughan, M. H., & Rich, A. (1963). The isolation and characterization of DNA associated with chloroplast preparations. *Journal of Molecular Biology*, *7*(2), 130–141. https://doi.org/https://doi.org/10.1016/S0022-2836(63)80042-X

Clark, A. G. (1988). Deterministic Theory of Heteroplasmy. *Evolution*, *42*(3), 621–626. https://doi.org/10.1007/BF00832686

Clay Montier, L. L., Deng, J. J., & Bai, Y. (2009). Number matters: control of mammalian mitochondrial DNA copy number. *Journal of Genetics and Genomics*, *36*(3), 125–131. https://doi.org/10.1016/S1673-8527(08)60099-5

Coleman, A. W., & Nerozzi, A. M. (1999). Temporal and Spatial Coordination of Cells with Their Plastid Component. *International Review of Cytology*, *193*, 125–164.

https://doi.org/https://doi.org/10.1016/S0074-7696(08)61780-5

Correns, C. (1909). Vererbungsversuche mit blass(gelb)grünen und buntblättrigen Sippen bei *Mirabilis jalapa, Urtica pilulifera* und *Lunaria annua*. *Ztschr f Ind Abstamm u Vererbungsl*, *1*, 291–329.

Corriveau, J. L., & Coleman, A. W. (1988). Rapid Screening Method to Detect Potential Biparental Inheritance of Plastid DNA and Results for Over 200 Angiosperm Species. *American Journal of Botany*, Vol. 75, p. 1443. https://doi.org/10.2307/2444695

Cupp, J. D., & Nielsen, B. L. (2014). Minireview: DNA replication in plant mitochondria. *Mitochondrion*, *19*(PB), 231–237. https://doi.org/10.1016/j.mito.2014.03.008

De Block, M., Debrouwer, D., & Moens, T. (1997). The development of a nuclear male sterility system in wheat. Expression of the barnase gene under the control of tapetum specific promoters. *Theoretical and Applied Genetics*, *95*(1–2), 125–131. https://doi.org/10.1007/s001220050540

Delph, L. F., & Kelly, J. K. (2014). On the importance of balancing selection in plants. *New Phytologist*, *201*(1), 45–56. https://doi.org/10.1111/nph.12441

Dornier, A., & Dufay, M. (2013). How selfing, inbreeding depression, and pollen limitation impact nuclear-cytoplasmic gynodioecy: A model. *Evolution*, *67*(9), 2674–2687. https://doi.org/10.1111/evo.12142

Doublet, V., Souty-Grosset, C., Bouchon, D., Cordaux, R., & Marcadé, I. (2008). A thirty million year-old inherited heteroplasmy. *PloS One*, *3*(8), e2938. https://doi.org/10.1371/journal.pone.0002938

Eberhard, W. G. (1980). Evolutionary Consequences of Intracellular Organelle Competition. *The Quarterly Review of Biology*, *55*(3), 231–249. https://doi.org/10.1086/411855

Ellis, J. R., Bentley, K. E., & McCauley, D. E. (2008). Detection of rare paternal chloroplast inheritance in controlled crosses of the endangered sunflower *Helianthus verticillatus*. *Heredity*, *100*(6), 574–580. https://doi.org/10.1038/hdy.2008.11

Feng, X., Kaur, a P., Mackenzie, S. a, & Dweikat, I. M. (2009). Substoichiometric shifting in the fertility reversion of cytoplasmic male sterile pearl millet. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *118*(7), 1361–1370. https://doi.org/10.1007/s00122-009-0986-5

Fields, P. D., McCauley, D. E., McAssey, E. V, & Taylor, D. R. (2014). Patterns of cyto-nuclear linkage disequilibrium in *Silene latifolia*: genomic heterogeneity and temporal stability. *Heredity*, *112*(2), 99–104. https://doi.org/10.1038/hdy.2013.79

Fishman, L., & Willis, J. H. (2006). A cytonuclear incompatibility causes anther sterility in *Mimulus* hybrids. *Evolution*, *60*(7), 1372–1381. https://doi.org/10.1111/j.0014-3820.2006.tb01216.x

Fitter, J. T., Thomas, M. R., Rose, R. J., & Steele-Scott, N. (1996). Heteroplasmy of the chloroplast genome of *Medicago sativa* L. cv 'Regen S'' confirmed by sequence analysis. *Theoretical and Applied Genetics*, *93*(5), 685–690. https://doi.org/10.1007/BF00224063

Floro, E. R. (2011). *Mitochondrial heteroplasmy in* Mimulus guttatus. Kent State University.

Forde, B. G., Oliver, R. J., & Leaver, C. J. (1978). Variation in mitochondrial translation

products associated with male-sterile cytoplasms in maize. *Proceedings of the National Academy of Sciences of the United States of America*, *75*(8), 3841–3845. https://doi.org/10.1073/pnas.75.8.3841

Frey, J. E., Frey, B., & Forcioli, D. (2005). Quantitative assessment of heteroplasmy levels in *Senecio vulgaris* chloroplast DNA. *Genetica*, *123*(3), 255–261. https://doi.org/10.1007/s10709-004-3711-y

Galtier, N. (2011). The intriguing evolutionary dynamics of plant mitochondrial DNA. *BMC Biology*, *9*(61), 1–3. https://doi.org/10.1186/1741-7007-9-61

García-Díaz, A., Oya, R., Sánchez, A., & Luque, F. (2003). Effect of prolonged vegetative reproduction of olive tree cultivars (*Olea europaea* L.) in mitochondrial homoplasmy and heteroplasmy. *Genome*, *46*(3), 377–381. https://doi.org/10.1139/g03-017

García, M. A., Nicholson, E. H., & Nickrent, D. L. (2004). Extensive Intraindividual Variation in Plastid rDNA Sequences from the Holoparasite *Cynomorium coccineum* (Cynomoriaceae). *Journal of Molecular Evolution*, *58*(3), 322–332. https://doi.org/10.1007/s00239-003-2554-y

Garraud, C. (2011). *Evolution of gynodioecy-gynomonoecy Experimental approaches in* Silene nutans.

Glick, L., Sabath, N., Ashman, T.-L., Goldberg, E., & Mayrose, I. (2016). Polyploidy and sexual system in angiosperms: Is there an association? *American Journal of Botany*, *103*(10), 1–13. https://doi.org/10.3732/ajb.1500424

Govindaraju, D. R., Wagner, D. B., Smith, G. P., & Dancik, B. P. (1988). Chloroplast DNA variation within individual trees of a *Pinus banksiana* - *Pinus conforta*

sympatric region. *Canadian Journal of Forest Research*, *18*(10), 1347–1350. https://doi.org/10.1139/x88-209

Griffiths, A. J. F. (1992). Fungal Senescence. *Annual Review of Genetics*, *26*, 351–372.

Hagemann, R. (2000). Erwin Baur or Carl Correns: who really created the theory of plastid inheritance? *The Journal of Heredity*, *91*, 435–440.

Hagström, E., Freyer, C., Battersby, B. J., Stewart, J. B., & Larsson, N. G. (2014). No recombination of mtDNA after heteroplasmy for 50 generations in the mouse maternal germline. *Nucleic Acids Research*, *42*(2), 1111–1116. https://doi.org/10.1093/nar/gkt969

Handa, H. (2003). The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): Comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. *Nucleic Acids Research*, *31*(20), 5907–5916. https://doi.org/10.1093/nar/gkg795

Hansen, A. K., Escobar, L. K., Gilbert, L. E., & Jansen, R. K. (2007). Paternal, maternal, and biparental inheritance of the chloroplast genome in Passiflora (Passifloraceae): Implications for phylogenetic studies. *American Journal of Botany*, *94*(1), 42–46. https://doi.org/10.3732/ajb.94.1.42

Hansen, A. K., Gilbert, L. E., Simpson, B. B., Downie, S. R., Cervi, A. C., & Jansen, R. K. (2006). Phylogenetic Relationships and Chromosome Number Evolution in *Passiflora*. *Systematic Botany*, *31*(1), 138–150. https://doi.org/10.1600/036364406775971769

Hanson, M. R., & Sattarzadeh, A. (2011). Stromules: Recent Insights into a Long Neglected Feature of Plastid Morphology and Function. *Plant Physiology*, *155*(4),

1486–1492. https://doi.org/10.1104/pp.110.170852

Havird, J. C., Whitehill, N. S., Snow, C. D., & Sloan, D. B. (2015). Conservative and compensatory evolution in oxidative phosphorylation complexes of angiosperms with highly divergent rates of mitochondrial genome evolution. *Evolution*, 3069–3081. https://doi.org/10.1111/evo.12808

Houliston, G. J., & Olson, M. S. (2006). Nonneutral evolution of organelle genes in *Silene vulgaris*. *Genetics*, *174*(4), 1983–1994. https://doi.org/10.1534/genetics.106.060202

Hu, Y., Zhang, Q., Rao, G., & Sodmergen. (2008). Occurrence of plastids in the sperm cells of Caprifoliaceae: Biparental plastid inheritance in angiosperms is unilaterally derived from maternal inheritance. *Plant and Cell Physiology*, *49*(6), 958–968. https://doi.org/10.1093/pcp/pcn069

Iida, K., Takishita, K., Ohshima, K., & Inagaki, Y. (2007). Assessing the monophyly of chlorophyll-c containing plastids by multi-gene phylogenies under the unlinked model conditions. *Molecular Phylogenetics and Evolution*, *45*(1), 227–238. https://doi.org/10.1016/J.YMPEV.2007.05.003

Jacobs, M. S., & Wade, M. J. (2003). A synthetic review of the theory of gynodioecy. *The American Naturalist*, *161*(6), 837–851. https://doi.org/10.1086/375174

Jenuth, J. P., Peterson, A. C., & Shoubridge, E. A. (1997). Tissue-specific selection for different mtDNA genotypes in heteroplasmic mice. *Nature Genetics*, *16*, 93–95.

Johnson, L. B., & Palmer, J. D. (1989). *Heteroplasmy of chloroplast DNA in* Medicago. *12*, 3–11.

Kitagawa, J., Posluszny, U., Gerrath, J. M., & Wolyn, D. J. (1994). Developmental and

morphological analyses of homeotic cytoplasmic male sterile and fertile carrot flowers. *Sexual Plant Reproduction*, *7*(1), 41–50. https://doi.org/10.1007/BF00241886

Kmiec, B., Woloszynska, M., & Janska, H. (2006). Heteroplasmy as a common state of mitochondrial genetic information in plants and animals. *Current Genetics*, *50*, 149–159. https://doi.org/10.1007/s00294-006-0082-1

Koboldt, D. C., Zhang, Q., Larson, D. E., Shen, D., Mclellan, M. D., Lin, L., … Wilson, R. K. (2012). VarScan 2 : Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Research*, *22*(3), 568–576. https://doi.org/10.1101/gr.129684.111

Konrad, A., Thompson, O., Waterston, R. H., Moerman, D. G., Keightley, P. D., Bergthorsson, U., & Katju, V. (2017). Mitochondrial mutation rate, spectrum and heteroplasmy in *Caenorhabditis elegans* spontaneous mutation accumulation lines of differing population size. *Molecular Biology and Evolution*, *34*(6), 1319–1334. https://doi.org/10.1093/molbev/msx051

Korpelainen, H. (2004). The evolutionary processes of mitochondrial and chloroplast genomes differ from those of nuclear genomes. *Naturwissenschaften*, *91*(11), 505–518. https://doi.org/10.1007/s00114-004-0571-3

Kumari, M., Clarke, H. J., des Francs-Small, C. C., Small, I., Khan, T. N., & Siddique, K. H. M. (2011). Albinism does not correlate with biparental inheritance of plastid DNA in interspecific hybrids in *Cicer* species. *Plant Science*, *180*(4), 628–633. https://doi.org/10.1016/j.plantsci.2011.01.003

Lavrov, D. V., & Pett, W. (2016). Animal mitochondrial DNA as we do not know it: Mt-

Genome Organization and Evolution in Nonbilaterian Lineages. *Genome Biology and Evolution*, *8*(9), 2896–2913. https://doi.org/10.1093/gbe/evw195

Lee, D. J., Blake, T. K., & Smith, S. E. (1988). Biparental inheritance of chloroplast DNA and the existence of heteroplasmic cells in alfalfa. *Theoretical and Applied Genetics*, *76*, 545–549.

Lei, W., Ni, D., Wang, Y., Shao, J., Wang, X., Yang, D., … Liu, C. (2016). Intraspecific and heteroplasmic variations, gene losses and inversions in the chloroplast genome of *Astragalus* membranaceus. *Scientific Reports*, *6*(May 2015), 1–13. https://doi.org/10.1038/srep21669

Levsen, N., Bergero, R., Charlesworth, D., & Wolff, K. (2016). Frequent, geographically structured heteroplasmy in the mitochondria of a flowering plant, ribwort plantain (*Plantago lanceolata*). *Heredity*, *117*(1), 1–7. https://doi.org/10.1038/hdy.2016.15

Lewontin, R. C. (1970). The Units of Selection. *Annual Review of Ecology and Systematics*, *1*(1), 1–18. https://doi.org/10.1146/annurev.es.01.110170.000245

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Lightowlers, R. N., Chinnery, P. F., Turnbull, D. M., & Howell, N. (1997). Mammalian mitochondrial genetics: heredity, heteroplasmy and disease. *Trends in Genetics*, *13*(11), 450–455. https://doi.org/https://doi.org/10.1016/S0168-9525(97)01266-3

Lilly, J. W., Bartoszewski, G., Malepszy, S., & Havey, M. J. (2001). A major deletion in the cucumber mitochondrial genome sorts with the MSC phenotype. *Current Genetics*, *40*(2), 144–151. https://doi.org/10.1007/s002940100238

Lilly, J. W., Havey, M. J., Jackson, S. a, & Jiang, J. (2001). Cytogenomic analyses reveal the structural plasticity of the chloroplast genome in higher plants. *The Plant Cell*, *13*(2), 245–254. https://doi.org/10.1105/tpc.13.2.245

Linke, B., Nothnagel, T., & Börner, T. (2003). Flower development in carrot CMS plants: Mitochondria affect the expression of MADS box genes homologous to GLOBOSA and DEFICIENS. *Plant Journal*, *34*(1), 27–37. https://doi.org/10.1046/j.1365-313X.2003.01703.x

Logan, D. C. (2006). Plant mitochondrial dynamics. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1763*(5–6), 430–441. https://doi.org/10.1016/j.bbamcr.2006.01.003

Logan, D. C. (2010). Mitochondrial fusion, division and positioning in plants. *Biochem. Soc. Trans.*, *38*(3), 789–795. https://doi.org/10.1042/BST0380789

Logan, D. C., & Leaver, C. J. (2000). Mitochondria-targeted GFP highlights the heterogeneity of mitochondrial shape, size and movement within living plant cells. *Journal of Experimental Botany*, *51*(346), 865–871. https://doi.org/10.1093/jexbot/51.346.865

Ma, H., & O'Farrell, P. H. (2015). Selections that isolate recombinant mitochondrial genomes in animals. *ELife*, *4*(AUGUST2015), 1–16. https://doi.org/10.7554/eLife.07247

Ma, H., Xu, H., & O'Farrell, P. H. (2014). Transmission of mitochondrial mutations and action of purifying selection in *Drosophila melanogaster*. *Nature Genetics*, *46*(4), 393–397. https://doi.org/10.1038/ng.2919

Mackenzie, S. A., & Chase, C. D. (1990). Fertility Restoration Is Associated with Loss of

a Portion of the Mitochondrial Genome in ?Cytoplasmic Male-Sterile Common Bean. *Plant Cell*, *2*(9), 905–912. https://doi.org/10.2307/3869326

Mandel, J. R., McAssey, E. V., Roland, K. M., & McCauley, D. E. (2012). Mitochondrial gene diversity associated with the atp9 stop codon in natural populations of wild carrot (*Daucus carota* ssp. *carota*). *Journal of Heredity*, *103*, 418–425. https://doi.org/10.1093/jhered/esr142

Mandel, J. R., & McCauley, D. E. (2015). Pervasive Mitochondrial Sequence Heteroplasmy in Natural Populations of Wild Carrot, *Daucus carota* spp. *carota* L. *Plos One*, *10*(8), 1–12. https://doi.org/10.1371/journal.pone.0136303

Mandel, J. R., Ramsey, A. J., Iorizzo, M., & Simon, P. W. (2016). Patterns of Gene Flow between Crop and Wild Carrot, *Daucus carota* (Apiaceae) in the United States. *Plos One*, *11*(9), 1–19. https://doi.org/10.1371/journal.pone.0161971

Marechal, A., & Brisson, N. (2010). Recombination and the maintenance of plant organelle genome stability. *New Phytologist*, *186*(2), 299–317. https://doi.org/10.1111/j.1469-8137.2010.03195.x

Mariani, C., Beuckeleer, M. D., Truettner, J., Leemans, J., & Goldberg, R. B. (1990). Induction of male sterility in plants by a chimaeric ribonuclease gene. *Nature*, *347*(6295), 737–741. https://doi.org/10.1038/347737a0

Masoud, S. A., Johnson, L. B., & Sorensen, E. L. (1990). High transmission of paternal plastid DNA in alfalfa plants demonstrated by restriction fragment polymorphic analysis. *Theoretical and Applied Genetics*, *79*(1), 49–55. https://doi.org/10.1007/BF00223786

McCauley, D. E. (2013). Paternal leakage, heteroplasmy, and the evolution of plant

mitochondrial genomes. *New Phytologist*, *200*, 966–977. https://doi.org/10.1111/nph.12431

McCauley, D. E., & Bailey, M. F. (2009). Recent advances in the study of gynodioecy: The interface of theory and empiricism. *Annals of Botany*, Vol. 104, pp. 611–620. https://doi.org/10.1093/aob/mcp141

McCauley, D. E., Bailey, M. F., Sherman, N. A., & Darnell, M. Z. (2005). Evidence for paternal transmission and heteroplasmy in the mitochondrial genome of *Silene vulgaris*, a gynodioecious plant. *Heredity*, *95*(1), 50–58. https://doi.org/10.1038/sj.hdy.6800676

McCauley, D. E., & Ellis, J. R. (2008). Recombination and linkage disequilibrium among mitochondrial genes in structured populations of the gynodioecious plant *Silene vulgaris*. *Evolution*, *62*(4), 823–832. https://doi.org/10.1111/j.1558-5646.2008.00331.x

McCauley, D. E., & Olson, M. S. (2008). Do recent findings in plant mitochondrial molecular and population genetics have implications for the study of gynodioecy and cytonuclear conflict? *Evolution*, *62*, 1013–1025. https://doi.org/10.1111/j.1558-5646.2008.00363.x

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., … DePristo, M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*, 1297–1303. https://doi.org/10.1101/gr.107524.110.20

Medgyesy, P., Fejes, E., & Maliga, P. (1985). Interspecific Chloroplast Recombination in a Nicotiana Somatic Hybrid. *Proc. Natl. Acad. Sci. USA*, *82*(20), 6960–6964.

https://doi.org/10.1073/pnas.82.20.6960

Mendel. (1866). Versuche über Pflanzen-Hybriden. *Verh. Naturforsch. Ver.*, *4*, 3–47.

Montgomery, B. R., Bailey, M. F., Brown, G. G., & Delph, L. F. (2014). Evaluation of the cost of restoration of male fertility in *Brassica napus*. *Botany*, *853*(ja), 1–27. https://doi.org/doi:10.1139/cjb-2014-0140

Moon, E., Kao, T. H., & Wu, R. (1987). Rice chloroplast DNA molecules are heterogeneous as revealed by DNA sequences of a cluster of genes. *Nucleic Acids Research*, *15*(2), 611–630. https://doi.org/10.1093/nar/15.2.611

Morley, S. A., & Nielsen, B. L. (2016). Chloroplast DNA Copy Number Changes during Plant Development in Organelle DNA Polymerase Mutants. *Frontiers in Plant Science*, *7*(February), 57. https://doi.org/10.3389/fpls.2016.00057

Mower, J. P., Case, A. L., Floro, E. R., & Willis, J. H. (2012). Evidence against equimolarity of large repeat arrangements and a predominant master circle structure of the mitochondrial genome from a monkeyflower (*Mimulus guttatus*) lineage with cryptic CMS. *Genome Biology and Evolution*, *4*(5), 670–686. https://doi.org/10.1093/gbe/evs042

Nagata, N. (2010). Mechanisms for independent cytoplasmic inheritance of mitochondria and plastids in angiosperms. *Journal of Plant Research*, Vol. 123, pp. 193–199. https://doi.org/10.1007/s10265-009-0293-x

Neal, P. R., & Anderson, G. J. (2005). Are “mating systems” “breeding systems” of inconsistent and confusing terminology in plant reproductive biology? Or is it the other way around? *Plant Systematics and Evolution*, Vol. 250, pp. 173–185. https://doi.org/10.1007/s00606-004-0229-9

Nishimura, Y., & Stern, D. B. (2010). Differential replication of two chloroplast genome forms in heteroplasmic *Chlamydomonas reinhardtii* gametes contributes to alternative inheritance patterns. *Genetics*, *185*(4), 1167–1181. https://doi.org/10.1534/genetics.110.118265

Noyszewski, A. K., Ghavami, F., Alnemer, L. M., Soltani, A., Gu, Y. Q., Huo, N., … Kianian, S. F. (2014). Accelerated evolution of the mitochondrial genome in an alloplasmic line of durum wheat. *BMC Genomics*, *15*(1), 67. https://doi.org/10.1186/1471-2164-15-67

Nunes, M. D. S., Dolezal, M., & Schlötterer, C. (2013). Extensive paternal mtDNA leakage in natural populations of Drosophila melanogaster. *Molecular Ecology*, *22*(8), 2106–2117. https://doi.org/10.1111/mec.12256

Oldenburg, D., & Bendich, A. (1996). Size and Structure of Replicating Mitochondrial DNA in Cultured Tobacco Cells. *The Plant Cell*, *8*(3), 447–461. https://doi.org/10.1105/tpc.8.3.447

Oldenburg, D J, & Bendich, a J. (1998). The structure of mitochondrial DNA from the liverwort, *Marchantia polymorpha*. *Journal of Molecular Biology*, *276*(4), 745–758. https://doi.org/10.1006/jmbi.1997.1581

Oldenburg, Delene J., & Bendich, A. J. (2015). DNA maintenance in plastids and mitochondria of plants. *Frontiers in Plant Science*, *6*(October), 1–15. https://doi.org/10.3389/fpls.2015.00883

Palmer, J. D., Adams, K. L., Cho, Y., Parkinson, C. L., Qiu, Y. L., & Song, K. (2000). Dynamic evolution of plant mitochondrial genomes: mobile genes and introns and highly variable mutation rates. *Proceedings of the National Academy of Sciences of*

the United States of America, *97*(13), 6960–6966. https://doi.org/10.1073/pnas.97.13.6960

Pannell, J. R. (2002). The Evolution and Maintenance of Androdioecy. *Annual Review of Ecology and Systematics*, *33*(1), 397–425. https://doi.org/10.1146/annurev.ecolsys.33.010802.150419

Pearl, S. A., Welch, M. E., & McCauley, D. E. (2009). Mitochondrial heteroplasmy and paternal leakage in natural populations of *Silene vulgaris*, a gynodioecious plant. *Molecular Biology and Evolution*, *26*(3), 537–545. https://doi.org/10.1093/molbev/msn273

Pelletier, G., & Budar, F. (2007). The molecular biology of cytoplasmically inherited male sterility and prospects for its engineering. *Current Opinion in Biotechnology*, *18*(2), 121–125. https://doi.org/10.1016/j.copbio.2006.12.002

Petit, R. J., Duminil, J., Fineschi, S., Hampe, A., Salvini, D., & Vendramin, G. G. (2005). Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology*, *14*(3), 689–701. https://doi.org/10.1111/j.1365-294X.2004.02410.x

Preuten, T., Cincu, E., Fuchs, J., Zoschke, R., Liere, K., & Börner, T. (2010). Fewer genes than organelles: Extremely low and variable gene copy numbers in mitochondria of somatic plant cells. *Plant Journal*, *64*(6), 948–959. https://doi.org/10.1111/j.1365-313X.2010.04389.x

Pupilli, F., Labombarda, P., & Arcioni, S. (2001). New mitochondrial genome organization in three interspecific somatic hybrids of Medicago sativa including the parent-specific amplification of substoichiometric mitochondrial DNA units.

Theoretical and Applied Genetics, *103*(6–7), 972–978. https://doi.org/10.1007/s001220100663

Rand, D. M. (2001). The Units of Selection of Mitochondrial DNA. *Annual Review of Ecology and Systematics*, Vol. 32, pp. 415–448. https://doi.org/10.1146/annurev.ecolsys.32.081501.114109

Reboud, X., & Zeyl, C. (1994). Organelle inheritance in plants. *Heredity*, *72*(2), 132–140. https://doi.org/10.1038/hdy.1994.19

Renner, S. S. (2014). The relative and absolute frequencies of angiosperm sexual systems: Dioecy, monoecy, gynodioecy, and an updated online database. *American Journal of Botany*, *101*(10), 1588–1596. https://doi.org/10.3732/ajb.1400196

Robison, G. A., Balvin, O., Schal, C., Vargo, E. L., & Booth, W. (2015). Extensive Mitochondrial Heteroplasmy in Natural Populations of a Resurging Human Pest, the Bed Bug (Hemiptera: Cimicidae). *Journal of Medical Entomology*, *52*(4), 734–738. https://doi.org/10.1093/jme/tjv055

Rodríguez-Moreno, L., González, V. M., Benjak, A., Martí, M. C., Puigdomènech, P., Aranda, M. a, & Garcia-Mas, J. (2011). Determination of the melon chloroplast and mitochondrial genome sequences reveals that the largest reported mitochondrial genome in plants contains a significant amount of DNA having a nuclear origin. *BMC Genomics*, *12*(1), 424. https://doi.org/10.1186/1471-2164-12-424

Rokas, A., Ladoukakis, E., & Zouros, E. (2003). Animal mitochondrial DNA recombination revisited. *Trends in Ecology and Evolution*, *18*(8), 411–417. https://doi.org/10.1016/S0169-5347(03)00125-3

Roy, S., & Schreiber, E. (2014). Detecting and Quantifying Low Level Gene Variants in

Sanger Sequencing Traces Using the ab1 Peak Reporter Too. *Journal of Biomolecular Technology*, *25*(May), 2014. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4162221/

Roze, D., Rousse, F., & Michalakis, Y. (2005). Germline bottlenecks, biparental inheritance and selection on mitochondrial variants: A two-level selection model. *Genetics*, *170*(3), 1385–1399. https://doi.org/10.1534/genetics.104.039495

Sabir, J. S. M., Arasappan, D., Bahieldin, A., Abo-Aba, S., Bafeel, S., Zari, T. a., … Jansen, R. K. (2014). Whole mitochondrial and plastid genome SNP analysis of nine date palm cultivars reveals plastid heteroplasmy and close phylogenetic relationships among cultivars. *PLoS ONE*, *9*(4). https://doi.org/10.1371/journal.pone.0094158

Sager, R. (1977). Genetic Analysis of Chloroplast Dna in Chlamydomonas. *Advances in Genetics*, *19*, 287–340. https://doi.org/10.1016/S0065-2660(08)60247-3

Sager, R., & Ishidat, M. R. (1963). Chloroplast DNA in *Chlamydomonas*. *Proceedings of the National Academy of Sciences of the United States of America*, *50*(4), 725–730. https://doi.org/10.1073/pnas.50.4.725

Sakamoto, W. (2003). Leaf-variegated mutations and their responsible genes in Arabidopsis thaliana. *Genes & Genetic Systems*, *78*, 1–9. https://doi.org/10.1266/ggs.78.1

Sakamoto, W., Miyagishima, S., & Jarvis, P. (2008). Chloroplast Biogenesis: Control of Plastid Development, Protein Import, Division and Inheritance. *The Arabidopsis Book*, *62*(1), 1. https://doi.org/10.1199/tab.0110

Scarcelli, N., Mariac, C., Couvreur, T. L. P., Faye, A., Richard, D., Sabot, F., …

Vigouroux, Y. (2016). Intra-individual polymorphism in chloroplasts from NGS data: Where does it come from and how to handle it? *Molecular Ecology Resources*, *16*(2), 434–445. https://doi.org/10.1111/1755-0998.12462

Schattat, M. H., Klösgen, R. B., & Mathur, J. (2012). New insights on stromules: Stroma filled tubules extended by independent plastids. *Plant Signaling & Behavior*, *7*(9), 1132–1137. https://doi.org/10.4161/psb.21342

Scheike, R., Gerold, E., Brennicke, A., Mehring-Lemper, M., & Wricke, G. (1992). Unique patterns of mitochondrial genes, transcripts and proteins in different male-sterile cytoplasms of *Daucus carota*. *Theoretical and Applied Genetics*, *83*(4), 419–427. https://doi.org/10.1007/BF00226529

Schnable, P. S. & Wise, R. P. (1998). The molecular basis of cytoplasmic male sterility and fertility restoration. *Trends in Plant Science*, *3*(5), 175–180. https://doi.org/https://doi.org/10.1016/S1360-1385(98)01235-7

Schultz, S. T. (2009). Leaf Gas Exchange, Water Status, Spatial Dispersion, and Gender in Gynodioecious *Bidens sandvicensis* (Asteraceae). *International Journal of Plant Sciences*, *170*(2), 200–209. https://doi.org/10.1086/593049

Sharpley, M. S., Marciniak, C., Eckel-Mahan, K., McManus, M., Crimi, M., Waymire, K., … Wallace, D. C. (2012). Heteroplasmy of mouse mtDNA is genetically unstable and results in altered behavior and cognition. *Cell*, *151*(2), 333–343. https://doi.org/10.1016/j.cell.2012.09.004

Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., … Sugiura, M. (1986). The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *The EMBO Journal*, *5*(9), 2043–

2049. https://doi.org/10.1002/j.1460-2075.1986.tb04464.x

Sloan, D. B. (2013). One ring to rule them all? Genome sequencing provides new insights into the ’master circle’model of plant mitochondrial DNA structure. *New Phytologist*. https://doi.org/10.1111/nph.12395

Sloan, D. B., Alverson, A. J., Chuckalovcak, J. P., Wu, M., McCauley, D. E., Palmer, J. D., & Taylor, D. R. (2012). Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. *PLoS Biology*, *10*(1), e1001241. https://doi.org/10.1371/journal.pbio.1001241

Sloan, D. B., Alverson, A. J., Wu, M., Palmer, J. D., & Taylor, D. R. (2012). Recent acceleration of plastid sequence and structural evolution coincides with extreme mitochondrial divergence in the angiosperm genus Silene. *Genome Biology and Evolution*, *4*(3), 294–306. https://doi.org/10.1093/gbe/evs006

Smith, S. E., Bingham, E. T., & Fulton, R. W. (1986). Transmission of chlorophyll deficiencies in Medicago sativaEvidence for biparental inheritance of plastids. *Journal of Heredity*, *77*(1), 35–38. Retrieved from http://dx.doi.org/10.1093/oxfordjournals.jhered.a110163

Spreitzer, R. J., & Chastain, C. J. (1987). Heteroplasmic suppression of an amber mutation in the Chlamydomonas chloroplast gene that encodes the large subunit of ribulosebisphosphate carboxylase/oxygenase. *Current Genetics*, *11*(8), 611–616.

Stewart, J. B., & Chinnery, P. F. (2015). The dynamics of mitochondrial DNA heteroplasmy: implications for human health and disease. *Nature Reviews Genetics*, *16*, 530. Retrieved from http://dx.doi.org/10.1038/nrg3966

Stoike, L. L., & Sears, B. B. (1998). Plastome mutator-induced alterations arise in

Oenothera chloroplast DNA through template slippage. *Genetics*, *149*(1), 347–353.

Szklarczyk, M., Szymański, M., Wójcik-Jagła, M., Simon, P. W., Weihe, A., & Börner, T. (2014). Mitochondrial atp9 genes from petaloid male-sterile and male-fertile carrots differ in their status of heteroplasmy, recombination involvement, post-transcriptional processing as well as accumulation of RNA and protein product. *Theoretical and Applied Genetics*, *127*, 1689–1701. https://doi.org/10.1007/s00122-014-2331-x

Takahata, N., & Maruyama, T. (1981). A mathematical model of extranuclear genes and the genetic variability maintained in a finite population. *Genetical Research*, *37*(3), 291–302. https://doi.org/DOI: 10.1017/S0016672300020292

Terasawa, K., Odahara, M., Kabeya, Y., Kikugawa, T., Sekine, Y., Fujiwara, M., & Sato, N. (2007). The mitochondrial genome of the moss physcomitrella patens sheds new light on mitochondrial evolution in land plants. *Molecular Biology and Evolution*, *24*(3), 699–709. https://doi.org/10.1093/molbev/msl198

Touzet, P. (2012). Mitochondrial Genome Evolution and Gynodioecy. *Advances in Botanical Research*, *63*(March), 71. https://doi.org/10.1016/B978-0-12-394279-1.00004-1

Touzet, P., & H. Meyer, E. (2014). Cytoplasmic male sterility and mitochondrial metabolism in plants. *Mitochondrion*, *19*, 166–171. https://doi.org/10.1016/j.mito.2014.04.009

Tsang, W. Y., & Lemire, B. D. (2002). Stable heteroplasmy but differential inheritance of a large mitochondrial DNA deletion in nematodes. *Biochemistry and Cell Biology*, *80*(5), 645–654. https://doi.org/10.1139/o02-135

Unseld, M., Marienfeld, J. R., Brandt, P., & Brennicke, A. (1997). The mitochondrial genome of Arabidopsis thaliana contains 57 genes in 366,924 nucleotides. *Nature Genetics*, *15*(1), 57–61. https://doi.org/10.1038/ng0197-57

van Damme, J. M. M., Hundscheid, M. P. J., Ivanovic, S., & Koelewijn, H. P. (2004). Multiple CMS-restorer gene polymorphism in gynodioecious *Plantago coronopus*. *Heredity*, *93*(2), 175–181. https://doi.org/10.1038/sj.hdy.6800490

Wade, M. J., & McCauley, D. E. (2005). Paternal leakage sustains the cytoplasmic polymorphism underlying gynodioecy but remains invasible by nuclear restorers. *The American Naturalist*, *166*(5), 592–602. https://doi.org/10.1086/491660

Wallace, D. C. (2005). A Mitochondrial Paradigm of Metabolic and Degenerative Diseases, Aging, and Cancer: A Dawn for Evolutionary Medicine. *Annual Review of Genetics*, *39*, 359–407. https://doi.org/10.1146/annurev.genet.39.110304.095751.A

Wallace, D. C., & Chalkia, D. (2013). Mitochondrial DNA genetics and the heteroplasmy conundrum in evolution and disease. *Cold Spring Harbor Perspectives in Medicine*, *3*(10), 1–48. https://doi.org/10.1101/cshperspect.a021220

Wang, C., Du, W., Su, Q. P., Zhu, M., Feng, P., Li, Y., … Yu, L. (2015). Dynamic tubulation of mitochondria drives mitochondrial network formation. *Cell Research*, 1–13. https://doi.org/10.1038/cr.2015.89

Weihe, A., Apitz, J., Pohlheim, F., Salinas-Hartwig, A., & Börner, T. (2009). Biparental inheritance of plastidial and mitochondrial DNA and hybrid variegation in Pelargonium. *Molecular Genetics and Genomics*, *282*(6), 587–593. https://doi.org/10.1007/s00438-009-0488-9

Welch, M. E., Darnell, M. Z., & McCauley, D. E. (2006). Variable populations within

variable populations: Quantifying mitochondrial heteroplasmy in natural populations of the gynodioecious plant Silene vulgaris. *Genetics*, *174*(2), 829–837. https://doi.org/10.1534/genetics.106.059246

Wheeler, G. L., Dorman, H. E., Buchanan, A., Challagundla, L., & Wallace, L. E. (2014). A Review of the Prevalence, Utility, and Caveats of Using Chloroplast Simple Sequence Repeats for Studies of Plant Biology. *Applications in Plant Sciences*, *2*(12), 1400059. https://doi.org/10.3732/apps.1400059

White, D. J., Wolff, J. N., Pierson, M., & Gemmell, N. J. (2008). Revealing the hidden complexities of mtDNA inheritance. *Molecular Ecology*, *17*(23), 4925–4942. https://doi.org/10.1111/j.1365-294X.2008.03982.x

Wolfe, A. D., & Randle, C. P. (2004). Recombination, Heteroplasmy, Haplotype Polymorphism, and Paralogy in Plastid Genes: Implications for Plant Molecular Systematics. *Systematic Botany*, *29*(4), 1011–1020. https://doi.org/10.1600/0363644042451008

Wolfe, K. H., Li, W.-H., & Sharp, P. M. (1987). Rates of Nucleotide Substitution Vary Greatly among Plant Mitochondrial, Chloroplast, and Nuclear DNAs. *Proceedings of the National Academy of Sciences of the United States of America*, *84*(24), 9054–9058. https://doi.org/10.1073/pnas.84.24.9054

Wolff, J. N., Ladoukakis, E. D., Enríquez, J. A., Dowling, D. K., Enriquez, J. A., & Dowling, D. K. (2014). Mitonuclear interactions: evolutionary consequences over multiple biological scales. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(May), 20130443–20130443. https://doi.org/10.1098/rstb.2013.0443

Woloszynska, M. (2010). Heteroplasmy and stoichiometric complexity of plant mitochondrial genomes-though this be madness, yet there's method in't. *Journal of Experimental Botany*, *61*(3), 657–671. https://doi.org/10.1093/jxb/erp361

Wu, Z., Cuthbert, J. M., Taylor, D. R., & Sloan, D. B. (2015). The massive mitochondrial genome of the angiosperm Silene noctiflora is evolving by gain or loss of entire chromosomes. *Proc. Natl. Acad. Sci. USA*, *112*(33), 10185–10191. https://doi.org/10.1073/pnas.1421397112

Xiong, H., Barker, S. C., Burger, T. D., Raoult, D., & Shao, R. (2013). Heteroplasmy in the Mitochondrial Genomes of Human Lice and Ticks Revealed by High Throughput Sequencing. *PLoS ONE*, *8*(9), 1–9. https://doi.org/10.1371/journal.pone.0073329

Xu, J. (2005). The inheritance of organelle genes and genomes: patterns and mechanisms. *Genome*, *48*(6), 951–958. https://doi.org/10.1139/g05-082

Yamagata, Y., Doi, K., Yasui, H., & Yoshimura, A. (2007). Identification of mutants for abnormal pollen development in rice. *Breeding Science*, *57*(4), 331–337. https://doi.org/10.1270/jsbbs.57.331

Yamato, K. T., & Newton, K. J. (1999). Heteroplasmy and Homoplasmy for Maize Mitochondrial Mutants: A Rare Homoplasmic nad4 Deletion Mutant Plant. *Journal of Heredity*, *90*(3), 369–373. https://doi.org/10.1093/jhered/90.3.369

Yampolsky, C., & Yampolsky, H. (1922). Distribution of sex forms in the phanerogamic flora. *Bibliotecha Genetica*, *3*, 1–62.

Zaegel, V., Guermann, B., Le Ret, M., Andrés, C., Meyer, D., Erhardt, M., … Imbault, P. (2006). The Plant-Specific ssDNA Binding Protein OSB1 Is Involved in the

Stoichiometric Transmission of Mitochondrial DNA in Arabidopsis. *The Plant Cell Online*, *18*(12), 3548–3563. https://doi.org/10.1105/tpc.106.042028

Zeng, L., Zhang, Q., Sun, R., Kong, H., Zhang, N., & Ma, H. (2014). Resolution of deep angiosperm phylogeny using conserved nuclear genes and estimates of early divergence times. *Nature Communications*, *5*, 1–12. https://doi.org/10.1038/ncomms5956

Zhang, Q., Liu, Y., & Sodmergen. (2003). Examination of the Cytoplasmic DNA in Male Reproductive Cells to Determine the Potential for Cytoplasmic Inheritance in 295 Angiosperm Species. *Plant and Cell Physiology*, *44*(9), 941–951. https://doi.org/10.1093/pcp/pcg121

Zhang, Q., & Sodmergen. (2010). Why does biparental plastid inheritance revive in angiosperms? *Journal of Plant Research*, *123*(2), 201–206. https://doi.org/10.1007/s10265-009-0291-z

Zhu, A., Guo, W., Jain, K., & Mower, J. P. (2014). Unprecedented heterogeneity in the synonymous substitution rate within a plant genome. *Molecular Biology and Evolution*, *31*(5), 1228–1236. https://doi.org/10.1093/molbev/msu079

Zoschke, R., Liere, K., & Börner, T. (2007). From seedling to mature plant: Arabidopsis plastidial genome copy number, RNA accumulation and transcription are differentially regulated during leaf development. *Plant Journal*, *50*(4), 710–722. https://doi.org/10.1111/j.1365-313X.2007.03084.x

CHAPTER 3: MITOCHONDRIAL AND PLASTID HETEROPLASMY, PHENOTYPIC DIFFERENTIATION, AND FITNESS IN WILD CARROT, *DAUCUS CAROTA*

INTRODUCTION

Organelles are traditionally considered to be uniparentally-inherited, yet this is often not the case in animals (Barr et al., 2005; Ladoukakis and Zouros, 2017) or in plants (Kubo and Newton, 2008; McCauley and Olson, 2008). Instances of some level of biparental inheritance of differing organellar genomes (see Figure 1 in Ramsey and Mandel, 2019) can create within individual organellar genetic variation, i.e., heteroplasmy. Research into the effects of heteroplasmy on phenotype have often been performed using heteroplasmic variants with known deleterious effects (Ye et al., 2014; Chen et al., 2015; Lin et al., 2016), particularly in humans (Chinnery and Turnbull, 1999). The putative negative effects of heteroplasmy are suggested to be an evolutionary mechanism to maintain uniparental inheritance (Birky, 1995); however, the occurrence of heteroplasmy across many diverse taxa and its persistence in populations and across generations (Chinnery et al., 2000; Taylor and Breden, 2002; Jayaprakash et al., 2014) may indicate that heteroplasmy could confer neutral or beneficial effects on phenotype (James and Ballard, 2003; McCauley, 2013).

While the mitochondrial and plastid genomes of plants contain genes essential for metabolic functions, synonymous heteroplasmic variants within genes or those within intergenic spaces should have neutral phenotypic effects. Given the large sizes of mitochondrial and plastid genomes in plants relative to their gene content, most

heteroplasmic sites might be expected to be found within intergenic regions. Furthermore, the high levels of recombination found within these genomes can create additional heteroplasmic variants via recombination between two haplotypes. This simulates the four-gamete rule of sexual genomes quite well (Wright et al., 2008; Mandel et al., 2012). Heteroplasmic variants are not necessarily purged through the germline bottleneck (Aanen et al., 2014), particularly if some level of biparental inheritance also occurs (often termed maternal or paternal "leakage") which would aid in the maintenance of heteroplasmy across generations (Roze et al., 2005). As such, heteroplasmy as an overall phenomenon may be more common than once believed.

A wealth of knowledge has been gained over the past several decades about the detrimental effects associated with specific heteroplasmic variants (Chinnery and Turnbull, 1999; Gorman et al., 2015), yet little is known about the possible beneficial effects of heteroplasmy when it occurs within genes and when variants are non-synonymous (Lane, 2011). An oft-cited example is found in gynodioecious species. Some individuals of these species have mitochondrial haplotypes which cause male sterility (Barr et al., 2005; Case et al., 2016). Theoretical (Delph et al., 2007; Touzet, 2012; McCauley, 2013) and empirical (De Cauwer et al., 2012) studies have demonstrated that females gain a selective advantage over hermaphrodites. Thus, male-sterile individuals, functionally female, gain an increase in fitness as 1) resources can be directed solely to seed production, rather than both seed and pollen production and 2) self-compatible hermaphrodites are prevented from selfing and potentially producing substandard offspring (Touzet, 2012).

However, when populations become pollen-limited, selection will favor hermaphrodites containing nuclear restorer alleles, and balancing selection between the CMS and male-restored hermaphrodite phenotypes becomes perceptible. It has been suggested that selection for paternal leakage may arise as leakage of organelles could restore male fertility in offspring. Pollen produced in male-restored hermaphrodites contain a complementary set of CMS-restorer alleles in the two genomes. Given uniparental inheritance, the association may be disrupted as organellar genes from the pollen would not be transmitted to offspring. Evolutionary forces acting on the nuclear genome could permit (or favor) paternal leakage (Burt and Trivers, 2006; McCauley, 2013). Offspring between CMS and male-restored hermaphrodites would be male-fertile and produce viable pollen, giving the nuclear genome a greater dispersal range than nuclear genomes found in seed alone. Thus, the nuclear restorer allele would increase in frequency until the population is no longer pollen-limited and selection will then again favor CMS (Burt and Trivers, 2006).

Although this example illustrates leakage of specific CMS loci, if leakage does occur, the entire genome should be leaked. Selection for organellar leakage and the generation of heteroplasmy of CMS loci requires the generation of heteroplasmy across the rest of the genome where haplotypes differ from one another. Creating potential genome-wide heteroplasmy in this way raises further questions about its effect on fitness-related traits. Positive and negative selection may indeed act on phenotypes driven by differences at other loci. Furthermore, when considering epistatic cytonuclear interactions and, particularly, the added complexity of interactions between all three plant genomes (Burton et al., 2013), it seems reasonable that each level of selection may occur.

Organellar heteroplasmy may be the evolutionary analog to nuclear heterozygosity. Just as some detrimental phenotypes are associated with nuclear heterozygosity (Prout, 1952; Piálek and Barton, 1997), heterozygosity in a general sense is not considered a predictor of disadvantageous fitness-related phenotypes (David, 1998; Szulkin et al., 2010). In a similar manner, heteroplasmy in a general sense need not be considered a detriment to fitness. Nuclear genomes contain at least two copies (potentially heterozygous) of each gene associated with organellar genes and this may necessitate a corresponding state of organellar heteroplasmy. The concept underlying heterozygosity-fitness correlations (HFC) may be also underlie heteroplasmy-fitness correlations (H_pFC).

To address whether heteroplasmy correlates with fitness-related traits, we performed a common-garden greenhouse study using the gynodioecious species *D. carota*. We used one mt locus and one plastid (pt) locus as proxies for heteroplasmy across each organellar genome. The heteroplasmic variants of each locus are presumably selectively neutral. The pt locus is intergenic, located between *trnS* and *trnG* (called StoG here). The mt gene *cox1* contains an intragenic synonymous substitution. We categorized heteroplasmy for both organellar loci, and quantified the level of *cox1* heteroplasmy. We suggest that heteroplasmy may benefit fitness. If heteroplasmy is correlated with certain fitness-related traits, trait differences will be apparent between heteroplasmic and non-heteroplasmic (homoplasmic) individuals. Additionally, as penetrance of a phenotype may depend on the relative level of heteroplasmy (threshold effect), fitness-related traits will have a relationship with the level of *cox1* heteroplasmy.

METHODS

Study Species–

One common gynodioecious species is *Daucus carota*, with a native distribution around the Mediterranean Sea and in Southwest Asia and Europe. It has been introduced the world over and is considered invasive in many locations where it has the potential to disrupt native ecology (Ramsey et al., in prep). Several mt haplotypes are associated with CMS in this species. Male-sterile phenotypes include aborted pollen, deformed anthers (brown anther), and anthers that develop into either petals (petaloid), carpels (carpeloid), or sepals (sepaloid), each of which may be associated with deformed petals (Nothnagel et al., 2000). CMS in this species has been associated with the mitochondrial genes *atp9* (Szklarczyk et al., 2014) and *atp6* (Tan et al., 2018). The species is often cited as being biennial, yet it may be more accurate to describe it as a monocarpic perennial, as annual, biennial, and triennial life histories have been documented, and a geographic pattern to its annuality may exist (Gross, 1981; de Jong et al., 2016). It forms a basal rosette during its first year of growth. After undergoing vernalization it develops an upright stem (bolts), flowers, and dies. The main stem ends in a terminal primary compound umbel. Secondary umbels branch from the main stem, and this pattern continues until death, sometimes resulting in sixth order umbels (pers. obs.). Sometimes the umbels contain a central purple floret (Polte and Reinhold, 2013), while all other florets are generally white (sometimes light pink). The height is often reported as 1-2 m (Stolarczyk and Janick, 2011; Stachler and Kells, 2014), but a large degree of variation exists (pers. obs.).

Sampling and experimental design–

We selected seeds from five field-collected accessions in the United States, in addition to 14 collections of *D. carota* seed accessioned at the United States Department of Agriculture's Germplasm Resources Information Network (USDA-GRIN; Table 1). These accessions were selected to encompass a large portion of its native range (Europe, North Africa, and Southwest Asia) and a portion of its introduced range, North America. Aside from geographic origin, we randomly sampled accessions, with no prior knowledge of their heteroplasmy status or phenotypic variation. In September 2015, 30 to 35 seeds of each accession were hand-rinsed and placed in petri dishes on filter paper. The seeds were kept at room temperature and moistened with a solution of 1% Banrot (Everris NA Inc., Dublin, Ohio) to prevent the growth of mold until germination occurred, at which point the seedlings were planted in generic all-purpose potting mix (BWI, Memphis, Tennessee) in 4-in pots. The plants were grown at the South Campus Greenhouse at the University of Memphis, in Memphis, Tennessee with automated daily watering. *Daucus carota* requires vernalization to bolt and flower, so the greenhouse was exposed to the ambient winter conditions as much as possible. Heaters were used on occasion to prevent freezing of pipes in sub-freezing temperatures. All tissue from seedlings that did not survive to the five-leaf stage was collected for DNA extraction. Once seedlings had reached the five-leaf stage, the fifth leaf was collected for DNA extraction, and the seedlings were transplanted into 9-in pots for the remainder of the study. DNA was extracted from the fifth leaf with the SQ Omega Biotek Plant Kit (Atlanta, Georgia, USA) with the addition of 1% Ascorbic Acid and 1% polyvinyl

propylene to the SQ1 Buffer provided in the kit and performed a cleanup step with the E.Z.N.A. Cycle Pure Kit from Omega Biotek.

Fitness traits–

The dates of germination and appearance of each true basal leaf were recorded until the plant had developed 50 leaves or had bolted. We calculated the number of days to germination from the date the seeds were placed in petri dishes and the number of days to each developmental stage from the date of germination (i.e., the date of planting). The average number of days between each basal leaf were calculated (i.e., not leaves on upright bolt) over time as a rate of growth. Of those individuals that bolted, we recorded overall height and the number of umbels and calculated the number of days to death of each individual based on whether they continued to develop new leaves once all others had died, or the individual had bolted and all leaves and stems had died.

To assess pollen viability, we first either recorded that the individual was male-sterile by the presence of clear CMS phenotypes (altered morphology of stamen) or, if male-sterility was suspected, collected an umbellet from the primary umbel once anthers were exserted and anthesis thought to have occurred, should the individual have been male-fertile. Such anthers were dissected and viewed under light microscopy. The samples were fixed in 95% ethanol and stored at -20°C. They were scored as having the brown anther phenotype of male-sterility if anthers and/or pollen were not fully developed. Umbellets from male-fertile individuals were collected once anthers were exserted and dehisced. Following the simplified version of the Alexander staining method detailed in Peterson et al. (2010), the pollen and anthers were stained with the modified Alexander stain (1 ml of 1% Malachite green [in 95% ethanol] + 5 ml of 1% Acid fuchsin

[in DI H_2O] + 0.5 ml of 1% Orange G [in DI H_2O] + 4 ml glacial acetic acid + 10 ml of 95% ethanol + 25 ml of 100% glycerol + 54.5 ml of distilled H_2O) on a glass slide and slowly heated it over an alcohol flame for 30 seconds. A cover slip was placed on the sample, and we applied pressure to ensure the sample was in one plane. The slides were sealed with nail varnish, observed under light microscopy, and the ratio of aborted (blue-green) to non-aborted (red) pollen was calculated.

Quantitative PCR mixture experiment and mitochondrial heteroplasmy–

All DNA samples were diluted to 20 ng/µl after quantifying the DNA concentration with a NanoDrop 2000 Spectrophotometer (Thermo Scientific). *cox1* heteroplasmy was assayed using a quantitative PCR (qPCR) assay that can detect heteroplasmy based on the amplitude of two separate fluorophores in the same assay (see Mandel & McCauley [2015] for details on the *cox1* assay and probe set). The qPCR was carried out in a total reaction volume of 12.5 µl (6.25 µl TaqMan Genotyping Master Mix [Applied Biosystems, Foster City, California, USA], 0.625 µl of *cox1* qPCR assay, 1 µl of DNA with the remainder of water. A CFX96 Real-Time System (Bio-Rad Laboratories, Inc., Hercules, California, USA) was used with conditions of 2 min at 50°C and 10 min at 95°C, prior to 40 cycles of 15 s at 95°C and 1 min at 60°C, with each cycle ending in a plate read.

A synonymous T/C single nucleotide polymorphism (SNP) in the coding region was assayed and scored for each individual as having the T or C allele based on their cycle threshold (Ct). The presence of both alleles amplifying sufficiently enough to cross their respective Ct values may indicate the presence of both T and C alleles. To determine the mathematical relationship between the relative quantity of each allele, and thereby

quantify the level of heteroplasmy, a mixture experiment was performed following the approach of Mandel and McCauley (2015). Individuals were selected that were apparently homoplasmic, based on the presence of Ct values for only the T or C allele. The DNA of these individuals were mixed together to generate known relative concentrations of C to T: 100:0, 95:5 90:10, 75:25, 50:50, 25:75, 10:90, 5:95, 0:100. The Ct value of each fluorophore was adjusted based on the observed absence of the T and C alleles in the concentrations of 100:0 and 0:100, respectively. A standard curve of the resulting Ct data was generated in a linear regression in JMP Pro 14.0.0 (SAS Institute, 2018). qPCR on all individuals was performed in duplicate, then the same cycle thresholds were applied to each run. The resulting Ct values were entered in the equation derived from the mixture experiment to determine the relative amount of the C or T allele, yielding a heteroplasmy score. The two heteroplasmy scores for each individual were averaged. Although some individuals had Ct values for each fluorophore (for instance, appearing late in the qPCR reaction, >37 cycles) and appeared to be heteroplasmic, once those Ct values were entered into the equation, only those individuals with heteroplasmy scores between 0.5% and 99.5% were classified as heteroplasmic. Heteroplasmy scores < 0.5% or higher than 99.5% were classified as homoplasmic. To determine whether the means and variances around the mean level of heteroplasmy varied between accessions, Levene's and Welch's tests were performed. Heteroplasmy scores were further classified into the haplotypes of C (0 – 0.5%), Ct (0.5 – 50%), Tc (50 – 99.5%), and T (99.5 – 100%), where Ct and Tc indicate major C and T and minor t and c alleles, respectively.

Plastid heteroplasmy–

A length variant in the intergenic space between the *trnS* and *trnG* genes ('StoG' region) of the chloroplast genome was amplified using the 'S' and 'G' primers from Hamilton (1999). The PCR was carried out in a total reaction volume of 17 µl (11 µl DI H_2O, 1.5 µl of 10X PCR buffer [100 mM KCl, 100 mM Tris HCl (pH 9.0), 80 mM $(NH_4)_2SO_4$, and 1.0% Triton X-100], 0.7 µl $MgCl_2$ (25 mM), 0.2 µl dNTPs (20 mM), 0.2 µl forward primer (5 µM), 0.2 µl reverse primer (20 µM), 0.2 µl fluorescently labeled M13 primer (10 µM), 1 µl of Taq DNA polymerase, and 2 µl template DNA (20 ng/µl). The PCR conditions were: 3 min at 95°C; 10 cycles of 30 s at 94°C, 30 s at 65°C and 45 s at 72°C, annealing temperature decreasing to 54°C by 1°C per cycle, followed by 30 cycles of 30 s at 94°C, 30 s at 55°C, 45 s at 72°C, followed by 10 min at 72°C. The PCR amplicons were then diluted 1:10 and visualized using an ABI 3130xl DNA sequencer (Applied Biosystems) with GeneScan 500 LIZ dye Size Standard (Applied Biosystems) included in each lane to allow for accurate fragment size determination. Alleles were scored using GeneMarker v. 2.6.3 (SoftGenetics, State College, Pennsylvania). Heteroplasmic individuals were scored based on the presence of more than one allele. Heteroplasmy scores were not calculated as we are unable to quantify the major and minor alleles based on fragment size analysis.

Correlation of heteroplasmy and fitness traits–

With a heteroplasmic StoG sample size of just 16, we were unable to perform comparisons between heteroplasmic and homoplasmic individuals. Therefore, heteroplasmy-fitness correlations were performed for only *cox1*. A Principle Components Analysis was performed to determine correlations between the nine traits we recorded or

calculated: days to germination, days to first leaf, days to bolt, days to death, growth rate, number of leaves, number of umbels, height, and percent pollen viability. We then performed a series of Kaplan-Meier survival analyses to the life stages of germination, first leaf, bolt, and death. Survival analyses perform Log-Rank and Wilcoxon tests to test the hypothesis that survival between groups is the same. First, individuals were grouped as whether they lived 1, 2, or 3 years (annuality), grouped as whether they were heteroplasmic or homoplasmic (heteroplasmic state), and grouped as their haplotype (either C, Ct, Tc, or T). Individuals were then grouped as their heteroplasmic state by accession, grouped as heteroplasmic state by annuality, grouped as haplotype by accession, and grouped as haplotype by annuality.

To determine if there is a threshold effect of heteroplasmy at which phenotypes change, we ran a series of linear regressions. Each trait (Y) was regressed by the heteroplasmy score (X) of all individuals combined, and a separate series of the same regressions were performed but with individuals divided into annual, biennial, and triennial. An additional curvilinear regression was performed in the same manner with all individuals combined.

Finally, to access any differences between heteroplasmic and homoplasmic individuals, a series of one-way analyses of variance (ANOVA) was performed with annuality, accession, male sterility, haplotype, heteroplasmy state, and heteroplasmy score as separate predictors of the nine fitness-related traits. We also performed two-way ANOVAs to test for interactions of annuality and heteroplasmy score, annuality and heteroplasmy state, and accession and heteroplasmy state. ANOVAs with significant p-values were followed by either Tukey's Honestly Significant Difference (HSD) post hoc

test or a Student's t test. We used JMP Pro 14.0.0 (SAS Institute, 2018) to perform all analyses.

RESULTS & DISCUSSION

Fitness Traits–

We plated approximately 600 seeds from 19 accessions on petri dishes (Table 1). The number of germinated seeds differed between accessions, from a low of 7 seeds of accession Gravel.2009 from New York State, to a high of 33 from accession Ames31580 from Nador, Morocco. In total, 514 (~87.5%) of all seeds successfully germinated and were planted in pots. Forty-five of these died before reaching the first true leaf stage, and 279 of the remaining 469 died before bolting. An additional 48 bolted, but did not survive to successfully flower. The remaining 142 individuals flowered, and we were able to record all nine fitness-related traits for them (Table 1).

Overall, the mean days to germination, first leaf, bolt, and death were 10.22 ± 0.69, 23.7 ± 0.8, 329.89 ± 12.17, and 311.57 ± 8.18, respectively. Mean growth rate as the number of days between each leaf was 17.1 ± 0.31. The mean numbers of leaves and umbels were 17.44 ± 0.58 and 7.08 ± 0.19, respectively, and the mean height of the individuals that bolted was 123.74 ± 2.72 cm. Of the individuals that successfully flowered, percent pollen viability averaged 60.41 ± 3.55. All nine fitness-related characteristics significantly varied by accession, heteroplasmy state, life history, and haplotype (Tables 1-4).

Heteroplasmy Assay and Mixture Experiment–

The *cox1* heteroplasmy mixture experiment resulted in the equation y = 2.384x - 0.8639 with an R^2 of 0.9659. Of 514 individuals that germinated (Table 1), 163 and 276

were scored as heteroplasmic and homoplasmic, respectively, based on our pre-defined value of 0.05%.

Results of the one-way ANOVA to test for equality of the means and variances of the *cox1* heteroplasmy score revealed that accessions Ames24710 and PI652290 did not have means and variances equal to the rest of the accessions (Welch's test, F = 7.4694, DF = 18, P = <0.0001; Levene's test, F =23.2420, DF = 18, *P* = <0.0001, respectively).

Correlations of Fitness Traits and ANOVA Summaries–

Nine fitness-related traits were recorded, and several of these show high correlations with one another, as well as a number of results in the ANOVAs (Table 5). Days to germination had a highly positive correlation coefficient with days to first leaf (*r* = 0.9598), yet it had slight to moderately negative correlation coefficients with all other dependent variables. Days to bolt was highly coordinated with days to death and number of leaves (*r* = 0.9912 and 0.8517, respectively). The most negative correlation coefficients were between days to germination and pollen viability and days to first leaf and pollen viability (*r* = -0.4546 and -0.4564, respectively. The number of leaves had a highly positive correlation coefficient with days to death (*r* = 0.8534), but a moderately negative coefficient with growth rate (*r* = -0.4568).

Table 1. Geographic origins and regions, sample sizes, and means ± standard errors of nine fitness-related traits in 19 accessions of *Daucus carota*.

Accession	Origin	Region	N	Days to Germination	Days to 1st Leaf	Days to Bolt	Days to Death	Growth Rate	No. Leaves	Umbel No.	Height (cm)	% Pollen Viability
Ames26378	Portalegre, Portugal	Europe	20	6.9 ± 2.61	21.21 ± 3.1	376.82 ± 47.06	301.45 ± 34.94	14.78 ± 1.4	18.05 ± 2.62	6.4 ± 0.58	120.76 ± 6.81	87.27 ± 12.0
Ames27410	Uzbekistan	Middle East	27	7.44 ± 2.51	20.11 ± 2.93	297.08 ± 43.29	299.22 ± 33.62	15.16 ± 1.27	18.59 ± 2.48	6.33 ± 0.61	143.39 ± 6.27	38.91 ± 12.65
Ames30239	Sousse, Tunisia	North Africa	28	7.29 ± 2.19	19.82 ± 2.55	353.54 ± 30.61	419.78 ± 29.65	16.02 ± 1.1	22.43 ± 2.15	7.86 ± 0.39	134.1 ± 4.43	75.82 ± 9.2
Ames30269	Beja, Tunisia	North Africa	30	9.53 ± 2.17	22.3 ± 2.53	373.91 ± 47.06	364.23 ± 29.0	19.14 ± 1.09	17.6 ± 2.14	10 ± 0.61	194.4 ± 6.81	61.26 ± 13.42
Ames31565	Sidi Kacem, Morocco	North Africa	23	9.83 ± 2.59	24.68 ± 3.2	402.67 ± 90.11	231.52 ± 34.69	22.11 ± 1.38	10.55 ± 2.7	6 ± 1.07	105.83 ± 13.05	77.85 ± 26.83
Ames31580	Nador, Morocco	North Africa	33	8.94 ± 2.07	23.52 ± 2.42	337.59 ± 28.98	400.87 ± 28.53	19.12 ± 1.05	20.21 ± 2.04	5.8 ± 0.37	80.01 ± 4.27	32.64 ± 7.91
Bowl.2009	New York State	North America	19	10.32 ± 2.72	23.21 ± 3.18	327.22 ± 52.02	363.25 ± 39.53	13.09 ± 1.37	31 ± 2.68	7.38 ± 0.65	103.01 ± 7.53	82.62 ± 13.42
CIBC.D	Memphis, TN	North America	24	7.21 ± 2.34	24.13 ± 2.73	344.15 ± 43.29	313.79 ± 31.25	18.43 ± 1.18	18.13 ± 2.3	8.38 ± 0.51	130.82 ± 6.27	74 ± 10.52
Gravel.2009	New York State	North America	7	32.57 ± 4.37	44.71 ± 5.1	-	295.5 ± 66.3	16.08 ± 2.21	16.71 ± 4.31	-	-	-
OCNJ.2009	Ocean City, NJ	North America	25	14.52 ± 2.68	29.92 ± 3.15	270 ± 69.8	260.08 ± 36.09	22.34 ± 1.36	11 ± 2.66	5.8 ± 0.82	110.74 ± 10.11	-

Table 1 (continued).

Accession	Origin	Region	N	Days to Germination	Days to 1st Leaf	Days to Bolt	Days to Death	Growth Rate	No. Leaves	Umbel No.	Height (cm)	% Pollen Viability
PI390896	Central Israel	Middle East	21	5 ± 2.5	19.9 ± 3.0	473.29 ± 41.71	425.5 ± 34.41	21.43 ± 1.3	17.75 ± 2.53	9.89 ± 0.61	148.32 ± 6.04	59.05 ± 12.65
PI478874	Sicily, Italy	Europe	16	7.69 ± 2.86	19.31 ± 3.34	286 ± 52.02	314.62 ± 42.59	16.93 ± 1.44	17.13 ± 2.82	7.38 ± 0.65	142.83 ± 7.53	76.15 ± 15.49
PI478877	Geneva, Switzerland	Europe	18	10 ± 2.86	20.33 ± 3.34	228 ± 156.07	255.4 ± 49.42	13.07 ± 1.44	20.94 ± 2.82	4.91 ± 0.56	157.5 ± 22.6	-
PI652213	Kazakhstan	Middle East	26	6.15 ± 2.27	16.38 ± 2.65	-	344.83 ± 36.09	15.33 ± 1.15	20 ± 2.24	-	-	-
PI652290	Szczecin, Poland	Europe	26	9.31 ± 2.66	22.81 ± 3.11	-	255.67 ± 48.42	12.89 ± 1.34	21.38 ± 2.62	-	-	-
PI652303	Central Greece	Europe	28	5.57 ± 2.19	20.67 ± 2.59	235.77 ± 37.85	276.15 ± 29.65	16.87 ± 1.12	12.89 ± 2.18	-	98.02 ± 5.48	51.62 ± 11.44
PI652341	Syria	Middle East	22	7.14 ± 2.74	19.5 ± 3.6	258 ± 58.99	232.13 ± 41.29	16.19 ± 1.56	10.56 ± 3.04	6.75 ± 0.92	116.49 ± 8.54	55.19 ± 18.97
PI652354	Izmir, Turkey	Middle East	25	5 ± 2.55	15.79 ± 3.0	239.9 ± 49.35	279.75 ± 34.38	14.34 ± 1.3	15.42 ± 2.53	6 ± 0.58	134.45 ± 7.15	54.51 ± 12.0
UT.2009	Knoxville, TN	North America	21	44.29 ± 2.5	59.05 ± 2.92	-	182.57 ± 33.45	19.07 ± 1.26	8.33 ± 2.46	-	-	-

Table 2. Sample sizes and means ± standard errors of fitness-related traits of annuals, biennials, and triennials.

Annuality	N	Days to Germination	Days to 1st Leaf	Days to Bolt	Days to Death	Growth Rate	No. Leaves	Umbel No.	Height (cm)	Pollen Viability
Annual	308	11.45 ± 0.97	25.4 ± 1.11	207.6 ± 4.82	231.93 ± 4.11	17.71 ± 0.38	11.7 ± 0.39	6.38 ± 0.21	114.97 ± 3.2	60.05 ± 4.82
Biennial	91	7.32 ± 0.43	20.4 ± 0.7	491.19 ± 8.08	542.62 ± 8.85	16.9 ± 0.61	29.24 ± 1.18	8.05 ± 0.32	135.16 ± 4.6	59.6 ± 5.38
Triennial	40	7.35 ± 0.46	18.65 ± 0.72	757.5 ± 26.93	853 ± 29.66	12.99 ± 0.59	33.15 ± 1.6	9 ± 1.53	159.7 ± 8.45	94.85 ± 0.11

Table 3. Sample sizes and means ± standard errors of fitness-related traits of heteroplasmic, homoplasmic, and all individuals combined.

Heteroplasmic State	N	Days to Germination	Days to 1st Leaf	Days to Bolt	Days to Death	Growth Rate	No. Leaves	Umbel No.	Height (cm)	Pollen Viability
Heteroplasmic	163	9.28 ± 0.74	22.57 ± 0.84	346.7 ± 22.95	313.62 ± 14.44	17.38 ± 0.48	17.09 ± 1	7.14 ± 0.34	121.73 ± 5.05	64.1 ± 5.88
Homoplasmic	276	10.78 ± 1.01	24.38 ± 1.17	319.98 ± 13.83	310.29 ± 9.78	16.93 ± 0.4	17.66 ± 0.71	7.04 ± 0.23	124.91 ± 3.17	58.45 ± 4.46
All	439	10.22 ± 0.69	23.7 ± 0.8	329.89 ± 12.17	311.57 ± 8.18	17.1 ± 0.31	17.44 ± 0.58	7.08 ± 0.19	123.74 ± 2.72	60.41 ± 3.55

Table 4. Sample sizes and means ± standard errors of fitness-related traits of the four mitochondrial haplotypes.

cox1 Haplotype	N	Days to Germination	Days to 1st Leaf	Days to Bolt	Days to Death	Growth Rate	No. Leaves	Umbel No.	Height (cm)	Pollen Viability
C	241	10.72 ± 1.05	24.35 ± 1.21	324.93 ± 14.53	317.5 ± 10.66	17.33 ± 0.44	17.34 ± 0.75	7.09 ± 0.24	125.04 ± 3.21	59.05 ± 4.59
Ct	153	9.29 ± 0.78	22.54 ± 0.9	344.97 ± 23.93	308.55 ± 15.13	17.3 ± 0.5	16.91 ± 1.04	7.15 ± 0.34	122.09 ± 5.24	63.59 ± 6.13
T	33	11.39 ± 3.65	24.91 ± 4.15	245.71 ± 27.2	256.96 ± 19.12	14.15 ± 0.78	20.25 ± 2.08	6.2 ± 0.86	123.01 ± 17.13	48.97 ± 20.2
Tc	12	8.83 ± 1.58	22.17 ± 1.59	373.5 ± 86.31	363.82 ± 42.7	17.98 ± 1.3	18.75 ± 3.33	7 ± 3	116.2 ± 21.44	75.14 ± 12.94

Table 5. ANOVA results from nine fitness traits.

Dependent Variable	Fixed Effect	df	*F*	*P*
Days to Germination	Accession	18	14.1048	**<.0001***
	Annuality	2	3.7696	**0.0238**
	Het. State	1	10.3675	**0.0014***
	Haplotype	3	0.4114	0.7449
	Accession x Het. State	18	2.9158	**<.0001***
Days to First Leaf	Accession	18	11.3809	**<.0001***
	Annuality	2	5.4127	**0.0048**
	Het. State	1	9.278	**0.0025**
	Haplotype	3	0.4591	0.711
	Accession x Het. State	18	2.9345	**<.0001***
Days to Bolt	Accession	12	2.447	**0.0061**
	Annuality	2	143.39	**<.0001***
	Het. State	1	0.0031	0.9553
	Haplotype	3	0.9331	0.426
	Accession x Het. State	12	0.7906	0.6596
Days to Death	Accession	18	4.1382	**<.0001***
	Annuality	2	326.15	**<.0001***
	Het. State	1	1.125	0.2895
	Haplotype	3	1.4765	0.2204
	Accession x Het. State	18	0.5755	0.9166
Growth Rate	Accession	18	4.9746	**<.0001***
	Annuality	2	11.3282	**<.0001***
	Het. State	1	1.3201	0.2513
	Haplotype	3	2.5539	**0.055**

Table 5 (continued).

Dependent Variable	Fixed Effect	df	*F*	*P*
	Accession x Het. State	18	0.9687	0.4955
Number of Leaves	Accession	18	4.3821	**<.0001***
	Annuality	2	241.12	**<.0001***
	Het. State	1	0.0306	0.8612
	Haplotype	3	0.7407	0.5282
	Accession x Het. State	18	0.3842	0.9902
Number of Umbels	Accession	12	7.3764	**<.0001***
	Annuality	2	35.4572	**<.0001***
	Het. State	1	1.7021	0.1928
	Haplotype	3	0.2564	0.8567
	Accession x Het. State	12	1.0023	0.4513
Height	Accession	12	24.0496	**<.0001***
	Annuality	2	43.7422	**<.0001***
	Het. State	1	0.6747	0.4119
	Haplotype	3	0.1433	0.9338
	Accession x Het. State	12	1.873	**0.042**
Pollen Viability	Accession	11	1.9702	**0.0386**
	Annuality	2	0.7339	0.4821
	Het. State	1	0.817	0.3678
	Haplotype	3	0.3349	0.8001
	Accession x Het. State	11	1.1552	0.3273

Results of survival analyses show that annuals take significantly longer to germinate (Log-Rank: $\chi^2 = 12.37$, DF = 2, $P = 0.0021$; Wilcoxon: $\chi^2 = 7.48$, DF = 2, $P =$

0.0238; Table 2) and develop their first true leaf (Log-Rank: χ^2 = 26.91, DF = 2, P < 0.0001; Wilcoxon: χ^2 = 27.90, DF = 2, P = 0.0238) than biennials or triennials. Heteroplasmic individuals in accession Ames26378 take significantly longer to germinate (Log-Rank: χ^2 = 7.94, DF = 2, P = 0.0189; Wilcoxon: χ^2 = 6.99, DF = 2, P = 0.0303; Table 3) and almost significantly longer to develop their first true leaf (Log-Rank: χ^2 = 2.53, DF = 1, P = 0.11; Wilcoxon: χ^2 = 2.73, DF = 1, P = 0.0986) than homoplasmic individuals. Homoplasmic individuals in accession Ames31580 almost take significantly longer to germinate (Log-Rank: χ^2 =3 .27, DF = 1, P = 0.0708; Wilcoxon: χ^2 = 3.27, DF = 2, P = 0.0708) but do take significantly longer to develop their first true leaf (Log-Rank: χ^2 = 4.49, DF = 1, P = 0.0342; Wilcoxon: χ^2 = 3.69, DF = 1, P = 0.0547) compared to heteroplasmic individuals. Homoplasmic individuals in accession PI652290 almost take significantly longer to germinate than heteroplasmic individuals (Log-Rank: χ^2 = 3.3333, DF = 1, P = 0.0679; Wilcoxon: χ^2 = 3.3333, DF = 1, P = 0.0679), whereas heteroplasmic individuals in accession PI652341 take significantly longer to develop their first true leaf compared to homoplasmic individuals (Log-Rank: χ^2 = 5.28, DF = 1, P = 0.0215; Wilcoxon: χ^2 = 3.55, DF = 1, P = 0.0596).

For the later life stages, homoplasmic individuals live significantly longer in accession Ames31565 than heteroplasmic individuals (Log-Rank: χ^2 = 5.85, DF = 1, P = 0.0156; Wilcoxon: χ^2 = 3.51, DF = 1, P = 0.0611). Heteroplasmic individuals in accession Gravel live significantly longer than homoplasmic individuals (Log-Rank: χ^2 = 5.63, DF = 1, P = 0.0177; Wilcoxon: χ^2 = 5.3333, DF = 1, P = 0.0209), while homoplasmic individuals in accession 478874 almost live significantly longer than

heteroplasmic individuals (Log-Rank: $\chi^2 = 2.83$, DF = 1, $P = 0.0923$; Wilcoxon: $\chi^2 = 2.94$, DF = 1, $P = 0.0865$).

Annual heteroplasmic individuals almost have significantly fewer days to death compared to biennials (Log-Rank: $\chi^2 = 3.28$, DF = 1, $P = 0.0702$; Wilcoxon: $\chi^2 = 3.58$, DF = 1, $P = 0.0584$; Table 2). By haplotype, C individuals takes significantly longer to germinate than either T or Ct (Log-Rank: $\chi^2 = 7.94$, DF = 2, $P = 0.0189$; Wilcoxon: $\chi^2 = 6.99$, DF = 2, $P = 0.0303$; Table 4).

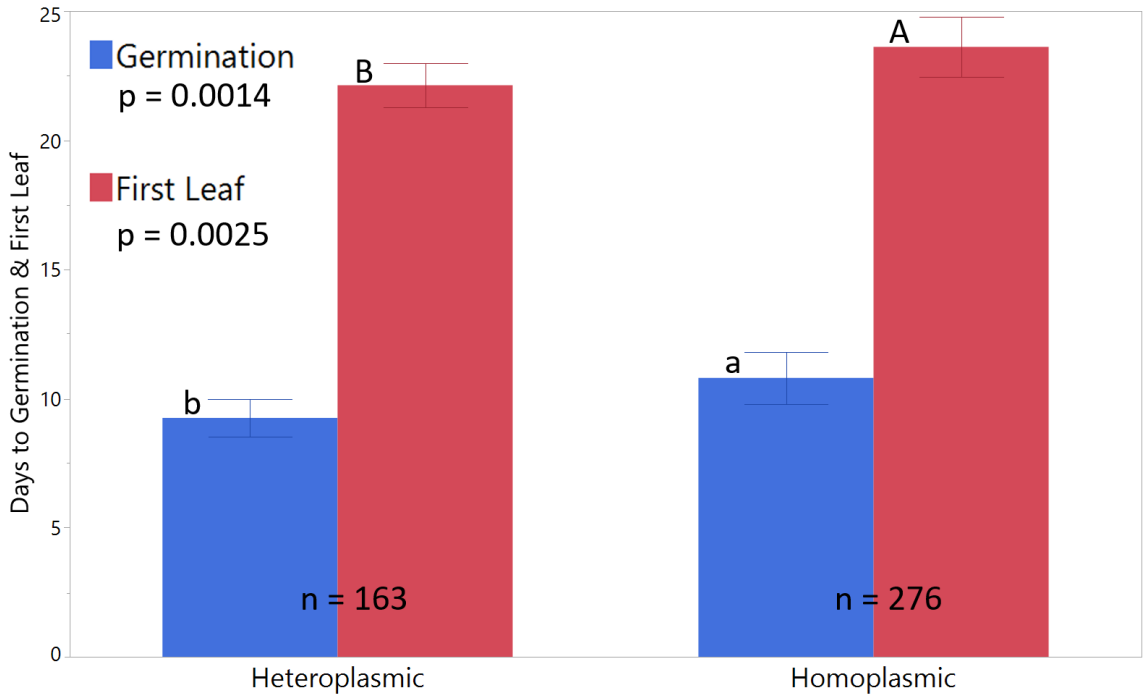


Fig. 1. Days to germination and first leaf for heteroplasmic and homoplasmic individuals. Letters different from one another indicate significantly different days to germination (lowercase) and first leaf (uppercase).

These findings support the notion that fitness-related traits are geographically, and potentially genetically, structured given that variation within accession was lower than among accession. There were several interesting outcomes. In our study annuals took longer to germinate, heteroplasmic individuals germinated faster than homoplasmic individuals, however, this was not an absolute trend, as there were exceptions. Although

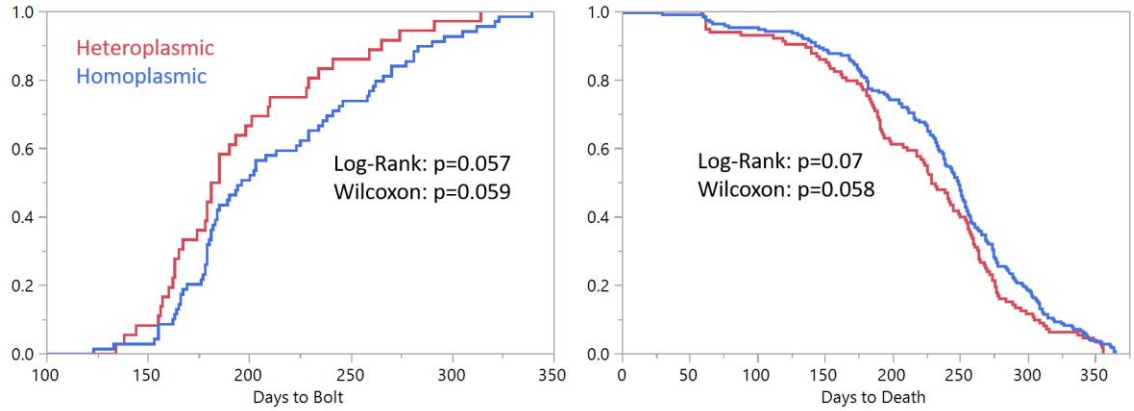


Fig. 2. Survival curves of heteroplasmic and homoplasmic individuals for days to bolt and days to death.

we attempted to grow these in a common garden (greenhouse at the University of Memphis) we cannot rule out the possibility some accessions may have been more adapted to this environment. In sum, this work has demonstrated the interactions among the phenome, mitochondrial genome, location of origin, and life history.

ACKNOWLEDGEMENTS

Daucus carota accessions were provided by USDA-ARS at Ames, Iowa. For greenhouse assistance and lab work, we thank Zachary Ramsey, Nathanael Prather, and Robert Coleman, in addition to several undergraduates: Nicholas Bailey, S. Mike Ballou, Jr., Matthew DeVazier, Amanda Gelsinger, Blair Combest, Jakyra Scaife, and Robby Schranze. We thank Wesley Boyd for microscopy work. We also thank Tom Cunningham at the Molecular Resource Center at the University of Tennessee Health Science Center for assistance collecting data. This work was supported by USDA-NIFA award 2014-33522-21826 to JRM and the W. Harry Feinstone Center for Genomic Research, and to AJR, an Earl Core Student Research Award from the Southern Appalachian Botanical Society, a Graduate Student Research Award from the Botanical

Society of America, and a Bill A. Simco Student Research Award from the University of Memphis Department of Biological Sciences.

REFERENCES

Aanen, D. K., J. N. Spelbrink, and M. Beekman. 2014. What cost mitochondria? The maintenance of functional mitochondrial DNA within and across generations. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369: 20130438.

Barr, C. M., M. Neiman, and D. R. Taylor. 2005. Inheritance and recombination of mitochondrial genomes in plants, fungi and animals. *The New Phytologist* 168: 39–50.

Birky, C. W. 1995. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proceedings of the National Academy of Sciences of the United States of America* 92: 11331–11338.

Burt, A., and R. Trivers. 2006. Genes in Conflict: The Biology of Selfish Genetic Elements. Harvard University Press.

Burton, R. S., R. J. Pereira, and F. S. Barreto. 2013. Cytonuclear Genomic Interactions and Hybrid Breakdown. *Annual Review of Ecology, Evolution, and Systematics* 44: 281–302.

Case, A. L., F. R. Finseth, C. M. Barr, and L. Fishman. 2016. Selfish evolution of cytonuclear hybrid incompatibility in Mimulus. *Proceedings of the Royal Society B: Biological Sciences* 283.

De Cauwer, I., J. F. Arnaud, E. K. Klein, M. Dufay, I. De Cauwer, J. F. Arnaud, E. K.

Klein, and M. Dufay. 2012. Disentangling the causes of heterogeneity in male fecundity in gynodioecious Beta vulgaris ssp. maritima. *New Phytologist* 195: 676–687.

Chen, Z., Y. Qi, S. French, G. Zhang, R. C. Garcia, R. Balaban, and H. Xu. 2015. Genetic mosaic analysis of a deleterious mitochondrial DNA mutation in *Drosophila* reveals novel aspects of mitochondrial regulation and function. *Molecular Biology of the Cell* 26: 674–684.

Chinnery, P. F., D. R. Thorburn, D. C. Samuels, S. L. White, H. H. M. Dahl, D. M. Turnbull, R. N. Lightowlers, et al. 2000. The inheritance of mitochondrial DNA heteroplasmy: random drift, selection or both? *Trends in Genetics* 16: 500–505.

Chinnery, P. F., and D. Turnbull. 1999. Mitochondrial DNA and disease. *The Lancet* 354: S17–S21.

David, P. 1998. Heterozygosity-fitness correlation: new perspectives on old problems. *Heredity* 80: 531–537.

Delph, L. F., P. Touzet, and M. F. Bailey. 2007. Merging theory and mechanism in studies of gynodioecy. *Trends in Ecology and Evolution* 22: 17–24.

Gorman, G. S., A. M. Schaefer, Y. Ng, N. Gomez, E. L. Blakely, C. L. Alston, C. Feeney, et al. 2015. Prevalence of nuclear and mitochondrial DNA mutations related to adult mitochondrial disease. *Annals of Neurology* 77: 753–759.

Gross, K. L. 1981. Predictions of fate from rosette size in four ‘biennial’ plant species: *Verbascum thapsus*, *Oenothera biennis*, *Daucus carota*, and *Tragopogon dubius*. *Oecologia* 48: 209–213.

Hamilton, M. 1999. Four primer pairs for the amplification of chloroplast intergenic

regions with intraspecific variation. *Molecular Ecology* 8: 521–523.

James, A. C., and J. W. O. Ballard. 2003. Mitochondrial genotype affects fitness in *Drosophila simulans*. *Genetics* 164: 187–194.

Jayaprakash, A., E. Benson, and R. Liang. 2014. Heteroplasmy is ubiquitous and stable at the single cell level. *bioRxiv*: 0–15.

de Jong, T. J., C. Grebenstein, and W. L. M. Tamis. 2016. Demography and life history of *Daucus carota* L. populations in the Netherlands. *Flora: Morphology, Distribution, Functional Ecology of Plants* 224: 154–158.

Kubo, T., and K. J. Newton. 2008. Angiosperm mitochondrial genomes and mutations. *Mitochondrion* 8: 5–14.

Ladoukakis, E. D., and E. Zouros. 2017. Evolution and inheritance of animal mitochondrial DNA: Rules and exceptions. *Journal of Biological Research-Thessaloniki* 24: 1–7.

Lane, N. 2011. Mitonuclear match: optimizing fitness and fertility over generations drives ageing within generations. *Bioessays*: 860–869.

Lin, Y.-F., A. M. Schulz, M. W. Pellegrino, Y. Lu, S. Shaham, and C. M. Haynes. 2016. Maintenance and propagation of a deleterious mitochondrial genome by the mitochondrial unfolded protein response. *Nature* 533: 1–8.

Mandel, J. R., E. V. McAssey, K. M. Roland, and D. E. McCauley. 2012. Mitochondrial gene diversity associated with the atp9 stop codon in natural populations of wild carrot (*Daucus carota* ssp. *carota*). *Journal of Heredity* 103: 418–425.

Mandel, J. R., and D. E. McCauley. 2015. Pervasive Mitochondrial Sequence Heteroplasmy in Natural Populations of Wild Carrot, *Daucus carota* spp. *carota* L.

Plos One 10: 1–12.

McCauley, D. E. 2013. Paternal leakage, heteroplasmy, and the evolution of plant mitochondrial genomes. *New Phytologist* 200: 966–977.

McCauley, D. E., and M. S. Olson. 2008. Do recent findings in plant mitochondrial molecular and population genetics have implications for the study of gynodioecy and cytonuclear conflict? *Evolution* 62: 1013–1025.

Nothnagel, T., P. Straka, and B. Linke. 2000. Male sterility in populations of *Daucus* and the development of alloplasmic male-sterile lines of carrot. *Plant Breeding* 119: 145–152.

Peterson, R., J. P. Slovin, and C. Chen. 2010. A simplified method for differential staining of aborted and non-aborted pollen grains. *International Journal of Plant Biology* 1: 66–69.

Piálek, J., and N. H. Barton. 1997. The Spread of an Advantageious Allele across a Barrier: The Effects of Random Drift and Selection against Heterozygotes. *Genetics* 145: 493–507.

Polte, S., and K. Reinhold. 2013. The function of the wild carrot's dark central floret: Attract, guide or deter? *Plant Species Biology* 28: 81–86.

Prout, T. 1952. Selection Against Heterozygotes for Autosomal Lethals in Natural Populations of *Drosophila willistoni*. *Proceedings of the National Academy of Sciences* 38: 478–481.

Ramsey, A. J., and J. R. Mandel. 2019. When One Genome Is Not Enough: Organellar Heteroplasmy in Plants. *Annual Plant Reviews Online* 2: 1–40.

Roze, D., F. Rousse, and Y. Michalakis. 2005. Germline bottlenecks, biparental

inheritance and selection on mitochondrial variants: A two-level selection model. *Genetics* 170: 1385–1399.

SAS Institute. 2018. Discovering JMP 14.0.0.

Stachler, J. M., and J. J. Kells. 2014. Wild Carrot (*Daucus carota*) Control in No-Tillage Cropping Systems. *Weed Technology* 11: 444–452.

Stolarczyk, J., and J. Janick. 2011. Carrot: History and Iconography. *Chronica Horticulturae* 51: 13–18.

Szklarczyk, M., M. Szymański, M. Wójcik-Jagła, P. W. Simon, A. Weihe, and T. Börner. 2014. Mitochondrial atp9 genes from petaloid male-sterile and male-fertile carrots differ in their status of heteroplasmy, recombination involvement, post-transcriptional processing as well as accumulation of RNA and protein product. *Theoretical and Applied Genetics* 127: 1689–1701.

Szulkin, M., N. Bierne, and P. David. 2010. Heterozygosity-fitness correlations: A time for reappraisal. *Evolution* 64: 1202–1217.

Tan, G.-F., F. Wang, X.-Y. Zhang, and A.-S. Xiong. 2018. Different lengths, copies and expression levels of the mitochondrial atp6 gene in male sterile and fertile lines of carrot (*Daucus carota* L.). *Mitochondrial DNA Part A* 29: 446–454.

Taylor, J. S., and F. Breden. 2002. The inheritance of heteroplasmy in guppies. *Journal of Fish Biology* 60: 1346–1350.

Touzet, P. 2012. Mitochondrial Genome Evolution and Gynodioecy. *Advances in Botanical Research* 63: 71.

Wright, S. I., N. Nano, J. P. Foxe, and V.-U. N. Dar. 2008. Effective population size and tests of neutrality at cytoplasmic genes in *Arabidopsis*. *Genetics research* 90: 119–

28.

Ye, K., J. Lu, F. Ma, A. Keinan, and Z. Gu. 2014. Extensive pathogenicity of mitochondrial heteroplasmy in healthy human individuals. *Proceedings of the National Academy of Sciences of the United States of America* 111: 10654–9.

CHAPTER 4: HETEROPLASMY AND PATTERNS OF CYTONUCLEAR LINKAGE DISEQUILIBRIUM IN WILD CARROT

Introduction

Studying interactions between organellar (mitochondrial, mt; plastid, pt) and nuclear genomes has the potential to yield new insights into evolutionary processes including selection, demography, admixture, and cytonuclear incompatibilities (Wade and Goodnight 2006; Sambatti et al. 2008; Brandvain and Wade 2009; Fields et al. 2014; Wade and Drown 2016). Phenotypic variation and fitness may be influenced by cytonuclear epistasis and has been shown in a variety of organisms including humans, yeast, *Drosophila*, and various plant species (Tao et al. 2004; Zeyl et al. 2005; Rand et al. 2006; Camus et al. 2017). For the mt genome (mtDNA), evidence for mitonuclear functional coordination has been demonstrated and is postulated to have evolved via cooperative evolution between nuclear genes that encode proteins targeted to the mt and mt genes (Blier et al. 2001; Sloan et al. 2015). Similarly, accelerated coevolution in pt-encoded ribosomal genes and plastid–nuclear interactions has been shown in some plant species, as evidenced by a survey of substitution rates in the phylogeny of Geraniaceae (Weng et al. 2016).

Cytonuclear Linkage Disequilibrium

Cytonuclear linkage disequilibrium (cnLD) describes the non-random association between alleles of loci residing separately in nuclear and organellar genomes (Asmussen et al. 1989; Latta et al. 2001; Jaramillo-Correa and Bousquet 2005; Fields et al. 2014; Sloan et al. 2015). The normalized measure of LD, *D'*, ranges from -1 to 1, where *D'* = 0

indicates random assortment and *D'* = -1 or 1 indicates non-random assortment (see methods; Hedrick 1987). As there is no physical linkage between nuclear and organellar genomes, the accumulation of cytonuclear linkage disequilibria could be due to inheritance patterns and be indicative of evolutionary processes. Additionally, high levels of selfing, non-random mating (Asmussen et al. 1989), and neutral demographic processes, including founder effects or range expansions, may lead to significant patterns of cnLD (*D'* -1 or 1). Positive or negative cytonuclear selection could also lead to non-random associations between nuclear and organellar alleles and has been shown mathematically (e.g., Asmussen and Clegg 1982; Wade and Goodnight 2006; Wade and Brandvain 2009).

Linkage of Mitochondrial and Plastid Genomes

In plants, statistical associations may arise in a more complicated fashion owing to the additional organellar genome of the plastid (pt; ptDNA). Therefore, cnLD may be assessed between the nuclear genome and mtDNA or ptDNA, separately. If the inheritance patterns of the organellar genomes are maternal, however, the two would be inherited as a single cytoplasmic unit within egg cells, a so-called plasmon (Tsujimura et al. 2013). Following this, one might predict that the two genomes exhibit similar values of cnLD, while being in complete LD with one another, just as if the genomes were physically linked and do not undergo random assortment or recombination. High levels of organellar LD were found in the plant species *Quercus glauca* (ringed cup oak; Lin et al. 2003), *Silene vulgaris* (bladder campion; Olson and McCauley 2000) and *Beta vulgaris* (beet; Desplanque et al. 2000). Lin et al. (2003) found strong levels of LD between pairs of mt and pt loci in *Q. glauca*, although they also reported a few instances

of random associations (*D'* ~0), thought to arise through rare paternal leakage of organelles and recurrent mutation. In *S. vulgaris*, complete LD was found between mt and pt haplotypes while the authors also recovered congruent mt and pt phylogenies, as would be expected given similar mutation rates and inheritance patterns (Olson and McCauley 2000). Similarly, Desplanque et al. (2000) found a normalized *D'* of 0.965 in *B. vulgaris* when comparing mt to pt haplotypes.

Yet another complicating factor for evaluating patterns of cnLD in plants, is that the inheritance patterns of the two organellar genomes may be different. For example, conifers inherit mt either maternally or paternally (i.e. seed or pollen; dependent on taxa), but pt are inherited paternally (Mogensen 1996). In *Pinus ponderosa* (ponderosa pine), where mt are inherited maternally and pt, paternally, cnLD was found to be slightly higher in mt-nuclear loci pairs than in pt-nuclear loci pairs. Out of 11 nuclear allozymes, three loci showed significant associations with mtDNA, whereas two loci showed significant associations with the ptDNA (Latta et al. 2001). This is particularly notable given the different dispersal ranges of pollen and seed in this species. Spatial structuring of the mt haplotypes, but not pt haplotypes, was revealed, highlighting the different inheritance patterns of the two organelles. Given the generally smaller dispersal range of seed, the authors suggest more instances of mt-nuclear LD would be found than pt-nuclear LD. In a hybridizing zone between two spruce species, no significant mt-nuclear associations were found (Jaramillo-Correa and Bousquet 2005). Tsujimura et al. (2013) suggested that disrupted linkage between the mtDNA and ptDNA in Emmer and Dinkel wheat were the result of paternal leakage in species with presumably maternally-inherited organelles.

Heteroplasmy and Its Effect on Linkage Disequilibrium

Both paternal and maternal lines can contribute organellar genomes when the two are biparentally-inherited together. Given that maternal and paternal organellar genomes have allelic variation sufficient to distinguish one from the other, non-uniparental inheritance of organellar genomes will create a potentially detectable mixture of genomes within offspring. This mixture is termed heteroplasmy (Ramsey and Mandel 2019) and has been studied in numerous taxa – from humans (Naue et al. 2015), *Caenorhabditis* (Liau et al. 2007), and *Drosophila* (Volz-Lingenhöhl et al. 1992) to *Nicotiana* (Yamato and Newton 1999), *Silene* (McCauley et al. 2005; McCauley and Ellis 2008; Bentley et al. 2010), and *Daucus* (Mandel et al. 2012, 2016). Evidence is beginning to demonstrate that recombination in animal mtDNA is possible, at least in certain taxa (reviewed in Rokas et al. 2003). However, recombination has long been accepted in plant organellar genomes (Medgyesy et al. 1985; Barr et al. 2005; Fritsch et al. 2014), although mtDNA is more often found to undergo recombination than ptDNA. The mtDNA in plants is much larger than their animal counterparts (at least an order of magnitude), contains many repeat regions, and is now known to exist as multiple, distinct chromosomes, and not as single, circular molecules (Alverson et al. 2011; termed 'cytoploid' in Ramsey and Mandel 2019; Sloan et al. 2012; Sloan 2013; Wu et al. 2015). These chromosomes also exist in various numbers in different cells and tissues, a phenomenon termed substoichiometric shifting and has been shown in both the mtDNA (Chen et al. 2011; Sloan 2013) and ptDNA (Guo et al. 2014). In carrot (*Daucus carota*), the mtDNA exists as three subgenomic circles, which are proposed to be the result of recombination events between direct and inverted repeats (Robison and Wolyn 2002; Iorizzo et al. 2012).

Multiple copies of organellar genomes exist per cell (Jarvis and López-Juez 2013; Gurdona et al. 2016), and they are thought to randomly sort during cell division (Birky 2001; McCauley 2013).

Recent mathematical modeling has shown that given any type of non-neutral mutation, selection against heteroplasmy can lead to the evolution of uniparental inheritance (Christie et al. 2015; Christie and Beekman 2017a, 2017b). It follows that there could be strong selection against the state of heteroplasmy in organisms (Rand 2001). However, others have postulated that heteroplasmy could be maintained by balancing selection (e.g., Doublet et al. 2008). Many have suggested that within-individual selection on organellar genomes provides a possible mechanism for the individual to rapidly adapt (e.g., Korpelainen 2004; Woloszynska 2010; Christie and Beekman 2017a, 2017b). In addition, the potential for genetic or genomic conflict with the nuclear genome is of much interest (Burt and Trivers 2006; McCauley 2013). Maintaining associations between the nuclear and organellar genomes may facilitate co-adaptation (Sharpley et al. 2012; Havird et al. 2015), for example to maintain 'matching' of nuclear- and mt-encoded gene products in the face of nuclear heterozygosity and/or gene duplicates (Wolff et al. 2014). Thus, the roles of adaptive selection in heteroplasmy, cytonuclear interactions, and the coevolution of nuclear and organellar genomes remain open questions.

Aims

Given that heteroplasmy appears to be a widespread phenomenon, we wish to address whether patterns of cnLD differ between heteroplasmy states. If maintaining associations between the nuclear and organellar genomes facilitates co-adaptation (e.g.,

Sharpley et al. 2012; Havird et al. 2015), one might expect tight linkage between alleles in the heteroplasmic state and those encoded in the nucleus. We address this hypothesis in populations of wild carrot, *D. carota* L. spp. *carota*, in which we have documented extensive heteroplasmy in mtDNA and ptDNA (Mandel et al. 2012, 2016; Mandel and McCauley 2015). In this study, we calculated LD between all combinations of two markers in the mtDNA, one in the ptDNA, and 15 simple sequence repeats (SSRs) located in the nuclear genome of wild carrot.

Materials and Methods

Sampling

In order to access cytonuclear and organellar LD, we used nuclear and plastid locus data from Mandel et al. (2016) for the 265 individuals studied here. Leaf material or seeds (grown in the University of Memphis Greenhouses to obtain leaf material) were collected from 136 and 129 *D. carota* individuals found on Nantucket Island, Massachusetts (Nantucket) or the Olympic Peninsula, Washington (Olympic) United States (Supplementary Table S1; which provides information on the numbers and percentages of heteroplasmic and homoplasmic individuals per population within each region), presenting us with natural replicates. These regions are known to contain intrapopulation organellar polymorphisms, and in the region of Nantucket, the presence of individuals with CMS (pers. obs.). Most leaves and seeds were collected from publicly available properties, and no specific permissions were required to collect. Some populations on Nantucket were collected from properties owned by private conservation organizations, and permission was obtained from those organizations before collecting. This study did not involve endangered or protected species. Genomic DNA was isolated

from a 2-cm portion of leaf tissue obtained from individual leaves using the SQ Omega Biotek Plant Kit (Atlanta, GA, USA), following manufacturers recommendations, with the addition of 1% polyvinyl propylene (PVP) and 1% Ascorbic Acid to Buffer SQ1 provided by the manufacturer. A cleanup step was performed when necessary using the E.Z.N.A. Cycle Pure Kit from Omega Biotek.

Mandel et al. (2016) genotyped these 265 individuals for a plastid locus and nuclear SSRs as a part of a population genetic study to assess crop-wild gene flow in carrot. Fifteen nuclear microsatellites (SSRs) were PCR amplified using primers developed by Cavagnaro et al. (2011), and the intergenic region between the plastid genes *trnS* and *trnG* (referred to as "StoG" here) was PCR amplified using primer pairs developed by Hamilton (Hamilton 1999; see Table S1 of Mandel et al. 2016 for PCR reaction details).

Heteroplasmy Assays

We assayed two mt genes, *cox1* and *atp9*, for heteroplasmy. For *cox1*, we assayed a synonymous T/C single nucleotide polymorphism (SNP) in the coding region and quantified the level of heteroplasmy using the qPCR strategy of Mandel and McCauley (2015). Following these methods, we developed an artificial heteroplasmy mixture experiment to establish the mathematical relationship between differences in Ct value and relative copy numbers of the *cox1* heteroplasmic variants. As in Mandel and McCauley (2015), we chose pairs of individuals that were homoplasmic for either the T or C SNP. To simulate levels of heteroplasmy we mixed together their DNA to generate the following T versus C percent ratios, respectively: 100:0, 95:5 90:10, 75:25, 50:50, 25:75, 10:90, 5:95, 0:100. We analyzed the resulting Ct data in a linear regression framework,

and the resulting standard curve equation was used to estimate the relative copy number within an individual. For this study, all individuals were genotyped twice, and the resulting Ct values were entered into the equation to quantify the relative amounts of the T or C SNP in the sample (heteroplasmy score). The two heteroplasmy scores for each individual were averaged; only individuals with heteroplasmy scores ≥ 0.5% were classified as heteroplasmic.

We also attempted to use the qPCR strategy of Mandel and McCauley (2015) to study a T/C SNP in the *atp9* gene; however, this locus was not polymorphic in our study populations. We therefore studied a 42-bp insertion:deletion (indel) variant present downstream of the coding region. Individuals may vary with the absence or presence of one or two copies of the repeat as shown in Mandel et al. (2012). PCR of the *atp9* locus was performed following the same methodology as Mandel et al. (2012). The resulting PCR products were visualized on 1% agarose gels, and individuals were scored as either not containing the repeat (short) or containing the repeat once or twice (medium and long, respectively). Although there was evidence of heteroplasmy, we did not attempt to classify any individual as heteroplasmic for *atp9*, as we were unable to calculate a heteroplasmy score and some individuals exhibited very faint minor allele bands.

Mandel et al. (2016) scored StoG data based on an individual's major haplotype (ignoring heteroplasmy). For this study, however, we called heteroplasmy based on the presence of major and minor alleles as determined in GeneMarker v.2.6.3 (SoftGenetics, State College, PA), following the methods of Ellis et al. (2008). We called minor alleles based on their relative peak above background fluorescence, yet we did not calculate a quantitative heteroplasmy score.

Heteroplasmy and Patterns of Linkage Disequilibrium

From some of our previous work, we know that there is population structure between each region, so we pooled each regions' populations together to perform LD calculations. Within each region, we investigated LD between all combinations of nuclear and organellar loci (i.e., amongst all pairs of nuclear SSRs, nucLD; between each nuclear SSR and each organellar locus, mt-nucLD and pt-nucLD; between the two mt genes, mtLD; and between each mt gene and StoG, mt-ptLD; Fig. 1). For the measures of LD, Hedrick's (1987) multi-allelic extension of Lewontin's (1964a, 1964b) normalized *D*' was employed. The value of *D*' is standardized by dividing *D* by D_{max}, the maximum value of *D* possible given the observed allele frequencies p_i, (1 - p_i), and q_j, (1 - q_j). Thus, the standardized value of *D'* ranges from -1 to 1, with *D'* = 0 indicating linkage equilibrium, or complete random associations between alleles.

For the nucLD *D'* and r^2 calculations, we employed the software program Multiallelic Interallelic Disequilibrium Analysis Software (MIDAS; Gaunt et al. 2006) which can handle data in a multiallelic (per locus) state. Using the software Arlequin, (Excoffier et al. 2010), we obtained significance values for each nucLD calculation using a Fisher's Exact Test approximated by Monte-Carlo simulations under a null hypothesis of random assortment. The Fisher's Exact Test performed by Arlequin is comparable to the traditional 2 by 2 contingency table but allows for a table of any size (Excoffier et al. 2010). For this analysis, the maximum number of alternative tables explored was set to

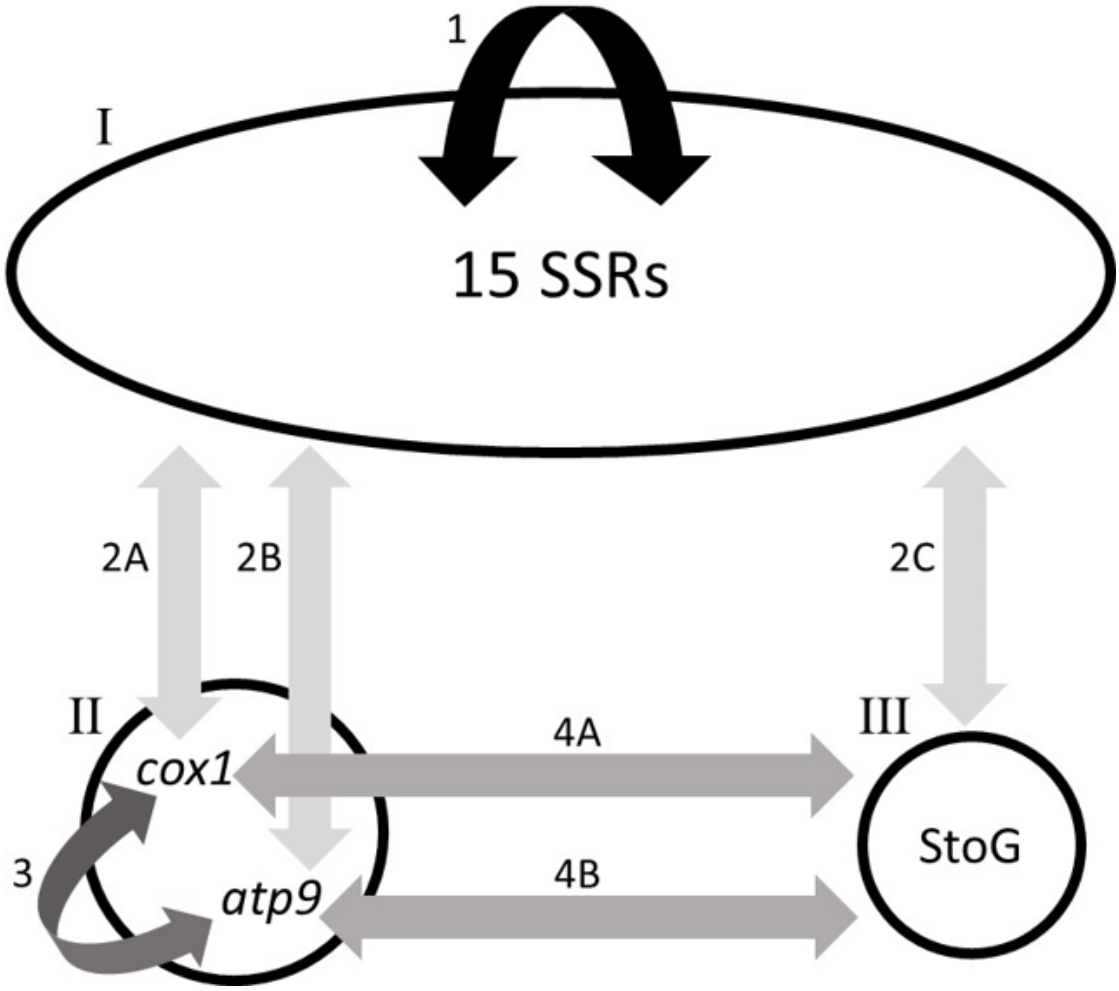


Fig. 1 Diagrammatic representation of the measures of LD calculated here within and between the nuclear (I), mitochondrial (II), and plastid (III) genomes. We calculated nucLD between 15 microsatellites (SSRs; 1); cytonuclear LD (cnLD) between the 15 nuclear SSRs and each of the mitochondrial genes cox1 (2A) and atp9 (2B) and the plastid locus StoG (2C); between cox1 and atp9 (3); and between StoG and each of cox1 (4A) and atp9 (4B).

100,000. We also used MIDAS to calculate organellar *D'* and r^2 values (within mtDNA and between mt-ptDNA) following the same procedures as above. The significance of multiple comparisons among locus pairs was assessed using the Benjamini-Hochberg correction for false discovery (Benjamini and Hochberg 1995) of 5%.

For calculations of cnLD, we used the program Cytonuclear Disequilibria for Multiallelic data (CNDm; Asmussen and Basten 1996; Basten and Asmussen 1997)

which calculates *D*' as well as significance using a Fisher's Exact Test approximated by Monte-Carlo simulations. We also corrected for multiple tests as above. For these measures, we calculated LD between the most common nuclear allele and all other nuclear alleles binned together (as CNDm requires biallelic data) and the mt SNP genotypes (*sensu* Fields et al. 2014). The Nantucket region dataset contained all 15 nuclear SSRs, yet one SSR failed to amplify in the Olympic region, so we used the remaining 14 for this region. In all cases above, the major organellar haplotype (as determined by qPCR) of heteroplasmic individuals was used as the mt or pt allele (i.e. all heteroplasmic individuals were treated as strictly haploid).

To determine if there is an influence of organellar heteroplasmy on LD, we separated our datasets into individuals which were heteroplasmic and homoplasmic for *cox1* and StoG, separately, within the Olympic region and *cox1* for the Nantucket region. We were unable to perform this experiment for StoG in the Nantucket region as only two individuals were heteroplasmic. Following the same procedures to analyze cytonuclear and organellar LD, as above, we thus carried out two additional sets of LD calculations (only heteroplasmic individuals and only homoplasmic individuals) for each cytonuclear comparison. We wished to know if measures of LD varied amongst these sets. To this end, we used a Z-transformation approach (Sokal and Rohlf 2012) to perform a pairwise test of heterogeneity of the *D'* value for each cytonuclear locus pair between heteroplasmic and homoplasmic individuals (since *D*' values are mathematically similar to the product-moment correlation employed in the Z-transformation approach), followed by a multiple test correction as above.

We then calculated within mtDNA (*cox1-atp9*) *D'* and r^2 again for the heteroplasmic and homoplasmic *cox1* datasets. CNDm does not calculate r^2 values for cnLD, so for the heteroplasmic and homoplasmic datasets we calculated these using the formula $D^2/(p_1p_2q_iq_2)$, following Hill and Robertson (1968), where D^2 is equal to the non-normalized value of linkage disequilibrium calculated in CNDm, p_1 *and* p_2 are equal to the frequencies of the major and minor *cox1* alleles (either 'C' or 'T'), respectively, and q_i, *and* q_2 are equal to the frequencies of the major and minor StoG alleles, respectively.

Lastly, for each marker type (nuclear, chloroplast, mitochondrial) we calculated unbiased nuclear gene diversity (uHe) separately in the two regions, chloroplast unbiased diversity (cpuH), and mitochondrial unbiased diversity (mtuH) (GenAlEx v. 6.5; Peakall and Smouse 2012). We also assessed population structure using analysis of molecular variation (AMOVA), as implemented in GenAlEx, to hierarchically partition genetic variation and estimate Wright's F_{ST}. Statistical significance (i.e., H_0 = no genetic differentiation among the populations) was determined by performing 1000 permutations.

Results

Heteroplasmy Assays

For the *cox1* qPCR mixture experiment, the R^2 value of the linear regression was 0.988, indicating a good fit of our artificial mixtures and Ct values from qPCR. As in previous work, a minimum threshold frequency of 0.5% of the minority SNP variant was used to define an individual as heteroplasmic (Mandel and McCauley 2015).

Overall, we found substantial levels of *cox1* T:C heteroplasmy in the populations sampled here. 113 of 265 individuals (42.6%) were heteroplasmic with each of the two *cox1* SNP variants serving as the minor variant in at least some heteroplasmic individuals

(Supplemental Table S1). Of the remaining 152 individuals, 103 (67.8%) were homoplasmic for the 'C' variant and 49 (32.2%) homoplasmic for the 'T' variant for the SNP. In the Olympic region, 39 individuals (28.7%) were heteroplasmic, and in the Nantucket region, 74 individuals (57.4%) were heteroplasmic (Supplementary Table S1).

Although there was indication of *atp9* heteroplasmy based on the presence of a second faint band in the correct size of one of the three alleles, we did not score any individual as heteroplasmic. In the Olympic region there were 83, 31 and 15 individuals with no insertion and with one and two copies of the insertion, respectively. In the Nantucket region there were 30, 91 and 3 individuals with no insertion and with one and two copies of the insertion, respectively.

A total of five alleles were present for StoG. Although some individuals contained one major and two minor alleles (tri-heteroplasmic; Supplementary Fig. S1; see Appendix 1 for all supplementary material), only the two highest allelic peaks were called as the major and minor haplotypes. Thirty-six individuals were heteroplasmic overall (13.6%). In the Olympic region all five alleles were present, and 102 were homoplasmic for either 432, 438, 440, 444, or 445. Thirty-four (25.0%) were heteroplasmic for either 440/445, 440/438, 438/440, or 445/438 (major/minor allele). However, in the Nantucket region, only three alleles were present (438, 440, and 444), and 127 individuals were homoplasmic. Two individuals (1.6%) were heteroplasmic, one for 440/444 and one for 440/438 (Supplementary Table S1).

Heteroplasmy and Patterns of Linkage Disequilibrium

In the Olympic region nuclear *D'* for the pooled dataset ranged from -0.63 to 0.36 (Supplementary Table S2; r^2 values can be found in Supplementary Table S3). After

calculating cnLD between the mtDNA and nuclear genome, *D'* for *atp9* ranged from -0.39 to 0.33, *D'* for *cox1* ranged from -0.41 to 0.36. The range of *D'* for StoG was -0.16 to 0.56 (Fig. 2; Supplementary Table S4). As heteroplasmy was present in both *cox1* and StoG, we divided these datasets into heteroplasmic and homoplasmic individuals for each locus. *D'* for the heteroplasmic *cox1* dataset ranged from -0.48 to 1, and for the homoplasmic *cox1* dataset *D'* ranged from -0.40 to 0.82. Four *cox1*-nuclear locus combinations exhibited *D'* values of 1 in the heteroplasmic dataset, but none were higher than 0.82 in the homoplasmic dataset. *D'* for the heteroplasmic StoG dataset ranged from -0.48 to 1, and it ranged from -0.39 to 0.48 in the homoplasmic StoG dataset. Three StoG-nuclear loci combinations had *D'* values of 1 in the heteroplasmic dataset, yet the highest *D'* value in the homoplasmic dataset was 0.48 (Fig. 2). Paired *D'* and r^2 values for each cytonuclear locus combination between heteroplasmic and homoplasmic datasets for each region are presented in Supplementary Table S5.

After a Z-transformation, a test of pairwise significance, and a Benjamini-Hochberg correction for false discovery between heteroplasmic and homoplasmic datasets, four *cox1*-nuclear allelic pairs had *D'* values significantly higher in the heteroplasmic *cox1* dataset, with each *D'* value equal to 1. Four StoG-nuclear allelic pairs had significantly higher *D'* values in the heteroplasmic StoG dataset, with three of their *D'* values equaling 1(Fig. 2). For between organellar LD, *D'* between *cox1*-StoG and *atp9*-StoG was 0.40 for each. Within the mtDNA (*cox1*-*atp9*) *D'* for the pooled dataset was -0.68, but for the heteroplasmic and homoplasmic *cox1* datasets *D'* was 0.49 and -0.79, respectively. Each value was significantly different from zero based on a Fisher's Exact Test (see Table 1 for paired *D'* and r^2 values for each organellar locus

combination).

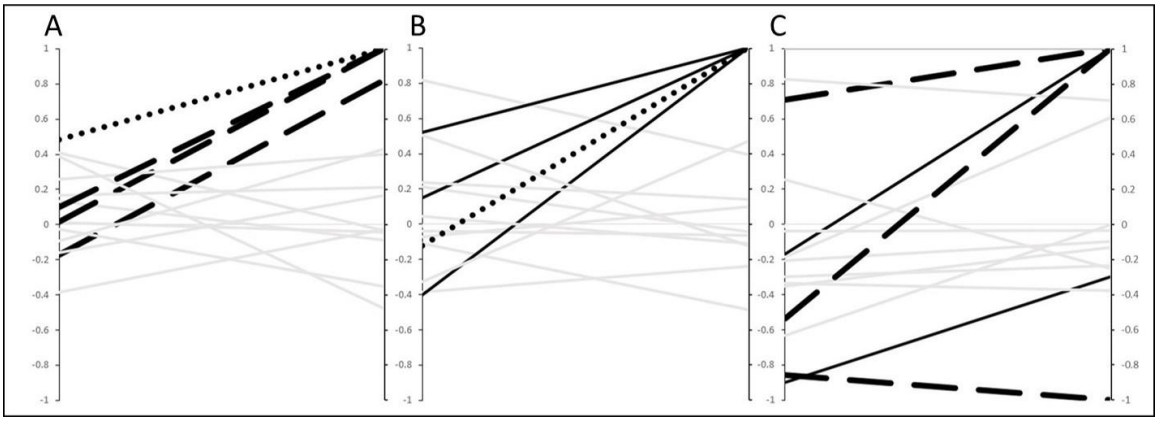


Fig. 2 cnLD values between heteroplasmic and homoplasmic StoG (A) and cox1 (B) datasets in the Olympic region and between heteroplasmic and homoplasmic cox1 datasets in the Nantucket region (C). Homoplasmic cnLD values are aligned along the left and heteroplasmic cnLD values are aligned along the right of each. Lines indicate paired cnLD values between homoplasmic and heteroplasmic datasets. Black lines indicate significantly different heteroplasmic and homoplasmic cnLD values. Of those, dotted lines indicate shared nuclear microsatellites (SSRs) in the Olympic region with significantly different cnLD values between heteroplasmic and homoplasmic datasets for StoG and cox1; dashed lines indicate shared nuclear SSRs with significantly different cnLD values between the Olympic StoG heteroplasmic and homoplasmic datasets and the Nantucket cox1 heteroplasmic and homoplasmic datasets.

In the Nantucket region nuclear *D'* for the pooled dataset ranged from -0.66 to 0.51 (Supplementary Table S6; r^2 values can be found in Supplementary Table S7). cnLD between *atp9* and *cox1*, separately, and the nuclear SSRs ranged from -0.28 to 0.34 and from -0.86 to 1, respectively. StoG cnLD ranged from -0.57 to 0.40 (Supplementary Table S4). We only scored heteroplasmy for *cox1* in the Nantucket region, so we were able to separate only this dataset into heteroplasmic and homoplasmic *cox1* individuals. *D'* of the heteroplasmic *cox1* dataset ranged from -1 to 1, and three *cox1*-nuclear allelic combinations had *D'* values of 1, and one other had a *D'* value of -1. *D'* of the homoplasmic *cox1* dataset ranged from -0.90 to 1, with only one *cox1*-nuclear allelic

combination with *D'* equal to 1 (Fig. 2). Each of these were significantly different from zero based on a Fisher's Exact test.

Table 1. Values of within and between organellar linkage disequilibrium (*D')* for the Olympic (left) and Nantucket (right) regions. *D'* of *cox1-atp9* are presented for each of the pooled, heteroplasmic, and homoplasmic datasets. The value of *D'* is standardized by dividing *D* by the maximum value of *D* possible based on the observed allelic frequencies p_i, $(1 - p_i)$, and q_j, $(1 - q_j)$ and ranges from -1 to 1, with *D'* = 0 signifying linkage equilibrium, random assortment, while *D'* on either extreme indicating linkage disequilibrium. mt = mitochondria; pt = plastid. All values except the one in bold are significantly different from zero based on a Fisher's Exact Test and after Benjamini-Hochberg correction for multiple tests.

Olympic				Nantucket			
Locus Association	**Organellar Association**	***D'***	***r²***	**Locus Association**	**Organellar Association**	***D'***	***r²***
cox1-atp9	mt-mt			*cox1-atp9*	mt-mt		
Pooled	mt-mt	-0.68	0.122	Pooled	mt-mt	0.81	0.207
Heteroplasmic	mt-mt	0.46	0.028	Heteroplasmic	mt-mt	**0.12**	0.009
Homoplasmic	mt-mt	-0.79	0.25	Homoplasmic	mt-mt	1	0.44
cox1-StoG	mt-pt	-0.4	0.019	*cox1*-StoG	mt-pt	-0.83	0.062
atp9-StoG	mt-pt	0.4	0.1	*atp9*-StoG	mt-pt	-0.22	0.013

After a Z-transformation and a test of pairwise significance, five *cox1*-nuclear allelic pairs showed significantly different *D'* values between the heteroplasmic and homoplasmic *cox1* datasets. Of these, four were higher in the heteroplasmic dataset, with three of them equaling 1, while one was lower with *D'* equal to -1 (Fig. 2). Between organellar LD varied by mt gene: *cox1*-StoG *D'* = 0.83, whereas *atp9*-StoG *D'* = 0.22. Within mtDNA also varied between each dataset. For the pooled dataset *cox1-atp9 D'* = 0.81 and for the heteroplasmic and homoplasmic *cox1* datasets *cox1-atp9 D'* = 0.12 and 1, respectively. All organellar LD values were significantly different from zero based on a Fisher's Exact test, except for the heteroplasmic *cox1-atp9* allelic combination (Table 1).

In total, 13 cytonuclear allelic combinations had significant differences in pairwise *D'* values between heteroplasmic and homoplasmic datasets (12 are higher, and one is lower, in the heteroplasmic datasets). Of four allelic pairs with significant differences in the Olympic StoG datasets, one is shared with the Olympic *cox1* dataset and the remaining three are shared with the Nantucket *cox1* dataset. No cytonuclear allelic combinations with significant differences in cnLD are shared between the *cox1* datasets of each region (Fig. 2).

It should be noted that after dividing the StoG and *cox1* datasets into heteroplasmic and homoplasmic subsets and necessarily reducing the sample sizes, fewer values of *D'* are significantly different from zero in the heteroplasmic subsets. Some population genetic statistics including measures of gene diversity and population structure can be found in Supplementary Table S8. The nuclear, chloroplast, and mitochondrial values of population structure suggest substantial gene flow via pollen and seed, and/or some degree of paternal inheritance of organellar genomes as previously reported (Mandel et al. 2012; 2016).

Discussion

Uniparental inheritance of organelles is considered typical and thus should result in offspring comprising a homogenous set of organellar DNA, i.e., homoplasmy. However, studies continue to reveal that patterns of organellar inheritance are more fluid than traditionally believed (e.g., Thyssen et al. 2012; see Fig. 1 in Ramsey and Mandel 2019). Investigations into organellar inheritance in animals is limited to mt, but studies in plant organellar inheritance benefit from the presence of the pt. Values of LD between organelles can offer indirect evidence of separate organellar inheritance patterns

(McCauley 2013), while values of LD within organelles may indicate recombination permitted by the underlying heteroplasmy generated through non-strict uniparental inheritance.

Patterns of Organellar Linkage Disequilibrium

Given paternal transmission of organelles, heteroplasmy may be generated. Either inherited maternally or augmented through additional paternal leakage events, such that it is maintained across generations, heteroplasmy may play a role in shaping patterns of within organellar LD. As within a heterozygous nuclear genome, recombination will act to limit linkage between genes, given allelic variants are present within organellar genomes (i.e., heteroplasmy; Mandel et al. 2012). In carrot recombination occurs within both organellar genomes (Ruhlman et al. 2006; Iorizzo et al. 2012). Interestingly, even though mtDNA LD varies between regions, here we show that within mtDNA LD varies between each of the pooled and *cox1* heteroplasmic and homoplasmic datasets. Pooled individuals in each region have the highest absolute *D'* magnitude while *cox1* heteroplasmic individuals have *D'* values closest to complete random assortment (Supplementary Table S9). This indicates recombination within the mtDNA, permitted by the presence, and possible maintenance across generations, of heteroplasmy.

Patterns of Cytonuclear Linkage Disequilibrium

Apart from *cox1*, the absolute magnitude of cnLD between each nuclear and organellar locus for pooled individuals is relatively modest and quite similar between two regions of the US, even though they are separated by more than 4,000 km (Supplementary Table S9). We also found this pattern in the nucLD datasets. Heteroplasmic individuals have more values of complete LD for *cox1*- and StoG-nuclear

associations than either pooled or homoplasmic individuals. Moreover, mean and median *cox1*- and StoG-nuclear *D'* for heteroplasmic individuals are each higher compared to homoplasmic individuals. McCauley (2013) suggests that paternal leakage and heteroplasmy may be higher in populations with histories of hybridization or admixture. Wild carrot in North America is introduced, most likely from multiple sources (e.g., different parts of Europe and the Mediterranean region) and in different parts of North America. Multiple introductions likely permit hybridization and admixture in new ranges. As this occurs, different cytonuclear allelic combinations may be formed, some of which cause cytonuclear incompatibility. Given this incompatibility impacts fitness, some combinations will not often be found together, impacting measures of LD. Furthermore, our previous work in these regions has shown hybridization between wild and locally-grown crop carrot, some of which were shown to be heteroplasmic for StoG (Mandel et al. 2016). Thus, processes of historic admixture, and contemporary hybridization with crop carrot, in North American populations of wild carrot, may play a role in shaping the patterns of cnLD displayed here.

While all four nuclear SSRs with significantly different associations with StoG were shared with the *cox1* dataset of either region, no such SSRs were shared between the *cox1* datasets between regions (Fig. 2). Given these results represent evolutionarily-relevant *cox1*-nuclear associations, one might predict that the same *cox1*-nuclear associations would be shared between regions, yet that is not what we found. With fewer StoG alleles present in the Nantucket region, we found fewer heteroplasmic individuals than in the Olympic region, as might be expected given the region-specific frequency of paternal leakage events between individuals containing differing ptDNA. This necessarily

limited calculating measures of StoG-nuclear cnLD in the Olympic region. The clear next step is to interrogate another polymorphic pt locus for these measures of LD to determine if the patterns we found here are consistent.

Summary and Implications for Cytonuclear Coevolution

In light of these findings, we suggest that heteroplasmy (potentially due to its generation from paternal leakage) aids in the maintenance of cnLD, and it has potential implications for cytonuclear interactions and coevolution. At the same time, heteroplasmy acts to disrupt organellar LD, indirectly, possibly by enabling the generation of new combinations of alleles during recombination. Although admixture and hybridization can disrupt linkage within and between nuclear and organellar genomes, and possibly cytonuclear compatibilities, we show that compared to homoplasmic individuals, heteroplasmic individuals demonstrate higher levels of cnLD. In contrast, each organellar LD value is closer to random assortment in heteroplasmic individuals which would be expected given heteroplasmy and recombination (Mandel et al. 2012; McCauley 2013). To test the hypothesis that heteroplasmy has maintained cnLD and disrupted organellar LD in these regions with admixture, these patterns should be compared to those in genetically-isolated populations from the native range of *D. carota*.

Homoplasmic individuals experiencing cytonuclear incompatibilities may have lower fitness, and this has been suggested in humans for individuals with nuclear and mtDNA genes of different ancestry (Zaidi and Makova 2018). However, heteroplasmy could mitigate the negative effects of incompatible nuclear and organellar genes and maintain individual fitness by providing alternative allelic variants to better match with nuclear alleles. If selection pressures acting upon cytonuclear interactions are stronger

than those for organellar interactions, cnLD would be maintained at the population level while organellar LD would be disrupted. Selection for heteroplasmy (possibly by the nuclear genome) could even mitigate cytonuclear incompatibles by allowing for better matching of organellar and nuclear alleles. These hypotheses await future testing; however, for now, we suggest that heteroplasmy is an important source of genetic variation within populations that should be considered.

Funding

This work was supported by the United States Department of Agriculture – National Institute of Food and Agriculture [grant number 2014-33522-21826 to JRM and DEM]. Travel funding for AJR to present in the symposium was provided by the supporters thereof.

Acknowledgements

AJR would like to thank Geoff Hill and Justin Havird for the opportunity to present in the symposium. We are grateful to the many conservation organizations on Nantucket, and private landowners on the Olympic Peninsula and on Nantucket Island for permission to collect. We would like to thank Zachary Ramsey, Jeremy Mandel, Mark Mandel, Michelle McCauley, Sara Samoray, and Robert Coleman for lab work and assistance. We also thank Tom Cunningham at the Molecular Resource Center at the University of Tennessee Health Science Center for assistance collecting data.

References

Alverson AJ, Rice DW, Dickinson S, Barry K, Palmer JD. 2011. Origins and Recombination of the Bacterial-sized Multichromosomal Mitochondrial Genome of Cucumber. Plant Cell 23:2499–2513.

Asai S, Shimoda C, Nishiumi I, Eguchi K, Yamagishi S. 1999. Isolation of microsatellite loci for paternity testing in the rufous vanga *Scheba rufa*. Mol Ecol 8:513–25.

Asmussen MA, Arnold J, Avise JC. 1989. The Effects of Assortative Mating and Migration on Cytonuclear Associations in Hybrid Zones. Genetics 122:923–34.

Asmussen MA, Basten CJ. 1996. Constraints and normalized measures for cytonuclear disequilibria. Heredity (Edinb) 76:207–14.

Asmussen MA, Clegg MT. 1982. Rates of Decay of Linkage Disequilibrium Under Two-Locus Models of Selection. J Math Biol 14:37–70.

Barr CM, Neiman M, Taylor DR. 2005. Inheritance and recombination of mitochondrial genomes in plants, fungi and animals. New Phytol 168:39–50.

Basten CJ, Asmussen MA. 1997. The Exact Test for Cytonuclear Disequilibria. Genetics 146:1165–71.

Benjamini Y, Hochberg Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B 57:289–300.

Bentley KE, Mandel JR, McCauley DE. 2010. Paternal Leakage and Heteroplasmy of Mitochondrial Genomes in *Silene vulgaris*: Evidence From Experimental Crosses. Genetics 185:961–68.

Birky CW. 2001. The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. Annu Rev Genet 35:125–48.

Blier PU, Dufresne F, Burton RS. 2001. Natural selection and the evolution of mtDNA-encoded peptides: Evidence for intergenomic co-adaptation. Trends Genet 17:400–406.

Brandvain Y, Wade MJ. 2009. The Functional Transfer of Genes From the Mitochondria to the Nucleus: The Effects of Selection, Mutation, Population Size and Rate of Self-Fertilization. Genetics 182:1129–39.

Burt A, Trivers R. 2006. Genes in Conflict: The Biology of Selfish Genetic Elements Harvard University Press.

Camus MF, Wolff JN, Sgrò CM, Dowling DK. 2017. Experimental evidence that thermal selection has shaped the latitudinal distribution of mitochondrial haplotypes in Australian *Drosophila melanogaster*. Mol Biol Evol 34:2600–2612.

Cavagnaro PF, Chung S-M, Manin S, Yildiz M, Ali A, Alessandro MS, Iorizzo M, Senalik D a, Simon PW. 2011. Microsatellite isolation and marker development in carrot - genomic distribution, linkage mapping, genetic diversity analysis and marker transferability across Apiaceae. BMC Genomics 12:386.

Chen J, Guan R, Chang S, Du T, Zhang H, Xing H. 2011. Substoichiometrically Different Mitotypes Coexist in Mitochondrial Genomes of *Brassica napus* L. PLoS One 6:e17662.

Christie JR, Beekman M. 2017a. Selective sweeps of mitochondrial DNA can drive the evolution of uniparental inheritance. *Evolution* 71:2090–99.

Christie JR, Beekman M. 2017b. Uniparental inheritance promotes adaptive evolution in cytoplasmic genomes. Mol Biol Evol 34:677–91.

Christie JR, Schaerf TM, Beekman M. 2015. Selection against Heteroplasmy Explains

the Evolution of Uniparental Inheritance of Mitochondria. PLoS Genet 11:e1005112.

Desplanque B, Viard F, Bernard J, Forcioli D, Saumitou-Laprade P, Cuguen J, Van Dijk H. 2000. The linkage disequilibrium between chloroplast DNA and mitochondrial DNA haplotypes in *Beta vulgaris* ssp. *maritima* (L.): The usefulness of both genomes for population genetic studies. Mol Ecol 9:141–54.

Doublet V, Souty-Grosset C, Bouchon D, Cordaux R, Marcadé I. 2008. A thirty million year-old inherited heteroplasmy. PLoS One 3:e2938.

Excoffier L, Lischer HEL, Note CP. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10:564–67.

Fields PD, McCauley DE, McAssey E V, Taylor DR. 2014. Patterns of cyto-nuclear linkage disequilibrium in *Silene latifolia*: genomic heterogeneity and temporal stability. Heredity (Edinb) 112:99–104.

Fritsch ES, Chabbert CD, Klaus B, Steinmetz LM. 2014. A genome-wide map of mitochondrial DNA recombination in yeast. Genetics 198:755–71.

Gaunt TR, Rodriguez S, Zapata C, Day INM. 2006. MIDAS: software for analysis and visualisation of interallelic disequilibrium between multiallelic markers. BMC Bioinformatics 7:227.

Guo W, Grewe F, Cobo-Clark A, Fan W, Duan Z, Adams RP, Schwarzbach AE, Mower JP. 2014. Predominant and Substoichiometric Isomers of the Plastid Genome Coexist within *Juniperus* Plants and Have Shifted Multiple Times during Cupressophyte Evolution. Genome Biol Evol 6:580–90.

Gurdona C, Svab Z, Feng Y, Kumar D, Maliga P. 2016. Cell-to-cell movement of mitochondria in plants. Proc Natl Acad Sci USA 113:3395–3400.

Havird JC, Hall MD, Dowling DK. 2015. The evolution of sex: A new hypothesis based on mitochondrial mutational erosion. BioEssays 37:951–58.

Hedrick PW. 1987. Gametic disequilibrium measures: proceed with caution. Genetics 117:331–41.

Hill WG, Robertson A. 1968. Linkage disequilibrium in finite populations. Theor Appl Genet 38:226–31.

Iorizzo M, Senalik D, Szklarczyk M, Grzebelus D, Spooner D, Simon P. 2012. De novo assembly of the carrot mitochondrial genome using next generation sequencing of whole genomic DNA provides first evidence of DNA transfer into an angiosperm plastid genome. BMC Plant Biol.

Jaramillo-Correa JP, Bousquet J. 2005. Mitochondrial genome recombination in the zone of contact between two hybridizing conifers. Genetics 171:1951–62.

Jarvis P, López-Juez E. 2013. Biogenesis and homeostasis of chloroplasts and other plastids. Nat Rev Mol Cell Biol 14:787–802.

Korpelainen H. 2004. The evolutionary processes of mitochondrial and chloroplast genomes differ from those of nuclear genomes. Naturwissenschaften 91:505–18.

Latta RG, Linhart YB, Mitton JB. 2001. Cytonuclear disequilibrium and genetic drift in a natural population of ponderosa pine. Genetics 158:843–50.

Lewontin RC. 1964a. The Interactions of Selection and Linkage II. Optimus Models. Genetics 50(4):49–67.

Lewontin RC. 1964b. The Interaction of Selection and Linkage. I. General

Considerations; Heterotic Models. Genetics 49:49–67.

Liau W, Gonzalez-Serricchio AS, Deshommes C, Chin K, LaMunyon CW. 2007. A persistent mitochondrial deletion reduces fitness and sperm performance in heteroplasmic populations of *C. elegans*. BMC Genet 8:1–11.

Lin TP, Chuang WJ, Huang SSF, Hwang SY. 2003. Evidence for the existence of some dissociation in an otherwise strong linkage disequilibrium between mitochondrial and chloroplastic genomes in *Cyclobalanopsis glauca*. Mol Ecol 12:2661–68.

Mandel JR, McAssey E V., Roland KM, McCauley DE. 2012. Mitochondrial gene diversity associated with the atp9 stop codon in natural populations of wild carrot (*Daucus carota* ssp. *carota*). J Hered 103:418–25.

Mandel JR, McCauley DE. 2015. Pervasive Mitochondrial Sequence Heteroplasmy in Natural Populations of Wild Carrot, *Daucus carota* spp. *carota* L. PLoS One 10:1–12.

Mandel JR, Ramsey AJ, Iorizzo M, Simon PW. 2016. Patterns of Gene Flow between Crop and Wild Carrot, *Daucus carota* (Apiaceae) in the United States. PLoS One 11:1–19.

McCauley DE. 2013. Paternal leakage, heteroplasmy, and the evolution of plant mitochondrial genomes. New Phytol 200:966–77.

McCauley DE, Bailey MF, Sherman NA, Darnell MZ. 2005. Evidence for paternal transmission and heteroplasmy in the mitochondrial genome of *Silene vulgaris*, a gynodioecious plant. Heredity (Edinb) 95:50–58.

McCauley DE, Ellis JR. 2008. Recombination and linkage disequilibrium among mitochondrial genes in structured populations of the gynodioecious plant *Silene*

vulgaris. *Evolution* 62:823–32.

Medgyesy P, Fejes E, Maliga P. 1985. Interspecific Chloroplast Recombination in a Nicotiana Somatic Hybrid. Proc Natl Acad Sci USA 82:6960–64.

Mogensen HL. 1996. The Hows and Whys of Cytoplasmic Inheritance in Seed Plants. Am J Bot 83:383–404.

Naue J, Hörer S, Sänger T, Strobl C, Hatzer-Grubwieser P, Parson W, Lutz-Bonengel S. 2015. Evidence for frequent and tissue-specific sequence heteroplasmy in human mitochondrial DNA. Mitochondrion 20:82–94.

Olson MS, McCauley DE. 2000. Linkage disequilibrium and phylogenetic congruence between chloroplast and mitochondrial haplotypes in *Silene vulgaris*. Proc R Soc London B Biol Sci 267:1801–8.

Peakall R, Smouse PE. 2012. GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28:2537–39.

Ramsey AJ, Mandel JR. 2019. When One Genome Is Not Enough: Organellar Heteroplasmy in Plants. Annu Plant Rev Online 2:1–40.

Rand DM. 2001. The Units of Selection of Mitochondrial DNA. Annu Rev Ecol Syst.

Rand DM, Fry A, Sheldahl L. 2006. Nuclear-Mitochondrial Epistasis and *Drosophila* Aging: Introgression of *Drosophila simulans* mtDNA Modifies Longevity in *D. melanogaster* Nuclear Backgrounds. Genetics 172:329–41.

Robison MM, Wolyn DJ. 2002. Complex organization of the mitochondrial genome of petaloid CMS carrot. Mol Genet Genomics 268:232–39.

Rokas A, Ladoukakis ED, Zouros E. 2003. Animal mitochondrial DNA recombination revisited. Trends Ecol Evol 18:411–17.

Ruhlman TA, Lee S, Jansen RK, Hostetler JB, Tallon LJ, Town CD, Daniell H. 2006. Complete plastid genome sequence of *Daucus carota*: implications for biotechnology and phylogeny of angiosperms. BMC Genomics 7:222.

Sambatti JBM, Ortiz-Barrientos D, Baack EJ, Rieseberg LH. 2008. Ecological selection maintains cytonuclear incompatibilities in hybridizing sunflowers. Ecol Lett 11:1082–91.

Sharpley MS, Marciniak C, Eckel-Mahan K, McManus M, Crimi M, Waymire K, Lin CS, Masubuchi S, Friend N, Koike M, Chalkia D, MacGregor G, Sassone-Corsi P, Wallace DC. 2012. Heteroplasmy of mouse mtDNA is genetically unstable and results in altered behavior and cognition. Cell 151:333–43.

Sloan DB. 2013. One ring to rule them all? Genome sequencing provides new insights into the 'master circle'model of plant mitochondrial DNA structure. New Phytol.

Sloan DB, Alverson AJ, Chuckalovcak JP, Wu M, McCauley DE, Palmer JD, Taylor DR. 2012. Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. PLoS Biol 10:e1001241.

Sloan DB, Fields PD, Havird JC. 2015. Mitonuclear linkage disequilibrium in human populations. Proc R Soc B Biol Sci 282:20151704.

Sokal R, Rohlf F. 2012. Biometry: the principles and practice of statistics in biological research. 4th ed New York: WH Freeman and co.

Tao D, Hu F, Yang J, Yang G, Yang Y, Xu P, Li J, Ye C, Dai L. 2004. Cytoplasm and cytoplasm-nucleus interactions affect agronomic traits in *japonica* rice. Euphytica 135:129–34.

Thyssen G, Svab Z, Maliga P. 2012. Exceptional inheritance of plastids via pollen in

Nicotiana sylvestris with no detectable paternal mitochondrial DNA in the progeny. Plant J 72:84–88.

Tsujimura M, Mori N, Yamagishi H, Terachi T. 2013. A possible breakage of linkage disequilibrium between mitochondrial and chloroplast genomes during Emmer and Dinkel wheat evolution. Genome 56:187–93.

Volz-Lingenhöhl A, Solignac M, Sperlich D. 1992. Stable heteroplasmy for a large-scale deletion in the coding region of *Drosophila subobscura* mitochondrial DNA. Proc Natl Acad Sci USA 89:11528–32.

Wade MJ, Brandvain Y. 2009. Reversing mother's curse: Selection on male mitochondrial fitness effects. *Evolution* 63:1084–89.

Wade MJ, Drown DM. 2016. Nuclear-mitochondrial epistasis: a gene's eye view of genomic conflict. Ecol Evol 6:6460–72.

Wade MJ, Goodnight CJ. 2006. Cyto-nuclear epistasis: two-locus random genetic drift in hermaphroditic and dioecious species. *Evolution* 60:643–59.

Weng ML, Ruhlman TA, Jansen RK. 2016. Plastid-Nuclear Interaction and Accelerated Coevolution in Plastid Ribosomal Genes in Geraniaceae. Genome Biol Evol 8:1824–38.

Wolff JN, Ladoukakis ED, Enríquez JA, Dowling DK, Enriquez JA. 2014. Mitonuclear interactions: evolutionary consequences over multiple biological scales. Philos Trans R Soc B Biol Sci 369:20130443–20130443.

Woloszynska M. 2010. Heteroplasmy and stoichiometric complexity of plant mitochondrial genomes-though this be madness, yet there's method in't. J Exp Bot 61:657–71.

Wu Z, Cuthbert JM, Taylor DR, Sloan DB. 2015. The massive mitochondrial genome of the angiosperm *Silene noctiflora* is evolving by gain or loss of entire chromosomes. Proc Natl Acad Sci USA 112:10185–91.

Yamato KT, Newton KJ. 1999. Heteroplasmy and Homoplasmy for Maize Mitochondrial Mutants: A Rare Homoplasmic nad4 Deletion Mutant Plant. J Hered 90:369–73.

Zaidi AA, Makova KD. 2018. Mito-nuclear effects uncovered in admixed populations. bioRxiv 349126.

Zeyl C, Andreson B, Weninck E. 2005. Nuclear-mitochondrial epistasis for fitness in *Saccharomyces cerevisiae*. Evolution 59:910–14.

CHAPTER 5: ALTERED POLLINATION ECOLOGY OF A NATIVE PLANT SPECIES IN THE PRESENCE OF AN ATTRACTIVE NON-NATIVE PLANT SPECIES

Introduction

Humans have facilitated the movement of species past their natural ranges for millennia. Within the last century, researchers have begun to document the effects of introduced species on the ecology of native flora. Introduced species often outcompete native species for limited resources including water, nutrients, space, or sunlight (Bartomeus et al. 2008; Lopezaraiza-Mikel et al. 2007; Morales and Traveset 2009). Well-known examples include *Taraxacum* spp. (dandelion, Kandori et al. 2009), *Lonicera japonica* Thunb. (Japanese honeysuckle, Schierenbeck 2004), *Pueraria montana* (Lour.) Merr. (kudzu, J.H. Miller and Edwards 1983), and *Opuntia* (L.) (prickly pear cactus; Novoa et al. 2015). Finally, given that plants are keystone species in many environments, the effects of non-native species may cascade to affect entire ecosystems. Valtonen et al. (2006) suggested a "bottom-up" effect of the presence of the invasive *Lupinus polyphyllus* Lindl. in areas of invasion. In addition to containing fewer butterflies with a higher proportion of flying individuals, road verges containing *L. polyphyllus* showed a decrease in species diversity of flora compared to road verges that contained no *L. polyphyllus*.

In some cases, the effects of these interactions appear to be facilitative. For example, non-native species with flowers that are attractive to pollinators may increase pollinator visits and enhance pollination of native species, the so-called "magnet" species

effect (Bartomeus et al. 2008; Muñoz and Cavieres 2008; Morales and Traveset 2009; Bruckman and Campbell 2016). In populations of native species that are pollen- or pollinator-limited, more pollinator visits could increase seed set. Albrecht et al. (2012) demonstrated increased seed set with increasing pollinator diversity in a study of radish, *Raphanus sativus* L. However, at the highest species richness investigated, there was no difference in mean fruits per flower and a slight reduction in mean number of seeds set per fruit compared to levels of moderate pollinator species richness.

A meta-analysis of 40 studies reported a significantly negative overall effect of non-native species on native species via pollination disruption (Morales and Traveset 2009). Non-native species may alter pollinator behavior, and a pollinator that visits both native and non-native species may follow one of three scenarios. The pollinator may initially visit the non-native species before visiting the native species, it may visit the non-native species after visiting the native species, or it may alternate from one to another. Depending on where pollen is deposited on the pollinator and where stigmas contact the body, stigmas of native species could be "contaminated" with heterospecific pollen (Brown et al. 2002; Jakobsson et al. 2008) through interspecific pollen transfer (Morales and Traveset 2008). Bruckman and Campbell (2016) showed that at low densities of the exotic *Brassica nigra* (L.) W.D.J. Koch, an increase in visits yielded an increase in seed set on the native species *Phacelia parryi* Torr., but at high densities exotic pollen deposition overwhelmed their stigmas and reduced seed set. Presumably, this amount of heterospecific pollen overloaded the stigma's capacity or reduced the amount of native pollen such that it interfered with pollen germination and/or pollen tube growth and ultimately, fertilization (Bruckman and Campbell 2016). In a study of nine

sympatric species, McLernon et al. (1996) found that they received different amounts of heterospecific pollen and that up to 100% of pollen grains on stigmas were heterospecific. Similarly, Larson et al. (2006) found no significant increase in conspecific pollen loads and significantly lower conspecific pollen loads on native species in six of eight communities studied. Under these conditions, pollinator behavior may reduce seed set in native species (Bjerknes et al. 2007; Kandori et al. 2009; Vanparys et al. 2008).

Although the effects of non-native species on native species are variable, there are some consistent patterns. Native flowers that morphologically resemble a non-native flower will be affected more than native flowers that do not resemble a non-native flower, even superficially (Kandori et al. 2009; Morales and Traveset 2009). Negative effects are most prevalent when non-native species are at high densities relative to the native species (Brown et al. 2002; Flanagan et al. 2010; Lopezaraiza-Mikel et al. 2007; Muñoz and Cavieres 2008). Likewise, when both species are generalist-pollinated, they may compete for the same pollinators, increasing the probability that the presence of the non-native may impact native pollination (Bjerknes et al. 2007; Mitchell et al. 2009; Morales and Traveset 2008, 2009).

We investigated whether the presence of a non-native plant, *Daucus carota* subsp. *carota* (L.), impacts the pollination ecology of a native species, *Sericocarpus asteroides* (L.) B.S.P. on Nantucket Island, Massachusetts (hereafter, Nantucket). No studies have examined effects of the presence of *D. carota* on native species, and we selected *S. asteroides* due to similarities with *D. carota*, which might make its pollination ecology susceptible to disturbance by *D. carota*. Our project incorporated an observational study

of both species' pollinator activity, a manipulative experiment, and an examination of *S. asteroides* inflorescences for heterospecific pollen.

Sericocarpus asteroides is common on sandplain grasslands (Farnsworth 2007), many of which are currently protected by conservation organizations. Undisturbed landscapes are typically resistant to invasion by *D. carota*, but human activities, particularly development, can dramatically increase its presence on the landscape (Buckley et al. 2007). Despite the high levels of conservation on Nantucket, human development over the last two centuries has increased, and roadways have dissected many sandplain grasslands. Many of the roadways near and within conservation properties are unpaved, with sand as substrate, the sides of which are suitable habitat for *D. carota*.

Our study consisted of three components: an observational study of *S. asteroides* in populations containing *D. carota* (sympatric) and those containing no *D. carota* (allopatric), a removal experiment where we restored sympatric *S. asteroides* populations to an allopatric state through the removal of *D. carota* inflorescences, and an observational study of the extent of heterospecific pollen deposition in various natural populations of *S. asteroides*. We hypothesized that 1) The presence of *D. carota* would increase the number and diversity of pollinators on *S. asteroides*; 2) The removal of *D. carota* from sympatric plots would restore pollinator visits and diversity to those found in plots with only *S. asteroides*; and 3) Higher levels of heterospecific pollen would be found in *S. asteroides* populations located close to *D. carota*.

Materials and Methods

Species Descriptions. *Daucus carota* (Queen Anne's Lace, or wild carrot; Apiaceae) is a common non-native species throughout North America. It is the wild relative of crop carrot (*D. carota* subsp. *sativus*), and it is considered a highly invasive species in many areas. In the United States, the species is typically biennial, as it grows a rosette during the first year and bolts (flowers) May-August in the second year before dying (Mitich 1996). The inflorescence of *D. carota* is a compound umbel composed of hundreds of white florets. The species can grow 1-2 meters tall (Mitich 1996), though there is considerable variation in height across populations (pers. obs.). *Daucus carota* is reportedly generalist-pollinated in its native range of central Asia, and Diptera (flies) are the most common pollinators (Ahmad and Aslam 2002; Westmoreland and Muntan 1996). In the United States individuals grow along roadsides and in disturbed landscapes, and the species is documented in this type of environment on Nantucket (Sorrie and Dunwiddie 1996). A single plant can produce several thousand spiny fruits (Lacey 1981; Mitich 1996). The species was introduced to New England with the earliest settlers but quickly escaped into the wild (Ahmad and Aslam 2002; Mitich 1996). It was described on Nantucket in 1888 as being "Too common, a great pest; over-running entire fields" (Owen 1888).

One native Nantucket Asteraceae (Compositae) species with a similar morphology and life history as *D. carota* is *Sericocarpus asteroides* (toothed whitetop aster). Like *D. carota*, this aster species produces a rosette of leaves during vegetative growth and an upright stem when bolting. *Sericocarpus asteroides* is rhizomatous (pers. obs.), although sexual reproduction does occur (Frey et al. 2007; Giblin 1997). It is generalist-pollinated, possessing a corymb-like inflorescence composed of a variable

number of heads (capitula), thus resembling a small compound umbel. Together with the white florets, these characteristics give its inflorescences an overall morphology similar to *D. carota*. On Nantucket, *S. asteroides* flowers from late June or early July through to September (Bicknell 1915). Neither species is known to exhibit ultraviolet reflectance, which may influence pollinator attractiveness.

Pollinator Observational Study. In June and July 2015 and June 2016, we selected *D. carota* plots (some in sympatry with *S. asteroides*) from a survey of *D. carota* and *S. asteroides* occurrences being undertaken concurrently (See Appendix 2 for all supplementary FigShare documents; FigShare; all FigShare documents are available at https://doi.org/10.6084/m9.figshare.7542191.v2). We relocated sites using their known GPS coordinates and Smart Compass Pro (Smart Tools Co., Daegu, Republic of Korea) installed on an Android smartphone with a GPS accuracy of <5m (van Diggelen et al. 2015). Peak pollinator activity for *D. carota* has been reported as 10:00-11:00 local time (Goyal et al. 1989; Westmoreland and Muntan 1996; Ahmad and Aslam 2002). Thus, we chose a 1-m^2 plot with the highest floral density and observed pollinator activity for three hours each morning (08:00-11:00). We observed 10 and seven separate *D. carota* populations in 2015 and 2016, respectively.

It is worth noting that estimates of diversity may vary due to the approach used to identify pollinator visitors. For example, Lamborn and Ollerton (2000) estimated 20 pollinator visitor species through visual inspection on wild carrot in central England, whereas Bohart and Nye (1960) identified 334 insect species representing 71 families on crop carrot inflorescences in Utah by collecting pollinator visitors. We visually identified visitors in an apparent pollinating position (i.e., not on the sides or bottom of umbels).

Although we were able to visually identify some pollinators to genus and/or species, we did not collect pollinators. Therefore, we recorded pollinators at family level, similar to Ahmad and Aslam (2002), Brown et al. (2002), and Parachnowitsch and Elle (2005). We recorded local weather conditions hourly.

To determine whether *D. carota* increases the number of pollinator visitors and pollinator diversity visiting *S. asteroides*, we performed one observational study (as for *D. carota*) and one removal experiment. First, we divided the 23 populations of *S. asteroides* recorded from a prior survey (FigShare) into two groups - those in sympatry (<5m from one another and visible) and allopatry (>5m away and out of visual sight) with respect to *D. carota*. We observed six sympatric and six allopatric populations in the same fashion as for the *D. carota pollinator* survey, and we selected a random population each morning based on their accessibility and size of the *S. asteroides* population (preferring larger ones), alternating between sympatric and allopatric populations from day to day.

Next, we returned to the five sympatric populations with the highest numbers of *D. carota* present and performed a removal experiment. On day one (pre-removal) we observed pollinator visitors as before. Immediately after observation, we searched for and removed umbels of *D. carota*, so none were visible at eye level within the patch or obscured from sight behind nearby bushes or curves in roads. We then observed pollinators the following morning (post-removal). Anthers of *Sericocarpus* are known to dehisce in the morning, and pollen be absent by day's end (Giblin 1997). Therefore, we observed *S. asteroides* between 08:00-11:00 in order to overlap with the highest period of *D. carota* pollinator activity.

We analyzed pollinator visitation at the family level. We calculated the Shannon's and Simpson diversity indices (H and D, respectively), the species richness and evenness (used as family richness and evenness here; S and E, respectively) in Microsoft Excel (2013) of total visits per plot, following Donaldson et al. (2002) and Fang and Huang (2013), for the observational study and removal experiment. We performed a two-tailed t-test on the total number of visits between the sympatric and allopatric observational study populations and another two-tailed t-test on the total number of visits between the pre- and post-removal plots. Additionally, it is known that floral density may impact pollinator visitation (e.g., in Mustajärvi et al. 2001), so we addressed the possibility of inflorescence density affecting the numbers of pollinators in three ways. First, to determine if the number of inflorescences differed between observation plots with *D. carota* and those that did not, we performed a two-tailed t-tests of the number of inflorescences (sum of *D. carota* and *S. asteroides*) between sympatric and allopatric populations and a two-tailed t-test between pre- and post-removal plots. Second, we standardized the total number of visits by dividing that number by the number of *S. asteroides* inflorescences within the 1-m^2 observation plot, similar to Kandori et al. (2009). Third, we standardized total visits in the sympatric populations and pre-removal plots by dividing total visits by the sum of both species' inflorescences. We again performed t-tests as before using these standardized visits. We performed all t-tests in JMP version 13.0.0 (SAS Institute, Cary, North Carolina).

Heterospecific Pollen Assay. To determine the amount of heterospecific pollen on *S. asteroides*, we collected 15 inflorescences from each of the 23 *S. asteroides* populations across Nantucket. We measured the distance of each *S. asteroides* population

to the nearest *D. carota* population using their GPS coordinates and the Measure Distance feature in Google Maps. We later verified and checked these values for accuracy by measuring the same distances using QGIS version 3.4.3-Madeira (Open Source Geospatial Foundation Project, http://qgis.osgeo.org/) and determining the average distance between *S. asteroides* and *D. carota* for Google Maps and QGIS (337.8 and 336 m, respectively) and the average deviation and average percent deviation across all populations (4.3 m and 6.5%, respectively). Once a population of *S. asteroides* was located, 15 inflorescences were collected and stored in 70% ethanol on ice in the field and during transfer to the University of Memphis, before being stored at -20°C in the lab. We took note of the local floral community of each population and collected pollen samples from flowering species across Nantucket to create a pollen key (see below); however, we did not perform a rigorous survey of the floral community.

We then dissected the inflorescences and mounted all receptive stigmas from one head of each inflorescence on a glass slide, which we covered with a cover slip, sealed using clear nail polish, and observed under light microscopy (following Peterson et al. 2010). We created a dichotomous key of pollen samples, collected from species flowering across Nantucket, based on their microscopic characteristics to generate a key of pollen samples identified to the lowest possible taxon (*D. carota*, *S. asteroides*, *Achillea millifolum* L., *Rosa* spp., *Leucanthemum* spp., Fabaceae spp.). We used this key to identify and count pollen grains. However, some pollen grains were unidentifiable to taxa in our key, so we scored these as one of four morphospecies (Unknowns 1, 2, 3, and LrgUknown), based on their microscopic characteristics, and counted them. The number of stigmas per head likely contributed to the amount of pollen per slide, so we

standardized each total pollen count by dividing that number by the number of stigmas per head. We used distance to the nearest *D. carota* population to predict total pollen, standardized total pollen, and total heterospecific pollen, separately, in a Generalized Linear Model with a Poisson distribution in JMP.

Results

Pollinator Observational Study. We observed a total of 654 pollinator visitors (477 in 2015 and 177 in 2016; FigShare). These represented 16 different insect families – three families of bees, two families of butterflies, five families of wasps, and six families of flies. Although weather seemed to affect pollinator visits during one morning of rain, which required a second visit, most days had sunny to moderate cloud cover and light winds.

Table 1. Sample sizes (N), mean number of *Sericocarpus asteroides* L. inflorescences, mean overall visits, and diversity indices for pollinator visitors within study plots. S = mean family richness; H = Shannon diversity index; D = Simpson's diversity index; E = family evenness. See text for details of analyses.

Study Plot	N	Mean *S. asteroides* Inflorescences (±SE)	Mean Overall Visits (±SE)	S	H	D	E
Sympatric	6	17.5 ± 4.45	15.5 ± 4.58	3.17	0.8	0.42	0.56
Allopatric	6	20.5 ± 4.95	14.5 ± 5.69	3.17	0.76	0.41	0.53
Pre-removal	5	12.4 ± 2.22	35 ± 6.23	4.17	1.03	0.57	0.78
Post-removal	5	11 ± 2.03	16 ± 5.15	3	0.75	0.42	0.63

Sericocarpus asteroides attracted only eight of the 16 insect families observed on *D. carota* – the three families of bees and two families of butterflies, one wasp family, and two families of flies (FigShare). The distribution of the number of visitors was

similar between species, with the majority comprised of sweat bees (Halictidae) followed by families of flies (Calliphoridae, Muscidae, and Syrphidae). Weather was fairly constant within and between years.

In the observational study, *S. asteroides* within sympatric and allopatric populations received 93 and 87 total visits, respectively (Table 1). Diversity indices of pollinator visitors were the same or higher for the sympatric populations, and the number of inflorescences and mean visits were not significantly different between sympatric and allopatric populations (Table 1). Total visits standardized by the number of *S. asteroides* inflorescences and by total inflorescences were not significantly different between sympatric and allopatric populations (Figure 1).

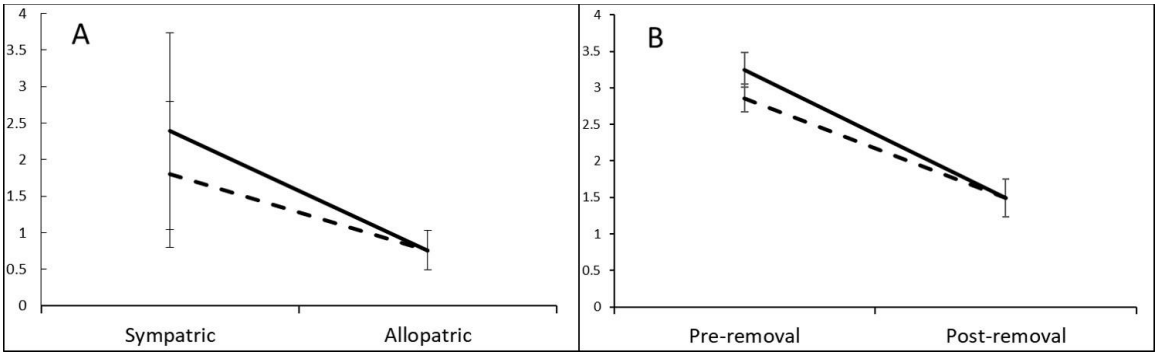


Figure 1. The number of pollinator visits to *Sericocarpus asteroides* L. A) Sympatric plots were not significantly different than allopatric plots, either after standardizing by the number of S. asteroides inflorescences (solid line) or by total inflorescences (dashed line). B) Plots of S. asteroides received significantly fewer visits following the removal of *Daucus carota* L. inflorescences after standardizing by both the number of S. asteroides inflorescences and total inflorescences ($p = 0.003$ and $p = 0.009$, respectively).

In the removal experiment, *S. asteroides* within the pre-removal plots received a total of 175 visits, and those within the post-removal plots received a total of 81 visits (Table 1; Figure 1). Diversity indices for pollinator visitors were higher for the pre-removal plots, and the mean number of inflorescences did not differ between pre- and post-removal plots. Total pollinator visits did not differ between pre- and post-removal

plots (Table 1). Total visits standardized by the number of *S. asteroides* inflorescences and by the total number of inflorescences were significantly different between pre- and post-removal plots ($p = 0.002$ and $p = 0.005$, respectively; Figure 1).

Heterospecific Pollen Assay. A total of 3292 stigmas from 325 inflorescences were assayed for heterospecific pollen from 23 *S. asteroides* populations across Nantucket (Table 2). Population distances from *D. carota* ranged from 10 m to 1172 m. Of the 83991 pollen grains counted, 7023 (8.4%) were scored as heterospecific pollen, with *S. asteroides* pollen comprising the remainder. Of the heterospecific pollen, only 91 (0.11% of total) were identified as *D. carota* (Table 2). *Sericocarpus asteroides* populations closer to *D. carota* had higher numbers of total pollen and standardized pollen per stigma compared to those farther from *D. carota* (each $p = 0.0001$). There was no significant relationship between the amount of heterospecific pollen and the closest *D. carota* population.

Table 2. *Sericocarpus asteroides* L. pollen counts and the percentage of the total number of pollen grains by taxa and overall.

Taxa	**Count**	**Percent of total**
Sericocarpus asteroides	76968	91.64
Daucus carota	91	0.11
Achillea millefolium	2599	3.09
Rosa virginiana	206	0.25
Fabaceae	146	0.17
Leucanthemum sp.	3869	4.61
Unknown1	73	0.09
Unknown2	11	0.01
Unknown3	18	0.02
LrgUnknown	10	0.01
Total Heterospecific	7023	8.36
Total	83991	100

Discussion

The shared pollination ecology of *D. carota* and *S. asteroides* poses a potential risk to the native species. Here, we have confirmed that both species are generalist-pollinated. By far the most common pollinators were sweat bees (Halictidae), followed by three families of flies (Muscidae, Calliphoridae, and Syrphidae). Presumably due to its smaller, less attractive inflorescences, *S. asteroides* received fewer pollinator visits and had a less diverse pollinator community overall. It is possible that *S. asteroides* has a pollinator visitation peak outside of the observed time frame. Although anecdotal, we were able to observe *S. asteroides* sites at various times throughout the day, but there was usually little pollinator activity during early to mid-afternoon (pers. obs.).

Following the magnet species hypothesis (Laverty 1992; Morales and Traveset 2009), one would predict that the number of pollinator visits and the diversity of pollinators found on *S. asteroides* would increase when in the presence of *D. carota*. Such was the case on native species in the presence of the invasive *Carpobrotus affine acinaciformis* (L.) L. Bolus in a study by Bartomeus et al. (2008). We also found this trend in our study with an increase in visits to *S. asteroides*, whereby *D. carota* acts as the magnet species. Moreover, *S. asteroides* received significantly fewer pollinator visits following the removal of *D. carota* and its magnet effect.

Although the diversity of pollinator visitors between sympatric and allopatric plots was essentially the same, pollinator visitor diversity decreased in plots post-removal (Table 1). We observed post-removal plots only one day after removal. It would be worthwhile to observe pollinator activity over an extended period of time. Subsequent observations could be made during which *D. carota* inflorescences would re-emerge in

the community. Based on our data, we predict that during the re-emergence of *D. carota* inflorescences, pollinator activity would increase as the magnet effect is restored. This aligns with Laverty (1992) in which higher fruit and seed set of *Podophyllum peltatum* were found when *Pedicularis canadensis* was permitted to flower following a removal experiment.

The greater difference in the diversity and number of pollinator visitors to *S. asteroides* in the removal experiment compared to the observational study is noteworthy (Table 1; Figure 1). This is likely a result of our inability to control for site, whereas the pre- and post-removal plots were paired from one day to the next. Thus, we effectively controlled for extraneous factors (e.g., the presence of human activity, local pollinator community composition, and particularly the presence of other flowering species). Given that the pollinator species composition in an area does not shift from one day to the next, *S. asteroides* should have received the same diversity of pollinator visitors after the removal of *D. carota* umbels, but with the removal of the magnet species, the diversity decreased.

Interestingly, given these findings, there is no evidence for an increase in heterospecific pollen. Yet it was present on nearly every inflorescence assayed and was essentially equally common at all distances from *D. carota*. Site-specific differences in the floral community existed, but in sites we observed, the floral communities typically consisted of a combination of *A. millefolium, R. virginiana,* two *Ranunculus* species, and two *Leucanthemum* species (pers. obs.). As we did not perform community-wide assessments, we did not factor this information into our analyses. It is possible the intervening topography contained additional flowering species for which we did not

consider. Furthermore, the most proximal *D. carota* population to any given *S. asteroides* population may not be the most ecologically-relevant. Pollinating insects are not likely to follow straight paths, but rather zig-zag between patches or engage in trapline foraging (Ohashi and Thomson 2009).

An increase in pollinator visitation would seem beneficial, though we did not evaluate fertilization or seed set. However, as outlined in Morales and Traveset (2008) fertilization and seed set may be impacted by the transfer of pollen in a frequency-dependent manner during foraging, and Dietzsch et al. (2011) found decreasing conspecific pollen with a corresponding increase in heterospecific pollen on the native *Digitalis purpurea* L. in the presence of an invasive. We observed many pollinators traveling between the two species (pers. obs.), potentially transferring pollen. Even though the magnitude of heterospecific pollen levels can vary greatly (Morales and Traveset 2008), the presence of only 91 *D. carota* pollen grains (~ 0.1%) is very striking and unexpected (Table 2). The number of *S. asteroides* inflorescences with *D. carota* pollen present was quite small (data not shown). We anticipated more interspecific pollen transfer from *D. carota*, yet almost all heterospecific pollen is from other taxa. We did find more overall pollen deposited on stigmas closer to *D. carota*, and as we found no increase in heterospecific pollen, the overall increase is due to an increase in conspecific pollen.

As plants are the foundation of almost all ecosystems, it is necessary to better understand how non-native species impact native species. In a set of observational studies, we showed that the presence of *D. carota* altered pollinator activity found on *S. asteroides*, principally after the removal of *D. carota* umbels and its magnet effect.

Second, the altered pollinator ecology increased conspecific, but not heterospecific, pollen on *S. asteroides*. These results are particularly noteworthy on Nantucket since it is known that plant species are more at risk from the impacts of invasive species (Pyšek et al. 2012), and as of 1996, introduced species comprise 39% (489 species) of the island's flora (Sorrie and Dunwiddie 1996). These results add to the growing body of knowledge about the various impacts non-native species have on native species and their communities and highlights the continued need for research in this field.

Acknowledgements

We thank Sarah Oktay, Jennifer Karberg, Karen Beattie, Sarah Bois, Kelly Omand, and Rachael Freeman for assistance on Nantucket. We also thank Bruce Mandel, Woody Lindley, and Neil Foley for transportation across Nantucket and to Noah Karberg for granting access to, and transportation on, the Nantucket Airport. We thank the Nantucket Islands Land Bank, Nantucket Conservation Foundation, and Linda Loring Nature Foundation for permission to perform this research. We thank E. Keith Bowers and Bernie J. Daigle, Jr. for assistance with statistical analyses. This project was funded in part by two Small Research Grants from the Nantucket Biodiversity Initiative, one Graduate Student Research Award from the New England Botanical Club, each awarded to AJR, and a United States Department of Agriculture-National Institute of Food and Agriculture Award 2014-33522-21826 awarded to JRM.

Literature Cited

Ahmad, M. and M. Aslam. 2002. Pollinators visiting carrot (*Daucus carota* L.) seed crop. J. Res. 13:31–35.

Albrecht, M., B. Schmid, Y. Hautier, and C. B. Müller. 2012. Diverse pollinator

communities enhance plant reproductive success. Proc. R. Soc. B Biol. Sci. 279:4845–4852.

Bartomeus, I., M. Vilà, and L. Santamaría. 2008. Contrasting effects of invasive plants in plant-pollinator networks. Oecologia. 155:761–770.

Bicknell, E. P. 1915. The Ferns and Flowering Plants of Nantucket-XVI. Bull. Torrey Bot. Club. 42:549–570.

Bjerknes, A. L., Ø. Totland, S. J. Hegland, and A. Nielsen. 2007. Do alien plant invasions really affect pollination success in native plant species? Biol. Conserv. 138:1–12.

Bohart, G. E. and W. P. Nye. 1960. Insect pollinators of carrots in Utah. Bull. Utah agric. Exp. Stn. 4:3–15.

Brown, B. J., R. J. Mitchell, and S. A. Graham. 2002. Competition for pollination between an invasive species (purple loosestrife) and a native congener. Ecology. 83:2328–2336.

Bruckman, D. and D. R. Campbell. 2016. Pollination of a native plant changes with distance and density of invasive plants in a simulated biological invasion. Am. J. Bot. 103:1458–1465.

Buckley, Y. M., B. M. Bolker, and M. Rees. 2007. Disturbance, invasion and re-invasion: Managing the weed-shaped hole in disturbed ecosystems. Ecol. Lett. 10:809–817.

Dietzsch, A. C., D. A. Stanley, and J. C. Stout. 2011. Relative abundance of an invasive alien plant affects native pollination processes. Oecologia. 167:469–479.

van Diggelen, Frank, Enge, P. 2015. The World's first GPS MOOC and Worldwide Laboratory using Smartphones., pp. 361–369. Proceedings of the 28th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION

GNSS+ 2015). Tampa, Florida.

Donaldson, J., I. Nänni, C. Zachariades, and J. Kemper. 2002. Effects of Habitat Fragmentation on Pollinator Diversity and Plant Reproductive Success in Renosterveld Shrublands of South Africa. Conserv. Biol. 16:1267–1276.

Fang, Q. and S. Q. Huang. 2013. A directed network analysis of heterospecific pollen transfer in a biodiverse community. Ecology. 94:1176–1185.

Farnsworth, E. J. 2007. Plant life history traits of rare versus frequent plant taxa of sandplains: Implications for research and management trials. Biol. Conserv. 136:44–52.

Flanagan, R. J., R. J. Mitchell, and J. D. Karron. 2010. Increased relative abundance of an invasive competitor for pollination, *Lythrum salicaria*, reduces seed number in *Mimulus ringens*. Oecologia. 164:445–54.

Frey, B., C. Kempler, and D. L. Ehret. 2007. Micro-propagation of White-top Aster, *Sericocarpus rigidus*, a threatened species from the Garry Oak Ecosystem in British Columbia. Can. Field-Naturalist. 121:40–45.

Giblin, D. 1997. *Aster curtus*: current knowledge of its biology and threats to its survival., pp. 93–101. *In* Dunn;P. and K. Ewing, eds. Ecology and Conservation of the South Puget Sound Prairie Landscape. The Nature Conservancy of Washington, Seattle, WA.

Goyal, N., M. Singh, and J. Kandoria. 1989. Role of insect pollination in seed production of carrot, *Daucus carota* L. Indian Bee J. 51:89–93.

Jakobsson, A., B. Padrón, and A. Traveset. 2008. Pollen transfer from invasive *Carpobrotus* spp. to natives–A study of pollinator behaviour and reproduction

success. Biol. Conserv. 141:136–145.

Kandori, I., T. Hirao, S. Matsunaga, and T. Kurosaki. 2009. An invasive dandelion unilaterally reduces the reproduction of a native congener through competition for pollination. Oecologia. 159:559–569.

Lacey, E. P. 1981. Seed Dispersal in Wild Carrot (*Daucus carota*). Michigan Bot. 20:15–20.

Lamborn, E. and J. Ollerton. 2000. Experimental assessment of the functional morphology of inflorescences of *Daucus carota* (Apiaceae): Testing the "fly catcher effect." Funct. Ecol. 14:445–454.

Larson, D. L., R. a. Royer, and M. R. Royer. 2006. Insect visitation and pollen deposition in an invaded prairie plant community. Biol. Conserv. 130:148–159.

Laverty, T. M. 1992. Plant Interactions for Pollinator Visits: A Test of the Magnet Species Effect. Oecologia. 89:502–508.

Lopezaraiza-Mikel, M. E., R. B. Hayes, M. R. Whalley, and J. Memmott. 2007. The impact of an alien plant on a native plant-pollinator network: An experimental approach. Ecol. Lett. 10:539–550.

McLernon, S. M., S. D. Murphy, and L. W. Aarssen. 1996. Heterospecific Pollen Transfer Between Sympatric Species in a Midsuccessional Old-Field Community. Am. J. Bot. 83:1168–1174.

Miller, J. H. J. H. and B. Edwards. 1983. Kudzu: Where did it come from and how can we stop it? South. J. Appl. For. 7:165–169.

Mitchell, R. J., R. J. Flanagan, B. J. Brown, N. M. Waser, and J. D. Karron. 2009. New frontiers in competition for pollination. Ann. Bot. 103:1403–1413.

Mitich, L. W. 1996. Wild carrot (*Daucus carota* L.). Weed Technol. 10:455–457.

Morales, C. L. and A. Traveset. 2008. Interspecific pollen transfer: Magnitude, prevalence and consequences for plant fitness. CRC. Crit. Rev. Plant Sci. 27:221–238.

———. 2009. A meta-analysis of impacts of alien vs. native plants on pollinator visitation and reproductive success of co-flowering native plants. Ecol. Lett. 12:716–728.

Muñoz, A. A. and L. A. Cavieres. 2008. The presence of a showy invasive plant disrupts pollinator service and reproductive output in native alpine species only at high densities. J. Ecol. 96:459–467.

Mustajärvi, K., P. Siikamäki, S. Rytkönen, and A. Lammi. 2001. Consequences of plant population size and density for plant-pollinator interactions and plant performance. J. Ecol. 89:80–87.

Novoa, A., J. J. Le Roux, M. P. Robertson, J. R. U. Wilson, and D. M. Richardson. 2015. Introduced and invasive cactus species: A global review. AoB Plants. 7. doi: 10.1093/aobpla/plu078.

Ohashi, K. and J. D. Thomson. 2009. Trapline foraging by pollinators: Its ontogeny, economics and possible consequences for plants. Ann. Bot. 103:1365–1378.

Owen, M. L. 1888. A Catalogue of Plants Growing Without Cultivation in the County of Nantucket, Mass., 1st edition. Gazette Printing Company, Northampton, Mass.

Parachnowitsch, A. L. and E. Elle. 2005. Insect visitation to wildflowers in the endangered Garry Oak, *Quercus garryana*, ecosystem of British Columbia. Can. Field-Naturalist. 119:245–253.

Peterson, R., J. P. Slovin, and C. Chen. 2010. A simplified method for differential staining of aborted and non-aborted pollen grains. Int. J. Plant Biol. 1:66–69.

Pyšek, P., V. Jarošík, P. E. Hulme, J. Pergl, M. Hejda, U. Schaffner, and M. Vilà. 2012. A global assessment of invasive plant impacts on resident species, communities and ecosystems: The interaction of impact measures, invading species' traits and environment. Glob. Chang. Biol. 18:1725–1737.

Schierenbeck, K. 2004. Japanese honeysuckle (*Lonicera japonica*) as an invasive species; history, ecology, and context. CRC. Crit. Rev. Plant Sci. 23:391–400.

Sorrie, B. A. and P. W. Dunwiddie. 1996. A flora of the vascular and non-vascular plants of Nantucket, Tuckernuck, and Muskeget Islands. Rhodora. 98:94–98.

Valtonen, A., J. Jantunen, and K. Saarinen. 2006. Flora and lepidoptera fauna adversely affected by invasive *Lupinus polyphyllus* along road verges. Biol. Conserv. 133:389–396.

Vanparys, V., P. Meerts, and A. Jacquemart. 2008. Plant–pollinator interactions: comparison between an invasive and a native congeneric species. Acta oecologica. 34:361–369.

Westmoreland, D. and C. Muntan. 1996. The Influence of Dark Central Florets on Insect Attraction and Fruit Production in Queen Anne's Lace (*Daucus carota* L.). Am. Midl Nat. 135:122–129.

CHAPTER 6: CONCLUSION

During a thorough review of the primary literature, I gained substantial knowledge of heteroplasmy from its initial discovery in 1909, to the development of modern detection methods and its apparent pervasiveness within individuals and across diverse taxa. Especially important to considerations of cytonuclear interactions is the increase in understanding of organellar genomes as not being strictly uniparentally-inherited and composed of one small molecule. They are neither purely haploid, in the strict sense, nor non-recombining. In Ramsey & Mandel (2019) we termed organellar genomes as cytoploid, defined as to encompass the various characteristics that distinguish it from both asexual (haploid) and sexual (diploid +) genomes. Plant organellar genomes, in particular, deviate from the traditional view of organellar genomes. Although the content of the mitochondrial (mt) genome (mtDNA) throughout eukaryotes and that of the plastid (pt) genome (ptDNA) within plants is highly conserved, the structure of organellar genomes is highly variable, complex, and dynamic in nature. The cytoploid nature of organellar genomes complicates the subject of cytonuclear interactions. In my research, I have narrowed the complexity of cytonuclear interactions to its interface with heteroplasmy.

Cytonuclear interactions are becoming increasingly recognized as playing a vital role in eukaryotic evolution (Havird et al., 2019). Traditionally, the study of genetically-heritable traits has focused exclusively on nuclear genes. Yet as emphasis has increased on the role of cytonuclear interactions and the methods to detect non-uniparental inheritance and heteroplasmy have advanced in modern times, the impact of these phenomena in determining phenotypic traits has been scrutinized. The field of

cytonuclear interactions has attracted much attention, and new questions have arisen through recent discoveries (McCauley, 2013). With my dissertation research presented here, I have attempted to address some of these questions, yet many new lines of inquiry have piqued my interest. I have approached my research from a bottoms-up approach. I investigated cytonuclear interactions, but with the additional consideration of heteroplasmy, at the individual phenotypic and population level in the gynodioecious species *Daucus carota*.

At the individual level, heteroplasmy has the potential, and is often cited, to affect phenotypes relating to fitness (Christie et al., 2011; Shen et al., 2019). Although the results of my research are not conclusive, I do provide some evidence of an association between heteroplasmy and specific phenotypic traits. In *D. carota*, heteroplasmic individuals from certain populations survive longer than homoplasmic individuals. However, in other populations, that pattern does not hold. It is probable that the environment of the greenhouse where these plants were grown impacted their overall phenomes (Wolff et al., 2014). Only one of the 19 accessions were collected from a Memphis location, an accession most likely to experience its native environmental conditions. Surprisingly, this accession faired among the worst in terms of germination, growth, and phenology. Other accessions performed much better, chief among those are the accessions collected from Europe. The experiment was conducted in a greenhouse setting where all plants experienced the same temperature and watering regimes, conditions which are not likely to be identical across the accessions' geographic origins. Complex nuclear gene by mt gene by pt gene (GxGxG) interactions may have produced confounding factors which were not anticipated, nor precisely accounted. Even so,

survivorship to flowering among some accessions was higher in heteroplasmic individuals.

Perhaps the most conclusive data with the least amount of ambiguity presented here are the results from the study of linkage disequilibrium between heteroplasmic and homoplasmic individuals. As cytonuclear interactions are thought to involve tight linkage between nuclear and organellar genes, levels of linkage disequilibrium (LD) between organellar and nuclear genomes (cnLD) is expected to be high. High cnLD would indicate non-random assortment of alleles presumably coevolved with one another, the pairs of which maximize metabolic functions within cells. Additionally, given both the mt and pt have the same pattern of uniparental (usually maternal) inheritance of organelles and no recombination within organellar genomes, within and between organellar LD is also expected to be high. Genes in each organellar genome are believed to exist on one haploid molecule and inherited as a single unit. Genes among organellar genomes with the same pattern of uniparental inheritance would also appear to be transmitted as if physically being on the same molecule. Under this type of linkage, the absolute magnitude of LD values is expected to approach 1.

However, in two geographic regions of North America (Olympic Peninsula, Washington and Nantucket Island, Massachusetts) we found that cnLD and organellar LD varied considerably (Ramsey et al., 2019). Linkage disequilibrium of 15 nuclear allelic combinations neither varied between themselves nor among regions, indicating near random assortment. Yet cytonuclear LD between the nuclear and three organellar loci varied by each organellar locus and between regions. When samples were pooled together, the absolute magnitude of cnLD was 1 for only *cox1* in the Nantucket region.

After dividing our dataset into *cox1* heteroplasmic (*n* = 113) and homoplasmic (*n* = 152) individuals in the Nantucket region and by the pt locus in the Olympic region, we found cnLD differed between those groups. Each heteroplasmic dataset contained at least one absolute cnLD value of 1, whereas only the *cox1* homoplasmic dataset in the Nantucket region contained that just one cnLD value of 1. After a Z-transformation test, 12 absolute cnLD values were significantly higher in the heteroplasmic datasets compared to the homoplasmic datasets, whereas only one was significantly lower in the homoplasmic datasets. The two mt loci also displayed different within organellar LD, and organellar LD between the two mt loci and the pt varied within and between regions and between *cox1* heteroplasmic and homoplasmic datasets.

These results offer some intriguing interpretations. First, linkage due to uniparental inheritance of both organelles does not seem to apply to the mtDNA and ptDNA in *D. carota*. At least rare biparental inheritance of these organelles has been documented (Boblenz et al., 1990; Mandel et al., 2016). If leakage of one or both organelles is more common in *D. carota* than previously supposed, it should be expected that between organellar LD values could approach near randomness (~0). Further, recombination has been demonstrated by the 4-gamete rule (Mandel et al., 2012). Our results also support this evidence as within mt LD varied between the *cox1* pooled, heteroplasmic, and homoplasmic datasets within and between regions, and between organellar LD varied with the same pattern. As the organelles may have differing inheritance patterns from one another, and recombination within their genomes, LD could thus appear to be unstructured.

A high level of organellar genetic variation may permit the existence of multiple metabolic states. Through the interaction of various combinations of alleles in each genome, slightly divergent protein complexes could exist. Coupled with environmental variation, selection on these protein complexes and the genes from which they are derived, may allow for better acclimation within generations and adaptation across generations.

Although these results do not definitively demonstrate that heteroplasmy drives patterns of cytonuclear and organellar LD, we do demonstrate that more work in this line of inquiry, both in carrot and beyond, is warranted. Of particular interest is investigating whether isolated populations of *Daucus carota* in its native range (populations with presumably no introgression) contain heteroplasmy and whether they have the same patterns of LD observed here. If heteroplasmy does exist, it may indicate that 1) inheritance and paternal leakage is sufficient, and introgression is not necessary, to maintain heteroplasmy across generations and within populations, and 2) heteroplasmy may be driving the patterns of LD we uncovered.

Daucus carota has many appealing characteristics that make it useful when considering the relationship between heteroplasmy and cytonuclear interactions. First, it has proven as a favorable species to use in the short time required of a dissertation. Cultivation is straightforward. It has a short generation time and high-quality DNA extractions and downstream applications are easy to perform. Second, nothing about its known genetic makeup would appear to present confounding factors when addressing these questions. The nuclear genome is relatively small, at 473 Mb (Iorizzo et al., 2016), and, in addition to having cytoploid characteristics (Ramsey and Mandel, 2019) which we

deemed of importance, the organellar genomes are not exceptionally noteworthy when put in the context of all characterized plant organellar genomes. Both organellar genomes are believed to undergo recombination. The ptDNA contains 136 coding genes in a circular molecule nearly 156 kb in size, typical of most other angiosperm species sequenced to date (Ruhlman et al., 2006). The mtDNA is also typical of most other angiosperm mtDNAs. It is 281 kb in size and contains 64 coding genes in a configuration of three circular molecules. There is also evidence of DNA transfer from the mtDNA into the ptDNA (Iorizzo et al., 2012). This latter fact has been known and accounted for in this research, as we verified that our mt loci of interest were not were present in the pt. The program used in Chapter 5, icHET, also incorporates this information as it is designed to omit sequences that align to more than one of the three genomes (Phan et al., 2019).

However, as my research progressed, it became evident that, like many other study species in various fields, *D. carota* is inadequate to fully address my curiosity in this field. In fact, a comprehensive survey of heteroplasmy and its role in cytonuclear interactions across an assortment of species is needed, as many interactions may be species-specific. Unsolved questions I sought to answer through my dissertation remain: To what degree is heteroplasmy maternally-inherited, and does paternal leakage contribute to the maintenance of heteroplasmy across generations? I was unable to answer these questions with *Daucus carota* in my limited time as they would require an investigation over the course of several generations, and therefore years. A comprehensive design would need to be structured such that controlled crosses could be

performed, crosses which I was unable to perform in *D. carota* during the time of my research.

Additional questions have also arisen. To what degree is the relationship between heteroplasmy and the environment determining cytonuclear interactions? Another approach in examining cytonuclear interactions is incorporating environmental effects, some of which have been discovered. In lab-reared *Drosophila melanogaster* with specific mt mutations, temperature has been shown to affect males more than females (Ma and O'Farrell, 2015), and in crosses of wild *D. melanogaster* designed to transplant one mtDNA onto the nuclear background of a separate geographic origin, (Camus et al., 2017), latitudinal effects were discovered in fitness characteristics. Although these studies have reported heteroplasmy (and sometimes necessitated it in their design model), none have explicitly investigated the interaction of heteroplasmy and the environment in determining an effect on cytonuclear interactions. The next step would be to design studies to interrogate that interface as well. In such a design utilizing plant species, complex GxGxGxE interactions would need to be carefully teased apart. Such interactions are likely to have been involved in the greenhouse fitness study presented here. While these phenomena are complex, my dissertation research has laid the foundation for future work in *D. carota*, as well as other plant species aimed at disentangling cytonuclear interactions, fitness, and heteroplasmy.

REFERENCES

Boblenz, K., T. Nothnagel, and M. Metzlaff. 1990. Paternal inheritance of plastids in the genus *Daucus*. *MGG Molecular & General Genetics* 220: 489–491.

Camus, M. F., J. N. Wolff, C. M. Sgrò, and D. K. Dowling. 2017. Experimental evidence that thermal selection has shaped the latitudinal distribution of mitochondrial haplotypes in Australian *Drosophila melanogaster*. *Molecular Biology and Evolution* 34: 2600–2612.

Christie, J., A. Picornell, A. Moya, M. Ramon, and J. Castro. 2011. Mitochondrial DNA effects on fitness in *Drosophila subobscura*. *Heredity* 107: 239–245.

Havird, J. C., R. J. Weaver, L. Milani, F. Ghiselli, R. Greenway, A. J. Ramsey, A. G. Jimenez, et al. 2019. Beyond the Powerhouse: Integrating Mitonuclear Evolution, Physiology, and Theory in Comparative Biology. *Integrative and Comparative Biology* icz132: 1–8.

Iorizzo, M., S. Ellison, D. Senalik, P. Zeng, P. Satapoomin, J. Huang, M. Bowman, et al. 2016. A high-quality carrot genome assembly provides new insights into carotenoid accumulation and asterid genome evolution. *Nature Genetics*.

Iorizzo, M., D. Senalik, M. Szklarczyk, D. Grzebelus, D. Spooner, and P. Simon. 2012. De novo assembly of the carrot mitochondrial genome using next generation sequencing of whole genomic DNA provides first evidence of DNA transfer into an angiosperm plastid genome. *BMC Plant Biology* 12: 1–17.

Ma, H., and P. H. O'Farrell. 2015. Selections that isolate recombinant mitochondrial genomes in animals. *eLife* 4: 1–16.

Mandel, J. R., E. V. McAssey, K. M. Roland, and D. E. McCauley. 2012. Mitochondrial gene diversity associated with the *atp9* stop codon in natural populations of wild carrot (*Daucus carota* ssp. carota). *Journal of Heredity* 103: 418–425.

Mandel, J. R., A. J. Ramsey, M. Iorizzo, and P. W. Simon. 2016. Patterns of Gene Flow between Crop and Wild Carrot, *Daucus carota* (Apiaceae) in the United States. *Plos One* 11: 1–19.

McCauley, D. E. 2013. Paternal leakage, heteroplasmy, and the evolution of plant mitochondrial genomes. *New Phytologist* 200: 966–977.

Phan, V., D.-T. Pham, C. Melton, A. J. Ramsey, B. J. Daigle, and J. R. Mandel. 2019. icHET: interactive visualization of cytoplasmic heteroplasmy. *Bioinformatics*: 1–2.

Ramsey, A. J., and J. R. Mandel. 2019. When One Genome Is Not Enough: Organellar Heteroplasmy in Plants. *Annual Plant Reviews Online* 2: 1–40.

Ramsey, A. J., D. E. McCauley, and J. R. Mandel. 2019. Heteroplasmy and Patterns of Cytonuclear Linkage Disequilibrium in Wild Carrot. *Integrative and Comparative Biology* icz102: 1–11.

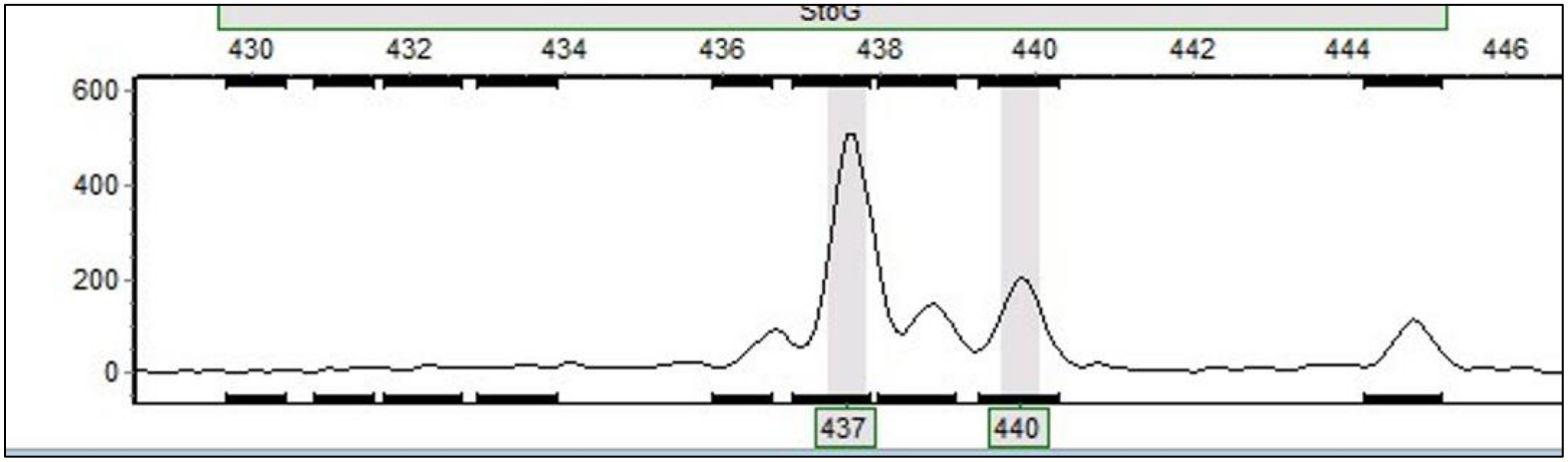
Ruhlman, T. A., S. Lee, R. K. Jansen, J. B. Hostetler, L. J. Tallon, C. D. Town, and H. Daniell. 2006. Complete plastid genome sequence of *Daucus carota*: implications for biotechnology and phylogeny of angiosperms. *BMC genomics* 7: 222.

Shen, J., W. Shou, Y. Zhang, G. Yuan, Y. Zhao, J. Chen, and M. J. Havey. 2019. Rare maternal and biparental transmission of the cucumber mitochondrial DNA reveals sorting of polymorphisms among progenies. *Theoretical and Applied Genetics* 132: 1223–1233.

Wolff, J. N., E. D. Ladoukakis, J. A. Enríquez, D. K. Dowling, J. A. Enriquez, and D. K. Dowling. 2014. Mitonuclear interactions: evolutionary consequences over multiple

biological scales. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369: 20130443–20130443.

APPENDIX 1: CHAPTER 4 SUPPLEMENTARY MATERIAL



Supplementary Figure 1. Image of GeneMarker StoG analysis of a potential tri-heteroplasmic individual from the Olympic region. Three alleles are present, a major 437 (called 438 in our analysis) and two minor, 440 and 445, alleles. Only the two largest peaks were retained in our analysis.

Supplementary Table 1. Samples sizes from 12 populations in two regions of the United States, Olympic and Nantucket. n = number of individuals assayed; n_{het} = number of individuals heteroplasmic for *cox1* and StoG, left and right of vertical line, respectively, per population and region; $\%_{het}$ = percent of individuals heteroplasmic for *cox1* and StoG, left and right of the vertical line, respectively, per population and region.

Population	State	County	*n*	n_{het}	$\%_{het}$
Eberle	Washington	Clallam	36	25 | 16	69.4 | 44.4
Fasola	Washington	Clallam	11	5 | 0	45.5 | 0.0
Fencebird	Washington	Clallam	31	4 | 17	12.9 | 54.8
Hemlock	Washington	Clallam	14	1 | 0	7.1 | 0.0
Kendall	Washington	Clallam	16	1 | 0	6.3 | 0.0
Medsker	Washington	Clallam	13	2 | 1	15.4 | 7.7
Prince	Washington	Clallam	15	1 | 0	6.7 | 0.0
Olympic Total			136	39 | 34	28.7 | 25.0
Bartlett Farms	Massachusetts	Nantucket	37	11 | 1	29.7 | 0.0
Moors End	Massachusetts	Nantucket	47	42 | 0	89.4 | 0.0
Tuckernuck	Massachusetts	Nantucket	16	8 | 0	50.0 | 0.0
Cliff	Massachusetts	Nantucket	19	6 | 0	31.6 | 0.0
Polpis	Massachusetts	Nantucket	10	7 | 1	70.0 | 0.0
Nantucket Total			129	74 | 2	57.4 | 1.6
Overall Total			**265**	**113 | 36**	**42.6 | 13.6**

Supplementary Table 2. Normalized nuclear linkage disequilibrium (*D*') among 14 simple sequence repeat loci in the Olympic region. All values, except those in bold are significantly different from zero based on Fisher's Exact Tests and after a Benjamini-Hochberg correction for multiple tests.

Locus	GSSR3	GSSR4	GSSR6	GSSR7	GSSR11	GSSR16	GSSR24	GSSR31	GSSR35	GSSR57	GSSR65	GSSR85	GSSR107	GSSR111
GSSR3	-													
GSSR4	0.037	-												
GSSR6	0.247	-0.096	-											
GSSR7	0.136	0.245	0.214	-										
GSSR11	-0.006	-0.144	0.057	-0.122	-									
GSSR16	-0.103	0.089	-0.193	-0.176	0.002	-								
GSSR24	0.050	-0.055	0.011	0.157	0.127	**0.047**	-							
GSSR31	0.017	-0.120	0.105	-0.105	0.089	-0.352	0.055	-						
GSSR35	**0.075**	0.031	0.005	-0.098	0.250	-0.626	-0.299	0.178	-					
GSSR57	-0.053	0.019	0.025	-0.080	0.033	0.069	-0.332	0.019	0.120	-				
GSSR65	**-0.033**	**-0.244**	**-0.131**	**0.011**	**-0.351**	**-0.019**	**0.017**	**0.110**	**-0.133**	-0.179	-			
GSSR85	**-0.334**	**0.026**	-0.168	0.136	-0.095	**-0.053**	0.086	-0.147	**0.005**	-0.136	-0.378	-		
GSSR107	-0.395	0.021	**0.053**	0.153	0.070	**0.033**	0.356	0.088	0.187	-0.260	0.009	0.184	-	
GSSR111	**-0.582**	-0.530	-0.169	-0.428	-0.177	-0.398	0.154	0.154	-0.066	-0.082	0.108	-0.397	0.093	-

Supplementary Table 3: Nuclear linkage disequilibrium r^2 values among 14 simple sequence repeat loci in the Olympic region.

Locus	**GSSR3**	**GSSR4**	**GSSR6**	**GSSR7**	**GSSR11**	**GSSR16**	**GSSR24**	**GSSR31**	**GSSR35**	**GSSR57**	**GSSR65**	**GSSR85**	**GSSR107**	**GSSR111**
GSSR3	-													
GSSR4	0.00084	-												
GSSR6	0.03602	0.00088	-											
GSSR7	0.01045	0.05633	0.04374	-										
GSSR11	1.00E-05	0.00526	0.00121	0.00355	-									
GSSR16	0.00227	0.00606	0.00469	0.00375	0	-								
GSSR24	0.00197	0.00038	9.00E-05	0.01778	0.00814	0.00217	-							
GSSR31	0.00027	0.00204	0.00741	0.00147	0.0044	0.02304	0.00274	-						
GSSR35	0.00221	0.00062	2.00E-05	0.00057	0.0157	0.03262	0.00745	0.01417	-					
GSSR57	0.00078	0.0002	0.00035	0.00103	0.00071	0.00356	0.0246	0.0003	0.00533	-				
GSSR65	0.00025	0.00836	0.00234	7.00E-05	0.04485	6.00E-05	0.00025	0.01197	0.0016	0.00777	-			
GSSR85	0.01895	0.00063	0.00282	0.01668	0.0024	0.00038	0.00584	0.00322	2.00E-05	0.00328	0.02083	-		
GSSR107	0.02251	0.0004	0.00253	0.02204	0.00168	0.00076	0.08575	0.0047	0.02568	0.01022	5.00E-05	0.02862	-	
GSSR111	0.02854	0.01434	0.00142	0.0088	0.00413	0.01054	0.00946	0.00846	0.00015	0.00059	0.00425	0.00835	0.00503	-

Supplementary Table 4. Values of normalized cytonuclear linkage disequilibrium (*D*') between 14 (Olympic) and 15 (Nantucket) nuclear and three organellar loci (mtDNA: a*tp9*, c*ox1*; ptDNA, StoG region) and the heteroplasmic and homoplasmic partitions for *cox1* and StoG. *atp9* within each region and StoG in the Nantucket region. Hom = homoplasmic dataset; Het = heteroplasmic dataset. * represent *D*' significantly different from zero based on Fisher's Exact Tests and after a Benjamini-Hochberg correction for multiple tests.

Nuclear locus	**Olympic**							**Nantucket**				
	atp9	***cox1 Pooled***	***cox1 Hom***	***cox1 Het***	**StoG *Pooled***	**StoG *Hom***	**StoG *Het***	***atp9***	***cox1 Pooled***	***cox1 Hom***	***cox1 Het***	**StoG**
GSSR3	.2133	-0.1598	-0.1107	-0.4848	-0.0513	0.1332	-0.0885	-0.109	-0.769*	-0.899*	-0.299	-0.0087
GSSR4	-.3865*	0.5986*	0.5222*	1	-0.155	-0.1594	0.1667	-0.014	-0.259	-0.352*	-0.129	0.1864
GSSR6	.1512	0.1945	0.2115	-0.0435	-0.0105	-0.027	-0.35	0.337	0.211	-0.184	0.609	-0.3283
GSSR7	.2549	0.2403	0.8167*	0.3968	-0.22	-0.1711	0.8239*	0.015	-0.860*	-0.857*	-1	-0.2301*
GSSR9	-	-	-	-	-	-	-	0.011	0.812	-0.633*	0	0.2062
GSSR11	.0808	-0.4139*	-0.3841*	-0.24	0.0234	-0.0938	0.4286	0.026	0.372	-0.17	1	-0.0392
GSSR16	.0145	-0.0286	-0.0704	0.1014	0.2285	0.4082*	-0.0333	0.077	0.364	-0.205	-0.097	-0.1673
GSSR24	-.1761	0.1248	0.1501	1	0.0359	0.0134	-0.0425	0.028	0.074	-0.037	-0.036	-0.1838
GSSR31	0.0522	-0.0614	-0.0385	-0.0617	0.1980	0.0149	1*	-0.075	0.715*	0.707*	1	0.4*
GSSR35	0.1603	-0.3987*	-0.4023*	1	-0.087	-0.3864	-0.026	-0.284*	-0.405*	-0.333	-0.377	0.0889
GSSR57	0.3294	-0.2045	-0.3284	0.4697	0.5574*	0.4831*	1	0.360	1*	1	1	0.1
GSSR65	-0.1676	0.2544	0.2378	0.1429	0.1429	0.1677	0.2105	-0.203	-0.231	-0.296	-0.236	-0.5701*
GSSR85	-0.0827	0.0024	0.0449	-0.1111	0.0303	0.2582	0.4	-0.039	0.112	0.253	-0.250	0.0333
GSSR107	0.2411	-0.1442	-0.1236	1	0.2917	0.3878*	-0.4766*	0.301*	0.804	0.826*	0.704	0.1467
GSSR111	-0.1047	0.3553	0.5046	-0.12	0.2152	0.1005	1	-0.018	-0.205	-0.536*	1	-0.0179*

Supplementary Table 5. Values of normalized cytonuclear linkage disequilibrium (*D*') and r^2 values between 14 (Olympic) and 15 (Nantucket) nuclear and the heteroplasmic and homoplasmic partitions for two organellar loci (mtDNA: c*ox1*; ptDNA, StoG region). Hom = homoplasmic dataset; Het = heteroplasmic dataset. * represent *D*' significantly different from zero based on Fisher's Exact Tests and after a Benjamini-Hochberg correction for multiple tests.

Nuclear locus	**Olympic**								**Nantucket**			
	cox1				**StoG**				***cox1***			
	Hom		***Het***		***Hom***		***Het***		***Hom***		***Het***	
	D'	r^2	*D'*	r^2	*D'*	r^2	*D'*	r^2	*D'*	r^2	*D'*	r^2
GSSR3	-0.1107	0.0074	-0.4848	0.0418	0.1332	0.0106	-0.0885	0.0044	-0.899*	0.1722	-0.299	0.0105
GSSR4	0.5222*	0.0005	1	0.0398	-0.1594	0.5436	0.1667	0.0018	-0.352*	0.1215	-0.129	0.0103
GSSR6	0.2115	0.0076	-0.0435	0.0004	-0.027	0.0002	-0.35	0.0466	-0.184	0.0331	0.609	0.0096
GSSR7	0.8167*	0.1279	0.3968	0.0052	-0.1711	0.0096	0.8239*	0.1297	-0.857*	0.2109	-1	0.0005
GSSR9	-	-	-	-	-	-	-	-	-0.633*	0.0425	0	0.0005
GSSR11	-0.3841*	0.0581	-0.24	0.0095	-0.0938	0.0038	0.4286	0.0520	-0.17	0.0509	1	0.0482
GSSR16	-0.0704	0.0034	0.1014	-0.0002	0.4082*	0.1046	-0.0333	0.0004	-0.205	0.0043	-0.097	0.0049
GSSR24	0.1501	0.6319	1	0.0285	0.0134	-0.0005	-0.0425	0.0008	-0.037	0.0001	-0.036	0.0013
GSSR31	-0.0385	0.0012	-0.0617	0.0009	0.0149	0.0001	1*	0.2352	0.707*	0.0772	1	0.0397
GSSR35	-0.4023*	0.0828	1	0.1021	-0.3864	0.0250	-0.026	0.0004	-0.333	0.0825	-0.377	0.0067
GSSR57	-0.3284	0.0159	0.4697	0.0120	0.4831*	0.0679	1	0.0014	1	0.2170	1	0.0001
GSSR65	0.2378	0.0210	0.1429	0.0012	0.1677	0.0100	0.2105	0.0051	-0.296	0.0444	-0.236	0.0035
GSSR85	0.0449	0.0005	-0.1111	0.0033	0.2582	0.0324	0.4	0.0242	0.253	0.0359	-0.250	0.0378
GSSR107	-0.1236	0.0098	1	0.0221	0.3878*	0.0878	-0.4766*	0.1595	0.826*	0.1441	0.704	0.0208
GSSR111	0.5046	0.0279	-0.12	0.0034	0.1005	0.0026	1	0.0421	-0.536*	0.2526	1	0.0007

Supplementary Table 6. Normalized nuclear linkage disequilibrium (*D*') among 14 simple sequence repeat loci in the Nantucket region. All values, except those in bold are significantly different from zero based on Fisher's Exact Tests and after a Benjamini-Hochberg correction for multiple tests.

Locus	GSSR3	GSSR4	GSSR6	GSSR7	GSSR9	GSSR11	GSSR16	GSSR24	GSSR31	GSSR35	GSSR57	GSSR65	GSSR85	GSSR107	GSSR111
GSSR3	-														
GSSR4	0.338	-													
GSSR6	0.078	-0.103	-												
GSSR7	-0.089	0.172	0.106	-											
GSSR9	-0.308	-0.615	-0.109	-0.164	-										
GSSR11	-0.102	-0.456	-0.102	-0.482	0.020	-									
GSSR16	-0.544	-0.005	-0.427	-0.005	0.140	-0.239	-								
GSSR24	**-0.225**	-0.663	-0.139	-0.494	0.014	0.455	0.068	-							
GSSR31	0.026	0.371	0.117	0.318	0.109	-0.493	0.363	-0.370	-						
GSSR35	-0.100	0.014	-0.375	-0.152	0.077	0.294	0.122	0.076	-0.176	-					
GSSR57	0.010	**0.236**	0.106	0.395	0.054	-0.522	-0.104	-0.297	0.513	-0.037	-				
GSSR65	0.093	-0.179	0.071	-0.015	-0.014	**0.049**	-0.198	**0.130**	0.003	0.068	0.021	-			
GSSR85	0.141	0.137	0.140	0.094	0.082	-0.253	-0.056	-0.114	0.128	-0.575	0.070	0.045	-		
GSSR107	0.001	0.244	0.233	0.038	0.371	-0.007	-0.051	**-0.198**	0.247	-0.314	0.003	-0.118	0.222	-	
GSSR111	-0.375	0.030	-0.447	0.079	0.106	0.032	0.415	0.121	0.273	0.130	-0.147	-0.313	-0.191	0.138	-

Supplementary Table 7: Nuclear linkage disequilibrium r^2 values among 15 simple sequence repeat loci in the Nantucket region.

Locus	GSSR3	GSSR4	GSSR6	GSSR7	GSSR9	GSSR11	GSSR16	GSSR24	GSSR31	GSSR35	GSSR57	GSSR65	GSSR85	GSSR107	GSSR111
GSSR3	-														
GSSR4	0.03589	-													
GSSR6	0.00296	0.00116	-												
GSSR7	0.00481	0.01086	0.00635	-											
GSSR9	0.01469	0.01836	0.00089	0.00355	-										
GSSR11	0.00758	0.04788	0.0037	0.1457	8.00E-05	-									
GSSR16	0.05712	0	0.01705	0	0.01573	0.01131	-								
GSSR24	0.02099	0.05723	0.00388	0.08637	8.00E-05	0.1169	0.00215	-							
GSSR31	0.00058	0.05234	0.0081	0.09825	0.00312	0.14734	0.0433	0.04692	-						
GSSR35	0.00268	0.00017	0.01825	0.00528	0.0034	0.03147	0.01083	0.00373	0.0068	-					
GSSR57	9.00E-05	0.0204	0.00635	0.15588	0.00074	0.17074	0.00179	0.03129	0.2551	0.00031	-				
GSSR65	0.00815	0.00675	0.00262	0.00013	3.00E-05	0.0022	0.00713	0.01047	1.00E-05	0.00185	0.00041	-			
GSSR85	0.00955	0.01227	0.01935	0.00488	0.00303	0.02241	0.00029	0.00259	0.0095	0.04208	0.00276	0.00105	-		
GSSR107	0	0.01838	0.02595	0.00119	0.02955	3.00E-05	0.00052	0.01643	0.04972	0.02668	1.00E-05	0.00936	0.02319	-	
GSSR111	0.03291	0.00087	0.02272	0.00238	0.00743	0.00032	0.14173	0.00829	0.02953	0.01488	0.0043	0.0215	0.00405	0.00619	-

Supplementary Table 8. Carrot population genetic diversity measures for nuclear, chloroplast, and mitochondrial loci. uHe = unbiased nuclear gene diversity; cpuH = chloroplast unbiased diversity and mtuH = mitochondrial unbiased diversity (Peakall and Smouse 2012).

Population	Nuclear uHe	cpuH	mtuH
Olympic			
Eberle	0.76	0.48	0.19
Fasola	0.62	0.33	0
Fencebird	0.70	0.40	0.26
Hemlock	0.80	0.62	0.26
Kendall	0.67	0.46	0.32
Medsker	0.78	0.69	0.33
Prince	0.69	0.25	0.06
Nantucket			
Bartlett Farms	0.73	0.47	0.47
Moors End	0.72	0.04	0.11
Tuckernuck	0.80	0.00	0.24
Cliff	0.60	0.53	0.31
Polpis	0.64	0.56	0.28
Olympic F_{ST}	0.12	0.30	0.31
Nantucket F_{ST}	0.12	0.43	0.15

Supplementary Table 9. Sorted and color coded absolute cytonuclear linkage disequilibrium (|D'|) between 14 (Olympic) and 15 (Nantucket) nuclear and three organellar loci (mtDNA: a*tp9*, c*ox1*; ptDNA, StoG region) and the heteroplasmic and homoplasmic partitions for *cox1* and StoG. *atp9* within each region and StoG in the Nantucket region. Hom = homoplasmic dataset; Het = heteroplasmic dataset; OLY = Olympic region; NAN = Nantucket region.

	Olympic							Nantucket				
	atp9	*cox1*	*cox1*	*cox1*	*StoG*	*StoG*	*StoG*	*atp9*	*cox1*	*cox1*	*cox1*	*StoG*
		Pooled	*Hom*	*Het*	*Pooled*	*Hom*	*Het*		*Pooled*	*Hom*	*Het*	
	-	-	-	-	-	-	-	0.01	0.07	0.04	0.00	0.01
	0.01	0.00	0.04	0.04	0.01	0.01	0.03	0.01	0.11	0.17	0.04	0.02
	0.05	0.03	0.04	0.06	0.02	0.01	0.03	0.02	0.21	0.18	0.10	0.03
	0.08	0.06	0.07	0.10	0.03	0.03	0.04	0.02	0.21	0.21	0.13	0.04
	0.08	0.12	0.11	0.11	0.04	0.09	0.09	0.03	0.23	0.25	0.24	0.09
	0.10	0.14	0.12	0.12	0.05	0.10	0.17	0.03	0.26	0.30	0.25	0.10
	0.15	0.16	0.15	0.14	0.09	0.13	0.21	0.04	0.36	0.33	0.30	0.15
	0.16	0.19	0.21	0.24	0.14	0.16	0.35	0.08	0.37	0.35	0.38	0.17
	0.17	0.20	0.24	0.40	0.16	0.17	0.40	0.08	0.41	0.54	0.61	0.18
	0.18	0.24	0.33	0.47	0.20	0.17	0.43	0.11	0.72	0.63	0.70	0.19
	0.21	0.25	0.38	0.48	0.22	0.26	0.48	0.20	0.77	0.71	1.00	0.21
	0.24	0.36	0.40	1.00	0.22	0.39	0.82	0.28	0.80	0.83	1.00	0.23
	0.25	0.40	0.50	1.00	0.23	0.39	1.00	0.30	0.81	0.86	1.00	0.33
	0.33	0.41	0.52	1.00	0.29	0.41	1.00	0.34	0.86	0.90	1.00	0.40
	0.39	0.60	0.82	1.00	0.56	0.48	1.00	0.36	1.00	1.00	1.00	0.57
Median	0.16	0.20	0.22	0.32	0.15	0.16	0.38	0.08	0.37	0.35	0.38	0.17
Mean	0.17	0.23	0.28	0.44	0.16	0.20	0.43	0.13	0.48	0.49	0.52	0.18
Standard Deviation	0.10551	0.17	0.23	0.39	0.15	0.16	0.38	0.13	0.31	0.31	0.4	0.16
Standard Error	0.0282	0.04	0.06	0.11	0.04	0.04	0.1	0.03	0.08	0.08	0.1	0.04

APPENDIX 2: CHAPTER 5 SUPPLEMENTARY MATERIAL

Appendix S1. Populations of *Daucus carota* located across Nantucket Island, Massachusetts with GPS coordinates, distance to the ocean and their average population density calculated from up to five density measurements.

Location No.	Location Site	Latitude	Longitude	Ocean Distance	Mean Density
1	Eel Point, LLF Property, Roadside.1	41.2932110	-70.1669590	98.15	6.2
2	Nantucket Field Station, Driveway	41.2950780	-70.0404140	115.76	21.2
3	Nantucket Field Station, field	41.2960030	-70.0400360	34.69	11.2
4	Nantucket Field Station, Sea Bluff	41.2967560	-70.0389810	6.77	8
5	M6 Condo, Backyard.1	41.2733780	-70.1012720	944.09	10
6	M6 Condo, Backyard.2	41.2730500	-70.1018060	1001	3
7	Eel Point, LLF Property, Roadside.2	41.2914060	-70.1761000	265.77	63.8
8	Eel Point, LLF Property, Roadside.3	41.2919690	-70.1748360	217.2	13.4
9	Eel Point, LLF Property, Roadside.4	41.2921060	-70.1742750	204.55	28.4
10	118 Bellevue, Eel Point Rd	41.2924690	-70.1722030	181.45	30.2
11	5 Little Neck Rd	41.2794970	-70.1931250	273.86	18.6
12	Entrance Rd, Little Neck	41.2793000	-70.1935720	250.08	17.6
13	Little Neck Parking Lot, Beachside	41.2797390	-70.1941250	6.57	53.5
14	10 Blue Heron Way	41.2816690	-70.1904830	111.26	8.4
15	NILB Parking Lot	41.2807920	-70.1919640	192.9	1.333333333
16	NILB, Pine Tree	41.2812810	-70.1920360	80.57	1
17	NCF, Sanford Farm Overlook	41.2828670	-70.1373640	1190	8.2
18	NCF, Sanford Farm, Tile Silo	41.2811580	-70.1381860	1381	7
19	Waste Options, Madaket Bike Path	41.2822930	-70.1654490	1294	4.8
20	D.P.W. Office Entrance	41.2832110	-70.1675390	1209	3
21	East of DPW Entrance	41.2822150	-70.1639630	1291	3.666666667
22	Madaket, 3 Brick Stone, Roadside	41.2818400	-70.1552360	1288	17.6

Appendix S1 (continued).

Location No.	Location Site	Latitude	Longitude	Ocean Distance	Mean Density
23	Madaket Rd @ Worth Rd	41.2819370	-70.1581760	1297	13.6
24	NCF, Madaket Rd, West of Worth Rd	41.2819700	-70.1594810	1291	16.4
25	Madaket Rd, Bridge	41.2820310	-70.1609540	1289	7.2
26	Miacomet Rd, Loop.1	41.2455100	-70.1117830	360.52	12.8
27	Miacomet Rd, Loop.2	41.2448130	-70.1114810	276.07	27
28	Miacomet Beach Parking Lot	41.2433910	-70.1109580	107.43	13.4
29	Miacomet Rock Walking Trail	41.2470510	-70.1139980	466.27	42.4
30	Starbuck Rd, South Loop	41.2694220	-70.1978960	73.07	4.8
31	Starbuck Rd, SE Loop	41.2691900	-70.1968550	68.15	14.8
32	Proprietor Rd, Madaket Beach Parking Lot	41.2630890	-70.1805810	75.7	3
33	48 Proprietors Rd	41.2703490	-70.1939430	176.02	8
34	Pinetops, 54 Proprietors Rd	41.2713430	-70.1936440	285.84	12
35	Arkansas Ave @ S. Cambridge St.	41.2767500	-70.1869140	519	6.4
36	Head of Plains @ Red Barn Rd	41.2702190	-70.1845650	678.35	1.5
37	Red Barn Rd Near Head of Plains	41.2697860	-70.1840420	647.23	8.4
38	Red Barn Rd @ WNTX Driveway	41.2683900	-70.1827010	567.55	10.8
39	Long Pond Sanctuary, Parking Lot	41.2721420	-70.1848300	844.7	16
40	Massasoit Bridge Rd, Fork	41.2730840	-70.1728990	1342.56	8
41	36 Washington St.	41.2812490	-70.0957510	50.5	1
42	2 Fairgrounds Rd	41.2804390	-70.0944620	12.1	10
43	N Beach Rd @ Brant Pt Rd, field	41.2910750	-70.1029120	255.78	6.6
44	Bathing Beach @ Hulbert Ave.	41.2933330	-70.1050770	193.68	24.4
45	Jetties	41.2941020	-70.1049370	179.64	6.2
46	Cliffside Beach Club	41.2937860	-70.1075840	192.29	30.6
47	Brant Point Beach	41.2900410	-70.0914530	0	5

Appendix S1 (continued).

Location No.	Location Site	Latitude	Longitude	Ocean Distance	Mean Density
48	New Ln @ Franklin St.	41.2860680	-70.1086500	955.1	3.25
49	68-70 W. Chester St	41.2868590	-70.1119840	873.76	38.4
50	96-98 W. Chester St	41.2864200	-70.1160570	876.96	7.4
51	Crooked Ln @ Cliff Rd	41.2891460	-70.1208740	552.21	9.333333333
52	The Tupancy Links	41.2904550	-70.1280810	365.44	13.8
53	The Sound Bluff, The Tupancy Links	41.2935410	-70.1295460	12.49	2.5
54	The Tupancy Links, Trail.1	41.2917610	-70.1283260	214.25	5.25
55	The Tupancy Links, Trail.2	41.2904350	-70.1267740	364.9	11
56	Cliff Rd Bike Path	41.2889980	-70.1234230	507.51	3
57	186 Cliff Rd	41.2877810	-70.1281670	656.31	6
58	7 Washing Pond Rd	41.2889160	-70.1318370	546.41	9
59	27 Washing Pond Rd	41.2916060	-70.1357080	220.8	32.8
60	26 Washing Pond Rd	41.2921920	-70.1340300	160.11	6.5
61	39 Madaket Rd	41.2815610	-70.1153860	1416.8	11
62	Christian Science Society	41.2813930	-70.1095250	1166.66	1.666666667
63	Nantucket Cottage Hospital	41.2758060	-70.1014620	777.51	15.4
64	Nantucket Tackle Center	41.2719400	-70.0938710	628.41	17
65	Milestone Bike Path, 67 Milestone Rd	41.2698900	-70.0647750	1947.31	2.5
66	Larsen Acres, 4M	41.2630900	-69.9821220	1625.44	2.5
67	Milesone Overlook Trial, P.L. Entr.	41.2624970	-70.0127889	2558.86	37.5
68	Milesone Bog	41.2640060	-70.0072980	2574.95	17.33333333
69	220 Milestone Rd	41.2622760	-70.0086780	2430.11	14
70	Milesone Bog, North.	41.2742970	-70.0052660	2655.42	3
71	199 Polpis Rd, Polpis Bike Path	41.2902890	-70.0380900	702.49	6.8
72	180 Polpis Rd, Polpis Bike Path	41.2903690	-70.0402390	568.9	18.8

Appendix S1 (continued).

Location No.	Location Site	Latitude	Longitude	Ocean Distance	Mean Density
73	153 Polpis Rd, Polpis Bike Path	41.2900290	-70.0457080	465.56	2
74	Polpis Bike Path	41.2887120	-70.0501700	744.56	22.4
75	82 Polpis Rd, Polpis Bike Path	41.2829280	-70.0624170	785.45	9.6
76	55R Polpis Rd, Polpis Bike Path	41.2788090	-70.0679430	958.94	15.6
77	6 Fair St	41.2825430	-70.0996300	341.96	1.5
78	20 Vesper Lane	41.2744710	-70.1019970	902.47	2.5
79	Madaket Beach	41.2523820	-70.1527070	27.3	8.4
80	Madaket Beach Place	41.2535570	-70.1516570	183.21	3.333333333
81	Hummock Pond Rd @ Hellers Way	41.2569510	-70.1447480	734.42	1.25
82	14 Hellers Way	41.2557380	-70.1403880	680.44	39.33333333
83	Cisco Beach Bike Path	41.2580340	-70.1433310	880.01	7.5
84	199R Hummock Pond Rd	41.2601750	-70.1406370	1195.74	1
85	167 Hummock Pond Rd	41.2638770	-70.1346820	1705.9	34.4
86	Advice 5Â¢	41.2688400	-70.1247260	2462.3	5.6
87	8 Somerset Ln	41.2685730	-70.1230260	2478.39	4.8
88	18 Somerset Ln	41.2671520	-70.1221920	2349.64	2.2
89	Raceway Dr @ Bartlett Rd	41.2612140	-70.1181080	1866.84	1.5
90	Tuckernuck.1	41.3008770	-70.2485920	70.07	24
91	Madaket Harbor	41.2755810	-70.1955700	7.94	14.6
92	Burnt Swamp Ln	41.2756000	-70.1231000	2011.68	9
93	Milestone Rd @ New St	41.2621410	-69.9730050	860.89	16
94	Captain Cabin, Baxter Ave	41.2690390	-69.9625770	136.56	9.666666667
95	36 Baxter Ave	41.2703800	-69.9623020	117.21	11.5
96	Owl's Nest	41.2726090	-69.9621460	109.88	13.25
97	Baxter Field Cliff	41.2780210	-69.9623570	15.33	109.4

Appendix S1 (continued).

Location No.	Location Site	Latitude	Longitude	Ocean Distance	Mean Density
98	Sankaty Lighthouse	41.2829530	-69.9649260	31.07	7
99	Sankaty Rd @ Bayberry Sias Ln	41.2757680	-69.9640790	233.6	11.8
100	Sankaty Rd @ Isobel's Way	41.2739820	-69.9640680	262.26	29.2
101	Wells Guest Cottage	41.2743990	-69.9661460	423.23	53.6
102	Meetinghouse Ln @ Sankaty Rd	41.2711420	-69.9639140	245.79	25.6
103	20 Sconset Ave	41.2592700	-69.9651650	119.05	4.5
104	South Rd	41.2661380	-69.9656460	365.27	6.333333333
105	W Sankaty Rd @ Coffin St	41.2656980	-69.9668670	449.09	2.333333333
106	New St @ Shell St	41.2629880	-69.9641310	159.12	4
107	One Ocean Ave	41.2612640	-69.9644000	127.84	3.666666667
108	Ocean Ave @ Carew Ln	41.2589310	-69.9652730	119.51	7.8
109	19 Ocean Ave	41.2554110	-69.9686370	217.34	15
110	Low Beach Rd Beach, NILB	41.2535920	-69.9682880	121.22	5
111	Low Beach Rd	41.2511350	-69.9759930	364.75	6.6
112	Underhill Ln @ Morey Ln	41.2578560	-69.9670210	210.42	15.75
113	44 New St	41.2627670	-69.9711460	726.17	15.8
114	15 1/2 Burnell	41.2646950	-69.9691640	624.81	10
115	Blackfish @ Burnell St	41.2668360	-69.9691980	667.51	6
116	Sconset Bike Path	41.2621700	-69.9796610	1391.15	3
117	Phillips Run Rd @ Milestone Rd	41.2625320	-69.9923940	2027.77	28.8
118	Milestone Rd @ Proprietors Way	41.2694400	-70.0615420	2027.77	11
119	69 Sparks Ave	41.2747740	-70.0975400	598.86	6
120	Ruddick Commons	41.2655020	-69.9764250	1245.87	4
121	Proprietors Way @ Hinsdale Rd	41.2645700	-70.0646400	2204.8	4
122	8 Pine Tree Rd	41.2630710	-70.0650340	2043.87	20.4

Appendix S1 (continued).

Location No.	Location Site	Latitude	Longitude	Ocean Distance	Mean Density
123	Pine Tree Rd @ Old South Rd	41.2612850	-70.0656750	1834.65	3.6
124	Old South Rd Bike Path	41.2616050	-70.0683280	1866.84	4.666666667
125	Consignment Shop	41.2617620	-70.0767130	1866.84	1
126	Eel Pt Rd @ Dionis Beach Rd	41.2893360	-70.1499290	464.01	4
127	Eel Pt Rd @ Primrose Ln	41.2874240	-70.1449310	686.29	3
128	7 Eel Pt Rd	41.2853730	-70.1416590	905.47	8.6
129	Surrey Ave @ Sandsbury Rd	41.2463320	-69.9923480	379.94	5.5
130	Surrey Ave @ Old Tom Nevers Rd	41.2451830	-69.9926830	254.72	13.4
131	Wanoma Way, Tom Nevers Beach	41.2430110	-69.9943970	63.45	9.6
132	Nichols Rd	41.2453970	-69.9948750	332.62	2.8
133	3 Bosworth Rd	41.2446660	-69.9973740	310.62	5.2
134	JFK Bomb Shelter	41.2408600	-70.0049280	71.79	1.5
135	7 New South Rd	41.2412810	-70.0165810	200.45	3
136	Forked Pond Valley	41.2416030	-70.0228180	125.33	1
137	Russells Way	41.2497150	-70.0358860	777.14	14
138	38 Wigwam Rd	41.2509220	-70.0336370	934.54	5
139	Wigwam Rd @ Russells Way	41.2528190	-70.0362690	1124.2	17.4
140	Milestone Bike Path, New South Rd	41.2676040	-70.0488720	2671.51	1
141	Surfside Beach Parking Lot, East	41.2442170	-70.0938150	118.78	79.2
142	Surfside Beach Parking Lot, West	41.2445410	-70.0947350	171.25	52.2
143	Farrell Beach, NILB	41.2438920	-70.0988950	284.09	17.4
144	27 Western Ave	41.2439470	-70.0973780	213.13	3.5
145	Uncatena St @ Nonantum Ave	41.2451410	-70.0879220	83.49	19.6
146	Pochick Ave @ Surside Rd	41.2494270	-70.0943630	678.53	7.5
147	Masaquett Ave @ Surside Rd	41.2509260	-70.0944300	837.33	4

Appendix S1 (continued).

Location No.	Location Site	Latitude	Longitude	Ocean Distance	Mean Density
148	124 Surfside Rd	41.2541860	-70.0938020	1190.77	6.333333333
149	South Shore Rd @ Surfside Rd	41.2585560	-70.0979400	1754.18	2.8
150	5 Wherowhero Ln	41.2576700	-70.1013550	1786.37	2.666666667
151	Zachary Way @ South Shore Rd	41.2505360	-70.1020150	1050.68	4
152	85-87 South Shore Rd	41.2455100	-70.1048390	480.81	5.4
153	34 W Miacomet Rd, Miacomet Beach	41.2438840	-70.1191130	62.65	6.4
154	28-34 W Miacomet Rd.1	41.2455130	-70.1179810	237.28	6.8
155	28-34 W Miacomet Rd.2	41.2490000	-70.1162560	649.56	1
156	28-34 W Miacomet Rd.3	41.2501470	-70.1156650	786.61	4
157	6-28 W Miacomet Rd	41.2534530	-70.1169910	1108.58	2.6
158	76 Millbrook Rd	41.2692160	-70.1302670	2365.74	6
159	Millbrook Rd, NILB Property	41.2767730	-70.1280030	1882.93	1
160	55 Millbrook Rd	41.2763360	-70.1264950	1947.31	8.6
161	19 Burnt Swamp Ln	41.2756460	-70.1226280	1995.59	1.333333333
162	Massasoit Bridge Rd	41.2747020	-70.1679300	1754.18	1
163	2 Massasoit Bridge Rd	41.2713660	-70.1787050	965.3	3.2
164	330 Madaket Rd	41.2712380	-70.2011040	123.93	13.6
165	Ames Ave	41.2733650	-70.2047280	106.34	2
166	Massachusetts Ave	41.2737200	-70.2093980	0	2
167	Rhode Island Ave	41.2734440	-70.2063850	118.94	10
168	Tennessee Ave	41.2761640	-70.1932950	103.55	11.8
169	Tennessee Ave @ N Cambridge St	41.2790190	-70.1875650	296.78	37.8
170	Vestal St @ Winn St	41.2887050	-70.1101210	714.18	30.4
171	74 Monomoy Rd	41.2813850	-70.0786900	96.64	4.6
172	Monomoy Harbor Path	41.2789410	-70.0832780	167.76	3

Appendix S1 (continued).

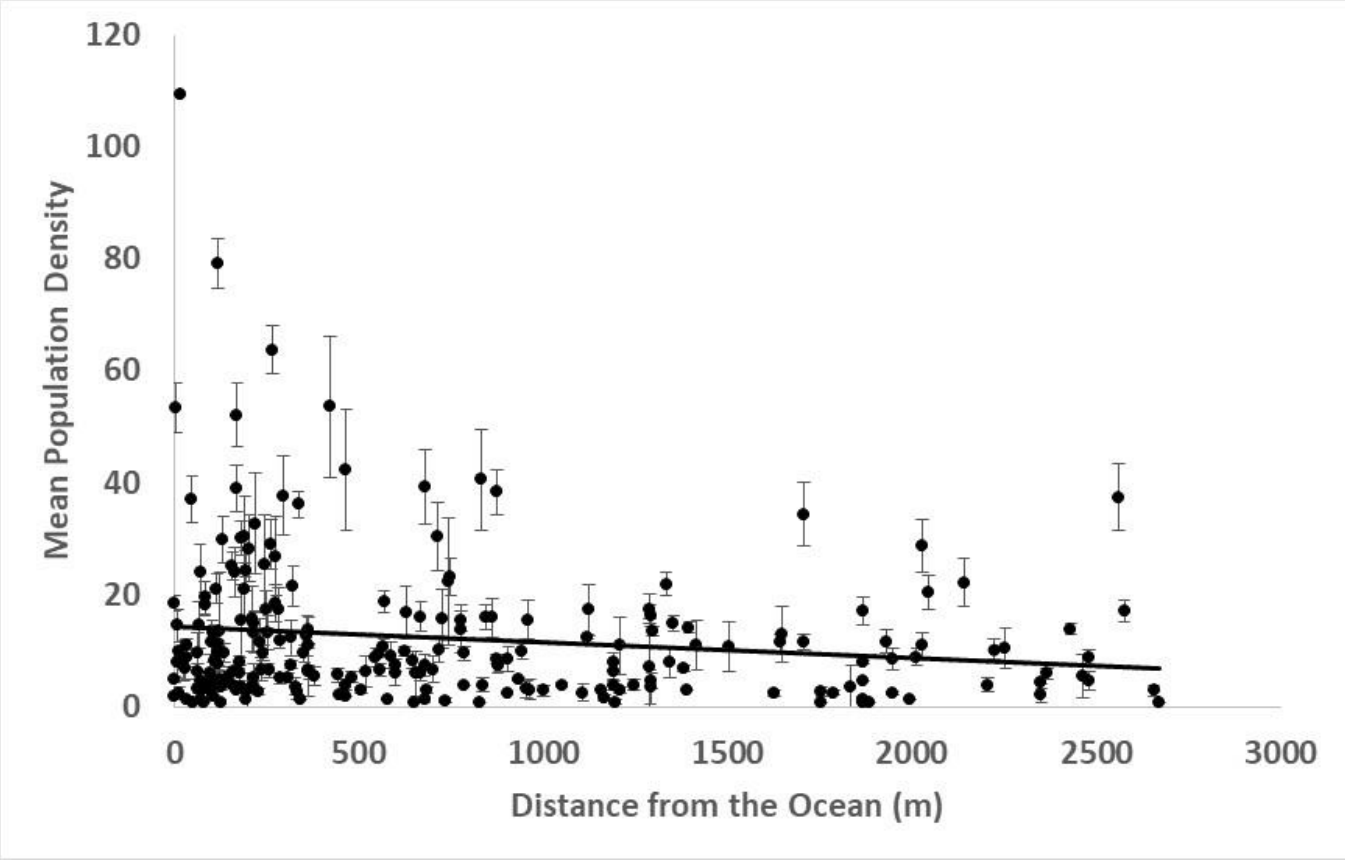
Location No.	Location Site	Latitude	Longitude	Ocean Distance	Mean Density
173	47 Monomoy Rd	41.2786650	-70.0821990	139.64	4.6
174	Eat Fire Spring Rd	41.3073590	-70.0068420	146.55	5.5
175	16 Eat Fire Spring Rd	41.3072060	-70.0016010	586.4	9.2
176	Wauwinet	41.3279320	-69.9966000	214.34	3.5
177	Squam Rd @ Wauwinet Rd	41.3256560	-69.9962050	316.9	12.4
178	Crow's Net Way @ Squam Rd	41.3256240	-69.9937860	132.31	30
179	Squam Rd.1	41.3234390	-69.9928110	170.16	39
180	Squam Rd.2	41.3183950	-69.9918720	336.63	36.2
181	Squam Rd.3	41.3068220	-69.9827280	190.87	21
182	Quidnet.1	41.3048150	-69.9798540	95.94	3
183	Sesachacha	41.3038950	-69.9793950	118.23	9.6
184	Windswept Cranberry Bog	41.2969980	-70.0058120	749.81	23.33333333
185	Almanack Pond Rd.1	41.2918320	-70.0078290	1156.53	3
186	Almanack Pond Rd.2	41.2970390	-70.0113130	559.64	6.666666667
187	Miacomet Golf Course, South	41.2555500	-70.1207230	1207.16	11
188	A Full House	41.2443260	-70.1205370	83.29	18.4
189	W Miacomet Rd.1	41.2462810	-70.1220660	228.96	2.8
190	W Miacomet Rd.2	41.2477250	-70.1229940	322.58	21.6
191	Proprietors Way @ Mioxes Pond Rd	41.2516190	-70.1279400	577.12	1.5
192	Proprietors Way	41.2494780	-70.1291740	316.32	7.6
193	Heller Way	41.2545940	-70.1300570	828.84	1
194	Bartlett Farm Rd	41.2537520	-70.1310500	717.13	10.4
195	Cisco Brewery.1	41.2629230	-70.1301240	1705.9	11.6
196	240 Polpis Rd	41.2905200	-70.0241600	444.06	5.8
197	9 Pocomo Rd	41.3147400	-70.0088330	599.64	7.5

Appendix S1 (continued).

Location No.	Location Site	Latitude	Longitude	Ocean Distance	Mean Density
198	Pocomo Rd.1	41.3152130	-70.0143830	328.23	3.666666667
199	43 Pocomo Rd	41.3147360	-70.0208300	181.32	3.5
200	79 Pocomo Rd	41.3146100	-70.0301340	115.77	4.5
201	Pocomo Head	41.3160470	-70.0323460	0	18.5
202	Polpis Harbor Rd	41.3022670	-70.0107010	35.54	1.5
203	Red Barn Rd.1	41.2644170	-70.1810130	158.45	25.2
204	Sheep Pond Rd	41.2646750	-70.1860610	65.3	3.333333333
205	Red Barn Rd.2	41.2626830	-70.1767690	241.86	9.75
206	Moors End.1	41.2793100	-70.0689670	874.03	8.666666667
207	Moors End.2	41.2778900	-70.0708730	831.93	40.6
208	Polpis Rd @ Sesachacha Pond	41.2901360	-69.9836180	1190.13	4
209	320 Polpis Rd	41.2974770	-69.9978600	1866.84	17.2
210	Rock Face	41.2931870	-70.0209490	164.99	24.2
211	Airport, gravel lot	41.2562610	-70.0662590	1333.75	22
212	Airport 1	41.2629430	-70.0560760	2140.43	22.2
213	Airport 2	41.2538350	-70.0530930	1121.01	12.6
214	Airport 3, Sand Bank	41.2645170	-70.0477730	2349.64	4.6
215	Airport 4, field	41.2633590	-70.0484990	2220.9	10.4
216	Airport 5	41.2610080	-70.0505150	1931.31	11.8
217	Airport 6, Bunker	41.2564100	-70.0547630	1395.06	14.25
218	Airport 7	41.2468510	-70.0698990	288.44	5.2
219	Airport 8, w/TWTA	41.2458980	-70.0705180	179.79	15.4
220	Airport 9, west fence	41.2503080	-70.0709310	669.07	16.2
221	Tom Nevers @ Norwood	41.2613330	-70.0135520	2478.39	8.8
222	Tom Nevers @ Kendrick	41.2596710	-70.0115950	2253.08	10.6

Appendix S1 (continued).

Location No.	Location Site	Latitude	Longitude	Ocean Distance	Mean Density
223	Tom Nevers @ Exeter	41.2566330	-70.0078590	1866.84	8
224	Tom Nevers @ Berkeley	41.2550520	-70.0059070	1641.53	11.6
225	Tom Nevers w/TWTA	41.2530440	-70.0034340	1353.05	15
226	Tom Nevers @ Marcus	41.2448370	-69.9976300	352.52	9.6
227	Tom Nevers field cliff	41.2403650	-70.0053130	48.32	37.2
228	16 Berkeley	41.2533260	-70.0069740	1503.37	10.8
229	22 Exeter	41.2542220	-70.0092270	1647.53	13
230	Wauwinet 1	41.3304300	-69.9972200	120.07	4.5
231	Wauwinet 2	41.3363340	-70.0010140	33.26	9



Appendix S2. Linear regression of the mean population density (± SE) by the distance from the ocean (m). Average density per population were calculated from up to five counts per population. Equation of line of best fit: $y = -0.002765x + 14.32$; Adjusted $R^2 = 0.01728$; $p = 0.0257$.

Appendix S3. Observed treatment sites with their GPS coordinates, the date of observation, population name, and local hourly weather conditions for 8:00 to 11:00 am.

Date	Treatment	Population	Latitude	Longitude	Temp. (°C)	% Humidity	Dew Point (°C)	Wind (kph)	Conditions	Time (Hr)
7/3/2015	Non-native(allopatric)	Nantucket Tackle Center	41.27194	-70.093871	17.2	75	17	5	Clear	8:00
7/3/2015	Non-native(allopatric)	Nantucket Tackle Center	41.27194	-70.093871	21.7	76	17.2	11	Clear	9:00
7/3/2015	Non-native(allopatric)	Nantucket Tackle Center	41.27194	-70.093871	22.8	52	12.2	20	Clear	10:00
7/3/2015	Non-native(allopatric)	Nantucket Tackle Center	41.27194	-70.093871	22.8	52	12.2	20	Clear	11:00
7/5/2015	Native(allopatric)	The Tupancy Links	41.290506	-70.128865	22.2	71	16.7	14	Clear	8:00
7/5/2015	Native(allopatric)	The Tupancy Links	41.290506	-70.128865	20.6	70	17	15	Clear	9:00
7/5/2015	Native(allopatric)	The Tupancy Links	41.290506	-70.128865	21.1	65	15	12	Clear	10:00
7/5/2015	Native(allopatric)	The Tupancy Links	41.290506	-70.128865	21.1	64	15	11	Clear	11:00
7/7/2015	Non-native(allopatric)	Washing Pond Rd	41.288479	-70.131772	21.1	79	17.2	11	Clear	8:00
7/7/2015	Non-native(allopatric)	Washing Pond Rd	41.288479	-70.131772	21.1	78	17	10	Clear	9:00
7/7/2015	Non-native(allopatric)	Washing Pond Rd	41.288479	-70.131772	22.8	84	20	13	Scattered Clouds	10:00
7/7/2015	Non-native(allopatric)	Washing Pond Rd	41.288479	-70.131772	22.8	84	20	13	Scattered Clouds	11:00
7/8/2015	Sympatric	Starbuck Rd	41.270231	-70.194051	18.9	80	19	10	Clear	8:00
7/8/2015	Sympatric	Starbuck Rd	41.270231	-70.194051	20	79	18	5	Clear	9:00
7/8/2015	Sympatric	Starbuck Rd	41.270231	-70.194051	20	79	18	5	Clear	10:00
7/8/2015	Sympatric	Starbuck Rd	41.270231	-70.194051	21.1	80	19	15	Clear	11:00
7/9/2015	Native(allopatric)	Ruddick Commons	41.265995	-69.976201	20.6	68	14.4	14	Overcast	8:00
7/9/2015	Native(allopatric)	Ruddick Commons	41.265995	-69.976201	18	72	14	14	Clear	9:00
7/9/2015	Native(allopatric)	Ruddick Commons	41.265995	-69.976201	19.4	57	12	10	Clear	10:00

Appendix S3 (continued).

Date	Treatment	Population	Latitude	Longitude	Temp. (°C)	% Humidity	Dew Point (°C)	Wind (kph)	Conditions	Time (Hr)
7/9/2015	Native(allopatric)	Ruddick Commons	41.265995	-69.976201	20.6	68	13	7	Clear	11:00
7/11/2015	Native(allopatric)	Linda Loring Foundation	41.29001	-70.171507	23.9	66	15.6	16	Clear	8:00
7/11/2015	Native(allopatric)	Linda Loring Foundation	41.29001	-70.171507	22.8	66	16.1	15	Clear	9:00
7/11/2015	Native(allopatric)	Linda Loring Foundation	41.29001	-70.171507	23.3	66	16.1	10	Clear	10:00
7/11/2015	Native(allopatric)	Linda Loring Foundation	41.29001	-70.171507	24.4	58	15.6	8	Clear	11:00
7/12/2015	Non-native(allopatric)	Pine Tree Rd @ Old South Rd	41.261379	-70.0659	24.4	67	17.8	13	Clear	8:00
7/12/2015	Non-native(allopatric)	Pine Tree Rd @ Old South Rd	41.261379	-70.0659	24.4	67	17.8	13	Clear	9:00
7/12/2015	Non-native(allopatric)	Pine Tree Rd @ Old South Rd	41.261379	-70.0659	24.4	67	17.8	13	Clear	10:00
7/12/2015	Non-native(allopatric)	Pine Tree Rd @ Old South Rd	41.261379	-70.0659	25.6	62	17.8	14	Clear	11:00
7/13/2015	Native(allopatric)	Middle Moors	41.265514	-70.012435	22.8	84	20	6	Clear	8:00
7/13/2015	Native(allopatric)	Middle Moors	41.265514	-70.012435	22.8	84	20	6	Clear	9:00
7/13/2015	Native(allopatric)	Middle Moors	41.265514	-70.012435	22.8	84	20	6	Clear	10:00
7/13/2015	Native(allopatric)	Middle Moors	41.265514	-70.012435	22.8	82	19.4	9	Clear	11:00
7/15/2015	Native(allopatric)	Ram Pasture	41.260507	-70.155101	26.7	94	22.8	9	Mostly Cloudy	8:00
7/15/2015	Native(allopatric)	Ram Pasture	41.260507	-70.155101	23.3	91	20	10	Foggy	9:00
7/15/2015	Native(allopatric)	Ram Pasture	41.260507	-70.155101	23.3	94	22.2	15	Clear	10:00
7/15/2015	Native(allopatric)	Ram Pasture	41.260507	-70.155101	26.7	82	22.8	12	Clear	11:00
7/17/2015	Sympatric	The Tupancy Links	41.290411	-70.12783	20	76	16.1	7	Clear	8:00
7/17/2015	Sympatric	The Tupancy Links	41.290411	-70.12783	20	75	16.1	7	Clear	9:00

Appendix S3 (continued).

Date	Treatment	Population	Latitude	Longitude	Temp. (°C)	% Humidity	Dew Point (°C)	Wind (kph)	Conditions	Time (Hr)
7/17/2015	Sympatric	The Tupancy Links	41.290411	-70.12783	21.1	68	17.2	9	Clear	10:00
7/17/2015	Sympatric	The Tupancy Links	41.290411	-70.12783	22.2	76	17.2	9	Clear	11:00
7/19/2015	Sympatric	Miacomet Rd	41.24479	-70.111335	21.7	87	19.4	8	Overcast	8:00
7/19/2015	Sympatric	Miacomet Rd	41.24479	-70.111335	21.7	87	19.4	8	Overcast	9:00
7/19/2015	Sympatric	Miacomet Rd	41.24479	-70.111335	22.2	87	20	11	Overcast	10:00
7/19/2015	Sympatric	Miacomet Rd	41.24479	-70.111335	22.2	93	21	10	Mist	11:00
7/20/2015	Sympatric	West Miacomet Rd	41.251299	-70.115314	23.3	90	21.7	10	Partly Cloudy	8:00
7/20/2015	Sympatric	West Miacomet Rd	41.251299	-70.115314	25	85	22	11	Partly Cloudy	9:00
7/20/2015	Sympatric	West Miacomet Rd	41.251299	-70.115314	25.6	79	21.7	11	Clear	10:00
7/20/2015	Sympatric	West Miacomet Rd	41.251299	-70.115314	23.3	90	22.2	12	Clear	11:00
7/21/2015	Sympatric	Red Barn Rd @ Head of Plains Rd	41.270194	-70.184774	25	85	22.2	0	Scattered Clouds	8:00
7/21/2015	Sympatric	Red Barn Rd @ Head of Plains Rd	41.270194	-70.184774	25	85	22.2	0	Scattered Clouds	9:00
7/21/2015	Sympatric	Red Barn Rd @ Head of Plains Rd	41.270194	-70.184774	25	85	22.2	0	Scattered Clouds	10:00
7/21/2015	Sympatric	Red Barn Rd @ Head of Plains Rd	41.270194	-70.184774	24.4	85	21.7	8	Clear	11:00
7/22/2015	Non-native(allopatric)	M6 Condo Backyard	41.2732	-70.101602	24.4	64	17.2	16	Clear	8:00
7/22/2015	Non-native(allopatric)	M6 Condo Backyard	41.2732	-70.101602	24.4	64	17.2	16	Clear	9:00
7/22/2015	Non-native(allopatric)	M6 Condo Backyard	41.2732	-70.101602	24.4	64	17.2	16	Clear	10:00
7/22/2015	Non-native(allopatric)	M6 Condo Backyard	41.2732	-70.101602	24.4	64	17.2	16	Clear	11:00

Appendix S3 (continued).

Date	Treatment	Population	Latitude	Longitude	Temp. (°C)	% Humidity	Dew Point (°C)	Wind (kph)	Conditions	Time (Hr)
7/14/2016	Non-native(allopatric)	The Sea Grille	41.271908	-70.0940741	20	93	18	12	Clear	8:00
7/14/2016	Non-native(allopatric)	The Sea Grille	41.271908	-70.0940741	21.7	87	17.2	10	Clear	9:00
7/14/2016	Non-native(allopatric)	The Sea Grille	41.271908	-70.0940741	21.7	87	17	12	Clear	10:00
7/14/2016	Non-native(allopatric)	The Sea Grille	41.271908	-70.0940741	21.7	87	17	12	Clear	11:00
7/15/2016	Native(allopatric)	Linda Loring Osprey Nest	41.290714	-70.170426	21.1	90	16.8	10	Clear	8:00
7/15/2016	Native(allopatric)	Linda Loring Osprey Nest	41.290714	-70.170426	21.7	90	15	16	Clear	9:00
7/15/2016	Native(allopatric)	Linda Loring Osprey Nest	41.290714	-70.170426	21.7	90	15.5	16	Clear	10:00
7/15/2016	Native(allopatric)	Linda Loring Osprey Nest	41.290714	-70.170426	21.7	90	15.5	16	Clear	11:00
7/16/2016	Pre-removal	The Tupancy Links.2	41.290508	-70.128057	25.6	76	20.1	0	Clear	8:00
7/16/2016	Pre-removal	The Tupancy Links.2	41.290508	-70.128057	23.3	85	19	10	Clear	9:00
7/16/2016	Pre-removal	The Tupancy Links.2	41.290508	-70.128057	23.3	85	19	11	Clear	10:00
7/16/2016	Pre-removal	The Tupancy Links.2	41.290508	-70.128057	23.3	80	18.5	11	Clear	11:00
7/17/2016	Post-removal	The Tupancy Links.2	41.290508	-70.128057	20.5	100	21	11	Overcast	8:00
7/17/2016	Post-removal	The Tupancy Links.2	41.290508	-70.128057	21.1	95	20.5	11	Overcast	9:00
7/17/2016	Post-removal	The Tupancy Links.2	41.290508	-70.128057	21.6	95	20.5	11	Overcast	10:00
7/17/2016	Post-removal	The Tupancy Links.2	41.290508	-70.128057	21.6	90	20	10	Overcast	11:00
7/18/2016	Pre-removal	Red Barn Rd @ Head of Plains Rd.2	41.27019	-70.184541	20.5	100	20	19	Partly Cloudy	8:00
7/18/2016	Pre-removal	Red Barn Rd @ Head of Plains Rd.2	41.27019	-70.184541	21.1	100	21	19	Partly Cloudy	9:00

Appendix S3 (continued).

Date	Treatment	Population	Latitude	Longitude	Temp. (°C)	% Humidity	Dew Point (°C)	Wind (kph)	Conditions	Time (Hr)
7/18/2016	Pre-removal	Red Barn Rd @ Head of Plains Rd.2	41.27019	-70.184541	21.1	97	21	28	Partly Cloudy	10:00
7/18/2016	Pre-removal	Red Barn Rd @ Head of Plains Rd.2	41.27019	-70.184541	21.1	95	20.5	20	Partly Cloudy	11:00
7/19/2016	Post-removal	Red Barn Rd @ Head of Plains Rd.2	41.27019	-70.184541	22.8	85	20	14	Clear	8:00
7/19/2016	Post-removal	Red Barn Rd @ Head of Plains Rd.2	41.27019	-70.184541	25	74	20	13	Clear	9:00
7/19/2016	Post-removal	Red Barn Rd @ Head of Plains Rd.2	41.27019	-70.184541	25	73	19.5	17	Clear	10:00
7/19/2016	Post-removal	Red Barn Rd @ Head of Plains Rd.2	41.27019	-70.184541	25.5	74	20	17	Clear	11:00
7/21/2016	Pre-removal	Miacomet Trail.2	41.246066	-70.111801	22.2	84	18	6	Clear	8:00
7/21/2016	Pre-removal	Miacomet Trail.2	41.246066	-70.111801	23.9	69	16	6	Clear	9:00
7/21/2016	Pre-removal	Miacomet Trail.2	41.246066	-70.111801	23.9	69	16	6	Clear	10:00
7/21/2016	Pre-removal	Miacomet Trail.2	41.246066	-70.111801	25	49	14	11	Clear	11:00
7/23/2016	Post-removal	Miacomet Trail.2	41.246066	-70.111801	22.2	94	21	16	Clear	8:00
7/23/2016	Post-removal	Miacomet Trail.2	41.246066	-70.111801	23.3	85	21	21	Clear	9:00
7/23/2016	Post-removal	Miacomet Trail.2	41.246066	-70.111801	24.4	82	21.5	27	Clear	10:00
7/23/2016	Post-removal	Miacomet Trail.2	41.246066	-70.111801	24.4	82	21.5	29	Clear	11:00
7/24/2016	Pre-removal	Tom Nevers	41.253104	-70.003545	20	80	17	15	Clear	8:00
7/24/2016	Pre-removal	Tom Nevers	41.253104	-70.003545	20.1	75	16	14	Clear	9:00
7/24/2016	Pre-removal	Tom Nevers	41.253104	-70.003545	22	90	20	16	Clear	10:00
7/24/2016	Pre-removal	Tom Nevers	41.253104	-70.003545	22	95	21	10	Clear	11:00
7/25/2016	Post-removal	Tom Nevers	41.253104	-70.003545	21.7	94	20	11	Clear	8:00
7/25/2016	Post-removal	Tom Nevers	41.253104	-70.003545	21.7	93	20	21	Clear	9:00
7/25/2016	Post-removal	Tom Nevers	41.253104	-70.003545	22.8	88	20.1	21	Clear	10:00

Appendix S3 (continued).

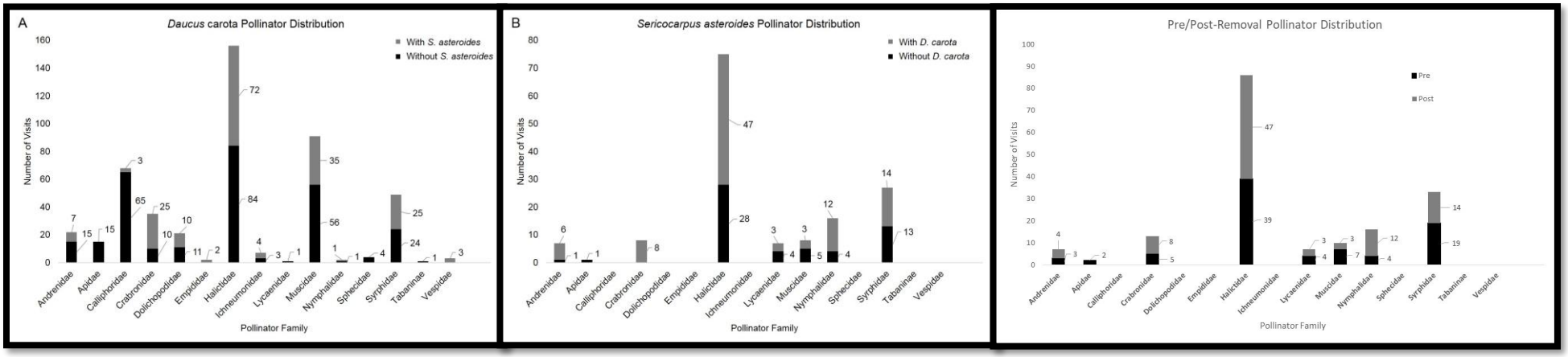
Date	Treatment	Population	Latitude	Longitude	Temp. (°C)	% Humidity	Dew Point (°C)	Wind (kph)	Conditions	Time (Hr)
7/25/2016	Post-removal	Tom Nevers	41.253104	-70.003545	22.8	87	20	21	Clear	11:00

Appendix S4. Populations of *Sericocarpus asteroides* located across Nantucket Island, Massachusetts with their GPS coordinates.

Location Number	Location Site	Latitude	Longitude
1	Linda Loring Foundation.1	41.289992	-70.171492
2	Linda Loring Foundation.2	41.289914	-70.172725
3	ACK Sparkle Clean	41.286922	-70.174961
4	Linda Loring Foundation.3	41.289061	-70.174649
5	Linda Loring Foundation.4	41.291503	-70.170817
6	LLF, Beach Trail, South end	41.287781	-70.177794
7	Sanford Farm	41.278289	-70.137069
8	North Head Hummock Pond, Sanford Farm	41.277894	-70.135561
9	Sanford Farm, W of deer exclosure	41.276303	-70.137775
10	NILB, Madaket Rd to Eel Pt property	41.283573	-70.156976
11	NILB, Worth Rd property	41.285976	-70.155407
12	Dionis	41.287827	-70.159307
13	Miacomet Rd	41.244914	-70.111420
14	Head of Prairies @ Red Barn Rd	41.270177	-70.184992
15	The Tupancy Links	41.290455	-70.128081
16	The Tupancy Links, West Entrance	41.288227	-70.131163
17	Milestone Trail	41.265514	-70.012435
18	The Tupancy Links.2	41.291761	-70.128326
19	Tuckernuck.1	41.300877	-70.248592
20	Middle Moors	41.265514	-70.012435
21	West Miacomet Rd	41.251299	-70.115314
22	Ram Pasture	41.260507	-70.155101
23	Ruddick Commons	41.265995	-69.976201
24	Starbuck Rd	41.270231	-70.194051
25	Linda Loring.1	41.290700	-70.170427
26	Linda Loring.2	41.290257	-70.171681
27	The Tupancy Links.1	41.290537	-70.128030
28	The Tupancy Links.2	41.290748	-70.128431
29	The Tupancy Links.3	41.292683	-70.129337
30	Red Barn @ Head of Plains	41.270166	-70.184936
31	Red Barn Rd	41.271229	-70.185237
32	225 Madaket Rd	41.277954	-70.182333
33	Madaket @ Worth	41.281781	-70.158223
34	Northern Loop @ Ram Pasture	41.279454	-70.138048
35	Ram Pasture Rd.1	41.275887	-70.140255
36	Ram Pasture Rd.2	41.273783	-70.140407
37	Ram Pasture @ Ocean Walk	41.260781	-70.154432

Appendix S4 (continued).

Location Number	Location Site	Latitude	Longitude
38	Ram Pasture Rd.3	41.260414	-70.155120
39	280 Miacomet Rd	41.244408	-70.110810
40	146 Miacomet Rd	41.248534	-70.114390
41	Middle Moors.1	41.265705	-70.012212
42	Middle Moors.2	41.264824	-70.013954
43	Milestone @ South Pasture	41.264464	-70.026957
44	Sconset Bike Path @ Phillips Run	41.262284	-69.990021
45	Milestone @ Skinners	41.263067	-69.977518
46	Wauwinet @ Fargo	41.317307	-70.006141
47	101 Polpis Rd	41.285003	-70.057695



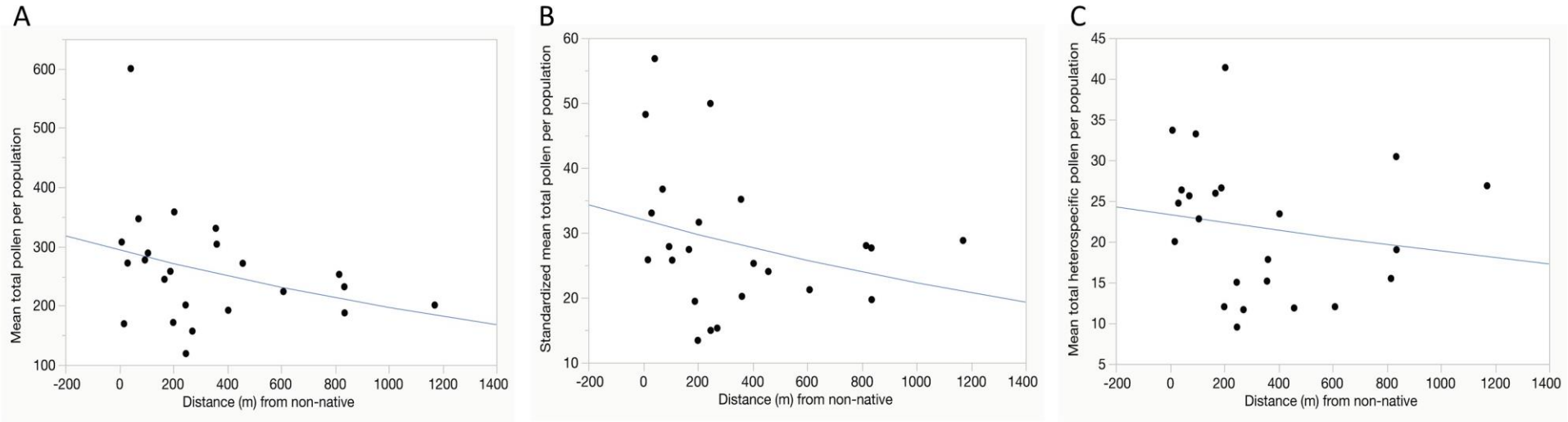
Appendix S5. The total number of visits by pollinator family observed on *Daucus carota* (A) and *Sericocarpus asteroides* (B) when in allopatry or sympatry with the other, and the pollinator distribution of the pre- and post-removal treatments for *S. asteroides* (C).

Appendix S6: Populations collected for the heterospecific pollen assay, their GPS coordinates, and the number, name, GPS coordinates, and distance to the nearest *Daucus carota* population. Fifteen inflorescences were collected from each population.

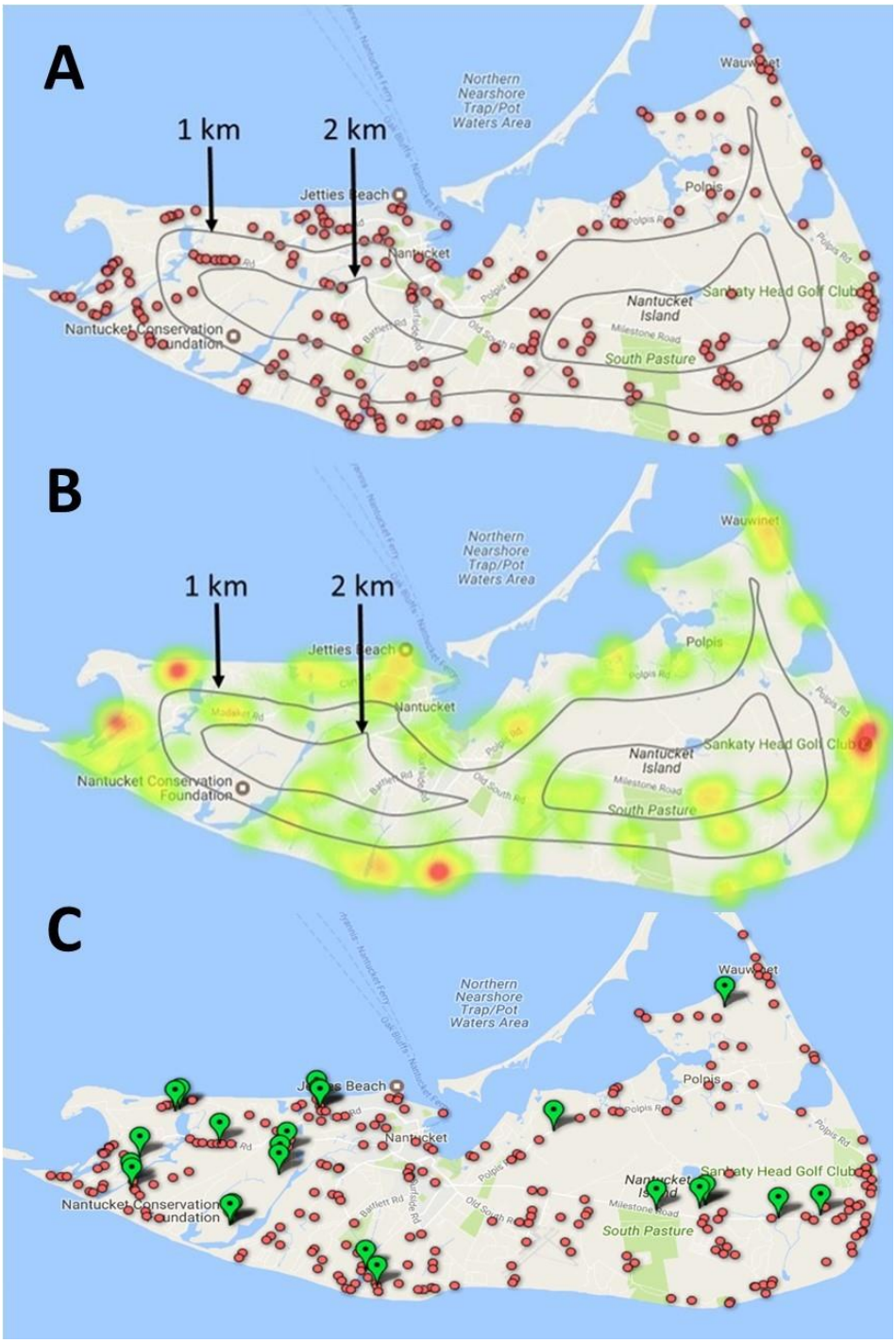
Site Number	Site Location	Latitude	Longitude	Distance	Closest QAL Pop. No.	Closest QAL Pop. Name	Latitude	Longitude
1a	Linda Loring.1	41.2907000	-70.1704270	248.12	10	118 Bellevue, Eel Point Rd	41.2924690	-70.1722030
2a	Linda Loring.2	41.2902570	-70.1716810	247.02	10	118 Bellevue, Eel Point Rd	41.2924690	-70.1722030
3a	The Tupancy Links.1	41.2905370	-70.1280300	9.77	52	The Tupancy Links	41.2904550	-70.1280810
4a	The Tupancy Links.2	41.2907480	-70.1284310	43.46	52	The Tupancy Links	41.2904550	-70.1280810
5a	The Tupancy Links.3	41.2926830	-70.1293370	95.98	53	The Sound Bluff, The Tupancy Links	41.2935410	-70.1295460
6a	Red Barn @ Head of Plains	41.2701660	-70.1849360	31.74	36	Head of Plains @ Red Barn Rd	41.2702190	-70.1845650
7a	Red Barn Rd	41.2712290	-70.1852370	107.32	39	Long Pond Sanctuary, Parking Lot	41.2721420	-70.1848300
8a	225 Madaket Rd	41.2779540	-70.1823330	405.15	35	Arkansas Ave @ S. Cambridge St.	41.2767500	-70.1869140
9a	Madaket @ Worth	41.2817810	-70.1582230	18.38	23	Madaket Rd @ Worth Rd	41.2819370	-70.1581760
10a	Northern Loop @ Ram Pasture	41.2794540	-70.1380480	190.37	18	NCF, Sanford Farm, Tile Silo	41.2811580	-70.1381860
11a	Ram Pasture Rd.1	41.2758870	-70.1402550	609.98	18	NCF, Sanford Farm, Tile Silo	41.2811580	-70.1381860
12a	Ram Pasture Rd.2	41.2737830	-70.1404070	837.17	18	NCF, Sanford Farm, Tile Silo	41.2811580	-70.1381860
13a	Ram Pasture @ Ocean Walk	41.2607810	-70.1544320	836.1	80	Madaket Beach Place	41.2535570	-70.1516570
14a	Ram Pasture Rd.3	41.2604140	-70.1551200	816.89	80	Madaket Beach Place	41.2535570	-70.1516570
15a	280 Miacomet Rd	41.2444080	-70.1108100	72.06	27	Miacomet Rd, Loop.2	41.2448130	-70.1114810
16a	146 Miacomet Rd	41.2485340	-70.1143900	168.49	29	Miacomet Rock Walking Trail	41.2470510	-70.1139980

Appendix S6 (continued).

Site Number	Site Location	Latitude	Longitude	Distance	Closest QAL Pop. No.	Closest QAL Pop. Name	Latitude	Longitude
17a	Middle Moors.1	41.2657049	-70.0122120	359.02	67	Milesone Overlook Trial, P.L. Entr.	41.2624970	-70.0127889
18a	Middle Moors.2	41.2648240	-70.0139540	272.06	67	Milesone Overlook Trial, P.L. Entr.	41.2624970	-70.0127889
19a	Milestone @ South Pasture	41.2644640	-70.0269570	1171.61	221	Tom Nevers @ Norwood	41.2613330	-70.0135520
20a	Sconset Bike Path @ Phillips Run	41.2622840	-69.9900210	201.01	117	Phillips Run Rd @ Milestone Rd	41.2625320	-69.9923940
21a	Milestone @ Skinners	41.2630670	-69.9775180	204.83	116	Sconset Bike Path	41.2621700	-69.9796610
22a	Wauwinet @ Fargo	41.3173070	-70.0061410	362.45	197	9 Pocomo Rd	41.3147400	-70.0088330
23a	101 Polpis Rd	41.2850030	-70.0576950	459.08	75	82 Polpis Rd, Polpis Bike Path	41.2829280	-70.0624170



Appendix S7. Results of generalized linear model performed in JMP (ver. 13.0.0) of A) mean total pollen, B) mean total standardized pollen, and C) mean total heterospecific pollen for 23 *Sericocarpus asteroides* populations at given distances from *Daucus carota*. Data were normalized by dividing total counts per slide by the number of stigmas per slide.



Appendix S8. *Daucus carota* occurrences (A), those occurrence densities (B), and *D. carota* and *Sericocarpus asteroides* occurrences across Nantucket Island, MA. Red points (A, C) indicate *D. carota* occurrences; in B, cool colors represent low densities while warm colors indicate high densities; green pins (C) indicate *S. asteroides* occurrences.