

DISSECTING PROCESSES UNDERLYING REINFORCEMENT AND ITS CONSEQUENCES FOR SPECIATION

by

EMILY RUTH BEWICK

(Under the Direction of Kelly Dyer)

ABSTRACT

A primary focus of evolutionary biology is the origin of biodiversity, particularly how new species arise and how they remain distinct. Hybrid zones, where recently diverged species come into contact, provide a window into the processes driving the early stages of speciation since the balance of selection and gene flow between species is crucial to determining whether species remain distinct or fuse back into one. When hybrids are sterile or unfit, selection can act to increase pre-mating barriers to reinforce species boundaries. Whether this reinforcement occurs will depend on the strength of this selection and gene flow both between and within species. This reinforcement can also have secondary effects, where selection in the contact zone drives divergence between populations in this zone and those outside it. In two species of fruit fly in North America, such a process is occurring. *Drosophila subquinaria* and *D. recens* share a contact zone east of the Rocky Mountains. In *D. subquinaria* reinforcing selection has strengthened pre-mating isolation, and *D. subquinaria* females from this zone reject *D. recens* males as potential mates. These females also reject males of their own species from populations

outside this zone. In contrast, *D. recens* does not show either of these patterns in mating behavior. Here, I dissect what forces are contributing to divergence both between and within species. I found that strong selection to avoid hybridization coupled with reduced gene flow from populations outside the contact zone has facilitated the evolution of premating isolation in *D. subquinaria*. Additionally, the genetic basis of mate choice between and within species may be shared or closely linked, which may have facilitated the evolution of isolation within this species. On the other hand, a lower cost to hybridization, coupled with abundant gene flow from populations outside the contact zone appears to have hindered the evolution of reinforced premating isolation in *D. recens*. Thus, the outcome of secondary contact can differ for each species involved, depending on the conditions of gene flow and selection, and divergence within species as a result of this contact may be facilitated by genetic linkage.

INDEX WORDS: *Drosophila*, speciation, hybridization, reinforcement, introgression, gene flow, mate choice

DISSECTING PROCESSES UNDERLYING REINFORCEMENT AND ITS
CONSEQUENCES FOR SPECIATION

by

EMILY RUTH BEWICK

BS, Vanderbilt University, 2008

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2014

© 2014

Emily R. Bewick

All Rights Reserved

DISSECTING PROCESSES UNDERLYING REINFORCEMENT AND ITS
CONSEQUENCES FOR SPECIATION

by

EMILY RUTH BEWICK

Major Professor: Kelly Dyer

Committee: Michael Arnold
John Burke
David Hall
Daniel Promislow

Electronic Version Approved:

Julie Coffield
Interim Dean of the Graduate School
The University of Georgia
August 2014

DEDICATION

I would like to dedicate this dissertation to my parents, who have supported me unconditionally. To my father: who first inspired my love for nature and science, encouraging my interest to the extent of dissecting catfish in our sink. To my mother: who kept me grounded and encouraged me to stick through the hard times and who never fails to be my biggest cheerleader.

ACKNOWLEDGEMENTS

First, I would like to acknowledge my husband, Adam Bewick, without whose support I could not have completed this task. I will be forever grateful for the love and support he gave to help me focus on my work and finish my degree. Second I would like to thank my advisor, Dr. Kelly Dyer, who took a chance and took me on as her first student, and all the members of my committee, whose helpful guidance has improved me. The strides I have made as a student are due to their thoughtful guidance and commitment to my education.

I particularly want to acknowledge the members of the Dyer lab. First, I want to thank Cheryl Pinzone, who has been a calming and empowering influence on my life. I would like to acknowledge all the members of the Dyer lab, both past and present, who have helped me complete this work, most especially Brooke White, Michael Bray, and Devon Gaydos. Additionally, I want to thank Katie Bockrath and Catherine Debban for field assistance. I would also like to thank all the members of the EvolGen lab discussion group for a series of fruitful discussions that helped me sharpen my critical thinking.

I would like to thank three graduate colleagues in particular: Louisa Carter Staton, Kerin Bentley, and Jared Lee. Their professional and emotional support was invaluable during my graduate career. Additionally, I would like to thank Eleanor Kuntz for encouraging me through the final, difficult stages of finishing this dissertation.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	v
CHAPTER	
1 INTRODUCTION AND LITERATURE REVIEW	1
Literature Cited	7
2 REINFORCEMENT SHAPES CLINES IN FEMALE MATING DISCRIMINATION IN <i>DROSOPHILA SUBQUINARIA</i>	11
Abstract	12
Introduction	13
Materials and Methods	17
Results	23
Discussion	28
Acknowledgements	37
Literature Cited	37
3 INTROGRESSION AND POPULATION GENETIC ANALYSES OF TWO SPECIES IN SECONDARY CONTACT: IMPLICATIONS FOR REINFORCEMENT AND ITS CONSEQUENCES	52
Introduction	52
Materials and Methods	57
Results	61

Discussion	66
Literature Cited	72
4 THE SAME CHROMOSOMES CONTRIBUTE TO FEMALE MATE DISCRIMINATION BOTH BETWEEN AND WITHIN SPECIES IN <i>DROSOPHILA</i> <i>SUBQUINARIA</i>	118
Introduction	118
Materials and Methods	121
Results	125
Discussion	129
Literature Cited	135
5 CONCLUSIONS AND FUTURE DIRECTIONS	145
Literature Cited	151
APPENDICES	
A Supplemental Material for Chapter 2, in submission at <i>Evolution</i>	154

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

The origin of biodiversity has been a primary focus of the study of evolution, and in order to understand diversification, we must understand how and why species remain distinct. Animal species are generally categorized as potentially or actually reproducing groups (Mayr 1940), thus how that ability to reproduce changes over time and space becomes central to defining the speciation process. A general model of speciation is one in which species are separated by a geographic barrier and begin to diverge at the genetic level. Over time, the genetic divergence between these species can result in incompatibilities between them, causing barriers to reproduction. These barriers can be generally categorized as those that act before or after mating (pre- and post-mating barriers, respectively). If two newly diverged populations come back into contact after divergence (secondary contact), the outcome of this contact could determine if these species fuse or continue diverging (Dobzhansky 1951). If isolating barriers are low, the species may fuse back together or this region may become a hybrid zone, where the formation of hybrids can result in gene exchange between species. Whether these species remain distinct will depend on the conflict between selection against hybrids and gene flow between the two genomes (Barton and Hewitt 1985; Coyne and Orr 2004).

When postmating barriers have accumulated in allopatry but pre-mating barriers have not, selection can act to increase pre-mating barriers and thus lead to higher

divergence between species in this hybrid zone (Dobzhansky (1940)). However, as pointed out by Felsenstein (1981), the very nature of gene flow in a hybrid zone could restrict the evolution of pre-mating isolation by recombining genes associated with mating traits and hybrid fitness. Theoretical and experimental work has now shown that this reinforcement of species boundaries can proceed under some circumstances of gene flow between species (reviewed in Coyne and Orr 2004), and we have seen evidence that introgression can occur in species undergoing reinforcement (Borge et al. 2005; Macholan et al. 2007; Kulathinal et al. 2009). Another obstacle to reinforcement in hybrid zones is gene flow from allopatric populations. By its nature, reinforcing selection only occurs in populations that have an opportunity for hybridization, so alleles beneficial in reducing maladaptive hybridization may not be beneficial outside of the zone of sympatry, and in fact may carry costs (Pfennig 2000; Higgie and Blows 2007). Too much gene flow from allopatric populations can swamp out the effects of selection for pre-mating isolation in sympatry; however reinforcement is possible as long as gene flow is not too high and reinforced alleles not too deleterious in allopatry (Caisse and Antonovics 1978; Sanderson 1989; Liou and Price 1994; Matute 2010). These results suggest that reinforcement may occur more easily when populations in contact zones have reduced gene flow from outside populations.

However, this points to a potential consequence of reinforcement, which is that it can also lead to divergence between conspecific sympatric and allopatric populations (Howard 1993; Nosil et al. 2003; Hoskin et al. 2005; Pfennig and Ryan 2006; Ortiz-Barrientos et al. 2009; Hoskin and Higgie 2010). Because reinforcing selection is expected to act on mating signals so that males and females can distinguish between

appropriate and inappropriate mates, male mating signals and female preferences for them can diverge from the rest of the range, creating a pattern of reproductive character displacement (RCD) (Howard 1993; Servedio and Noor 2003; Servedio 2004). If the signals or preferences diverge enough, females from populations sympatric with the other species may no longer recognize conspecific allopatric individuals as potential mates. Cascade reinforcement or “RCD speciation” describes this process where reinforcing selection on assortative mating in sympatry leads to sexual isolation between populations within a species (Ortiz-Barrientos et al. 2009; Hoskin and Higgie 2010). In this case pre-mating barriers arise before post-mating barriers, and the cessation of gene flow between sympatric and allopatric populations could allow genetic incompatibilities to accumulate (Coyne and Orr 2004). Eventually, this divergence may become complete, leading to two “good species” in a series of events that began with selection against hybridization.

While we understand that reinforcement does occur in nature, we are still in the early stages of understanding the natural conditions that facilitate reinforcement and what role reinforcement itself may have in generating biodiversity. First, an assessment of variation within and among populations of traits under reinforcing selection is necessary for a complete understanding of how strong selection may be, where it is acting, and what the potential consequences of reinforcement are. Second, patterns of gene flow between and within species under reinforcement have been thoroughly examined in only a few cases (e.g., Borge et al. 2005; Macholan et al. 2007; Kulathinal et al. 2009), and better understanding of gene flow both within and between species will allow us to identify what universal patterns and processes promote/hinder the evolution of reinforcement. Lastly, the genetic architecture of reinforced traits is known in only a few systems (Ortiz-

Barrientos et al. 2004; Saether et al. 2007; Hopkins and Rausher 2011), so we do not yet know if universal patterns exist in the genetics of reinforcement, though we have better expectations based on these systems. Additionally, the genetic architecture of reinforced traits may determine whether reinforcement can lead to a second divergence event within species.

In order to answer these questions about how reinforcement proceeds in nature and whether the consequences of reinforcement can lead to divergence within species, an appropriate empirical system is needed. Few cases of reinforcement in nature have been identified, and of those cases, a system of sister species of fruit fly in North America suits the requirements of an appropriate model. *Drosophila subquinaria* and *D. recens* are sister species of *Drosophila* that are distributed across western and eastern North America, respectively (Fig. 1). They have a zone of contact to the east of the Canadian Rocky Mountains, and they utilize mushrooms for all life stages, including courtship and oviposition. Both species can be found on the same mushrooms.

In the zone of contact, they have the opportunity to hybridize. However, hybridization is often costly because almost all *D. recens* are infected with a strain of the bacterial endosymbiont *Wolbachia*. When an infected male mates with an uninfected female, a cytoplasmic incompatibility occurs between sperm and egg, resulting in the death of almost all embryos (Werren and Jaenike 1995). Because *D. subquinaria* is not infected with *Wolbachia*, when an uninfected *D. subquinaria* female mates with an infected *D. recens* male, almost all of her offspring die as embryos (Werren and Jaenike 1995). This constitutes a high cost for *D. subquinaria* females who mate with heterospecific males. In the reciprocal cross (female *D. recens* x male *D. subquinaria*),

cytoplasmic incompatibility does not occur, and offspring are viable but hybrid males are sterile (Shoemaker et al. 1999). Female *D. subquinaria* in sympatry with *D. recens* show highly choosy behavior and reject *D. recens* males (Jaenike et al. 2006), whereas *D. subquinaria* females from populations allopatric with *D. recens* only reject *D. recens* males at moderate levels. *D. recens* females on the other hand, do not show higher discrimination in sympatric populations, despite the cost of sterile hybrid males. *D. recens* females mate with *D. subquinaria* at a moderate rate regardless of contact with *D. subquinaria*, and thus show no reproductive character displacement.

Sympatric *D. subquinaria* females not only reject *D. recens* males, but they also will not mate with their allopatric conspecifics (Jaenike et al. 2006). This may be a byproduct of reinforcing selection in sympatry, consistent with the predictions of cascade reinforcement. In contrast, there is no reproductive character displacement or sexual isolation among populations of *D. recens*. Previous genetic work suggests that reinforcement and cascade reinforcement have evolved in the face of gene flow from allopatric populations, though the assessments of genetic structure have used only mtDNA (Jaenike et al. 2006). Additionally, there appears to be some genetic differentiation between inland populations of *D. subquinaria* (both allopatric and sympatric) and allopatric populations located to the west of the Coast Mountains in the Pacific Northwest. Thus, this system is ripe for asking not only what forces contribute to reinforcement but also how reinforcement may generate further divergence within species.

In order to understand if any universal patterns underlie the speciation process, we must identify the forces acting to generate or maintain divergence in nature. In this

dissertation, I dissect the conditions under which reinforcement has occurred in *D. subquinaria* and how those conditions may have facilitated divergence in *D. subquinaria* as a consequence of reinforcement. In my first chapter, I survey variation in female mating behavior across the range of *D. subquinaria*, both against *D. recens* and among allopatric conspecific populations. I use a clinal analysis to quantify the strength of selection on discrimination behaviors and ask whether there is evidence that discrimination within and between species may share a genetic basis. In my second chapter, I assess introgression between *D. subquinaria* and *D. recens* and genetic structure within species to illuminate how gene flow and selection may be interacting to promote/hinder divergence between species. I also ask whether different regions of the genome show different patterns of introgression, possibly indicating regions under selection. Finally I use a simple assessment of genetic architecture to ask whether female discrimination within and between species is dominant or recessive, which chromosomes contribute most, and whether this architecture differs among sympatric populations. Ultimately, I aim to uncover the processes promoting divergence both between and within species.

Literature Cited

- Barton, N. H. and G. M. Hewitt. 1985. Analysis of hybrid zones. *Annu. Rev. Ecol. S.* 16:113-148.
- Borge, T., M. T. Webster, G. Andersson, and G. P. Saetre. 2005. Contrasting patterns of polymorphism and divergence on the Z chromosome and autosomes in two *Ficedula* flycatcher species. *Genetics* 171:1861-1873.
- Caisse, M. and J. Antonovics. 1978. Evolution in closely adjacent plant populations .IX. Evolution of reproductive isolation in clinal populations. *Heredity* 40:371-384.
- Coyne, J. A. and H. A. Orr. 2004. Speciation. Sinauer Associates, Sunderland (Massachusetts).
- Dobzhansky, T. 1940. Speciation as a stage in evolutionary divergence. *Am. Nat.* 74:312-321.
- Dobzhansky, T. 1951. Genetics and the origin of species. Columbia Univ. Press, New York.
- Felsenstein, J. 1981. Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution* 35:124-138.
- Higgie, M. and M. W. Blows. 2007. Are traits that experience reinforcement also under sexual selection? *Am. Nat.* 170:409-420.
- Hopkins, R. and M. D. Rausher. 2011. Identification of two genes causing reinforcement in the Texas wildflower *Phlox drummondii*. *Nature* 469:411-414.
- Hoskin, C. J. and M. Higgie. 2010. Speciation via species interactions: the divergence of mating traits within species. *Ecol. Lett.* 13:409-420.

- Hoskin, C. J., M. Higgie, K. R. McDonald, and C. Moritz. 2005. Reinforcement drives rapid allopatric speciation. *Nature* 437:1353-1356.
- Howard, D. J. 1993. Reinforcement: Origin, dynamics, and fate of an evolutionary hypothesis. Pp. 46-69 in H. Rg, ed. *Hybrid zones and the evolutionary process*. Oxford University Press, New York.
- Jaenike, J., K. A. Dyer, C. Cornish, and M. S. Minhas. 2006. Asymmetrical reinforcement and *Wolbachia* infection in *Drosophila*. *PLoS Biol.* 4:1852-1862.
- Kulathinal, R. J., L. S. Stevison, and M. A. Noor. 2009. The genomics of speciation in *Drosophila*: diversity, divergence, and introgression estimated using low-coverage genome sequencing. *PLoS Genet.* 5:e1000550.
- Liou, L. W. and T. D. Price. 1994. Speciation by reinforcement of premating isolation. *Evolution* 48:1451-1459.
- Macholan, M., P. Munclinger, M. Sugerikova, P. Dufkova, B. Bimova, E. Bozikova, J. Zima, and J. Pialek. 2007. Genetic analysis of autosomal and X-linked markers across a mouse hybrid zone. *Evolution* 61:746-771.
- Matute, D. R. 2010. Reinforcement can overcome gene flow during speciation in *Drosophila*. *Current Biology* 20:2229-2233.
- Mayr, E. 1940. Speciation phenomena in birds. *Am. Nat.* 74:249-278.
- Nosil, P., B. J. Crespi, and C. P. Sandoval. 2003. Reproductive isolation driven by the combined effects of ecological adaptation and reinforcement. *P. Roy. Soc. B-Biol. Sci.* 270:1911-1918.
- Ortiz-Barrientos, D., B. A. Counterman, and M. A. F. Noor. 2004. The genetics of speciation by reinforcement. *PLoS Biol.* 2:2256-2263.

- Ortiz-Barrientos, D., A. Greal, and P. Nosil. 2009. The genetics and ecology of reinforcement: implications for the evolution of prezygotic isolation in sympatry and beyond. Pp. 156-182 in C. D. Schlichting, and T. A. Mousseau, eds. Year in Evolutionary Biology 2009. Wiley-Blackwell, Malden.
- Pfennig, K. S. 2000. Female spadefoot toads compromise on mate quality to ensure conspecific matings. *Behav. Ecol.* 11:220-227.
- Pfennig, K. S. and M. J. Ryan. 2006. Reproductive character displacement generates reproductive isolation among conspecific populations: an artificial neural network study. *P. Roy. Soc. B-Biol. Sci.* 273:1361-1368.
- Saether, S. A., G. P. Saetre, T. Borge, C. Wiley, N. Svedin, G. Andersson, T. Veen, J. Haavie, M. R. Servedio, S. Bures, M. Kral, M. B. Hjernquist, L. Gustafsson, J. Traff, and A. Qvarnstrom. 2007. Sex chromosome-linked species recognition and evolution of reproductive isolation in flycatchers. *Science* 318:95-97.
- Sanderson, N. 1989. Can gene flow prevent reinforcement? *Evolution* 43:1223-1235.
- Servedio, M. R. 2004. The what and why of research on reinforcement. *PLoS Biol.* 2:e420.
- Servedio, M. R. and M. A. F. Noor. 2003. The role of reinforcement in speciation: theory and data. *Annu. Rev. Ecol. Evol. S.* 34:339-364.
- Shoemaker, D. D., V. Katju, and J. Jaenike. 1999. *Wolbachia* and the evolution of reproductive isolation between *Drosophila recens* and *Drosophila subquinaria*. *Evolution* 53:1157-1164.
- Werren, J. H. and J. Jaenike. 1995. *Wolbachia* and cytoplasmic incompatibility in mycophagous *Drosophila* and their relatives. *Heredity* 75:320-326.

Figure 1.1.

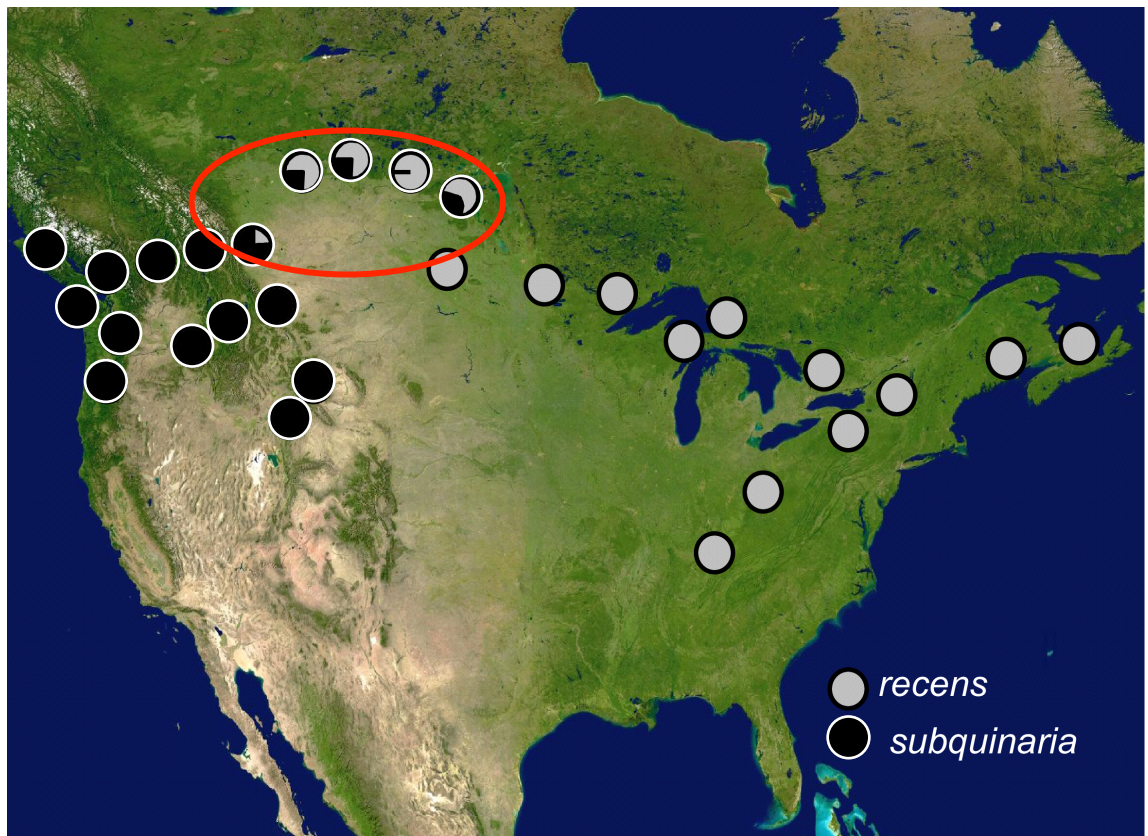


Fig. 1.1. Ranges of *D. subquinaria* and *D. recens* adapted from Jaenike et al. (2006). *D. recens* is shown in gray; *D. subquinaria* is shown in black. The hybrid zone east of the Rocky Mountains is indicated with a red circle.

CHAPTER 2

REINFORCEMENT SHAPES CLINES IN FEMALE MATE DISCRIMINATION IN

*DROSOPHILA SUBQUINARIA*¹

¹ Bewick, Emily R. and K. A. Dyer. Submitted to *Evolution*, 5/27/14.

Abstract

Reinforcement of species boundaries may alter mate recognition in a way that also affects patterns of mate preference among conspecific populations. In the fly *Drosophila subquinaria*, females sympatric with the closely related species *D. recens* reject mating with heterospecific males as well as with conspecific males from allopatric populations. Here we assess geographic variation in behavioral isolation within and among populations of *D. subquinaria* and use cline theory to understand patterns of selection on reinforced discrimination and its consequences for sexual isolation within species. We find that selection has fixed rejection of *D. recens* males in sympatry, while significant genetic variation in this behavior occurs within allopatric populations. In conspecific matings, patterns of sexual isolation are also asymmetric and divergent in populations that are sympatric with *D. recens*. The clines in behavioral discrimination within and between species are similar in shape and are maintained by strong selection in the face of gene flow, and we show that some of their genetic basis may be either shared or linked. Thus, while reinforcement can drive extremely strong phenotypic divergence, the long-term consequences for incipient speciation depend on gene flow, genetic linkage of discrimination traits, and the cost of these behaviors in allopatry.

Introduction

When closely related species come into secondary contact following divergence, the opportunity for hybridization may arise. Natural selection can strengthen pre-mating isolation in zones of sympatry in response to reduced hybrid fitness (Dobzhansky 1951), and this reinforcement of species boundaries may be a key step to completing the speciation process when post-mating barriers are incomplete (Howard 1993; Noor 1999; Servedio and Noor 2003; Coyne and Orr 2004). There is now considerable theoretical and empirical evidence for reinforcement (reviewed in Coyne and Orr 2004), and recent work suggests it may play an important role in the diversification of certain taxa, for example fruit flies (e.g. Yukilevich 2012).

Natural selection to avoid mating with the wrong species may alter mate recognition systems in a way that introduces effects on mate choice within species (Howard 1993; Coyne and Orr 2004; Ortiz-Barrientos et al. 2009; Hoskin and Higgie 2010; Mendelson and Shaw 2012). This has been observed in some frogs and *Drosophila* species, in which sympatric females reject heterospecific males but also will not mate with conspecific allopatric males (Hoskin et al. 2005; Jaenike et al. 2006). For instance, this pattern may result if females that co-occur with the other species evolve to use population-specific rather than species-specific cues in mate discrimination (Hoskin and Higgie 2010). Reinforcing selection may also alter sympatric male mating signals in a way that makes them less attractive to allopatric females, which has been found in *D. serrata* (Higgie and Blows 2007). Thus, natural and sexual selection may interact during reinforcement, with a possible outcome of secondary isolation or ‘cascade reinforcement,’ whereby the processes that complete speciation between two species also

initiate isolation within a species (Howard 1993; Ortiz-Barrientos et al. 2009; Hoskin and Higgie 2010). We are in the early stages of understanding how widespread these effects are, but if they are common their consequences for diversity may be significant.

Thus far, the genetic basis of reinforced discrimination has been identified in only a few systems, and there is no case of cascade reinforcement in which the genes that underlie both reinforced and within-species discrimination are known. Even without the genetic loci in hand, assaying the phenotypic distribution of reinforced discrimination through a cline from allopatry to sympatry with the other species may give valuable information about the nature and strength of selection on mate discrimination (Jiggins and Mallet 2000; Gay et al. 2008), inform what maintains differences in mate discrimination between sympatric and allopatric populations, and provide information about the trait's genetic basis (Barton and Hewitt 1985; Barton and Gale 1993). This approach has not been used to understand the process of reinforcement, but mirrors studies that have assayed variation in traits related to reproductive isolation across the hybrid zone of two diverging species (e.g. Orr 1996; Gay et al. 2008; Gompert et al. 2010; Bimova et al. 2011). The shape of the cline for a given trait is determined by the strength of selection against hybrids, the extent of gene flow across the cline, and the amount of time since contact (Endler 1977). A narrow cline compared to neutral expectations suggests that selection maintains the cline despite gene flow (Lewontin and Krakauer 1973). For example, in the case of a tension zone, selection against hybrids is balanced by dispersal of parental types into the zone (Barton and Hewitt 1989).

Comparing clines of different traits can also inform our understanding of how these traits are related to each other. If secondary isolation (i.e., mate discrimination

among populations within a species) is a byproduct of reinforcement of species boundaries, we would expect that clines and patterns of variation in intra- and interspecific discrimination would mirror each other. Clines for different traits may become coupled if the same genes underlie the two traits or if different but linked genes are involved. Even without a genetic relationship between the traits, cline coupling can result from similar strength and type of selection on each cline. On the other hand, clines may also become coupled because different isolating barriers tend to attract each other. For instance, tension zones move toward environmental barriers, and different types of incompatibilities can become coupled together to strengthen overall isolation (Barton and Bengtsson 1986; Barton and Gale 1993; Barton and de Cara 2009; Bierne et al. 2011).

To date, there has been very little work to understand how much genetic variation segregates within sympatric and allopatric populations for reinforced discrimination against another species. Most studies focus on patterns of mate discrimination of sympatric versus allopatric populations, and often either combine multiple genetic lines to obtain a snapshot of a population or assay only one or a few lines from a population. In order to better understand the causes and consequences of selection for reinforced discrimination as well as its downstream effects on patterns of sexual isolation within species, in this study we use a cline theory approach to study the process of reinforcement in the fly *D. subquinaria*. This species shows a pattern of reinforcement against its sister species, *D. recens*, as well as behavioral isolation among conspecific populations (Jaenike et al. 2006). *D. subquinaria* and *D. recens* occur in western and eastern North America, respectively, and are in secondary contact for about 1200 km east of the Rocky Mountain range in Canada. These species are morphologically nearly identical and have

no known ecological differences; in sympatric populations they can be found on the same mushrooms that are utilized for all life stages. Where these species co-occur, *D. subquinaria* females discriminate strongly against *D. recens* males, whereas *D. subquinaria* females from allopatric populations will mate with *D. recens* males at a moderate frequency (Jaenike et al. 2006). Reinforcing selection on *D. subquinaria* may be due to a *Wolbachia* infection in *D. recens*, which causes offspring to die as a result of cytoplasmic incompatibility when a *D. subquinaria* female mates with a *D. recens* male (Werren and Jaenike 1995; Shoemaker et al. 1999). When *Wolbachia* is cured from *D. recens* the resulting hybrid male offspring are sterile, indicating these are distinct species.

D. subquinaria females from the region of sympatry with *D. recens* also strongly discriminate against conspecific males from allopatric populations (Jaenike et al. 2006). This isolation is primarily behavioral, as there are no known post-zygotic effects based on laboratory assays of fertility of F1 and F2 individuals from intra-specific crosses (Jaenike et al. 2006; K. Dyer, unpublished results). The post-zygotic effects in between-species crosses but not within species suggests that the behavioral isolation between conspecific populations may be a byproduct of reinforcement (i.e., cascade reinforcement), rather than the other way around. The basis for this female mate discrimination may be epicuticular compounds, which have been shown to be important signals in mate recognition both within and between species, and which show reproductive character displacement among populations that are sympatric vs. allopatric with *D. recens* (Curtis et al. 2013; Giglio and Dyer 2013; Dyer et al. 2014).

In this study, we explore geographic patterns of variation of behavioral isolation in *D. subquinaria* to investigate the relationship between inter- and intraspecific

reproductive isolation and patterns of selection on these traits. We measure female discrimination of *D. subquinaria* both within and among populations against *D. recens* males and *D. subquinaria* males, and ask: (1) How much variation is there in female mate discrimination, both within and among populations? (2) How do these behaviors change across the geographic range, from allopatry into sympatry with *D. recens*? (3) How steep are the geographic clines in discrimination behaviors, and what is the strength of selection on these traits? and (4) Is there evidence of a shared genetic basis or genetic linkage between these behaviors? Throughout, we address these questions within the framework of cline theory and the genetic structure among populations of *D. subquinaria*.

Materials and Methods

Fly collections and species diagnosis

Flies were collected from natural populations during the summers of 2009-2011 by sweep netting over baits of commercial mushrooms (Fig. 2.1, Table 2.1). We established laboratory isofemale lines, and maintained cultures at 20°C on a 12hr light/dark cycle on Instant *Drosophila* Medium (Carolina Biological, Cary NC) supplemented with commercial *Agaricus bisporus* mushrooms. To identify flies as *D. recens* or *D. subquinaria* we used multiple molecular markers. Nearly all *D. recens* harbor *Wolbachia* (Shoemaker et al. 1999; Jaenike et al. 2006), which we screened for by PCR of the *wsp* gene (Zhou et al. 1998). Using population genetic data of individuals of known species composition (Jaenike et al. 2006; Dyer et al. 2011; K. Dyer, unpubl. data),

we designed restriction fragment assays for the mtDNA gene *Cytochrome Oxidase I* (*COI*) and the Y-linked gene *kl-3* based on fixed differences between species.

Patterns of population genetic differentiation

We characterized patterns of genetic diversity and population differentiation of *D. subquinaria* at the mtDNA *COI* gene and at eight autosomal unmapped microsatellites. In most cases we used wild-caught individuals; for a few populations we increased the sample size by adding single flies from isofemale lines, and for these individuals we randomly sampled one allele per microsatellite locus (Table 2.1). We sequenced 549 bp of *COI* as described previously (Jaenike et al. 2006), and we did not sequence any *D. subquinaria* that carried a *D. recens* mtDNA haplotype. The eight microsatellite markers are listed in Appendix A, Table S1 and were initially developed from *D. recens* (Ross et al. 2003). To genotype microsatellite loci, we used the three-primer fluorescent amplification method of Shuelke (2000) as implemented in Miles et al. (2008). Amplicons were run concurrent with a size standard, and fragments were scored with GeneMarker (SoftGenetics, State College, PA).

We calculated allelic richness, observed and expected heterozygosity, and tested for departures from Hardy-Weinberg Equilibrium (HWE) and for linkage disequilibrium (LD) using Arlequin version 3.5.1.2 (Excoffier and Lischer 2010), with significance determined by 10,000 permutations of the data. Based on these results, we used all eight microsatellite loci in our analyses because we found no consistent deviation from HWE or LD between pairs of loci across populations (Appendix A, Table S2). We estimated pairwise F_{ST} using Arlequin and pairwise S_{nn} using DnaSP version 5.10.01 (Librado and

Rozas 2009). We resolved a mtDNA haplotype network using SplitsTree (Huson and Bryant 2006), implementing a neighbor-joining algorithm and a Kimura two-parameter model. We also investigated patterns of genetic differentiation among geographic regions using an analysis of molecular variance (AMOVA) and the associated Φ statistics, as implemented in Arlequin. For the mtDNA, we also repeated the analyses including *COI* sequences from flies collected between 1997-2005 (Jaenike et al. 2006). Finally, to infer the number of genetic clusters (K) in the microsatellite data, we used the program Structure version 2.3.4 (Pritchard et al. 2000; Hubisz et al. 2009). We used an admixture model with correlated allele frequencies, and ran the program five times at each value of K with a burn-in of 100,000 and run length of 250,000 steps. We identified the most likely K based on the mean $\ln(\text{Pr}|K)$ across values of K and also using the ΔK method of Evanno et al. (2005).

Mating assays for female discrimination

We used no choice mating assays to measure mate discrimination by each of 55 isofemale lines from the eight sampled populations of *D. subquinaria*, which included 2-16 isofemale lines/population (Table 2.1, Fig. 2.1). Lines were maintained in the laboratory for at least ten generations before being used in mating assays. Females from each line were tested against each of four male types: coastal allopatric, inland allopatric, and sympatric *D. subquinaria*, and sympatric *D. recens* (Fig. 2.1). The subdivision of allopatric populations into coastal and inland regions is based on previous work in this system (Jaenike et al. 2006; Dyer et al. 2014). All the males used in mate trials were taken from large mixed stocks; each was created by combining virgin females and males

in equal numbers from 10-25 isofemale lines. The sympatric stock combined 25 lines from Kawtikh, Hinton, and Canmore, the inland allopatric stock combined ten lines from Deary, Missoula, and Shuswap, and the coastal allopatric stock combined ten lines from Portland and Seattle. A mixed stock of *D. recens* was created in the same way, using eight lines from the same three sympatric populations. Each mixed stock was maintained for at least four generations, mixing flies across vials each generation, before being used in a mate trial.

Virgins were collected under CO₂ anesthesia and stored 10-15 flies/vial, and all flies used in mating experiments were seven days old. Mating assays took place in 4 mL vials that contained a blended mushroom-agar food, and commenced within an hour of the incubator lights turning on. Flies were placed into vials by air aspiration, and we recorded whether each pair copulated within two hours. Isofemale lines were randomized among blocks, and within a block each isofemale line was tested against every male type. Each isofemale line was tested in at least three blocks for a minimum of 18 mating assays per male type.

Variation in mate discrimination within and among populations

Each isofemale line was tested in several blocks but not in every block; thus, we first tested for a block effect using a logistic regression for each individual isofemale line, considering block and male type as fixed effects. Because only one line had a significant block effect (Shuswap 3) we combined data across blocks, and the results did not differ if we excluded this line from the analyses. We tested for large-scale variation in mate discrimination using a logistic regression with male type, female region, and their

interaction as fixed effects in the model. To investigate fine-scale geographic variation in mate discrimination we used logistic regressions with female population and isofemale line nested within population as fixed effects in the model. We completed this analysis separately for each combination of male type and female region. This approach results in more statistical tests than a single model with all of the different factors; however, in spite of extensive efforts we were unable to obtain consistent convergence when we used fewer but more complicated models. This is likely because large portions of the mating data show no variation. Statistical analyses used JMP version 10 (SAS Institutes, Cary, NC).

We also asked whether females from lines that were less likely to mate with *D. recens* were less likely to mate with coastal allopatric *D. subquinaria*. As we found no variation for female mating rate of sympatric *D. subquinaria* against *D. recens*, we used the 20 lines from the nearby inland allopatric *D. subquinaria* populations (Deary, Missoula, Shuswap). We used linear regression to test for a correlation of mating rate between male types, and used the propensity of these lines to mate with their own males as a control for general level of discrimination.

Cline and selection analyses on mate discrimination

We analyzed cline shapes using the simulated annealing algorithm implemented in Cfit7 (Gay et al. 2008). We fit sigmoidal curves to four phenotypic clines (discrimination against *D. subquinaria* coastal allopatric males, inland allopatric males, sympatric males, and *D. recens* sympatric males) and two genetic clines (multilocus microsatellite and mtDNA *COI*). For each behavioral trait we treated each mean

isofemale line mating rate as an individual trait value. For the microsatellites, we used the multilocus assignment values of each genotyped individual to one of the clusters under $K=2$, and treated the data as a quantitative trait in the analyses. For the mtDNA, we scored whether each individual carried a coastal or inland haplotype based on the minimum spanning network (Appendix A, Fig. S1), and fit a simple sigmoid curve with right tail.

We used the phenotypic variation to test the fit of three cline models that correspond to different levels of introgression across the range, as implemented in Cfit7 (Jiggins and Mallet 2000; Gay et al. 2008). In the null unimodal model, there is high migration across the cline center with all individual phenotypic values drawn from a single distribution. In this model the distribution of phenotypes in the center resembles a hybrid swarm. In the bimodal model, phenotypic values come from two separate distributions that are on either side of the cline center, with very few intermediates. The trimodal model allows a third distribution for individuals in the center of the cline, thus these individuals have mixed phenotypes from the two opposite distributions. We used the Akaike Information Criterion (AIC) to determine the best fitting model for each cline and checked for convergence via multiple runs with different random seeds. In a few cases, our runs did not converge despite increasing iterations, thus we also report the standard deviation across at least five runs. This lack of convergence was likely due to high variation and low sample sizes within some populations.

We compared the clines in female discrimination against coastal allopatric males and *D. recens* males by asking whether they share a cline center (i.e., are coincident) and width (i.e., are concordant). We used Cfit7 to constrain the slopes, centers, or both slope

and center of these clines. We assumed the bimodal distribution for each cline since it is preferred for both (see Results), and used the AIC to determine the best fitting model.

We used the ecotone model to estimate the strength of selection acting on each cline. This model assumes a step-like change in the selection environment and that dispersal distance is similar to or larger than the cline width (Slatkin 1973; Endler 1977). Selection, s , is estimated following $w=\sigma/\sqrt{s}$, where w is the cline width (estimated as $w=4/\text{slope}$, as in Gay et al. 2008) and σ is the standard deviation of adult-offspring dispersal distance (Haldane 1948; Bazykin 1969; Slatkin 1973). We estimated the variance in dispersal to be about 1.5 km based on a mark-recapture study of *D. falleni*, a sympatric mushroom-feeding member of the quinaria group (Montague 1985). We calculated selection under a range of dispersal variances, including 1 km, 2km, 5km, and 10km, as longer range wind dispersal is likely to occur occasionally.

Results

Species composition and patterns of gene flow

D. recens and *D. subquinaria* were sympatric in the three sampled populations east of the Rocky Mountains, where the percent *D. recens* ranged from 53% to 92% of the total composition of the two species (Table 2.1). Of the total 208 sampled *D. subquinaria* individuals, three (1.4%) carried a *D. recens* mtDNA haplotype, indicative of hybridization and subsequent introgression of the mtDNA from *D. recens* into *D. subquinaria*. We also identified one individual of *D. subquinaria* from Missoula infected with *Wolbachia*, though sequencing revealed this strain is not closely related to the one found in *D. recens* (data not shown).

Using the mtDNA *COI* gene, the primary axis of genetic differentiation in *D. subquinnaria* is to the east and west of the Coast Mountains, and not between sympatry and allopatry with *D. recens*. The largest percentage of variation among groups is explained when populations are separated into two groups, consisting of inland (allopatric inland and sympatric populations) and coastal populations ($\Phi_{CT-coast\ vs\ inland}=0.38$, $P=0.04$; Appendix A, Table S3). An AMOVA that excludes coastal populations indicates there is substantial gene flow between sympatric and nearby allopatric populations ($\Phi_{CT-sym\ vs\ inland\ allo}=0.012$, $P=0.30$; Appendix A, Table S3). Similar patterns were obtained using pairwise F_{ST} and S_{nn} (Appendix A, Table S4), and also visually from a haplotype network (Appendix A, Fig. S1). When we combine our data with those from an earlier study (Jaenike et al. 2006), we find consistent patterns of differentiation, indicating stable genetic structure across years (Appendix A, Table S3).

In contrast to the mtDNA, the strongest axis of population differentiation for the eight microsatellite loci is on either side of the Rocky Mountains, consistent with the division between allopatry and sympatry with *D. recens*. First, seven of nine pairwise F_{ST} comparisons between sympatric and inland allopatric populations were significant, whereas two of six comparisons between coastal and inland allopatric populations were, suggesting a stronger break between sympatry and allopatric regions (Appendix A, Table S5). Other measurements of genetic distance gave similar results (results not shown). Second, using the program Structure, the most probable number of genetic clusters was $K=2$ using either the highest averaged likelihood (average $LnL=-3502.52$) or ΔK ($\Delta K=256.64$; Appendix A, Table S6). The two clusters primarily separate the sympatric and allopatric populations (Appendix A, Table S7; Fig. S2A). We also ran Structure

excluding the two coastal allopatric populations, with consistent results (Appendix A, Tables S6 and S7; Fig. S2b). Finally, using an AMOVA with all eight populations, the largest percentage of variation among groups (13%) is explained when populations are separated into two groups based on sympatry with *D. recens* ($\Phi_{CT-sym\ vs\ allo}=0.13$, $P=0.02$). However, excluding the coastal allopatric populations from the AMOVA, as these are the most distant from sympatry with *D. recens*, indicates the presence of gene flow between sympatric and nearby allopatric inland populations ($\Phi_{CT-sym\ vs\ inland\ allo}=0.14$, $P=0.10$; Appendix A, Table S3).

Variation in mate discrimination within and among populations

There is tremendous genetic variation in mating rate that depends on the geographic origin of both the male and the female (Fig. 2.2A-D, Appendix A, Fig. S3). Using a model that included all of the data and tested for the effects of male type, female region, and their interaction, we found that all three effects were highly significant (male type: $\chi^2=1440$, $P<0.0001$; female region $\chi^2=897$, $P<0.0001$; interaction $\chi^2=814$, $P<0.0001$). We subdivided the mating assay data by the three female regions and the four male types, and for each male-female combination we tested for variation in mating rate among populations and among lines (nested within population) (Table 2.2). We did not perform this analysis for the sympatric females with *D. recens* males, as none of the 689 pairs of this type copulated.

Within species, the mating rates for “home” combinations where the female and male are from the same geographic region are generally high and show no significant genetic variation (Table 2.2; Appendix A, Fig. S3), as expected if female preferences and

male traits are at regional optima. In contrast, there is significant genetic variation in mating rates when conspecific females and males from different geographic regions are paired. Most of the genetic variation in mating rate appears to be among lines rather than among populations within a region, though the number of populations surveyed in each region is small (Table 2.2).

Between species, while no sympatric *D. subquinaria* female mated with a *D. recens* male, *D. subquinaria* females from outside of this region were willing to mate with these heterospecific males. Across the 37 allopatric lines we surveyed, females mated with *D. recens* males on average 26% of the time (range 0.03-0.83 across lines). The genetic variation among lines in both the inland and coastal parts of allopatry is highly significant (coastal: $\chi^2=64$, $P<0.0001$; inland: $\chi^2=140$, $P<0.0001$; Table 2.2). The lack of any mating within sympatry coupled with the high level of variation in allopatry suggest that selection against mating with *D. recens* in sympatry may be very strong.

Comparing across lines, our data indicate there may be some genetic linkage or shared genetic basis of female discrimination against both *D. recens* males and *D. subquinaria* coastal males. Inland allopatric females that mate less often with *D. recens* males also mate less often with *D. subquinaria* coastal males ($r^2=0.25$, $F_{1,18}=6.1$, $P=0.024$; Fig. 2.3). This is not because these females are generally more choosy, as there is no correlation between mating rate between coastal allopatric males and their own inland allopatric males ($r^2=0.09$, $F_{1,18}=1.8$, $P=0.20$; Fig 3).

Cline and selection analysis of mate discrimination

A single model did not fit for all clines (Tables 3, S8), suggesting that the biological basis of the clines differs. For discrimination against *D. recens* males the bimodal model was the best fit ($\Delta AICc > 2.0$). The bimodal model was also the best fit for the cline in discrimination against coastal allopatric males, though it was not a statistically significant improvement over the unimodal model ($\Delta AICc = 1.19$). Both the microsatellite cline and the cline in discrimination against inland allopatric males were best described by the trimodal model ($\Delta AICc > 2.0$). Finally, the sympatric cline was best fit by the unimodal model ($\Delta AICc > 2.0$).

Using the best-fitting models, all four of the behavior clines as well as the microsatellite cline have centers that fall in in the middle of the Rocky Mountains and also in the region where allopatry turns into sympatry with *D. recens* (Fig. 2.2A-E). The cline of discrimination against sympatric males is weaker than the others and in the opposite direction to the other behavior clines, consistent with slightly lower mating rates of allopatric females with sympatric males. The center of the mtDNA cline, in contrast, falls at the Coast Mountains (Fig. 2.2F). The clines vary in width, with the discrimination cline against *D. recens* males being the steepest at 4 km (Table 2.3). The other phenotypic cline widths are generally wider, for instance the cline in discrimination against inland allopatric males and the microsatellite cline range from 62-110 km wide (Table 2.3). The cline in discrimination against allopatric coastal males differs in width depending on the model, with 4 km under the bimodal model up to 159 km under the unimodal model (Table 2.3). For this cline we use the average of the slopes of these two models (Appendix A, Table S9) to calculate the width in Fig. 2.2B and in the selection

analysis (Table 2.4). Finally, we find that the clines in discrimination against *D. recens* males and coastal allopatric males are both concordant and coincident. Using a bimodal model, the best-fitting model for these clines constrains both width and center simultaneously ($\Delta AICc > 3$; Appendix A, Table S9). This indicates that these clines in discrimination may be coupled.

Using a range of dispersal values (variance 1-10 km) and the width from each best fitting cline model, the selection estimates vary from very weak to extremely strong across clines (Table 2.4). For the behavioral traits, discrimination against *D. recens* and coastal allopatric males is strong ($s=0.06-0.62$ and $0.014-0.14$, respectively). In contrast, selection against inland allopatric males and sympatric males is weak, with the highest values of s at 0.0008 and 0.02 , respectively. At the molecular markers, the barrier to gene flow at the Coast Mountains for the mtDNA is strong, with s as high as 0.6 . In contrast, selection on the microsatellite cline is weak (s up to 0.003), supporting some gene flow among populations of *D. subquinnaria* across the Rocky Mountains and also across the boundary between sympatry and allopatry with *D. recens*.

Discussion

Reinforcement involves the interaction of selection and gene flow to strengthen prezygotic isolation between hybridizing species (Servedio and Kirkpatrick 1997; Noor 1999; Servedio and Noor 2003; Coyne and Orr 2004). As a consequence of reinforcement, changes in mate recognition systems may create isolation among populations within a species (Howard 1993; Coyne and Orr 2004; Hoskin et al. 2005; Ortiz-Barrientos et al. 2009). How natural and sexual selection interact to alter mate

recognition systems and subsequently drive reproductive isolation is not well understood (Andersson 1994; Coyne and Orr 2004; Hoskin and Higgie 2010; Mendelson and Shaw 2012). Here we utilize a cline theory approach to provide a window into the processes that underlie behavioral isolation between *D. recens* and *D. subquinaria* and among populations of *D. subquinaria*. Studying the variation in mate discrimination both within and among populations may help illuminate both the causes and consequences of selection for reinforced species boundaries, including how selection drives the divergence of mate recognition systems and thus ultimately speciation. We place this variation in phenotypic traits within the context of background genetic variation to reveal where selection acts and how strong that selection is.

Genetic structure of the sympatric-allopatric boundary

For the nuclear genome, genetic differentiation among populations of *D. subquinaria* is moderate and primarily between populations on either side of the Rocky Mountains, coincident with the boundary for risk of hybridization with *D. recens* (Fig. 2.2E). This pattern of population differentiation is in contrast to what has been found using the mtDNA, which identifies a major phylogeographic break at the Coast Mountains (this study; Jaenike et al. 2006). Why the mtDNA and nuclear markers have such contrasting patterns of genetic differentiation is unknown. Ecological selection may promote divergence on either side of the Coast Mountains (Dyer et al. 2014; M. Stribos and K. Dyer, unpubl. data), but this is hard to interpret with respect to mating patterns since female preferences and male traits generally map to nuclear loci.

Tension zones can be pushed toward regions of low population density, and cause otherwise unlinked loci to accumulate genetic differences on either side of this barrier (reviewed in Barton and Hewitt 1985; Abbott et al. 2013). Coincident with this, divergent selection in mate discrimination may be contributing to differentiation between sympatric and allopatric populations of *D. subquinaria*. However, this differentiation appears to be in the early stages. Specifically, in the AMOVAs that consider only sympatric populations and the nearby inland allopatric populations, the Φ_{CT} values are low and non-significant (Appendix A, Table S3). Furthermore, the microsatellite cline fits a trimodal distribution, indicating differentiation with the potential for introgression of alleles across the cline (Gay et al. 2008). Thus, we find that across the sympatry-allopatry boundary there is some gene flow, though not too much, as is necessary for reinforcing selection to operate (Kirkpatrick 2000; reviewed in Coyne and Orr 2004).

Reinforced discrimination against *D. recens*

A striking result of our study is that we find no genetic variation in rejection of *D. recens* within and among sympatric populations of *D. subquinaria*. This is independent of the frequency of *D. recens*, which ranged from ~50-90% across populations at the time of sampling. Others have suggested that the frequency of the sister species may affect the strength of reinforcing selection (Yukilevich 2012; Nosil 2013), and thus in the future it would be useful to sample populations where *D. recens* is rare relative to *D. subquinaria*. Our estimates of selection to discriminate against *D. recens* males are very strong (s up to 0.62), which is consistent with the death of all hybrid offspring of *D. subquinaria* females. The upper values are among the highest selection coefficients estimated in

natural populations (e.g., Hoekstra et al. 2001; Kingsolver et al. 2001). Importantly, the width of this cline is much narrower than for the microsatellites, even though the centers are near each other, suggesting that selection is operating in the face of gene flow to maintain it. Our cline analysis also indicates that a bimodal distribution is the best fit for the phenotypic distribution of this trait. In the tension and hybrid zone literature, bimodality is considered a strong indicator of nearly complete speciation and occurs when prezygotic barriers are present and/or postzygotic selection are strong (Jiggins and Mallet 2000). We note that in this case we are analyzing a cline within a single species and not across two species, which highlights the potential for reinforcement to generate strong phenotypic divergence within *D. subquinaria*.

In contrast to the area of sympatry, we find substantial genetic variation among allopatric *D. subquinaria* lines for rejection of *D. recens*, from complete rejection to nearly complete acceptance, with a range of values in between (Fig. 2.2A; Appendix A, Fig. S3). Without knowing the genetic basis of the reinforced discrimination we cannot determine whether this variation is a result of segregating ancestral variation or due to migration from sympatric populations. However, we suggest it is more likely the latter given that our genetic markers indicate there is migration between regions that are sympatric and allopatric with *D. recens*.

The finding of segregating genetic variation among allopatric lines for reinforced discrimination suggests three things about reinforcement in this system. First, the broad distribution of discrimination found among lines in allopatry suggests that reinforced discrimination may not be a simple all-or-nothing trait. Instead, the genetic basis may be more complicated than a single locus of large effect and, as has been found in *D.*

pseudoobscura, there may be multiple genetic routes to reinforced discrimination in *D. subquinaria* (Ortiz-Barrientos et al. 2004). Second, alleles that confer reinforced discrimination against *D. recens* may not be very costly to harbor outside of sympatry. However, they do not seem to be spreading to fixation, which suggests that they are probably not beneficial and there are likely tradeoffs between species recognition and mate quality recognition, as has been suggested in other systems (Pfennig 2000; Higgie and Blows 2007). Third, the presence of these alleles in allopatry suggests that these populations harbor the genetic material to respond to selection to increase discrimination and maintain species boundaries. Thus, if *D. recens* were to expand its range westward or if levels of gene flow were high, the combination of strong selection due to the death of hybrid offspring and the presence of genetic variation in discrimination against *D. recens* suggests these alleles would increase in frequency to prevent hybridization between species. These findings are only evident because we assayed within-population variation in discrimination; in contrast, most other studies of reinforcement combine across lines to estimate the population level of discrimination.

Behavioral isolation within *D. subquinaria*

Within *D. subquinaria*, our analyses indicate that female preferences are locally adapted, differ across the geographic range of the species, and are maintained in the face of gene flow. Consistent with previous work (Jaenike et al. 2006), sympatric females show the highest preference for their own males and discriminate against males from other geographic regions. Sympatric females mate with coastal allopatric males less than 10% of the time, though there is variation across lines in the level of discrimination

(Table 2.2). This creates a steep cline and strong selection (s up to 0.14). There is a slight preference for the bimodal model, as for the *D. recens* cline, which indicates strong phenotypic divergence in mate discrimination across the cline and reflects the presence of premating barriers. In contrast, these sympatric females do not discriminate nearly as strongly against inland allopatric males as against coastal males. This results in a cline against inland allopatric males that is much wider than for coastal males, and which supports a trimodal phenotypic distribution such that there are abundant hybrid phenotypes in the contact zone (Barton and Hewitt 1985; Gay et al. 2008). This inland allopatric region may be a ‘melting pot’ where alleles from the coastal and sympatric regions come together by migration.

Females from throughout the geographic range of *D. subquinaria* will mate with males from the region of sympatry with *D. recens*. Selection against these males is weak ($s < 0.02$), and a unimodal phenotypic distribution is preferred, indicative of significant introgression across the cline. There is a slight decrease in mating rates of allopatric females with sympatric males relative to their own males (Fig. 2.2D), which could result from a change in the mating signals of these males due to reinforced selection by *D. recens*, similar to what is seen in *D. serrata* (Higgie and Blows 2007). Indeed, in *D. subquinaria* the epicuticular profiles of sympatric individuals differ from allopatric individuals, consistent with reinforcement driving differences in male mating signals across populations (Dyer et al. 2014). Even a low cost conditions of strong sexual selection may be sufficient to prevent the sympatric male signals from spreading to high frequency in allopatric populations.

Cascade reinforcement and the potential for incipient speciation

One way that cascade reinforcement can be facilitated is if the genes responsible for reinforced pre-zygotic isolation between species also cause sexual isolation within species. The genetic basis of reinforced discrimination is known in only a few systems (e.g. Ortiz-Barrientos et al. 2004; Saether et al. 2007; Hopkins and Rausher 2011), and to our knowledge the genetic basis of cascade reinforcement has not been identified in any system. Complicating this, there may be multiple genetic mechanisms of reinforced female mate discrimination within a species, as is the case in *D. pseudoobscura* (Ortiz-Barrientos et al. 2004), and the same may be true for within-species discrimination due to cascade reinforcement. This underscores the importance of accounting for variation not only among populations, but within them as well. If the genetic basis of reinforced and within species discrimination is the same or tightly linked, we predict that at the broad scale the between and within species discrimination clines should be tightly correlated, and that at the fine scale, we should see an association of between and within species discrimination among isofemale lines. We find support for both predictions. First, the bimodal model is the best fitting phenotypic distribution for both the *D. recens* and coastal allopatric clines, whereas this model does not fit any other cline (Table 2.3). Furthermore, these clines share the same shape and are both centered at the same location (Appendix A, Table S9). Second, females from inland allopatric lines that are more likely to mate with *D. recens* males are also more likely to mate with allopatric coastal conspecific males (Fig. 2.3). Because the lines that are more choosy against both types of males do not also mate less against their own males (Fig. 2.3), this suggests at least some of the genetic basis of discrimination between and within species is shared or linked.

Further support for an association of discrimination phenotypes is that females use epicuticular hydrocarbons to discriminate against both conspecific and heterospecific males (Dyer et al. 2014). However, while these results are suggestive, genetic mapping of both behaviors is necessary before any firm conclusions can be made.

Alternatively, selection on these traits may be independent, but clines are coupled through means other than reinforcement of mate discrimination. With the exception of the mtDNA, all of our cline centers fall within the Rocky Mountain range. It is possible that the physical barrier of the mountain range is recruiting clines to it, which acts to sharpen clines and enhance barriers to gene flow (Barton and Hewitt 1985; Bierne et al. 2011; Abbott et al. 2013). This can allow the build up of reproductive barriers (Felsenstein 1981; Barton and de Cara 2009). On the other hand, if a single physical barrier to dispersal is coupling otherwise independent clines, we would predict the shapes and phenotypic distribution models to be the same across clines. Instead, we see that some traits move across the physical barrier somewhat easily (i.e. discrimination against sympatric and inland allopatric males), while others appear to be maintained by selection and not solely by reduced dispersal (i.e., discrimination against coastal and *D. recens* males).

In conclusion, cascade reinforcement involves the interaction of selection, migration, and the genetic relationship between phenotypes, whose relative roles will ultimately determine the outcome of incipient speciation. Our results suggest that genetic variation in mate discrimination within populations can be high in certain circumstances, and reinforcing selection can be strong enough to drive phenotypes to fixation in sympatry. Allopatric females will accept the male signal trait from both allopatric and

sympatric populations of *D. subquinaria*, and many will accept heterospecific *D. recens* males as well. In contrast, sympatric *D. subquinaria* females have narrowed the range in the male signals they find acceptable to only include the conspecific males they co-occur with. This asymmetry could have significant consequences for the direction of gene flow among populations and the potential for divergence within *D. subquinaria*. Because behavioral isolation among conspecific populations is asymmetric and gene flow occurs among populations, this may slow the process of reproductive isolation to maintain the tension zone indefinitely. On the other hand, the coupling of clines and genetic association of discrimination traits may allow for genetic divergence to accumulate across the genome and thus promote eventual complete isolation even in the face of ongoing gene flow. Ultimately, to distinguish whether these clines have been coupled by exogenous factors or are acted upon by the same forces of selection we need to uncover the genetic basis of the within and between-species female discrimination traits as well as of the male signal trait. This, in combination with identifying the pattern of fine-scale genomic differentiation both among populations and between species, will provide a robust understanding of the mechanisms of ongoing divergence.

Acknowledgements

We thank A. Bewick, K. Bockrath, C. Debban, J. Jaenike, C. Pinzone, and A. Swiegart for assistance with fly collecting; M. Bray, B. White, J. Lopez, and R. Webster for laboratory assistance; and J. Good, D. Hall, J. Jaenike, members of the Dyer lab, and three anonymous reviewers for useful comments on the manuscript. This work was funded by NSF grants 1149350 and 1132807 to KAD and 1110462 to ERB and KAD, by the NIGMS of the NIH under award T32GM007103 to ERB, and by a UGA Alton Fellowship to ERB.

Literature Cited

- Andersson, M. B. 1994. Sexual selection. Princeton University Press, Princeton, N.J.
- Barton, N. and B. O. Bengtsson. 1986. The barrier to genetic exchange between hybridizing populations. *Genetics* 57:357-376.
- Barton, N. H. and M. A. R. de Cara. 2009. The evolution of strong reproductive isolation. *Evolution* 63:1171-1190.
- Barton, N. H. and K. S. Gale. 1993. Genetic analysis of hybrid zones *in* R. G. Harrison, ed. Hybrid zones and the evolutionary process. Oxford Univ. Press, Oxford, UK.
- Barton, N. H. and G. M. Hewitt. 1985. Analysis of hybrid zones. *Annu. Rev. Ecol. Syst.* 16:113-148.
- Barton, N. H. and G. M. Hewitt. 1989. Adaptation, speciation and hybrid zones. *Nature* 341:497-503.
- Bazykin, A. D. 1969. Hypothetical mechanism of speciation. *Evolution* 23:685-&.

- Bierne, N., J. Welch, E. Loire, F. Bonhomme, and P. David. 2011. The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Mol. Ecol.* 20:2044-2072.
- Bimova, B. V., M. Macholan, S. J. E. Baird, P. Munclinger, P. Dufkova, C. M. Laukaitis, R. C. Karn, K. Luzynski, P. K. Tucker, and J. Pialek. 2011. Reinforcement selection acting on the European house mouse hybrid zone. *Mol. Ecol.* 20:2403-2424.
- Coyne, J. A. and H. A. Orr. 2004. *Speciation*. Sinauer Associates, Sunderland (Massachusetts).
- Curtis, S., J. L. Sztepanacz, B. E. White, K. A. Dyer, H. D. Rundle, and P. Mayer. 2013. Epicuticular compounds of *Drosophila subquinaria* and *D. recens*: identification, quantification, and their role in female mate choice. *J. Chem. Ecol.* 39:579-590.
- Dobzhansky, T. 1951. *Genetics and the origin of species*. Columbia Univ. Press, New York.
- Dyer, K. A., C. Burke, and J. Jaenike. 2011. *Wolbachia*-mediated persistence of mtDNA from a potentially extinct species. *Mol. Ecol.* 20:2805-2817.
- Dyer, K. A., B. E. White, J. L. Sztepanacz, E. R. Bewick, and H. D. Rundle. 2014. Reproductive character displacement of epicuticular compounds and their contribution to mate choice in *Drosophila subquinaria* and *D. recens*. *Evolution* 68:1163–1175.
- Endler, J. A. 1977. *Geographic variation, speciation, and clines*. Princeton University Press, Princeton, N.J.

- Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14:2611-2620.
- Excoffier, L. and H. E. L. Lischer. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10:564-567.
- Felsenstein, J. 1981. Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution* 35:124-138.
- Gay, L., P. A. Crochet, D. A. Bell, and T. Lenormand. 2008. Comparing clines on molecular and phenotypic traits in hybrid zones: a window on tension zone models. *Evolution* 62:2789-2806.
- Giglio, E. M. and K. A. Dyer. 2013. Divergence of premating behaviors in the closely related species *Drosophila subquinaria* and *D. recens*. *Ecol. Evol.* 3:365-374.
- Gompert, Z., L. K. Lucas, J. A. Fordyce, M. L. Forister, and C. C. Nice. 2010. Secondary contact between *Lycaeides idas* and *L. melissa* in the Rocky Mountains: extensive admixture and a patchy hybrid zone. *Mol. Ecol.* 19:3171-3192.
- Haldane, J. B. S. 1948. The theory of a cline. *J. Genet.* 48:277-284.
- Higgie, M. and M. W. Blows. 2007. Are traits that experience reinforcement also under sexual selection? *Am. Nat.* 170:409-420.
- Hoekstra, H. E., J. M. Hoekstra, D. Berrigan, S. N. Vignieri, A. Hoang, C. E. Hill, P. Beerli, and J. G. Kingsolver. 2001. Strength and tempo of directional selection in the wild. *P. Natl. Acad. Sci. USA* 98:9157-9160.

- Hopkins, R. and M. D. Rausher. 2011. Identification of two genes causing reinforcement in the Texas wildflower *Phlox drummondii*. *Nature* 469:411-414.
- Hoskin, C. J. and M. Higgie. 2010. Speciation via species interactions: the divergence of mating traits within species. *Ecol. Lett.* 13:409-420.
- Hoskin, C. J., M. Higgie, K. R. McDonald, and C. Moritz. 2005. Reinforcement drives rapid allopatric speciation. *Nature* 437:1353-1356.
- Howard, D. J. 1993. Reinforcement: Origin, dynamics, and fate of an evolutionary hypothesis. Pp. 46-69 in H. RG, ed. *Hybrid zones and the evolutionary process*. Oxford University Press, New York.
- Hubisz, M. J., D. Falush, M. Stephens, and J. K. Pritchard. 2009. Inferring weak population structure with the assistance of sample group information. *Mol. Ecol. Resour.* 9:1322-1332.
- Huson, D. H. and D. Bryant. 2006. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* 23:254-267.
- Jaenike, J., K. A. Dyer, C. Cornish, and M. S. Minhas. 2006. Asymmetrical reinforcement and *Wolbachia* infection in *Drosophila*. *Plos Biol.* 4:1852-1862.
- Jiggins, C. D. and J. Mallet. 2000. Bimodal hybrid zones and speciation. *Trends Ecol. Evol.* 15:250-255.
- Kingsolver, J. G., H. E. Hoekstra, J. M. Hoekstra, D. Berrigan, S. N. Vignieri, C. E. Hill, A. Hoang, P. Gibert, and P. Beerli. 2001. The strength of phenotypic selection in natural populations. *Am. Nat.* 157:245-261.
- Kirkpatrick, M. 2000. Reinforcement and divergence under assortative mating. *P. Roy. Soc. Lond. B. Bio.* 267:1649-1655.

- Lewontin, R. C. and J. Krakauer. 1973. Distribution of gene frequency as a test of theory of selective neutrality of polymorphisms. *Genetics* 74:175-195.
- Librado, P. and J. Rozas. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451-1452.
- Mendelson, T. C. and K. L. Shaw. 2012. The (mis)concept of species recognition. *Trends Ecol. Evol.* 27:421-427.
- Miles, L. G., S. R. Isberg, C. Moran, C. Hagen, and T. C. Glenn. 2008. 253 Novel polymorphic microsatellites for the saltwater crocodile (*Crocodylus porosus*). *Conserv. Gen.* 10:963-980.
- Montague, J. R. 1985. Body size, reproductive biology, and dispersal behavior among artificial baits in *Drosophila falleni*. *Dros. Inf. Serv.* 61:123-126.
- Noor, M. A. F. 1999. Reinforcement and other consequences of sympatry. *Genetics* 83:503-508.
- Nosil, P. 2013. Degree of sympatry affects reinforcement in *Drosophila*. *Evolution* 67:868-872.
- Orr, M. R. 1996. Life-history adaptation and reproductive isolation in a grasshopper hybrid zone. *Evolution* 50:704-716.
- Ortiz-Barrientos, D., B. A. Counterman, and M. A. F. Noor. 2004. The genetics of speciation by reinforcement. *Plos Biol.* 2:2256-2263.
- Ortiz-Barrientos, D., A. Greal, and P. Nosil. 2009. The genetics and ecology of reinforcement implications for the evolution of prezygotic isolation in sympatry and beyond. Pp. 156-182 in C. D. Schlichting, and T. A. Mousseau, eds. *Year in Evolutionary Biology* 2009. Wiley-Blackwell, Malden.

- Pfennig, K. S. 2000. Female spadefoot toads compromise on mate quality to ensure conspecific matings. *Behav. Ecol.* 11:220-227.
- Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945-959.
- Ross, C. L., K. A. Dyer, T. Erez, S. J. Miller, J. Jaenike, and T. A. Markow. 2003. Rapid divergence of microsatellite abundance among species of *Drosophila*. *Mol. Biol. Evol.* 20:1143-1157.
- Saether, S. A., G. P. Saetre, T. Borge, C. Wiley, N. Svedin, G. Andersson, T. Veen, J. Haavie, M. R. Servedio, S. Bures, M. Kral, M. B. Hjernquist, L. Gustafsson, J. Traff, and A. Qvarnstrom. 2007. Sex chromosome-linked species recognition and evolution of reproductive isolation in flycatchers. *Science* 318:95-97.
- Servedio, M. R. and M. Kirkpatrick. 1997. The effects of gene flow on reinforcement. *Evolution* 51:1764-1772.
- Servedio, M. R. and M. A. F. Noor. 2003. The role of reinforcement in speciation: theory and data. *Annu. Rev. Ecol. Syst.* 34:339-364.
- Shoemaker, D. D., V. Katju, and J. Jaenike. 1999. *Wolbachia* and the evolution of reproductive isolation between *Drosophila recens* and *Drosophila subquinaria*. *Evolution* 53:1157-1164.
- Shuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nat. Biotechnol.* 18:233-234.
- Slatkin, M. 1973. Gene flow and selection in a cline. *Genetics* 75:733-756.
- Werren, J. H. and J. Jaenike. 1995. *Wolbachia* and cytoplasmic incompatibility in mycophagous *Drosophila* and their relatives. *Genetics* 75:320-326.

- Yukilevich, R. 2012. Asymmetrical patterns of speciation uniquely support reinforcement in *Drosophila*. *Evolution* 66:1430-1446.
- Zhou, W. G., F. Rousset, and S. O'Neill. 1998. Phylogeny and PCR-based classification of *Wolbachia* strains using wsp gene sequences. *P. Roy. Soc. Lond. B. Bio.* 265:509-515.

Table 2.1. Summary of the populations used in this study. For each population, the location, abbreviation (Abbr.), and region with respect to overlap with *D. recens* is indicated. For each allopatric (allo) population we indicate whether it is to the west (coast) or east (inland) of the Coast Mountains. Also shown is the abundance of *D. recens* relative to *D. subquinaria* at the time of collection with the total number of flies collected of both species, the number of alleles sampled for mtDNA sequencing and microsatellite genotyping, and the number of isofemale lines assayed for behavior.

Site	Abbr.	Region	Longitude	Latitude	Year	% <i>D. recens</i> among wild flies (total <i>recens</i> and <i>subquinaria</i>)	N mtDNA	N microsats	N mating trial lines
Portland, OR	POR	allo-coast	-122.68	45.52	2010	0 (70)	18	56	5
Seattle, WA	SEA	allo-coast	-122.33	47.61	2010	0 (40)	11	53	5
Deary, ID	DEA	allo-inland	-116.56	46.80	2009/2011	0 (7)	4	4	3
Missoula, MT	MIS	allo-inland	-113.99	46.87	2010/2011	0 (13)	12	56	13
Shuswap, BC	SHU	allo-inland	-119.71	50.79	2010	0 (4)	4	8	4
Canmore, AB	CAN	sympatric	-115.35	51.09	2010	53 (20)	6	19	2
Hinton, AB	HIN	sympatric	-117.57	53.41	2010	73 (118)	18	34	16
Kawtikh, AB	KAW	sympatric	-112.93	53.43	2010	92 (390)	13	32	7

Table 2.2. Geographic variation in female discrimination against each male type. For each female geographic region and male type, shown are the likelihood ratio test statistic χ^2 for the whole model and for variation among populations and lines, and the associated *P*-values. The Bonferroni corrected *P*-value is $0.05/11 = 0.0045$; values in bold are significant at this level. Sympatric females with *D. recens* males were not analyzed because none of the pairs of this combination mated.

Female region	Model effect	df	Male type			
			Coastal Allopatric	Inland Allopatric	Sympatric	<i>D. recens</i>
Allopatric coastal	Whole model	9	11.2 (<i>P</i> = 0.26)	12.5 (<i>P</i> = 0.19)	36.7 (<i>P</i> < 0.0001)	63.9 (<i>P</i> < 0.0001)
	Population	1	0.02 (<i>P</i> = 0.88)	0.43 (<i>P</i> = 0.51)	0.019 (<i>P</i> = 0.89)	1.1 (<i>P</i> = 0.29)
	Line (population)	8	11.2 (<i>P</i> = 0.19)	12.0 (<i>P</i> = 0.15)	36.6 (<i>P</i> < 0.0001)	62.1 (<i>P</i> < 0.0001)
Allopatric inland	Whole model	19	61.2 (<i>P</i> < 0.0001)	21.9 (<i>P</i> = 0.29)	53.5 (<i>P</i> < 0.0001)	139.5 (<i>P</i> < 0.0001)
	Population	2	22.2 (<i>P</i> < 0.0001)	0.023 (<i>P</i> = 0.99)	0.76 (<i>P</i> = 0.68)	8.4 (<i>P</i> = 0.015)
	Line (population)	17	39.8 (<i>P</i> = 0.0014)	21.5 (<i>P</i> = 0.21)	53.4 (<i>P</i> < 0.0001)	124.9 (<i>P</i> < 0.0001)
Sympatric	Whole model	24	52.3 (<i>P</i> = 0.0007)	76.1 (<i>P</i> < 0.0001)	39.7 (<i>P</i> = 0.0375)	na
	Population	2	0.000003 (<i>P</i> = 1.0)	2.293 (<i>P</i> = 0.23)	1.447 (<i>P</i> = 0.49)	na
	Line (population)	22	46.4 (<i>P</i> = 0.0018)	73.2 (<i>P</i> < 0.0001)	37.2 (<i>P</i> = 0.0227)	na

Table 2.3. Summary of the characteristics of clines in female mate discrimination. The cline centers and widths are with respect to the 50° transect angle in Figure 1, with the distance from the westernmost sampled point. The best-fitting models based on $\Delta AICc$ are shown in bold. See the text for a description of the three different models, and see Gay et al. (2008) for a description of the parameters included in each model.

Cline	Model	N Parameters	llmax	Center (km)	Width (km)	$\Delta AICc$
Allopatric coastal	Unimodal	7	64.35	501	159	1.19
	Bimodal	8	66.32	485	4	0
	Trimodal	12	69.52	469	183	6.91
Allopatric inland	Unimodal	7	52.99	622	10	2.62
	Bimodal	8	52.76	623	5	5.84
	Trimodal	12	62.19	589	110	0
Sympatric	Unimodal	7	44.29	618	22	0
	Bimodal	8	44.25	486	10	2.82
	Trimodal	12	49.15	86	971	5.86
<i>D. recens</i>	Unimodal	7	154.86	484	4	4.93
	Bimodal	8	158.09	484	4	0
	Trimodal	12	160.10	482	4	8.12
Microsatellites	Unimodal	7	128.21	484	4	84.66
	Bimodal	8	165.68	512	89	11.73
	Trimodal	12	175.15	509	62	0
mtDNA	Sigmoid	4	-28.05	171	4	na

Table 2.4. Summary of selection estimates. The selection coefficient (s) was inferred from estimated the cline width (w) under four estimates of variance in dispersal (σ^2), using the formula $\sqrt{s} = \sigma/w$. For the Allopatric coastal cline, the unimodal and bimodal models were indistinguishable statistically, thus we used the mean of the slope parameter estimates of both models to calculate width.

Cline	Width (km)	Dispersal (σ^2)			
	(= $\sigma/\sqrt{s}$)	1km	2km	5km	10km
Allopatric coastal	8.37	0.0143	0.0285	0.0714	0.1425
Allopatric inland	109.87	0.0001	0.0002	0.0004	0.0008
Sympatric	22.74	0.0019	0.0039	0.0043	0.0193
<i>D. recens</i>	4.00	0.0625	0.1250	0.3125	0.6241
Microsatellites	62.48	0.0003	0.0005	0.0013	0.0026
mtDNA	4.06	0.0606	0.1212	0.3033	0.6056

Fig. 2.1.

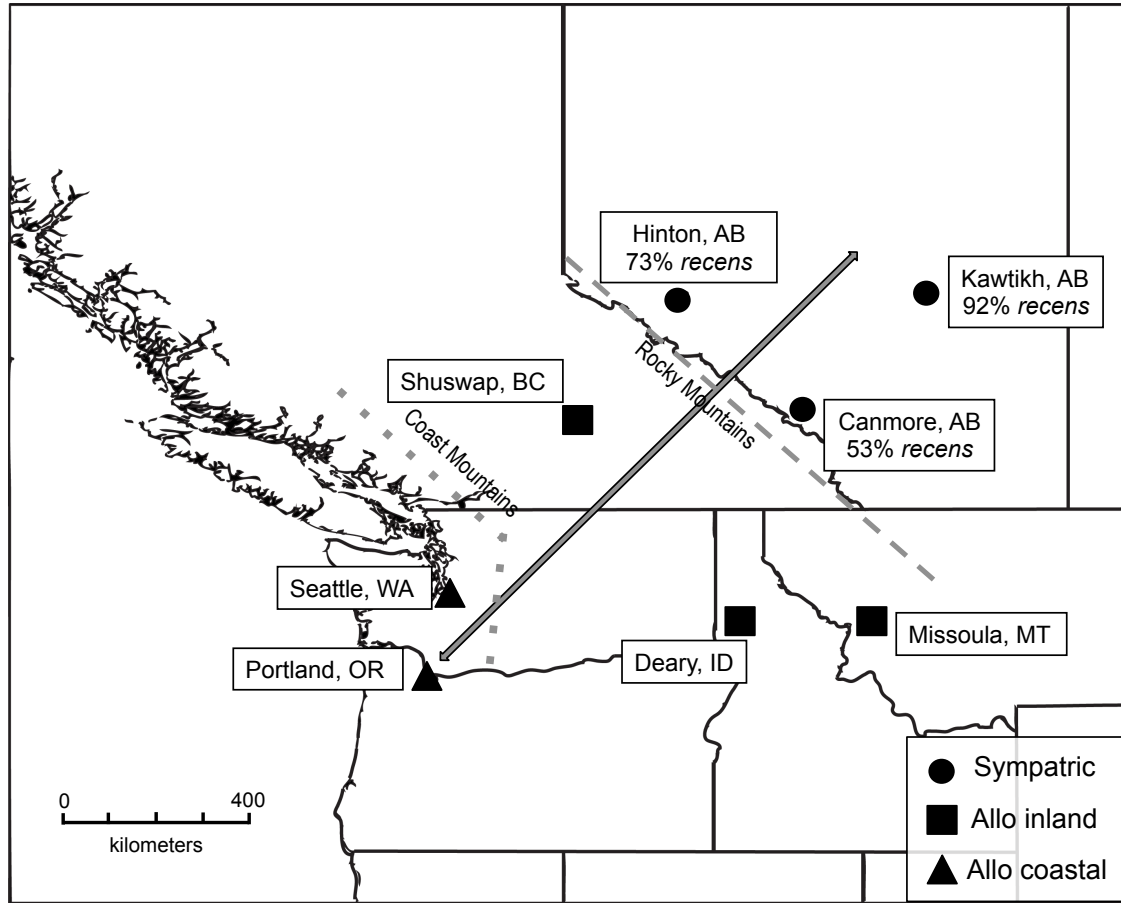


Fig. 2.1. Relative locations of populations used in this study. Sympatric populations are represented by circles, with the percent of each collection that was *D. recens* noted within the label. Inland allopatric and coastal allopatric sites are represented by squares and triangles, respectively. Gray dotted and dashed lines indicate significant population genetic breaks, which correspond with the Coast and Rocky Mountains, respectively. The grey arrow indicates the 50° transect angle used to construct clines.

Fig. 2.2.

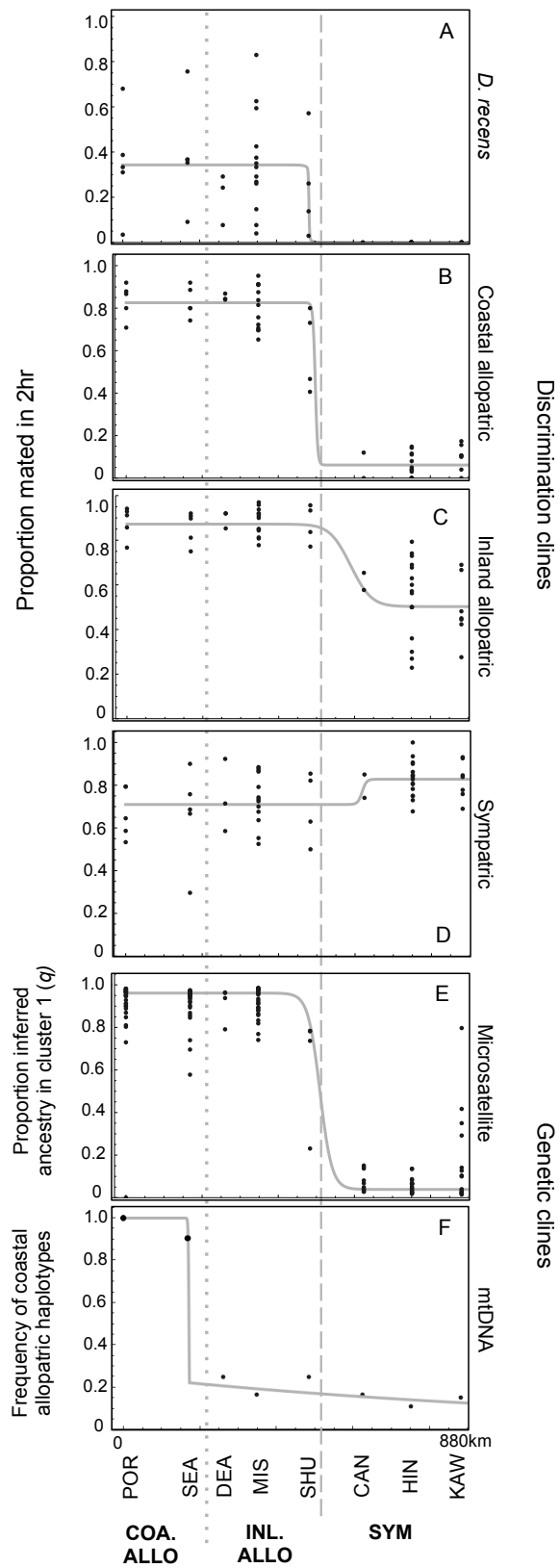


Fig. 2.2. Clines in behavioral discrimination by male type (panels A-D) and in genetic distance (panels E-F). Source populations of isofemale lines used in mate trials are indicated along the *X*-axis, with the abbreviations from Table 2.1 used and their placement proportional to their distance along the cline from west to east. Black dots indicate the mean fraction mated of each isofemale line in the phenotypic clines (A-D), the individual assignment values to cluster 1 in the multilocus microsatellite cline (E), and the proportion of coastal allopatric haplotypes in each population for the mtDNA cline (F). Each cline is drawn from the best fitting model of the phenotypic distribution (*D. recens*: bimodal; allopatric coastal: mean of bimodal and unimodal estimates; allopatric inland: trimodal; sympatric: unimodal; microsatellite: trimodal; mtDNA: simple sigmoid). The dotted and dashed vertical lines indicate the relative locations of the Coast and Rocky Mountains, respectively.

Fig. 2.3.

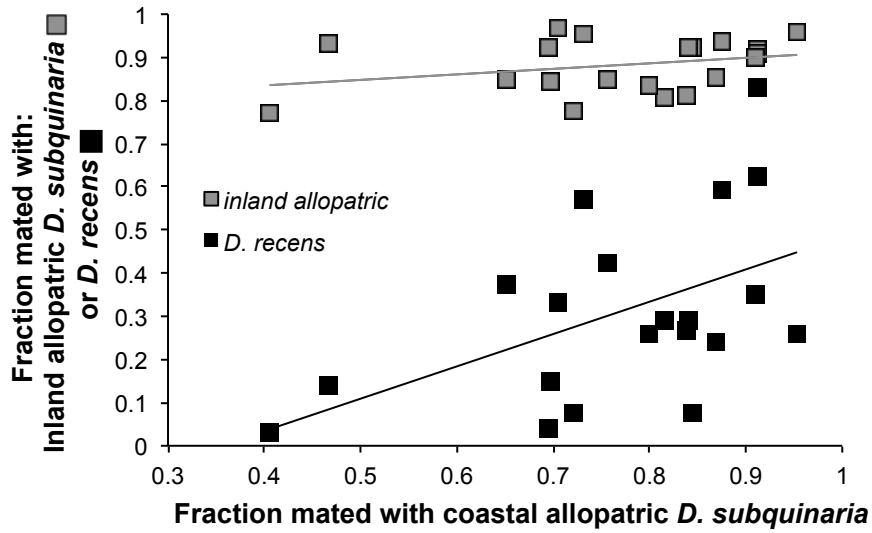


Fig. 2.3. Correlations of mate discrimination between coastal *D. subquinaria* and *D. recens* males (black squares and black line) and between coastal *D. subquinaria* and inland allopatric *D. subquinaria* males (grey squares and grey line). Only isofemale lines from inland allopatric populations (Missoula, Deary, Shuswap) are included. The lines indicate the regressions (black: $r^2 = 0.25$; $F_{1,18} = 6.1$, $P = 0.024$; grey: $r^2 = 0.09$; $F_{1,18} = 1.8$, $P = 0.20$)

CHAPTER 3

INTROGRESSION AND POPULATION GENETIC ANALYSES OF TWO SPECIES
IN SECONDARY CONTACT: IMPLICATIONS FOR REINFORCEMENT AND ITS
CONSEQUENCES

Introduction

When two recently diverged species come into secondary contact, the opportunity for hybridization arises. If reproductive barriers are incomplete, hybrid zones can provide a window into the processes that influence speciation at the earliest stages. Whether or not species remain distinct when they come into secondary contact depends on the extent of gene flow between species and the strength of selection against hybrids (Barton and Hewitt 1985). When hybrids are unfit due to postzygotic incompatibilities that accumulated during divergence, selection can act to strengthen pre-mating barriers and thus reinforce species boundaries (Dobzhansky 1951; Noor 1999). This selection relies on a delicate balance of gene flow between and within species: not enough and selection will not act; too much gene flow will swamp the effects of selection and impede the evolution of increased pre-mating isolation (Felsenstein (1981). We now see that reinforcement can occur with some gene flow between and within species (Noor 1999; Nosil et al. 2003; Borge et al. 2005), as demonstrated empirically by Matute (2010) and theoretically by others (Liou and Price 1994; Servedio and Kirkpatrick 1997). Analyzing gene flow between and within recently diverged species in secondary contact will help us

understand the influence of gene flow on reinforcement in nature. It can also provide insight on how reinforcement proceeded or was inhibited and may point to a role of geography in that process.

Despite strong selection against hybrids or high pre-mating isolation, in many hybrid zones we see that introgression does occur, and the patterns of differential introgression across the genome can point to regions under selection (Saetre et al. 2003; Payseur et al. 2004; Borge et al. 2005). Regions of the genome that harbor reproductively or ecologically isolating factors may be less likely to introgress across a hybrid zone (Barton and Bengtsson 1986; Borge et al. 2005; Maroja et al. 2009; Bimova et al. 2011). The genomes can be viewed as in conflict, i.e., resisting invasion from outside alleles, however, this is not always the case as some introgression may actually be favored (Arnold 1997). In addition, the introgression of autosomal and X-linked genes may differ for reasons related to both selection and effective population size, and patterns across any given chromosome may differ (Payseur et al. 2004). Because of the large X effect and both theoretical and empirical work in reinforcement (Servedio and Noor 2003; Coyne and Orr 2004; Ortiz-Barrientos et al. 2004; Saether et al. 2007; Presgraves 2008), it is likely that X chromosomes may harbor isolating factors and thus X-linked loci are less likely to introgress easily across species boundaries. It is also possible that there are factors on autosomes that interact epistatically with the X that may also resist introgression (Teeter et al. 2008).

The amount of introgression will also be determined by the occurrence of hybridization, and patterns of reproductive isolation are expected to be reflected in the patterns of genetic differentiation between species (e.g., in flycatchers, Borge et al.

2005). For any given system, we can make predictions about the directions and strength of introgression based on what we know about patterns of hybrid fitness and pre-mating isolation. For instance, when hybrid fitness is very low and pre-mating isolation is strong, introgression should also be low. Additionally, when one sex suffers reduced fitness or sterility, introgression between species may be biased in favor of the fertile or fitter sex (Saetre et al. 1999).

Not only can introgression between species affect divergence in a hybrid zone, but so can gene flow among populations within species. Within a species undergoing reinforcement, too much gene flow from allopatric populations into sympatric populations can swamp out the effects of reinforcement such that pre-mating isolation between species does not evolve (Liou and Price 1994; Matute 2010). Both empirical and theoretical work suggests that in order for the process of reinforcement to begin, gene flow cannot be too high between or within species. Because of this swamping effect of gene flow between sympatric and allopatric populations, some models and experiments have addressed the conditions under which reinforcement can proceed under gene flow from allopatric populations (Caisse and Antonovics 1978; Sanderson 1989; Liou and Price 1994; Matute 2010). This suggests that barriers to gene flow between sympatric and allopatric populations may be beneficial to the evolution of reinforcement, and such barriers may occur when populations within species are already fragmented or ecologically divergent (Servedio and Noor 2003).

Reinforcement itself may also generate barriers between populations within a species through the differences created by selection in sympatry (described as “cascade reinforcement” or “reproductive character displacement (RCD) speciation”) (Hoskin et

al. 2005; Ortiz-Barrientos et al. 2009). During this process, females from populations sympatric with the other species no longer recognize conspecific allopatric males as suitable mates, and behavioral isolation among conspecific populations is initiated. It has been suggested that cascade reinforcement may proceed more easily in fragmented populations or between populations where gene flow is already limited (Coyne and Orr 2004; Hoskin and Higgie 2010). However it may be difficult to determine if gene flow has slowed as a result of reinforcing selection, in conjunction with it, or if these populations were already distinct before reinforcement. Ultimately, in order to identify what forces drive divergence with gene flow, gene flow among species and among populations within species must be examined, and we must compare conditions found in nature to the theoretical and experimental expectations surrounding speciation.

A good model to approach these questions are the closely related species *Drosophila subquinaria* and *D. recens*, which are in secondary contact east of the Rocky Mountains in Canada (Fig. 3.1). These species may have diverged in allopatric refugia during the last glacial maximum, about 85-50 kya, and came back into secondary contact 20-10 kya since the glaciers receded. They can be found feeding and breeding on the same mushrooms, and have no known differences in ecology. Other than male genitalia, they are morphologically identical.

One known difference between these species is that *D. recens* is almost always infected with the endosymbiont *Wolbachia* while *D. subquinaria* is not infected (Werren and Jaenike 1995). *Wolbachia* infections can cause cytoplasmic incompatibility (CI) between sperm and eggs, such that uninfected eggs cannot utilize sperm from infected males (Werren and Jaenike 1995). In this system, when uninfected *D. subquinaria*

females mate with infected *D. recens* males, almost all offspring die as embryos (Shoemaker et al. 1999). On the other hand, a mating between *D. recens* females and *D. subquinaria* males produces viable offspring, but hybrid males are sterile. Consistent with a high cost of heterospecific matings for *D. subquinaria* females, these females from populations sympatric with *D. recens* are incredibly choosy (Jaenike et al. 2006). Furthermore, consistent with predictions of cascade reinforcement, these sympatric *D. subquinaria* also reject allopatric conspecifics (Jaenike et al. 2006). In contrast, despite also suffering a cost of hybridization through sterile male offspring, *D. recens* females from sympatric populations do not show any patterns of reproductive character displacement. Instead, *D. recens* females will mate with *D. subquinaria* at a low rate regardless of their geographic origin, and there is the potential for gene flow between species through these hybrid and fertile daughters.

Previous work in *D. subquinaria* and *D. recens* indicated that gene flow occurs both between species and among sympatric and allopatric populations within each species (Jaenike et al. 2006), Bewick and Dyer). First, about 3% of *D. subquinaria* individuals carry a *D. recens* mtDNA haplotype, suggesting that some hybridization does occur in nature (no *D. recens* individuals carry a *D. subquinaria* haplotype since CI prevents viable offspring between *D. subquinaria* females and *D. recens* males) (Jaenike et al. 2006). In *D. subquinaria*, the mtDNA shows high gene flow among populations sympatric with *D. recens* and neighboring allopatric populations, though there is strong differentiation among allopatric populations found on the Pacific coast of North America and all other inland populations. In contrast, eight autosomal microsatellite markers showed moderate differentiation between sympatric and allopatric populations on either

side of the Rocky Mountains and no differentiation between coastal and inland populations. In *D. recens*, to date only the mtDNA has been surveyed, and this marker showed high gene flow among all populations surveyed, which included populations both sympatric and allopatric with *D. subquinaria* (Jaenike 2006, Shoemaker and Jaenike 1997).

Here we address a set of questions to illuminate whether the patterns of gene flow between these two diverging species as well as among populations within each species reflect patterns of sexual isolation. First, is there evidence that introgression occurs between *D. subquinaria* and *D. recens*, and do different regions of the genome show different patterns of introgression? Does this correspond to what is known about reinforcement between these species? Second, what is the population structure within species, and does it reflect reproductive character displacement? Ultimately we hope to better understand how the process of gene flow may contribute to divergence between and within species.

Materials and Methods

Collections and species identification

Flies were collected by sweep netting over white button mushroom (*Agaricus bisporus*) baits, and collection years span from 2001-2012, with some populations sampled across multiple years (Table 3.1) and including some flies sampled from Jaenike et al. (Jaenike et al. 2006). Sampling sites included both allopatric and sympatric populations of both *D. subquinaria* and *D. recens* (Fig. 3.1). How these flies are grouped, and relevant geographic boundaries are shown in Table 3.1. Wild males were sorted by

potential species and frozen. Wild females were isolated in individual food vials where each female was allowed to oviposit to set up isofemale lines. Food vials are standard *Drosophila* vials with Formula 4-24 Standard *Drosophila* Medium from Carolina Biological, supplemented with *Agaricus bisporus* mushrooms. When F1 larvae were observed, wild females were frozen for DNA extraction. From these lines we established isofemale lines which are maintained at low density at 20°C on a 12:12hr light:dark cycle.

D. subquinaria and *D. recens* are nearly morphologically indistinguishable, thus we used molecular markers to diagnose species of wild flies following the procedure described in Bewick and Dyer. Where possible, we selected at least 25 wild females of each species from each population for genotyping. For some populations, sample sizes of collected flies were small, so we included some wild males, or if isofemale lines had been established but the corresponding wild female sample was lost, we used female flies taken from the lines themselves. A total of 684 individuals were used in this study, which included 354 *D. subquinaria* and 330 *D. recens*.

Microsatellite genotyping

To assay population differentiation, we used two sets of microsatellites. The first was developed from *D. recens* (Ross et al. 2003), and the second set were developed from *D. transversa*, a closely related species from the quinaria group (Raisanen et al 2009). We were able to amplify a subset of each of these published marker sets in both *D. subquinaria* and *D. recens*, and any markers that did not reliably amplify in both species were excluded. Our final dataset included 30 markers; a subset of these samples and

markers were included in Bewick and Dyer (in review). We amplified microsatellites following the protocol of (Miles et al. 2008) using the three-primer fluorescent method described in (Shuelke 2000). Amplified products were run alongside a size standard (GGF Rox 500) at the Georgia Genomics Facility, and alleles were scored and confirmed by eye three times using GeneMarker software (SoftGenetics State College, PA). After allele confirmation, we looked for uniform homozygosity across wild males as an indicator of X-linkage. Based on this assessment, nine markers are X-linked and twenty-one are autosomal. In the case of males, we converted all X-linked markers to haploid status by replacing one allele with “0”. In the case of flies taken directly from laboratory lines, we replaced one random allele from each locus (both X-linked and autosomal) with “0” in order to avoid the increase in homozygosity due to inbreeding.

Genetic diversity

We used tests for linkage disequilibrium (LD) in Genepop (citation) and Hardy-Weinberg equilibrium (HWE) as implemented in Arlequin v3.5 (Excoffier and Lischer 2010). We also used Arlequin to calculate population pairwise F_{ST} among populations. For X-linked markers, only female samples were used to determine HWE and LD. We tested for evidence of isolation by distance within each species using a Mantel test using Slatkin’s pairwise F_{ST} (Slatkin 1995) and geographic distance matrix (in km, generated from GPS coordinates using Geographic Distance Matrix Generator, American Museum of Natural History). We used Arlequin to compute allelic richness. We used Microchecker (Van Oosterhout et al. 2004) to identify if null alleles were present.

Analysis of molecular variance (AMOVA)

We explored higher level patterns of population structure using AMOVA as implemented in Arlequin v3.5. We divided *D. subquinaria* populations into groups based on our preliminary results of genetic structure found in Bewick and Dyer. We grouped them by sympatric versus allopatric (two groups), all inland populations versus coastal populations (two groups), sympatric versus inland allopatric versus coastal allopatric (three groups), and east or west of the Rocky Mountains, as loosely defined by the British Columbia-Alberta border (two groups). We also looked for structure between sympatric and allopatric populations of *D. recens*, and between the Great Smoky Mountains and other populations of *D. recens*. To identify if the X-chromosome and the autosomes show different patterns of gene flow between and within these species, we therefore performed each of these analyses on three datasets: a dataset including all markers, a dataset including only autosomal markers, and a dataset including only X-linked markers.

Structure Analysis

We used the program Structure (Pritchard et al. 2000; Hubisz et al. 2009) to infer genetic clusters within and between species. We used an admixture model assuming correlated allele frequencies. We used a burnin of 100,000 in 1,000,000 MCMC steps. Each run was repeated ten times for each possible K . We compared average $\ln L$ s of putative K s and used the ΔK method described in Evanno et al (2005) to identify best K s, and used the “greedy” algorithm in CLUMPP (Jakobsson and Rosenberg 2007) to summarize q -values across runs. We used Structure to infer clusters in four scenarios: 1) using all samples of both *D. subquinaria* and *D. recens*, 2) using all samples of *D.*

subquinaria, 3) using only allopatric samples of *D. subquinaria*, and 4) using all samples of *D. recens*. We used again used the three separate datasets (all markers, auto only, and X only) to ask whether the X versus the autosomes exhibit different patterns of introgression. For Structure results investigating introgression between *D. subquinaria* and *D. recens* we used one-way ANOVA to test whether individual q-values (an individual's proportion of ancestry from each of the *K* clusters) indicated equal or unequal admixture between species. Figures were rendered using the program Distruct (Rosenberg et al. 2002).

Results

Gene diversity, HWE, and LD

The average number of alleles in each population are shown in Table 3.2; within *D. subquinaria* average number of alleles per marker is 30.17 (range: 8 to 76) and within *D. recens* the average number of alleles per locus is 32.2 (range 7 to 75). The observed and expected heterozygosity for each locus and population are shown in Table 3.3. After correction only two autosomal loci (3091 and 3047) were found to be out of Hardy-Weinberg equilibrium in a majority of populations (Table 3.4). AMOVA results were unchanged when these markers were excluded; however Structure analyses, which are sensitive to violations of HWE, did change within species, thus we report only results with these markers excluded. Additionally, Microchecker results indicated that null alleles for many of the loci may exist in some populations, however patterns across loci and populations were not consistent. *D. recens* and *D. subquinaria* are hyper-diverse species (Ross et al. 2003; Shoemaker et al. 2004; Cutter et al. 2013) and thus have large

effective population sizes, so null alleles are likely unavoidable in microsatellite markers in these species. After correction for multiple tests, there were no consistent patterns in LD between loci across populations.

Genetic structure and introgression between species

When all markers are considered, cluster assignment through Structure analyses including both species support $K = 2$ as the main genetic groups ($\Delta K = 7021.24$), with each group corresponding to either *D. subquinaria* or *D. recens* (Table 3.5, Fig. 3.2). $\ln P(K)$ values support $K = 11$, but even at this high number of clusters, the two major groups of *D. subquinaria* and *D. recens* remain distinct. This result of ΔK support for two clusters that are maintained at higher K s held when the complete dataset was broken into autosomal and X-linked markers (autosomal: $\Delta K(2) = 11683.76$, $\ln P(K = 11) = -5262.22$; X-linked: $\Delta K(2) = 8567.42$, $\ln P(K = 19) = -19901.28$) (Figs. 3.3 and 3.4). Based on inspection of individual q-values, it appears that introgression is biased from *D. recens* into *D. subquinaria* than vice versa in the complete and autosomal datasets, though it is less clear in the X-linked dataset (Figs. 3.2 – 3.4). One-way ANOVA of q-values shows that introgression at autosomal markers is significantly biased in the direction of *D. recens* into *D. subquinaria* ($F = 17.09$, $P < 0.0001$), while introgression is unbiased in X-linked markers ($F = 1.113$, $P = 0.918$). This may indicate that selection acts against introgression at X-linked markers, while autosomal markers are more permeable to introgression. Additionally, the biased introgression from *D. recens* into *D. subquinaria* at autosomal markers is significantly higher in sympatric populations compared to allopatric populations ($F = 31.86$, $P < 0.0001$).

Genetic differentiation within *D. recens*

Overall, it appears that there is little, if any, genetic structure within *D. recens*. In *D. recens* there was no evidence of isolation by distance based on Mantel tests of pairwise geographic distance and Slatkin's F_{ST} ($r = 0.084$, $P = 0.184$). With the complete dataset, Structure analysis supports $K = 1$ ($\ln P(K) = -39114.26$) and $K = 4$ ($\Delta K = 22.91$) (Table 3.6), and the four genetic clusters do not appear to correspond to either geography or allopatry/sympatry boundaries (Fig. 3.5A). K s of two and three also find some support with high ΔK values, and it is worth noting that if $K = 1$ is the true K , then we would not detect it with ΔK . F_{ST} measures with all loci also indicate no genetic structure that corresponds with allopatry/sympatry or known geographic barriers (Table 3.7). However, AMOVA results indicate that there is some genetic structure between populations sympatric with *D. subquinaria* and those that are allopatric (1.36% of variation is explained by this grouping ($\Phi_{CT} = 0.014$, $P = 0.0015$))(Table 3.8). Twice as much variation (2.55%) is explained when populations are grouped into the Great Smoky Mountains versus all others ($\Phi_{CT} = 0.0255$, $P = 0.1402$) (Table 3.8).

Only considering autosomal markers, structure analysis again supports $K = 1$ ($\ln P(K) = -30622.03$) and ΔK method suggests $K = 5$ ($\Delta K = 152.06$), and patterns also appear to be unrelated to sympatry/allopatry with *D. subquinaria* or known geographic barriers (Fig. 3.5B). Although ΔK suggests $K = 5$, it is worth noting again that we cannot use this method to assess support for $K = 1$. The distribution of clusters at $K = 5$ corresponds to other work in *D. recens* mtDNA which indicates some haplotype differentiation between populations in the Great Smoky Mountains (GSM) and those distributed in the north and west of North America (Jaenike et al. 2006), which could

result if GSM was an ancestral population or glacial refugia. Interestingly, when only X-linked markers are analyzed, $K = 4$ is preferred (ΔK), and there appears to be some distinction between sympatric and allopatric populations (Fig. 3.5C). In AMOVA, more variation is explained by the sympatry/allopatry grouping in X-linked compared to autosomal markers (for autosomal markers: $\Phi_{CT} = 0.0121$ ($P = 0.0002$); for X-linked markers: $\Phi_{CT} = 0.02045$ ($P = 0.0696$), Table 3.8). Overall, AMOVAs that consider GSM as a separate group explain more variation in each of these datasets; however, none of these phi-statistics are significant (Table 3.8).

Genetic differentiation within *D. subquinaria*

There was no evidence of isolation by distance based on Mantel tests of population pairwise geographic distance and Slatkin's F_{ST} ($r = 0.082$, $P = 0.367$). Within *D. subquinaria*, there is major genetic structure between populations that are sympatric versus allopatric with *D. recens*. When all markers are considered in a Structure analysis, the ΔK method supports two genetic clusters ($\Delta K = 130.17$) and these clusters remain somewhat distinct at $K = 4$ supported by $\ln P(K) = -37562.08$ (Table 3.9, Fig. 3.6A). This is supported by measures of F_{ST} , which shows strong differentiation between allopatric and sympatric populations (Table 3.10). The two main clusters correspond roughly to the populations on the west and east of the Rocky Mountains, though there is also a major difference in the proportion of sister species between these populations on either side of the range.

Analysis with autosomal markers alone suggests three or four major genetic clusters ($\Delta K(3) = 8.53$, $\ln P(K = 4) = -25319.87$) (Table 3.9, Fig. 3.6B). Inspection of the

results indicates that populations on the east side of the Rockies and in contact with high proportions of with *D. recens* remain distinct from populations on the west side at $K > 2$. Analysis with only X-linked markers suggests five or seven major clusters ($\Delta K(5) = 130.02$, $\ln P(K = 7) = -12022.12$, Table 3.9, Fig. 3.6C). Again, inspection of plots of q values indicates that populations on either side of the Rockies compose the major groups even at these higher K values (Fig. 3.6C). Interestingly, it appears that there may be some differentiation between coastal allopatric populations and inland populations, as is seen in the mtDNA (Fig. 3.6B and 3.6C). A structure analysis of populations found only in the allopatric/west genetic cluster showed some differentiation of coastal populations from inland populations in X-linked markers (Fig. 3.7).

Additional investigation through AMOVA and F_{ST} clarifies the genetic relationship between allopatric and sympatric populations. When we apply a strict definition of sympatry (i.e. whether any *D. recens* individuals were identified in *D. subquinaria* populations), then we find sympatric populations on both the western and eastern side of the Rocky Mountains (Table 3.11). In an AMOVA based on two groups (sympatric versus allopatric) using the complete dataset, 1.62% of variation is explained by this partition ($\Phi_{CT} = 0.0162$, $P = 0.0141$, Table 3.11). However, partitioning groups based on lying east or west of the Rocky Mountain barrier (roughly defined by the British Columbia/Alberta provincial line) explained three times the genetic variation (3.54%, $\Phi_{CT} = 0.0354$, $P < 0.0001$, Table 3.11). The increase in variation explained by this partition holds when the dataset is broken into autosomal and X-linked loci (Table 3.11). This indicates that the ecological/geographic barrier of the Rocky Mountains may be playing a strong role in the genetic differentiation of *D. subquinaria*.

Discussion

When two recently diverged species come into secondary contact, there is a new opportunity for gene flow in this hybrid zone. The balance between selection against hybrids and gene flow between species in this zone may determine if divergence continues, stalls, or if the two species fuse. Reinforcement of pre-mating isolation cannot proceed with too much gene flow between species or between sympatric and allopatric populations within species, and by studying gene flow between species in these hybrid zones, we can identify how gene flow may have hindered or promoted the evolution of pre-mating isolation. Here we dissect population structure and gene flow within and between two species in secondary contact. We find that introgression is occurring occasionally between *D. subquinaria* and *D. recens*, and that introgression at autosomal markers is biased into *D. subquinaria*, but unbiased at X-linked markers. Additionally, we see that there is significant structure among *D. subquinaria* populations, while no strong patterns exist within *D. recens*. Selection and gene flow may have interacted in such a way to facilitate reinforcement in *D. subquinaria* while hindering it in *D. recens*.

Introgression between species

Our structure results indicate two major genetic groups that correspond to each species (Figs. 3.2 – 3.4). These groups are upheld whether we examine only autosomal loci, X-linked loci, or both. Many more *D. subquinaria* individuals have a high proportion of admixture from the *D. recens* genetic cluster than individuals of *D. recens* have from the *D. subquinaria* cluster ($F=17.09$, $P < 0.0001$). The potential reasons for

this asymmetry in introgression become clearer in the context of cytoplasmic incompatibility, the potential directions of backcrossing, and female preferences. First, because of *Wolbachia* infection in *D. recens* almost all offspring of *D. recens* males who mate with *D. subquinaria* females die as embryos, resulting in a genetic dead end (Werren and Jaenike 1995; Shoemaker et al. 1999; Jaenike et al. 2006). On the other hand, *D. recens* females produce viable offspring with *D. subquinaria* males, however their male offspring are sterile (Shoemaker et al. 1999). This leaves only female offspring of this cross as a mechanism of introgression. If female discrimination of hybrids against *D. recens* is partially dominant, as unpublished data suggests (K. Dyer, unpublished data) and as has been seen in another *Drosophila* reinforcement system (Ortiz-Barrientos et al. 2004), then sympatric females would be expected to preferentially backcross with *D. subquinaria* males. Thus, even in the event of hybridization between species, hybrid females will backcross to *D. subquinaria* most often, resulting in more introgression from *D. recens* to *D. subquinaria* than the other way around. This is similar to what is seen in flycatchers, where introgression in hybrid zones is much higher from pied to collared than collared to pied. In this case as well, biased introgression could be related to how hybrids mate as well as whether gene flow from allopatric populations is high (Borge et al. 2005). What this also suggests is that selection on reinforcement in *D. subquinaria* must be very high such that it can occur despite some gene flow from its sister species (Matute 2010)

Gene flow within species

Structure results suggest that there is little if any population structure between sympatric and allopatric populations of *D. recens*. AMOVA results suggest that there may be some differentiation between sympatric/allopatric populations, but taken together with the results of Structure and F_{ST} analysis, this division is inconsistent and probably weak. This result corresponds with the lack of reproductive character displacement seen in *D. recens*. In the context of models of reinforcement, we expect that too much gene flow among allopatric and sympatric populations will swamp the effects of reinforcement selection (Liou and Price 1994; Matute 2010). This may be the case in *D. recens*. Despite the cost of hybrid male sterility, which is sufficient to drive reinforcement in other systems (Noor 1995; Kelly and Noor 1996; Noor 1999), *D. recens* females show no reproductive character displacement. Given our results, the lack of reinforcement within *D. recens* may be explained by a lower cost to hybridization (compared to that for *D. subquinaria* females) and swamping gene flow from allopatric populations.

Patterns of gene flow within *D. subquinaria* also correspond with the patterns of sexual isolation seen within species. Sympatric populations on the east of the Rocky Mountains are differentiated at both autosomal and X-linked markers from populations on the west of the mountain range. One possible reason for this differentiation may be that introgression from *D. recens* is changing allele frequencies in sympatric populations. Populations located on the eastern part of the range have a much higher percentage of *D. recens* in their community, and are likely to experience higher encounter rates and higher opportunities for hybridization, but will also have higher selection to avoid hybridization and fewer migrants from allopatric populations, so the source of this structure is

uncertain. Like introgression with *D. recens*, gene flow between sympatric populations on the east of the mountain range and populations on the west appears to be asymmetric (Fig. 3.6) with more allopatric contributions to sympatric genomes than the other way around. This may be due to a lack of migration from the eastern side of the mountain to the west, or sympatric alleles may carry a cost in allopatry. Additionally, there appears to be some structure in the X chromosome between allopatric populations to the east and west of the Coast Mountains that is not seen in the autosomes (Fig. 3.6C and 3.7C). This is the same pattern seen among mtDNA haplotypes in previous work (Jaenike et al. 2006; Bewick and Dyer in review). While it is speculative, this result could be the result of some unknown X-mtDNA epistatic interaction (Montooth et al. 2010) or ecological adaptations of these populations.

The role of the Rocky Mountains in inhibiting gene flow

While it appears that a substantial amount of the genetic structure within *D. subquinaria* is explained by the Rocky Mountains, the true nature of the differentiation may be more complex. If we use a strict definition of sympatry as any population of *D. subquinaria* with at least one *D. recens* individual, then there are two sympatric populations on the west side of the mountains (Revelstoke and Golden, BC) that each have a very low prevalence of *D. recens*. We note that because these populations have been sampled only once we do not know whether they consistently have a low prevalence of *D. recens*, or whether these are recent migrants from populations where *D. recens* is more common. *D. subquinaria* from these populations group more closely with allopatric populations, and preliminary data suggests that these females do not exhibit high mate

discrimination against *D. recens* like they do in sympatric populations where *D. recens* is at a higher prevalence (K. Dyer, unpublished data).

The sympatric population at Lake Louise, Alberta, also represents an interesting case. If we define the Rocky Mountains based on the British Columbia/Alberta boundary where the mountains are at their highest, then this population falls on the east side of the mountains. However, it is clearly more genetically similar to those on the west (Fig. 3.6). Additionally, the percentage of *D. recens* in this population is quite low compared to other sympatric populations on the east of the mountains (Table 3.1). This population is located at high elevation in the range, but it is only 94 km away from the sympatric population Canmore, AB, and it also is upstream along the same river valley. It is possible that elevation is a barrier not only to the migration of *D. recens* but also to the migration of *D. subquinaria*. It is difficult to determine whether the boundary of the mountains or a certain percentage of *D. recens* (and thus certain strength of selection/opportunity for introgression of *D. recens* alleles) is what is driving patterns of genetic differentiation. Perhaps reinforcing selection works in concert with ecological barriers to result in stronger reproductive isolation and genetic differentiation among these groups (i.e. they reinforce each other) (Abbott et al. 2013). This could be evidence of coupling of intrinsic and extrinsic barriers around the Rocky Mountain Range (Barton and de Cara 2009; Hoskin and Higgie 2010; Bierne et al. 2011; Abbott et al. 2013). These results emphasize the importance of considering the potential role of ecology and geography in facilitating both reinforcement and divergence within species (Ortiz-Barrientos et al. 2009; Hoskin and Higgie 2010), and brings up the question of whether

cascade reinforcement could proceed at all with free gene flow between populations, if it can hinder the evolution of reinforcement itself.

In summary, this system has given us an opportunity to investigate how gene flow may promote or hinder the evolution of pre-mating isolation both between and within species. Given theoretical and empirical work on reinforcement under gene flow (Liou and Price 1994; Servedio and Kirkpatrick 1997; Matute 2010), it seems likely that differences in conditions of gene flow have shaped the outcome of secondary contact in this system. For instance, *D. recens* does not have the geographic barrier that *D. subquinaria* has, and shows much less genetic structure within species, as well as no reproductive character displacement (Jaenike et al. 2006). The lack of reinforcement in sympatric *D. recens* populations could be a result of high gene flow between sympatric and allopatric species, coupled with a lower cost to hybridization for these females compared to *D. subquinaria* females. On the other hand, the introgression of alleles from *D. recens* into *D. subquinaria* suggests that reinforcing selection must be very strong in order for pre-mating isolation to arise and persist in the face of this gene flow. Additionally, the Rocky Mountains may be acting as a barrier to gene flow within *D. subquinaria*, thus ameliorating the swamping effect of gene flow from allopatric populations. This reduced gene flow could also be facilitating cascade reinforcement within *D. subquinaria*. Thus in the same secondary contact event, one species undergoes too much gene flow and lower selection that may have inhibited reinforcement, yet in the sister species, strong selection in the face of gene flow, coupled with population fragmentation, facilitated reinforcement and may have promoted a secondary cascade isolation event.

Literature Cited

- Abbott, R., D. Albach, S. Ansell, J. W. Arntzen, S. J. E. Baird, N. Bierne, J. W. Boughman, A. Brelsford, C. A. Buerkle, R. Buggs, R. K. Butlin, U. Dieckmann, F. Eroukhmanoff, A. Grill, S. H. Cahan, J. S. Hermansen, G. Hewitt, A. G. Hudson, C. Jiggins, J. Jones, B. Keller, T. Marczewski, J. Mallet, P. Martinez-Rodriguez, M. Most, S. Mullen, R. Nichols, A. W. Nolte, C. Parisod, K. Pfennig, A. M. Rice, M. G. Ritchie, B. Seifert, C. M. Smadja, R. Stelkens, J. M. Szymura, R. Vainola, J. B. W. Wolf, and D. Zinner. 2013. Hybridization and speciation. *J. Evol. Biol.* 26:229-246.
- Arnold, M. L. 1997. Natural hybridization and evolution. Oxford University Press, New York.
- Barton, N. and B. O. Bengtsson. 1986. The barrier to genetic exchange between hybridizing populations. *Heredity* 57:357-376.
- Barton, N. H. and M. A. R. de Cara. 2009. The evolution of strong reproductive isolation. *Evolution* 63:1171-1190.
- Barton, N. H. and G. M. Hewitt. 1985. Analysis of hybrid zones. *Annu. Rev. Ecol. S.* 16:113-148.
- Bewick, E. R. and K. A. Dyer. in review. Reinforcement shapes clines in female mate discrimination in *D. subquinaria*. *Evolution*.
- Bierne, N., J. Welch, E. Loire, F. Bonhomme, and P. David. 2011. The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Mol. Ecol.* 20:2044-2072.

- Bimova, B. V., M. Macholan, S. J. E. Baird, P. Munclinger, P. Dufkova, C. M. Laukaitis, R. C. Karn, K. Luzynski, P. K. Tucker, and J. Pialek. 2011. Reinforcement selection acting on the European house mouse hybrid zone. *Mol. Ecol.* 20:2403-2424.
- Borge, T., M. T. Webster, G. Andersson, and G. P. Saetre. 2005. Contrasting patterns of polymorphism and divergence on the Z chromosome and autosomes in two *Ficedula* flycatcher species. *Genetics* 171:1861-1873.
- Caisse, M. and J. Antonovics. 1978. Evolution in closely adjacent plant populations .IX. Evolution of reproductive isolation in clinal populations. *Heredity* 40:371-384.
- Coyne, J. A. and H. A. Orr. 2004. Speciation. Sinauer Associates, Sunderland (Massachusetts).
- Cutter, A. D., R. Jovelin, and A. Dey. 2013. Molecular hyperdiversity and evolution in very large populations. *Mol Ecol* 22:2074-2095.
- Dobzhansky, T. 1951. Genetics and the origin of species. Columbia Univ. Press, New York.
- Excoffier, L. and H. E. L. Lischer. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Res.* 10:564-567.
- Felsenstein, J. 1981. Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution* 35:124-138.
- Hoskin, C. J. and M. Higgie. 2010. Speciation via species interactions: the divergence of mating traits within species. *Ecol. Lett.* 13:409-420.

- Hoskin, C. J., M. Higgie, K. R. McDonald, and C. Moritz. 2005. Reinforcement drives rapid allopatric speciation. *Nature* 437:1353-1356.
- Hubisz, M. J., D. Falush, M. Stephens, and J. K. Pritchard. 2009. Inferring weak population structure with the assistance of sample group information. *Mol. Ecol. Res.* 9:1322-1332.
- Jaenike, J., K. A. Dyer, C. Cornish, and M. S. Minhas. 2006. Asymmetrical reinforcement and *Wolbachia* infection in *Drosophila*. *PLoS Biol.* 4:1852-1862.
- Jakobsson, M. and N. A. Rosenberg. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23:1801-1806.
- Kelly, J. K. and M. A. F. Noor. 1996. Speciation by reinforcement: A model derived from studies of *Drosophila*. *Genetics* 143:1485-1497.
- Liou, L. W. and T. D. Price. 1994. Speciation by reinforcement of premating isolation. *Evolution* 48:1451-1459.
- Maroja, L. S., J. A. Andres, and R. G. Harrison. 2009. Genealogical discordance and patterns of introgression and selection across a cricket hybrid zone. *Evolution* 63:2999-3015.
- Matute, D. R. 2010. Reinforcement can overcome gene flow during speciation in *Drosophila*. *Curr. Biol.* 20:2229-2233.
- Miles, L. G., S. R. Isberg, C. Moran, C. Hagen, and T. C. Glenn. 2008. 253 Novel polymorphic microsatellites for the saltwater crocodile (*Crocodylus porosus*). *Conservation Genetics* 10:963-980.

- Montooth, K. L., C. D. Meiklejohn, D. N. Abt, and D. M. Rand. 2010. Mitochondrial-nuclear epistasis affects fitness within species but does not contribute to fixed incompatibilities between species of *Drosophila*. *Evolution* 64:3364-3379.
- Noor, M. A. 1995. Reinforcement in speciation. *Trends Ecol. Evol.* 10:492.
- Noor, M. A. F. 1999. Reinforcement and other consequences of sympatry. *Heredity* 83:503-508.
- Nosil, P., B. J. Crespi, and C. P. Sandoval. 2003. Reproductive isolation driven by the combined effects of ecological adaptation and reinforcement. *P. Roy. Soc. B-Biol. Sci.* 270:1911-1918.
- Ortiz-Barrientos, D., B. A. Counterman, and M. A. F. Noor. 2004. The genetics of speciation by reinforcement. *PLoS Biol.* 2:2256-2263.
- Ortiz-Barrientos, D., A. Greal, and P. Nosil. 2009. The genetics and ecology of reinforcement: implications for the evolution of prezygotic isolation in sympatry and beyond. Pp. 156-182 in C. D. Schlichting, and T. A. Mousseau, eds. *Year in Evolutionary Biology 2009*. Wiley-Blackwell, Malden.
- Payseur, B. A., J. G. Krenz, and M. W. Nachman. 2004. Differential patterns of introgression across the X chromosome in a hybrid zone between two species of house mice. *Evolution* 58:2064-2078.
- Presgraves, D. C. 2008. Sex chromosomes and speciation in *Drosophila*. *Trends Genet.* 24:336-343.
- Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945-959.

- Rosenberg, N. A., J. K. Pritchard, J. L. Weber, H. M. Cann, K. K. Kidd, L. A. Zhivotovsky, and M. W. Feldman. 2002. Genetic structure of human populations. *Science* 298:2381-2385.
- Ross, C. L., K. A. Dyer, T. Erez, S. J. Miller, J. Jaenike, and T. A. Markow. 2003. Rapid divergence of microsatellite abundance among species of *Drosophila*. *Mol. Biol. Evol.* 20:1143-1157.
- Saether, S. A., G. P. Saetre, T. Borge, C. Wiley, N. Svedin, G. Andersson, T. Veen, J. Haavie, M. R. Servedio, S. Bures, M. Kral, M. B. Hjernquist, L. Gustafsson, J. Traff, and A. Qvarnstrom. 2007. Sex chromosome-linked species recognition and evolution of reproductive isolation in flycatchers. *Science* 318:95-97.
- Saetre, G. P., T. Borge, K. Lindroos, J. Haavie, B. C. Sheldon, C. Primmer, and A. C. Syvanen. 2003. Sex chromosome evolution and speciation in *Ficedula* flycatchers. *P. Roy. Soc. B-Biol. Sci.* 270:53-59.
- Saetre, G. P., M. Kral, S. Bures, and R. A. Ims. 1999. Dynamics of a clinal hybrid zone and a comparison with island hybrid zones of flycatchers (*Ficedula hypoleuca* and *F. albicollis*). *J. Zool.* 247:53-64.
- Sanderson, N. 1989. Can gene flow prevent reinforcement? *Evolution* 43:1223-1235.
- Servedio, M. R. and M. Kirkpatrick. 1997. The effects of gene flow on reinforcement. *Evolution* 51:1764-1772.
- Servedio, M. R. and M. A. F. Noor. 2003. The role of reinforcement in speciation: theory and data. *Annu. Rev. Ecol. Evol. S.* 34:339-364.

- Shoemaker, D. D., K. A. Dyer, M. Ahrens, K. McAbee, and J. Jaenike. 2004. Decreased diversity but increased substitution rate in host mtDNA as a consequence of *Wolbachia* endosymbiont infection. *Genetics* 168:2049-2058.
- Shoemaker, D. D., V. Katju, and J. Jaenike. 1999. *Wolbachia* and the evolution of reproductive isolation between *Drosophila recens* and *Drosophila subquinaria*. *Evolution* 53:1157-1164.
- Shuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18:233-234.
- Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139:1463-1463.
- Teeter, K. C., B. A. Payseur, L. W. Harris, M. A. Bakewell, L. M. Thibodeau, J. E. O'Brien, J. G. Krenz, M. A. Sans-Fuentes, M. W. Nachman, and P. K. Tucker. 2008. Genome-wide patterns of gene flow across a house mouse hybrid zone. *Genome Res.* 18:67-76.
- Van Oosterhout, C., W. F. Hutchinson, D. P. M. Wills, and P. Shipley. 2004. Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* 4:535-538.
- Werren, J. H. and J. Jaenike. 1995. *Wolbachia* and cytoplasmic incompatibility in mycophagous *Drosophila* and their relatives. *Heredity* 75:320-326.

Table 3.1. Sampling locations. Numbers correspond to location identifiers on the map in Figure #. Species, regions within species, and sites are listed. Dotted gray and dashed gray lines indicate boundaries of the Coast and Rocky Mountains respectively. Abbreviations used throughout the paper are also indicated, as well as the year flies were collected from the wild. Asterisks indicate DNA samples taken from flies used in Jaenike et al. (2006). The number of flies used in this study are listed as well as the estimated % of sister species found in sympatric ranges. Estimations were made based on diagnosed wild flies that may include more flies than used in this paper. Only two flies were collected from Nordegg, AB and both were diagnosed as *D. recens*. However, since this population is well inside the range of *D. subquinaria*, it is likely that this population is actually sympatric.

Species	Region	Site	Abbreviation	Year collected	Longitude	Latitude	N flies	% sister species
<i>D. subquinaria</i>	Coastal allopatric	1 Portland, OR	subqPort	2010	-122.6762	45.5235	29	0%
		2 Seattle, WA	subqSea	2010/2011	-122.3321	47.6062	28	0%
	Inland allopatric	3 Deary, ID	subqDea	2011	-116.5560	46.7993	2	0%
		4 Missoula, MT	subqMiss	2010/2011	-113.9940	46.8722	32	0%
		5 Shuswap, BC	subqShu	2010	-119.7110	50.7857	4	0%
		6 Confluence, BC	subqCon	2012	-116.4209	51.4364	15	0%
		7 Takakkaw Falls, BC	subqTak	2012	-116.4741	51.5001	15	0%
	Sympatric	8 Wapta, BC	subqWap	2012	-116.3500	51.4400	31	0%
		9 Sicamous, BC	subqSic	2012	-118.9805	50.8372	29	0%
		10 Revelstoke, BC	subqRev	2012	-118.1957	50.9981	24	4%
		11 Golden, BC	subqGold	2012	-116.9631	51.2961	24	13%
		19 Lake Louise, AB	subqLL	2012	-116.1773	51.4254	34	3%
		10 Canmore, AB	subqCan	2010	-115.3468	51.0870	10	53%
		11 Hinton, AB	subqHin	2010	-117.5700	53.4065	19	73%
		12 Jasper, AB*	subqJas	2002	-118.1027	52.8702	31	79%
		13 Prince Albert, AB*	subqPA	2002	-113.4909	53.5444	4	87%
		14 Winston Churchill Park, AB*	subqWC	2002	-112.1067	54.7199	2	44%

Table 3.1. Continued.

		15	Calgary, AB	subqCal	2012	-114.0581	51.0453	3	81%
		16	Kawtikh, AB	subqKaw	2010	-112.9326	53.4247	18	92%
							total	354	
<i>D. recens</i>	Sympatric	10	Revelstoke, BC	recRev	2012	-118.1957	50.9981	1	96%
		11	Golden, BC	recGold	2012	-116.9631	51.2961	4	87%
		12	Lake Louise, AB	recLL	2012	-116.1773	51.4254	1	97%
		13	Canmore, AB	recCan	2010	-115.3468	51.0870	10	47%
		14	Hinton, AB	recHin	2010	-117.5700	53.4065	25	27%
		15	Jasper, AB*	recJas	2002	-118.1027	52.8702	8	21%
		26	Nordeg, AB	recNor	2010	-116.0815	52.4680	2	-
		16	Prince Albert, AB*	recPA	2002	-113.4909	53.5444	31	13%
		17	Winston Churchill Park, AB*	recWC	2002	-112.1067	54.7199	7	56%
		18	Calgary, AB	recCal	2012	-114.0581	51.0453	26	9%
		19	Kawtikh, AB	recKaw	2010	-112.9326	53.4247	25	8%
		20	Edmonton, AB*	recEd	2002	-113.4909	53.5444	11	25%
	Allopatric	21	Minot, ND*	recMin	2002	-101.2963	48.2325	7	0%
		22	Bemidji, MN*	recBem	2002	-94.8803	47.4736	7	0%
		23	Munising, MI*	recMun	2002	-86.6479	46.4111	7	0%
		24	Great Smoky Mountains, TN*	recGSM	2001	-83.5497	35.6119	36	0%
		25	Samuel de Champlain Provincial Park, ON*	recSdC	2004	-78.8741	46.2931	33	0%
		26	Rochester, NY*	recRoc	2002	-77.6109	43.161	23	0%
		27	Peru, NY	recPeru		-73.5268	44.5784	24	0%
		28	Bethlehem, NH*	recNH	2002	-71.6875	44.2804	7	0%
		30	Charlottetown, PEI*	recPEI	2004	-63.1311	46.2382	4	0%
		34	Brunswick Pines, ME*	recME	2004	-69.9577	43.9092	31	0%
							total	330	0%

Table 3.2. A, number of alleles, is listed for each of the populations. Mean and standard deviation within populations are summarized.

Locus	Populations												
	subqPort	subqSea	subqDea	subqMiss	subqShu	subqCon	subqTak	subqWap	subqSic	subqCan	subqHin	subqJas	subqPA
<i>Drec1007</i>	5	6	-	5	4	6	5	6	6	7	7	8	4
<i>Drec1008</i>	28	18	2	19	8	11	14	23	27	9	14	21	3
<i>Drec1024</i>	11	16	2	12	5	11	11	17	17	14	16	23	8
<i>Drec1079</i>	7	6	-	7	4	5	6	6	5	10	5	8	4
<i>Drec1081</i>	6	6	2	8	3	8	9	7	8	5	15	12	2
<i>Drec1083</i>	6	6	-	10	3	6	5	11	9	2	6	10	3
<i>Drec2002</i>	17	18	2	16	6	10	11	13	21	9	17	22	4
<i>Drec2004</i>	5	12	-	15	3	4	9	11	9	9	8	12	-
<i>Drec2084</i>	9	8	2	5	5	4	7	9	11	4	8	20	-
<i>Drec2085</i>	17	16	2	14	6	9	8	16	18	10	14	21	2
<i>Drec3001</i>	7	8	-	9	4	5	5	9	7	9	10	13	5
<i>Drec3028</i>	3	2	-	3	3	5	5	4	5	7	7	7	3
<i>Drec3038</i>	17	18	2	17	7	11	15	15	17	13	15	22	5
<i>Drec3047</i>	22	11	2	10	3	11	9	12	19	9	12	20	5
<i>Drec3089</i>	4	4	2	5	3	4	4	8	8	10	12	10	6
<i>Drec3090</i>	13	10	2	11	5	10	10	10	13	5	5	9	3
<i>Drec3091</i>	11	9	-	9	4	7	6	11	7	7	7	8	4
<i>Drec4069</i>	3	5	-	2	2	4	4	7	8	6	7	8	2
<i>Dtra31</i>	10	11	-	14	4	11	12	16	14	8	15	19	6
<i>Dtra40</i>	12	18	2	14	5	11	15	14	15	5	10	17	5
<i>Dtra86</i>	4	4	-	10	6	4	4	4	5	5	7	6	4
Mean	10.333	10.095	2	10.238	4.429	7.476	8.286	10.905	11.857	7.762	10.333	14.095	4.105
Std.Dev.	6.703	5.328	0	4.721	1.535	2.943	3.649	4.784	6.142	2.931	3.979	6.066	1.56

Table 3.2. Continued.

Locus	Populations													
	subqWC	subqCal	subqKaw	subqRev	subqGold	subqLL	recJas	recEd	recWC	recCan	recHin	recKaw	recNor	recPA
<i>Drec1007</i>	4	3	8	5	6	8	8	6	5	5	11	8	3	10
<i>Drec1008</i>	3	4	19	23	21	27	9	14	8	13	26	21	3	23
<i>Drec1024</i>	4	4	16	19	15	22	13	15	8	14	22	19	-	21
<i>Drec1079</i>	2	3	6	6	7	10	3	7	4	7	10	12	3	8
<i>Drec1081</i>	-	3	13	10	7	8	6	9	9	13	18	17	2	5
<i>Drec1083</i>	-	-	9	7	7	9	7	13	6	4	10	13	2	16
<i>Drec2002</i>	3	5	18	19	18	19	9	13	11	10	25	22	3	21
<i>Drec2004</i>	3	4	10	9	11	14	11	12	11	11	10	16	-	4
<i>Drec2084</i>	4	6	15	8	10	12	9	4	3	8	8	15	3	7
<i>Drec2085</i>	4	5	11	14	16	22	10	15	13	15	27	23	2	12
<i>Drec3001</i>	3	5	14	5	6	10	10	8	6	11	17	17	4	19
<i>Drec3028</i>	3	3	4	4	4	5	5	5	4	5	6	6	2	7
<i>Drec3038</i>	3	5	13	12	17	16	11	12	10	16	26	19	4	24
<i>Drec3047</i>	2	3	9	12	12	17	13	14	9	13	23	18	2	22
<i>Drec3089</i>	3	3	12	5	6	9	9	13	6	10	14	13	2	16
<i>Drec3090</i>	4	5	8	11	9	10	7	8	7	11	14	14	4	11
<i>Drec3091</i>	3	2	10	9	10	11	5	4	7	8	8	11	3	9
<i>Drec4069</i>	2	-	6	5	6	10	5	8	6	5	8	8	2	10
<i>Dtra31</i>	2	5	13	16	15	19	8	16	10	13	18	14	-	22
<i>Dtra40</i>	3	3	11	12	13	15	6	7	6	10	13	9	3	10
<i>Dtra86</i>	3	3	4	3	7	6	5	8	5	5	6	7	2	8
Mean	3.053	3.895	10.905	10.19	10.619	13.286	7.483	8.966	6.643	9.379	14.467	13.633	2.619	13.034
Std. Dev.	0.705	1.1	4.194	5.528	4.811	5.892	2.836	3.877	2.67	3.812	6.699	5.493	0.74	6.196

Table 3.2. Continued.

	Populations														Mean
	recRev	recGold	recLL	recCal	recPeru	recGSM	recME	recSdC	recBem	recMin	recMun	recPEI	recNH	recRoc	
<i>Drec1007</i>	2	5	-	7	8	8	11	11	6	5	4	3	6	8	6.23
<i>Drec1008</i>	2	6	2	24	24	26	20	23	10	12	11	5	10	18	14.73
<i>Drec1024</i>	2	5	2	24	26	29	22	27	11	11	11	6	8	18	13.93
<i>Drec1079</i>	2	3	-	10	10	13	12	9	7	4	5	4	7	12	6.64
<i>Drec1081</i>	-	7	2	16	19	17	15	15	8	8	8	6	7	13	9.03
<i>Drec1083</i>	2	6	2	15	10	18	10	16	8	7	6	6	4	15	8.03
<i>Drec2002</i>	-	6	2	23	25	21	23	25	10	7	10	5	9	21	13.73
<i>Drec2004</i>	2	6	2	18	17	18	20	19	7	8	6	4	5	16	9.76
<i>Drec2084</i>	2	5	-	18	10	14	16	17	2	5	7	2	2	16	8.21
<i>Drec2085</i>	-	7	-	23	24	24	27	24	11	11	10	6	10	20	13.82
<i>Drec3001</i>	2	3	-	17	17	20	12	17	7	10	6	8	11	13	9.46
<i>Drec3028</i>	2	2	2	4	5	6	7	6	4	4	4	2	4	6	4.38
<i>Drec3038</i>	2	7	2	22	26	29	27	26	8	9	9	6	10	22	13.83
<i>Drec3047</i>	2	5	2	23	20	21	23	27	8	11	9	4	10	21	12.20
<i>Drec3089</i>	2	5	-	14	14	19	17	14	9	9	9	6	12	8	8.48
<i>Drec3090</i>	2	5	2	16	14	16	15	16	9	9	7	5	8	12	8.90
<i>Drec3091</i>	-	5	2	10	12	12	9	8	5	5	6	5	7	10	7.33
<i>Drec4069</i>	2	3	-	12	11	9	8	9	6	7	4	4	4	8	5.95
<i>Dtra31</i>	2	7	2	21	22	12	24	27	7	8	6	3	10	15	12.23
<i>Dtra40</i>	2	4	2	10	8	6	6	9	6	6	7	4	3	8	8.41
<i>Dtra86</i>	2	4	2	9	9	6	7	9	6	6	5	4	6	9	5.48
Mean	2	5.179	2	14.931	14.552	14.655	14.034	14.667	6.586	6.833	6.2	4.069	6.483	12.429	
Std.Dev.	0	1.467	0	6.233	6.727	7.296	6.982	7.604	2.413	2.718	2.483	1.624	2.836	5.385	

Table 3.3. Observed and expected heterozygosity by locus and population. Mean and standard deviation within populations is shown. Bold values indicate significant deviations from Hardy-Weinberg equilibrium after Bonferroni correction ($P < 0.0001$).

Locus	Populations																	
	subqPort		subqSea		subqDea		subqMiss		subqShu		subqCon		subqTak		subqWap		subqSic	
	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E
<i>Drec1007</i>	0.655	0.541	0.560	0.605	-	-	0.519	0.544	0.750	0.821	0.733	0.586	0.714	0.598	0.387	0.527	0.621	0.650
<i>Drec1008</i>	0.692	0.956	0.920	0.894	0.000	1.000	0.481	0.875	1.000	1.000	0.800	0.844	0.929	0.902	0.935	0.940	0.897	0.952
<i>Drec1024</i>	0.821	0.797	0.741	0.877	0.000	1.000	0.741	0.875	0.750	0.786	0.800	0.846	0.714	0.770	0.933	0.890	0.862	0.920
<i>Drec1079</i>	0.321	0.323	0.444	0.502	-	-	0.444	0.534	0.750	0.643	0.286	0.481	0.429	0.571	0.407	0.553	0.448	0.560
<i>Drec1081</i>	0.759	0.811	0.923	0.803	0.000	1.000	0.667	0.794	0.750	0.607	0.867	0.848	0.800	0.828	0.767	0.766	0.690	0.810
<i>Drec1083</i>	0.438	0.677	0.318	0.672	-	-	0.353	0.751	0.750	0.607	0.600	0.687	0.533	0.602	0.567	0.711	0.517	0.730
<i>Drec2002</i>	0.821	0.860	0.769	0.890	0.000	1.000	0.652	0.889	0.750	0.893	0.733	0.855	0.857	0.870	0.700	0.821	0.690	0.903
<i>Drec2004</i>	0.500	0.574	0.857	0.875	-	-	0.500	0.855	0.500	0.833	0.333	0.572	0.538	0.794	0.714	0.769	0.667	0.766
<i>Drec2084</i>	0.679	0.753	0.630	0.770	0.000	1.000	0.400	0.749	0.750	0.786	0.214	0.648	0.308	0.738	0.429	0.782	0.643	0.773
<i>Drec2085</i>	0.750	0.868	0.741	0.857	0.000	1.000	0.667	0.812	1.000	0.893	0.692	0.868	0.786	0.743	0.862	0.852	0.724	0.880
<i>Drec3001</i>	0.852	0.755	0.852	0.806	-	-	0.724	0.792	0.750	0.750	0.600	0.743	0.733	0.692	0.742	0.758	0.828	0.803
<i>Drec3028</i>	0.577	0.446	0.231	0.205	-	-	0.231	0.207	1.000	0.679	0.467	0.405	0.533	0.605	0.581	0.491	0.345	0.496
<i>Drec3038</i>	0.897	0.929	0.741	0.905	0.000	1.000	0.552	0.914	1.000	0.964	1.000	0.880	0.857	0.905	0.828	0.888	0.821	0.888
<i>Drec3047</i>	0.786	0.936	0.630	0.708	0.000	1.000	0.704	0.799	0.000	0.714	0.692	0.782	0.714	0.765	0.467	0.809	0.552	0.850
<i>Drec3089</i>	0.448	0.471	0.519	0.468	0.000	1.000	0.296	0.376	0.250	0.607	0.400	0.395	0.500	0.426	0.387	0.418	0.464	0.673
<i>Drec3090</i>	0.897	0.881	0.885	0.827	0.000	1.000	0.615	0.853	0.750	0.857	0.800	0.880	0.786	0.902	0.833	0.857	0.704	0.815
<i>Drec3091</i>	0.759	0.858	0.800	0.817	-	-	0.556	0.821	0.500	0.750	0.333	0.841	0.333	0.778	0.600	0.838	0.045	0.815
<i>Drec4069</i>	0.286	0.309	0.185	0.238	-	-	0.148	0.195	0.250	0.250	0.154	0.403	0.357	0.323	0.500	0.579	0.207	0.370
<i>Dtra31</i>	0.450	0.845	0.750	0.848	-	-	0.519	0.882	0.500	0.786	0.733	0.828	0.714	0.876	0.733	0.900	0.655	0.868
<i>Dtra40</i>	0.538	0.779	0.542	0.885	0.000	1.000	0.643	0.870	0.500	0.857	0.667	0.869	0.733	0.922	0.839	0.893	0.690	0.888
<i>Dtra86</i>	0.250	0.233	0.381	0.372	-	-	0.630	0.801	0.750	0.893	0.500	0.458	0.467	0.480	0.533	0.455	0.552	0.513
Mean	0.627	0.695	0.639	0.706	0	1	0.526	0.723	0.667	0.761	0.591	0.701	0.635	0.719	0.654	0.738	0.601	0.758
Std. Dev.	0.204	0.225	0.225	0.222	0	0	0.165	0.22	0.266	0.164	0.232	0.182	0.187	0.17	0.183	0.164	0.212	0.159

Table 3.3. Continued.

Locus	Populations																	
	subqRev		subqGold		subqLL		subqCan		subqHin		subqJas		subqPA		subqWC		subqCal	
	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E
<i>Drec1007</i>	0.583	0.560	0.542	0.691	0.576	0.673	0.857	0.838	0.625	0.745	0.548	0.736	0.250	0.821	1.000	1.000	0.667	0.600
<i>Drec1008</i>	0.818	0.962	0.750	0.935	0.735	0.946	0.444	0.889	0.625	0.889	0.733	0.824	0.250	0.464	0.500	0.833	0.667	0.800
<i>Drec1024</i>	0.870	0.917	0.875	0.897	0.871	0.913	0.889	0.954	0.938	0.877	0.871	0.940	1.000	1.000	1.000	1.000	0.667	0.867
<i>Drec1079</i>	0.364	0.457	0.500	0.580	0.333	0.456	1.000	0.906	0.625	0.647	0.690	0.714	0.500	0.821	1.000	0.667	0.333	0.733
<i>Drec1081</i>	0.833	0.842	0.917	0.809	0.750	0.769	0.333	0.731	0.800	0.831	0.750	0.815	1.000	1.000	-	-	0.667	0.733
<i>Drec1083</i>	0.417	0.742	0.333	0.723	0.438	0.739	0.500	0.400	0.375	0.627	0.333	0.730	0.000	0.714	-	-	-	-
<i>Drec2002</i>	0.792	0.902	0.870	0.929	0.903	0.884	0.667	0.949	0.938	0.939	0.774	0.947	1.000	1.000	0.500	0.833	1.000	0.933
<i>Drec2004</i>	0.600	0.773	0.478	0.678	0.700	0.790	1.000	0.926	0.571	0.707	0.808	0.850	-	-	0.500	0.833	0.667	0.867
<i>Drec2084</i>	0.500	0.770	0.667	0.786	0.516	0.841	0.333	0.810	0.500	0.830	0.833	0.916	-	-	1.000	1.000	1.000	1.000
<i>Drec2085</i>	0.667	0.878	0.818	0.869	0.594	0.842	0.857	0.943	0.867	0.909	0.839	0.922	1.000	1.000	1.000	1.000	0.667	0.933
<i>Drec3001</i>	0.667	0.754	0.917	0.781	0.676	0.784	0.875	0.897	0.786	0.844	0.871	0.885	1.000	0.857	1.000	0.833	1.000	0.933
<i>Drec3028</i>	0.625	0.581	0.417	0.396	0.303	0.392	0.556	0.749	0.667	0.649	0.677	0.668	0.750	0.607	1.000	0.833	0.667	0.733
<i>Drec3038</i>	0.750	0.873	0.917	0.917	0.581	0.901	0.889	0.953	0.933	0.925	0.862	0.935	0.750	0.786	0.500	0.833	1.000	0.933
<i>Drec3047</i>	0.458	0.793	0.625	0.806	0.688	0.813	0.222	0.901	0.533	0.899	0.484	0.869	0.250	0.893	0.000	0.667	0.667	0.600
<i>Drec3089</i>	0.458	0.483	0.417	0.427	0.333	0.441	1.000	0.934	0.750	0.907	0.548	0.751	0.500	0.929	0.500	0.833	0.667	0.733
<i>Drec3090</i>	0.875	0.847	0.783	0.856	0.750	0.835	0.500	0.757	0.500	0.702	0.742	0.729	1.000	0.833	1.000	1.000	1.000	0.933
<i>Drec3091</i>	0.458	0.832	0.292	0.832	0.387	0.861	0.667	0.854	0.529	0.759	0.387	0.811	0.250	0.750	0.500	0.833	0.667	0.533
<i>Drec4069</i>	0.391	0.349	0.208	0.236	0.406	0.605	1.000	0.800	0.750	0.699	0.516	0.620	0.500	0.571	0.500	0.500	-	-
<i>Dtra31</i>	0.458	0.873	0.478	0.872	0.909	0.873	0.500	0.772	0.941	0.889	0.581	0.909	0.750	0.929	0.000	0.667	1.000	0.933
<i>Dtra40</i>	0.833	0.871	0.667	0.792	0.824	0.879	0.556	0.690	0.765	0.775	0.677	0.831	0.750	0.786	1.000	0.833	0.667	0.733
<i>Dtra86</i>	0.304	0.335	0.417	0.453	0.559	0.594	0.444	0.405	0.813	0.784	0.484	0.721	0.750	0.750	1.000	0.833	0.333	0.733
Mean	0.606	0.733	0.614	0.727	0.611	0.754	0.671	0.812	0.706	0.802	0.667	0.815	0.645	0.816	0.711	0.833	0.737	0.804
Std. Dev.	0.185	0.192	0.226	0.198	0.193	0.165	0.254	0.158	0.169	0.101	0.165	0.098	0.326	0.152	0.346	0.136	0.21	0.136

Table 3.3. Continued.

Locus	Populations																	
	recRev		recGold		recLL		recCan		recHin		recJas		recNor		recPA		recWC	
	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E
<i>Drec1007</i>	1.000	1.000	0.750	0.464	-	-	0.444	0.719	0.760	0.842	0.750	0.825	0.500	0.833	0.742	0.720	1.000	0.725
<i>Drec1008</i>	1.000	1.000	1.000	0.607	1.000	1.000	0.800	0.958	0.833	0.967	0.750	0.917	0.500	0.833	0.828	0.930	0.857	0.824
<i>Drec1024</i>	1.000	1.000	1.000	0.929	1.000	1.000	0.900	0.958	0.960	0.961	1.000	0.975	-	-	0.862	0.938	0.857	0.923
<i>Drec1079</i>	1.000	1.000	0.750	0.929	-	-	1.000	0.853	0.750	0.779	0.429	0.648	0.500	0.833	0.667	0.720	0.571	0.582
<i>Drec1081</i>	-	-	1.000	0.933	1.000	1.000	0.800	0.932	0.880	0.925	1.000	0.929	1.000	1.000	0.600	0.844	1.000	0.955
<i>Drec1083</i>	1.000	1.000	0.500	0.929	1.000	1.000	0.375	0.725	0.294	0.861	0.750	0.775	1.000	1.000	0.433	0.915	0.571	0.846
<i>Drec2002</i>	-	-	1.000	0.893	1.000	1.000	0.667	0.850	0.760	0.966	0.625	0.917	0.500	0.833	0.792	0.954	0.857	0.967
<i>Drec2004</i>	1.000	1.000	1.000	0.964	1.000	1.000	0.889	0.915	0.778	0.778	0.750	0.925	-	-	1.000	1.000	0.857	0.956
<i>Drec2084</i>	1.000	1.000	1.000	0.933	-	-	0.375	0.917	0.357	0.828	0.857	0.945	0.500	0.833	0.714	0.879	0.000	0.800
<i>Drec2085</i>	-	-	0.750	0.857	-	-	0.900	0.958	0.800	0.971	0.857	0.956	1.000	1.000	0.818	0.935	1.000	0.989
<i>Drec3001</i>	1.000	1.000	0.667	0.929	-	-	1.000	0.941	0.800	0.922	0.714	0.956	1.000	1.000	0.839	0.921	0.857	0.835
<i>Drec3028</i>	1.000	1.000	1.000	0.964	1.000	1.000	0.600	0.711	0.680	0.691	0.625	0.767	1.000	1.000	0.630	0.667	0.857	0.659
<i>Drec3038</i>	1.000	1.000	0.750	0.857	1.000	1.000	1.000	0.974	0.875	0.967	0.750	0.950	1.000	1.000	0.815	0.957	0.857	0.923
<i>Drec3047</i>	1.000	1.000	0.667	0.964	1.000	1.000	0.700	0.911	0.708	0.958	0.875	0.975	0.500	0.500	0.593	0.959	0.714	0.945
<i>Drec3089</i>	1.000	1.000	0.500	0.600	-	-	0.700	0.916	0.680	0.931	0.750	0.900	1.000	1.000	0.556	0.906	0.857	0.857
<i>Drec3090</i>	1.000	1.000	0.750	0.750	1.000	1.000	0.875	0.925	0.680	0.928	0.750	0.892	1.000	1.000	0.727	0.909	0.667	0.879
<i>Drec3091</i>	-	-	0.750	0.964	1.000	1.000	0.556	0.876	0.520	0.852	0.000	0.833	0.500	0.833	0.467	0.844	0.500	0.909
<i>Drec4069</i>	1.000	1.000	0.250	0.857	-	-	0.750	0.792	0.609	0.697	0.571	0.791	1.000	1.000	0.538	0.787	0.667	0.758
<i>Dtra31</i>	1.000	1.000	1.000	0.750	1.000	1.000	1.000	0.902	0.870	0.938	0.875	0.808	-	-	0.828	0.930	0.857	0.890
<i>Dtra40</i>	1.000	1.000	1.000	-	1.000	1.000	1.000	0.884	0.583	0.856	0.875	0.717	0.500	0.833	0.800	0.823	1.000	0.835
<i>Dtra86</i>	1.000	1.000	0.750	0.839	1.000	1.000	0.500	0.653	0.708	0.726	0.625	0.683	1.000	1.000	0.808	0.692	1.000	0.758
Mean	1	1	0.802	0.846	1	1	0.754	0.87	0.709	0.873	0.723	0.861	0.778	0.907	0.717	0.868	0.781	0.848
Std. Dev.	0	0	0.212	0.142	0	0	0.213	0.094	0.166	0.093	0.215	0.1	0.256	0.131	0.145	0.098	0.234	0.106

Table 3.3. Continued.

Locus	Populations																	
	recCal		recKaw		recEd		recMin		recBem		recMun		recGSM		recSdC		recRoc	
	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E
<i>Drec1007</i>	0.800	0.705	0.720	0.694	0.727	0.671	0.750	0.725	0.857	0.769	0.857	0.670	0.583	0.684	0.788	0.743	0.652	0.729
<i>Drec1008</i>	0.846	0.951	0.727	0.920	1.000	0.952	0.875	0.958	1.000	0.945	0.857	0.967	0.765	0.953	0.848	0.942	0.957	0.898
<i>Drec1024</i>	0.885	0.955	0.880	0.953	0.909	0.952	0.625	0.942	0.857	0.956	1.000	0.985	0.848	0.959	0.909	0.949	0.696	0.928
<i>Drec1079</i>	0.619	0.792	0.667	0.761	0.727	0.784	0.857	0.626	0.714	0.802	0.833	0.833	0.706	0.872	0.613	0.815	0.773	0.876
<i>Drec1081</i>	0.917	0.918	0.920	0.925	0.714	0.912	1.000	0.909	1.000	0.956	0.714	0.912	0.829	0.914	0.783	0.926	0.824	0.859
<i>Drec1083</i>	0.318	0.933	0.400	0.906	0.636	0.939	0.625	0.817	0.571	0.923	0.500	0.848	0.613	0.890	0.452	0.888	0.783	0.912
<i>Drec2002</i>	0.913	0.941	0.720	0.941	0.818	0.944	0.875	0.875	0.667	0.970	0.571	0.945	0.788	0.915	0.719	0.948	0.818	0.939
<i>Drec2004</i>	0.870	0.929	0.818	0.910	0.727	0.926	1.000	0.956	0.500	0.879	0.750	0.929	0.750	0.906	0.690	0.894	0.682	0.802
<i>Drec2084</i>	0.560	0.891	0.600	0.858	0.800	0.778	0.400	0.867	0.000	0.667	0.600	0.933	0.379	0.884	0.844	0.915	0.941	0.907
<i>Drec2085</i>	0.875	0.947	0.833	0.954	0.900	0.958	1.000	0.985	0.833	0.985	0.857	0.956	0.788	0.918	0.750	0.956	0.864	0.930
<i>Drec3001</i>	0.885	0.910	0.800	0.900	0.818	0.879	0.875	0.925	0.857	0.890	1.000	0.864	0.861	0.937	0.970	0.917	0.818	0.893
<i>Drec3028</i>	0.667	0.602	0.760	0.651	0.700	0.716	0.750	0.692	1.000	0.758	0.429	0.736	0.647	0.693	0.594	0.714	0.696	0.688
<i>Drec3038</i>	0.783	0.964	0.880	0.947	1.000	0.942	0.750	0.925	0.500	0.924	1.000	0.978	0.771	0.958	0.688	0.966	1.000	0.954
<i>Drec3047</i>	0.652	0.966	0.560	0.910	0.818	0.961	0.625	0.908	0.833	0.924	0.714	0.901	0.714	0.867	0.750	0.962	0.773	0.958
<i>Drec3089</i>	0.792	0.900	0.708	0.890	0.545	0.939	0.857	0.912	0.714	0.945	0.714	0.879	0.600	0.924	0.758	0.902	0.565	0.858
<i>Drec3090</i>	0.920	0.928	0.880	0.917	0.714	0.923	0.714	0.912	0.571	0.934	1.000	0.911	0.882	0.912	0.806	0.931	0.773	0.897
<i>Drec3091</i>	0.640	0.873	0.600	0.890	0.300	0.784	0.500	0.650	0.714	0.802	0.286	0.835	0.333	0.833	0.485	0.835	0.238	0.847
<i>Drec4069</i>	0.750	0.852	0.542	0.787	0.455	0.857	0.875	0.842	0.714	0.813	0.857	0.714	0.686	0.758	0.576	0.818	0.565	0.740
<i>Dtra31</i>	0.875	0.948	0.647	0.820	0.909	0.948	0.875	0.842	0.571	0.879	0.857	0.802	0.471	0.914	0.909	0.917	0.870	0.886
<i>Dtra40</i>	0.769	0.845	0.957	0.824	0.545	0.732	0.625	0.775	0.714	0.791	1.000	0.857	0.343	0.392	0.576	0.651	0.696	0.564
<i>Dtra86</i>	0.696	0.793	0.640	0.759	0.636	0.745	0.714	0.747	0.571	0.824	1.000	0.758	0.706	0.655	0.667	0.787	0.818	0.785
Mean	0.763	0.883	0.727	0.863	0.733	0.869	0.77	0.847	0.703	0.873	0.781	0.867	0.67	0.845	0.722	0.875	0.752	0.85
Std. Dev.	0.15	0.093	0.142	0.087	0.175	0.096	0.163	0.106	0.226	0.086	0.207	0.09	0.167	0.138	0.139	0.089	0.165	0.1

Table 3.3. Continued.

Locus	Populations										Mean	
	recPeru		recNH		recPeru		recME		recPEI		<i>subq</i>	<i>recens</i>
	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _O
<i>Drec1007</i>	0.696	0.792	0.714	0.780	0.696	0.792	0.733	0.793	0.500	0.679	0.593	0.719
<i>Drec1008</i>	0.792	0.966	1.000	0.956	0.792	0.966	0.967	0.933	0.750	0.786	0.753	0.851
<i>Drec1024</i>	0.913	0.968	0.667	0.894	0.913	0.968	0.867	0.948	0.750	0.929	0.844	0.866
<i>Drec1079</i>	0.625	0.822	0.833	0.894	0.625	0.822	0.833	0.843	0.750	0.786	0.483	0.704
<i>Drec1081</i>	0.958	0.936	0.500	0.894	0.958	0.936	0.810	0.922	1.000	0.929	0.765	0.854
<i>Drec1083</i>	0.462	0.886	0.333	0.712	0.462	0.886	0.571	0.813	0.750	0.893	0.427	0.516
<i>Drec2002</i>	0.875	0.965	0.714	0.912	0.875	0.965	0.857	0.943	0.750	0.857	0.790	0.785
<i>Drec2004</i>	0.650	0.892	0.800	0.822	0.650	0.892	0.889	0.903	0.500	0.750	0.638	0.773
<i>Drec2084</i>	0.450	0.873	1.000	1.000	0.450	0.873	0.741	0.899	0.000	0.667	0.583	0.607
<i>Drec2085</i>	0.826	0.958	0.667	0.970	0.826	0.958	0.815	0.957	1.000	0.929	0.756	0.825
<i>Drec3001</i>	0.864	0.922	1.000	0.956	0.864	0.922	0.900	0.871	1.000	1.000	0.794	0.867
<i>Drec3028</i>	0.609	0.625	0.857	0.758	0.609	0.625	0.621	0.784	0.250	0.536	0.483	0.669
<i>Drec3038</i>	0.958	0.969	1.000	0.970	0.958	0.969	0.964	0.961	0.750	0.893	0.799	0.853
<i>Drec3047</i>	0.667	0.954	0.857	0.945	0.667	0.954	0.586	0.940	0.250	0.821	0.565	0.684
<i>Drec3089</i>	0.636	0.914	0.857	0.978	0.636	0.914	0.700	0.908	0.750	0.893	0.475	0.679
<i>Drec3090</i>	0.818	0.922	0.667	0.924	0.818	0.922	0.724	0.921	1.000	0.933	0.765	0.799
<i>Drec3091</i>	0.478	0.882	0.429	0.901	0.478	0.882	0.393	0.853	0.250	0.893	0.488	0.450
<i>Drec4069</i>	0.762	0.861	0.714	0.714	0.762	0.861	0.655	0.778	0.750	0.643	0.366	0.639
<i>Dtra31</i>	0.833	0.879	1.000	0.956	0.833	0.879	0.900	0.879	0.750	0.607	0.655	0.834
<i>Dtra40</i>	0.583	0.593	0.143	0.385	0.583	0.593	0.567	0.610	0.750	0.643	0.702	0.665
<i>Dtra86</i>	0.727	0.790	0.857	0.857	0.727	0.790	0.621	0.763	0.750	0.643	0.497	0.712
Mean	0.723	0.875	0.743	0.866	0.723	0.875	0.748	0.868	0.667	0.796		
Std. Dev.	0.156	0.104	0.234	0.14	0.156	0.104	0.152	0.087	0.278	0.134		

Table 3.3. Observed and expected heterozygosity for X-linked loci. Calculated using only female samples. Monomorphic and all male populations excluded. Bold values significant after Bonferroni correction ($P=0.00035$).

Locus	Populations																	
	subqPort		subqSea		subqMiss		subqShu		subqCon		subqTak		subqWap		subqSic		subqCan	
	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E
<i>Drec1019</i>	0.679	0.584	0.741	0.621	0.520	0.597	0.500	0.786	0.750	0.742	0.889	0.791	0.677	0.741	0.880	0.767	0.800	0.911
<i>Drec1074</i>	0.897	0.829	0.815	0.796	0.654	0.647	0.750	0.964	0.750	0.767	0.556	0.693	0.710	0.722	0.793	0.717	1.000	0.933
<i>Drec2083</i>	0.893	0.969	0.815	0.975	0.769	0.915	1.000	1.000	0.857	0.967	0.889	0.967	0.867	0.949	0.966	0.940	0.250	0.750
<i>Drec3079</i>	0.880	0.921	0.708	0.885	0.607	0.795	0.750	0.893	0.750	0.792	0.778	0.909	0.581	0.887	0.536	0.877	0.600	0.911
<i>Dtra135</i>	0.300	0.689	0.273	0.857	0.211	0.785	0.000	0.800	0.625	0.933	0.778	0.935	0.806	0.926	0.750	0.916	0.400	0.889
<i>Dtra26A</i>	0.444	0.534	0.520	0.536	0.280	0.285	0.500	0.571	0.375	0.458	0.444	0.471	0.567	0.536	0.586	0.565	0.400	0.644
<i>Dtra62</i>	0.893	0.936	1.000	0.914	0.926	0.901	1.000	0.964	0.875	0.875	1.000	0.948	0.833	0.929	0.931	0.883	1.000	0.964
<i>Dtra68</i>	0.750	0.799	0.870	0.782	0.692	0.675	0.750	0.714	0.750	0.792	0.778	0.856	0.742	0.759	0.793	0.796	0.400	0.844
<i>Dtra89</i>	0.464	0.470	0.357	0.323	0.083	0.083	0.333	0.333	0.200	0.200	0.333	0.318	0.231	0.434	0.542	0.482	0.000	0.533
Mean	0.689	0.748	0.678	0.743	0.527	0.632	0.62	0.781	0.659	0.725	0.716	0.765	0.668	0.765	0.753	0.771	0.539	0.82
Std. Dev.	0.231	0.186	0.244	0.211	0.28	0.28	0.323	0.217	0.227	0.245	0.223	0.23	0.194	0.182	0.164	0.159	0.341	0.147

Table 3.3. continued. Observed and expected heterozygosity for X-linked loci.

Locus	Populations																	
	subqHin		subqJas		subqPA		subqCal		subqKaw		subqRev		subqGold		subqLL		recRev	
	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E
<i>Drec1019</i>	0.375	0.847	0.621	0.836	0.500	0.500	0.667	0.867	0.375	0.733	0.792	0.745	0.625	0.810	0.652	0.733	0.000	0.167
<i>Drec1074</i>	0.938	0.875	0.933	0.891	0.500	0.833	1.000	0.933	0.286	0.747	0.739	0.694	0.708	0.715	0.818	0.770	0.149	0.840
<i>Drec2083</i>	0.692	0.945	0.867	0.958	-	-	1.000	1.000	1.000	0.985	0.917	0.926	0.783	0.959	0.870	0.947	0.078	0.224
<i>Drec3079</i>	0.813	0.863	0.800	0.914	1.000	1.000	1.000	0.933	0.375	0.858	0.708	0.880	0.708	0.921	0.696	0.888	0.605	0.063
<i>Dtra135</i>	0.333	0.935	0.567	0.956	0.500	0.833	1.000	1.000	0.286	0.945	0.833	0.887	0.750	0.934	0.818	0.897	0.102	0.288
<i>Dtra26A</i>	0.400	0.522	0.586	0.583	1.000	0.667	1.000	0.800	0.857	0.747	0.417	0.513	0.500	0.434	0.500	0.500	0.950	0.644
<i>Dtra62</i>	1.000	0.927	0.900	0.922	1.000	1.000	1.000	1.000	1.000	0.978	0.875	0.909	0.833	0.926	0.870	0.892	0.048	0.376
<i>Dtra68</i>	0.688	0.696	0.800	0.753	1.000	0.833	1.000	0.867	0.571	0.747	0.917	0.804	0.583	0.832	0.783	0.776	0.286	-
<i>Dtra89</i>	0.438	0.837	0.200	0.739	-	-	0.667	0.867	0.286	0.868	0.278	0.422	0.500	0.518	0.286	0.626	0.006	0.002
Mean	0.631	0.827	0.697	0.839	0.786	0.81	0.926	0.919	0.56	0.845	0.719	0.753	0.666	0.783	0.699	0.781	0.247	0.326
Std. Dev.	0.254	0.137	0.232	0.125	0.267	0.178	0.147	0.073	0.31	0.106	0.226	0.181	0.12	0.191	0.196	0.146	0.324	0.287

Table 3.3. continued. Observed and expected heterozygosity for X-linked loci.

Locus	Populations													
	recGold		recCan		recHin		recJas		recPA		recEd		recKaw	
	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E
<i>Drec1019</i>	1.000	0.139	0.007	0.000	0.000	0.167	0.086	1.000	0.007	0.000	0.086	1.000	1.000	0.139
<i>Drec1074</i>	1.000	0.313	0.532	0.171	0.149	0.840	0.120	1.000	0.532	0.171	0.120	1.000	1.000	0.313
<i>Drec2083</i>	-	0.139	0.802	0.028	0.078	0.224	1.000	-	0.802	0.028	1.000	-	-	0.139
<i>Drec3079</i>	1.000	0.307	0.214	0.014	0.605	0.063	0.312	1.000	0.214	0.014	0.312	1.000	1.000	0.307
<i>Dtra135</i>	-	1.000	0.360	0.066	0.102	0.288	1.000	-	0.360	0.066	1.000	-	-	1.000
<i>Dtra26A</i>	-	1.000	1.000	0.458	0.950	0.644	0.545	1.000	1.000	0.458	0.545	1.000	-	1.000
<i>Dtra62</i>	1.000	1.000	0.552	0.958	0.048	0.376	0.659	1.000	0.552	0.958	0.659	1.000	1.000	1.000
<i>Dtra68</i>	-	-	-	0.018	0.286	-	-	-	-	0.018	-	-	-	-
<i>Dtra89</i>	-	-	1.000	0.001	0.006	0.002	0.429	-	1.000	0.001	0.429	-	-	-
Mean	1	0.557	0.559	0.19	0.247	0.326	0.519	1	0.559	0.19	0.519	1	1	0.557
Std. Dev.	0	0.42	0.361	0.323	0.324	0.287	0.355	0	0.361	0.323	0.355	0	0	0.42

Table 3.3. Continued. Observed and expected heterozygosity for X-linked loci.

Locus	Populations								Mean	
	recCal		recMin		recGSM		recPeru		<i>subq</i>	<i>recens</i>
	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _E	H _O	H _O
<i>Drec1019</i>	0.000	0.000	0.003	1.000	0.000	0.000	0.003	1.000	0.650	0.658
<i>Drec1074</i>	0.249	0.045	0.532	1.000	0.249	0.045	0.532	1.000	0.756	0.855
<i>Drec2083</i>	0.710	0.000	0.000	1.000	0.710	0.000	0.000	1.000	0.840	0.808
<i>Drec3079</i>	0.780	0.571	0.001	1.000	0.780	0.571	0.001	1.000	0.723	0.873
<i>Dtra135</i>	0.533	0.180	0.547	1.000	0.533	0.180	0.547	1.000	0.543	0.765
<i>Dtra26A</i>	0.039	0.108	0.166	1.000	0.039	0.108	0.166	1.000	0.552	0.794
<i>Dtra62</i>	0.243	0.441	0.260	1.000	0.243	0.441	0.260	1.000	0.937	0.855
<i>Dtra68</i>	-	-	-	1.000	-	-	-	1.000	0.757	0.500
<i>Dtra89</i>	0.076	0.002	0.006	-	0.076	0.002	0.006	-	0.325	0.220
Mean	0.329	0.168	0.189	1	0.329	0.168	0.189	1		
Std. Dev.	0.307	0.22	0.236	0	0.307	0.22	0.236	0		

Table 3.4. *P*-values of Hardy-Weinberg exact tests of autosomal loci performed in Arlequin (10,000 permutations) for *D. recens* populations. Bolded values are significant after Bonferroni correction ($P < 0.0001$). Dashes indicate monomorphic populations which were not tested.

Locus	Populations													
	recJas	recEd	recWC	recCan	recHin	recKaw	recNor	recPA	recRev	recGold	recLL	recCal	recPeru	recGSM
<i>Drec1007</i>	0.169	1.000	0.349	0.050	0.217	0.563	0.335	0.977	1.000	0.768	-	0.469	0.102	0.184
<i>Drec1008</i>	0.228	1.000	0.933	0.116	0.001	0.032	0.333	0.068	1.000	1.000	1.000	0.115	0.009	0.003
<i>Drec1024</i>	1.000	0.541	0.573	0.486	0.243	0.090	-	0.097	1.000	1.000	1.000	0.075	0.358	0.028
<i>Drec1079</i>	0.184	0.740	0.660	0.925	0.106	0.223	0.331	0.411	1.000	1.000	-	0.139	0.128	0.075
<i>Drec1081</i>	1.000	0.137	1.000	0.197	0.214	0.536	1.000	0.356	-	1.000	1.000	0.739	0.467	0.493
<i>Drec1083</i>	1.000	0.008	0.078	0.045	0.000	0.000	1.000	0.000	1.000	0.028	1.000	0.000	0.001	0.001
<i>Drec2002</i>	0.051	0.228	0.253	0.032	0.002	0.006	0.335	0.040	-	1.000	1.000	0.251	0.037	0.024
<i>Drec2004</i>	0.027	0.110	0.371	0.236	0.115	0.329	-	1.000	1.000	1.000	1.000	0.787	0.018	0.310
<i>Drec2084</i>	0.425	0.695	0.067	0.001	0.002	0.020	0.333	0.408	1.000	0.464	-	0.000	0.000	0.000
<i>Drec2085</i>	0.347	0.443	1.000	0.461	0.000	0.148	1.000	0.314	-	0.153	-	0.224	0.076	0.000
<i>Drec3001</i>	0.060	0.253	0.664	0.582	0.017	0.172	1.000	0.608	1.000	1.000	-	0.860	0.567	0.040
<i>Drec3028</i>	0.289	0.948	0.535	0.092	0.614	0.540	1.000	0.096	1.000	0.315	1.000	0.081	0.187	0.781
<i>Drec3038</i>	0.087	1.000	0.528	1.000	0.126	0.120	1.000	0.050	1.000	0.146	1.000	0.000	0.337	0.000
<i>Drec3047</i>	0.221	0.130	0.090	0.019	0.001	0.000	1.000	0.000	1.000	0.200	1.000	0.000	0.000	0.013
<i>Drec3089</i>	0.342	0.000	0.193	0.131	0.000	0.092	1.000	0.000	1.000	0.086	-	0.505	0.006	0.000
<i>Drec3090</i>	0.449	0.177	0.206	0.626	0.005	0.880	1.000	0.216	1.000	0.658	1.000	0.695	0.382	0.409
<i>Drec3091</i>	0.000	0.002	0.024	0.021	0.003	0.001	0.333	0.000	-	0.659	1.000	0.137	0.000	0.000
<i>Drec4069</i>	0.183	0.015	0.409	0.803	0.389	0.012	1.000	0.020	1.000	0.143	-	0.916	0.785	0.342
<i>Dtra31</i>	0.990	0.559	0.671	1.000	0.642	0.100	-	0.010	1.000	1.000	1.000	0.222	0.442	0.000
<i>Dtra40</i>	0.940	0.199	0.978	0.329	0.001	0.044	0.331	0.078	1.000	1.000	1.000	0.019	0.439	0.549
<i>Dtra86</i>	0.900	0.443	0.516	0.194	0.430	0.451	1.000	0.479	1.000	0.315	1.000	0.231	0.236	0.887

Table 3.4. Continued.

Locus	Populations							
	recME	recSdC	recBem	recMin	recMun	recPEI	recNH	recRoc
<i>Drec1007</i>	0.778	0.629	1.000	0.781	0.442	1.000	1.000	0.030
<i>Drec1008</i>	0.816	0.021	0.455	0.346	0.246	0.769	1.000	0.543
<i>Drec1024</i>	0.267	0.316	0.323	0.007	1.000	0.314	0.027	0.001
<i>Drec1079</i>	0.188	0.009	0.746	0.776	0.947	1.000	0.711	0.004
<i>Drec1081</i>	0.129	0.050	1.000	1.000	0.217	1.000	0.014	0.001
<i>Drec1083</i>	0.002	0.000	0.033	0.215	0.020	0.423	0.102	0.022
<i>Drec2002</i>	0.043	0.011	0.010	0.495	0.004	0.660	0.135	0.008
<i>Drec2004</i>	0.687	0.019	0.044	1.000	0.310	0.088	0.897	0.083
<i>Drec2084</i>	0.102	0.064	0.333	0.060	0.061	0.333	1.000	0.924
<i>Drec2085</i>	0.006	0.000	0.093	1.000	0.326	0.310	0.010	0.195
<i>Drec3001</i>	0.826	0.751	0.794	0.648	0.831	1.000	1.000	0.009
<i>Drec3028</i>	0.004	0.085	0.806	0.737	0.265	0.428	0.824	0.699
<i>Drec3038</i>	0.651	0.000	0.004	0.217	1.000	0.435	1.000	0.303
<i>Drec3047</i>	0.000	0.000	0.491	0.010	0.028	0.028	0.410	0.000
<i>Drec3089</i>	0.000	0.022	0.100	0.655	0.197	0.426	0.155	0.007
<i>Drec3090</i>	0.043	0.105	0.011	0.140	0.496	1.000	0.099	0.028
<i>Drec3091</i>	0.000	0.000	0.203	0.501	0.002	0.009	0.008	0.000
<i>Drec4069</i>	0.174	0.007	0.852	1.000	0.204	1.000	1.000	0.010
<i>Dtra31</i>	0.610	0.545	0.088	0.675	0.738	1.000	0.320	0.327
<i>Dtra40</i>	0.287	0.445	0.182	0.455	0.947	1.000	0.077	1.000
<i>Dtra86</i>	0.146	0.573	0.090	0.286	0.806	1.000	0.938	0.413

Table 3.4. Continued.

Locus	Populations											
	subqPort	subqSea	subqMiss	subqShu	subqCon	subqTak	subqWap	subqSic	subqCan	subqHin	subqJas	subqPA
<i>Drec1007</i>	0.011	0.830	0.542	0.319	1.000	0.530	0.136	0.704	0.946	0.060	0.105	0.028
<i>Drec1008</i>	0.000	0.566	0.000	1.000	0.572	0.966	0.294	0.172	0.001	0.020	0.247	0.146
<i>Drec1024</i>	0.839	0.145	0.006	0.770	0.317	0.387	0.923	0.364	0.431	0.810	0.295	1.000
<i>Drec1079</i>	0.328	0.211	0.028	1.000	0.160	0.443	0.173	0.001	0.781	0.436	0.062	0.314
<i>Drec1081</i>	0.048	0.684	0.000	1.000	0.949	0.723	0.642	0.128	0.006	0.413	0.045	1.000
<i>Drec1083</i>	0.051	0.000	0.001	1.000	0.845	0.598	0.098	0.012	1.000	0.282	0.000	0.029
<i>Drec2002</i>	0.605	0.124	0.002	0.432	0.373	0.802	0.078	0.029	0.035	0.534	0.004	1.000
<i>Drec2004</i>	0.267	0.386	0.000	0.333	0.010	0.088	0.022	0.007	1.000	0.015	0.962	-
<i>Drec2084</i>	0.645	0.038	0.016	0.767	0.000	0.001	0.000	0.297	0.064	0.022	0.157	-
<i>Drec2085</i>	0.103	0.089	0.054	1.000	0.109	0.431	0.682	0.004	0.508	0.517	0.263	1.000
<i>Drec3001</i>	0.559	0.747	0.000	0.316	0.125	0.794	0.862	0.963	0.767	0.698	0.252	0.660
<i>Drec3028</i>	0.272	1.000	1.000	0.313	1.000	0.379	1.000	0.039	0.108	0.191	0.720	1.000
<i>Drec3038</i>	0.457	0.086	0.000	1.000	0.647	0.382	0.621	0.360	0.461	0.673	0.260	0.771
<i>Drec3047</i>	0.000	0.640	0.095	0.029	0.153	0.619	0.000	0.000	0.000	0.010	0.000	0.010
<i>Drec3089</i>	0.894	1.000	0.197	0.143	0.366	1.000	0.181	0.001	1.000	0.043	0.002	0.030
<i>Drec3090</i>	0.805	0.421	0.000	0.657	0.672	0.605	0.744	0.191	0.024	0.251	0.705	1.000
<i>Drec3091</i>	0.205	0.127	0.002	0.312	0.000	0.002	0.004	0.000	0.173	0.034	0.000	0.030
<i>Drec4069</i>	0.621	0.066	0.267	1.000	0.017	1.000	0.148	0.001	0.788	0.681	0.328	1.000
<i>Dtra31</i>	0.000	0.396	0.000	0.317	0.148	0.231	0.007	0.001	0.086	0.309	0.000	0.318
<i>Dtra40</i>	0.005	0.000	0.000	0.084	0.111	0.008	0.317	0.002	0.817	0.404	0.088	0.770
<i>Dtra86</i>	1.000	0.652	0.001	0.435	0.488	0.059	0.725	0.453	1.000	0.039	0.029	1.000

Table 3.4. Continued.

Locus	Populations					
	subqWC	subqCal	subqKaw	subqRev	subqGold	subqLL
<i>Drec1007</i>	1.000	1.000	0.275	1.000	0.092	0.156
<i>Drec1008</i>	0.332	0.601	0.002	0.004	0.016	0.010
<i>Drec1024</i>	1.000	0.467	0.240	0.231	0.421	0.105
<i>Drec1079</i>	1.000	0.201	0.188	0.340	0.186	0.045
<i>Drec1081</i>	-	1.000	0.027	0.797	0.814	0.792
<i>Drec1083</i>	-	-	0.003	0.000	0.000	0.001
<i>Drec2002</i>	0.335	1.000	0.448	0.286	0.117	0.883
<i>Drec2004</i>	0.336	0.465	0.032	0.004	0.004	0.307
<i>Drec2084</i>	1.000	1.000	0.003	0.002	0.180	0.000
<i>Drec2085</i>	1.000	0.198	0.535	0.002	0.239	0.001
<i>Drec3001</i>	1.000	1.000	0.751	0.849	0.610	0.526
<i>Drec3028</i>	1.000	1.000	0.745	0.072	0.641	0.012
<i>Drec3038</i>	0.332	1.000	0.042	0.389	0.803	0.000
<i>Drec3047</i>	0.334	1.000	0.000	0.000	0.099	0.020
<i>Drec3089</i>	0.332	1.000	0.017	0.250	0.067	0.182
<i>Drec3090</i>	1.000	1.000	0.235	0.926	0.716	0.211
<i>Drec3091</i>	0.334	1.000	0.045	0.000	0.000	0.000
<i>Drec4069</i>	1.000	-	0.752	1.000	0.166	0.017
<i>Dtra31</i>	0.332	1.000	0.023	0.000	0.001	0.853
<i>Dtra40</i>	1.000	1.000	0.199	0.659	0.067	0.075
<i>Dtra86</i>	1.000	0.200	0.347	0.618	0.071	0.378

Table 3.4. *P*-values of Hardy-Weinberg exact tests of X-linked loci performed in Arlequin (10,000 permutations) for *D. recens* populations. Only female individuals were used to test for HWE; populations composed entirely of males are therefore excluded from analysis. Bolded values are significant after Bonferroni correction ($P = 0.0004$). Dashes indicate monomorphic populations which were not tested.

Locus	Populations											
	recJas	recEd	recCan	recHin	recKaw	recPA	recRev	recGold	recCal	recPeru	recGSM	recMin
<i>Drec1019</i>	0.086	1.000	0.007	0.000	0.000	0.167	1.000	0.139	0.000	0.000	0.003	1.000
<i>Drec1074</i>	0.120	1.000	0.532	0.172	0.149	0.840	1.000	0.313	0.249	0.045	0.532	1.000
<i>Drec2083</i>	1.000	-	0.802	0.028	0.078	0.224	-	0.139	0.710	0.000	0.000	1.000
<i>Drec3079</i>	0.312	1.000	0.214	0.014	0.605	0.063	1.000	0.307	0.780	0.571	0.001	1.000
<i>Dtra135</i>	1.000	-	0.360	0.066	0.102	0.288	-	1.000	0.533	0.180	0.547	1.000
<i>Dtra26A</i>	0.545	1.000	1.000	0.458	0.950	0.644	-	1.000	0.039	0.108	0.166	1.000
<i>Dtra62</i>	0.660	1.000	0.553	0.958	0.048	0.376	1.000	1.000	0.243	0.442	0.260	1.000
<i>Dtra68</i>	-	-	-	0.018	0.286	-	-	-	-	-	-	1.000
<i>Dtra89</i>	0.429	-	1.000	0.001	0.006	0.003	-	-	0.076	0.002	0.006	-

Table 3.4. *P*-values of Hardy-Weinberg exact tests of X-linked loci performed in Arlequin (10,000 permutations) for *D. subquinaria* populations. Only female individuals were used to test for HWE; populations composed entirely of males are therefore excluded from analysis. Bolded values are significant after Bonferroni correction ($P < 0.0004$). subqDea is not included as all individuals were haplotypic, and thus not tested. Dashes indicate monomorphic populations which were not tested.

Locus	Populations									
	subqPort	subqSea	subqMiss	subqShu	subqCon	subqTak	subqWap	subqSic	subqCan	subqHin
<i>Drec1019</i>	0.945	0.992	0.116	0.313	0.807	1.000	0.054	0.920	0.487	0.000
<i>Drec1074</i>	0.248	0.146	0.270	0.142	1.000	0.357	0.618	0.959	1.000	0.471
<i>Drec2083</i>	0.202	0.000	0.000	1.000	0.269	0.301	0.498	0.596	0.027	0.001
<i>Drec3079</i>	0.291	0.018	0.001	0.424	0.622	0.111	0.000	0.000	0.043	0.419
<i>Dtra135</i>	0.002	0.000	0.000	0.067	0.016	0.213	0.024	0.047	0.010	0.000
<i>Dtra26A</i>	0.033	0.805	0.035	1.000	1.000	1.000	1.000	0.915	0.240	0.157
<i>Dtra62</i>	0.112	0.759	0.002	1.000	0.504	1.000	0.059	0.938	1.000	0.723
<i>Dtra68</i>	0.064	0.190	0.186	1.000	0.302	0.385	0.816	0.179	0.036	0.812
<i>Dtra89</i>	0.416	1.000	1.000	1.000	1.000	1.000	0.004	1.000	0.048	0.010

Table 3.4. Continued. *P*-values of Hardy-Weinberg exact tests of X-linked loci performed in Arlequin (10,000 permutations) for *D. subquinaria* populations. Only female individuals were used to test for HWE; populations composed entirely of males are therefore excluded from analysis. Bolded values are significant after Bonferroni correction ($P < 0.0004$). Dashes indicate monomorphic populations which were not tested.

Locus	Populations						
	subqJas	subqPA	subqCal	subqKaw	subqRev	subqGold	subqLL
<i>Drec1019</i>	0.006	1.000	0.464	0.027	0.225	0.028	0.297
<i>Drec1074</i>	0.143	0.332	1.000	0.003	0.745	0.020	0.433
<i>Drec2083</i>	0.015	-	1.000	1.000	0.275	0.017	0.492
<i>Drec3079</i>	0.016	1.000	1.000	0.001	0.054	0.052	0.038
<i>Dtra135</i>	0.001	0.332	1.000	0.000	0.419	0.002	0.564
<i>Dtra26A</i>	0.557	1.000	1.000	1.000	0.512	0.728	0.450
<i>Dtra62</i>	0.334	1.000	1.000	1.000	0.432	0.447	0.879
<i>Dtra68</i>	0.782	1.000	1.000	0.115	0.322	0.014	0.549
<i>Dtra89</i>	0.000	-	0.463	0.007	0.130	0.836	0.001

Table 3.5. Structure results for analyses that included both species. Mean $\text{LnP}(K)$ was calculated across ten replicate runs. Highest values of $\text{LnP}(K)$ and ΔK are bolded.

Complete dataset			Autosomal dataset		X-linked dataset	
K	Mean $\text{LnP}(K)$	ΔK	Mean $\text{LnP}(K)$	ΔK	Mean $\text{LnP}(K)$	ΔK
1	-86715.14	NA	-62853.78	NA	-23861.62	NA
2	-78607.16	7021.24	-57316.85	11683.76	-21266.63	8567.42
3	-77153.53	439.38	-56330.09	312.68	-20885.58	1.41
4	-76713.14	38.22	-56085.08	2.47	-20559.1	2.65
5	-76396.01	0.16	-55926.91	0.10	-20312.46	1.28
6	-76138.92	2.79	-55781.94	0.92	-20141.06	1.84
7	-76229.13	0.15	-55737.77	0.84	-20009.25	12.00
8	-76368.75	0.81	-55567.19	1.04	-20122.53	0.39
9	-75879.13	7.50	-55634.09	0.83	-20551.01	1.08
10	-75933.56	0.38	-55336.35	1.58	-20257.61	2.60
11	-75871.44	2.68	-55262.22	0.63	-20603.82	1.38
12	-76253.92	0.68	-55275.63	0.27	-20155.77	4.64
13	-77678.50	1.27	-55344.67	0.19	-20127.04	0.31
14	-75899.19	35.68	-55455.5	0.25	-20144.33	0.46
15	-77795.32	0.11	-55487.83	0.33	-20095.59	1.14
16	-79388.47	0.18	-55650.03	0.91	-20126.55	0.75
17	-80368.44	0.39	-56112.25	0.75	-20044.45	0.30
18	-79778.66	0.75	-55776.17	3.29	-20008.25	2.25
19	-81530.56	1.39	-56223.69	0.59	-19901.28	7.69
20	-78665.73	1.58	-56276.21	0.31	-20312.66	0.83
21	-79594.20	1.08	-56526.11	0.08	-20373.2	1.23
22	-83057.57	1.68	-56835.46	0.19	-20851.98	1.84
23	-82508.23	0.04	-56983.65	0.81	-20632.57	1.84
24	-81799.74	0.26	-58052.81	1.04	-20776.03	0.39
25	-80183.87	NA	-57044.18	NA	-20799.63	NA

Table 3.6. Structure results for *D. recens*. Mean $\text{LnP}(K)$ was calculated across ten replicate runs. Highest values of $\text{LnP}(K)$ and ΔK are bolded.

K	Complete dataset		Autosomal dataset		X-linked dataset	
	Mean $\text{LnP}(K)$	ΔK	Mean $\text{LnP}(K)$	ΔK	Mean $\text{LnP}(K)$	ΔK
1	-39114.26	NA	-30622.03	NA	-8493.00	NA
2	-39493.68	18.85	-30814.46	0.26	-8330.37	2.60
3	-39227.13	13.09	-31015.31	4.92	-8350.05	2.43
4	-39276.59	22.91	-31000.58	2.15	-8330.54	6.63
5	-40267.86	0.06	-30863.58	152.02	-8197.23	4.25
6	-41066.36	0.55	-35333.42	2.75	-8148.46	5.14
7	-40160.24	1.45	-32338.02	1.33	-8168.28	4.36
8	-41217.71	0.53	-32238.06	0.35	-8259.17	3.38
9	-40945.11	0.44	-31458.82	2.01	-8571.50	1.27
10	-39781.52	1.87	-31771.41	0.57	-8681.03	0.50
11	-39623.92	0.14	-31613.60	0.44	-8870.16	1.08
12	-39506.95	0.95	-31686.35	0.73	-9130.17	0.93
13	-39638.82	0.17	-32051.81	1.02	-9178.36	0.64
14	-39806.61	0.02	-31989.81	0.88	-9297.50	0.17
15	-39968.31	0.37	-32172.19	1.24	-9388.13	0.48
16	-40240.50	0.72	-32633.51	0.93	-9398.30	1.97
17	-40175.47	0.86	-32521.89	1.79	-9681.55	1.59
18	-40362.65	0.01	-33007.82	2.07	-9627.67	1.64
19	-40546.89	0.65	-32651.60	5.14	-9813.41	1.05
20	-40943.91	0.92	-33295.13	1.47	-9749.92	1.95
21	-41007.28	0.67	-32905.94	18.40	-10081.24	1.30
22	-41398.66	NA	-33585.76	NA	-10123.74	0.28

Table 3.7. Population pairwise F_{ST} within *D. recens*. F_{ST} values are shown below the diagonal, P -values at 10,000 permutations are shown above the diagonal. Values significant at $P < 0.01$ are shown in bold; values significant at $P < 0.0001$ are italicized.

	recRev	recGold	recLL	recCan	recHin	recJas	recNor	recPA	recWC	recCal	recKaw	recEd
recRev	0.000											
recGold	0.071	0.000										
recLL	0.169	0.005	0.000									
recCan	0.059	0.022	0.021	0.000								
recHin	0.079	0.006	0.012	0.018	0.000							
recJas	0.104	0.005	0.094	-0.007	-0.008	0.000						
recNor	0.119	-0.103	0.256	-0.134	-0.148	-0.064	0.000					
recPA	0.118	-0.097	0.100	-0.078	-0.076	-0.062	-0.051	0.000				
recWC	0.116	-0.045	0.157	-0.027	-0.030	-0.012	-0.052	-0.066	0.000			
recCal	0.100	0.008	0.081	0.005	-0.005	0.000	-0.084	-0.082	-0.011	0.000		
recKaw	0.064	0.004	0.007	0.008	-0.002	-0.010	-0.133	-0.106	-0.050	-0.008	0.000	
recEd	0.112	-0.034	0.133	-0.031	-0.040	-0.003	-0.075	-0.062	0.005	-0.021	-0.052	0.000
recMin	0.127	-0.010	0.154	-0.013	-0.018	-0.004	-0.035	-0.033	0.010	-0.005	-0.031	0.002
recBem	0.074	-0.048	0.132	-0.035	-0.040	-0.013	-0.071	-0.054	0.010	-0.020	-0.058	0.001
recMun	0.116	-0.012	0.132	-0.021	-0.024	-0.009	-0.042	-0.047	0.021	-0.003	-0.036	0.010
recGSM	0.152	0.045	0.091	0.043	0.038	0.028	-0.046	-0.036	0.036	0.030	0.035	0.022
recSdC	0.096	-0.010	0.102	-0.016	-0.022	0.005	-0.051	-0.063	0.022	-0.002	-0.026	0.013
recRoc	0.117	-0.004	0.108	-0.009	-0.014	0.015	-0.032	-0.047	0.035	0.007	-0.026	0.032
recPeru	0.097	0.025	0.019	0.024	0.029	-0.006	-0.119	-0.077	-0.017	0.006	0.023	-0.026
recNH	0.140	-0.017	0.148	-0.023	-0.017	-0.011	-0.048	-0.012	0.035	-0.002	-0.035	0.034
recME	0.118	-0.004	0.114	-0.016	-0.014	0.007	-0.042	-0.051	0.029	0.008	-0.018	0.023
recPEI	0.172	-0.024	0.149	-0.010	-0.013	-0.008	-0.006	-0.031	0.041	0.012	-0.029	0.038

Table 3.7 continued. Population pairwise F_{ST} within *D. recens*. F_{ST} values are shown below the diagonal, P -values at 10,000 permutations are shown above the diagonal. Values significant at $P < 0.05$ are shown in bold; values significant at $P < 0.01$ are italicized.

	recMin	recBem	recMun	recGSM	recSdC	recRoc	recPeru	recNH	recME	recPEI
recMin	0.000									
recBem	0.004	0.000								
recMun	0.024	0.009	0.000							
recGSM	0.022	0.002	0.003	0.000						
recSdC	0.014	0.001	-0.001	-0.010	0.000					
recRoc	0.035	0.014	0.012	-0.004	0.009	0.000				
recPeru	-0.015	-0.044	-0.039	0.003	-0.045	-0.038	0.000			
recNH	0.021	0.020	0.013	-0.017	-0.010	-0.002	-0.049	0.000		
recME	0.028	0.005	0.000	-0.007	0.004	0.014	-0.045	-0.007	0.000	
recPEI	0.032	0.029	0.025	0.002	0.007	0.033	-0.040	0.036	0.018	0.000

Table 3.8. Results of AMOVA including both autosomal and X-linked markers for *D. recens*. *P*-values in bold are significant at the Bonferroni-corrected level of $P < 0.005$.

Grouping:		All loci		Autosomal loci		X-linked	
		%var	<i>F</i>	%var	<i>F</i>	%var	<i>F</i>
Sympatric v. Allopatric	Among groups	1.2	$\Phi_{CT} = 0.0120$ ($P = 0.0005$)	1.03	$\Phi_{CT} = 0.0103$ ($P = 0.0002$)	2.05	$\Phi_{CT} = 0.02045$ ($P = 0.0696$)
	Among pops	-1.35	$\Phi_{SC} = -0.0137$ ($P = 1.0000$)	-0.68	$\Phi_{SC} = -0.0069$ ($P = 1.0000$)	-4.7	-0.04803 ($P = 1.0000$)
	Within pops	100.16	$\Phi_{ST} = -0.0016$ ($P = 1.0000$)	99.65	$\Phi_{ST} = 0.0035$ ($P = 0.9968$)	102.66	-0.0266 ($P = 1.0000$)
GSM v. all others	Among groups	2.54	$\Phi_{CT} = 0.0254$ ($P = 0.1329$)	1.08	$\Phi_{CT} = 0.0108$ ($P = 0.3206$)	9.42	$\Phi_{CT} = 0.09424$ ($P = 0.08772$)
	Among pops	-1.23	$\Phi_{SC} = -0.0127$ ($P = 1.0000$)	-0.36	$\Phi_{SC} = -0.0036$ ($P = 0.9999$)	-5.35	$\Phi_{SC} = -0.05905$ ($P = 1.0000$)
	Within pops	98.7	$\Phi_{ST} = 0.0130$ ($P = 1.0000$)	99.27	$\Phi_{ST} = 0.0073$ ($P = 0.99614$)	95.92	$\Phi_{ST} = 0.04076$ ($P = 1.0000$)

Table 3.9. Structure results for *D. subquinaria*. Mean $\text{LnP}(K)$ was calculated across ten replicate runs. Highest values of $\text{LnP}(K)$ and ΔK are bolded.

K	Complete dataset		Autosomal dataset		X-linked dataset	
	Mean $\text{LnP}(K)$	ΔK	Mean $\text{LnP}(K)$	ΔK	Mean $\text{LnP}(K)$	ΔK
1	-39677.13	NA	-26814.25	NA	-12861.83	NA
2	-38229.05	130.17	-26844.61	1.61	-12587.40	46.56
3	-37679.68	112.75	-25880.86	8.53	-12428.18	76.41
4	-37562.08	0.27	-25319.87	5.84	-12884.65	2.82
5	-37631.48	0.61	-25620.06	0.34	-12167.61	130.02
6	-38251.10	0.59	-26071.18	0.48	-12115.62	1.47
7	-37664.02	6.20	-25536.90	6.22	-12022.12	12.83
8	-38466.36	3.51	-25402.53	3.98	-12152.73	1.42
9	-44005.11	2.02	-25462.11	2.22	-12412.58	0.31
10	-41340.80	0.23	-25644.45	3.19	-12764.55	0.55
11	-39116.74	1.69	-26199.27	3.74	-12947.88	0.73
12	-38963.55	0.37	-27955.53	1.21	-13003.12	0.51
13	-38225.12	0.36	-27589.21	0.17	-12881.66	0.35
14	-38072.36	1.96	-26772.41	0.16	-12837.38	0.31
15	-38684.51	0.71	-26173.63	0.46	-12835.13	0.71
16	-38730.64	0.43	-25821.33	1.71	-12948.43	0.05
17	-39084.02	0.92	-26385.93	0.25	-13067.79	0.55
18	-38852.00	1.85	-26686.46	0.38	-13083.01	1.11
19	-39428.39	0.60	-27283.43	0.62	-13334.88	1.02
20	-39322.09	NA	-26726.36	NA	-13395.30	NA

Table 3.10. Population pairwise F_{ST} within *D. subquinaria*. F_{ST} values are shown below the diagonal. Values significant at $P < 0.01$ after 10,000 permutations are shown in bold; values significant at $P < 0.0001$ are italicized.

	subqPort	subqSea	subqDea	subqMiss	subqShu	subqCon	subqTak	subqWap	subqSic	subqRev	subqGold
subqPort	0.000										
subqSea	0.001	0.000									
subqDea	-0.261	-0.286	0.000								
subqMiss	0.018	0.016	-0.282	0.000							
subqShu	0.017	0.025	-0.352	0.031	0.000						
subqCon	-0.011	-0.002	-0.318	0.010	0.020	0.000					
subqTak	-0.003	0.007	-0.310	0.020	0.008	0.000	0.000				
subqWap	-0.001	-0.002	-0.371	0.001	0.016	-0.013	-0.005	0.000			
subqSic	0.005	0.006	-0.370	0.004	-0.003	-0.008	-0.005	0.007	0.000		
subqRev	-0.006	-0.001	-0.362	0.005	0.006	-0.006	-0.007	0.002	0.001	0.000	
subqGold	-0.009	-0.004	-0.383	-0.004	0.003	-0.015	-0.007	0.004	0.001	0.003	0.000
subqLL	0.004	0.007	-0.328	0.007	0.009	0.001	0.002	0.000	0.003	0.001	-0.001
subqCan	0.102	0.093	-0.155	0.123	0.053	0.088	0.081	0.052	0.036	0.047	0.047
subqHin	0.118	0.105	-0.239	0.103	0.078	0.090	0.090	0.082	0.061	0.080	0.072
subqJas	0.054	0.044	-0.361	0.023	0.022	0.027	0.022	0.044	0.030	0.040	0.037
subqPA	0.050	0.042	-0.185	0.076	0.016	0.060	0.044	0.016	-0.003	0.015	0.009
subqWC	0.022	0.021	-0.198	0.038	-0.056	0.047	0.030	-0.037	-0.055	-0.038	-0.045
subqCal	0.128	0.106	-0.397	0.104	0.069	0.111	0.093	0.097	0.077	0.096	0.092
subqKaw	0.078	0.077	-0.222	0.064	0.049	0.081	0.069	0.042	0.029	0.043	0.043

Table 3.10 Continued. Population pairwise F_{ST} within *D. subquinaria*. F_{ST} values are shown below the diagonal, P-values at 10,000 permutations are shown above the diagonal. Values significant at $P < 0.01$ are shown in bold; values significant at $P < 0.0001$ are italicized.

	subqLL	subqCan	subqHin	subqJas	subqPA	subqWC	subqCal	subqKaw
subqLL	0.000							
subqCan	<i>0.058</i>	0.000						
subqHin	<i>0.076</i>	0.022	0.000					
subqJas	<i>0.031</i>	-0.016	-0.004	0.000				
subqPA	0.019	-0.013	-0.030	-0.064	0.000			
subqWC	-0.005	-0.011	-0.047	-0.088	-0.091	0.000		
subqCal	<i>0.094</i>	0.035	0.014	0.013	-0.064	-0.118	0.000	
subqKaw	<i>0.049</i>	-0.003	-0.014	-0.026	-0.055	-0.026	0.012	0.000

Table 3.11. Results of AMOVA including both autosomal and X-linked markers for *D. subquinnaria*. *P*-values in bold are significant at $P < 0.05$; Italicized values are significant at the Bonferroni-corrected level of $P < 0.0014$.

Grouping		All loci		Autosomal loci		X-linked loci	
		%var	<i>F</i>	%var	<i>F</i>	%var	<i>F</i>
Sympatric v. Allopatric	Among groups	1.48	$\Phi_{CT} = 0.01481$ ($P = 0.0137$)	1.38	$\Phi_{CT} = 0.0138$ ($P = 0.0159$)	1.74	$\Phi_{CT} = 0.0174$ ($P = 0.0139$)
	Among pops	1.25	$\Phi_{SC} = 0.01266$ ($P < 0.0001$)	1.92	$\Phi_{SC} = 0.0195$ ($P < 0.0001$)	-0.51	$\Phi_{SC} = -0.0052$ ($P = 0.9983$)
	Within pops	97.27	$\Phi_{ST} = 0.02728$ ($P < 0.0001$)	96.7	$\Phi_{ST} = 0.0330$ ($P < 0.0001$)	98.76	$\Phi_{ST} = 0.0124$ ($P = 0.2122$)
Inland v. Coastal	Among groups	0.78	$\Phi_{CT} = 0.0078$ ($P = 0.1549$)	1.14	$\Phi_{CT} = 0.0114$ ($P = 0.1528$)	-0.17	$\Phi_{CT} = -0.0017$ ($P = 0.5561$)
	Among pops	1.82	$\Phi_{SC} = 0.0183$ ($P < 0.0001$)	2.33	$\Phi_{SC} = 0.0235$ ($P < 0.0001$)	0.49	$\Phi_{SC} = 0.0048$ ($P = 0.2616$)
	Within pops	97.4	$\Phi_{ST} = 0.0260$ ($P < 0.0001$)	96.54	$\Phi_{ST} = 0.0346$ ($P < 0.0001$)	99.69	$\Phi_{ST} = 0.0031$ ($P = 0.2196$)
Sympatric v. Inland Allopatric v. Coastal Allopatric	Among groups	1.21	$\Phi_{CT} = 0.0121$ ($P = 0.0439$)	1.22	$\Phi_{CT} = 0.0122$ ($P = 0.0455$)	1.17	$\Phi_{CT} = 0.0117$ ($P = 0.0765$)
	Among pops	1.25	$\Phi_{SC} = 0.0127$ ($P < 0.0001$)	1.86	$\Phi_{SC} = 0.0188$ ($P < 0.0001$)	-0.34	$\Phi_{SC} = -0.0035$ ($P = 0.9901$)
	Within pops	97.55	$\Phi_{ST} = 0.0246$ ($P < 0.0001$)	96.92	$\Phi_{ST} = 0.0308$ ($P < 0.0001$)	99.17	$\Phi_{ST} = 0.0083$ ($P = 0.0083$)
West v. East of Rockies	Among groups	3.46	$\Phi_{CT} = 0.0346$ ($P = 0.0069$)	3.54	$\Phi_{CT} = 0.0354$ ($P < 0.0001$)	3.26	$\Phi_{CT} = 0.0326$ ($P = 0.0059$)
	Among pops	0.35	$\Phi_{SC} = 0.0037$ ($P = 0.2609$)	1.00	$\Phi_{SC} = 0.0104$ ($P < 0.0001$)	-1.15	$\Phi_{SC} = -0.0119$ ($P = 1.0000$)
	Within pops	96.19	$\Phi_{ST} = 0.0381$ ($P < 0.0001$)	95.46	$\Phi_{ST} = 0.0454$ ($P < 0.0001$)	97.89	$\Phi_{ST} = 0.0211$ ($P = 0.2197$)

Fig. 3.1.

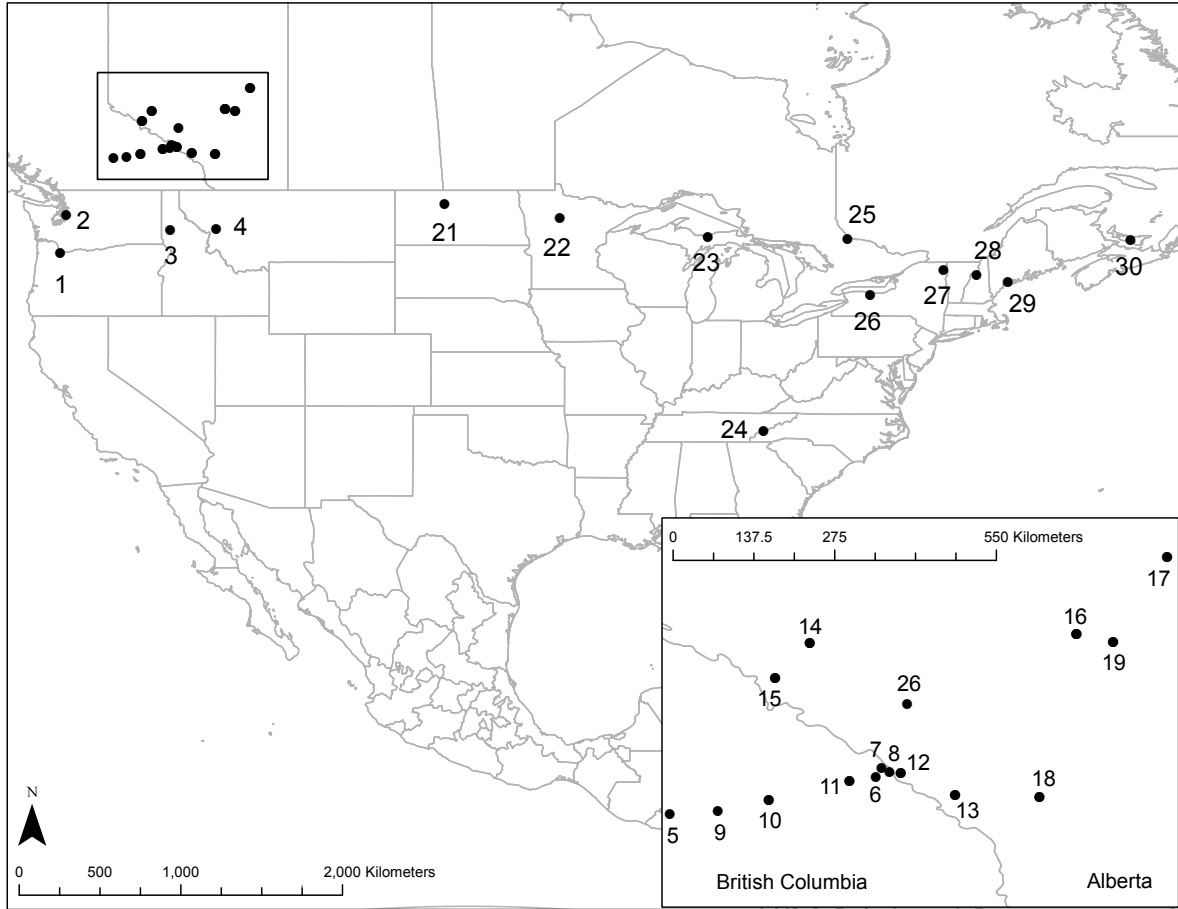


Fig. 3.1. Sampling locations of flies used in this study. Location numbers correspond to numbers in Table 1. Inset shows detail of dense sampling area around the Rocky Mountains

Fig. 3.2.

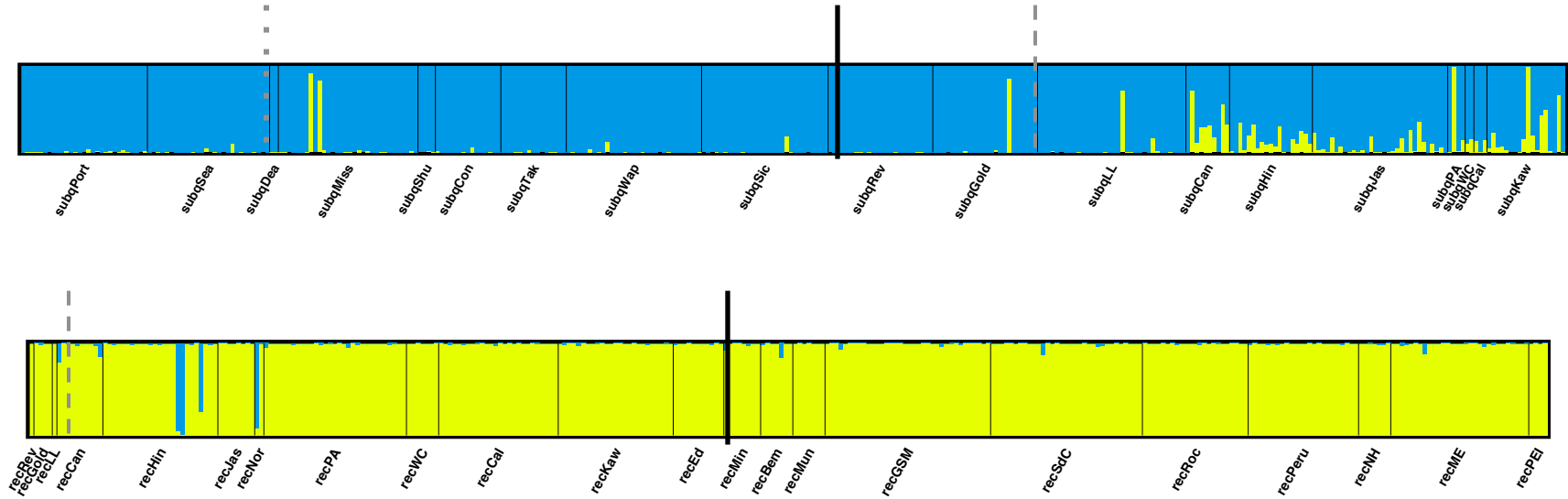


Fig. 3.2. Structure results of individual q values summarized with CLUMPP across runs where $K=2$ using all markers. These runs included both species (all 684 individuals). Gray dotted and dashed lines represent boundaries of the Coast and Rocky Mountains, respectively. Black lines indicate allopatry/sympatry borders. Populations are shown under the plot.

Fig. 3.3.

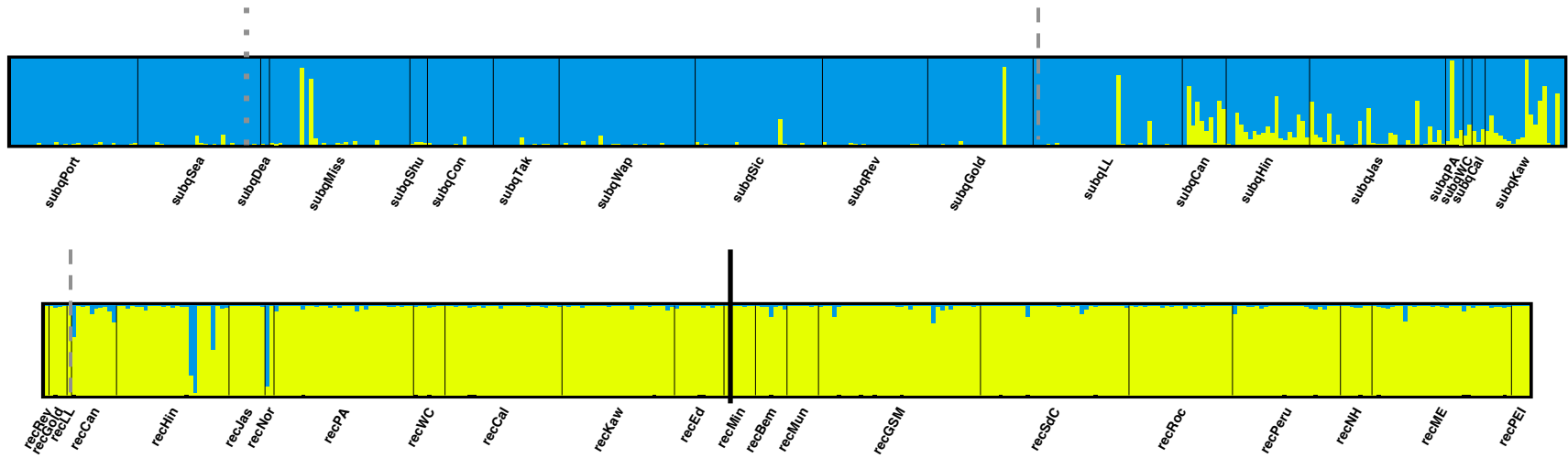


Fig. 3.3. Structure results of individual q values summarized with CLUMPP across runs where $K=2$ using autosomal markers. These runs included both species (all 684 individuals). Gray dotted and dashed lines represent boundaries of the Coast and Rocky Mountains, respectively. Black lines indicate allopatry/sympatry borders. Populations are shown under the plot.

Fig. 3.4.

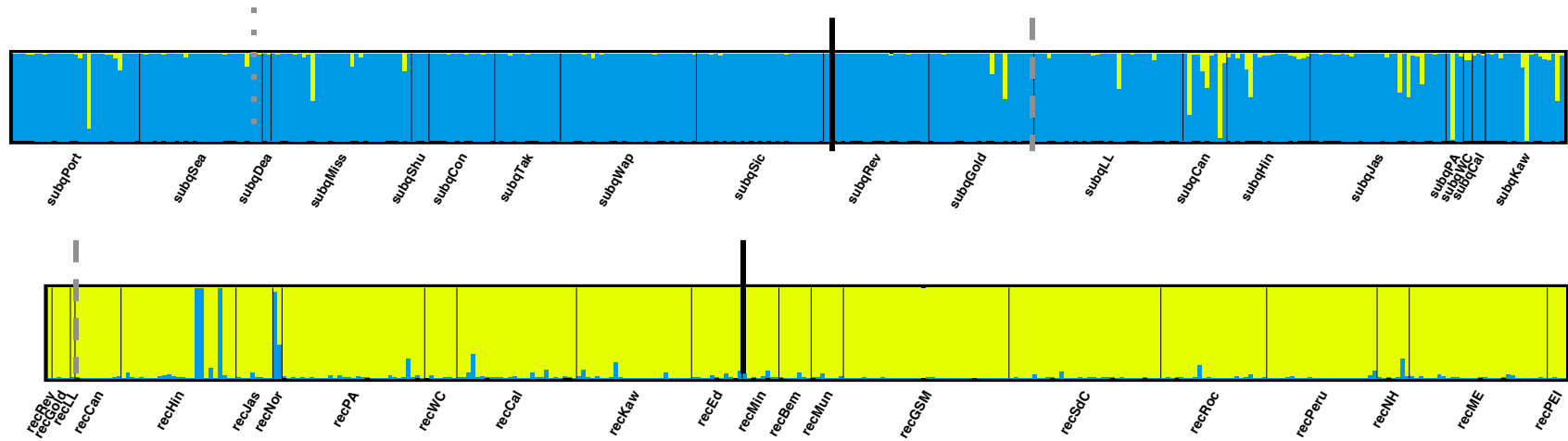
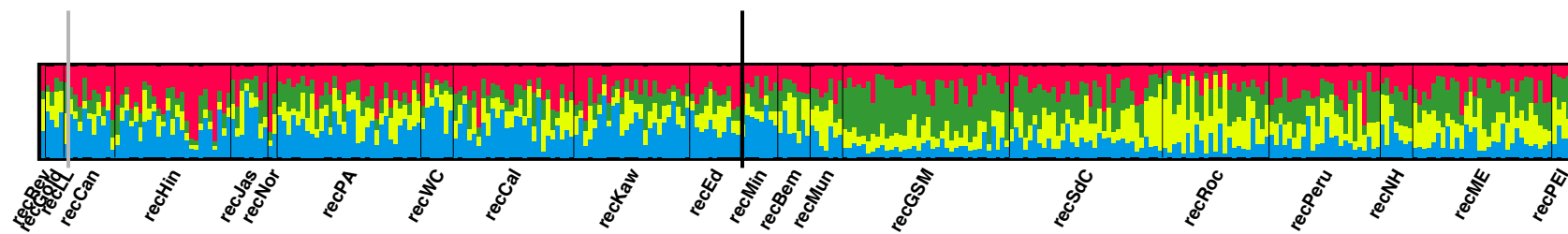


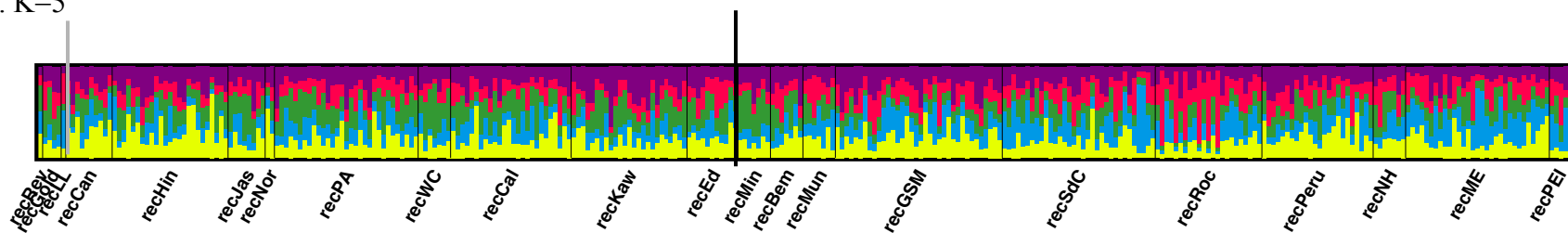
Fig. 3.4. Structure results of individual q values summarized with CLUMPP across runs where $K=2$ using X-linked markers. These runs included both species (all 684 individuals). Gray dotted and dashed lines represent boundaries of the Coast and Rocky Mountains, respectively. Black lines indicate allopatry/sympatry borders. Populations are shown under the plot.

Fig. 3.5.

A. K=4



B. K=5



C. K=4

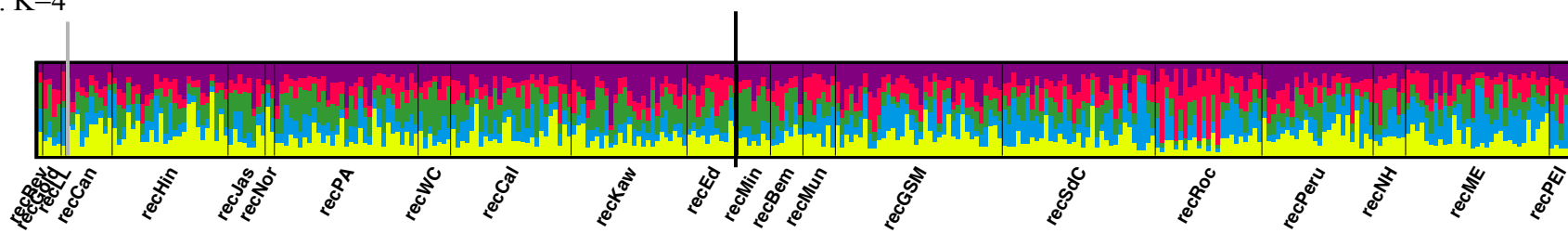
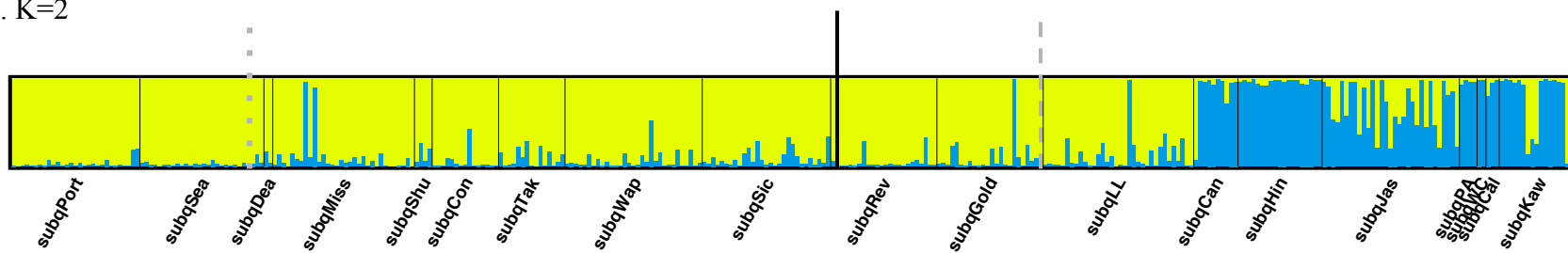


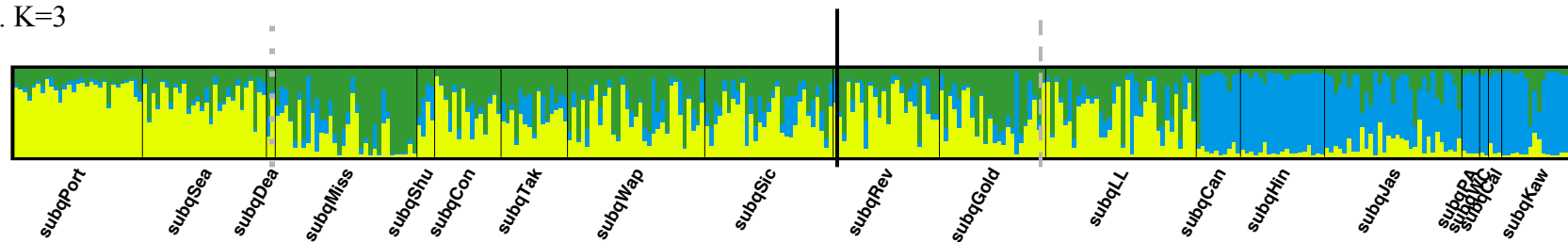
Fig. 3.5. Structure results of individual q values summarized with CLUMPP across runs using only *D. recens* individuals A) Structure results of summarized q -values across runs where $K=4$ using all markers. B) Structure results of summarized q -values across runs where $K=5$ using autosomal markers. C) Structure results of summarized q -values across runs where $K=4$ using X-linked markers. Gray and black lines indicate the Rocky Mountains and the sympatry/allopatry boundary, respectively.

Fig. 3.6.

A. $K=2$



B. $K=3$



C. $K=5$

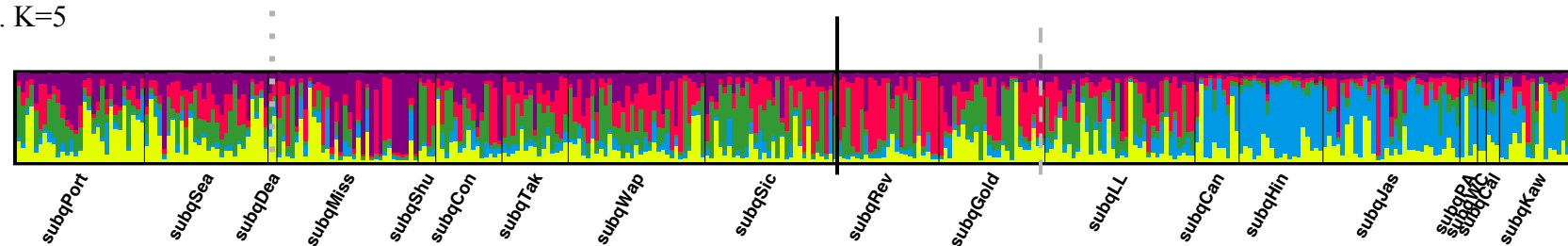
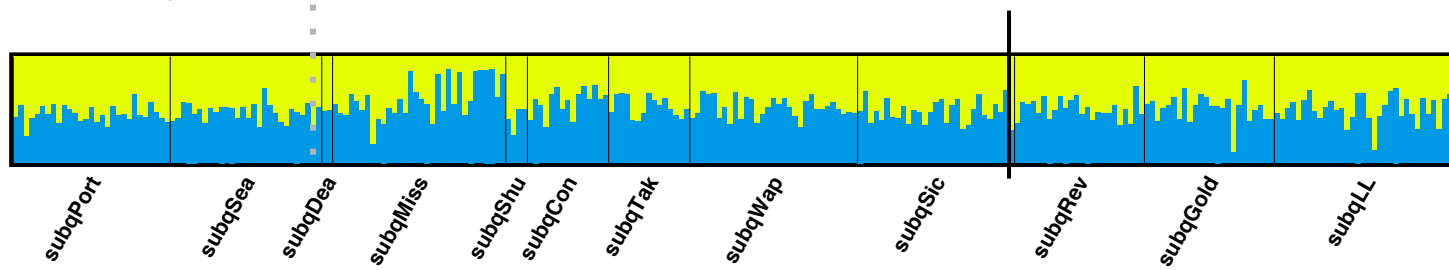


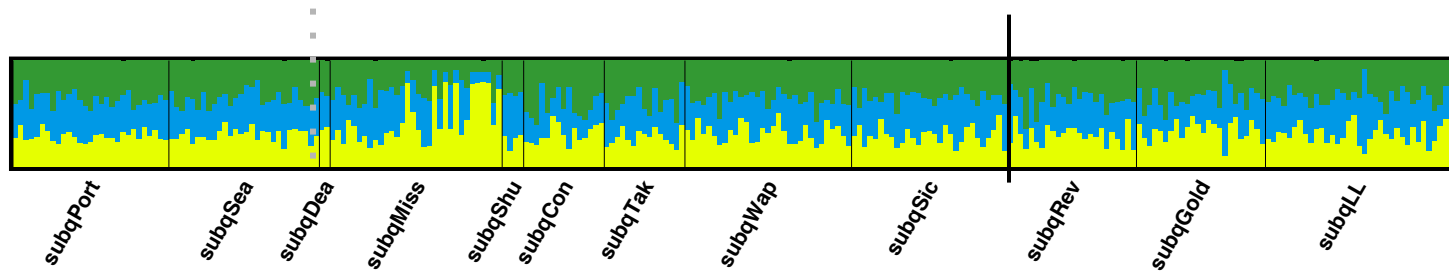
Fig. 3.6. Structure results of individual q values summarized with CLUMPP across runs using only *D. subquinaria* individuals. A) Structure results of summarized q-values across runs where $K=2$ using all markers. B) Structure results of summarized q-values across runs where $K=3$ using autosomal markers. C) Structure results of summarized q-values across runs where $K=5$ using X-linked markers. Gray dotted and dashed lines represent the Coast and Rocky Mountains, respectively. The black lines represent the sympatry/allopatry boundary.

Fig 3.7.

A. All markers, K=2



B. Autosomal markers, K=3



C. X-linked markers, K=4

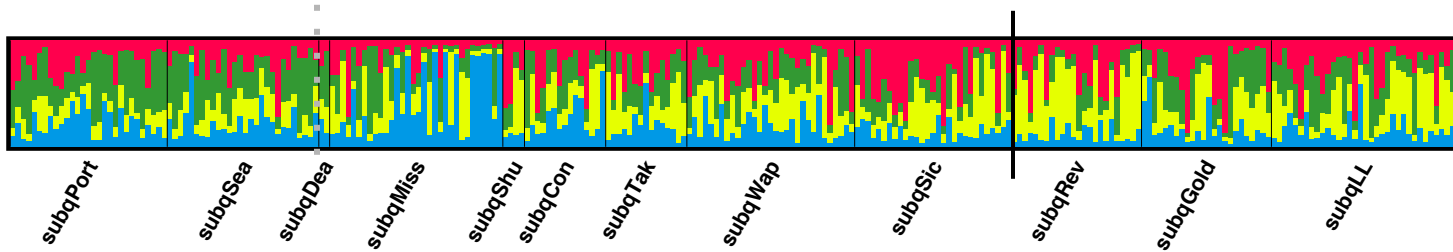


Fig 3.7. Structure results of individual q values summarized with CLUMPP across runs using only *D. subquinaria* individuals from populations with percent of *D. recens* in community < 40%. A) Structure results of summarized q values across runs where $K=2$ using all markers. B) Structure results of summarized q values across runs where $K=3$ using autosomal markers. C) Structure results of summarized q values across runs where $K=4$ using X-linked markers.

CHAPTER 4

THE SAME CHROMOSOMES CONTRIBUTE TO FEMALE MATE
DISCRIMINATION BOTH BETWEEN AND WITHIN SPECIES IN *DROSOPHILA*
SUBQUINARIA

Introduction

When species come back into contact following divergence, there is an opportunity for hybridization. If hybrids are unfit, selection can act to strengthen pre-mating barriers in a process known as reinforcement (Dobzhansky 1940; Noor 1999; Coyne and Orr 2004). There is substantial empirical and theoretical work supporting reinforcement (reviewed in Coyne and Orr 2004), and multiple cases have been observed in nature (Noor 1995; Saetre et al. 1997; Pfennig 2003; Jaenike et al. 2006; Gregorio et al. 2012). One obstacle in the evolution of reinforcement is how alleles that confer stronger premating isolation can spread in a population that is experiencing gene flow from not only another species but also allopatric populations of its own species. Multiple theoretical models have addressed what genetic architectures of reinforced discrimination, in particular for the evolution of female preference, would facilitate its evolution in the face of this gene flow (reviewed in Coyne and Orr 2004). This work suggests that dominance and sex chromosome linkage, especially when post-zygotic isolating factors are sex-linked, promote the evolution of reinforcement (Servedio and Noor 2003; Hall and Kirkpatrick 2006; Lemmon and Kirkpatrick 2006).

To date, the genetic basis of reinforced female discrimination has been mapped in two animal species. First, the pied and collared flycatchers of Europe undergo reinforcement in a hybrid zone in the Czech Republic and Sweden (Saetre and Saether 2010). In this system, female sex chromosomes are ZW and males are ZZ. Introgression and cross-rearing experiments showed that hybrid females mated with the species of males from which they inherited their Z chromosome (Saether et al. 2007). Because females in these species are heterogametic, the preference trait behaves as dominant. Additionally, both post-zygotic isolating factors and male plumage traits map to the Z chromosome (Saetre et al. 2003). Thus in this system it appears that dominance and the linkage of these traits on the Z has facilitated the evolution of reinforcement. Second, in *Drosophila pseudoobscura*, which shares a hybrid zone with *D. persimilis* in California and the Pacific Northwest, *D. pseudoobscura* females show behavior consistent with reinforced female preference (Noor 1995). Again, traits involved in reinforced female preference are dominant and map to a sex chromosome (X), which interacts epistatically with autosomal loci to confer female preference (Ortiz-Barrientos et al. 2004).

In some species undergoing reinforcement, secondary isolation between allopatric and sympatric populations has been identified (Hoskin et al. 2005; Jaenike et al. 2006; Higgie and Blows 2007). This secondary isolation may be a byproduct of selection in sympatry for reinforced traits, and this may be facilitated if the same or linked genes determine mate recognition both between species and within species. This process is described as “cascade reinforcement” or “reproductive character displacement (RCD) speciation” (Ortiz-Barrientos et al. 2009; Hoskin and Higgie 2010). This type of cascade reinforcement has been identified in a few species, for example toads and *Drosophila*

(Hoskin et al. 2005; Jaenike et al. 2006). However, to date, the genetic basis of this second isolation event have not been assessed in any system. If the genes involved in both types of preferences are the same, then we can hold the same predictions (sex-linkage and dominance) for both reinforced and cascade traits.

Complicating this picture is that different sympatric populations of a species may have different responses to reinforcement, as is seen in *D. pseudoobscura* (Ortiz-Barrientos et al. 2004). In this system, different populations have different QTL that contribute to the same phenotype of reinforced female mating discrimination. The same could be true of sexual isolation within species, i.e. there is variation in which genes contribute to the same mate choice phenotype. Only genetic dissection of the traits involved in reinforcement will allow us to identify whether different populations share a genetic response to reinforcement (Ortiz-Barrientos et al. 2004).

In a North American species of fruit fly, *Drosophila subquinaria*, patterns of both reinforcement and cascade reinforcement have been identified (Jaenike et al. 2006). *Drosophila subquinaria* shares a hybrid zone east of the Rocky Mountains with its sister species, *D. recens*. Because of a *Wolbachia* infection in *D. recens*, ~98% of embryos die when a *D. subquinaria* female mates with a *D. recens* male. Consistent with selection to avoid mating with heterospecific males, Jaenike et al. (2006) showed that sympatric *D. subquinaria* females reject *D. recens* males while allopatric *D. subquinaria* females mate with them at low rates. These sympatric females also reject allopatric conspecifics, suggesting that cascade reinforcement may be in progress. We showed in Bewick and Dyer (in review) that three sympatric populations (Canmore, Alberta; Hinton, Alberta; and Kawtikh, Alberta) are fixed in rejection of heterospecific males (i.e. no females

mated with *D. recens* males in 689 single-pair mate trials from 25 lines). These females also reject allopatric males, with mating rates between 0% and 17% in 2hr. Also, though these mating rates are low, there is significant variation among lines in mating rates with coastal allopatric males.

Here we assess the broad-scale genetic architecture of both reinforced discrimination between species and sexual isolation within species in each of three sympatric populations of *D. subquinaria*. Specifically, we first ask whether female discrimination is dominant or recessive, and then we use a classical *Drosophila* whole-chromosome substitution approach to ask which chromosomes underlie these discrimination behaviors. With these data we ask whether the same chromosomes are involved in conferring isolation both between and within species, and whether there is variation in which chromosomes contribute across lines.

Materials and Methods

All flies were maintained at 20°C on a 12hr light/dark cycle on Instant *Drosophila* Medium (Carolina Biological, Cary NC) supplemented with white mushrooms (*Agaricus bisporus*). All of our crosses involved virgin males and females; virgins were collected under light CO₂ anesthesia and maintained 10-15 flies/vial. We used one line from each of three sympatric populations: Canmore, AB (Can12), Hinton, AB (Hin25), Kawtikh, AB (Kaw12). The allopatric line used for each cross came from Seattle, WA (Sea5). These sympatric populations are located to the east of the Rocky Mountains and differ in the percent of *D. recens* present (Fig. 4.1). Tester males in mate discrimination trials were from mixed stocks. *D. recens* males came from a naturally *Wolbachia* uninfected *D.*

recens mixed stock composed of five isofemale lines. Sympatric *D. subquinaria* discriminate against uninfected *D. recens* males the same as they do infected *D. recens* males (Dyer et al. 2014), but using uninfected flies allows us to leave flies together for longer and then assay for the presence of larvae to determine if a mating occurred. Coastal *D. subquinaria* males came from a mixed stock of ten isofemale lines used previously in chapter one.

We first determined whether female discrimination was dominant or recessive. We used single parent crosses of sympatric males (discrimination is fixed in sympatry, so males carry the discrimination allele(s)) with allopatric females to generate F1 females. We screened F1 females for mating rates with *D. recens* males and coastal allopatric males (average N = 27, N $\geq$ 20 per F1 cross / test male type). We asked whether these F1 females mate with coastal allopatric *D. subquinaria* and *D. recens*, and did these mate trials along side control lines, which consisted of parental strains and coastal *D. subquinaria* and *D. recens* females (average N = 47, N $\geq$ 20 for each parent or control/test male type). In these and all mate trials, males were added by mouth aspiration to 1-dram vials with blended mushroom agar food the day before a trial, and females were added within an hour of lights on. We used a 1:3 female:male sex ratio in mating trials, thus females have a choice of three potential male mates. Vials were observed for two hours, and the beginning and ending of copulation were recorded. Females were left with all males for an additional 24hrs, after which males were removed. Females were allowed to oviposit for three additional days, and we checked each vial for the presence of larvae a week later.

As the results will show, female discrimination is dominant, and thus to dissect the genetic basis of these traits we backcrossed sympatric chromosomes from our three sympatric lines (Canmore, Hinton, Kawtikh) into the Seattle, WA allopatric background. Because recombination does not occur in *Drosophila* males, backcrossing through F1 male offspring allows the movement of whole chromosomes into a genomic background. Each parental cross was a sympatric male by allopatric female (Fig. 4.1), since sympatric female by allopatric male crosses proved to be too unreliable to generate a large enough backcross population to screen (5:1 male to female ratio left to mate for a week generated success rates < 5%). Half of our F1 males were backcrossed to allopatric parental females, and the daughters of this cross carry two allopatric alleles (allo/allo) or one allopatric and one sympatric allele (allo/sym) at each autosomal locus (Fig. 4.1). Because all daughters of this cross are homozygous allo/allo at the X-chromosome, in order to screen for a role of the X chromosome, we crossed the other half of the F1 males to sympatric females and then crossed their sons to allopatric females (Fig. 4.1A and B, respectively). In this way, the autosomes are the same composition as the first set of crosses (allo/allo or allo/sym) and all the X-chromosomes are allo/sym (Fig. 4.1B). Crosses were timed so that BC1 females of cross scheme A (Fig. 4.1A) emerged at the same time as BC2 females of cross scheme B (Fig. 4.1B). Parents of each cross were frozen for later genetic analyses.

Backcrossed females (total N=2304) were screened in mate trials against either *D. recens* males or *D. subquinaria* males, following the scheme of our F1 trials. A mold infection in stocks kept some parents from being screened in every trial; however using a logistic regression of each parental or control type with trial and male as fixed effects

showed no difference across trials in mating rates of all control or parental females (all $P > 0.23$), so environmental variation is negligible. Backcross females were frozen for genetic analyses.

The parents of each cross and the backcross females were genotyped at each of the four major autosomes. *D. subquinaria* has the ancestral *Drosophila* karyotype and thus six chromosomes: an X chromosome (also known as Muller Element A), four major autosomes (Muller elements B through E), and a non-recombining dot chromosome (Muller Element F), which we did not consider here (Patterson 1952). The genotype of the X chromosomes is known from the direction of the backcross, so markers were not necessary for this chromosome. In order to identify microsatellite markers on each of the major autosomes, we created an F2 mapping population of 92 individuals using one Can12 male and one Sea5 female. We first sequenced a gene from each major autosome (there is high conservation of chromosomal content in *Drosophila* species (Schaeffer et al. 2008)) for each of the parents of this cross, and identified parent specific restriction digest cut-sites. At each gene we performed restriction enzyme digests on the parent specific SNPs on the 92 F2 females. To link microsatellite markers to these genes, we genotyped the F2 population at each of nine polymorphic microsatellites. We amplified microsatellites following the protocol of Miles et al. (2008) using the three-primer fluorescent method described in Shuelke (2000). Amplified products were run alongside a size standard (GGF Rox 500) at the Georgia Genomics Facility. Alleles were scored and confirmed by eye using GeneMarker software (SoftGenetics, State College, PA). Finally, we used the program MapDisto (Lorieux 2012) to identify linkage groups, using a minimum LOD of 3 and an r_{max} of 0.3. We chose four microsatellites to determine

chromosome origin in our backcrossing experiment, including one associated with each major autosome that had potentially fixed differences between allopatric and sympatric populations. While for most crosses we used this same set of four microsatellites, but in one series of crosses for Canmore, shared alleles with Seattle made it necessary to genotype an additional microsatellite marker to verify chromosomal origins in backcross females.

Since the only possible genotypes at any locus are allo/allo or sym/allo, genotypes for each female were scored as either allo/allo (“A”) or sym/allo (“S”) corresponding to whether a sympatric or allopatric allele was present. This binomial S/A notation was used to denote presence/absence of a sympatric allele. These data were then combined with mating data and used to test whether the presence of a sympatric allele affected mating. We used contingency tests and Pearson’s X^2 statistic to test for significant differences in proportion mated among groups of females. We used a nominal logistic regression to test for significant effects of each chromosome as well as potential interactions in BC females, beginning with a model where each chromosome is a fixed effect, and followed by models testing interactions where they were suspected.

Results

F1 mating rates

We found that F1 females rejected mating with *D. recens* males and thus behaved like sympatric females, suggesting that reinforced discrimination has a dominant genetic basis. Specifically, during 2 hr mate trials with *D. recens* males, none of 92 F1 females from across populations mated. In all crosses, F1 female mating rates with *D. recens*

males differed significantly from allopatric parents (Fisher's exact test, all $P < 0.0001$) since allopatric females mate with *D. recens* in 2hr at a rate of 62.5% but sympatric females also do not mate at all (Fig. 4.2).

In contrast, mate discrimination within species appears to be more complicated. While Kawtikh F1 mating rates in 2hr with coastal allopatric males do not differ from sympatric mating rates (0/28), mating rates of F1 females with a Hinton or Canmore parent are higher. F1 females from Hinton mate in 2hr at an average rate of 24%, which differs significantly from both sympatric females (0%; Fisher's exact test, $P = 0.0247$) and allopatric females (84%; Fisher's exact test, $P < 0.0001$) (Fig. 4.2). Half of Canmore F1 females mated in 2hr with coastal males, which differed significantly from both pure sympatric females (Fisher's exact test, $P < 0.0001$) and allopatric females (Fisher's exact test, $P = 0.0045$). A binomial test indicates that this rate does not differ significantly from the expected rate of 42% if the trait were completely additive ($P = 1.0$). F1 female mating rates with coastal males are significantly different across populations (Pearson's $\chi^2 = 10.45$, $P = 0.0054$), which may indicate that the genetic basis of this trait is different among these populations.

In general, backcross females (hereafter "BC") mate at low rates with *D. recens*, but mate at a much higher rate with coastal *D. subquinaria* males. Within the Canmore BCs, females mate at 50.1% with coastal males and 9.6% with recens males (Fig. 4.2). Within the Hinton BCs, females mate at 41.3% with coastal males and 6.0% with recens males (Fig. 4.2). Within the Kawtikh BC population, females mate at 33.8% with coastal males and 4.5% with *D. recens* males (Fig. 4.2). These mating rates are significantly

different among populations (Pearson's $\chi^2=20.987$, $P<0.0001$ for mating with coastal males; Pearson's $\chi^2=8.972$, $P=0.113$ for mating with *D. recens* males; Fig. 4.2).

Effect of the X chromosome on mating rates

Overall, across sympatric populations it appears that the X chromosome has an effect on 2hr mating rates with both *D. recens* and coastal allopatric males. When we consider only the X chromosome sympatric/allopatric status (disregarding the status of other chromosomes), this decrease is significant for Hinton and Kawtikh BCs mating with coastal males ($\chi^2=8.336$, $P=0.0039$ and $\chi^2=19.037$, $P<0.0001$, respectively), and though there is a decrease in mating for Canmore backcross females with sympatric X's, this decrease is not significant ($\chi^2=1.798$, $P=0.1799$) (Fig. 4.3). For mating with *D. recens* males, the effect of the X chromosome is significant for Hinton and Canmore BC females ($\chi^2=11.4267$, $P=0.0007$, $\chi^2=8.83712$, $P=0.003$), but it is not significant in Kawtikh BC females ($\chi^2=0.6778$, $P=0.4104$) (Fig. 4.3). These trends hold when we increase the potential mating time from 2 to 26hrs, where we again see a general decrease in female mating rates when the X chromosome is sympatric in origin (results not shown).

Effect of autosomes and interactions with X chromosome

We used a nominal logistic regression with each population of backcross females and each male type to identify significant effects of individual chromosomes on mating rates. Unlike our analysis of X chromosome effect discussed above, we only used backcross females for which we have complete genotype information for all four

autosomes, and due to missing data for some failed genotypes our power is lower (total sample sizes of 437, 411, 547 for Canmore, Hinton, and Kawtikh respectively). First, in Hinton and Kawtikh, there is a significant effect of the chromosome that corresponds to the C Muller element in 2hr mating rates with coastal allopatric males ($X^2=13.5$, $P=0.0002$; $X^2=11.11$, $P=0.0009$, respectively, Table 4.2, Fig. 4.3), and this chromosome also has an effect on mating rates with *D. recens* males ($X^2=12.62$, $P=0.0004$; $X^2=4.92$, $P=0.0265$, respectively, Table 4.2). The effect of the X chromosome is significant in Hinton and Kawtikh for mating rates with coastal allopatric males ($X^2=5.98$, $P=0.0145$; $X^2=7.50$, $P=0.0062$, respectively, Table 4.2), but this effect disappears in mating with *D. recens* males (Table 4.2). This may be due to the reduced power of this dataset, along with the overall low percentage of backcross mating to *D. recens* males. Overall this suggests that in these two populations, the same chromosome (Muller element C) is contributing to discrimination both between and within species.

The chromosomes contributing to mating behavior in Canmore BC females differ from those contributing to discrimination in Hinton and Kawtikh populations. First, the X chromosome does not appear to contribute to either between or within species discrimination. The C element again has a significant effect on mating with coastal allopatric males in 2hr ($X^2=7.06$, $P=0.0079$), and its effect is marginally significant in 2hr mating with *D. recens* males ($X^2=3.60$, $P=0.0577$) (Table 4.2, Fig. 4.3). Interestingly, the E element shares significant effects on mating with both coastal allopatric males and *D. recens* males ($X^2=4.49$, $P=0.0341$; $X^2=5.11$, $P=0.0238$) (Table 4.2, Fig. 4.3). This indicates that in this population the same chromosome does contribute to both types of discrimination, however it appears to be a different chromosome than in Hinton and

Kawtikh. The D element also has a significant effect on mating rates with *D. recens* ($X^2=4.52$, $P=0.0238$) (Table 4.2, Fig. 4.3).

We also tested interactions among the chromosomes indicated as having a significant effect. In Kawtikh BC females, the interaction between the X chromosome and C element has a significant effect on mating with coastal allopatric males ($X^2=5.37$, $P=0.0204$); however, that interaction is not significant in Hinton BC females ($X^2=0.46$, $P=0.4969$) (Table 4.2).

Discussion

When species come into contact in a hybrid zone, selection can act to increase premating barriers to prevent maladaptive hybridization. The genetic basis of this reinforcement has been mapped in few systems (Ortiz-Barrientos et al. 2004; Saether et al. 2007; Hopkins and Rausher 2011), and here we add to this knowledge by identifying chromosomes with a large effect on reinforced mating discrimination in a species undergoing reinforcement. We also ask, for the first time in any system, whether these same chromosomes affect discrimination within species. Finally, we ask whether there is variation among lines in the genetic basis of these behaviors. We find evidence that reinforced discrimination between species is dominant for all lines, and that within species discrimination is dominant for two of three lines. Additionally, we find that discrimination within and between species are driven by the same chromosomes, though the identity of these chromosomes appears to be line-dependent. In sum, our results are consistent with the hypothesis that chromosomal linkage of discrimination within and

between species has facilitated divergence within species as a result of reinforcing selection.

First, we find that female discrimination against *D. recens* males is dominant. Regardless of sympatric population origin, all F1 females reject *D. recens* males completely (Fig. 4.2) just as pure sympatric females do. This result follows the prediction of “Haldane’s sieve” that adaptive mutations that are dominant will be more likely to spread in a population since the benefit of the adaptation will not be hidden in heterozygotes (Noor 1999). On the other hand, the dominance of discrimination against coastal allopatric males is line-dependent. F1 females from Kawtikh completely reject coastal allopatric males, but F1 females from Hinton and Canmore mate at low rates with coastal allopatric females (Fig. 4.2). However these rates are significantly different from the rates of both the sympatric and allopatric parents, suggesting that this trait is not strictly dominant and may be more complex. The differences in mating rates among these populations when interacting with coastal allopatric males may indicate that the genetic basis for discrimination within species is more complex than discrimination between species, or that the genetic basis of this discrimination is different among these populations, which has been seen in reinforced preference and could hold true for discrimination within species (Ortiz-Barrientos et al. 2004).

Empirical and theoretical work suggests that sex chromosomes may be involved in reproductive isolation and reinforcement (Saetre et al. 2003; Servedio and Noor 2003; Coyne and Orr 2004; Ortiz-Barrientos et al. 2004; Servedio 2004; Saether et al. 2007), and our own population genetic work suggested that this chromosome may be particularly resistant to introgression both between and within species. We therefore

focused initially on what effect the X chromosome has on mating rates. When we ignore the status of other chromosomes, we see that the X chromosome has a significant effect on mating rates with *D. recens*, though this is not significant for every line. When we consider a smaller dataset with complete genotypes for all chromosomes, the effect of the X chromosome disappears, and we see an effect of the chromosome that corresponds to the C Muller element in the lines from Hinton and Kawtikh (Table 4.2). However, the C has no effect in Canmore populations, but D and E do (Table 4.2). Again, this supports earlier work that suggests that the genetic basis of reinforced traits can differ within a species undergoing reinforcement (Ortiz-Barrientos et al. 2004). The involvement of the X chromosome in discrimination between species follows expectations that sex chromosomes are disproportionately involved in reproductive isolation (reviewed in Coyne and Orr 2004; Presgraves 2008), which has been confirmed in other cases of reinforcement (Saetre et al. 2003; Ortiz-Barrientos et al. 2004; Saether et al. 2007). Additionally, as seen in another case of reinforcement in *Drosophila*, the X chromosome plays an important role in reinforcement, but does not act alone (Ortiz-Barrientos et al. 2004), and in our case, it may not have the strongest effect on mating compared to the C element. These results suggest that more than one locus is required for reinforcement to proceed in *D. subquinaria*.

The presence of sympatric X chromosome also reduces backcross female mating rates with coastal allopatric males in some lines. The X chromosome, along with the C element, has a significant effect on mating rates with coastal males for BC females from Hinton and Kawtikh lines. However in Canmore, the effect of the X disappears, though the effect of the C element remains. Additionally, the E element contributes significantly

to variation in mating rates with coastal allopatric males for these females. This indicates that much like the genetic basis of reinforced female preference, the genetic basis of cascade preference may vary in the populations undergoing cascade reinforcement.

A major question underlying cascade reinforcement is whether the same genes underlie discrimination within and between species. We find evidence that the same chromosomes may underlie discrimination both between and within species, though importantly, these chromosomes may vary among lines. Within lines, we find that one chromosome is involved in both discriminations: C in Hinton and Kawtikh and E in Canmore. The line from Canmore is an interesting case compared to discrimination in other lines. Within species discrimination in Canmore females is more complex and reinforced discrimination involves different chromosomes, suggesting that responses to reinforcement are different in different populations (Ortiz-Barrientos et al. 2004). In particular we see that the C and E elements have effects on mating rates with coastal allopatric males, and D and E affect mating with *D. recens* males. The significant effect of the E element on mating with both male types, again supports the hypothesis that the same chromosome and potentially the same genes are involved in both discriminations. However, the involvement of two different chromosomes (D and E) in determining mating also suggests that variation exists among populations in the genetic basis of reinforced discrimination and discrimination within species.

There are two main avenues to move forward from this work. First, while our results support the hypothesis that the same genes may confer both discrimination within and between species, fine-mapping will confirm if the same or linked genes are involved in determining each discrimination phenotype. Based on work by Giglio and Dyer (2013)

and Dyer et al. (2014), we have a reasonable expectation that sympatric *D. subquinaria* females use cuticular hydrocarbons as mating cues. Sympatric *D. subquinaria* females mate much less frequently when they cannot smell, and cuticular hydrocarbon profiles are distinct among populations of *D. subquinaria* and between *D. subquinaria* and *D. recens*. Fine-mapping may reveal olfactory genes are involved in mate discrimination. Also, since this assay is only focused on a few lines, it would be worth assessing how much variation exists in the genetic basis of these discriminations at a large scale across more lines and populations.

Second, while we have identified which chromosomes contribute the most to female discrimination, we do not know what chromosomes contribute to variation in the male trait that females use as a mating cue. Reinforcement in flycatchers may have been facilitated by such linkage with both female preference and male plumage trait residing on the Z chromosome (Saether et al. 2007). A similar assessment of variation in backcross male mating success would answer whether such linkage is occurring in this system.

To our knowledge the genetic basis of cascade reinforcement has not been mapped in any system, and with this work we have support that the genetic basis of discrimination between and within species may be the same, and fine-mapping and functional assays will confirm whether this is the case. Much like another *Drosophila* reinforcement system (Ortiz-Barrientos et al. 2004), we find evidence that the genetic basis of reinforced discrimination may vary within sympatric populations. We find similar evidence that the genetic basis of cascade reinforcement may vary among lines as well. Ultimately it appears that there can be multiple genetic responses to reinforcing

selection in sympatry, and though these responses differ, they may determine the genetic basis of divergence within species.

Literature Cited

- Coyne, J. A. and H. A. Orr. 2004. Speciation. Sinauer Associates, Sunderland (Massachusetts).
- Dobzhansky, T. 1940. Speciation as a stage in evolutionary divergence. *Am. Nat.* 74:312-321.
- Dyer, K. A., B. E. White, J. L. Sztepanacz, E. R. Bewick, and H. D. Rundle. 2014. Reproductive character displacement of epicuticular compounds and their contribution to mate choice in *Drosophila subquinaria* and *D. recens*. *Evolution* 68:1163–1175.
- Giglio, E. M. and K. A. Dyer. 2013. Divergence of premating behaviors in the closely related species *Drosophila subquinaria* and *D. recens*. *Ecol. Evol.* 3:365-374.
- Gregorio, O., E. L. Berdan, G. M. Kozak, and R. C. Fuller. 2012. Reinforcement of male mate preferences in sympatric killifish species *Lucania goodei* and *Lucania parva*. *Behav. Ecol. Sociobiol.* 66:1429-1436.
- Hall, D. W. and M. Kirkpatrick. 2006. Reinforcement and sex linkage. *Evolution* 60:908-921.
- Higgie, M. and M. W. Blows. 2007. Are traits that experience reinforcement also under sexual selection? *Am. Nat.* 170:409-420.
- Hopkins, R. and M. D. Rausher. 2011. Identification of two genes causing reinforcement in the Texas wildflower *Phlox drummondii*. *Nature* 469:411-414.
- Hoskin, C. J. and M. Higgie. 2010. Speciation via species interactions: the divergence of mating traits within species. *Ecol. Lett.* 13:409-420.

- Hoskin, C. J., M. Higgie, K. R. McDonald, and C. Moritz. 2005. Reinforcement drives rapid allopatric speciation. *Nature* 437:1353-1356.
- Jaenike, J., K. A. Dyer, C. Cornish, and M. S. Minhas. 2006. Asymmetrical reinforcement and *Wolbachia* infection in *Drosophila*. *PLoS Biol.* 4:1852-1862.
- Lemmon, A. R. and M. Kirkpatrick. 2006. Reinforcement and the genetics of hybrid incompatibilities. *Genetics* 173:1145-1155.
- Lorieux, M. 2012. MapDisto: fast and efficient computation of genetic linkage maps. *Mol. Breeding* 30:1231-1235.
- Miles, L. G., S. R. Isberg, C. Moran, C. Hagen, and T. C. Glenn. 2008. 253 Novel polymorphic microsatellites for the saltwater crocodile (*Crocodylus porosus*). *Conserv. Genet.* 10:963-980.
- Noor, M. A. 1995. Reinforcement in speciation. *Trends Ecol. Evol.* 10:492.
- Noor, M. A. F. 1999. Reinforcement and other consequences of sympatry. *Heredity* 83:503-508.
- Ortiz-Barrientos, D., B. A. Counterman, and M. A. F. Noor. 2004. The genetics of speciation by reinforcement. *PLoS Biol.* 2:2256-2263.
- Ortiz-Barrientos, D., A. Grealy, and P. Nosil. 2009. The genetics and ecology of reinforcement: implications for the evolution of prezygotic isolation in sympatry and beyond. Pp. 156-182 in C. D. Schlichting, and T. A. Mousseau, eds. *Year in Evolutionary Biology* 2009. Wiley-Blackwell, Malden.
- Patterson, J. T. 1952. *Evolution in the genus Drosophila*. Macmillan, New York,.

- Pfennig, K. S. 2003. A test of alternative hypotheses for the evolution of reproductive isolation between spadefoot toads: support for the reinforcement hypothesis. *Evolution* 57:2842-2851.
- Presgraves, D. C. 2008. Sex chromosomes and speciation in *Drosophila*. *Trends Genet.* 24:336-343.
- Saether, S. A., G. P. Saetre, T. Borge, C. Wiley, N. Svedin, G. Andersson, T. Veen, J. Haavie, M. R. Servedio, S. Bures, M. Kral, M. B. Hjernquist, L. Gustafsson, J. Traff, and A. Qvarnstrom. 2007. Sex chromosome-linked species recognition and evolution of reproductive isolation in flycatchers. *Science* 318:95-97.
- Saetre, G. P., T. Borge, K. Lindroos, J. Haavie, B. C. Sheldon, C. Primmer, and A. C. Syvanen. 2003. Sex chromosome evolution and speciation in *Ficedula* flycatchers. *P. Roy. Soc. B-Biol. Sci.* 270:53-59.
- Saetre, G. P., T. Moum, S. Bures, M. Kral, M. Adamjan, and J. Moreno. 1997. A sexually selected character displacement in flycatchers reinforces premating isolation. *Nature* 387:589-592.
- Saetre, G. P. and S. A. Saether. 2010. Ecology and genetics of speciation in *Ficedula* flycatchers. *Mol. Ecol.* 19:1091-1106.
- Schaeffer, S. W., A. Bhutkar, B. F. McAllister, M. Matsuda, L. M. Matzkin, P. M. O'Grady, C. Rohde, V. L. Valente, M. Aguade, W. W. Anderson, K. Edwards, A. C. Garcia, J. Goodman, J. Hartigan, E. Kataoka, R. T. Lapoint, E. R. Lozovsky, C. A. Machado, M. A. Noor, M. Papaceit, L. K. Reed, S. Richards, T. T. Rieger, S. M. Russo, H. Sato, C. Segarra, D. R. Smith, T. F. Smith, V. Strelets, Y. N. Tobari, Y. Tomimura, M. Wasserman, T. Watts, R. Wilson, K. Yoshida, T. A.

- Markow, W. M. Gelbart, and T. C. Kaufman. 2008. Polytene chromosomal maps of 11 *Drosophila* species: the order of genomic scaffolds inferred from genetic and physical maps. *Genetics* 179:1601-1655.
- Servedio, M. R. 2004. The evolution of premating isolation: local adaptation and natural and sexual selection against hybrids. *Evolution* 58:913-924.
- Servedio, M. R. and M. A. F. Noor. 2003. The role of reinforcement in speciation: theory and data. *Annu. Rev. Ecol. Evol. S.* 34:339-364.
- Shuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18:233-234.

Table 4.1. Genes and chromosomal markers associated with the autosomal Muller elements. Genes are listed first in italics, followed by microsatellite markers.

Element			
B	C	D	E
<i>black</i>	<i>mag</i>	<i>pale</i>	<i>ebony</i>
Dtra31	1008	1024	3091
	1079	1084	
	Dtra40	3038	

Table 4.2. Results of nominal logistic regression of effect of chromosomes on mating in 2hr using only backcross females for which complete genotypes were available. X^2 values and P values of effect likelihood ratio tests are reported. Values significant at $P<0.05$ are in bold, values significant after Bonferroni correction ($P<0.0014$) are italicized.

Male type	Model/ Muller element	Backcross female sympatric source population		
		Canmore	Hinton	Kawtikh
Coastal	Whole model (df=5)	14.36 <i>(P=0.0135)</i>	20.57 <i>(P=0.0010)</i>	25.66 <i>(P=0.0001)</i>
	A (X chromosome) (df=1)	0.51 $P=0.4743$	5.98 <i>P=0.0145</i>	7.50 <i>P=0.0062</i>
	B (df=1)	0.76 $P=0.3832$	0.03 $P=0.813$	2.08 $P=0.1497$
	C (df=1)	7.06 <i>P=0.0079</i>	13.5 <i>P=0.0002</i>	11.11 <i>P=0.0009</i>
	D (df=1)	0.36 $P=0.5489$	0.001 $P=0.9814$	0.03 $P=0.8588$
	E (df=1)	4.49 <i>P=0.0341</i>	0.9 $P=0.3426$	0.36 $P=0.5512$
<i>D. recens</i>	Whole model (df=5)	13.71 <i>P=0.0175</i>	19.20 <i>P=0.0018</i>	11.95 <i>P=0.0355</i>
	A (X chromosome) (df=1)	0.33 $P=0.5641$	1.87 $P=0.1717$	0.55 $P=0.4574$
	B (df=1)	0.49 $P=0.4827$	0.36 $P=0.5484$	1.54 $P=0.2153$
	C (df=1)	3.60 $P=0.0577$	12.62 <i>P=0.0004</i>	4.92 <i>P=0.0265</i>
	D (df=1)	4.52 <i>P=0.0335</i>	0.62 $P=0.4298$	3.73 $P=0.0534$
	E (df=1)	5.11 <i>P=0.0238</i>	2.97 $P=0.5484$	0.77 $P=0.3803$

Table 4.3. Results of nominal logistic regressions for interactions among chromosomes showing a significant effect on mating in 2hr. X^2 values and P values of effect likelihood ratio tests are reported. Values significant at $P<0.05$ are in bold.

Male type-		Canmore	Hinton	Kawtikh
Coastal allopatric <i>D. subquinaria</i>	Whole Model	-	20.07 $P=0.0002$	28.69 $P<0.0001$
	X	-	6.51 $P=0.0107$	5.41 $P=0.0200$
	C	-	13.66 $P=0.0002$	14.57 $P=0.0001$
	X*C	-	0.46 $P=0.4969$	5.37 $P=0.0204$
<i>D. recens</i>	Whole Model	17.40 $P=0.0006$	-	-
	C	10.74 $P=0.0010$	-	-
	E	6.14 $P=0.0132$	-	-
	CxE	0.03 $P=0.8615$	-	-
	Whole model	13.95 $P=0.0030$	-	-
	D	7.42 $P=0.0064$	-	-
	E	6.98 $P=0.0082$	-	-
	DxE	4.86 $P=0.0276$	-	-

Fig.1.

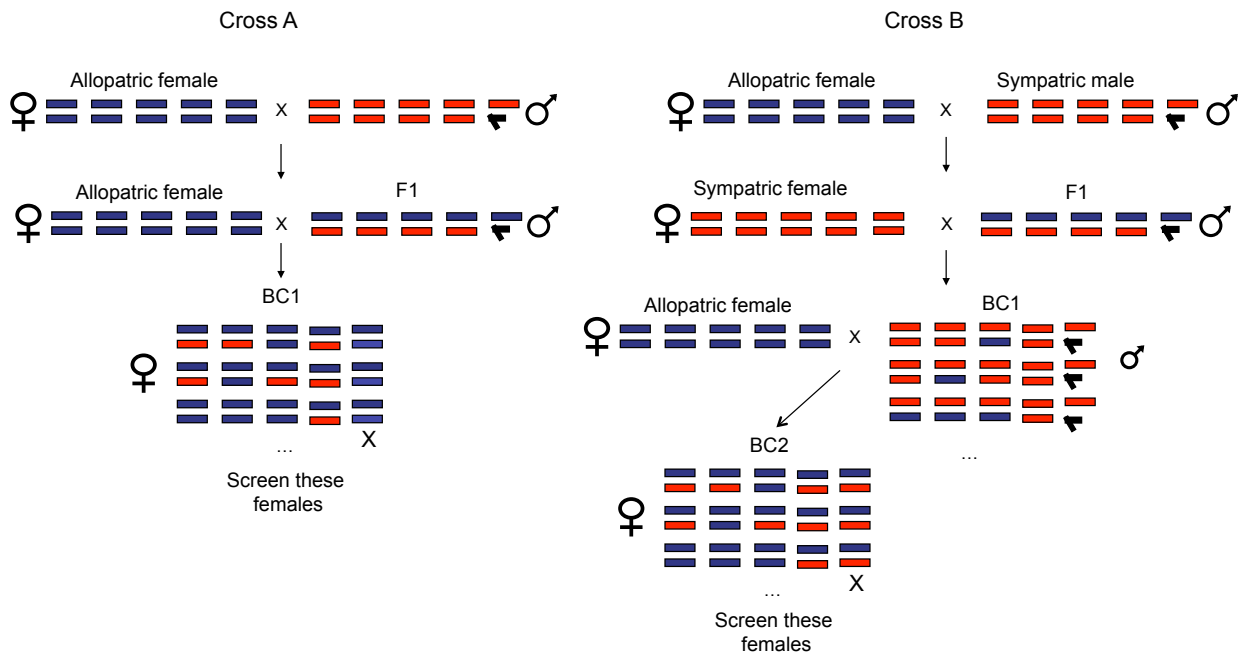


Fig. 4.1. Backcross scheme. Cross A depicts scheme for making backcross females with allopatric X chromosomes. Cross B depicts scheme for making backcross females with one sympatric X chromosome.

Fig. 4.2.

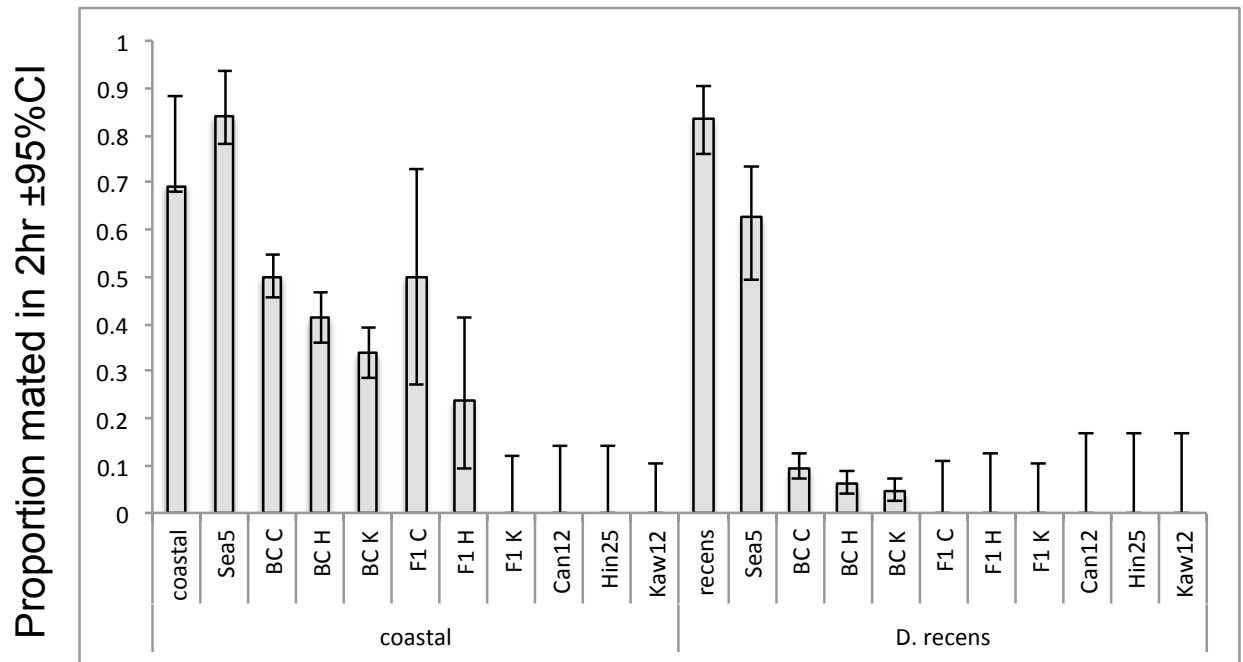


Fig. 4.2. Mating rates in 2hr with both *D. recens* and coastal allopatric *D. subquinaria* males across female types. Mating rates in 2hr shown on Y-axis. Female type within tester male type listed along the X-axis.

Fig. 4.3.

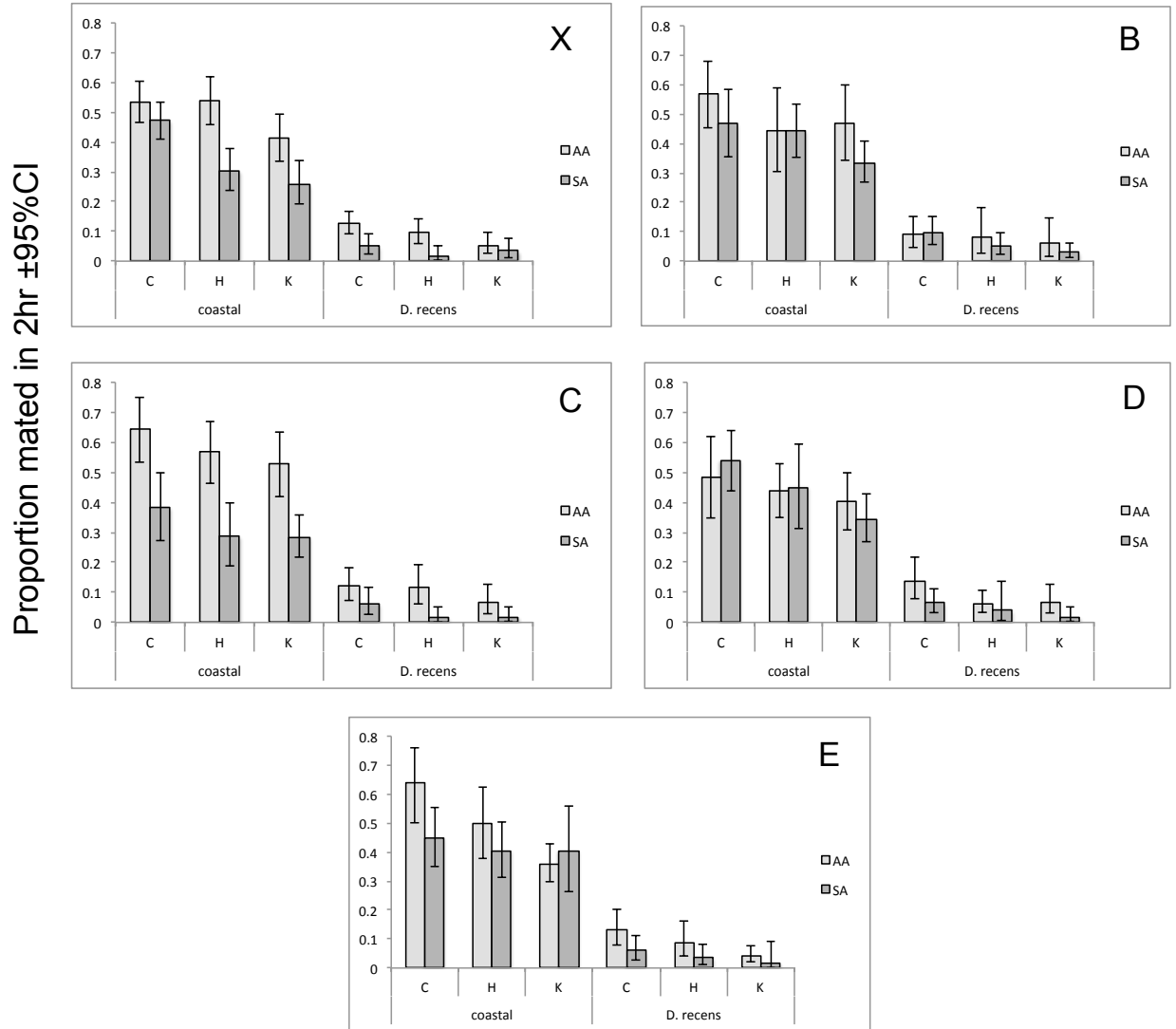


Fig. 4.3. Effect of each chromosome on mating rates in backcross females. Mating rates in 2hr shown on the Y-axis. Backcross female type within male tester type shown along Y-axis. Letter in upper right panel indicates Muller element. Color of bar indicates backcross female genotypes.

CHAPTER 5

CONCLUSIONS AND FUTURE DIRECTIONS

A goal of evolutionary biology is to identify how biodiversity arises and is maintained. Hybrid zones, where two recently diverged species come back into contact and hybridize, provide a window into the earliest stages of speciation. The outcome of this contact will be determined by the conflict between gene flow and selection, and we now understand that selection in this zone can have consequences for divergence outside of this zone (Howard 1993; Coyne and Orr 2004; Ortiz-Barrientos et al. 2009; Hoskin and Higgie 2010). For instance, barriers between species can be reinforced in response to maladaptive hybridization, and reinforcing selection can have consequences for sexual isolation within species as well as between (Hoskin et al. 2005; Jaenike et al. 2006; Hoskin and Higgie 2010). To understand how speciation proceeds in hybrid zones and the consequences of this selection, we must assess what forces and processes promote divergence.

The fruit fly *Drosophila subquinaria* is a good study system to approach these questions. Previous work showed evidence of reinforcement in *D. subquinaria* following secondary contact with its sister species, *D. recens*, while *D. recens* shows no evidence of reinforcement despite suffering a cost to heterospecific mating (Jaenike et al. 2006). Additionally, *D. subquinaria* females from populations sympatric with *D. recens* show discrimination against both *D. recens* males and allopatric conspecifics, while *D. recens*

females show no sexual isolation among populations. In order to fully understand why we see these patterns, we must assess what forces are promoting or hindering divergence in these species. Ultimately, patterns identified here may illuminate general patterns in the evolution of divergence both between and within species, and how they are related to each other.

In this dissertation I set out to examine what role gene flow, selection, and genetic architecture played in the evolution of reinforcement and the consequences of reinforcement in this system. In chapter two, I showed that selection for reinforced discrimination in *D. subquinaria* is strong enough to fix the phenotype in some sympatric populations, and there is some standing variation in allopatric populations for the reinforced discrimination phenotype. Additionally, it appears that populations of *D. subquinaria* are locally adapted, which could be contributing to sexual isolation. The existence of local sexual selection optima supports the theory that sympatric populations may evolve away from allopatric optima in response to reinforcement, and this may be a source of sexual isolation (Higgie and Blows 2007). Among inland allopatric lines where natural recombination occurs between sympatric and allopatric genomes, lines less likely to mate with *D. recens* males were also less likely to mate with coastal *D. subquinaria* males, indicating a potential genetic relationship between the two traits. This supports the idea that reinforcement has led to isolation within species due to a linked genetic basis between discrimination within and between species. Finally, the existence of some allopatric lines with sympatric-like discrimination against *D. recens* may indicate that the trait either moved out of sympatry into allopatry or the trait under reinforcement exists as

standing genetic variation outside of sympatry. Though it is important to note that these females may reject *D. recens* males for different reasons than females in sympatry.

In chapter three, I found evidence of introgression between *D. recens* and *D. subquinaria*, and this gene flow is asymmetric at autosomal loci with *D. recens* contributing more alleles to *D. subquinaria* populations than the other way around. This suggests that selection for reinforcement in *D. subquinaria* has been strong to overcome the effects of gene flow from *D. recens*. The X chromosome does not show biased introgression, which could suggest a role for this chromosome in reproductive isolation. Within *D. recens* there is substantial gene flow between sympatric and allopatric populations, which could have hindered the evolution of reinforcement in sympatric populations (consistent with results shown in Matute 2010). Thus, despite the cost of hybrid male sterility, gene flow between and within species may have stalled the evolution of reinforcement in *D. recens*. Within *D. subquinaria*, we find strong genetic structure between sympatric populations on the east of the Rocky Mountains and populations on the west of the mountains. Overall, we see that the Rocky Mountains may be slowing gene flow among allopatric and sympatric populations, which would support the idea that reinforcing and sexual selection may be strengthened when coupled with reduced gene flow between sympatric and allopatric populations (Servedio and Kirkpatrick 1997; Ritchie 2007; Hoskin and Higgie 2010; Matute 2010; Abbott et al. 2013).

In chapter four, I found that sympatric *D. subquinaria* discrimination against *D. recens* is dominant, and discrimination within species is complex. Generally, the presence of a sympatric X chromosome reduces mating rates both between and within species, and

it is likely that the autosome corresponding to the C Muller element also contributes to isolation both between and within species. This follows results of other work that suggests sex chromosomes in general harbor genes involved in reproductive isolation and reinforcement, though they may not be sufficient to generate total isolation and may interact epistatically with autosomes (Saetre et al. 2003; Servedio and Saetre 2003; Ortiz-Barrientos et al. 2004; Hall and Kirkpatrick 2006; Saether et al. 2007; Presgraves 2008).

It is clear that in nature, the same secondary contact in a hybrid zone will not necessarily result in the same outcome for each species. In this case, for *D. subquinaria* secondary contact has resulted in increased divergence between species and driven the evolution of sexual isolation within species, while its sister species *D. recens* has undergone neither of these changes. We now understand that in *D. subquinaria*, selection appears to have overcome gene flow, while in *D. recens* gene flow has overcome selection.

Moving forward, it will be important to dissect the genetics underlying this event in order to completely understand how divergence between and within species has occurred. Fine-mapping will confirm whether the genes involved in female discrimination within and between species are truly the same. The genetic basis of reinforcement has been mapped in few systems (Ortiz-Barrientos et al. 2004; Hopkins and Rausher 2011) and this study represents the first investigation into of the genetic basis of cascade reinforcement. If we are to know whether the same genetic basis is necessary for cascade reinforcement, we must identify the genes that contribute to isolation within species, and based on other work in this system we have a better idea of what those genes might be. In *D. subquinaria* it seems likely that cuticular hydrocarbons

are the main traits used in female mate choice (Curtis et al. 2013; Giglio and Dyer 2013; Dyer et al. 2014). Cuticular hydrocarbon profiles are distinct between *D. subquinaria* and *D. recens* and among sympatric and allopatric populations of *D. subquinaria*, and sympatric females that cannot smell are less likely to mate. Mapping of female discrimination in this species may reveal genes in the olfaction pathway that underlie sexual isolation between and within species.

Furthermore, while I focused on female discrimination in this dissertation, it is possible that male mate choice may contribute to some of the variation in mating behavior. Male mate choice has been observed to be under reinforcing selection in other species (Gregorio et al. 2012), and based on results in Giglio and Dyer (2013), sympatric *D. subquinaria* males who were unable to detect olfactory cues did not mate as readily. This could indicate an unknown male component to the sexual isolation seen both within and between species.

Additionally, ecological selection could play a role in divergence within *D. subquinaria*. We see phenotypic differences among populations of *D. subquinaria* from across the range. Cuticular hydrocarbons are distinct among regions of *D. subquinaria* (Dyer et al. 2014), and coastal individuals are significantly larger than inland individuals (Stribos and Dyer, unpublished results). The range of *D. subquinaria* covers coastal British Columbia east to the plains of Alberta, north to Alaska and south to Montana and Idaho. There is substantial variation in ecology and climate across this range, so it is possible that local adaptation and ecological selection may also be driving divergence along with reinforcement. For instance, we found some evidence that the X chromosome is differentiated between coastal allopatric and inland allopatric species of *D.*

subquinaria, which corresponds to patterns seen in mtDNA. It is possible that ecological selection is acting on an X-mtDNA epistatic interaction (Montooth et al. 2010). Thus, ecological selection could be acting in concert with reinforcing selection to promote divergence among populations.

Disentangling the processes driving reinforcement and its consequences will require a thorough evolutionary genetic dissection of multiple natural systems; however, some broad patterns emerge from this system with regard to divergence both between and within species. Too much gene flow can inhibit the evolution of reinforcement by swamping out reinforced traits. Thus a logical step is that too much gene flow between populations within species will also inhibit the evolution of cascade reinforcement. Additionally, a shared genetic basis of discrimination may facilitate secondary divergence, but an important consideration is the costs associated with rejecting potential viable mates. If migrants from allopatry are frequent, these costs may be high. This suggests that populations in sympatry will be under the conflicting costs of maladaptive hybridization versus the costs of rejecting viable mates. Whether the balance tips in favor of one of the forces over the other could determine the outcome of secondary contact.

Literature Cited

- Abbott, R., D. Albach, S. Ansell, J. W. Arntzen, S. J. E. Baird, N. Bierne, J. W. Boughman, A. Brelsford, C. A. Buerkle, R. Buggs, R. K. Butlin, U. Dieckmann, F. Eroukhmanoff, A. Grill, S. H. Cahan, J. S. Hermansen, G. Hewitt, A. G. Hudson, C. Jiggins, J. Jones, B. Keller, T. Marczewski, J. Mallet, P. Martinez-Rodriguez, M. Most, S. Mullen, R. Nichols, A. W. Nolte, C. Parisod, K. Pfennig, A. M. Rice, M. G. Ritchie, B. Seifert, C. M. Smadja, R. Stelkens, J. M. Szymura, R. Vainola, J. B. W. Wolf, and D. Zinner. 2013. Hybridization and speciation. *J. Evol. Biol.* 26:229-246.
- Coyne, J. A. and H. A. Orr. 2004. Speciation. Sinauer Associates, Sunderland (Massachusetts).
- Curtis, S., J. L. Sztepanacz, B. E. White, K. A. Dyer, H. D. Rundle, and P. Mayer. 2013. Epicuticular compounds of *Drosophila subquinaria* and *D. recens*: identification, quantification, and their role in female mate choice. *J Chem Ecol* 39:579-590.
- Dyer, K. A., B. E. White, J. L. Sztepanacz, E. R. Bewick, and H. D. Rundle. 2014. Reproductive character displacement of epicuticular compounds and their contribution to mate choice in *Drosophila subquinaria* and *D. recens*. *Evolution* 68:1163–1175.
- Giglio, E. M. and K. A. Dyer. 2013. Divergence of premating behaviors in the closely related species *Drosophila subquinaria* and *D. recens*. *Ecol. Evol.* 3:365-374.
- Gregorio, O., E. L. Berdan, G. M. Kozak, and R. C. Fuller. 2012. Reinforcement of male mate preferences in sympatric killifish species *Lucania goodei* and *Lucania parva*. *Behav. Ecol. Sociobiol.* 66:1429-1436.

- Hall, D. W. and M. Kirkpatrick. 2006. Reinforcement and sex linkage. *Evolution* 60:908-921.
- Higgie, M. and M. W. Blows. 2007. Are traits that experience reinforcement also under sexual selection? *Am. Nat.* 170:409-420.
- Hopkins, R. and M. D. Rausher. 2011. Identification of two genes causing reinforcement in the Texas wildflower *Phlox drummondii*. *Nature* 469:411-414.
- Hoskin, C. J. and M. Higgie. 2010. Speciation via species interactions: the divergence of mating traits within species. *Ecol. Lett.* 13:409-420.
- Hoskin, C. J., M. Higgie, K. R. McDonald, and C. Moritz. 2005. Reinforcement drives rapid allopatric speciation. *Nature* 437:1353-1356.
- Howard, D. J. 1993. Reinforcement: Origin, dynamics, and fate of an evolutionary hypothesis. Pp. 46-69 in H. Rg, ed. *Hybrid zones and the evolutionary process*. Oxford University Press, New York.
- Jaenike, J., K. A. Dyer, C. Cornish, and M. S. Minhas. 2006. Asymmetrical reinforcement and *Wolbachia* infection in *Drosophila*. *PLoS Biol.* 4:1852-1862.
- Matute, D. R. 2010. Reinforcement can overcome gene flow during speciation in *Drosophila*. *Curr. Biol.* 20:2229-2233.
- Montooth, K. L., C. D. Meiklejohn, D. N. Abt, and D. M. Rand. 2010. Mitochondrial-nuclear epistasis affects fitness within species but does not contribute to fixed incompatibilities between species of *Drosophila*. *Evolution* 64:3364-3379.
- Ortiz-Barrientos, D., B. A. Counterman, and M. A. F. Noor. 2004. The genetics of speciation by reinforcement. *PLoS Biol.* 2:2256-2263.

- Ortiz-Barrientos, D., A. Greal, and P. Nosil. 2009. The genetics and ecology of reinforcement implications for the evolution of prezygotic isolation in sympatry and beyond. Pp. 156-182 in C. D. Schlichting, and T. A. Mousseau, eds. Year in Evolutionary Biology 2009. Wiley-Blackwell, Malden.
- Presgraves, D. C. 2008. Sex chromosomes and speciation in *Drosophila*. Trends in genetics : TIG 24:336-343.
- Ritchie, M. G. 2007. Sexual selection and speciation. Annu. Rev. Ecol. Evol. S. 38:79-102.
- Saether, S. A., G. P. Saetre, T. Borge, C. Wiley, N. Svedin, G. Andersson, T. Veen, J. Haavie, M. R. Servedio, S. Bures, M. Kral, M. B. Hjernquist, L. Gustafsson, J. Traff, and A. Qvarnstrom. 2007. Sex chromosome-linked species recognition and evolution of reproductive isolation in flycatchers. Science 318:95-97.
- Saetre, G. P., T. Borge, K. Lindroos, J. Haavie, B. C. Sheldon, C. Primmer, and A. C. Syvanen. 2003. Sex chromosome evolution and speciation in *Ficedula* flycatchers. P. Roy. Soc. B-Biol. Sci. 270:53-59.
- Servedio, M. R. and M. Kirkpatrick. 1997. The effects of gene flow on reinforcement. Evolution 51:1764-1772.
- Servedio, M. R. and G. P. Saetre. 2003. Speciation as a positive feedback loop between postzygotic and prezygotic barriers to gene flow. P. Roy. Soc. B-Biol. Sci. 270:1473-1479.

APPENDIX A

Supplemental Material for Chapter 2, in submission at *Evolution*

Figure S1.

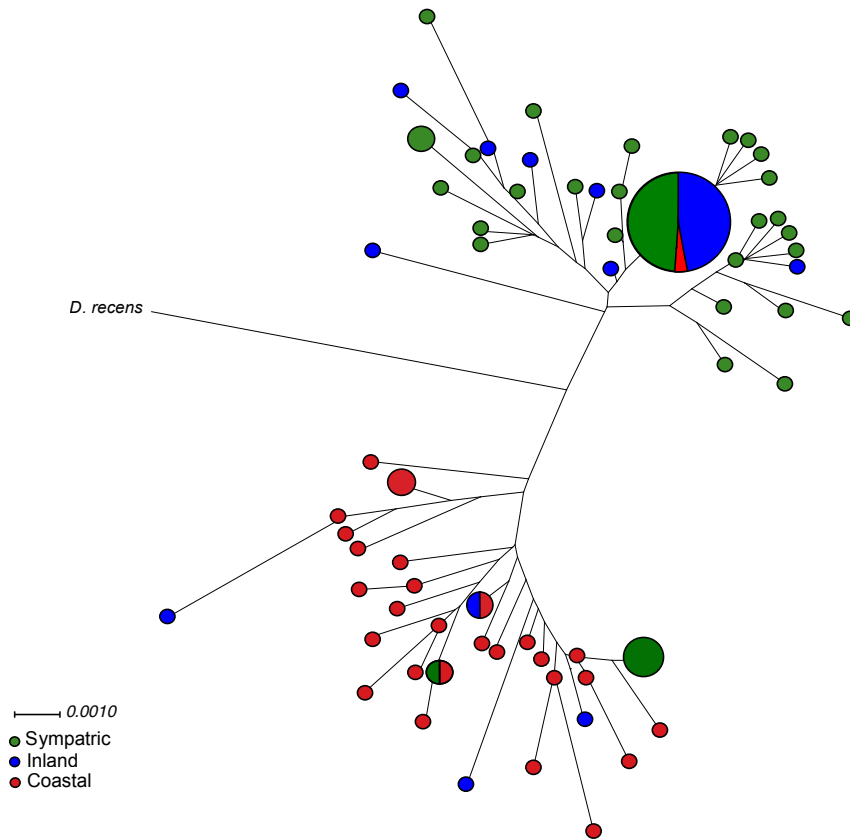
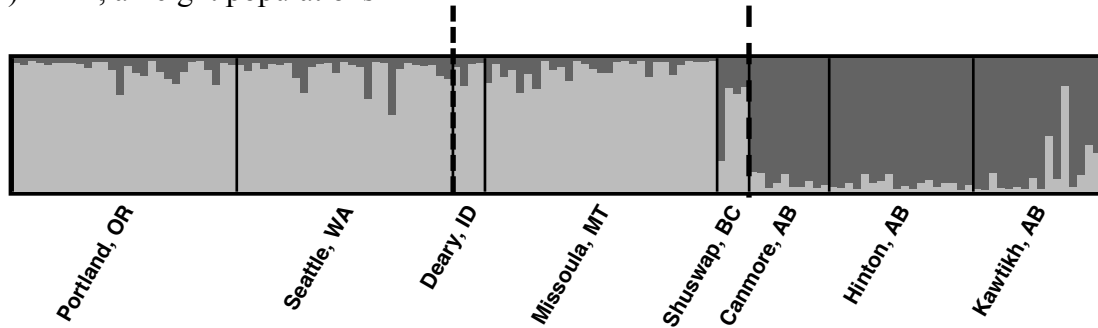


Figure S1. Minimum spanning tree for *D. subquinnaria* mtDNA *COI*, using samples collected in 2010-2011. The location of each sample is indicated by the color (red for coast allopatric, blue for inland allopatric, and green for populations sympatric with *D. recens*), and the size of each circle indicates the number of samples sharing a haplotype. The three *D. subquinnaria* samples with a *D. recens* mtDNA haplotype were not included. This figure visually shows the pattern of differentiation of populations to the east and west of the Coast Mountains.

Figure S2.

A) $K = 2$, all eight populations



B) $K = 2$, including only the inland six populations

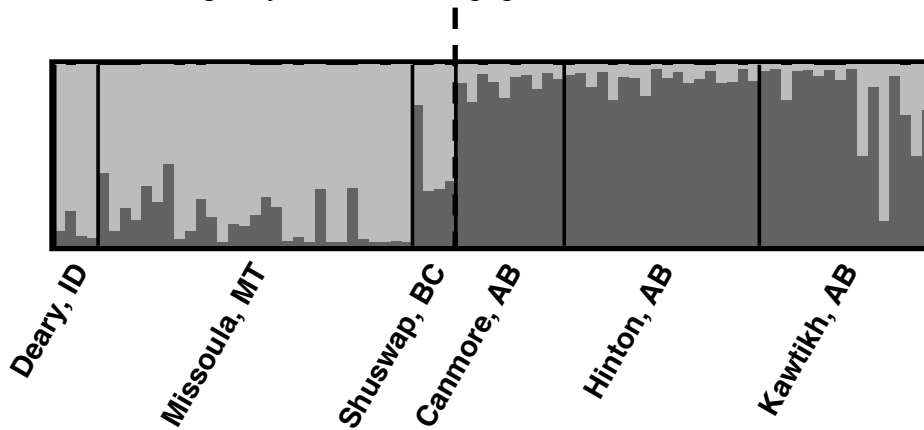


Figure S2. Results from Structure showing differentiation of sympatric and allopatric populations using the microsatellite markers. Each column is an individual, which are grouped by population. Populations are shown from west to east, and the proportion ancestry of each individual into each genetic cluster is shown as light and dark grey. The dotted and dashed lines respectively show the divisions between the coastal allopatric, inland allopatric, and sympatric populations

Figure S3.

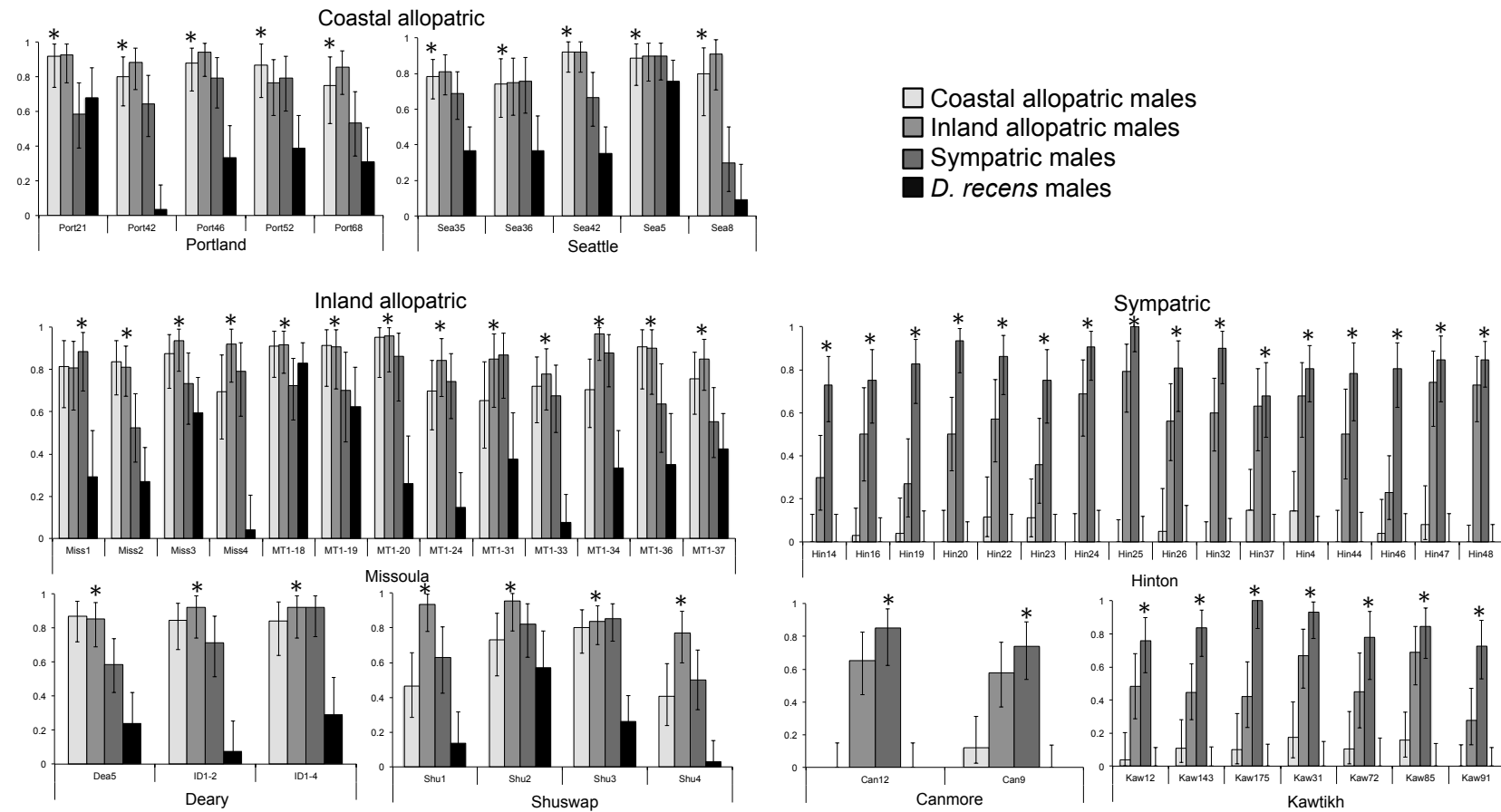


Figure S3. For each isofemale line, the mean proportion of pairs mated in two hours against each male type. Isofemale lines are grouped by population and geographic region, and each male type is shown in a different shade of grey. The error bars indicate the 95% confidence intervals calculated using a binomial distribution. The stars indicate the 'home' male type for each line. Within species, the mating rates for “home” combinations where the female and male are from the same geographic region are generally high and show no significant genetic variation.

Table S1. Primers used to amplify microsatellite loci.

Locus	Repeat motif	Primer	Sequence (5' to 3')
<i>Drec1007</i>	CA	1007_F1	TTTCATTAGCACGCGCAAC
		1007_R1	CGATTGCTTTTTGCCTTTG
<i>Drec1081</i>	CA	1081_F2	GCAATAAAACTAGCAAAACAAATAGC
		1081_R2	GCGTGTGCAGCTTGTAATG
<i>Drec3001</i>	CA	3001_F2	AGTGGAAAGGGGATTTTTCC
		3001_R1	GTGCCTTGAAAGCTCCAGTG
<i>Drec3028</i>	CAG	3028_F2	GTGGCATCGTTGGCAAATC
		3028_R1	AACATTGATGCCAGCAAGC
<i>Drec3038</i>	CAGT	3038_F1	TTTCAACTCGGCTGCTTTTC
		3038_R1	CTGACTGTCCGGCTCAATG
<i>Drec3089</i>	CA	3089_F1	CGATGTCTGAACACATAAAAAGTG
		3089_R1	AAGGCAAATTCTATTCAACTAAATG
<i>Drec3091</i>	GT	3091_F1	TATGTGCTCGGGCGTGTG
		3091_R1	GGCGGGTACGTATTTGTTTC
<i>Drec4069</i>	CAT	4069_F1	CCCCACATCGTCATCAAAG
		4069_R2	TGCACACCGATGAGTACGAG

Table S2. Variability measures for each of the eight microsatellites within each population and across all populations. Shown are the number of different alleles (A), the observed heterozygosity (H_O), and the expected heterozygosity (H_E). Populations are listed in the same order and with the same abbreviations as in Table 1. The dotted line indicates the boundary between coastal and inland allopatric populations, and the dashed line indicates the boundary between sympatric and allopatric populations. H_O values in bold show a deficiency of heterozygotes at $P < 0.001$. Deviations from Hardy Weinberg Equilibrium only occur when data are summed across populations, which are likely a result of the Wahlund effect.

Locus		POR	SEA	DEA	MIS	SHU	CAN	HIN	KAW	mean	sd	total
<i>Drec1007</i>	A	5	6	2	5	4	6	7	8	5.38	1.85	10
	Ho	0.65	0.56	0.50	0.48	0.75	0.86	0.62	0.80	0.65	0.14	0.62
	He	0.55	0.60	0.50	0.54	0.82	0.81	0.74	0.86	0.68	0.15	0.69
<i>Drec1081</i>	A	6	6	3	6	3	5	15	13	7.13	4.45	20
	Ho	0.74	0.92	1.00	0.67	0.75	0.33	0.80	0.69	0.74	0.20	0.73
	He	0.81	0.81	0.83	0.76	0.61	0.79	0.81	0.79	0.78	0.07	0.84
<i>Drec3001</i>	A	7	8	3	9	4	6	9	14	7.5	3.42	20
	Ho	0.85	0.85	0.50	0.74	0.75	0.80	0.72	0.93	0.77	0.13	0.81
	He	0.76	0.81	0.83	0.78	0.75	0.84	0.84	0.88	0.81	0.05	0.84
<i>Drec3028</i>	A	3	2	1	3	3	6	6	4	3.5	1.77	7
	Ho	0.58	0.24	na	0.22	1.00	0.50	0.64	0.57	0.54	0.26	0.43
	He	0.45	0.22	na	0.21	0.68	0.77	0.64	0.71	0.52	0.27	0.56
<i>Drec3038</i>	A	17	18	4	14	6	13	15	13	12.5	4.99	34
	Ho	0.92	0.73	1.00	0.58	1.00	0.89	0.93	0.81	0.86	0.15	0.80
	He	0.93	0.90	1.00	0.90	1.00	0.95	0.91	0.92	0.94	0.04	0.93
<i>Drec3089</i>	A	4	4	2	5	3	10	12	12	6.5	4.14	18
	Ho	0.43	0.54	0.50	0.31	0.33	1.00	0.75	0.67	0.57	0.23	0.53
	He	0.44	0.47	0.50	0.37	0.73	0.93	0.90	0.86	0.65	0.23	0.66

(continued)

Locus		POR	SEA	DEA	MIS	SHU	CAN	HIN	KAW	mean	sd	total
<i>Drec3091</i>	A	11	9	2	9	4	7	7	9	7.25	2.96	13
	Ho	0.79	0.80	0.00	0.58	0.50	0.75	0.60	0.67	0.59	0.26	0.68
	He	0.86	0.82	0.68	0.83	0.75	0.85	0.80	0.82	0.80	0.06	0.84
<i>Drec4069</i>	A	3	4	1	2	2	6	7	6	3.88	2.23	10
	Ho	0.29	0.15	na	0.15	0.25	1.00	0.75	0.56	0.45	0.33	0.36
	He	0.31	0.22	na	0.21	0.25	0.79	0.68	0.67	0.45	0.26	0.43
all loci	A mean	7	7.13	2.25	6.63	3.63	7.38	9.75	9.88	6.70	2.66	
	A sd	5.07	5.20	1.23	4.26	1.64	3.54	4.77	4.76	3.81	1.56	
	Ho mean	0.57	0.53	0.55	0.42	0.59	0.69	0.66	0.63	0.58	0.08	
	Ho sd	0.30	0.33	0.41	0.25	0.35	0.34	0.26	0.27	0.31	0.05	
	He mean	0.57	0.54	0.62	0.51	0.62	0.75	0.70	0.72	0.63	0.09	
	He sd	0.30	0.33	0.33	0.33	0.31	0.29	0.28	0.28	0.31	0.02	

Table S3. Molecular analysis of variance (AMOVA) of the mtDNA *COI* locus and eight microsatellites using different groupings of *D. subquinnaria* populations. The 2009-2011 samples include samples from this study only, and the 2001-2011 mtDNA samples also include this study and data from Jaenike et al. (2006). *P*-values in bold are significant at the Bonferroni-corrected level of $P < 0.004$. These data show that at the mtDNA most regional variation is distributed on either side of the Coast Mountains, and at the microsatellites most regional variation is distributed on either side of the Rocky Mountains.

Grouping	Partition	mtDNA		mtDNA		microsatellites	
		2009-2011		2001-2011		2009-2011	
		% of variation	<i>F</i>	% of variation	<i>F</i>	% of variation	<i>F</i>
Two groups: Sympatric, Allopatric (Allo inland + allo coast)	Among groups	12.63	$\Phi_{CT} = 0.13$ ($P = 0.13$)	10.06	$\Phi_{CT} = 0.10$ ($P = 0.03$)	12.87	$\Phi_{CT} = 0.13$ ($P = 0.02$)
	Among populations within groups	12.64	$\Phi_{SC} = 0.14$ ($P < 0.0001$)	14.78	$\Phi_{SC} = 0.16$ ($P < 0.0001$)	1.45	$\Phi_{SC} = 0.02$ ($P = 0.0006$)
	Within populations	74.73	$\Phi_{ST} = 0.25$ ($P < 0.0001$)	75.17	$\Phi_{ST} = 0.25$ ($P < 0.0001$)	85.68	$\Phi_{ST} = 0.14$ ($P < 0.0001$)
Two groups: Sympatric, Allo inland (POR and SEA excluded)	Among groups	1.19	$\Phi_{CT} = 0.012$ ($P = 0.30$)	0.32	$\Phi_{CT} = 0.003$ ($P = 0.33$)	13.60	$\Phi_{CT} = 0.14$ ($P = 0.10$)
	Among populations within groups	-5.50	$\Phi_{SC} = -0.05$ ($P = 0.99$)	-1.19	$\Phi_{SC} = -0.01$ ($P = 0.74$)	1.70	$\Phi_{SC} = 0.02$ ($P = 0.05$)
	Within populations	104.31	$\Phi_{ST} = -0.04$ ($P = 0.99$)	100.87	$\Phi_{ST} = -0.01$ ($P = 0.75$)	84.70	$\Phi_{ST} = 0.15$ ($P < 0.0001$)
Two groups: Inland (Allo inland + Sympatric), Allo Coast	Among groups	37.56	$\Phi_{CT} = 0.38$ ($P = 0.04$)	39.50	$\Phi_{CT} = 0.39$ ($P < 0.0001$)	1.21	$\Phi_{CT} = 0.01$ ($P = 0.32$)
	Among populations within groups	-2.52	$\Phi_{SC} = -0.04$ ($P = 1.0$)	-0.57	$\Phi_{SC} = -0.01$ ($P = 0.77$)	8.07	$\Phi_{SC} = 0.08$ ($P < 0.0001$)
	Within populations	64.96	$\Phi_{ST} = 0.35$ ($P < 0.0001$)	61.08	$\Phi_{ST} = 0.39$ ($P < 0.0001$)	90.72	$\Phi_{ST} = 0.09$ ($P < 0.0001$)
Three groups: Sympatric, Allo inland, Allo coast	Among groups	29.66	$\Phi_{CT} = 0.30$ ($P = 0.007$)	29.04	$\Phi_{CT} = 0.29$ ($P < 0.0001$)	9.31	$\Phi_{CT} = 0.09$ ($P = 0.03$)
	Among populations within groups	-3.10	$\Phi_{SC} = -0.04$ ($P = 0.99$)	-0.76	$\Phi_{SC} = -0.01$ ($P = 0.77$)	1.29	$\Phi_{SC} = 0.01$ ($P = 0.03$)
	Within populations	73.44	$\Phi_{ST} = 0.27$ ($P < 0.0001$)	71.72	$\Phi_{ST} = 0.28$ ($P < 0.0001$)	89.39	$\Phi_{ST} = 0.11$ ($P < 0.0001$)

Table S4. Estimates of pairwise population differentiation at the mtDNA *COI* locus of *D. subquinaria*. The pairwise F_{ST} is below diagonal, and pairwise S_{nn} is above the diagonal. Values in bold are significant at the Bonferroni-corrected level of $P < 0.002$, based on 10,000 permutations. Populations are listed in the same order and with the same abbreviations as in Table 1. The dashed line indicates the boundary between sympatric and allopatric populations, and the dotted line indicates the boundary between coastal and inland allopatric populations. In sum, 11 of 12 significant F_{ST} comparisons and all five of five significant S_{nn} comparisons are between populations to either side of the Coast Mountains, and there is no differentiation between populations sympatric with *D. recens* and those just outside this region.

	POR	SEA	DEA	MIS	SHU	CAN	HIN	KAW
POR		0.514	0.924	0.933	0.955	0.930	0.970	0.879
SEA	-0.018		0.611	0.783	0.767	0.759	0.897	0.677
DEA	0.286	0.269		0.519	0.292	0.439	0.732	0.467
MIS	0.367	0.330	-0.043		0.594	0.483	0.470	0.321
SHU	0.307	0.284	-0.132	-0.120		0.424	0.712	0.574
CAN	0.346	0.328	-0.087	-0.069	-0.105		0.540	0.388
HIN	0.407	0.378	-0.009	-0.025	-0.076	-0.055		0.475
KAW	0.361	0.316	-0.046	-0.043	-0.089	-0.089	-0.021	

Table S5. Pairwise population differentiation at the eight microsatellite loci using F_{ST} . Values in bold are significant at the $P < 0.01$ level, and values in bold are significant at the Bonferroni-corrected level of $P < 0.002$, based on 10000 permutations. Populations are listed in the same order and with the same abbreviations as in Table 1. The dashed line indicates the separation between allopatric and sympatric populations, and the dotted line indicates the boundary between coastal and inland allopatric populations. Of the 15 significant comparisons, 13 are between sympatric vs. allopatric populations. Comparing populations in neighboring regions, two of six comparisons are significant between coastal and inland allopatric populations, and seven of nine comparisons between sympatric and inland allopatric regions are significant, suggesting the break between sympatry and allopatry is more substantial.

	POR	SEA	DEA	MIS	SHU	CAN	HIN
SEA	0.005						
DEA	0.011	-0.012					
MIS	0.025	0.019	-0.024				
SHU	0.043	0.047	0.064	0.083			
CAN	0.137	0.156	0.199	0.179	0.083		
HIN	0.143	0.148	0.176	0.185	0.089	0.022	
KAW	0.102	0.111	0.107	0.147	0.059	0.000	0.009

Table S7. Proportion of membership of each population into each cluster, as defined by the Structure analyses. For each dataset one of five runs with the most probable values of $K=2$ is shown here; the results were consistent among runs. Results are shown for the analyses of all populations (8 populations) and only the Inland populations (6 populations, excluding Seattle and Portland). The dashed line indicates the separation between allopatric and sympatric populations, and the dotted line indicates the boundary between coastal and inland allopatric populations. These results show that the major clustering is between sympatric and allopatric populations.

Population	All populations		Inland only	
	Cluster 1	Cluster 2	Cluster 1	Cluster 2
POR	0.926	0.074	-	-
SEA	0.907	0.093	-	-
DEA	0.914	0.086	0.906	0.094
MIS	0.922	0.078	0.854	0.146
SHU	0.634	0.366	0.560	0.44
CAN	0.077	0.923	0.101	0.899
HIN	0.059	0.941	0.079	0.921
KAW	0.159	0.841	0.193	0.807

Table S8. Comparison of three models to describe each phenotypic and genotypic cline. Shown are the mean and standard deviations (in parentheses) of the maximum likelihood (llmax) across five runs. The best supported model(s) based on the AIC values are indicated in bold.

Cline	Model	N Parameters	llmax	Center (km/10)	Slope
Allo Coast	Unimodal	7	64.35	50.09	-0.252
			(0.004)	(0.24)	(0.01)
	Bimodal	8	66.32	48.46	-9.306
			(0)	(0.01)	(0.90)
Allo Inland	Trimodal	12	69.52	46.91	-0.227
			(1.10)	(0.40)	(0.02)
	Unimodal	7	52.99	62.22	-3.867
			(0)	(0.12)	(2.42)
Sympatric	Bimodal	8	52.76	62.32	-8.299
			(0.004)	(0.02)	(1.64)
	Trimodal	12	62.19	58.95	-0.36
			(0.97)	(2.17)	(0.23)
<i>D. recens</i>	Unimodal	7	44.29	61.85	1.76
			(0.07)	(1.64)	(1.68)
	Bimodal	8	44.25	48.55	3.898
			(0.001)	(0.14)	(2.169)
Microsat	Trimodal	12	49.15	8.63	0.047
			(0.86)	(8.44)	(0.01)
	Unimodal	7	154.86	48.43	-9.77
			(0)	(0.002)	(0.18)
mtDNA	Bimodal	8	158.09	48.43	-9.99
			(0.001)	(0.003)	(0.001)
	Trimodal	12	160.10	48.20	-8.02
			(0.75)	(0.347)	(4.05)
Microsat	Unimodal	7	128.21	48.40	-9.52
			(0.001)	(0.005)	(0.65)
	Bimodal	8	165.68	51.23	-0.45
			(0.001)	(0.023)	(0.003)
mtDNA	Trimodal	12	175.15	50.59	-0.64
			(1.16)	(1.16)	(0.39)
	Sigmoid	4	-28.06	17.06	-9.97
			(0)	(0.001)	(0.05)

Table S9. Constraints between clines in discrimination against *D. recens* and coastal allopatric males. Cline width was estimated as 4/slope. The bimodal model was used, and clines were constrained to have the same width (slope), center, width and center together, or no constraint. Number of parameters (N_p) is shown, and the best fitting model determined by ΔAIC is indicated in bold. This shows that the clines in discrimination against *D. recens* and coastal allopatric males are coupled, meaning they share the same center and width.

Constraint	N_p	llmax	Center (km)	Width (km)	$\Delta AICc$
Width	15	224.31	na	5	3.15
Center	15	223.98	485	na	3.81
Width and Center	14	223.98	484	4	0.00
No constraint	16	224.46	na	na	6.85