THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

Genetic Dissection of a Major Polymorphism Underlying Population Divergence in Sexually Selected Pheromones in *Drosophila serrata*

Bosco Benedict Rusuwa
B.Ed., MSc.

A thesis submitted for the degree of Doctor of Philosophy at

The University of Queensland in 2013

School of Biological Sciences

Abstract

With the advent of QTL mapping in the late 1980s, evolutionary geneticists have actively dissected the genetic basis of adaptive phenotypes. In recent times, advances in whole genome sequencing have allowed for finer scale mapping of traits that have diverged due to natural selection. These relatively inexpensive sequencing technologies have also greatly facilitated adaptive trait dissection in non-model species that exhibit especially interesting patterns of adaptation. One class of trait that has largely escaped the attention of the next generation ecological and evolutionary genomics research programme, are sexually-selected traits. Sexual selection is a strong form of directional selection in nature and is thought to be responsible for the evolution of elaborate sexual ornaments and armaments. However we still know comparatively little of the genetic basis of such traits.

The cuticular hydrocarbons (CHCs) of *Drosophila serrata* provide an opportunity for dissecting the divergence of sexually selected traits. Populations along the eastern Australian coast exhibit latitudinal variation in these traits in association with climatic factors and consistent with the action of divergent natural selection. Under experimental settings, divergent sexual as well as natural selection has been implicated in the evolution of CHCs. Natural populations of *D. serrata* along the northern part of the Eastern Australian coast have recently been observed to exhibit a polymorphism in three of their CHC compounds representing the shortest carbon chains; 5,9-tetracosadiene (5,9-C_{24}), 5,9-pentacosadiene (5,9-C_{25}) and 9-pentacosene (9-C_{25}). One class of phenotype only expresses these three compounds in trace amounts ('low' phenotype) whereas the other has normal levels ('high' phenotype). These short-chained CHCs also exhibit strong genetically-based latitudinal clines up to, but not beyond, 20 degrees south of the equator. Based on both traditional QTL and modern sequence-based genomic mapping approaches, the aim of this study was to dissect the genetic basis of the polymorphism in *D. serrata* CHCs and to examine its adaptive significance within the context of sexual and natural selection.

Through F2 QTL mapping based on a cross between two inbred lines, from the opposite ends of the *D. serrata* CHC cline, one a 'low' phenotype and the other a 'high' phenotype, the genetic basis of all CHCs in this species was mapped to twenty two overlapping QTLs on chromosomes 2 and 3. The short-chain CHC polymorphism was traced to two major effect recessive QTLs on the right arm of chromosome 3, which including their interaction, accounted for more than 70% of the variance in the polymorphism (Chapter 2). Fine mapping of this major polymorphism was then conducted using next generation DNA sequencing and bulk segregant analysis of an advanced F60 cross of the same founding lines. Bulk segregant analysis revealed a single peak of genetic differentiation on chromosome 3R (~20kb), harbouring three adjacent fatty acyl-CoA reductase

genes (Chapter 3). Other candidate genes already reported in the literature as underlying CHC variation in *Drosophila* were also detected nearby the reductases but appear unlikely responsible for the polymorphism. An analysis of genome sequences for nine independent wild-derived inbred lines, fixed for either the 'low' or 'high' phenotype, replicated the results of the bulk segregant analysis and uncovered a hotspot of fixed nucleotide differences within three adjacent reductase genes, particularly in gene CG17560. This gene has recently been confirmed to be expressed in *D. serrata* oenocytes in both CTN42 and FORS4 lines (Chapter 4).

In an attempt to identify the likely sources of natural and sexual selection acting on and maintaining this polymorphism, I assayed desiccation resistance and heat stress survival on multiple lines of contrasting phenotype from a single population of flies in Cooktown, far north Queensland. I also assayed male and female mate choice on the same lines (Chapter 5). The fitness effects of the polymorphism appeared sex-specific. Natural selection seemed to operate on this polymorphism through females, with 'low' individuals being superior to 'high' ones in their survival to desiccation. By contrast, the effect of sexual selection on this polymorphism was evident in males but not females. 'Low' males were half as likely to succeed in copulation as their 'high' counterparts.

This is the first study to trace the genetic basis of CHC variation to gene level in *D. serrata* and exposes a small genomic region where a polymorphism may be maintained by an antagonistic relationship between natural and sexual selection. Although complex traits are generally polygenic, mutations within specific segments of a biosynthetic pathway may have a pervasive effect on trait divergence if targeted by selection. Further tissue–specific gene expression analysis is likely to pinpoint the gene and ultimately the actual causal mutation(s) underlying this polymorphism. Combining traditional QTL and modern genomic approaches may greatly accelerate the fine-scale genetic dissection of adaptive trait divergence in non-model species.

Declaration by author

This thesis *is composed of my original work, and contains* no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my research higher degree candidature and does not include a substantial part of work that has been submitted *to qualify for the award of any* other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the General Award Rules of The University of Queensland, immediately made available for research and study in accordance with the *Copyright Act 1968*.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis.

Publications during candidature

Peer-reviewed paper

STOCKER, A. J., B. B. RUSUWA, M. J. BLACKET, F. D. FRENTIU, M. SULLIVAN *et al.*, 2012 Physical and Linkage Maps for *Drosophila serrata*, a Model Species for Studies of Clinal Adaptation and Sexual Selection. G3 (Bethesda) **2:** 287-297.

Publications included in this thesis

No publications included.

Contributions by others to the thesis

Dr Stephen Chenoweth was instrumental in the conception and design phase of this project as well as in advising on data analysis. Dr Francesca Frentiu helped with the setting up of the F2 intercoss and running the gas chromatograph for the QTL mapping flies presented in Chapter 2, Dr James Hereward's advice was crucial in testing the primers for the markers used in the F2 QTL cross. Mitchell Sullivan and Dr Scott Beatson were instrumental in the development of ESTs from CTN42 and FORS4 RNA-seq data which I used to design the PCR primers for the QTL mapping experiment in Chapter 2. Nick Appleton and Camille Latimer helped with the laboratory bench work during the desiccation and mating assays while Scott Allen helped a lot with bioinformatics during DNA sequence data analyses for chapters 3 and 4. Although it is used in the thesis, but does not form a specific chapter, my contribution to the *D. serrata* genome sequencing effort was to maintain the reference fly line and prepare all DNA samples required for sequencing. Stephen Chenoweth performed the draft genome assembly. In Chapter 4, the *in situ* hybridizations I refer to were the result of a recent collaboration between Stephen Chenoweth, Prof. Sean Carroll and Dr Henry Chung at the University of Wisconsin.

Statement of parts of the thesis submitted to qualify for the award of another degree

None

Acknowledgements

I would like to thank my supervisor Stephen Chenoweth for his academic mentoring and moral support during the entire period of my studies. His sense of optimism and encouragement proved the most vital at times when the science seemed stuck. It was through him that I got introduced to working with *Drosophila* and the basics of quantitative genetics that were pivotal for my appreciation of evolutionary biology and particularly the workings of trait divergence in nature. I owe a lot of thanks to my second supervisor Daniel Ortiz-Barrientos for his encouragement. I would also like to sincerely acknowledge the generous technical and social support that I enjoyed amongst the Chenoweth laboratory members as I went through my experiments. James Hereward was hugely supportive during molecular laboratory work and Scott Allen was an asset when it came to analysis of DNA sequence data. I got a lot of help with long hours of laboratory bench work with flies from Tom Gosden, Camille and Nick.

My studies would never have been possible without financial support from the Australian Government through the International Postgraduate Research Scholarship and the University of Queensland Research scholarship. I owe a lot of gratitude to the Australian people for this honourable gesture. Stephen Chenoweth was very kind in allowing me access his research resources for my studies. I also wish to thank Alan Webb and Jane Thompson for their material and moral support during my study period. Their generosity and understanding were of the highest order. I would also like to acknowledge the University of Malawi for endorsing my study leave and their support during my stay in Australia.

My wife Anga, our sons Wiwa and Nohakera and the larger extended family were a huge support and their unwavering love and patience as I stayed away from home for years during my studies were phenomenal.

Lastly, I would like to acknowledge the huge role my deceased parents played in encouraging me to forge ahead with school although, sadly, they never lived to see me come this far. I simply cannot thank them enough.

Keywords

Drosophila serrata, polymorphism, cuticular hydrocarbon, fatty acyl-coa reductase, sexually antagonistic, mate choice

Australian and New Zealand Standard Research Classifications (ANZSRC)

ANZSRC code: 060411, Population, Ecological and Evolutionary Genetics, 30%

ANZSRC code: 060303, Biological Adaptation, 30%

ANZSRC code: 060412, Quantitative Genetics, 40%

Fields of Research (FoR) Classification

FoR code: 0603, Evolutionary Biology, 30%

FoR code: 0604, Genetics, 70%

TABLE OF CONTENTS

List of Figures & Tables

List of Abbreviations used in the Thesis

CHC: Cuticular hydrocarbon

QTL: Quantitative trait locus

DNA: Deoxyribonucleic acid

BSA: Bulk segregant analysis

FAS: Fatty acid synthetase

CTN42: Cooktown 42

FORS4: Forster 4

cDNA: Cloned Deoxyribonucleic acid

EST: Expressed sequence tag

BLAST: Basic Local Alignment Search Tool

PCR: Polymerase chain reaction

SNP: Single nucleotide polymorphism

AGRF: Australian genome research facility

TRD: Transmission ratio

LOD: Logarithm of the odds

MAF: Minor allele frequency

UTR: Untranslated region

mRNA: messenger ribonucleic acid

BWA: Burrows Wheeler Alignment

SAM: Sequence Alignment/Map

CRISPR: Clustered regularly interspaced palindromic repeats

CHAPTER 1 GENERAL INTRODUCTION

1.1. NATURAL SELECTION AND THE OCCURRENCE OF PHENOTYPIC DIFFERENCES IN NATURE

Across a species range and under different environmental contexts, natural populations often differ in mean phenotype (JENKINS and HOFFMANN 2000; GRANT and GRANT 2002; COLLAR *et al.* 2011; RAIA and MEIRI 2011). For example, clinal variation in heat and cold resistance occurs along a latitudinal gradient in *Drosophila melanogaster* (HOFFMANN *et al.* 2002) while body size changes along an altitudinal axis in the Tibetan frog, *Nanorana parkeri* (MA *et al.* 2009). Osmotic tolerance in Atlantic killifish (*Fundulus heteroclitus*) varies along a salinity gradient spanning the Chesapeake Bay along the USA Atlantic coastline (WHITEHEAD *et al.* 2011). The process by which these types of differences arise is commonly referred to as phenotypic divergence. Phenotypic divergence can occur through deterministic processes such as natural or sexual selection and also stochastic processes, such as genetic drift, mutation and founder events (TRAVISANO *et al.* 1995; KELLER and TAYLOR 2008; SANCHEZ-GUILLEN *et al.* 2011); however the relative importance of each may vary (LEINONEN *et al.* 2008; SANCHEZ-GUILLEN *et al.* 2011). Empirical evidence from a variety of sources suggests that phenotypic divergence is often adaptive and occurs through selection acting on functionally divergent alleles in different environments (ENDLER 1986; MULLEN and HOEKSTRA 2008). Understanding the frequency with which divergence is an adaptive process requires formally distinguishing drift from selection, an empirical enterprise that has been a long-term focus for evolutionary geneticists.

Null model comparison techniques, which compare levels of observed population differentiation with those that would be expected purely under genetic drift (neutral expectation) such as the Q_{ST}-F_{ST} test (SPITZE 1993), provide a correlative means to infer the relative importance of each process (SCHLUTER 2000; MERILA and CRNOKRAK 2001; RIESEBERG *et al.* 2002; LEINONEN *et al.* 2008). Adaptive divergence is implied if the extent of among population differentiation in a quantitative trait, Q_{ST}, exceeds its neutral marker analogue, $F_{ST,}$ which provides a measure of divergence due to drift (LEINONEN *et al.* 2008). A meta-analysis of Q_{ST}-F_{ST} studies spanning a broad range of trait types has confirmed that in the majority of published studies, differentiation in quantitative traits exceeds that for neutral

marker loci (LEINONEN *et al.* 2008), thereby suggesting a clear role for selection in driving phenotypic divergence.

A survey of studies using quantitative trait locus (QTL) data also arrives at a similar conclusion about the role of selection. The QTL sign test can detect whether trait divergence may be due to divergence under selection rather than genetic drift (ORR 1998b). When divergence occurs under directional selection, it is expected that most QTLs from a divergent population cross will share a common direction of effect whereas drift, which does not exhibit directional bias, will produce no such pattern (SCHLUTER 2000; MERILA and CRNOKRAK 2001; RIESEBERG *et al.* 2002; LEINONEN *et al.* 2008). Following this reasoning, Rieseberg and colleagues (RIESEBERG *et al.* 2002) detected a consistent pattern of directional bias in a survey of QTL effects for over 500 traits from different studies (RIESEBERG *et al.* 2002). Both this survey and the Q_{ST}-F_{ST} meta-analysis of Leinonen and colleagues (LEINONEN *et al.* 2008) strongly implicate divergent natural selection in shaping divergence in morphological (MERILA and CRNOKRAK 2001) and life history traits (RIESEBERG *et al.* 2002).

Reciprocal transplant experiments compare the performance of individuals across different environments and provide a powerful experimental complement to null model comparisons in assessing the role of selection in phenotypic divergence (SCHLUTER 2000). Reciprocal transplant studies reveal significant genetically underlain trait divergence consistent with selection acting on fitness trade-offs between 'native' and 'away' habitats. For example, across a depth gradient of marine habitats, native transplants of the brooding coral *Seriatopora hystrix* survive best in their native habitat, leading to habitat and genetic partitioning of this species across reef environments (BONGAERTS *et al.* 2011). A similar experimental approach revealed adaptive geographic divergence in growth rates of fence lizards, *Sceloporus undulates* in response to the native thermal biophysical environment (NIEWIAROWSKI and ROOSENBURG 1993). The ubiquity and magnitude of such performance trade-offs of phenotypes transplanted across environments further points to selection as a major driver of trait divergence (SCHLUTER 2000).

1.2. THE GENETIC ARCHITECTURE OF ADAPTIVE PHENOTYPES

Since the advent of QTL mapping in the late 1980s (LANDER and BOTSTEIN 1989) evolutionary geneticists have enthusiastically dissected the genetic architectures of adaptive

traits. For complex traits, these comprise descriptions in terms of gene and allele number, patterns of allelic and mutational effects as well as pleiotropy, dominance and epistasis, which underlie the genotype-phenotype map (HANSEN 2006). Genetic architectures for adaptive phenotypic differences in quantitative traits vary from those involving few major effect loci (BAERWALD *et al.* 2011) to those implicating many loci of small effect (BUCKLER *et al.* 2009) and are context-dependent, hinging on effects of the genetic background, sex and environment (MACKAY 2009). Extensive epistatic and pleiotropic gene effects also characterise genetic architectures of many quantitative traits (BUCKLER *et al.* 2009). Studies that discover functional genetic polymorphisms for ecologically significant traits segregating in natural populations may help to identify the actual genes underlying trait divergence as well as the evolutionary forces that maintain such polymorphisms (FEDER and MITCHELL-OLDS 2003).

QTL mapping has been an important initial tool for dissecting the genetic basis of phenotypic variation and understanding the broad scale basis of the adaptive evolution of complex traits (VOSS and SHAFFER 1997; HALL *et al.* 2006). Although its findings may be biased in terms of overestimation of effect sizes (BEAVIS *et al.* 1994)and underestimation of gene number (TILLMAN-WALL *et al.* 1992), QTL mapping is often the first step towards the discovery of the actual genes for the traits in question (SINHA *et al.* 2008; STUDER and DOEBLEY 2011). QTLs of large effect have been reported in a number of studies (BRADSHAW *et al.* 1995; VOSS and SHAFFER 1997; HALL *et al.* 2006). Single QTLs of major effect may however actually consist of a cluster of multiple closely linked QTLs each with small individual effects (MCINNES and QUIGG 2010; YEAMAN and WHITLOCK 2011) and this has been demonstrated for inflorescence in the rose, *R. wichurana*, (KAWAMURA *et al.* 2011) and mandible and tail bone length in the mouse, *Mus musculus* (CHRISTIANS and SENGER 2007).

Changes in recombination rate and physical reorganization of the genome through translocations, inversions or transposable elements may re-distribute locally adaptive alleles into tight linkage groups comprising several alleles (YEAMAN and OTTO 2011; YEAMAN and WHITLOCK 2011). When multiple adaptive variants of individually small effect occur in tight linkage, their combined fitness effects can become sufficiently large to overcome the homogenising effects of migration (HILL and ROBERTSON 1966; SLATKIN 2008; YEAMAN and OTTO 2011; YEAMAN and WHITLOCK 2011). Clusters of linked alleles are thus likely to be shielded from loss by migration and drift so that their existence after successive episodes of local adaptation will surpass that expected for any random group of genes (YEAMAN and

WHITLOCK 2011). Many small effect alleles involved in adaptation at the outset will thus tend to be later swapped with a few seemingly large effect QTLs, effectively creating concentrated architectures of such loci (YEAMAN and WHITLOCK 2011). For instance, recent work on mapping sequence polymorphisms underlying the major effect QTL for cuticular pigmentation, a trait known to evolve under selection (HOLLOCHER *et al.* 2000), at the *bric-a-brac* locus in *Drosophila melanogaster* found that the major effect on the trait is actually due to the cumulative influence of multiple polymorphisms in three different, but proximate, functional regions (BICKEL *et al.* 2011).

Fine mapping strategies based on whole genome sequence data may be useful in refining findings from traditional QTL mapping. Genome-wide analysis of DNA variants correlated with phenotypic differences in complex traits potentially fast-track the identification of alleles that underlie variation in adaptive traits (GOCKEL *et al.* 2001; CAO *et al.* 2011) and reveal common polymorphisms that have potentially been driven to a high frequency by natural selection (MAGWIRE *et al.* 2012). By using a combination of both traditional QTL mapping plus whole genome sequence based strategies it is possible to map the genetic basis of trait variation down to the relevant gene. A recent compilation of more than 300 published experimental studies shows evidence of morphological and physiological trait variation within and between species being mapped to mutations in both coding and *cis*-regulatory regions of individual genes (STERN and ORGOGOZO 2008). For example, changes in melanocortin-1 receptor (*Mc1r*) gene (HOEKSTRA *et al.* 2006) and in the *cis*-regulatory region of the Agouti pigmentation gene (MULLEN AND HOEKSTRA 2008) have been linked to adaptive cryptic colouration in natural populations of mice (*Peromyscus polionotus*).

1.3. SEXUALLY SELECTED TRAIT VARIATION

1.3.1 *The role of sexual and natural selection in phenotypic evolution*

Darwin recognised two types of selective processes that might contribute to phenotypic evolution; natural selection and sexual selection (DARWIN 1859; DARWIN 1871). Natural selection is the disparity in individual survival and/or fecundity of individuals with different traits within a population, while sexual selection is the disproportionate mating success among different individuals on account of their divergent phenotypes (DARWIN 1871; ANDERSSON 1994). Sexually selected traits comprise a huge proportion of phenotypic diversity in nature (CHENOWETH and MCGUIGAN 2010). Male traits in natural populations of

different taxa including pinnipeds, birds, frogs, crickets, fish, *Anolis* lizards and insects, for instance, have been empirically demonstrated to have undergone phenotypic changes through sexual selection (LAZELL 1972; LANDE 1982; ENDLER 1983; WESTEBERHARD 1983; KIRKPATRICK 1987; CHENOWETH and BLOWS 2003; FITZPATRICK *et al.* 2012), a process that may also lead to the evolution and maintenance of sexual reproductive isolation and speciation (RITCHIE 2007).

Organisms are simultaneously under the influence of both natural and sexual selection (MAAN and SEEHAUSEN 2011). Sexually selected traits evolve in an ecological context and many attributes of trait phenotypes are a combination of both sexual and viability functions (BONDURIANSKY 2011; MAAN and SEEHAUSEN 2011). A strict delineation of roles of sexual and viability selection in long-term trait evolution is thus extremely difficult (BONDURIANSKY 2011). Within the context of adaptive speciation, for instance, natural and sexual selection reinforce each other's actions and the latter process may be regarded as a subset of the former (VAN DOORN *et al.* 2009). This notwithstanding, interaction between these two processes may substantially influence the direction and extent to which traits evolve in nature (BLOWS 2002; ONEAL and KNOWLES 2013). Sexual selection, renowned for aiding the persistence and exaggeration of seemingly detrimental traits (DARWIN 1871), may help in the purging of deleterious mutations (AGRAWAL 2001), sometimes favouring genotypes associated with enhanced survival or viability (ZAHAVI 1975; HAMILTON and ZUK 1982; KIRKPATRICK 1987) and lead to adaptive trait divergence (CANDOLIN and HEUSCHELE 2008). The divergence of body size in different populations of *Parafontaria tonominea* millipedes (SOTA and TANABE 2010) and the complex colour patterns of Australian dragon lizards (Agamidae) (CHEN *et al.* 2012) are examples where sexually selected traits have evolved in a direction consistent with the action of natural selection.

When trait values favoured by sexual selection coincide with local adaptation, natural and sexual selection may work in unison (HINE *et al.* 2011); divergent ecological selection pressures that enhance survival will also favour the evolution of unique sexual traits even in the face of gene flow (LANDE and KIRKPATRICK 1988; VAN DOORN *et al.* 2009; ONEAL and KNOWLES 2013). Sexual selection can, likewise, also enhance adaptation (LORCH *et al.* 2003; FRICKE and ARNQVIST 2007; CANDOLIN and HEUSCHELE 2008). Although sexual selection may sometimes be stronger than natural selection (HOEKSTRA *et al.* 2001), its effects on trait

divergence are most prominent when the process acts in concert with ecological or niche specialization (RITCHIE 2007).

Sexual and natural selection processes may also sometimes oppose each other (RUNDLE *et al.* 2006; CANDOLIN and HEUSCHELE 2008). For example, in *Drosophila serrata,* the interaction between sexual and natural selection leads to the divergence of pheromonal cuticular hydrocarbons (CHCs) in a manner and direction that is opposed to their evolution under solitary natural selection (BLOWS and HIGGIE 2002). Sexual selection may drive changes in a male trait that are associated with substantial direct fitness costs (LANDE 1981; SVENSSON *et al.* 2004; CANDOLIN and HEUSCHELE 2008). Many sexually selected traits such as bright colours and conspicuous displays tend to do poorly under natural selection because of their associated disadvantages (DARWIN 1871; CANDOLIN and HEUSCHELE 2008). The sexually selected male dark-mane of the African lion *Panthera leo*, for instance, leads to higher mating success but its colour also increases body surface temperatures and leads to lower food intake and abnormal sperm (WEST and PACKER 2002; CANDOLIN and HEUSCHELE 2008).

1.3.2 *Genetics of sexually selected traits*

There is a wealth of data on the genetics of habitat-specific selection that drives adaptive trait divergence but empirical data on the genetic basis of sexually selected traits is less extensive (MERILA and SHELDON 1999; CHENOWETH and MCGUIGAN 2010). This is surprising given that many phenotypic traits may covary with both sexual and nonsexual fitness (BONDURIANSKY and CHENOWETH 2009; MAAN and CUMMINGS 2009). Understanding the full extent of the evolutionary mechanisms of sexual selection, one of the strongest forms of directional selection in nature (KINGSOLVER *et al.* 2001), may require an appreciation of the sexually selected genetic polymorphisms underlying this process (CHENOWETH and MCGUIGAN 2010).

Sexually selected traits possess abundant additive genetic variance (POMIANKOWSKI and MOLLER 1995), are particularly responsive to artificial selection (AHUJA and SINGH 2008), exhibit rapid rates of divergence and attain extreme elaboration (KIRKPATRICK 1982; EMLEN 2008). Multivariate sexual selection however commonly exhausts genetic variance for trait combinations under selection and constrains further evolution of individual sexually selected traits (CHENOWETH and MCGUIGAN 2010). Antagonistic pleiotropic covariation of sexual

traits with nonsexual fitness also can limit their evolution despite the existence of standing genetic variance for such traits (VONSCHANTZ *et al.* 1995).

From the perspective of molecular genetics, the dissection of sexually selected traits is interesting because the loci under adaptation via natural selection may evolve under very different processes compared with those also affected by sexual selection. For example, although balancing selection can occur under both viability (HEDRICK 2006) and sexual selection (BROOKS 2002; CHUNCO *et al.* 2007), genetic polymorphism may develop at loci that are under opposing natural and sexual selection or even sexually antagonistic selection (CHENOWETH and MCGUIGAN 2010). For example, in the Soay sheep, male horn length variation is controlled by a single locus and a polymorphism is maintained by a trade-off between sexual selection and viability selection (JOHNSTON *et al.* 2013). In *D. serrata,* male attractiveness genes for cuticular hydrocarbon (CHC) blends that evolve through sexual selection via female choice are not favoured by natural selection for enhanced productivity (SKROBLIN and BLOWS 2006). Changes in genes and gene pathways that coordinate sex-specific development may lead to the divergent evolution of sexual traits (CHENOWETH and MCGUIGAN 2010). The fruitless (*fru*) and double-sex (*dsx*) *genes* in *Drosophila* are vital for the development of the brain region associated with male sensory processing and courtship behaviour (KIMURA *et al.* 2008). The *fru* gene affects the neural coordination of wing song and its male-specific expression leads to females lacking the song trait (CLYNE and MIESENBOCK 2008). In *D. melanogaster,* male courtship relies on CHC diene pheromones that are underlain by *desatF*; female-specific expression of this gene is *cis*-activated by the female *dsx* transcript (SHIRANGI *et al.* 2009). Changes in *dsx* regulation of *desatF* correlate with interspecific variation in sexual dimorphism of dienes (SHIRANGI *et al.* 2009).

One potential model system for dissecting divergence in sexually selected traits is the CHCs of *Drosophila serrata*. *D. serrata* and *D. melanogaster* populations along the eastern Australian coast exhibit clinal variation in their cuticular hydrocarbon profiles that is associated with variation in temperature and humidity (FRENTIU and CHENOWETH 2010). The occurrence of such parallel geographic patterns in CHC traits in different species in shared environments is indicative of the action of selection (GUO *et al.* 1991; WU *et al.* 1998). Q_{ST}-F_{ST} comparison has shown that divergent selection more so than drift has likely caused changes in *D. serrata* CHC traits along the cline (CHENOWETH and BLOWS 2008).

The molecular genetic basis of *D. serrata* CHCs is currently unknown but the system presents a rare opportunity for dissecting the genetic basis of a geographically variable set of sexually-selected traits that are also under the influence of natural selection. We have recently discovered a major polymorphism affecting these sexually selected traits that is only present in northern, tropical populations, but through its phenotypic effects steepens latitudinal clines in these traits. Below, I briefly discuss the nature of cuticular hydrocarbons in insects and what is known about their biosynthesis before outlining my general approach to dissecting the genetic basis of the major polymorphism in *D. serrata* CHCs and determining its fitness effects.

1.4. ECOLOGICAL ROLES, BIOSYNTHESIS AND GENETIC BASIS OF CUTICULAR HYDROCARBONS IN INSECTS

1.4.1 *Ecological roles*

Insect hydrocarbons are waxy compounds deposited on the body surface of insects that have diverse biological roles (SCHAL *et al.* 1998; HOWARD and BLOMQUIST 2005). CHCs are important in maintaining water balance through reducing water loss and keeping water out of spiracles (BLOMQUIST *et al.* 1993; SCHAL *et al.* 1998; MONTOOTH and GIBBS 2003) but also function as pheromones, kairomones or allomones and mediate predator-prey, parasitoid-host and defence interactions (SCHAL *et al.* 1998; HOWARD and BLOMQUIST 2005; LANG and MENZEL 2011). They are also involved in chemical mimicry and camouflage (HOWARD and BLOMQUIST 2005).

Specific components of the complex blends of CHCs serve as both intra- and inter-specific recognition signals between individuals, species, genders, colonies, nests, castes and broods (HOWARD and BLOMQUIST 2005). Mate recognition and gender delineation roles are played out through sexual dimorphism in these hydrocarbons (GINZEL *et al.* 2003) so that only one gender of the relevant species produces a particular blend of hydrocarbons (NELSON and CARLSON 1986; GINZEL *et al.* 2003). When the compounds produced are not gender-specific, it is the sex-differences in relative quantities of compounds that are important (JALLON and COBB 1987; HOWARD 1998; HOWARD and BLOMQUIST 2005).

1.4.2 *Biosynthesis*

Insect CHCs and CHC-derived sex pheromones, a complex mixture of straight-chain, methyl-branched and unsaturated components (LOCKEY 1988), are synthesised by large epidermal cells called oenocytes that are rich in smooth endoplasmic reticulum and mitochondria (BLOMQUIST *et al.* 1993; MONTOOTH and GIBBS 2003). The CHCs are then selectively transported through an aqueous medium to the epicuticle, fat body and other internal deposition sites, including the ovaries and gonads by a multi-functional high-density haemolymph lipoprotein called lipophorin that serves as a juvenile hormone binding protein in many insects (GU *et al.* 1995; SCHAL *et al.* 1998).

CHCs are synthesised through elongation of fatty acyl-CoAs when fatty acids condense head-to-head, reduction of resulting fatty acid chains to aldehydes and eventual fatty aldehyde conversion to one-carbon-shorter n-Alkanes and n-Alkenes by loss of the carboxyl group (JURENKA *et al.* 1987; BLOMQUIST *et al.* 1993; MPURU *et al.* 1996; HOWARD and BLOMQUIST 2005). The reduction of long-chain acyl-CoAs to aldehydes and their decarbonylation to form alkanes occurs either under anaerobic conditions without any cofactors or involves a cytochrome P450-mediated reaction that requires nicotinamide adenine dinucleotide phosphate (NADPH) and Oxygen (REED *et al.* 1994; MPURU *et al.* 1996). The resulting hydrocarbons have chain lengths ranging from 21 to over 40 carbons (MPURU *et al.* 1996), the majority of which have an odd number of carbons due to the decarboxylation of even-chain fatty acid precursors (RULE and ROELOFS 1989). CHCs may further be modified into more polar constituents like 3,11-dimethylnonacosane or hydroxylated to an alcohol and then oxidized into methyl ketones (CHASE *et al.* 1990). Fatty acyl-CoA elongases and chain-shortening enzymes determine the chain length while desaturases impact on the degree of saturation of resulting hydrocarbons (RULE and ROELOFS 1989; TILLMAN-WALL *et al.* 1992; CHERTEMPS *et al.* 2007b). Changes in chain-shortening enzymes may affect proportions of CHC components, sometimes leading to the divergence of hydrocarbon profiles among insect populations (WU *et al.* 1998).

In all insects, the pathway for producing fatty acyl-CoAs involves the metabolism of propionate and acetate (DILLWITH *et al.* 1982; VAZ *et al.* 1988); propionyl-CoA is initially dehydrogenated to acryl-CoA, then converted to 3-hydroxypropionyl-CoA before being eventually oxidised to form acetyl-CoA (HALARNKAR *et al.* 1986; BLOMQUIST *et al.* 1989;

HALARNKAR and BLOMQUIST 1989). Amino acids valine, isoleucine, methionine and succinate serve as the precursors to the propionate in the biosynthesis of methyl-branched hydrocarbons and their associated pheromone components (CHU and BLOMQUIST 1980; CHASE *et al.* 1990). In *D. melanogaster*, hydrocarbon biosynthesis involves the desaturation of palmitic acid to palmitoleic acid which is then elongated to dienic fatty acid carbon chains, direct precursors of hydrocarbons (CHERTEMPS *et al.* 2007b). In *Ostrinia* moths, all CHC precursors are also derived *de novo* from palmitic acid through enzymatic steps that include desaturation and chain-shortening (LASSANCE *et al.* 2013).

Elongation of 18:1-CoAs and 24:1-CoAs produces fatty acids that are precursors of methyl-branched alkanes (BLOMQUIST *et al.* 1980; HALARNKAR and BLOMQUIST 1989). The replacement of an acetyl-CoA group with a propionyl-CoA group (as methylmalonyl-CoA) at specific points during chain elongation may also yield methyl branched hydrocarbons as when happens when, for instance, malonyl-CoA, the two-carbon donor for chain elongation is replaced with methylmalonyl-CoA (GUO *et al.* 1991; MPURU *et al.* 1996). The methylmalonyl-CoA comes from amino acids valine, isoleucine and methionine or from the conversion of succinate (succinyl-CoA) into methylmalonyl-CoA by alimentary canal microorganisms (GUO *et al.* 1991; HOWARD and BLOMQUIST 2005). Leucine is a source of carbons for odd chain length 2-methyl hydrocarbons (CHARLTON and ROELOFS 1991).

The branching methyl group may be inserted early in chain elongation and not as the penultimate unit (HOWARD and BLOMQUIST 2005). While stearyl-CoA is elongated to a 26 carbon acyl-CoA that serves as a precursor to n-pentacosene, linoleoyl-CoA gets elongated to a 28 carbon-chain that is a precursor to the 27:2 hydrocarbons (VAZ *et al.* 1988). Although both 18:1-CoAs and 24:1-CoAs are elongated to a 28 carbon-chain backbone (TILLMAN-WALL *et al.* 1992), even-numbered carbon-chain methylalkanes tend to arise from the elongation of a valine carbon skeleton while odd numbered methylalkanes result from the elongation of isoleucine carbon backbones (BLAILOCK *et al.* 1976). A microsomal fatty acid synthetase (FAS) from the abdominal integument is involved in the production of methyl-branched fatty acid precursors while the cytosolic fatty acid synthetase aids in the production of straight chain hydrocarbons (BLOMQUIST *et al.* 1989; TILLMAN *et al.* 1999; HOWARD and BLOMQUIST 2005). Figure 1.1 below shows a typical insect hydrocarbon biosynthesis pathway as depicted in TILLMAN *et al.* (1999).

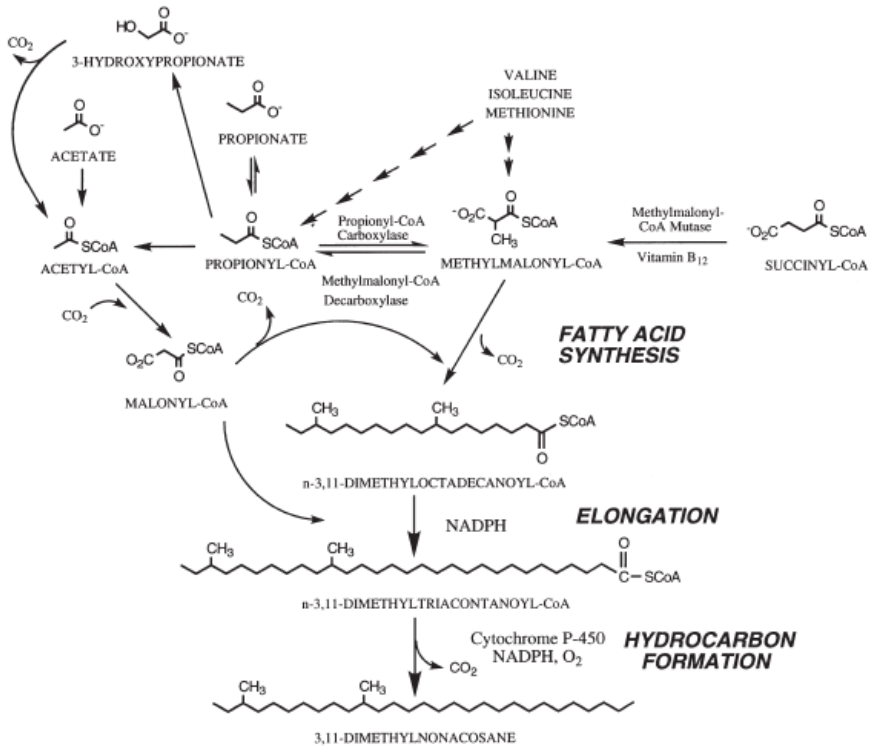


Figure 1.1

Hydrocarbon biosynthesis pathway as exemplified through production pathway of a 3,11-Dimethylnonacosane hydrocarbon in the German cockroach, depicting the roles of propionate, acetate, succinate as well as amino acids valine, Isoleucine and Methionine in the production of hydrocarbons (Adapted from Tillman et al. 1999).

1.4.3 *Genetic basis of CHCs in Drosophila*

A lot of work has been done on the genetic basis of CHC production in *Drosophila*. Although some studies report few QTLs that underlie substantial proportions of phenotypic variance for individual CHCs (GLEASON *et al.* 2005), others point to the possibility that the genetic architecture of CHC variation is highly polygenic (RULE and ROELOFS 1989; FOLEY *et al.* 2007; ETGES *et al.* 2009). Pleiotropic and epistatic interactions as well as QTL x environment interactions also feature in the genetic architectures of these traits (GLEASON *et al.* 2005; FOLEY *et al.* 2007; ETGES *et al.* 2009). Almost all of the four chromosomes of the Drosophila genome (X, 2, 3 and 4) have been implicated in CHC variation (RULE and ROELOFS 1989; ETGES *et al.* 2009; GLEASON *et al.* 2009) but the involvement of chromosome 3 is seemingly paramount.

In the *Drosophila melanogaster* complex, inter-specific variation in female-specific diene CHCs maps to five non-overlapping loci on the left arm of chromosome 3 (MARCILLAC *et al.* 2005a) while reduced levels of 7,11-heptacosadiene and elevated amounts of its isomer 5,9-heptacosadiene is driven by a single locus on the right arm of this same chromosome (COYNE *et al.* 1999). A locus on chromosome 3 also affects the balance between diene and monoene CHC production in female *Drosophila* (HOOPER *et al.* 1992; WICKER-THOMAS and JALLON 2001). The majority of QTLs for interspecific differences in male CHC pheromone concentration between *Drosophila simulans* and *D. sechellia* are on the third chromosome (CIVETTA and CANTOR 2003; GLEASON *et al.* 2005). There is also significant epistasis between QTLs on chromosome 3 for variation in heptacosadiene (GLEASON *et al.* 2005). The main changes in the proportions of 7-monoene hydrocarbons in *D. simulans* are, however, controlled by a locus on chromosome 2 (FERVEUR 1991). Epistatic interactions between chromosomes X and 2 mostly drive major disparities in the methyl-branched hexacosane: heptacosadiene ratios in both sexes of *D. pseudoobscura* and *D. persimilis* (NOOR and COYNE 1996). Male monoenes map to several loci spread across five major chromosomal arms in the *Drosophila melanogaster* complex (MARCILLAC *et al.* 2005a).

A number of genes have been implicated in the phenotypic variation of CHCs in *Drosophila* through changes in both their coding regions (DALLERAC *et al.* 2000) and their *cis*-regulatory regions (LEGENDRE *et al.* 2008; SHIRANGI *et al.* 2009). Two desaturase genes, *desat1* and *desat2* have been implicated in hydrocarbon polymorphism in *D. melanogaster* (DALLERAC *et al.* 2000) and *D. mojavensis* (ETGES *et al.* 2009). Changes in the *cis*-regulatory regions of a female-biased *desatF* gene that implicate the double-sex (*dsx*) gene affect CHC production in *D. melanogaster* (LEGENDRE *et al.* 2008; SHIRANGI *et al.* 2009). Sex determination candidate genes *doublesex* and *Voila* are associated with inter-specific differences in CHCs between *D. simulans* and *D. sechellia* (CIVETTA and CANTOR 2003) while the *Ngbo* gene locus affects changes in 7-monoene hydrocarbons in *D. simulans* (FERVEUR 1991). QTLs underlying CHC variation in *D. mojavensis* are linked to the *fruitless* gene (ETGES *et al.* 2009). The *Antennapediea* (*Antp*) homeotic gene drives the balance between diene and monoene CHC quantities in female *Drosophila* (HOOPER *et al.* 1992; WICKER-THOMAS and JALLON 2001).

1.5. AIMS OF THE STUDY

Within the northern populations of its Eastern Australian geographical distribution, *D. serrata* has recently been observed to exhibit extensive variation in three of its CHC compounds representing the shortest carbon chains; 5,9-tetracosadiene (5,9-C_{24}), 5,9-pentacosadiene (5,9-C_{25}) and 9-pentacosene (9-C_{25}) (FRENTIU AND CHENOWETH 2010). One class of phenotype only expresses these three compounds in trace amounts whereas the other has normal levels (further details provided in Chapter 2). These short-chained CHCs themselves exhibit strong genetically-based latitudinal clines up to but not beyond 20 degrees south of the equator (CHENOWETH AND BLOWS 2008). The genetic basis of this variation is presently unknown. Based on both traditional QTL and modern genomic mapping tools, the aim of this study is to dissect the genetic basis of the segregating factor contributing to clinal variation in 5,9-tetracosadiene, 5,9-pentacosadiene and 9-pentacosene and to examine its adaptive significance within the context of sexual and natural selection processes in *D. serrata*.

Chapter two traces the genetic basis of this trait difference through broad QTL mapping based in an F2 population resulting from a cross between two inbred lines, one a 'low' phenotype, and another a 'high' phenotype, from the opposite ends of the *D. serrata* CHC cline. This study sets the general framework for further fine mapping of the QTL affecting the major CHC polymorphism and also examines the possible pleiotropic effects these QTLs have on the longer chained, methy-branched alkanes, monoenes and dienes that are expressed in *D. serrata*.

The third chapter is based on one of the earliest protocols for mapping quantitative traits; bulk segregant analysis (BSA) (MICHELMORE *ET AL.* 1991), but employs next generation DNA sequencing for genotyping. I perform BSA at F60 in an advanced hybrid population founded from the same 'low' and 'high' inbred lines that were used for the QTL work in chapter 2. DNA pools of low and high CHC phenotypes were sequenced with Illumina HiSeq technology and SNP frequency differences analysed to locate potential candidate regions coinciding with these disparities.

The fourth chapter attempts to validate a 20kb QTL region implicated in chapter 3 as harbouring the major effect polymorphism. I use an independent sample of genetic variation from the northernmost population along the cline where the 'low' phenotype occurs with

highest frequency. First, I tested the hypothesis that the 'low' phenotype is likely to be underlain by a common mutation. As the mutation was recessive, I performed complementation tests by making F1 crosses between multiple independent 'low' lines sourced from the same local population and the 'low' line used to found all mapping crosses. Confirming that they did not complement the 'low' phenotype, I then conducted genome re-sequencing of multiple 'low' and 'high' inbred *D. serrata* lines in a bid to further validate findings of the bulk segregant analysis.

Chapter 5 begins to identify the likely sources of natural and sexual selection acting on the major effect QTL that could lead to its maintenance in far-northern populations. I assayed desiccation resistance and heat stress survival on multiple lines of contrasting phenotypes from the same far northern population. I further assayed male and female mate choice on the same lines to probe the potential role of sexual selection in influencing divergence.

The final chapter discusses the general findings of the study and suggests further avenues of research in light of these findings. By combining both established traditional QTL mapping protocols and modern sequence-based genomic approaches, this study provides a detailed account of the genetic underpinnings of natural divergence in a trait subject to contrasting natural and sexual selection.

CHAPTER 2 GENETIC ARCHITECTURE OF A POLYMORPHISM AFFECTING CLINAL DIVERGENCE IN THE SEXUALLY-SELECTED CUTICULAR HYDROCARBONS OF *DROSOPHILA SERRATA*

2.1. ABSTRACT

In the fruit fly *Drosophila serrata,* three short-carbon-chained cuticular hydrocarbons (CHCs), 5,9-tetracosadiene, 5,9- pentacosadiene and 9-pentacosene show a genetically-based latitudinal cline along the eastern coast of Australia with trait expression decreasing dramatically towards the equator. The cline in these traits has previously been shown to be a consequence of divergent natural selection. Recently, a major polymorphism was discovered in these CHCs that dramatically reduces their expression. As the polymorphism only occurs in northernmost populations it effectively steepens the existing latitudinal cline. The genetic basis of this polymorphism, which splits individuals into two distinct phenotypic classes, has not yet been established. In this chapter, QTL mapping of these and a further five CHCs was conducted on an F2 mapping population produced from a cross between inbred 'low' (Cooktown, CTN42) and 'high' phenotype (Forster, FORS4) lines spanning the CHC cline. When analysed as a binary trait, the polymorphism mapped to two major-effect QTLs on chromosome 3R, with the 'low' alleles having recessive effects. The same two QTL locations were implicated when the three constituent short-chained CHCs were analysed individually. Univariate QTL mapping of all eight individual *D. serrata* CHCs revealed a total of 22 QTLs, with a maximum of three QTLs per trait. The QTLs were individually responsible for 5 to 31% of total phenotypic variance in the eight CHCs. QTL intervals for most traits overlapped; likely suggesting that pleiotropy may be a common element to CHC genetic architecture. While the 'low' allele at the QTLs underlying the polymorphism in short-chained CHCs decreased trait values, co-localising QTLs for longer-chained CHCs had opposite effects, effectively increasing trait expression. Although few genomic regions of major effect seem to drive CHC divergence along the cline, more fine mapping work is needed to confirm whether these effects accrue from pleiotropy or from clusters of multiple linked loci.

2.2. INTRODUCTION

Investigating the mechanisms that lead to and sustain phenotypic diversity is a central preoccupation in evolutionary ecology (TOBLER and CARSON 2010). Trait variation underlain by within-population genetic polymorphisms may lead to character divergence among populations, potentially making it an important component of species formation (SMITH *et al.* 2001; MULLEN *et al.* 2009; WAGNER *et al.* 2012). Environmentally-driven divergent selection sometimes overrides the homogenizing effects of gene flow to cause population divergence that culminates in gradual variations in phenotype along geographical gradients, called clines (HUXLEY 1938; HALDANE 1948; SALOMON 2002). Clinal patterns of population divergence may also be spurred by neutral secondary contacts among populations that previously differentiated in allopatry, the extent of the cline being indicative of the degree of such differentiation (ENDLER 1973; TURGEON and BERNATCHEZ 2001; BERMOND *et al.* 2012). For instance, clinal polymorphism at neutral nuclear loci of admixed freshwater *Coregonus artedi* fish races persists across a range of postglacial lakes that once existed as multiple isolated refugia in north America (TURGEON and BERNATCHEZ 2001). Multiple successive founder events from genetically distinct demes may also lead to genetic drift overriding migration to maintain patterns of isolation-by-distance across geographical or environmental gradients (VASEMAGI 2006; KELLER *et al.* 2009). Where clines are adaptive, one trait value is favoured in one part of the species range while a different value prevails in another (ENDLER 1977; MULLEN and HOEKSTRA 2008). Clinal phenotypic variations of this nature are common and have been observed in a diverse range of phenotypes such as thermal tolerance, phenological traits, dispersal capacity and life-history traits (HOFFMANN *et al.* 2002). For example, migration timing in the Chinook salmon, *Oncorhynchus tshawytscha*, varies latitudinally through selection and adaptation to seasonal ecological factors (O'MALLEY and BANKS 2008). Adaptive latitudinal variations in leaf angle and morphology, seedling growth and flowering time occur in *Arabidopsis thaliana* (HOPKINS *et al.* 2008) as do clinal polymorphisms for metabolic enzymes in *Drosophila melanogaster* (OAKESHOTT *et al.* 1982).

Drosophila species are well known to exhibit parallel latitudinal body size clines across multiple continents; an observation that provides a strong inference of divergent natural selection between latitudes (VAZ *et al.* 1988; GILCHRIST *et al.* 2001; HALLAS *et al.* 2002; DE JONG and BOCHDANOVITS 2003). Clinal variation in other different morphological, physiological, chemical and behavioural traits are also common in *Drosophila* (BEAVIS *et al.*

1994; HALLAS *et al.* 2002; SCHMIDT *et al.* 2005; HOFFMANN and WEEKS 2007). One such pattern of repeated geographical variation occurs for cuticular hydrocarbons (CHC) (LOCKEY 1988; WU *et al.* 1998; FRENTIU and CHENOWETH 2010). Clinal geographical variation for *D. melanogaster* CHC profiles occurs in Europe, Africa and America, where the abundance of longer carbon chain compounds increases towards the equator, both in the northern and the southern hemispheres (WU *et al.* 1998). Short chained hydrocarbons decrease towards the equator whereas the production of longer chained CHCs increases. In Australia, similar long-chain/short-chain latitudinal variations in CHCs have been detected in co-distributed populations of *D. serrata* and *D. melanogaster* (FRENTIU and CHENOWETH 2010) spanning both temperate and tropical climates (JENKINS and HOFFMANN 2000; LIEFTING *et al.* 2009). Given their independent population histories, *D. serrata* is endemic whereas *D. melanogaster* colonised Australia only 200 years ago, a shared axis of divergence strongly suggests a common response to climatic selection of some form.

Insect cuticular hydrocarbons are synthesised in large epidermal cells called oenocytes through elongation desaturation and reduction of fatty acyl-CoAs and the eventual conversion of the resulting fatty aldehydes into one carbon shorter n-Alkanes and n-Alkenes of chain length ranging from 21 to over 40 carbons (JURENKA *et al.* 1987; BLOMQUIST *et al.* 1993; MPURU *et al.* 1996). These CHCs are then selectively transported to internal deposition sites and the epicuticle where they mediate some ecological functions (GU *et al.* 1995; SCHAL *et al.* 1998). In *Drosophila,* CHCs have been known to play a part in the elicitation of courtship as pheromones (FERVEUR 1997) and in the maintenance of water balance through reducing water loss (IRAQI *et al.* 2000). Clinal changes in *Drosophila* CHC blends correlate with changes in environmental variables such as humidity and temperature (WU *et al.* 1998; FRENTIU and CHENOWETH 2010) and have been associated with desiccation tolerance in a number of correlational and experimental studies (BLAILOCK *et al.* 1976; LOCKEY 1988; BERRY and KREITMAN 1993; TILLMAN *et al.* 1999).

CHCs are by far the most heavily studied traits in *D. serrata,* (BLOWS and ALLAN 1998;CHENOWETH and BLOWS 2005;HIGGIE and BLOWS 2008;HINE *et al.* 2011). A Q_{ST}-F_{ST} study (SPITZE 1993) has ruled out genetic drift as the sole cause of clinal divergence in *D. serrata* CHCs implicating divergent selection (CHENOWETH and BLOWS 2008; FRENTIU and CHENOWETH 2010). The main CHCs that constitute the mate and species recognition system in *D. serrata* are 5,9-tetracosadiene (5,9-C_{24}), 5,9-pentacosadiene (5,9-C_{25}), 9-pentacosene (9-

C_{25}), 9-hexacosene (9-C_{26}), 2-methyl-hexacosane (2-Me-C_{26}), 5,9-heptacosadiene (5,9-C_{27}), 2-methyl-octacosane (2-Me-C_{28}), 5,9-nonacosadiene (5,9-C_{29}) and 2-methyl-triacontane (2-Me-C_{30}) (BLOWS and ALLAN 1998; HOWARD *et al.* 2003; HIGGIE and BLOWS 2007) (Figure 2.1). Specific blends of these CHCs confer high male attractiveness and respond rapidly to experimental manipulation of both sexual and natural selection (BLOWS 2002; RUNDLE *et al.* 2009). These compounds also evolve differently in correlation with variation in environmental temperature, rainfall and humidity (FRENTIU and CHENOWETH 2010) as well novel larval food environments under lab conditions (RUNDLE *et al.* 2005; FRENTIU and CHENOWETH 2010).

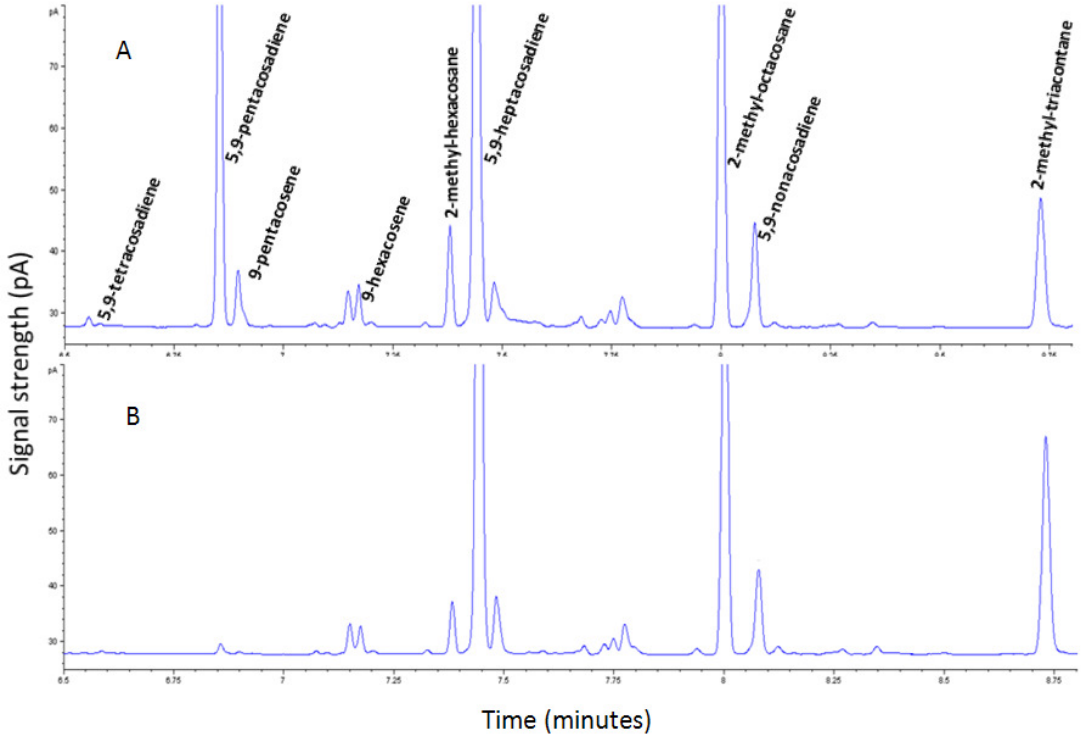


Figure 2.1

Gas chromatograph of nine cuticular hydrocarbons of increasing linear carbon chain length (from 24C to 30C) in *Drosophila serrata,* showing the high (A) and low (B) phenotype profiles. The X-axis shows the relative retention time (minutes) it took for a compound to pass through the gas-filled column. The Y-axis is a measure of signal strength with the area under the peaks proportional to the amount of each compound on a fly.

The species has recently been observed to exhibit contrasting clinal patterns in its CHC blend; three shortest-chained CHC compounds (5,9-tetracosadiene 5,9-pentacosadiene and 9-pentacosene) decline towards the equator while two longer chain compounds (5,9-heptacosadiene and 5,9-nonacosadiene) show the opposite trend, increasing towards the equator and decreasing away from it (Figure 2.2). Intriguingly, a polymorphism has been discovered in northern populations that decreases the population means for the short carbon chained 5,9-tetracosadiene, 5,9-pentacosadiene and 9-pentacosene, splitting them into two distinct phenotypes, one with dramatically reduced quantities of these compounds relative to the other (Figure 2.3). The frequency of this polymorphism itself also changes with latitude, rising toward the equator (Figure 2.4), an observation that suggests it may be sensitive to environmental climatic factors that vary with latitude.

Genetically-based clines offer an opportunity for unravelling the linkages between genetic variation on the one hand and quantitative trait variation on the other (WEEKS *et al.* 2002). A recent QTL study found that desiccation resistance QTL significantly colocalise with CHC QTL in *D. melanogaster* (RULE and ROELOFS 1989). Although genetically based clinal patterns in desiccation resistance (BLOWS 1993) and development time (HALLAS *et al.* 2002; SGRO and BLOWS 2004) have been reported in *D. serrata* populations, neither the molecular genetic basis nor the potential adaptive value of the clinal variation in the balance of short versus long chained CHCs as exhibited by this polymorphism is presently known. Here I use a QTL mapping approach to characterise the broad-scale genetic architecture of the major effect polymorphism in short-chained (less than 26 carbons) 5,9-tetracosadiene, 5,9-pentacosadiene and 9-pentacosene CHCs derived from a cross between two lines differing in these CHC phenotypes and also spanning the latitudinal cline for these CHCs. I also map QTLs for the longer carbon chain CHCs. This study represents an initial effort to understand the genetic basis of these interesting traits with dual biological functions that make them targets of both natural and sexual selection.

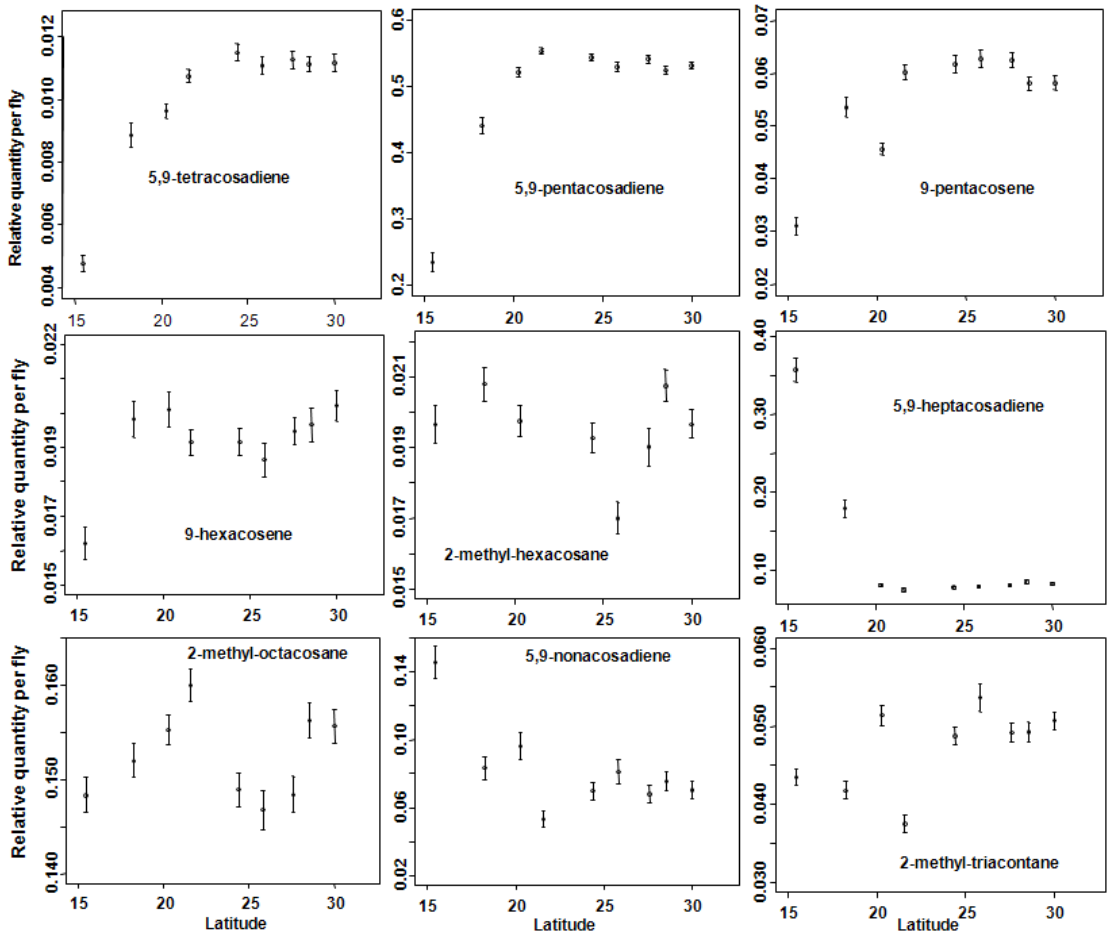


Figure 2.2

Latitudinal distribution of nine cuticular hydrocarbons of *Drosophila serrata* along the eastern coast of Australia showing opposing clinal patterns between three short-chained CHCs (5,9-tetracosadiene 5,9-pentacosadiene and 9-pentacosene) and other long carbon chain hydrocarbons (e.g. 5,9-heptacosadiene and 5,9-nonacosadiene). The Y-axis expresses the relative amount of each CHC per fly (with respect to the total amount of its CHCs). Means are shown +/- 1 standard error. Raw data were sourced from Chenoweth and Blows (2008) and traits were measured under common laboratory conditions and therefore represent genetic differences between populations.

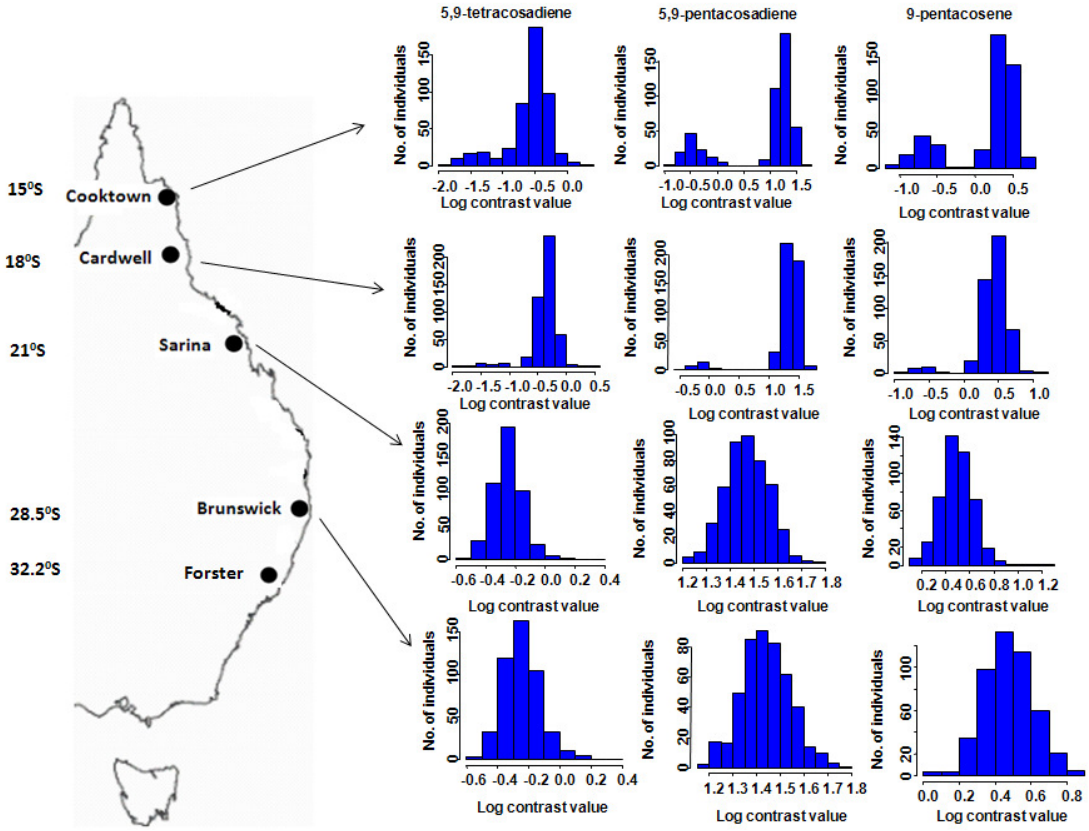


Figure 2.3

Changes in the phenotypic distribution for three short-chained CHCs along a latitudinal gradient in *D. serrata:* 5,9-tetracosadiene (left column), 5,9-pentacosadiene (central column) and 9-pentacosene (right column). A bimodal distribution was observed in the two northernmost populations, Cooktown and Cardwell whereas the distribution is Gaussian beyond 20^{0}S of the equator at Sarina and Brunswick Heads. The X-axis shows log-contrast values for these traits as defined in equation (1) of section 2.3.2. Raw data sourced from Chenoweth and Blows (2008).

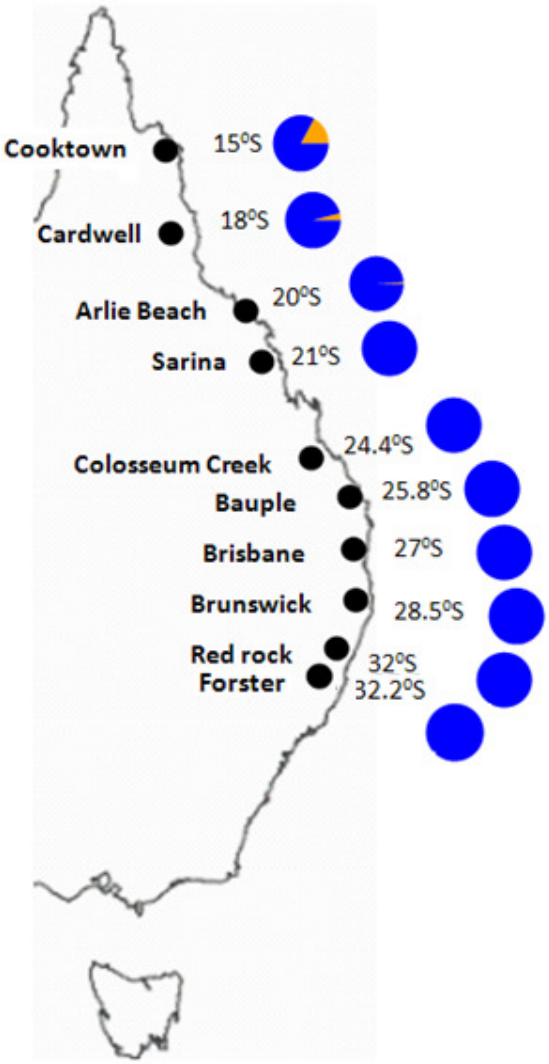


Figure 2.4

Latitudinal changes in the frequency of 'low' (orange) and 'high' (blue) short-chained CHC phenotypes in *Drosophila serrata* along the eastern coast of Australia. The 'low' phenotype has not been found beyond 20^0S of the equator. Raw data re-analysed from Chenoweth and Blows (2008).

2.3. METHODS AND MATERIALS

2.3.1 *Study populations and F2 cross design*

To create a mapping population, a reciprocal F2 intercross was made between two highly inbred (20 generations full sib mating) lines of *D. serrata* derived from two natural populations at two extremes of the species' eastern Australian distribution (Cooktown: CTN42, Forster: FORS4). CTN42 expresses a drastically reduced amount of 5,9-tetracosadiene, 5,9-pentacosadiene and 9-pentacosene relative to the FORS4 line. A large number of chromosomal inversions have been documented in this species (SLATKIN 2008) and so to minimise their potential effects in the QTL mapping, the founder lines were first confirmed to be homosequential using protocols based on polytene chromosome squashes as described by Stocker and colleagues (SLATKIN 2008).

2.3.2 *Phenotyping: CHC extraction and quantification*

Newly emerged virgin F2 flies were sexed (180 males, 203 females) and held singly in freshly yeasted vials for 5 days before being assayed for their CHC profiles. Cuticular hydrocarbons were extracted from individual flies by washing each fly in 100mL of hexane in a microvial insert for three minutes and then vortexing for one minute (HIGGIE and BLOWS 2007). The CHC samples were run on an Agilent Technologies 6890N gas Chromatograph (Wilmington, Delaware, United States). Individual fly CHC profiles were derived by integrating the area under the following nine peaks using the Agilent Chem Station Software (version Rev B.04.02), in the order of their retention times: 5,9-tetracosadiene, 5,9-pentacosadiene, 9-pentacosene, 9-hexacosene, 2-methyl-hexacosane, 5,9-heptacosadiene, 2-methyl-octacosane, 5,9-nonacosadiene and 2-methyl-triacontane (RUNDLE *et al.* 2009). Relative quantities of each of these nine chemicals were determined for each fly by dividing the area under each peak by the total area under all peaks to minimise sample-to-sample variation in total CHC extraction efficiency.

The CHC data were log-contrast transformed to remove the unit-sum constraint introduced by the compositional nature of the data (FISHER 1937; RUNDLE *et al.* 2005):

$$\text{Logcontrast } CHC_n = \log_{10}[\text{Prop}(CHC_n) / \text{Prop}(CHC_{\text{common divisor}})] \quad (1)$$

The proportional area of 9-hexacosene was used as a common divisor in the log-contrast transformation to finally produce eight log-contrast peak values for each individual fly. The use of log-contrasts here serves to aid comparisons with previous work in this species. One trait is effectively lost through the transformation (the common divisor) and so all further analyses are based of eight rather than nine CHCs. Statistical analyses were performed in R (TEAM 2008). Specific R packages used are cited when mentioned.

2.3.3 *SNP discovery and genotyping*

Illumina RNA-seq (35-bp reads) on cDNA created from RNA extracted from three-day-old adult flies from each of the two parental lines (CTN42, FORS4) was used for SNP discovery. RNA-seq data and library preparation was performed by GeneWorks Pty. Ltd. at Thebarton, Southern Australia (www.geneworks.com.au). Briefly, ESTs from the two lines were separately assembled using Edena 2.0 (ENDLER 1973) in "strict" mode with a minimum overlap value of 21 bp. This yielded 23,081 ESTs from CTN42 (N50 = 322 bp) and 922 from FORS4 (N50 = 276 bp). These ESTs were annotated using BLAST (ALTSCHUL *et al.* 1997) against the *D. melanogaster* genome according to the procedure followed by Frentiu and colleagues (FRENTIU *et al.* 2009). SNPs were then discovered by aligning the ESTs to each other using MUMmer version 3.2 (KURTZ *et al.* 2004). Locus-specific oligonucleotides were then designed for a subset of 65 SNPs using Primer3Plus (UNTERGASSER *et al.* 2007) and I validated these SNPs via bidirectional Sanger sequencing of PCR amplicons from CTN42 and FORS4 lines.

PCR reactions for the amplicons were performed in 25µl reactions on a BIO-RAD DNAEngine Peltier Thermal Cycler. The reaction mixture was made up of 5µl of 5M PCR buffer, 0.5µl of 10mM DNTP, 1.25µl of 2.5M MgCl2, 0.5µl of 10µM of forward and reverse primers respectively, 0.2µl of Taq polymerase, 16.05µl of ultra-pure water and 1.0µl of template genomic DNA extracted using a standard phenol-chloroform procedure (whole fly bodies). Amplification was performed with initial denaturing at 95 °C degrees for 10 min then 35 cycles of 30 seconds at 94°C for denaturing, 30 seconds at between 55 and 58°C (depending on primer pair) for annealing, 2 minutes at 72°C for elongation and a final step of 5 minutes at 72°C for elongation. The amplicons were sequenced at Macrogen (South Korea) and at the Australian Genome Research Facility (AGRF, Australia). SNP genotyping of individual flies on 61 SNPS was carried on a SEQUENOM© MASS ARRAY platform (AGRF) using two multiplexes of 30 and 31 SNPs, with approximately 10ng of genomic

DNA for each multiplex assay. Out of the 61 markers, three were discarded (s10, s36, and s37) because no difference in genotype calls was detected between the parental lines.

2.3.4 *Linkage map construction*

Joinmap 4.0 (VAN OOIJEN 2006) was used to assign markers to linkage groups beginning at LOD = 4.0 and ending at 10.0. Maintaining all other parameters as default, a least squares approach through the regression option in Joinmap was adopted for map construction with a Kosambi mapping function (CHUNCO *et al.* 2007). Linkage maps were plotted in MapChart (BROOKS 2002). Chi-square (χ^2) tests based on a significance threshold of $\alpha = 0.05$ were used to test for transmission ratio distortion (TRD) but no Bonferroni corrections were made because physical linkage between markers on the same chromosome nullifies the supposition of independence between tests.

Markers that exhibited TRD on the basis of χ^2 test results alone were not automatically excluded because many of the typed autosomal SNPs had nominally significant distortions from the Mendelian expectation of 1:2:1. It was also impossible to map large portions of the genome with the $\alpha = 0.05$ significance criterion without ending up with unusually short maps. Moreover, QTL mapping on the F2 cross was the goal of this study. QTL analysis is not necessarily negatively affected by distortion; in some cases distortion may actually increase mapping power (GU *et al.* 1995). The default method for assessing linkage in Joinmap 4.0 ("Independence LOD") is also not sensitive to TRD. Accordingly, a strategy that hinges on adopting a minor allele frequency (MAF) cut-off criterion for each SNP as a basis for marker exclusion, similar to that of Muchero and colleagues (MUCHERO *et al.* 2009), was used. In this study, 0.25 was the minor allele frequency cut-off threshold. Appendix 1 contains details of the identities of the 61 markers in question, with the excluded markers in shown bold.

2.3.5 *QTL mapping*

QTL mapping was performed in two ways, first, by analysing the CHC polymorphism as a binary trait and second, as most of the CHC traits in this study were correlated (Table 2.1), with some exhibiting contrasting clinal patterns (between short-chained and long-chained traits), I also performed univariate QTL mapping on all eight logcontrast CHC traits.

For the binary QTL analysis, I first recoded the data, classifying flies as either low phenotype (1) or high phenotype (0) on the basis of their 5,9-pentacosadiene phenotype. Of the three

CHC traits that diverged clinally towards the equator, 5,9-pentacosadiene revealed a dichotomous phenotypic pattern in the F2 mapping population with two distinct phenotypes; one with a smaller mean trait value of 0.07 ± 0.02 (n=84, range: from -0.32 to 0.53) than the other with a mean trait value of 1.07 ± 0.01(n=299, range: from 0.73 to 1.38)(t=51.2, p<0.0001) (Table 2.1). A 3:1 segregation ratio could not be rejected using Chi Square analysis (χ^2 = 2.181, p = 0.14), which is consistent with a single recessive segregating factor. The F2 progeny came from a reciprocal cross and contained male and female flies. I therefore first tested whether sex and cross had any effect on this polymorphism based on the phenotypic data of 5,9-pentacosadiene to determine if these were potential covariates for QTL mapping. In no case were these effects significant and so they were not fitted in subsequent analyses. Standard interval mapping in R/qtl was then performed using the scanone function and the model="binary" option. I performed QTL detection and then fitted multiple QTL models including interactions between them. Model simplification was performed following BROMAN and SEN (2009).

QTL mapping for the eight logcontrast CHC traits was performed using the composite interval mapping (CIM) method in R/qtl (BROMAN and SEN 2009). QTL scanning was conducted with a maximum of 4 marker covariates at a window of 2.5cM. I also mapped the three short-chained 5,9-tetracosadiene, 5,9-pentacosadiene and 9-pentacosene that were the basis of the CHC polymorphism using a subset of the data from high individuals only, i.e. after removing low flies from the dataset. This was done to assess whether major QTLs affecting the polymorphism might also affect regular trait variation, beyond the "knock down" effect. In both binary and trait-based analyses, significance thresholds (95%) for the detected QTLs were determined through permutation (n=1000) tests (DOERGE and CHURCHILL 1996). As with the binary analysis for each trait I performed QTL detection and then fitted multiple QTL models including interactions between them. Model simplification was performed following BROMAN and SEN (2009). The parental lines used to create an F2 intercross mapping population differed in mean for five of the eight CHCs mapped. No significant differences were found for the methyl-branched alkane traits (Table 2.2); however these were still mapped.

Table 2.1

Phenotypic correlation matrix of eight log contrast CHCs in an F2 mapping population of *D. serrata*. Shown are Pearson *r* values (n=383). Significant values (p<0.05) are indicated an asterisk.

	1	2	3	4	5	6	7	8
1. 5,9-tetracosadiene	1.00							
2. 5,9-pentacosadiene	0.87*	1.00						
3. 9-pentacosene	0.84*	0.93*	1.00					
4. 2-methyl-hexacosane	0.57*	0.56*	0.60*	1.00				
5. 5,9-heptacosadiene	-0.35*	-0.41*	-0.32*	0.16*	1.00			
6. 2-methyl-octacosane	0.41*	0.37*	0.46*	0.86*	0.29*	1.00		
7. 5,9-nonacosadiene	0.35*	0.38*	0.38*	0.55*	0.15*	0.65*	1.00	
8. 2-methyl-triacontane	0.07	-0.03	0.10	0.51*	0.40*	0.81*	0.49*	1.00

Table 2.2

Differences in CHC expression between two parental lines of *D. serrata* (Cooktown, CTN42 and Forster, FORS4) intercrossed to create an F2 mapping population. Values are expressed as log contrasts. The three short-chained compounds involved in the major polymorphism are shaded in grey.

	POPULATION				
Log contrast CHC	**CTN42 (n=97)**	**FORS4 (n=114)**	**Df**	**t**	**p value**
5,9-tetracosadiene	-1.39 ± 0.08	-0.64 ± 0.04	156	-11.06	p<0.0001
5,9-pentacosadiene	-0.09 ± 0.10	1.17 ± 0.05	140	-8.39	p=0.0001
9-pentacosene	-0.25 ± 0.06	0.56 ± 0.03	150	-11.39	p<0.0001
2-methyl-hexacosane	0.45 ± 0.04	0.36 ± 0.06	197	1.15	p=0.2529
5,9-heptacosadiene	1.52 ± 0.03	0.83 ± 0.04	210	13.64	p<0.0001
2-methyl-octacosane	1.07 ± 0.03	1.06 ± 0.04	200	0.10	p=0.9195
5,9-nonacosadiene	0.83 ± 0.04	1.03 ± 0.06	184	-2.76	p=0.0063
2-methyl-triacontane	0.51 ± 0.04	0.63 ± 0.05	207	-1.70	p=0.0902

2.4. RESULTS

2.4.1. *Linkage map*

The linkage map was 245.3cM long and consisted of 58 markers spread across four linkage groups, namely chromosomes X, 2, 3 and 4 (dot chromosome)(Figure 2.5). The chromosome number and arms were in agreement with that of *D. melanogaster* and for almost all the BLASTed marker regions the associated genes indicated strong conservation of chromosome arm gene content. A physical map of the *D. serrata* genome also confirmed strong conservation of chromosome arm level gene content between *D. melanogaster* and *D. serrata* (STOCKER *et al.* 2012).

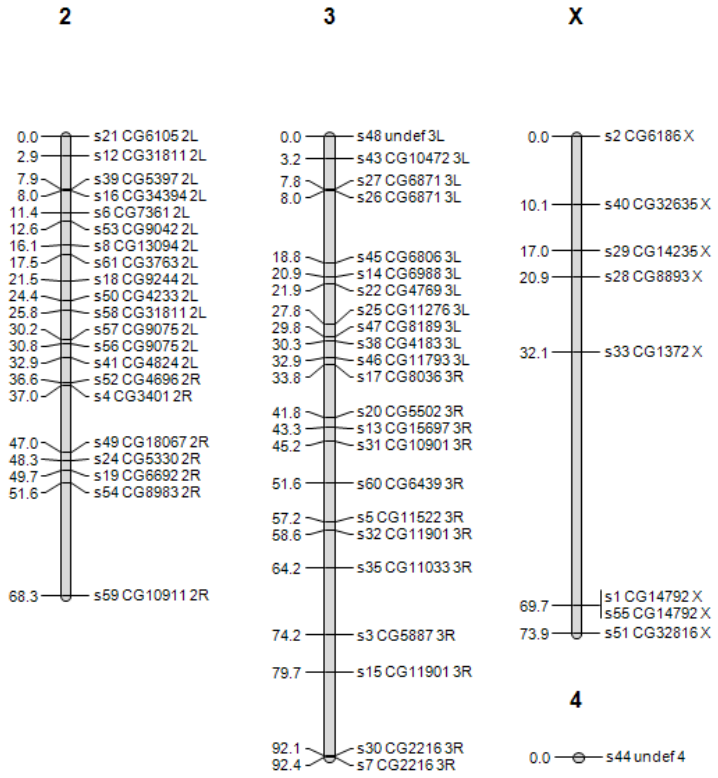


Figure 2.5

***Drosophila serrata* F2 linkage map showing the positions of EST-derived SNP markers spread across four chromosomes. Orthologous *D. melanogaster* gene names (inferred from the highest BLAST hit of the EST used for the SNP assay) and their chromosomal arm locations are indicated alongside the SNP labels used as marker names.**

2.4.2. *Binary CHC QTL analysis*

Sex and reciprocal cross had no effect on CHC polymorphism and were therefore excluded as covariates for QTL mapping (Sex: $F_{1, 379} = 0.27$, $p = 0.6020$, Cross: $F_{1, 379} = 0.26$, $p = 0.6121$, Sex x Cross: $F_{1,379} = 0.20$, $p = 0.6511$). Interval mapping of the CHC polymorphism as a binary trait detected two major QTL peaks on the right arm of chromosome 3 at positions 54cM and 74cM although the whole chromosome had a generally significant signal (Figure 2.6). These two QTLs were non-overlapping in their confidence intervals (Table 2.3). In the full model, the two QTLs and the interaction between them explained 71% of the phenotypic variance in the CHC polymorphism (Tables 2.3 and 2.4).

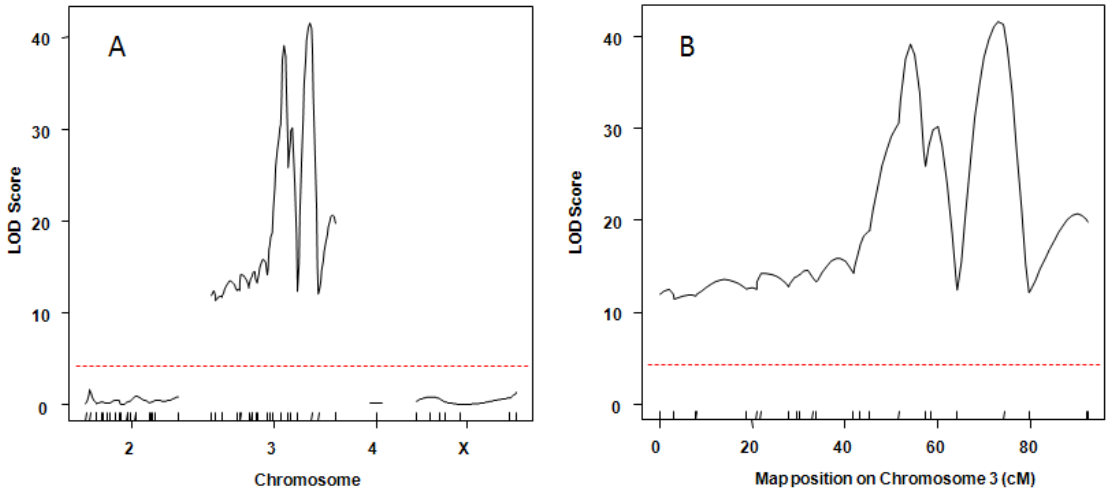


Figure 2.6

Interval mapping LOD profile for the short-chained CHC polymorphism modelled as a binary trait in an F2 intercross mapping population of *Drosophila serrata*. The red line on the graphs indicates permutation-based significance thresholds for QTL detection. Panel A shows the entire genome whereas panel B is for chromosome 3 only.

Table 2.3

QTL positions and effect sizes of a short-chained CHC polymorphism in an F2 mapping population of *D. serrata* detected via interval mapping as a binary trait. LOD peaks and ranges are indicated as 1.5 LOD intervals. %Vp is the percentage of the phenotypic variance explained by each QTL. Note that the %Vp values reported here are sourced from the multiple QTL model rather than a single-QTL model. The estimated variance explained by a QTL was higher for single QTLs model because multiple QTL models correct for linkage between QTLs whereas single QTL models cannot.

			MAP POSITION					
Trait	QTL	Chromosome	LOD Peak Genetic (cM)	LOD Peak Range Genetic (cM)	Nearest Marker	LOD	%VP	Effect of 'high' allele
Polymorphism	1	3R	54	53.0 – 55.0	s60	39.15	22.56	0.21
	2	3R	74	72.0 – 74.0	s3	41.65	18.74	0.18

Table 2.4

Significant interactions between two autosomal (Chromosomes 3R) QTL positions of a major short-chained CHC polymorphism in an F2 mapping population of *D. serrata*. An asterisk indicates significant QTL interactions ($p<0.05$). For this polymorphism, the phenotypic variance explained by the best multiple QTL model (listed interactions plus main effects) is also provided.

Trait	QTLs involved	F value	p-value	% Variance explained
Polymorphism	QTL1 x QTL2	$F_{4, 382} = 31.04$	0.0001*	9.48
	Full model	$F_{8, 382} = 117.00$	0.0001*	71.43

The QTL linked to the s3 marker was recessive, with the homozygote AA for the CTN42 allele expressing the 'low' phenotype while both the heterozygote AB and homozyote BB FORS4 allele had the 'high' phenotype. The double CTN42 homozygote appeared to have the greatest chance of expressing the 'low' phenotype. The effect of the QTL linked to the s60 marker appeared slightly less recessive in action (Figure 2.7).

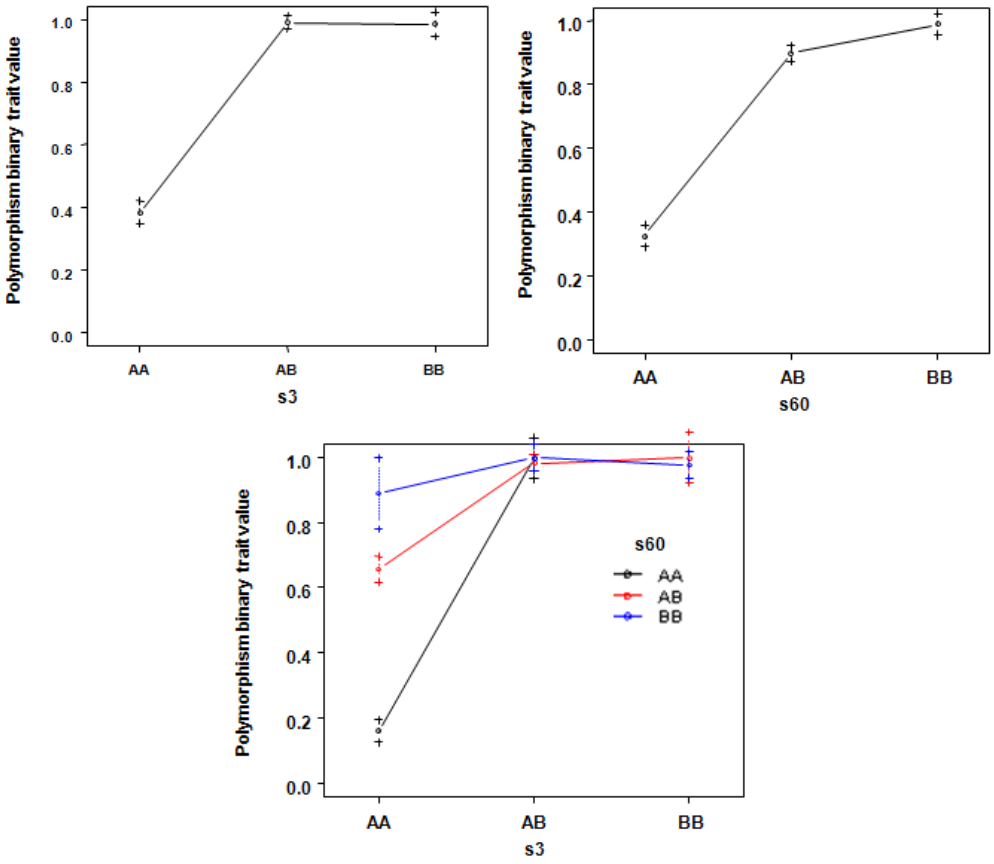


Figure 2.7

QTL marker effects of s3 and s60 on the short-chained CHC polymorphism modelled as a binary trait in an F2 intercross mapping population of *Drosophila serrata.* On the X-axis, A denotes the CTN42 ('low') allele and B denotes the FORS4 allele ('high').

2.4.3. *Univariate CHC QTL analysis*

QTL detection

A total of twenty two QTLs were detected for the eight CHCs and a maximum of three QTLs were detected per trait. These fell on chromosomes 2 and 3 and there were no QTL detected on chromosomes X and 4 (Figure 2.8 and Figure 2.9). The QTL locations plus their 1.5 LOD confidence intervals are shown in Table 2.5 and Table 2.6.

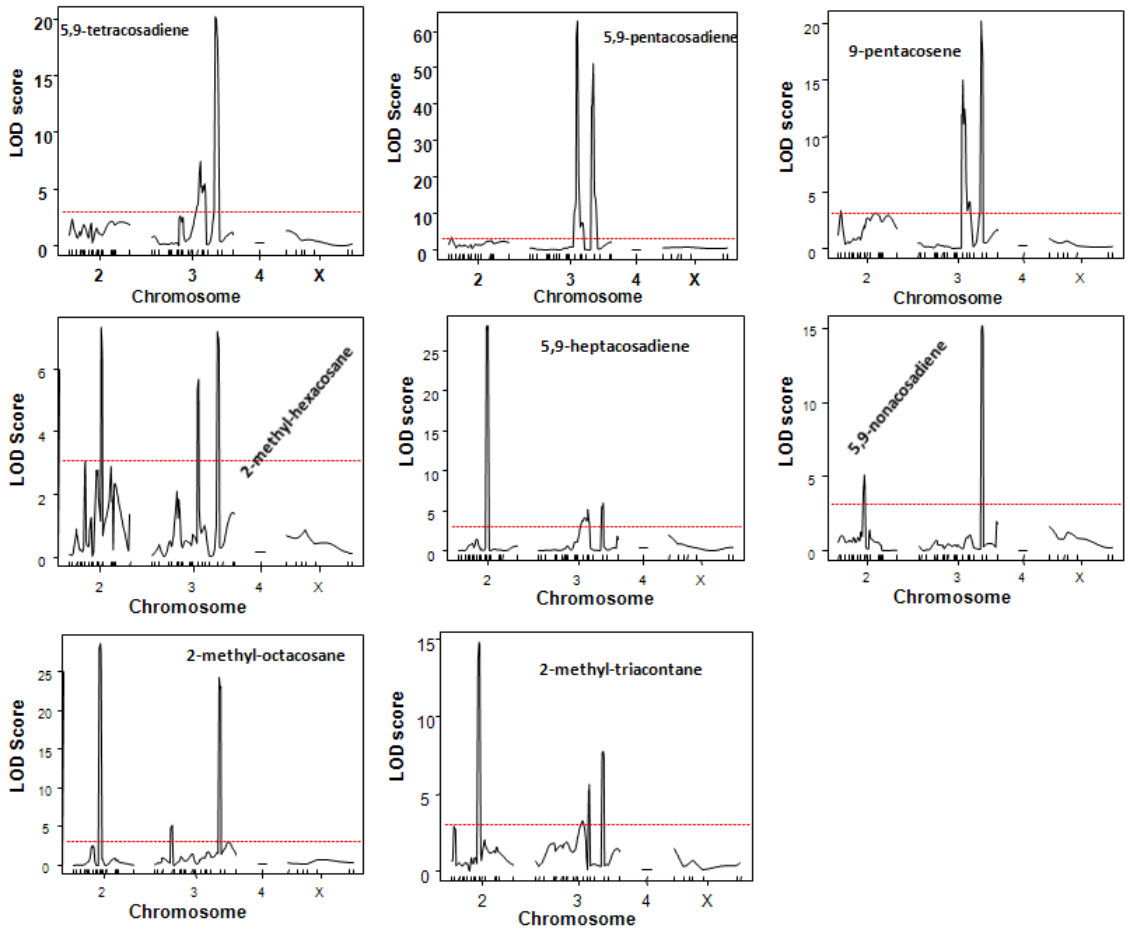


Figure 2.8

Composite interval QTL mapping LOD profiles for eight log contrast CHCs on four linkage groups (2, 3, 4 and X) in an F2 intercross mapping population of *Drosophila serrata*. The red line on the graphs indicates permutation-based significance thresholds for QTL detection. Each panel is a different cuticular hydrocarbon.

Table 2.5

QTL positions and effect sizes of eight log contrast CHCs in an F2 mapping population of *D. serrata* detected via composite interval mapping. LOD peaks and ranges are indicated as 1.5 LOD intervals. %Vp is the percentage of the phenotypic variance explained by each QTL for a trait. Note that the %Vp values reported here are sourced from the multiple QTL model for each trait rather than single-QTL models. The estimated variance explained by a QTL was higher for a single QTL models because multiple QTL models correct for linkage between QTLs whereas single QTL models cannot.

			MAP POSITION					
Trait	**QTL**	**Chromosome**	**LOD Peak Genetic (cM)**	**LOD Peak Range Genetic (cM)**	**Nearest marker**	**LOD**	**%Vp**	**Effect of high allele**
5,9-tetracosadiene	1	3R	55.0	53.0 – 56.0	s5	8.27	10.8	0.15
5,9-tetracosadiene	2	3R	74.0	72.0 – 75.0	s3	20.27	16.1	0.19
5,9- pentacosadiene	3	2L	3.0	1.0 - 4.0	s12	3.41	9.0	-0.07
5,9- pentacosadiene	4	3R	54.0	53.0 – 56.0	s60	62.75	21.9	0.24
5,9- pentacosadiene	5	3R	74.0	73.0 - 74.2	s3	51.08	20.4	0.21
9-pentacosene	6	2L	2.9	1.0 - 4.0	s12	3.72	9.8	-0.06
9-pentacosene	7	3R	55.0	54.0 – 56.0	s5	16.86	16.1	0.15
9-pentacosene	8	3R	74.0	72.0 – 75.0	s3	14.75	20.3	0.17
2-methyl-hexacosane	9	2L	36.0	35.0 – 38.0	s52	7.34	30.9	-0.14
2-methyl-hexacosane	10	3R	52.0	48.0 – 61.0	s60	4.77	5.3	0.05
2-methyl-hexacosane	11	3R	73.0	72.0 – 76.0	s3	7.17	5.7	0.07
5,9-heptacosadiene	12	2L	32.0	31.0 – 34.0	s41	28.14	26.0	-0.14
5,9-heptacosadiene	13	3R	54.0	51.6 - 56.0	s60	6.72	10.8	-0.08
5,9-heptacosadiene	14	3R	75.0	74.2 - 78.0	s3	6.06	11.7	-0.08
5,9-nonacosadiene	15	2L	30.2	28.0 - 32.9	s57	5.17	26.3	-0.12
5,9-nonacosadiene	16	3R	74.2	72.0 – 76.0	s3	15.19	16.3	0.10
2-methyl-octacosane	17	2L	30.2	28.0 – 32.0	s57	28.67	24.6	-0.27
2-methyl-octacosane	18	3L	20.9	18.0 – 25.0	s14	5.08	6.4	0.11
2-methyl-octacosane	19	3R	74.0	72.0 – 76.0	s3	24.42	19.2	0.26
2-methyl-triacontane	20	2L	32.0	21.0 - 36.6	s41	14.88	19.1	-0.11
2-methyl-triacontane	21	3R	54.0	51.6 – 56.0	s60	5.70	9.3	-0.06
2-methyl-triacontane	22	3R	74.0	71.0 – 76.0	s3	7.76	12.2	0.10

QTL effects

For each of the CHCs, the best multiple QTL model (QTL main effects plus significant interactions) explained between 32 and 73% of the phenotypic variance. There were significant interactions among QTLs for five of the eight traits that accounted for between 3 and 20% of the phenotypic variance (Table 2.6). Individually, the 22 QTLs were responsible for between 5 and 31% of phenotypic variance (Table 2.5). Most of the QTLs (16 out of 22) had major effects, accounting for at least 10% of the phenotypic variance in a trait.

There were differences between QTLs in the direction of their phenotypic effects. In slightly over half of the cases (12 out of 22), the 'high' FORS4 allele had positive additive effects on the CHCs. However, on chromosome 2, the 'high' allele had negative effects for all QTLs detected. By contrast, for the majority of the QTLs detected on chromosome 3 (12 out of 15) the FORS4 allele had positive additive genetic effects on their associated traits and the CTN42 allele was associated with trait decreases.

Table 2.6. Significant interactions among five autosomal (Chromosomes 2 and 3) QTL positions of eight CHCs in an F2 mapping population of *D. serrata*. For each trait the results shown are from the best multiple QTL model after model simplification. Significant QTL interactions (p<0.05) are indicated with an asterisk. For each trait the phenotypic variance explained by the best multiple QTL model is also provided.

TRAIT	QTLS INVOLVED	F VALUE	P-VALUE	% VARIANCE EXPLAINED
5,9-tetracosadiene	QTL1× QTL2	$F_{4, 382} = 9.34$	0.00001*	5.03
	Full model	$F_{8, 382} = 40.00$	0.00001*	49.71
5,9-pentacosadiene	QTL3× QTL4	$F_{12, 382} = 3.03$	0.0004*	2.78
	QTL3× QTL5	$F_{12, 382} = 6.37$	0.0001*	5.83
	QTL4× QTL5	$F_{12, 382} = 10.44$	0.0001*	9.56
	QTL3 × QTL4 × QTL5	$F_{8, 382} = 2.76$	0.0057*	1.69
	Full model	$F_{26, 382} = 35.50$	0.0001*	72.84
9-pentacosene	QTL6× QTL7	$F_{12, 382} = 1.97$	0.0261*	2.50
	QTL6× QTL8	$F_{12, 382} = 4.07$	0.0001*	5.10
	QTL7× QTL8	$F_{12, 382} = 5.96$	0.0001*	7.50
	QTL6× QTL7 × QTL8	$F_{8, 382} = 1.99$	0.0472*	1.70
	Full model	$F_{26, 382} = 24.75$	0.0001*	62.69
2-methyl-hexacosane	QTL9× QTL10	$F_{12, 382} = 0.93$	0.5123	1.59
	QTL9 ×QTL11	$F_{12, 382} = 0.53$	0.8984	0.89
	QTL10× QTL11	$F_{12, 382} = 1.06$	0.3937	1.80
	QTL9 × QTL10× QTL11	$F_{8, 382} = 0.50$	0.8559	0.57
	Full model	$F_{26, 382} = 11.00$	0.0001*	49.57
5,9-heptacosadiene	QTL12 × QTL13	$F_{12, 382} = 2.72$	0.0016*	3.90
	QTL12× QTL14	$F_{12, 382} = 3.25$	0.0002*	4.70
	QTL13× QTL14	$F_{12, 382} = 4.36$	0.0001*	6.32
	QTL12× QTL13 × QTL14	$F_{8, 382} = 3.30$	0.0012*	3.20
	Full model	$F_{26, 382} = 18.00$	0.0001*	56.98
5,9-nonacosadiene	QTL15 × QTL16	$F_{4, 382} = 0.06$	0.9930	0.04
	Full model	$F_{8, 382} = 47.00$	0.0001*	41.51
2-methyl-octacosane	QTL17 × QTL18	$F_{4, 382} = 0.86$	0.5910	1.38
	QTL17 × QTL19	$F_{12, 382} = 1.33$	0.1997	2.14
	QTL18 × QTL19	$F_{12, 382} = 1.86$	0.0383*	3.00
	QTL17 ×QTL18 × QTL19	$F_{8, 382} = 1.05$	0.3998	1.13
	Full model	$F_{26\ 382} = 14.88$	0.0001*	52.15
2-methyl-triacontane	QTL21×QTL22	$F_{12, 382} = 1.43$	0.1498	3.24
	QTL20×QTL22	$F_{12, 382} = 1.19$	0.2886	2.70
	QTL20× QTL21	$F_{12, 382} = 1.47$	0.1316	3.43
	QTL20×QTL21 × QTL22	$F_{8, 382} = 1.57$	0.1331	2.37
	Full model	$F_{26\ 382} = 8.00$	0.0001*	32.73

Overlap of CHC-QTL locations

There was considerable overlap in the genomic location of QTLs affecting different CHCs. For the major polymorphism the three correlated traits that are affected, 5,9-tetracosadiene, 5,9-pentacosadiene and 9-pentacosene, each had two QTLs that mapped to the same locations on chromosome 3R. This was true for both the full data set (Table 2.5) and the subset of data that excluded 'low' phenotypes for a major polymorphism in these traits (Table 2.6). Interestingly, these two locations coincided with QTL locations for the other long-carbon diene, 5,9-heptacosadiene (QTL 13 and QTL 14), a trait that exhibits an opposing clinal pattern to the three short-chain CHCs (Figure 2.9). Beyond the QTLs affecting the major polymorphism, only one QTL for 2-methyl-octacosane on chromosome 3L (QTL 18), stood out as truly independent. Each of the eight traits in this study had a QTL (QTL2, QTL5, QTL8, QTL11, QTL14, QTL16, QTL19 and QTL22) that mapped to the same overlapping region on chromosome 3R centred on 73cM (interval 72.0 – 78.0cM). Six of the eight traits also had a QTL each overlapping within the interval of 48.0 – 61.0cM on chromosome 3R (QTL1, QTL4, QTL7, QTL10, QTL13 and QTL21). There was a co-localisation of QTLs for 5,9-pentacosadiene and 9-pentacosene (QTL3 and QTL6) within the region 1.0 – 4.0cM on chromosome 2L. QTLs for four traits, 5,9-heptacosadiene (QTL12), 2-methyl-octacosane (QTL15), 5,9-nonacosadiene (QTL18) and 2-methyl-tracontane (QTL20) had overlapping intervals and co-localised on a region spanning from 21.0 to 36.6cM on chromosome 2L. There was also a slight overlap in the intervals of QTL20 for 2-methyl-tracontane (21.0 – 36.6cM) and QTL9 for 2-methyl-hexacosane (35.0 – 38.0cM).

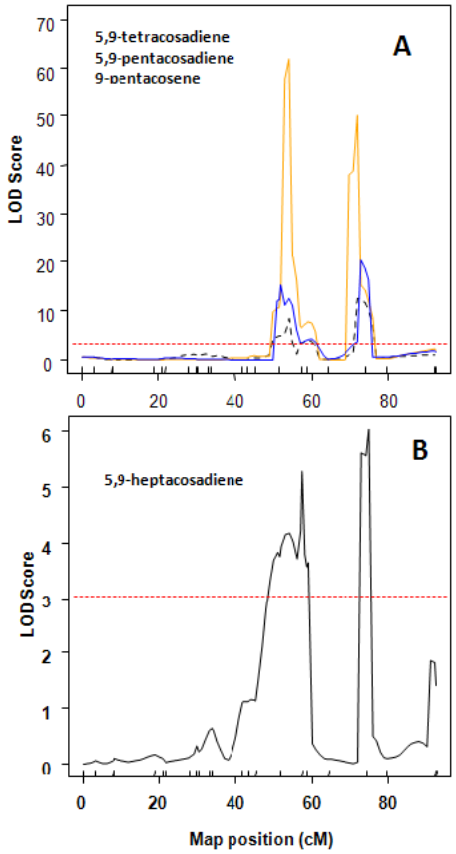


Figure 2.9

A). Co-localisation of QTLs on 3R for three short-chained CHCs. Shown above are 5,9-tetracosadiene (black dotted line), 5,9-pentacosadiene (yellow line) and 9-pentacosene (blue) which all underlie the 'low' phenotype. B). QTLs for the diene 5,9-heptacosadiene which shows an opposing clinal pattern to the three shorter-chained compounds also co-localised on 3R. The red line on the graphs indicates permutation-based significance thresholds for QTL detection.

2.4.4. *Remapping of CHC-QTLs after removing the 'low' phenotype individuals*

The short-chained CHC-QTL analyses were likely to be strongly dominated (very high LOD scores) by the major polymorphism leading to a bimodal distribution for the traits 5,9-tetracosadiene, 5,9-pentacosadiene and 9-pentacosane. For this reason, I conducted a complementary QTL analysis for these three traits with the 'low' phenotypes removed from the dataset. This analysis was used to investigate the possibility that the same QTL region

harbour mutations with both major (knockdown) and more minor to moderate effects. Six QTLs were detected for the three short-chained CHC on chromosomes 2 and 3 when these traits were mapped based on data from 'high' individuals only (Figure 2.10, Table 2.7). Five of these QTLs had overlapping ranges on two loci on chromosome 3R and contributed between 6 and 16% of phenotypic variance in these traits. The 'high' allele on all the QTLs on chromosome 3R had a positive additive effect on these CHCs. Their interactions, though significant, only contributed less than 2% of variance in the traits (Table 2.8).

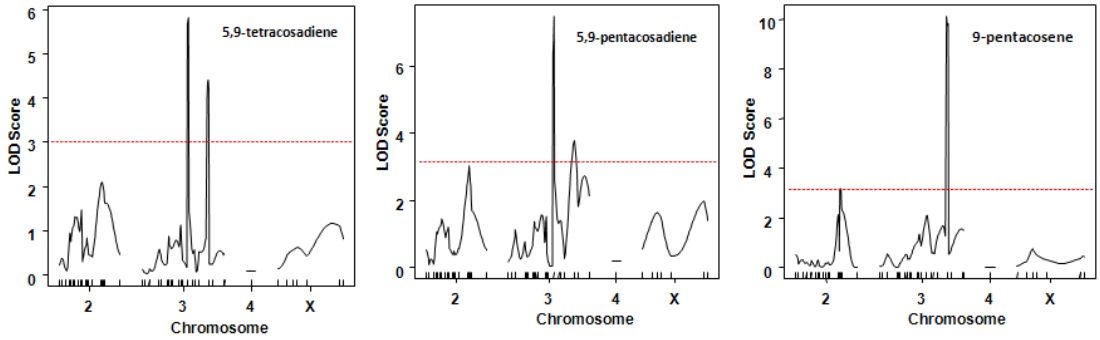


Figure 2.10

Composite interval QTL mapping LOD profiles for three log contrast short-chained CHCs on two linkage groups (2 and 3) in an F2 intercross mapping population of *Drosophila serrata* based on a subset of data that excluded 'low' phenotypes for a major polymorphism in these traits. The red line on the graphs indicates permutation-based significance thresholds for QTL detection. Each panel is a different cuticular hydrocarbon.

Table 2. 7. QTL positions and effect sizes of three log contrast short-carbon chained CHCs in an F2 mapping population of *D. serrata* based on composite interval mapping of a subset of data that excluded low phenotypes for a major polymorphism in these three traits. LOD peaks and ranges are indicated as 1.5 LOD intervals. %Vp is the percentage of the phenotypic variance explained by each QTL for a trait. Note that the %Vp values reported here are sourced from the multiple QTL model for each trait rather than single-QTL models.

			MAP POSITION					
Trait	QTL	Chromosome	LOD peak Genetic (cM)	LOD peak Range Genetic (cM)	Nearest marker	LOD	%VP	Additive effect of high allele
5,9-tetracosadiene	1	3R	52.0	46-56	s60	5.82	6.70	0.05
5,9-tetracosadiene	2	3R	74.2	69-78	s3	4.14	5.71	0.07
5,9- pentacosadiene	3	3R	52.0	48-55	s60	7.48	9.53	0.05
5,9- pentacosadiene	4	3R	74.0	68-89	s3	3.56	5.49	0.04
9-pentacosene	5	2	49.7	45-56	s54	3.21	4.41	-0.03
9-pentacosene	6	3R	73.0	72-76	s3	10.15	15.99	0.09

Table 2.8. Interactions between autosomal (Chromosomes 2 and 3) QTL positions of three CHCs in an F2 mapping population of *D. serrata* mapped with a subset of data that excluded 'low' phenotypes for a major polymorphism in these traits. For each trait the results shown are from the best multiple QTL model after model simplification. Significant QTL interactions ($p<0.05$) are indicated with an asterisk. For each trait the phenotypic variance explained by the best multiple QTL model is also provided.

Trait	QTLs involved	F value	p-value	% Variance explained
5,9-tetracosadiene	QTL25 x QTL26	$F_{4, 298} = 0.83$	0.5030	0.88
	Full model	$F_{8, 298} = 12.00$	0.0001	23.61
5,9-pentacosadiene	QTL27x QTL28	$F_{4, 298} = 1.25$	0.2887	1.28
	Full model	$F_{8, 298} = 35.50$	0.0001	26.08
9-pentacosene	QTL8 x QTL9	$F_{8, 298} = 0.43$	0.0261	0.49
	Full model	$F_{8, 298} = 7.50$	0.0001	17.26

2.5. DISCUSSION

The goals of this QTL study were two-fold. The first was to examine the broad-scale genetic architecture of a major CHC polymorphism in *D. serrata* that knocks down expression of three short-chained CHCs that themselves exhibit strong latitudinal clinal variation among far northern populations. The second goal was to explore the genetic basis of eight individual cuticular hydrocarbons that comprise the *D. serrata* CHC phenotype. I discuss each of these points separately below.

2.5.1 *Genetic basis of the CHC polymorphism*

With only two major effect QTLs detected through these analyses, comprising over 70% of the phenotypic variance in the binary-coded trait, the observed CHC polymorphism in this study seems to have a relatively simple genetic basis. There were slight differences in the observed level of dominance in the QTLs. The QTL at 74cM showed a near classic pattern of complete dominance, whereas the second at 54cM had slightly more additive effects. While it is still far too early to speculate about the specific genes involved, it is of note that similar major gene effects have been reported for comparable traits in *D. melanogaster*. For example, the divergence of 7,11-heptacosadiene and its isomer 5,9-heptacosadiene between African and Caribbean *D. melanogaster* populations appear to reside entirely on chromosome 3 and mapped to a single segregating factor (COYNE *et al.* 1999). This polymorphism was eventually traced to the desaturase gene, *desat*2 (DALLERAC *et al.* 2000). Interestingly, the nearest marker, s3 to the QTL at position 74.2cM resides in a *desaturase* gene, *desat*1 which is adjacent to the *desat*2 locus. However, the nature of the *D. serrata* polymorphism is quite different to that observed in *D. melanogaster*. There the polymorphism involved conversion of one diene to another (from 5,9- to 7,11-heptacosadiene) which is a clear signal of the involvement of a fatty acid *desaturase*. However in *D. serrata* the polymorphism affects CHCs depending on the chain length; both dienes and monoenes are affected. Elongases have been documented to be involved in the elongation of both dienic fatty acid precursors and, to a lesser extent, monoenes to very long fatty acids of up to C30 (CHERTEMPS *et al.* 2007b; WICKER-THOMAS *et al.* 2009). It may not necessarily be the case, therefore, that a *desaturase* is involved in this polymorphism because it is fatty acyl-CoA elongases and chain-shortening enzymes that affect hydrocarbon chain length (RULE and ROELOFS 1989; TILLMAN-WALL *et al.* 1992; CHERTEMPS *et al.* 2007b). Changes in biosynthetic pathways of such enzymes may affect the nature of CHC composition, occasionally leading to the divergence of hydrocarbon

profiles among insect populations (WU *et al.* 1998). It might be reasonable to expect that the genomic loci affecting this CHC divergence reside either within elongase genes or in other loci tightly linked to these genes.

When I analysed the three constituent short chain CHCs, 5,9-tetracosadiene, 5,9-pentacosadiene and 9-pentacosene, the same pattern was evident. Each of the traits had a QTL at the same genomic positions. One question of interest is whether the two QTL regions identified on 3R are affecting simply the "knock down" component of these traits or whether they also harbour variants that affect variation among high individuals. I explored this by removing all low phenotypes from the analysis and remapping QTLs for the three short chain traits. When this was done I recovered the same two major QTL locations, which suggests strongly that these genomic regions may harbour multiple CHC-affecting variants in addition to the knockdown mutation. It is not unusual for a single QTL region to contain multiple mutations that affect a phenotype (WANG *et al.* 2007). For example a major effect *desatF* locus affecting CHC diene polymorphism in *D. melanogaster* was observed to be due to different kinds of regulatory mutations when dissected upon finer mapping (SHIRANGI *et al.* 2009). It will be interesting upon fine mapping to see whether these major QTLs fragment into multiple smaller effect QTLs or map to different genes.

When spatial patterns of genetic polymorphisms mirror those of advantageous phenotypes, adaptations to continuous local environments may involve gradual changes in allele frequency across space (NOVEMBRE and DI RIENZO 2009). Spatially varying selection underlies a substantial part of the clinal phenotypic and genetic differentiation in *Drosophila* populations along the eastern Australian coast (OAKESHOTT *et al.* 1982; GOCKEL *et al.* 2001; HOFFMANN and WEEKS 2007; KOLACZKOWSKI *et al.* 2011). It is not yet known whether this clinal CHC polymorphism is underlain by a recent north-originating mutation that is spreading southwards. A population's phenotypic responses to various selective forces depend on the combined effects of both new mutations and its standing genetic variation (RIESEBERG *et al.* 2003; MASEL 2006; BARRETT and SCHLUTER 2008). Although standing genetic variation may confer large-effect adaptive genetic changes associated with rapid phenotypic divergence (COLOSIMO *et al.* 2005; BARRETT and SCHLUTER 2008; MILLER and VINCENT 2008; TOBLER and CARSON 2010), over the long term evolutionary innovations would be impossible without a continuous supply of new mutations (BARRETT and SCHLUTER 2008). The accumulation of nonsynonymous amino acid substitutions in a biosynthetic fatty-acyl reductase gene by

positive selection in *Ostrinia* moths, for instance, results in the divergence of species-specific CHC ratios (LASSANCE *et al.* 2013).

It is possible that this CHC polymorphism may be underlain by a novel mutation and that its geographic distribution has been shaped by recent migration. Local adaptation and establishment of clines have been linked with *de novo* genetic variance in a variety of organisms (FELDMAN *et al.* 2009; LINNEN *et al.* 2009). An advantageous phenotype underlain by a novel beneficial mutation, initially established within a single geographic location, may thus proliferate outwards in a 'wave of advance' of that favourable mutation (FISHER 1937; NOVEMBRE and DI RIENZO 2009). Some adaptive geographic phenotypic clines in *Drosophila* flies have been reported to evolve and spread rapidly (HUEY *et al.* 2000; GILCHRIST *et al.* 2001). A globally prevalent latitudinal cline in a selected polymorphism in the alcohol dehydrogenase (*Adh*) locus of *D. melanogaster* (OAKESHOTT *et al.* 1982; BERRY and KREITMAN 1993) that corresponds with adaptation to climate, for instance, has shifted within the past two decades in resonance with climate change (UMINA *et al.* 2005). It would be interesting to track changes in the extent of the cline in this CHC polymorphism over time.

2.5. 2. *A shared genetic basis to multiple D. serrata CHCs*

Variation in the eight CHCs mapped to a total of 22 largely overlapping QTLs, with all but one QTL-location affecting more than one trait. Genomic clustering of QTLs affecting functionally-related traits is a common phenomenon and has been reported for a diversity of traits in both plants (BURKE *et al.* 2002; CAI and MORISHIMA 2002; JUENGER *et al.* 2005; PEREZ-VEGA *et al.* 2010) and animals (PROTAS *et al.* 2008; ALBERT *et al.* 2009; WRIGHT *et al.* 2010). Such co-localisation of different trait QTLs to roughly the same sites on a linkage map may indicate a shared developmental or genetic architecture through either tight linkage of distinct trait loci or pleiotropy of QTLs whose traits are functionally associated (HERDER *et al.* 2006; EDWARDS and WEINIG 2011). As the traits in this study were highly correlated, their evolution would be expected to be inter-twined through indirect effects of the selection on correlated traits (LANDE 1980; FALCONER and MACKAY 1996).

When covariation of traits is coupled with extensive QTL co-localization, pleiotropy may be a real possibility (KENNEY-HUNT *et al.* 2008; EDWARDS and WEINIG 2011). Although inferences of the extent of pleiotropy may be exaggerated if they are solely based on overlapping confidence intervals of coincident QTLs (WRIGHT *et al.* 2010), it is difficult to

totally rule out the possibility of QTL pleiotropy for the CHC traits in this study. From the perspective of CHCs, pleiotropy for their biosynthesis has been implicated in the clustering of methyl-branched CHCs QTLs in genomic regions with high recombination rate in the jewel wasps, *Nasonia giraulti* and *Nasonia vitripennis*, (NIEHUIS *et al.* 2011). The *desat1* gene pleiotropically influences different pathways responsible for both the processing of sex-specific CHC pheromones and other physiological or ecological aspects that determine the overall levels of CHCs production in *D. melanogaster* (MARCILLAC *et al.* 2005a). Foley and colleagues showed that several QTLs pleiotropically influence CHC expression in both male and female *D. melanogaster* (FOLEY *et al.* 2007). Moreover, Etges *et al.* (2009) found that in *D. mojavensis*, where epicuticular hydrocarbon expression is underlain by multigenic influences and some epistasis, multiple QTLs behaved pleiotropically to effect CHC variation.

While the effects of Chromosome 3R QTLs on 5,9-tetracosadiene, 5,9- pentacosadiene and 9-pentacosene were positive, their effects on phenotypically correlated 5,9-heptacosadiene and 2-methyl-triacontane were negative. Although pleiotropy may sometimes compromise adaptive trait divergence for correlated phenotypes when a mutation that is beneficial for one trait is detrimental to another (LANDE 1980; SCARCELLI *et al.* 2007; EDWARDS and WEINIG 2011; WAGNER and ZHANG 2011), traits underlain by similar genes with correlated effects will respond more rapidly together relative to either of them separately (LANDE 1980; VIA 1984). Because pleiotropic linkages among traits may enhance their co-selection, pleiotropic effects are more likely among traits that evolve together under selection relative to those that are under independent selection (CHEVERUD *et al.* 2004; KENNEY-HUNT *et al.* 2008). Whether this assumption resonates with the three positively correlated short chain CHCs that diverged in a similar latitudinal manner and are the focus of this study remains to be empirically verified. The increase in the long chain CHCs that was associated with the knock down polymorphism in the three short chain CHCs could be adaptive as well.

Although major effects were detected for the majority of QTLs in this study, fine-scale high power QTL mapping sometimes reveals previously mapped major effect QTLs as clusters of multiple closely linked QTLs (CHRISTIANS and SENGER 2007; KAWAMURA *et al.* 2011). Changes in the recombination rate and reorganization of the physical distribution of locally adaptive alleles in the genome may establish concentrated architectures of tight linkage groups of several small-effect alleles that may when combined play a significant role in trait

changes (WRIGHT *et al.* 2010; YEAMAN and WHITLOCK 2011). Seemingly pleiotropic effects of QTLs may also reflect the existence of linked loci when QTL mapping is based on weakly resolved genetic maps that tend to bias mapping results towards pleiotropy simply because of lack power to reject its existence (LATTA and GARDNER 2009).

The mapping of CHCs to many QTLs on chromosome 3 is in accord with findings of several similar studies in *Drosophila* (COYNE *et al.* 1999; DALLERAC *et al.* 2000; TAKAHASHI *et al.* 2001; GLEASON *et al.* 2009; SHIRANGI *et al.* 2009). Hydrocarbon differences in *Drosophila* seem to be largely influenced by chromosome 3 loci (FERVEUR *et al.* 1996; WICKER-THOMAS and JALLON 2001; MARCILLAC *et al.* 2005a; WICKER-THOMAS *et al.* 2009) and their epistatic interactions (GLEASON *et al.* 2005). The majority of QTLs for interspecific differences in male CHC pheromone concentration between *Drosophila simulans* and *D. sechellia* have been reported to occur on this same chromosome (CIVETTA and CANTOR 2003; GLEASON *et al.* 2005). Although studies of gene expression, genome-wide association mapping in humans and genomic selection in agriculture show that the genetic basis of many complex traits is exceptionally polygenic (MCINNES and QUIGG 2010; ROCKMAN 2012), a large proportion of phenotypic differences among local populations in heterogeneous environments is often accounted for by only a few loci of moderate to large-effects (VOSS and SHAFFER 1997; COLOSIMO *et al.* 2005; BAERWALD *et al.* 2011). It appears the CHCs of *D. serrata* are similar in this regard.

The results of this study have shown potentially extensive pleiotropy and major effect QTLs underlying *D. serrata* clinal divergence along the East coast of Australia. However, because pleiotropy and linkage may both be implicated in major gene QTL effects, whether it is one or both of these factors driving the effects detected here is not presently clear. Higher-resolution mapping will be required to settle this question (CHRISTIANS and SENGER 2007; LATTA and GARDNER 2009). In the next chapter I apply such a high resolution mapping approach to the major CHC polymorphism affecting short-chained CHCs.

CHAPTER 3. FINE MAPPING A MAJOR CHC POLYMORPHISM IN *DROSOPHILA SERRATA* USING BULK SEGREGANT ANALYSIS OF WHOLE GENOME NEXT-GENERATION DNA SEQUENCE DATA IN AN ADVANCED INTERCROSS POPULATION

3.1. ABSTRACT

Parallel patterns of trait divergence are common in nature and may reflect shared adaptations to similar environmental challenges. Along the eastern Australian coast, parallel latitudinal clines occur in the cuticular hydrocarbons of *Drosophila serrata* and *Drosophila melanogaster* that are associated with environmental temperature, rainfall and humidity, suggesting a role for common environmental selection. These waxy compounds traits are vital for reducing water loss and also serve as contact sex pheromones. In *D. serrata*, a major polymorphism has been recently observed in 5,9-tetracosadiene, 5,9-pentacosadiene and 9-pentacosene that steepens the existing cline in these traits northwards. An initial F2 intercross QTL mapping investigation into the genetic basis of this polymorphism identified two major QTLs on the right arm of the third chromosome. In this experiment, the polymorphism was fine-mapped using bulk segregant analysis (BSA) on a highly advanced intercross population. Whole genome next generation DNA sequencing was used to genotype multiple DNA bulks of 'low' and 'high' flies sampled at F60. Alignment of sequence reads to a draft *D. serrata* genome sequence, revealed a single large QTL peak on the right arm of the third chromosome. Functional annotation of this region indicated that the extreme peak of the QTL (~20kb) contains three adjacent fatty acyl-CoA reductase genes. Nearby candidate desaturase and elongase loci appear excluded as major contributors to the CHC polymorphism. Reductases are involved in the terminal conversion of aldehydes into CHCs and their knock-down has been reported to result in substantial declines in cuticular hydrocarbon production, susceptibility to desiccation and reduced viability in *Drosophila*. Whole genome sequencing of advanced crosses provides a very powerful tool for fine mapping quantitative traits in non-model organisms.

3.2. INTRODUCTION

Genotypic changes in trait values may reflect adaptations to local environmental challenges (ENDLER 1986). While the genetic basis to most quantitative traits includes a large number of loci (MCINNES and QUIGG 2010), understanding the specific genes and mutations that are preferentially selected for during adaptation is fundamental in evolutionary genetics because it may shed light on how different types of molecular genetic variation underlying a trait fuel evolutionary change (MACKAY 2004; HOEKSTRA *et al.* 2006; STERN and ORGOGOZO 2008). Quantitative trait locus (QTL) analysis is often the first step in the dissection of genetic architecture (VOSS and SHAFFER 1997; HERDER *et al.* 2006). Although a useful strategy for the initial identification of genomic regions associated with trait variation, the detected QTL regions often span large portions of the genome and potentially harbour multiple causal mutations for the traits in question (STUDER and DOEBLEY 2011). This causes two problems, effect size overestimation and loss of detection of linked alleles with opposing effects. In early generation QTL mapping crosses (e.g. F2 intercross, and backcrosses), linked QTLs of small-effect can hardly be distinguished from a single QTL of large-effect because potentially linked alleles affecting the phenotype in the same direction, will manifest as one large-effect QTL, exaggerating estimates of individual QTL effect sizes in the process (BERTIN *et al.* 2012; HOBBS and MACRAE 2012). The cumulative influences of these less discoverable smaller effect QTLs may be particularly severe when QTLs reside in genomic regions of low recombination rate (TILLMAN-WALL *et al.* 1992). By contrast, multiple alleles which influence a phenotype, but in opposite directions within a QTL region, may be obscured because their effects essentially cancel each other out (BERTIN *et al.* 2012). Linkage maps that lack sufficient marker density may also seriously compromise the resolution of the genetic map and its power for locating QTLs (YOUNG 1996; HALL *et al.* 2010). It is therefore unexpected that that gene identification will result from a QTL study based on early generation intercrosses alone (FLINT 2011).

A potential solution to the problem of limited mapping resolution is the use of advanced intercrossing coupled with selective genotyping of individuals with divergent phenotypes (CHARLTON and ROELOFS 1991; HOBBS and MACRAE 2012; IRVING *et al.* 2012; OLSON *et al.* 2013). Furthermore the application of genome-wide genotyping-by-sequencing to such designs may further facilitate gene–level detection QTL by providing near-perfect marker coverage (GILCHRIST *et al.* 2001; CAO *et al.* 2011). Recent developments in next generation

sequencing technology allow for this approach to fine mapping and can deliver enhanced whole-genome marker coverage unthinkable under traditional QTL mapping protocols (HUEY *et al.* 2000; GOCKEL *et al.* 2001), speeding up the mapping of complex traits (GOCKEL *et al.* 2001; LASSANCE *et al.* 2010).

A particular application of this approach is the use of whole-genome re-sequence data (KOLACZKOWSKI *et al.* 2011) in the classic bulk segregant design (MICHELMORE *et al.* 1991). Analysis of allele frequency differences between DNA pools of individuals at opposing extremes of a traits' distribution can reveal the underlying genetic architecture; in genomic regions containing no QTL, allele frequencies should roughly be equal between the two DNA bulks but should otherwise differ in those regions harbouring causal loci (MICHELMORE *et al.* 1991). With large DNA bulks and sufficient sequencing depth, this technique is useful for detecting even small-effect QTLs (KOLACZKOWSKI *et al.* 2011). Bulk segregant analysis based on whole-genome sequence data may be useful for identifying causal polymorphisms underpinning adaptive change (KOLACZKOWSKI *et al.* 2011). For example, using this approach, Van Leeuwen and colleagues have recently revealed a nonsynonymous mutation in the major chitin synthase (CHS1) gene underlying pesticide resistance in the two-spotted spider mite, *Tetranychus urticae* (VAN LEEUWEN *et al.* 2012). The approach can also uncover differential gene expression patterns and has been useful for locating genes within QTLs for resistance to leaf rust in barley (CHEN *et al.* 2011).

Drosophila serrata Malloch is a fruit fly of the *montium* subgroup of the *melanogaster* species group and is endemic to both the temperate and tropical areas of eastern coastal Australia (MALLOCH 1927; DOBZHANSKY and MATHER 1961). Its cuticular hydrocarbons (CHCs) have been extensively studied and comprise 5,9-tetracosadiene, 5,9-pentacosadiene, 9-pentacosene, 9-hexacosene, 2-methyl-hexacosane, 5,9-heptacosadiene, 2-methyl-octacosane, 5,9-nonacosadiene and 2-methyl-triacontane (BLOWS and ALLAN 1998; HOWARD *et al.* 2003). In northern Queensland, *D. serrata* CHCs show a major polymorphism which splits 5,9-tetracosadiene, 5,9-pentacosadiene and 9-pentacosene into two distinct phenotypes; in one group these three CHCs are expressed at normal levels whereas in the other they are dramatically knocked down in expression. Although *D. serrata* CHCs have been known to evolve through both natural and sexual selection under experimental conditions (BLOWS 2002) and possibly in response to environmental factors in the wild (FRENTIU and CHENOWETH 2010), the genetic basis of this recently discovered major polymorphism remains

unknown. An initial QTL mapping experiment, of the genetic basis of this major polymorphism based on an F2 intercross, identified two major-effect QTLs on chromosome 3R (Chapter 2, this thesis). In this experiment, I fine map this polymorphism using bulk segregant analysis of whole genome Illumina DNA sequence data in DNA pools of phenotypically contrasting 'low' and 'high' flies.

3.3. METHODS AND MATERIALS

3.3.1 *Fly crosses and maintenance*

Two inbred lines of *Drosophila serrata* that differ in their CHC compounds were used to found the advanced intercross population. These are the same parental lines used to found the F2 intercross presented in Chapter 2. The FORS4 line shows the 'high' phenotype with normal levels of 5,9-tetracosadiene, 5,9-pentacosadiene and 9-pentacosene whereas the 'low' line, CTN42, expresses these compounds in only trace amounts. I performed reciprocal crosses to generate one very large population. The population was maintained in 32 glass bottles (500mL) for 60 generations with discrete generations. I favoured the "large outbred" approach over random crossing of pairs of flies every generation because a larger population size could be maintained under this set up which does not requires sexing of flies at each generation. Flies were maintained on a standard laboratory yeast medium, at 25^0C with a 12hours:12hours light: dark photoperiod for all 60 generations.

3.3. 2 *Phenotyping for CHCs*

I sampled flies for CHC phenotyping and subsequent bulked segregant analysis at F60. Newly emerged virgin flies were sexed (n = 600 female flies) under light CO_2 anaesthesia and held singly in food vials for 5 days before being assayed for their CHC profile through gas chromatography. Cuticular hydrocarbons were extracted from individual flies by washing each fly in 100mL of hexane in a microvial insert for 3 minutes and then vortexing for 1 minute (HIGGIE and BLOWS 2007). The CHC samples were run on an Agilent Technologies 6890N gas Chromatograph (Wilmington, Delaware, United States). Individual fly CHC profiles were derived by integrating the area under the following nine peaks, in the order of their retention times: 5,9-tetracosadiene, 5,9-pentacosadiene, 9-pentacosene, 9-hexacosene, 2-methyl-hexacosane, 5,9-heptacosadiene, 2-methyl-octacosane, 5,9-nonacosadiene and 2-methyl-triacontane (RUNDLE *et al.* 2005). Relative quantities of each of these nine chemicals were determined for each fly by dividing the area under each peak by the total area under all

peaks to minimise sample-to-sample variation in total CHC extraction efficiency. The CHC data were log-contrast transformed to remove the unit-sum constraint introduced by the compositional nature of the data (FISHER 1937; RUNDLE *et al.* 2005). The proportional area of 9-hexacosene, was used as a common divisor in the log-contrast transformation to finally produce eight log-contrast peak values for each individual fly. All the flies phenotyped were individually preserved in 95% alcohol and kept in a -80^{o}C freezer pending DNA extraction

3.3. 3 *DNA sequence pool selection*

A total of 85 'low' flies from the lower tail of the 5,9-pentacosadiene distribution and 85 'high' flies from the upper tail were chosen for DNA extraction (representing top and bottom 14%). The trait 5,9-pentacosadiene was chosen for delineating the 'low' and 'high' pools; it is the largest CHC peak in *D. serrata* and is the easiest to score. As with the F2 data, this trait was bimodal in the F60 sample indicating the presence of the major polymorphism. Two DNA bulks were made up each from the 'low' and 'high' flies, effectively resulting in two types of pooled DNA samples that were expected to differ genetically only for the loci of greatest effect on their phenotype status but theoretically undifferentiated for all other regions (MICHELMORE *et al.* 1991). The first technical replicate set of DNA pools was made using DNA extracted from fly heads while the second used DNA extracted from the rest of the bodies.

3.3.4 *DNA extraction and pooling*

Phenol-chroloform extraction was used to obtain fly genomic DNA. Briefly, 10 flies were put into one eppendorf tube containing 250uL of a solution made up of 10mL 1M stock of 0.1M Tris HCl (pH9.0), 20mL 0.5M stock of 0.1M EDTA and 10mL of 10% stock of 1% SDS solution. The flies were ground with a micro pestle and then incubated in water bath at 70^{o}C for one hour. After incubation, 35uL of Potassium Acetate was added to each eppendorf tube and the tubes were incubated on ice for 30 minutes. The tubes were then spun in a centrifuge at 13,000rpm speed for 15minutes after which the supernatant fluid was carefully moved into a new eppendorf tube, leaving behind all the precipitate. About 250uL of Phenol Chloroform Isoamyl Alcohol (25mL Phenol + 24mL Chloroform + 1mL Isoamyl Alcohol) were then added to and mixed with the extracted supernatant by shaking. The mixture was spun in a centrifuge for 5 minutes at 13,000rpm. The aqueous top layer of this mixture (about 240uL) was moved into another new eppendorf tube into which was added 240uL of Isopropanol. This mixture was centrifuged at 13,000rpm for 5 minutes, resulting in DNA pellets deposited

at the bottom of the tube. After the supernatant was removed, the remaining pellets were centrifuged in 500uL of 70% Ethanol at 13,000rpm for 5 minutes. After the removal of the alcohol, the DNA pellets were air-dried overnight before being finally resuspended in 20uL of ultra-pure water. Equi-molar quantities of DNA from each of the 10-fly extracted DNA samples were mixed per phenotype to form one 'low' DNA and one 'high' DNA pool per replicate for sequencing.

3.3.5 *DNA Sequencing for BSA*

Paired-end Illumina DNA sequencing (HiSeq 2000) was used to genotype the DNA bulks. I also sequenced the CTN42 line so that the parentage of different SNPs could be determined for analysis. The *D. serrata* reference genome is based on one of the parents of the cross in this experiment, FORS4, so I did not need to re-sequence this line. The parental CTN42 line was sequenced at the Australian Genome Research Facility (AGRF), whereas the F60 samples were sequenced at BGI-Honk Kong Co. Limited (China). Table 3.1 gives details of the sequencing libraries, number of reads and raw coverage for each pool.

Table 3.1

Details of sequencing libraries for pooled DNA samples of a CTN42-FORS4 *Drosophila serrata* intercross (F60) and CTN42 parental line used for bulk segregant analysis of CHC polymorphism in the species.

Generation	CHC Polymorphism Phenotype	Sequence Type (insert size)	Read Length	Sequencing Provider	No. of Reads	Coverage
F60(1)	'High'	Paired-end (500bp)	100bp	BGI	66,631,462	31.4566
F60(1)	'Low'	Paired-end (500bp)	100bp	BGI	62,418,518	29.2918
F60(2)	'High'	Paired-end (500bp)	100bp	BGI	59,914,718	27.5689
F60(2)	'Low'	Paired-end (500bp)	100bp	BGI	60,005,830	27.6818
CTN42	'Low'	Paired-end (500bp)	100bp	AGRF	116,022,310	57.7446

3.3.6 *Genome-wide QTL screen*

The analysis pipeline for genetic mapping of the CHC major polymorphism in this experiment is shown in Figure 3.1. A recently produced *D. serrata* reference genome based on the FORS4 line (S. F. Chenoweth, unpublished) was used as the main framework for

assessing differences between the 'high' and 'low' DNA pools. The genome presently has an N50 scaffold size of 1.6Mbp, consists of 5799 scaffolds and has a total sequence length of ~178Mbp. 'High' and 'low' pool sequence reads from this experiment were aligned to sequence scaffolds of this reference genome using the bwa-short algorithm of the Burrows Wheeler Alignment (BWA) software, a tool that efficiently aligns relatively short nucleotide sequences against a long reference sequence (LI and DURBIN 2009). A set of utilities in the Sequence Alignment/Map (SAM) tools (LI *et al.* 2009) were used to convert the reference genome-mapped reads of the DNA sequence pools (generated as output of BWA) into 'mpileup' file formats that were then the basis for calling SNPs. I then used the PoPoolation2 set of Perl scripts (KOFLER *et al.* 2011) to generate a list of SNPs differentiating the 'low' and 'high' pools for each replicate. The final output files contained read counts for each variable site in the 'low' and 'high' pools. These flat files were then used as inputs to R where I used a customised script to calculate measures of genetic differentiation.

To eliminate SNPs due to potential sequencing errors, I first generated a set of informative SNPs representing fixed differences between CTN42 and FORS4 lines (1,592,398) by aligning the CTN42 reads to the *D. serrata* reference (FORS4) genome. This set of informative SNPs was then used to filter the sites considered in subsequent analyses. All the SNPs between 'low' and 'high' pools from this experiment were first cross-checked against this list of SNPs and those that did not match these sites were discarded as possible sequencing errors.

As a measure of genetic differentiation between the 'low' and 'high' pools, I then calculated Fisher's G test statistic for all SNPs. The G estimates are based on single SNPs and it is known that there can be significant site-to-site variation in allele frequency estimates due to random variation in the sequencing processes (KOLACZKOWSKI *et al.* 2011). Therefore, for the initial genome-wide QTL scan I calculated the mean value of G in non-overlapping 50 SNP windows (within scaffolds) across the *D. serrata* genome. I set a highly conservative QTL peak detection threshold by ranking the 50-SNP mean G values in each replicate and taking the value corresponding to the top 0.5% of these as the threshold for each analysis.

3.3.7 *Local assembly of the genome underlying QTL peaks*

Because the *D. serrata* genome scaffolds were unordered, in the QTL detection scans there appeared to be multiple genomic regions of strong genetic differentiation between 'low' and 'high' pools. It was not known if this represented a single or many QTLs. Therefore, I needed to better assemble the genome around these enriched regions. To achieve this, I identified all the *D. serrata* scaffolds containing hits above the QTL detection threshold and used them to start a local assembly based on their alignment with the closely related *Drosophila kikkawai* genome. The scaffolds were aligned to the *D. kikkawai* genome using MUMmer 3.0 (KURTZ *et al.* 2004). Initial attempts to align the *D. serrata* scaffolds directly to the *D. melanogaster* genome were not useful for genome assembly because macrosynteny is poorly conserved between these two species (STOCKER *et al.* 2012). It should be noted that, strong conservation of chromosomal arm sequence content was observed nonetheless which is consistent with previous comparative physical mapping work in *D. serrata* and *D. melanogaster* (STOCKER et al. 2012). *Drosophila kikkawai* is the only other montium species that has been sequenced and, fortunately, aligned to *D. serrata* scaffolds with much longer runs of conserved sequence. In flies there is strong conservation of microsynteny between species whereas synteny of larger regions is poorly conserved (RANZ *et al.* 2001). The MUMmer 3.0 software program has the capacity to evaluate and compare sequence assemblies of species whose genomes are largely similar but may have undergone significant rearrangements (KURTZ *et al.* 2004). Similar approaches have taken advantage of this feature of *Drosophila* genome diversification to order and align scaffolds (NOVEMBRE *et al.* 2005). Using this approach I was able to assemble over approximately 1.7Mbp of the right arm of the *D. serrata* 3rd chromosome into a single superscaffold. Sequence alignment of the s3 marker from the F2 QTL analysis, which was linked to the QTL at position 74.2cM on 3R, to this superscaffold confirmed its 3R origin. This superscaffold alone accounted for all QTL-containing regions in the two replicate analyses.

To better localise the positions of the QTL peaks on the 3R superscaffold, I estimated G`, which is a smoothed version of Fisher's G test (KOLACZKOWSKI *et al.* 2011). G` is a more powerful metric for QTL detection via next-gen bulk segregant analysis because it effectively minimises site-to-site variance in sequencing error which can randomly exaggerate or obscure the overall mapping signal. To estimate G` I applied the LOWESS smoother available in R (CLEVELAND 1981) to the single-SNP G estimates. Because Fisher's G test is based on a

contingency table, which can be highly sensitive to low cell counts, G cannot be estimated at sites where an allele was missing/fixed in one of the pools. This property consequently created sparse regions of G estimates along the local assembly in the areas of highest differentiation. Therefore as a complementary approach, I also determined the 'low' (CTN42) allele frequency difference between 'low' and 'high' DNA bulks across the local assembly.

3.3.8 *Annotation of QTL regions*

To identify the potential candidate genes within the QTL intervals on the superscaffold, I first used the assembly sequence as a query in TBLASTX (default settings) against the *D. melanogaster* gene-extended and transcript databases (genome version r5.50 FB201302). Using the FlyBase gene identification names from the TBLASTX query I then performed a functional annotation using DAVID (Database for Annotation, Visualization and Integrated Discovery) bioinformatics resources (HUANG *et al.* 2009). CHC synthesis involves elongation of fatty acyl-CoAs through head-to-head condensation of fatty acids, fatty acid reduction to aldehydes and fatty aldehyde conversion to one carbon shorter n-Alkanes and n-Alkenes (JURENKA *et al.* 1987; BLOMQUIST *et al.* 1993; MPURU *et al.* 1996). I searched the string "fatty-acid" to identify potential candidates that may be affecting CHC biosynthesis. Three classes of genes were identified; fatty acid *desaturases*, *elongases* and *reductases*.

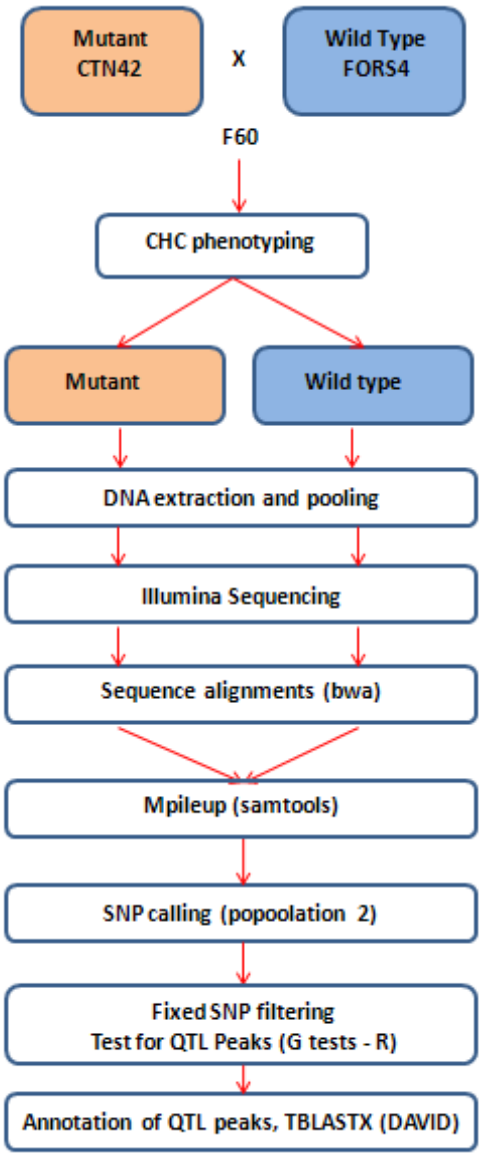


Figure 3.1

Pipeline outlining the analysis for genetic mapping of a major polymorphism in three CHCs of *D. serrata* using bulk segregant analysis of whole genome sequence data in a F60 advanced intercross between parental 'low' (CTN42) and 'high' (FORS4) lines.

3.4 RESULTS

3.4.1 *Genome-wide DNA Sequence Analysis*

There were 1,592,398 fixed SNP differences between the two parental lines (FORS4 and CTN42). Because of differences in sequence coverage, samples differed in the number of informative SNPs available for analysis (Table 3.2). Although, over 2.7 million SNPs were detected for each of the replicates, I restricted any further analysis to those SNPs verified as fixed between the parental lines only.

Table 3.2

Details of SNP detection between 'low' and 'high' pooled DNA samples from the CTN42 FORS4 intercross used for bulk segregant analysis. Data are shown for both technical replicates (A and B).

Replicate	No. of SNPs detected	SNPs of parental origin
F60 (A)	3,122,797	1,240,845
F60 (B)	2,705,692	1,351,057

Applying a 0.05% cut off for 50-SNP mean G estimates, an alignment of the *D. serrata* scaffolds in an unordered manner produced four clear peaks of genetic differentiation in both replicates (Figure 3.2). When the genome scans were then replotted using the super scaffold only a single QTL peak was observed, suggesting that the appearance of multiple peaks was simply an artefact of a limited initial genome assembly rather than multiple major QTL (Figure 3.2).

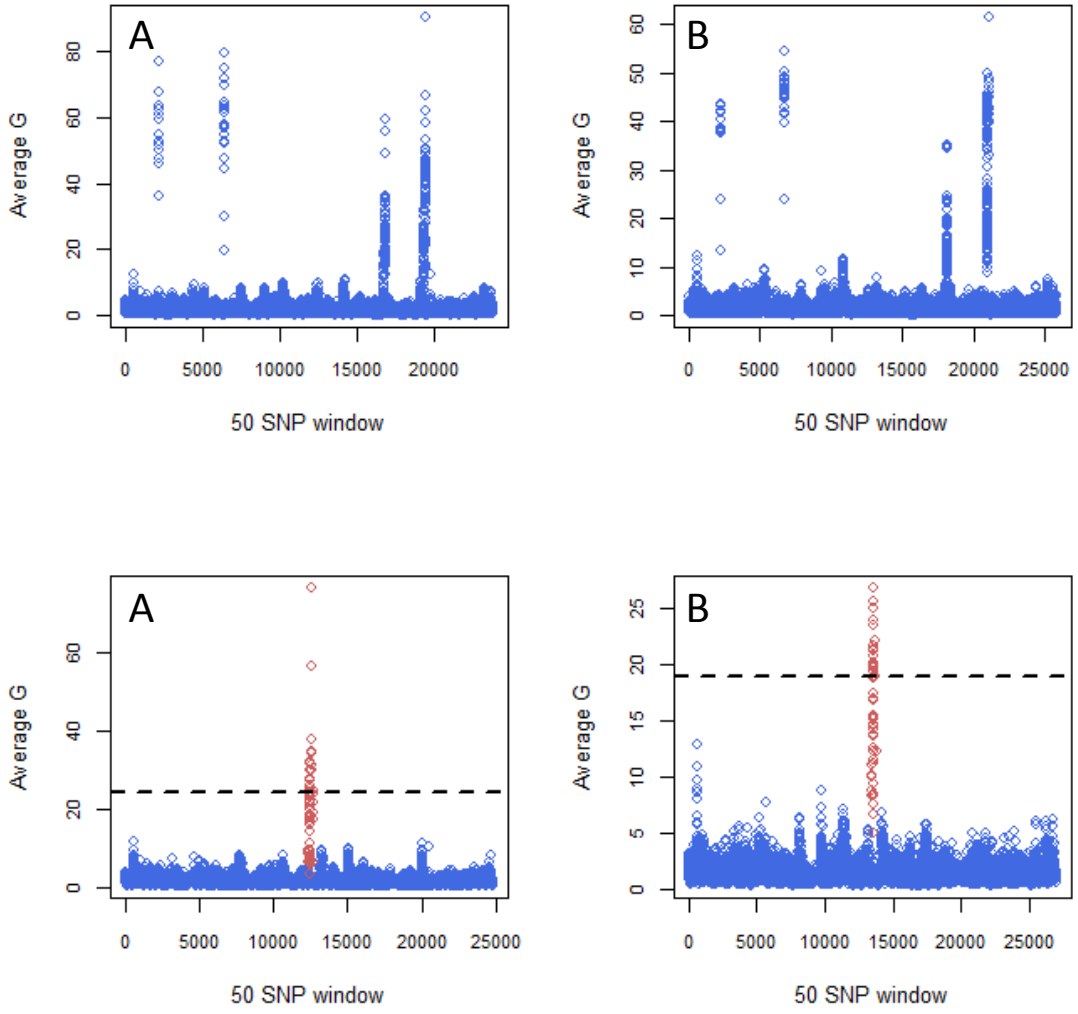


Figure 3.2

Genetic differentiation between 'low' and 'high' DNA pools of *D. serrata* measured through the mean G applied to 50 SNP windows in two F60 samples. The two columns (F60A, F60B) are results from two technical replicate samples. The top row depicts results based on unordered scaffolds while the bottom row shows the same analysis with the 3R superscaffold assembled.

3.4.2 *Super Scaffold Assembly*

Four scaffolds with QTL peaks above the threshold (scaffolds 2.1, 8.1, 42.1 and 59.1) were used to create a local assembly based on their alignment with *D. kikkawai* scaffolds using nucmer in MUMmer 3.0. The regions aligned with two *D. kikkawai* scaffolds. All four *D. serrata* scaffolds, aligned with gi|343973783|gb|JH111225.1 (Figure 3.3). Part of scaffold 59.1 aligned with *D. kikkawai* scaffold gi|343973922|gb|JH111086.1 and effectively facilitated joining the two *D. kikkawai* scaffolds (Figure 3.3). Although it did not contain a QTL peak, to verify that I was on the edge of a QTL region I also included a segment of *D. serrata* scaffold 18.1 which aligned well with scaffold gi|343973922|gb|JH111086.1 (Figure 3.3). The resulting superscaffold aligned exclusively with *D. melanogaster* chromosome 3R (Figure 3.4) and was 1,767,125 bp in length.

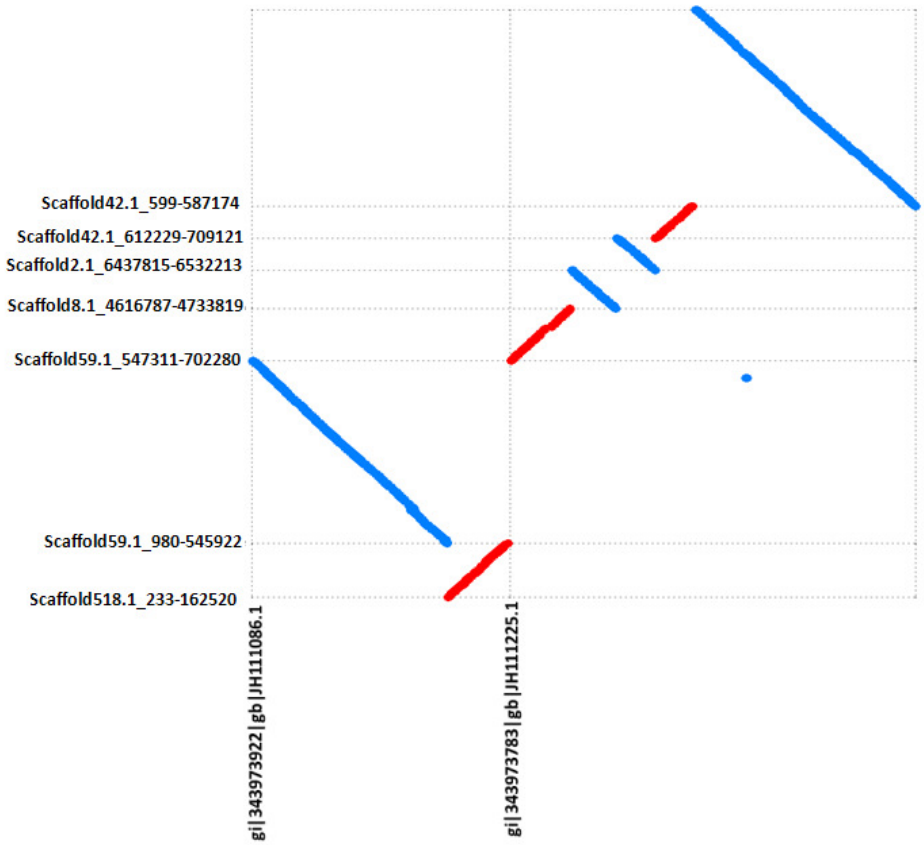


Figure 3.3

Mummer plot showing DNA sequence scaffold alignment between *D. serrata* sequence data (Y-axis) and *D. kikkawai* sequence data (X-axis), showing overlapping of several *D. serrata* scaffolds on some *D. kikkawai* scaffolds. The red segments show regions where *D. serrata* and *D. kikkawai* scaffolds were in perfect alignment while the blue segments show regions where, in spite of matching up with *D. kikkawai*, *D. serrata* scaffolds were inverted, which necessitated the re-arrangement and inversion of some scaffolds in the alignment to produce a contiguous *D. serrata* local assembly. The small blue dot indicates a segment of *D. serrata* scaffold 59.1 that did not align properly with the rest of this scaffold.

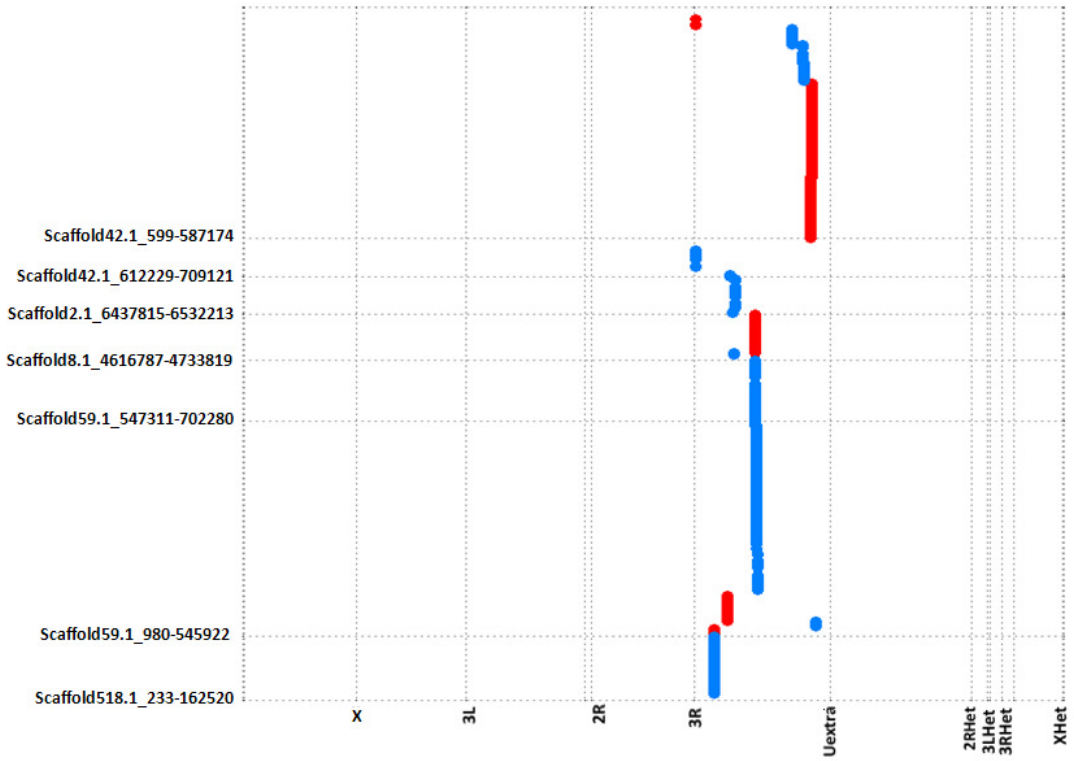


Figure 3.4

DNA sequence alignment between the *D. serrata* 1.7 Mbp superscaffold and the entire *D. melanogaster* genome (X-axis) showing conservation of sequence content with *D. melanogaster* chromosome 3R.

3.4.3 *Finer scale QTL localisation and annotation of the QTL-containing regions*

The G` estimates across the super scaffold resulted in one major peak (Figure 3.5). In both replicates there was a corresponding spike in individual G values with a maximum at the exact same position in each replicate (Position 811,136; Figure 3.5). These peaks corresponded tightly to the region of maximal 'low' allele frequency difference as well (Figure 3.5).

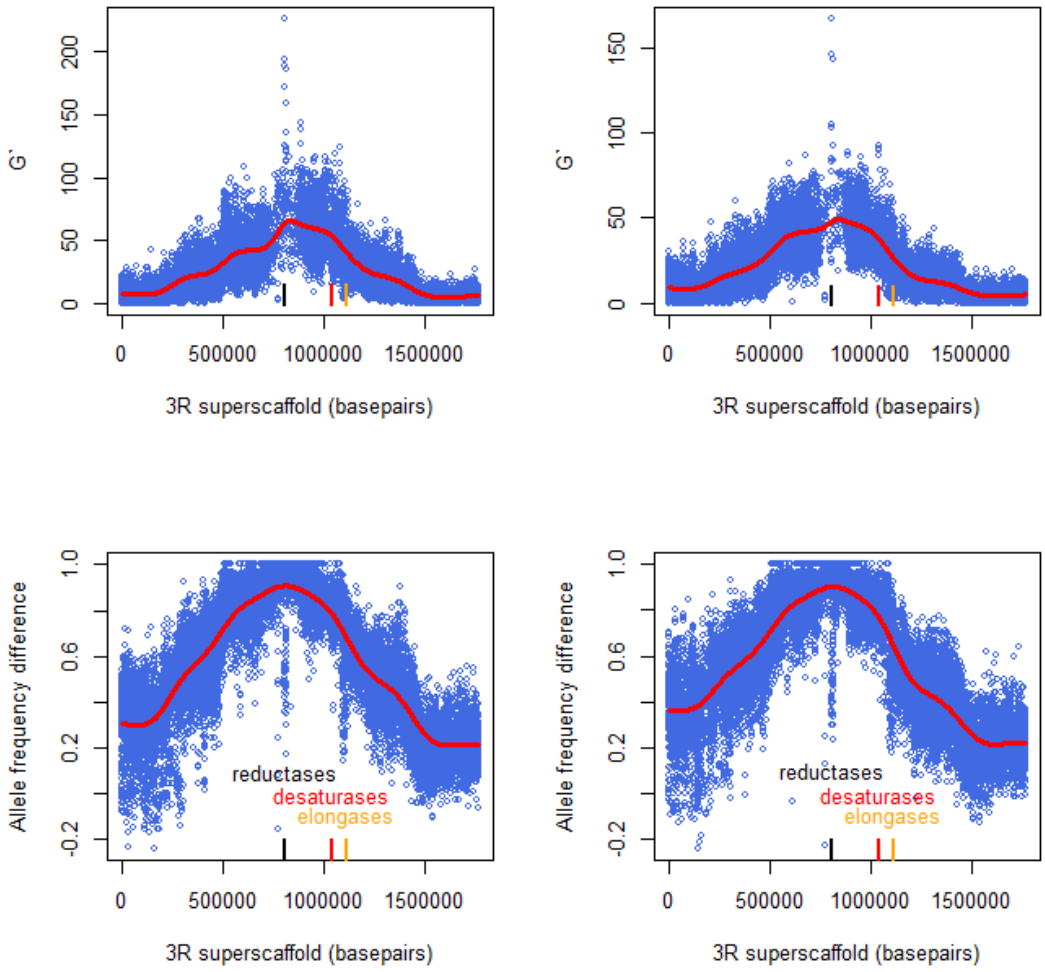


Figure 3.5

Genetic differentiation between F60 'low' and 'high' DNA pools along the local 3R assembly. Shown are the individual SNP estimates in blue with smoothed estimates indicated by the red line. Plots are shown for both G` (top row) and the 'low' (CTN42) allele frequency difference (bottom row) for each technical replicate. Indicated on the X-axis are also are the locations of three classes of candidate gene (three fatty-acid reductases, two desaturases, and two elongases) with predicted or known functions in CHC biosynthesis.

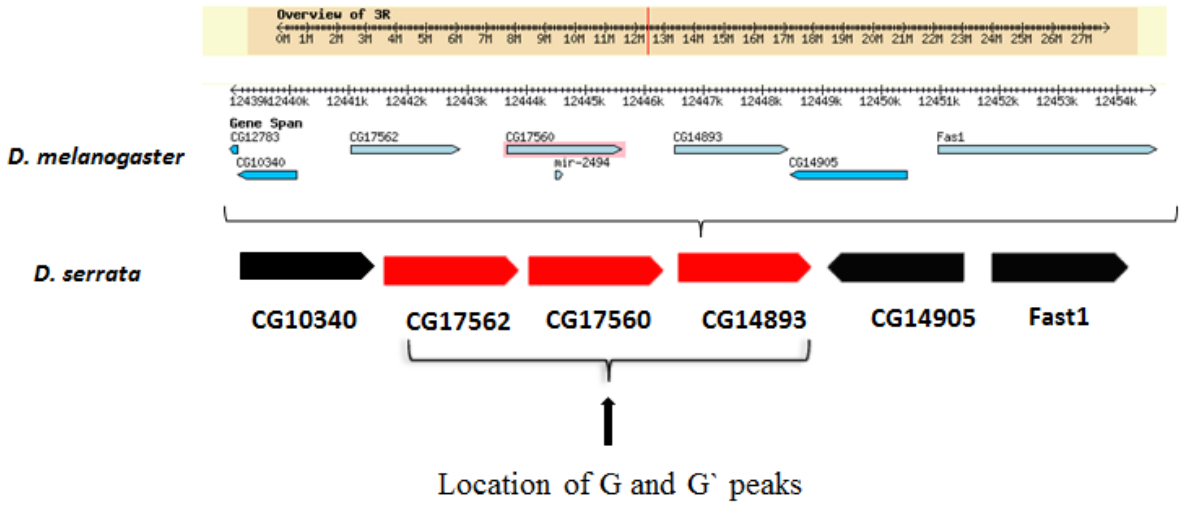


Figure 3.6

Genomic region of the QTL peak is homologous with a region of 3R in *D. melanogaster* and contains three fatty acid reductase genes (shown in red). *D. melanogaster* image and data sourced from Flybase.

Annotation of the superscaffold revealed that the QTL peak was situated in the gene, Dmel\CG17560 which is a *fatty Acyl-CoA reductase* gene. Immediately either side of this gene there were two other *fatty Acyl-CoA reductase* genes (Dmel\CG17562 and Dmel\CG14893). This region was syntenic with *D. melanogaster* (Figure 3.6). Using the whole adult fly RNA-seq data used for SNP marker discovery in Chapter 2, I was able to verify these three loci are indeed expressed in the CTN42 and the FORS4 lines (Figure 3.7). More broadly across the superscaffold but outside the region of maximal differentiation, there were two other classes of candidate genes implicated in CHC biosynthesis; desaturases and elongases (Figure 3.5). These genes included two adjacent desaturases, Dmel\desat1 (CG5887) and Dmel\desat2 (CG5925) and also two adjacent predicted elongases, Dmel\CG31522 and Dmel\CG31523.

Because of their possible role in CHC biosynthesis, I also verified that the three reductases at the QTL peak, Dmel\CG17560, Dmel\CG17562 and Dmel\CG14893 were indeed expressed in *D. serrata.* To do this I used the whole adult RNA-Seq data used for SNP marker development and EST discovery in CTN42 and FORS4 lines (Chapter 2 methods). While this data was not of sufficient quality to support a quantitative analysis of expression differences between the parental lines, it was sufficient to verify that the genes are indeed expressed in adult *D. serrata.*

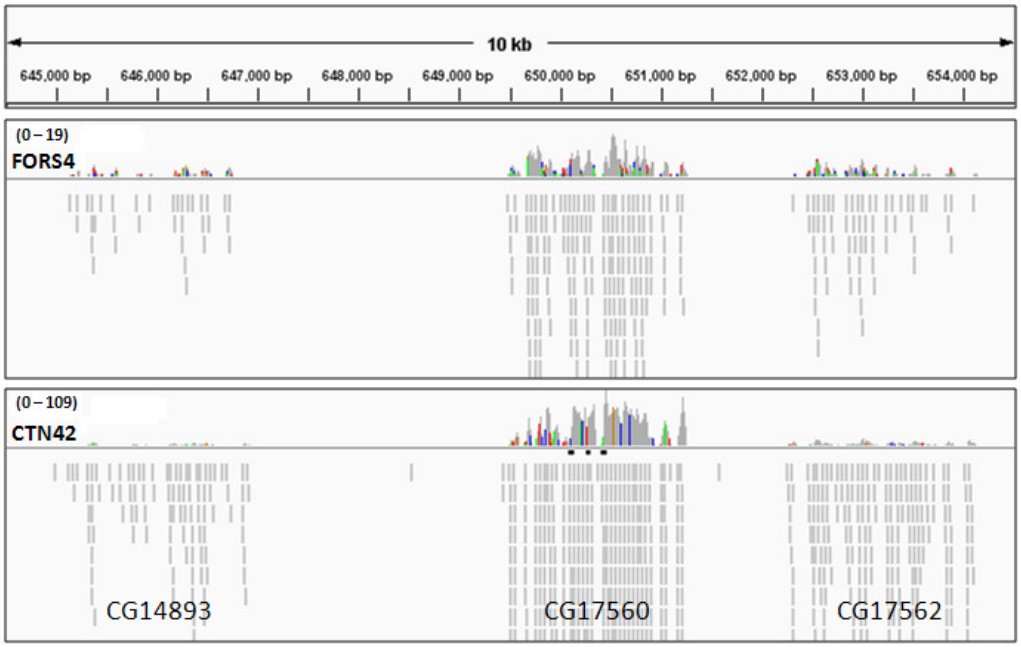


Figure 3.7

Plot of RNA-seq reads from whole body adult samples of parental FORS4 (above) and CTN42 (below) mapping lines showing the expression of three reductase genes (Dmel\CG14893, Dmel\CG17560 and Dmel\CG17562) in *D. serrata.* The read count ranges are shown within brackets above the name of each sample. RNA-Seq data are those used for SNP discovery and EST assembly in Chapter 2.

3.5. DISCUSSION

Studies that focus on dissecting the genetic basis of trait variation are vital for understanding the number and nature of DNA sequence variants that drive adaptive phenotypic change (MACKAY 2004; GLEASON *et al.* 2009). CHCs are a particularly versatile suite of traits in *Drosophila* that are responsive to forces of both natural and sexual selection (CHENOWETH and BLOWS 2005; CHENOWETH *et al.* 2008; RUNDLE *et al.* 2009; HINE *et al.* 2011). Although the role of selection in the divergence of these traits has been established (CHENOWETH and BLOWS 2008) the actual genetic polymorphisms targeted by selection remain unclear. This study used a fine-mapping strategy to identify three adjacent and related protein coding Fatty-acyl-CoA reductase genes (Dmel\CG17562, Dmel\CG17560 and Dmel\CG14893) potentially underlying a major polymorphism in 5,9-tetracosadiene, 5,9-pentacosadiene and 9-pentacosene in *D. serrata.* Results of an initial F2 QTL mapping experiment on this polymorphism indicated two major chromosome 3R loci that predominantly accounted for phenotypic variation in 5,9-tetracosadiene, 5,9-pentacosadiene and 9-pentacosene (Chapter 2, this chapter). Consistent with this finding, regions of strong genetic differentiation between the 'low' and 'high' phenotype DNA pools were also located on 3R.

CHC biosynthesis involves elongation of fatty acyl-CoAs with the aid of fatty acid elongases (MPURU *et al.* 1996). Fatty acid reductases facilitate the decarbonylation of the resulting fatty acyl-CoA chains to form aldehyde intermediates that are finally converted into n-Alkanes, n-Alkenes, secondary alcohols, and ketones (CHEESBROUGH and KOLATTUKUDY 1984; BLOMQUIST *et al.* 1993; PARK 2005). A number of genes, principally elongases (CHERTEMPS *et al.* 2007a), desaturases (MARCILLAC *et al.* 2005b) and reductases (TURGEON and BERNATCHEZ 2001) have been implicated in CHC biosynthesis in *Drosophila* through changes in their coding regions (DALLERAC *et al.* 2000), changes in *cis*-regulatory regions (LEGENDRE *et al.* 2008; SHIRANGI *et al.* 2009) or at other adjacent loci (COYNE *et al.* 1999; GLEASON *et al.* 2009).

This work is the first of its kind to fine map changes in sexually and naturally selected CHC traits in *D. serrata* to loci within fatty-acyl-CoA reductase genes. Although closely positioned desaturase and elongase genes were also implicated in this experiment, it was the reductase candidate genes that were nestled at the very peak of genetic differentiation between the 'low'

and the 'high' DNA pools. Two groups of reductases, an NADH-dependent fatty acyl reductase and an NADPH-dependent aldehyde reductase facilitate the production of cuticular hydrocarbons and primary alcohols from the fatty acid intermediates (ROWLAND *et al.* 2006). Many plant and animal taxa possess fatty aldehyde-generating fatty acyl-CoA reductase enzymes (WANG and KOLATTUKUDY 1995) and changes in their associated genes may have impacts on hydrocarbon production. RNAi knock-down of decarbonylating reductase genes involved in the conversion of aldehydes during CHC biosynthesis in *Drosophila* oenocytes leads to substantial deficiencies in cuticular hydrocarbon production (TURGEON and BERNATCHEZ 2001).

The finding of reductase genes potentially affecting CHC production as revealed in this study corroborate results found in other animal and plant groups. Genetic variation in a fatty-acyl reductase gene drives phenotypic differences in female hydrocarbon pheromone biosynthesis in *Ostrinia nubilalis* moths (LASSANCE *et al.* 2010). A pheromone-gland-specific long-chain fatty-acyl reductase gene that controls the conversion of fatty acid pheromone precursors has also been reported in the silkworm *Bombyx mori* (MOTO *et al.* 2003). A fatty acyl-CoA reductase (AmFAR1) gene capable of driving the conversion of a wide range of fatty acids is also implicated in the synthesis of a bouquet of hydrocarbon communication pheromones in honey bees, *Apis mellifera* (TEERAWANICHPAN *et al.* 2010). In *Arabidopsis thaliana*, a gene, CER4, encoding fatty acyl reductase is specifically involved in the production of C_{24} to C_{28} long-chain primary alcohols, a major cuticular wax component (ROWLAND *et al.* 2006).

Desaturase and elongase genes were positioned quite close to the peak of maximum genetic differentiation between 'low' and 'high' sequence pools. These genes are also involved in CHC biosynthesis and genetic polymorphisms in the cis-regulatory regions and other loci adjacent to the desaturase genes have a pervasive effect on the most abundant CHCs in *D. simulans*, *D. sechellia* and *D. melanogaster* (COYNE *et al.* 1999; GLEASON *et al.* 2009). Mutations within these genes drive dienic hydrocarbon polymorphism (DALLERAC *et al.* 2000; LEGENDRE *et al.* 2008; SHIRANGI *et al.* 2009), drastic depression of unsaturated hydrocarbon production and the obliteration of pheromonal CHC sexual dimorphism in *Drosophila* (MARCILLAC *et al.* 2005a; WICKER-THOMAS *et al.* 2009). Elongase genes are involved in the elongation of short unsaturated fatty acids and play a role in vaccenyl acetate biosynthesis (CHERTEMPS *et al.* 2005). The fact that a signal of differentiation was also detected on these two families of genes demonstrates that genetic polymorphisms

accumulating at different points in a single biosynthetic pathway may lead to substantial differences in CHC pheromone expression (LASSANCE *et al.* 2010). It is interesting that two types of candidate genes that were detected and are already documented to affect changes in CHC*s* in *Drosophila* did not occupy the region of maximum genetic differentiation and are thus effectively ruled out as the major driver of the short-chained CHC polymorphism in this study. This is understandable because, unlike the monoene-diene CHC polymorphisms that implicated these other candidate genes, the polymorphism in this experiment was not dienic but involved a loss of two dienes and a monoene. In *Ostrinia* moths, the accumulation of nonsynonymous amino acid substitutions in a fatty-acyl reductase gene results in the divergence of species-specific pheromonal hydrocarbon ratios (LASSANCE *et al.* 2013). The key mode of action of this gene appears to be that individual amino acid changes confer large changes in substrate conversion efficiencies. Thus the polymorphism may not affect precursor chain length *per se* but rather whether precursors of differing chain lengths are ultimately converted into pheromones. The involvement of reductases rather than desaturases and elongases is thus more plausible in the case of *D. serrata*.

It is not yet known which of the three loci are involved in the polymorphism. The RNA-seq data used for marker development in Chapter 2 confirmed their expression in whole adult flies in both 'low' and 'high' founder lines. This observation, while confirming the genes are active in *D. serrata* does not preclude a loss of expression mutation underlying the phenotype. This is because CHCs are expressed in oenocyte cells and the genes may be expressed in multiple other tissue types such as the fat body. Thus a mutation that occurs within an oenocyte enhancer could potentially knockout expression in those cells whilst preserving its expression in the rest of the body. An alternative explanation is that a change in the protein coding sequence may be involved. In the present data set there were too many different types of polymorphisms differentiating 'low' and 'high' lines to infer the causal variant or mode of action. Inspection of Fly Atlas confirmed that two of the genes (CG17562, CG17560) exhibit near identical expression profiles across different tissues. Their expression profile is consistent with adult oenocyte expression. Interestingly, Flyexpress images show that CG17562 is indeed expressed in larval oenocytes (Stages 13-16: http://www.flyexpress.net/search.php?type=gene_images&source=1&search=CG17562&gene=37530). However, CG14893 is expressed mainly in the accessory glands in *D. melanogaster* and so may not be as important for CHCs.

There was a discrepancy in the number of QTLs detected for this CHC polymorphism between the initial F2 QTL mapping and the BSA; two broad major effect QTL were identified in the former while BSA detected only one of these regions. This resonates with several empirical studies that report a reduction in the number of detected QTLs upon fine mapping. For example, in a nine-generation advanced intercross used for fine mapping body weight in chickens, BESNIER *et al.* (2011) could only detect five of the nine QTL that were earlier found with the original F2 population. Similar results have been found in advanced crosses of maize (HUANG *et al.* 2010) and mice (WANG *et al.* 2003; LIONIKAS *et al.* 2010).

Discrepancies between QTL positions located using conventional mapping and those found through fine mapping of advanced intercrosses may be due to a number of factors. Cosegregation between a QTL and its distant markers due to chance or selection effects in early generation experimental populations may create an artifactual QTL, particularly in the vicinity of large QTL effects, creating an appearance of two QTL when only one exists especially when marker coverage is low (KNAPP *et al.* 1990; HALEY and KNOTT 1992; DOERGE and CHURCHILL 1996; BESNIER *et al.* 2011). Fine mapping with advanced crosses may eliminate these 'ghost' QTLs. Markers within the same vicinity and linked to the two QTLs (s3 and s5) did not concur with the expected Mendelian ratios. Tightly linked marker loci may not be easily ordered on a linkage map (GORING and TERWILLIGER 2000). Although map positions of most markers on the linkage map used for the F2 mapping were comparable with their relative positions on the physical map, there were a few cases where discrepancies occurred in the location of SNPs between the physical and linkage maps (STOCKER *et al.* 2012). It is possible that some markers in the F2 QTL mapping were incorrectly placed, resulting in weaknesses in the QTL detection analysis. This might have contributed to the discrepancy in the number of detected QTLs between the F2 cross and the BSA experiment. Although the use of advanced intercross populations for mapping the genetic basis of traits improves accuracy by reducing confidence intervals around QTLs (HUANG *et al.* 2010; LIONIKAS *et al.* 2010), the accumulation of recombination events and subsequent break down of linkage disequilibrium among loci in later generations of an advanced intercross (DARVASI and SOLLER 1995), may lead to significant deviation of QTL effect positions from those observed in earlier (e.g. F2) generations (BESNIER *et al.* 2011). The disruption of linkage within clusters of QTL in coupling phase (easily detectable in early cross conventional mapping population), may fractionate them into several linked loci of small effects that are beyond detection and thus less discoverable in the advanced intercross population (HUANG *et*

al. 2010; LIONIKAS *et al.* 2010). Thirdly, even with randomised matings and large effective population sizes, QTLs may also be lost by chance in the development of advanced intercoss lines (IRAQI *et al.* 2000).

It is presently unclear why only one QTL remained in the F60 experiment compared with two in the F2 mapping results in chapter 2. One possibility may be that the unreplicated QTL may have been actually a cluster of multiple linked small-effect QTLs whose linkage was heavily disrupted in the creation of the advanced cross to the point where the resulting individual small QTLs became undetectable. The second possibility is that the region containing the second QTL on 3R was fixed during the advancement of the F2 population. Unfortunately attempts to find the region that may have become fixed were unsuccessful because the nearest marker (s60) was quite distant from the QTL peak in the F2 and would have no longer been in linkage disequilibrium with the QTL by generation 60; this marker was not fixed in the F60 analysis. Thirdly, perhaps owing to the limited density of markers in the F2, the two QTLs implicated in Chapter 2 were in fact only ever one QTL. The possibility of inversions may also not be totally ruled out. Although the two lines that were used to create a mapping cross at F2 were initially scanned for inversions, the existence of inversions may potentially be greatly underestimated (KIRKPATRICK and BARTON 2006) and it is possible that some inversions that were too small could not be detected by the polytene chromosome squashes. Parallel latitudinal clinal inversions abound in *Drosophila* across continents (HOFFMANN *et al.* 2004). Inversions drastically suppress recombination in the sections of the genome where they occur as well as in those loci to which they are tightly linked (HOFFMANN *et al.* 2004; KIRKPATRICK and BARTON 2006). This has been reported to occur rather strongly in *Drosophila* (PEGUEROLES *et al.* 2010). The existence of inversions results in enhanced association among loci that may 'lock up' co-adapted sets of alleles (HOFFMANN *et al.* 2004; PEGUEROLES *et al.* 2010; AYALA *et al.* 2013). If, for some reason, a missed inversion occurred in the region where the two QTLs in the F2 cross were detected, these two QTLs would have been tightly linked and would have passed down the generations as one unit.

These uncertainties notwithstanding, QTL mapping using both early generation cross (e.g. F2) and an advanced cross of the same lineage effectively combines the detection power of the former and accuracy of the latter to deliver a superior trait mapping strategy (LIONIKAS *et al.* 2010). This BSA experiment, involving more than a million SNP markers across the whole genome, discounted the possibility of two QTLs and mapped the CHC polymorphism to one

region harbouring three adjacent candidate genes. This finding underscores the power of genome wide scanning in the context of advanced crosses, selective genotyping and bulked segregant analysis for the dissection of the genetic basis of quantitative traits.

The mapping of a major effect polymorphism in a pheromonal trait to a single very small QTL region lends support to the fact that although the genetic basis of many ecologically important complex traits is polygenic (FLINT and MACKAY 2009; ROCKMAN 2012), a large proportion of trait differences among local populations may sometimes be genuinely explained by only a few large-effect loci (VOSS and SHAFFER 1997; COLOSIMO *et al.* 2005; TOBLER and CARSON 2010; BAERWALD *et al.* 2011; YEAMAN and WHITLOCK 2011). Similarly, a single major effect locus underlies variation in a sex pheromone in the moth *Heliothis subflexa* (UMINA *et al.* 2005). The fact that three genetically correlated CHC traits mapped to these few, shared loci, may indicate that pleiotropy rather than gene linkage underlie their correlations. Pleiotropic effects will tend to manifest among traits that evolve together under selection relative to those that are under independent selection (CHEVERUD *et al.* 2004; WAGNER *et al.* 2007; KENNEY-HUNT *et al.* 2008). Selection may thus be driving changes in these CHC traits by targeting a few shared genetic polymorphisms uncovered in this study.

The polymorphism under study here involved the reduction of small chain and a correlated increase of long chain CHCs, a phenotypic effect that is reflected in the broad scale latitudinal cline in these traits. The reductase genes mapped in this study could thus have significant fitness effects in the northern tropical range of the species where its frequency is highest. Further studies will be required to understand fully the ecological context that both maintains the polymorphism and drives its clinal divergence. Further molecular work will also be required to determine if one or more of the three reductase genes are involved in CHC biosynthesis and by what molecular mechanisms (coding vs. regulatory) the polymorphism is generated.

CHAPTER 4. REPLICATION OF AN ASSOCIATION BETWEEN FATTY-ACID REDUCTASE LOCI AND A MAJOR CHC POLYMORPHISM IN AN INDEPENDENT SAMPLE OF GENETIC VARIATION FROM A NATURAL POPULATION OF *D. SERRATA*

4.1 ABSTRACT

Cuticular hydrocarbons (CHCs) are a suite of ecologically important traits implicated in both stress resistance and also mate and species recognition in *Drosophila*. Populations of *D. serrata* in the northern part of the Eastern Australia exhibit a major phenotypic polymorphism in three of their CHC compounds representing the shortest carbon chains; 5,9-tetracosadiene (5,9-C_{24}), 5,9-pentacosadiene (5,9-C_{25}) and 9-pentacosene (9-C_{25}). One phenotype expresses these three compounds only in trace amounts whereas the other expresses normal levels. Earlier F2 QTL mapping using a geographically divergent cross traced this polymorphism to major effect QTLs on chromosome 3R. Subsequent bulk segregant analysis and whole genome resequencing of an advanced F60 intercross from the same founding lines detected a single QTL harbouring three adjacent fatty acid reductase genes (CG15760, CG15762 and CG14893). In this chapter I independently validate and test the generality of this finding, focusing on a sample of within-population genetic variation from a natural population of flies where the polymorphism is at its highest frequency. First, exploiting the recessive mode of action of the QTL, I set out to confirm whether the same genetic polymorphism is responsible for the phenotype by crossing the original 'low' parent line to five additional inbred lines also expressing the 'low' phenotype. In all the F1 progeny no complementation of the 'low' phenotype occurred, suggesting a common genetic basis. Second, by resequencing the QTL region in nine inbred lines (5 'low' and 4 'high') I validated the association between the fatty acid reductase loci and the CHC phenotype and broadened the generality of this finding. Across the genomic region containing the major QTL in the bulk segregant analysis (Chapter 3), there was a hotspot of fixed nucleotide differences between 'low' and 'high' lines at these reductase loci with a particularly large number in the central gene, CG15760. The majority of potentially causal SNPs discovered by this study involved changes in amino acids in the gene CG15760, which has recently been confirmed to be expressed in *D. serrata* oenocytes in both CTN42 and FORS4 lines.

4.2 INTRODUCTION

The nature of genetic polymorphisms in a population underlies its potential for divergence (BARRETT and SCHLUTER 2008). While theory has described the mutational effects relevant for adaptive trait transitions (FISHER 1930; ORR 1998a) and empirical work on geographic patterns of phenotypic variation in the wild serves as indirect evidence for selection (ENDLER 1986; MAYR 1963), finding actual genetic polymorphisms that are segregating in natural populations which account for phenotypic variation may provide a fuller picture of the causal basis of adaptive trait divergence (NACHMAN 2005).

Cuticular hydrocarbons (CHCs) produced from fatty acids are an ecologically important suite of traits that serve as pheromonal, kairomonal or allomonal chemical messengers in insects (HOWARD and BLOMQUIST 2005). Pheromonal bouquets of these compounds are ecologically significant in Drosophilid flies where they sometimes mediate species-specific mating and reproductive isolation between different species (COYNE 1996) as well as desiccation resistance via reduced water loss (AGRAWAL 2001; HOFFMANN *et al.* 2004; PUNZALAN *et al.* 2005). In the montium Drosophilid, *Drosophila serrata,* CHCs are known to evolve under the influence of both natural and sexual selection (CHENOWETH and BLOWS 2008; RUNDLE *et al.* 2009). A major phenotypic polymorphism which exacerbates clinal variation has recently been observed for 5,9-tetracosadiene, 5,9-pentacosadiene and 9-pentacosene.

Initial attempts at tracing the genetic basis of this divergent phenotype using second generation (F2) inter-cross quantitative trait loci (QTL) mapping identified two major-effect recessive QTLs on chromosome 3R (Chapter 2). A combination of advanced inter-crossing of this population for 60 generations (F60) followed by bulked segregant analysis of whole-genome resequence data allowed me to fine-map the polymorphism to three adjacent fatty Acy-CoA reductase genes within an 8kb region of 3R. These results were obtained by crossing only two lines that were divergent for this CHC polymorphism and from opposite extremes of the species geographical distribution. The use of just a single pair of inbred lines for mapping the genetic basis for a trait may lead to results that are not reproducible across a wider sampling of the population and hence sweeping generalisations about the genetic architecture of the relevant trait may be misleading (BARNWELL and NOOR 2008a). The validity and generality of previous mapping results for this CHC polymorphism are therefore unconfirmed. Replicating previous genotype-phenotype associations in an independent

sample of genetic variation is a powerful means to validate the existence of a QTL when direct transgenic manipulation is not possible (PALSSON *et al.* 2005). In this experiment, I use this approach to validate the existence of a major QTL in three fatty acid reductase genes underlying the short-chained CHC polymorphism in *D. serrata*. Furthermore, by analysing patterns of phenotypic and DNA sequence variation in nine independent wild-derived inbred lines of *D. serrata* (five 'low' and four 'high') from the same population (Cooktown, Queensland, Australia) I extend the generality of this finding. I first use complementation tests with the 'low' parent line used in earlier mapping experiments to confirm that the same polymorphism underlies the CHC trait. Then, sequencing these nine genomes, I tested for fixed genetic differences within the major QTL region from Chapter 3 to test the prediction that the fatty acid reductase genes would form the strongest axis of genetic differentiation between 'low' and 'high' lines.

4.3 METHODS AND MATERIALS

4.3.1 *Fly crosses and maintenance*

Nine wild-derived inbred lines of *Drosophila serrata,* five 'low's (CTN10, CTN18, CTN32, CTN34 and 180) and four 'high's (CTN44, CTN21, 134, 145), were used for this experiment. Each of the lines was inbred through full-sib mating for at least 17 generations prior to phenotyping. Flies were maintained at 25^0C with a 12hours:12hours light:dark photoperiod on a diet of standard laboratory fly food made up of 36g agar, 108g raw sugar and 74g yeast, cooked in 2L of water and mixed with 24mL propionic acid and 12mL nipagen. The food media surface was always lightly dashed with live yeast to improve oviposition (DOERGE and CHURCHILL 1996).

4.3.2 *Complementation tests for a common polymorphism*

F2 inter-cross QTL mapping of the CHC polymorphism under investigation established that the QTL responsible for the CHC difference between these two phenotypes is recessive (Chapter 2). The recovery of a phenotype of interest from a cross between one line possessing a particular known recessive phenotype and another independent line of similar phenotype may indicate that their phenotypic similarity is underlain by the same recessive gene (WALTERS *et al.* 1997). To confirm whether the same gene was implicated across multiple 'low' phenotype expressing individuals, each of five 'low' lines were first crossed against the

'low' parental line (CTN42) which was used to found the F2 intercross and the F60 bulk segregant analysis flies of previous chapters. To also verify that there were no differences between the new 'low' inbred lines I analysed all pairwise crosses possible between these lines (total of 21 crosses).

4.3.3 *CHC phenotyping of parental and F1 flies*

In each parental line of F1 cross, newly eclosed virgin flies were sexed (female n = 3, male n = 3) under light CO_2 anaesthesia. These flies were kept singly in 7mL food vials for five days before CHCs were extracted for profiling. To extract the CHCs, each individual fly was washed by placing it in 100mL of hexane in a microvial insert for 3 minutes before being vortexed for 1 minute (HIGGIE and BLOWS 2007). The CHC samples were then run through an Agilent Technologies 6890N gas Chromatograph (Wilmington, Delaware, United States).

Lines were scored as either 'low' or 'high' using previously established approaches (Chapter 2 methods). Qualitative complementation tests were conducted simply by comparing the 'low' phenotype of all offspring with the parents, with complementation confirmed when all offspring expressed the 'low' phenotype. I also statistically compared the amount of 5,9-pentacosadiene (expressed as a proportion of the total CHC content) among parents and F1 progeny using ANOVA conducted on each of the 21 pairwise crosses. Statistical analyses were conducted in R (TEAM 2008).

4.3.4 *DNA extraction*

I used a phenol-chloroform extraction protocol to obtain whole-fly genomic DNA from the nine inbred lines. For each line, genomic DNA was extracted from a total of 180 flies in groups of twenty. Each group of flies was ground with a micro pestle in a single eppendorf tube containing 250uL of a solution made up of 10mL 1M stock of 0.1M Tris HCl (pH9.0), 20mL 0.5M stock of 0.1M EDTA and 10mL of 10% stock of 1% SDS solution. The eppendorf was then incubated in a water bath at 70^{o}C for one hour after which 35uL of potassium acetate was added. Thereafter, the tubes were incubated on ice for 30 minutes followed by their centrifugation at 13,000rpm speed for 15minutes. The supernatant fluid after this process was moved into a new eppendorf tube, ensuring that all the precipitate was left behind. To the supernatant, 250uL of phenol chloroform isoamyl alcohol (25mL phenol + 24mL chloroform + 1mL isoamyl alcohol) was added and the combination was thoroughly mixed by gentle shaking. After a further spinning of this mixture in a centrifuge for 5 minutes

at 13,000rpm, the resultant aqueous top layer (about 240uL) was transferred into another new eppendorf tube to which 240uL of isopropanol was added. The centrifugation of this tube at 13,000rpm for 5 minutes resulted in DNA pellet deposits at the bottom of the tube. Following the decanting of the supernatant, the residual pellets were spun in 500uL of 70% ethanol at 13,000rpm for 5 minutes after which the alcohol was removed and the DNA pellets were air dried overnight. The DNA was finally resuspended in 20uL of water. For each of the nine lines, 3ug of genomic DNA was sent for sequencing.

4.3.5 *DNA sequencing and analysis*

Paired-end Illumina HiSeq DNA sequencing was done for each of the DNA samples from the nine lines at BGI-Honk Kong Co. Limited (China); subsequent analysis of their re-sequenced genomes was conducted as follows. Reads for each line were independently aligned to the *D. serrata* reference genome using the bwa-short algorithm of the Burrows Wheeler Alignment (BWA) software (LI and DURBIN 2009). Sequence Alignment/Map (SAM) tools (LI *et al.* 2009) were used to convert the reference genome-mapped reads for each line (generated as output of BWA) into 'mpileup' file formats that were then the basis for calling SNPs. A list of SNPs differentiating all nine lines from the reference genome was then generated using the PoPoolation2 set of Perl scripts (KOFLER *et al.* 2011). The final output files contained read counts for each variable site in each of the nine lines.

As my goal was to independently validate the association detected in Chapter 3, I focussed analysis on the 1.7Mbp super-scaffold that contained the major QTL. As the mutation is recessive and I was working with inbred lines, I developed a diagnostic criterion for whether a SNP could be the causal variant. First, SNPs had to be fixed and identical across all 'low' lines. Second, the major allele of all 'low' lines must differ from the 'high' lines. Third, in the cases of a tri-allelic SNP, it was permissible for multiple 'high' lines to possess different bases as long as the major allele was not the same base as in the 'low's. I applied this rule using a custom perl script to classify SNPs as either a zero (not a possible causal variant) or 1 (potential causal variant) score. I then counted the number of 1 coded SNPs within 5000bp windows along the 1.7Mbp superscaffold testing for an overepresentation of positive hits. Orthologous amino acid sequences of *D. melanogaster* and other *Drosophila* species were used to deduce the open reading frame of the *D. serrata* sequences in order to facilitate the formal classification of the SNPs as synonymous or non-synonymous.

4.4 RESULTS

4.4.1 *Complementation tests*

In all of the crosses between CTN42 and the 'low' inbred lines the offspring exhibited the same 'low' CHC phenotype (Table 1). These results suggest that the same genetic polymorphism(s) are responsible for the 'low' phenotype within the Cooktown population. I additionally examined the potential for quantitative variation in the amount of 5,9-pentacosadiene. In all but one of the resulting 21 crosses, (CTN42 x CTN18) there was no significant difference between parents and F1 offspring (Table 1). In addition to sharing a common genetic polymorphism(s) affecting the 'low' phenotype, these results indicate limited quantitative differences between lines in any genetic polymorphisms mediating smaller effects on quantitative variation in 5,9-pentacosadiene expression in 'low' individuals.

4.4.2 *Fixed differences delineating 'low' and 'high' phenotypes*

Across the nine lines there were 138,429 SNPs detected in the 1.7Mbp 3R superscaffold. Potential causal SNPs were defined as points where all 'low' lines shared the same fixed allele and where the 'high' lines did not contain the same nucleotide as the 'low' lines. Across the superscaffold, there were 166 SNPs meeting this criterion. The largest number of fixed differences (71) occurred in one region (~10kbp) containing three adjacent fatty acyl-CoA reductases (Figure 4.1). No fixed differences were found within the 10kbp region containing CHC candidate desaturases and only 4 SNPs were fixed within the 20kbp region containing the elongases.

Table 4.1

Results of qualitative and quantitative complementation tests in crosses between multiple 'low' lines. Sex of the parental lines is indicated as either F (female) or M (male) against the identity of the respective line. The percentage 5,9-pentacosadiene (mean ± standard deviation) of total CHC production per fly is shown for parents and pooled sex offspring.

Cross	'Low' Parent 1	'Low' Parent 2	Phenotype status of offspring	% of 5,9-pentacosadiene in offspring	% of 5,9-pentacosadiene 'Low' Parent 1	% of 5,9-pentacosadiene 'Low' Parent 2	F	d.f.	P-value
1	CTN42(F)	CTN32(M)	All 'low'	0.58 ± 0.31	0.41 ± 0.12	0.35 ± 0.10	2.00	2,15	0.1702
2	CTN42(F)	CTN34(M)	All 'low'	0.59 ± 0.19	0.41 ± 0.12	0.52 ± 0.20	1.43	2,15	0.2711
3	CTN42(F)	CTN10(M)	All 'low'	0.48 ± 0.24	0.41 ± 0.12	0.57 ± 0.56	0.402	2,15	0.7933
4	CTN42(F)	CTN18(M)	All 'low'	0.48 ± 0.13	0.41 ± 0.12	0.64 ± 0.28	3.86	2,15	0.0440
5	CTN42(F)	152(M)	All 'low'	0.59 ± 0.34	0.41 ± 0.12	0.51 ± 0.36	0.56	2,15	0.5460
6	CTN42(F)	180(M)	All 'low'	0.49 ± 0.21	0.41 ± 0.12	0.47 ± 0.22	0.25	2,15	0.6806
7	CTN32(F)	CTN34(M)	All 'low'	0.40 ± 0.13	0.35 ± 0.10	0.52 ± 0.20	0.73	2,15	0.4969
8	CTN32(F)	CTN10(M)	All 'low'	0.46 ± 0.29	0.35 ± 0.10	0.57 ± 0.56	0.40	2,15	0.6798
9	CTN32(F)	CTN18(M)	All 'low'	0.54 ± 0.28	0.35 ± 0.10	0.64 ± 0.28	3.06	2,15	0.0770
10	CTN32(F)	152(M)	All 'low'	0.43 ± 0.11	0.35 ± 0.10	0.51 ± 0.36	0.60	2,15	0.5600
11	CTN32(F)	180(M)	All 'low'	0.46 ± 0.22	0.35 ± 0.10	0.47 ± 0.22	0.44	2,15	0.6524
12	CTN34(F)	CTN10(M)	All 'low'	0.67 ± 0.42	0.52 ± 0.20	0.57 ± 0.56	0.30	2,15	0.7455
13	CTN34(F)	CTN18(M)	All 'low'	0.64 ± 0.23	0.52 ± 0.20	0.64 ± 0.28	1.15	2,15	0.3448
14	CTN34(F)	152(M)	All 'low'	0.65 ± 0.39	0.52 ± 0.20	0.51 ± 0.36	0.42	2,15	0.6620
15	CTN34(F)	180(M)	All 'low'	0.49 ± 0.15	0.52 ± 0.20	0.47 ± 0.22	0.15	2,15	0.8619
16	CTN10(F)	CTN18(M)	All 'low'	0.71 ± 0.49	0.57 ± 0.56	0.64 ± 0.28	0.19	2,15	0.8260
17	CTN10(F)	152(M)	All 'low'	0.48 ± 0.31	0.57 ± 0.56	0.51 ± 0.36	0.04	2,15	0.9581
18	CTN10(F)	180(M)	All 'low'	0.34 ± 0.12	0.57 ± 0.56	0.47 ± 0.22	0.54	2,15	0.5927
19	CTN18(F)	152(M)	All 'low'	0.53 ± 0.18	0.64 ± 0.28	0.51 ± 0.36	0.58	2,15	0.5727
20	CTN18(F)	180(M)	All 'low'	0.59 ± 0.38	0.64 ± 0.28	0.47 ± 0.22	1.45	2,15	0.2663
21	152(F)	180(M)	All 'low'	0.53 ± 0.30	0.51 ± 0.36	0.47 ± 0.22	0.05	2,15	0.9501

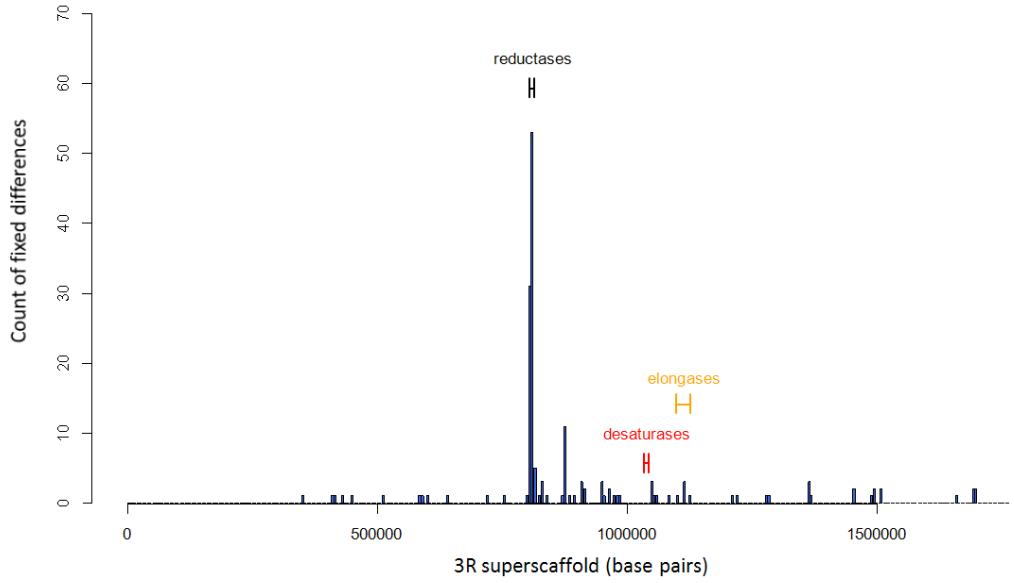


Figure 4.1

A plot of the frequency of fixed differences between five 'low' and four 'high' lines of *D. serrata* based on the alignment of their respective genomes against a reference *D. serrata* genome scaffold.

An analysis of fixed differences between the parental 'high' FORS4 line and 'low' CTN42 line that were also supported as fixed between the multiple 'low' and 'high' resequenced lines revealed a total of thirty fixed mutations within the region containing three reductase genes. Fifteen potentially causal SNPs were found within exons 2, 3 and 4 of the 1.5kbp coding region (CDS plus introns) of CG17560, the majority of which were non-synonymous. A total of thirteen fixed SNPs were found within the 2Kbp coding region of gene CG14893 and six of these mutations in exons 3, 4 and 5 were non-synonymous. Only 2 fixed SNPs occurred within the 2kbp coding region of CG17562, both of which were non-synonymous (Table 4.2). Amino acid Clustal alignment of CG17560 orthologs from 8 *Drosophila* species are shown in figure 4.3. On-going i*n situ* hybridization work on this polymorphism shows that gene GC17560 is expressed in adult oenocyte cells in male and female *D. serrata* in both high (FORS4) and low (CTN42) CHC phenotype lines (Figure 4.4)(Henry Chung, Sean Carroll, Steve Chenoweth, unpublished data).

Table 4.2

A list of SNPs fixed between 'low' and 'high' phenotype genomic pools of parental 'high' (FORS4) and 'low' (CTN42) phenotype lines used for F2 QTL mapping and bulk segregant mapping of a major-effect CHC polymorphism in *D. serrata* (Chapters 2 and 3). These genetic differences were also found to be fixed between contrasting genome groups of 'low' (5 lines) and 'high' (4 lines) phenotypes.

Gene Name	Gene Position (offset to start codon)	Nucleotide Change FORS4/CTN42	Type of Mutation	Amino Acid Change
CG17562	860 (exon3)	A/G	Non-synonymous	Isoleucine to Valine
	866 (exon3)	A/G	Non-synonymous	Lysine to Glutamic acid
CG17560	182 (exon2)	C/T	Non-synonymous	Threonine to Isoleucine
	280 (intron2)	T/C	-	-
	290 (intron2)	A/C	-	-
	308 (exon3)	A/G	Synonymous	None
	626 (exon3)	G/C	Non-synonymous	Lysine to Asparagine
	824 (intron3)	G/C	-	-
	971 (exon4)	G/A	Non-synonymous	Valine to Isoleucine
	1120 (exon4)	G/T	Non-synonymous	Lysine to Asparagine
	1280 (exon4)	C/T	Non-synonymous	Leucine to Phenyalanine
	1316 (exon4)	T/A	Non-synonymous	Phenyalanine to Isoleucine
	1448 (exon4)	T/A	Non-synonymous	Leucine to Isoleucine
	1471(exon4)	C/T	Synonymous	T
	1527 (exon4)	C/G	Non-synonymous	Glycine to Alanine
	1596 (intron4)	T/A	-	-
	1633 (intron4)	T/A	-	-
CG14893	355 (exon3)	A/C	Synonymous	None
	485 (exon3)	G/A	Non-synonymous	Valine to Isoleucine
	595 (exon3)	C/T	Synonymous	None
	601 (exon3)	C/T	Synonymous	None
	688 (exon3)	A/G	Non-synonymous	Isoleucine to Methionine
	856 (intron3)	A/G	-	-
	1379 (exon4)	A/G	Non-synonymous	Isoleucine to Valine
	1609 (exon5)	A/T	Synonymous	None
	1611 (exon5)	A/G	Non-synonymous	Lysine to Arginine
	1654 (intron5)	C/T	-	-
	1702 (exon5)	T/C	Non-synonymous	Valine to Alanine
	1754 (exon5)	C/T	Synonymous	None
	1762 (exon5)	G/A	Non-synonymous	Glycine to Glutamic acid

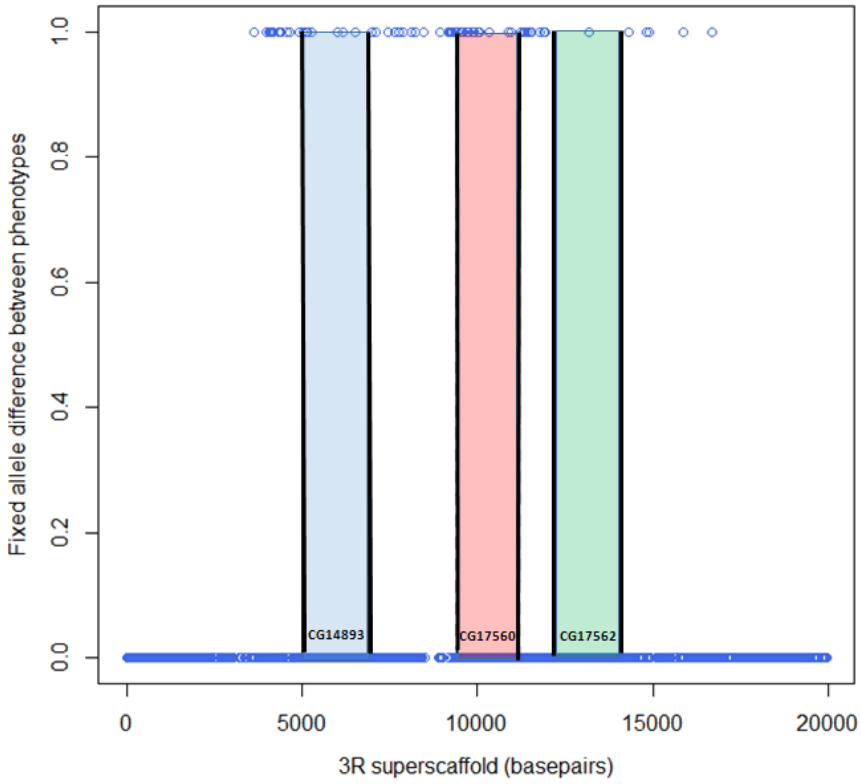
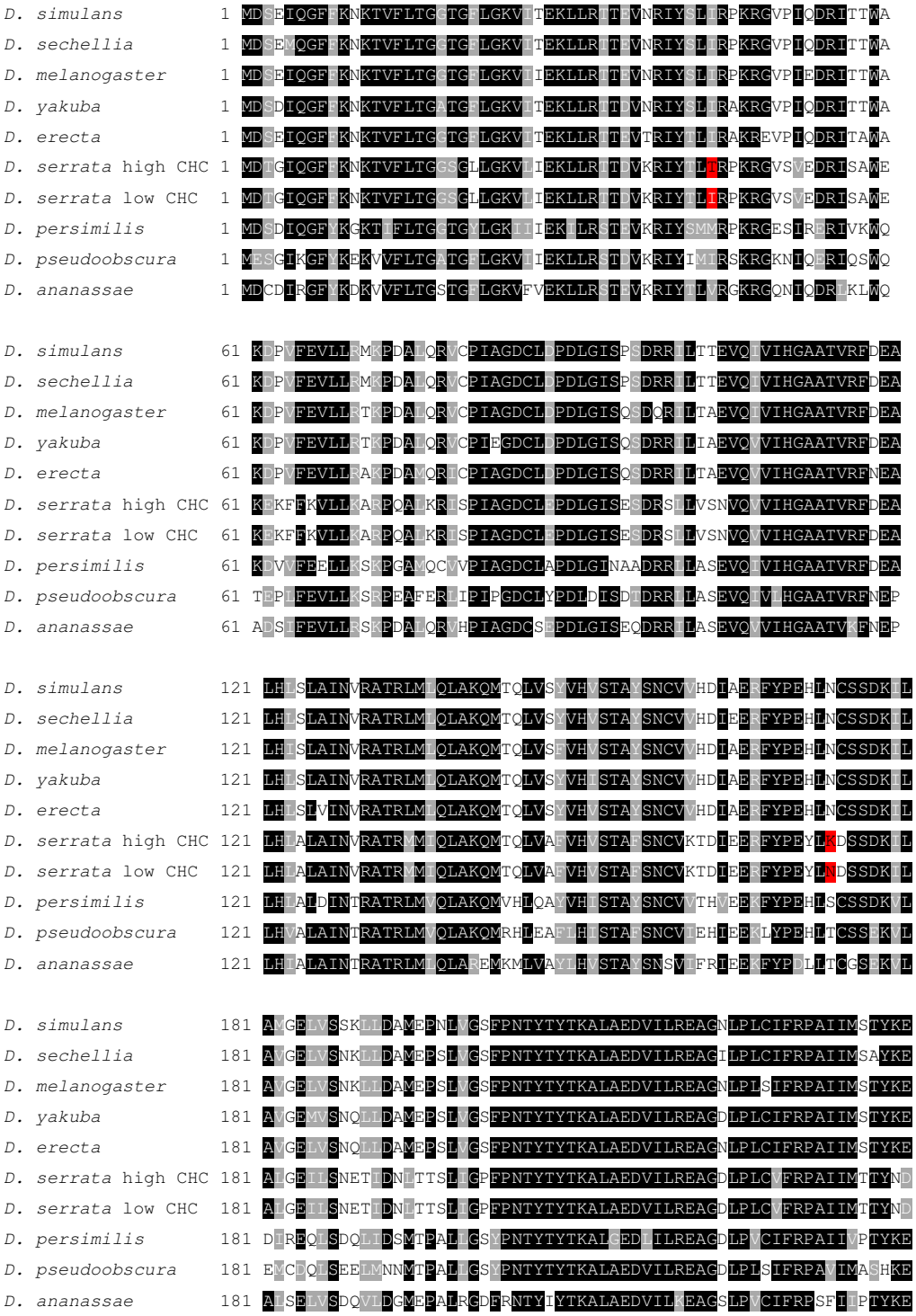
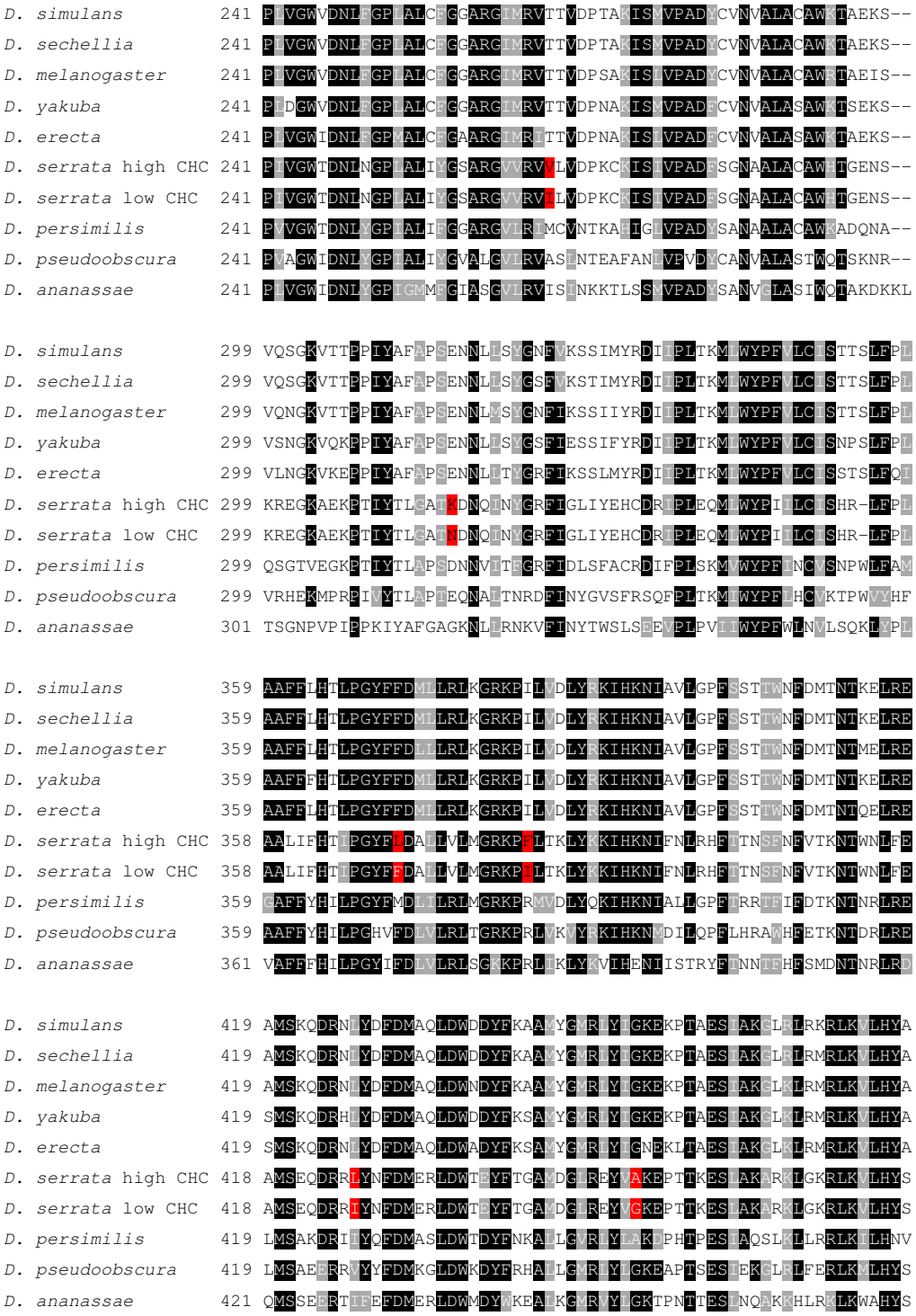


Figure 4.2

Possible causal SNPs differentiating 'low' and 'high' lines in the region of three adjacent reductase genes in *D. serrata*. Shaded areas correspond to estimated region of coding sequence including introns but does not include UTRs.


D. simulans 1 MDSEIQGFFKNKTVFLTGGTGFLGKVITEKLLRTTPVNRIYSLIRPKRGVPIQDRITTWA
D. sechellia 1 MDSEMQGFFKNKTVFLTGGTGFLGKVITEKLLRTTPVNRIYSLIRPKRGVPIQDRITTWA
D. melanogaster 1 MDSEIQGFFKNKTVFLTGGTGFLGKVIIEKLLRTTPVNRIYSLIRPKRGVPIEDRITTWA
D. yakuba 1 MDSDIQGFFKNKTVFLTGATGFLGKVITEKLLRTTPVNRIYSLIRAKRGVPIQDRITTWA
D. erecta 1 MDSEIQGFFKNKTVFLTGGTGFLGKVITEKLLRTTPVTRIYILIRAKREVPIQDRITAWA
D. serrata high CHC 1 MDTGIQGFFKNKTVFLTGGSGLLGKVLIEKLLRTTPVKRIYTLIRPKRGVSVEDRISAWE
D. serrata low CHC 1 MDTGIQGFFKNKTVFLTGGSGLLGKVLIEKLLRTTPVKRIYTLIRPKRGVSVEDRISAWE
D. persimilis 1 MDSDIQGFFKGKTIFLTGGTGFLGKIIIEKILRSTPVKRIYSMMRPKRGESIRERIVKWQ
D. pseudoobscura 1 MESGIKGFFKEKVVFLTGATGFLGKVIIEKLLRSTPVKRIYIMIRSKRGKNIQERIQSWQ
D. ananassae 1 MDCDIRGFFKDKVVFLTGSTGFLGKVFVEKLLRSTPVKRIYILVRGKRGQNIQDRLKLWQ
D. simulans 61 KDPVFEVLLRMKPDALQRVCPIAGDCLGPDLGISPSDRRILTTEVQIVIHGAATVRFDEA
D. sechellia 61 KDPVFEVLLRMKPDALQRVCPIAGDCLGPDLGISPSDRRILTTEVQIVIHGAATVRFDEA
D. melanogaster 61 KDPVFEVLLRTKPDALQRVCPIAGDCLGPDLGISQSDQRILTAEVQIVIHGAATVRFDEA
D. yakuba 61 KDPVFEVLLRTKPDALQRVCPIEGDCLGPDLGISQSDRRILIAEVQVVIHGAATVRFDEA
D. erecta 61 KDPVFEVLLRAKPDAMQRICPIAGDCLGPDLGISQSDRRILTAEVQVVIHGAATVRFNEA
D. serrata high CHC 61 KDKFFKVLLKARPQALKRISPIAGDCLGPDLGISESDRSLLVSNVQVVIHGAATVRFDEA
D. serrata low CHC 61 KDKFFKVLLKARPQALKRISPIAGDCLGPDLGISESDRSLLVSNVQVVIHGAATVRFDEA
D. persimilis 61 KDVVFEELLKSKPGAMQCVVPIAGDCLAPDLGINAADRRLLASEVQIVIHGAATVRFDEA
D. pseudoobscura 61 TGPLFEVLLKSRPEAFERLIPIPGDCLYPDLDISDTDRRILASEVQIVLHGAATVRFNEP
D. ananassae 61 ADSIFEVLLRSKPDALQRVHPIAGDCSGPDLGISEQDRRILASEVQVVIHGAATVKFNEP
D. simulans 121 LHLSLAINVRATRLMLQLAKQMTQLVSYVHVSTAYSNCVVHDIAERFYPEHLNCSSDKIL
D. sechellia 121 LHLSLAINVRATRLMLQLAKQMTQLVSYVHVSTAYSNCVVHDIEERFYPEHLNCSSDKIL
D. melanogaster 121 LHISLAINVRATRLMLQLAKQMTQLVSYVHVSTAYSNCVVHDIAERFYPEHLNCSSDKIL
D. yakuba 121 LHLSLAINVRATRLMLQLAKQMTQLVSYVHISTAYSNCVVHDIAERFYPEHLNCSSDKIL
D. erecta 121 LHLSLVINVRATRLMLQLAKQMTQLVSYVHVSTAYSNCVVHDIAERFYPEHLNCSSDKIL
D. serrata high CHC 121 LHLALAINVRATRVMIQLAKQMTQLVAFVHVSTAFSNCVKTDIEERFYPEYLKDSSDKIL
D. serrata low CHC 121 LHLALAINVRATRVMIQLAKQMTQLVAFVHVSTAFSNCVKTDIEERFYPEYLNDSSDKIL
D. persimilis 121 LHLALDINTRATRLMVQLAKQMVHLQAYVHISTAYSNCVYTHVEEKFYPEHLSCSSDKVL
D. pseudoobscura 121 LHVALAINTRATRLMVQLAKQMRHLEAFIHISTAFSNCVTEHIEEKLYPEHLTCSSFKVL
D. ananassae 121 LHIALAINTRATRLMLQLAREMKMLVAYIHVSTAYSNSVTFRIEEKFYPDLLTCGSFKVL
D. simulans 181 AVGELVSSKLLDAMEPNLVGSFPNTYTYTKALAEDVILREAGNLPLCIFRPAIIMSTYKE
D. sechellia 181 AVGELVSNKLLDAMEPSLVGSFPNTYTYTKALAEDVILREAGILPLCIFRPAIIMSAYKE
D. melanogaster 181 AVGELVSNKLLDAMEPSLVGSFPNTYTYTKALAEDVILREAGNLPLSIFRPAIIMSTYKE
D. yakuba 181 AVGEMVSNQLLDAMEPSLVGSFPNTYTYTKALAEDVILREAGDLPLCIFRPAIIMSTYKE
D. erecta 181 AVGELVSNQLLDAMEPSLVGSFPNTYTYTKALAEDVILREAGNLPLCIFRPAIIMSTYKE
D. serrata high CHC 181 ALGEIISNETLDNLTTSLVGPFPNTYTYTKALAEDVILREAGDLPLCVFRPAIIMTTYND
D. serrata low CHC 181 ALGEIISNETLDNLTTSLVGPFPNTYTYTKALAEDVILREAGDLPLCVFRPAIIMTTYND
D. persimilis 181 DIREQLSDQLLDSMTPALVGSYPNTYTYTKALGEDLILREAGDLPVCIFRPAIIVPTYKE
D. pseudoobscura 181 EMCDQLSEELVNNMTPALVGSYPNTYTYTKALAEDVILREAGDLPLSIFRPAVIMASHKE
D. ananassae 181 ALSELVSDQVLDGMEPALRGDFRNTYIYTKALAEDVILKEAGSLPVCIFRPSFIIPTYKE


D. simulans 241 PLVGWVDNLFGPFALCFGGARGIMRVTTVDPTAKISMVPADYCVNVALACAWKTAEKS--
D. sechellia 241 PLVGWVDNLFGPFALCFGGARGIMRVTTVDPTAKISMVPADYCVNVALACAWKTAEKS--
D. melanogaster 241 PLVGWVDNLFGPFALCFGGARGIMRVTTVDPSAKISLVPADYCVNVALACAWKTAEIS--
D. yakuba 241 PLDGWVDNLFGPFALCFGGARGIMRVTTVDPNAKISMVPADYCVNVALASAWKTSEKS--
D. erecta 241 PLVGWIDNLFGPFALCFGAARGIMRITTVDPNAKISLVPADFCVNVALASAWKTAEKS--
D. serrata high CHC 241 PIVGWTDNLNGPFALIYGSARGVVRVVLVDPKCKISIVPADFSGNAALACAWHTGENS--
D. serrata low CHC 241 PIVGWTDNLNGPFALIYGSARGVVRVHLVDPKCKISIVPADFSGNAALACAWHTGENS--
D. persimilis 241 PVVGWTDNLYGPFALIFGGARGVIRIMCVNTKAKIGLVPADYSANAALACAWKADQNA--
D. pseudoobscura 241 PVAGWIDNLYGPFALIYGVALGVLRVASLNTEAFANLVPVDYCANVALASTWQTSKNR--
D. ananassae 241 PLVGWIDNLYGPIGMMFGIASGVLRVISLNKKTLSSMVPADYSANVGLASIWQTAKDKKL
D. simulans 299 VQSGKVTTPPIYAFAPSENNLLSYGNFVKSSIMYRDIIPLTKMLWYPFYLCISTTSLFPL
D. sechellia 299 VQSGKVTTPPIYAFAPSENNLLSYGSFVKSTIMYRDIIPLTKMLWYPFYLCISTTSLFPL
D. melanogaster 299 VQNGKVTTPPIYAFAPSENNLLSYGNFIKSSIIYRDIIPLTKMLWYPFYLCISTTSLFPL
D. yakuba 299 VSNGKVQKPPIYAFAPSENNLLSYGSFIESSIFYRDIIPLTKMLWYPFYLCISNPSLFPL
D. erecta 299 VLNGKVKEPPIYAFAPSENNLLTYGRFIKSSLMYRDIIPLTKMLWYPFYLCISSTSLFQI
D. serrata high CHC 299 KREGKAAEKPTIYTLGATKDNQLNYGRFIGLIYEHCDRIPLEQMLWYPIFLCISHR-LFPL
D. serrata low CHC 299 KREGKAAEKPTIYTLGATNDNQLNYGRFIGLIYEHCDRIPLEQMLWYPIFLCISHR-LFPL
D. persimilis 299 QSGTVEGKPTIYTLAPSDNNVLTFGRFIDLSFACRDIFPLSKMVWYPFFNCVSNPWLFAM
D. pseudoobscura 299 VRHEKMPRPIVYTLAPTEQNALTNRDFINYGVSFRSQFPLTKMLWYPFFHCVKTPWYYHF
D. ananassae 301 TSGNPVPIPPKIYAFGAGKNLLRNKVFINYTWSLSEEVPLPVIIWYPFWLNVLSQKLYPL
D. simulans 359 AAFFLHILPGYFFDMLLRLKGRKPILVDLYRKIHKNIAVLGPFSSTTWNFDMTNTKELRE
D. sechellia 359 AAFFLHILPGYFFDMLLRLKGRKPILVDLYRKIHKNIAVLGPFSSTTWNFDMTNTKELRE
D. melanogaster 359 AAFFLHILPGYFFDLLLRLKGRKPILVDLYRKIHKNIAVLGPFSSTTWNFDMTNTMELRE
D. yakuba 359 AAFFFHILPGYFFDMLLRLKGRKPILVDLYRKIHKNIAVLGPFSSTTWNFDMTNTKELRE
D. erecta 359 AAFFLHILPGYFFDMLLRLKGRKPILVDLYRKIHKNIAVLGPFSSTTWNFDMTNTQELRE
D. serrata high CHC 358 AALIFHILPGYFFDALLVLMGRKPFLTKLYKKIHKNIFNLRHFITNSFNFVTKNTWNLFE
D. serrata low CHC 358 AALIFHILPGYFFDALLVLMGRKPILTKLYKKIHKNIFNLRHFITNSFNFVTKNTWNLFE
D. persimilis 359 GAFFYHILPGYFMDLILRLMGRKPRMVDLYQKIHKNIALLGPFTRRTFIFDTKNTNRLRE
D. pseudoobscura 359 AAFFYHILPGHVFDLVLRLTGRKPRLVKVYRKIHKNMDILQPFLHRAMHFETKNTDRLRE
D. ananassae 361 VAFFFHILPGYIFDLVLRLSGKKPRLIKLYKVIFENIISTRYFINNTFHFSMDNTNRLRD
D. simulans 419 AMSKQDRNLYDFDMAQLDWDDYFKAAVYGMRLYIGKEKPTAESIAKGLSLRKRLKVLHYA
D. sechellia 419 AMSKQDRNLYDFDMAQLDWDDYFKAAVYGMRLYIGKEKPTAESIAKGLSLRMRLKVLHYA
D. melanogaster 419 AMSKQDRNLYDFDMAQLDWNDYFKAAVYGMRLYIGKEKPTAESIAKGLSLRMRLKVLHYA
D. yakuba 419 SMSKQDRHLYDFDMAQLDWDDYFKSAVYGMRLYIGKEKPTAESIAKGLSLRMRLKVLHYA
D. erecta 419 SMSKQDRNLYDFDMAQLDWADYFKSAVYGMRLYIGNEKLTAESIAKGLSLRMRLKVLHYA
D. serrata high CHC 418 AMSEQDRRLYNFDMERLDWTEYFTGAVDGLREYVAKEPTTKESLAKARALGKRLKVLHYS
D. serrata low CHC 418 AMSEQDRRIYNFDMERLDWTEYFTGAVDGLREYVGKEPTTKESLAKARALGKRLKVLHYS
D. persimilis 419 LMSAKDRIIYQFDMASLDWTDYFNKAVLGVRLYLAKDPHTPESIAQSLSLLRRLKSLHNV
D. pseudoobscura 419 LMSAEDRRVYYFDMKGLDWKDYFRHAVLGMRLYIGKEAPTSESIEKGLSLFERLKVLHYS
D. ananassae 421 QMSSEDRTIFEFDMERLDWMDYWKEAVKGMRVYIGKTPNTTESLNQAKSHLRKLKWAHYS

```
D. simulans          479 FASSLVSLAGYVLMSLARLVV---
D. sechellia         479 FASSLVSLAGYVLMSLARLVV---
D. melanogaster      479 FASSLVALAGYILMSLARLVV---
D. yakuba            479 FASSLASLAGYALMSLGKLVV---
D. erecta            479 FASSLVSAAAYALMSLAKLVV---
D. serrata high CHC  478 FMASLGSLAAYGLMILAKFI----
D. serrata low CHC   478 FMASLGSLAAYGLMILAKFI----
D. persimilis        479 LKASLACGAGAILMSLSRLLIN--
D. pseudoobscura     479 LQVALCCVAGLLLMWLLKFIVFSI
D. ananassae         481 LVAVLIFLAGYVLMIVIKVFFT--
```

Figure 4.3

Amino acid alignment of *D. serrata* CG17560 against known *Drosophila* orthologs. Indicated in red for *D. serrata* are consensus 'low' and 'high' sequences based on the known fixed differences in table 4.2. Alignment was produced using CLUSTALW and shading was done using BOXSHADE.

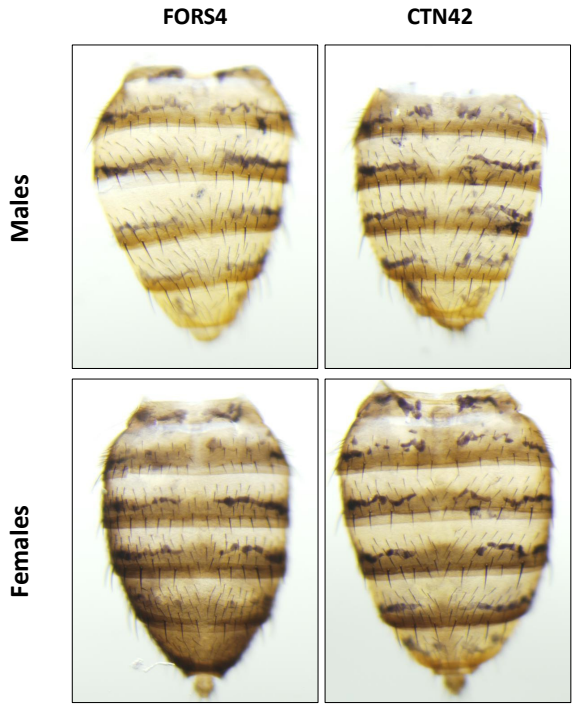


Figure 4.4

***In situ* hybridization showing expression of the GC17560 gene in adult oenocyte cells in male and female *D. serrata* in both high (FORS4) and low (CTN42) CHC phenotype lines. (Henry Chung, Sean Carroll, Steve Chenoweth, unpublished data).**

4.5 DISCUSSION

The aim of this study was to test the hypothesis that the same mutation underlies the 'low' CHC phenotype among nine random inbred lines originating from the same population (Cooktown), and to validate the results of the fine-scale QTL mapping experiment (Chapter 3) which implicated mutations within and around three adjacent fatty acid reductase loci. Qualitative complementation tests using all pairwise crosses between the five 'low' lines and the 'low' parental line from previous chapters were unequivocal in their findings, all 'low' phenotype crosses did not complement to restore the 'high' phenotype, indicating that the same mutation or set of mutations gives rise to this 'low' phenotype.

Most of the trait loci identified by a variety of QTL mapping studies are based on a sample of individuals from a specific inbred cross and may not be widely applicable to their natural populations because they only capture a partial picture of the full repertoire of genetic variation available in the source populations (BERTIN *et al.* 2012). In some cases, rare allelic variants with minimal phenotypic impact in the original population may incidentally account for a large fraction of trait variance in any particular mapping population (BEAVIS *et al.* 1994; SCOVILLE *et al.* 2011). Although QTL mapping of traits based on a single cross from a pair of inbred lines can reveal the genetic basis of trait variation, the use of a limited sample of lines lacks replication and may not reveal the exact nature of the importance of such QTLs to natural population variation and may potentially result in misleading generalisations about the genetic architecture of the relevant traits (BARNWELL and NOOR 2008a). GWAS studies have been used to various degrees of success to circumvent this weakness but the huge number of loci detected, typically accounting for little of the variation, impose particularly stringent significance levels that make solid conclusions about particular polymorphisms rather elusive (SCOVILLE *et al.* 2011).

Earlier in this study (Chapter 2 and 3), an F2 QTL endeavour was followed up with bulk segregant analysis to refine the mapping of CHC variation using a highly reshuffled genome of an advanced cross (YOUNG 1996; HOBBS and MACRAE 2012; OLSON *et al.* 2013). These QTL mapping strategies were, however, restricted to a single cross. Following up QTL-mapping work with a set of random replicate inbred lines to scan for the presence of earlier detected loci may be useful for discounting QTLs of fortuitous importance in single mapping crosses and thus validating those of actual importance in the wider natural setting (SCOVILLE *et al.*

2011). Therefore it was important to independently validate the findings of the fine-mapping bulk segregant analysis of Chapter 3. The number of fixed SNP differences distinguishing replicate 'low' from replicate 'high' lines was quite obviously maximal around the reductase genes. The consistency of results across these two independent experiments strongly implicates involvement of fatty acid reductase genes in the CHC polymorphism and also appears to rule out the involvement of elongases and desaturases as having major effects.

The mapping of this polymorphism to one major locus harbouring many mutations within three adjacent and functionally-related genes may indicate that a hot spot of genetic variation is involved in the polymorphism. The finding is especially compelling here because in most natural populations of *Drosophila,* linkage disequilibrium is very low which would under normal circumstances rule out finding a 10kbp block of differentiation between 'low' and 'high' lines sourced from a wild population. It may well be the case that the causal polymorphism is actually a combination of mutations in a haplotype rather than a single SNP. For example, a phenotypic polymorphism in respect to alcohol dehydrogenase expression levels in *Drosophila melanogaster* was found to be driven by multiple related genetic polymorphisms within the *Adh* gene (STAM and LAURIE 1996). Furthermore, the major effect QTL for cuticular pigmentation, at the *bric-a-brac* locus in *D. melanogaster* arises through cumulative effects of multiple polymorphisms in three different functional regions of chromosome 3L (BICKEL *et al.* 2011). However these regions are separated in some cases by up to 50kbp whereas the regions implicated in *D. serrata* are far smaller (<4kbp).

Interestingly, recent work using *in situ* hybridizations has confirmed that of the three fatty acyl-CoAs identified here, only CG17560 is expressed in *D. serrata* oeneocyte cells (Fig 4.4: Dr Henry Chung, Prof. Sean Carroll, A/Prof Steve Chenoweth, unpublished data). This is an important finding for two reasons. Firstly, it effectively rules out CG17562 and CG14893 as strong candidates underlying the CHC polymorphism in *D. serrata.* Secondly, because CG17560 is highly expressed in both low (CTN42) and high (FORS4) lines, it is more likely that protein-coding changes are involved than regulatory changes. It is remarkable that the majority (8 out of ten) SNPs in the coding region of CG17560 involved amino acid changes suggesting a history of positive selection.

It is very likely that these fixed mutations may be of functional consequence, although further studies, perhaps involving transformation into yeast (LASSANCE *et al.* 2010; LIENARD *et al.*

2010; LASSANCE *et al.* 2013), would be required to confirm this. This might involve comparison of the levels of three CHCs as expressed by the 'low' and 'high' CHC phenotype allelic strains of CG17560-inoculated yeast cells like those of *Saccharomyces cerevisiae*. Both strains would be incubated with galactose and fatty acids methyl ester precursors for one day followed by CHC extraction with n-hexane and GC-MS analyses. The Amino acid sequences of CG17560 and its orthologs contain the same Pfam domains as the Lepidopteran fatty acid reductase genes of *Sterile* and *NAD-binding* (Figure 4.3). In moths, amino acid changes in these genes confer both intra and interspecific variation in pheromone blends by directly affecting the degree to which different carbon-chain length precursors are converted into pheromones during the final step of biosynthesis. It has been demonstrated using transformation experiments in yeast that even when different precursors are available in equal amounts, different ratios of pheromone hydrocarbons are produced by different alleles. There may be a very similar mode of operation in *D. serrata* whereby one or more of the low genotype amino acid changes affect the conversion efficiency of the precursors of C24 and C25 cuticular hydrocarbons from their specific mono and di-saturated precursors.

While this experiment successfully validated the association between the fatty acid reductase loci and the CHC phenotype and broadened the generality of previous findings, future research efforts should focus on understanding whether the specific amino acid changes observed between low and high phenotype lines mediate this CHC polymorphism. This could be achieved via transformation of different *D. serrata* CG17560 alleles into yeast and testing their efficiency in the conversion of CHC precursors into CHCs. In addition, sequencing of a larger sample of natural genetic variation, perhaps from several other natural populations where the polymorphism occurs, may be effective in further excluding some of the candidate polymorphisms detected in this study.

CHAPTER 5. THE EFFECTS OF A MAJOR CUTICULAR HYDROCARBON POLYMORPHISM ON DESICCATION RESISTANCE, THERMAL TOLERANCE AND MATING SUCCESS IN *D. SERRATA*

5.1 ABSTRACT

Understanding the sources of selection on individual polymorphisms is critical to understanding their persistence in natural populations. Cuticular hydrocarbons (CHCs), a suite of waxy compounds deposited on the body surface, are remarkably diverse in insects where they enhance desiccation resistance and also mediate courtship. *Drosophila serrata*, an endemic Australian fruit fly species, has recently been observed to segregate a major effect polymorphism in three of its shortest carbon-chained CHCs, 5,9-tetracosadiene, 5,9-pentacosadiene and 9-pentacosene, in the far northern populations of its Australian range; one phenotype has vastly reduced levels of the chemicals while another has normal levels. This polymorphism has recently been fine-mapped as a single QTL comprising three adjacent fatty acid reductase genes. The ecological factors associated with this drastic phenotypic change are presently unclear. This study investigated fitness differences between the two QTL genotypes through its effects on desiccation resistance, heat shock stress and mating success. For most traits, the effects were sex-specific; desiccation resistance was significantly higher in 'low' females than 'high' females but there was no significant effect on males. No differences between the two phenotypes were found for survival to heat stress. Intriguingly, the polymorphism had a significant effect on male mating success; in competitive trials, 'low' males were only half as successful as 'high' males. Being of 'low' phenotype did not influence the chooser males' and females' probability of their mating with a 'low' of the opposite sex. Within the context of these phenotypes, the polymorphism has sexually antagonistic effects; enhancing survival in females at a cost to male mating success. These opposing natural and sexual selection processes may therefore explain why the polymorphism is maintained at intermediate frequencies in far northern populations despite large effects on either sex.

5.2 INTRODUCTION

Phenotypic differences are a common occurrence in nature (DARWIN 1859) and understanding the evolutionary processes that drive and maintain these trait disparities is a fundamental goal of evolutionary biology (TRAVISANO *et al.* 1995; KELLER and TAYLOR 2008; SANCHEZ-GUILLEN *et al.* 2011). Phenotypic differentiation takes place at many scales, between species and often among population within a single species. Sometimes a single population may harbour divergent phenotypes leading to a polymorphism. Polymorphisms can persist in a population despite strong fitness effects when selection acts antagonistically between heterogenous environments, individual fitness components, different phases of a lifecycle (haploid vs. diploid) or between males and females (FOERSTER *et al.* 2007; AHUJA and SINGH 2008; ANDERSEN *et al.* 2010; IMMLER *et al.* 2012). Other mechanisms like secondary contacts among populations that previously differentiated in allopatry (TURGEON and BERNATCHEZ 2001; BERMOND *et al.* 2012) and range expansion involving multiple successive founder events from genetically distinct source demes may also maintain genetic variation patterns of isolation by distance across geographical or environmental gradients (CLEGG *et al.* 2002; VASEMAGI 2006; KELLER *et al.* 2009). Such non-antagonistic evolutionary processes are, however, not predicted to enhance the same spectrum of additive genetic variance for fitness or its components to a degree comparable with antagonistically driven selection (RICHMAN 2000; CONNALLON and CLARK 2013). Antagonistic selection also maintains a wider diversity of alleles in the population for much longer than would be possible under selectively neutral genetic variation (RICHMAN 2000). More additive genetic variance for fitness may thus accrue from balanced genetic polymorphism acquired through antagonistically selected alleles than from alleles maintained by non-antagonistic mechanisms (FOERSTER *et al.* 2007). Selection acting in this manner and targeting genes of major effect sustains allelic variation for a diversity of traits in nature (BHUTKAR *et al.* 2007) including alternate social organizations in ants (KRIEGER and ROSS 2002), colour polymorphisms in pepper moths (M'BALAKA *et al.* 2013), mice (BARNWELL and NOOR 2008b), humans and fish (LAMASON *et al.* 2005) and sickle cell anaemia genes in vertebrates (WOOD *et al.* 2005).

Cuticular hydrocarbons (CHCs) are a suite of traits that exhibit remarkable divergence and broadly function in promoting survival to abiotic stress (VAZ *et al.* 1988) and mediating courtship in insects (HOWARD *et al.* 2003). These waxy compounds, usually deposited on the

insect cuticle, are produced by oenocyte cells in the abdomen through elongation of fatty acids and eventual decarbonylation of the intermediate aldehydes (HOWARD and BLOMQUIST 2005). The hydrocarbons are important in maintaining water balance through reducing water loss (HILL and ROBERTSON 1966; MONTOOTH and GIBBS 2003), a particularly significant role in an insect life given that susceptibility to dehydration is one of the most important environmental stresses that terrestrial arthropods are exposed to in nature (PUNZALAN *et al.* 2005). Physiological capacity for desiccation and temperature tolerance may often determine insects' ability to cope with changing climatic conditions and hence their persistence and relative abundance in some habitats (IMMLER *et al.* 2012; CONNALLON and CLARK 2013). Insects from warmer, drier environments exhibit a variety of adaptive differences in water balance including reduced cuticular permeability, reduced excretory water loss and differences in the quantity and composition of cuticular lipids (PUNZALAN *et al.* 2005; BASSETT *et al.* 2013). QTLs for desiccation resistance tend to localise with QTLs for CHCs suggesting a pleiotropic genetic basis (RULE and ROELOFS 1989).

In *Drosophila*, CHCs, are also known to be sensitive to temperature (IRAQI *et al.* 2000; DILLON *et al.* 2009; PEGUEROLES *et al.* 2010). Temperature-driven changes in CHC composition alter cuticular permeability and have effects on the desiccation resistance of these insects (PEGUEROLES *et al.* 2010). Genetic variation for desiccation and temperature resistance may impact on the species range of *Drosophila* species (KNAPP *et al.* 1990; WANG *et al.* 2003; KELLERMANN *et al.* 2009). Geographical changes in *Drosophila* CHC compounds have been reported across the world (LOCKEY 1988; VAZ *et al.* 1988; WU *et al.* 1998; FRENTIU and CHENOWETH 2010; MAAN and SEEHAUSEN 2011), some of which correlate well with desiccation and/or temperature tolerance (LOCKEY 1988; BERRY and KREITMAN 1993; TILLMAN *et al.* 1999; PUNZALAN *et al.* 2005).

In addition to their likely impacts on abiotic stress resistance, CHCs profiles also affect mate and species recognition in *Drosophila* (JALLON 1984; HOWARD *et al.* 2003; CHENOWETH and BLOWS 2005; MCGUIGAN *et al.* 2008; ROUGHGARDEN and AKCAY 2010b). Non-volatile long-chain CHCs function as gustatory chemical cues that promote sex-specific mating behaviour, enhancing male courtship and female sexual receptivity but blocking non-productive male-male courtship (ROUGHGARDEN and AKCAY 2010b; ROUGHGARDEN and AKCAY 2010a). CHC phenotype affects the likelihood of heterospecific mating attempts between *D. serrata* and *D.*

birchii (HIGGIE *et al.* 2000). CHCs evolve under experimental sexual selection in both *D. pseudoobscura* (HUNT *et al.* 2012) and *D. serrata* (CHENOWETH and BLOWS 2008).

Drosophila serrata Malloch (MALLOCH 1927; DOBZHANSKY and MATHER 1961) is a fruit fly of the *montium* subgroup of the *melanogaster* species group endemic to the eastern Australian coast where its distribution traverses both temperate and tropical climates (JENKINS and HOFFMANN 2000; LIEFTING *et al.* 2009). The role of both sexual and natural selection in the evolution of its CHC traits, which constitute its mate and species recognition system, has been experimentally demonstrated (BLOWS 2002; CHENOWETH and BLOWS 2005; RUNDLE *et al.* 2009). In its natural settings, environment-related latitudinal clines in its CHCs that may be a consequence of divergent natural selection associated with variation in temperature, rainfall and humidity have also been documented (FRENTIU and CHENOWETH 2010).

CHCs in *D. serrata* comprise 5,9-tetracosadiene (5,9-C_{24}), 5,9-pentacosadiene (5,9-C_{25}), 9-pentacosene (9-C_{25}), 9-hexacosene (9-C_{26}), 2-methyl-hexacosane (2-Me-C_{26}), 5,9-heptacosadiene (5,9-C_{27}), 2-methyl-octacosane (2-Me-C_{28}), 5,9-nonacosadiene (5,9-C_{29}) and 2-methyl-triacontane (2-Me-C_{30})(BLOWS and ALLAN 1998; HOWARD *et al.* 2003; HIGGIE and BLOWS 2007). The species has recently been reported to exhibit a polymorphism in its three shortest carbon-chain CHCs (Chapter 2, this thesis) and two distinct phenotypes exist; one phenotype has reduced levels of 5,9-tetracosadiene, 5,9-pentacosadiene and 9-pentacosene while another has normal amounts (Figure 5.1). The frequency of this major-effect polymorphism is highest in the northern stretches of the Australian Eastern coast (Chapter 2, Figure 2.4). The environmental factors and processes associated with this drastic phenotypic change are yet known. Because of the possible influences of CHCs on abiotic stress as well as mate choice, this set of experiments was aimed at investigating any potential differences in the performance of these two phenotypes and thus shed light on the extent to which natural and/or sexual selection may be implicated in the maintenance of this polymorphism.

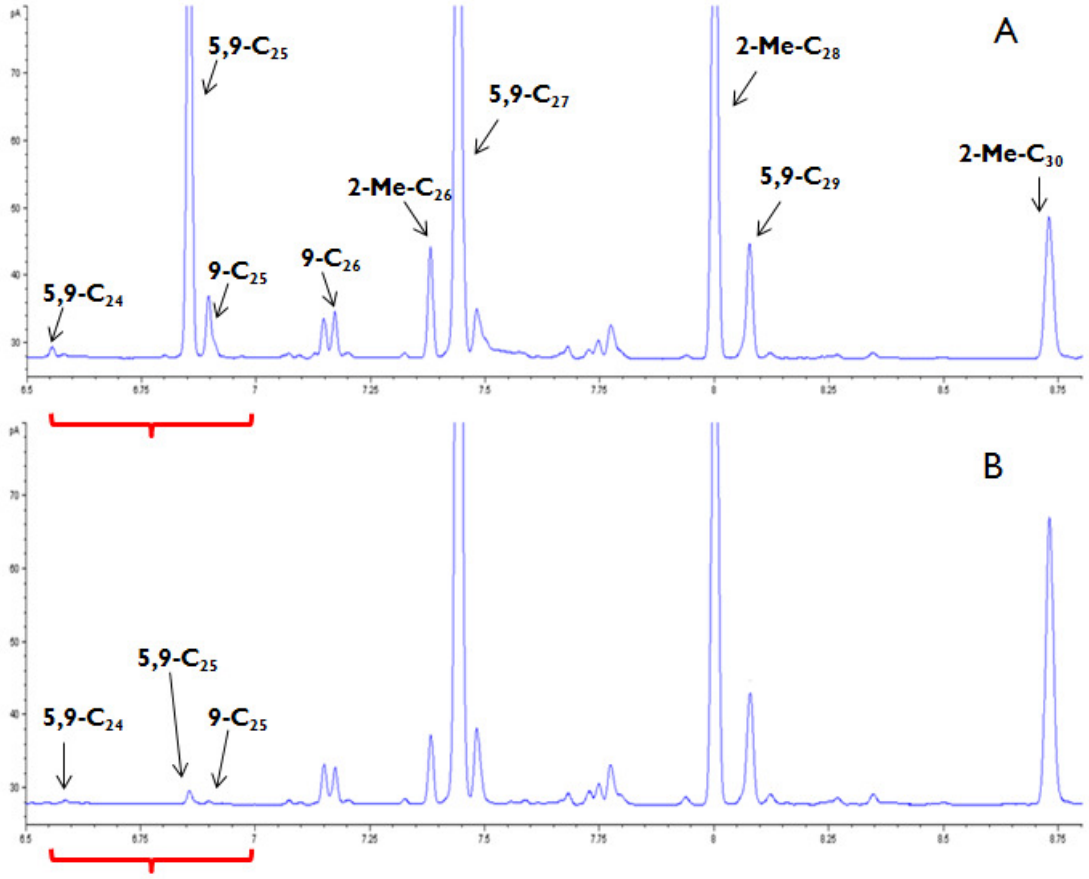


Figure 5.1

Cuticular hydrocarbon profile of nine traits of increasing linear carbon chain length in *Drosophila serrata,* showing 'high' (A) and 'low' (B) polymorphism for 5,9-tetracosadine (5,9-C_{24}), 5,9-pentacosadiene (5,9-C_{25}) and 9-pentacosene (9-C_{25}). The X-axis shows the relative time (increasing from left to right) it took for a trait to be detected on the gas chromatograph and hence its relative size while the Y-axis depicts the signal strength from the detector and is proportional to the abundance of that compound.

5.3 MATERIALS AND METHODS

5.3.1 *Study populations*

I used a panel of wild-derived inbred lines from Cooktown ($15^0$28'11.62"S, $145^0$14'55.27"E) northern Australia for this study that were fixed for the 'low' (four lines) and 'high' (four lines) CHC phenotypes. These lines were established from wild caught females and inbred

for 17 generations for full-sib mating post capture. To avoid effects of inbreeding on the traits I created a series of reciprocal intercrosses within each phenotypic class (the CHC mutation is recessive and so inter-crosses still retain the phenotype status of the two parental lines, as shown in Chapter 4) from four 'low' and four 'high' lines resulting in a total of 12 'low' and 12 'high' genotypes for testing. Although the number of lines that could be established from these populations was limited, I include analyses at the genotype (F1) line level here to allow me to correct for residual genotypic effects that may influence phenotype differences over and beyond their CHC phenotype status. This is a conservative approach.

The lines were reared at a constant temperature of 25^0C on a 12:12 hr light: dark cycle. Flies were maintained on the same standard fly diet comprising 36g agar, 108g raw sugar and 74g yeast, cooked in 2L of water combined with 24mL propionic acid and 12mL nipagen. Live yeast was sprinkled on the media surface to enhance oviposition (DOERGE and CHURCHILL 1996). For both desiccation resistance and heat stress tolerance, fly vial replicates from each line were made up from different culture vials to offset the confounding effects of the common environment on the variation among the lines (SCOVILLE *et al.* 2011).

5.3.2 *Experimental flies*

At less than one day post-eclosion, flies were separated into males and females under light CO_2 anaesthesia and then held individually in 7mL food vials until they were 4 days old. Thirty virgin males and thirty virgin females per line were then randomly selected from all possible holding vials and held as three mating pairs (3 males - 3 females) per vial containing 10mL of food media sprinkled with live yeast. To standardise larval density per vial and hence mitigate against related effects of larval environment on the offspring phenotypes (SCOVILLE *et al.* 2011), the pairs were allowed to mate for 4 days after which the vials were cleared of the flies. The offspring from these vials were pooled per cross and kept separate by sex in vials of five before the mating, desiccation and heat tolerance trials.

5.3.3 *Desiccation resistance*

A modified version of the procedure for desiccation assay in PUNZALAN *et al.* (2005) and CONNALLON and CLARK (2013) was followed in this experiment. Desiccation tolerance was measured by enclosing 20 flies of each sex (from each of the six 'low' and six 'high' line crosses) in a standard 30ml *Drosophila* vial without any medium. Approximately 3g of fresh Drierite desiccant was suspended on the top end of the vial on cotton gauze. The vial was then

sealed off with parafilm to sustain low humidity. Trial runs indicated that this set up reduced humidity in the vial to 10% in 30 minutes. Flies were observed at hourly intervals for death, as indicated by failure to right themselves or to move their legs when their vials were tapped or inverted. The time elapsed until 10 flies were dead (completely immobile) was recorded and defined as LT_{50}. As the number of dead flies neared 10, vials were checked for death every 20 min. Because temperature parameters alone can account for up to 97% of variability in desiccation tolerance (BEAVIS *et al.* 1994), all tests were carried out under a uniform temperature of 25^0C.

5.3.4 *Heat shock stress survival*

Heat tolerance was measured as survival after a high (potentially lethal) temperature stress (GREENBERG *et al.* 2003) at 4-6 days post-eclosion. Twelve replicates of twenty flies for each sex per line cross were enclosed in an empty stoppered vial and placed in a constant temperature cabinet set at 38^0C for 30 minutes. Flies were then transferred to fresh media vials and left to recover at 25◦C for 24 hours. The number of survivors per replicate vial was then scored.

5.3.5 *Mating success assay*

In *D. serrata*, it is known that both female and males choose mates on the basis of their CHCs (CHENOWETH and BLOWS 2005). I therefore also assayed female mating success in addition to male mating success. Disparities in mating success between phenotypes were assessed by determining whether a focal 'low' or 'high' chooser fly preferred to mate with a 'low' or 'high' suitor when potential mates of both types of phenotypes were availed to them. Time taken before mating was also recorded. Briefly, individuals were collected as virgins soon after eclosion and held separately by sex in yeasted food vials for four days. On the fourth day, a focal fly ('low'/'high') was presented with two flies of the opposite sex, one from its own phenotype (not from the same line as the focal fly) and another from a contrasting phenotype. Once the focal fly had chosen its mate and copulation commenced, the unsuccessful fly was taken out of the vial and identified. To facilitate identification of the two potential suitors, one of either fly phenotype was slightly wing-clipped. The number of wing-clipped competing flies (left wing clip and right wing clip) was kept even between the two phenotypes to balance out any effects of wing clipping on courtship success.

5.3.6 *Statistical Analyses*

Experimental data were analysed using mixed effects linear models fitted using the MIXED (Gaussian response variables) or GLIMMIX (binary responses) procedures of SAS (version 9.3, SAS Inst. Cary, NC). Desiccation resistance was modelled as:

$$Y = \mu + S + M + [S \times M] + G_{(M)} + [S \times G_{(M)}] + \varepsilon \quad [1]$$

where, Y was the time taken for 50% of the test flies to die, μ is the intercept, S is the sex of the flies, M is the 'low' status of the flies, $G_{(M)}$ is the random genotype nested within 'low' status of the flies and ε is the random error. When the sexes are analysed together, the experiment is effectively a split-plot design and so line nested within 'low' status and an interaction between sex and line within 'low' status were also fitted. The error term for the 'low' effect is $G_{(M)}$.

A generalized linear mixed effects model was fitted to the heat shock stress survival data as:

$$Y= \mu + S + M + [S \times M] + G_{(M)} + [S \times G_{(M)}] + \varepsilon \quad [2]$$

where Y was survival (dead or alive) after 24 hours post-exposure to heat shock, μ is the intercept, S is the sex of the flies, M is the 'low' status of the flies, G_M is the genotype for the 'low' status of the flies and ε is the random error.

The mating assay data was analysed in males and females separately. Mating success was assessed using a binary mixed model through Proc GLIMMIX as:

$$Y = \mu + SM + CM + [SM \times CM] + G_{(SM)} + G_{(CM)} + \varepsilon \quad [3]$$

where Y was mating success of the focal fly (accepted or rejected), μ is the intercept, SM is the 'low' status of the suitor, CM is the 'low' status of the chooser, $G_{(SM)}$ is the random genotype nested within the 'low' status of the suitor, $G_{(CM)}$ is the random genotype nested within the 'low' status of the chooser and ε is random error.

5.4 RESULTS

5.4.1 *Desiccation resistance*

The CHC polymorphism affected desiccation resistance in a sex dependent manner (Table 1). A significant sex and 'low' status interaction indicated that for female flies, 'low's had higher desiccation resistance than 'high's. However there was no detectable effect of the mutation on male desiccation resistance. Both 'low' and 'high' females survived longer than their 'low' and 'high' male counterparts (Table 5.1 and Figure 5.2).

Table 5.1

The effects of sex and 'low' status on desiccation resistance, survival to heat shock stress and mating success (male and female) across 'high' and 'low' short-chained CHC phenotypes in Cooktown lines of *Drosophila serrata*.

Trait	Effect	d.f.	F-value	P-value
Desiccation Resistance	Sex	1,10	357.22	p<0.0001
	'Low' Status	1,10	2.41	p=0.1513
	Sex × 'Low' Status	1,10	5.29	p=0.0442
Heat Shock Stress Survival	Sex	1,10	19.4	p=0.0006
	'Low' Status	1,10	2.24	p=0.4819
	Sex × 'Low' Status	1,10	5.7	p=0.0381
Male Mating Success	Chooser 'Low' Status	1,10	2.33	p=0.1582
	Suitor 'Low' Status	1,10	26.8	p=0.0004
	Chooser 'Low' Status × Suitor 'Low' Status	1,275	2.58	p=0.4545
Female Mating Success	Chooser 'Low' Status	1,10	0.44	p=0.5203
	Suitor 'Low' Status	1,10	1.42	p=0.1093
	Chooser 'Low' Status × Suitor 'Low' Status	1,343	1.59	p=0.2081

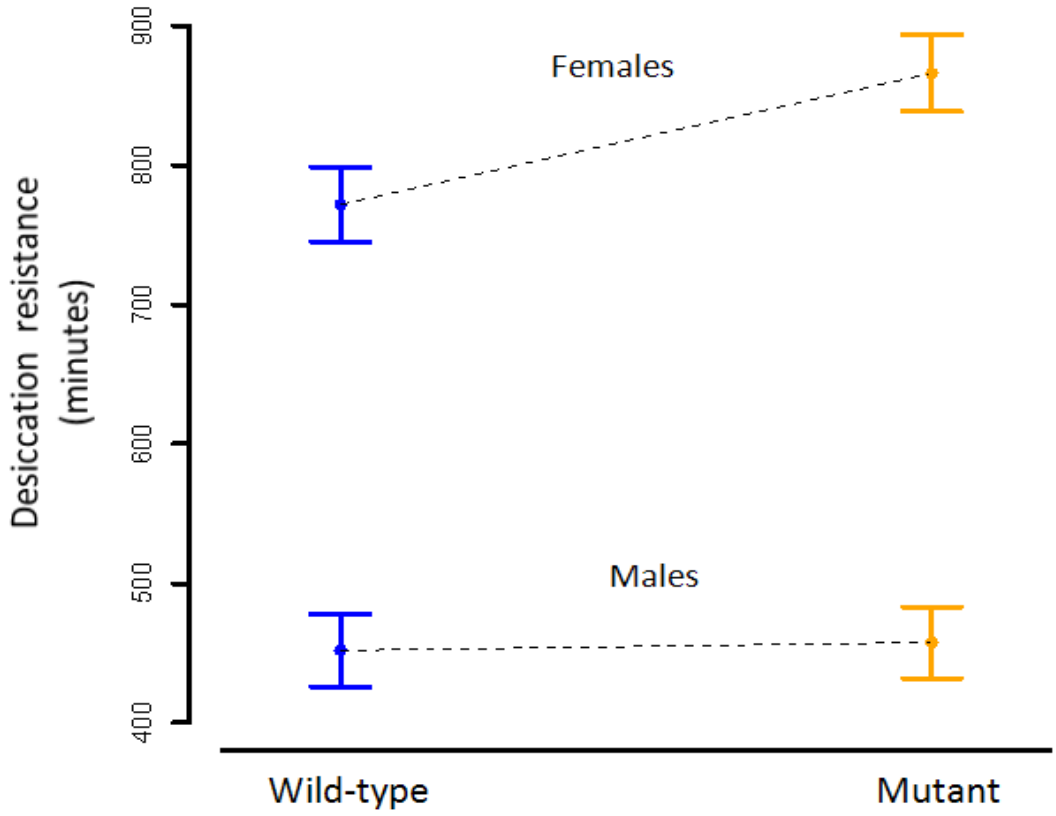


Figure 5.2

Differences in survival to desiccation between short-chained CHC 'high' (blue) and 'low' (orange) phenotypes in *Drosophila serrata* from Cooktown measured as the time taken for half of the tested flies to succumb to desiccation (Y-axis). Error bars are standard errors. In total 78 trials were conducted for the 'low' phenotype (n= 33 for females, n = 45 for males) and 79 trials for the 'high' phenotype (n = 37 for females, n = 42 for males).

5.4.2 *Heat shock stress survival*

The effect of CHC polymorphism on survival to heat stress was also sex dependent as indicated with a significant sex by 'low' status interaction (Table 5.1). The direction of the genotype effect was different between sexes with 'low' females having higher survival than 'high' females but 'low' males having lower resistance than 'high' males. However none of

the within-sex among genotype comparisons were significant in post hoc tests conducted at line mean level (these were highly significant for males when post-hoc tests were conducted at the level of individual vials). The mutation appeared to increase the overall level of sexual dimorphism in heat shock resistance with sexual dimorphism present in 'low's but absent in 'high's (Table 5.2, figure 5.3).

Table 5.2

Differences in desiccation resistance, survival to heat shock stress and mating success between 'low' and 'high' short-chained CHC phenotypes in *Drosophila serrata* lines from Cooktown (Significant post-hoc comparisons based on t-values are indicated by an asterisk).

Desiccation resistance

'Low' Female	'High' Female	'Low' Male	'High' Male	df	t-value	p-value
866.03 ± 27.58	771.72 ±26.83			10	2.45	0.0342*
866.03 ± 27.58		457.64 ±25.63		10	14.89	0.0001*
866.03 ± 27.58			451.95± 26.01	10	10.92	0.0001*
	771.72 ±26.83	457.64 ±25.63		10	8.46	0.0001*
	771.72 ±26.83		451.95± 26.01	10	11.82	0.0001*
		457.64 ±25.63	451.95± 26.01	10	0.16	0.8793

Heat Shock Stress Survival

'Low' Female	'High' Female	'Low' Male	'High' Male	df	t-value	p-value
0.28 ±0.04	0.21 ±0.04			10	1.23	0.2471
0.28 ±0.04		0.06 ±0.04		10	4.88	0.0006*
0.28 ±0.04			0.14 ±0.04	10	2.24	0.0349
	0.21 ±0.04	0.06 ±0.04		10	2.59	0.0268*
	0.21 ±0.04		0.14 ±0.04	10	1.63	0.1339
		0.06 ±0.04	0.14 ±0.04	10	-1.29	0.2265

Mating Success: Proportion of successful matings

'Low' Female	'High' Female	'Low' Male	'High' Male	df	t-value	p-value
0.42 ±0.04	0.49 ±0.04			10	1.42	0.2603
		0.36± 0.04	0.67 ±0.04	10	10	0.0004*

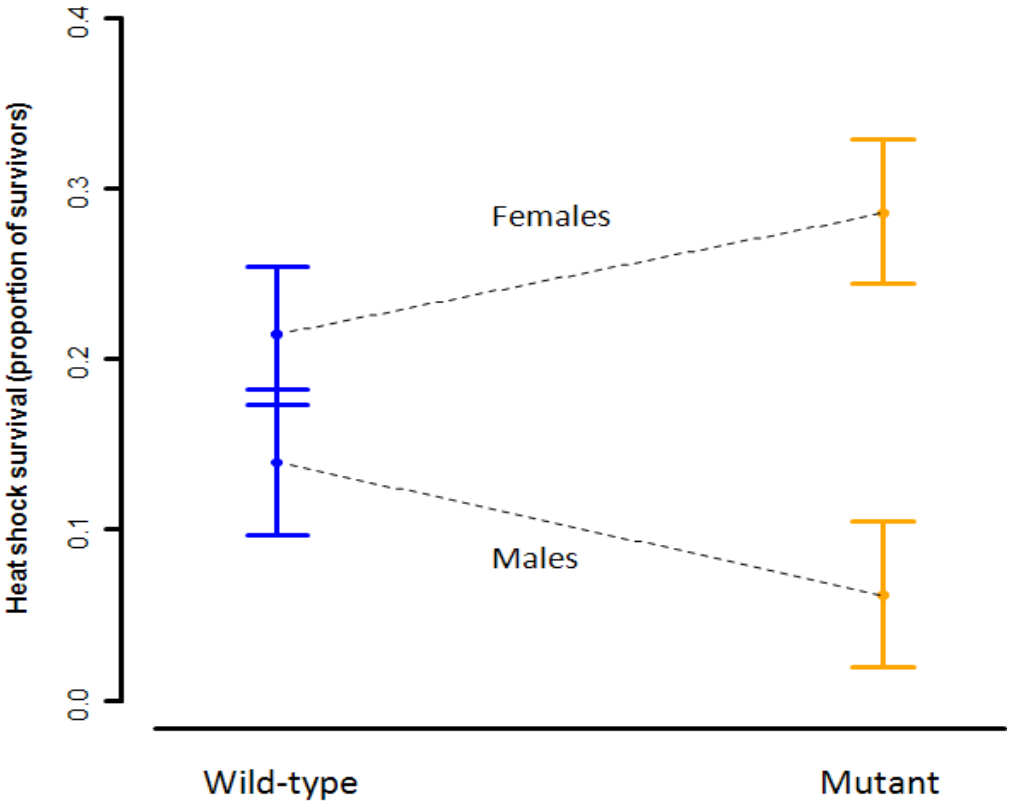


Figure 5.3

Differences in survival to heat shock stress between 5,9-pentacosadiene CHC 'high' (blue) and 'low' (orange) phenotypes in *Drosophila serrata* measured as the fraction of the number of the tested flies that recovered 24 hours after exposure to lethal temperature (Y-axis). Error bars are standard errors. In total 118 trials were conducted for the 'low' phenotype (n= 71 for females, n = 47 for males) and 130 trials for the 'high' phenotype (n = 69 for females, n = 61 for males)

5.4.3 *Mating success*

A total of 585 mating trials were conducted, involving 158 female and 187 male 'low' choosers as well as 107 female and 131 male 'high' choosers. About 98% of these trials resulted in copulation. The CHC 'low' status of the suitors had an impact on male mating success; 'high' males were almost twice more likely to succeed in mating than 'low' males

(Figure 5.4). The interaction between 'low' status of male suitor and female chooser flies was not significant, indicating that there was no change in the 'low' effect on male mating success when courting 'low' females as opposed to 'high' females (Table 5.1). In females, the CHC 'low' status of the suitor did not significantly affect female mating success and both 'high' and 'low' females had an equal probability of mating. The interaction between 'low' status of the chooser and suitor was non-significant in the male mate choice assays (Table 5.1).

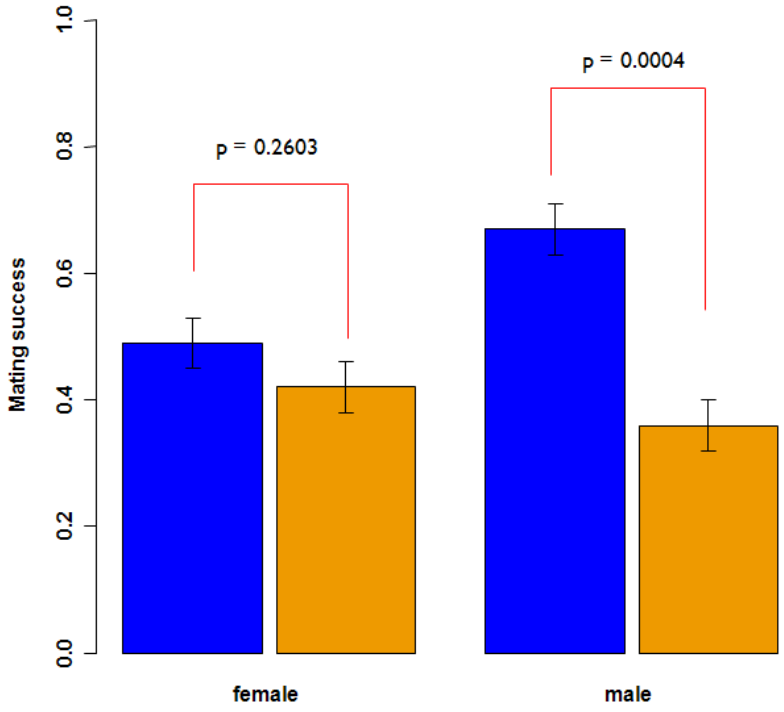


Figure 5.4

Differences in mating success between short-chained CHC 'high' (blue) and 'low' (orange) phenotypes in *Drosophila serrata* lines from Cooktown measured as the fraction of total courtship attempts that resulted in copulation (Y-axis). Error bars are standard errors. A total of 583 trials were conducted.

5.5 DISCUSSION

The goal of this study was to begin to investigate the possible sources of natural and sexual selection acting on a short-chained CHC polymorphism with an unusual geographical distribution. The CHC polymorphism increases in frequency northwards and would appear to be adaptive. CHCs are known to vary in correspondence with latitudinal climatic factors like temperature and humidity (TILLMAN *et al.* 1999; FRENTIU and CHENOWETH 2010) and are implicated in insect responses to desiccation stress (LOCKEY 1988; PUNZALAN *et al.* 2005; MAAN and SEEHAUSEN 2011). Owing to the role that CHCs play in sexual selection in this species (CHENOWETH and BLOWS 2005), I also assayed male and female mating success in addition to the classic environmental stress traits of heat and desiccation resistance. I found effects of the mutation on both stress traits and mating success, but the results were sex-dependent, suggesting a complex interplay between selection on males and females may be maintaining this polymorphism.

In this study, 'low' females were found to cope better than 'high' females under desiccating conditions. Desiccation resistance can be achieved physiologically through either increasing the quantity of water in the body, or through minimising rates of water loss (DARVASI and SOLLER 1995). It is this water loss component that CHCs are thought to mediate (PEGUEROLES *et al.* 2010). The major polymorphism in this study involved changes in 5,9-pentacosadiene. It is interesting that another study on the links between desiccation resistance and CHCs in *D. pseudoobscura* found that differences in cuticular permeability were largely driven by variance in the proportion of n-pentacosadiene and that high transcuticular water loss rates were generally correlated with higher proportions of relatively short-chain CHCs because of their role in enhancing cuticle permeability (PEGUEROLES *et al.* 2010). The fact that the 'low's for this polymorphism, which involved the loss of short-chained CHCs, were more resistant to desiccation may indicate lower rates of water loss through the cuticle.

It is puzzling why the desiccation resistance of females actually increased in a genotype that has effectively lost three of its compounds including its most abundant; 5,9-pentacosadiene. However the effect may be due to a concomitant increase in the longer chained compounds. Higher proportions of longer chain lengths of both methyl-branched alkanes and alkadienes are associated with lower cuticular permeabilities and lower water loss rates in *D. pseudoobscura* (PEGUEROLES *et al.* 2010). Female *Drosophila melanogaster* selected for high

desiccation stress resistance have significantly higher relative concentrations of long chain-length CHCs and concomitant higher lipid melting temperatures than control populations (BERRY and KREITMAN 1993; PUNZALAN *et al.* 2005). Furthermore, in some arthropods, acclimation to desiccation stress involves the mobilisation and deposition of long methyl-branched alkanes on the epicuticle (COYNE and ELWYN 2006; CLARK and WORLAND 2008; PEGUEROLES *et al.* 2010). The better survival of female 'low' *D. serrata,* with more long carbon chain CHCs, found in this study resonates well with these findings.

Within either phenotype, females survived longer under desiccating conditions than males. Apart from the effects of differences in environment (LOCKEY 1988; MORGAN and MACKAY 2006), desiccation resistance may also vary with sex (MCMILLAN and ROBERTSON 1974; AYALA *et al.* 2013). This pattern of sexual dimorphism in desiccation resistance occurs commonly in *Drosophila* species (MCMILLAN and ROBERTSON 1974; GREENBERG *et al.* 2003; MORGAN and MACKAY 2006; AYALA *et al.* 2013). Differences in desiccation tolerance correlate well with differences in body size; small sizes tend to lead to lower desiccation stress tolerance (AGRAWAL 2001; GREENBERG *et al.* 2003; NORRY *et al.* 2004). Although sizes were not measured in this study, male-female size dimorphism is similar for *D. serrata* as it is with other *Drosophila* with males 30% smaller than females (MAGIAFOGLOU and HOFFMANN 2003). Desiccation resistance in *D. serrata* is positively genetically correlated with wing length, a proxy for overall size (GILLESPIE 1984). The observed sex-based disparities in survival to desiccation in this study may reflect the effect of size differences between the sexes.

This study focused on a major CHC clinal polymorphism that is strongest towards the equator and is likely under the influence of environmental factors (FRENTIU and CHENOWETH 2010). Changes in insect CHCs have been implicated in temperature stress traits (TOOLSON and KUPERSIMBRON 1989). It is intriguing that although *Drosophila* females are more resistant to heat shocks and tend to have an extended lifespan if repeatedly exposed to a mild heat stress relative to their male counterparts (WADA *et al.* 1990; GREENBERG *et al.* 2003; BASSETT *et al.* 2013), there were no differences in survival to lethal temperatures between the sexes among 'high' flies. However, males and females differed in survival if they were 'low', with females surviving longer than 'low' males. Although not statistically significant, the direction of change in survival between 'low's and 'high's in females (positive) was opposite that in males (negative). The short-chained CHC polymorphism seems to bring about sexual dimorphism in

survival to heat stress. The CHC polymorphism in this study was mapped to a single genomic location containing three reductase genes on the right arm of the third chromosome (Chapters 2, 3 and 4). The right arm of chromosome 3 in *Drosophila* has also been reported to contain a dominant large-effect QTL for knockdown resistance to high temperature (NORRY *et al.* 2004). The findings here suggest the potential for sex-specific selection for heat shock survival acting on this genomic region. Most thermotolerance QTLs in *D. melanogaster* are sex-specific, probably pointing to possibility of the existence of unique genetic architectures for resistance to temperature extremes (MORGAN and MACKAY 2006).

It is interesting that the locus implicated in this CHC polymorphism, which affects survival to extreme heat, sits close to a region harbouring desaturase genes whose expression is part of the molecular basis for cold tolerance and cold acclimation response of organisms under changing environmental temperatures (LOS *et al.* 1997; LOS and MURATA 1998; VEGA *et al.* 2004). In a study that used knockout manipulation, alleles of *Desat2*, a gene known to affect CHC variation (TAKAHASHI *et al.* 2001), were implicated in cold tolerance in *D. melanogaster* (GREENBERG *et al.* 2003). Although the generality of this finding has been challenged (COYNE and ELWYN 2006), cold tolerance and acclimation response are known to correlate with increased expression of Δ6, Δ9, Δ12 and ω3 desaturase genes and concomitant shifts in fatty acid production that favour di-unsaturated fatty acids relative to monounsaturated fatty acids (WADA *et al.* 1990; MURATA and WADA 1995; SAKAMOTO *et al.* 1997; VEGA *et al.* 2004). This may not be surprising because levels of unsaturated fatty acids in membrane lipids may determine survival to low temperatures (TASAKA *et al.* 1996; LOS and MURATA 1998). The results of this experiment may be in agreement with studies that find QTL co-localization for cold- and heat-stress resistance traits on a single QTL (NORRY *et al.* 2004), that may lead to antagonistic effects between the two traits; decreased resistance to cold leading to increased resistance to heat (MORGAN and MACKAY 2006; SINCLAIR *et al.* 2007; CLARK and WORLAND 2008). Such genetic trade-offs between heat- and cold-stress have been reported to occur on the right arm of chromosome 3 in *D. melanogaster* (ANDERSON *et al.* 2003; MORGAN and MACKAY 2006).

In this study, males 'low' for a polymorphism that enhanced desiccation tolerance in females had far lower mating success than 'high' males. Mating seldom occurs randomly in nature and may be driven by mate preferences based on a diversity of traits (ANDERSSON 1994).

Shorter chained male CHCs are vital for stimulating female mating in *Drosophila* (SHARMA *et al.* 2012), allowing females to assess male genetic quality through their CHC phenotype (CHENOWETH and BLOWS 2003). One of the CHCs impacted by this CHC polymorphism, 5,9-pentacosadiene, has been shown to influence mate choice in *D. serrata* (CHENOWETH and BLOWS 2003). The combined mating success and desiccation resistance effects observed in this study resonate with recent experimental evolution work on CHC traits that found that *D. melanogaster* males selected for increased desiccation resistance achieved significantly fewer matings than controls (BERRY and KREITMAN 1993). My analysis considered a sample of different genotypes within each phenotypic class and therefore provided a control for other genetic factors that may also explain variation in mating success and desiccation resistance. The fact that the effect of the 'low' status was detectable after accounting for this variation among different genotypes, suggests that the polymorphism may have an effect above and beyond regular genotype-to-genotype differences.

The findings of this study indicate that the major CHC polymorphism appears to have sex-specific fitness affects; enhancing survival in females and depressing mating success in males. Both sexual selection through female preference for males with particular CHC profiles and natural selection are known to drive the evolution of these traits in *D. serrata* (BLOWS 2002; CHENOWETH and BLOWS 2003; RUNDLE *et al.* 2009). The mutation at hand is seemingly sexually antagonistic, affected by natural selection in females but through opposing sexual selection in males. The interaction between natural and sexual selection may significantly affect the manner in which traits evolve in nature (BLOWS 2002; RUNDLE *et al.* 2009; HINE *et al.* 2011; ONEAL and KNOWLES 2013). These two processes have been known to work in opposition under certain conditions (RUNDLE *et al.* 2006; CANDOLIN and HEUSCHELE 2008) but usually within the same sex (DARWIN 1871; CANDOLIN and HEUSCHELE 2008). This study documents an interesting scenario where these antagonistic effects work across the sexes.

Sexually homologous traits are underlain by the same genome and are thus genetically correlated (FISHER 1930; LANDE and ARNOLD 1983). Sex-specific selection acting on this shared genetic variation may lead to intralocus sexual conflict manifested through opposite fitness effects in males and females (KWAN *et al.* 2008; BONDURIANSKY and CHENOWETH 2009; ANDERSEN *et al.* 2010). In *Drosophila melanogaster,* differences in selection between females and males acting on the same genotypes result in high fitness in females but low fitness in males and vice versa (CHIPPINDALE *et al.* 2001; PRASAD *et al.* 2007). The higher

survival of ‘low’ females in this study may counteract the negative effects of reduced mating success of ‘low’ males and may explain why the polymorphism is maintained in the Cooktown population. It is interesting that further south this polymorphism is not observed. Indeed in the Cardwell population which is some 300kms south ($18^0$15'58.53"S, $146^0$01'45.64"E) the frequency of the mutation is much lower (Chapter 2, Figure 2.4). It could be the case that the equilibrium between natural and sexual selection differs between this population and Cooktown. If the importance of desiccation resistance is lower in this population, the negative cost of reduced male mating success may be offset to a lesser degree. More studies with wild populations where this polymorphism abounds may be useful for further understanding its maintenance and geographical distribution.

CHAPTER 6: GENERAL DISCUSSION

Ever since Charles Darwin documented and highlighted the ubiquity of phenotypic differences between species and sexes (DARWIN 1859), the actual processes that lead to phenotypic evolution have been of primary interest to evolutionary biologists. Darwin recognised that heritable phenotypic variation that was linked to survival (DARWIN 1859) and/or mating success (DARWIN 1871) could explain many repeated patterns of phenotypic differentiation. Although it is now recognised that phenotypic evolution can occur not only through Darwinian selection but also stochastic processes such as genetic drift (TRAVISANO *et al.* 1995; KELLER and TAYLOR 2008; SANCHEZ-GUILLEN *et al.* 2011), a plethora of empirical studies nonetheless strongly implicate divergent selection in shaping trait divergence (SCHLUTER 2000; MERILA and CRNOKRAK 2001; RIESEBERG *et al.* 2002; LEINONEN *et al.* 2008).

Insect hydrocarbons are an ecologically important suite of traits implicated in maintaining water balance through reducing water loss and keeping water out of spiracles (BLOMQUIST *et al.* 1993; SCHAL *et al.* 1998; MONTOOTH and GIBBS 2003) as well as serving as both intra- and inter-specific mate recognition signals (GINZEL and HANKS 2003; HOWARD and BLOMQUIST 2005). Although the role of selection in the divergence of CHCs in *D. serrata* has been established (CHENOWETH and BLOWS 2008; HINE *et al.* 2011), the actual genetic polymorphisms targeted by selection have until now been unknown. Unravelling functional genetic polymorphisms for ecologically significant traits segregating in natural populations may help to identify both the actual genes for the trait as well as the evolutionary forces that maintain variation (FEDER and MITCHELL-OLDS 2003). This study focused on mapping the genetic basis of a recently discovered polymorphism in three CHCs of *Drosophila serrata* that steepens an existing latitudinal cline. The study also examined the potential fitness effects of this polymorphism through an investigation on disparities in desiccation resistance, heat shock stress and mating success between the two phenotypes.

F2 QTL mapping traced the polymorphism to two major effect QTLs on chromosome 3R. Bulk segregant analysis of whole genome sequence data from an advanced intercross population revealed a single region of genetic differentiation on the same chromosomal arm and whose functional annotation revealed three adjacent fatty acyl-CoA reductase genes.

Although nearby candidate desaturase and elongase loci were evident, they did not appear involved in the major polymorphism. A follow up analysis of fixed genetic differences between independent replicate 'low' and 'high' lines, from a single population, confirmed the results of the bulk segregant study and replicated the association. Parallel fixed single nucleotide substitutions and indels in the coding and untranslated regions of three reductase genes clearly distinguished the 'low' from the 'high' lines. One of the fatty acyl-CoA reductase genes, CG17560, was confirmed to be expressed in CHC-producing oeneocyte cells in 'high' and 'low' lines strongly implicating the involvement of amino acid changes in this gene. The polymorphism had sex-specific fitness effects; desiccation resistance was significantly higher in 'low' than 'high' females but no significant effects were detected in males. 'Low' males had dramatically reduced mating success compared with their 'high' counterparts. I discuss these findings in the light of the main trends that were detected.

6.1 *The genetic basis of D. serrata* CHCs

The short–chained CHC polymorphism mapped to two major-effect QTLs that overlapped in location with QTLs for five other CHCs. In general all eight CHCs mapped to shared QTLs. Clustering of different trait QTLs on a linkage map, commonly reported for multiple traits in both plants (BURKE *et al.* 2002; CAI and MORISHIMA 2002; JUENGER *et al.* 2005; PEREZ-VEGA *et al.* 2010) and animals (PROTAS *et al.* 2008; ALBERT *et al.* 2009; WRIGHT *et al.* 2010) may indicate a shared genetic architecture through either tight linkage or pleiotropy (HALL *et al.* 2006; EDWARDS and WEINIG 2011). Upon finer dissection of the major polymorphism, this study effectively ruled out the presence of many, linked smaller-effect QTLs within the earlier detected QTLs and posits pleiotropy as a likely feature of the genetic architecture of the CHCs investigated. QTLs pleiotropically influence CHC expression in both *D. melanogaster* (FOLEY *et al.* 2007) and *D. mojavensis* (ETGES *et al.* 2009). Further, pleiotropy may also occur for methyl-branched CHC biosynthesis in the jewel wasps, *Nasonia giraulti* and *Nasonia vitripennis* (NIEHUIS *et al.* 2011). However, these QTL mapping studies, like the F2 analysis in the present study do lack sufficient mapping resolution to distinguish linkage from pleiotropy. However, in perhaps a more compelling example of pleiotropy, an experimental insertion of a transposon into the *desat1* gene affected production of multiple different CHCs in *D. melanogaster* (MARCILLAC *et al.* 2005a). Given that different CHCs share a common biosynthetic pathway, only differing in chain length, the position and number of double bonds

and degree of methyl branching, pleiotropy was not an entirely unexpected finding in the present study.

Most of the CHCs mapped to QTLs on chromosome 3, a finding that is in accord with those of similar studies in *Drosophila* which find that CHC variation is largely driven by loci on chromosome 3 (FERVEUR *et al.* 1996; COYNE *et al.* 1999; DALLERAC *et al.* 2000; WICKER-THOMAS and JALLON 2001; GLEASON *et al.* 2009). By using several fine-mapping strategies I identified fatty-acyl-CoA reductase genes on chromosome 3R as underlying the CHC polymorphism in *D. serrata*. Two groups of reductases, an NADH-dependent fatty acyl reductase and an NADPH-dependent aldehyde reductase facilitate the production of cuticular hydrocarbons and primary alcohols from the fatty acid intermediates (ROWLAND *et al.* 2006). Finding reductase genes potentially affecting CHC production resonates with other studies in both plants (WANG and KOLATTUKUDY 1995; ROWLAND *et al.* 2006) and animals (TURGEON and BERNATCHEZ 2001; MOTO *et al.* 2003; LASSANCE *et al.* 2010; TEERAWANICHPAN *et al.* 2010). Except for a recent study that used an RNAi knock-down approach to confirm the role of decarbonylating reductase genes in CHC biosynthesis in *Drosophila* (TURGEON and BERNATCHEZ 2001), most of the studies that relate fatty-acyl reductase genes as affecting hydrocarbon biosynthesis in animals are not in *Drosophila* but in *Ostrinia nubilalis* moths (LASSANCE *et al.* 2010), silkworm *Bombyx mori* (MOTO *et al.* 2003) and honey bees, *Apis mellifera* (TEERAWANICHPAN *et al.* 2010). This is the first study that has mapped changes in CHCs in a *Drosophila* species to natural variation in biosynthetically relevant fatty-acyl-CoA reductase genes; most of the prior research in *Drosophila* has largely focused on desaturases and elongases (TAKAHASHI *et al.* 2001). This discovery brings a new dimension to the characterisation of the genetic architecture of CHCs in *Drosophila*. It will be interesting to find out how this locus impacts on CHC polymorphisms in other well studied *Drosophila* species like *D. melanogaster* through such techniques as RNAi knockouts.

The polymorphism under study here affected the production of 24 Carbon and 25 Carbon chain CHCs. Although the actual point on the biosynthetic pathway of these hydrocarbons that is the target of these mutations in *D. serrata* is not yet elucidated, it perhaps involves a defect at the chain-reduction stage when aldehyde intermediates are being decarboxylated into hydrocarbons. It might be the case that the elongation of fatty acids proceeds in manner similar to the pathway reported for *D. melanogaster* by LEGENDRE *et al.* (2008);14C and 16C Acyl-CoAs may be elongated by the addition of either one or two carbons at each elongation

step to yield a longer carbon chain backbone, perhaps a 26C chain. Removal of one or two carbons in the final decarboylation stage would result in the formation of 5,9-tetracosadiene, 5,9-pentacosadiene and 9-pentacosene. I illustrate how this might happen in figure 6.1 based on the work of Tillman *et al.* (1999).

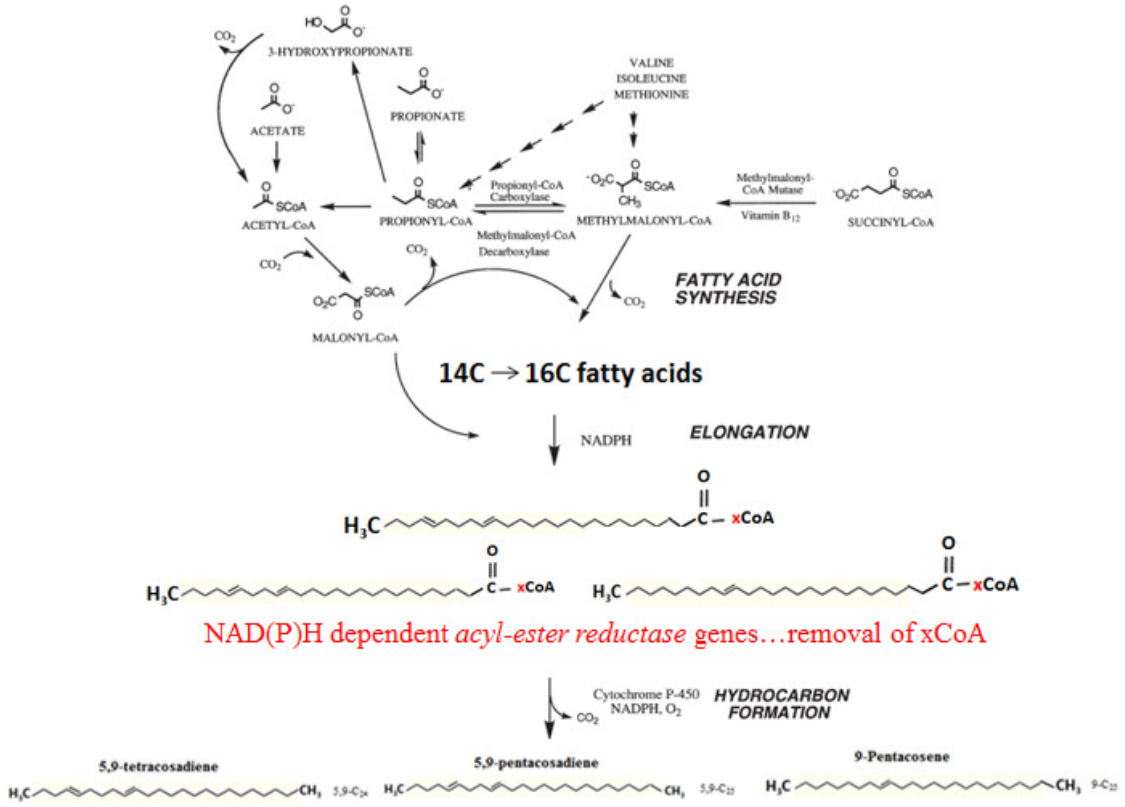


Figure 6.1

A schematic representation of a hypothetical biosynthetic pathway for the production (from known CHC substrates) of three CHCs that depict a major effect polymorphism in *D. serrata*. The x in the xCoA represents an unknown type of acetylCoA chain that is an immediate precursor of the three CHCs in this study. The figure is adapted and modified from CHC biosynthetic pathways depicted by TILLMAN *et al.* (1999).

Although many phenotypic traits have both sexual and viability-related roles (MAAN and CUMMINGS 2009; BONDURIANSKY 2011) and despite the fact that sexual selection is one of the strongest forms of directional selection in nature (KINGSOLVER *et al.* 2001), empirical data on the genetic basis of sexually selected traits are less extensive (MERILA and SHELDON 1999; CHENOWETH and MCGUIGAN 2010). This study uncovered a single small genomic region as

underlying the major polymorphism in sexually selected CHCs, suggesting a simple genetic basis. The findings run counter to the expectation of an exceptionally polygenic basis of many complex traits (McINNES and QUIGG 2010; ROCKMAN 2012). The possibility of a highly polygenic genetic architecture for phenotypic divergence of *Drosophila* CHCs has been reported (RULE and ROELOFS 1989; FOLEY *et al.* 2007; ETGES *et al.* 2009) although other studies find few QTLs of large effect for interspecific CHC variation (GLEASON *et al.* 2005). A CHC polymorphism in *D. melanogaster,* has also been mapped to a single segregating factor (COYNE *et al.* 1999).

A single major-effect locus for variation in a sex pheromone has also been reported in the moth *Heliothis subflexa* (UMINA *et al.* 2005). Similar substantial phenotypic differences in other traits driven by only a few loci of moderate to large effects have also been recorded in other taxa (VOSS and SHAFFER 1997; COLOSIMO *et al.* 2005; BAERWALD *et al.* 2011). Although this study supports the possibility that the *D. serrata* CHC polymorphism is based on a common major-effect mutation or haplotype, the fact that when I removed 'low' individuals and remapped QTLs in the F2 intercross, QTLs were found at the same locations, may indicate that the regions also harbour smaller effect loci.

6.2 *A naturally occurring polymorphism under antagonistic natural and sexual selection*

The interaction between sexual and natural selection processes may have implications on the evolution and maintenance of genetic variation in sexually selected traits. Sexual selection may sometimes be consistent with the action of natural selection in promoting survival or viability (ZAHAVI 1975; HAMILTON and ZUK 1982; KIRKPATRICK 1987), resulting in adaptive trait divergence (CANDOLIN and HEUSCHELE 2008; CHEN *et al.* 2012). Through this study and within the context of the phenotypes I investigated, however, I found that the CHC polymorphism had sexually antagonistic effects; enhancing survival in females at a cost to male mating success, suggestive of opposing natural and sexual selection processes. Sexual selection, often associated with the evolution of seemingly detrimental traits (DARWIN 1871), may counteract natural selection (RUNDLE *et al.* 2006; CANDOLIN and HEUSCHELE 2008), and cause changes in traits that bring about direct nonsexual fitness costs (LANDE 1981; SVENSSON *et al.* 2004; RUNDLE *et al.* 2006; CANDOLIN and HEUSCHELE 2008). In this study, sexual selection through female mate choice on the major polymorphism negatively impacted on male mating success.

I found that female 'low's had higher survival under desiccation stress. Natural selection then seems to favour this mutation in females while sexual selection acts against it in males. In another study of the same species studied here, male attractiveness genes for CHC blends that evolve through sexual selection via female choice are not favoured by natural selection for enhanced productivity in *D. serrata* (SKROBLIN and BLOWS 2006) . Such opposing natural and sexual selection may lead to balancing selection (CHENOWETH and MCGUIGAN 2010). Despite its significant fitness effects on either sex, sexually antagonistic selection processes may be maintaining the polymorphism at intermediate frequencies in far northern populations of the species range.

6.3 *Utility of next generation sequencing approaches for adaptive trait dissection in non-model organisms*

Understanding the genetic basis of complex traits is pivotal for dissecting the genetic architectures of adaptive trait divergence (HANSEN 2006). QTL mapping is a useful tool for initial investigating the broad scale basis of the adaptive evolution of complex traits (VOSS and SHAFFER 1997; HERDER *et al.* 2006), although concerns have been raised about potential bias due to overestimation of effect sizes (BEAVIS *et al.* 1994) and underestimation of gene number (TILLMAN-WALL *et al.* 1992). In this study, I exploited a range of supporting approaches including the creation of advanced crosses, bulk segregant analysis and whole genome resequencing to fine tune results from F2 QTL mapping. I was also able to validate the QTL by resequencing the QTL region in multiple 'low' and 'high' lines.

Through these approaches, I uncovered several fixed genetic differences within fatty Acyl-CoA reducatse genes that clearly delineated these two phenotypic groups of lines. The success of this study illustrates the power of recent whole genome re-sequencing protocols that offer higher marker coverage to address shortfalls of traditional QTL mapping. Analysis of genome-wide DNA variants correlated with phenotypic differences in complex traits potentially fast-tracks the identification of alleles and common polymorphisms that are driven to a high frequency by natural selection and thus underlying variation in adaptive traits (GOCKEL *et al.* 2001; CAO *et al.* 2011; MAGWIRE *et al.* 2012). Comparative genome-wide data analysis of multiple populations may uncover allele frequency variation in genes that drive phenotypic changes (QIAN *et al.* 2013) as well as pinpoint genomic hotspots of mutational

changes targeted by selection (LITI *et al.* 2009; HUFFORD *et al.* 2012; MICHAELSON *et al.* 2012; SUPPLY *et al.* 2013)**.** In *Drosophila melanogaster*, this approach has recently helped to dissect the genetic basis of susceptibility to viral infection (MAGWIRE *et al.* 2012). With increasing affordability of DNA sequencing, the new genomic approaches are likely to dominate genetic characterisation of adaptive trait transitions in non-model species.

6.4 *Directions for future research*

Although this study is the first case of a fatty acyl-CoA gene being implicated in a *D. serrata* CHC polymorphism, the actual causal mutations remain to be elucidated. In some sense, my study is just the beginning. It is just becoming apparent that the CG17560 may indeed be the causal locus as it is the only one of the three reductases to be expressed in oeneocytes. Although it is more likely that protein coding sequence changes have led to the polymorphism, expression level changes should be investigated using quantitative rtPCR experiments. The question of whether the multiple SNPs identified act independently as single mutations or as haploytypic clusters of SNPs needs to be investigated but will require access to either novel recombinants or transgenic tools. Site-directed mutagenesis, especially using the recently developed CRISPR/Cas9 system (clustered regularly interspaced palindromic repeats/CRISPR associated) technique (BASSETT *et al.* 2013) on *D. serrata* embryos to induce mutations that are transmissible through the germline has the potential to create stable lines that may be highly efficient for refining the actual causal variants underlying this CHC polymorphism. Further molecular population genetic analysis of the QTL region identified in this study will likely uncover patterns of molecular evolution underlying this genomic regions and reveal its evolutionary history. To better understand the fitness effects of the polymorphism experimental evolution studies that manipulate the presence of natural and sexual selection may be a useful tool.

7. **LIST OF REFERENCES**

AGRAWAL, A. F., 2001 Sexual selection and the maintenance of sexual reproduction. Nature **411:** 692-695.

AHUJA, A., and R. S. SINGH, 2008 Variation and evolution of male sex combs in *Drosophila*: Nature of selection response and theories of genetic variation for sexual traits. Genetics **179:** 503-509.

ALBERT, F. W., O. CARLBORG, I. PLYUSNINA, F. BESNIER, D. HEDWIG *et al.*, 2009 Genetic Architecture of Tameness in a Rat Model of Animal Domestication. Genetics **182:** 541-554.

ALTSCHUL, S. F., T. L. MADDEN, A. A. SCHAFFER, J. ZHANG, Z. ZHANG *et al.*, 1997 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25:** 3389-3402.

ANDERSEN, L. H., T. N. KRISTENSEN, V. LOESCHCKE, S. TOFT and D. MAYNTZ, 2010 Protein and carbohydrate composition of larval food affects tolerance to thermal stress and desiccation in adult *Drosophila melanogaster*. Journal of Insect Physiology **56:** 336-340.

ANDERSON, A. R., J. E. COLLINGE, A. A. HOFFMANN, M. KELLETT and S. W. MCKECHNIE, 2003 Thermal tolerance trade-offs associated with the right arm of chromosome 3 and marked by the hsr-omega gene in *Drosophila melanogaster*. Heredity **90:** 195-202.

ANDERSSON, M. B., 1994 *Sexual Selection*. Princeton University Press.

AYALA, D., R. F. GUERRERO and M. KIRKPATRICK, 2013 Reproductive isolation and local adaptation quantified for a chromosome inversion in a malaria mosquito. Evolution **67:** 946-958.

BAERWALD, M. R., J. L. PETERSEN, R. P. HEDRICK, G. J. SCHISLER and B. MAY, 2011 A major effect quantitative trait locus for whirling disease resistance identified in rainbow trout (*Oncorhynchus mykiss*). Heredity **106:** 920-926.

BARNWELL, C. V., and M. A. NOOR, 2008a Failure to replicate two mate preference QTLs across multiple strains of *Drosophila pseudoobscura*. J Hered **99:** 653-656.

BARNWELL, C. V., and M. A. F. NOOR, 2008b Failure to Replicate Two Mate Preference QTLs across Multiple Strains of *Drosophila pseudoobscura*. Journal of Heredity **99:** 653-656.

BARRETT, R. D. H., and D. SCHLUTER, 2008 Adaptation from standing genetic variation. Trends in Ecology & Evolution **23:** 38-44.

BASSETT, A. R., C. TIBBIT, C. P. PONTING and J.-L. LIU, 2013 Highly Efficient Targeted Mutagenesis of *Drosophila* with the CRISPR/Cas9 System. Cell Reports **4:** 220-228.

BEAVIS, W. D., O. S. SMITH, D. GRANT and R. FINCHER, 1994 Identification of Quantitative Trait Loci Using a Small Sample of Topcrossed and F4 Progeny from Maize. Crop Science **34:** 882-896.

BERMOND, G., M. CIOSI, E. LOMBAERT, A. BLIN, M. BORIANI *et al.*, 2012 Secondary Contact and Admixture between Independently Invading Populations of the Western Corn Rootworm, *Diabrotica virgifera virgifera* in Europe. Plos One **7**.

BERRY, A., and M. KREITMAN, 1993 Molecular analysis of an allozyme cline - Alcohol-Dehydrogenase in *Drosophila melanogaster* on the East-coast of North-America. Genetics **134:** 869-893.

BERTIN, M. J., P. V. ZIMBA, K. R. BEAUCHESNE, K. M. HUNCIK and P. D. R. MOELLER, 2012 Identification of toxic fatty acid amides isolated from the harmful alga *Prymnesium parvum* carter. Harmful Algae **20:** 111-116.

BESNIER, F., P. WAHLBERG, L. RONNEGARD, W. EK, L. ANDERSSON *et al.*, 2011 Fine mapping and replication of QTL in outbred chicken advanced intercross lines. Genetics Selection Evolution **43**.

BHUTKAR, A., S. M. RUSSO, T. F. SMITH and W. M. GELBART, 2007 Genome-scale analysis of positionally relocated genes. Genome Research **17:** 1880-1887.

BICKEL, R. D., A. KOPP and S. V. NUZHDIN, 2011 Composite Effects of Polymorphisms near Multiple Regulatory Elements Create a Major-Effect QTL. Plos Genetics **7**.

BLAILOCK, T. T., G. J. BLOMQUIST and L. L. JACKSON, 1976 Biosynthesis of 2-Methylalkanes in Crickets - *Nemobius fasciatus* and *Gryllus pennsylvanicus*. Biochemical and Biophysical Research Communications **68:** 841-849.

BLOMQUIST, G. J., A. J. CHU, J. H. NELSON and J. G. POMONIS, 1980 Incorporation of [2,3-13C]succinate into methyl-branched alkanes in a termite. Archives of Biochemistry and Biophysics **204:** 648-650.

BLOMQUIST, G. J., J. A. TILLMANWALL, L. GUO, D. R. QUILICI, P. D. GU *et al.*, 1993 Hydrocarbon and hydrocarbon derived sex-pheromones in insects - biochemistry and endocrine regulation.

BLOMQUIST, G. J., A. H. VAZ, R. A. JURENKA and R. C. REITZ, 1989 Elongation Reactions Involved in Hydrocarbon Biosynthesis in Insects. Acs Symposium Series **389:** 314-322.

BLOWS, M. W., 1993 The Genetics of Central and Marginal Populations of *Drosophila serrata* .2. Hybrid Breakdown in Fitness Components as a Correlated Response to Selection for Desiccation Resistance. Evolution **47:** 1271-1285.

BLOWS, M. W., 2002 Interaction between natural and sexual selection during the evolution of mate recognition. Proceedings of the Royal Society B-Biological Sciences **269:** 1113-1118.

BLOWS, M. W., and R. A. ALLAN, 1998 Levels of mate recognition within and between two *Drosophila* species and their hybrids. American Naturalist **152:** 826-837.

BLOWS, M. W., and M. HIGGIE, 2002 Evolutionary experiments on mate recognition in the *Drosophila serrata* species complex. Genetica **116:** 239-250.

BONDURIANSKY, R., 2011 Sexual Selection and Conflict as Engines of Ecological Diversification. American Naturalist **178:** 729-745.

BONDURIANSKY, R., and S. F. CHENOWETH, 2009 Intralocus sexual conflict. Trends Ecol Evol **24:** 280-288.

BONGAERTS, P., C. RIGINOS, K. B. HAY, M. J. H. VAN OPPEN, O. HOEGH-GULDBERG *et al.*, 2011 Adaptive divergence in a scleractinian coral: physiological adaptation of *Seriatopora hystrix* to shallow and deep reef habitats. Bmc Evolutionary Biology **11**.

BRADSHAW, H. D., S. M. WILBERT, K. G. OTTO and D. W. SCHEMSKE, 1995 Genetic-Mapping of Floral Traits Associated with Reproductive Isolation in Monkeyflowers (*Mimulus*). Nature **376:** 762-765.

BROMAN, K. W., and S. SEN, 2009 Guide to QTL Mapping with R/qtl. Guide to QTL Mapping with R/qtl.

BROOKS, R., 2002 Variation in female mate choice within guppy populations: population divergence, multiple ornaments and the maintenance of polymorphism. Genetica **116:** 343-358.

BUCKLER, E. S., J. B. HOLLAND, P. J. BRADBURY, C. B. ACHARYA, P. J. BROWN *et al.*, 2009 The Genetic Architecture of Maize Flowering Time. Science **325:** 714-718.

BURKE, J. M., S. TANG, S. J. KNAPP and L. H. RIESEBERG, 2002 Genetic analysis of sunflower domestication. Genetics **161:** 1257-1267.

CAI, W., and H. MORISHIMA, 2002 QTL clusters reflect character associations in wild and cultivated rice. Theor Appl Genet **104:** 1217-1228.

CANDOLIN, U., and J. HEUSCHELE, 2008 Is sexual selection beneficial during adaptation to environmental change? Trends in Ecology & Evolution **23:** 446-452.

CAO, J., K. SCHNEEBERGER, S. OSSOWSKI, T. GUNTHER, S. BENDER *et al.*, 2011 Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. Nat Genet **43:** 956-963.

CHARLTON, R. E., and W. L. ROELOFS, 1991 Biosynthesis of a Volatile, Methyl-Branched Hydrocarbon Sex-Pheromone from Leucine by Arctiid Moths (Holomelina Spp). Archives of Insect Biochemistry and Physiology **18:** 81-97.

CHASE, J., R. A. JURENKA, C. SCHAL, P. P. HALARNKAR and G. J. BLOMQUIST, 1990 Biosynthesis of Methyl Branched Hydrocarbons of the German Cockroach *Blattella germanica* (L) (Orthoptera, Blattellidae). Insect Biochemistry **20:** 149-156.

CHEESBROUGH, T. M., and P. E. KOLATTUKUDY, 1984 Alkane biosynthesis by decarbonylation of aldehydes catalyzed by a particulate preparation from *Pisum sativum*. Proc Natl Acad Sci U S A **81:** 6613-6617.

CHEN, I. P., D. STUART-FOX, A. F. HUGALL and M. R. E. SYMONDS, 2012 Sexual selection and the evolution of complex color patterns in dragon lizards. Evolution; international journal of organic evolution **66:** 3605-3614.

CHEN, X., P. E. HEDLEY, J. MORRIS, H. LIU, R. E. NIKS *et al.*, 2011 Combining genetical genomics and bulked segregant analysis-based differential expression: an approach to gene localization. Theor Appl Genet **122:** 1375-1383.

CHENOWETH, S. F., and M. W. BLOWS, 2003 Signal trait sexual dimorphism and mutual sexual selection in *Drosophila serrata*. Evolution **57:** 2326-2334.

CHENOWETH, S. F., and M. W. BLOWS, 2005 Contrasting mutual sexual selection on homologous signal traits in *Drosophila serrata*. American Naturalist **165:** 281-289.

CHENOWETH, S. F., and M. W. BLOWS, 2008 Q(st) meets the G matrix: The dimensionality of adaptive divergence in multiple correlated quantitative traits. Evolution **62:** 1437-1449.

CHENOWETH, S. F., and K. MCGUIGAN, 2010 The Genetic Basis of Sexually Selected Variation. Annu. Rev. Ecol. Evol. Syst. **41:** 81-101.

CHENOWETH, S. F., H. D. RUNDLE and M. W. BLOWS, 2008 Genetic constraints and the evolution of display trait sexual dimorphism by natural and sexual selection. Am Nat **171:** 22-34.

CHERTEMPS, T., L. DUPORTETS, C. LABEUR, R. UEDA, K. TAKAHASHI *et al.*, 2007a A female-biased expressed elongase involved in long-chain hydrocarbon biosynthesis and courtship behavior in *Drosophila melanogaster*. Proc Natl Acad Sci U S A **104:** 4273-4278.

CHERTEMPS, T., L. DUPORTETS, C. LABEUR, R. UEDA, K. TAKAHASHI *et al.*, 2007b A female-biased expressed elongase involved in long-chain hydrocarbon biosynthesis and courtship behavior in *Drosophila melanogaster*. Proceedings of the National Academy of Sciences of the United States of America **104:** 4273-4278.

CHERTEMPS, T., L. DUPORTETS, C. LABEUR and C. WICKER-THOMAS, 2005 A new elongase selectively expressed in *Drosophila* male reproductive system. Biochem Biophys Res Commun **333:** 1066-1072.

CHEVERUD, J. M., T. H. EHRICH, T. T. VAUGHN, S. F. KOREISHI, R. B. LINSEY *et al.*, 2004 Pleiotropic effects on mandibular morphology II: Differential epistasis and genetic variation in morphological integration. Journal of Experimental Zoology Part B-Molecular and Developmental Evolution **302B:** 424-435.

CHIPPINDALE, A. K., J. R. GIBSON and W. R. RICE, 2001 Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. Proceedings of the National Academy of Sciences of the United States of America **98:** 1671-1675.

CHRISTIANS, J. K., and L. K. SENGER, 2007 Fine mapping dissects pleiotropic growth quantitative trait locus into linked loci. Mammalian Genome **18:** 240-245.

CHU, A. J., and G. J. BLOMQUIST, 1980 Biosynthesis of Hydrocarbons in Insects - Succinate Is a Precursor of the Methyl Branched Alkanes. Archives of Biochemistry and Biophysics **201:** 304-312.

CHUNCO, A. J., J. S. MCKINNON and M. R. SERVEDIO, 2007 Microhabitat variation and sexual selection can maintain male color polymorphisms. Evolution **61:** 2504-2515.

CIVETTA, A., and E. J. F. CANTOR, 2003 The genetics of mating recognition between *Drosophila simulans* and *D. sechellia*. Genetical Research **82:** 117-126.

CLARK, M. S., and M. R. WORLAND, 2008 How insects survive the cold: molecular mechanisms-a review. Journal of Comparative Physiology B-Biochemical Systemic and Environmental Physiology **178:** 917-933.

CLEGG, S. M., S. M. DEGNAN, J. KIKKAWA, C. MORITZ, A. ESTOUP *et al.*, 2002 Genetic consequences of sequential founder events by an island-colonizing bird. Proceedings of the National Academy of Sciences of the United States of America **99:** 8127-8132.

CLEVELAND, W. S., 1981 Lowess - a Program for Smoothing Scatterplots by Robust Locally Weighted Regression. American Statistician **35:** 54-54.

CLUTTON-BROCK, T., 2010 We do not need a Sexual Selection 2.0-nor a theory of Genial Selection. Animal Behaviour **79:** E7-E10.

CLYNE, J. D., and G. MIESENBOCK, 2008 Sex-specific control and tuning of the pattern generator for courtship song in *Drosophila*. Cell **133:** 354-363.

COLLAR, D. C., J. A. SCHULTE, 2ND and J. B. LOSOS, 2011 Evolution of extreme body size disparity in monitor lizards (Varanus). Evolution **65:** 2664-2680.

COLOSIMO, P. F., K. E. HOSEMANN, S. BALABHADRA, G. VILLARREAL, M. DICKSON *et al.*, 2005 Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. Science **307:** 1928-1933.

CONNALLON, T., and A. G. CLARK, 2013 Antagonistic versus nonantagonistic models of balancing selection: characterizing the relative timescales and hitchhiking effects of partial selective sweeps. Evolution **67:** 908-917.

COYNE, J. A., 1996 Genetics of differences in pheromonal hydrocarbons between *Drosophila melanogaster* and *D. simulans*. Genetics **143:** 353-364.

COYNE, J. A., and S. ELWYN, 2006 Desaturase-2, environmental adaptation, and sexual isolation in *Drosophila melanogaster*. Evolution **60:** 626-627.

COYNE, J. A., C. WICKER-THOMAS and J. M. JALLON, 1999 A gene responsible for a cuticular hydrocarbon polymorphism in *Drosophila melanogaster*. Genetical Research **73:** 189-203.

DALLERAC, R., C. LABEUR, J. M. JALLON, D. C. KNIPPIE, W. L. ROELOFS *et al.*, 2000 A Delta 9 desaturase gene with a different substrate specificity is responsible for the cuticular diene hydrocarbon polymorphism in *Drosophila melanogaster*. Proceedings of the National Academy of Sciences of the United States of America **97:** 9449-9454.

DARVASI, A., and M. SOLLER, 1995 Advanced intercross lines, an experimental population for fine genetic mapping. Genetics **141:** 1199-1207.

DARWIN, C., 1859 *On the Origin of Species by Means of Natural Selection*. Murray, London.

DARWIN, C., 1871 *The Descent of Man and Selection in Relation to Sex*. Murray London.

DE JONG, G., and Z. BOCHDANOVITS, 2003 Latitudinal clines in *Drosophila melanogaster*: body size, allozyme frequencies, inversion frequencies, and the insulin-signalling pathway. Journal of Genetics **82:** 207-223.

DILLON, M. E., G. WANG, P. A. GARRITY and R. B. HUEY, 2009 Thermal preference in *Drosophila*. Journal of Thermal Biology **34:** 109-119.

DILLWITH, J. W., J. H. NELSON, J. G. POMONIS, D. R. NELSON and G. J. BLOMQUIST, 1982 A C-13 Nmr-Study of Methyl-Branched Hydrocarbon Biosynthesis in the Housefly. Journal of Biological Chemistry **257:** 1305-1314.

DOBZHANSKY, T., and W. B. MATHER, 1961 Evolutionary Status of *Drosophila Serrata*. Evolution **15:** 461-&.

DOERGE, R. W., and G. A. CHURCHILL, 1996 Permutation tests for multiple loci affecting a quantitative character. Genetics **142:** 285-294.

EDWARDS, C. E., and C. WEINIG, 2011 The quantitative-genetic and QTL architecture of trait integration and modularity in Brassica rapa across simulated seasonal settings. Heredity **106:** 661-677.

EMLEN, D. J., 2008 The Evolution of Animal Weapons, pp. 387-413 in *Annual Review of Ecology Evolution and Systematics*. Annual Reviews, Palo Alto.

ENDLER, J. A., 1973 Gene flow and population differentiation. Science **179:** 243-250.

ENDLER, J. A., 1977 Speciation, geographical variation and clines. Princeton University Press, Princeton, NJ.

ENDLER, J. A., 1983 Natural and Sexual Selection on Color Patterns in Poeciliid Fishes. Environmental Biology of Fishes **9:** 173-190.

ENDLER, J. A., 1986 Natural selection in the wild. Monographs in Population Biology**:** viii-xiii, 1-336.

ETGES, W. J., C. C. DE OLIVEIRA, M. G. RITCHIE and M. A. NOOR, 2009 Genetics of incipient speciation in *Drosophila mojavensis*: II. Host plants and mating status influence cuticular hydrocarbon QTL expression and G x E interactions. Evolution **63:** 1712-1730.

FALCONER, D. S., and T. F. C. MACKAY, 1996 *Introduction to Quantitative Genetics*. Longman London.

FEDER, M. E., and T. MITCHELL-OLDS, 2003 Evolutionary and ecological functional genomics. Nat Rev Genet **4:** 651-657.

FELDMAN, C. R., E. D. BRODIE, JR., E. D. BRODIE, III and M. E. PFRENDER, 2009 The evolutionary origins of beneficial alleles during the repeated adaptation of garter snakes to deadly prey. Proceedings of the National Academy of Sciences of the United States of America **106:** 13415-13420.

FERVEUR, J. F., 1991 Genetic-Control of Pheromones in *Drosophila simulans* .1. Ngbo, a Locus on the 2nd Chromosome. Genetics **128:** 293-301.

FERVEUR, J. F., 1997 The pheromonal role of cuticular hydrocarbons in Drosophila melanogaster. Bioessays **19:** 353-358.

FERVEUR, J. F., M. COBB, H. BOUKELLA and J. M. JALLON, 1996 World-wide variation in Drosophila melanogaster sex pheromone: Behavioural effects, genetic bases and potential evolutionary consequences. Genetica **97:** 73-80.

FISHER, R. A., 1930 *The Genetical Theory of Natural Selection* Oxford Univ. Press Oxford

FISHER, R. A., 1937 The wave of advance of advantageous genes. Annals of Eugenics **7:** 355-369.

FITZPATRICK, J. L., M. ALMBRO, A. GONZALEZ-VOYER, N. KOLM and L. W. SIMMONS, 2012 Male contest competition and the coevolution of weaponry and testes in pinnipeds. Evolution **66:** 3595-3604.

FLINT, J., 2011 Mapping quantitative traits and strategies to find quantitative trait genes. Methods **53:** 163-174.

FLINT, J., and T. F. C. MACKAY, 2009 Genetic architecture of quantitative traits in mice, flies, and humans. Genome Research **19:** 723-733.

FOERSTER, K., T. COULSON, B. C. SHELDON, J. M. PEMBERTON, T. H. CLUTTON-BROCK *et al.*, 2007 Sexually antagonistic genetic variation for fitness in red deer. Nature **447:** 1107-U1109.

FOLEY, B., S. F. CHENOWETH, S. V. NUZHDIN and M. W. BLOWS, 2007 Natural genetic variation in cuticular hydrocarbon expression in male and female *Drosophila melanogaster*. Genetics **175:** 1465-1477.

FRENTIU, F. D., M. ADAMSKI, E. A. MCGRAW, M. W. BLOWS and S. F. CHENOWETH, 2009 An expressed sequence tag (EST) library for *Drosophila serrata*, a model system for sexual selection and climatic adaptation studies. Bmc Genomics **10**.

FRENTIU, F. D., and S. F. CHENOWETH, 2010 Clines in cuticular hydrocarbons in two *Drosophila* species with independent population histories. Evolution **64:** 1784-1794.

FRICKE, C., and G. ARNQVIST, 2007 Rapid adaptation to a novel host in a seed beetle (*Callosobruchus maculatus*): The role of sexual selection. Evolution **61:** 440-454.

GILCHRIST, G. W., R. B. HUEY and L. SERRA, 2001 Rapid evolution of wing size clines in *Drosophila subobscura*. Genetica **112:** 273-286.

GILLESPIE, J. H., 1984 Pleiotropic overdominance and the maintenance of genetic variation in polygenic characters. Genetics **107:** 321-330.

GINZEL, M. D., G. J. BLOMQUIST, J. G. MILLAR and L. M. HANKS, 2003 Role of contact pheromones in mate recognition in *Xylotrechus colonus*. Journal of Chemical Ecology **29:** 533-545.

GINZEL, M. D., and L. M. HANKS, 2003 Contact pheromones as mate recognition cues of four species of longhorned beetles (Coleoptera : Cerambycidae). Journal of Insect Behavior **16:** 181-187.

GLEASON, J. M., J. M. JALLON, J. D. ROUAULT and M. G. RITCHIE, 2005 Quantitative trait loci for cuticular hydrocarbons associated with sexual isolation between *Drosophila simulans* and *D. sechellia*. Genetics **171:** 1789-1798.

GLEASON, J. M., R. A. JAMES, C. WICKER-THOMAS and M. G. RITCHIE, 2009 Identification of quantitative trait loci function through analysis of multiple cuticular hydrocarbons differing between *Drosophila simulans* and *Drosophila sechellia* females. Heredity **103:** 416-424.

GOCKEL, J., W. J. KENNINGTON, A. HOFFMANN, D. B. GOLDSTEIN and L. PARTRIDGE, 2001 Nonclinality of molecular variation implicates selection in maintaining a morphological cline of *Drosophila melanogaster*. Genetics **158:** 319-323.

GORING, H. H. H., and J. D. TERWILLIGER, 2000 Linkage analysis in the presence of errors III: Marker loci and their map as nuisance parameters. American Journal of Human Genetics **66:** 1298-1309.

GRANT, P. R., and B. R. GRANT, 2002 Unpredictable evolution in a 30-year study of Darwin's finches. Science **296:** 707-711.

GREENBERG, A. J., J. R. MORAN, J. A. COYNE and C. I. WU, 2003 Ecological adaptation during incipient speciation revealed by precise gene replacement. Science **302:** 1754-1757.

GU, X., D. QUILICI, P. JUAREZ, G. J. BLOMQUIST and C. SCHAL, 1995 Biosynthesis of Hydrocarbons and Contact Sex-Pheromone and Their Transport by Lipophorin in Females of the German-Cockroach (*Blattella germanica*). Journal of Insect Physiology **41:** 257-267.

GUO, L., D. R. QUILICI, J. CHASE and G. J. BLOMQUIST, 1991 Gut Tract Microorganisms Supply the Precursors for Methyl-Branched Hydrocarbon Biosynthesis in the Termite, *Zootermopsis nevadensis*. Insect Biochemistry **21:** 327-333.

HALARNKAR, P. P., and G. J. BLOMQUIST, 1989 Comparative aspects of propionate metabolism. Comp Biochem Physiol B **92:** 227-231.

HALARNKAR, P. P., J. D. CHAMBERS and G. J. BLOMQUIST, 1986 Metabolism of Propionate to Acetate in 9 Insect Species. Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology **84:** 469-472.

HALDANE, J. B., 1948 The theory of a cline. J Genet **48:** 277-284.

HALEY, C. S., and S. A. KNOTT, 1992 A Simple Regression Method for Mapping Quantitative Trait Loci in Line Crosses Using Flanking Markers. Heredity **69:** 315-324.

HALL, D., C. TEGSTROM and P. K. INGVARSSON, 2010 Using association mapping to dissect the genetic basis of complex traits in plants. Briefings in Functional Genomics **9:** 157-165.

HALL, M. C., C. J. BASTEN and J. H. WILLIS, 2006 Pleiotropic quantitative trait loci contribute to population divergence in traits associated with life-history variation in *Mimulus guttatus*. Genetics **172:** 1829-1844.

HALLAS, R., M. SCHIFFER and A. A. HOFFMANN, 2002 Clinal variation in *Drosophila serrata* for stress resistance and body size. Genetical Research **79:** 141-148.

HAMILTON, W. D., and M. ZUK, 1982 Heritable true fitness and bright birds: a role for parasites? Science **218:** 384-387.

HANSEN, T. F., 2006 The evolution of genetic architecture, pp. 123-157 in *Annual Review of Ecology Evolution and Systematics*.

HEDRICK, P. W., 2006 Genetic polymorphism in heterogeneous environments: The age of genomics, pp. 67-93 in *Annual Review of Ecology Evolution and Systematics*.

HERDER, F., A. W. NOLTE, J. PFAENDER, J. SCHWARZER, R. K. HADIATY *et al.*, 2006 Adaptive radiation and hybridization in Wallace's Dreamponds: evidence from sailfin silversides in the Malili Lakes of Sulawesi. Proceedings of the Royal Society B-Biological Sciences **273:** 2209-2217.

HIGGIE, M., and M. W. BLOWS, 2007 Are traits that experience reinforcement also under sexual selection? American Naturalist **170:** 409-420.

HIGGIE, M., and M. W. BLOWS, 2008 The evolution of reproductive character displacement conflicts with how sexual selection operates within a species. Evolution **62:** 1192-1203.

HIGGIE, M., S. CHENOWETH and M. W. BLOWS, 2000 Natural selection and the reinforcement of mate recognition. Science **290:** 519-521.

HILL, W. G., and A. ROBERTSON, 1966 The effect of linkage on limits to artificial selection. Genet Res **8:** 269-294.

HINE, E., K. MCGUIGAN and M. W. BLOWS, 2011 Natural selection stops the evolution of male attractiveness. Proceedings of the National Academy of Sciences of the United States of America **108:** 3659-3664.

HOBBS, J. P. A., and H. MACRAE, 2012 Unusual weather and trapped coral spawn lead to fish kill at a remote coral atoll. Coral Reefs **31:** 961-961.

HOEKSTRA, H. E., R. J. HIRSCHMANN, R. A. BUNDEY, P. A. INSEL and J. P. CROSSLAND, 2006 A single amino acid mutation contributes to adaptive beach mouse color pattern. Science **313:** 101-104.

HOEKSTRA, H. E., J. M. HOEKSTRA, D. BERRIGAN, S. N. VIGNIERI, A. HOANG *et al.*, 2001 Strength and tempo of directional selection in the wild. Proceedings of the National Academy of Sciences of the United States of America **98:** 9157-9160.

HOFFMANN, A. A., A. ANDERSON and R. HALLAS, 2002 Opposing clines for high and low temperature resistance in *Drosophila melanogaster*. Ecology Letters **5:** 614-618.

HOFFMANN, A. A., C. M. SGRO and A. R. WEEKS, 2004 Chromosomal inversion polymorphisms and adaptation. Trends in Ecology & Evolution **19:** 482-488.

HOFFMANN, A. A., and A. R. WEEKS, 2007 Climatic selection on genes and traits after a 100 year-old invasion: a critical look at the temperate-tropical clines in *Drosophila melanogaster* from eastern Australia. Genetica **129:** 133-147.

HOLLOCHER, H., J. L. HATCHER and E. G. DYRESON, 2000 Evolution of abdominal pigmentation differences across species in the *Drosophila dunni* subgroup. Evolution **54:** 2046-2056.

HOOPER, J. E., M. PEREZ-ALONSO, J. R. BERMINGHAM, M. PROUT, B. A. ROCKLEIN *et al.*, 1992 Comparative studies of *Drosophila* Antennapedia genes. Genetics **132:** 453-469.

HOPKINS, R., J. SCHMITT and J. R. STINCHCOMBE, 2008 A latitudinal cline and response to vernalization in leaf angle and morphology in *Arabidopsis thaliana* (Brassicaceae). New Phytologist **179:** 155-164.

HOWARD, R. W., 1998 Ontogenetic, reproductive, and nutritional effects on the cuticular hydrocarbons of the host-specific ectoparasitoid Cephalonomia tarsalis (Hymenoptera : Bethylidae). Annals of the Entomological Society of America **91:** 101-112.

HOWARD, R. W., and G. J. BLOMQUIST, 2005 Ecological, behavioral, andbiochemical aspects of insect hydrocarbons. Annual Review of Entomology **50:** 371-393.

HOWARD, R. W., L. L. JACKSON, H. BANSE and M. W. BLOWS, 2003 Cuticular hydrocarbons of *Drosophila birchii* and *D. serrata*: Identification and role in mate choice in D-serrata. Journal of Chemical Ecology **29:** 961-976.

HUANG, D. W., B. T. SHERMAN and R. A. LEMPICKI, 2009 Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature Protocols **4:** 44-57.

HUANG, Y.-F., D. MADUR, V. COMBES, C. L. KY, D. COUBRICHE *et al.*, 2010 The Genetic Architecture of Grain Yield and Related Traits in *Zea* maize L. Revealed by Comparing Intermated and Conventional Populations. Genetics **186:** 395-U612.

HUEY, R. B., G. W. GILCHRIST, M. L. CARLSON, D. BERRIGAN and L. SERRA, 2000 Rapid evolution of a geographic cline in size in an introduced fly. Science **287:** 308-309.

HUFFORD, M. B., X. XU, J. VAN HEERWAARDEN, T. PYHAJARVI, J. M. CHIA *et al.*, 2012 Comparative population genomics of maize domestication and improvement. Nature Genetics **44:** 808-U118.

HUNT, J., R. R. SNOOK, C. MITCHELL, H. S. CRUDGINGTON and A. J. MOORE, 2012 Sexual selection and experimental evolution of chemical signals in *Drosophila pseudoobscura*. Journal of Evolutionary Biology **25:** 2232-2241.

HUXLEY, J. S., 1938 Clines: an auxiliary taxonomic principle. Nature Reviews Genetics **42** 219.

IMMLER, S., G. ARNQVIST and S. P. OTTO, 2012 Ploidally antagonistic selection maintains stable genetic polymorphism. Evolution **66:** 55-65.

IRAQI, F., S. J. CLAPCOTT, P. KUMARI, C. S. HALEY, S. J. KEMP *et al.*, 2000 Fine mapping of trypanosomiasis resistance loci in murine advanced intercross lines. Mammalian Genome **11:** 645-648.

IRVING, H., J. M. RIVERON, S. S. IBRAHIM, N. F. LOBO and C. S. WONDJI, 2012 Positional cloning of rp2 QTL associates the P450 genes CYP6Z1, CYP6Z3 and CYP6M7 with pyrethroid resistance in the malaria vector *Anopheles funestus*. Heredity **109:** 383-392.

JALLON, J. M., 1984 A few chemical words exchanged by *Drosophila* during courtship and mating. Behav Genet **14:** 441-478.

JALLON, J. M., and M. COBB, 1987 Variations in Potentially Aphrodisiac Cuticular Hydrocarbons among the 8 Species of the *Drosophila melanogaster* Subgroup. Journal of Neurogenetics **4:** 140-141.

JENKINS, N. L., and A. A. HOFFMANN, 2000 Variation in morphological traits and trait asymmetry in field *Drosophila serrata* from marginal populations. Journal of Evolutionary Biology **13:** 113-130.

JOHNSTON, S. E., J. GRATTEN, C. BERENOS, J. G. PILKINGTON, T. H. CLUTTON-BROCK *et al.*, 2013 Life history trade-offs at a single locus maintain sexually selected genetic variation. Nature **502:** 93-95.

JUENGER, T., J. M. PEREZ-PEREZ, S. BERNAL and J. L. MICOL, 2005 Quantitative trait loci mapping of floral and leaf morphology traits in *Arabidopsis thaliana*: evidence for modular genetic architecture. Evolution & Development **7:** 259-271.

JURENKA, R. A., M. DERENOBALES and G. J. BLOMQUIST, 1987 Denovo Biosynthesis of Polyunsaturated Fatty-Acids in the Cockroach *Periplaneta americana*. Archives of Biochemistry and Biophysics **255:** 184-193.

KAWAMURA, K., L. H.-S. OYANT, L. CRESPEL, T. THOUROUDE, D. LALANNE *et al.*, 2011 Quantitative trait loci for flowering time and inflorescence architecture in rose. Theoretical and Applied Genetics **122:** 661-675.

KELLER, S. R., D. R. SOWELL, M. NEIMAN, L. M. WOLFE and D. R. TAYLOR, 2009 Adaptation and colonization history affect the evolution of clines in two introduced species. New Phytol **183:** 678-690.

KELLER, S. R., and D. R. TAYLOR, 2008 History, chance and adaptation during biological invasion: separating stochastic phenotypic evolution from response to selection. Ecology Letters **11:** 852-866.

KELLERMANN, V., B. VAN HEERWAARDEN, C. M. SGRO and A. A. HOFFMANN, 2009 Fundamental Evolutionary Limits in Ecological Traits Drive *Drosophila* Species Distributions. Science **325:** 1244-1246.

KENNEY-HUNT, J. P., B. WANG, E. A. NORGARD, G. FAWCETT, D. FALK *et al.*, 2008 Pleiotropic patterns of quantitative trait loci for 70 murine skeletal traits. Genetics **178:** 2275-2288.

KIMURA, K., T. HACHIYA, M. KOGANEZAWA, T. TAZAWA and D. YAMAMOTO, 2008 Fruitless and Doublesex Coordinate to Generate Male-Specific Neurons that Can Initiate Courtship (vol 59, pg 759, 2008). Neuron **60:** 390-391.

KINGSOLVER, J. G., H. E. HOEKSTRA, J. M. HOEKSTRA, D. BERRIGAN, S. N. VIGNIERI *et al.*, 2001 The strength of phenotypic selection in natural populations. American Naturalist **157:** 245-261.

KIRKPATRICK, M., 1982 Sexual Selection and the Evolution of Female Choice. Evolution **36:** 1-12.

KIRKPATRICK, M., 1987 Sexual Selection by Female Choice in Polygynous Animals. Annual Review of Ecology and Systematics **18:** 43-70.

KIRKPATRICK, M., and N. BARTON, 2006 Chromosome inversions, local adaptation and speciation. Genetics **173:** 419-434.

KNAPP, S. J., W. C. BRIDGES and D. BIRKES, 1990 Mapping Quantitative Trait Loci Using Molecular Marker Linkage Maps. Theoretical and Applied Genetics **79:** 583-592.

KOFLER, R., R. V. PANDEY and C. SCHLOETTERER, 2011 PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics **27:** 3435-3436.

KOLACZKOWSKI, B., A. D. KERN, A. K. HOLLOWAY and D. J. BEGUN, 2011 Genomic Differentiation Between Temperate and Tropical Australian Populations of *Drosophila melanogaster*. Genetics **187:** 245-260.

KRIEGER, M. J. B., and K. G. ROSS, 2002 Identification of a major gene regulating complex social behavior. Science **295:** 328-332.

KURTZ, S., A. PHILLIPPY, A. L. DELCHER, M. SMOOT, M. SHUMWAY *et al.*, 2004 Versatile and open software for comparing large genomes. Genome Biology **5**.

KWAN, L., S. BEDHOMME, N. G. PRASAD and A. K. CHIPPINDALE, 2008 Sexual conflict and environmental change: trade-offs within and between the sexes during the evolution of desiccation resistance. Journal of Genetics **87:** 383-394.

LAMASON, R. L., M. MOHIDEEN, J. R. MEST, A. C. WONG, H. L. NORTON *et al.*, 2005 SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. Science **310:** 1782-1786.

LANDE, R., 1980 The Genetic Covariance between Characters Maintained by Pleiotropic Mutations. Genetics **94:** 203-215.

LANDE, R., 1981 Models of speciation by sexual selection on polygenic traits. Proc Natl Acad Sci U S A **78:** 3721-3725.

LANDE, R., 1982 Rapid Origin of Sexual Isolation and Character Divergence in a Cline. Evolution **36:** 213-223.

LANDE, R., and S. J. ARNOLD, 1983 The Measurement of Selection on Correlated Characters. Evolution **37:** 1210-1226.

LANDE, R., and M. KIRKPATRICK, 1988 Ecological speciation by sexual selection. J Theor Biol **133:** 85-98.

LANDER, E. S., and D. BOTSTEIN, 1989 Mapping Mendelian Factors Underlying Quantitative Traits Using Rflp Linkage Maps. Genetics **121:** 185-199.

LANG, C., and F. MENZEL, 2011 *Lasius niger* ants discriminate aphids based on their cuticular hydrocarbons. Animal Behaviour **82:** 1245-1254.

LASSANCE, J.-M., A. T. GROOT, M. A. LIENARD, B. ANTONY, C. BORGWARDT *et al.*, 2010 Allelic variation in a fatty-acyl reductase gene causes divergence in moth sex pheromones. Nature **466:** 486-U487.

LASSANCE, J.-M., M. A. LIENARD, B. ANTONY, S. QIAN, T. FUJII *et al.*, 2013 Functional consequences of sequence variation in the pheromone biosynthetic gene pgFAR for Ostrinia moths. Proceedings of the National Academy of Sciences of the United States of America **110:** 3967-3972.

LATTA, R. G., and K. M. GARDNER, 2009 Natural selection on pleiotropic quantitative trait Loci affecting a life-history trade-off in *Avena barbata*. Evolution **63:** 2153-2163.

LAZELL, J. D., 1972 The Anoles (Sauria, Iguanidae) of the Lesser Antilles Bull. Mus. Comp. Zool., Harvard Univ. **143:** 1-115.

LEGENDRE, A., X.-X. MIAO, J.-L. DA LAGE and C. WICKER-THOMAS, 2008 Evolution of a desaturase involved in female pheromonal cuticular hydrocarbon biosynthesis and courtship behavior in *Drosophila*. Insect Biochemistry and Molecular Biology **38:** 244-255.

LEINONEN, T., R. B. O'HARA, J. M. CANO and J. MERILA, 2008 Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. Journal of Evolutionary Biology **21:** 1-17.

LI, H., and R. DURBIN, 2009 Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics **25:** 1754-1760.

LI, H., B. HANDSAKER, A. WYSOKER, T. FENNELL, J. RUAN *et al.*, 2009 The Sequence Alignment/Map format and SAMtools. Bioinformatics **25:** 2078-2079.

LIEFTING, M., A. A. HOFFMANN and J. ELLERS, 2009 Plasticity versus environmental canalization: population differences in thermal responses along a latitudinal gradient in *Drosophila serrata*. Evolution **63:** 1954-1963.

LIENARD, M. A., A. K. HAGSTROM, J.-M. LASSANCE and C. LOFSTEDT, 2010 Evolution of multicomponent pheromone signals in small ermine moths involves a single fatty-acyl reductase gene. Proceedings of the National Academy of Sciences of the United States of America **107:** 10955-10960.

LINNEN, C. R., E. P. KINGSLEY, J. D. JENSEN and H. E. HOEKSTRA, 2009 On the Origin and Spread of an Adaptive Allele in Deer Mice. Science **325:** 1095-1098.

LIONIKAS, A., R. CHENG, J. E. LIM, A. A. PALMER and D. A. BLIZARD, 2010 Fine-mapping of muscle weight QTL in LG/J and SM/J intercrosses. Physiological Genomics **42A:** 33-38.

LITI, G., D. M. CARTER, A. M. MOSES, J. WARRINGER, L. PARTS *et al.*, 2009 Population genomics of domestic and wild yeasts. Nature **458:** 337-341.

LOCKEY, K. H., 1988 Lipids of the Insect Cuticle - Origin, Composition and Function. Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology **89:** 595-645.

LORCH, P. D., S. PROULX, L. ROWE and T. DAY, 2003 Condition-dependent sexual selection can accelerate adaptation. Evolutionary Ecology Research **5:** 867-881.

LOS, D. A., and N. MURATA, 1998 Structure and expression of fatty acid desaturases. Biochimica Et Biophysica Acta-Lipids and Lipid Metabolism **1394:** 3-15.

LOS, D. A., M. K. RAY and N. MURATA, 1997 Differences in the control of the temperature-dependent expression of four genes for desaturases in *Synechocystis* sp. PCC 6803. Molecular Microbiology **25:** 1167-1175.

MA, X., X. LU and J. MERILA, 2009 Altitudinal decline of body size in a Tibetan frog. Journal of Zoology **279:** 364-371.

MAAN, M. E., and M. E. CUMMINGS, 2009 Sexual dimorphism and directional sexual selection on aposematic signals in a poison frog. Proc Natl Acad Sci U S A **106:** 19072-19077.

MAAN, M. E., and O. SEEHAUSEN, 2011 Ecology, sexual selection and speciation. Ecology Letters **14:** 591-602.

MACKAY, T. F. C., 2004 Genetic dissection of quantitative traits. Evolution of Population Biology**:** 51-73.

MACKAY, T. F. C., 2009 Mutations and quantitative genetic variation: lessons from *Drosophila*. Philosophical Transactions of the Royal Society B-Biological Sciences **365:** 1229-1239.

MAGIAFOGLOU, A., and A. A. HOFFMANN, 2003 Cross-generation effects due to cold exposure in *Drosophila serrata*. Functional Ecology **17:** 664-672.

MAGWIRE, M. M., D. K. FABIAN, H. SCHWEYEN, C. CAO, B. LONGDON *et al.*, 2012 Genome-wide association studies reveal a simple genetic basis of resistance to naturally coevolving viruses in *Drosophila melanogaster*. PLoS Genet **8:** e1003057.

MALLOCH, J. R., 1927 Notes on Australian Diptera. X. . Proc. Linnean Soc. N. S. Wales **52, Pt. II:** 1-16.

MARCILLAC, F., F. BOUSQUET, J. ALABOUVETTE, F. SAVARIT and J. F. FERVEUR, 2005a A mutation with major effects on *Drosophila melanogaster* sex pheromones. Genetics **171:** 1617-1628.

MARCILLAC, F., Y. GROSJEAN and J. F. FERVEUR, 2005b A single mutation alters production and discrimination of *Drosophila* sex pheromones. Proceedings of the Royal Society B-Biological Sciences **272:** 303-309.

MASEL, J., 2006 Cryptic genetic variation is enriched for potential adaptations. Genetics **172:** 1985-1991.

MAYR, E., 1963 *Animal Species and Evolution*. Harvard University Press, Cambridge, Massachusetts.

MCGUIGAN, K., A. VAN HOMRIGH and M. W. BLOWS, 2008 An evolutionary limit to male mating success. Evolution **62:** 1528-1537.

MCINNES, A. S., and A. QUIGG, 2010 Near-Annual Fish Kills in Small Embayments: Casual vs. Causal Factors. Journal of Coastal Research **26:** 957-966.

MCMILLAN, I., and A. ROBERTSON, 1974 The power of methods for the detection of major genes affecting quantitative characters. Heredity (Edinb) **32:** 349-356.

MERILA, J., and P. CRNOKRAK, 2001 Comparison of genetic differentiation at marker loci and quantitative traits. Journal of Evolutionary Biology **14:** 892-903.

MERILA, J., and B. C. SHELDON, 1999 Genetic architecture of fitness and nonfitness traits: empirical patterns and development of ideas. Heredity **83:** 103-109.

MICHAELSON, J. J., Y. J. SHI, M. GUJRAL, H. C. ZHENG, D. MALHOTRA *et al.*, 2012 Whole-Genome Sequencing in Autism Identifies Hot Spots for De Novo Germline Mutation. Cell **151:** 1431-1442.

MICHELMORE, R. W., I. PARAN and R. V. KESSELI, 1991 Identification of Markers Linked to Disease-Resistance Genes by Bulked Segregant Analysis - a Rapid Method to Detect Markers in Specific Genomic Regions by Using Segregating Populations. Proceedings of the National Academy of Sciences of the United States of America **88:** 9828-9832.

MILLER, M. P., and E. R. VINCENT, 2008 Rapid natural selection for resistance to an introduced parasite of rainbow trout. Evolutionary Applications **1:** 336-341.

MONTOOTH, K. L., and A. G. GIBBS, 2003 Cuticular pheromones and water balance in the house fly, Musca domestica. Comparative Biochemistry and Physiology a-Molecular & Integrative Physiology **135:** 457-465.

MORGAN, T. J., and T. F. C. MACKAY, 2006 Quantitative trait loci for thermotolerance phenotypes in *Drosophila melanogaster*. Heredity **96:** 232-242.

MOTO, K., T. YOSHIGA, M. YAMAMOTO, S. TAKAHASHI, K. OKANO *et al.*, 2003 Pheromone gland-specific fatty-acyl reductase of the silkmoth, *Bombyx mori*. Proc Natl Acad Sci U S A **100:** 9156-9161.

MPURU, S., J. R. REED, R. C. REITZ and G. J. BLOMQUIST, 1996 Mechanism of hydrocarbon biosynthesis from aldehyde in selected insect species: Requirement for O-2 and

NADPH and carbonyl group released as CO2. Insect Biochemistry and Molecular Biology **26:** 203-208.

MUCHERO, W., N. N. DIOP, P. R. BHAT, R. D. FENTON, S. WANAMAKER *et al.*, 2009 A consensus genetic map of cowpea *Vigna unguiculata* (L) Walp. and synteny based on EST-derived SNPs. Proceedings of the National Academy of Sciences of the United States of America **106:** 18159-18164.

MULLEN, L. M., and H. E. HOEKSTRA, 2008 Natural selection along an environmental gradient: A classic cline in mouse pigmentation. Evolution **62:** 1555-1569.

MULLEN, L. M., S. N. VIGNIERI, J. A. GORE and H. E. HOEKSTRA, 2009 Adaptive basis of geographic variation: genetic, phenotypic and environmental differences among beach mouse populations. Proceedings of the Royal Society B-Biological Sciences **276:** 3809-3818.

MURATA, N., and H. WADA, 1995 Acyl-Lipid Desaturases and Their Importance in the Tolerance and Acclimatization to Cold of Cyanobacteria. Biochemical Journal **308:** 1-8.

NACHMAN, M. W., 2005 The genetic basis of adaptation: lessons from concealing coloration in pocket mice. Genetica **123:** 125-136.

NELSON, D. R., and D. A. CARLSON, 1986 Cuticular Hydrocarbons of the Tsetse-Flies Glossina-Morsitans-Morsitans, *Glossina austeni* and *Glossina pallidipes*. Insect Biochemistry **16:** 403-416.

NIEHUIS, O., J. BUELLESBACH, A. K. JUDSON, T. SCHMITT and J. GADAU, 2011 Genetics of cuticular hydrocarbon differences between males of the parasitoid wasps *Nasonia giraulti* and *Nasonia vitripennis*. Heredity **107:** 61-70.

NIEWIAROWSKI, P. H., and W. ROOSENBURG, 1993 Reciprocal Transplant Reveals Sources of Variation in Growth-Rates of the Lizard *Sceloporus undulatus*. Ecology **74:** 1992-2002.

NOOR, M. A. F., and J. A. COYNE, 1996 Genetics of a difference in cuticular hydrocarbons between *Drosophila pseudoobscura* and *D. persimilis*. Genetical Research **68:** 117-123.

NORRY, F. M., J. DAHLGAARD and V. LOESCHCKE, 2004 Quantitative trait loci affecting knockdown resistance to high temperature in *Drosophila melanogaster*. Molecular Ecology **13:** 3585-3594.

NOVEMBRE, J., and A. DI RIENZO, 2009 Spatial patterns of variation due to natural selection in humans. Nature Reviews Genetics **10:** 745-755.

NOVEMBRE, J., A. P. GALVANI and M. SLATKIN, 2005 The geographic spread of the CCR5 Delta 32 HIV-resistance allele. Plos Biology **3:** 1954-1962.

O'MALLEY, K. G., and M. A. BANKS, 2008 A latitudinal cline in the Chinook salmon (*Oncorhynchus tshawytscha*) Clock gene: evidence for selection on PolyQ length variants. Proceedings of the Royal Society B-Biological Sciences **275:** 2813-2821.

OAKESHOTT, J. G., J. B. GIBSON, P. R. ANDERSON, W. R. KNIBB, D. G. ANDERSON *et al.*, 1982 Alcohol-Dehydrogenase and Glycerol-3-Phosphate Dehydrogenase Clines in *Drosophila melanogaster* on Different Continents. Evolution **36:** 86-96.

OLSON, W., E. EMMENEGGER, J. GLENN, J. WINTON and F. GOETZ, 2013 Comparative susceptibility among three stocks of yellow perch, *Perca flavescens* (Mitchill), to viral haemorrhagic septicaemia virus strain IVb from the Great Lakes. Journal of fish diseases **36:** 711-719.

ONEAL, E., and L. L. KNOWLES, 2013 Ecological selection as the cause and sexual differentiation as the consequence of species divergence? Proceedings of the Royal Society B-Biological Sciences **280:** 22236-22236.

ORR, H. A., 1998a The population genetics of adaptation: The distribution of factors fixed during adaptive evolution. Evolution **52:** 935-949.

ORR, H. A., 1998b Testing natural selection vs. genetic drift in phenotypic evolution using quantitative trait locus data. Genetics **149:** 2099-2104.

PALSSON, A., J. DODGSON, I. DWORKIN and G. GIBSON, 2005 Tests for the replication of an association between Egfr and natural variation in *Drosophila melanogaster* wing morphology. BMC Genet **6:** 44.

PARK, M. O., 2005 New pathway for long-chain n-alkane synthesis via 1-alcohol in *Vibrio furnissii* M1. Journal of Bacteriology **187:** 1426-1429.

PEGUEROLES, C., V. ORDONEZ, F. MESTRES and M. PASCUAL, 2010 Recombination and selection in the maintenance of the adaptive value of inversions. Journal of Evolutionary Biology **23:** 2709-2717.

PEREZ-VEGA, E., A. PANEDA, C. RODRIGUEZ-SUAREZ, A. CAMPA, R. GIRALDEZ *et al.*, 2010 Mapping of QTLs for morpho-agronomic and seed quality traits in a RIL population of common bean (*Phaseolus vulgaris* L.). Theoretical and Applied Genetics **120:** 1367-1380.

POMIANKOWSKI, A., and A. P. MOLLER, 1995 A Resolution of the Lek Paradox. Proceedings of the Royal Society B-Biological Sciences **260:** 21-29.

PRASAD, N. G., S. BEDHOMME, T. DAY and A. K. CHIPPINDALE, 2007 An evolutionary cost of separate genders revealed by male-limited evolution. American Naturalist **169:** 29-37.

PROTAS, M., I. TABANSKY, M. CONRAD, J. B. GROSS, O. VIDAL *et al.*, 2008 Multi-trait evolution in a cave fish, *Astyanax mexicanus*. Evolution & Development **10:** 196-209.

PUNZALAN, D., F. H. RODD and K. A. HUGHES, 2005 Perceptual processes and the maintenance of polymorphism through frequency-dependent predation. Evolutionary Ecology **19:** 303-320.

QIAN, W., L. DENG, D. S. LU and S. H. XU, 2013 Genome-Wide Landscapes of Human Local Adaptation in Asia. Plos One **8**.

RAIA, P., and S. MEIRI, 2011 The tempo and mode of evolution: body sizes of island mammals. Evolution **65:** 1927-1934.

RANZ, J. M., F. CASALS and A. RUIZ, 2001 How malleable is the eukaryotic genome? Extreme rate of chromosomal rearrangement in the genus *Drosophila*. Genome Research **11:** 230-239.

REED, J. R., D. VANDERWEL, S. CHOI, J. G. POMONIS, R. C. REITZ *et al.*, 1994 Unusual mechanism of hydrocarbon formation in the housefly: cytochrome P450 converts aldehyde to the sex pheromone component (Z)-9-tricosene and CO2. Proc Natl Acad Sci U S A **91:** 10000-10004.

RICHMAN, A., 2000 Evolution of balanced genetic polymorphism. Molecular Ecology **9:** 1953-1963.

RIESEBERG, L. H., O. RAYMOND, D. M. ROSENTHAL, Z. LAI, K. LIVINGSTONE *et al.*, 2003 Major ecological transitions in wild sunflowers facilitated by hybridization. Science **301:** 1211-1216.

RIESEBERG, L. H., A. WIDMER, A. M. ARNTZ and J. M. BURKE, 2002 Directional selection is the primary cause of phenotypic diversification. Proceedings of the National Academy of Sciences of the United States of America **99:** 12242-12245.

RITCHIE, M. G., 2007 Sexual selection and speciation. Annual Review of Ecology Evolution and Systematics **38:** 79-102.

ROCKMAN, M. V., 2012 The QTN program and the alleles that matter for evolution: all that's gold does not glitter. Evolution **66:** 1-17.

ROWLAND, O., H. ZHENG, S. R. HEPWORTH, P. LAM, R. JETTER *et al.*, 2006 CER4 encodes an alcohol-forming fatty acyl-coenzyme A reductase involved in cuticular wax production in *Arabidopsis*. Plant Physiol **142:** 866-877.

RULE, G. S., and W. L. ROELOFS, 1989 Biosynthesis of Sex-Pheromone Components from Linolenic Acid in Arctiid Moths. Archives of Insect Biochemistry and Physiology **12:** 89-97.

RUNDLE, H. D., S. F. CHENOWETH and M. W. BLOWS, 2006 The roles of natural and sexual selection during adaptation to a novel environment. Evolution **60:** 2218-2225.

RUNDLE, H. D., S. F. CHENOWETH and M. W. BLOWS, 2009 The diversification of mate preferences by natural and sexual selection. Journal of Evolutionary Biology **22:** 1608-1615.

RUNDLE, H. D., S. F. CHENOWETH, P. DOUGHTY and M. W. BLOWS, 2005 Divergent selection and the evolution of signal traits and mating preferences. Plos Biology **3:** 1988-1995.

SAKAMOTO, T., S. HIGASHI, H. WADA, N. MURATA and D. A. BRYANT, 1997 Low-temperature-induced desaturation of fatty acids and expression of desaturase genes in the cyanobacterium *Synechococcus* sp. PCC 7002. Fems Microbiology Letters **152:** 313-320.

SALOMON, M., 2002 A revised cline theory that can be used for quantified analyses of evolutionary processes without parapatric speciation. Journal of Biogeography **29:** 509-517.

SANCHEZ-GUILLEN, R. A., B. HANSSON, M. WELLENREUTHER, E. I. SVENSSON and A. CORDERO-RIVERA, 2011 The influence of stochastic and selective forces in the population divergence of female colour polymorphism in damselflies of the genus *Ischnura*. Heredity **107:** 513-522.

SCARCELLI, N., J. M. CHEVERUD, B. A. SCHAAL and P. X. KOVER, 2007 Antagonistic pleiotropic effects reduce the potential adaptive value of the *FRIGIDA* locus. Proceedings of the National Academy of Sciences of the United States of America **104:** 16986-16991.

SCHAL, C., V. L. SEVALA, H. P. YOUNG and J. A. S. BACHMANN, 1998 Sites of synthesis and transport pathways of insect hydrocarbons: Cuticle and ovary as target tissues. American Zoologist **38:** 382-393.

SCHLUTER, D., 2000 *The ecology of adaptive radiation*. Oxford University Press.

SCHMIDT, P. S., L. MATZKIN, M. IPPOLITO and W. F. EANES, 2005 Geographic variation in diapause incidence, life-history traits, and climatic adaptation in *Drosophila melanogaster*. Evolution **59:** 1721-1732.

SCOVILLE, A. G., Y. W. LEE, J. H. WILLIS and J. K. KELLY, 2011 Explaining the heritability of an ecologically significant trait in terms of individual quantitative trait loci. Biology Letters **7:** 896-898.

SGRO, C. M., and M. W. BLOWS, 2004 The genetic covariance among clinal environments after adaptation to an environmental gradient in *Drosophila serrata*. Genetics **167:** 1281-1291.

SHARMA, M. D., J. HUNT and D. J. HOSKEN, 2012 Antagonistic responses to natural and sexual selection and the sex-specific evolution of cuticular hydrocarbons in *Drosophila simulans*. Evolution **66:** 665-677.

SHIRANGI, T. R., H. D. DUFOUR, T. M. WILLIAMS and S. B. CARROLL, 2009 Rapid Evolution of Sex Pheromone-Producing Enzyme Expression in *Drosophila*. Plos Biology **7**.

SINCLAIR, B. J., A. G. GIBBS and S. P. ROBERTS, 2007 Gene transcription during exposure to, and recovery from, cold and desiccation stress in *Drosophila melanogaster*. Insect Molecular Biology **16:** 435-443.

SINHA, H., L. DAVID, R. C. PASCON, S. CLAUDER-MUNSTER, S. KRISHNAKUMAR *et al.*, 2008 Sequential Elimination of Major-Effect Contributors Identifies Additional Quantitative Trait Loci Conditioning High-Temperature Growth in Yeast. Genetics **180:** 1661-1670.

SKROBLIN, A., and M. W. BLOWS, 2006 Measuring natural and sexual selection on breeding values of male display traits in *Drosophila serrata*. Journal of Evolutionary Biology **19:** 35-41.

SLATKIN, M., 2008 Linkage disequilibrium - understanding the evolutionary past and mapping the medical future. Nature Reviews Genetics **9:** 477-485.

SMITH, T. B., C. J. SCHNEIDER and K. HOLDER, 2001 Refugial isolation versus ecological gradients. Genetica **112:** 383-398.

SOTA, T., and T. TANABE, 2010 Multiple speciation events in an arthropod with divergent evolution in sexual morphology. Proceedings of the Royal Society B-Biological Sciences **277:** 689-696.

SPITZE, K., 1993 Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. Genetics **135:** 367-374.

STAM, L. F., and C. C. LAURIE, 1996 Molecular dissection of a major gene effect on a quantitative trait: The level of alcohol dehydrogenase expression in *Drosophila melanogaster*. Genetics **144:** 1559-1564.

STERN, D. L., and V. ORGOGOZO, 2008 The loci of evolution: How predictable is genetic evolution ? Evolution **62:** 2155-2177.

STOCKER, A. J., B. B. RUSUWA, M. J. BLACKET, F. D. FRENTIU, M. SULLIVAN *et al.*, 2012 Physical and Linkage Maps for *Drosophila serrata*, a Model Species for Studies of Clinal Adaptation and Sexual Selection. G3 (Bethesda, Md.) **2:** 287-297.

STUDER, A. J., and J. F. DOEBLEY, 2011 Do Large Effect QTL Fractionate? A Case Study at the Maize Domestication QTL teosinte branched1. Genetics **188:** 673-U257.

SUPPLY, P., M. MARCEAU, S. MANGENOT, D. ROCHE, C. ROUANET *et al.*, 2013 Genomic analysis of smooth tubercle bacilli provides insights into ancestry and pathoadaptation of *Mycobacterium tuberculosis*. Nature Genetics **45:** 172-179.

SVENSSON, O., A. NYMAN and C. KVARNEMO, 2004 Costly courtship or dishonest display? Intensely displaying sand goby males have lower lipid content. Journal of Fish Biology **64:** 1425-1429.

TAKAHASHI, A., S. C. TSAUR, J. A. COYNE and C. I. WU, 2001 The nucleotide changes governing cuticular hydrocarbon variation and their evolution in *Drosophila melanogaster*. Proceedings of the National Academy of Sciences of the United States of America **98:** 3920-3925.

TASAKA, Y., Z. GOMBOS, Y. NISHIYAMA, P. MOHANTY, T. OHBA *et al.*, 1996 Targeted mutagenesis of acyl-lipid desaturases in *Synechocystis*: Evidence for the important roles of polyunsaturated membrane lipids in growth, respiration and photosynthesis. Embo Journal **15:** 6416-6425.

TEAM, R. D. C., 2008 R: A language and environment for statistical computing, pp. R Foundation for Statistical Computing, Vienna, Austria.

TEERAWANICHPAN, P., A. J. ROBERTSON and X. QIU, 2010 A fatty acyl-CoA reductase highly expressed in the head of honey bee (*Apis mellifera*) involves biosynthesis of a wide range of aliphatic fatty alcohols. Insect Biochem Mol Biol **40:** 641-649.

TILLMAN-WALL, J. A., D. VANDERWEL, M. E. KUENZLI, R. C. REITZ and G. J. BLOMQUIST, 1992 Regulation of sex pheromone biosynthesis in the housefly, *Musca domestica*: relative contribution of the elongation and reductive steps. Archives of Biochemistry and Biophysics **299:** 92-99.

TILLMAN, J. A., S. J. SEYBOLD, R. A. JURENKA and G. J. BLOMQUIST, 1999 Insect pheromones - an overview of biosynthesis and endocrine regulation. Insect Biochemistry and Molecular Biology **29:** 481-514.

TOBLER, M., and E. W. CARSON, 2010 Environmental variation, hybridization, and phenotypic diversification in Cuatro Cienegas pupfishes. Journal of Evolutionary Biology **23:** 1475-1489.

TOOLSON, E. C., and R. KUPERSIMBRON, 1989 Laboratory Evolution of Epicuticular Hydrocarbon Composition and Cuticular Permeability in *Drosophila pseudoobscura* - Effects on Sexual Dimorphism and Thermal-Acclimation Ability. Evolution **43:** 468-473.

TRAVISANO, M., J. A. MONGOLD, A. F. BENNETT and R. E. LENSKI, 1995 Experimental tests of the roles of adaptation, chance, and history in evolution. Science **267:** 87-90.

TURGEON, J., and L. BERNATCHEZ, 2001 Clinal variation at microsatellite loci reveals historical secondary intergradation between glacial races of *Coregonus artedi* (Teleostei : Coregoninae). Evolution **55:** 2274-2286.

UMINA, P. A., A. R. WEEKS, M. R. KEARNEY, S. W. MCKECHNIE and A. A. HOFFMANN, 2005 A rapid shift in a classic clinal pattern in *Drosophila* reflecting climate change. Science **308:** 691-693.

UNTERGASSER, A., H. NIJVEEN, X. RAO, T. BISSELING, R. GEURTS *et al.*, 2007 Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Research **35:** W71-W74.

VAN DOORN, G. S., P. EDELAAR and F. J. WEISSING, 2009 On the Origin of Species by Natural and Sexual Selection. Science **326:** 1704-1707.

VAN LEEUWEN, T., P. DEMAEGHT, E. J. OSBORNE, W. DERMAUW, S. GOHLKE *et al.*, 2012 Population bulk segregant mapping uncovers resistance mutations and the mode of action of a chitin synthesis inhibitor in arthropods. Proc Natl Acad Sci U S A **109:** 4407-4412.

VAN OOIJEN, J. W., 2006 JoinMap 4, Software for the calculation of genetic linkage maps in experimental populations. Kyazma, B. V. Wageningen, Netherlands.

VASEMAGI, A., 2006 The adaptive hypothesis of clinal variation revisited: Single-locus clines as a result of spatially restricted gene flow. Genetics **173:** 2411-2414.

VAZ, A. H., R. A. JURENKA, G. J. BLOMQUIST and R. C. REITZ, 1988 Tissue and chain length specificity of the fatty acyl-CoA elongation system in the American cockroach. Arch Biochem Biophys **267:** 551-557.

VEGA, S. E., A. H. DEL RIO, J. B. BAMBERG and J. P. PALTA, 2004 Evidence for the up-regulation of stearoyl-ACP (A9) desaturase gene expression during cold acclimation. American Journal of Potato Research **81:** 125-135.

VIA, S., 1984 The Quantitative Genetics of Polyphagy in an Insect Herbivore .2. Genetic Correlations in Larval Performance within and among Host Plants. Evolution **38:** 896-905.

VONSCHANTZ, T., M. TUFVESSON, G. GORANSSON, M. GRAHN, M. WILHELMSON *et al.*, 1995 Artificial Selection for Increased Comb Size and Its Effects on Other Sexual Characters and Viability in *Gallus domesticus* (the Domestic Chicken). Heredity **75:** 518-529.

VOSS, S. R., and H. B. SHAFFER, 1997 Adaptive evolution via a major gene effect: Paedomorphosis in the Mexican axolotl. Proceedings of the National Academy of Sciences of the United States of America **94:** 14185-14189.

WADA, H., Z. GOMBOS and N. MURATA, 1990 Enhancement of chilling tolerance of a cyanobacterium by genetic manipulation of fatty acid desaturation. Nature **347:** 200-203.

WAGNER, C. E., A. R. MCCUNE and I. J. LOVETTE, 2012 Recent speciation between sympatric Tanganyikan cichlid colour morphs. Molecular Ecology **21:** 3283-3292.

WAGNER, G. P., M. PAVLICEV and J. M. CHEVERUD, 2007 The road to modularity. Nature Reviews Genetics **8:** 921-931.

WAGNER, G. P., and J. ZHANG, 2011 The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms. Nature Reviews Genetics **12:** 204-213.

WALTERS, S. A., T. C. WEHNER and K. R. BARKER, 1997 A single recessive gene for resistance to the root-knot nematode (*Meloidogyne javanica*) in *Cucumis sativus* var hardwickii. Journal of Heredity **88:** 66-69.

WANG, J., K. D. GONZALEZ, W. A. SCARINGE, K. TSAI, N. LIU *et al.*, 2007 Evidence for mutation showers. Proc Natl Acad Sci U S A **104:** 8403-8408.

WANG, X., and P. E. KOLATTUKUDY, 1995 Solubilization and purification of aldehyde-generating fatty acyl-CoA reductase from green alga *Botryococcus braunii*. FEBS Lett **370:** 15-18.

WANG, X. S., I. LE ROY, E. NICODEME, R. H. LI, R. WAGNER *et al.*, 2003 Using advanced intercross line for high-resolution mapping of HDL cholesterol quantitative trait loci. Genome Research **13:** 1654-1664.

WEEKS, A. R., S. W. MCKECHNIE and A. A. HOFFMANN, 2002 Dissecting adaptive clinal variation: markers, inversions and size/stress associations in *Drosophila melanogaster* from a central field population. Ecology Letters **5:** 756-763.

WEST, P. M., and C. PACKER, 2002 Sexual selection, temperature, and the lion's mane. Science **297:** 1339-1343.

WESTEBERHARD, M. J., 1983 Sexual Selection, Social Competition, and Speciation. Quarterly Review of Biology **58:** 155-183.

WHITEHEAD, A., J. L. ROACH, S. ZHANG and F. GALVEZ, 2011 Genomic mechanisms of evolved physiological plasticity in killifish distributed along an environmental salinity gradient. Proc Natl Acad Sci U S A **108:** 6193-6198.

WICKER-THOMAS, C., I. GUENACHI and Y. F. KEITA, 2009 Contribution of oenocytes and pheromones to courtship behaviour in *Drosophila*. BMC Biochem **10:** 21.

WICKER-THOMAS, C., and J. M. JALLON, 2001 Control of female pheromones in *Drosophila melanogaster* by homeotic genes. Genet Res **78:** 235-242.

WOOD, E. T., D. A. STOVER, M. SLATKIN, M. W. NACHMAN and M. F. HAMMER, 2005 The beta -globin recombinational hotspot reduces the effects of strong selection around HbC, a recently arisen mutation providing resistance to malaria. Am J Hum Genet **77:** 637-642.

WRIGHT, D., C. J. RUBIN, A. M. BARRIO, K. SCHUTZ, S. KERJE *et al.*, 2010 The genetic architecture of domestication in the chicken: effects of pleiotropy and linkage. Molecular Ecology **19:** 5140-5156.

WU, W. Q., J. W. ZHU, J. MILLAR and C. LOFSTEDT, 1998 A comparative study of sex pheromone biosynthesis in two strains of the turnip moth, *Agrotis segetum*, producing different ratios of sex pheromone components. Insect Biochemistry and Molecular Biology **28:** 895-900.

YEAMAN, S., and S. P. OTTO, 2011 Establishment and maintenance of adaptive genetic divergence under migration, selection, and drift. Evolution **65:** 2123-2129.

YEAMAN, S., and M. C. WHITLOCK, 2011 The genetic architecture of adaptation under migration-selection balance. Evolution **65:** 1897-1911.

YOUNG, N. D., 1996 QTL mapping and quantitative disease resistance in plants. Annual Review of Phytopathology **34:** 479-501.

ZAHAVI, A., 1975 Mate selection-a selection for a handicap. J Theor Biol **53:** 205-214.

8. APPENDIX

A list of EST-derived SNP markers for the construction of a *D. serrata* linkage map. Sequenced PCR products from these markers were BLASTed onto the *D. melanogaster* genome on FlyBase to validate their corresponding chromosome arm locations. The P-value derives from the χ^2 test of allele frequency distortion from the expected Mendelian ratio of 1:2:1 for the autosomal markers and 1:1 for X-linked markers.

Marker	Linkage Group	*D. melanogaster* Chromosome Arm	CG No.	FlyBase Locus Name	FlyBase ID	MAF	Pvalue
s1	X	X	CG14792	Dmel\sta	FBgn0003517	0.45	0.000
s2	X	X	CG6182	Dmel\Tsf1	FBgn0022355	0.40	0.000
s3	3	3R	CG5887	Dmel\desat1	FBgn0086687	0.49	0.000
s4	2	2R	CG3401	Dmel\BTub60D	FBgn0003888	0.35	0.107
s5	3	3R	CG11522	Dmel\RpL6	FBgn0039857	0.26	0.012
s6	2	2L	CG7361	Dmel\RFeSp	FBgn0021906	0.43	0.001
s7	3	3R	CG2216	Dmel\Fer1HCH	FBgn0015222	0.43	0.000
s8	2	2L	CG13092	Dmel\Db31	FBgn0032048	0.37	0.290
s9	2	2R	CG3124	Dmel\CG3124	FBgn0034840	0.29	0.000
s10	**Not informative**	**2L**	**CG9042**	**Dmel\Gpdh**	**FBgn0001128**	**0.00**	**0.000**
s11	**2**	**2L**	**CG31811**	**Dmel\cenG1A**	**FBgn0028509**	**0.12**	**0.000**
s12	2	2L	CG31811	Dmel\cenG1A	FBgn0028509	0.29	0.000
s13	3	3R	CG15697	Dmel\RpS30	FBgn0038834	0.31	0.546
s14	3	3L	CG6988	Dmel\Pdi	FBgn0014002	0.31	0.345
s15	3	3R	CG11901	Dmel\EF1y	FBgn0029176	0.33	0.000
s16	2	2L	CG34394	Dmel\CG34394	FBgn0085423	0.39	0.042
s17	3	3R	CG8036	Dmel\CG8036	FBgn0037607	0.29	0.046
s18	2	2L	CG9244	Dmel\Acon	FBgn0010100	0.46	0.000
s19	2	2R	CG6692	Dmel\Cp1	FBgn0013770	0.38	0.084
s20	3	3R	CG5502	Dmel\RpL4	FBgn0003279	0.29	0.138
s21	2	2L	CG6105	Dmel\(2)06225	FBgn0010612	0.49	0.000
s22	3	3L	CG4769	Dmel\CG4769	FBgn0035600	0.31	0.478
s23	**2**	**2R**	**CG5330**	**Dmel\Nap1**	**FBgn0015268**	**0.18**	**0.000**
s24	2	2R	CG5330	Dmel\Nap1	FBgn0015268	0.37	0.338
s25	3	3L	CG11276	RpS4	FBgn0011284	0.28	0.026
s26	3	3L	CG6871	Dmel\Cat	FBgn0000261	0.29	0.210
s27	3	3L	CG6871	Dmel\Cat	FBgn0000261	0.30	0.347
s28	X	X	CG8893	Dmel\Gapdh2	FBgn0001091	0.43	0.000
s29	X	X	CG14235	Dmel\CG14235	FBgn0031066	0.42	0.000
s30	3	3R	CG2216	Dmel\Fer1HCH	FBgn0015222	0.44	0.000
s31	3	3R	CG10901	Dmel\osk	FBgn0003015	0.30	0.482
s32	3	3R	CG11901	Dmel\EF1 γ	FBgn0029176	0.30	0.364
s33	X	X	CG1372	Dmel\yl	FBgn0004649	0.43	0.000
s34	2	2R	CG3161	Dmel\Vha16-1	FBgn0262736	**0.17**	0.000
s35	3	3R	CG11033	Dmel\Kdm2	FBgn0037659	0.48	0.000
s36	**Not informative**	**3R**	**CG7610**	**Dmel\ATPsyn-γ**	**FBgn0020235**	**0.00**	**0.000**
s37	**Not informative**	**3L**	**CG6988**	**Dmel\Pdi**	**FBgn0014002**	**0.00**	**0.000**
s38	3	3L	CG4183	Hsp26	FBgn0001225	0.29	0.061
s39	2	2L	CG5397	Dmel\CG5397	FBgn0031327	0.40	0.021
s40	X	X	CG32635	Dmel\CG32635	FBgn0052635	0.41	0.000
s41	2	2L	CG4824	Dmel\BicC	FBgn0000182	0.37	0.205
s42	**2**	**2R**	**CG9364**	**Dmel\Treh**	**FBgn0003748**	**0.22**	**0.000**
s43	3	3L	CG10472	Dmel\CG10472	FBgn0035670	0.30	0.421
s44	4	4	No definition line found	1095281-1095303		0.32	0.226
s45	3	3L	CG6806	Dmel\Lsp2	FBgn0002565	0.31	0.507
s46	3	3L	CG11793	Dmel\Sod	FBgn0003462	0.29	0.065
s47	3	3L	CG8189	Dmel\ATPsyn-b	FBgn0019644	0.28	0.032
s48	3	3L	No definition line found	18976884-18976901		0.33	0.485
s49	2	2R	CG18067	Dmel\CG18067	FBgn0034512	0.37	0.113
s50	2	2L	CG4233	Dmel\Got2	FBgn0001125	0.46	0.000
s51	X	X	CG32816	Dmel\CG32816	FBgn0052816	0.46	0.000
s52	2	2R	CG4696	Dmel\Mp20	FBgn0002789	0.35	0.081
s53	2	2L	CG9042	Dmel\Gpdh	FBgn0001128	0.40	0.025
s54	2	2R	CG8983	Dmel\ERp60	FBgn0033663	0.40	0.017
s55	X	X	CG14792	Dmel\sta	FBgn0003517	0.45	0.000
s56	2	2L	CG9075	Dmel\eIF-4a	FBgn0001942	0.43	0.000
s57	2	2L	CG9075	Dmel\eIF-4a	FBgn0001942	0.36	0.213
s58	2	2L	CG31811	Dmel\cenG1A	FBgn0028509	0.47	0.000
s59	2	2R	CG10911	Dmel\CG10911	FBgn0034295	0.43	0.000
s60	3	3R	CG6439	Dmel\CG6439	FBgn0038922	0.32	0.807
s61	2	2L	CG3763	Dmel\Fbp2	FBgn0000640	0.37	0.279