Stony Brook University

The genetic mapping of reproductive diapuase in *Drosophila melanogaster* (Diptera: Drosophilidae)

A Dissertation Presented

by

Lei Zhu

To

The Graduate School

In Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Ecology and Evolution

Stony Brook University

May 2009

Stony Brook University

The Graduate School

Lei Zhu

We, the dissertation committee for the above candidate for the Doctor of Philosophy degree, hereby recommend acceptance of this dissertation.

Walter F. Eanes – Dissertation Advisor
Professor, Ecology and Evolution, Stony Brook University

John R. True – Chairperson of Defense
Associate Professor, Ecology and Evolution, Stony Brook University

R. Geeta
Associate Professor, Ecology and Evolution, Stony Brook University

Paul S. Schmidt
Assistant Professor, Department of Biology, University of Pennsylvania

This dissertation is accepted by the Graduate School

Lawrence Martin
Dean of the Graduate School

Abstract of the Dissertation
The genetic mapping of reproductive diapuase in
Drosophila melanogaster (Diptera: Drosophilidae)

by

Lei Zhu
Doctor of Philosophy
in
Ecology and Evolution
Stony Brook University
2009

A sub-Saharan species *Drosophila melanogaster* has historically spread into temperate regions very recently. Among many traits potentially contributing to this recent adaptation event, reproductive diapause is of great interest. The aim of this study is to map the gene(s) responsible for reproductive diapause in *Drosophila melanogaster*. This trait enables females to arrest oocyte development in previtellognensis stage. Meanwhile, the aging process is greatly reduced and stress resistance enhanced. Altogether, this syndrome enhances overwintering ability. The development biology of reproductive diapause is also an intriguing topic. In insects, the mechanism of diapause has been extensively studied and different molecular signaling mechanisms have been suggested. We have come to the point that mapping the trait in a model species and studying the population genetics of the causal gene will provide novel information about selection and adaptation of overwintering.

The first goal of the proposed research was to map the causal genetic locus. Utilizing a comprehensive approach integrating recombinational and deletion complementation mapping, *P*-element excision knock-out, linkage disequilibrium association mapping, the author identified the locus and the allele producing reproductive diapause in gene *couch potato*. By studying the sequence divergence of this allele and linked sequence polymorphism, the author examined the molecular signature of nature selection together with the origin and genealogy of reproductive diapause.

In particular, the reproductive diapause locus (or loci) is expected to be under selection during adaptation to temperate environments. It is already known that *Drosophila simulans,* a sibling species, does not exhibit a diapause response. The

D. melanogaster African populations, thought to be ancestral, are also non-diapausing. These observations suggest that the diapause allele(s) is likely to have arisen recently. It is also possible that different diapause alleles have independently arisen multiple times. This putative locus is expected to show a certain degree of evidence for directional selection, especially in temperate region populations. This should lead to a latitudinal cline in the frequency of diapause allele as observed for the trait itself. Diapause is associated with a number of life history trade-offs subjecting to balancing selection, which is expected to be reflected in the underlying loci.

The author studied the molecular evolution of *couch potato* in a population sample consisting of a large sample of North American populations and a sample of African populations. *D. simulans* and *D. yakuba* are included as out. The results reflect the action of balancing and directional selection. These results are finally discussed in the context of seasonality adaptation.

Dedicated to

Common Sense

TABLE OF CONTENTS

List of Figures

List of Tables

Acknowledgments

To:

Dr. Prof. James F. Rohlf and Dr. Prof. Walter F. Eanes.

Dr. Shian-Ren Liou and Dr. John R. True, Dr. David Wang, Dr. Efe Sezgin, Dr. R. Geeta and Mr. Kamalraj Mohanraj for their support during this last six years.

Chapter one: The introduction

Diapause and Reproductive Diapause in the Genus of *Drosophila*

Diapause in insects is a state in which development and senescence are halted. In plants, a similar process is usually referred to as dormancy and in vertebrates, hibernation, though whether those processes are mechanistically equivalent is yet to be explored. It can occur in one or more stages of life cycle, embryonic, larval or reproduction diapause. In some species, diapause is a stage of the life cycle and the entering and breaking diapause seems to be controlled by an intrinsic development program rather then environmental stimuli. In many other cases, diapause is a result of a certain interactions between several environmental stimuli, including food availability, population density, temperature and photoperiod. Diapause is usually associated with elevated stress resistance, which enhances the survivability in the stressful environment.

Many *Drosophila* species are known to diapause. Table 1 summarizes currently known diapausing Drosophilinae species. Different species diapause in different life stages and by far, reproductive diapause seems to be the most common strategy. Diapause arose multiple times in different lineages. Nevertheless, due to our limited knowledge of the ecology of many *Drosophila* species, it is possible that some species may diapause in stages yet to be discovered. In reproductive diapause, also known as ovarian diapause, females stop vitellogenesis and shift energy allocation from reproduction to storage. It was first described in *Drosophila nitens* (BERTANI 1947) and subsequently it has been discovered in many other Drosphilids from many different species groups. Reproductive diapause in *Drosophila* is generally thought to be a trait favorable for overwintering in temperate and cold regions. To a great extend, the duration and intensity of diapause is variable at both the population and species levels.

For a species of tropical African origin, *D. melanogaster*, reproductive diapause appears to be a recent adaptation and variable among populations (SCHMIDT *et al.* 2005a; SCHMIDT *et al.* 2005b). Northern temperate species, on the other hand, display strong diapause. A large proportion of Drosophilae are tropical species, and whether many of these species could diapause in response to cues such as heat or desiccation remains unexplored. In the distantly related species *Chymomyza costata*, larval diapause is a response to overcrowding (BOTELLA and MENSUA 1987). However, reproductive diapause is not necessarily an overwintering strategy. For example, *D. robusta* females are known to enter diapause in August-September and most if not all females diapause by October in Missouri, during which time the weather is still considerably warm (CARSON AND

STALKER 1948). Second, reproductive diapause is possibly not the only method for female overwintering. *D. pseudoobscura* and *D. subobscura* overwinter as adults and continue to reproduce during the winter (EPLING *et al.* 1955). Other adaptations, like egg retention, are also possible female overwintering strategies (BOULETREAU-MERLE and FOUILLET 2002).

Reproductive diapause is not only a combination of previtellogenic arrest of oocyte development, elevated energy storage, and stress resistance. The ability to diapause is associated with life history tradeoffs in non-diapausing inducing environments as well (SCHMIDT *et al.* 2005a; SCHMIDT *et al.* 2005b). *D. melanogaster* with diapausing ability have longer life span, and better resistance to both cold stress and starvation. They have longer egg-to-adult development. The female reproductive output per day is also lower and reproduction value is more evenly distributed over time. When taking in to account their longer reproduction time and life span, the total reproduction value is similar. These phenotypes could not only result from the pleiotropic effects of diapause genes, but also beause these traits are subjected to similar selection. For example, only in newly isolated strains from natural populations do diapause-able flies live longer. In strains maintained in lab for many generations (like *Canton-S*), there is no relationship between diapause ability and longevity (personal communication with Dr. Marc Tatar, KHAZAELI and CURTSINGER 2005). These observations imply that the associations between diapause ability and life history tradeoffs can be decoupled and the association is probably not of a functional nature. Understanding the molecular and genetic details of reproductive diapause will contribute greatly to our knowledge of insect adaptation and life history evolution.

The Role of Photoperiod in Reproductive Diapause

Reproductive diapause in *Drosophila* has been shown to be under temperature, photoperiod and endocrine regulation. The relative importance of these conditions varies across species. In *D. robusta*, females will enter diapause when the temperature is still warm and it appears that photoperiod is the determinant cue (CARSON AND STALKER 1948). In many other species, geographic clines of diapause critical day length have been reported. Generally, northern populations enter diapause under a critical day length that is longer than southern populations (Table 1). In *D. melanogaster*, it is unclear whether photoperiod-temperature (SAUNDERS *et al.* 1989) or low temperature alone are sufficient for diapause induction (TATAR and YIN 2001). It is worth noting that in *D. nitens*, females will enter reproductive diapause in November, even when they are maintained under constant thermo-photoperiod lab environment. Diapause in *D. nitens* remains from November until March of the next year, but can be broken using cold treatment (10 days of 5 °C). This feature implies a unique mechanism of diapause in this species.

In most diapausing *Drosophila* species, temperature and photoperiod induce diapause. Different species have different thermo-photoperiod thresholds. In Japan, *D. lummei* will enter diapause when day length is shorter than 17 hours (WATABE 1995), in contrast to 10-12 hours for *D. triauraria* (KIMURA 1983) in the same region. The day length threshold is variable not only interspecificly but also among populations of the same species. In *D. littoralis*, the critical day length ranges from 12 to 20 hours (LANKINEN 1986) and the northern strains have longer critical day lengths than the southern strains. Similar observations have been discovered in other species (Table 1). *D. lacertosa* shows a latitudinal cline of critical day length on Honshu (South) and Hokkaido (North) Island, Japan, but the cline is disjunctive. The cline in Hokkaido is not an extension of that in Honshu, but instead, it has a much longer critical day length. As a result, the clines on the two islands have similar slopes, but different *y*-intercepts (ICHIJO 1986). This is apparently an indication of the usual complexity of the genetic architecture in the pathways involved in photoperiod measure. The duration of diapause of *D. triauraria* from the north is longer than its southern strains (KIMURA 1983). There probably is a cline in diapause critical day length (YOSHIDA and KIMURA 1994a), and in inbred lines of this species, critical day length and diapause duration are correlated (KIMURA and YOSHIDA 1995). In lower temperatures, the critical day length may become longer in some species like *D. testacea* (KIMURA 1982), but not in all species. In *D. melanogaster*, diapause induction was shown to require shorter than 14 hours of day length at 14°C (SAUNDERS *et al.* 1989). As temperature decreases, this day length threshold fades and most *Canton-S* females

will diapause at 10°C regardless of photoperiod (except for 16L: 6D). In fact, several researchers have questioned the role of photoperiod in inducing reproductive diapause in *D. melanogaster*, since it is possible to induce diapause by low temperature alone (TATAR and YIN 2001).

Considerable understanding of circadian cycles and their relationships to photoperiodism has accumulated (reviewed by HARDIN 2005). Since reproductive diapause is generally believed to be under photoperiod control and requires a critical photoperiod trigger, whether and how circadian cycle or clock genes are involved in reproductive diapause has became an intriguing question. The centerpiece of *Drosophila* circadian rhythms and photoperiodism is the Per and Tim interaction (MYERS *et al.* 1995). The Per-Tim hetero-dimer is transported into nucleus and prevents the Clk-Cyc complex from binding to their target region and acting as a transcriptional factor. Clk and Cyc are products of gene *clock* and *cycle* respectively. Clk-Cyc complex binds to E-box and start not only *period* (*Per*) and *Timless* (*Tim*), but the transcription of other clock genes (DARLINGTON *et al.* 1998; GLOSSOP *et al.* 2003; LIM *et al.* 2007). In this way, Per and Tim provide the negative feedback that makes circadian oscillation possible. It was recently discovered that the Tim-Per complex disintegrates after six hours. Per rather than the Per-Tim complex is then thought to depress the transcription factor activity of Clk-Cyc complex (MEYER *et al.* 2006).

There is a weak association between *per* and diapause. In *D. melanogaster*, both diapause frequency and *per* are latitudinal clinal. The *Thr-Gly* repeat length in *per* is clinal in North-Africa and Europe (COSTA *et al.* 1992; SAWYER *et al.* 1997). There are two major allele classes, a long allele $(TG)_{20}$ and a short allele $(TG)_{17}$ (SAWYER *et al.* 1997). The short allele generates a more precise 24 hour clock at 29°C, but can be greatly affected by temperature. At lower temperature, this allele leads to a clock much shorter than 24 hours. The long allele can generate a more temperature independent clock, but is less precise at 29°C than the shorter allele. This makes the long allele favored and more common in the north. This also makes it possible for *per* to act as a modifier of the reproductive diapause pathway by providing a more precise 24 hour clock under diapause inducing temperature. However, to our current knowledge, it is unlikely that the critical day length of diapause *D. melanogaster* has a latitudinal cline. In fact, in many strains of *D. melanogaster*, diapause is not a photoperiodic response. On the other hand, *D. littoralis* has a strong latitudinal cline of diapause critical day length but there are no coding difference in this *Thr-Gly* repeat region (LANKINEN 1986; LANKINEN and FORSMAN 2006). It appears that the cline of $(TG)_{20}$ and $(TG)_{17}$ does not causally affect the diapause cline and the cline in *per* alleles is not driven by the selection on diapause. Additionally, *per* mutants are unable to affect

either the ability to enter reproductive diapause or the photoperiodism of diapause (SAUNDERS 1990; SAUNDERS *et al.* 1989). However, the mutants can slightly alter the critical day length. In Saunders *et al.* (1989), the authors investigated the roles of *per* by using three different *per* mutations and *per* deletion heterozygotes. The photoperiod-diapause response curves of these mutants were compared to *Canton-S*. The *per* mutations in this study were all isolated from per^{O} allele (O, not 0, as abolished), a loss of function mutation generated by ethyl methanesulfonate mutagenesis (KONOPKA and BENZER 1971) and they all have nonsense mutations at codon 464 truncating *per* mRNA (YU *et al.* 1987). The *per* double deletion heterozygote will produce no *per* mRNA. All mutations and the double deletion were known to be arrhythmic for behavioral circadian rhythms. However, all the tested genotypes can still enter diapause under short day condition, like *Canton-S*; an indication that the circadian clock is not necessary for diapause induction. In fact, the critical day length might be two hours shorter than *Canton-S*, indicating that *per* still may have certain effect on diapause photoperiodism but such effects are independent of circadian clock. Further studies confirmed this view (SAUNDERS 1990). Therefore, it is unlikely that *per* is causally involved in diapause induction.

It has been reported that a single nucleotide insertion in gene *Timeless* (the product of which forms hetero-dimer with *per*) affects reproductive diapause in Europe and Near East populations of *D. melanogaster* (TAUBER *et al.* 2007). A single nucleotide insertion generates an alternative start codon slightly upstream of the ancestral allele. As a consequence, the *ls-tim* allele produces two isoforms, a longer form L-Tim_{1421} and short form S-Tim_{1398} (SANDRELLI *et al.* 2007). In every natural population surveyed, the *ls-tim* allele possesses a high diapause proportion regardless of photoperiod (TAUBER *et al.* 2007). This may be the result of a difference of stability in Tim protein (SANDRELLI *et al.* 2007). L-Tim protein is more stable than S-Tim due to reduced responsiveness to light. Tim is produced mainly in the dark and both Per and Tim protein levels peak at late night (HARDIN *et al.* 1990). In daylight, Tim level drops via degradation dependent on the photoreceptor Cry. Cry will bind to Tim or Tim complex (Tim-Per-Dbt in cytoplasm) under stimulation of light (CASHMORE *et al.* 1999; CERIANI *et al.* 1999). The existence of a Cry-Tim interaction promotes tyrosine phosphorylation and degradation of Tim (NAIDOO *et al.* 1999). This will reduce the formation of Tim-Per-Dbt complex. Since the Tim-Per-Dbt complex was thought to be required to remove the Clk-Cyc complex from the E-box (DARLINGTON *et al.* 1998), the negative feedback control by Per on the *tim* and *per* loci will be weakened. The authors found L-Tim has reduced binding affinity to Cry, but both isoforms have the same affinity to Per (SANDRELLI *et al.* 2007). This means L-Tim will render the Tim-Per feedback loop less sensitive because it will more slowly degrade in

the presence of light than S-Tim. When African *D. melanogaster* was first subjected to temperate environments, it experienced longer day lengths in the summer, which in tropical regions the day length is more constant all year around. Therefore the reduced circadian cycle light responsiveness is thought to be adaptive to the long summer days. This is suggested to be the cause of the signature of positive selection observed for the *ls-tim* allele (TAUBER *et al.* 2007). In addition, the authors suggested that reduction in Tim-Per feedback results in increase in diapause induction. The null allele tim^{01}, which should completely lack Tim-Per feedback, possesses a 10~20% higher proportion of diapause than the *ls-tim* wild-type the introgressed in *Canton-S* background (TAUBER *et al.* 2007). Finally, using transformants of *p[s-tim]*, *p[l-tim]* and *p[ls-tim]* in an otherwise *tim* null genetic background, the authors reported a similar finding; that the *p[s-tim]* transformant has a weaker diapause response (TAUBER *et al.* 2007). Based on those two results, it seems that the diapause enhancing effect is due to the *tim* locus rather genetic the background. The authors suggested that by reducing photo-responsiveness, *ls-tim* females express a stronger diapause phenotype (TAUBER *et al.* 2007). Exactly how this happens is unknown.

However, there is also strong evidence of genetic background effects. In natural populations, *ls-tim* females show a 10~20 percent higher proportion of diapause (TAUBER *et al.* 2007). There is a similar difference between *tim-null* and *ls-tim*. In both cases, the difference is independent of photoperiod; diapause-photoperiod relationships doesn't depend on a critical day length. However, the phenotypic differences among transformants *p[s-tim]*, *p[l-tim]* and *p[ls-tim]*, though significant, are much less than that between *s-tim* and *ls-tim* females and the diapause-photoperiod response shows a dependency of critical day length (TAUBER *et al.* 2007). The photoperiod dependency is similar to that of *Canton-S* reported by Saunders *et al.* (1989); showing critical daylength of 14 hours. These observations imply that first, other loci in the genome account for the difference in shape of the diapause-photoperiod response curve and consequently, the critical daylength for diapause. Second, alleles at other loci co-segregating with alleles at the *tim* locus, can magnify the difference between *ls-tim* and *s-tim*.

The cline at the *tim* locus generally shows a counter-intuitive decline of *ls-tim* frequency with latitude. The *ls-tim* allele, due to its stronger diapause phenotype, is expected to be favored in the north, where the overwintering requirement makes diapause critical. We do not know whether diapausing frequency in Europe also exhibits a latitudinal cline as in North America (SCHMIDT *et al.* 2005a). However, comparing the Bitetto, Salice and Houten populations, it is interesting to note that Salice has the highest diapause frequency and the north-most

population, Houten, falls between Bitetto and Salice (Figure 1). Between Bitetto and Salice, there is a 20% difference in diapause frequency, although they are only 123km apart. It is also interesting to note that both *ls-tim* and *s-tim* females from the Houten population tend to diapause more strongly than southern counterparts. This implies that other diapause enhancing loci may be acting independently of *tim* and alleles at these loci are likely to have high frequencies in the northern populations.

Tauber *et al.* (2007) also suggested that under closer examination the *tim* cline is not latitudinal, but rather a isolation-by-distance cline. *ls-tim* is a recently derived allele that is absent in African populations and shows a signature of directional selection (TAUBER *et al.* 2007). It appears that the *ls-tim* allele originated in Europe shortly postdated the colonization, has been under directional selection, and subsequently has reached a high frequency in southern populations. The migration out of southern populations has possibly generated an overland distance cline. With no information about the historical migration of *D. melanogaster* in southern Europe or a selectively neutral cline to compare with, it is unknown whether the *tim* cline is indistinguishable from a neutral expectation or how this cline is the consequence of selection rather than population structure. However, from Tauber *et al.* (2007), it seems that the frequency of the *ls-tim* allele drops at a similar rate over distance in both southern and northern directions, which implies that this locus is not under selection by overwintering stress. If the *ls-tim* allele is selected for diapause and consequently overwintering ability, one should see its frequency dropping more slowly north of Italy (where it has the highest frequency) than south rather than at the same rate. It is possible that the *ls-tim* allele has reached a high frequency not due to selection for diapause, but rather, due to the selection for lower light sensitivity of the circadian cycle as suggested in Tauber *et al.* (2007). However, certain difficulties still persist in this theory. If reduced light sensitivity is an adaptation for an African species recently colonizing the temperate zone, this adaptation will be surely due to the around-the-year variation of day length in the new environment, especially the increase of daylength in the summer. If this is true and since daylength increase will be much larger in the north, one should expect the *ls-tim* allele, which has less light sensitivity, to be favored in the northern region. This would be reflected in a latitudinal rather than distance cline.

Even if an isolation-by-distance model provides the best statistical explanation for geographic variation in *tim* (cline under this model has the highest regression regression coefficient), it is questionable whether the model is biological meaningful. As shown in Figure 1, the population samples for this study were collected all across Europe and they are separated by potential

geographic barriers. If the *tim* cline is generated by migration one does not expect to see a distance cline because of migration barriers. The rationale for choosing an overland distance model is used to explain the frequency of diapause in the Crete and Haifa populations, which are both separated from the European mainland. In short, it is more likely that the *tim* cline is a negative-latitude correlated and both the Crete and Haifa are not southern extensions of the same cline. It is also worth mentioning that *D. melanogaster* urban populations can overwinter by finding refuge in human dwellings and show a lower diapause frequency than rural population (SCHMIDT and CONDE 2006). Some populations in Tauber *et al.* (2007) may have been collected in urban while others from rural populations (TAUBER *et al.* 2007 Table S1). Therefore, it is possible that some of the geographic pattern of *ls-tim* is due to sampling issue.

The relationship between reproductive diapause and the clock genes *per* and *tim*, if any, seems to be a weak one. Both genes slow geographic variation, but these are unlikely to be caused by selection for diapause. Both genes are essential for circadian cycles, but the null mutants, despite being arrhythmic, can still diapause and can even respond to critical inducing photoperiods. This is puzzling since these two proteins are the centerpieces of circadian cycles and downstream genes. If diapause is a photoperiod-dependent trait, how can *D. melanogaster* measure photoperiod without a running clock? There are several alternative explanations. First, reproductive diapause is mainly controlled by temperature and only, at best, weakly responds to photoperiod. Second, clock genes other than *per* and *tim* link circadian rhythms and photoperiod. Third, *per* and *tim* are involved in the diapause pathway as modifiers, affecting the strength of the diapause phenotype.

It is worth mentioning that diapause does not involve a photoperiod measure, but a measurement of dark hours or night length instead. Work in *D. triauraria* support this hypothesis (YOSHIDA and KIMURA 1993; YOSHIDA and KIMURA 1994b). This species is a member of *D. melanogaster* species group. Unlike the relatively shallow diapause in *D. melanogaster*, the strains of *D. triauraria* collected from north Japan firmly diapause. Whether the same mechanism exists in *D. melanogaster* is currently unknown. Even though *D. melanogaster* and *D. triauaria* belong to the same species group, there is at least 41.3 MYA divergence between them (TAMURA *et al.* 2004) and it is possible that diapause has different genetic architectures in the two species. Nevertheless, this alternative model provides us with the insight of studying a complete different diapause mechanism besides the circadian cycle and photoperiodism.

In summary, reproductive diapause in *D. melanogaster* is under both

temperature and photoperiod control. There is little doubt about the importance of temperature, while the significance of critical day length is unclear. One issue is that the degree of photoperiod dependency might be variable in different natural populations. It is also possible that diapause is induced independent of the circadian cycle, but by a measurement of night length alone, which can be temperature dependent and involving *tim*, but not *per*. However, we have limited knowledge of the mechanism of either photoperiod/night-length dependence or the link between this and the endocrine regulation of reproductive diapause.

The Role of Endocrine in Reproductive Diapause

It appears that a central part of reproductive diapause regulation is the insulin signaling pathway. During reproductive diapause, oogenesis is halted in pre-vitellogenesis before stage 8 and the aging process is also retarded (TATAR *et al.* 2001a). This draws attention to the insulin signaling pathway, which plays an important role in both reproduction and aging (TATAR *et al.* 2003). Diapausing individuals are more stress resistant, which is another biological response linked to insulin signaling (TATAR *et al.* 2003). Certainly, delayed reproduction and retarded aging in the diapause state might be related in different ways than commonly controlled by the insulin signaling pathway. Halting reproduction can have a direct affect on aging. For example, it has been shown that abating the germline can extend life span, while abating the gonad somatic cells will not, which suggests gonad somatic cells can produce aging-resistant signaling when reproduction is reduced (FLATT *et al.* 2008; PATEL *et al.* 2002). Similarity, stress resistance and longevity are also related (PATEL *et al.* 2002).

Based on several lines of evidences, in *D. melanogaster*, reproductive diapause appears to be caused by downregulation of insulin signaling pathway via juvenile hormone (JH) and ecdysone (20E). Insulin receptor (*InR*) mutants are dwarf, stress resistant, long-lived and have reduced ovary development; similar to the diapause phenotype (TATAR *et al.* 2001b). However, losing the insulin receptor substrate protein (CHICO) results in life span extension without reducing oogenesis or enhancing resistance (CLANCY *et al.* 2001). The phenotype of *InR* mutants can be restored to normal by applying JH (TATAR *et al.* 2001b). It was shown earlier that applying JH hormone breaks diapause (SAUNDERS *et al.* 1990). Altogether, these observations suggest that diapause is caused by reduced JH levels, which in turn are caused by downregulation of insulin signaling. JH is also involved in enhancing oocyte maturation and production of yolk proteins. It also has been shown, that upon mating sex peptide in seminal fluid can stimulate JH synthesis, which is consistent with the observation that mated females are harder to diapause (MOSHITZKY *et al.* 1996).

However, it is possible that some of these effects of JH on reproduction may be indirect. For example, vitellogenesis can still occur without JH, but never without 20-Hydroxyecdysone (20E). Applying 20E can also break diapause and may be more efficient than JH treatment (RICHARD *et al.* 2001; RICHARD *et al.* 1998). Diapause in *Canton-S* can be terminated by simple transfer the flies from diapausing temperature to 25°C, during which all animals start vitellogenesis 8-12 hours after transfer (RICHARD *et al.* 1998). During this process, 20E level increases (RICHARD *et al.* 1998). These observations imply that diapause might

also be directly controlled by the level of 20E. However, mated females have reduced levels of ecdysteroids compared to virgin females, and females mated with accessory gland ablated males have even lower levels (HARSHMAN *et al.* 1999). This appears contradictory since mated females are less likely to diapause.

High levels of JH and 20E both appear to inhibit diapause. However, the regulation of JH and 20E is complex and not fully understood (GRUNTENKO and RAUSCHENBACH 2008). It is also possibly that the hormones have different regulatory functions in the different stages of oogenesis. Increasing 20E level leads to apoptosis of developing oocytes of stage 8 and 9 by directly regulating early ecdysone response genes, such as BR-C (TERASHIMA and BOWNES 2004; TERASHIMA and BOWNES 2005). This process is involved in adjusting egg production according to available nutrients (TERASHIMA and BOWNES 2004). When females are subjected to restricted food, insulin signaling is downregulated (PARTRIDGE *et al.* 2005), which JH level reduces. JH is shown to suppress 20E induced apoptosis of the stage 9 egg chamber (TERASHIMA and BOWNES 2005). It is unclear, however, whether this is due to JH direct regulation of E20 level or JH's blocking apoptosis indirectly.

The apoptosis of the stage 9 egg chambers under high 20E level is sometimes referred to as oosorption (BELL and BOHM 1975). It is important to recognize the difference between reproductive diapause and oosorption. Diapause is development arrest in pre-vitellogenic stage and oosorption is apoptosis of egg chambers in stage 9, which is an early vitellogenic stage. The effects of 20E and JH on these two processes are also different. It is suggested that cell death in stage 8 is dependent on a mid-oogenesis checkpoint mechanism (reviewed in MCCALL 2004). At present, we don't know if diapause is also due to apoptosis of egg chambers but depends on a different development check point. In fact, it is not clear whether apoptosis is involved at all in reproductive diapause. This is certainly a direction for further research.

It is also important to recognize that the regulation of JH and 20E are dependent on the age of the fly and the oocyte development stage. In the larval stage, ecdysteroids are produced in prothoracic cells and ring gland (DELBECQUE *et al.* 1990). In adult females, ecdysteroids are mainly produced by follicle cells (TU *et al.* 2002). As insulin signaling is able to affect ovary development, it is expected that the ecdysteroids production of *InR* mutations to be lowered, at least during a short post-eclosion period. However, experiments show that the post-eclosion ecdysteroids pulse occurs in both *InR* wide-type and females heterozygous for *InR* mutations, but at 24 hour post-eclosion in the *InR* mutant heterozygotes compared to 18 hours for wide-type. Females homozygous for *InR*

mutations lack this pulse (TU *et al.* 2002). This might be of importance since it is known that diapause is inducible in young virgin females and the inducibility declines with age. Immediate post-eclosion hormonal control might play an important role. The reason that young virgin females are sensitive to diapause inducing conditions may lie in the specific hormone pool and interactions at a specific development stage. The induction of diapause might require this particular endocrine environment.

Diapause may require downregulation of insulin signaling in certain neuronal cells (WILLIAMS *et al.* 2006). Just as it was observed in *InR* mutants, the diapause enhancing effect of the down-regulation of insulin signaling has been recently observed for *dp110*, which encodes the catalytic subunit of *Drosophila* type I_A PI3K (LEEVERS *et al.* 1996). In Williams *et al.* 2006, the authors used a GAL4-UAS drive system to over-express *dp110* in different tissues in a strain capable of diapause under short day-low temperature conditions. Under inducing conditions, over-expression by using a pan-neuronal GAL4 drive prevents this strain from entering diapause but over-expression using a visual system drive will not. This result implies that diapause requires downregulation of insulin signaling in certain neuronal cells. Though a visual system presumably is responsible for diapause by providing a sensor for diapause cues, its role is more likely to be upstream than downstream of insulin signaling.

In *D. melanogaster*, type I_A PI3K is a hetero-dimer consisting of Dp110 and Dp60, which is the regulatory subunit similar to p85 or p55 in humans (ALBERT *et al.* 1997; WEINKOVE *et al.* 1997). Dp60 is also known as PI3K21B (due to its physical position in cytological band 21B on the second chromosome). It is the regulatory subunit since it has two SH2 phosphopeptide affinity domains. Insulin activates PI3K possibly though a SH2-phosphopeptide interaction of the regulatory domain of Dp60 and subsequently recruiting Dp110. Dp110 generates phosphatidylinositol 3,4,5-trisphosphate (PtdIns(3,4,5)P_3) and PtdIns are important precursors to a number of messenger molecules. In this way, Dp110 is directly or indirectly involved in many biological processes via Ptdlns. However, null mutations of Dp110 and Dp60 are phenotypicly similar, except the Dp110 mutation has stronger phenotype, which indicates that Dp110 can still fulfill at least part of its function in the absence of Dp60 (WEINKOVE *et al.* 1999). Over expressing *dp110*, for these reasons, should still provide a valid way to study the consequences of the effects of up-regulating insulin signaling to type IA PI3K complex. The implication of dp110 in diapause expression in natural population is elusive since no coding variation is found in *dp110* (WILLIAMS *et al.* 2006). Investigating the natural molecular variation of PI3K21B may become an interesting future direction.

Aside from the lack of variation, it is questionable whether *dp110* is the cause of phenotypic diapause variation. Under close examination, the role of *dp110* is not supported by genotypic differences in the *dp110* locus between the W and C lines used in the mapping. The W and C lines are isofemale lines collected from Windsor, ON and Cartersville, GA, respectively. When newly enclosed females are treated in a 12°C 10L:14D inducing condition, the W and C lines have ~62% and ~18% diapause frequency respectively (WILLIAMS and SOKOLOWSKI 1993). It was suggested that the difference in diapause frequency is due to a single recessive autosomal factor. The segregation ratios of the F2 crosses and F1 backcross to the recessive parent (W line) were not significantly different from 3:1 and 1:1, respectively (WILLIAMS and SOKOLOWSKI 1993). Williams *et al.* (2006) claim that QTL mapping experiments determine that the recessive factor is located within the span of deficiency *Df(3R)H-B79* on the third chromosome. The diapause difference between W and C lines is larger when heterozygous over deficiencies of *dp110* than over the same deficiencies with a *P*-element insertion containing a *dp110* rescue fragment. This genetic compensation is consistent with the hypothesis that diapause is controlled by a single recessive allele on the third chromosome and that diapause is recessive. It is also consistent with the hypothesis that diapause is due to the downregulation of the insulin signaling pathway. However, this experiment lacks a proper control. Since both *dp110* deficiencies are generated by imperfect excision of P-element insertion H^{D179} (WEINKOVE *et al.* 1999), the strain which H^{D179} was placed into should be an best control. If not available, a precise excision of H^{D179} would also serve the purpose. In either case, this experiment can be better controlled than using the balance chromosome (*TM3, Ser, y+*). Without a proper control, one can only claim that adding an extra copy of *dp110* using *P*-element transgenic constructions reduces the phenotypic difference between W and C lines rather than that diapause variation between these two lines is due to differences in *dp110* expression.

It is also possible that other genes within deficiency *Df(3R)H-B79* are causally related to the diapause difference between W and C lines. It is also unclear in Williams *et al.* (2006) how the candidate gene was narrowed down to *dp110* in this 979 kb genomic region with over 100 genes. This conclusion is further complicated by the observation that between W and C lines there are no coding differences in *dp110*. If *dp110* is the cause of the phenotypic difference, it is possibly due to a transcriptional difference. However, the mRNA levels of W and C strains are reported to be similar under diapausing condition. Since neither the expression nor the sequences of *dp110* differ between these two lines, it is more appropriate to conclude that *dp110* is probably involved in diapause but is not the gene causing difference between W and C lines. Further work is required

to identify whether other genes in the span of deficiency *Df(3R)H-B79* have causal effects on reproductive diapause. The identification of *dp110* as the underlying gene, to some degree, may reflect an *a priori* expectation because of its connection to insulin signaling.

A relationship between insulin signaling pathway and diapause has also been demonstrated in the mosquito (SIM and DENLINGER 2008). In *Culex pipiens*, downregulating insulin signaling pathway by knocking down the expression of *InR* by using RNAi induces females to diapause under non-diapausing conditions. When this diapause is induced by RNAi, it can be broken by JH treatment. It is believed that other aspects of diapause such as stress resistance and energy storage are caused by a downstream signal transduction of the fork head transcription factor FOXO. In *C. pipiens*, FOXO is activated when insulin signaling is shut down (SIM and DENLINGER 2008). In general, when insulin is present, FOXO are phosphorylated by Akt and Skg, preventing its transport into nucleus (reviewed in VAN DER HEIDE *et al.* 2004). When insulin is removed, FOXO is transported into the nucleus and promotes transcription of target genes. Since FOXO RNAi in diapausing females reduces survival and fat storage, the authors suggest that FOXO is responsible for the extended life span and evaluated energy storage in diapause state. However whether FOXO RNAi treatment can restore the ovary development is not known and it is also unknown if FOXO is a downstream or upstream of JH. The authors suggested that JH synthesis and FOXO regulation are independently controlled by *InR* in such a way that JH controls ovarian development and FOXO independently controls energy storage and stress resistance (and longevity). FOXO is known to be modified by phosphorylation, acetylation and ubiquitination and different modifications can result in binding to different sets of partner proteins and turning on gene sets of different processes (reviewed in CALNAN and BRUNET 2008). Diapause may be only related to a small subset. Particularly, it will require further work to understand how FOXO is regulated in diapause and which gene(s) are turned on by FOXO.

In summary, our understanding of the regulation of diapause in *D. melanogaster* is still limited. We have some understanding of how insulin signaling, FOXO, JH and 20E are involved; however, those pathways affect a range of traits and some of them are phenotypically similar to diapause, yet fundamentally different. For example, reduced ovarian development is not only associated with diapause, it can also arise from inadequate nutrition. Therefore, it is necessary to determine whether the termination of reproduction observed in mutants of the insulin signaling pathway is due to diapause. We also need to be fully aware of the interspecific differences in diapause. Diapause in *D.*

melanogaster is shallow at best and one can't assume the mechanisms to be common to all insects. For example, reproductive diapause in *C. pipiens* is related to food condition, a blood meal prevents diapause, but a strict carbohydrate meal promotes it (ROBICH and DENLINGER 2005). In *D. melanogaster*, rearing the adults in the absence of protein will extend life span under dietary restriction (MIN and TATAR 2006), but it will not induce diapause. While *InR* mutants mimic diapause, *CHICO* mutants will extent life span without reproductive tradeoff. These contradictions require further research to understand the role of the insulin signaling pathway in reproductive diapause.

The Role of Gene *couch potato* in Reproductive Diapause

Schmidt *et al.* (2005) found that in *D. melanogaster* diapause usually appears dominant over nondiapause and is due to genetic factors on the third chromosome (SCHMIDT *et al.* 2005a). This contradicts early work indicating diapause is due to a recessive factor on the third chromosome (WILLIAMS and SOKOLOWSKI 1993). As reproductive diapause is most likely polygenic, these observations could be due to different underlying genes among lines. By using QTL mapping, this dominant factor was eventually mapped to the *couch potato* gene (SCHMIDT *et al.* 2008b). A Lys/Ile replacement has the strongest association with diapause/nondiapause in a natural population and its frequency is correlated with the diapause frequency in eastern US populations. I show in this dissertation that sequence variation in this gene shows a pattern consistent with recent balancing selection and in comparison with African *D. melanogaster* populations, diapause is a recent adaptation.

The *couch potato* gene was known for being required for normal behavior; it was identified by mutations associated with general sluggishness (BELLEN *et al.* 1992a; BELLEN *et al.* 1992b; GLASSCOCK and TANOUYE 2005). The Cpo protein has a RNA recognition motif (RRM), also found in a human gene called RNA-binding protein with multiple splicing (*RBPMS*) (SHIMAMOTO *et al.* 1996). It was recently reported that RBPMS of *Xenopus* physically interact with Smad proteins: Smad2, Smad3, and Smad4 and the presence of TGF-beta increases these binding affinities (SUN *et al.* 2006). Smad is a group of proteins involved in the TGF-beta signaling pathway by forming protein complexes and regulating gene expression by acting as transcription factors (MASSAGUE and WOTTON 2000). Their nucleus localization is promoted by phosphorylation of C-terminals. RBPMS is able to enhance this phosphorylation and upregulate downstream events. However, compared to human RBPMS, Cpo is considerable larger and contains several hundred additional N-terminal amino acids. This part is highly variable and possibly contains a coiled-coil motif, while the RRM domain is highly conserved. Besides encoding proteins isoforms that contain the RRM, the gene also encodes for one isoform without RRM. A tract of 21 amino acids, including the Lys/Ile polymorphism which is associated with diapause, is utilized only by this isoform. It is not known whether Cpo can be involved in transcription regulatory roles, similar to RBPMS, or whether the N-terminal amino acid domain is of functional importance.

The genetic mapping of reproductive diapause, which leads to the indentification of *cpo*, will be addressed in the second chapter. The population genetics study of *cpo*, which examines the signature of selection associated with

this locus, will be discussed in the third chapter.

Table 1.

Known diapausing species in the Diptera: Drosophilae. For phylogenetic comparison, a number of non-diapausing species are also included. It is likely that diapause arose independently in different lineages and diapausing ability is likely to be ancestral.

Notes:

a, Very limited overwintering activity, they migrate to warm regions during winter

b, Aestivo-hibernal reproductive diapause controlled by flight activity

c, Latitudinal cline of critical day length

d, Start in early in August/September, when temperatures are still high

e, Latitudinal cline of diapause frequency in natural populations

f, Sister species to *D. melanogaster*, non-diapausing

g, *D. auraria* species complex members of China and Japan. Diapause phenotype ranges from none, weak to strong dependents on the geographic origin of species. *D. quadraira* is a sister species to *D. triauraria*, yet lack diapause. Duration of diapause is longer the northern strains in *D. triauraria*.

h, Subtropical species, diapause inducible in the lab, probably not diapausing in nature due to absence of inducing condition.

i, Overwinter as adults, probably reproduce in winter

j, Diapause from November to March even if kept in constant lab condition.

k, Probably the same spices

Genus	Sub-genus		Species-group	Subgroup	Species	Diapuase	Note	Reference
Drosophila	Drosophila	quinaraia section	immigrans	curviceps	*D. curviceps*	NO		
				immigrans	*D. immigrans*	NO	*a*	(KIMURA and BEPPU 1993)
				nasuta	*D. albomicans*	NO		
			quiniaria	transversa	*D. transversa*	Reproductive diapause		(MUONA and LUMME 1981)
				-	*D. phalerata*	Reproductive diapause		
		virilis repleta radiation	melanica	-	*D. moriwakii*	Reproductive diapause	*b*	(ICHIJO *et al.* 1992)
			robusta	lacertosa	*D. lacertosa*	Reproductive diapause	*c*	(ICHIJO 1986)
				robusta	*D. robusta*	Reproductive diapause	*d*	(CARSON and STALKER 1948)
			virilis	-	*D. lummei*	Reproductive diapause	*c*	(LUMME and KERANEN 1978)
					D. littoralis	Reproductive diapause	*c*	(LUMME and OIKARINEN 1977)
		-	-	testacea	*D. testacea*	Reproductive diapause		(KIMURA 1982)
		-	-	polychaeta	*D. daruma*	Reproductive diapause		(HIRAI *et al.* 2000)
	Sophophora	melanogaster		melanogaster	*D. melanogaster*	Reproductive diapause	*e*	(SAUNDERS *et al.* 1989)
					D. simulans	NO	*f*	(SCHMIDT *et al.* 2005b)
				montium	*D. lutescens*	Reproductive diapause / Weak		
					D. auraria	Reproductive diapause		(OHTSU *et al.* 1992)
					D. subauraria	Reproductive diapause / Strong	*b, g*	(KIMURA *et al.* 1992)
					D. quadraria	NO		(KIMURA 1983)
					D. triauraria	Reproductive diapause		(MINAMI and KIMURA 1980)
					D. rufa	Reproductive diapause / Weak		
				takahashii	*D. takahashii*	Reproductive diapause / Weak	*h*	(KIMURA *et al.* 1994)
					D. lutescens	Reproductive diapause / Weak		
			obscura	obscura	*D. subsilvestris*	Pupal diapause		
					D. alpina	Pupal diapause		(GOTO *et al.* 1999)
					D. bifasciata	Reproductive diapause		
					D. tristis	Reproductive diapause		(BASDEN 1954)
					D. subobscura	NO		(GOTO *et al.* 1999)
				pseudoobscua	*D. pseudoobscura*	NO	*i*	(EPLING *et al.* 1955)
				subobscura	*D. guanche*	NO		(GOTO *et al.* 1999)
Hitodrosophila	-		quadrivittata	confusa	*D. confusa*	Reproductive diapause		(CHARLESWORTH and SHORROCKS 1980)
			melanderi	-	*D. cameraira*	Reproductive diapause		
Scaptodrosophila	-		victoria	-	*(D. nitens) Scaptodrosophila rufifron*	Reproductive diapause	*j*	(BERTANI 1947)
					(D. deflexa) (D. guyenoti) Scaptodrosophila deflexa	Larval diapause	*k*	(BASDEN 1954)
						Larval diapause		(BURLA 1951)
Chymomyza	Chymomyza		costata	-	*Chymomyza costata*	Larval diapause		(RIIHIMAA 1984)

Figure 1.

D. melanogaster populations investigated for *ls-tim* frequency by Tauber *et al.* (2007) (TAUBER *et al.* 2007).

A: Above, the geographic location of populations. Dark color of the physical map shows high elevations. Umea is located 4 degrees further north. Dash line represents latitude and the vertical fine dash line stands for Greenwich Meridian. These populations are potentially separated by mountain ranges and water bodies, which would confound either the latitudinal cline or over-land distance cline from the Novoli hypothesis. See detailed discussion in the text.

B: Lower right, the diapause frequency in response to photoperiod in three different populations. Dash lines; *s-sim* females, continuous lines; *ls-sim* females. Within the same population, *ls-tim* females are more likely to diapause than *s-tim* females. Among different populations, northern *ls-tim* and *s-tim* females may be slightly more strongly diapausing than their southern counterparts. However, this does not necessarily translate into that northern population is more diapausing than southern one because of the low frequency of *ls-tim* in Houten population. Lower left, expected diapause photoperiod. Mean proportions, together with ±acrsin(SEM) when applicable, are plotted in acrsin scale with proportion of diapause in percentage as reference on the right. Redrawn using Figure 3 and table S1 from Tauber *et al*. 2007. Map is of Mercator projection.

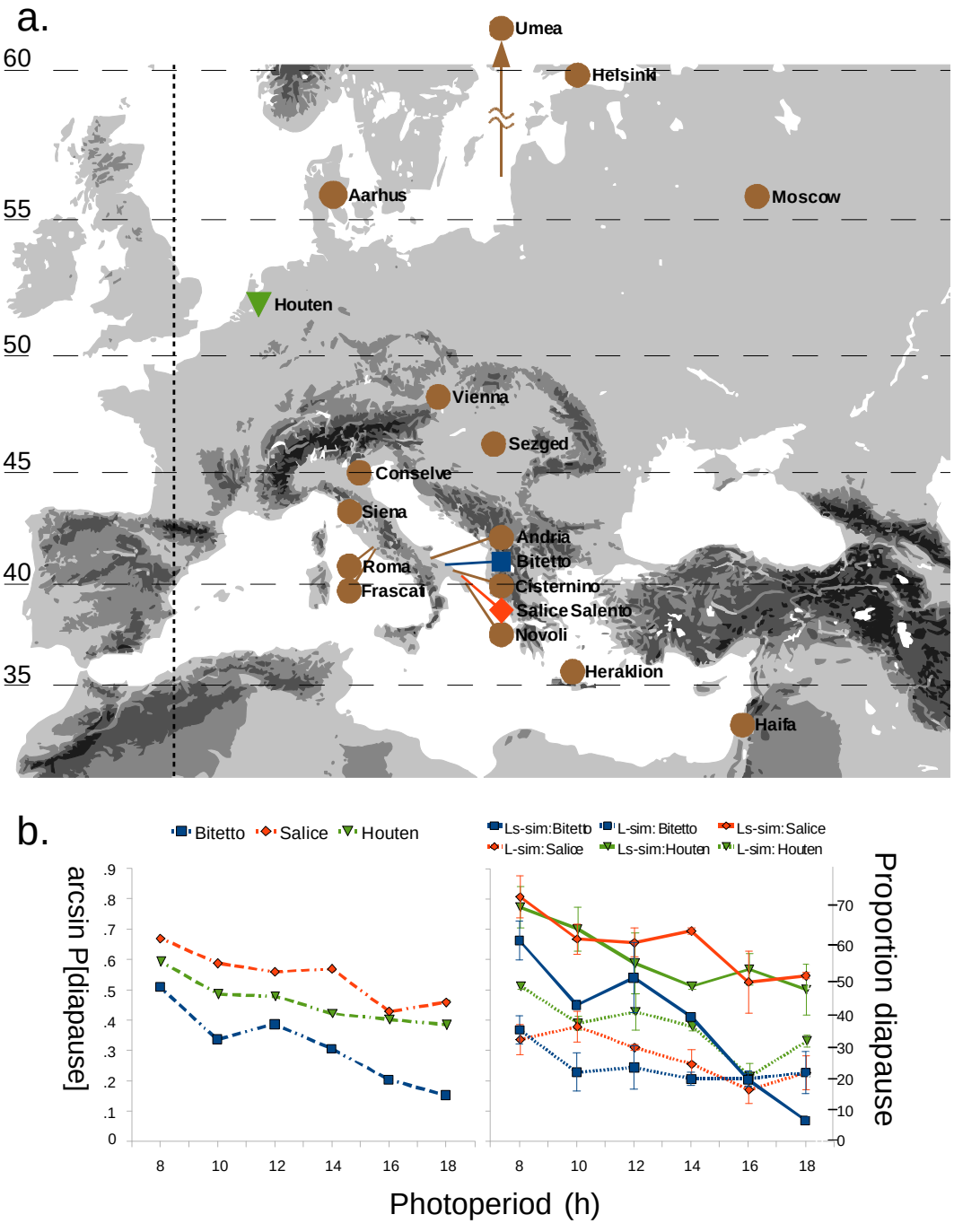
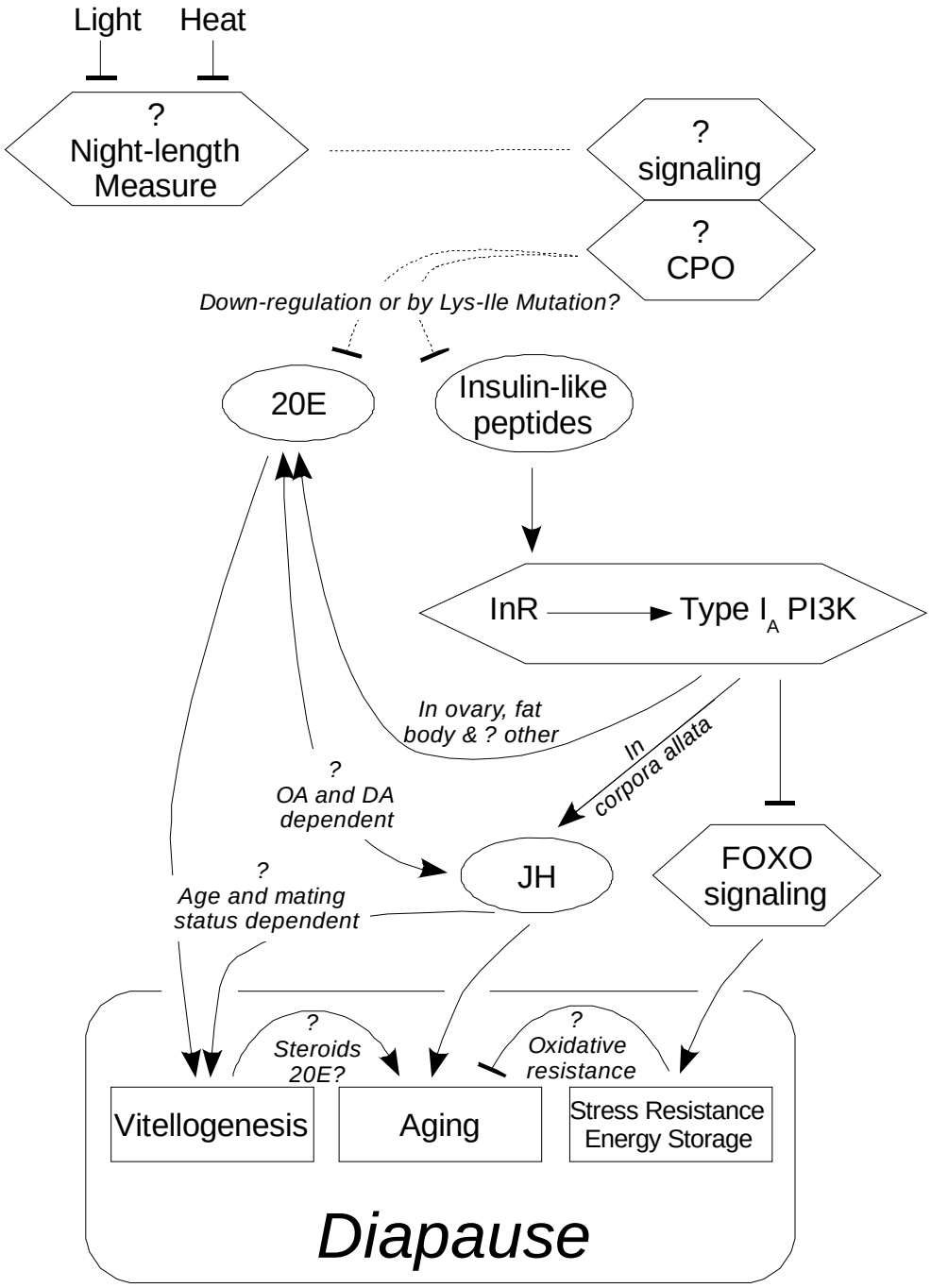

a.
60
55
50
45
40
35
Umea
Helsinki
Aarhus
Moscow
Houten
Vienna
Sezged
Conselve
Siena
Andria
Bitetto
Roma
Frascati
Cisternino
Salice Salento
Novoli
Heraklion
Haifa
b.
Bitetto
Salice
Houten
Ls-sim:Bitetto
L-sim: Bitetto
Ls-sim:Salice
L-sim:Salice
Ls-sim:Houten
L-sim: Houten
arcsin P[diapause]
Proportion diapause
Photoperiod (h)

Figure 2.

Diagram of major signal transductions proposed in reproductive diapause. Hexagon boxes: signaling transduction cascades. Oval boxes: hormonal molecules. Square boxes: phenotypic response. Lines: direction of regulation, with arrow ends indicate positive regulations and dimension ends indicate negative regulations. Dash lines indicate unknown interactions. Under diapause conditions, insulin signaling is downregulated. This downregulates the synthesis of JH (juvenile hormones) in *corpora allata* and 20E (20-Hydroxyecdysone) in ovary and fat body cells, both via *InR*-type IA PI3K complex and their downstream insulin signaling cascade. JH and 20E both promotes vitellogenesis and the balance between the two is probably maintained by OA (octopamine) and DA (dopamine). Downregulation of insulin signaling results in arrest of vitellogenesis. It will also slow aging via JH. FOXO signaling will be upregulated which leads to expression of stress resistance and energy storage related genes. We have no information about how *couch potato* involves in diapause.


Light
Heat
?
Night-length
Measure
?
signaling
?
CPO
Down-regulation or by Lys-Ile Mutation?
20E
Insulin-like
peptides
InR
Type I A PI3K
In ovary, fat
body & ? other
In
corpora allata
?
OA and DA
dependent
?
Age and mating
status dependent
JH
FOXO
signaling
?
Steroids
20E?
?
Oxidative
resistance
Vitellogenesis
Aging
Stress Resistance
Energy Storage
Diapause

Chapter two: The genetic mapping of reproductive diapause in *Drosophila melanogaster,* gene *couch potato* forms the basis of latitudinal cline of reproductive diapause.

Introduction

Natural populations encounter environmental stresses that diminish individual fitness, and these stresses are variable across space and time. It is predicted that the resulting natural selection often leads to a situation in which no genotype has the highest fitness across all environments, and polymorphism is maintained. This concept is pervasive in arguments about the evolution of life history variation and associated genetic tradeoffs (ZERA and HARSHMAN 2001). However, the expected molecular polymorphism associated with life history tradeoffs and adaptation remains elusive. Although we would expect these phenomena to be universal, the complexities of genetic dissection of such variation suggest that the best opportunity to identify the genetic basis for life history variation lies in the study of the genetic models in their natural populations (FEDER and MITCHELL-OLDS 2003).

Drosophila melanogaster is a human commensal that has spread from areas of Sub-Saharan Africa to Europe and Asia, possibly over the last 5,000 to 16,000 years, and into the Western Hemisphere and Australia in the past several hundred years (BAUDRY *et al.* 2006; DAVID and CAPY 1988; THORNTON and ANDOLFATTO 2006). This worldwide expansion from the tropics has required adaptation to the pronounced seasonality present in temperate habitats, and there are many examples of both single-gene polymorphism and quantitative trait variation that show geographic patterns associated with the transition from tropical to temperate climates in this species (BERRY and KREITMAN 1993; SEZGIN *et al.* 2004). There also is good evidence that D. melanogaster overwinters at the adult stage in temperate habitats (BOULETREAU-MERLE *et al.* 2003; MITROVSKI and HOFFMANN 2001), and that temperate populations do not merely reflect recurrent migration from more moderate climates (IVES 1954; IVES 1970). This overwintering survivorship clearly presents a variety of challenges, including the need for lifespan extension well beyond that typically measured in the laboratory, as well as increased physiological tolerance of extended exposure to suboptimal conditions (DENLINGER 2002).

The best-studied adaptation to seasonality in insects is the expression of a diapause syndrome. This phenotype is analogous, and potentially in part homologous, to the dauer stage in *Caenorhabditis elegans*; unlike the genes and pathways underlying dauer formation, however, very little is known about the

genetic basis of diapause in *Drosophila*. Although many aspects of insect diapause vary across taxa, the expression of diapause is associated with a suite of physiological changes that allow persistence during periods of stress exposure (DENLINGER 2002). Once thought absent in this tropical species, reproductive diapause occurs in D. melanogaster and is cued by exposure of adults to short days and low temperatures (SAUNDERS *et al.* 1989). The expression of this diapause is under neuroendocrine control (SAUNDERS and GILBERT 1990) and results in lifespan extension, delayed senescence, and increased stress resistance (TATAR *et al.* 2001a). The incidence of reproductive diapause exhibits a strong latitudinal cline in eastern North American populations (SCHMIDT *et al.* 2005a), varying from ≈30% in southern Florida to 90% in New England. Diapause incidence also varies predictably with the season (high incidence in spring, lower in fall) in temperate orchard populations in Pennsylvania and New Jersey (SCHMIDT and CONDE 2006). The spatial and temporal patterns of diapause variation appear to reflect a robust series of life history tradeoffs between genotypes with a high and a low propensity to express diapause (SCHMIDT and CONDE 2006; SCHMIDT and PAABY 2008; SCHMIDT *et al.* 2005b). These properties of reproductive diapause in D. melanogaster offer a unique opportunity to elucidate the genetic architecture of an important fitness trait with pleiotropic effects on life histories. Similarly, the global spread of a sophisticated genetic model organism presents us with a superior opportunity to study adaptation to novel environments that has emerged over a relatively short evolutionary period, possibly as recent as 2,000 generations in the New World.

Previous analyses in a set of isogenic *D. melanogaster* laboratory stocks and inbred lines had demonstrated that the genetic factors affecting diapause expression were entirely associated with the third chromosome (SCHMIDT *et al.* 2005a). In this report, we use third-chromosome recombinants in a standardized genetic background, genetic complementation, and linkage association in a natural population to identify the gene responsible for diapause variation in natural populations.

Results

A set of 15 third-chromosome SNP markers was used to initially map the reproductive diapause phenotype in 250 recombinant inbred lines (RILs) generated from recombination between the parental lines *w;6326;VT46* and *w;6326;6326*. These lines were selected because the inbred line *VT46* does not exhibit diapause, whereas the isogenic line *6326* expresses diapause (SCHMIDT *et al.* 2005a) and also forms a reference genome in other studies (PARKS *et al.* 2004). Using these markers, multi-interval mapping (LI *et al.* 2006) places a single identified quantitative trait locus (QTL) between markers at cytological band positions 90D1 and 92E8 (Figure 3). A subset of 20 RILs possessed informative recombination events within the identified interval; the placement of an additional eight SNP markers in these lines provided a higher-resolution map (Figure 4). This refined analysis eliminated a number of genes in the initial 248-kb region and identified a final candidate gene as couch potato (cpo).

To independently test the hypothesis that variation in diapause expression is associated with the *cpo* gene, we carried out four sets of genetic complementation studies using cpo P element and piggyBac transposon-derived constructs in a standardized genetic background. These analyses used the alleles $cpo^{BG02810}$ (BELLEN *et al.* 2004) and cpo^{P3} (the latter is a precise excision of the *P[w+GT]* element in $cpo^{BG02810}$), as well as the homozygous viable *cpo* hypomorphs cpo^{v3}, cpo^{cp1}, and cpo^{cp2} *(23)*. A second set of crosses used FLP-FRT site-specific recombination (PARKS *et al.* 2004) of *piggyBac* elements to create duplications and deletions of both the cpo gene region (90C6–90E1) and the 5' flanking region outside cpo (90B7–90C1). These studies all confirm that genetic modifications of the cpo gene alone cause pronounced and repeatable effects on diapause expression (Figure 8).

Since *cpo* dosage appears to influence diapause, we measured the cpo transcript levels in the progenitor *VT46* (RIL line 201) and *6326* (RIL line 107) lines. The results of the RT-PCR that includes *dp110* and *Gapdh* controls showed that *6326* third instar larvae and adults possess ≈15% of the level of cpo transcript as *VT46* (Figure 9).

We next determined whether diapause phenotype in natural populations was associated with molecular polymorphism within the *cpo* locus. The *cpo* gene spans 84 kb and encodes six discrete transcripts. The major coding region (exon five of the *cpo*-RA transcript) is 449 aa separated by 37.5 kb from a scattered 3' set of seven small exons encoding another 286 aa (depending on the transcript and annotation). Preliminary sequence analysis of exon five in a diverse set of lines

identified extensive silent and amino acid polymorphism, as well as two major haplotypes. The central portion of the exon also contains several short tracts of polyglutamine repeats that vary in copy number. We began our linkage disequilibrium association analysis by examining this major exon and its variation.

We sequenced 3.5 kb spanning exon five from 35 third chromosomes from a single natural population (Davis Peach Farm; DPF) that were placed into the standardized genetic background (*w;6326*) used throughout this study. The DPF third-chromosome lines were then assayed for diapause expression under the standard induction conditions. The association between diapause phenotype and each of 192 polymorphic sites over the 3.5-kb region is depicted in Figure 5. The only polymorphic sites that were significantly associated with diapause phenotype are located in the 3' end of *cpo* exon five, and these reflect the aforementioned major haplotypes. Two of these polymorphisms are nonconservative amino acid changes. Amino acid position 363 possesses an alanine-to-valine substitution; high-diapause lines tend to possess the valine codon that is derived with respect to other Drosophila species. However, the strongest association was observed for residue 472. High-diapause alleles, without exception, were characterized by the radical, charge-changing substitution of a lysine for isoleucine.

A series of nested regressions were used to evaluate the relative contribution of the four significantly associated nucleotide polymorphisms. When the nucleotide state at residue was taken into consideration, no single other site accounted for any additional variance in diapause phenotype in the sequenced population [site 2069: $F_{1,34} = 0.83$, $P > 0.368$; site 2145 (amino acid position 363): $F_{1,34} = 0.083$, $P > 0.775$; site 2163: $F_{1,34} = 1.53$, $P > 0.226$]. In contrast, adding residue 472 did explain a significant additional variance in regressions for all other SNP sites. Thus, although the linkage group in the 3' portion of exon five is significantly associated with diapause phenotype, this is driven exclusively by the polymorphic Lys/Ile change and disequilibrium between this and other polymorphic sites.

We have shown that diapause incidence varies with latitude in the eastern United States (SCHMIDT *et al.* 2005a). Therefore, any putative quantitative trait nucleotide for diapause would be predicted to also exhibit a latitudinal cline in allele frequency, and this should extend to any polymorphisms that are in significant linkage disequilibrium with the causal site or sites. To address this expectation, we used restriction enzymes to screen five *cpo* SNPs (in *cpo* exons 1, 5, and 6 that span ≈70 kb) from 11 natural populations across the latitudinal gradient in the eastern United States (Figure 6). Between exons 1, 5, and 6 there is

only weak linkage disequilibrium among the SNPs, and within exon five the strength of association among sites becomes nonsignificant at distances of 1–2 kb, which is typical of other studies in *D. melanogaster* (LANGLEY *et al.* 2000; LONG *et al.* 1998; MCDONALD and KREITMAN 1991; MIYASHITA *et al.* 1993). The SNP showing the strongest association with latitude is the BsiE1 site in exon five (R^2 = 0.95; F = 82.97, $P < 1.2 \times 10^{-5}$). This corresponds to the Ala/Val nonsynonymous polymorphism at residue 363, and the derived allele (Val) increases in frequency with latitude. This pattern is consistent with the observation that alleles increasing in frequency with spread of *D. melanogaster* into temperate regions are most often derived (SEZGIN *et al.* 2004). Diapause incidence was previously observed to segregate as an autosomal dominant and varied from 33–80% in these populations (SCHMIDT *et al.* 2005a). This relationship predicts an associated allele frequency cline in *cpo* of 0.18 to 0.55; the observed linear allele frequency cline of 0.21 to 0.53 is strikingly similar.

To further evaluate whether the two identified amino acid variants have differential effects on patterns of diapause expression, we used 12 cpo^{472Ile} and 12 cpo^{472Lys} alleles in two sets of complementation analyses (Figure 7). It should be noted that the only nucleotide position that was identical among allelic replicates was the candidate site (residue 472). The *cpo* hypomorph, $cpo^{BG02810}$, was associated with higher diapause expression than the wild-type revertant, cpo^{P3}, created by the precise excision of the P element in $cpo^{BG02810}$ (χ^2 = 31.93, $P <$ 0.0001). In this background, there was a pronounced difference in the level of diapause expression between cpo^{472Ile} and cpo^{472Lys} (χ2 = 74.51, $P < 0.0001$). The complementation of these alleles, as clearly indicated by a lack of a significant interaction term (χ^2 = 0.0071, $P > 0.93$), demonstrated that the magnitude of the difference between the two natural *cpo* alleles was similar across all genetic backgrounds. Identical results were obtained in complementation studies using the *cpo* region piggyBac-derived deletion ($cpo^{3.42}$) and duplication ($cpo^{3.20}$). Again, the manipulation of the *cpo* locus predictably affected diapause expression (χ^2 = 64.18, $P < 0.0001$). Patterns of diapause expression were distinct between cpo^{472Ile} and cpo^{472Lys} alleles (χ^2 = 61.73, $P < 0.0001$), and this difference was maintained in both the *cpo* deletion and duplication backgrounds (interaction term χ^2 = 0.93, $P > 0.76$).

Discussion

Our study shows that the variation in reproductive diapause expression is associated with the *cpo* gene and, moreover, a single linkage group in the 3' end of *cpo* exon five. The most likely cause is an isoleucine-to-lysine mutation in residue 472 that also shows a latitudinal cline consistent with that previously observed for diapause incidence. This does not rule out other genes contributing to the variation in diapause in natural populations, as these analyses were conducted in a standardized genetic background where the second and X chromosomes were isogenic. The manipulations of levels of *cpo* expression by hypomorphic alleles and gene duplications/deletions support the hypothesis that variation for diapause phenotype in *D. melanogaster* is associated with *cpo* expression level. This hypothesis was further supported by initial RT-PCR expression analyses of the parental lines used to map diapause to *cpo*. Given these results, it is interesting that the candidate site is an amino acid polymorphism in the coding region of a single transcript: the change in primary protein sequence may be of direct functional significance and/or affect some aspect of dosage. At present, there are no data to test these hypotheses.

The *couch potato* gene was first identified in a screen for genes expressed in sensory organ precursor cells during peripheral nervous development (BELLEN *et al.* 1992b). It was shown subsequently to produce a nuclear protein and encodes an RNA-binding domain that is expressed in the peripheral and central nervous system of embryos, larvae, and adults, and in such tissues as the midgut, glia, and salivary glands (BELLEN *et al.* 1992a). Also, *cpo* is highly expressed in the ring gland (HARVIE *et al.* 1998), the primary endocrine structure in *Drosophila*; this is particularly interesting because diapause is under neuroendocrine control. We also identified a series of *cpo* ecdysone response elements, suggesting that the effects of *cpo* on diapause in *D. melanogaster* may be mediated by ecdysteroids (RICHARD *et al.* 2001). Loss-of-function *cpo* mutations are lethal, and partial loss-of-function mutations generate a variety of behavioral phenotypes (BELLEN *et al.* 1992b) and neurological abnormalities (GLASSCOCK and TANOUYE 2005). The three predicted ORFs contain a nuclear localization sequence, polyglutamine OPA repeat regions, and an RRM-RNA recognition motif in five of the six transcripts (BELLEN *et al.* 1992a). We see that the RRM region is very highly conserved across the 12 sequenced *Drosophila* genomes, but the OPA repeats show extensive copy number variation across taxa. Of particular note is that this alternate splicing predicts that the Ile/Lys amino acid polymorphism in residue 472 lies in the only protein lacking the RNA-binding domain. Our RT-PCR was not designed to target this transcript specifically.

Two recent studies have also investigated the genetic architecture for the reproductive diapause trait in *D. melanogaster*. Williams *et al.* (WILLIAMS *et al.* 2006) linked natural variation in reproductive diapause to the insulin-regulated phosphatidylinositol 3-kinase gene, *dp110*, also found close to *cpo* on the third chromosome at band 92F3. This association with diapause was further supported by manipulations of *dp110* using transgene constructions. However, no naturally occurring amino acid sequence variation was observed in the *dp110* gene, and no significant differences in *dp110* expression were detected between the high- and low-diapause lines. In *C. elegans* the homologue of dp110, age-1, is associated with dauer formation, and this concurrence suggests a common role for this gene in countering environmental stress across species. The identification of *dp110* as a diapause gene also suggests that aspects of diapause in *Drosophila* may be regulated by insulin signaling, and that *C. elegans* dauer and *D. melanogaster* diapause may be more than analogous phenotypes (TATAR and YIN 2001). Insulin signaling also has been shown to affect diapause expression in the mosquito *Culex pipiens* (SIM and DENLINGER 2008).

Diapause induction was also studied in association with allelic variation in the gene *timeless* (*tim*) (SANDRELLI *et al.* 2007), which encodes a light-responsive component of the circadian clock (TAUBER *et al.* 2007). As diapause expression is dependent on photoperiod, the authors hypothesized that this gene might impact reproductive diapause. In particular, a newly derived allele, *ls-tim*, increases in frequency across Europe, raising a possible association between temperate habitats, diapause incidence, and *tim* allele frequencies. Tauber *et al.* (TAUBER *et al.* 2007) showed that in three populations there was a significant relationship between the proportion expressing diapause, the testing photoperiod, and homozygous genotypes for the timeless alleles *s-tim* and *ls-tim*. It will be interesting to determine whether a similar association exists for the timeless alleles and diapause in North American populations, and in particular whether there is a latitudinal cline, as would be predicted.

In *D. melanogaster* the genetic variance associated with reproductive diapause has been shown to have profound pleiotropic effects on a number of other fitness-related phenotypes. These include lifespan, rates of senescence, fecundity profiles, development time, lipid content, and resistance to a variety of stressors (SCHMIDT and CONDE 2006; SCHMIDT *et al.* 2005a; SCHMIDT *et al.* 2008b). The absence of diapause induction in *Drosophila simulans* and African populations of *D. melanogaster* (SCHMIDT *et al.* 2005b) suggests that this trait is of recent evolutionary origin (SAUNDERS and GILBERT 1990) or is very rare in these populations. Our analysis identified a single gene and single nucleotide polymorphism that explains the observed variance in diapause expression in

natural populations. However, it is likely that the diapause trait is polygenic in *D. melanogaster*, and that more genes will be discovered that modify its expression and variation. Nevertheless, the combined quantitative mapping analysis, complementation studies, detailed evaluation of molecular variation, and the predicted association of geographic variation are compelling in singling out *cpo*. A number of important questions remain. These concern the function of *cpo*, and in particular that of the splicing product *cpo*-RH that replaces the putative RNA-binding domain with a highly basic lysine/arginine rich terminus of 41 aa. Whether this product in turn indirectly impacts many downstream genes or acts more directly remains to be deciphered to explain the many pleiotropic effects that variation in this gene has on life history traits.

Materials and Methods

Stocks.

All lines and cpo alleles were placed into a common genetic background of *w;6326* using marker-assisted introgression (MERRITT *et al.* 2006). The complementation analyses used $cpo^{BG02810}$ (BELLEN *et al.* 2004) and cpo^{P3} (a precise excision of the *P[w+GT]* element in $cpo^{BG02810}$). Mobilization of the P element used standard crosses to a male stock carrying a transposase source; excisions were recovered in males, extracted, and the background replaced. Excision was confirmed by PCR fragment analysis and sequencing of the gene region. The other cpo hypomorphs used (cpo^{v3}, cpo^{cp1}, and cpo^{cp2}) have been described (BELLEN *et al.* 1992b). Line *w;6326;VT46* (low-diapause parental genotype) is derived from an inbred line, *VT46*, collected in 1997 from Whiting, Vt (SCHMIDT *et al.* 2005a). Line *w;6326;6326* (high-diapause parental genotype) is a derivative of Bloomington line *6326* with the white-marked chromosome from Bloomington stock 2475.

To create site-specific duplications and deletions across the cpo gene region, piggyBac FLP-FRT-facilitated recombination (PARKS *et al.* 2004) was carried out between insertions *PBac(WH)CG7357[f00521]* and *PBac(WH)CG7785[f06154]*. More than 100 third recombinant chromosome lines were recovered over the *TM3* balancer chromosome, screened, and characterized. The manipulation of the *cpo* gene region resulted in a 214-kb duplication (designated $cpo^{3.20}$) and deletion ($cpo^{3.42}$) that span polytene bands 90C6 to 90E1. This region also covers four other genes (*CG31246*, *tinc*, *Rim*, and *Dnase2*). The selected gene deletion ($cpo^{3.42}$) is homozygous, lethal, and is clearly visible in polytene preparations. The selected gene duplication ($cpo^{3.20}$) was confirmed by PCR fragment analysis (PARKS *et al.* 2004). Polytene chromosome preparations showed $cpo^{3.20}$ as a duplication. Duplications and deletions of the 5' region immediately flanking cpo were created by *piggyBac* FLP-FRT-facilitated recombination between insertions *PBac(WH)CG14325[f03448]* and *PBac(WH)CG31249[f07289]*. This region covers 58.14 kb and includes eight genes (*alt*, *ald*, *CG31360*, *CG31249*, *CG31251*, *CG7655*, *CG7523*, and *CG14322*) and 10 tRNAs. Allele $cpo^{1.36}$ was confirmed by PCR and fragment analysis as a genomic duplication and was homozygous viable; allele $cpo^{1.45}$ was confirmed as a deletion that was homozygous lethal. All third-chromosome lines were placed into the common *w;6326* genetic background.

Mapping.

To map diapause QTL, a single F1 female from a cross between the high-diapause (*w;6326;6326*) and low-diapause (*w;6326;VT46*) parental lines was mated to *w;6326;TM3/Dr* males, and more than 250 recombinant male progeny were recovered. Each *w;6326;TM3/+RIL* male was mated to the *w;6326;TM3/Dr* balancer line to directly establish RILs for the third chromosome. All of the lines had identical X and second chromosomes. Using restriction site polymorphisms designed from Hoskins et al. (HOSKINS *et al.* 2001) and from direct sequencing, 15 SNP sites were used initially to map 275 recombination events in 201 homozygous fertile lines. All lines were phenotyped for diapause, and we recovered 112 high-diapause and 79 low-diapause RILs. Each RIL was discretely partitioned into a high- or low-diapause class (see methods below). Thus, the phenotype data were binary in nature and were analyzed by multiple-interval mapping for ordinal traits (LI *et al.* 2006). This method is a maximum likelihood-based approach that uses multiple marker intervals to determine significance for models that contain variable numbers of putative QTLs. Data were analyzed in Windows QTL Cartographer 2.5. Subsequently, an additional eight SNP markers were designed within the interval between cytological bands 90D1 and 92E8 by direct sequencing of parental lines.

Diapause Phenotyping.

The diapause induction phenotype of the RILs was tested in the homozygous condition (*+RIL/+RIL*), as well as over the *TM3* balancer (*+RIL/TM3*), which expresses a low-diapause phenotype (SCHMIDT *et al.* 2005a). These results indicated codominance of high- and low-diapause alleles: *+HD/+HD* flies were nonvitellogenic, but *+HD/TM3* flies contained stage 8 oocytes and a mean of 1.2 stage 14 oocytes per set of ovaries. In contrast, *+LD/+LD* flies, as well as *+LD/TM3* flies, were strongly vitellogenic, with an average of 14.3 stage 14 oocytes per ovary set. For purposes of mapping analysis, all lines were scored in the homozygous state. All lines and crosses that were phenotyped were maintained at low density in vial cultures on standard cornmeal-molasses medium at 25°C, 12L:12D. Females were collected within 2 h of eclosion from replicate vial cultures and placed at 11°C, 10L:14D, in Percival I36VL incubators. These females were dissected 4 weeks later, and the developmental status of the ovaries was assessed according to King (KING 1970). A female was scored as diapausing if egg development was arrested before vitellogenesis (before stage 8); a female was scored as nondiapausing if vitellogenin was observed in either ovary (stage 8 or later). At least 100 females were scored from 10 replicate cultures for each RIL and line cross. A bimodal distribution was observed: the incidence of diapause in a given RIL was either high (>70%) or low (<25%), with no intermediate frequencies observed. Thus, the trait was treated as binary and each RIL scored as

a high- or low-diapause genotype.

Allele Frequency Clines.

Eleven populations from the US East Coast were included in this study. Ten were collected by Brian Verrelli in 1997 and have been described previously (SEZGIN *et al.* 2004). The DPF population was collected in 2005 in Mount Sinai, NY, by Thomas Merritt. Isofemale lines were immediately established in the field from these populations, and the third chromosomes were extracted using a *TM3/Dr* stock. A subset (n = 46 for DPF, n = 48 for all other populations) of homozygous, viable, third-chromosome lines was used in the restriction fragment length polymorphism assay. The primer pairs and restriction enzymes are: exon 1 reference SNP, TspRI: 5'-GTCAAAGCGGGGAAAATATAGC-3' and 5'-AAATGTGTGGTAAAACCTCTGCG-3'; exon five SNPs, AfeI: 5'-ACAGCAACACCAGTGCAGGAG-3' and 5'-TCCATGCTCTGCGAAAGTCC-3', and BsiEI (residue 363 Ala/Val): 5'-ACAGCAACACCAGTGCAGGAG-3' and 5'-TCCATGCTCTGCGAAAGTCC-3'; and both exon 6 SNPs, DdeI: 5'-CGCTCAAAAGTAACGCTCGC-3' and 5'-CTCACCGATGCAGTTTTGCC-3'.

Complementation Studies.

To test the effects of the cpo gene on diapause phenotype, four sets of complementation analyses were conducted using $cpo^{BG02810}$ and cpo^{P3}. These stocks are genetically identical, except at the *cpo* locus specifically. Figure 8A shows the results of paired crosses between these two lines and three lab-generated hypomorphic *cpo P* element insertion alleles, cpo^{v3}, cpo^{cp1}, and cpo^{cp2} (BELLEN *et al.* 1992b). The analysis clearly shows that the precise excision allele cpo^{P3} always possesses significantly lower incidence of reproductive diapause against each hypomorphic allele (Wald $\chi^2 = 103.48$, 1 df, $P < 0.0001$). Because the complementation was done using paired comparisons in which the genetic background is identical except for the gene of interest, the data confirm a role for cpo in determining diapause expression.

A second set of complementation crosses used duplications and deletions of both the cpo gene region (90C6–90E1) and the 5' flanking region outside cpo (90B7–90C1) that were created through FLP-FRT site-specific recombination (SCHMIDT *et al.* 2005a). Figure 8B depicts the effect of flanking duplication and deletion on patterns of diapause expression. In both the cpoBG02810 and cpoP3 backgrounds, there was equivalency between the duplication and deletion (Wald

χ^2 = 0.061, 1 df, P = 0.806); in the complementation analysis, the only significant effect was attributed to the difference between the hypomorph and the restorative excision (Wald χ^2 = 22.57, 1 df, $P < 0.0001$). Thus, the genomic interval between cytological bands 90B7–90C1 (outside of *cpo*) does not affect diapause phenotype. In contrast, the deletion ($cpo^{3.42}$) and duplication ($cpo^{3.20}$) of the genomic region containing the functional *cpo* locus had a significant impact on diapause expression (Figure 8C). In the complementation analysis, significant effects were observed between the duplication/deletion (Wald χ^2 = 90.96, 1 df, $P < 0.0001$), the cpo backgrounds of $cpo^{BG02810}$ and cpoP3 (Wald χ^2 = 27.83, 1 df, $P < 0.0001$), and the interaction term (Wald χ^2 = 8.12, 1 df, $P < 0.01$). The significant interaction term demonstrated failure to complement. The duplication of the high-diapause *cpo* allele (in *6326*) resulted in the rescue of the low-diapause phenotype, even when over the $cpo^{BG02810}$ hypomorph. These results strongly implicate the *cpo* gene and its dosage level with the expression of diapause, and do so in a predictable fashion.

The last complementation experiment involved crosses of the parental lines and informative recombinants to the paired $cpo^{BG02810}$ and cpo^{P3} alleles (Figure 8D). In both backgrounds, high-diapause alleles exhibited a significantly higher level of diapause expression than did the low-diapause alleles (Wald χ^2 = 131.65, 1 df, $P < 0.0001$). The patterns of expression were highly distinct between $cpo^{BG02810}$ and cpo^{P3} (Wald χ^2 = 213.08, 1 df, $P < 0.0001$), but the line by *cpo* allele interaction term was nonsignificant (Wald χ^2 = 0.060, 1 df, $P < 0.807$). These results demonstrate that the difference in diapause phenotype between both the parental lines and the informative recombinants within the *cpo* locus is due to the gene specifically. The complementation also indicates additive effects of the substitution of high-diapause alleles. These complementation analyses confirmed a direct role for *cpo* in determining diapause expression and showed that the high-diapause phenotype is associated with hypomorphic expression of cpo.

Expression Analyses.

Two of the RILs were selected that showed no recombination events across the third chromosome. Based on multi-SNP genotype across the third chromosome, line 201 was fully *6326* (high diapause) and line 107 was fully *VT46* (low diapause). These lines were replicated 10 times in vial cultures and maintained at a constant low density for three generations on standard medium and culture conditions. In the subsequent generation, eight third instar larvae and eight freshly eclosed adult females were randomly collected across replicate vials for each line; the larvae were frozen for RNA isolation, and adult females were aged to 5 days. At this time, RNA samples were prepared from each line using the

Qiagen RNeasy Mini Kit. Total RNAs were further treated with Promega RQ1 RNase Free DNase. First-strand cDNAs were synthesized by using an Invitrogen SuperScript First-Strand Synthesis System for RT-PCR Kit. Gene-specific RT primer and oligo(dT) were mixed together with RNA samples in the first step according to the manufacturer's protocol. GAPDH was used as an internal control. Both lines were sequenced for all three genes across the appropriate region to ensure that the nucleotide and primer sequences were identical between the lines. The same PCR conditions were used to amplify RNA samples without reverse transcription. No DNA contamination was detected in either line. Primers and oligos used are: Cpo-5E-3F: 5'-CAGCACAAGTTGCCTACGCAG-3'; Cpo-5E-3R: 5'-TGCTCCTGCACTGGTGTTGCT-3'; Cpo-RT3: 5'-AGTCCATGCTCTGCGAAAGT-3'; Gapdh-F1: 5'-CGACCCGGCCAACATCAACT-3'; Gapdh-R1: 5'-TGGGCAGCGCCACGTCCATC-3'; Gapdh-RT: 5'-AATCCTTGCTCTGCATGTAC-3'; Dp110F1: 5'-CAACGCATCGGCCACTACTTCTTC-3'; Dp110R: 5'-TGGCGTTGACGTCAGCATTCTC-3'; Dp110RT: 5'-AAGTTTCCCGTAGGTAGTC-3'. Both lines were run using identical PCR conditions for each primer combination, and were passed through 30, 35, and/or 40 PCR cycles. Expression results were visualized on 0.8% agarose gels and f luorescence intensity determined on a Typhoon scanner (General Electric).

Notes:

A report based on this chapter appeared as:

An amino acid polymorphism in the *couch potato* gene forms the basis for climatic adaptation in *Drosophila melanogaster*

In the Proceedings of the National Academy of Sciences of the United States of America, volume 105, issue 42, page 16207-16211, October 21, 2008.

CT Zhu (Lei Zhu) played an important role in this research and contributed to a large part of this report:

1, WF Eanes, E Segzin, CT Zhu generated the RIL's
2, CT Zhu developed the genetic markers for recombination mapping.
3, CT Zhu genotyped the RIL's.
4, CT Zhu performed the multiple-interval mapping for ordinal trait.
5, CT Zhu developed the genetic markers and performed the Phenotyping for detailed mapping in 90D-92E region
6, CT Zhu developed genotyping method for *cpo* gene region and performed genotyping of DPF lines for genetic association mapping.
7, CT Zhu developed genotyping method for detecting the allele frequency variation in five different regions of *cpo*, as well as performed the genotyping for those alleles in all the population included in this study.
8, CT Zhu generated the *P*-element excision *cpo* alleles and determined the sequence features for all alleles included in the study.
9, CT Zhu generated the *piggyBac* deletion and duplication alleles included in this study and confirmed the existence of deletion/duplication using PCR amplification method.

Figure 3.

Likelihood plot for the QTL analysis of reproductive diapause. The likelihood ratio, as calculated by multiple-interval mapping for ordinal traits (Li *et al.* 2006) and implemented in QTL Cartographer v.2.5 (Category Trait Mapping, Forward Model), is plotted as a function of location in cM units on chromosome 3. The positions of the markers used in the analysis are indicated by arrows, and they correspond to cytological positions 61C, 64E, 65D, 68C, 75F, 83B, 85D, 87E, 89A, 90A, 90D, 92D, 94D, 98B, and 99A. The two markers that flank the identified QTL are indicated. Drawn using data from Table 3. Marker positions are taken from Table 4.

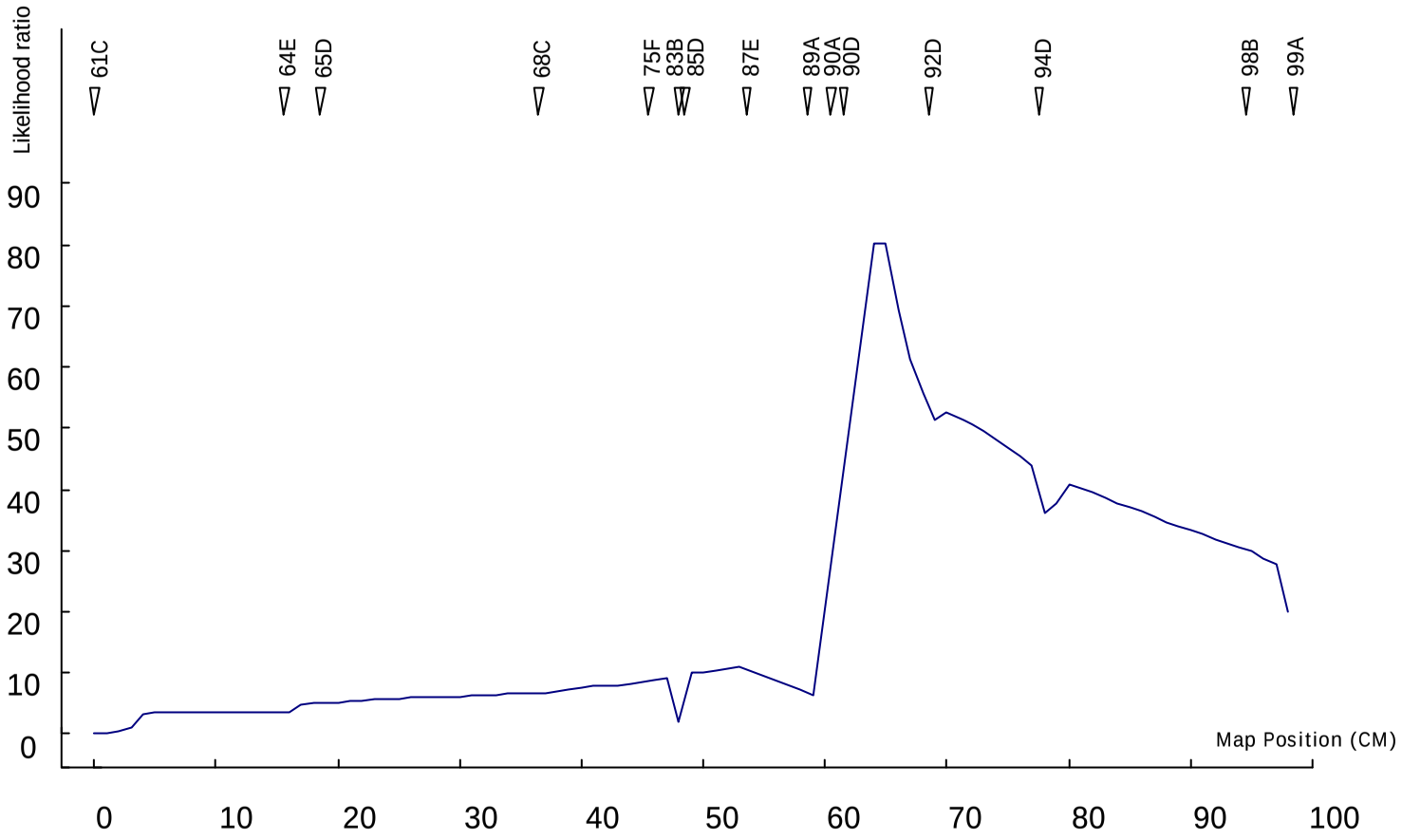

Likelihood ratio
61C
64E
65D
68C
75F
83B
85D
87E
89A
90A
90D
92D
94D
98B
99A
90
80
70
60
50
40
30
20
10
0
Map Position (CM)
0
10
20
30
40
50
60
70
80
90
100

Figure 4.

SNP genotypes across the right arm of the third chromosome and diapause incidence in the 20 RILs that were recombinant in the interval to which diapause mapped. The cytological position of each SNP marker is given (Left, top) and ranged from band 83B to 99A. Four SNPs were placed in the *cpo* gene and are listed as *cpo1–4*. The variation in diapause phenotype clearly maps only to the interval between SNP markers *cpo3* and *cpo4*; this corresponds to the 3' end of the *cpo* locus containing all of the coding sequence in exon five. Drawn using data from Table 5.

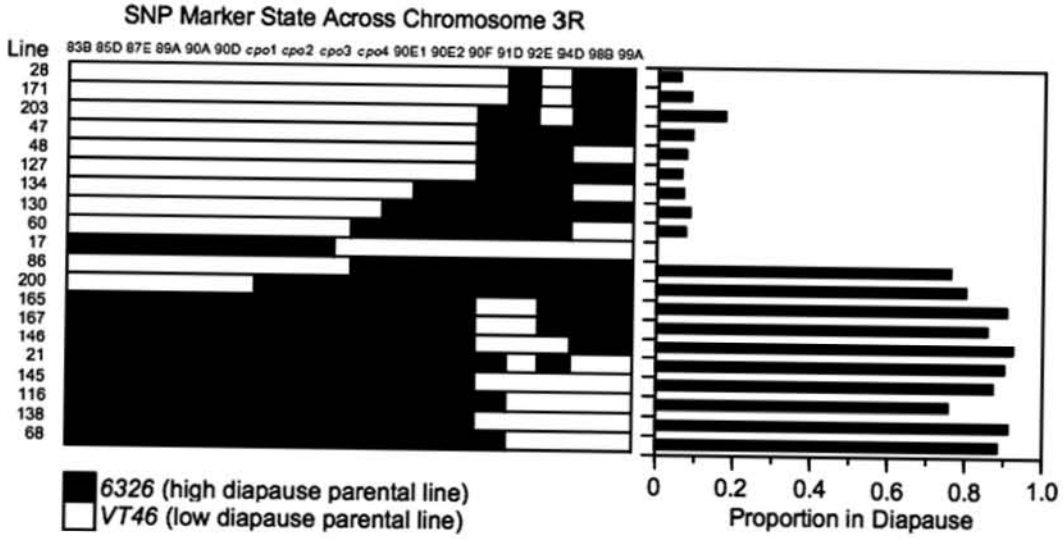

SNP Marker State Across Chromosome 3R
Line
83B 85D 87E 89A 90A 90D cpo1 cpo2 cpo3 cpo4 90E1 90E2 90F 91D 92E 94D 98B 99A
28
171
203
47
48
127
134
130
60
17
86
200
165
167
146
21
145
116
138
68
6326 (high diapause parental line)
VT46 (low diapause parental line)
0
0.2
0.4
0.6
0.8
1.0
Proportion in Diapause

Figure 5.

The association between diapause phenotype and each of 192 identified polymorphisms in the 3.5-kb region encompassing *cpo* exon five. Each point represents the transformed P value resulting from a nominal logistic regression of nucleotide state on diapause incidence in the 35 extracted third chromosomes from the DPF population. The dashed line indicates significance threshold based on Bonferroni adjustment for multiple testing. The four sites (two synonymous and two nonsynonymous substitutions) that are significantly associated with diapause phenotype span 322 bp and are in significant linkage disequilibrium with one another (D ranges from 0.162 to 0.238, $P < 0.0001$ for each). The amino acid polymorphism at residue 472 is present in only the smallest *cpo* transcript (*cpo*-RH).

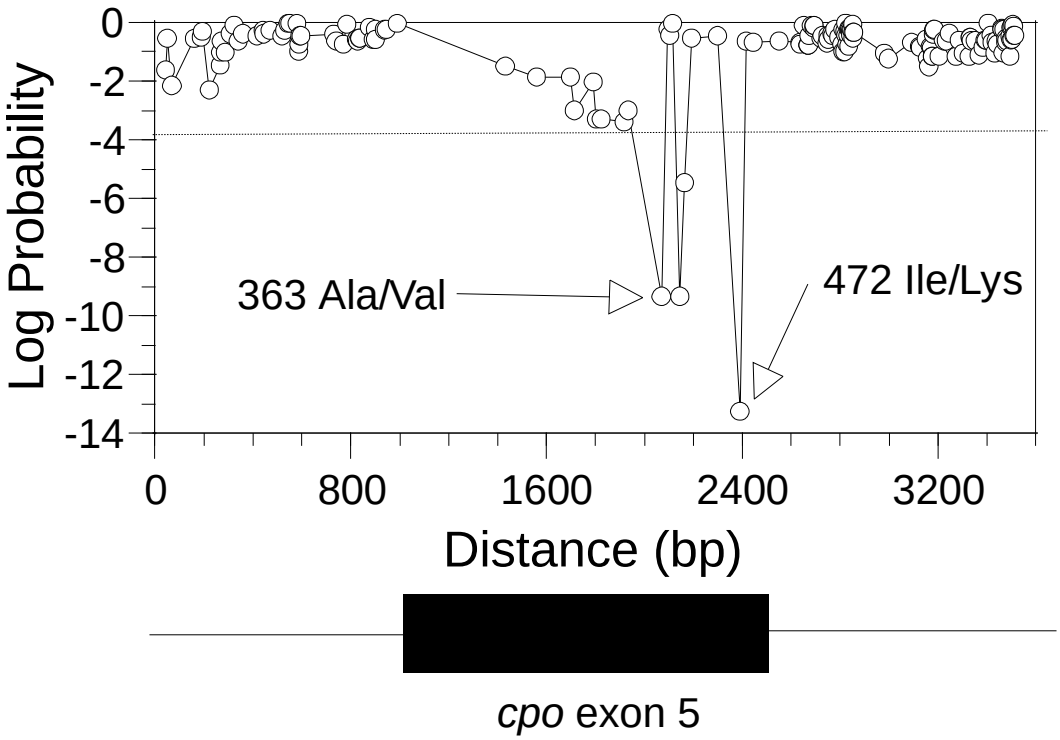

0
-2
-4
-6
-8
-10
-12
-14
Log Probability
363 Ala/Val
472 Ile/Lys
0
800
1600
2400
3200
Distance (bp)
cpo exon 5

Figure 6.

Variation in allele frequency for the five assayed *cpo* SNPs as a function of latitude of the populations. (A) Latitudinal clines for four reference SNPs in exons 1 (diamonds), 5 (circles), and 6 (5' SNP, triangles; 3' SNP, squares). The SNP in the 3' end of exon 6 exhibits a significant cline in frequency ($F_{1,10}$ = 11.82, P < 0.007, R^2 = 0.568), whereas patterns of geographic allele frequency variation are homogeneous for the other three markers. (B) The SNP corresponding to the polymorphism at amino acid residue 363 exhibits an allele frequency cline ($F_{1,10}$ = 82.97, P < 0.00001, R^2 = 0.92), and it is the derived allele (Val) that increases positively with latitude. This polymorphism is in strong linkage disequilibrium with the amino acid polymorphism at residue 472 (D = 0.216, χ^2 = 37.99, P < 0.0001). This suggests that the $cpo^{472Ile}/cpo^{472Lys}$ polymorphism also varies significantly with geography. This is further supported by direct sequencing of *cpo* alleles in third chromosomes from the northernmost and southernmost populations sampled. The frequency of the derived cpo^{472Lys} allele was 0.15 in the southern (n = 23 sequences) and 0.61 in the northern (n = 24 sequences) population (data not shown).

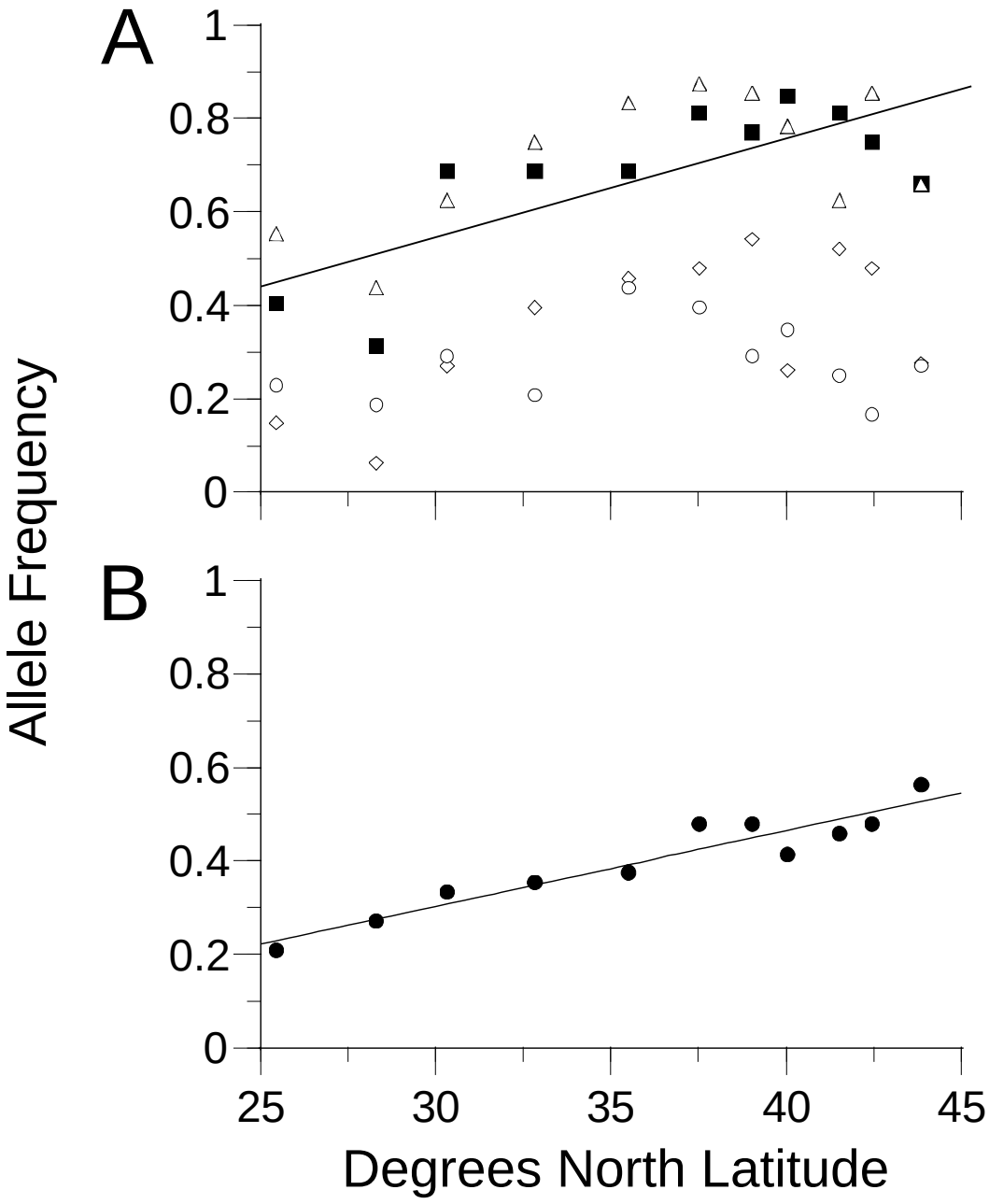

A
B
1
0.8
0.6
0.4
0.2
0
25
30
35
40
45
Allele Frequency
Degrees North Latitude

Figure 7.

Complementation analyses of 24 *cpo* alleles (12 cpo^{472Ile}, 12 cpo^{472Lys}) derived from the DPF population. (A) Diapause incidence in F1 progeny from crosses to the *cpo* hypomorphic allele $cpo^{BG02810}$ and the wild-type revertant (cpo^{P3}) created by the precise excision. Patterns of diapause expression exhibited codominance and stepwise effects of allelic substitution. Genotypes possessing two high-diapause alleles ($cpo^{BG02810}$ or cpo^{472Lys}) exhibited a strong diapause phenotype; the substitution of a low-diapause allele (cpo^{P3} or cpo^{472Ile}) resulted in intermediate diapause expression, and substituting a second low-diapause allele resulted in a low-diapause phenotype. In both genetic backgrounds, the cpo^{472Lys} allele is associated with a higher diapause expression than the cpo^{472Ile} allele. (B) Patterns of diapause incidence in F1 progeny from crosses to the *cpo* deletion ($cpo^{3.42}$) and duplication ($cpo^{3.20}$) were qualitatively identical. The observation that the duplication of the *cpo* allele from the high-diapause parental strain (stock 6326) results in a reduction in diapause incidence again indicates that diapause phenotype may be governed by cpo expression.

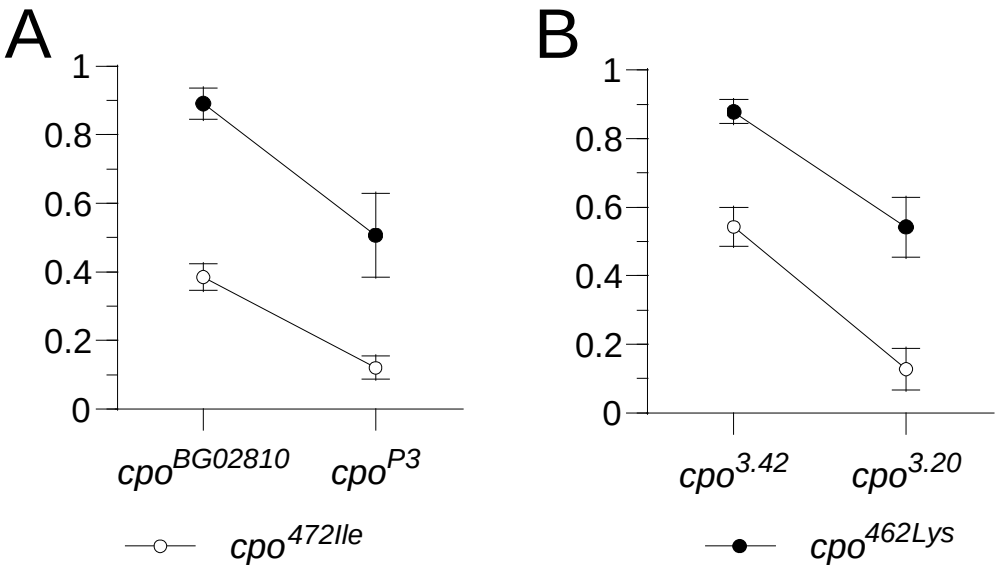

A
1
0.8
0.6
0.4
0.2
0
cpoBG02810
cpoP3
cpo472Ile
B
1
0.8
0.6
0.4
0.2
0
cpo3.42
cpo3.20
cpo462Lys

Figure 8.

The proportion of flies in diapause (95% score confidence intervals) when crossed to the $cpo^{BG02810}$ hypomorph and its precise excision, cpo^{P3}.(A) The cpo hypomorphic alleles cpo^{v3} (open circles), cpo^{cp1} (closed circles), and cpo^{cp2} (squares) exhibited similar patterns of increased diapause expression over $cpo^{BG02810}$ relative to cpo^{P3}.(B) Patterns of diapause expression in the $cpo^{BG02810}$ and cpo^{P3} backgrounds were indistinguishable between duplications (open) and deletions (filled) of the 5' flanking region of *cpo*.(C) Duplication (open) of the region containing the cpo gene resulted in low-diapause expression in both the $cpo^{BG02810}$ and cpo^{P3} backgrounds. Because the duplication was in the 6326 background, which was the high-diapause parental line, this indicates that the duplication of the diapause allele rescues the low-diapause phenotype. The deletion (filled) of this genomic region resulted in higher diapause expression relative to the duplication when crossed to both $cpo^{BG02810}$ and cpo^{P3}. The significant line / allele interaction indicates that manipulation of the *cpo* region causes informative changes in patterns of diapause expression. (D) Patterns of diapause expression for high-diapause (circles) parental (filled symbols), and informative recombinant inbred (open symbols) lines were consistently higher than for low-diapause lines (triangles) in both the $cpo^{BG02810}$ and cpo^{P3} backgrounds. The largely parallel lines, reflecting a nonsignificant line / allele interaction term, indicate that diapause expression acts in a codominant fashion in these lines.

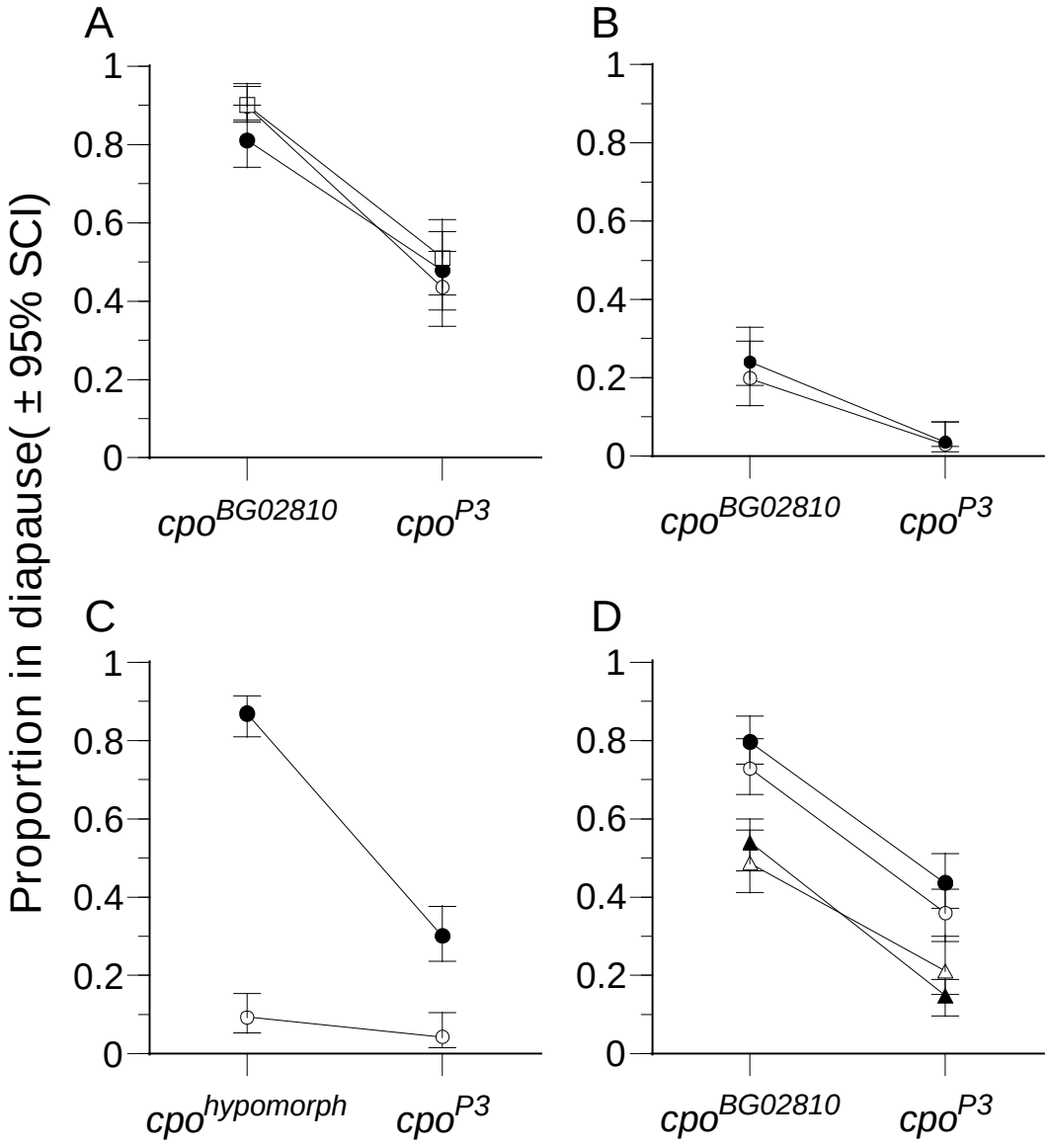
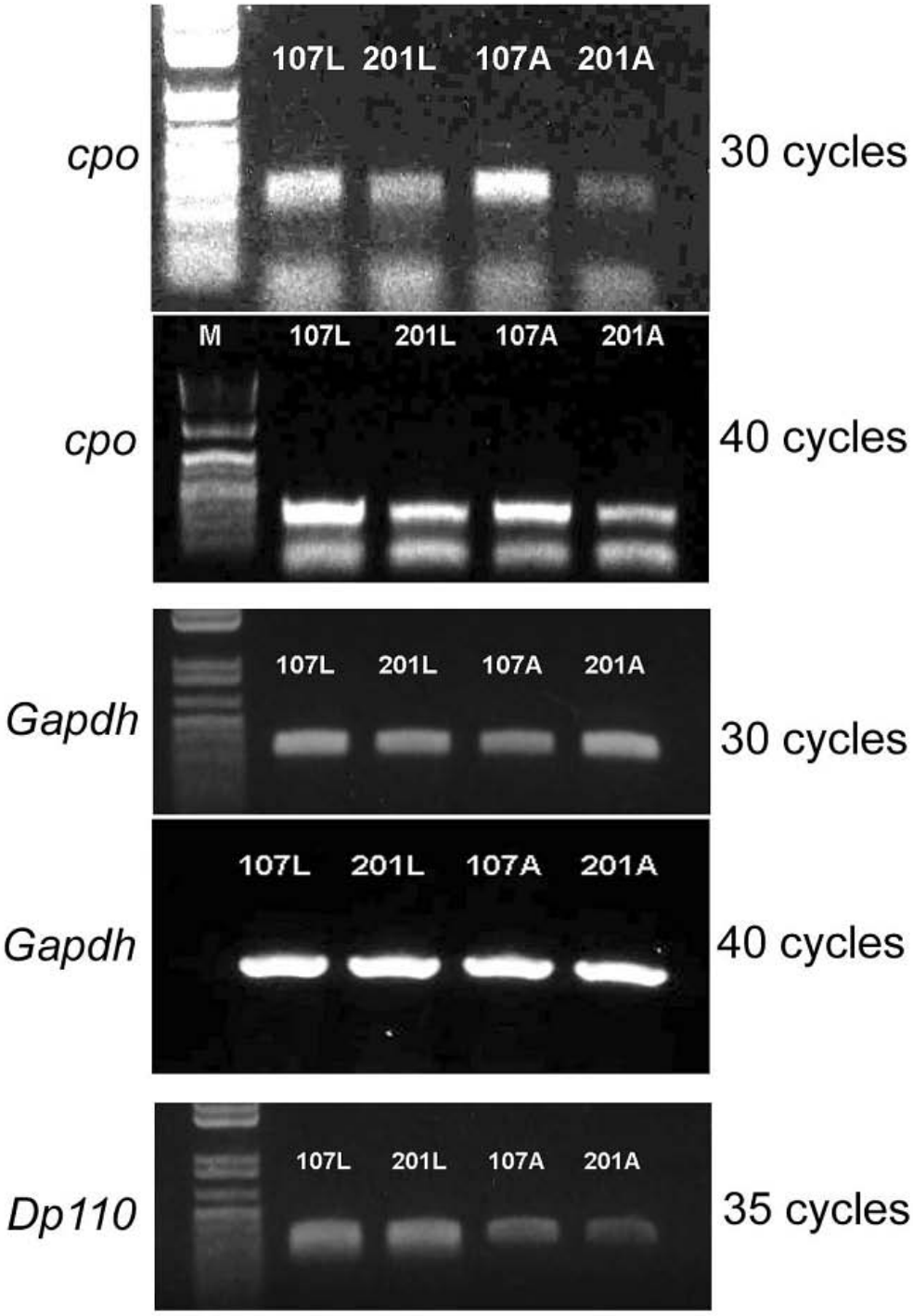

A
B
C
D
Proportion in diapause(± 95% SCI)
1
0.8
0.6
0.4
0.2
0
cpoBG02810
cpoP3
cpoBG02810
cpoP3
cpohypomorph
cpoP3
cpoBG02810
cpoP3

Figure 9.

Results of semiquantitative PCR for *cpo*, *Gapdh*, and *dp110* in the high-diapause (line 201) and low-diapause (line 107) parental lines. Expression patterns are depicted for third instar larvae (L) and 5-day-old adult (A) flies reared at 25°C, 12L:12D. The results show a greater than 5-fold decrease in *cpo* expression associated with line 201 or the *6326* (D) allele at 30 cycles; expression patterns also remained distinct after 40 PCR cycles. Gapdh was used as an internal control and did not demonstrate any differential expression between the lines. The identified candidate gene for diapause *dp110* was also included; third instar larvae exhibited equivalent levels of expression, but adult profiles were distinct between the lines.


cpo
107L 201L 107A 201A
30 cycles
cpo
M 107L 201L 107A 201A
40 cycles
Gapdh
107L 201L 107A 201A
30 cycles
Gapdh
107L 201L 107A 201A
40 cycles
Dp110
107L 201L 107A 201A
35 cycles

Chapter three: the population genetics of reproductive diapause in *Drosophila melanogaster*, the *couch potato* gene is under balancing selection

Introduction

Over several hundred years, the African species, *Drosophila melanogaster* has adapted to the temperate regions of the New World where it has encountered a number of life history challenges (DAVID and CAPY 1988). Among these, reproductive diapause is a important adaptation where females are able to stop reproduction and enhance their overwintering ability. Ancestral populations in Africa, as well as the sister species *Drosophila simulans*, lack of this ability (SCHMIDT *et al.* 2005a). The diapause trait seasonally cycles in the eastern US and is associated with other life history strategies and trade-offs in both males and females, suggesting that the underlying genes may be under balancing selection (SCHMIDT and CONDE 2006; SCHMIDT *et al.* 2005b). Using QTL mapping, we have identified the *couch potato* (*cpo*) gene as a major locus contributing to variation in reproductive diapause expression, and we found a molecular polymorphism in this gene that can largely explain the observed phenotypic variation seen in a North American population (SCHMIDT *et al.* 2008b). In this chapter, we show that two specific *cpo* haplotypes, that are relatively rare in Africa, have increased substantially, and in a latitudinal fashion in North America paralleling reproductive diapause.

The large *couch potato* gene is previously known for being essential for normal adult behavior (BELLEN *et al.* 1992b). Exons 1-4 are not translated. Exon five encodes the largest portion of its coding sequence. Exons 6-12, encode a RNA binding domain and are relatively conserved across the entire *Drosophila* genus. The first half of *cpo* exon five is highly conserved between and within *D. melanogaster* and *D. simulans*, suggesting it is under strong purifying selection. There are several fixed differences between species, most of which are synonymous (Figure 2b). In contrast, the second half of exon five is highly polymorphic and encodes most of the amino acid sequence segregating for the changes that are diapause-associated. In the Davis Peach Farm (DPF) sample (New York, USA), this exon possesses 13 synonymous and 7 non-synonymous polymorphisms, as well as two regions of poly-glutamine and poly-serine repeat length variation (Figure 10a). We did find some low frequency deletion-insertion polymorphism in this region. The most common one is an 18 base pair deletion 472-489 bp, which has roughly a frequency of 10% in North America. This deletion doesn't cause a frameshift. The first part of the exon may encode a helix-loop-helix structure, which may be shortened by this deletion.

In the following report, I will present the features of molecular polymorphism in and around *cpo* exon five. In the North American populations, the polymorphisms segregate mainly as two major haplotype groups. The haplotype structures may be a signature of balancing selection associated with reproductive diapause in the New World. There is also evidence that the same region that is under balancing selection in the North American population is the target of positive selection in *D. melanogaster* and *D. simulans* lineages.

Material and Methods

The North American *D. melanogaster* populations were collected as isofemale lines. Samples from Maine (RR lines) were collected form Rocky Ridge Orchards, Bowdoinham, ME in 2006 by Dr. Paul Schmidt. The New York population (DPF lines) was sampled from David's Peach Farm in Wading River, Long Island, New York during the late summer in 2005 by Dr. Thomas Merritt. Vermont (VT lines) and Florida (HFL lines) samples were collected in fall 1997 by Dr. Brian Verrelli from Whiting, Vermont and Homestead, Florida respectively. The third chromosomes were extracted using a balance stock of *TM8/TM6* and placed in a common genetic background *w; w1118* (Bloomington Stock No. 6326) for the X and the second chromosomes. They are subsequently maintained over the third chromosome balancer *TM8*. The diapause phenotype assay is carried out according to Saunders *et al* 1989, but see Schmidt *et al.* 2008a.

African isofemale inbred lines are from Kenya (KY, by Dr. B. Ballard and S. Charlat), Cameron (MD, by Dr. P. Konje and J. Pool), Gabon (GN, by Dr. B. Ballard and S. Charlat) and Zimbabwe (ZW, by Dr. R. Ramey and L. Brown, ZK, by Dr. T. Mutangadura and ZS, by Dr. R. Ramey). The third chromosomes were extracted using a balance stock of *TM8/TM6*. The extracted lines are subsequently maintained over the third chromosome balancer *TM8*.

The third chromosome extracted lines are used for DNA sequencing. *D. simulans* lines were collected in 1996 as isofemale lines and inbred for more than 20 generations. DP96-4SB, DP96-2SB and DP96-2SA were collected from New York. CT96-4SB, CT96-6SA were collected from Connecticut and GA96-5SA was collected from Georgia, U.S. by Dr. Brain Verrelli.

For the *D. melanogaster* samples, a single homozygote from each extracted lines was used for DNA sequence. For the *D. simulans* samples, a single individual from ethanol-preserved samples for each inbred line is used for DNA sequencing. DNA amplification is done using *Pfu Ultra II* (Stratagene) to minimize the polymerase error. Amplicons are directly sequenced following the *Sap-Exo1* treatment. Sequence alignment and phylogeny reconstruction is conducted using MEGA 3.1(KUMAR *et al.* 2004). Sequence analysis is conducted using DnaSP (ROZAS and ROZAS 1995).

Results

Diapause is strongly associated with a T/A transversion in exon five that results in an Ile/Lys (residue 472) replacement (SCHMIDT *et al.* 2008b). Many of the polymorphisms in exon five co-segregate with this causal polymorphism, forming two haplotypes covering the distal polymorphic part of this exon (Figure 10a). The haplotypes are designated here by their glutamine and serine amino acid repeat length; Haplotype Q5S7 contains His242, a five glutamine repeat starting at 291, Thr318, Ala363, a seven poly-serine repeat starting at 418 and Il472. Haplotype Q7S5 contains Glu242, a seven glutamine repeat, Asn318, Val363, a five serine repeat and Lys472 (Figure 10a). Both haplotypes Q5S7 and Q7S5 possess intermediate frequencies in the DPF sample and comprise two thirds of the gene copies with the remaining haplotypes seen in deep branches (Figure 11). Lys472 is always associated with diapause and Ile472 is always non-diapause. Neither amino acid repeat length nor any other amino acid replacements are sufficient to cause diapause. Lines DPF072 and DPF085 show exceptional non-diapause phenotypes, suggesting there are modifiers on the third chromosome (the X and second chromosomes are replaced for the assay) that can override the effect of *cpo* gene.

The C-terminus of *cpo* is also undergoing novel evolution in *D. melanogaster* subgroup. Transcript analysis of *cpo* predicts several alternatively spliced transcripts, but only one utilizes the full length exon five and this lacks exons 6-12 that encode the RNA binding domain. All other transcripts possess the RNA binding exons, but lack the last 41 amino acids coded by exon five including the I472K polymorphism (BELLEN *et al.* 1992b). Between *D. melanogaster* and *D. simulans*, we see a frame shift in this part of exon five. The *D. simulans* protein has an additional 26 amino acids, completely different from the last 11 amino acids encoded in *D. melanogaster*. It is impossible to distinguish either *D. melanogaster* or *D. simulans* is the derived state for the frame shifting mutation. Considering the lack of diapause in *D. simulans* (SCHMIDT *et al.* 2005a), this may have opened a new adaptative landscape for the *D. melanogaster* lineage,. The next closely related species, *Drosophila yakuba*, is fixed for yet another frame shifting difference besides other sequence variations. This type of extreme variability has not been seen before for a gene essential for normal viability in *Drosophila*.

The DPF population differs from the African populations not only in the level of polymorphism, but more importantly in the haplotype structure. The diapause mutation Lys472 is found in Africa, although in low frequency. One important difference between African and American sequences is the extensive haplotype

polymorphisms in ancestral African populations. The less common haplotypes in North American are simply nested in the deep genealogy of African alleles (Figure 12a). Both Q5S7 and Q7S5 are relatively infrequent in the African populations (4 each/53).

Is the North American population a simple random sample of the ancestral population and if not, how does it differ? We use simulations to answer this question. Estimating population parameters (θ and recombination rate) of exon five from the African samples, we use coalescent simulation function in DnaSP (ROZAS and ROZAS 1995) to generate 100000 pseudo-sequences. We then calculate the probability of statistics observed in DPF population given the sample size of DPF population in these 100000 pseudo-sequences. The model of coalescent simulation we used assumes a constant population size and the absence of selection. For exon five, the African sample has similar nucleotide diversity to DPF population [π, by Nei (NEI 1987): 0.00576 vs. 0.00486 per site, $P = 0.3025$], but has marginally more of segregating sites [θ-W, by Watterson (WATTERSON 1975): 0.00663 vs. 0.00387 per site, $P < 0.056$]. In contrast, the DPF population processes fewer haplotypes and lower haplotype diversity (numbers of haplotypes, 14 vs. 53 for African samples, $P < 10^{-4}$. Haplotype diversity, 0.877 vs. 0.989 for African samples, $P < 0.0003$). It shows a much stronger haplotype structure [*ZnS*, (KELLY 1997): 0.290 vs. 0.0707 for African population, $P < 10^{-4}$]. It appears that selection in the DPF population, favors the expansion of two unique combinations of polymorphisms as the Q7S5 and Q5S7 haplotypes. The remaining haplotypes represent ancestral variation rather than recombinants between these two major haplotypes. It appears that the Lys472 polymorphism, perhaps with the entire Q7S5 haplotype, originated in Africa and became the target of strong selection upon entry into temperate climates. The emergence of two haplotypes in the DPF populations is the likely result of balancing selection acting across a geographic range where there is increasing selection on diapause. No other gene yet surveyed in *D. melanogaster* shows this dramatic pattern.

There are other statistical supports for balancing selection in exon five. Tajima's *D* test is a test that compares the nucleotide polymorphism frequency spectrum against a neutral expectation (TAJIMA 1989). When *D* is positive, it means that there is an excess intermediate frequency mutations. Balancing selection is expected to lead to excess numbers of intermediate frequency mutations and positive values of *D*. Across exon five, *D* is slightly negative and does not statistically significantly differ from zero in the DPF or African samples. However, between the poly-glutamine and poly-serine repeats, Tajima's *D* for the American population is significantly positive (Figure 10b), apparently driven by the intermediate frequencies of the two diverged haplotypes. The African sample

does not show this pattern (Figure 12b).

In our four North America samples, we observed a marked change in the frequency of the diapausing haplotype, Q7S5. It reaches the highest frequency in northern populations (Maine: 0.499, Vermont: 0.304, New York: 0.341 and Florida: 0.087). This is consistent with our early findings that the SNP for A363V (in strong linkage disequilibrium with I472K and haplotype Q7S5) is strongly clinal with the frequency of the Val363 allele highest in the north (SCHMIDT *et al.* 2008b). The Lys472 allele is even more clinal (Maine: 0.559, Vermont: 0.609, New York: 0.383 and Florida: 0.087). However, the frequency of Q5S7 haplotype is not clinal (Maine: 0.206, Vermont: 0.130, New York: 0.362 and Florida: 0.217). In Homestead, FL, most alleles do not belong to either Q5S7 or Q7S5 haplotype. In our survey, none of the chromosomes carrying cosmopolitan inversion *In(3R)P* process either Q7S5 and Q5S7 haplotypes. I propose that the balancing selection responsible for maintaining the two major haplotypes is stronger in the north where diapausing ability is expected to be adaptive. Therefore, the pattern of polymorphism in *cpo* in different populations reflects both balancing and directional selection for reproductive diapause.

It is desirable to identify the extent of the *cpo* region possibly affected by balancing selection. Linkage disequilibrium analysis shows that in DPF population, the haplotype in exon five extends for at least 500bp upstream, but not significantly downstream of the exon. Among many segregating sites in the region upstream, seven sites form a separate cosegregating block. The Q7S5 and Q5S7 haplotypes are strongly associated with only one common upstream haplotype block, while the minor haplotypes are associated with a number of different upstream alleles. In *Drosophila*, linkage disequilibrium generally decays within 1kb between segregating sites, which is true downstream of exon five. The long-range linkage pattern upstream, therefore, suggests the presence of selection, not only in the amino acid sequence level, but possibly the transcriptional regulation as well (Figure 13).

There is also evidence that part of *cpo* has been under positive selection in the *D. melanogaster* and *D. simulans* linages. We can use the Donald-Kreitman test (MCDONALD and KREITMAN 1991) to investigate this question for exon five, but necessarily exclude the small frame shifted portion between *D. melanogaster* and *D. simulans*. The entire exon five does not display positive selection between DPF *D. melanogaster* and *D. simulans* between African *D. melanogaster* and *D. simulans* (*P*-value for *G* test: 0.068, with William's Correction: 0.072, Table 2). Nevertheless, if we consider the first half of the exon is highly conserved and likely to be under purifying selection, when analyzing only the second half of

exon five, we detect an excess of fixed nonsynonymous substitutions between African *D. melanogaster* and *D. simulans* (*P*-value for G test: 0.042, with William's Correction: 0.047, Table 2). Therefore, there is evidence that the same region under balancing selection in *D. melanogaster* is also the target of adaptive change between species.

Discussion

The cosmopolitan populations of *D. melanogaster* are of tropical African origin. Its colonization of the New World is generally thought to have happened in the past a few hundred years. Since the reproductive diapause ability appears to be absent in African populations, it is remarkable that the important adaptation to temperate environment took place in a very brief time. We previously identified a Lys-Ile mutation in the gene *cpo* that is strongly involved in reproductive diapause. As an apparent adaptation, the frequency of diapausing-associated allele increases with latitude. This latitudinal cline is consistent with the hypothesis that *cpo* alone seems to explain a large proportion of variation of diapause frequency in natural populations. Unlike in African populations, in North America the diapause associated polymorphism co-segregates as two haplotypes consistent with balancing selection. The seasonality in the north leads to balancing selection arising from the requirement for surviving longer.

The haplotype genealogy of *cpo* and the different pattern of polymorphism between African and North American populations shows the signature of recent selection. This is consistent with our independent findings that *cpo* is a important gene in reproductive diapause. We found that the diapausing-associated *cpo* mutations in North American are not recently derived, instead they exist in the ancestral variation in the African samples at low frequency. It is possible that additional recently derived alleles at other loci are required for full diapause expression in North American populations. Recent work on diapause-related genes could provide us with some promising candidate genes for modifiers (SIM and DENLINGER 2008; TAUBER *et al.* 2007; WILLIAMS *et al.* 2006).

This apparent selection on *cpo* may also partially be due to selection on traits other than reproductive diapause. For most of the polymorphisms in this haplotype, only Lys472 is clearly associated with reproductive diapause. Hitchhiking may generate a haplotype structure but it is not likely to be the explanation of the pattern we see. We shall see a haplotype around Lys472, but in reality the haplotype doesn't extend further downstream of Lys472. It is possible that selection on other polymorphisms, although not associated with diapause, is driven by the similar ecological factors. Those polymorphisms, together with diapause causal locus, could generate the haplotype pattern.

The *cpo* gene contains a RNA binding domain, which is coded by exons six to twelve. This may imply that *cpo* is involved in transcriptional regulation. We have no clear information about the biological function of exon five. From its amino acid sequence, we are able to tell that the conserved first half of exon

encodes amino acid sequence of low complexity and possibly a helix domain. We have no knowledge regarding the structure and possible function of the second half of exon five. Because of multiple splicing, it is possible that the polymorphism in exon five has pleiotropic effects on other biological functions as well.

It is unkown whether the large intronic region of this gene has any effect on reproductive diapause except that several polymorphisms within 500 base pair upstream exon five are associated with the Q7S5 and Q5S7 haplotypes. Polymorphisms in the intronic regions are worthy of further research, especially insertion-deletion and transposable elements. Also, it is unkown whether this gene has any temperature sensitive alternative splicing, and if diapause is related to alternative splicing features under diapause inducing conditions.

In *D. melanogaster*, *cpo* is involved in reproductive diapause, but this role may be limited to species closely related to *D. melanogaster*. Parts of *cpo* are fast evolving. However, the RNA binding domain is relatively conserved and its homologue can be identified in some other species. Even its homologue can be found in many insects including mosquito, silk worm and honey bee, the homologue of *D. melanogaster* exon five cannot be identified with high confidence. In the genus of *Drosophila*, it is possible to identify orthlogues of *cpo* in other sequenced *Drosophila* genomes by sequence similiarity and gene order. However, the homologue of *D. melanogaster* exon five seems to exist only in *D. melanogaster* species subgroup (*D. melanogaster*, *D. simulans*, *D. sechellia*, *D. erecta*, and *D. yakuba*). In *D. melanogaster*, diapause is strongly associated with polymorphism in exon five. Though it is unkown functionally how this exon causes diapause, it is likely to mechanism specific to *D. melanogaster* species subgroup since this exon is only found in this group.

Figure 10.

Figure 10a shows the amino acid polymorphisms in the two most abundant haplotypes in North America, haplotypes 5Q7S and 7Q5S. Amino acid polymorphisms are shown in red italic font. Each amino acid polymorphism is shown with its codon and the casual nucleotide polymorphism in blue. The position of each polymorphism is also indicated. Numbers on the top are the amino acid sequence positions for each polymorphism. Numbers on the bottom are the nucleotide sequence positions with nucleotide number 1 as the start of this exon. Figure 10b shows the Tajima's *D* value from approximately 1 kb upstream to the end of this exon. Tajima's *D* is calculated in 150 bp windows with 25 bp steps. The North American estimates, based on DPF lines, are shown in a red solid line and a blue dash line represent the value of African population. Regions 724-873 bp and 814-963 bp shows significant positive Tajima's *D* for the North American population ($0.01 < P < 0.05$).

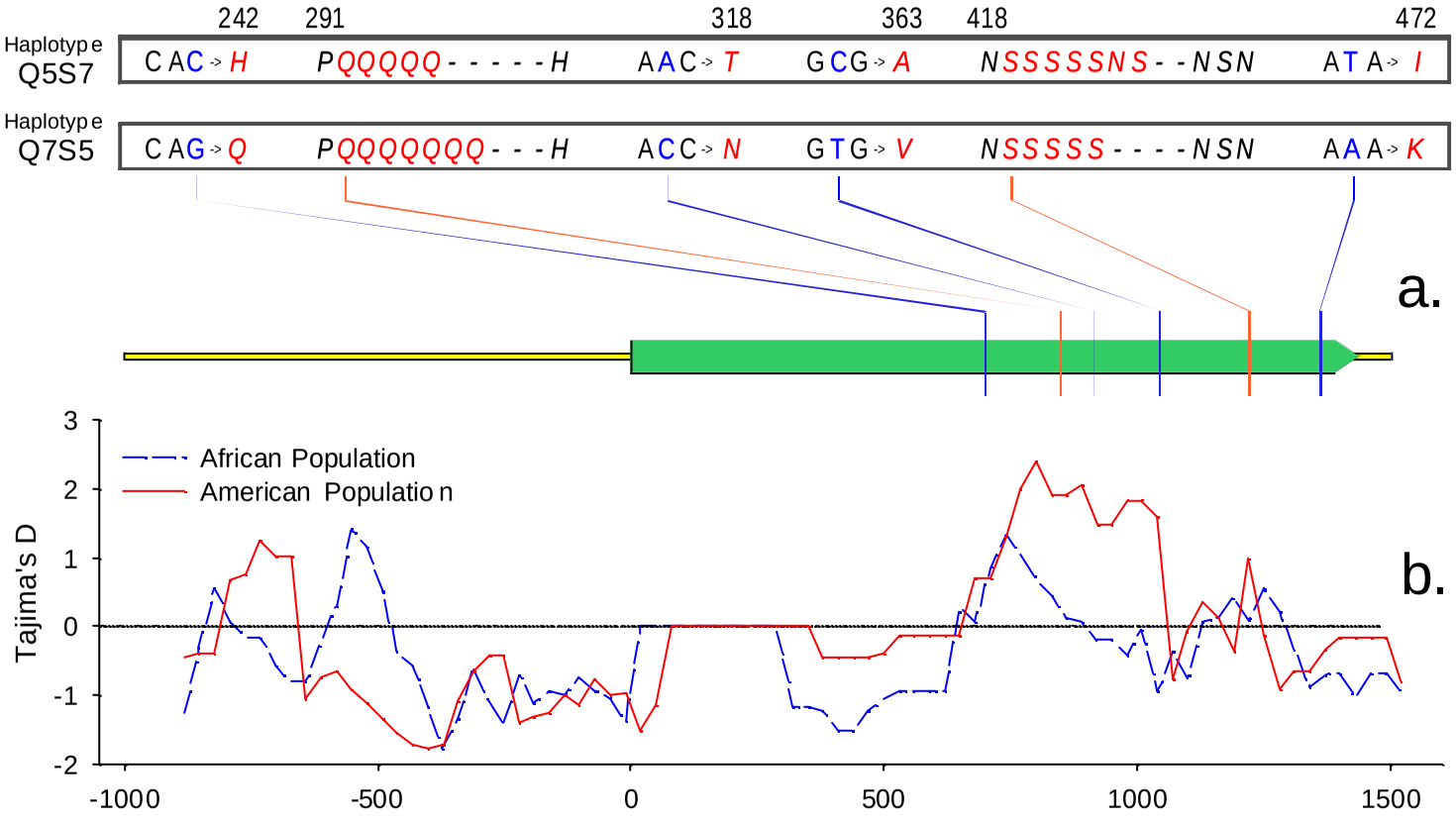

242
291
318
363
418
472
Haplotype Q5S7
CAC -> H
PQQQQQ - - - - - H
AAC -> T
GCG -> A
NSSSSSNS - - NSN
ATA -> I
Haplotype Q7S5
CAG -> Q
PQQQQQQQ - - - H
ACC -> N
GTG -> V
NSSSSS - - - - NSN
AAA -> K
a.
African Population
American Population
Tajima's D
3
2
1
0
-1
-2
-1000
-500
0
500
1000
1500
b.

Figure 11.

Figure 11a is a maximum parsimony phylogeny for the Davis Peach Farm lines (DPF) with *D. simulans* as outgroup. ST-DM is the standard *D. melanogaster* genome reference sequence. The phylogeny is based on the sequence of *cpo* exon five excluding the insertion-deletion and the region affected by the fixed frame-shifting mutation between *D. melanogaster* and *D. simulans*. For DPF lines, line numbers are followed by their haplotype classes. The values on the branches are bootstrapping values. The scale bar shows the relative branch length of two mutations. Figure 11b shows the polymorphisms of the region from base pair 87 to base pair 1472 of *cpo* exon5, upon which Figure 11a is based upon. Replacement polymorphisms are shown in bold. With reference to the *D. melanogaster* standard genome sequence, identical sequences are shown in dots. The nucleotide position number is based on a line having 10 Glu and 9 Ser residues in the amino acid repeat regions, which is the longest observed allele for this exon. Non-diapause and diapause phenotypes are labeled using two different colors.

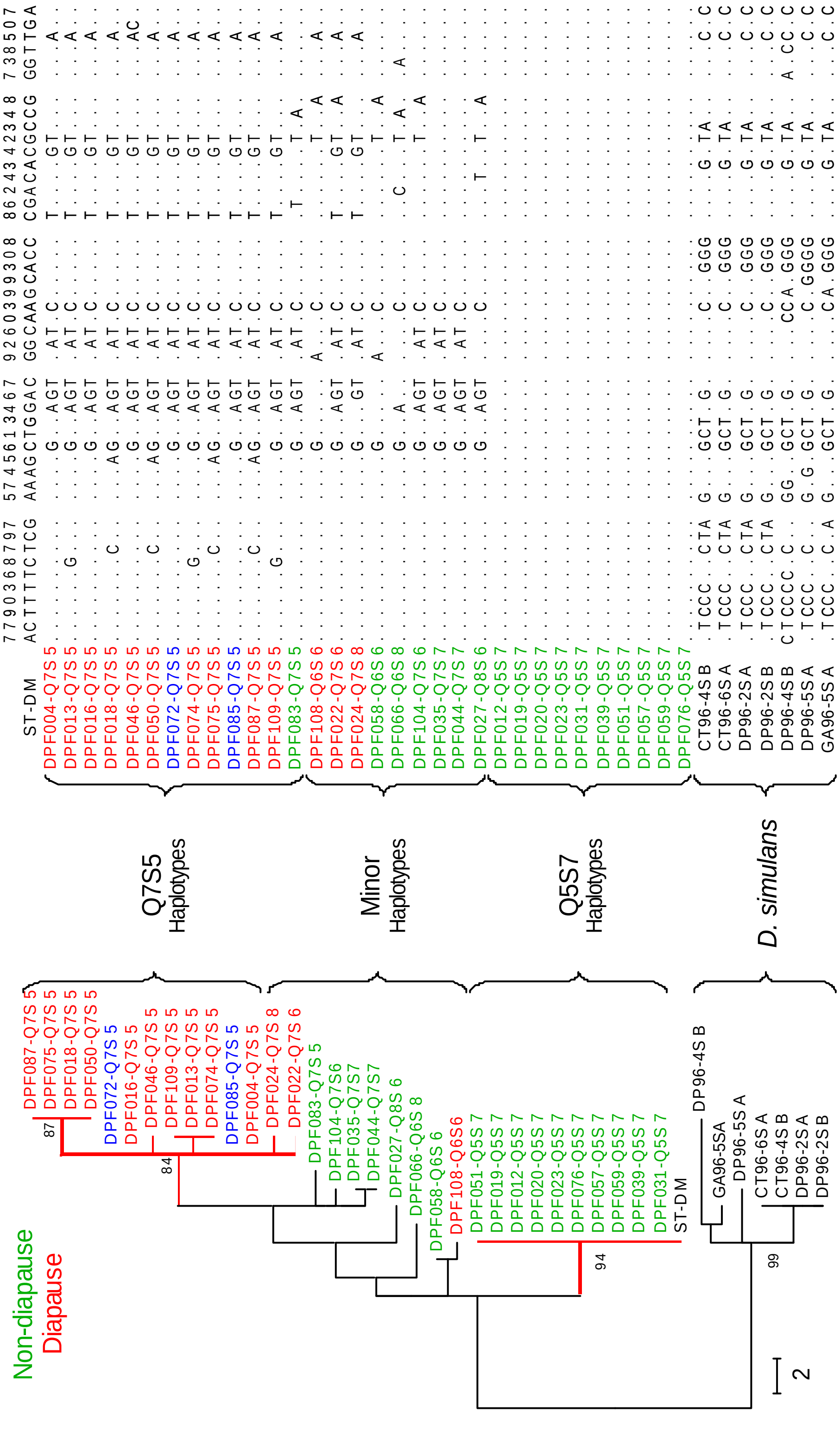

Non-diapause
Diapause
Q7S5
Haplotypes
Minor
Haplotypes
Q5S7
Haplotypes
D. simulans
a.
b.

Figure 12.

The neighbor-join phylogeny for Davis Peach Farm and African lines (Figure 12a) and for North American lines only (Figure 12b). *D. simulans* lines are the outgroup. DPF is the Davis Peach Farm line. North American lines are followed by their haplotype labels. RR, VT, DPF and HFL lines are North American population samples. KY, MD, GN, ZW, ZK and ZS are African population samples. ST-DM is the standard *D. melanogaster* genome sequence. Lines are color labeled by origin. The neighbor-join tree is based on the sequence of *cpo* exon five excluding insertion-deletion polymorphisms and the region affected the fixed frame-shifting mutation between *D. melanogaster* and *D. simulans*. The neighbor-join phylogeny is constructed using Kimura 2-parameter model for nucleotide substitution. Both transitions and transversions are included. Scale bar shows the branch length of an estimated mean number of 0.002 substitutions per site. This figure is drawn using data from Figure 17.

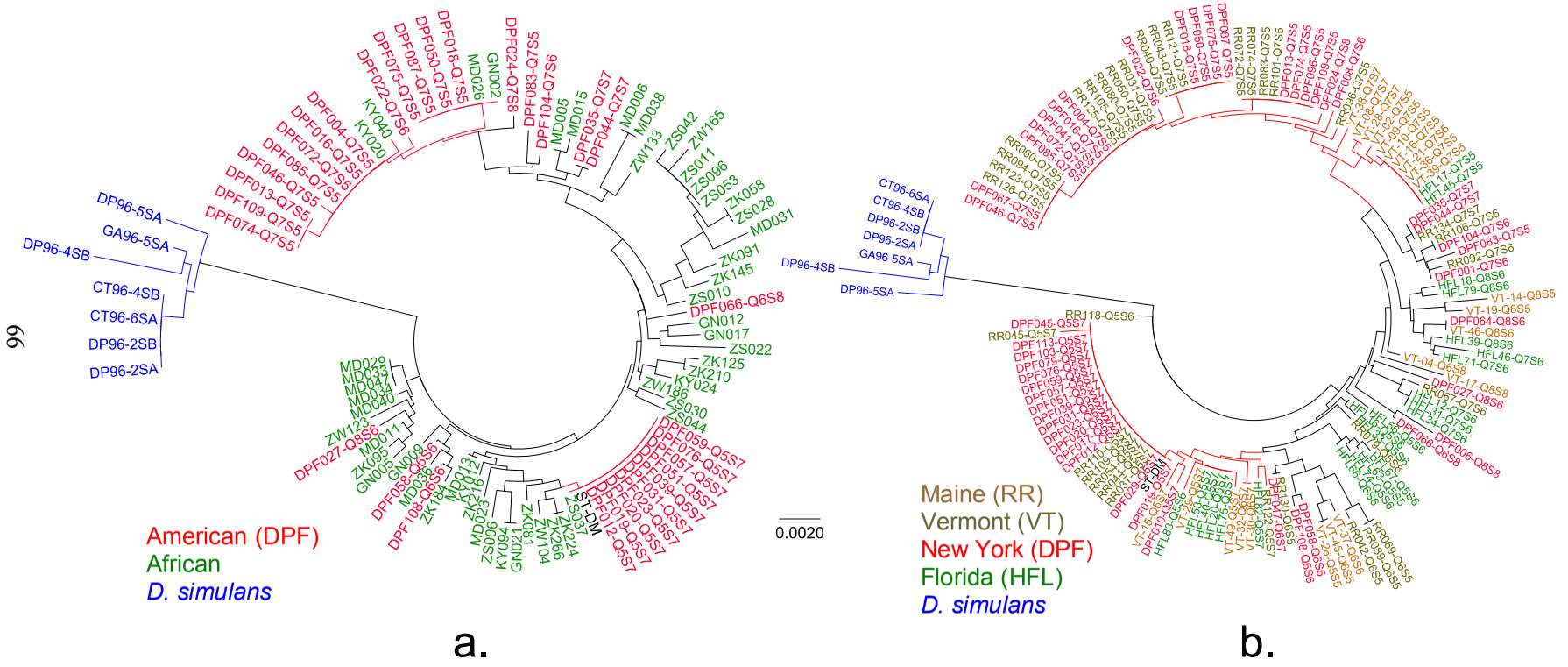

99
DP96-5SA
GA96-5SA
DP96-4SB
CT96-4SB
CT96-6SA
DP96-2SB
DP96-2SA
American (DPF)
African
D. simulans
ST-DM
0.0020
a.
Maine (RR)
Vermont (VT)
New York (DPF)
Florida (HFL)
D. simulans
b.

Figure 13.

The haplotype blocks in *cpo* exon five and 1kb upstream and downstream. DPF lines are used in this analysis. The vertical axis indices the linkage disequilibrium measure R^2 and the horizontal axis indicates the *P* values from permutation tests. Haplotype blocks are assigned visually. For each haplotype block, the population frequency and its connections between adjacent block are indicated. For simplicity, haplotypes having a population frequency less then 0.10 are omitted from this figure. The analysis is done using JLIN (CARTER *et al.* 2006) and Haploview (BARRETT *et al.* 2005). Drawn using data from Figure 17.

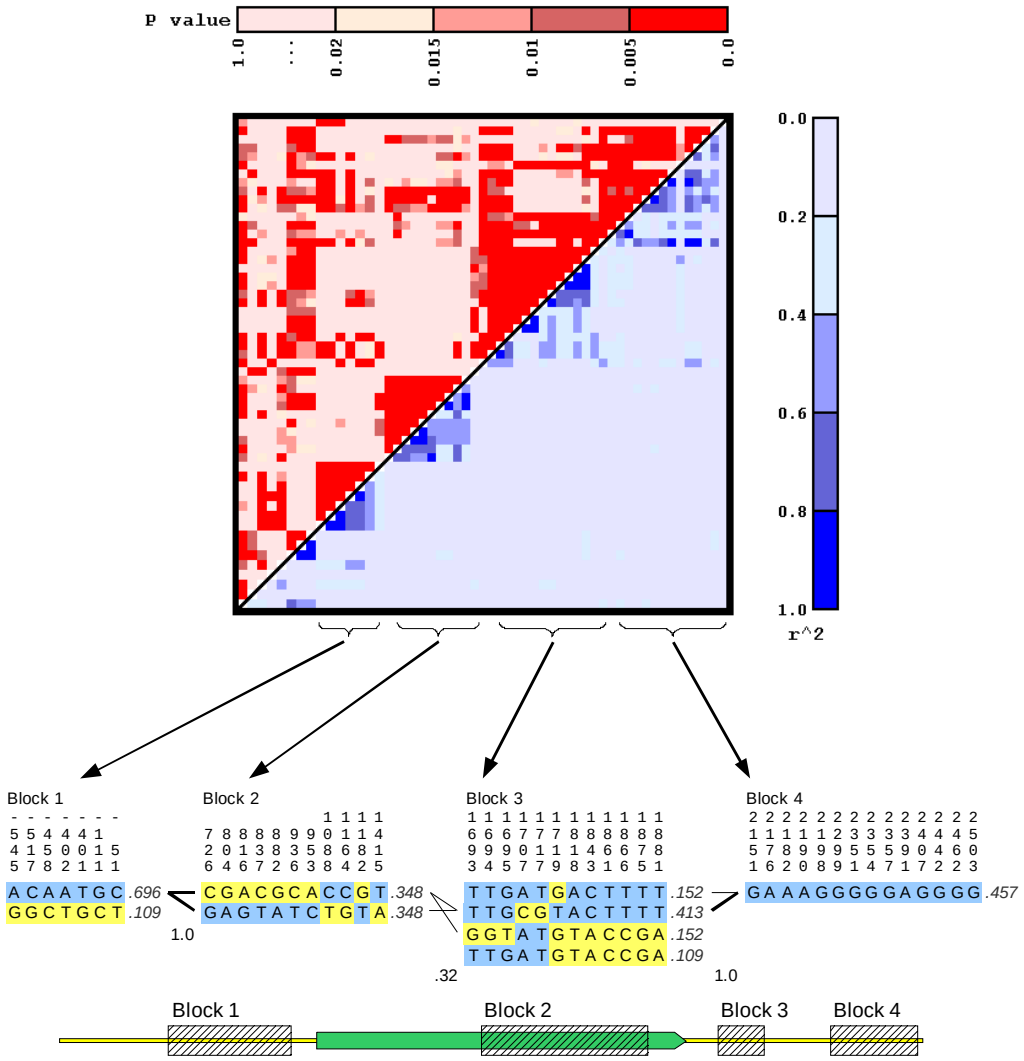

p value
1.0
...
0.02
0.015
0.01
0.005
0.0
0.0
0.2
0.4
0.6
0.8
1.0
r^2
Block 1
Block 2
Block 3
Block 4
ACAATGC .696
GGCTGCT .109
CGACGCACCGT .348
GAGTATCTGTA .348
TTGATGACTTTT .152
TTGCGTACTTTT .413
GGTATGTACCGA .152
TTGATGTACCGA .109
GAAAGGGGGAGGGG .457
1.0
.32
1.0
Block 1
Block 2
Block 3
Block 4

Table 2.

The McDonald and Kreitman test for African *D. melanogaster* samples and *D. simulans* samples. It is calculated from Figure 17 and 7 excluding insertion/deletion differences, repeat length polymorphisms and frame shift.

For the entire exon five:

	Fixed differences between species	Polymorphic sites
Synonymous Substitutions	6	36
Nonsynonymous Substitutions	10	21
G value	3.34	$P = 0.068$
G value with Williams' correction	3.23	$P = 0.072$

For the second half of exon five:

	Fixed differences between species	Polymorphic sites
Synonymous Substitutions	2	21
Nonsynonymous Substitutions	9	20
G value	4.15	$P = 0.042$
G value with Williams' correction	3.96	$P = 0.047$

Chapter four: Additional results and discussions

In this part, I summarize the main data that have been referred to in Chapters two and three. Tables 3 to 5 present the data and the molecular markers used in the QTL mapping. The results of QTL mapping are presented and discussed in the Chapter two. Tables 6 and 7 present the molecular polymorphisms in *cpo* exon five, from which the population genetic studies of the chapter three are based upon.

Figures 14 and 15 summarize the data of five different SNP across *cpo* exon five and the observed latitudinal cline in the allele frequency of these loci. The observed allele frequency of SNP BsiE1 shows a strong agreement of the diapause frequency among those populations reported in Schmidt *et al*. 2005. Not only is the slope of the cline but also the y–intercept highly similar between the observed and expected. This indicates that polymorphisms in *cpo* are the major source of variation of reproductive diapause in these surveyed US East Coast populations.

Figure 16 is a visual presentation of *cpo* alleles generated by using *P*-element excision. These alleles are used in the genetic complementation experiments reported in the chapter two.

Table 3.

The data of QTL mapping. RIL (Recombinant Inbreed Line). ND = non-diapause. D = Diapause. 1 and 0 indicate the genotype of parental lines *6326* and *VT46*, respectively.

RIL	Marker															Phenotype
	61C	64E	65D	68C	75F	83B	85D	87E	89A	90A	90D	92D	94D	98B	99A	
RIL006	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	ND
RIL007	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL009	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	ND
RIL010	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	ND
RIL015	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	ND
RIL017	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	ND
RIL019	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	ND
RIL028	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	ND
RIL029	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL030	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	ND
RIL034	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	ND
RIL036	1	0	0	0	0	1	1	1	0	0	0	0	1	0	0	ND
RIL040	1	1	1	1	1	1	1	1	0	0	0	0	0	0	1	ND
RIL041	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	ND
RIL043	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	ND
RIL044	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	ND
RIL057	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	ND
RIL061	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL064	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	ND
RIL065	1	0	1	1	1	1	1	1	0	0	0	0	0	0	0	ND
RIL069	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL071	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL073	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL078	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	ND
RIL079	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL080	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL083	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL087	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL088	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	ND
RIL089	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL090	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	ND
RIL092	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL095	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL097	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL100	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	ND

RIL	Marker															Phenotype
	61C	64E	65D	68C	75F	83B	85D	87E	89A	90A	90D	92D	94D	98B	99A	
RIL102	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	ND
RIL107	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL108	1	0	0	0	0	0	0	0	0	0	0	1	1	1	1	ND
RIL110	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	ND
RIL114	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	ND
RIL115	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	ND
RIL116	0	0	1	1	1	1	1	1	1	1	1	0	0	0	0	ND
RIL121	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL122	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	ND
RIL123	1	1	1	1	0	0	0	0	0	0	0	0	0	1	1	ND
RIL127	1	0	0	0	0	0	0	0	0	0	0	1	1	1	1	ND
RIL128	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL130	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	ND
RIL131	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	ND
RIL132	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	ND
RIL133	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL134	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	ND
RIL135	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	ND
RIL141	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL146	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	ND
RIL149	0	1	1	1	0	0	0	0	0	0	0	0	0	1	1	ND
RIL151	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	ND
RIL159	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	ND
RIL164	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL168	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	ND
RIL170	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL172	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL174	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	ND
RIL175	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	ND
RIL177	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL180	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	ND
RIL181	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	ND
RIL182	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	ND
RIL183	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	ND
RIL185	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL187	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ND

RIL	Marker															Phenotype
	61C	64E	65D	68C	75F	83B	85D	87E	89A	90A	90D	92D	94D	98B	99A	
RIL188	1	0	1	0	0	1	0	1	0	0	0	0	0	0	0	ND
RIL189	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL191	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL192	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL194	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1	ND
RIL195	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	ND
RIL203	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	ND
RIL208	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ND
RIL001	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL002	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	D
RIL008	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL011	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL013	1	1	1	0	0	0	0	0	1	1	1	1	1	1	0	D
RIL014	0	0	0	1	1	1	1	1	1	1	1	1	0	0	0	D
RIL018	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	D
RIL020	1	0	0	0	1	1	1	1	1	1	1	1	1	0	0	D
RIL021	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	D
RIL022	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL023	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL024	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	D
RIL025	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL026	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL027	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL031	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL033	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	D
RIL035	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	D
RIL038	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL042	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL046	0	0	0	1	1	1	1	1	1	1	1	1	0	0	0	D
RIL049	0	1	0	0	0	1	1	1	1	1	1	1	1	0	0	D
RIL056	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL070	0	0	0	0	0	1	0	1	1	1	1	1	0	0	0	D
RIL074	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL077	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL081	0	0	0	1	0	1	1	1	1	1	1	1	1	0	0	D
RIL084	1	1	0	1	1	1	1	1	1	1	1	1	1	0	0	D

RIL	Marker															Phenotype
	61C	64E	65D	68C	75F	83B	85D	87E	89A	90A	90D	92D	94D	98B	99A	
RIL096	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	D
RIL101	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL103	0	1	1	1	1	0	1	0	0	0	0	0	1	1	0	D
RIL105	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL109	1	0	0	0	1	0	1	1	1	1	1	1	1	1	0	D
RIL111	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL112	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL113	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	D
RIL117	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL118	1	0	0	1	1	1	1	1	1	1	1	1	1	0	0	D
RIL119	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL120	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	D
RIL126	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	D
RIL137	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL138	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	D
RIL140	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	D
RIL142	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	D
RIL143	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	D
RIL144	0	0	0	0	1	1	1	1	1	1	1	1	1	0	0	D
RIL150	0	0	0	1	1	1	1	1	1	1	1	1	1	1	0	D
RIL152	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	D
RIL153	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	D
RIL154	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL158	1	1	0	1	1	1	1	1	1	1	1	1	0	0	0	D
RIL160	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	D
RIL161	1	1	1	1	0	1	1	1	1	1	1	1	1	0	0	D
RIL162	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL163	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	D
RIL167	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	D
RIL169	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL173	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL184	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	D
RIL186	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	D
RIL190	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	D
RIL193	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL196	1	0	0	0	1	1	1	1	1	1	1	1	1	1	0	D

RIL	Marker														Phenotype
	61C	64E	65D	68C	75F	83B	85D	87E	89A	90A	90D	92D	94D	98B	99A	
RIL198	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	D
RIL199	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	D
RIL200	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	D
RIL201	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL202	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL209	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	D
RIL211	1	0	0	0	0	1	1	1	1	1	1	1	1	1	0	D
RIL212	1	1	0	0	0	0	0	0	0	0	1	1	1	1	1	D
RIL157	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL039	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	D
RIL155	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL139	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL124	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL136	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL148	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL178	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL179	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL094	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL154	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL156	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL166	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL147	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL091	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL197	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL005	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	D
RIL012	1	0	0	1	0	0	1	1	1	1	1	1	1	0	0	D
RIL037	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL050	0	1	0	0	0	1	1	1	1	1	1	1	1	1	0	D
RIL052	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	D
RIL053	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	D
RIL054	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL055	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	D
RIL059	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	D
RIL063	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL066	1	1	1	1	0	0	0	0	0	1	1	1	1	1	1	D
RIL067	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	D

RIL	Marker														Phenotype
	61C	64E	65D	68C	75F	83B	85D	87E	89A	90A	90D	92D	94D	98B	99A	
RIL068	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	D
RIL072	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	D
RIL075	1	1	1	0	0	0	0	0	1	1	1	1	1	1	0	D
RIL076	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	D
RIL082	1	0	1	1	1	1	1	1	1	1	1	1	1	0	0	D
RIL085	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	D
RIL125	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	D
RIL158	1	1	0	1	1	1	1	1	1	1	1	1	0	0	0	D
RIL165	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	D
RIL176	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	D
RIL210	1	0	0	0	0	0	0	0	0	0	1	1	1	1	1	D

Table 4.

The sequences of QTL mapping marker primers, their genomic and cytologic position on the chromosome. If the marker is of a restriction endonuclease digestion site polymorphism, the type of restriction endonuclease is indicated in the ‘type’ column.

Name	Genomic Position		Primer Pair		Type	Map Position
Ptpmeg	release4:3L	336376	CCGCAGATTAAGCTTATAGTAGCC	CTGAATGTCATACTCGTCGCTC	NsiI	61C
Roo64	AE003634	306913			Insertion	64E12-65A6
Roo65	AE003630	157164			Insertion	65D3-E11
3L098	release4:3L	11393279	CAGCCGCTGGAAATGCGATTATTGTCACGGTCAT	GGTCGGGGGCCCTTCAATCAGCTGTCAAATG	ClaI	68C7
3L164	release4:3L	19075392	CCTCCTGCTCGCTGCCACTTTTG	CGCAACTGAACGAAGCCCTCCTG	ClaI	75F6-F7
3R012	release4:3R	1761010	AAGGAAACGAATTAAAGGCAGACCCA	TGGGAAAAGGGAACGTTAAAGAGCA	Size	83B-83C
3R034	release4:3R	4985049	CGAGAACCGAGAAATGTTGGGGAAA	CGATCCGAAAGAGAGGGCTGCTAAAA	HindIII	85D-85D
3R065	release4:3R	9181732	GAGACGTGATAGCATTCGACCGACAC	TCCGCCCACTGAAGACACAATTACAC	In-Del	87E-87F
3R082	release4:3R	11589546	CGGACAATGAGTGGGACAGAGAGACA	AAGCGGAGCAGATGGGCGATAAA	DraI	89A1-A5
3R092	release4:3R	13035415	GCGACAGCGCAAAAACTCCTGT	AAGATCATTCTCACGTTCCTCACGATG	In-Del	89E-90A
3R105	release4:3R	13770615	GGGAGTCGCTTGGCTGATGAAGAAG	TCGGGCAATGGGCATGAAATAAAAG	XbaI	90D-90D
3R151	release4:3R	16241996	CAGCGGCAGGCAAAGTCATAAAAGTC	CAAACGGGACAAAAGTGAGAGCGAAA	ClaI	92D-92D
3R170	release4:3R	16878096	TTGCAGATGCTCACTTCGATTAAGAGG	GAATCTCACCCAACGAAACCAACTAAGA	SpeI	94D-94D
3R214	release4:3R	23783244	TCGCTGACGGAGAGGATTGAGAC	GATAGCCCCGAAAACGCACAAAA	BglII	98B-98C
3R223	release4:3R	25168765	TGAGAGTTTCTTTGGGCGGTAGTGGA	ACGGGAGCCTGACGATTGAGGA	AviII	99A-99B5

Table 5.

The data of QTL mapping in 90D-91D region. RIL (Recombinant Inbred Line). ND = non-diapause. D = Diapause. 1 and 0 indicate the genotype of parental diapausing line *6326* and non-diapausing *VT46*, respectively. 1/0in RIL 60 at marker *cpo*+50kb represents the actual recombination event was detected within this marker. Markers in *cpo* gene region are named by their distance from the beginning of the gene.

Phenotype	RIL	61C	64E	65D	68C	75F	83B	85D	87E	89A	90A	90D	*Cpo* +13kb	*Cpo* +34kb	*Cpo* +50kb	*Cpo* +54kb	*Cpo* +57kb	*Cpo* +64kb	*Cpo* +68kb	*Cpo* +71kb	*Cpo* +77kb	*DNAse Htl* 90F	*DNAse Htl* 91D	92D	94D	98B	99A
ND	60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0
D	86	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
ND	17	0	0	0	1	1	1	1	1	1	1	1	1	1	1/0						0	0	0	0	0	0	0
ND	130	0	0	0	0	0	0	0	0	0	0	0	0	0	0						0	1	1	1	1	1	1
ND	108	1	0	0	0	0	0	0	0	0	0	0	1									1	1	1	1	1	1
D	200	0	0	0	0	0	0	0	0	0	0	0	1									1	1	1	1	1	1
ND	134	0	0	0	0	0	0	0	0	0	0	0										0	1	1	1	0	0
ND	28	0	0	0	0	0	0	0	0	0	0	0											0	1	0	1	1
ND	48	1	1	1	0	0	0	0	0	0	0	0											0	1	1	0	0
ND	47	1	0	0	0	0	0	0	0	0	0	0											0	1	1	1	1
ND	127	1	0	0	0	0	0	0	0	0	0	0											0	1	1	1	1
ND	171	0	0	0	0	0	0	0	0	0	0	0											0	0	1	0	1
ND	203	0	0	0	0	0	0	0	0	0	0	0											0	1	1	0	1
D	68	1	1	1	1	1	1	1	1	1	1	1											1	1	0	0	0
D	146	1	1	1	1	1	1	1	1	1	1	1											1	0	0	0	1
D	116	0	0	1	1	1	1	1	1	1	1	1											1	1	0	0	0
D	145	1	0	0	1	1	1	1	1	1	1	1											1	0	0	0	0
D	165	0	0	0	1	1	1	1	1	1	1	1											1	0	0	1	1
D	167	0	0	0	1	1	1	1	1	1	1	1											1	0	0	1	1
D	138	0	0	0	1	1	1	1	1	1	1	1											1	0	0	0	0
D	21	0	0	0	0	1	1	1	1	1	1	1											1	1	0	1	0

Figure 14.

The cline of frequencies in 5 different SNP's *cpo*. The population samples are from: HFL (Homestead, FL; latitude 25° 2'), MFL (Merrit Island, FL; 28° 3'), JFL (Jacksonville, FL; 30° 2'), SC (Eutawville, SC; 33° 2'), NC (Smithfield, NC; 35° 3'), VA (Richmond, VA; 37° 3'), MD (Churchville, MD; 39° 3'), CT (Middlefield, CT; 41° 3'), MA (Concord, MA; 42° 0'), DPF (Wading River, NY, 40° 57') and VT (Whiting, VT 43° 6'). For each population, 48 extracted chromosomes were genotyped expect for DPF (n = 46). The P values of regression coeffients are: Exon1-TspR1, $P < 0.040$; Exon5-Afe1, $P < 0.621$, Exon5-BsiE1, $P < 0.001$, Exon6-Dde1-Upstream, $P < 0.371$; and Exon6-Dde1-downstream, $P < 0.005$.

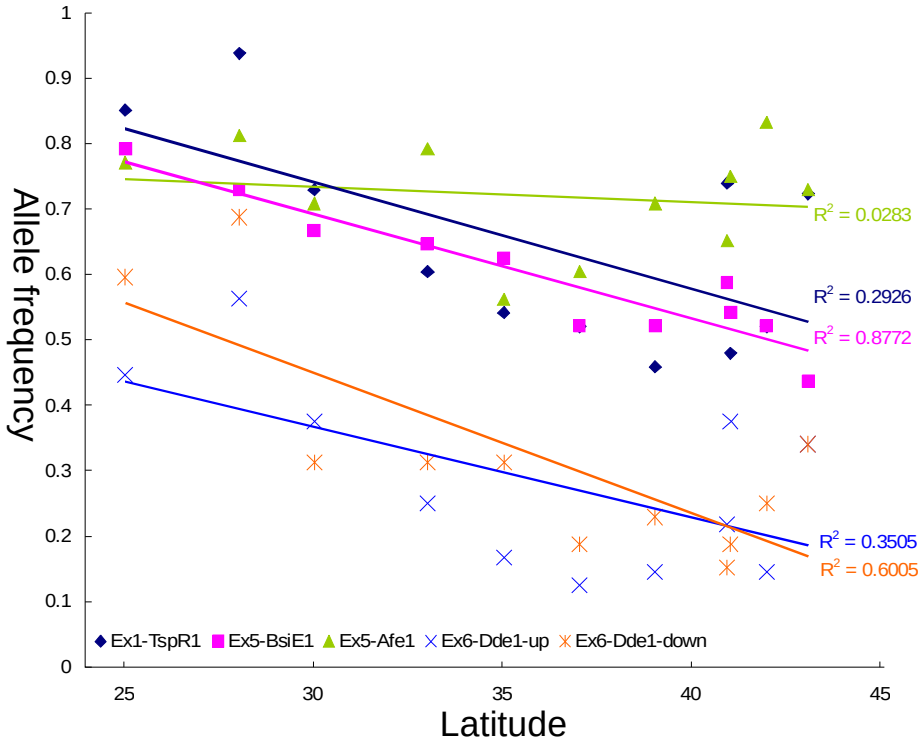

Allele frequency
Latitude
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
25
30
35
40
45
Ex1-TspR1
Ex5-BsiE1
Ex5-Afe1
Ex6-Dde1-up
Ex6-Dde1-down
R2 = 0.0283
R2 = 0.2926
R2 = 0.8772
R2 = 0.3505
R2 = 0.6005

Figure 15.

Expected and observed latitudinal clines of diapause frequency. The expected diapause frequency is calculated from observed allele frequency of Val 363 with the following three assumptions: Exon five BsiE1 (Val 363) and Lys 472 are in complete linkage disequilibrium, diapause is completely dominant over non-diapause and populations are in Hardy-Weinberg equilibrium. The observed diapause frequencies are from Schmidt *et al.* 2005a.

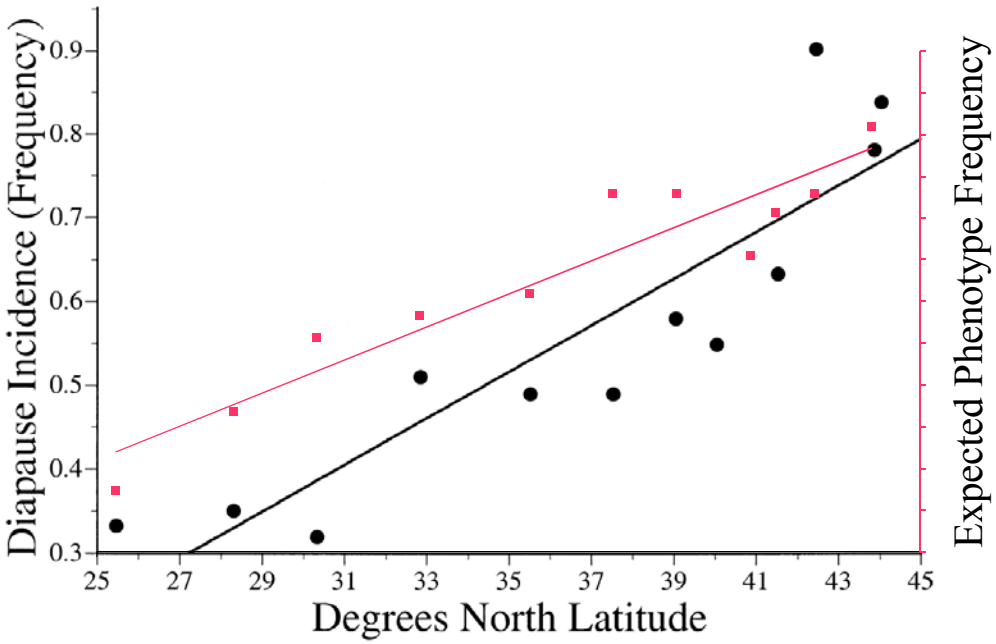

Diapause Incidence (Frequency)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
25
27
29
31
33
35
37
39
41
43
45
Degrees North Latitude
Expected Phenotype Frequency

Figure 16.

The *P-element* excision alleles generated for genetic complementation studies of Chapter two. The figure is draw to scale. Line *P1*, *P2* and *P10* are insertions of a 38 bp small DNA fragment: CAGGCCATGATGAAATAACATGTTATTTCATCATGGT. Line *D7* is an insertion of a 60bp small DNA fragment: CAGGCCATGATGAAATAACATGTTATATATATAACATGTTATGTTATTTCA TCATGGT. Line *P7* is an insertion of a 167bp small DNA fragment: CAGGCCATGATGAAATAACATAAGGTGGTCCCGTCGGCAAGAGACATC CACTTAACGTATGCTTGCAATAAGTGCGAGTGAAAGGAATAGTATTCTG AGTGTCGTATTGAGTCTGAGTGAGACAGCGATATGATTGTTGATTAACC CTTAGCATGTCCGTGGGGT. Lines *P5* and *P8* are insertions of 744 bp and 816 bp respectively. Origins of the insertion sequences are unkown. Line *P13* and *P14* are large insertions. Part of the inserted seqeuence are probably truncated *P-elements*. DNA sequences confirmed the absence of *P-element* repeat sequence in those two alleles, thus they should be stable alleles without *P-element* activity and lack of the complete *mini-white* gene. Lines *P3*, *P4*, *P6*, *P11* and *P12* are perfect excisions of $cpo^{BG02810}$. Lines *P1* to *P14* are homozygous viable lines and line *D1* to *D12* are lethal lines. The sequence features for *D1* to *D12* are unknown except for *D7*.

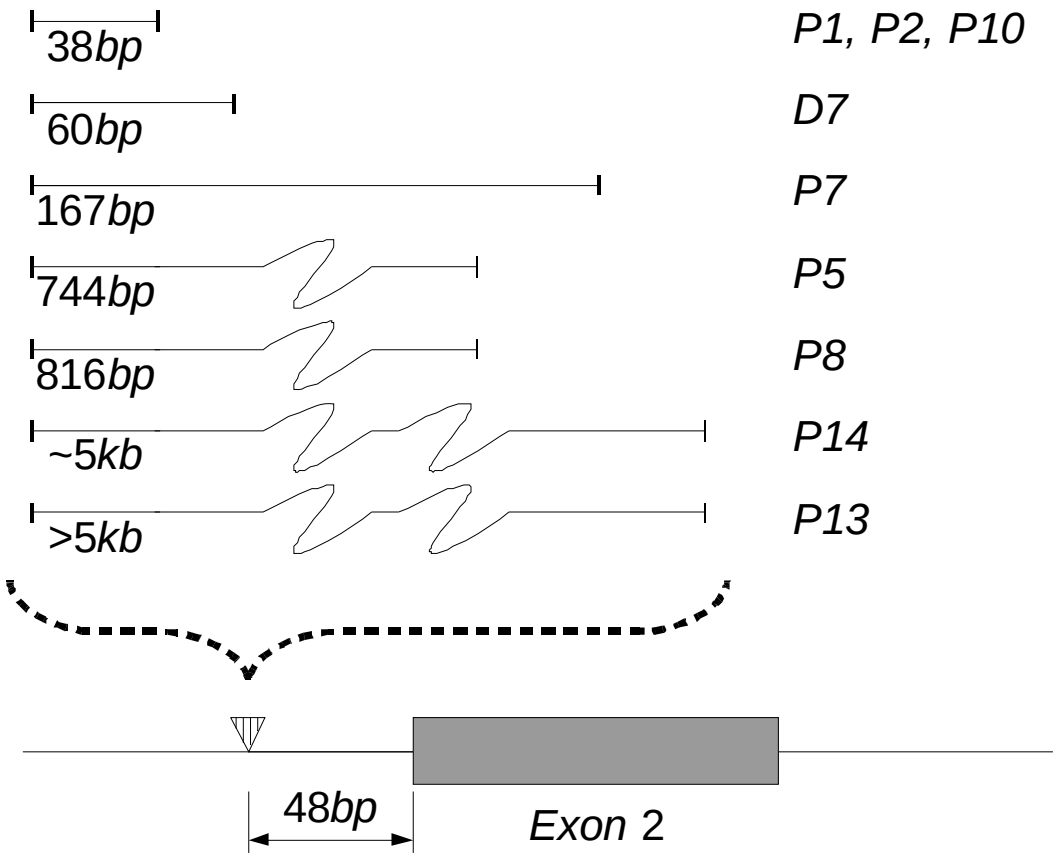
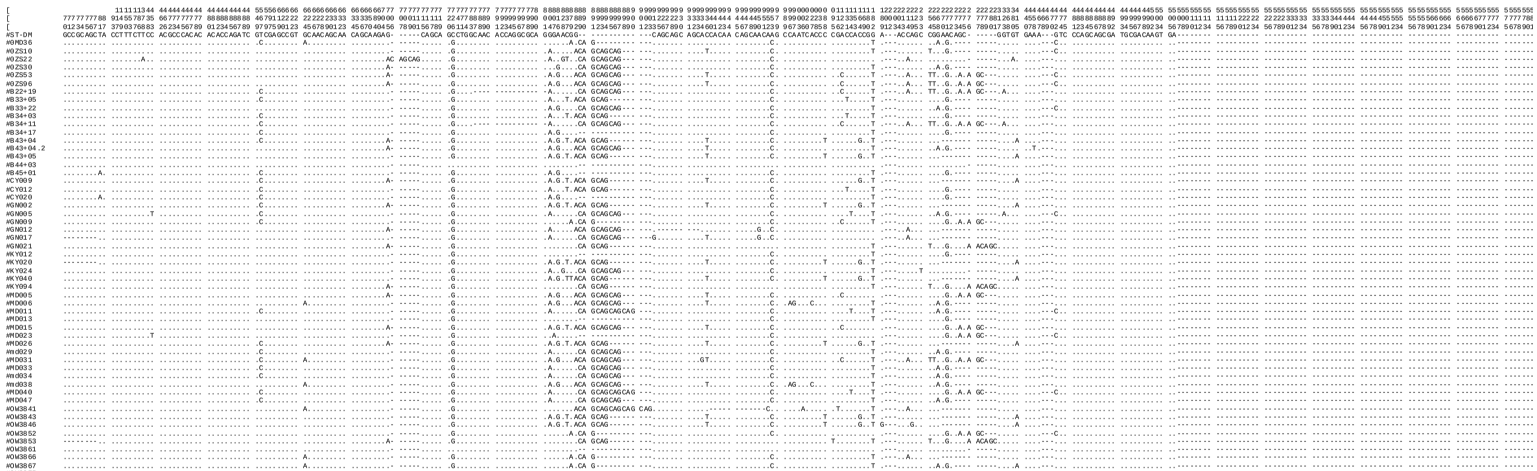

38bp
P1, P2, P10
60bp
D7
167bp
P7
744bp
P5
816bp
P8
~5kb
P14
>5kb
P13
48bp
Exon 2

Figure 17.

The DNA polymorphisms of *cpo* exon five. The lines: MD (Cameron, Africa), ZS, ZK and ZW (Zimbabwe, Africa), CY (Taiwan, China), B (the Bahamas), KY (Kenya, Africa), GN (Gabon, Africa), DP (DPF, New York, US), VT (Vermont, US), HFL (Florida, US), RR (Maine, US), 6326 (*w1118*, Bloomington Stock 6326), OW (Old World lines in Bloomington Stock, 'OW" followed by line number). DP96, CT96 and GA96 (*D. simulans* inbred lines) and ST-DM (*D. melanogaster* genome sequence). Line numbers starting with lower case indicate the lines were sequenced by cloning. Line numbers starting with 0 indicate an inbred line was sequenced for that strain. Of the rest of the *D. melanogaster* lines, the third chromosomes were first extracted using *TM6/TM8* and one single homozygote is used for sequencing. '.' indicates identical sequence. '-' indicates detection. Base pair 1 is the start of translation for *cpo* isoform PH.



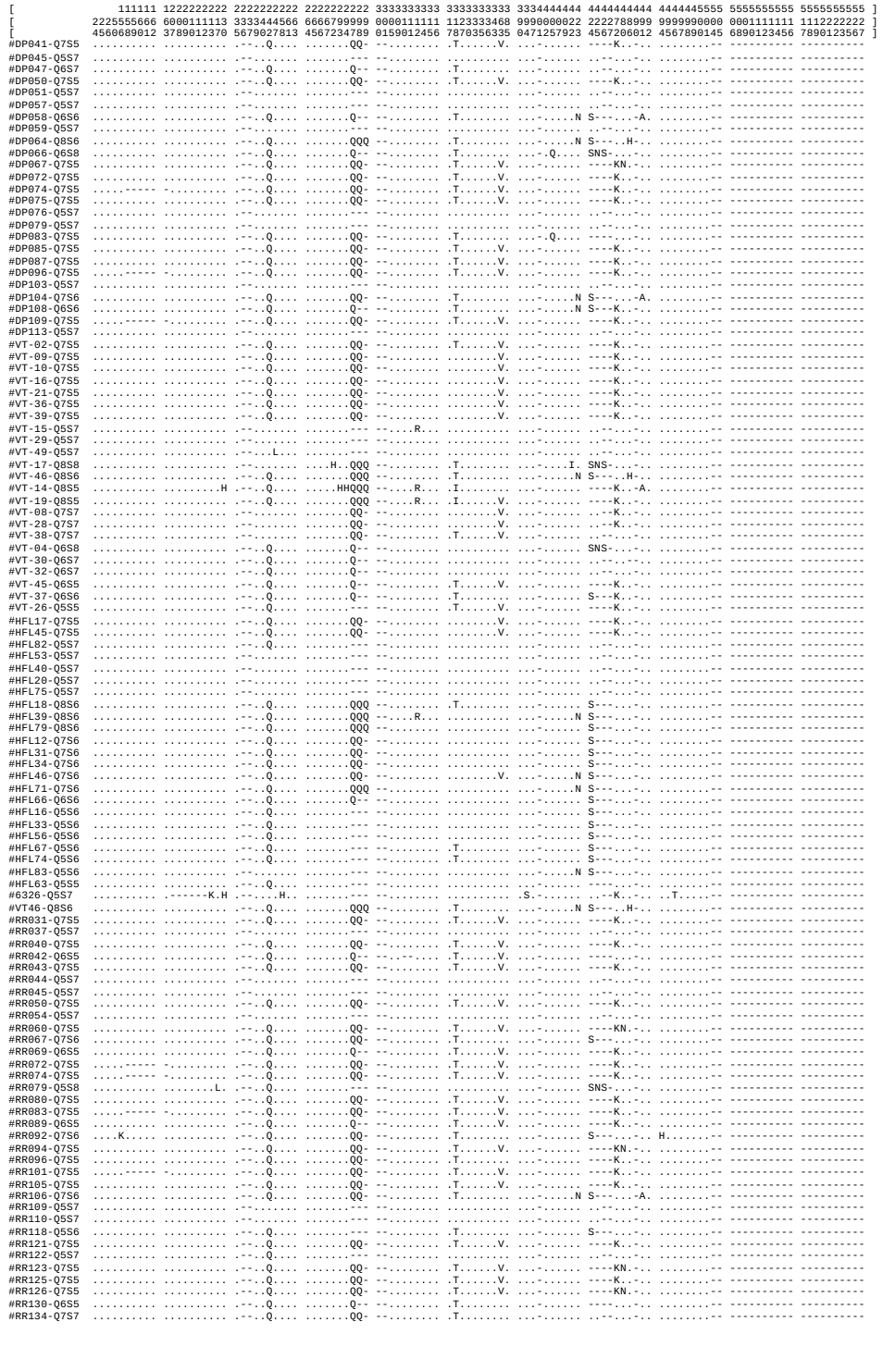
Continue...

Figure 18.

The amino acid polymorphisms of *cpo* exon five. Generated by translation the DNA polymorphism table (Figure 17) using the standard genetic code.

```
[              111111 1222222222 2222222222 2222222222 3333333333 3333333333 3334444444 4444444444 4444445555 5555555555 5555555555
[           2225555666 6000111113 3333444566 6666799999 0000111111 1123333468 9990000022 2222788999 9999990000 0001111111 1112222222
[           4560689012 3789012370 5679027813 4567234789 0159012456 7870356335 0471257923 4567206012 4567890145 6890123456 7890123567
#ST-DM      AAAVQAHTHQ IAVQQQQQQQ Q--QQHFQQQ HQAQQQQ--- --LSSSSHNS NNAAITSTAT AAT-TPTASS NS--IKQ-GS PSSEMRQK-- ---------- ----------
#0MD36      .......... .......... ....Q.... .......Q-- --........ .T........ ...-.....N S---...-A. ......-- ---------- ----------
#0ZS10      .......... .......... ....Q.... .......QQQ --........ .T........ ...-...... S---...-A. ......-- ---------- ----------
#0ZS22      .......... .......... .QQ..Q.... .......QQQ --........ .T........ ....-.Q.... ..---...-. ......-- ---------- ----------
#0ZS30      .......... .......... ....Q.... .......QQQ --........ .T........ ...-.....N S---...-. ......-- ---------- ----------
#0ZS53      .......... .......... ....Q.... .......QQQ --........ .T........ ....-.Q..I. SNS-....A. ......-- ---------- ----------
#0ZS96      .......... .......... ....Q.... .......QQQ --........ .T........ ....-.Q..I. SNS-....A. ......-- ---------- ----------
#B22+19     .......... .......... ....Q.... -------QQQ --........ .T........ ....-.Q..I. SNS-...-. ......-- ---------- ----------
#B33+05     .......... .......... ....Q.... .......QQ- --........ .T........ ...-...... S---...-. ......-- ---------- ----------
#B33+22     .......... .......... ....Q.... .......QQQ --........ .T........ ...-.....N S---...-A. ......-- ---------- ----------
#B34+03     .......... .......... ....Q.... .......QQ- --........ .T........ ...-...... S---...-. ......-- ---------- ----------
#B34+11     .......... .......... ....Q.... ------.QQQ --........ .T........ ....-.Q..I. SNS-...-. ......-- ---------- ----------
#B34+17     .......... .......... ....Q.... ........--- --........ .T........ ...-...... S---...-. ......-- ---------- ----------
#B43+04     .......... .......... ....Q.... .......QQ- --........ .T......V. ...-...... ----K..-. ......-- ---------- ----------
#B43+04.2   .......... .......... ....Q.... .......QQQ --........ .T........ ...-.....N S---..H-. ......-- ---------- ----------
#B43+05     .......... .......... ....Q.... .......QQ- --........ .T......V. ...-...... ----K..-. ......-- ---------- ----------
#B44+03     .......... .......... ....Q.... ........--- --........ .T........ ...-...... ....-..-. ......-- ---------- ----------
#B45+01     .......... .......... ....Q.... ........--- --........ .T........ ...-...... S---...-. ......-- ---------- ----------
#CY009      .......... .......... ....Q.... .......QQ- --........ .T........ ...-...... ----K..-. ......-- ---------- ----------
#CY012      .......... .......... ....Q.... .......QQ- --........ .T........ ...-...... S---...-. ......-- ---------- ----------
#CY020      .......... .......... ....Q.... ........--- --........ .T........ ...-...... S---...-. ......-- ---------- ----------
#GN002      .......... .......... ....Q.... .......QQ- --........ .T......V. ...-...... ----K..-. ......-- ---------- ----------
#GN005      .......... .......... ....Q.... .......QQQ --........ .T........ ...-.....N S-------A. ......-- ---------- ----------
#GN009      .......... .......... ....Q.... .......Q-- --........ .T........ ...-...... SNS-...-. ......-- ---------- ----------
#GN012      .......... .......... ....Q.... .......QQQ --.---.... ST........ ....-.Q.... ----...-. ......-- ---------- ----------
#GN017      .......... .......... ....Q.... .......QQQ --V..... ST........ ....-.Q.... ----...-. ......-- ---------- ----------
#GN021      .......... .......... ....Q.... .......QQ- --........ .......... ...-...... S.NS...-. ......-- ---------- ----------
#KY012      .......... .......... ....Q.... ........--- --........ .T........ ...-...... S---...-. ......-- ---------- ----------
#KY020      .......... .......... ....Q.... .......QQ- --........ .T......V. ...-...... ----K..-. ......-- ---------- ----------
#KY024      .......... .......... ....Q.... .......QQQ --........ .T........ ...-...... ....-..-. ......-- ---------- ----------
#KY040      .......... .......... ....Q.... .....H.QQ- --........ .T......V. ...-...... ----K..-. ......-- ---------- ----------
#KY094      .......... .......... ....Q.... .......QQ- --........ .......... ...-...... S.NS...-. ......-- ---------- ----------
#MD005      .......... .......... ....Q.... .......QQQ --........ .T........ ...-...... SNS-...-. ......-- ---------- ----------
#MD006      .......... .......... ....Q.... .......QQQ --........ .T..V..P.. ...-.....N S---...-. ......-- ---------- ----------
#MD011      .......... .......... ....Q.... .......QQQ Q-...... .T........ ...-.....N S------A. ......-- ---------- ----------
#MD013      .......... .......... ....Q.... ........--- --........ .T........ ...-...... S---...-. ......-- ---------- ----------
#MD015      .......... .......... ....Q.... .......QQQ --........ .T........ ...-...... SNS-...-. ......-- ---------- ----------
#MD023      .......... .......... ....Q.... ........--- --........ .T........ ...-...... SNS-....A. ......-- ---------- ----------
#MD026      .......... .......... ....Q.... .......QQ- --........ .T......V. ...-...... ----K..-. ......-- ---------- ----------
#md029      .......... .......... ....Q.... .......QQQ --........ .T........ ...-.....N S---...-. ......-- ---------- ----------
#MD031      .......... .......... ....Q.... .......QQQ --.....G... .T........ ....-.Q..I. SNS-...-. ......-- ---------- ----------
#MD033      .......... .......... ....Q.... .......QQQ --........ .T........ ...-.....N S---...-. ......-- ---------- ----------
#md034      .......... .......... ....Q.... .......QQQ --........ .T....... ...-.....N S---...-. ......-- ---------- ----------
#md038      .......... .......... ....Q.... .......QQQ --........ .T..V..P.. ...-.....N S---...-. ......-- ---------- ----------
#MD040      .......... .......... ....Q.... .......QQQ Q-...... .T........ ...-...... SNS-...-A. ......-- ---------- ----------
#MD047      .......... .......... ....Q.... .......QQQ --........ .T........ ...-.....N S---...-. ......-- ---------- ----------
#OW3841     .......... .......... ....Q.... .......QQQ QQ..---.. -H....T... ....-.Q.... ----...-. ......-- ---------- ----------
#OW3843     .......... .......... ....Q.... .......QQ- --........ .T......V. ...-...... ----K..-. ......-- ---------- ----------
#OW3846     .......... .......... ....Q.... .......QQ- --........ .T......V. ..A--..A... ----K..-. ......-- ---------- ----------
#OW3852     .......... .......... ....Q.... .......Q-- --........ .T........ ..-...... SNS-...-A. ......-- ---------- ----------
#OW3853     .......... .......... ....Q.... .......QQ- --........ .......... ...-...... S.NS...-. ......-- ---------- ----------
#OW3861     .......... .......... ........ ........--- --........ .......... ...-...... ..-...-. ......-- ---------- ----------
#OW3866     .......... .......... ....Q.... .......Q-- --........ .T........ ....-.Q.... ----...-. ......-- ---------- ----------
#OW3867     .......... .......... ....Q.... .......Q-- --........ .T........ ...-.....N S---K..-. ......-- ---------- ----------
#OW3875     .......... .......... ........ ........--- --........ .......... ...-...... ..-...-. ......-- ---------- ----------
#OW3876     .......... .......... ........ ........--- --........ .......... ..A-AA.... ..-...-. ......-- ---------- ----------
#OW4268     .......... .......... ........ ........--- --........ .......... ...-...... ...-K..-. ......-- ---------- ----------
#OW4271     .......... .......... ....Q.... .......Q-- --........ .T........ ...-.....N S---K..-. ......-- ---------- ----------
#ZK058      .......... .......... ....Q.... .......QQQ --........ .T........ ....-.Q..I. SNS-...-. ......-- ---------- ----------
#ZK081      .......... .......... ....Q.... ........--- --........ .......... ...-...... ..-...-. ......-- ---------- ----------
#ZK091      .......... .......... ....Q.... .......QQQ --........ .T........ ...-.....I. SNS-...-. ......-- ---------- ----------
#ZK095      .......... .......... ....Q.... .......Q-- --........ .T........ ...-.....N S---...-. ......-- ---------- ----------
#ZK125      .......... .......... ....Q.... .......QQQ --........ .T........ ...-...... ..-...-. ......-- ---------- ----------
#ZK145      .......... .......... ....Q.... .......QQQ --........ .T........ ....-.Q..I. SNS-...-. ......-- ---------- ----------
#ZK184      .......... .......... ....Q.... .......Q-- --........ .T........ ...-.....N S---K..-. ......-- ---------- ----------
#ZK210      .......... .......... ....Q.... .......QQQ --........ .T........ ...-...... ..-...-.F ......-- ---------- ----------
#ZK216      .......... .......... ....Q.... .......QQQ Q-...... .T........ ...-...... SNS-...-. ......-- ---------- ----------
#ZK224      ..........- -......... ........ ........--- --........ .......... ...-...... ..-...-. ......-- ---------- ----------
#ZK266      ..........- -......... ........ ........--- --........ .......... ...-...... ..-...-. ......-- ---------- ----------
#ZS006      .......... .......... ....Q.... ........--- --........ .......... ...-...... S---...-. ......-- ---------- ----------
#ZS011      .......... .......... ....Q.... .......QQQ --........ .T........ ....-.Q..I. SNS-...-. ......-- ---------- ----------
#ZS028      .......... .......... ....Q.... .......QQQ --........ .T........ ....-.Q..I. SNS-...-. ......-- ---------- ----------
#ZS037      .......... .......... ........ ........--- --........ .......... ...-...... ..-...-. ......-- ---------- ----------
#ZS042      .......... .......... ....Q.... .......QQQ --........ .T........ ....-.Q.PI. SNS-...-. ......-- ---------- ----------
#ZS044      .......... .......... ....Q.... .......Q-- --........ .T........ ...-...... SNS-....A. ......-- ---------- ----------
#ZW104      ..........- -......... ........ ........--- --........ .......... ...-...... ..-...-. ......-- ---------- ----------
#ZW123      .......... .......... ....Q.... .......QQQ --........ .T........ ...-.....N S---K..-A. ......-- ---------- ----------
#ZW133      .......... .......... ....Q.... .......QQQ --........ .T........ ...-.....N S----..H-. ......-- ---------- ----------
#ZW165      .......... .......... ....Q.... .......QQQ --........ .T........ V..-.Q..I. SNS-...-. ......-- ---------- ----------
#ZW186      .......... .......... ....Q.... .......QQQ --........ .T........ ...-...... SNS-...-. ......-- ---------- ----------
#DP96-2SA   ...A..... .......... ....Q..H. .......QQ- --.-.-.... .T...A.S.A ...A.T.... ----T.P-EA .AAKCDISLG NESCILKYCV NRWKVAQSTG
#DP96-4SB   ...A..... .......... .....Q..H. .......QQ- --.---.... TTT..A.S.A .---.T.... ----T.P-EA .AAKCDKSLG NESCILKYCI NRWKVAQSRG
#DP96-2SB   ...A..... .......... ....Q..H. .......QQ- --.---.... .T...A.S.A ...A.T.... ----T.P-EA .AAKCDISLG NESCILKYCV NRWKVAESRG
#CT96-4SB   ...A..... .......... ....Q..H. .......QQ- --.---.... .T...A.S.A ...A.T.... ----T.P-EA .AAKCDISLG NESCILKYCI NRWKVAESRG
#CT96-6SA   ...A..... .......... ....Q..H. .......QQ- --.---.... .T...A.S.A ...A.T.... ----T.P-EA .AAKCDISLG NESCILKYCI NRWKVAQSTG
#DP96-5SA   ...A..... .......... ....Q..H. .......QQ- --.---.... .T.G.A.S.A ...A.T.... ----T.PKEA .AAKCDKSLG NESCILKYCI NRWKVAQSPG
#GA96-5SA   ...A..... .......... ....Q..H. .......QQ- --.---.... .TT..A.S.A ...A.T.... ----T.P-EA .AAKCDISLG NESCILKYCI NRWKVAQSTG
#DP001-Q7S6 .......... .......... ....Q.... .......QQ- --........ .T........ ...-...... S---..... H.......... ---------- ----------
#DP004-Q7S5 .......... .......... ....Q.... .......QQ- --........ .T......V. ...-...... ----K..-. ......-- ---------- ----------
#DP006-Q8S8 .......... .......... ....Q.... ------.QQQ --........ .T........ ....-.Q..I. SNS-...-. ......-- ---------- ----------
#DP008-Q7S6 .......... .......... R---Q.... .......QQ- --........ .T......V. ...-...... S---...-. ......-- ---------- ----------
#DP010-Q5S7 .......... .......... ........ ........--- --........ .......... ...-...... ..-...-. ......-- ---------- ----------
#DP012-Q5S7 .......... .......... ....Q.... ........--- --........ .......... ...-...... ..-...-. ......-- ---------- ----------
#DP013-Q7S5 ..........- -......... ....Q.... .......QQ- --........ .T......V. ...-...... ----K..-. ......-- ---------- ----------
#DP016-Q7S5 .......... .......... ....Q.... .......QQ- --........ .T......V. ...-...... ----K..-. ......-- ---------- ----------
#DP017-Q5S7 .......... .......... ........ ........--- --........ .......... ...-...... ..-...-.Y ......-- ---------- ----------
#DP018-Q7S5 .......... .......... ....Q.... .......QQ- --........ .T......V. ...-...... ----K..-. ......-- ---------- ----------
#DP019-Q5S7 .......... .......... ....Q.... ........--- --........ .......... ...-...... ..-...-. ......-- ---------- ----------
#DP020-Q5S7 .......... .......... ........ ........--- --........ .......... ...-...... ..-...-. ......-- ---------- ----------
#DP022-Q7S6 .......... .......... ....Q.... .......QQ- --........ .T......V. ...-.....N S---K..-. ......-- ---------- ----------
#DP023-Q5S7 .......... .......... ....Q.... ........--- --........ .......... ...-...... ..-...-. ......-- ---------- ----------
#DP024-Q7S8 .......... .......... ....Q.... .......QQ- --........ .T......V. ...-...... SNS-K..-. ......-- ---------- ----------
#DP027-Q8S6 .......... .......... ....Q.... .......QQQ --........ .T........ ...-.....N S---...-. ......-- ---------- ----------
#DP029-Q5S7 .......... .......... ....Q.... ........--- --........ .......... ...-...... ..-...-. ......-- ---------- ----------
#DP031-Q5S7 .......... .......... ........ ........--- --........ .......... ...-...... ..-...-. ......-- ---------- ----------
#DP035-Q7S7 .......... .......... ....Q.... .......QQ- --........ .T........ ...-...... ..-...-. ......-- ---------- ----------
#DP039-Q5S7 .......... .......... ........ ........--- --........ .......... ...-...... ..-...-. ......-- ---------- ----------
#DP044-Q7S7 .......... .......... ....Q.... .......QQ- --........ .T........ ...-...... ..-...-. ......-- ---------- ----------
#DP046-Q7S5 .......... .......... ....Q.... .......QQ- --........ .T......V. ...-...... ----KN.-. ......-- ---------- ----------
```

Continue...



References

ALBERT, S., T. TWARDZIK, M. HEISENBERG and S. SCHNEUWLY, 1997 Isolation and characterization of the droPIK57 gene encoding a new regulatory subunit of phosphatidylinositol 3-kinase from *Drosophila melanogaster*. Gene **198:** 181-189.

BARRETT, J. C., B. FRY, J. MALLER and M. J. DALY, 2005 Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics **21:** 263-265.

BASDEN, E. B., 1954 Diapause in *Drosophila* (Diptera: Drosophilidae). Proceedings of the Royal Entomology Society of London **29:** 114-118.

BAUDRY, E., N. DEROME, M. HUET and M. VEUILLE, 2006 Contrasted polymorphism patterns in a large sample of populations from the evolutionary genetics model *Drosophila simulans*. Genetics **173:** 759-767.

BELL, W. J., and M. K. BOHM, 1975 Oosorption in Insects. Biological Reviews of the Cambridge Philosophical Society **50:** 373-396.

BELLEN, H. J., S. KOOYER, D. DEVELYN and J. PEARLMAN, 1992a The *Drosophila* Couch Potato Protein Is Expressed in Nuclei of Peripheral Neuronal Precursors and Shows Homology to Rna-Binding Proteins. Genes & Development **6:** 2125-2136.

BELLEN, H. J., R. W. LEVIS, G. C. LIAO, Y. C. HE, J. W. CARLSON *et al.*, 2004 The BDGP gene disruption project: Single transposon insertions associated with 40% of *Drosophila* genes. Genetics **167:** 761-781.

BELLEN, H. J., H. VAESSIN, E. BIER, A. KOLODKIN, D. DEVELYN *et al.*, 1992b The *Drosophila Couch*-Potato Gene - an Essential Gene Required for Normal Adult Behavior. Genetics **131:** 365-375.

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.

BERTANI, G., 1947 Artificial Breaking of the Diapause in *Drosophila nitens*. Nature **159:** 309-309.

BOTELLA, L., and J. L. MENSUA, 1987 Larval Diapause Induced by Crowding in *Chymomyza costata* (Diptera, Drosophilidae). Annales Entomologici Fennici **53:** 41-47.

BOULETREAU-MERLE, J., and P. FOUILLET, 2002 How to overwinter and be a founder: egg-retention phenotypes and mating status in *Drosophila melanogaster*. Evolutionary Ecology **16:** 309-332.

BOULETREAU-MERLE, J., P. FOUILLET and J. VARALDI, 2003 Divergent strategies in low temperature environment for the sibling species *Drosophila melanogaster* and *D. simulans*: overwintering in extension border areas of France and comparison with African populations. Evolutionary Ecology **17:** 523-548.

BURLA, H., 1951 Systematik, Verbreitung und Oekolohie der *Drosophila*-Arten der Schweiz. Rev. Suisse. Zool. **58:** 23-175.
CALNAN, D. R., and A. BRUNET, 2008 The FoxO code. Oncogene **27:** 2276-2288.
CARSON, H. L., and H. D. STALKER, 1948 Reproductive Diapause in *Drosophila robusta*. Proceedings of the National Academy of Sciences of the United States of America **34:** 124-129.
CARTER, K. W., P. A. MCCASKIE and L. J. PALMER, 2006 JLIN: A java based linkage disequilibrium plotter. Bmc Bioinformatics **7**.
CASHMORE, A. R., J. A. JARILLO, Y. J. WU and D. M. LIU, 1999 Cryptochromes: Blue light receptors for plants and animals. Science **284:** 760-765.
CERIANI, M. F., T. K. DARLINGTON, D. STAKNIS, P. MAS, A. A. PETTI *et al.*, 1999 Light-dependent sequestration of TIMELESS by CRYPTOCHROME. Science **285:** 553-556.
CHARLESWORTH, P., and B. SHORROCKS, 1980 The Reproductive-Biology and Diapause of the British Fungal-Breeding *Drosophila*. Ecological Entomology **5:** 315-326.
CLANCY, D. J., D. GEMS, L. G. HARSHMAN, S. OLDHAM, H. STOCKER *et al.*, 2001 Extension of life-span by loss of CHICO, a *Drosophila* insulin receptor substrate protein. Science **292:** 104-106.
COSTA, R., A. A. PEIXOTO, G. BARBUJANI and C. P. KYRIACOU, 1992 A Latitudinal Cline in a *Drosophila* Clock Gene. Proceedings of the Royal Society of London Series B-Biological Sciences **250:** 43-49.
DARLINGTON, T. K., K. WAGER-SMITH, M. F. CERIANI, D. STAKNIS, N. GEKAKIS *et al.*, 1998 Closing the circadian loop: CLOCK-induced transcription of its own inhibitors per and tim. Science **280:** 1599-1603.
DAVID, J. R., and P. CAPY, 1988 Genetic-Variation of *Drosophila melanogaster* Natural-Populations. Trends in Genetics **4:** 106-111.
DELBECQUE, J. P., K. WEIDNER and K. H. HOFFMANN, 1990 Alternative Sites for Ecdysteroid Production in Insects. Invertebrate Reproduction & Development **18:** 29-42.
DENLINGER, D. L., 2002 Regulation of diapause. Annual Review of Entomology **47:** 93-122.
EPLING, C., D. F. MITCHELL and R. H. T. MATTONI, 1955 Frequencies of Inversion Combinations in the 3rd Chromosome of Wild Males of *Drosophila pseudoobscura*. Proceedings of the National Academy of Sciences of the United States of America **41:** 915-921.
FEDER, M. E., and T. MITCHELL-OLDS, 2003 Evolutionary and ecological functional genomics. Nature Reviews Genetics **4:** 651-657.
FLATT, T., K. J. MIN, C. D'ALTERIO, E. VILLA-CUESTA, J. CUMBERS *et al.*, 2008 Drosophila, germ-line modulation of insulin signaling and lifespan. Proceedings of the National Academy of Sciences of the United States of

America **105:** 6368-6373.
GLASSCOCK, E., and M. A. TANOUYE, 2005 *Drosophila* couch potato mutants exhibit complex neurological abnormalities including epilepsy phenotypes. Genetics **169:** 2137-2149.
GLOSSOP, N. R. J., J. H. HOUL, H. ZHENG, F. S. NG, S. M. DUDEK *et al.*, 2003 VRILLE feeds back to control circadian transcription of Clock in the *Drosophila* circadian oscillator. Neuron **37:** 249-261.
GOTO, S. G., T. YOSHIDA, K. BEPPU and M. T. KIMURA, 1999 Evolution of overwintering strategies in Eurasian species of the *Drosophila obscura* species group. Biological Journal of the Linnean Society **68:** 429-441.
GRUNTENKO, N. E., and I. Y. RAUSCHENBACH, 2008 Interplay of JH, 20E and biogenic amines under normal and stress conditions and its effect on reproduction. Journal of Insect Physiology **54:** 902-908.
HARDIN, P. E., 2005 The circadian timekeeping system of *Drosophila*. Current Biology **15:** R714-R722.
HARDIN, P. E., J. C. HALL and M. ROSBASH, 1990 Feedback of the *Drosophila* Period Gene-Product on Circadian Cycling of Its Messenger-Rna Levels. Nature **343:** 536-540.
HARSHMAN, L. G., A. M. LOEB and B. A. JOHNSON, 1999 Ecdysteroid titers in mated and unmated *Drosophila melanogaster* females. Journal of Insect Physiology **45:** 571-577.
HARVIE, P. D., M. FILIPPOVA and P. J. BRYANT, 1998 Genes expressed in the ring gland, the major endocrine organ of *Drosophila melanogaster*. Genetics **149:** 217-231.
HIRAI, Y., S. G. GOTO, T. YOSHIDA and M. T. KIMURA, 2000 Faunal and Ecological Surveys on Drosophilid Flies in Iriomote-jima, a Subtropical Island of Japan. Entomological Science **3:** 273-284.
HOSKINS, R. A., A. C. PHAN, M. NAEEMUDDIN, F. A. MAPA, D. A. RUDDY *et al.*, 2001 Single nucleotide polymorphism markers for genetic mapping in *Drosophila melanogaster*. Genome Research **11:** 1100-1113.
ICHIJO, N., 1986 Disjunctive Cline of Critical Photoperiod in the Reproductive Diapause of *Drosophila lacertosa*. Evolution **40:** 418-421.
ICHIJO, N., K. BEPPU and M. T. KIMURA, 1992 Aestivo-Hibernal Reproductive Diapause in *Drosophila moriwakii* - Flight as a Controlling Factor. Entomologia Experimentalis Et Applicata **62:** 23-28.
IVES, P. T., 1954 Genetic Changes in American Populations of *Drosophila melanogaster*. Proceedings of the National Academy of Sciences of the United States of America **40:** 87-92.
IVES, P. T., 1970 Further Genetic Studies of South Amherst Population of *Drosophila melanogaster*. Evolution **24:** 507-&.
KELLY, J. K., 1997 A test of neutrality based on interlocus associations. Genetics

146: 1197-1206.
KHAZAELI, A. A., and J. W. CURTSINGER, 2005 Lack of correlation between reproductive diapause and life span in *Drosophila melanogaster*. Drosophila Information Serials **88:** 51-52.
KIMURA, M. T., 1982 Effects of Photoperiod and Temperature on Reproductive Diapause in *Drosophila testacea*. Experientia **38:** 371-372.
KIMURA, M. T., 1983 Geographic-Variation and Genetic-Aspects of Reproductive Diapause in *Drosophila triauraria* and *Drosophila quadraria*. Physiological Entomology **8:** 181-186.
KIMURA, M. T., T. AWASAKI, T. OHTSU and K. SHIMADA, 1992 Seasonal-Changes in Glycogen and Trehalose Content in Relation to Winter Survival of 4 Temperate Species of *Drosophila*. Journal of Insect Physiology **38:** 871-875.
KIMURA, M. T., and K. BEPPU, 1993 Climatic Adaptations in the *Drosophila immigrans* Species Group - Seasonal Migration and Thermal Tolerance. Ecological Entomology **18:** 141-149.
KIMURA, M. T., T. OHTSU, T. YOSHIDA, T. AWASAKI and F. J. LIN, 1994 Climatic Adaptations and Distributions in the *Drosophila takahashii* Species Subgroup (Diptera, Drosophilidae). Journal of Natural History **28:** 401-409.
KIMURA, M. T., and T. YOSHIDA, 1995 A Genetic-Analysis of Photoperiodic Reproductive Diapause in *Drosophila triauraria*. Physiological Entomology **20:** 253-256.
KING, R. C., 1970 *Overian Development in Drosophila melanogaster*. Academic, New York.
KONOPKA, R. J., and S. BENZER, 1971 Clock Mutants of *Drosophila melanogaster*. Proceedings of the National Academy of Sciences of the United States of America **68:** 2112-&.
KUMAR, S., K. TAMURA and M. NEI, 2004 MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Briefings in Bioinformatics **5:** 150-163.
LANGLEY, C. H., B. P. LAZZARO, W. PHILLIPS, E. HEIKKINEN and J. M. BRAVERMAN, 2000 Linkage disequilibria and the site frequency spectra in the su(s) and su(w(a)) regions of the *Drosophila melanogaster* X chromosome. Genetics **156:** 1837-1852.
LANKINEN, P., 1986 Geographical Variation in Circadian Eclosion Rhythm and Photoperiodic Adult Diapause in *Drosophila littoralis*. Journal of Comparative Physiology a-Sensory Neural and Behavioral Physiology **159:** 123-142.
LANKINEN, P., and P. FORSMAN, 2006 Independence of genetic geographical variation between photoperiodic diapause, circadian eclosion rhythm, and

thr-gly repeat region of the period gene in *Drosophila littoralis*. Journal of Biological Rhythms **21:** 3-12.

LEEVERS, S. J., D. WEINKOVE, L. K. MACDOUGALL, E. HAFEN and M. D. WATERFIELD, 1996 The *Drosophila* phosphoinositide 3-kinase Dp110 promotes cell growth. Embo Journal **15:** 6584-6594.

LI, J., S. C. WANG and Z. B. ZENG, 2006 Multiple-interval mapping for ordinal traits. Genetics **173:** 1649-1663.

LIM, C., B. Y. CHUNG, J. L. PITMAN, J. J. MCGILL, S. PRADHAN *et al.*, 2007 clockword orange encodes a transcriptional repressor important for circadian-clock amplitude in *Drosophila*. Current Biology **17:** 1082-1089.

LONG, A. D., R. F. LYMAN, C. H. LANGLEY and T. F. C. MACKAY, 1998 Two sites in the Delta gene region contribute to naturally occurring variation in bristle number in *Drosophila melanogaster*. Genetics **149:** 999-1017.

LUMME, J., and L. KERANEN, 1978 Photoperiodic Diapause in *Drosophila lummei* Hackman Is Controlled by an X-Chromosomal Factor. Hereditas **89:** 261-262.

LUMME, J., and A. OIKARINEN, 1977 Genetic Basis of Geographically Variable Photoperiodic Diapause in *Drosophila littoralis*. Hereditas **86:** 129-141.

MASSAGUE, J., and D. WOTTON, 2000 Transcriptional control by the TGF-beta/Smad signaling system. Embo Journal **19:** 1745-1754.

MCCALL, K., 2004 Eggs over easy: cell death in the *Drosophila* ovary. Developmental Biology **274:** 3-14.

MCDONALD, J. H., and M. KREITMAN, 1991 Adaptive Protein Evolution at the Adh Locus in *Drosophila*. Nature **351:** 652-654.

MERRITT, T. J. S., E. SEZGIN, C. T. ZHU and W. F. EANES, 2006 Triglyceride pools, flight and activity variation at the Gpdh locus in *Drosophila melanogaster*. Genetics **172:** 293-304.

MEYER, P., L. SAEZ and M. W. YOUNG, 2006 PER-TIM interactions in living Drosophila cells: An interval timer for the circadian clock. Science **311:** 226-229.

MIN, K. J., and M. TATAR, 2006 Restriction of amino acids extends lifespan in *Drosophila melanogaster*. Mechanisms of Ageing and Development **127:** 643-646.

MINAMI, N., and M. T. KIMURA, 1980 Geographical Variation of Photoperiodic Adult Diapause in *Drosophila auraria*. Japanese Journal of Genetics **55:** 319-324.

MITROVSKI, P., and A. A. HOFFMANN, 2001 Postponed reproduction as an adaptation to winter conditions in *Drosophila melanogaster*: evidence for clinal variation under semi-natural conditions. Proceedings of the Royal Society of London Series B-Biological Sciences **268:** 2163-2168.

MIYASHITA, N. T., M. AGUADE and C. H. LANGLEY, 1993 Linkage Disequilibrium

in the White Locus Region of *Drosophila melanogaster*. Genetical Research **62:** 101-109.

MOSHITZKY, P., I. FLEISCHMANN, N. CHAIMOV, P. SAUDAN, S. KLAUSER *et al.*, 1996 Sex-peptide activates juvenile hormone biosynthesis in the *Drosophila melanogaster* corpus allatum. Archives of Insect Biochemistry and Physiology **32:** 363-374.

MUONA, O., and J. LUMME, 1981 Geographical Variation in the Reproductive-Cycle and Photoperiodic Diapause of *Drosophila phalerata* and *Drosophila transversa* (Drosophilidae, Diptera). Evolution **35:** 158-167.

MYERS, M. P., K. WAGERSMITH, C. S. WESLEY, M. W. YOUNG and A. SEHGAL, 1995 Positional Cloning and Sequence-Analysis of the *Drosophila* Clock Gene, Timeless. Science **270:** 805-808.

NAIDOO, N., W. SONG, M. HUNTER-ENSOR and A. SEHGAL, 1999 A role for the proteasome in the light response of the timeless clock protein. Science **285:** 1737-1741.

NEI, M., 1987 *Molecular Evolutionary Genetics*. Columbia University Press, New York.

OHTSU, T., M. T. KIMURA and S. H. HORI, 1992 Energy-Storage During Reproductive Diapause in the *Drosophila melanogaster* Species Group. Journal of Comparative Physiology B-Biochemical Systemic and Environmental Physiology **162:** 203-208.

PARKS, A. L., K. R. COOK, M. BELVIN, N. A. DOMPE, R. FAWCETT *et al.*, 2004 Systematic generation of high-resolution deletion coverage of the *Drosophila melanogaster* genome. Nature Genetics **36:** 288-292.

PARTRIDGE, L., D. GEMS and D. J. WITHERS, 2005 Sex and death: What is the connection? Cell **120:** 461-472.

PATEL, M. N., C. G. KNIGHT, C. KARAGEORGI and A. M. LEROI, 2002 Evolution of germ-line signals that regulate growth and aging in nematodes. Proceedings of the National Academy of Sciences of the United States of America **99:** 769-774.

RICHARD, D. S., J. M. JONES, M. R. BARBARITO, S. CERULA, J. P. DETWEILER *et al.*, 2001 Vitellogenesis in diapausing and mutant *Drosophila melanogaster*: further evidence for the relative roles of ecdysteroids and juvenile hormones. Journal of Insect Physiology **47:** 905-913.

RICHARD, D. S., N. L. WATKINS, R. B. SERAFIN and L. I. GILBERT, 1998 Ecdysteroids regulate yolk protein uptake by *Drosophila melanogaster* oocytes. Journal of Insect Physiology **44:** 637-644.

RIIHIMAA, A., 1984 The Inheritance of the Facultative Diapause of *Chymomyza costata*. Hereditas **101:** 283-283.

ROBICH, R. M., and D. L. DENLINGER, 2005 Diapause in the mosquito *Culex*

pipiens evokes a metabolic switch from blood feeding to sugar gluttony. Proceedings of the National Academy of Sciences of the United States of America **102:** 15912-15917.

ROZAS, J., and R. ROZAS, 1995 DnaSP, DNA sequence polymorphism: An interactive program for estimating population genetics parameters from DNA sequence data. Computer Applications in the Biosciences **11:** 621-625.

SANDRELLI, F., E. TAUBER, M. PEGORARO, G. MAZZOTTA, P. CISOTTO *et al.*, 2007 A molecular basis for natural selection at the timeless locus in *Drosophila melanogaster*. Science **316:** 1898-1900.

SAUNDERS, D. S., 1990 The Circadian Basis of Ovarian Diapause Regulation in *Drosophila melanogaster* - Is the Period Gene Causally Involved in Photoperiodic Time Measurement. Journal of Biological Rhythms **5:** 315-331.

SAUNDERS, D. S., and L. I. GILBERT, 1990 Regulation of Ovarian Diapause in *Drosophila melanogaster* by Photoperiod and Moderately Low-Temperature. Journal of Insect Physiology **36:** 195-200.

SAUNDERS, D. S., V. C. HENRICH and L. I. GILBERT, 1989 Induction of Diapause in *Drosophila melanogaster* - Photoperiodic Regulation and the Impact of Arrhythmic Clock Mutations on Time Measurement. Proceedings of the National Academy of Sciences of the United States of America **86:** 3748-3752.

SAUNDERS, D. S., D. S. RICHARD, S. W. APPLEBAUM, M. MA and L. I. GILBERT, 1990 Photoperiodic Diapause in *Drosophila melanogaster* Involves a Block to the Juvenile-Hormone Regulation of Ovarian Maturation. General and Comparative Endocrinology **79:** 174-184.

SAWYER, L. A., J. M. HENNESSY, A. A. PEIXOTO, E. ROSATO, H. PARKINSON *et al.*, 1997 Natural variation in a *Drosophila* clock gene and temperature compensation. Science **278:** 2117-2120.

SCHMIDT, P., C. ZHU, D. JAYATRI, B. MARISKA, Y. LI *et al.*, 2008a An amino acid polymorphism in the couch potato gene forms the basis for climatic adaptation in Drosophila melanogaster. In Press.

SCHMIDT, P. S., and D. R. CONDE, 2006 Environmental heterogeneity and the maintenance of genetic variation for reproductive diapause in *Drosophila melanogaster*. Evolution **60:** 1602-1611.

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

SCHMIDT, P. S., and A. B. PAABY, 2008 Reproductive diapause and life-history clines in North American populations of *Drosophila melanogaster*. Evolution **62:** 1204-1215.

SCHMIDT, P. S., A. B. PAABY and M. S. HESCHEL, 2005b Genetic variance for diapause expression and associated life histories in *Drosophila melanogaster*. Evolution **59:** 2616-2625.

SCHMIDT, P. S., C. T. ZHU, J. DAS, M. BATAVIA, L. YANG *et al.*, 2008b An amino acid polymorphism in the couch potato gene forms the basis for climatic adaptation in *Drosophila melanogaster*. Proceedings of the National Academy of Sciences of the United States of America **105:** 16207-16211.

SEZGIN, E., D. D. DUVERNELL, L. M. MATZKIN, Y. H. DUAN, C. T. ZHU *et al.*, 2004 Single-locus latitudinal clines and their relationship to temperate adaptation in metabolic genes and derived alleles in *Drosophila melanogaster*. Genetics **168:** 923-931.

SHIMAMOTO, A., S. KITAO, K. ICHIKAWA, N. SUZUKI, Y. YAMABE *et al.*, 1996 A unique human gene that spans over 230 kb in the human chromosome 8p11-12 and codes multiple family proteins sharing RNA-binding motifs. Proceedings of the National Academy of Sciences of the United States of America **93:** 10913-10917.

SIM, C., and D. L. DENLINGER, 2008 Insulin signaling and FOXO regulate the overwintering diapause of the mosquito *Culex pipiens*. Proceedings of the National Academy of Sciences of the United States of America **105:** 6777-6781.

SUN, Y., L. H. DING, H. ZHANG, J. Q. HAN, X. YANG *et al.*, 2006 Potentiation of Smad-mediated transcriptional activation by the RNA-binding protein RBPMS. Nucleic Acids Research **34:** 6314-6326.

TAJIMA, F., 1989 Statistical-Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. Genetics **123:** 585-595.

TAMURA, K., S. SUBRAMANIAN and S. KUMAR, 2004 Temporal patterns of fruit fly (Drosophila) evolution revealed by mutation clocks. Molecular Biology and Evolution **21:** 36-44.

TATAR, M., A. BARTKE and A. ANTEBI, 2003 The endocrine regulation of aging by insulin-like signals. Science **299:** 1346-1351.

TATAR, M., S. A. CHIEN and N. K. PRIEST, 2001a Negligible senescence during reproductive dormancy in *Drosophila melanogaster*. American Naturalist **158:** 248-258.

TATAR, M., A. KOPELMAN, D. EPSTEIN, M. P. TU, C. M. YIN *et al.*, 2001b A mutant Drosophila insulin receptor homolog that extends life-span and impairs neuroendocrine function. Science **292:** 107-110.

TATAR, M., and C. M. YIN, 2001 Slow aging during insect reproductive diapause: why butterflies, grasshoppers and flies are like worms. Experimental Gerontology **36:** 723-738.

TAUBER, E., M. ZORDAN, F. SANDRELLI, M. PEGORARO, N. OSTERWALDER *et al.*, 2007 Natural selection favors a newly derived timeless allele in

Drosophila melanogaster. Science **316:** 1895-1898.
TERASHIMA, J., and M. BOWNES, 2004 Translating available food into the number of eggs laid by *Drosphila melanogaster*. Genetics **167:** 1711-1719.
TERASHIMA, J., and M. BOWNES, 2005 A microarray analysis of genes involved in relating egg production to nutritional intake in *Drosophila melanogaster*. Cell Death and Differentiation **12:** 429-440.
THORNTON, K., and P. ANDOLFATTO, 2006 Approximate Bayesian inference reveals evidence for a recent, severe bottleneck in a Netherlands population of *Drosophila melanogaster*. Genetics **172:** 1607-1619.
TU, M. P., C. M. YIN and M. TATAR, 2002 Impaired ovarian ecdysone synthesis of *Drosophila melanogaster* insulin receptor mutants. Aging Cell **1:** 158-160.
VAN DER HEIDE, L. P., M. F. M. HOEKMAN and M. P. SMIDT, 2004 The ins and outs of FoxO shuttling: mechanisms of FoxO translocation and transcriptional regulation. Biochemical Journal **380:** 297-309.
WATABE, H., 1995 Genetic Analyses of the Photoperiodic Diapause of *Drosophila lummei* Hackman, 1972 (Diptera : Drosophilae). Journal of Hokkaido University of Eduction (Section II B) **45:** 37-41.
WATTERSON, G. A., 1975 Number of Segregating Sites in Genetic Models without Recombination. Theoretical Population Biology **7:** 256-276.
WEINKOVE, D., S. J. LEEVERS, L. K. MACDOUGALL and M. D. WATERFIELD, 1997 p60 is an adaptor for the *Drosophila* phosphoinositide 3-kinase, Dp110. Journal of Biological Chemistry **272:** 14606-14610.
WEINKOVE, D., T. P. NEUFELD, T. TWARDZIK, M. D. WATERFIELD and S. J. LEEVERS, 1999 Regulation of imaginal disc cell size, cell number and organ site by *Drosophila* class I-A phosphoinositide 3-kinase and its adaptor. Current Biology **9:** 1019-1029.
WILLIAMS, K. D., M. BUSTO, M. L. SUSTER, A. K. C. SO, Y. BEN-SHAHAR *et al.*, 2006 Natural variation in *Drosophila melanogaster* diapause due to the insulin-regulated PI3-kinase. Proceedings of the National Academy of Sciences of the United States of America **103:** 15911-15915.
WILLIAMS, K. D., and M. B. SOKOLOWSKI, 1993 Diapause in *Drosophila melanogaster* Females - a Genetic-Analysis. Heredity **71:** 312-317.
YOSHIDA, T., and M. T. KIMURA, 1993 The Photoperiodic Clock of *Drosophila triauraria* - Involvement of 2 Processes in the Night-Length Measurement System. Journal of Insect Physiology **39:** 101-106.
YOSHIDA, T., and M. T. KIMURA, 1994a Relation of the Circadian System to the Photoperiodic Clock in *Drosophila triauraria* (Diptera, Drosophilidae) - an Approach from Analysis of Geographic-Variation. Applied Entomology and Zoology **29:** 499-505.
YOSHIDA, T., and M. T. KIMURA, 1994b Some Effects of Low-Temperature on the Photoperiodic Clock in *Drosophila triauraria*. Physiological Entomology

19: 83-87.

YU, Q., A. C. JACQUIER, Y. CITRI, M. HAMBLEN, J. C. HALL *et al.*, 1987 Molecular Mapping of Point Mutations in the Period Gene That Stop or Speed up Biological Clocks in *Drosophila melanogaster*. Proceedings of the National Academy of Sciences of the United States of America **84:** 784-788.

ZERA, A. J., and L. G. HARSHMAN, 2001 The physiology of life history trade-offs in animals. Annual Review of Ecology and Systematics **32:** 95-126.