The molecular basis of cold tolerance in *Drosophila ananassae*

Annabella Königer

München 2019

The molecular basis of cold tolerance in *Drosophila ananassae*

Annabella Königer

Dissertation
an der Fakultät für Biologie
der Ludwig–Maximilians–Universität
München

vorgelegt von
Annabella Königer
aus Augsburg

München, den 28.03.2019

Erklärung

Diese Dissertation wurde im Sinne von §12 der Promotionsordnung von Dr. Sonja Grath betreut. Ich erkläre hiermit, dass die Dissertation nicht einer anderen Prüfungskommission vorgelegt worden ist und dass ich mich nicht anderweitig einer Doktorprüfung unterzogen habe.

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbstständig und ohne unerlaubte Hilfe angefertigt worden ist. Im Abschnitt „Note“ ist mein Anteil an der Arbeit im Einzelnen dargelegt.

München, den 16.03.2019
Datum, Ort

Annabella Königer
Unterschrift

Erstgutachter: Dr. Sonja Grath

Zweitgutachter: Prof. Dr. Dirk Metzler

Tag der Abgabe: 28.03.2019

Tag der mündlichen Prüfung: 03.09.2019

Contents

Note

In this thesis, I present the results of my doctoral research from May 2015 to March 2019. The presented work was undertaken in order to identify candidate genes for cold tolerance in *Drosophila ananassae*. To do so, I performed a transcriptome analysis for differential gene expression in multiple *D. ananassae* strains before and after a cold shock. Moreover, I mapped causal genomic regions in a population of fly strains with recombinant genotypes and generated transgenic Cas9 strains to set the groundwork for future loss-of-function studies of the identified candidate genes. Some parts of the presented work were carried out in collaboration with other scientists:

1. The high-throughput phenotyping assays (chapters 2.2, 3.1) of the Bangkok strains were carried out and analyzed by myself.

2. The transcriptome analysis (chapters 2.3, 3.2 and 4.1) is based on the following publication:
   Königer A. and Grath S. (2018) Transcriptome Analysis Reveals Candidate Genes for Cold Tolerance in *Drosophila ananassae*. *Genes* 9(12), 624.
   `https://doi.org/10.3390/genes9120624`
   The data for this publication was generated and analyzed by myself. For read mapping, I adapted a script provided by Ann Kathrin Huylmans and John Parsch, for differential gene expression analysis, I adapted a script provided by Korbinian von Heckel. The manuscript for the original publication was written by myself with input from my supervisor Sonja Grath. Many text passages in the mentioned chapters have been adapted (with minor changes) but without explicit reference. The original Figures and Tables were all created by myself and labeled with a citation.

3. The QTL mapping (chapters 2.4, 3.3 and 4.2) is based on an unpublished manuscript:

Königer A., Arif S. and Grath S. (2019) Three QTL influence chill coma recovery time in *Drosophila ananassae*.
I generated and analyzed the data myself and did the writing. For some parts of the computational work, I received input from Saad Arif, in particular for the marker calling, genetic map construction and the multiple-QTL model (chapters 2.4.3, 2.4.4 and 2.4.5).

4. The genome engineering project was designed by myself. I generated the Cas9 strains (chapters 2.5.1, 3.4.1, selected the target sites (chapters 2.5.3 and supervised Selina Mußgnug who did the cloning for all gRNA and donor plasmids (chapters 2.5.3 and 3.4.2) and subsequent marker screening (chapter 2.6).

5. Eslam Katab who worked as a student assistant in our group helped with general fly maintenance and DNA extractions for Sanger sequencing (chapters 2.1 and 2.5.2).

6. This thesis was written in LaTeX. I adapted the LMU layout template provided by Robert Dahlke and Sigmund Stintzing via the University Library of the LMU Munich:
`https://edoc.ub.uni-muenchen.de/help/#property_latex`

Abstract

For ectothermic organisms such as insects, temperature is one of the major factors that influence their geographical distribution and abundance. Their body temperature follows the external environment and resilience towards thermal extremes often determines a species fate upon climate change or range expansion. Nevertheless, many insect species that evolved and diversified in the tropics have expanded to temperate environments. This dissertation examines the molecular basis of cold tolerance in *Drosophila ananassae*, using fly strains from a population in Bangkok, Thailand as a model system. The sampling site in Bangkok is located within the ancestral range of *D. ananassae*. From here, the species expanded during the last 18,000 years and colonized temperate regions all over the globe.

I measured cold tolerance in fly strains from Bangkok by means of a test for chill coma recovery time (CCRT). There was substantial variation in the phenotype; I identified cold-sensitive strains ("Slow") and cold-tolerant strains ("Fast"). This finding was consistent with previous work and allowed me to study the genetic basis of the phenotypic variation without confounding effects of population divergence. Furthermore, population sub-structure was reportedly absent in the Bangkok strains, suggesting that the causal loci are few in number but have large effects on the phenotype. Thus, to uncover genes and genetic elements that underlie the difference in cold tolerance, I carried out two genome-wide screens for candidates:

First, using high-throughput mRNA-sequencing, I identified genes with differential expression in Slow and Fast strains in response to a cold shock. Generally, about 13% of all protein-coding genes responded to the cold shock with significant up- or downregulation. Two genes with unknown function, *GF14647* and *GF15058* showed a significant interaction of phenotype and cold shock. Further, in the Fast phenotype, transcript abundance of six genes involved in actin polymerization was elevated already before the cold shock, whereas expression in strains with the Slow phenotype was up-

regulated to comparable levels only after the cold shock. Surprisingly, upregulation of heat shock proteins (hsps), which act as chaperones upon exposure to various stresses, was stronger in the Slow strains. Reanalysis of recently published transcriptome data of cold-tolerant and cold-sensitive populations of *D. melanogaster* revealed a similar pattern, which contradicts previous findings that linked upregulation of hsps to a faster recovery from cold exposure.

Second, using a hierarchical quantitative trait locus (QTL) mapping approach, I identified genomic regions that underlie the difference in CCRT. To do so, a mapping population of recombinant inbred advanced intercross lines (RIAILs) was created, using the most cold-tolerant and the most cold-sensitive Bangkok strain as parental founders. The RIAILs were phenotyped for their CCRT and, together with the parental strains, genotyped for polymorphic markers with double-digest restriction site-associated DNA (ddRAD) sequencing. Combining standard interval mapping and a multiple-QTL model, I identified three QTL and epistasis among two of them. In total, 58 differentially expressed genes were located within the three QTL confidence intervals. Most prominently, *GF15058* was one of them. Furthermore, the orthologs of four *D. melanogaster* genes that had been associated with thermal tolerance before were located within the identified QTL regions: *MtnA*, *klarsicht*, *GF17132* and *GF14829*.

In conclusion, I identified a concrete list of 12 candidate genes for cold tolerance in *D. ananassae*. Hence, to pave the way for future functional analyses of the identified candidates, I generated *D. ananassae* Cas9 strains. The PiggyBac transposon system was used to generate stable, homozygous insertions of a Cas9 construct into the genomes of the most cold-tolerant and the most cold-sensitive strain from Bangkok and an additional strain from a derived, temperate population in Kathmandu. The Cas9 strains will facilitate CRISPR/Cas9-mediated homology-directed repair in future experiments and help to examine the functional contribution of each candidate gene to the cold tolerance phenotype.

Zusammenfassung

Als ektotherme Lebewesen sind Insekten besonders bedroht durch klimatische Veränderungen. Ihre Körpertemperatur entspricht in der Regel der Umgebungstemperatur, weshalb die geografische Verbreitung und das Vorkommen einer Art maßgeblich durch ihre Widerstandsfähigkeit gegenüber thermischen Extremen bestimmt ist. Dennoch haben sich viele ursprünglich in den Tropen beheimatete Arten an kältere Temperaturen angepasst und bevölkern heute auch gemäßigte Klimazonen. Die vorliegende Dissertation untersucht die molekularen Grundlagen der Kältetoleranz in *Drosophila ananassae*, wobei Fliegenlinien aus Bangkok, Thailand als Modellsystem verwendet wurden. Bangkok liegt im ursprünglichen Verbreitungsgebiet von *D. ananassae*. Von hier aus expandierte die Spezies im Laufe der letzten 18.000 Jahre und bevölkerte gemäßigte Regionen auf der ganzen Welt.

Ich habe zunächst die Kältetoleranz der Bangkok-Linien mithilfe eines Tests für „Chill coma recovery time“ (CCRT) gemessen, wobei sich erhebliche Unterschiede im Phänotyp zeigten. Das Vorhandensein von kältesensitiven Linien (langsame CCR) und kältetoleranten Linien (schnelle CCR) ermöglichte es mir, die genetische Basis für Kältetoleranz innerhalb einer Population zu untersuchen. Mit diesem Ansatz konnten potentiell verfälschende Effekte vermieden werden, welche durch die natürliche Divergenz zweier unterschiedlicher Populationen auftreten können. Des Weiteren wiesen die Bangkok-Linien in einer vorangegangenen Studie keinerlei Populationsstruktur auf, was darauf hindeutete, dass der Unterschied im Phänotyp in nur wenigen Genen mit jeweils großem Effekt begründet liegt. Um diese Gene zu identifizieren, führte ich zwei genomweite Analysen durch:

Als Erstes ermittelte ich durch mRNA-Sequenzierung Gene, die nach einem Kälteschock differentiell zwischen kältesensitiven und kältetoleranten Linien exprimiert sind. Insgesamt waren 13% aller Protein-kodierenden Gene vom Kälteschock betroffen. Lediglich zwei Gene mit bisher unbekannter Funktion, *GF14647* und *GF15058*, zeigten eine

signifikante Interaktion zwischen Phänotyp und Kälteschock. Darüber hinaus waren die Transkriptmengen von sechs Genen, die in die Aktinpolymerisation involviert sind, im kältetoleranten Phänotyp bereits vor dem Kälteschock erhöht, während sie im kältesensitiven Phänotyp erst nach dem Kälteschock vergleichbare Mengen erreichten. Des weiteren konnte ich aufzeigen, dass die Hochregulierung der Genexpression von Hitzeschockproteinen (Hsps), die grundsätzlich mit einer allgemeinen Stressreaktion verbunden ist, in den kältesensitiven Linien stärker ausgeprägt war als in den kältetoleranten Linien. Ein ähnliches Muster beobachteten wir auch bei der erneuten Auswertung eines publizierten *Drosophila melanogaster*-Datensatzes. Diese Erkenntnis steht im Widerspruch zu vorangegangenen Studien, die einen Zusammenhang zwischen der Expression von Hsps und einer schnelleren CCR nachwiesen.

Die zweite Analyse umfasste eine QTL-Kartierung (Quantitative Trait Locus mapping), mit der ich kausale Genom-Regionen identifizierte. Für die Kartierung stellte ich eine Population aus rekombinanten Linien her, wofür die kältesensitivste und die kältetoleranteste Bangkok-Linie miteinander gekreuzt wurden. Die rekombinanten Linien wurden phänotypisiert und zusammen mit den Gründerlinien durch double-digest restriction site-associated DNA (ddRAD)-Sequenzierung für polymorphe Marker genotypisiert. Durch eine Kombination aus Standard-Interval-Mapping und einem Multiple-QTL-Modell identifizierte ich drei QTL, von denen zwei miteinander interagierten. Innerhalb der drei Konfidenzintervalle befanden sich 58 differentiell exprimierte Gene und insgesamt fünf Gene, die in vorangegangen Studien mit Thermotoleranz assoziiert wurden: *MtnA*, *klarsicht*, *GF17132*, *GF14829* und, interessanterweise, *GF15058*.

Zusammenfassend ergibt sich eine konkrete Liste aus zwölf Kandidatengenen für Kältetoleranz in *D. ananassae*. Um zukünftige funktionelle Analysen in *D. ananassae* zu ermöglichen, stellte ich Cas9-Linien her. Mithilfe des PiggyBac-Transposon-Systems generierte ich stabile, homozygote Insertionen eines Cas9-Konstrukts in den Genomen von drei bestimmten Linien: Die kältesensitivste und die kältetoleranteste Linie aus der Bangkok-Population und eine zusätzliche Linie aus einer gemäßigten Region in Kathmandu, Nepal. Die Cas9-Linien werden in Zukunft präzise Genom-Editierung in *D. ananassae* ermöglichen. Durch CRISPR/Cas9-vermittelten Gen-knockout kann dann der funktionelle Beitrag eines jeden Kandidatengens zum Phänotyp Kältetoleranz genauer untersucht werden.

List of Abbreviations

bp	base pair
BP	biological process
CC	cellular component
CCRT	chill coma recovery time
cDNA	copy DNA
cM	centimorgan
CRISPR	clustered regularly interspaced short palindromic repeat
CRISPR/Cas	CRISPR-associated
CrRNA	CRISPR RNA
DAVID	database for visualization and integrated discovery
ddRAD	double-digest restriction site-associated
dNTP	deoxy-nucleoside triphosphate
DSB	double-strand break
dsRed	red fluorescent protein, isolated from *Discosoma*
EASE	expression analysis systematic explorer
EM	expectation maximization
EYFP	enhanced yellow fluorescent protein
F1	first filial generation
F	Fast allele
FDR	false discovery rate
GO	gene ontology
gRNA	guide RNA
HDR	homology-directed repair
Hsp	heat shock protein
indel	insertion and deletion
ITR	inverted terminal repeat

kb	kilo bases
KEGG	kyoto encyclopedia of genes and genomes
KRHT	knock-down resistance to high temperature
LB	lysogeny broth
LOD	logarithm of the odds
Mb	mega bases
MF	molecular function
MQM	multiple QTL mapping
mRNA	messenger RNA
NHEJ	non-homologous end joining
P	parental generation
PAM	protospacer adjacent motif
PBS	phosphate buffered saline
PCA	principal component analysis
PCR	polymerase chain reaction
QTL	quantitative trait locus
RIAIL	recombinant inbred advanced intercross line
RNAi	RNA interference
S	Slow allele
SNP	single nucleotide polymorphism
StDev	standard deviation
TALEN	transcription activator-like effector nuclease
tracrRNA	trans-activating RNA
UDP	uridine diphosphate
UGT	UDP-glycosyltransferases
UTP	uridine triphosphate
UTR	untranslated region
ZNF	zinc-finger nuclease

List of Figures

List of Tables

Chapter 1

Introduction

When organisms are faced with new environmental conditions, such as those caused by climate change or range expansion, they must adapt in order to survive. From an evolutionary perspective, studying how the genetic makeup of species is formed and how they adapt to their environment are central questions. Local adaptation is an evolutionary process by which individuals have their highest fitness in their local environment by means of natural selection (Dobzhansky, 1968). This means that a specific trait will be selected in a local population as a consequence of local habitat characteristics. One of the major factors that shape the characteristics of a habitat is temperature. Essentially, temperature determines state phases of molecules and catalytic rates of enzymes and therefore affects every animal on Earth (Angilletta, 2009). Insects which constitute the largest group of animals are especially vulnerable to thermal extremes, particularly towards cold (Araújo et al., 2013; Warren and Chick, 2013). As ectotherms, their body temperature follows the external environment as they are generally not able to metabolically produce heat. Therefore, their resilience towards low temperatures often determines the species fate upon climate change or range expansion. Indeed, most insect species evolved and diversified in warm climates (Throckmorton et al., 1975; Bale and Hayward, 2010). Nevertheless, insects have successfully colonized all ecosystems on Earth (reviewed in (Overgaard and MacMillan, 2016)).

Some species are able to remain active at even sub-zero temperatures, such as the antarctic midge *Belgica antarctica* (Teets et al., 2012) or the alpine beetle *Pytho deplanatus* (Ring, 1982). To cope with such extreme conditions, they rely on one of two strategies: they either tolerate the formation of ice crystals in their body fluids (freeze-tolerant insects (Sinclair et al., 2003)) or they suppress their supercooling point (SCP;

onset of the spontaneous formation of ice crystals) to avoid freezing (freeze-avoiding insects (Sformo et al., 2010)). However, a wide range of insect species suffer from chilling injuries and eventually death when they are exposed to cold temperatures above their SCP (chill-susceptible insects (Sinclair, 1999)). Chill-susceptibility has been reported for several orders of insects, including Coleoptera, Dictyoptera, Diptera, Hemiptera, Hymenoptera and Orthoptera (reviewed in (Overgaard and MacMillan, 2016)). Among them are species of great economic importance, such as the honey bee *Apis mellifera* (Free and Spencer-Booth, 1960; Villa and Rinderer, 1993), dreaded pests such as the migratory locust *Locusta migratoria* (Bayley et al., 2018), and dangerous disease vectors such as the common house mosquito *Culex pipiens* (Rinehart et al., 2006).

Another prominent example for chill-susceptible insects are species from the genus *Drosophila*, which serve as an important model system for genetic and physiological studies since the early 1900s (Kohler, 1994). The physiology of cold stress resistance in insects is extremely complex. Cold exposure disrupts ion and water homeostasis (MacMillan et al., 2015b) and changes metabolic profiles (MacMillan et al., 2016). In parallel, chilling injuries affect the organism at the cell level and are initiated by depolarization of the resting membrane potential and lipid phase transition (Quinn, 1985; Drobnis et al., 1993; Hosler et al., 2000; Andersen et al., 2015), eventually causing the cell to die (Overgaard and MacMillan, 2016). While a growing body of literature attempts to resolve this complexity, it remains challenging to identify specific genes and genetic elements that contribute to the diversity of cold stress-resistant phenotypes (Udaka et al., 2010; von Heckel et al., 2016; MacMillan et al., 2016). Drosophilids having a worldwide distribution are excellent models for studying these questions.

1.1 *Drosophila ananassae*: genetic features and demographic history

Drosophila ananassae belongs to the *ananassae* subgroup that is part of the *melanogaster* species group (Bock and Wheeler, 1972), (Figure 1.1) and shares a common ancestor with *Drosophila melanogaster* from which they split about 20 million years ago (Russo et al., 1995). It was one of the 12 species whose genomes were sequenced as part of the 12 genomes consortium (Drosophila 12 Genomes Consortium, 2007). *D. ananassae* possesses several unique genetic features when compared to closely related organisms. For example, meiotic recombination in males is frequent in *D. ananassae* (Kikkawa,

1938; Kale, 1969; Hinton, 1974) but occurs only rarely in other drosophilids (Morgan, 1910; Henderson, 1978). Moreover, populations of *D. ananassae* tend to be highly structured (Stephan and Langley, 1989; Singh and Singh, 2010) and possess a high degree of chromosomal polymorphisms such as inversions and translocations (Singh, 1983, 1984; Singh and Singh, 2007).

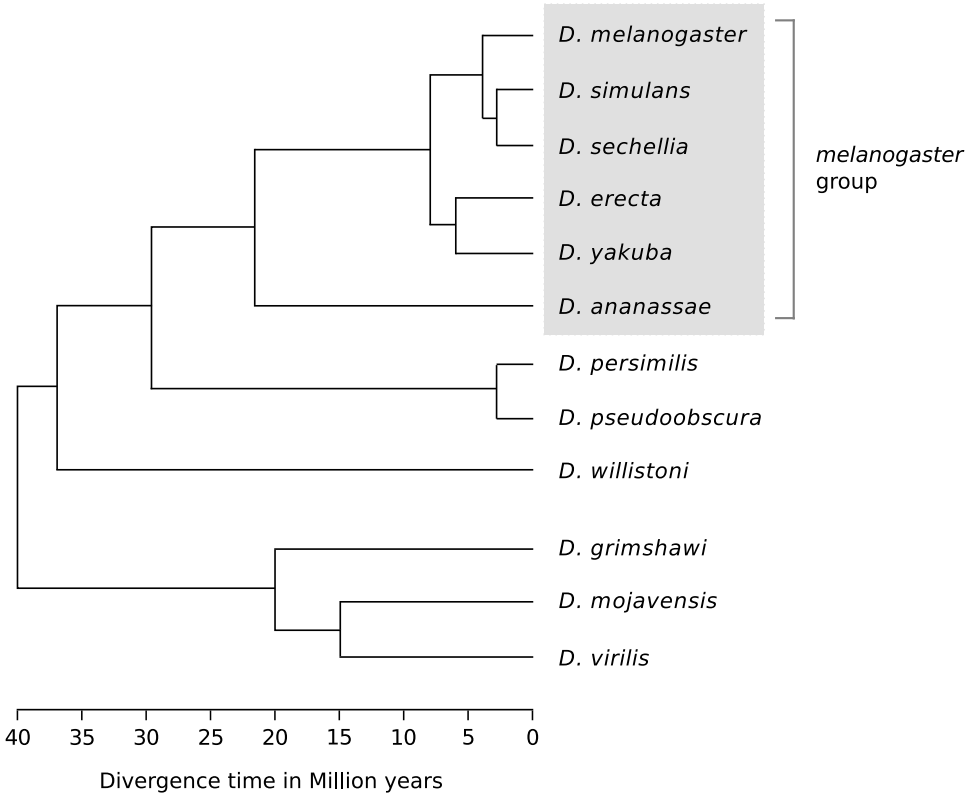


Figure 1.1: *Drosophila* phylogeny. Displayed are all species that were sequenced as part of the *Drosophila* 12 genomes consortium. Divergence times are indicated under the tree. Figure adapted from (Paris et al., 2013).

D. ananassae is a cosmopolitan, domestic species which is present in all zoogeographic regions of the globe except Antarctica and most abundant in tropical and subtropical regions of Asia (Tobari, 1993). The demographic history of *D. ananassae* was unravelled by Aparup Das and colleagues (Das et al., 2004). The authors analyzed DNA polymorphisms at ten intronic loci across the *D. ananassae* genome in 160 isofemale strains from 16 worldwide sampling locations. Among them, five populations from South-East Asia were identified as part of the ancestral species range: Bangkok, Kuala Lumpur, Java, Borneo and Manila (Figure 1.2). All strains from these locations showed a large number of polymorphic sites and alleles, high nucleotide diversity and

low levels of linkage disequilibrium (LD) – characteristics of ancestral populations. A novel and more advanced approach further corroborated this finding: a Bayesian-based clustering algorithm implemented through Markov chain Monte Carlo (MCMC) simulations identified the five South-East Asian populations as ancestral and the other eleven populations as derived. Furthermore, the ancestral populations showed negative values for Tajima's D, consistent with population growth. Today, all five sampling locations in South-East Asia are isolated by water. However, at the peak of the last glacial period during the late Pleistocene (about 18,000 years ago), sea levels fell below the margins of the Sunda shelf which encompasses an extension of the South-East Asian continental shelf (Johnson, 1964; Holloway, 1997). The exposed area, "Sundaland" connected the five sampling locations at that time which today represent the ancestral species range of *D. ananassae* (Figure 1.2).

1.2 Quantifying cold tolerance with chill coma recovery time

When being exposed to their critical thermal minimum, chill-susceptible insects suffer from an impaired synaptic transmission and muscular failure. As a consequence, they lose neuromuscular coordination and enter a comatose condition (chill coma) in which they are not able to move, feed, mate or escape predators (Kelty et al., 1996; Macmillan and Sinclair, 2011; Findsen et al., 2014). Notwithstanding, they are able to recover if the time of exposure and the specific temperature do not exceed a certain threshold. The time needed for recovery from a cold-induced coma (chill coma recovery time, CCRT) is a widely used measure to determine cold hardiness of insects in the laboratory. A test for CCRT comprises two phases: in phase one, the insect is subjected to a standardized cold shock (usually at 0°C) for a controlled amount of time. In phase two, it is returned to room temperature, and CCRT is recorded as the time the insect needs to stand back on its legs (David et al., 1998).

In *Drosophila* species, CCRT was shown to be a good predictor for the climatic condition at the geographical sampling site of the tested species ((Gibert et al., 2001) but see (Andersen et al., 2015) for a discussion of other measures). In their study, Gibert and colleagues demonstrated that across more than 70 tested species of *Drosophila*, tropical and subtropical species have a significantly longer CCRT than temperate species. Furthermore, CCRT follows latitudinal clines within species, as shown for Australian

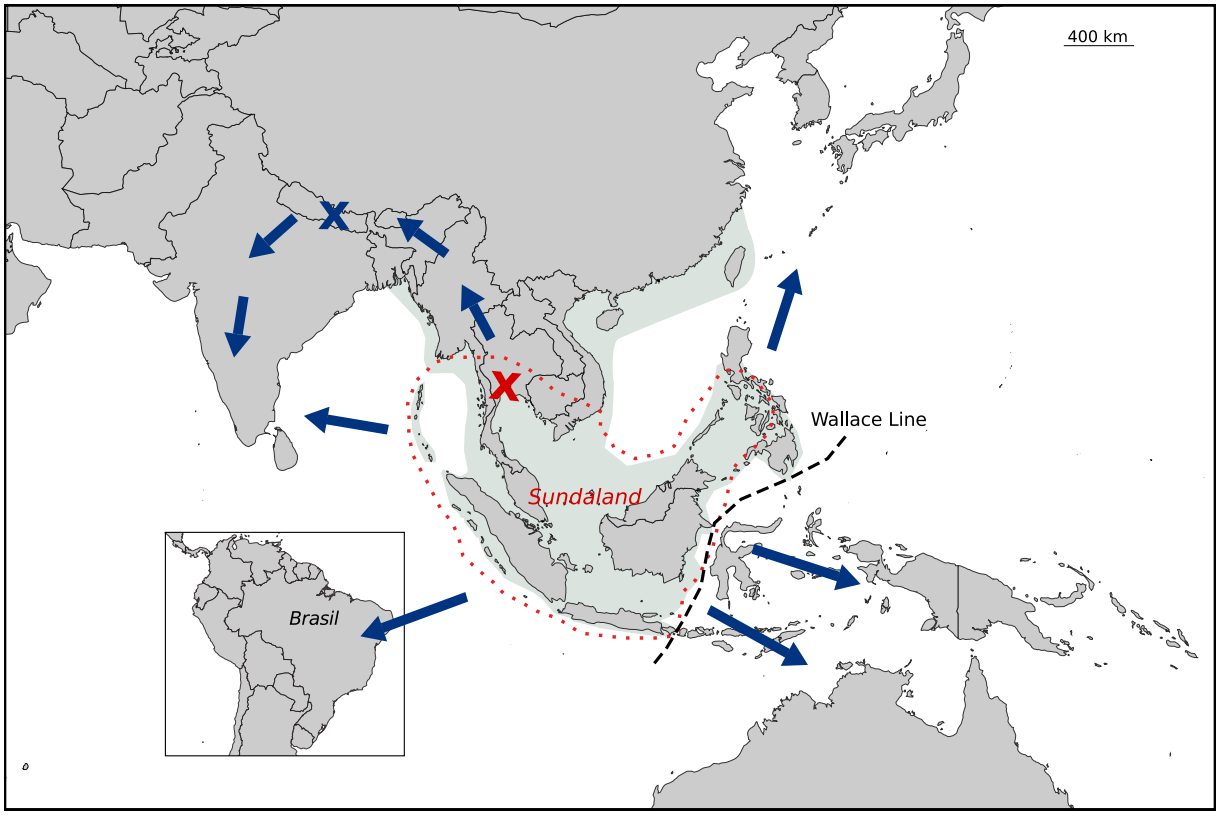


Figure 1.2: *Drosophila ananassae* evolved out of Sundaland. Inferred origin (red) and migration routes (blue) of *Drosophila ananassae* are based on the model of (Das et al., 2004). Exposure of the Sunda shelf during the late Pleistocene (18,000 years ago) is represented in green (based on the models of (Brandão et al., 2016) and (Voris, 2000)), modern coastlines are represented in gray (the world map was obtained from `http://mapsvg.com/maps/world/`). In the present work, fly strains from two sampling locations were used: Bangkok, Thailand in the ancestral species range (red X) and Kathmandu, Nepal in the derived species range (blue X).

populations of *D. melanogaster* (Hoffmann et al., 2002; Hoffmann and Weeks, 2007) and *D. serreata* (Hallas et al., 2002), in a global set of *D. melanogaster* populations (Ayrinhac et al., 2014), in European and African populations of *D. subobscura* (David et al., 2003) and in Indian populations of *D. ananassae* (Sisodia and Singh, 2010). These data suggest that CCRT is not only an ecologically relevant phenotype but also under natural selection in drosophilids. Furthermore, CCRT is influenced by age (David et al., 1998; Colinet et al., 2013), humidity (Kobey and Montooth, 2013), nutrition (Sisodia and Singh, 2012; Andersen et al., 2010, 2013), anesthesia (Nilson et al., 2006) and sex ((David et al., 1998) but see (von Heckel et al., 2016)). Controlling for these factors within the experimental setup is therefore crucial.

1.3 Local adaptation to cold in *Drosophila ananassae*: previous work and rationale

To date, there exist only a handful of studies reporting about cold tolerance investigations in *D. ananassae* (Singh and Yadav, 2015). The most extensive one was carried out by Sisodia and Singh (Sisodia and Singh, 2010). The authors compared CCRT among 45 Indian populations from different latitudes and found the phenotype to be associated with the climatic conditions at the sampling sites, consistent with local adaptation. In direct comparison, *D. ananassae* is less cold tolerant than *D. melanogaster* (Morin et al., 1997; Gibert et al., 2001).

However, a previous analysis suggests that *D. ananassae* is worthy of a more thorough examination and may be a more promising model to identify candidate genes for cold tolerance and potential targets for adaptive changes in the genome (Grath, 2010). In this previous analysis, fly strains of *D. ananassae* from a derived, temperate population in Kathmandu, Nepal and a tropical population of the ancestral species range in Bangkok, Thailand were phenotyped (the sampling sites are shown in Figure 1.2). The average annual temperature in Bangkok ranges from 30°C in summer to 25°C in winter. Kathmandu features a temperate climate where temperatures reach 28°C in summer but drop down to 3°C in winter (Figure A1). Consistent with local adaptation to the temperate climate, flies from Kathmandu recovered significantly faster from chill coma than the flies from Bangkok.

Interestingly, when comparing the single fly strains within each population, a bimodal distribution of CCRT within the Bangkok population came to light. Some strains recovered faster than others, with recovery times similar to the strains from Kathmandu (Figure 1.3, (Poxleitner, 2010)). Most strikingly, the phenotypic divergence between these two Bangkok groups was large if compared to within-population variation in *D. melanogaster* (von Heckel et al., 2016).

However, Das and colleagues reported population substructure as absent in the Bangkok population when they analyzed the same fly strains for their genetic diversity in order to infer the species' demographic history (Das et al., 2004). As the Bangkok population was identified as ancestral, the genetic variants that underlie the diverging CCRT phenotype likely represent standing genetic variation that was selected for when the species expanded to more temperate regions such as Kathmandu. Thus, exploring these genetic variants in the Bangkok population should allow for the identification

of causal genes and genetic elements without confounding effects of population divergence. Furthermore, the bimodal distribution of CCRT suggests that the causal loci are small in number but have large effects on the phenotype.

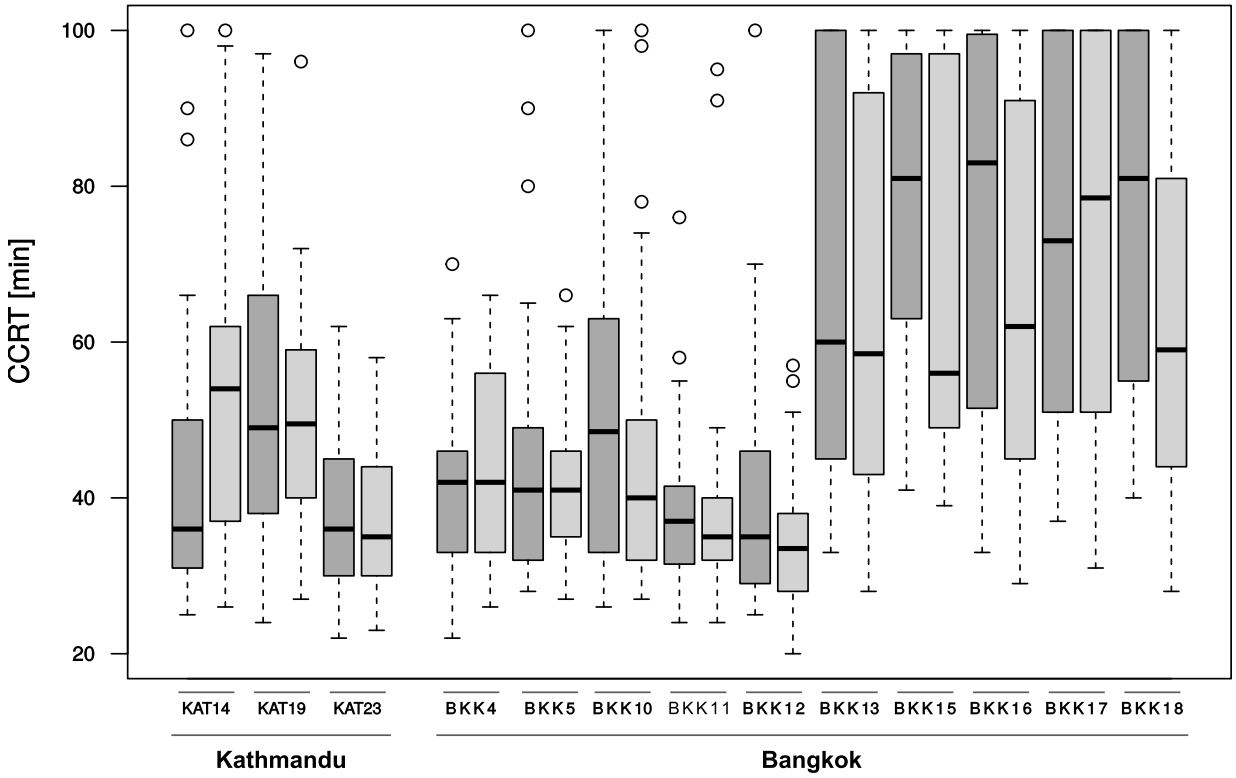


Figure 1.3: Previous tests for Chill Coma Recovery Time (CCRT) revealed a bimodal distribution among fly strains within the Bangkok population. CCRT was assessed for males and females separately and scored after a cold shock of 3 hours at 0°C. Females are shown in dark gray; males are shown in light gray. On average, 25 flies per strain and sex were tested. The data was taken from (Poxleitner, 2010).

1.4 From phenotype to genotype

Heritable genetic variants among individuals provide the molecular basis upon which adaptive evolution operates. The consequences of this process manifest on different timescales: in the short term, an allele that affects fitness in a local environment may change in frequency within a population from one generation to the next. In the long

term, this process leads to the formation of new species from a common ancestor over many generations. The nature of the underlying genetic variation is multitudinous, from single nucleotide polymorphisms (SNPs) to insertion or deletion polymorphisms (indels) to translocations or inversions. Genetic polymorphisms that underlie variation in the phenotype can occur in both coding or regulatory sequences.

Considering adaptation to low temperature, a non-synonymous mutation in the coding sequence of a gene alters the amino acid sequence, and potentially the structural conformation of its protein in a beneficial way, so that, for example, the molecular function of the protein is preserved upon cold exposure. However, pleiotropic effects are common, especially in genes with ubiquitous expression. Thus, structural mutations that are beneficial for one trait are often deleterious for another and therefore selected against (reviewed in (Paaby and Rockman, 2013)).

In contrast, mutations in regulatory sequences may have less drastic effects as, by definition, they alter expression regulation of a given gene but not the gene product itself. Such alterations are not necessarily global but can be tissue-specific or only relevant during a particular developmental stage or when exposed to a specific stimulus. A mutation in a promoter, for example, may lead to better accessibility of transcription factors, consequently speeding up transcription which may be beneficial upon recovery from chill coma. However, such patterns are generally subtle and therefore challenging to detect, especially for quantitative phenotypes with a complex genetic architecture. Disentangling to what extent both regulatory and structural changes contribute to adaptive phenotypes is a major objective of evolutionary genetics (Hoekstra and Coyne, 2007; Carroll, 2008; Romero et al., 2012) and Next Generation Sequencing (NGS) techniques allow us to address these questions.

High-throughput messenger RNA (mRNA)-sequencing has become instrumental in quantifying differential gene expression across experimental conditions. The general mRNA-sequencing workflow is shown in Figure 1.4A. Briefly, for this type of analysis, RNA is extracted from experimental samples (*e.g.*, cold-tolerant and cold-sensitive flies), reverse transcribed into copy DNA (cDNA) and then sequenced. For each gene, the normalized number of sequence reads is then analyzed across samples to highlight genes with differential regulation that are associated with the difference in the phenotype (Wang et al., 2009). However, to understand the genetic architecture of a complex trait such as CCRT in its entirety, it is necessary to consider all possible factors – structural mutations, non-coding RNAs and genomic interactions, for example, may not be

revealed by mRNA-sequencing but likewise influence the phenotype. For a quantitative phenotype, such causal variants can be mapped to chromosomal regions by statistical linkage. Quantitative trait locus (QTL) mapping requires a population of recombinant individuals, created by crossing parents that differ for the phenotype of interest. The combination of phenotype data with a genetic marker map obtained from these individuals allows one to narrow down the genomic regions that underlie the phenotypic difference, whereby the mapping power and resolution increase with the size of the mapping population, the number of crossover events in the recombinant genomes and the number and distribution of the available markers (Miles and Wayne, 2008).

A suitable method to develop a large set of markers from a large number of samples is double-digest restriction site-associated DNA (ddRAD)-sequencing (Figure 1.4B). With this method, genomic DNA is first digested, and only regions that directly flank the restriction sites are sequenced. Consequently, the complexity of the genome and the sequencing costs are greatly reduced and yet thousands of polymorphisms can be identified across samples with high accuracy (Davey et al., 2011; Peterson et al., 2012). Based on these markers, it is possible to estimate QTL positions, effects and their interactions across the whole genome (Lander and Botstein, 1989; Falconer and Mackay, 1996).

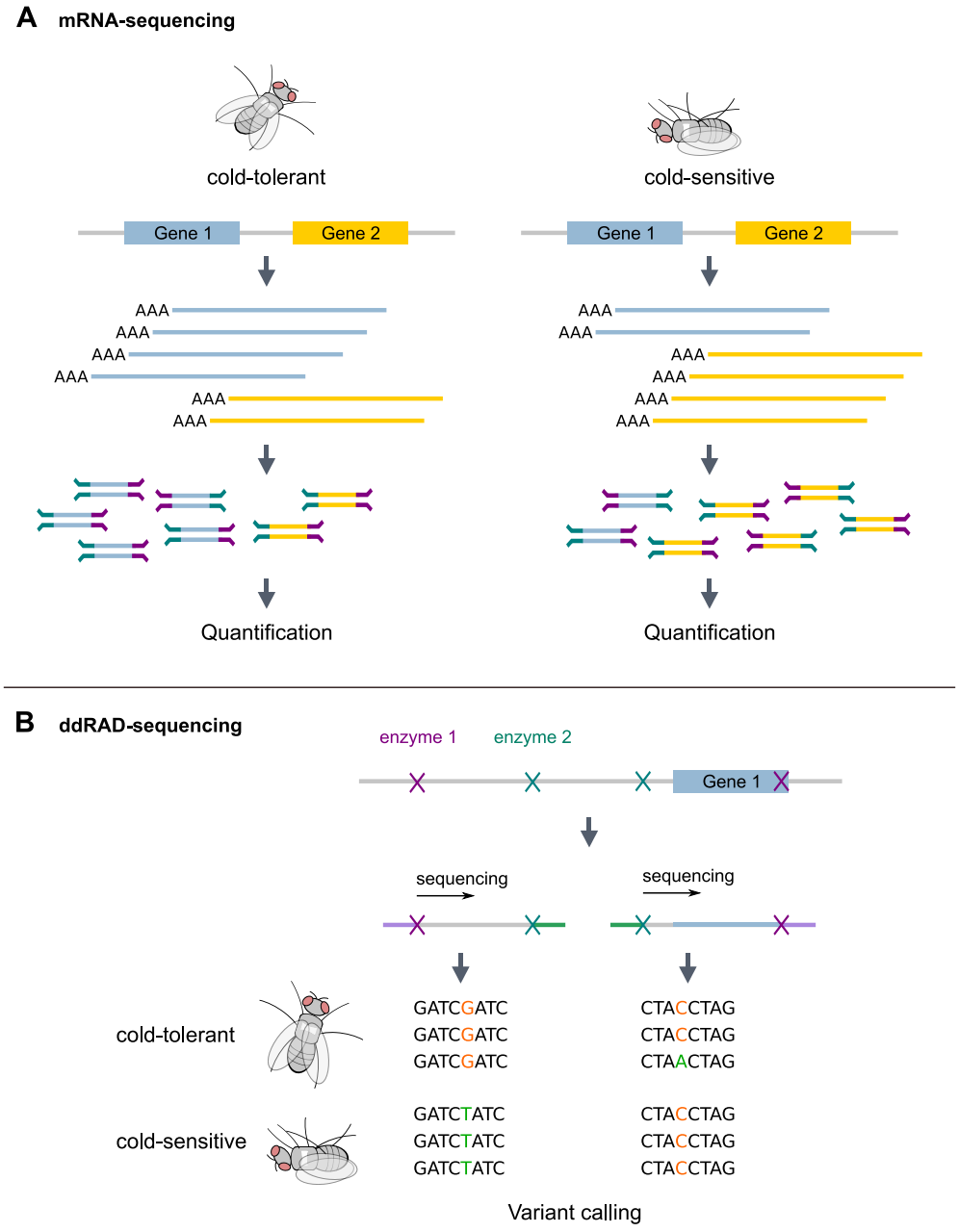


Figure 1.4: Sequencing techniques that were applied in the course of this dissertation.
A) Messenger RNA (mRNA)-sequencing is used to quantify differential gene expression across samples. First, poly(A)-tailed mRNA transcripts are enriched and reverse-transcribed to copy DNA (cDNA). The cDNA molecules are then fragmented and ligated to short adapter and barcode sequences (illustrated in green and purple). The cDNA library is sequenced using high-through techniques, resulting in short sequences which can be mapped to the reference genome or transcriptome to quantify reads per gene and sample.
B) Double-digest restriction-site associated DNA (ddRAD) markers are developed by digesting genomic DNA with two enzymes and subsequent sequencing of the adapter-ligated fragments. Combining a frequently-cutting enzyme with a rare-cutting enzyme allows for precise size selection and increases coverage of shared regions across samples.

1.5 From genotype to phenotype

Forward genetic screens such as transcriptome analyses and QTL mappings are typically used to associate genotype and phenotype. The logical next step is to establish a functional link to validate the identified candidate genes. One way to assess the contribution of a candidate gene to a trait of interest is to inhibit its expression and observe the effect on the phenotype. Drosophilids are excellent models to carry out such loss-of-function studies. They are easy to grow and mate in the laboratory, have a short generation time and, most importantly, genome editing is possible via germline transformation (reviewed in (Handler and Atkinson, 2006)). The insertion of engineered DNA sequences into the *Drosophila* genome is well established through the application of transposable element-based vector systems, such as the P-system (Rubin and Spradling, 1982; Spradling and Rubin, 1982) or the PiggyBac system (Fraser et al., 1983; Handler, 2002) (Figure 1.5B).

However, the insertion locations are always random, and targeted genome editing was not possible until the emergence of tools that build on site-specific nucleases such as Zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) (Bibikova et al., 2002; Liu et al., 2012). A less costly and time-consuming alternative for loss-of-function studies is gene silencing via RNA interference (RNAi) (Boettcher and McManus, 2015). In *D. melanogaster*, functional testing of candidate genes via RNAi is streamlined by the availability of a genome-wide RNAi library that currently covers as much as 91% of all protein-coding genes (Dietzl et al., 2007). Unfortunately, this option is not available for non-model organisms such as *D. ananassae*. Hence, the recently established CRISPR/Cas9 system promises to be a universal and powerful tool for precise and targeted genome editing of any species (Gaj et al., 2013).

Clustered regularly interspaced short palindromic repeats (CRISPRs) and CRISPR associated (Cas) proteins were initially identified as an adaptive defense system that is present in many bacteria and archaea. The system is triggered upon invasion with nucleic acids (*e.g.*, phages or plasmids) and reacts in a three-stage process: first, alien DNA sequences are incorporated into the CRISPR array in the host genome. Second, the array is transcribed and processed into CRISPR RNA (crRNA). Third, invasive sequences are recognized as complementary to the crRNAs and destroyed by DNA cleavage, whereby a short motif termed PAM (protospacer adjacent motif) is crucial for the discrimination between innate and invasive molecules (reviewed in (Al-Attar et al., 2011; Kryštofová, 2016)). In species with type I or III CRISPR/Cas systems, a multi-Cas protein complex

is involved in the recognition and cleavage of invasive DNA sequences. In contrast, the type II system mediates this process by a single ribonuclease, Cas9, which in turn requires a double-stranded crRNA molecule (trans-activating crRNA = tracrRNA) for recognition. Thus, Cas9 on its own is able to introduce a DNA double-strand break (DSB) at a specific site when guided by a complementary tracrRNA molecule (Jinek et al., 2012).

The type II CRISPR/Cas9 system was quickly adapted for targeted genome editing in a multitude of eukaryotic species, including human and mouse cells (Cong et al., 2013) and *D. melanogaster* (Gratz et al., 2013, 2014). Programmed to target genomic DNA, Cas9 introduces a DSB that triggers the cells' repair machinery. The majority of DSBs are repaired by re-ligating the break ends (non-homologous end-joining, NHEJ). NHEJ is efficient but also error-prone and often results in indels and frame-shifts (Bassett et al., 2013). Alternatively, precise recovery of the DSB is possible if a homologous sequence is available as a repair template (= homology-directed repair, HDR). This event occurs much less frequent than NHEJ but can be adapted to insert a sequence of interest such as a phenotypic marker into the DSB (Figure 1.5A). In doing so, effective gene knock-outs are possible as well as precise edits such as allele-replacements (Port et al., 2014; Gratz et al., 2014).

The components necessary for CRISPR/Cas9-mediated HDR can be delivered in various forms. In *D. melanogaster*, the most promising results for the insertion of a marker gene have been achieved with the injection of plasmids that express guide RNA (gRNA = crRNA + tracrRNA) *in vivo*, together with a double-stranded HDR repair template (= donor) into embryos with an endogenous, germline-specific source of Cas9. Using transgenic Cas9 strains greatly improved the efficiency of HDR (Kondo and Ueda, 2013; Ren et al., 2013; Gratz et al., 2014, 2015). Thus, the generation of Cas9 strains for *D. ananassae* would not only facilitate the production of knock-out strains for a large collection of cold tolerance candidate genes but also add significantly to future research on this species.

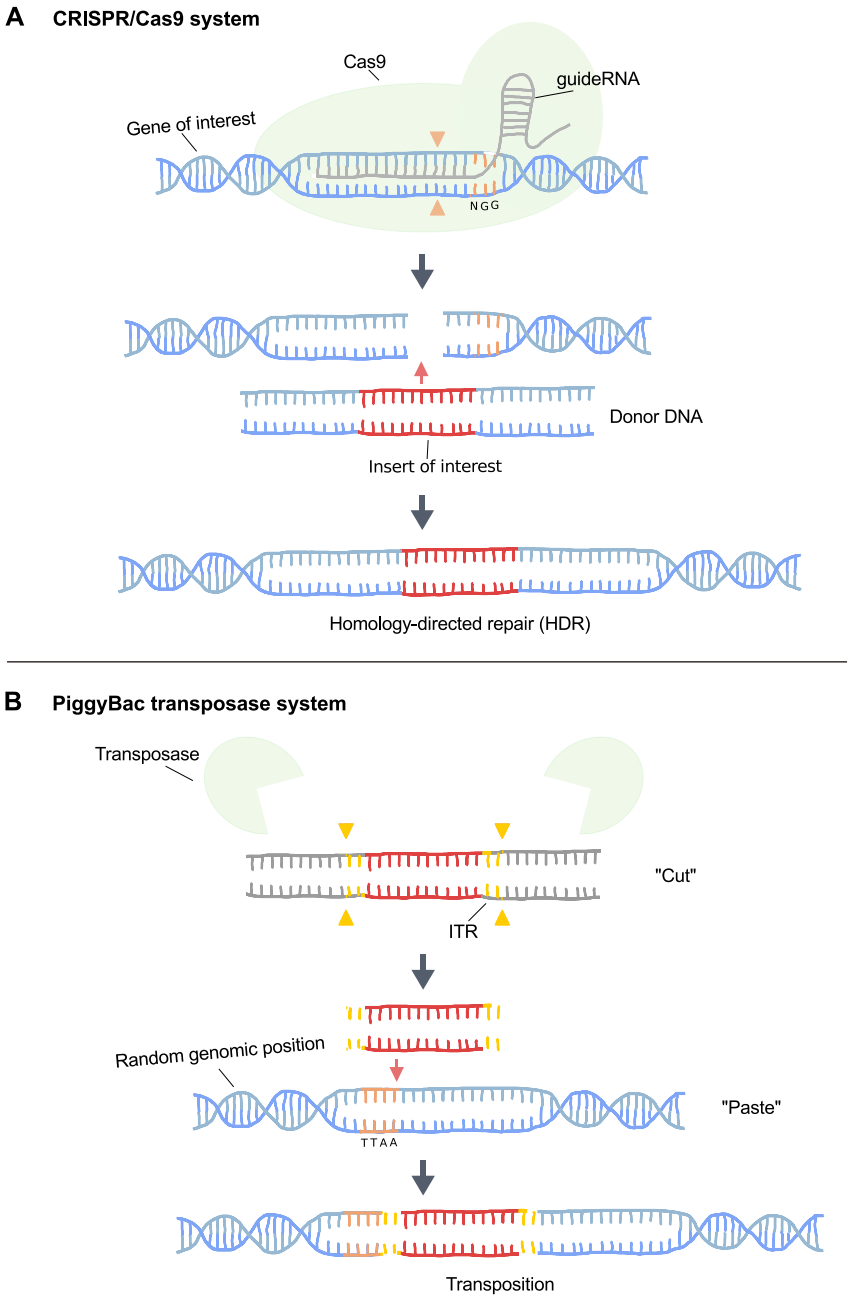


Figure 1.5: Genome editing techniques that were applied in the course of this dissertation.
A) The CRISPR/Cas9 system allows for precise insertion of DNA into a specific target site. A chimeric guide RNA is designed to match a 20 bp target sequence next to an NGG protospacer adjacent motif (PAM) and recruits Cas9 for DNA cleavage. The homology-directed repair (HDR) pathway can be adapted to insert a sequence of interest (indicated in red). For this purpose, a donor template is necessary, in which the insert of interest is flanked by sequences that are homologous to the genomic sequences flanking the target site.
B) The PiggyBac transposase system allows for insertion of DNA into TTAA landing sites that are randomly dispersed throughout the genome. The insert of interest is designed with short flanking sequences = inverted terminal repeats (ITR). The ITRs are recognized by the transposase, which then mobilizes the insert and integrates it into the genome. The integration is efficient, but it is not possible to target a specific site in the genome.

1.6 Aims and objectives

The overall aim of this thesis was to determine the molecular basis of cold tolerance in *D. ananassae*, a species that expanded from its tropical home range to temperate environments. Previous work suggests that the Bangkok population of *D. ananassae* is a promising model system to examine cold tolerance because it exhibits substantial variation in the phenotype, supposedly caused by variation at only a few causal loci. The present work comprises four main objectives, which altogether aim to identify candidate genes for cold tolerance and thus, potential targets for adaptive evolution in fly strains from the *D. ananassae* Bangkok population. The first objective was to perform high-throughput assays for Chill Coma Recovery Time (CCRT) in order to score a large sample size of individual flies for each Bangkok strain. In particular, I addressed the following questions:

- Can we confirm the bimodal distribution of the phenotype as previously identified by (Poxleitner, 2010)?
- With a larger sample size of individual flies per strain tested, which strains show the biggest difference for the phenotype?

The second objective was to identify genes with differential expression among cold-tolerant strains (fast CCRT) and cold-sensitive strains (slow CCRT) in response to cold exposure. Four cold-tolerant strains and four cold-sensitive strains were subjected to a cold shock of 3 hours at 0°C, and total mRNA profiles were extracted and sequenced at three different timepoints: 1) before the cold shock (baseline control), 2) at 15 min after the cold shock (early recovery phase) and 3) at 90 min after the cold shock (late recovery phase). The study design also matched a previously conducted transcriptome analysis in *D. melanogaster* (von Heckel et al., 2016), allowing us to compare the datasets. With this analysis, the following questions were addressed:

- Which genes are differentially expressed in response to the cold shock?
- Is there a detectable difference in transcript abundance between the two phenotypic groups already present before the cold shock?
- Are the same genes differentially expressed in *D. ananassae* as compared to *D. melanogaster*?

The third objective was to map causal genomic regions that influence the cold tolerance phenotype. The most cold-tolerant and the most cold-sensitive fly strains of the Bangkok population were used to create a panel of recombinant inbred advanced intercross lines (RIAILs) as a mapping population. Numerous molecular markers were established by ddRAD-sequencing, and a hierarchical approach was used for QTL mapping, combining standard interval mapping to identify loci of major effect and a multiple-QTL model to identify interactions among loci and estimate QTL effects. The following questions were addressed:

- Which genomic regions are associated with a faster recovery?
- What is their architecture (*i.e.*, are there epistatic or dominant effects)?
- Which differentially expressed genes are located in these regions?

The fourth objective was to initiate and test genetic tools and protocols for the first-time application of genome engineering in *D. ananassae*. The most cold-tolerant and the most cold-sensitive strains from the Bangkok population and an additional strain from the derived Kathmandu population were used to generate transgenic strains with a germline-specific source of Cas9. Further, plasmids for functional knock-outs of cold-tolerance candidate genes were constructed, and preliminary tests for CRISPR/Cas9-mediated homology-directed repair were carried out. Here, the following questions were addressed:

- To date, no such experiments have been carried out in this species. Is it possible to induce heritable genome alterations in *D. ananassae* via germline-mediated transformation?
- Can we transfer protocols and plasmids designed for transgenesis in *D. melanogaster* for functional analysis of (cold tolerance) candidate genes in *D. ananassae*?

Chapter 2

Material and Methods

2.1 Fly strains and fly maintenance

The fly strains that were used in this study (see Table B1) originate from single female flies that were previously collected from two populations: Bangkok, Thailand (coordinates = 13:50 N, 100:29 E), and Kathmandu, Nepal (coordinates = 27:49 N, 85:21 E) in 2002 (Das et al., 2004). Their offspring had been established as isofemale strains and maintained under standard laboratory conditions: all flies were kept at low density and raised in 50 ml vials on standard cornmeal molasses medium containing propionic acid and nipagin as preservatives, at a constant room temperature (22°C $\pm$ 1°C) and a constant 14 to 10 hours light to dark cycle.

2.2 Tests for chill coma recovery time

Chill coma recovery time (CCRT) was measured for male and female flies separately. At the age of 0 – 2 days, the flies were sex-separated under light CO_2-anesthesia, and ten individuals from the same strain were collected into a 50 ml vial containing 10 ml of cornmeal molasses medium. At the age of 4 – 6 days, the flies were transferred without anesthesia into new vials without food. For the cold shock, the vials were placed in an ice water bath (0°C $\pm$ 0.5°C) for exactly three hours. Back at room temperature (22°C $\pm$ 1°C), CCRT, which is defined as the time the flies needed to stand on their legs again

(David et al., 1998), was monitored in 2 min intervals for the duration of 90 min. Flies that were not awake after 90 min were assigned a recovery time of 92 min. Flies that died during the experiment ($< 1\%$) were excluded from the analysis. The generated data was analyzed in R (version 3.3.0) (R Core Team, 2018) (the script is provided in the supplementary folder CCRT.gz (see Appendix D).

2.3 Transcriptome analysis

In order to identify cold tolerance candidate genes with differential expression, total mRNA profiles were analyzed for eight fly strains of the Bangkok population (four fast-recovering and four slow-recovering strains, see Figure 3.1) at three different time-points: before the cold shock (control) and at 15 min and 90 min after a cold shock of three hours at 0°C (Figure 2.1). Additional laboratory protocols referred to in the text are provided in Appendix C. R-, python- and shell-scripts are provided in the supplementary folder transcriptome_analysis.gz (see Appendix D).

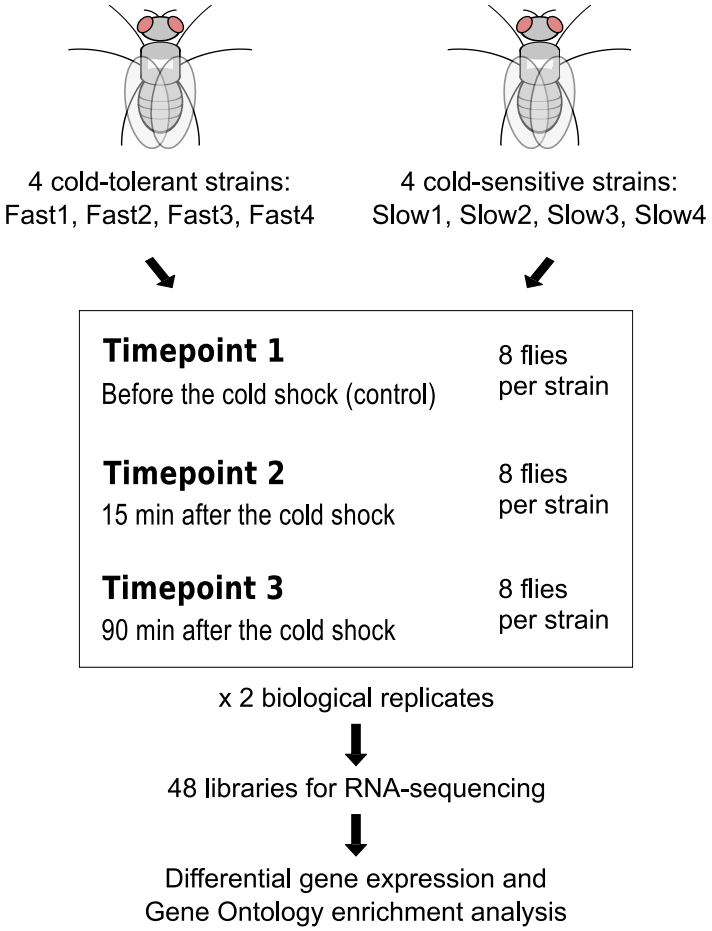


Figure 2.1: Transcriptome analysis workflow. As samples for the RNA-extractions, we used four fast-recovering and four slow-recovering fly strains. For each timepoint x fly strain combination, two biological replicates of eight pooled male flies were used. Figure modified from (Königer and Grath, 2018).

2.3.1 RNA-extraction and sequencing

RNA was extracted using the MasterPure RNA Purification Kit (Epicentre®, Illumina Company) (protocol 1, Appendix C). For each sample, eight whole male flies at 4 – 6 days of age were pooled. The cold shock was performed as described in section 2.2 (CCRT phenotyping). RNA quality was confirmed with a NanoDrop© (NanoDrop 1000, Thermo Fisher Scientific, Waltham, MA, USA) and a bioanalyzer (Bioanalyzer 2001, Agilent Technologies, provided by the LMU genomics service unit) and then sent to an external sequencing facility (GATC, Konstanz, Germany) which carried out poly(A) enrichment, fragmentation by sonication, 3'-cDNA synthesis and single-end sequencing of 50 bp reads on seven lanes of a HiSeq 2500 Illumina sequencer.

2.3.2 Read mapping and differential gene expression analysis

The raw reads were mapped to the *D. ananassae* transcriptome (including non-coding RNAs) using the annotation of FlyBase release 1.05 (Attrill et al., 2016). Mapping of the raw reads was done with NextGenMap (version 0.4.12) (Sedlazeck et al., 2013), which has been shown to produce reliable alignments in *Drosophila* (Gerken et al., 2015; Huylmans and Parsch, 2014). Differentially expressed genes were called with DESeq2 (version 1.16.1) (Love et al., 2014) as implemented in R (version 3.3.0) (R Core Team, 2018). A Wald test was used to test for significance in log2-fold changes. *P*-values were subsequently corrected for multiple testing according to Benjamini-Hochberg (Benjamini and Hochberg, 1995) and the false discovery rate (FDR) was set to 5%, *i.e.*, all genes with a corrected *P*-value ≤ 0.05 were reported as differentially expressed. DESeq2 corrects for library size and library composition using size factors that are calculated based on the given expression data. The geometric mean is calculated for each gene across all samples. The read counts for a gene in each sample are then divided by this mean. The median of these ratios in a given sample is the size factor for that sample. Further, DESeq2 uses a generalized linear model and shrinkage estimators for dispersion and fold change and thereby accounts for genes with low read counts and high dispersion. We used a two-factor design plus an interaction term ($\sim$ phenotype + timepoint + phenotype:timepoint) to analyze the effects of phenotype (with two levels: Slow and Fast) and timepoint (with three levels: control at room temperature, 15 min and 90 min after the cold shock) on gene expression levels. While using corrected *P*-values rather than

a fixed fold change cutoff for calling differential expression reflects the actual impact of the cold shock on gene expression better than an arbitrary fold change cutoff, this approach might be more prone to calling false positives.

2.3.3 Gene ontology enrichment analysis

We used the Database for Annotation, Visualization and Integrated Discovery (DAVID) (version 6.8) (Huang et al., 2009) to get an overview of Gene Ontology (GO) terms that were associated with lists of differentially expressed genes. Enrichment analysis was performed against the background of all annotated FlyBase gene IDs (Attrill et al., 2016) for three categories: biological process (BP), molecular function (MF) and cellular component (CC), with a minimum count of two genes per category to be reported. GO terms were counted as significant with an Expression Analysis Systematic Explorer (EASE) score of 0.05 after multiple testing correction according to Benjamini-Hochberg (Benjamini and Hochberg, 1995). Lists with significant GO terms were then submitted to a web server that REduces and VIsualizes Gene Ontology terms (REVIGO) (Supek et al., 2011) to remove redundant terms.

2.3.4 Comparison with *Drosophila melanogaster*

Previously, RNA-sequencing was used to analyze differential gene expression in *D. melanogaster* before and at 15 min and 90 min after a cold shock in cold-sensitive populations with slow recovery (Africa) and cold-tolerant populations with fast recovery (Europe) (von Heckel et al., 2016). It needs to be noted that in their study they applied a cold shock of seven hours, whereas we applied a cold shock with a duration of three hours only. There are two reasons for this deviation in the experimental protocol: first, using a preliminary test, we found that *D. ananassae* is more cold-sensitive and does not survive seven hours at $0°$C. Second, a three hour exposure to $0°$C for *D. ananassae* leads to similar CCRT as a seven-hour exposure to $0°$C for *D. melanogaster*. Thus, synchronizing CCRT allowed us to create a data set that can be used to compare gene expression in response to a cold shock across both species. To do so, we used the mapped *D. melanogaster* reads that were kindly provided by (von Heckel et al., 2016), and analyzed them with the same DESeq2 model as the *D. ananassae* reads.

2.4 Quantitative trait locus mapping

For QTL mapping, a panel of 94 Recombinant Inbred Advanced Intercross Lines (RIAILs) was created using the most cold-tolerant (Fast4) strain and the most cold-sensitive (Slow1) strain from the Bangkok population as founders. The RIAILs were phenotyped and, together with the founder strains, genotyped by double digest restriction site-associated DNA-sequencing (ddRAD-seq), and a hierarchical mapping approach was used to identify regions that influence the CCRT phenotype (see Figure 2.2). To obtain equal coverage of the X-chromosome but account for potential effects of pregnancy, only virgin female flies were used for phenotyping and genotyping. Additional laboratory protocols referred to in the text are provided in Appendix C. R-, python- and shell-scripts are provided in the supplementary folder QTL_mapping.gz (see Appendix D).

2.4.1 Mapping population

RIAILs were generated as follows: two initial crosses between the two parental strains were set up (Fast4 males x Slow1 females and Slow1 males x Fast4 females). Individuals from both F1 generations were mixed and allowed to mate freely with each other. Up to generation F4, intercrossing was continued in the form of mass breedings. In generation F4, 360 mating pairs were set up in separate vials to allow for one more generation of intercrossing and to initiate the inbred strains. From generation F5, full-sibling inbreeding was carried out by mating brother-sister pairs for five subsequent generations. Throughout all generations (P – F10), the parents were removed before the offspring emerged to avoid back-crosses. From generation F10 on, RIAILs were kept at low density in 50 ml vials (Figure 2.3).

2.4.2 DNA-extraction and sequencing

DNA was extracted from 94 RIAILs and the two parental strains with the DNeasy® Blood Tissue Kit (QIAGEN, Hilden, Germany) (protocol 3, Appendix C). For each fly strain, ten virgin female individuals were pooled. DNA concentrations and purity were assessed with a spectrophotometer (NanoDrop® ND 1000, VWR International, Radnor,

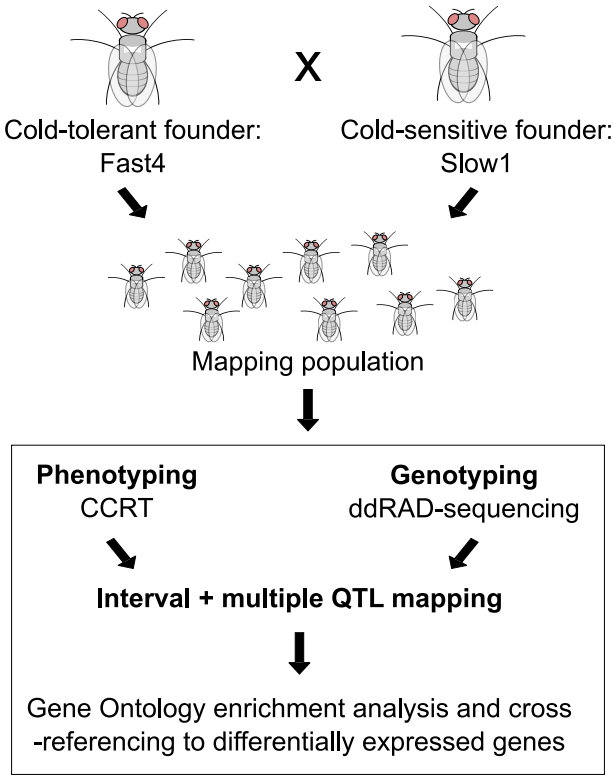


Figure 2.2: Quantitative trait locus (QTL) mapping workflow. A panel of Recombinant Inbred Advanced Intercross Lines (RIAILs) was generated from the parental strains Fast4 and Slow1. Virgin female flies of the RIAILs were phenotyped for their chill coma recovery time (CCRT) and genotyped for double-digest restriction site-associated (ddRAD) markers. Causal loci were identified with a hierarchical mapping approach, including standard interval mapping and a multiple-QTL model. QTL intervals were screened for enriched gene ontology terms and differentially expressed genes that were identified in the transcriptome analysis.

PA, USA). Library preparation and ddRAD-sequencing was carried out by an external sequencing service (ecogenics GmbH, Balgach, Switzerland) in the following way: DNA was double-digested with EcoRI and MseI and ligated to respective adapters comprising EcoRI and MseI restriction overhangs. Molecular identifier tags were added by polymerase chain reaction (PCR). The individual sample libraries were pooled and the resulting library pools were size selected for fragments between 500 – 600 bp with gel electrophoresis and extraction of the respective size range. The resulting size-selected library pools were sequenced on a NextSeqTM 500 Sequencing System (Illumina, San Diego, CA), producing single-ended reads of 75 bp length. Demultiplexing and trimming from Illumina adapter residuals were also carried out by the external service.

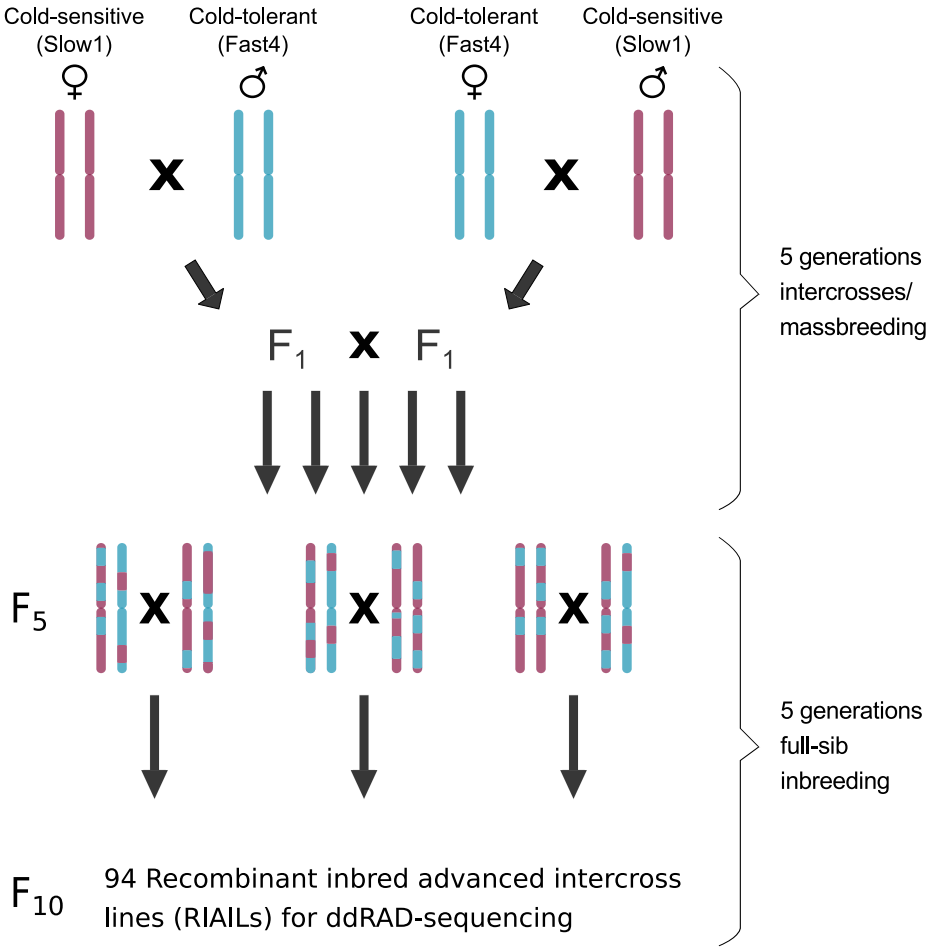


Figure 2.3: Crossing scheme for the generation of the Recombinant Inbred Advanced Intercross Lines (RAILs). Initial crosses between the parental strains Fast4 and Slow1 produced heterozygous F1 individuals, which were intercrossed up to generation F5. Subsequently, inbred strains were established by five generations of full-sibling inbreeding. Drawings of single chromosome pairs were used as representatives for the full genome.

2.4.3 Marker catalog construction

The software pipeline Stacks (version 1.45) (Catchen et al., 2011) was used to analyze the sequence data and to identify markers. First, to examine the quality of the sequence reads, the `process_radtags` program was run in Stacks, applying a sliding window size of 50% of the read length (`-w` 0.5) to filter out reads which dropped below a 99% probability of being correct (`-s` 20). Second, the processed reads of each sample were

mapped to the *D. ananassae* reference genome with NextGenMap (version 0.5.0) (Sedlazeck et al., 2013) using the annotation of FlyBase release 1.05 (Attrill et al., 2016). Third, the mapped reads were converted to bam format, sorted and indexed with samtools (version 0.1.18) (Li et al., 2009).

Fourth, the `ref_map.pl` wrapper program was run in Stacks, which executes the Stacks core pipeline by running each of the Stacks components individually. Briefly: `pstacks` assembled ddRAD loci for each sample, `cstacks` created a catalog of ddRAD loci from the two parental samples to create a set of all possible alleles expected in the mapping population and `sstacks` matched all RIAIL samples against the catalog. The `genotypes` program was executed last, applying automated corrections to the data (`-c`) to correct for false negative heterozygote alleles. Only those loci which were present in at least 75% of the samples were exported (`-r` 75).

Fifth, we applied additional corrections to the catalog by running the `rxstacks` program with the following filtering settings: non-biological haplotypes unlikely to occur in the population were pruned out (`--prune_haplo`), SNPs were recalled once sequencing errors were removed using the bounded SNP model (`--model_type_bounded`) with an error rate of 10% (`--bound_high` 0.1), and catalog loci with an average log likelihood less than –200 were removed (`--lnl_lim` -200.00). Sixth, `cstacks`, `sstacks` and `genotypes` (`-r` 75) were rerun to rebuild, match and export a new catalog with the filtered SNPs. `Load_radtags.pl` and `index_radtags.pl` were used to upload and index the new catalog to a MySQL database. Seventh, a custom R script was used for to remove markers with extreme values of residual heterozygosity within RIAILs, using cutoffs based on our inbreeding scheme ($\leq 15\%$ and $\geq 35\%$) (Falconer and Mackay, 1996) and to remove markers with an allele frequency drift $\geq 10\%$ from further analysis. Eighth, the MySQL database was used to check the markers for errors manually. A total of 1,400 markers were included in the downstream analysis.

2.4.4 Genetic map construction

Genetic map construction was conducted with R/qtl (Broman et al., 2003) as implemented in R (version 3.2.3) ((R Core Team, 2018). The function `countXO` was used to remove seven RAILs with > 200 crossover events, and one more RIAIL was removed due to a low number of genotyped markers (< 700). The downstream analysis then included 1,400 markers and 86 RIAIL-samples which were analyzed in an F5-intercross environment. Markers were partitioned into linkage groups based on a logarithm of

the odds score (LOD score) threshold of 8 and a maximum recombination frequency of 0.35, assuming a sequencing error rate of 1%. Map distances were calculated using the Haldane map function.

2.4.5 Analysis of quantitative trait loci (QTL)

QTL mapping was conducted with R/qtl (Broman et al., 2003). Prior to mapping, the genotype probabilities between marker positions were calculated with the function `calc.genoprob` on a grid size of 1 cM. Standard interval mapping was performed to identify major QTL, using the Expectation Maximization (EM) algorithm as implemented with the `scanone` function. The results are expressed as LOD score (Sen and Churchill, 2001). Significance thresholds were calculated with 1,000 genome-wide permutations. The single-QTL scan was extended with a more complex, two-dimensional scan using Haley-Knott regression as implemented with the `scantwo` function. Significance thresholds were again calculated with 1,000 genome-wide permutations.

To screen for additional QTL, estimate QTL effects and refine QTL positions, multiple-QTL mapping was performed (Arends et al., 2010). Here, missing genotypes were simulated from the joint distribution using a Hidden Markov model with 1,000 simulation replicates and an assumed error rate of 1% as implemented with the function `sim.geno`. The MQM model was identified with a forward selection/backward elimination search algorithm as implemented with the `stepwise` function, with the model choice criterion being penalized LOD scores. The penalties were derived from the significance permutations of the two-dimensional genome scan. To estimate the support interval for each identified QTL, an approximate 95% Bayesian credible interval was calculated as implemented by the `bayesint` function. Gene annotations for QTL intervals were downloaded from FlyBase (Attrill et al., 2016) and screened for enriched GO terms and enrichmed Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways with DAVID (version 6.8) (Huang et al., 2009) as described in chapter 2.3.3. In addition, we cross-referenced the QTL gene lists with lists of differentially expressed genes obtained from the transcriptome analysis and cold tolerance candidate genes from the literature.

2.5 Genome engineering

The general workflow for loss-of-function studies of cold tolerance candidate genes in *D. ananassae* is shown in Figure 2.4. In order to set the groundwork for these experiments, Cas9 strains were generated from the wild-type strains Fast4 and Slow1 from Bangkok and an additional strain from Kathmandu, KAT1, using the PiggyBac system for transgenesis (Figure 2.4A). Furthermore, preliminary work was done towards testing the functionality of the Cas9 strains for homology-directed repair (HDR). To do so, protocols initially developed for usage in *D. melanogaster* were adapted for the construction of gRNA and donor plasmids. The plasmids were designed to target two of the identified cold tolerance candidate genes: *GF14647* and *GF15058*. For the latter gene, the plasmids were injected into the germline of all three Cas9 strains (Figure 2.4B). Additional laboratory protocols referred to in the text are provided in Appendix C.

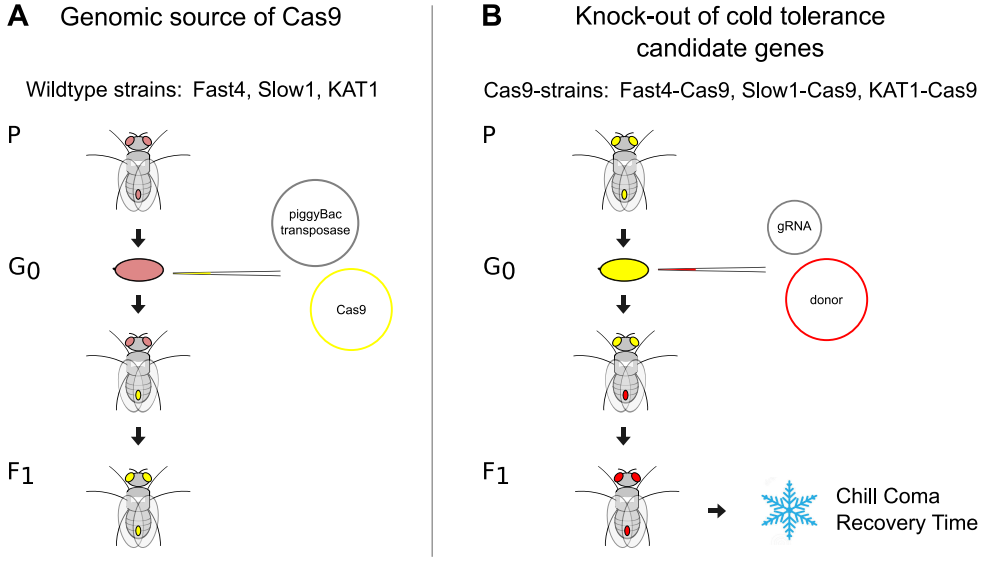


Figure 2.4: General workflow for loss-of-function studies of cold tolerance candidate genes. **A**) An endogenous source of Cas9 is inserted into the genome via germline transformation with the PiggyBac transposase. Successful transformation can be recognized by EYFP expression in the F1 generation (flies with yellow eyes). **B**) gRNA and donor plasmids for homology-directed repair are injected into the germline of the Cas9 strains, targeting the coding sequence of a cold tolerance candidate gene. Successful gene knock-out can be recognized by dsRed expression in the F1 generation (flies with bright red eyes).

2.5.1 Generation of Cas9 strains

The PiggyBac system was used to integrate Cas9 into the genomes of Fast4, Slow1 and KAT1. The PiggyBac vector pBac-nos:Cas9-3XP3:EYFP was provided by Prof. Alistair McGregor, Oxford Brookes University, Oxford, UK, and initially developed by David Stern, Janelia Research Campus, Ashburn, VA, USA (Stern et al., 2017). It contains the codon-optimized *Streptococcus pyogenes* Cas9 endonuclease which is under the control of the *D. melanogaster nanos* promoter and 3'UTR. *Nanos* is expressed in the male and female germline and the 3'UTR targets protein synthesis to the posterior pole of the embryo (Doren et al., 1998). Additionally, the construct contains the Enhanced Yellow Fluorescent Protein (EYFP) marker which is under the control of the 3XP3 promoter. 3XP3 drives expression in adult and larval eyes of drosophilids (Berghammer et al., 1999; Horn et al., 2000). The *nanos* sequence, Cas9, EYFP and 3XP3 (from now on referred to as Cas9 construct or Cas9 insert) are flanked by two inverted terminal repeat sequences which are recognized by the piggyBac transposase (Figure 2.5, Figure A6).

Figure 2.5: Schematic illustration of the PiggyBac vector pBac-nos:Cas9-3XP3:EYFP. The inverted terminal repeat sequences that serve as recognition signals for the PiggyBac transposase are denoted with P. The total construct size is 9,381bp, see Figure A6 for a detailed vector map.

The pBac-nos:Cas9-3XP3:EYFP plasmid was isolated from a bacterial culture with the QIAGEN® Plasmid Midi Kit (protocol 11, Appendix C), as well as a helper plasmid coding for the PiggyBac transposase (provided by Helène Hinaux, LMU Munich, Planegg-Martinsried, GER). For subsequent germline transformation, the plasmids were mixed in the following concentrations: 0.6 µg/µl pBac-nos:Cas9-3XP3:EYFP and 0.4 µg/µl helper plasmid.

Germline transformation:
Two different approaches were used for germline transformation: **1)** the plasmids and fly strains were sent to an external facility, Rainbow Transgenic Flies, Inc (Camarillo, CA, USA) which carried out microinjections and **2)** the plasmids were injected into pre-blastoderm embryos following the protocol of (Gompel and Schröder, 2015). Briefly, the procedure was as follows: three to four days prior to the injections, approximately 600 young flies (2 – 5 days old) of the strain to be injected were distributed into three mating cages. During this acclimatization period, egg-laying plates (protocol 16, Appendix C) were prepared with fresh yeast paste and changed several times per day. From two to three hours before the injections, plates were changed every 30 min. Microcapillary needles with an inside diameter of 0.5 µm (Eppendorf, Hamburg, Germany) were loaded with 2 µl injection mix (0.6 µg/µl pBac-nos:Cas9-3XP3:EYFP and 0.4 µg/µl helper plasmid). For each new cycle of injections, eggs were collected after 30 min, rinsed with H_2O and aligned with a brush on a cover slide (preliminary experiments showed that the chorion is translucent enough to allow for stage determination; therefore, embryos were not washed with 96% Ethanol as suggested by (Gompel and Schröder, 2015)). Aligned eggs were desiccated in a chamber with silica beads for 1 – 3 min, covered with extra virgin olive oil and the cover slide was mounted on a microscope (Axiovert 25, Zeiss, Oberkochen, Germany).

The plasmid mix was injected into the posterior end of the embryos, where in early developmental stages (younger than 1 hour) the precursor germ cells build a syncytium (Venken and Bellen, 2007). A FemtoJet express microinjector (Eppendorf, Hamburg, Germany) was used to adjust injection time and injection and compensation pressures as needed to release a volume of approximately 1/5 of the embryos' size. After the injections, as much as possible of the olive oil was removed by gravity flow and the cover slides were placed vertically into food vials with the posterior end of the embryo facing upwards and the anterior end slightly touching the surface of the food.

Marker screening and generation of homozygous Cas9 strains:
Injected embryos that developed into pupae were transferred into separate food vials. After emerging, adult flies were mated to either male or virgin female flies of the corresponding wild-type strain. Crosses with viable offspring were numbered consecutively to initiate a separate strain for each potential insertion. Larvae of the F1 generation were collected from the food vials with a wet brush and transferred to Petri dishes with 1X

phosphate buffered saline (PBS). Each Petri dish was screened for larvae with EYFP expression on a stereoscope (Leica M 165 FC, Leica, Wetzlar, Germany) using a 491 nm filter. Larvae that were positive for EYFP expression were rescued into fresh food vials. Once they had reached adulthood, males and females from the same cross number were allowed to mate freely with each other.

Flies from the next generation (F2) were used to map the Cas9 insert location with inverse PCRs as follows: DNA was extracted from five flies (males and females mixed) using the MasterPure Kit (Epicentre) (protocol 2, Appendix C) and digested with either NsiI or MspI. Digested DNA fragments were self-ligated with T4 DNA Ligase (New England Biolabs, Ipswich, MA, USA), precipitated and used as templates for PCR amplification with Taq Polymerase (Invitrogen, Waltham, MA, USA) (protocol 17, Appendix C) using primers #3 (B7) that was designed to match to the Cas9 insert as described in `https://fruitfly.org/about/methods/inverse.pcr.html`. PCR products were analyzed on a 1% agarose gel and samples with single bands were cleaned from primers and residual dNTPs and sequenced using BigDye® Terminator v1.1 chemistry and primers #3 (Table B7) on an ABI 3730 automated sequencer (Applied Biosystems, Foster City, CA, USA) (protocol 6, Appendix C). The sequence reads were assembled to the FlyBase reference sequence (Attrill et al., 2016) using CLC Main Workbench (version 7.9.1) (QIAGEN Aarhus, Aarhus, Denmark). The exact genomic location of the insert was obtained by blasting (Altschul et al., 1990) the genomic part of the sequence against the *D. ananassae* reference genome, release 1.05 (Attrill et al., 2016).

Subsequently, homozygous Cas9 strains were generated as follows: from generation F4, single males and virgin females were collected and mated in pairs. They were allowed to mate and produce offspring for about two weeks before they were used for genotyping. Their DNA was extracted using the MasterPure Kit (Epicentre) (protocol 2, Appendix C). Endogenous Cas9 inserts were PCR-amplified using LongAmp® Taq Polymerase (New England Biolabs, Ipswich, MA, USA) (Appendix X, protocol 5) using primers #4 – #6 (Table B7) that were designed to flank the insertion. PCR products were analyzed on a 1% agarose gel. Offspring from parents that were homozygous for the insertion were retained and established as homozygous strains.

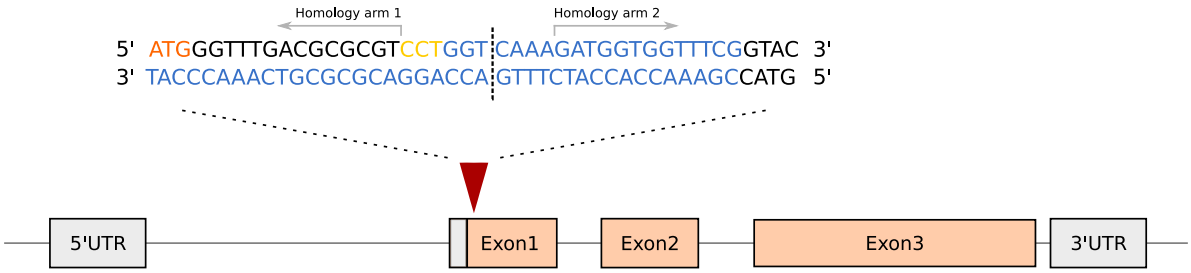
2.5.2 Sanger sequencing of *GF14647* and *GF15058*

Genomic DNA was extracted from single male flies of eight strains from Bangkok and three strains from Kathmandu using the Epicentre Kit (Epicentre®, Illumina Company) (protocol 2, Appendix C). After concentration and purity of the samples had been affirmed with a spectrophotometric measurement (NanoDrop 1000, Thermo Fisher Scientific, Waltham, MA, USA), polymerase chain reactions (PCRs) were conducted to amplify the genomic regions of interest using a protocol for Phusion® High-Fidelity DNA Polymerase (Thermo Fisher Scientific, Waltham, MA, USA) (protocol 4, Appendix C) with primers #1 and #2 (Table B7). The resulting PCR fragments were analyzed on a 1% agarose gel, cleaned from primers and residual dNTPs and sequenced using BigDye® Terminator v1.1 chemistry on an ABI 3730 automated sequencer (Applied Biosystems, Foster City, CA, USA) (protocol 6, Appendix C) and primers #1 – #31 (Table B8). The sequence reads were assembled to the *D. ananassae* reference sequence (release 1.05) that was obtained from FlyBase (Attrill et al., 2016) using CLC Main Workbench (version 7.9.1) (QIAGEN Aarhus, Aarhus, Denmark).

2.5.3 Guide RNA and donor plasmid construction for knock-out of two cold tolerance candidate genes

The choice of the target site for CRISPR/Cas9-mediated DSB within the coding sequence of the two cold tolerance candidate genes *GF14647* and *GF15058* was based on the following criteria: no predicted off-target effects as affirmed with the CRISPR optimal target finder `http://tools.flycrispr.molbio.wisc.edu/targetFinder/` (Gratz et al., 2014), no polymorphisms present in the gRNA sequence as affirmed by sequence analysis of the wild-type strains (see chapter 2.3.3), high predicted cleavage efficiency (score > 7) as affirmed with the CRISPR efficiency predictor `flyrnai.org/evaluateCrispr/` (Housden et al., 2015). The cleavage site for *GF14647* was located 20 bp downstream of the start codon, at position 5.868.772 – 5.868.773 on scaffold 12916. The homology arms for the donor plasmids flank the cleavage site with a total distance of 10 bp. The cleavage site for *GF15058* was located 38 bp downstream of the start codon, at position 2.420.775 – 2.420.776 on scaffold 12916. Here, the homology arms for the donor plasmids flank the cleavage site with a distance of 11 bp total (Figure 2.6).

GF14647:



GF15058:

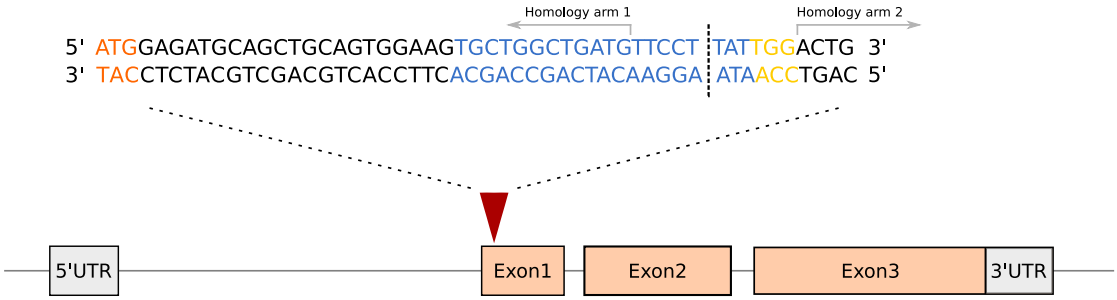


Figure 2.6: Design of guide RNA (gRNA) and homology arms for *GF14647* and *GF15058*. Transcription start codons are highlighted in orange, protospacer adjacent motif (PAM) sites are highlighted in yellow, DNA sequences complementary to the gRNAs are highlighted in blue, and the predicted double-strand break sites are indicated by the vertical dotted lines. UTR = untranslated region.

Construction of gRNA plasmids:

The gRNA insert containing BbsI restriction site overhangs was synthesized using T4 Polynucleotide Kinase (New England Biolabs, Ipswich, MA, USA) from single-stranded oligonucleotides (Table 2.1) (protocol 15, C). The plasmid pCFD3-dU6:3gRNA (which was provided by Simon Bullock via Addgene (plasmid #49410, `http://n2t.net/addgene:49410`) expresses gRNA under the control of the ubiquitous *Drosophila* U6:3 promoter (Port et al., 2014). It was digested with BbsI (New England Biolabs, Ipswich, MA, USA), purified from a 1% agarose gel with the QIAquick Gel Extraction Kit (protocol 13, Appendix C) and then ligated with a 1:200 dilution of the gRNA insert using T4 DNA Ligase (New England Biolabs, Ipswich, MA, USA) in two separate reactions

for *GF14647* and *GF15058* (Figure 2.7A). 5 µl of each ligated plasmid was transformed into One Shot® TOP10 Competent Cells (Invitrogen, Waltham, MA, USA). DNA was isolated with the QIAprep Spin® Miniprep Kit (QIAGEN, Venlo, Netherlands) from liquid lysogeny broth (LB) cultures (100 µg/ml Ampicillin) which had been inoculated with selected colonies. Isolated plasmids were sequenced using BigDye® Terminator v1.1 chemistry on an ABI 3730 automated sequencer (Applied Biosystems, Foster City, CA, USA) with primers #32 and #33 (Table B8). The sequence reads were examined for successful ligation in CLC Main Workbench (version 7.9.1) (QIAGEN Aarhus, Aarhus, Denmark) and both plasmids were grown in 50 ml LB cultures to be isolated in larger quantities with the QIAGEN Plasmid Midi Kit (QIAGEN, Venlo, Netherlands) (protocol 11, Appendix C).

	GF14647	*GF15058*
sense	GTCGGTCAAAGATGGTGGTTTCG	GTCGTGCTGGCTGATGTTCCTTAT
antisense	AAACCGAAACCACCATCTTTGACC	AAACATAAGGAACATCAGCCAGCA

Table 2.1: Oligonucleotide sequences for guide RNA (gRNA) synthesis. BbsI restriction site overhangs are shown in black, the sequences highlighted in blue are expressed as gRNA.

Construction of donor plasmids:

The donor plasmid template pDsRed-attP was provided by Melissa Harrison, Kate O'Connor-Giles and Jill Wildonger via Addgene (plasmid #51019 `http://n2t.net/addgene:51019`). It contains two cloning sites for homology arms that flank the fluorescent dsRed marker which is under the control of the 3XP3 promoter. In combination with Cas9 and gRNA, the 3XP3-dsRed construct gets inserted into the target site. In total, six different versions of pDsRed-attP were constructed, one for each gene and fly strain combination. Homology arms were amplified directly from the genome. The exact position and length of each homology arm was chosen based on the following criteria: length of approximately 1 kb, starting within 20 bp distance to the predicted target cleavage site but excluding the PAM site (see Figure 2.6), no polymorphisms present in the forward and reverse primer sequences as affirmed by sequence analysis of the wild-type strains, no secondary structures within primer sequences as affirmed at `http://biotools.nubic.northwestern.edu/OligoCalc.html` (Kibbe, 2007) and no AarI or SapI restriction sites within the homology arms sequences.

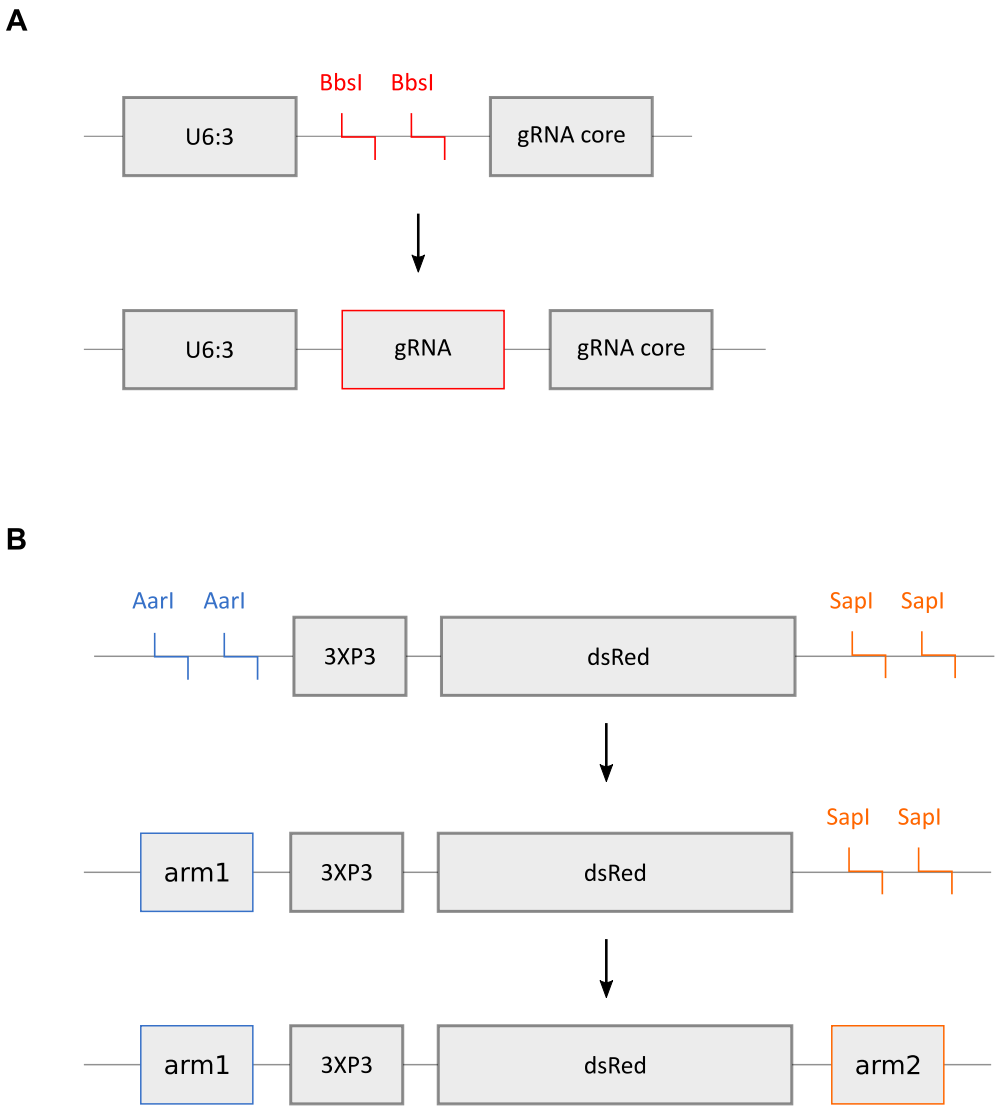


Figure 2.7: Cloning scheme for guide RNA (gRNA) (**A**) and donor (**B**) plasmids. The double-stranded gRNA is cloned into the BbsI restriction site of the pcFD3:U63 plasmid. Homology arms are cloned in two steps into the pHD-DsRed-attP plasmid: homology arm 1 is cloned into the AarI site, and homology arm 2 is cloned into the SapI site.

Forward and reverse primers were designed with restriction site overhangs according to the cloning plan: AarI for homology arm 1 and SapI for homology arm 2 (primers #7 – #10 in Table B7 and see Figure 2.7). Genomic DNA of Fast4, Slow1 and KAT1 was extracted from five pooled flies using the MasterPure DNA Purification Kit (Epicentre Madison, WI, USA) (protocol 2, Appendix C). In total, 12 different homology arms were amplified from genomic DNA using Phusion ® High-Fidelity DNA Poly-

merase (Thermo Fisher Scientific, Waltham, MA, USA) (protocol 4, Appendix C) and then cleaned with the QIAquick PCR Purification Kit (protocol 12, Appendix C). The initial pHD-DsRed-attP plasmid was digested with AarI (New England Biolabs, Ipswich, MA, USA), purified from a 1% agarose gel with the QIAquick Gel Extraction Kit (protocol 13, Appendix C) and then ligated in six separate reactions with homology arm 1 for *GF14647* and *GF15058* of Fast4, Slow1 and KAT1 (protocols 7 and 8, Appendix C). 5 µl of each ligated plasmid was transformed into One Shot® TOP10 Competent Cells (Invitrogen, Waltham, MA, USA) (protocol 9, Appendix C). Four colonies per plasmid were selected for colony PCR to confirm successful ligation, using Taq DNA Polymerase (Invitrogen, Waltham, MA, USA) and primers #11 (Table B7, protocol 14, Appendix C).

For each gene and fly strain, one successfully ligated plasmid was selected and isolated from an inoculated liquid LB culture with the ZyppyTM Plasmid Miniprep Kit (ZYMO Research, Irvine, Ca, USA) (protocol 10, Appendix C). Each of the six plasmids with homology arm 1 was digested with SapI (New England Biolabs, Ipswich, MA, USA), purified from a 1% agarose gel and then ligated with the corresponding homology arm 2. Again, colony PCRs were performed with primers #12 (Table B7) to confirm successful ligation. For each gene and fly strain, one successfully ligated plasmid was selected and isolated from an inoculated liquid LB culture with the ZyppyTM Plasmid Miniprep Kit and sequenced using BigDye® Terminator v1.1 chemistry on an ABI 3730 automated sequencer (Applied Biosystems, Foster City, CA, USA) with primers #32 and #33 (Table B8) to confirm integration of both homology arms. The sequence reads were examined for successful ligation in CLC Main Workbench (version 7.9.1) (QIAGEN Aarhus, Aarhus, Denmark). The final donor plasmids were grown in 50 ml LB cultures to be isolated in larger quantities with the QIAGEN Plasmid Midi Kit (QIAGEN, Venlo, Netherlands) (protocol 11, Appendix C).

2.6 CRISPR/Cas9-mediated homology-directed repair to knock out *GF15058*

Injections for germline transformation of the gRNA plasmid and the corresponding donor plasmids constructed to target the candidate gene *GF15058* were carried out by Rainbow Transgenic Flies, Inc (Camarillo, CA, USA) (see chapter 2.5.1). For this purpose, all three Cas9 strains were sent to the facility, together with the gRNA plasmid and the three donor plasmids. Injections were carried out with a mix of 0.1 µg/µl gRNA and 0.5 µg/µl donor plasmid. Injected embryos that developed into larvae were sent back from the external facility. Larvae that developed into adults were mated to the corresponding wild-type strain, and the F1 larvae were screened for dsRed expression.

Chapter 3

Results

3.1 Chill coma recovery time in fly strains from Bangkok

The high-throughput CCRT assays revealed four fast strains and four slow strains within the Bangkok population (Figure 3.1, Table 3.1), which is in accordance with previous results (Poxleitner, 2010). The strains were renamed according to their CCRT as either Fast or Slow (see Table B1). Across all tested strains, CCRT ranged from a minimum of 12 min to a maximum of $>$ 92 min. For each sex, the effect of the fly strain on CCRT was highly significant (Kruskal-Wallis test, $P <$ 2.2e-16). We compared CCRT among strains within and across the two phenotypic groups. Pairwise differences were significant for each FastxSlow combination, *i.e.*, non-significant differences were only found within the Fast group and within the Slow groups but not across groups (Tables 3.2 and 3.3).

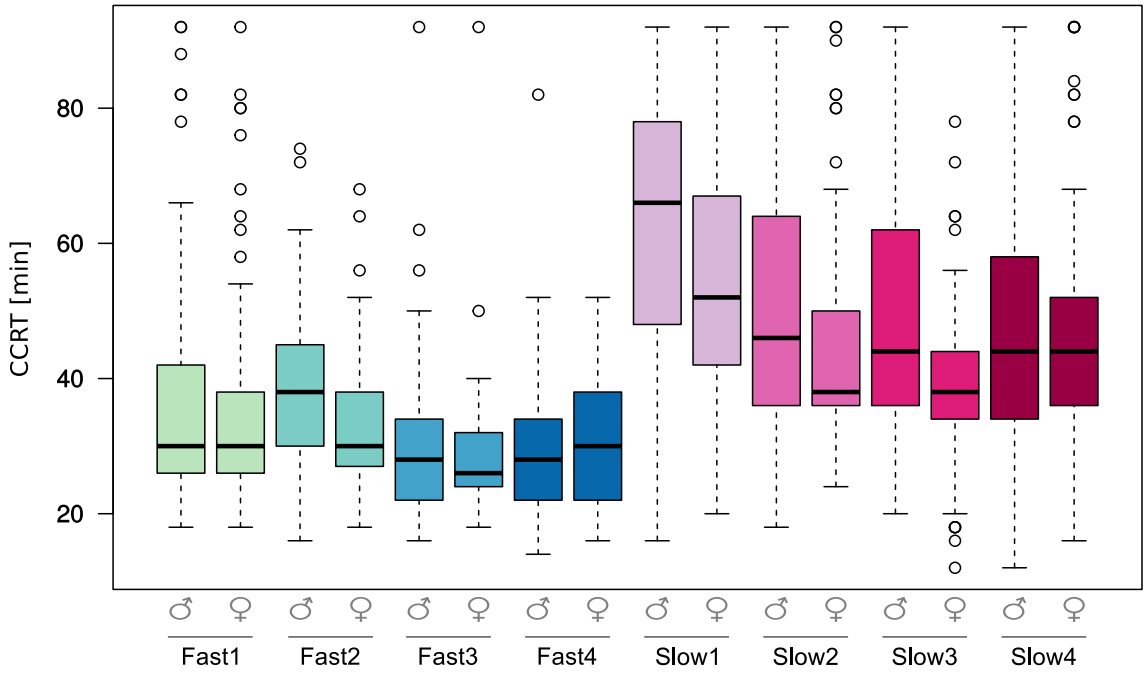


Figure 3.1: Chill Coma Recovery Time (CCRT) in eight isofemale fly strains from Bangkok. CCRT was scored for male and female flies separately at the age of 4 – 6 days after a cold shock of 3 hours at 0°C.

Table 3.1: Chill coma recovery time in fly strains from Bangkok

	Males			Females		
Strain	**CCRT [min]**	**N**	**StDev**	**CCRT [min]**	**N**	**StDev**
Fast1	35.05	99	14.89	37.28	100	18.25
Fast2	33.18	88	9.61	38.30	88	11.69
Fast3	28.42	100	8.61	30.20	100	11.14
Fast4	30.13	110	8.65	29.29	110	9.90
Slow1	53.92	100	18.40	63.70	100	20.70
Slow2	45.25	88	15.57	51.50	100	21.10
Slow3	38.28	86	11.93	53.27	110	19.87
Slow4	46.00	109	16.33	47.20	110	17.62

N = number of tested individuals (flies that died during the experiment (< 1%) were not included), StDev = standard deviation. The raw data for each individual are provided in a supplementary folder, CCRT.gz, see Appendix D.

Table 3.2: Significance of pairwise comparisons of the Bangkok strains for chill coma recovery time in male flies

	Fast1	**Fast2**	**Fast3**	**Fast4**	**Slow1**	**Slow2**	**Slow4**
Fast2	0.83221						
Fast3	0.00049	7.7e-05					
Fast4	0.04493	0.03849	0.13974				
Slow1	1.0e-12	1.6e-14	$<$ 2e-16	$<$ 2e-16			
Slow2	1.5e-09	8.7e-10	$<$ 2e-16	3.3e-14	0.00032		
Slow3	0.00089	0.00061	1.3e-11	1.0e-06	3.7e-09	0.01549	
Slow4	3.7e-09	6.3e-10	$<$ 2e-16	4.1e-15	0.00113	0.56283	0.00155

P-values were calculated with a Wilcoxon Rank Sum test and corrected for multiple testing according to (Benjamini and Hochberg, 1995). Non-significant *P*-values are highlighted in red.

Table 3.3: Significance of pairwise comparisons of the Bangkok strains for chill coma recovery time in female flies

	Fast1	**Fast2**	**Fast3**	**Fast4**	**Slow1**	**Slow2**	**Slow3**
Fast2	0.02622						
Fast3	0.00446	2.0e-07					
Fast4	0.00217	2.6e-08	0.80742				
Slow1	3.4e-15	2.3e-15	$<$ 2e-16	$<$ 2e-16			
Slow2	2.1e-08	3.5e-05	2.2e-16	$<$ 2e-16	7.3e-05		
Slow3	1.3e-08	3.3e-05	$<$ 2e-16	$<$ 2e-16	9.4e-06	0.85640	
Slow4	2.4e-07	0.00048	1.4e-15	$<$ 2e-16	1.9e-08	0.32045	0.35120

P-values were calculated with a Wilcoxon Rank Sum Text and corrected for multiple testing according to (Benjamini and Hochberg, 1995). Non-significant *P*-values are highlighted in red.

3.2 Transcriptome analysis

3.2.1 Transcriptome overview

We obtained transcriptome data from 48 cDNA libraries with an average of 49 Mio. reads and Phred quality scores above 30. The sequence quality was confirmed with FastQC (version 0.11.4) (Andrews, 2014). The proportion of reads that could be mapped to the reference was $> 94\%$ for each sample. A principal component analysis (PCA) based on the 500 most variable genes revealed tight clustering of the biological replicates (Figure 3.2). The first principal component accounted for 22% of the variance and clearly separated the three timepoints. The second principal component accounted for 19% of the variance and separated the fly strains from each other. Note that the data grouped according to strain and timepoint irrespective of subdividing the data (see also Appendix A, Figure A2 for PCA plots based on the 250 most variable genes and based on all genes).

3.2.2 Analysis of differential gene expression

DESeq2 (Love et al., 2014) was used to determine differential gene expression at room temperature and 15 min and 90 min after a cold shock of three hours among Fast and Slow fly strains. Of the 14,365 protein-coding genes annotated in FlyBase release 1.05, 14,250 (99.2%) had at least one read in at least one of the libraries. For 13,562 genes, at least one read could be mapped in all libraries. Numbers of differentially expressed genes in each category are shown in Table 3.4 for a false discovery rate (FDR) of 5%.

Expression differences before the cold shock.
As a reference for analysis of differential gene expression, RNA was extracted at room temperature from fly strains of both phenotypes. When raised under common garden conditions without being subjected to cold stress, 3.87% of all genes had higher expression in the Fast strains than in the Slow strains and about the same proportion (3.91%) had higher expression in the Slow strains than in the Fast strains (Table 3.4). After multiple testing correction, there was no significant GO enrichment in either of the two categories (Appendix D, Additional file 4: Table S1 and S2).

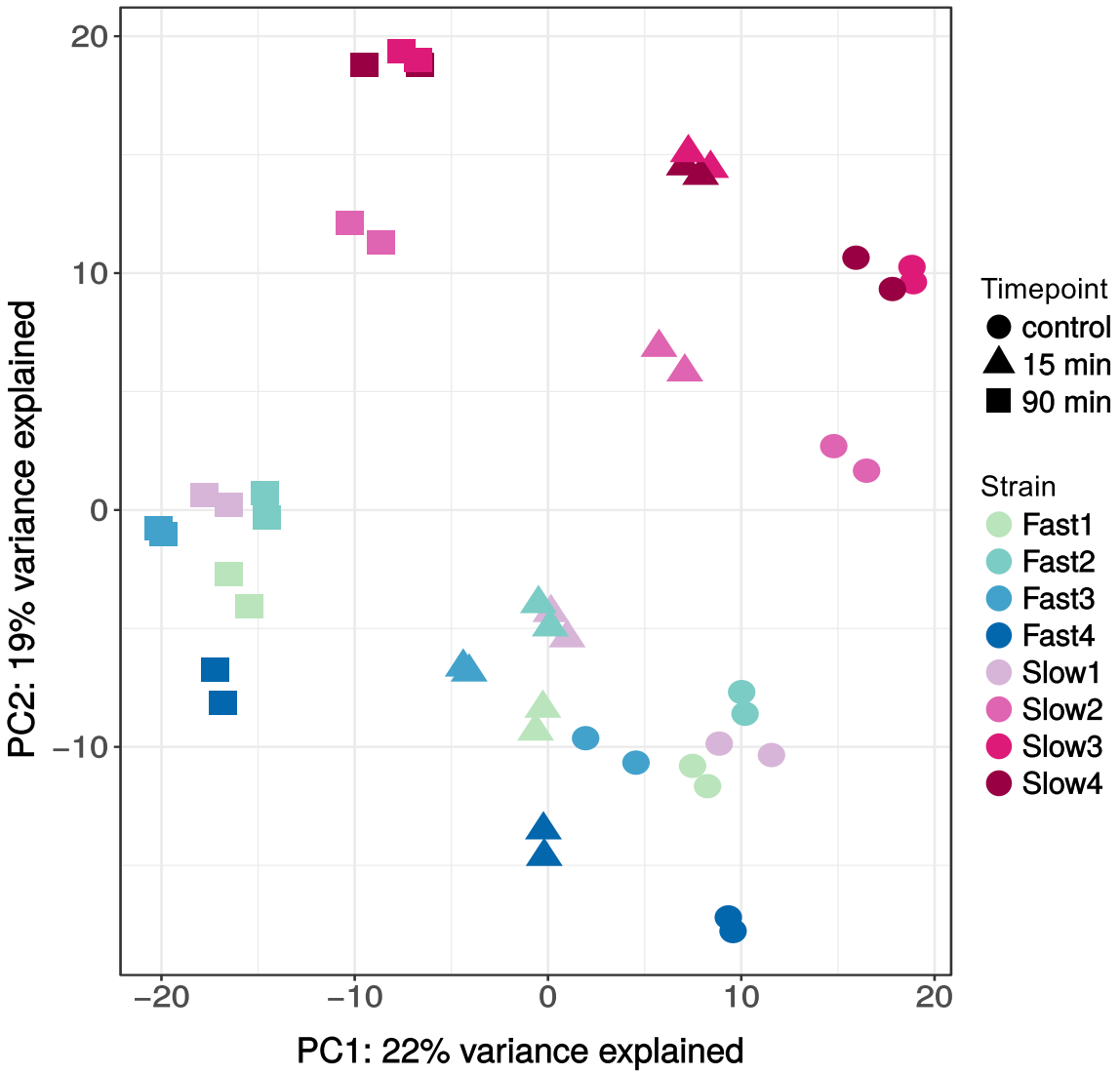


Figure 3.2: Principal Component Analysis (PCA) based on the 500 most variable genes. Reduced to two dimensions, the two biological replicates for each sample clustered tightly together and grouped samples according to timepoint and strain. Figure adapted from (Königer and Grath, 2018).

Table 3.4: Differentially expressed genes in *D. ananassae*

Timepoint	Direction	Fast CCRT	Slow CCRT	Fast vs Slow
Control	up	-	-	552 (3.87%)
	down	-	-	557 (3.91%)
15 min vs control	up	46 (0.32%)	82 (0.67%)	0
	down	11 (0.08%)	96 (0.78%)	0
90 min vs control	up	1096 (8.00%)	1086 (7.80%)	3 (0.02%)
	down	653 (4.80%)	986 (7.10%)	0
90 min vs 15 min	up	1114 (8.00%)	1145 (8.20%)	-
	down	733 (5.20%)	666 (4.80%)	-

Differentially expressed genes (FDR = 5%) in *D. ananassae* in control samples (without cold shock) and at 15 min and 90 min after the cold shock among Fast and Slow fly strains. The percentages were calculated according to the number of genes with non-zero read counts in the respective category (14,250 in total). Table adapted from (Königer and Grath, 2018).

Expression differences in the recovery phase.

At 15 min after the cold shock, there were more genes differentially expressed in the Slow phenotype than in the Fast phenotype (178 genes vs. 57 genes, Table 3.4, Appendix D, Additional file 4). About twice as many genes were upregulated and roughly ten times more genes are downregulated in the Slow phenotype compared to the Fast phenotype. No genes were significantly upregulated in one phenotype and downregulated in the other one. After multiple testing correction, there was no significant GO enrichment in either of the two phenotypes among up- or downregulated genes (Appendix D, Additional file 4: Table S3, S4, S6 and S7).

To identify general characteristics of the response to a cold shock in this population, we investigated the genes that were significantly up- or downregulated compared to the control in both phenotypes. In the set of upregulated genes, 33 genes overlapped between the Slow phenotype and the Fast phenotype. They were enriched in ten molecular function categories related to nucleotide binding (Appendix D, Additional file 4: Table S5). The gene with the highest fold change was *Hsp70* in both phenotypes. This is in line with previous studies on *D. melanogaster* (*e.g.*, (Gerken et al., 2015)). In the set of

downregulated genes, six genes overlapped between the Slow phenotype and the Fast phenotype (Appendix D, Additional file 4: Table S5 and S6). Among them were the orthologs of *odd skipped* (a transcription factor (Coulter et al., 1990)), *granny smith* which has metallopeptidase activity and is involved in proteolysis (Gramates et al., 2017) and *mulet* which is involved in spermatoid development (Castrillon et al., 1993). Among the five genes that were downregulated exclusively in the Fast strains was the ortholog of *Senescence marker protein-30* (*smp-30*). In *D. melanogaster*, *smp-30* transcription levels have been reported to increase following cold acclimation, and the protein is thought to play a role in the cytosolic maintenance of Ca^{2+} levels (Goto, 2000). Overall, we saw a downregulation of *smp-30* at both timepoints after the cold shock in *D. ananassae*, as well as in our reanalysis of the *D. melanogaster* data.

At 90 min after the cold shock, there were 2,072 differentially expressed genes in the Slow phenotype and 1,749 genes differentially expressed in the Fast phenotype (Table 3.4, Appendix D, Additional file 4). No gene was significantly upregulated in one phenotype and downregulated in the other one. In contrast to the 15 min timepoint, the same proportion of genes (about 8%) was upregulated in both phenotypes (Table 3.4). Among all upregulated genes, 888 were shared between the phenotypes, 198 genes were upregulated in the Slow strains only, and 208 genes are upregulated in the Fast strains only. Again, the gene with the highest fold change in expression in both phenotypes was *Hsp70*.

After removing redundant GO terms, the set of genes that were upregulated in both phenotypes was significantly enriched for several categories (Figure 3.3). There were 23 terms that were common between the phenotypes, among which the strongest enrichment was seen in BP terms related to signal reception and response, such as “signaling”, “cell communication” and “response to stimulus” (Appendix D, Additional file 4: Table S9 and S10). Terms that were exclusively enriched in the Fast phenotype included four CC terms related to “membrane” (Appendix D, Additional file 4: Table S9) whereas terms that were exclusively enriched in the Slow phenotype included “cell adhesion” but also “regulation of apoptotic process”. Enrichment in the term “regulation of apoptotic process” was about five-fold (Appendix D, Additional file 4: Table S10) and driven by two genes (*CIAPIN1* and *BI-1*), which are apoptosis inhibitors (Chae et al., 2003; The FlyBase Consortium, 2009) and by two genes (*GF24126* and *Dronc*), which positively regulate apoptosis (The FlyBase Consortium, 2009; Yuan et al., 2011).

Similar to the 15 min timepoint, more genes were downregulated in the Slow strains (7.1%) than in the Fast strains (4.8%) at 90 min after the cold shock (Table 3.4). Among all downregulated genes, 489 were shared between the phenotypes, 163 were only downregulated in the Fast strains and 498 genes were only downregulated in the Slow strains (Appendix D, Additional file 4, Table S9 and S10). Again, after removing redundant GO terms, the downregulated genes of both phenotypes were significantly enriched for several categories, most of which included terms related to metabolism (Figure 3.4). Four of them were commonly enriched in both phenotypes (Appendix D, Additional file 4: Table S12 and S13). Terms that were exclusively enriched in the Fast phenotype included further "metabolic process" terms but also terms related to "biosynthetic process" (Appendix D, Additional file 4: Table S12) whereas terms that were exclusively enriched in the Slow phenotype included terms related to "catabolic process" (Appendix D, Additional file 4: Table S13).

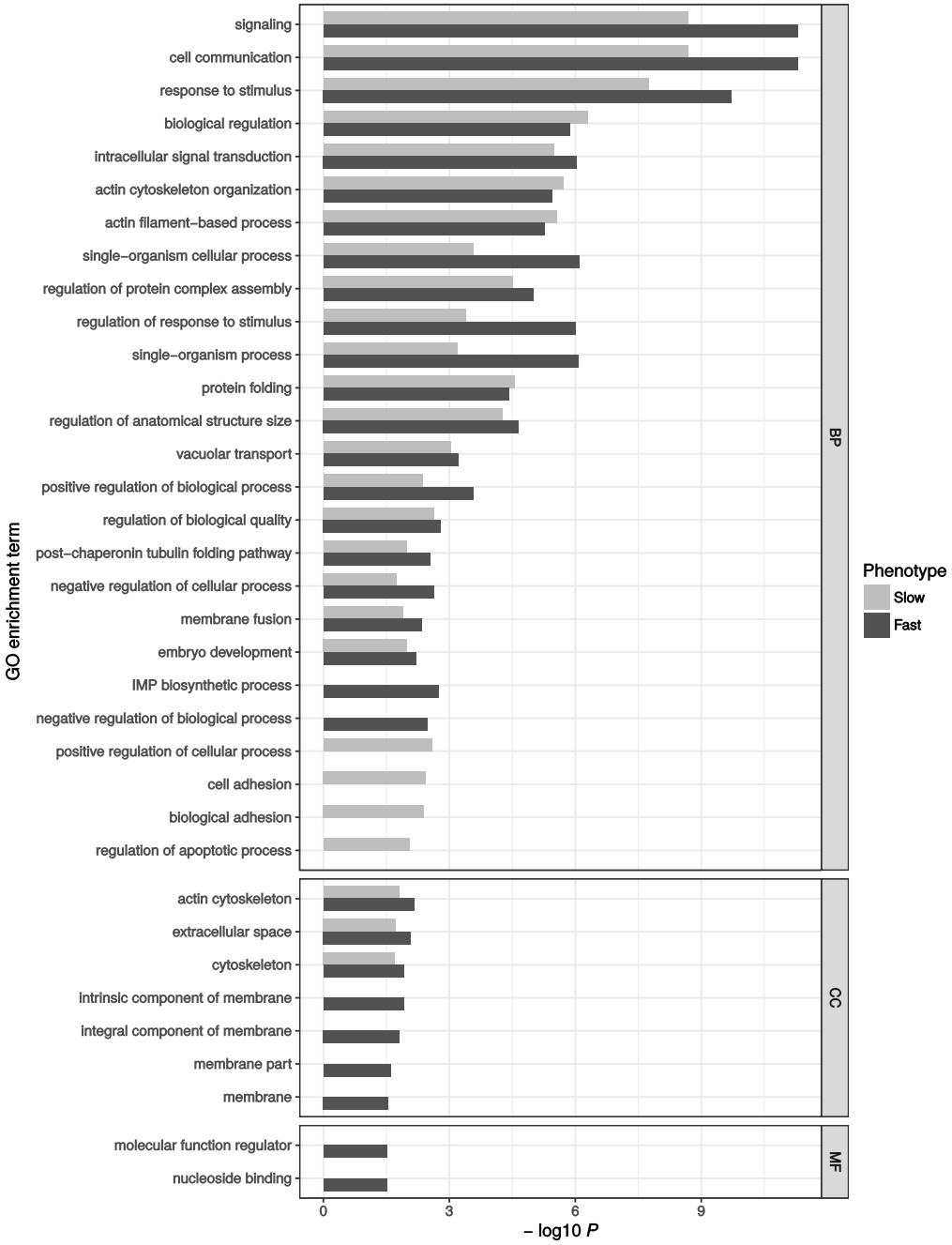


Figure 3.3: Gene Ontology (GO) enrichment analysis for genes that were significantly upregulated ($P < 0.05$) at 90 min after the cold shock in Slow fly strains (light bars) and Fast fly strains (dark bars). GO terms are plotted according to the significance of their enrichment (-log10 P-value after Benjamini-Hochberg correction). Shown are terms in three different categories: biological process (BP), cellular component (CC) and molecular function (MF). Figure adapted from (Königer and Grath, 2018).

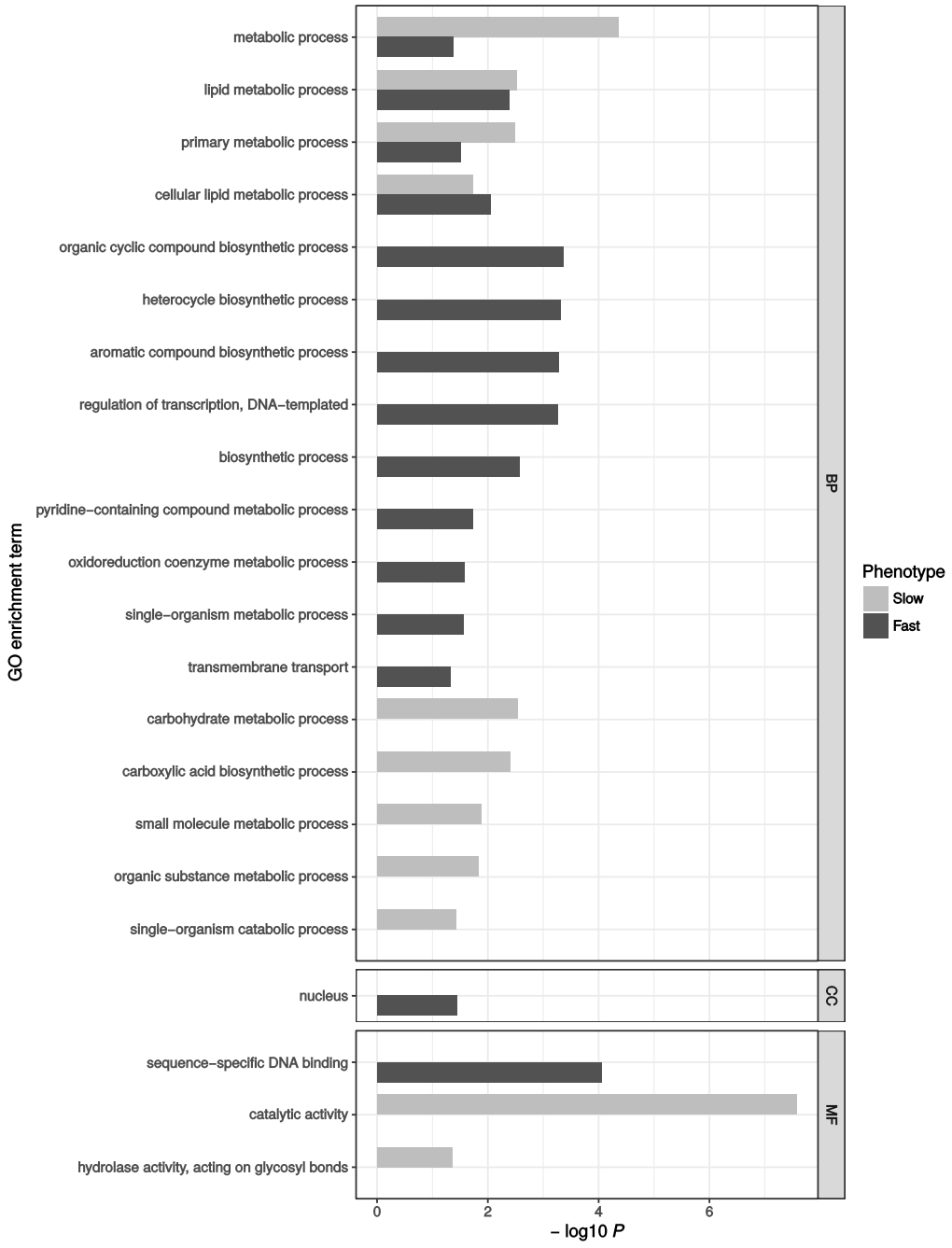


Figure 3.4: Gene Ontology (GO) enrichment analysis for genes that were significantly downregulated ($P < 0.05$) at 90 min after the cold shock in Slow fly strains (light bars) and Fast fly strains (dark bars). GO terms are plotted according to the significance of their enrichment (-log10 P-value after Benjamini-Hochberg correction). Shown are terms in three different categories: biological process (BP), cellular component (CC) and molecular function (MF). Figure adapted from (Königer and Grath, 2018).

3.2.3 Comparison with *Drosophila melanogaster*

For comparison with *D. melanogaster*, we re-analyzed the mapped read counts obtained from von Heckel and colleagues (von Heckel et al., 2016) with the same DESeq2 model as the *D. ananassae* data (Table 3.5). Compared to *D. ananassae*, we saw a higher proportion of genes with a significant interaction between phenotype and timepoint, a higher proportion of genes that responded to the cold shock and about four times more genes that were differentially expressed at room temperature (control samples) between the two phenotypes (Table 3.4 and 3.5). This finding is in line with our expectations since the two phenotypes in *D. melanogaster* were derived from two distinct populations (Africa and Europe) known to show genetic differentiation (Stephan and Li, 2007) and the two phenotypes in *D. ananassae* were from a single population.

Table 3.5: Differentially expressed genes in *D. melanogaster*

Timepoint	Direction	Fast CCRT	Slow CCRT	Fast vs Slow
Control	up	-	-	1080 (8.10%)
	down	-	-	1075 (8.10%)
15 min vs control	up	96 (0.73%)	142 (1.00%)	0
	down	86 (0.65%)	97 (0.72%)	1 (0.01%)
90 min vs control	up	1258 (9.30%)	1257 (9.46%)	10 (0.08%)
	down	1342 (9.90%)	1401 (11.00%)	13 (0.10%)
90 min vs 15 min	up	1060 (7.80%)	1086 (8.00%)	-
	down	582 (4.30%)	837 (6.20%)	-

Differentially expressed genes (FDR = 5%) in *D. melanogaster* in control samples (without cold shock) and 15 min and 90 min after a cold shock of 7 hours at 0 °C among Fast and Slow fly strains. Read counts were obtained from von Heckel and colleagues (von Heckel et al., 2016) and differentially expressed genes were analyzed with Deseq2, percentages were calculated according to the number of genes with non-zero read counts in the respective category (13,285 in total). Table adapted from (Königer and Grath, 2018).

3.2.4 Expression of heat shock proteins

Heat shock proteins (Hsps) were the genes with the strongest transcriptional response in both phenotypic groups and at both timepoints after the cold shock. Overall, we identified *D. ananassae* orthologs of the following ten Hsps as cold-inducible: *Hsp22*, *Hsp23*, *Hsp26*, *Hsp27*, *Hsp40*, *Hsp67Ba*, *Hsp67Bc*, *Hsp68*, *Hsp83* and *Hsp70*. For most of the Hsps, the transcriptional response after the cold shock was stronger in the Slow phenotype than in the Fast phenotype (Figure 3.5A). The only exceptions were *Hsp40*, *Hsp67Bc* and *Hsp83*, which showed a stronger effect in the Fast phenotype. We then looked at the expression of Hsps in *D. melanogaster* and found that expression patterns are very similar among the species, except for *Hsp68* which in *D. ananassae* showed stronger upregulation in the Slow phenotype than in the Fast phenotype, but in *D. melanogaster* showed stronger upregulation in the Fast phenotype (from Europe) than in the Slow phenotype (from Africa) (Figure 3.5B).

3.2.5 Identification of candidate genes

To identify the candidate genes that might underlie the phenotypic difference in CCRT for populations of *D. ananassae*, we applied two different approaches. First, we used an interaction term within our DESeq2 formula to identify genes that respond to the cold shock in a phenotype-specific way. Second, we aimed to identify those genes that could contribute to increased cold tolerance, because they are already expressed at higher levels in the control samples before the cold shock.

With the first approach, we identified three genes with a significant interaction between phenotype and the timepoint 90 min after the cold shock 3.4. As it is possible that the significance is driven by only a few strains or replicates within each of the phenotypes, we visualized the counts in each sample to verify the candidates. One of the three genes (*GF25091*) turned out to be such a false-positive case. The significance seemed to be driven by only four samples with a high number of read counts (Figure A3E). Therefore, we excluded that gene from our list of candidates. Thus, two candidate genes with significant interaction remained: *GF14647* and *GF15058*. The first candidate gene (*GF14647*) was upregulated in the Slow strains at 15 min but downregulated again at 90 min. In the Fast strains, expression levels for this gene were lower than in the Slow strains before the cold shock, but expression increased after the cold shock (Figure A3A). The *D. melanogaster* ortholog of *GF14647* is *CG10621*. In our analysis of the *D.*

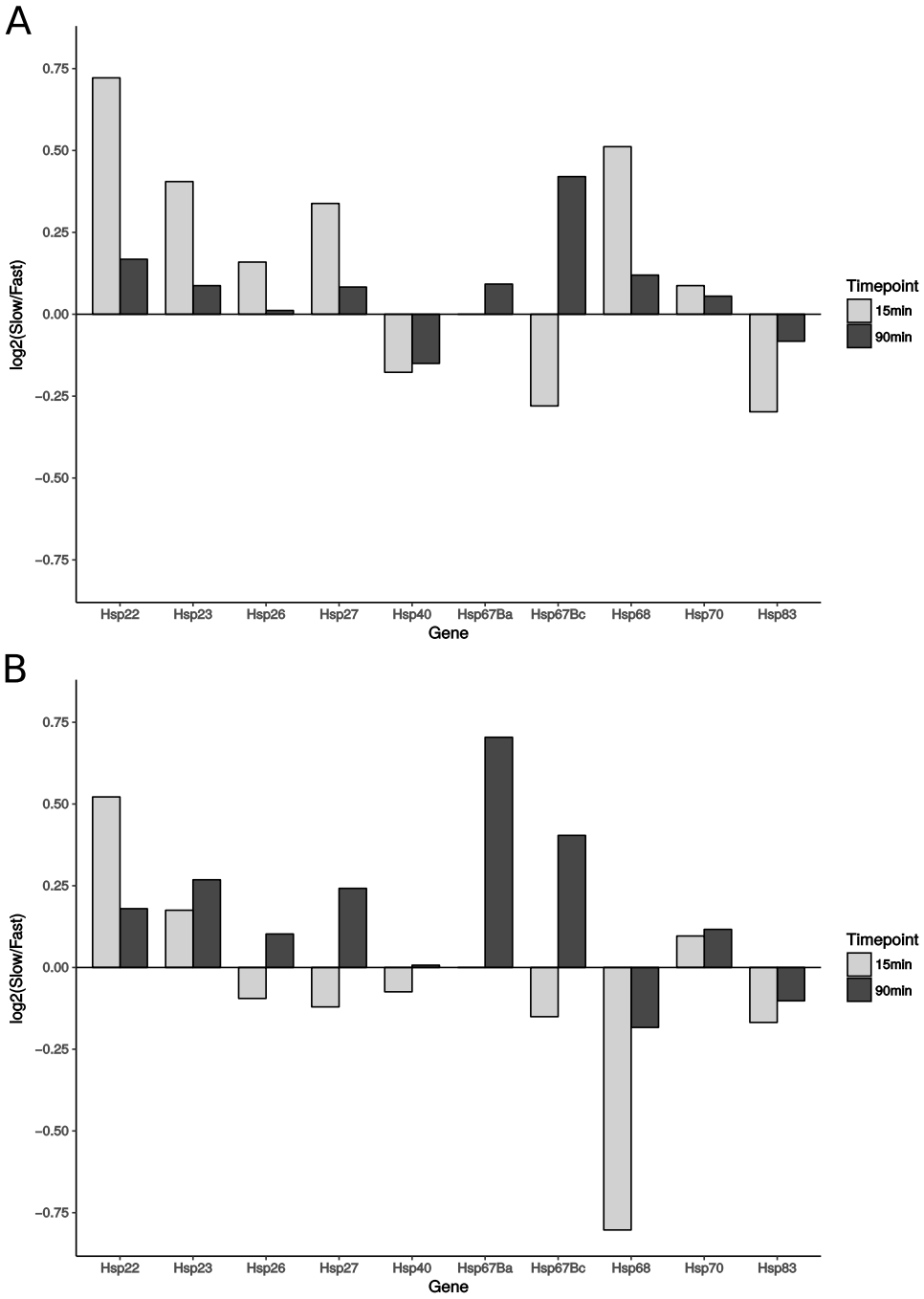


Figure 3.5: Comparison of transcriptional log2 fold changes of Hsps in Slow and Fast fly strains at 15 min (light bars) and 90 min (dark bars) in *D. ananassae* (**A**) and *D. melanogaster* (**B**). Bars represent the Slow:Fast log2 ratios of the log2 fold changes for each gene. A stronger transcriptional response in the Slow strains compared to the Fast strains can be seen as a positive bar, a stronger transcriptional response in the Fast strains compared to the Slow strains can be seen as a negative bar. Log2 fold changes for each gene and sample can be found in Appendix D, Additional file 4. Figure adapted from (Königer and Grath, 2018).

melanogaster data, we saw a different pattern: in both *D. melanogaster* phenotypes (*i.e.* populations), *CG10621* was downregulated after the cold shock. The downregulation was stronger in the Fast (European) strains, in which expression levels in the control samples were higher (Figure A3B). The function of the *CG10621* protein has not been characterized yet, but it contains conserved domains that belong to the "selenocysteine methyltransferase activity" family (Marchler-Bauer et al., 2015), suggesting that it catalyzes the methylation of selenocysteine into Se-methylselenocysteine (MSC). The second candidate gene (*GF15058*) was slightly downregulated at 15 min and upregulated again at 90 min in the Fast strains of *D. ananassae*. In the Slow strains, downregulation at 15 min was stronger and increased at 90 min back to control sample levels (Figure A3C). The *D. melanogaster* ortholog of *GF15058* is *CG10178*. We found *CG10178* to be subsequently downregulated after the cold shock (Figure A3D). The function of the gene was inferred from sequence or structural similarity to be "uridine diphosphate (UDP) glycosyltransferase activity" (Marchler-Bauer et al., 2015).

With the second approach, we analyzed the overlap of genes that were upregulated in the Fast phenotype already before the cold shock with those genes that were upregulated in the Slow phenotype only after the cold shock. We identified 552 genes with higher expression in the Fast phenotype compared to the Slow phenotype in the control samples, 82 genes were upregulated in the Slow phenotype at 15 min. There was an overlap of 13 genes between these two groups. There was no significant GO enrichment in this overlap. Among these genes was the ortholog of *Jun-related antigen* (*Jra*). *Jra* showed a significant interaction in *D. melanogaster* at 90 min (von Heckel et al., 2016). In the late recovery phase (90 min after the cold shock), 1,086 genes were upregulated in the Slow phenotype compared to the Fast phenotype. Among these genes were 100 genes that overlapped with the 552 genes that were upregulated in the Fast phenotype compared to the Slow phenotype at in the control samples. This set of overlapping genes was significantly enriched for seven GO terms (Figure 3.6, Appendix D, Additional file 4: Table S25). The strongest enrichment was seen in the category "actin polymerization or depolymerization", which was driven by six genes: *capping protein alpha* (*CPA*, *GF11927*), *capping protein beta* (*CPB*, *GF20820*), *Actin-related protein 2/3 complex, subunit 3B* (*Arcp3B*, *GF21827*), *Actin-related protein 2/3 complex, subunit 4* (*Arcp4*, *GF14506*), *twinstar* (*GF13484*) and *twinfilin* (*GF16237*).

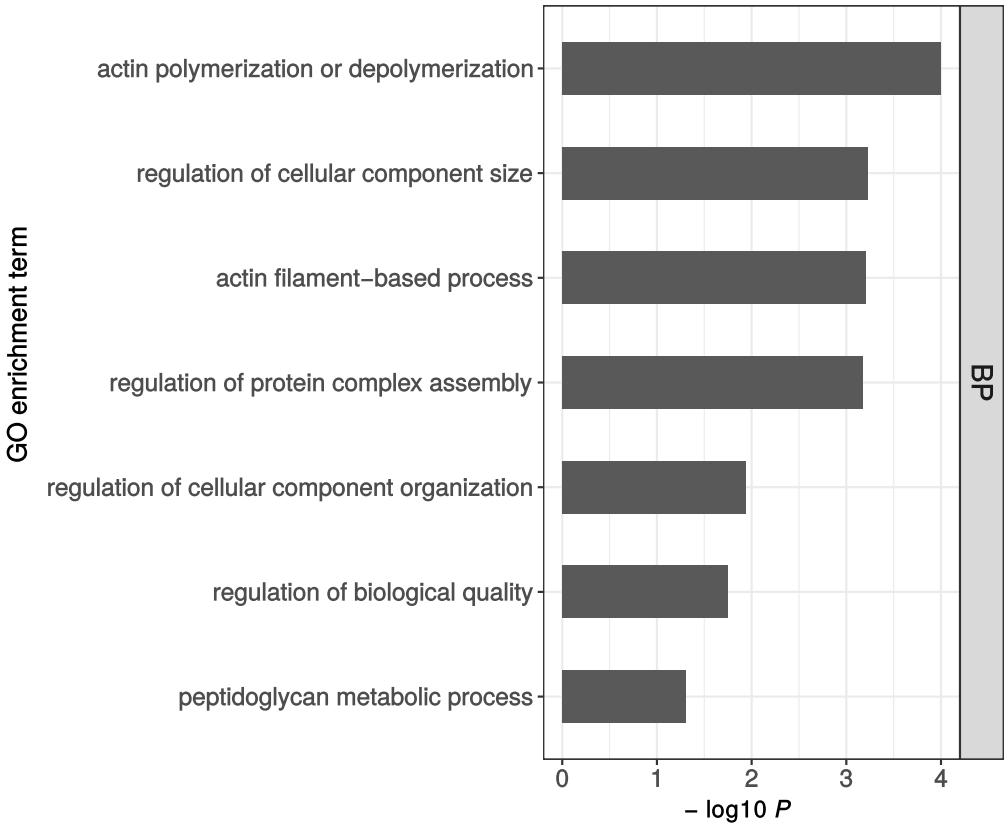


Figure 3.6: Gene Ontology (GO) enrichment in genes that were higher expressed in the Fast strains before the cold shock and also upregulated in the Slow strains at 90 min after the cold shock. GO terms are plotted according to the significance of their enrichment (log10 *P*-value after Benjamini–Hochberg correction). Enrichment was only significant in terms of the category biological process (BP). Figure adapted from (Königer and Grath, 2018).

3.3 Quantitative trait locus mapping

3.3.1 Chill coma recovery time of recombinant inbred advanced intercross lines

CCRT was scored for virgin female flies for each of the 94 RIAILs. The average time needed to recover ranged from 27.60 min for the fastest RIAIL (strain 1) to 83.03 min for the slowest RIAIL (line 94) (Table B2). Three RIAILs were faster than the parental founder Fast4, and 19 RIAILs were slower than the parental founder Slow1 (Figure 3.7).

3.3.2 Sequencing and genetic map

331.867.133 sequence reads were obtained with on average 3.281.450 reads per sample. 0.6% of the total reads (2.074.057) failed the Stacks `process_radtags` program quality check and were excluded from the analysis. In each of the samples, $>$ 94% of all reads mapped to the *D. ananassae* reference genome. After filtering, a total of 1,400 markers and 86 RIAILs were used for genetic map construction (see Table B2). They were partitioned into eight linkage groups (Table 3.6). The total map length was 962.0 cM, with an average marker spacing of 0.7 cM and a maximum marker spacing of 55.5 cM (Figure 3.8). Across all samples, 91.6% of the genotypes were available of which 37.4% were homozygous for the Fast allele (FF), 27.9% were heterozygous (FS), and 34.7% were homozygous for the Slow allele (SS).

Table 3.6: Marker linkage groups

Group	1	2	3	4	5	6	7	8
Markers	986	180	125	64	27	15	2	1

Markers were partitioned based on a LOD score threshold of 8 and a maximum recombination frequency of 0.35, assuming a sequencing error rate of 1%.

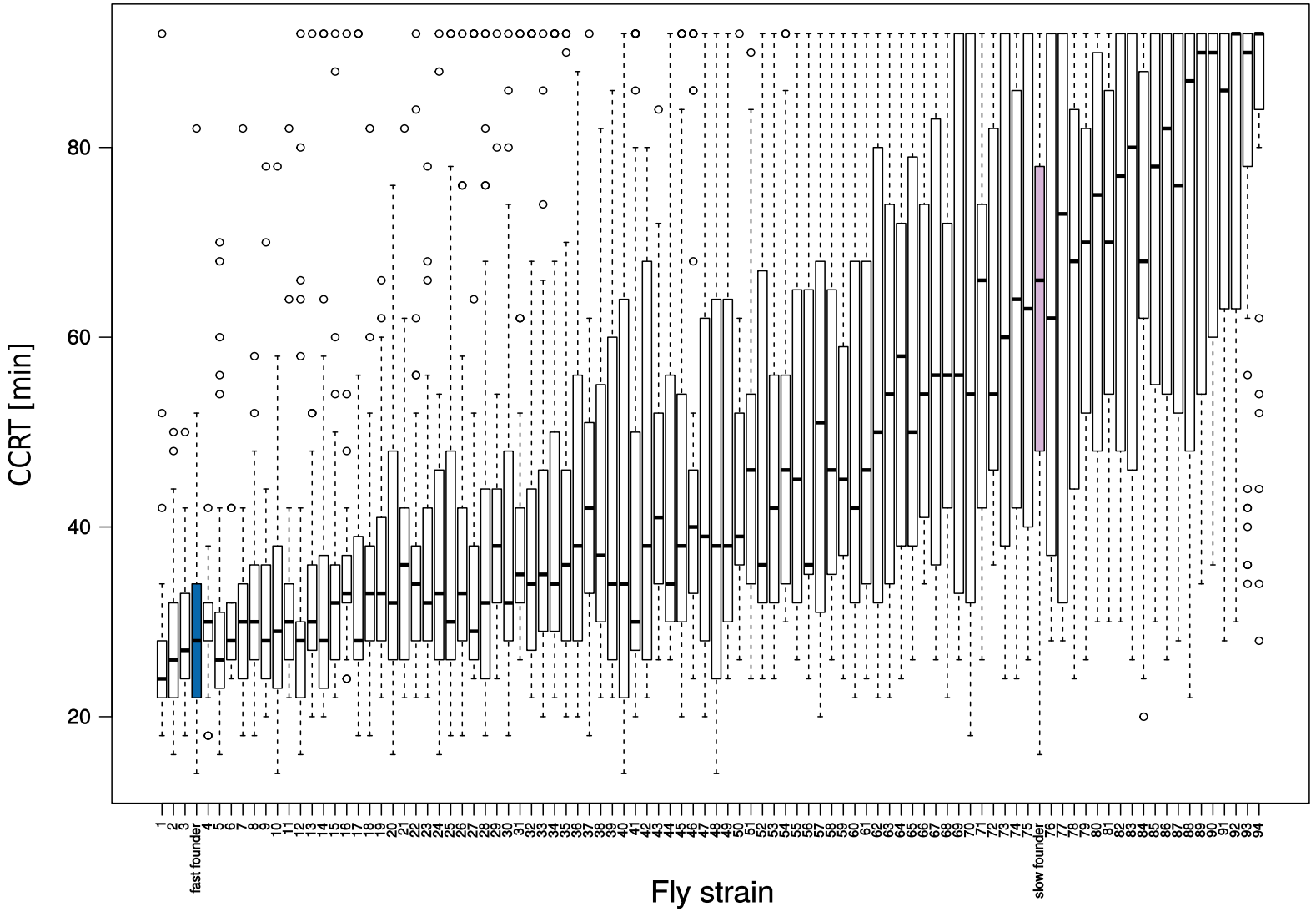


Figure 3.7: Chill Coma Recovery Time (CCRT) of 4 – 6 day old virgin female flies of the 94 Recombinant Inbred Advanced Intercross Lines (RIAILs) is displayed with white bars. CCRT of the two founder strains is displayed in blue (Fast4) and pink (Slow1) (see also Figure 3.1). The RIAILs were numbered in ascending order according to their CCRT (see Table B2).

3.3.3 One- and two-dimensional interval mapping

Interval mapping in the context of a single-QTL model revealed two major areas with LOD peaks which exceeded the permuted 5% significance level (LOD 3.53), one on scaffold 13337 and one on scaffold 13340 (Figure 3.8). The highest peak on scaffold 13337 was at 6.08 cM (LOD 5.80) and the highest peak on scaffold 13340 was at 80.05 cM (LOD 4.08).

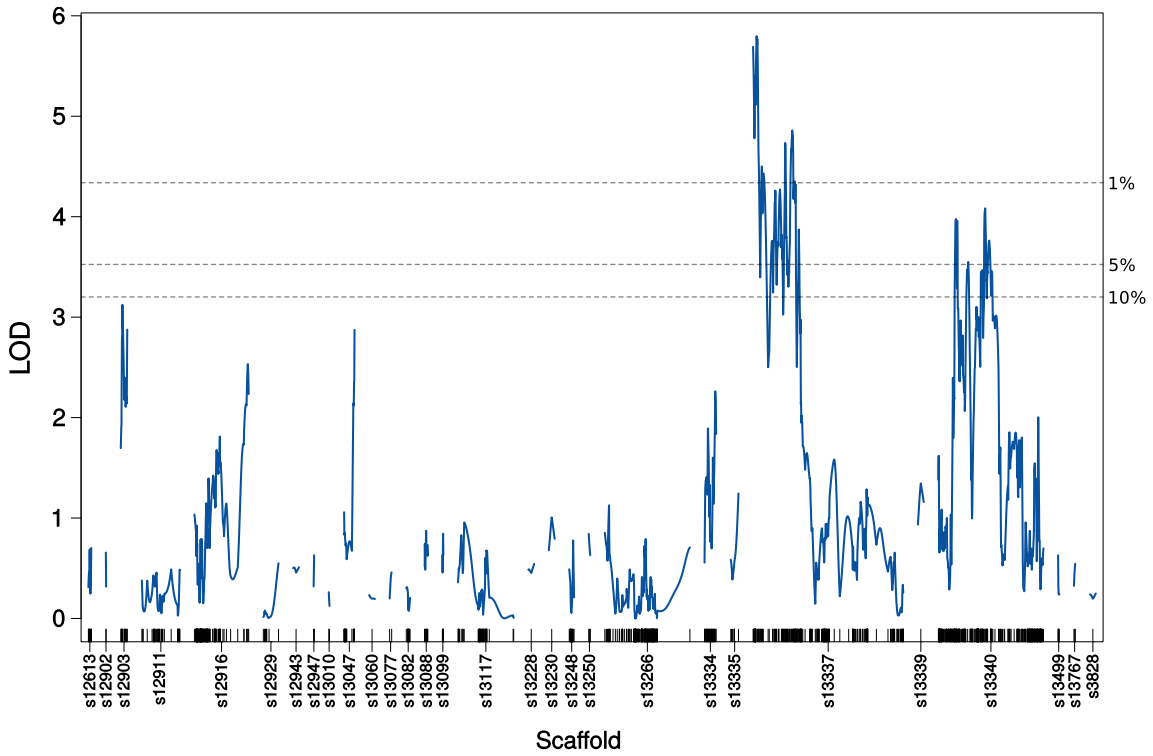


Figure 3.8: LOD-curves obtained with standard interval mapping. Significance thresholds (dotted lines) were calculated with 1,000 genome-wide permutations. The short vertical lines on the X-axis correspond to the marker positions.

The next step was to extend the initial, single-QTL scan with a two-dimensional scan, where we compared two possible models: the full (epistatic) model (H_{f1}) which allowed for the possibility of a second QTL and interactions among QTL versus the additive model (H_{a1}) which allowed for the possibility of a second QTL without interaction. Both, the full and the additive model reached maximum LOD scores at the same positions, 7.08 cM on scaffold 13337 and 30.1 cM on scaffold 13340 and LOD scores are significant considering the permutations (Table 3.7).

In comparison to the single-QTL model, we found supporting evidence for the presence of a second QTL under the additive model (lod.av1 *P*-value = 0.006), but not under the full model (lod.fv1 *P*-value = 0.668). There is no evidence for interaction among the two loci (lod.int *P*-value = 1).

Table 3.7: Results of the two-QTL genome scan

	pos1f [1]	pos2f [1]	lod.full[1]	*P*-value[1]*	lod.fv1[2]	*P*-value[2]*		
s13337:s13340	7.08	30.1	12.6	0	5.77	0.668		
	pos1a[3]	pos2a[3]	lod.add[3]	*P*-value[3]*	lod.av1[4]	*P*-value[4]*	lod.int[5]	*P*-value[5]*
s13337:s13340	7.08	30.1	11.4	0	4.54	0.006	1.24	1

1) QTL positions in [cM], LOD score and *P*-value for the full (epistatic) model versus the null model
2) LOD score and *P*-value for the full (epistatic) model versus the single-QTL model
3) QTL positions in [cM], LOD score and *P*-value for the additive model versus the null model
4) LOD score and *P*-value for the additive model versus the single-QTL model
5) LOD-score and *P*-value of (full model – additive model) = evidence for interaction
∗) *P*-values represent the proportion of permutation replicates with LOD scores $\geq$ the observed

3.3.4 Multiple-QTL mapping

In order to identify possible additional QTL of moderate effect, refine positions, separate linked loci and to estimate QTL effects, we applied a forward selection/backward elimination algorithm with penalized LOD scores and identified a model with three main terms and one interaction term (Table 3.8). The overall fit of the model had a LOD score of 19.26 (see also Figure A4) and explained 64.34% of the phenotypic variance.
In comparison to the one- and two-dimensional genome-scans, there is an additional locus on scaffold 12916 at position 16.7 cM (QTL1) which interacts with one of the previously identified loci, on scaffold 13340 (QTL3) (Table 3.8 and Figure 3.9, 3.10, 3.11).

The results of a drop one term at a time analysis of variance (ANOVA) indicated strong evidence for all three loci and the interaction of QTL1 and QTL3: for each QTL, the model with the QTL of interest at that particular position was compared to the model with the QTL of interest omitted, while all other QTL positions were fixed at their maximum likelihood estimates (Table 3.9, Figure 3.10).

Table 3.8: Summary table for the multiple-QTL model

Model formula: y ∼ QTL1 + QTL2 + QTL3 + QTL1:QTL3							
	df	SS	MS	LOD	%Var	*P*-value(χ^2)	*P*-value(F)
Model	10	11864.778	1186.47777	19.2575	64.34284	9.77e-15	3.11e-13
Error	75	6575.259	87.67012				
Total	85	18440.037					

df = degrees of freedom, SS = sums of squares, MS = mean squares, LOD = relative to the null model, %Var = proportion of variance in the phenotype explained by all terms in the model, *P*-value(χ^2) = based on LOD score following a χ^2-distribution, *P*-value(F) = F-statistic.

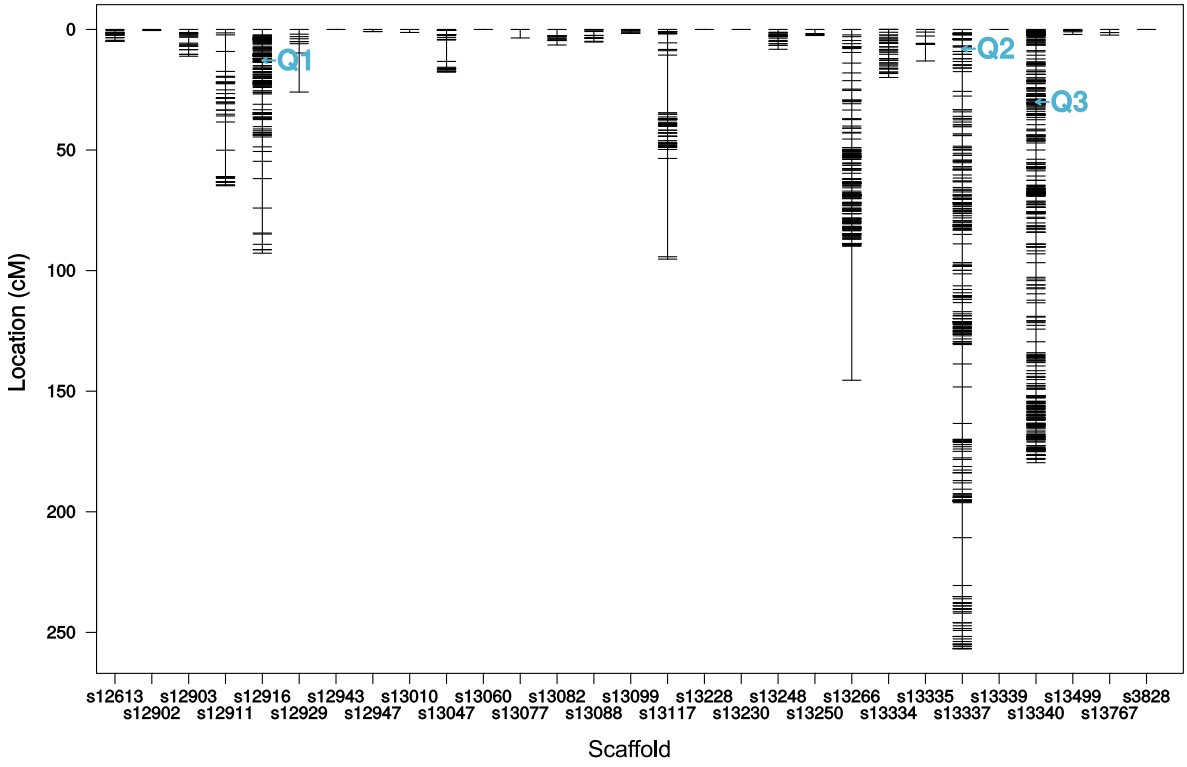


Figure 3.9: Genetic map with QTL positions as refined with the multiple-QTL model. X-axis = genomic scaffolds. Y-axis = genetic distances in centimorgan for markers (short horizontal lines) and positions of the three QTL (Q = QTL).

Table 3.9: Summary table for a drop one term ANOVA

	cytologic position	df	Type III SS	LOD	%Var	F value	*P*-value(χ^2)	*P*-value(F)
QTL1	s12916:16.7	6	3667	8.277	19.89	6.972	0	6.47e-06
QTL2	s13337:0.1	2	4903	10.404	26.59	27.962	0	8.45e-10
QTL3	s13340:30.1	6	5613	11.526	30.44	10.671	0	1.52e-08
QTL1:QTL3	s12916:16.7:s13340:30.1	4	2790	6.606	15.13	7.957	0	2.11e-05

S12916:16.7 = QTL on scaffold 12916 at position 16.7 cM. df = degrees of freedom, SS= sums of squares, MS = mean squares, LOD = relative to the null model, %Var = proportion of variance in the phenotype explained by all terms in the model, *P*-value(χ^2 = based on LOD score following a χ^2-distribution, *P*-value(F) = F-statistic.

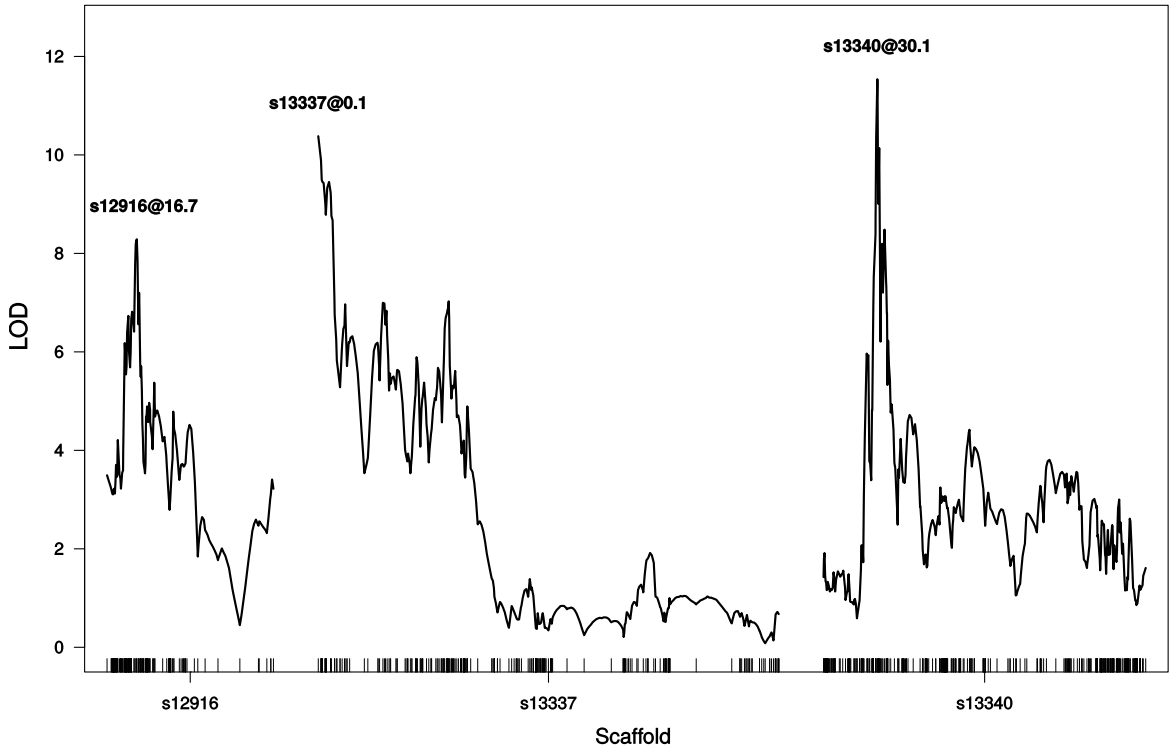


Figure 3.10: Profile LOD scores for the multiple-QTL model. LOD scores are calculated from the drop one at a time ANOVA (Table 3.9). The short vertical lines on the X-axis correspond to the marker positions.

Additive effect estimates ((SS-FF)/2) and deviation from dominance (FS-(FF+SS)/2) are shown in Table 3.10, where F denotes the Fast allele and S denotes the Slow allele. QTL1 on scaffold 12916 was a transgressive QTL as the Fast allele was associated with slower recovery, resulting in a negative effect size (Figure A5A). For QTL2, the estimated additive effect was positive while the estimated dominance effect was negative. RIAILs homozygous for the Fast allele had the fastest recovery, RIAILs homozygous for the Slow allele had the slowest recovery and heterozygote RIAILs had an intermediate phenotype (Figure A5B). The effect estimates for QTL3 went in the same direction as for QTL2. Here, however, the heterozygous phenotype was associated with the fastest recovery (Figure A5C). The more complex relationships of additive and dominance effects for the interaction of QTL1 and QTL3 can be understood best by plotting the interaction of the phenotype and the genotype at both marker positions (Figure 3.11).

Table 3.10: Estimated QTL effects

	cytologic position	est	SE	t
Intercept		48.1178	1.4540	33.093
QTL1	s12916:16.7a	-0.6307	1.2435	-0.507
QTL1	s12916:16.7d	-0.8054	2.7103	-0.297
QTL2	s13337:0.1a	9.2082	1.2598	7.309
QTL2	s13337:0.1d	-2.1719	2.4363	-0.891
QTL3	s13340:30.1a	1.7317	1.3830	1.252
QTL3	s13340:30.1d	-5.2343	2.7464	-1.906
QTL1:QTL3	s12916:16.7a:s13340:30.1a	3.7104	1.4818	2.504
QTL1:QTL3	s12916:16.7d:s13340:30.1a	-6.1786	2.7255	-2.267
QTL1:QTL3	s12916:16.7a:s13340:30.1d	2.5751	2.5695	1.002
QTL1:QTL3	s12916:16.7d:s13340:30.1d	22.5674	5.4525	4.139

a = additive effects, d = dominance deviation, s12916:16.7 = QTL on scaffold 12916 at position 16.7 cM.

As revealed by the interaction plot (Figure 3.11), RIAILs homozygous for the Slow allele at both QTL also had the slowest recovery. The fastest recovery was reached by those RIAILs which were homozygous for the Fast allele at QTL3 but homozygous for the Slow allele at (the transgressive) QTL1. Interestingly, cold tolerance of RIAILs which were heterozygous at QTL1 seemed to be independent of their genotype at QTL3.

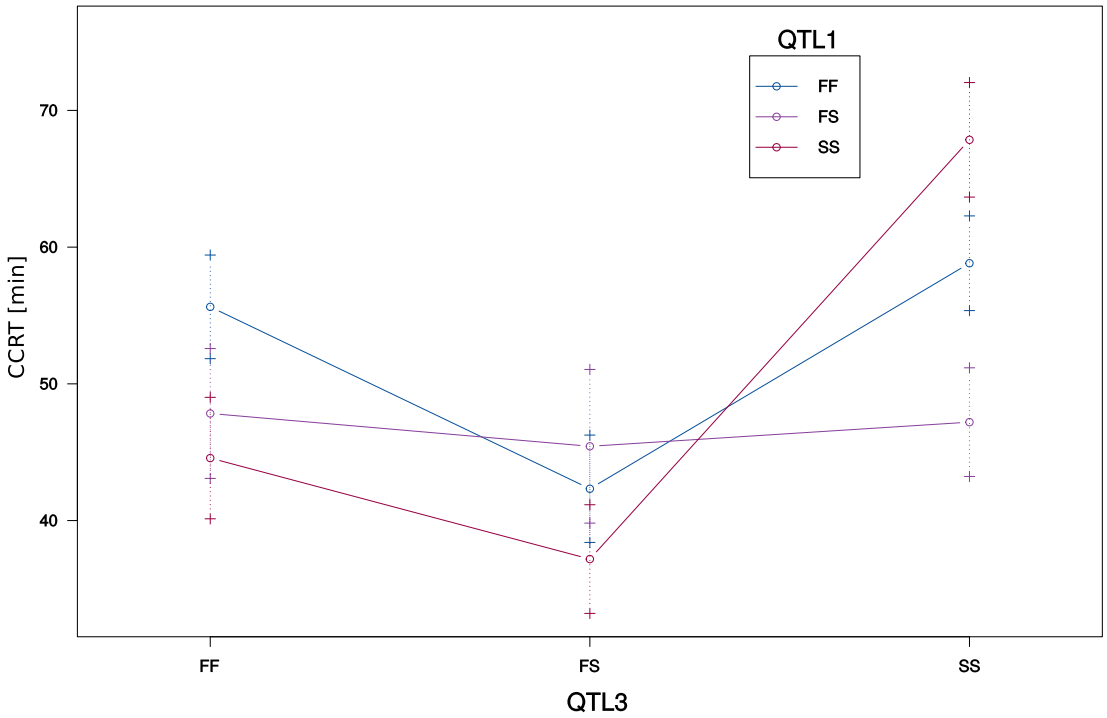


Figure 3.11: Interaction of QTL1 and QTL3. X-axis = genotypes for QTL3. The genotypes for QTL1 are represented by lines in different colors. Error bars are plotted at $\pm$ 1 standard error (SE). F = Fast parental allele, S = Slow parental allele.

3.3.5 Candidate genes within quantitative trait locus confidence intervals

We estimated the 95% Bayesian credible intervals for each of the three QTL (Table 3.11). Taken together, we identified 259 protein-coding genes within these intervals, of which 58 were differentially expressed in response to the cold shock and five genes were previously associated with thermo-tolerance in *Drosophila* (Tables 3.12 and B3, B4 and B5).

Table 3.11: 95% confidence intervals for QTL regions

	Scaffold	Cytologic position [cM]	Cytologic intervals [cM]	Physical interval [bp]	Genes*
QTL1	12916	16.747634	7.103214 – 92.933043	1,514,827 – 2,696,582	110
QTL2	13337	0.083871	0.083871 – 9.233870	83,871 – 226,785	11
QTL3	13340	30.053110	27.51427 – 36.52024	5,542,036 – 6,544,039	138

* Number of protein-coding genes within QTL intervals.

Table 3.12: Cold tolerance candidate genes within QTL regions

	DE genes	Cold tolerance candidate genes
QTL1	29	*GF15058* (Königer and Grath, 2018), *GF14829* (D.mel/*CG10383*, Norry et al., 2008)
QTL2	3	*GF24896* (D.mel/*klarsicht*, MacMillan et al., 2016)
QTL3	26	*MtnA* (D.mel/*MtnA*, Catalán et al., 2016), *GF17132* (D.mel/*CG5246*, von Heckel et al., 2018)

DE = differentially expressed genes in response to the cold shock as identified in chapter 2.3. DE genes are listed in Tables B3, B4 and B5. Cold tolerance candidate genes = genes previously identified as candidates for cold tolerance in *Drosophila*.

QTL1 spanned 1.2 Mb and contained 110 protein-coding genes. Among them, 29 genes were previously identified as differentially expressed in response to a cold shock (Table 3.7) of which 12 genes were upregulated and seven genes were downregulated in both phenotypic groups, Fast and Slow (Table B3). In the Fast phenotype, two genes, *GF15043* (ortholog of *CG31974*) and *GF14846* (ortholog of *bicoid stability factor*) were exclusively upregulated, and two genes, *GF15020* (ortholog of *ABC transporter expressed in trachea*) and *GF14865* (ortholog of *CG11454*) were exclusively downregulated. In the Slow phenotype, five genes were exclusively downregulated, but there were no exclusively upregulated genes. One of the five downregulated genes was *GF15058* (ortholog of *CG10178*), which was one out of two genes with a significant interaction of phenotype and cold shock. The function of *GF15058* is unknown but it is predicted to have UDP-glycosyl-transferase-activity (Marchler-Bauer et al., 2015).

QTL2 spanned 140 kb and contained eleven protein-coding genes (Table B4). There was no enrichment of KEGG pathways or GO terms. However, three of the eleven genes were differentially expressed in response to the cold shock. Two of them,

GF24884 (ortholog of *p130CAS*) and *GF24880* (ortholog of *Phosphoinositide-dependent kinase 1*) were upregulated in both phenotypes after the cold shock and one of them, *GF24896* (ortholog of *klarsicht*) was exclusively upregulated in the Fast phenotype only. *Klarsicht* was previously reported as upregulated in cold-acclimated flies of *D. melanogaster* (MacMillan et al., 2016).

QTL3 spanned 1.0 Mb and contained 138 protein-coding genes which were enriched in one molecular function, "serine-type endopeptidase activity" (GO:0004252) (Table B5 and Table B6). Out of the 138 genes, 26 were previously identified as differentially expressed in response to the cold shock. Among them, nine genes were upregulated, and five genes were downregulated in both phenotypes. In the Fast phenotype, one gene, *GF17809* (ortholog of *Archease*) was exclusively upregulated, and one gene, *GF17856* (ortholog of *Niemann-Pick type C-2c*) was exclusively downregulated. In the Slow phenotype, one gene, *GF17176* (ortholog of *aluminum tubes*) was exclusively upregulated and nine genes were exclusively downregulated. Nine genes drove the enrichment in the GO category "serine-type endopeptidase activity" (Table B6). All of them were located in the downstream region of QTL3 at 6,515,565 – 6,527,729 bp and adjacent to one another (Figure 3.12). Seven of these genes were also differentially expressed in response to the cold shock. Among them was *GF17132*, which was upregulated in both phenotypes and its ortholog in *D. melanogaster* showed a significant interaction of phenotype and cold shock (Table B5, (von Heckel et al., 2016)).

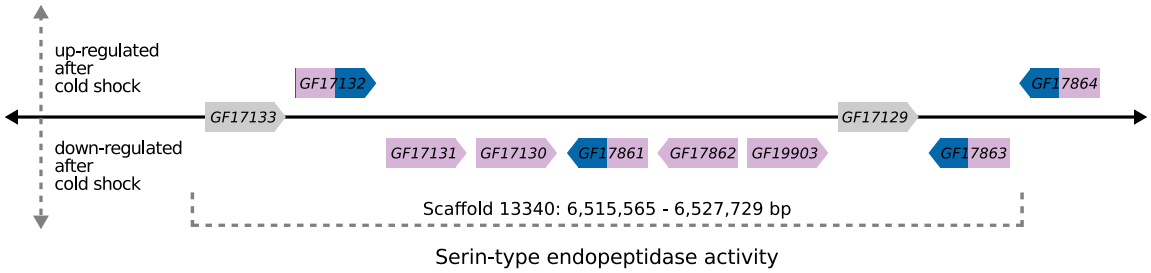


Figure 3.12: Schematic illustration of a genomic region within QTL3 that contains nine genes of the enriched GO category "serine-type endopeptidase activity". Genes were differentially expressed in response to the cold shock in either the Slow phenotype alone (genes shown in pink color) or in both phenotypes, Fast and Slow (genes shown in blue and pink color). Genes that were not differentially expressed are shown in gray. Gene lengths and distances between genes are not drawn to scale.

QTL3 also contained the gene *Metallothionein A* (*MtnA*) which caught our attention because it is involved in metal ion homeostasis and in its *D. melanogaster* ortholog, an indel polymorphism is associated with local adaptation to oxidative stress upon migration out of Sub-Saharan Africa into Europe (Catalán et al., 2016). *MtnA* was downregulated in response to cold in *D. melanogaster* (von Heckel et al., 2016) but not in *D. ananassae*.

3.4 Genome engineering

3.4.1 Cas9 strains

Successful germline transformation of the Cas9 construct into all three fly strains Fast4, Slow1 and KAT1 was achieved with embryo injections carried out by the external facility Rainbow Transgenic Flies, Inc (Camarillo, CA, USA). Transgenesis could be verified by EYFP expression in F1 larvae (Figure 3.13). Transformation efficiency was $< 1\%$ (Table 3.13). The injections that were in parallel carried out by myself for two of the strains following the protocol of (Gompel and Schröder, 2015) were not successful (Table 3.14).

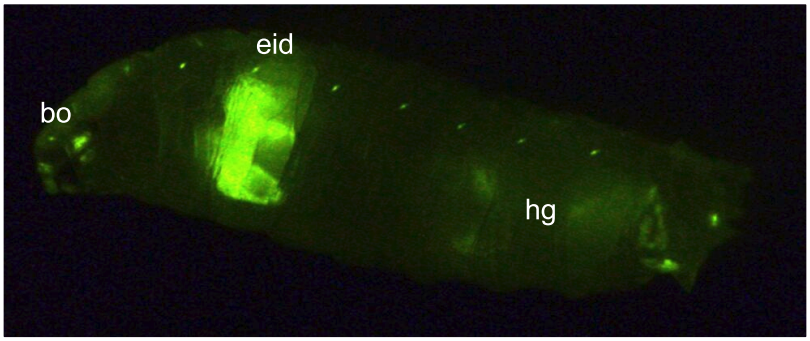


Figure 3.13: Transgenic *D. ananassae* larvae. The photo shows larvae of the F1 generation (see Figure 2.4). Transformation was successful for all three strains, Slow1, Fast4 and KAT1 (Slow1 is shown here as an example). Typical 3xP3 expression of Enhanced Yellow Fluorescent Protein (EYFP) was observed in larval eyes (bo = Bolwig's organ), the central nervous system (eid = eye imaginal discs) and hindgut (hg). The photo was taken by Selina Mußgnug.

Insertions of the Cas9 construct could be mapped to single locations for the strains Fas4-Cas9, Slow1-Cas9 and KAT1-Cas9 using inverse PCR (Table 3.15). The insertion for the strain Slow1-Cas9-2 could not be mapped because there was no PCR product (Figure A8). Fast4-Cas9, Slow1-Cas9 and KAT1-Cas9 could be established as strains homozygous for the insertion as confirmed by genotyping (Figure A9).

Table 3.13: PiggyBac transformation efficiency for insertion of the Cas9 construct achieved by approach 1: Rainbow Transgenic Flies, Inc.

Strain	Injected embryos	Larvae	Pupae	Adults	Egg-to-adult viability	Fertile adults	Transformants	Efficiency
Fast4	310	215	116	88	28.39%	52	1	1.14%
Slow1	300	170	59	52	17.33%	30	2	3.85%
KAT1	220	90	53	46	20.91%	35	1	2.17%
Total	830	475	228	186	22.41%	117	4	2.15%

The efficiency is calculated as the percentage of successful transformants in all hatched adults.

Table 3.14: PiggyBac transformation for attempted insertion of the Cas9 construct with approach 2: following the protocol of Gompel and Schröder, 2015

Strain	Injected embryos	Larvae	Pupae	Adults	Egg-to-adult viability	Fertile adults	Transformants	Efficiency
Fast4	601	15	12	11	1.83%	7	0	0%
KAT1	715	21	14	7	0.98%	5	0	0%
Total	1316	36	26	18	1.37%	12	0	0%

The efficiency is calculated as the percentage of successful transformants in all hatched adults.

Table 3.15: Cas9 insert locations

Strain	Scaffold	Position [bp]	Genomic site
Fast4-Cas9	13230	2,816,409	intronic
Slow1-Cas9	13335	1,951,624	intergenic
KAT1-Cas9	12948	356,455	intronic

Inverse PCR results for Cas9 insertions. Insertion sites were either between two genes (intergenic) or within introns.

3.4.2 CRISPR/Cas9-mediated knock-out

In total, eight different plasmids were constructed to target *GF14647* and *GF15058* for insertion of the dsRed marker via HDR. The plasmids are summarized in Table 3.16. PCR results for the amplification of homology arms and colony PCR products are shown in Figure A10. The plasmids for *GF14647* were generated for later use in future experiments and not tested yet. The plasmids for *GF15058*, however, were injected into all three Cas9 strains. The survival statistics for the attempted germline transformation are given in Table 3.17. The approach was not successful as we did not observe dsRed expression in the F1 generation.

Table 3.16: Summary table of all constructed plasmids for CRISPR/Cas9-mediated homology-directed repair to knock-out *GF14647* and *GF15058*

plasmids for gRNA expression	plasmids for HDR donor templates
pCFD3-dU6:3gRNA-GF14647	pDsRed-attP-GF14647-Fast4
	pDsRed-attP-GF14647-Slow1
	pDsRed-attP-GF14647-KAT1
pCFD3-dU6:3gRNA-GF15058	pdsRed-attP-GF15058-Fast4
	pdsRed-attP-GF15058-Slow1
	pdsRed-attP-GF15058-KAT1

Table 3.17: Efficiency and survival rates for CRISPR/Cas9-mediated homology-directed repair for knock-out of *GF15058*

Strain	Injected embryos	Larvae	Pupae	Adults	Egg-to-adult viability	Fertile adults	Transformants
Fast4-Cas9	540	95	38	28	5.19%	4	0
Slow1-Cas9	280	120	44	52	18.57%	13	0
KAT1-Cas9	510	62	28	20	3.92%	6	0
Total	1330	277	110	100	7.52%	23	0

Injections were carried out by Rainbow Transgenic Flies, Inc. Egg-to-adult viability was calculated as the percentage of surviving adults from the number of injected embryos.

Chapter 4

Discussion

Temperature is one of the major factors that shape habitat characteristics and influence the geographical distribution and abundance of all species on earth. For ectotherms, resilience towards low temperatures often determines the species fate upon climate change or range expansion, as physiological mechanisms for metabolic heat production are absent. The overall aim of this dissertation was to shed light on the molecular basis of cold tolerance in *D. ananassae*, a species of tropical origin that successfully colonized many temperate regions all over the globe (Tobari, 1993). In particular, we chose fly strains from a population in Bangkok, Thailand as models for our investigations, which represent the ancestral species range (Das et al., 2004). Consistent with previous results (Poxleitner, 2010), we found substantial variation in CCRT within this population and identified four strains with fast recovery and four strains with slow recovery (Figure 3.1). With regard to the demographic history of the species, we hypothesized that the causal loci are few in number but have large effects on the phenotype.

4.1 Transcriptome analysis

There was a detectable transcriptional response already early in the recovery phase (15 min after the cold shock). In both phenotypes, the strongest upregulation was seen in Hsps, headed by *Hsp70*. This is consistent with previous results reporting *Hsp70* to be upregulated in response to cold stress in *D. melanogaster* (Sejerkilde et al., 2003; Sinclair et al., 2007; von Heckel et al., 2016) and other insects (Kimura et al., 1998), and emphasizes an early onset of a general stress response in both of the phenotypes. However, there was a remarkable difference in the number of differentially expressed genes, as

about ten times more genes were downregulated in the Slow phenotype. There was no significant GO enrichment at this set of differentially expressed genes, but it included several genes associated with metal ion binding (Appendix D, Additional File 4: Table S6). As metal ion homeostasis is crucial for insect cold hardiness (Storey and Storey, 2012; MacMillan et al., 2015a; von Heckel et al., 2016), a lag in the transcription of such genes may have contributed to a prolonged CCR in the Slow strains.

However, the reduced variation in gene expression in the Fast strains suggests that the adaptive potential arises from the elimination of maladaptive plasticity. While we did not see a general pattern of canalization in gene expression in the Fast strains as it was found in other drosophilids (*D. melanogaster* (von Heckel et al., 2016), *D. serrata* (Liefting et al., 2009)), such a pattern might become more apparent when comparing gene expression in response to cold across multiple populations.

Later in the recovery phase (90 min after the cold shock), about 8% of all of the protein-coding genes were upregulated, and 5–7% were downregulated in both Fast and Slow (Table 3.4), which was slightly less than in *D. melanogaster* (9% and 10–11%, respectively (Table 3.5)). Note that we used the same FDR cutoff (0.05) for both datasets (see also section 2.3.2). In both phenotypes, the transcriptional response at this time-point is characterized by a continued upregulation of Hsps, an upregulation of genes involved in signaling and cell communication, and a downregulation of multiple GO terms that are related to various metabolic processes, which is, again, consistent with previous findings (von Heckel et al., 2016; MacMillan et al., 2016).

One of our central questions was: which genes differ in their response to the cold shock between the two phenotypic groups? The two distinct phenotypes were present as standing genetic variation within a single population, which is why we expected to find only a small number of candidate genes to underlie the difference in CCRT. Indeed, we found only two genes with a significant interaction between phenotype and cold shock: *GF14647* and *GF15058*.

The function of the first gene *GF14647* is inferred from sequence or structural similarity to code for a protein that catalyzes the methylation of selenocysteine into Se-methylselenocysteine (Marchler-Bauer et al., 2015). Se-methylselenocysteine has recently attracted attention in cancer research, as it was shown to provide organ-specific protection against chemotherapy-induced toxicity in rats (Chae et al., 2003). Interestingly, it was found to inhibit apoptosis in human neuroblastoma cells (Wang et al., 2014). We found *GF14647* transcripts to be cold-inducible in both phenotypes of *D. ananassae*,

but its upregulation persisted longer in the Fast strains (Figure A3A). This might be of special interest because we found evidence for the enrichment of apoptosis genes in the Slow strains at 90 min after the cold shock. Thus, we propose *GF14647* as an interesting new candidate gene whose role in cold tolerance should be investigated in more detail.

The function of the second gene that showed a significant interaction (*GF15058*) is predicted to code for a protein with UDP-glycosyltransferase activity (Marchler-Bauer et al., 2015). UDP-glycosyltransferases (UGTs) are membrane-bound enzymes that are located in the endoplasmatic reticulum. They catalyze the addition of a glycosyl group from a uridine triphosphate (UTP) sugar to a small hydrophobic molecule. Therefore, UGTs play an important role in maintaining homeostatic function and detoxification and are known as major members of phase II drug metabolizing enzymes (Bock, 2015). The cold shock led to a downregulation of *GF1505* in the Slow strains, but not in the Fast strains. Thus, we propose *GF1505* as a second new candidate gene for cold tolerance in *D. ananassae*.

4.1.1 Delayed onset of apoptosis may improve recovery from chill coma

Comparing differential gene expression between Fast and Slow strains at the 90 min time point, the most striking pattern was the roughly fivefold enrichment for the GO term "regulation of apoptotic process" in the Slow strains, but not in the Fast strains. This may be a consequence of the weaker cold hardiness in the Slow strains and an effort of the organism to safely eliminate affected cells, thereby protecting the fly from further damage. However, it has been proposed that a cold-induced onset of apoptosis could, in fact, be the cause for a reduced tolerance toward cold stress. In this case, the apoptotic pathway is overactive, leading to an unnecessary death of cells that could have survived (Yi et al., 2007; Yi and Lee, 2011; Teets and Denlinger, 2013). Thus, the reduced expression of genes that positively regulate the apoptotic pathway could provide an advantage that leads to faster CCR.

4.1.2 Increased expression of Hsps is unlikely to improve chill coma recovery time in *Drosophila*

Hsps are found in all organisms and are known for their essential role in maintaining protein homeostasis in various stress situations. They act as chaperones and thereby assist in folding newly synthesized proteins but also assist in the repair or degradation of proteins (Hendrick and Hartl, 1993; Parsell and Lindquist, 1993). Hsps are subdivided into different classes based on their molecular weight. Small Hsps (sHsps) have been suggested to protect cells from oxidative stress, as their expression levels not only increase with heat or cold stress but also with age (Kawasaki et al., 2016; Morrow and Tanguay, 2015).

Further, the knockdowns of two sHsps (*Hsp22*, *Hsp23*) have been linked to prolonged CCR (Colinet et al., 2010*a*,*b*). In our analysis, we found Hsps to be the most strongly upregulated genes in both phenotypes and at both time points in the recovery phase. The transcriptional response after cold shock was stronger in the Slow strains than in the Fast strains for all of the sHsps. We saw this effect for both *D. ananassae* (Figure 3.5A) and *D. melanogaster* (Figure 3.5B). Among the heavier Hsp classes, we saw the same effect for *Hsp70* in both species. In *D. ananassae*, we saw this effect also for most of the remaining Hsps. Transcriptional upregulation of *Hsp70* in response to cold stress has been reported for *D. melanogaster* (Sejerkilde et al., 2003; Sinclair et al., 2007) and other species (Kimura et al., 1998). Indeed, *D. melanogaster Hsp70* null mutant larvae show a decreased survival rate after severe cold exposure (Stětina et al., 2015), and knockdown of *Hsp70* is associated with reduced cold hardiness in the pupae of the flesh fly *Sarcophaga crassipalpis* (Rinehart et al., 2007). However, recent studies in adult flies suggest that the upregulation of *Hsp70* does not affect CCRT, but is rather related to repairing cell damage (Nielsen et al., 2005; Udaka et al., 2010; von Heckel et al., 2016). This is in line with our findings, which show that the upregulation of this gene is stronger in the Slow strains of both species, *D. melanogaster* and *D. ananassae*. Overall, our results suggest that an increase in expression of Hsps does not lead to an increase in cold tolerance, but rather reflects the degree of damage in the flies.

4.1.3 Genes involved in actin polymerization may have a preemptive role in cold tolerance

We also considered transcriptional differences between Fast and Slow strains that were already present before the cold shock. In total, 100 genes overlap between significantly upregulated genes before the cold shock in the Fast strains and significantly upregulated genes at 90 min after the cold shock in the Slow strains (Appendix D, Additional File 4: Table S11). The group of these genes was significantly enriched for "actin polymerization or depolymerization". The enrichment was driven by six genes: *CPA*, *CPB*, *Arpc3B*, *Arpc4*, *twinstar*, and *twinfilin*.

CPA and *CPB* code for two subunits of the actin capping protein heterodimer, which binds at actin filaments and regulates their polymerization by inhibiting the addition and loss of actin (Cooper and Sept, 2008). In *D. melanogaster*, there is a slightly higher expression of *CPB*, but not *CPA*, in the Fast phenotype before the cold shock. The actin-related 2/3 protein complex is responsible for the de novo nucleation of actin filaments (Mullins et al., 1998). Genes related to that complex were recently shown to be upregulated in the coral species *Acropora muricata* when subjected to cold stress (Lee et al., 2018). *Twinfilin* was found to play a crucial role in the maintenance of cytoskeleton actin dynamics in *D. melanogaster* (Kawasaki et al., 2016). *Twinstar* is also involved in actin polymerization. Recently, it was shown that knockdown of *twinstar* is associated with cell death in *D. melanogaster* (Morrow and Tanguay, 2015) In the porcelain crab *Petrolisthes cinctipes*, the homolog of *twinstar* (*cofilin*) was upregulated in response to heat stress (Garland et al., 2015).

The actin cytoskeleton was shown to play a central role in the cold hardiness of various species, including animals, *e.g. D. melanogaster* (Cottam et al., 2006; von Heckel et al., 2016), mosquitos (Kim et al., 2006), crickets (Des Marteaux et al., 2017, 2018), zooplankton (Bowman et al., 2018), silkworm (Chen et al., 2017) and plants, *e.g.* alfalfa (Oervar et al., 2000) and pear (Wu et al., 2012). Our results in *D. ananassae* support the vital role of actin polymerization in cold hardiness. The identification of these six genes involved in actin polymerization and other genes with a significantly higher *a priori* expression in the Fast flies compared to the Slow flies (Appendix D, Additional File 4: Table S11) suggests that genes in this set may have a preemptive role. Increased expression levels in these genes could have a direct effect on CCRT in this species.

4.2 Quantitative trait locus mapping

We created a panel of recombinant inbred advanced intercross lines (RIAILs) to map QTL that underlie the difference in CCRT between the most cold-tolerant and the most cold-sensitive strain of the Bangkok population, Fast4 and Slow1. The CCRT phenotypes of the RIAILs were distributed on a continuum (Figure 3.7), which was a first important indicator that we were looking at more than one dominant causal allele. Furthermore, for 19 of the RIAILs, CCRT exceeded the measurements for the parental strain Slow1, indicating that there was interaction among the parental alleles or loci in the recombinant genotypes of the mapping population. Genetic distances were estimated *de novo* from ddRAD-markers. The application of reduced representation libraries for genome-wide marker development is one of the more recent advances in QTL mapping studies (Peterson et al., 2012). The resulting reduction of genome complexity has two major benefits: first, it is more cost-effective than whole-genome sequencing of individual samples, allowing for a larger number of samples to be analyzed. Second, it is more accurate than whole-genome sequencing of pooled samples (Pool-sequencing) (Cutler and Jensen, 2010; Catchen et al., 2017).

However, RAD-based approaches come at the cost of marker density, especially in crossing designs with low genetic differentiation between the founder strains and low levels of linkage disequilibrium (Futschik and Schlötterer, 2010). Thus, to increase the mapping resolution and to expand the genetic map, we generated a mapping population in which five generations of intercrosses allowed for a sufficient number of crossover events (Pollard, 2012). Subsequently, we used stringent cutoffs for potential sequencing errors and distorted loci (see Chapter 2.4.3). This step certainly increased the robustness of the identified loci, but came at the cost of chromosomal coverage, as many smaller genomic scaffolds were excluded from the analysis at this step. It is therefore possible that our results do not cover all causal loci.

4.2.1 Three genomic regions influence chill coma recovery time

We identified two major loci, QTL2 on scaffold 13337 and QTL3 on scaffold 13340 with standard interval mapping, for which we used the EM algorithm to estimate conditional genotype probabilities and established significance thresholds based on genome-wide permutations to record the presence of a QTL. Such hypothesis-testing approaches are generally a solid first step in QTL mapping and are considered robust in data sets with

a large number of densely spaced markers (Broman and Sen, 2009). However, standard interval mapping analyzes loci only individually, one at a time, and is therefore not well suited to resolve complex genetic architectures of more than two contributing loci. The fit and exploration of multiple-QTL models overcomes this issue. The available genetic information is included collectively by treating markers as cofactors, thus increasing power to detect QTL and their interactions, separate linked QTL and refine QTL positions (Jansen and Stam, 1994; Arends et al., 2010). This is exactly what we observed in our data. A third locus, QTL1 on scaffold 12916 was identified, that interacted with QTL3.

4.2.2 Candidate genes within quantitative trait locus confidence intervals

We estimated 95% confidence intervals for each of the three identified QTL. It needs to be noted that, in general, these intervals should be considered as support regions rather than absolute boundaries (Broman and Sen, 2009). Further, the causal genetic variants may be located anywhere within these intervals. However, to narrow down the list of candidates, we screened each interval for genes that were identified as differentially expressed in response to the cold shock in the transcriptome analysis and for orthologs of genes in *D. melanogaster* that were previously associated with cold tolerance. Altogether, the QTL intervals contained 58 differentially expressed genes, five orthologs of thermotolerance candidate genes in *D. melanogaster*, and one significantly enriched GO term (Table 3.12 and Figure 3.12).

QTL2 explained the biggest proportion of phenotypic variation. It was clearly additive with intermediate effects of the two parental alleles (Figure A5B and Table 3.10). The support interval for QTL2 was rather small and contained only 11 protein-coding genes, three of which were also differentially expressed. One of them, *GF24896*, is the ortholog of the *D. melanogaster klarsicht* gene. *Klarsicht* is upregulated with cold-acclimation in *D. melanogaster* (MacMillan et al., 2016) and we found *GF24896* upregulated at 90 min after the cold shock in the Fast strains but not in the Slow strains. *Klarsicht* is expressed in a wide range of tissues, where it interacts with microtubules and promotes evenly spaced positioning of nuclei. Knock-out of *klarsicht* in muscle cells impairs locomotion and flight (Elhanany-Tamir et al., 2012) – and so does cold exposure. Thus, the functional analysis of *klarsicht* in future experiments may not only help

to better understand its specific contribution to cold tolerance but also give more insight on the connection between cold hardiness and impaired locomotion during chill coma.

At QTL3, heterozygous individuals were slightly faster than the ones homozygous for the Fast4 allele, and we identified a significant interaction with QTL1. QTL3 contained 26 differentially expressed genes, some of which belonged to a cluster that drove enrichment in the GO term serine-type endopeptidase activity (Figure 3.12). Serine peptidases and are involved in proteolysis, *i.e.*, they catalyze the hydrolysis of peptide bonds (Ross et al., 2003; Attrill et al., 2016). Among those genes, *GF17132* caught our attention because it was differentially expressed in response to the cold shock in *D. ananassae* and in addition, showed a significant interaction of phenotype and cold shock in *D. melanogaster* (von Heckel et al., 2016). QTL3 also contained the gene *Metallothionein A* (*MtnA*). Metallothioneins promote resistance to oxidative stress. They bind heavy metals and neutralize reactive oxygen and nitrogen species (Ruttkay-Nedecky et al., 2013). Exposure to cold leads to an increased abundance of free radicals, thereby inducing oxidative stress (Williams et al., 2014). In *D. melanogaster*, a 49 bp deletion in the 3'UTR of *MtnA* is associated with its transcriptional upregulation and with increased tolerance to oxidative stress (Catalán et al., 2016). The frequency of this polymorphism in natural populations follows latitudinal clines, suggesting that upregulation of *MtnA* is favored in temperate environments (Ramnarine et al., 2019). *MtnA* was downregulated after the cold shock in both phenotypes in our re-analysis of *D. melanogaster* (von Heckel et al., 2016), but not differentially expressed in *D. ananassae*. However, a direct link between cold stress and oxidative stress is yet to be established in drosophilids (Plantampa et al., 2016). Thus, a functional analysis of *MtnA* would be interesting in both species.

The QTL interval with the most interesting candidate genes was QTL1: one of the two genes of interest was *GF15058*, which was also one of two genes that showed a significant interaction of phenotype and cold shock in the transcriptome analysis (see chapter 3.2.5), corroborating its status as candidate for cold tolerance in *D. ananassae*. The function of *GF15058* was inferred to be uridine diphosphate (UDP) glycosyltransferase activity. UDP-glycosyltransferases (UGTs) are membrane-bound enzymes that are located in the endoplasmatic reticulum and catalyze the addition of a glycosyl group from a uridine triphosphate (UTP) sugar to a small hydrophobic molecule. Therefore, UGTs play an essential role in maintaining homeostatic function and detoxification and are known as major members of phase II drug metabolizing enzymes (Bock, 2015). The

cold shock led to a downregulation of *GF15058* in the Slow strains but not in the Fast strains. However, the Fast genotype at QTL1 is transgressive, *i.e.*, it increases CCRT. Thus, if *GF15058* was indeed one of the causal factors, our results suggest that keeping transcript abundance at a constant level after the cold shock is so costly for the organism that it slows down recovery.

The second gene of interest in QTL1 was *GF14829*. *GF14829* was the only gene in all three QTL intervals for which an ortholog was mapped to a thermotolerance QTL in *D. melanogaster* (Norry et al., 2008). Norry and colleagues mapped QTL for CCRT and knock-down resistance to high temperature (KRHT) in recombinant inbred lines that were derived from *D. melanogaster* strains from Australia and Denmark. Interestingly, the ortholog of *GF14829*, *CG10383* was located in a KRHT QTL – not in one of the CCRT QTL – suggesting a potential trade-off. *CG10383* is involved in the regulation of glycosylphosphatidylinositol metabolism and influences longevity and sleep (Donggi et al., 2012). *GF14829* was upregulated in response to the cold shock in both, Fast and Slow strains.

4.3 Genome engineering

In this thesis, microinjections into the germline of *D. ananassae* were carried out using two different protocols and two different systems for transgenesis, PiggyBac and CRISPR/Cas9. Thus, different results were realized:
The first step was to transform the germline-specific Cas9 construct into the wild-type genetic background, which was successfully achieved for all three fly strains. Homozygous Cas9 strains with single, stable insertions could be established for Fast4, Slow1 and KAT1. These strains offer the possibility for future loss-of-function studies, not only in both cold-tolerance phenotypes of the Bangkok population but also in the derived Kathmandu population. Transformants were obtained from approach 1 only, *i.e.* from embryos that were processed at the external facility, and transformation efficiency was low for all three strains ($< 4\%$) when compared to previous reports from other species (Handler, 2002; Gompel and Schröder, 2015).

However, such comparisons must be interpreted with caution, as transformation efficiency depends considerably on the size of the insert. The Cas9 insert was with 9.4 kb relatively large (even though transposition of 100 kb fragments have been reported for mouse cells (Li et al., 2011)). Other important factors that influence the success rate relate to the handling of the embryos during and after the injection. This becomes obvious with a comparison of the egg-to-adult viability between approach 1 and 2 (*i.e.*, injections that were carried out by myself). While 22.41% of all injected embryos developed into adults with approach 1, only 1.37% did with approach 2. The data also shows that most embryos died before the larval stage (Table 3.14), suggesting that they were killed during the injection process. In the end, obtaining transformants from only 12 fertile crosses was unlikely, assuming similar efficiencies for both approaches.

The established Cas9 strains were then used in the second part of the project for a pilot loss-of-function study in *D. ananassae* for the gene *GF15058*, which we identified as candidate from both the transcriptome analysis and the QTL mapping. Based on the previous results, we decided to carry out microinjections for the CRISPR/Cas9-mediated knock out with approach 1 (the external facility). This time, however, no transformants were observed.

There are two possible scenarios why we did not see dsRed expression: first, the transformation of the marker may have failed because there was no homology-directed repair (HDR). HDR could have failed due to the low efficiency of the pathway itself. The egg-to-adult viability was low and additionally, the proportion of fertile adults among

those that survived was low as well. Thus, only 23 crosses produced F1 offspring (Table 3.17). Previous experiments with *D. melanogaster*, in which gRNA and donor plasmids were injected into transgenic Cas9 strains reported an HDR efficiency of 16% (Gratz et al., 2014). Thus, it would be reasonable to expect a positive result, assuming the same success rate. As stated earlier, however, comparisons across experiments are difficult. For example, even when targeting the same gene, cleavage efficiencies can vary greatly between two gRNAs (Bier et al., 2018).

Second, knock-out of the candidate gene may have been lethal. Presumably, this would affect not only successful transformations but also the alternative, more frequently occurring NHEJ pathway or potential off-target effects. Indeed, this could explain why we obtained only 23 fertile adults from 1,330 injected embryos. Taken together, the low egg-to-adult viability and the low fertility of the survivors impaired the attempted knock-out and point at potential cytotoxic effects, induced by the injections. It is not possible to determine the underlying cause from our data. Therefore, experiments on a much larger scale are necessary.

Future experiments should also aim at targeting a second genomic site. The results will allow for a direct comparison and thereby help to locate and rectify potential problems in the general workflow. Further, sequencing the target site in offspring from all fertile F1 crosses will allow to determine if there was a double strand break that got repaired by the alternative NHEJ pathway. The gRNA and donor plasmids that were constructed but not injected yet for the knock-out of *GF14647* should be of great use for this purpose.

4.4 Conclusion and perspectives

This thesis investigated the molecular basis of cold tolerance in *D. ananassae*, using fly strains from a population within the ancestral species range in Bangkok, Thailand as a model system. The presented evidence extends our knowledge of how ectothermic organisms cope with cold stress, and highlights candidate genes and genomic regions with adaptive potential in *D. ananassae*. The successful PiggyBac transposon-mediated germline transformation provides the first proof-of-principle for genome editing in *D.ananassae*.

In summary, I identified a concrete list of 12 cold tolerance candidate genes (Table 4.1) which can be partitioned into three groups:
The first group comprises genes that have not been associated with cold tolerance in other species before: *GF15058*, *GF14647* and *GF14829*.
The second group comprises genes which are assigned to biological functions that relate to the actin cytoskeleton. The actin cytoskeleton is crucial for insect cold hardiness (Kim et al., 2006; Des Marteaux et al., 2018) and the presented results suggest that differential regulation of *CPA*, *CPB*, *Arcp3B*, *Arcp4*, *twinstar* and *twinfilin* plays a central role during recovery from cold exposure in *D. ananassae*.
The third group comprises genes that were located in QTL intervals and have been associated with thermotolerance in *D. melanogaster*: *klarsicht*, *MtnA* and *GF17132*.
Additionally, the following conclusions can be drawn from the present results:

- A small number of genes and genomic regions underlies the difference in cold tolerance among the Bangkok strains: Examination of transcript levels before and after a cold shock revealed that only two genes showed a significant interaction of phenotype and cold shock. Moreover, variation at only three genomic regions explained as much as 64% of the phenotypic variation between the most cold-tolerant and the most cold-sensitive strain.
- Most strikingly, one of the two genes that showed a significant interaction of phenotype and cold shock, *GF15058* was located within a genomic region that was identified as causal (QTL1). Thus, *GF15058* is the most promising candidate gene for cold tolerance in *D. ananassae*.
- In direct comparison to *D. melanogaster*, the most salient overlap is seen in the expression of heat shock proteins (hsps). In both species, stronger upregulation of

hsps is associated with increased cold-sensitivity (*i.e.*, slower CCRT). This finding suggests that hsps play a crucial in repairing cold-induced damages but do not prevent them.

- The PiggyBac transposable element system is an effective tool to induce heritable genome alterations in *D. ananassae*.

Table 4.1: List of candidate genes for cold tolerance in *Drosophila ananassae*

Candidate gene	Ortholog	Function	Experiment
GF15058	*CG10178*	UDP-glucosyltransferase [a]	DE + QTL
GF14647	*CG10621*	Homocysteine-binding [a]	DE
GF14829	*CG10383*	Determination of adult lifespan and sleep [b]	QTL
GF11927	*CPA*	Actin cytoskeleton organization [b]	DE
GF20820	*CPB*	Actin cytoskeleton organization [b]	DE
GF21827	*Arpc3B*	Actin binding [a]	DE
GF14506	*Arpc4*	Cortical actin cytoskeleton organization [b]	DE
GF13484	*twinstar*	Actin binding [b]	DE
GF16237	*twinfilin*	Actin binding [b]	DE
MtnA	*MtnA*	Metal ion homeostasis [b]	QTL
GF24896	*klarsicht*	Flight and locomotion [b]	QTL
GF17132	*CG5246*	Serine-type endopeptidase activity [a]	QTL

Candidate genes were identified in the transcriptome analysis (DE = differentially expressed, see chapter 3.2) and in the QTL mapping (QTL, see chapter 3.3). Evidence for gene function was inferred from either electronic annotation (**a**) or a mutant phenotype (**b**) in *D. melanogaster* orthologs on FlyBase (Attrill et al., 2016).

In order to turn candidate genes into causal genes, functional evidence is required. To begin with, loss-of-function experiments will help to establish a link between each candidate gene and CCRT. For this purpose, the germline transformation protocols described in chapters 2.5.3 and 2.6 can be resumed and eventually be adapted for the remaining candidate genes. Furthermore, detailed sequence analysis of each candidate gene, including its upstream and downstream regulatory regions may also reveal specific variants that distinguish Fast and Slow strains. Eventually, CRISPR/Cas9-mediated homology-directed repair could be used to insert Fast alleles into the Slow genomic background and *vice versa*. Expanding these analyses to additional, derived populations is a crucial next step in order to understand which genes were under selection upon range expansion to temperate regions. For this purpose, it would be of great help to sample fly strains from additional populations.

My research on the *D. ananassae* Bangkok strains fills a gap in the current literature on the molecular basis of cold tolerance in non-model organisms. Moreover, the generated Cas9 strains have set the groundwork for future experiments that require precise genomic modifications, including loss-of-function analyses of the identified candidate genes.

Appendix A

Additional Figures

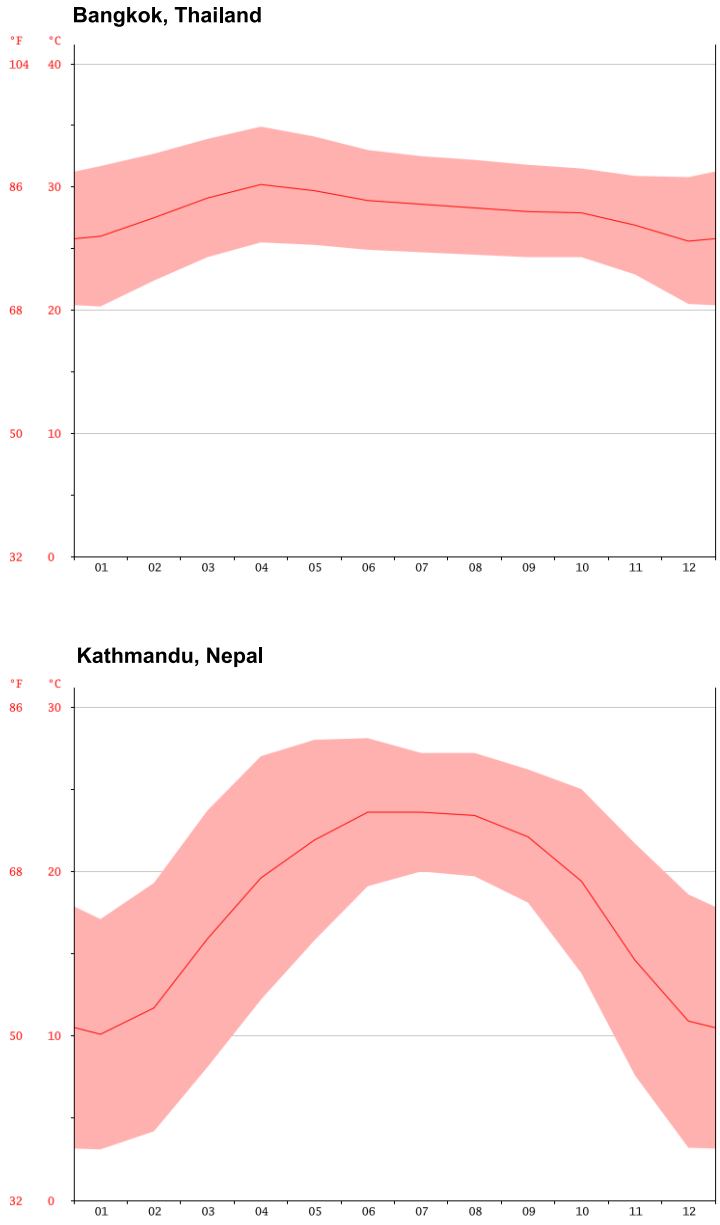


Figure A1: Climate diagrams for Bangkok, Thailand and Kathmandu, Nepal. The average annual temperature in Bangkok ranges from 30°C in April to 25°C in December. The average annual temperature in Kathmandu ranges from 28°C in June to 10°C in January. Data and diagrams were taken from `https://en.climate-data.org/asia/thailand/bangkok/bangkok-6313/` and `https://en.climate-data.org/asia/nepal/central-development-region/kathmandu-1137/`.

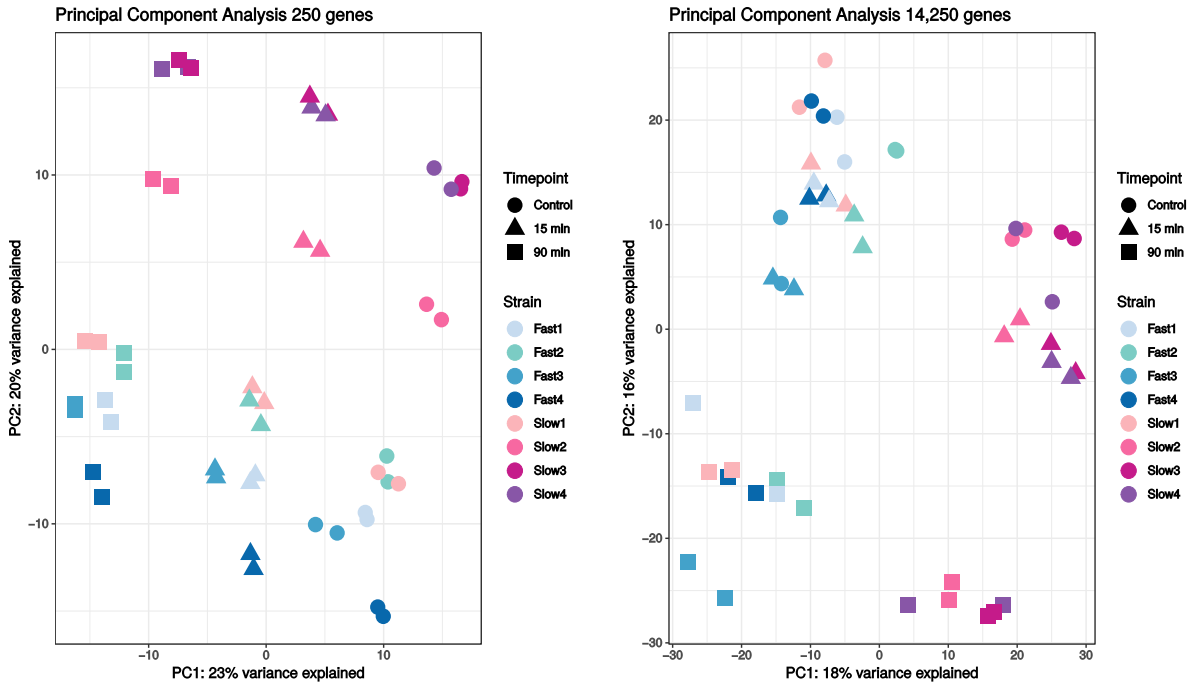


Figure A2: Principal Component Analysis (PCA) for differential gene expression before and after the cold shock in strains with fast recovery and strains with slow recovery.
Left: PCA based on the 250 most variable genes.
Right: PCA based on all 14,250 genes.
Reduced to two dimensions, the two biological replicates for each sample clustered tightly together, and the data grouped according to timepoint and strain irrespective of how many genes were included in the analysis. Figure adapted from (Königer and Grath, 2018).

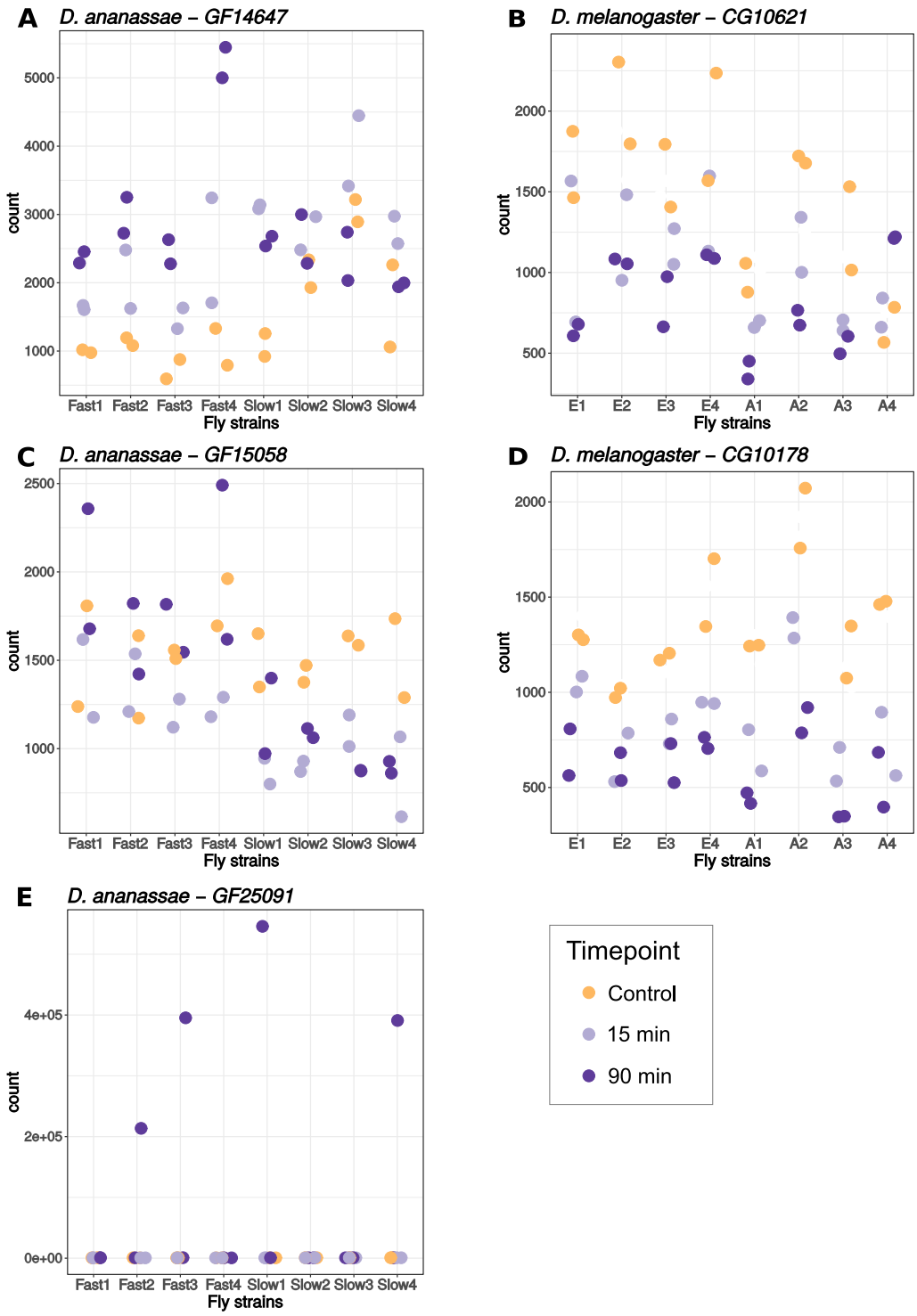


Figure A3: Read counts for genes with a significant interaction of phenotype and timepoint in *D. ananassae* and their orthologs in *D. melanogaster*. *GF14647* in *D. ananassae* (**A**) and its ortholog in *D. melanogaster CG10621* (**B**). *GF15058* in *D. ananassae* (**C**) and its ortholog in *D. melanogaster CG10178* (**D** and *GF25091* which was a false positive in *D. ananassae*. Read counts for the two *D. melanogaster* genes were taken from (von Heckel et al., 2016) wherein the four fast fly strains were from Europe (E1-E4) and the four slow strains were from Africa (A1-A4). Figure adapted from (Königer and Grath, 2018).

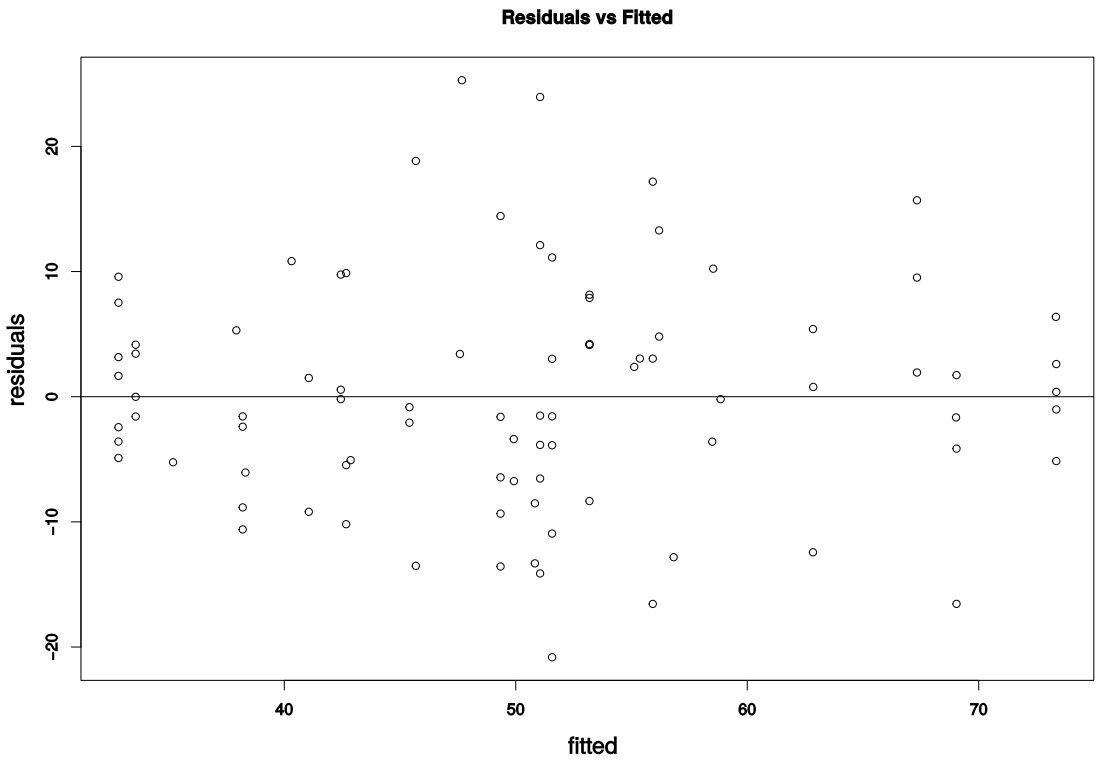


Figure A4: Residuals versus fitted plot of the multiple QTL model y ∼ QTL1 + QTL2 + QTL3 + QTL1:QTL3. The plot suggests that there were no apparent problems with the model fit, such as non-constant variance or outliers. The overall fit of the model was good (LOD score of 19.26).

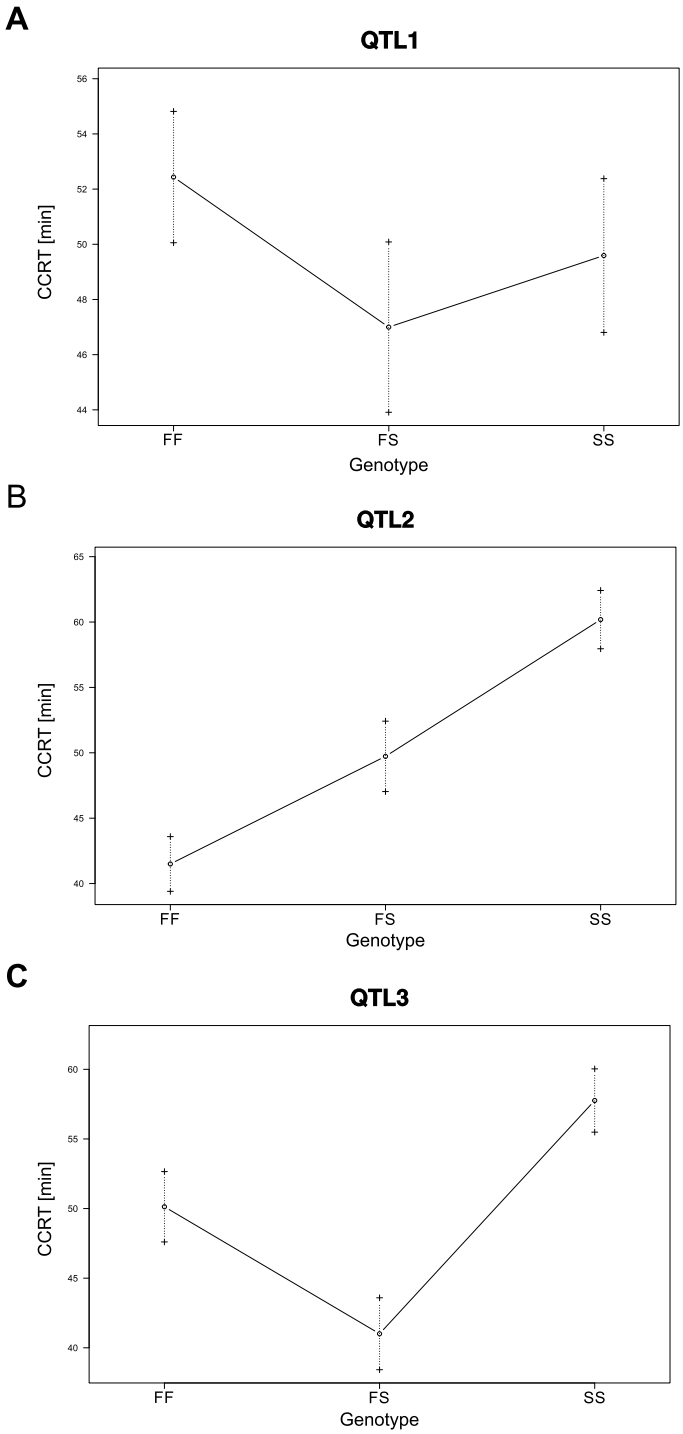


Figure A5: Effectplots for each of the three identified quantitative trait loci (QTL). F = cold-tolerant parental allele (Fast4), S = cold-sensitive parental allele (Slow1). The plots show the effect of the genotype on chill coma recovery time (CCRT) at each QTL separately:
QTL1 (**A**) is transgressive, *i.e.* individuals with the Fast4 allele at QTL1 had a slower recovery. The genotype effect at QTL2 (**B**) was intermediate, *i.e.* Fast4-homozygotes had the fastest recovery, Slow1-homozygotes had the slowest recovery and heterozygotes had intermediate recovery time. At QTL3 (**C**), the heterozygous genotype resulted in the fastest recovery.
Error bars are plotted at $\pm$ 1 standard error (SE).

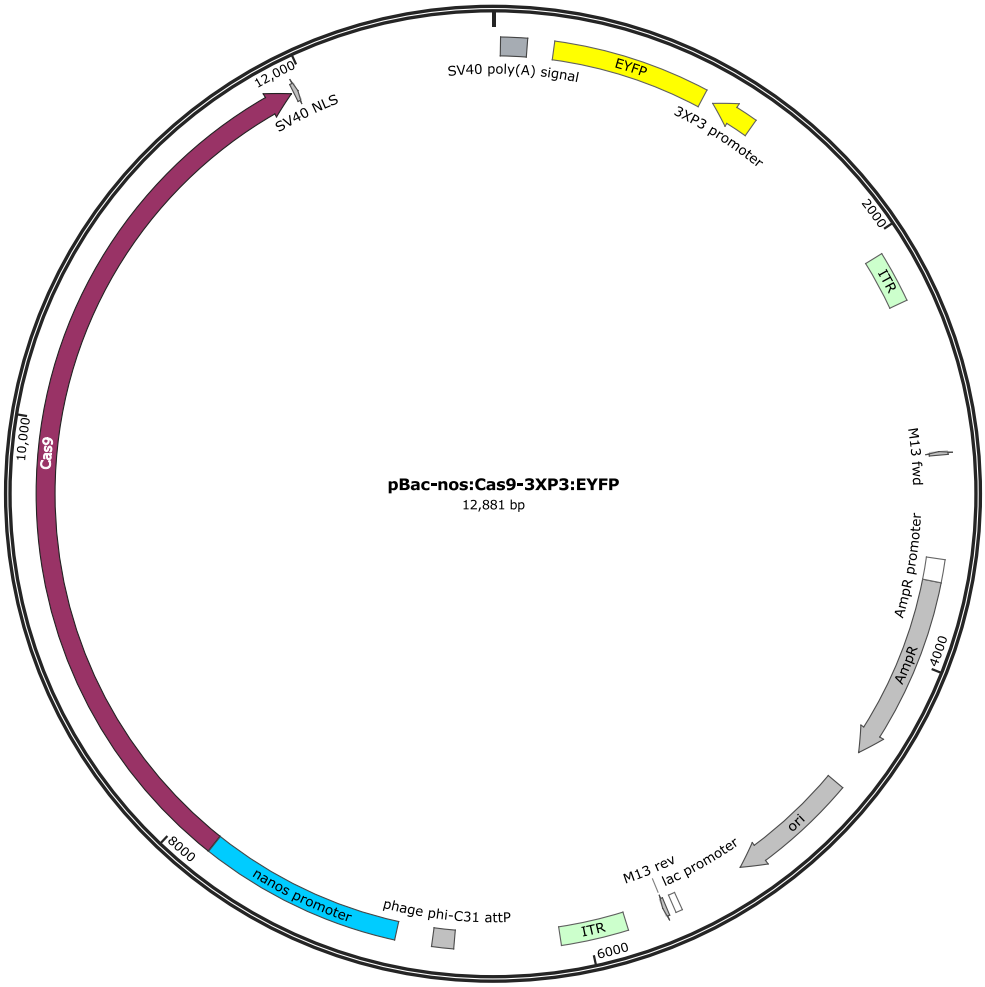


Figure A6: The PiggyBac-nanosCas9-vector was used to insert the Cas9 construct into the genome of *D. ananassae*. The vector map shows the Cas9 construct, consisting of the *nanos* promoter, the Cas9 coding sequence, the Enhanced Yellow Fluorescence Protein (EYFP) coding sequence and its 3XP3 promoter. The construct is flanked by two inverted terminal repeat (ITR) sequences. The co-injected PiggyBac transposase recognizes the ITR sequences, cuts the construct and inserts it into genomic TTAA landing sites.

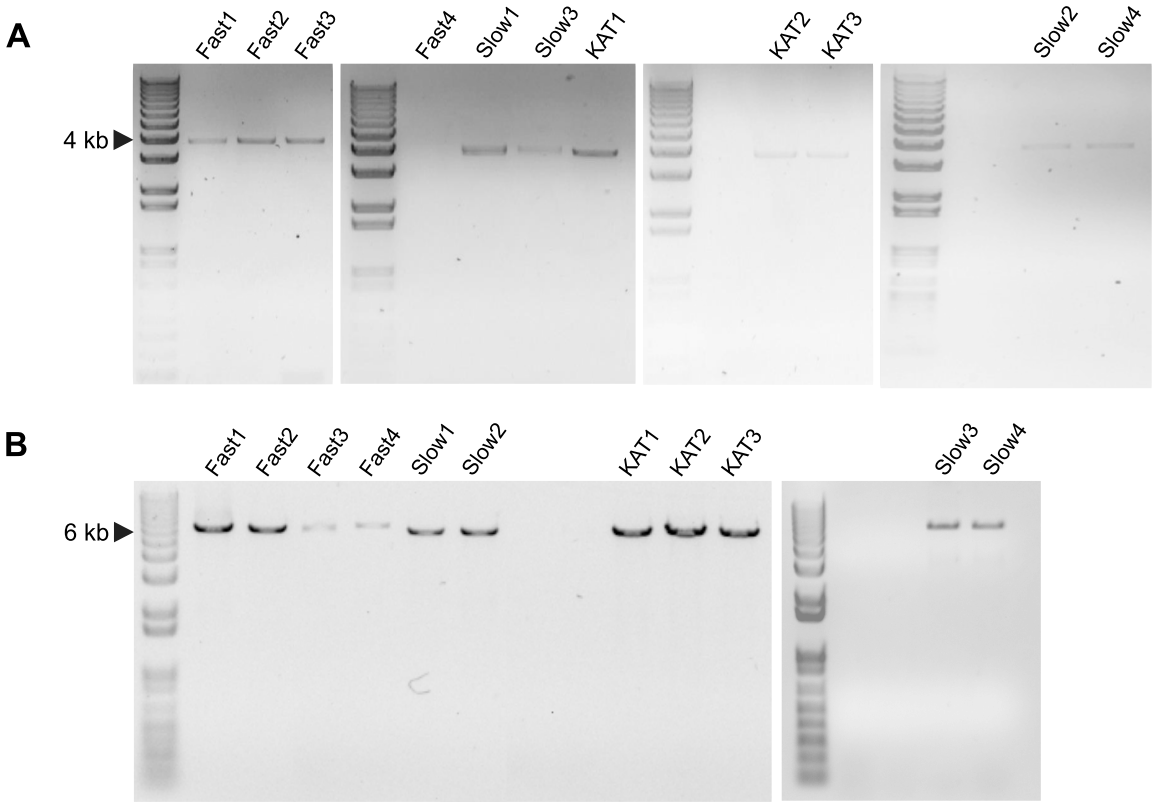


Figure A7: PCR amplification of *GF14647* and *GF15058*. The annealing temperature for *GF14647* (**A**) was 58.1°C, 62.0°C, 58.1°C and 62.5°C (from left to right). The annealing temperature for *GF15058* (**B**) was 62.0°C (left picture) and 68.0°C(right picture).

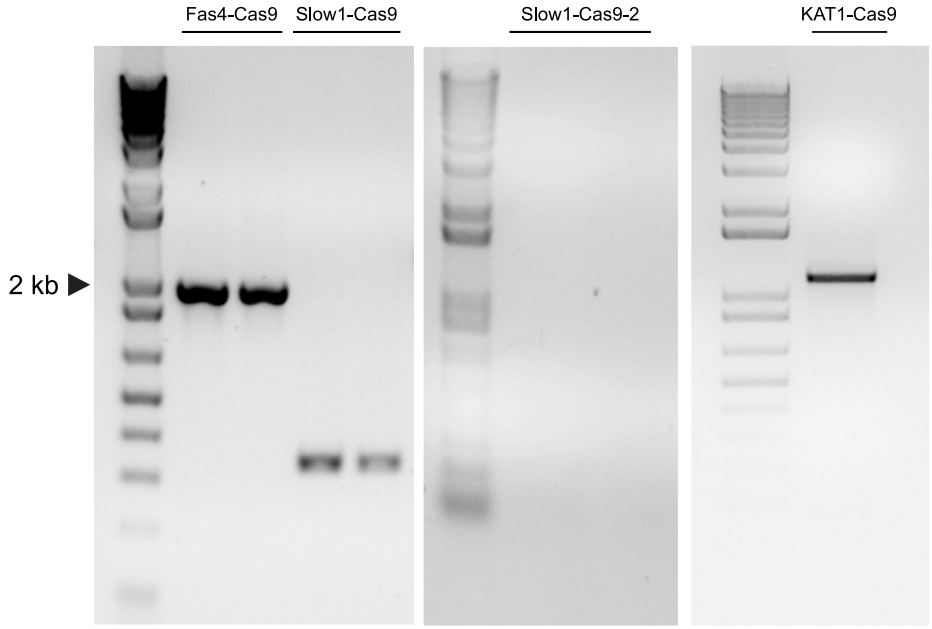


Figure A8: Gel pictures of the inverse PCR to map the Cas9 insert locations. Fast4-Cas9, Slow1-Cas9 and Slow1-Cas9-2 were analyzed in two technical replicates (two lanes next to each other).

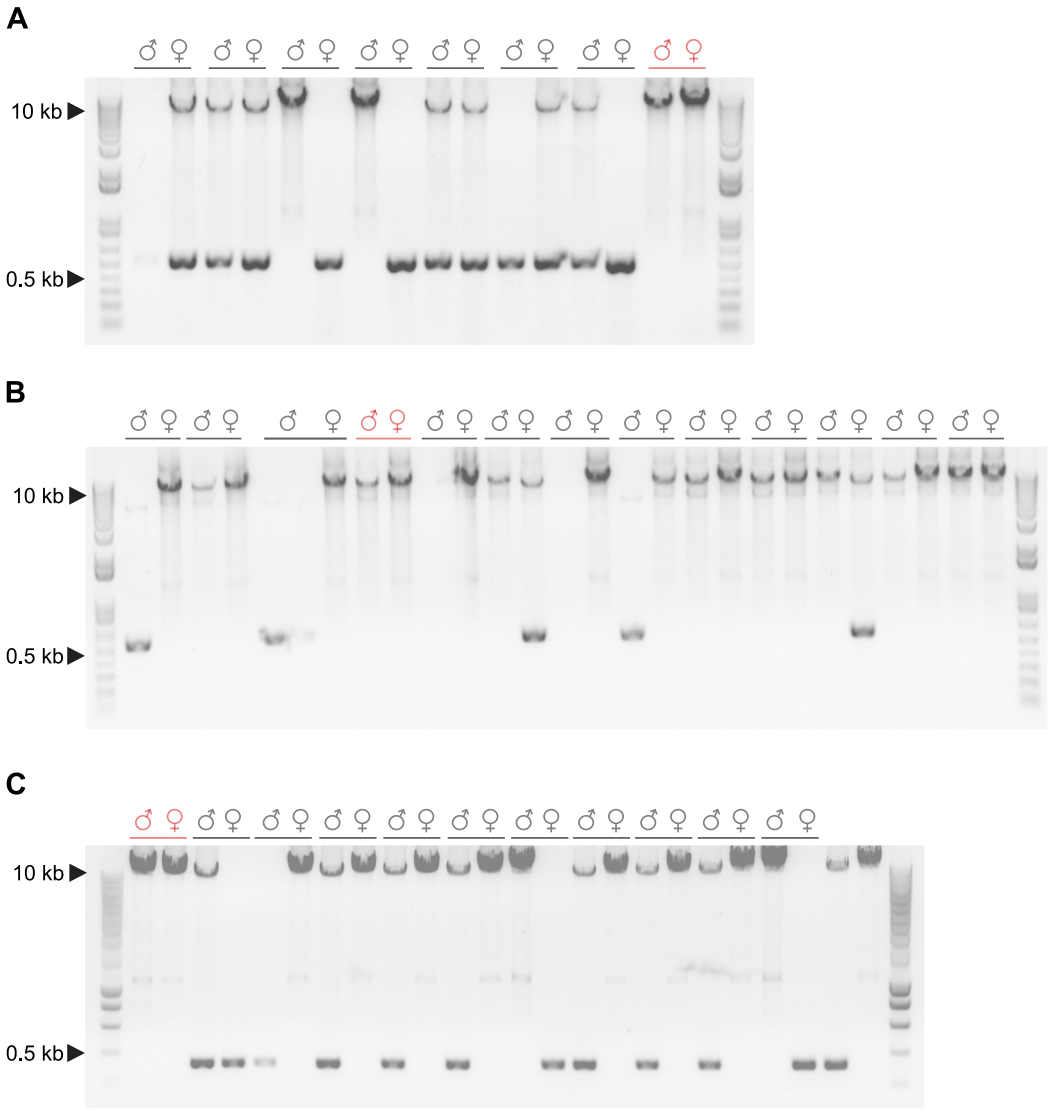


Figure A9: PCR amplification of Cas9 in Fast4-Cas9 (**A**), Slow1-Cas9 (**B**) and KAT1-Cas9 (**C**). Offspring from mating pairs highlighted in red were used to establish homozygous lines.

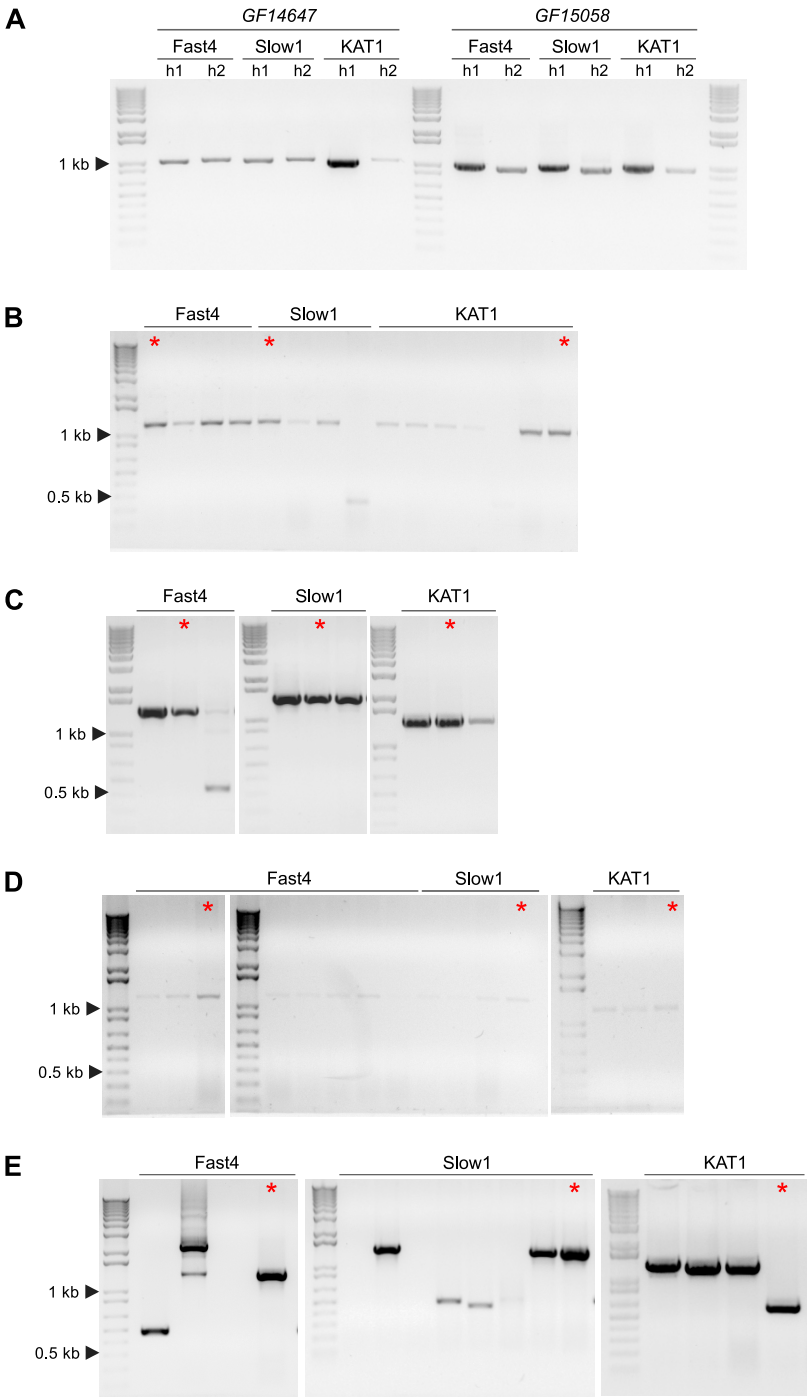


Figure A10: Gel pictures of amplified homology arms for pHD-DsRed-attP cloning. **A**: amplification of homology arm 1 (h1) and homology arm 2 (h2) from genomic DNA of Fast4, Slow1 and KAT1. **B**: colony PCR for arm 1 of *GF14647*. **C**: colony PCR for arm 2 of *GF14647*. **D**: colony PCR for arm 1 of *GF15058*. **E**: colony PCR for arm 2 of *GF15058*. Red asterisks denote samples that were used for subsequent cloning and analysis. Pictures were adapted from (Mußgnug, 2018).

Appendix B

Additional Tables

Table B1: Designation of the fly strains from Bangkok

Fly strain[a]	Original designation[b]
Fast1	BKK5
Fast2	BKK6
Fast3	BKK10
Fast4	BKK12
Slow1	BKK13
Slow2	BKK16
Slow3	BKK17
Slow4	BKK18
KAT1	KATH14

Fly strains from the Bangkok population were renamed according to their chill coma recovery time (Fast or Slow). The original designation is given as a reference to the documentation in my laboratory book.

Table B2: Chill Coma Recovery Time (CCRT) in Recombinant Inbred Advanced Intercross Lines (RIAILs)

RIAIL[a]	CCRT [min]	N	StDev	original designation[b]
1	27.60000	40	12.14612	180
2	27.94118	34	9.17817	160
3	29.25000	40	6.51133	256
4*	29.36000	25	5.49909	320
5	29.95833	48	12.43529	208
6*	30.40000	10	6.78561	108
7	30.75000	48	9.50364	18
8*	31.86364	44	8.17708	91
9*	32.00000	38	11.95487	73
10	32.16667	36	13.90889	329
11*	32.26316	38	11.03480	316
12*	32.48649	37	17.29101	4
13*	33.57447	47	11.73576	159
14	34.50000	40	17.37520	61
15	35.58974	39	15.45233	6
16	35.77778	36	11.35181	28
17	35.80645	31	18.12810	117
18	36.00000	18	15.71810	252
19	36.60000	40	11.14197	20
20	36.63158	38	16.20214	1
21	36.92857	28	13.55900	111
22	37.02041	49	14.08973	88
23	37.03448	29	14.79016	35
24	37.21053	38	16.76895	318
25	37.51020	49	19.60668	58
26	37.73684	38	15.90629	254
27	37.80000	40	20.21830	142
28	38.76000	50	18.62581	34
29	39.36170	47	12.61542	322
30	39.56000	50	18.64508	27
31	40.00000	38	15.11756	107

Table B2 continued

RIAIL[a]	CCRT [min]	N	StDev	original designation[b]
32	40.35000	40	20.40557	273
33	40.63636	44	16.17792	324
34	42.25000	40	20.06751	9
35	42.30769	39	18.98220	109
36	42.42105	38	18.76833	227
37	42.55556	36	14.01722	194
38	42.90000	40	17.08471	250
39	43.00000	40	20.08188	162
40	43.17949	39	26.20453	222
41	43.23077	39	24.61328	3
42	43.33333	30	20.64032	321
43	44.00000	38	13.40432	293
44	44.51282	39	19.82406	174
45	44.57143	49	21.32682	57
46	44.85000	40	19.02704	225
47	45.25000	40	21.94019	10
48	45.30612	49	24.41499	218
49	46.53061	49	21.98020	186
50	46.80000	10	18.90796	32
51	47.20000	50	15.88736	99
52	47.69231	39	22.16939	67
53	47.73333	45	21.40348	276
54	49.53846	39	16.66677	134
55	50.00000	36	20.80659	115
56*	50.41026	39	22.43202	226
57	51.00000	40	22.22957	70
58	51.15000	40	21.03422	82
59	52.20000	40	20.93114	236
60	52.48889	45	24.63610	175
61	52.55000	40	22.55927	72
62	54.60000	30	25.12397	289
63	54.89474	38	24.10651	281

Table B2 continued

RIAIL[a]	CCRT [min]	N	StDev	original designation[b]
64	57.31579	38	22.45609	274
65	57.40000	40	23.06824	303
66	57.50000	36	18.45690	272
67	58.41026	39	23.40571	264
68	58.65000	40	18.76925	241
69	58.96296	27	26.38980	150
70	61.00000	36	27.62814	52
71	61.08000	50	19.81634	290
72	61.33333	30	18.65353	314
73	62.69388	49	26.20926	185
74	63.16000	50	22.64577	292
75	63.63158	38	24.21014	172
76	63.77778	36	24.52650	177
77	64.52174	46	26.50093	229
78	64.89474	38	20.91784	195
79	67.36585	41	20.36388	251
80	68.20833	48	22.08394	92
81	68.24490	49	18.32409	64
82	68.76000	50	22.97075	138
83	69.26531	49	24.68027	228
84	69.48000	50	18.77823	199
85	70.76596	47	22.28102	53
86	72.33333	48	21.77382	233
87	72.96000	50	18.49982	245
88	73.10000	40	23.76789	258
89	73.73684	38	20.59216	327
90	75.00000	38	19.73199	90
91	75.96078	51	19.57750	294
92	76.85106	47	21.47899	143
93	79.71429	49	18.71274	51
94	83.02703	37	17.69257	263

Table B2 continued

RIAIL[a]	CCRT [min]	N	StDev	original designation[b]

Chill Coma Recovery Time (CCRT) in Recombinant Inbred Advanced Intercross Lines (RIAILs).
a) For QTL mapping, RIAILs were numbered in ascending order according to their CCRT. Samples marked with an asterisk were excluded from genetic map construction and mapping.
b) The original designation of each strain is given as a reference to the documentation in my laboratory book. N = number of individual flies tested. StDev = standard deviation.

Table B3: Differentially expressed genes in QTL1

Location on s12916 [bp]	Strand	Symbol	DE in *D. ana*	*D. mel* ortho	DE in *D. mel*
1,495,548..1,514,869	+	*GF15018*	-	*tutl*	-
1,520,121..1,525,756	+	*GF15019*	-	*bdl*	-
1,527,714..1,536,415	+	*GF15020*	↓ F-90	*Atet*	-
1,538,106..1,560,206	+	*GF15021*	↓ F-90, ↓ S-90	*ft*	-
1,560,809..1,561,765	+	*GF15022*	-	*Tspo*	↑ S-90
1,561,928..1,564,626	-	*GF14881*	-	*CG11835*	-
1,565,081..1,566,174	-	*GF14880*	-	*Mis12*	-
1,566,367..1,568,672	+	*GF15023*	-	*CG2794*	-
1,568,587..1,570,386	-	*GF14879*	-	*Nle*	-
1,570,426..1,574,826	+	*GF15024*	-	*Sf3b1*	↓ S-90
1,575,640..1,577,002	-	*GF14878*	-	*Ipk2*	↑ S-90
1,577,308..1,578,361	+	*GF15025*	↓ S-90	*coiled*	↓ F-15, ↓ F-90
1,578,516..1,579,295	-	*GF14877*	-	*Ptth*	-
1,579,402..1,582,025	+	*GF15026*	-	*Pph13*	-
1,587,275..1,597,582	-	*GF14876*	-	*Gsc*	-
1,605,145..1,605,921	+	*GF15027*	-	*CG13689*	-
1,632,499..1,686,835	-	*GF14875*	↓ F-90, ↓ S-90	*ds*	-
1,691,717..1,728,775	+	*GF15030*	-	*GABA-B-R3*	-
1,707,410..1,713,276	-	*GF14874*	-	*Eaat2*	-
1,736,428..1,746,637	+	*GF27242*	-	-	-
1,744,893..1,745,387	+	*GF15033*	-	-	-
1,746,589..1,748,233	+	*GF15034*	-	*Gr21a*	-
1,748,744..1,750,640	+	*GF15035*	-	*CG3544*	-
1,750,708..1,755,003	-	*GF14873*	↓ F-90, ↓ S-90	*Pkg21D*	↓ F-90, ↓ S-90
1,772,051..1,780,258	-	*GF14872*	↑ F-90, ↑ S-90	*l(2)gl*	↑ S-90
1,780,578..1,787,807	-	*GF14871*	↓ S-90	*Zir*	-
1,790,957..1,808,185	+	*GF15036*	↑ F-90, ↑ S-90	*bark*	↑ S-90
1,808,516..1,814,746	+	*GF15037*	↓ F-90, ↓ S-90	*Reph*	↓ S-90
1,815,847..1,818,675	-	*GF14870*	↓ S-90	*CG3407*	-
1,831,795..1,833,121	+	*GF15038*	-	*slp1*	-
1,842,578..1,844,263	+	*GF15039*	-	*slp2*	↑ S-90
1,867,040..1,872,083	+	*GF15040*	-	*TTLL4B*	-

Table B3 continued

Location on s12916 [bp]	Strand	Symbol	DE in *D. ana*	*D. mel* ortho	DE in *D. mel*
1,872,700..1,877,042	-	*GF14868*	-	*RpIIIC160*	-
1,878,589..1,897,459	-	*GF14867*	-	*tup*	-
1,925,977..1,951,286	-	*GF14866*	-	*ssp3*	-
1,951,897..1,965,251	+	*GF15042*	-	*CG42399*	-
1,965,276..1,966,136	-	*GF14865*	↓ F-90	*CG11454*	↓ F-90, ↓ S-90
1,967,249..1,968,732	+	*GF15043*	↑ F-90	*CG31974*	↓ F-90, ↓ S-90
1,968,916..1,970,585	+	*GF15044*	-	*CG31975*	-
1,970,666..1,971,316	+	*GF15045*	-	*ovm*	-
1,974,344..1,976,670	-	*GF14864*	-	*Gs1*	-
1,979,660..1,986,848	+	*GF15046*	-	*CG3164*	-
1,989,043..1,992,555	+	*GF15047*	-	*CG4822*	-
1,994,062..2,000,702	-	*GF14863*	-	*Sam-S*	↓ F-90, ↓ S-90
2,001,882..2,004,466	+	*GF15048*	-	*Nhe1*	↓ F-90
2,004,512..2,006,138	-	*GF14862*	-	*CG11377*	-
2,006,365..2,009,141	-	*GF14861*	-	*Ir21a*	-
2,009,406..2,045,078	-	*GF14860*	-	*Cda5*	↑ S-15
2,052,601..2,056,909	+	*GF15049*	↓ F-90, ↓ S-90	*dbr*	-
2,057,204..2,060,842	+	*GF15050*	-	*galectin*	-
2,061,411..2,062,952	+	*GF15051*	-	*CG11374*	-
2,070,927..2,076,302	-	*GF14859*	↓ F-90, ↓ S-90	*net*	↓ S-90
2,084,748..2,087,250	-	*GF14857*	-	*CG3702*	-
2,087,513..2,088,388	+	*GF15052*	-	*RpL40*	-
2,089,356..2,091,050	-	*GF14855*	-	*JMJD4*	-
2,091,644..2,200,143	-	*GF14854*	↓ S-90	*CG42750*	-
2,110,526..2,113,677	-	*GF27104*	↑ F-90, ↑ S-90	*Ptp36E*	-
2,136,053..2,136,982	-	*GF19690*	-	*CG31802*	-
2,177,783..2,178,132	+	*GF27517*	-	-	-
2,183,121..2,183,791	-	*GF26785*	-	*CG31788*	-
2,208,831..2,209,726	-	*GF14850*	-	*CG31787*	-
2,217,574..2,220,545	+	*GF15053*	-	*CG5790*	-
2,228,820..2,229,385	-	*GF26562*	-	-	-
2,238,351..2,305,035	+	*GF15054*	↑ F-90, ↑ S-90	*Fas3*	↑ S-90

Table B3 continued

Location on s12916 [bp]	Strand	Symbol	DE in *D. ana*	*D. mel* ortho	DE in *D. mel*
2,292,092..2,293,459	+	*GF26595*	-	*CG34171*	-
2,350,978..2,352,475	+	*GF15055*	-	-	-
2,354,253..2,354,922	-	*GF14848*	-	*RpS26*	-
2,355,501..2,359,916	-	*GF14847*	-	*ncm*	-
2,360,731..2,365,671	-	*GF14846*	↑ F-90	*bsf*	-
2,378,771..2,378,852	+	*mir-100*	-	-	-
2,379,181..2,379,258	+	*let-7*	-	-	-
2,379,476..2,379,585	+	*mir-125*	-	-	-
2,379,791..2,388,952	-	*GF14843*	↑ F-90, ↑ S-90	*CG10283*	-
2,390,014..2,392,184	+	*GF15056*	↑ F-90, ↑ S-90	*grnd*	↑ S-90
2,392,681..2,414,594	-	*GF14840*	↑ F-90, ↑ S-90	*kon*	↑ S-90
2,398,918..2,440,302	+	*GF15059*	↑ F-90, ↑ S-90	*CG10211*	-
2,404,008..2,404,654	+	*GF15057*	-	*CG33795*	-
2,405,457..2,407,729	+	*GF27373*	-	*CG33795*	-
2,411,958..2,412,305	+	*GF27163*	-	-	-
2,419,135..2,422,591	+	*GF15058*	↓ S-15, ↓ S-90	*CG10178*	↓ F-90, ↓ S-90
2,441,098..2,455,099	-	*GF14841*	-	*Pde11*	↑ S-90
2,494,015..2,498,307	+	*GF15060*	-	*CG15160*	-
2,499,026..2,499,720	-	*GF14839*	-	*amos*	-
2,502,281..2,506,868	-	*GF14838*	↑ F-90, ↑ S-90	*CG10413*	-
2,507,906..2,508,405	+	*GF15061*	-	*CG31789*	-
2,508,742..2,511,416	+	*GF15062*	-	*CG10333*	-
2,511,422..2,514,048	-	*GF14837*	-	*Atac2*	-
2,515,796..2,517,143	-	*GF14836*	-	*IFT46*	-
2,518,575..2,552,873	+	*GF15063*	↑ F-90, ↑ S-90	*MESR3*	↑ S-15, ↑ S-90
2,543,259..2,545,540	-	*GF14835*	-	*Cyp310a1*	-
2,555,430..2,556,305	-	*GF27716*	-	-	-
2,556,351..2,556,943	+	*GF15064*	-	-	-
2,556,918..2,561,366	-	*GF14834*	-	*CG43338*	-
2,561,447..2,563,516	-	*GF14833*	-	*CG43339*	-
2,563,622..2,565,397	-	*GF14832*	-	*CG43339*	-
2,565,573..2,571,881	+	*GF15065*	↓ F-90, ↓ S-90	*CG31751*	↓ F-90, ↓ S-90

Table B3 continued

Location on s12916 [bp]	Strand	Symbol	DE in *D. ana*	*D. mel* ortho	DE in *D. mel*
2,567,443..2,569,985	-	*GF14831*	-	*tos*	-
2,572,777..2,577,104	-	*GF26668*	-	*msl-1*	-
2,577,393..2,578,666	+	*GF15066*	-	*CG10336*	-
2,578,789..2,582,801	-	*GF14829*	↑ F-90, ↑ S-90	*CG10383*	↑ S-15, ↑ S-90
2,583,124..2,585,545	+	*GF15067*	-	*CG10338*	-
2,584,454..2,600,466	+	*GF15071*	-	*ninaA*	-
2,585,794..2,587,695	+	*GF15068*	-	*CG10341*	-
2,587,684..2,589,809	-	*GF14828*	↑ F-90, ↑ S-90	*CG10376*	↑ S-90
2,589,914..2,590,993	+	*GF15069*	-	*CG10343*	-
2,591,026..2,592,271	-	*GF14827*	-	*dbe*	↓ F-90
2,592,374..2,598,545	+	*GF15070*	↓ S-15	*PNUTS*	-
2,600,354..2,607,461	-	*GF14826*	-	*CG15824*	-
2,636,157..2,640,509	+	*GF15073*	-	*GluRIIC*	-
2,652,207..2,678,093	+	*GF15075*	-	*CG4341*	-

Differentially expressed (DE) genes located in the 95% confidence interval of QTL1 on scaffold 12916.
D.ana = D. ananassae, *D. mel = D. melanogaster*, ortho = orthologous gene
F = differentially expressed in strains with Fast chill coma recovery time
S = differentially expressed in strains with Slow chill coma recovery time
↑ = upregulated, ↓ = downregulated
90 = 90 min after the cold shock, 15 = 15 min after the cold shock

Table B4: Differentially expressed genes in QTL2

Location on s13337 [bp]	Strand	*Symbol*	DE in *D. ana*	*D. mel* ortho	DE in *D. mel*
121,240..122,441	+	*GF24897*	-	*CG7049*	↓ F-90, ↓ S-90
122,461..138,056	-	*GF24884*	↑ F-90, ↑ S-90	*p130CAS*	↑ S-90
148,097..151,194	+	*GF24898*	-	*Dic61B*	-
151,249..154,350	-	*GF24881*	-	*CG6845*	↓ F-90
154,478..155,444	+	*GF26299*	-	*CG30383*	-
160,850..179,359	-	*GF24880*	↑ F-90, ↑ S-90	*Pdk1*	↑ S-90
199,080..199,772	+	*GF24899*	-	-	-
222,421..223,261	+	*GF24900*	-	*CG13405*	-
80,650..119,476	+	*GF24896*	↑ F-90	*klar*	-
87,559..87,906	-	*GF24885*	-	-	-
88,547..88,909	+	*GF27787*	-	-	-

Differentially expressed (DE) genes located in the 95% confidence interval of QTL2 on scaffold 13337.
D.ana = *D. ananassae*, *D. mel* = *D. melanogaster*, ortho = orthologous gene
F = differentially expressed in strains with Fast chill coma recovery time
S = differentially expressed in strains with Slow chill coma recovery time
↑ = upregulated, ↓ = downregulated
90 = 90 min after the cold shock, 15 = 15 min after the cold shock

Table B5: Differentially expressed genes in QTL3

Location on s13340 [bp]	Strand	Symbol	DE in *D. ana*	*D. mel* ortho	DE in *D. mel*
5,532,124..5,546,898	+	*GF17803*	↓ F-90, ↓ S-90	*lab*	↓ S-90
5,548,646..5,549,289	+	*GF17804*	-	*agt*	↓ F-90, ↓ S-90
5,549,261..5,550,377	-	*GF17183*	-	*twr*	-
5,550,653..5,551,985	+	*GF17805*	-	*CG1307*	-
5,552,088..5,560,942	-	*GF17182*	-	*Taf1*	-
5,561,229..5,564,545	+	*GF17806*	↑ F-90, ↑ S-90	*unc-45*	↑ S-90
5,582,047..5,628,124	+	*GF17807*	↑ F-90, ↑ S-90	*CG31176*	↑ S-90
5,629,154..5,634,034	+	*GF17808*	-	*CG6349*	-
5,634,019..5,643,637	-	*GF17181*	-	*CG15497*	-
5,638,525..5,639,444	+	*GF17809*	↑ F-90	*Archease*	-
5,640,656..5,676,687	+	*GF17810*	-	*E2f1*	-
5,675,686..5,675,764	+	*mir-11*	-	-	-
5,704,357..5,722,361	+	*GF17811*	↑ F-90, ↑ S-90	*InR*	-
5,724,210..5,726,429	+	*GF17812*	-	*sll*	↓ F-90, ↓ S-90
5,728,577..5,729,553	+	*GF17813*	-	*CG14329*	-
5,729,978..5,730,576	-	*GF17180*	-	-	-
5,731,916..5,732,354	+	*GF17815*	-	-	-
5,733,123..5,733,826	+	*GF17816*	-	*CG14327*	↑ S-90
5,734,176..5,734,248	+	*GF25778*	-	-	-
5,734,541..5,734,612	-	*GF25779*	-	-	-
5,734,798..5,735,210	-	*GF17179*	-	-	-
5,735,577..5,736,050	+	*GF17818*	-	-	-
5,736,374..5,736,445	+	*GF25780*	-	-	-
5,736,720..5,736,792	-	*GF25781*	-	-	-
5,738,075..5,738,647	-	*GF17178*	-	*CG14323*	-
5,739,003..5,739,075	+	*GF25782*	-	-	-
5,741,247..5,741,320	-	*GF25783*	-	-	-
5,741,826..5,741,899	-	*GF25784*	-	-	-
5,742,395..5,743,072	+	*GF17819*	-	*AttD*	-
5,744,230..5,746,331	+	*GF17820*	-	*Toll-4*	-
5,747,468..5,749,360	+	*GF17821*	-	*CG14325*	-
5,749,579..5,749,651	-	*GF25785*	-	-	-

Table B5 continued

Location on s13340 [bp]	Strand	Symbol	DE in *D. ana*	*D. mel* ortho	DE in *D. mel*
5,751,423..5,751,495	-	*GF25786*	-	-	-
5,764,250..5,764,322	-	*GF25787*	-	-	-
5,764,555..5,764,626	+	*GF25788*	-	-	-
5,765,517..5,765,589	-	*GF25789*	-	-	-
5,766,660..5,786,042	+	*GF17823*	-	*CG14322*	-
5,774,987..5,775,059	+	*GF25790*	-	-	-
5,775,976..5,776,048	-	*GF25791*	-	-	-
5,785,200..5,785,272	+	*GF25792*	-	-	-
5,786,274..5,787,698	-	*GF17177*	-	*CG7523*	-
5,787,953..5,790,299	+	*GF17824*	-	*Mps1*	-
5,791,041..5,799,009	-	*GF17176*	↑ S-90	*alt*	↑ S-90
5,799,081..5,800,017	+	*GF17825*	↑ F-90, ↑ S-90	*CG7655*	↑ S-90
5,800,028..5,801,200	-	*GF17174*	-	*CG31251*	-
5,801,274..5,801,796	+	*GF17826*	-	-	-
5,801,967..5,803,023	-	*GF17173*	-	*CG31249*	-
5,803,255..5,804,289	+	*GF17827*	-	*CG3817*	-
5,804,166..5,804,986	-	*GF17172*	-	*VhaPPA1-2*	-
5,805,147..5,805,899	-	*GF17171*	-	-	-
5,806,020..5,806,960	-	*GF17170*	-	*RpS5b*	-
5,807,530..5,809,271	-	*GF17169*	-	*VhaPPA1-1*	-
5,823,352..5,826,619	+	*GF17828*	-	*Sdr*	↓ F-90, ↓ S-90
5,921,743..5,923,233	+	*GF17829*	-	*CG14861*	-
5,945,410..5,946,278	+	*GF17830*	-	*RpL10Aa*	-
5,959,834..5,979,566	-	*GF17168*	-	*dpr9*	-
5,981,561..5,982,481	-	*GF17167*	-	*CG6974*	-
5,982,749..5,983,501	-	*GF17166*	-	*CG6974*	-
5,984,725..5,995,799	-	*GF17165*	-	*CG6966*	-
5,997,984..5,998,290	-	*GF26293*	-	-	-
5,997,984..6,024,351	-	*GF17163*	↓ F-90, ↓ S-90	*CG42788*	-
6,024,559..6,035,088	-	*GF17162*	-	*CG6912*	-
6,035,297..6,037,317	+	*GF17831*	-	*CG3984*	↓ S-90
6,037,746..6,039,603	+	*GF17832*	-	*CG3987*	-

Table B5 continued

Location on s13340 [bp]	Strand	Symbol	DE in *D. ana*	*D. mel* ortho	DE in *D. mel*
6,039,992..6,043,368	-	*GF17161*	↓ S-90	*GlyS*	-
6,043,616..6,045,338	-	*GF17160*	-	*CG6912*	-
6,051,807..6,052,359	-	*GF26426*	-	*CG43089*	-
6,054,144..6,056,054	+	*GF17834*	-	-	-
6,056,473..6,059,815	-	*GF27827*	-	-	-
6,066,496..6,106,084	+	*GF17835*	-	*Gyc88E*	-
6,080,150..6,089,482	-	*GF17159*	-	*Mf*	-
6,106,481..6,111,679	-	*GF17158*	-	*CG6752*	-
6,112,044..6,118,272	+	*GF17836*	↓ S-90	*CG42542*	-
6,118,278..6,121,872	-	*GF17157*	-	*Spn88Ea*	↑ S-90
6,122,295..6,124,563	+	*GF17837*	↓ S-90	*Mau2*	-
6,124,597..6,126,876	-	*GF17156*	↓ S-15, ↓ S-90	*CG6654*	-
6,127,324..6,128,687	-	*GF17155*	-	*CG8319*	-
6,128,690..6,129,858	+	*GF27458*	-	-	-
6,130,890..6,132,033	+	*GF17838*	-	*CG9427*	↑ S-15
6,132,169..6,163,176	-	*GF17154*	↑ F-90, ↑ S-90	*CG8312*	↑ S-90
6,163,567..6,165,278	+	*GF17839*	-	*aurA*	-
6,165,218..6,167,967	-	*GF17153*	-	*CG12213*	↓ F-90
6,168,250..6,169,333	+	*GF17840*	-	*RpS29*	-
6,168,563..6,168,702	+	*snoRNA*	-	-	-
6,170,379..6,173,176	+	*GF17841*	-	*CG12947*	-
6,174,005..6,174,422	-	*MtnA*	-	*MtnA*	↓ F-90, ↓ S-90
6,174,653..6,178,608	+	*GF17842*	-	*CG8500*	-
6,178,646..6,180,836	-	*GF17151*	-	*CG12945*	-
6,181,095..6,182,723	+	*GF17843*	-	*CG8507*	↓ F-90, ↓ S-90
6,183,370..6,187,434	+	*GF17844*	-	*CG8516*	↓ S-90
6,187,694..6,190,969	-	*GF17150*	↑ F-90, ↑ S-90	*CG9467*	-
6,191,369..6,193,459	+	*GF17845*	-	*CG8526*	-
6,193,512..6,199,377	-	*GF17149*	-	*FBXO11*	-
6,204,017..6,204,881	-	*GF17148*	-	*CG9458*	↑ S-15
6,206,278..6,206,705	-	*GF17147*	-	*CG42857*	-
6,208,461..6,209,062	-	*GF17146*	-	*CG34302*	-

Table B5 continued

Location on s13340 [bp]	Strand	Symbol	DE in *D. ana*	*D. mel* ortho	DE in *D. mel*
6,215,599..6,242,856	-	*GF17145*	-	*Teh1*	-
6,246,751..6,253,807	+	*GF26557*	-	-	-
6,248,172..6,252,153	-	*GF26742*	-	-	-
6,260,806..6,374,519	+	*GF17848*	-	*Glut4EF*	↑ S-90
6,369,074..6,371,988	-	*GF17144*	-	*Art4*	↓ S-90
6,383,440..6,386,743	+	*GF17850*	-	*Spn85F*	-
6,386,781..6,387,646	-	*GF17142*	-	*CG5359*	-
6,388,349..6,424,107	-	*GF17141*	-	-	-
6,409,097..6,410,053	-	*GF17140*	-	*CG31407*	-
6,429,594..6,430,884	+	*GF17851*	-	*CG3909*	-
6,430,877..6,431,839	-	*GF17139*	-	*CG11722*	-
6,431,886..6,433,477	+	*GF17852*	↓ S-90	*mtTFB2*	↓ F-90, ↓ S-90
6,433,616..6,435,061	-	*GF17138*	-	*CG12811*	↓ F-90, ↓ S-90
6,435,436..6,437,938	+	*GF17853*	-	*ohgt*	-
6,437,903..6,442,658	-	*GF17137*	-	*Fancl*	↑ S-90
6,439,802..6,441,157	+	*GF17854*	-	*Npc2e*	↑ S-90
6,442,278..6,443,352	+	*GF17855*	-	*Npc2e*	-
6,443,837..6,450,986	-	*GF17136*	↓ F-90, ↓ S-90	*Npc2d*	↓ F-90, ↓ S-90
6,448,951..6,449,702	+	*GF17856*	↓ F-15	*Npc2c*	↑ S-90
6,450,588..6,451,255	+	*GF19922*		*CG33631*	-
6,451,480..6,452,111	+	*GF19923*		*CG33630*	-
6,456,902..6,459,402	+	*GF17857*		*nerfin-2*	-
6,459,498..6,461,052	-	*GF17135*	↑ F-90, ↑ S-90	*Alp13*	↓ F-90
6,479,993..6,480,083	+	*mir-317*	-	-	-
6,488,996..6,489,093	+	*mir-277*	-	-	-
6,489,985..6,490,081	+	*mir-34*	-	-	-
6,491,822..6,496,983	-	*GF17134*	-	*Fmr1*	-
6,498,364..6,499,621	-	*GF27525*	-	-	-
6,501,822..6,507,827	+	*GF17859*	-	*CG3940*	↓ F-15, ↓ F-90, ↓ S-90
6,514,342..6,515,352	+	*GF17860*	-	*TwdlW*	-
6,515,565..6,516,455	-	*GF17133*	-	*CG5255*	-

Table B5 continued

Location on s13340 [bp]	Strand	Symbol	DE in *D. ana*	*D. mel* ortho	DE in *D. mel*
6,517,010..6,518,056	-	*GF17132*	↑ F-90, ↑ S-90	*CG5246*	↑ F-15, ↓ S-15, ↓ S-90
6,518,793..6,519,882	-	*GF17131*	↓ S-90	*CG31267*	↓ F-90, ↓ S-90
6,520,464..6,521,437	-	*GF17130*	↓ S-90	*CG31266*	↓ F-90
6,521,839..6,522,838	+	*GF17861*	↓ F-90, ↓ S-90	*CG4053*	-
6,523,493..6,524,550	+	*GF17862*	↓ S-90	*CG17475*	↓ F-90, ↓ S-90
6,524,876..6,525,732	-	*GF19903*	↓ S-90	*CG31265*	↓ F-90, ↓ S-90
6,525,775..6,526,717	-	*GF17129*	-	*CG31269*	↓ S-90
6,526,920..6,527,729	+	*GF17863*	↓ F-90, ↓ S-90	*CG17477*	↓ S-90
6,528,337..6,564,381	+	*GF17864*	↑ F-90, ↑ S-90	*cher*	↑ S-90
6,534,112..6,534,336	-	*GF27237*	-	-	-
6,534,431..6,534,699	-	*GF28051*	-	-	-

Differentially expressed (DE) genes located in the 95% confidence interval of QTL3 on scaffold 13340.
D.ana = D. ananassae, *D. mel = D. melanogaster*, ortho = orthologous gene
F = differentially expressed in strains with Fast chill coma recovery time
S = differentially expressed in strains with Slow chill coma recovery time
↑ = upregulated, ↓ = downregulated
90 = 90 min after the cold shock, 15 = 15 min after the cold shock

Table B6: Enriched GO terms and pathways in QTL3

GO Category	Term	*P*-value	corrected *P*-value	Genes	Fold Enrichment
MF	serine-type endopeptidase activity GO:0004252	0.00038	0.01521	*GF17129* *GF19903* *GF17861* *GF17131* *GF17132* *GF17862* *GF17863* *GF17133* *GF17130*	4.83418
BP	intracellular cholesterol transport GO:0032367	0.04212	0.74766	*GF17136* *GF17856*	45.32941

P-values were corrected according to (Benjamini and Hochberg, 1995), MF = molecular function, BP = biological process.

Table B7: PCR primers

Primer pair #	Amplification of	Product size [bp]	Orientation	Sequence 5′->3′
1	*GF15058*	6,034	F	CCTCTAGCCACTTGAGATGAGT
			R	CACCCATAGGGAGTCTTAGTTC
2	*GF14647*	4,096	F	TGGAATCGCAGCAATCTCTGA
			R	GGGGCCATATTGATCCTCCAAT
3	inverse PCR	-	F	TAGCCGAGTCTCTGACTGA
	+ sequencing		R	CGGCGACTGAGATGTCCTAAA
4	Fast4	763 [A]	F	TGACCAGACCAAAATCGCCC
	Cas9 genotyping	10,144 [B]	R	ATGGCGGACTTAGAAGCTGT
5	Slow1	749 [A]	F	CGCAGGAGCGGAATGATAGG
	Cas9 genotyping	10,130 [B]	R	GGCAAGGTGACAAGTTAGCG
6	KAT1	328 [A]	F	GGTGGAATGGGGAGCTTCTT
	Cas9 genotyping	9,709 [B]	R	TCTCGGCACTATGAGCTGAC
7	*GF15058*	947	F	ATTACACCTGCTTATTCGCCACGTCTCAATGATGACAAG
	homology arm 1		R	AAATCACCTGCATATCTACCATCAGCCAGCACTTCCACT
8	*GF15058*	913	F	AATTGCTCTTCCTATGACTGTCGAGCTATAGCTCT
	homology arm 2		R	TATTGCTCTTCTGACCAGGCTGAATGTGGGCTCCT
9	*GF14647*	1,001	F	ATTACACCTGCTTATTCGCTGCGGGACAAACTTTGTTTA
	homology arm 1		R	AAATCACCTGCATATCTACGACGCGCGTCAAACCCATGA
10	*GF14647*	1,000	F	AATTGCTCTTCCTATGATGGTGGTTTCGGTACCCA
	homology arm 2		R	TATTGCTCTTCTGACCGTAGGTCTCGCCACTGTTT
11	Colony PCR	292 [A]	F	ACGAAAGGCTCAGTCGAAAG
	homology arm 1	1,292 or 1,239 [B]	R	CGGTCGAGGGTTCGAAATCGATAAG
12	Colony PCR	486 [A]	F	CCACCACCTGTTCCTGTAG
	homology arm 2	1,399 or 1,486 [B]	R	TGATATCAAAATTATACATGTCAACG

A) without insert, **B)** with insert

Table B8: Sequencing primers

Primer number #	Amplification of	Sequence 5′->3′
1	*GF15058*	GGCTTCTGATCATTGTAT
2	*GF15058*	CAGGTCTTAGAACATACT
3	*GF15058*	TCGGCGTCTTGTGTTGTG
4	*GF15058*	TTCACTTGTTCTGATAAG
5	*GF15058*	TGCGCGTTCTAATTAATG
6	*GF15058*	TGATTGCATGTGTTGCCA
7	*GF15058*	TAACACCAGAGCCCAGGA
8	*GF15058*	TAGGCTTAGGACTATTGA
9	*GF15058*	ACAGTTAGCTTGTGAGAA
10	*GF15058*	TCGAGAACTGCGCATTGT
11	*GF15058*	CACATCCAATTGGCAGTC
12	*GF15058*	CATGATTGAAGTGGAGTC
13	*GF15058*	ACTGGCTCTGGCTGGAGC
14	*GF15058*	AACTCCACTACTCCTTAG
15	*GF15058*	AATGACCTTCAGCCAGCG
16	*GF15058*	ACCTAGGAGTTGCAATAT
17	*GF15058*	AACGTCCTGATTCGGAAG
18	*GF15058*	CCTCACCGACTTAATGGT
19	*GF15058*	GAAGCCAGCTCCAAACTG
20	*GF15058*	AGTCTCGGTAGTCACCCG
21	*GF15058*	CTTGTATCAGATAGTGTT
22	*GF14647*	CGCTGGTACCCTCTGGCT
23	*GF14647*	GGTCGTCCGGGTTGAGTG
24	*GF14647*	AAGGTGAGTGCGGGACAA
25	*GF14647*	AAGTTCTTCAAATACTCA
26	*GF14647*	AACACTCACGTTGCAAGT
27	*GF14647*	TATATAGATTCAGGGCTT
28	*GF14647*	CACGTGCCGGATTACTAG
29	*GF14647*	ACACGGTGCCAGTCGGTG
30	*GF14647*	GCCCGAGATATTCGCCAC
31	*GF14647*	CCGAAACTCGACTACAAA
32	M13 forward	GTAAAACGACGGCCAG
33	M13 reverse	CAGGAAACAGCTATGAC

Appendix C

Laboratory Protocols

Protocol 1: RNA-extraction from whole flies with MasterPureTM RNA Purification Kit (Epicentre Madison, WI, USA)

All samples and working steps should be handled on ice. Volumes are given for RNA extraction from 1 single fly per sample.

1. Clean all surfaces and pipettes with RNAse AWAY® (Thermo Fisher Scientific) and 70% Ethanol.
2. Collect 10 flies per sample into a 1.5 ml reaction tube and freeze the samples in liquid nitrogen.
3. Dilute 1 µl of Proteinase K (50 µg/µl) into 300 µl of Tissue and Cell Lysis Solution for each sample.
4. Grind frozen tissues thoroughly, leave pestle inside till after step 4.
5. Add 300 µl of Tissue and Cell Lysis Solution containing the Proteinase K, remove pestle and vortex thoroughly.
6. Incubate at 65°C for 15 min; vortex every 5 min.
7. Place the samples on ice for 3 – 5 min.
8. Add 175 µl of MPC Protein Precipitation Reagent and vortex vigorously for 10 sec.
9. Pellet the debris by centrifugation at 4°C for 10 min at 10,000g (= 14,000rpm).
10. Transfer the supernatant to a clean 1.5ml reaction tube and discard the pellet.
11. Add 500 µl of Isopropanol to the recovered supernatant. Invert the tube 30 – 40 times.
12. Pellet the total nucleic acids by centrifugation at 4°C for 10 min at 14,000 rpm.
13. Remove all of the residual Isopropanol with a pipette. Steps 14 – 16 are optional (DNAseI digestion):

14. Prepare 200 µl of DNase I solution for each sample by diluting 15 µl of RNase-free DNase I up to 200 µl with 1X DNase Buffer. (Pipette up and down – do not vortex)
15. Completely resuspend the total nucleic acid pellet in 200 µl of DNase I solution.
16. Incubate at 37°C for at least 30 min.
17. Add 200 µl of 2X T and C Lysis Solution and vortex for 5 sec.
18. Add 200 µl of MPC Protein Precipitation Reagent, vortex 10 sec, place on ice for 3 – 5 min.
19. Pellet the debris by centrifugation at 4°C for 10 min at 14,000 rpm.
20. Transfer the supernatant into a clean 1.5 ml reaction tube and discard the pellet.
21. Add 500 µl of Isopropanol to the supernatant. Invert the tube 30 – 40 times.
22. Pellet the purified RNA by centrifugation at 4°C for 10 min at 14,000 rpm.
23. Carefully remove the Isopropanol without dislodging the RNA pellet (twice).
24. Add 500 µl 70% Ethanol, invert tube carefully, centrifuge at 4°C for 10 min at 14,000 rpm.
25. Remove all of the residual Ethanol very carefully with a pipette (twice). Air dry pellet for 15 min at RT (not on ice).
26. Resuspend (up and down many times) the RNA in 30 µl RiboGuard™ RNase Inhibitor-mastermix. Keep the mastermix as blank for spectrophotometric measurement.
27. Measure concentration and purity with a spectrophotometer. Store RNA samples at -80°C.

Protocol 2: DNA-extraction from whole flies with MasterPureTM DNA Purification Kit (Epicentre Madison, WI, USA)

Volumes are given for DNA extraction from one single fly [and five flies] per sample.

1. Collect 1 fly per sample into a 1.5 ml reaction tube and freeze the samples in liquid nitrogen.
2. Dilute 0.3 µl [1 µl] of Proteinase K (50 µg/µl) into 300 µl of Tissue and Cell Lysis Solution for each sample.
3. Grind frozen tissues thoroughly, leave pestle inside till after step 4.
4. Add 300µl of Tissue and Cell Lysis Solution containing the Proteinase K, remove pestle and vortex thoroughly.
5. Incubate at 65°C for 30 – 60 min at 300 rpm.
6. Place the samples on ice for 1 min (let cool to 37°C).
7. Add 1µl [5µl] RNAse A and mix thoroughly (invert 25x).
8. Incubate at 37°C for 30 min. Cool to 4°C (on ice 3 – 5 min.)
9. Add 58 µl [175 µl] MPC Protein Precipitation solution and vortex vigorously for 10 sec.
10. Place the samples on ice for 3 – 5 min.
11. Pellet the debris by centrifugation at 4°C for 10 min at 13,200 rpm
12. Transfer the supernatant to a clean 1.5ml reaction tube and discard the pellet.
13. Add 170 µl [500 µl] of Isopropanol to the recovered supernatant. Invert the tube 30 – 40 times.
14. Pellet the total nucleic acids by centrifugation at 4°C for 10 min at 13,200 rpm.
15. Remove all of the residual Isopropanol with a pipette.
16. Add 170 µl [500 µl] of 70% Ethanol and centrifuge at 4°C for 2 min at 12,500 rpm.
17. Remove all of the residual Ethanol with a pipette (twice). For sequencing, repeat steps 16 – 17
18. Air dry pellet for 15 – 30 min at RT (not on ice).
19. Resuspend pellet in 25 µl H_2Odd.
20. Measure concentration and purity with a spectrophotometer. Store DNA samples at -20°C.

Protocol 3: DNA-extraction from whole flies with the QIAGEN® Dneasy Blood Tissue Kit (QIAGEN, Venlo, Netherlands)

Volumes are given for DNA extraction from 10 pooled flies per sample.

1. Collect 10 flies per sample into a 1.5ml reaction tube and freeze the samples in liquid nitrogen.
2. Grind frozen tissues thoroughly and add 180 µl Buffer ATL.
3. Add 20 µl Proteinase K, remove pestle and vortex thoroughly.
4. Incubate at 56°C for 60 min at 300 rpm.
5. Vortex for 15 sec. Add 200 µl Buffer AL and vortex again. Quickly add 200 µl 96% Ethanol and vortex again.
6. Transfer each sample into a spin column (placed into a 2 ml collection tube). Centrifuge for 1 min at 8,000 rpm. Discard the collection tube and flow-through.
7. Place the spin column into a new 2 ml collection tube. Add 500 µl Buffer AW1 and centrifuge for 1 min at 8,000 rpm.
8. Add 500 µl Buffer AW2 and centrifuge for 1 min at 8,000 rpm. Discard the collection tube and flow-through.
9. Place the spin column into a new 1,5 ml collection tube, pipet 50 µl H_2Odd directly onto the spin column membrane. Incubate for 1 min at room temperature.
10. Elute the DNA by centrifugation for 1 min at 8,000 rpm.
11. Measure concentration and purity with a spectrophotometer. Store DNA samples at -20°C.

Protocol 4: PCR with Phusion® High-Fidelity DNA Polymerase (Thermo Fisher Scientific, Waltham, MA, USA)

H_2Odd	add to 50 µl
5X Phusion HF or GC Buffer	10 µl
10 mM dNTPs	1 µl
10 µM Forward Primer	2.5 µl
10 µM Reverse Primer	2.5 µl
Template DNA	50 – 250 ng
Phusion DNA Polymerase	0.5 µl

Thermocycling conditions for a routine PCR:

1.	98°C	30 sec
2.	98°C	10 sec
3.	45 – 72°C	30 sec
4.	72°C	15 – 30 sec per kb
Repeat steps 2 – 4	24 – 34 times	
5.	72°C	10 min
6.	8°C	hold

Protocol 5: for PCR with LongAmp® Taq DNA Polymerase (New England Biolabs, Ipswich, MA, USA)

H_2Odd	add to 50 µl
5X Phusion HF or GC Buffer	10 µl
10 mM dNTPs	1.5 µl
10 µM Forward Primer	2.0 µl
10 µM Reverse Primer	2.0 µl
Template DNA	1 – 500 ng
Phusion DNA Polymerase	2.0 µl

Thermocycling conditions for a routine PCR:

1.	94°C	30 sec
2.	94°C	30 sec
3.	45 – 72°C	45 sec
4.	60 – 65°C	50 sec per kb
Repeat steps 2 – 4	29 times	
5.	65°C	10 min
6.	8°C	hold

Protocol 6: Cycle sequencing reaction for BigDye® Terminator v1.1 (Thermo Fisher Scientific)

1. Add 1.5 µl ExoSAP-IT™ PCR Product Cleanup Reagent to the PCR reaction.
2. Incubate for 30 min at 37°C; heat inactivate for 15 min at 80°C.
   Steps 1 and 2 are required for sequencing of PCR products. For sequencing of purified plasmids, start from step 3.
3. Assemble the following reaction:

H_2Odd	add to 10 µl
5X BigDye Sequencing Buffer	1 µl
Ready Reaction Premix	2 µl
Primer	1 µl
DNA	1 µl

Thermocycling conditions:

1.	96°C	1 min
2.	96°C	10 sec
3.	50°C	15 sec
4.	60°C	4 min
Repeat 2 – 4	39 times	
5.	12°C	hold

Subsequent electrophoresis was run on an ABI 3730 automated sequencer (Applied Biosystems, Foster City, CA, USA) by the Genomics Service Unit (LMU Munich, Faculty of Biology, Division of Genetics, Planegg-Martinsried, DE).

Protocol 7: Restriction digest with NEB enzymes (New England Biolabs, Ipswich, MA, USA)

H_2Odd	add to 50 µl
CutSmart Buffer	5 µl
DNA	variable
Enzyme	1 unit/ 1 µg DNA

Thermocycling conditions for a routine digest:

1. 37°C 15 h
2. 65°C 20 min
3. 8°C hold

Protocol 8: Ligation with T4 DNA Ligase (New England Biolabs, Ipswich, MA, USA)

H_2Odd	add to 10 µl
10X T4 DNA Ligase Buffer	1 µl
Vector DNA	50 ng
Insert DNA	37.5 ng
T4 DNA Ligase	1 µl

Thermocycling conditions for a routine ligation:

1. 37°C 15 h
2. 65°C 10 min
3. 8°C hold

Protocol 9: Transformation of One Shot® TOP10 Competent Cells (Invitrogen, Waltham, MA, USA)

Competent cells should be handled on ice. Do not vortex. Equilibrate 42°C water bath and incubate S.O.C medium at 37°C beforehand.

1. Carefully thaw 50 µl Top10 cells on ice.
2. Add 5 µl of the ligation reaction (or 1 µl of purified plasmid) and mix by tapping gently.
3. Incubate on ice for 30 min.
4. Incubate for exactly 30 sec in a 42°C water bath, place back on ice.
5. Add 250 µl pre-warmed S.O.C medium.
6. Incubate for 1 hour at 37°C and 325 rpm.
7. Spread the cells on 2 LB plates (+ 100 µg/ml Ampicillin).
8. Incubate at 37°C overnight (15 h).
9. Select colonies for PCR or plasmid isolation.

Protocol 10: Plasmid isolation with ZyppyTM Plasmid Miniprep Kit (ZYMO Research, Irvine, Ca, USA)

1. Select a colony and inoculate a 5 ml liquid LB (+ 100 µg/ml Ampicillin) culture. Incubate overnight (15 h) at 37°C.
2. Transfer 1.5 ml of cell culture to a 1.5 ml reaction tube and centrifuge for 30 sec at 15,000 rpm.
3. Discard the supernatant and repeat step 2.
4. Resuspend the pellet in 600 µl H_2Odd.
5. Add 100 µl of 7X Lysis Buffer and mix by inverting the tube 4 – 6 times and proceed to the next step within 2 min.
6. Add 350 µl pre-cooled Neutralization Buffer and mix thoroughly. Invert the samples 2 – 3 times.
7. Centrifuge at 16,000 g for 4 min.
8. Transfer the supernatant into the provided spin column.
9. Place the spin column into a collection tube and centrifuge for 30 sec at 16,000 g.
10. Discard the flow-through and add 200 µl Endo-Wash Buffer. Centrifuge for 30 sec at 16,000 g.
11. Add 400 µl Zyppy™Wash Buffer to the column. Centrifuge for 1 min at 16,000 g.
12. Place the spin column into a new 1,5 ml collection tube, pipet 30 µl H_2Odd directly onto the spin column membrane. Incubate for 1 min at room temperature.
13. Elute the DNA by centrifugation for 1 min at 16,000 g.
14. Measure concentration and purity with a spectrophotometer. Store DNA samples at -20°C.

Protocol 11: Plasmid isolation with the QIAGEN® Plasmid Midi Kit (QIAGEN, Venlo, Netherlands)

1. Select a colony and inoculate an 8 ml liquid LB (+ 100 µg/ml Ampicillin) starter culture. Incubate for 8 h at 37°C.
2. Use 4 ml of the starter culture to inoculate a 50 ml liquid LB (+ 100 µg/ml Ampicillin) culture. Incubate overnight (15 h) at 37°C.
3. Transfer the cell culture into a 50 ml falcon and centrifuge for 15 min at 4°C and 6,000 g.
4. Discard the supernatant and resuspend the pellet in 4 ml Buffer P1.
5. Add 4 ml Buffer P2 and mix thoroughly by vigorously inverting the falcon and incubate for 5 min at room temperature.
6. Add 4 ml of pre-cooled Buffer P3 and mix immediately and thoroughly by vigorously inverting 4 – 6 times, and incubate on ice for 15 min.
7. Centrifuge at 20,000 g for 30 min at 4°C.
8. In the meantime, equilibrate the filter tip by applying 4 ml Buffer QBT and allow it to empty by gravity flow.
9. Apply the supernatant from step 7. to the barrel of the provided filter cartridge and incubate it at room temperature for 10 min.
10. Remove the cap from the filter cartridge outlet nozzle, install the cartridge on top of the filter tip and insert the plunger into the cartridge. Filter the lysate from the cartridge into the equilibrated filter tip and allow it to enter the resin by gravity flow.
11. Wash the filter tip twice with 10 ml QC Buffer. Discard the flow-through.
12. Elute DNA with 5 ml Buffer QF into a new 15 ml falcon.
13. Distribute the elute into 4 separate 1,5 ml reaction tubes.
14. Precipitate DNA in each tube with 500 µl Isopropanol and centrifuge at 15,000 g for 30 min at 4°C.
15. Discard the supernatant and wash the pellets with 500 µl 70% Ethanol. Centrifuge at 15,000 g for 10 min at 4°C. Remove all Ethanol with a pipette tip.
16. Air-dry the pellets for 15 min at room temperature.
17. Resuspend the pellets in 30 µl H_2Odd and combine all 4 samples into 1 tube.
18. Measure concentration and purity with a spectrophotometer. Store DNA samples at -20°C.

Protocol 12: PCR clean-up with the QIAquick® PCR Purification Kit (QIAGEN, Venlo, Netherlands)

1. Add 5 volumes of Buffer PB to 1 volume of the PCR reaction and mix.
2. Apply the mix to a provided spin column (placed in a 2 ml collection tube).
3. Centrifuge for 1 min at 13,000 rpm.
4. Discard the flow-through and add 750 µl Buffer PE.
5. Centrifuge for 1 min at 13,000 rpm.
6. Discard the flow-through and centrifuge again for 1 min at 13,000 rpm.
7. Place the spin column into a new 1.5 ml reaction tube and apply 30 µl H_2Odd to the center of the column.
8. Incubate at room temperature for 4 min and elute DNA by centrifuging for 1 min at 13,000 rpm.
9. Measure concentration and purity with a spectrophotometer. Store DNA samples at -20°C.

Protocol 13: Gel extraction with the QIAquick® Gel Extraction Kit (QIAGEN, Venlo, Netherlands)

1. Excise the DNA band from the agarose gel with a provided cutter and weigh it.
2. Add 3 volumes of QG Buffer to 1 volume of gel.
3. Incubate for 45 min at 50°C and 350 rpm, vortex every 15 min.
4. Add 1 volume of Isopropanol and mix.
5. Apply the sample to a provided spin column (placed into a 2 ml collection tube).
6. Centrifuge for 1 min at 13,000 rpm.
7. Discard the flow-through and centrifuge again for 1 min at 13,000 rpm.
8. Place the spin column into a new 1.5 ml reaction tube and apply 30 µl H_2Odd to the center of the column.
9. Incubate at room temperature for 4 min and elute DNA by centrifuging for 1 min at 13,000 rpm.
10. Measure concentration and purity with a spectrophotometer. Store DNA samples at -20°C.

Protocol 14: Colony-PCR with Taq DNA Polymerase (Invitrogen, Waltham, MA, USA)

H_2Odd	add to 20 µl
10X PCR Rxn Buffer (-$MgCl_2$)	2 µl
dNTPs (10 mM)	0.8 µl
MgCl2 (50 mM)	0.8 µl
Forward primer (10 mM)	0.4 µl
Reverse primer (10 mM)	0.4 µl
Taq Polymerase	0.2 µl

Each selected colony was touched with a pipette tip which was then dipped into the PCR mix and finally used to inoculate a numbered grid field on an LB (+ 100 µg/ml Ampicillin) plate by gently touching the surface. The plate was incubated for 3 – 4 h at 37°C.

PCR thermocycling conditions:

1.	94°C	30 sec
2.	94°C	30 sec
3.	55°C	30 sec
4.	68°C	30 sec
Repeat steps 2 – 4	29 times	
5.	68°C	5 min
6.	8°C	hold

Analyse PCR products on a 1% agarose gel. Selected colonies can be picked from the incubated plate and be used to inoculate 5 ml liquid LB (+ 100 µg/ml Ampicillin) cultures for plasmid isolation (Miniprep).

Protocol 15: Phosphorylation and annealing with T4 Polynucleotide Kinase (New England Biolabs, Ipswich, MA, USA)

H_2Odd	6.5 µl
10 x T4 Ligase Buffer	1.0 µl
Sense Strand (100 mM)	1.0 µl
Antisense Strand (100 mM)	1.0 µl
T4 Polynucleotide Kinase	0.5 µl

Cycling Conditions:

1.	37 °C	30 min
2.	95 °C	5 min
3.	Ramp down to 25°C at 5°C/min	
4.	8°C	hold

Protocol 16: Recipe for egg-laying plates
Mix the following ingredients in a 500 ml bottle:

Sucrose	6 g
Agar	11 g
H_2Odd	375 ml

Autoclave the mix and allow it to cool down to 60°C. Quickly add the following ingredients:

Grapejuice	125 ml
Propionic acid	2.4 ml
Nipagin	1 g dissolved in 3.4 ml 96% Ethanol

Mix and pour plates immediately. Store the plates for 24 h at room temperature, then at 8°C.

Protocol 17: Inverse PCR

1. Digest genomic DNA:

H_2Odd	39 µl
DNA	5 µl
CutSmart Buffer	5 µl
Enzyme (New England Biolabs)	1 µl

Incubate overnight (15 h) at 37°C; heat inactivate 20 min at 65°C.

2. Ligate digested DNA:

H_2Odd	39 µl
DNA	5 µl
10X Ligation Buffer	5 µl
T4 DNA Ligase (New England Biolabs)	1 µl

Incubate 10 min on ice, then overnight (15 h) at 16°C.

3. Precipitate DNA:

Add 40 µl 3 M NaOAc and 1 ml 96% Ethanol to each sample, invert the tube and incubate for at least 1 h at -80°C. Centrifuge for 30 min at 15,000 rpm and 4°C. Remove the supernatant, wash the pellet with 500 µl pre-chilled 70% Ethanol and centrifuge again for 15 min at 15,000 rpm and 4°C. Remove all Ethanol carefully with a pipette tip and allow the pellet to dry at room temperature. Resuspend in 40 µl H_2Odd.

4. inverse PCR:

H_2Odd	10.4 µl
10X Buffer	2 µl
10 mM dNTPs	0.8 µl
MgCl	0.8 µl
10 µM Forward Primer	0.4 µl
10 µM Reverse Primer	0.4 µl
DNA	5 µl
Taq Polymerase (Invitrogen)	0.2 µl

Thermocycling conditions for a routine inverse PCR:

1. 95°C	4 min
2. 95°C	30 sec
3. 55°C	45 sec
4. 72°C	1 min 30 sec
Repeat steps 2 – 4	29 times
5. 72°C	7 min
6. 8°C	hold

Appendix D

Supplementary Material

The following files and data are available as supplementary material:

The folder **thesis.gz** contains all documents and files for the writing of this thesis, including:

- The original LaTeX files
- All figures as separate .pdf files
- All tables in .csv format
- Additional file 4 with lists of differentially expressed genes

The folder **CCRT.gz** contains files for the phenotypic analysis, including:

- The raw CCRT data for Bangkok strains and RIAILs as .txt files
- The file CCRT.R, which includes the R code for the statistical analysis (see Tables 3.1, 3.2 and 3.3) and generation of the figures 1.3, 3.1, and 3.7

The folder **transcriptome_analysis.gz** contains the following files:

- The text file Transcriptome_analysis_documentation.txt, which describes the computational analysis of the raw sequence reads prior to analysis in R
- All shell and python scripts referred to in the Transcriptome_analysis_documentation.txt file
- The table gCounts.csv, which contains the gene counts of all fly strain and timepoint samples and was used as an input file for the analysis with DESeq2

- The file Transcriptome_analysis.R includes the R code for differential gene expression analysis with DESeq2, including the codes for the generation of the figures 3.2, 3.3, 3.4, 3.5, 3.6, A2 and A3

The folder **QTL_mapping.gz** contains the following files:

- The text file QTL_mapping_documentation.txt, which describes the computational analysis pipeline of the raw sequence reads prior to analysis in R
- All shell and python scripts referred to in the QTL_mapping_documentation.txt file
- The table Cold_r75_crossfile6.csv, which contains the genotype and marker information of all RIAILs and was used as an input file for the analysis with R/qtl
- The file QTL_mapping.R, which includes the R code for genetic map construction and QTL mapping, including the codes for the generation of figures 3.8, 3.9, 3.10, 3.11, A4 and A5

Bibliography

Al-Attar, S., Westra, E. R., van der Oost, J. and Brouns, S. J. (2011), 'Clustered regularly interspaced short palindromic repeats (CRISPRs): the hallmark of an ingenious antiviral defense mechanism in prokaryotes', *Biological Chemistry* **392**(4), 277–289.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990), 'Basic local alignment search tool', *Journal of Molecular Biology* **215**(3), 403–410.

Andersen, J., Findsen, A. and Overgaard, J. (2013), 'Feeding impairs chill coma recovery in the migratory locust (Locusta migratoria)', *Journal of insect physiology* **59**, 1041–1048.

Andersen, J. L., Manenti, T., Sørensen, J. G., MacMillan, H. A., Loeschcke, V. and Overgaard, J. (2015), 'How to assess Drosophila cold tolerance: chill coma temperature and lower lethal temperature are the best predictors of cold distribution limits', *Functional Ecology* **29**(1), 55–65.

Andersen, L. H., Kristensen, T. N., Loeschcke, V., Toft, S. and Mayntz, D. (2010), 'Protein and carbohydrate composition of larval food affects tolerance to thermal stress and desiccation in adult Drosophila melanogaster', *Journal of Insect Physiology* **56**(4), 336–340.

Andrews, S. (2014), 'FastQC A Quality Control tool for High Throughput Sequence Data', *http://www.bioinformatics.babraham.ac.uk/projects/fastqc* .

Angilletta, M. J. (2009), *Thermal Adaptation: A Theoretical and Empirical Synthesis*, Oxford University Press.

Araújo, M. B., Ferri-Yáñez, F., Bozinovic, F., Marquet, P. A., Valladares, F. and Chown, S. L. (2013), 'Heat freezes niche evolution', *Ecology Letters* **16**(9), 1206–1219.

Arends, D., Prins, P., Jansen, R. C. and Broman, K. W. (2010), 'R/qtl: high-throughput multiple QTL mapping', *Bioinformatics* **26**(23), 2990–2992.

Attrill, H., Falls, K., Goodman, J. L., Millburn, G. H., Antonazzo, G., Rey, A. J. and Marygold, S. J. (2016), 'FlyBase: establishing a Gene Group resource for Drosophila melanogaster', *Nucleic Acids Research* **44**(Database issue), D786–D792.

Ayrinhac, A., Debat, V., Gibert, P., Kister, A.-G., Legout, H., Moreteau, B., Vergilino, R. and David, J. R. (2014), 'Cold adaptation in geographical populations of Drosophila melanogaster: phenotypic plasticity is more important than genetic variability', *Functional Ecology* **18**(5), 700–706.

Bale, J. S. and Hayward, S. a. L. (2010), 'Insect overwintering in a changing climate', *The Journal of Experimental Biology* **213**(6), 980–994.

Bassett, A. R., Tibbit, C., Ponting, C. P. and Liu, J.-L. (2013), 'Highly Efficient Targeted Mutagenesis of Drosophila with the CRISPR/Cas9 System', *Cell Reports* **4**(1), 220–228.

Bayley, J. S., Winther, C. B., Andersen, M. K., Grønkjær, C., Nielsen, O. B., Pedersen, T. H. and Overgaard, J. (2018), 'Cold exposure causes cell death by depolarization-mediated Ca2+ overload in a chill-susceptible insect', *Proceedings of the National Academy of Sciences* **115**(41), E9737–E9744.

Benjamini, Y. and Hochberg, Y. (1995), 'Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing', *Journal of the Royal Statistical Society. Series B (Methodological)* **57**(1), 289–300.

Berghammer, A. J., Klingler, M. and Wimmer, E. A. (1999), 'A universal marker for transgenic insects', *Nature* **402**(6760), 370–371.

Bibikova, M., Golic, M., Golic, K. G. and Carroll, D. (2002), 'Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases.', *Genetics* **161**(3), 1169–1175.

Bier, E., Harrison, M. M., O'Connor-Giles, K. M. and Wildonger, J. (2018), 'Advances in Engineering the Fly Genome with the CRISPR-Cas System', *Genetics* **208**(1), 1–18.

Bock, I. and Wheeler, M. (1972), 'The Drosophila melanogaster species group', *University of Texas Publications* **7213**, 1–102.

Bock, K. W. (2015), 'The UDP-glycosyltransferase (UGT) superfamily expressed in humans, insects and plants: Animal-plant arms-race and co-evolution', *Biochemical pharmacology* **99**, 11–17.

Boettcher, M. and McManus, M. T. (2015), 'Choosing the Right Tool for the Job: RNAi, TALEN, or CRISPR', *Molecular Cell* **58**(4), 575–585.

Bowman, L. L., Kondrateva, E. S., Timofeyev, M. A. and Yampolsky, L. Y. (2018), 'Temperature gradient affects differentiation of gene expression and SNP allele frequencies in the dominant Lake Baikal zooplankton species', *Molecular Ecology* **27**(11), 2544–2559.

Brandão, A., Eng, K. K., Rito, T., Cavadas, B., Bulbeck, D., Gandini, F., Pala, M., Mormina, M., Hudson, B., White, J., Ko, T.-M., Saidin, M., Zafarina, Z., Oppenheimer, S., Richards, M. B., Pereira, L. and Soares, P. (2016), 'Quantifying the legacy of the Chinese Neolithic on the maternal genetic heritage of Taiwan and Island Southeast Asia', *Human Genetics* **135**, 363–376.

Broman, K. and Sen, S. (2009), *A Guide to QTL Mapping with R/qtl*, Statistics for Biology and Health, Springer, New York City, NY.

Broman, K. W., Wu, H., Sen, S. and Churchill, G. A. (2003), 'R/qtl: QTL mapping in experimental crosses', *Bioinformatics* **19**(7), 889–890.

Carroll, S. B. (2008), 'Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution', *Cell* **134**(1), 25–36.

Castrillon, D. H., Gonczy, P., Alexander, S., Rawson, R., Eberhart, C. G., Viswanathan, S., DiNardo, S. and Wasserman, S. A. (1993), 'Toward a Molecular Genetic Analysis of Spermatogenesis in Drosophila Melanogaster: Characterization of Male-Sterile Mutants Generated by Single P Element Mutagenesis', *Genetics* **135**(2), 489–505.

Catalán, A., Glaser-Schmitt, A., Argyridou, E., Duchen, P. and Parsch, J. (2016), 'An Indel Polymorphism in the MtnA 3' Untranslated Region Is Associated with Gene Expression Variation and Local Adaptation in Drosophila melanogaster', *PLoS Genetics* **12**(4), e1005987.

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W. and Postlethwait, J. H. (2011), 'Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences', *G3: Genes|Genomes|Genetics* **1**(3), 171–182.

Catchen, J. M., Hohenlohe, P. A., Bernatchez, L., Funk, W. C., Andrews, K. R. and Allendorf, F. W. (2017), 'Unbroken: RADseq remains a powerful tool for understanding the genetics of adaptation in natural populations', *Molecular Ecology Resources* **17**, 362–365.

Chae, H.-J., Ke, N., Kim, H.-R., Chen, S., Godzik, A., Dickman, M. and Reed, J. C. (2003), 'Evolutionarily conserved cytoprotection provided by Bax Inhibitor-1 homologs from animals, plants, and yeast', *Gene* **323**, 101–113.

Chen, Y.-R., Jiang, T., Zhu, J., Xie, Y.-C., Tan, Z.-C., Chen, Y.-H., Tang, S.-M., Hao, B.-F., Wang, S.-P., Huang, J.-S. and Shen, X.-J. (2017), 'Transcriptome sequencing reveals potential mechanisms of diapause preparation in bivoltine silkworm Bombyx mori (Lepidoptera: Bombycidae)', *Comparative Biochemistry and Physiology. Part D, Genomics & Proteomics* **24**, 68–78.

Colinet, H., Lee, S. F. and Hoffmann, A. (2010*a*), 'Knocking down expression of Hsp22 and Hsp23 by RNA interference affects recovery from chill coma in Drosophila melanogaster', *Journal of Experimental Biology* **213**(24), 4146–4150.

Colinet, H., Lee, S. F. and Hoffmann, A. (2010*b*), 'Temporal expression of heat shock genes during cold stress and recovery from chill coma in adult Drosophila melanogaster', *The FEBS Journal* **277**(1), 174–185.

Colinet, H., Siaussat, D., Bozzolan, F. and Bowler, K. (2013), 'Rapid decline of cold tolerance at young age is associated with expression of stress genes in Drosophila melanogaster', *Journal of Experimental Biology* **216**(2), 253–259.

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A. and Zhang, F. (2013), 'Multiplex Genome Engineering Using CRISPR/Cas Systems', *Science* **339**(6121), 819–823.

Cooper, J. A. and Sept, D. (2008), 'New Insights into Mechanism and Regulation of Actin Capping Protein', *International review of cell and molecular biology* **267**, 183–206.

Cottam, D. M., Tucker, J. B., RogersBald, M. M., Mackie, J. B., Macintyre, J., Scarborough, J. A., Ohkura, H. and Milner, M. J. (2006), 'Non-centrosomal microtubule-organising centres in cold-treated cultured Drosophila cells', *Cell Motility* **63**(2), 88–100.

Coulter, D. E., Swaykus, E. A., Beran-Koehn, M. A., Goldberg, D., Wieschaus, E. and Schedl, P. (1990), 'Molecular analysis of odd-skipped, a zinc finger encoding segmentation gene with a novel pair-rule expression pattern', *The EMBO journal* **9**(11), 3795–3804.

Cutler, D. J. and Jensen, J. D. (2010), 'To pool, or not to pool?', *Genetics* **186**(1), 41–3.

Das, A., Mohanty, S. and Stephan, W. (2004), 'Inferring the Population Structure and Demography of Drosophila ananassae From Multilocus Data', *Genetics* **168**(4), 1975–1985.

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M. and Blaxter, M. L. (2011), 'Genome-wide genetic marker discovery and genotyping using next-generation sequencing', *Nature Reviews Genetics* **12**(7), 499–510.

David, J., Gibert, P., Pla, E., Petavy, G., Karan, D. and Moreteau, B. (1998), 'Cold stress tolerance in Drosophila: Analysis of chill coma recovery in D. Melanogaster', *Journal of Thermal Biology* **23**, 291–299.

David, J. R., Gibert, P., Moreteau, B., Gilchrist, G. W. and Huey, R. B. (2003), 'The fly that came in from the cold: geographic variation of recovery time from low-temperature exposure in Drosophila subobscura', *Functional Ecology* **17**(4), 425–430.

Des Marteaux, L. E., McKinnon, A. H., Udaka, H., Toxopeus, J. and Sinclair, B. J. (2017), 'Effects of cold-acclimation on gene expression in Fall field cricket (Gryllus pennsylvanicus) ionoregulatory tissues', *BMC Genomics* **18**(1), 357.

Des Marteaux, L. E., Stinziano, J. R. and Sinclair, B. J. (2018), 'Effects of cold acclimation on rectal macromorphology, ultrastructure, and cytoskeletal stability in Gryllus pennsylvanicus crickets', *Journal of Insect Physiology* **104**, 15–24.

Dietzl, G., Chen, D., Schnorrer, F., Su, K.-C., Barinova, Y., Fellner, M., Gasser, B., Kinsey, K., Oppel, S., Scheiblauer, S., Couto, A., Marra, V., Keleman, K. and Dickson, B. J. (2007), 'A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila', *Nature* **448**(7150), 151–156.

Dobzhansky, T. (1968), On Some Fundamental Concepts of Darwinian Biology, *in* T. Dobzhansky, M. K. Hecht and W. C. Steere, eds, 'Evolutionary Biology: Volume 2', Springer, Boston, MA, pp. 1–34.

Donggi, P., Yeo, G. J., Young, E. L., Young, N. L., Rochelle, Y., Heon, Y. G., Seungmin, Y., Eunkyung, B., Kyung-Jin, M., Marc, T. and Joong-Jean, P. (2012), 'Misexpression screen delineates novel genes controlling Drosophila lifespan', *Mechanisms of Ageing and Development* **133**(5), 234–245.

Doren, M. V., Williamson, A. L. and Lehmann, R. (1998), 'Regulation of zygotic gene expression in Drosophila primordial germ cells', *Current Biology* **8**(4), 243–246.

Drobnis, E. Z., Crowe, L. M., Berger, T., Anchordoguy, T. J., Overstreet, J. W. and Crowe, J. H. (1993), 'Cold shock damage is due to lipid phase transitions in cell membranes: A demonstration using sperm as a model', *Journal of Experimental Zoology* **265**(4), 432–437.

Drosophila 12 Genomes Consortium (2007), 'Evolution of genes and genomes on the *Drosophila* phylogeny', *Nature* **450**(7167), 203–218.

Elhanany-Tamir, H., Yu, Y. V., Shnayder, M., Jain, A., Welte, M. and Volk, T. (2012), 'Organelle positioning in muscles requires cooperation between two KASH proteins and microtubules', *The Journal of Cell Biology* **198**(5), 833–846.

Falconer and Mackay (1996), *Introduction to Quantitative Genetics*, Oliver and Boyd, Edinburgh/London.

Findsen, A., Pedersen, T. H., Petersen, A. G., Nielsen, O. B. and Overgaard, J. (2014), 'Why do insects enter and recover from chill coma? Low temperature and high extracellular potassium compromise muscle function in Locusta migratoria', *Journal of Experimental Biology* **217**(8), 1297–1306.

Fraser, M. J., Smith, G. E. and Summers, M. D. (1983), 'Acquisition of Host Cell DNA Sequences by Baculoviruses: Relationship Between Host DNA Insertions and FP Mutants of Autographa californica and Galleria mellonella Nuclear Polyhedrosis Viruses', *Journal of Virology* **47**(2), 287–300.

Free, J. B. and Spencer-Booth, Y. (1960), 'Chill-Coma and Cold Death Temperatures of Apis Mellifera', *Entomologia Experimentalis et Applicata* **3**(3), 222–230.

Futschik, A. and Schlötterer, C. (2010), 'The next generation of molecular markers from massively parallel sequencing of pooled DNA samples', *Genetics* **186**(1), 207–18.

Gaj, T., Gersbach, C. A. and Barbas, C. F. (2013), 'ZFN, TALEN and CRISPR/Cas-based methods for genome engineering', *Trends in biotechnology* **31**(7), 397–405.

Garland, M. A., Stillman, J. H. and Tomanek, L. (2015), 'The proteomic response of cheliped myofibril tissue in the eurythermal porcelain crab Petrolisthes cinctipes to heat shock following acclimation to daily temperature fluctuations', *The Journal of Experimental Biology* **218**(Pt 3), 388–403.

Gerken, A. R., Eller, O. C., Hahn, D. A. and Morgan, T. J. (2015), 'Constraints, independence, and evolution of thermal plasticity: Probing genetic architecture of long- and short-term thermal acclimation', *Proceedings of the National Academy of Sciences of the United States of America* **112**(14), 4399–4404.

Gibert, P., Moreteau, B., Pétavy, G., Karan, D. and David, J. R. (2001), 'Chill-coma tolerance, a major climatic adaptation among Drosophila species', *Evolution; International Journal of Organic Evolution* **55**(5), 1063–1068.

Gompel, N. and Schröder, E. A. (2015), 'Drosophila germline transformation'. Available online at: http://gompel.org/wp-content/uploads/2015/12/Drosophila-transformation-with-chorion.pdf.

Goto, S. (2000), 'Expression of Drosophila homologue of senescence marker protein-30 during cold acclimation', *Journal of insect physiology* **46**, 1111–1120.

Gramates, L. S., Marygold, S. J., Santos, G. d., Urbano, J.-M., Antonazzo, G., Matthews, B. B., Rey, A. J., Tabone, C. J., Crosby, M. A., Emmert, D. B., Falls, K., Goodman, J. L., Hu, Y., Ponting, L., Schroeder, A. J., Strelets, V. B., Thurmond, J. and Zhou, P. (2017), 'FlyBase at 25: looking to the future', *Nucleic Acids Research* **45**(D1), D663–D671.

Grath, S. (2010), 'Molecular Evolution of Sex-Biased Genes in Drosophila ananassae'. Dissertation, LMU München: Fakultät für Biologie. https://edoc.ub.uni-muenchen.de/11651/.

Gratz, S. J., Cummings, A. M., Nguyen, J. N., Hamm, D. C., Donohue, L. K., Harrison, M. M., Wildonger, J. and O'Connor-Giles, K. M. (2013), 'Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease', *Genetics* **194**(4), 1029–1035.

Gratz, S. J., Rubinstein, C. D., Harrison, M. M., Wildonger, J. and O'Connor-Giles, K. M. (2015), 'CRISPR-Cas9 genome editing in Drosophila', *Current protocols in molecular biology* **111**, 31.2.1–31.2.20.

Gratz, S. J., Ukken, F. P., Rubinstein, C. D., Thiede, G., Donohue, L. K., Cummings, A. M. and O'Connor-Giles, K. M. (2014), 'Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in Drosophila', *Genetics* **196**(4), 961–971.

Hallas, R., Schiffer, M. and Hoffmann, A. A. (2002), 'Clinal variation in Drosophila serrata for stress resistance and body size', *Genetics Research* **79**(2), 141–148.

Handler, A. M. (2002), 'Use of the piggybac transposon for germ-line transformation of insects', *Insect Biochemistry and Molecular Biology* **32**, 1211–1220.

Handler, A. M. and Atkinson, P. W. (2006), 'Insect transgenesis: mechanisms, applications, and ecological safety', *Biotechnology & Genetic Engineering Reviews* **23**, 129–156.

Henderson, A. (1978), 'Spontaneous Chromosome Breakage At Male Meiosis Associated With Male Recombination In Drosophila Melanogaster', **88**, 93–107.

Hendrick, J. P. and Hartl, F.-U. (1993), 'Molecular Chaperone Functions of Heat-Shock Proteins', *Annual Review of Biochemistry* **62**(1), 349–384.

Hinton, C. W. (1974), An Extrachromosomal Suppressor of Male Crossing-Over in Drosophila ananassae, *in* R. F. Grell, ed., 'Mechanisms in Recombination', Springer, Boston, MA, pp. 391–397.

Hoekstra, H. E. and Coyne, J. A. (2007), 'The locus of evolution: evo devo and the genetics of adaptation', *Evolution; International Journal of Organic Evolution* **61**(5), 995–1016.

Hoffmann, A. A. and Weeks, A. R. (2007), 'Climatic selection on genes and traits after a 100 year-old invasion: a critical look at the temperate-tropical clines in Drosophila melanogaster from eastern Australia', *Genetica* **129**(2), 133–147.

Hoffmann, A., Anderson, A. and Hallas, R. (2002), 'Opposing clines for high and low temperature resistance in Drosophila melanogaster', *Ecology Letters* **5**(5), 614–618.

Holloway, J. (1997), 'Sundaland, Sulawesi and eastwards: a zoogeographic perspective.', *Malayan Nature Journal* **50**, 207–227.

Horn, C., Jaunich, B. and Wimmer, E. A. (2000), 'Highly sensitive, fluorescent transformation marker for Drosophila transgenesis', *Development Genes and Evolution* **210**(12), 623–629.

Hosler, J. S., Burns, J. E. and Esch, H. E. (2000), 'Flight muscle resting potential and species-specific differences in chill-coma', *Journal of Insect Physiology* **46**(5), 621–627.

Housden, B. E., Valvezan, A. J., Kelley, C., Sopko, R., Hu, Y., Roesel, C., Lin, S., Buckner, M., Tao, R., Yilmazel, B., Mohr, S. E., Manning, B. D. and Perrimon, N. (2015), 'Identification of potential drug targets for tuberous sclerosis complex by synthetic screens combining CRISPR-based knockouts with RNAi', *Science Signaling* **8**(393), rs9.

Huang, D. W., Sherman, B. T. and Lempicki, R. A. (2009), 'Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources', *Nature Protocols* **4**(1), 44–57.

Huylmans, A. K. and Parsch, J. (2014), 'Population- and Sex-Biased Gene Expression in the Excretion Organs of Drosophila melanogaster', *G3: Genes|Genomes|Genetics* **4**(12), 2307–2315.

Jansen, R. C. and Stam, P. (1994), 'High resolution of quantitative traits into multiple loci via interval mapping', *Genetics* **136**(4), 1447–1455.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. and Charpentier, E. (2012), 'A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity', *Science* **337**(6096), 816–821.

Johnson, D. S. (1964), 'A question of nomenclature.', *Malayan Nature Journal* **18**, 68–69.

Kale, P. G. (1969), 'The Meiotic Origin of Spontaneous Crossovers in Drosophila Ananassae Males', *Genetics* **62**(1), 123–133.

Kawasaki, F., L. Koonce, N., Guo, L., Fatima, S., Qiu, C., T. Moon, M., Zheng, Y. and W. Ordway, R. (2016), 'Small heat shock protein-mediated cell-autonomous and nonautonomous protection in a Drosophila model for environmental stress-induced degeneration', *Disease Models & Mechanisms* **9**(9), 953–964.

Kelty, J. D., Killian, K. and E. LEE, R. (1996), 'Cold shock and rapid cold-hardening of pharate adult flesh flies (Sarcophaga crassipalpis): Effects on behaviour and neuromuscular function following eclosion', *Physiological Entomology* **21**, 283–288.

Kibbe, W. A. (2007), 'OligoCalc: an online oligonucleotide properties calculator', *Nucleic Acids Research* **35**(Web Server issue), W43–46.

Kikkawa, H. (1938), 'Studies on the genetics and cytology of Drosophila ananassae', *Genetica* **20**(5), 458–516.

Kim, M., Robich, R. M., Rinehart, J. P. and Denlinger, D. L. (2006), 'Upregulation of two actin genes and redistribution of actin during diapause and cold stress in the northern house mosquito, Culex pipiens.', *Journal of insect physiology* **52**(11-12), 1226–1233.

Kimura, M. T., Yoshida, K. M. and Goto, S. G. (1998), 'Accumulation of Hsp70 mRNA under environmental stresses in diapausing and nondiapausing adults of Drosophila triauraria', *Journal of Insect Physiology* **44**(10), 1009–1015.

Kobey, R. L. and Montooth, K. L. (2013), 'Mortality from desiccation contributes to a genotype–temperature interaction for cold survival in Drosophila melanogaster', *Journal of Experimental Biology* **216**(7), 1174–1182.

Kohler, R. E. (1994), *Lords of the Fly: Drosophila Genetics and the Experimental Life*, University of Chicago Press.

Kondo, S. and Ueda, R. (2013), 'Highly Improved Gene Targeting by Germline-Specific Cas9 Expression in Drosophila', *Genetics* **195**(3), 715–721.

Kryštofová, S. (2016), 'CRISPR/Cas in genome defense and gene editing', *Acta Chimica Slovaca* **9**(1), 68–74.

Königer, A. and Grath, S. (2018), 'Transcriptome Analysis Reveals Candidate Genes for Cold Tolerance in Drosophila ananassae', *Genes* **9**(12), 624.

Lander, E. S. and Botstein, D. (1989), 'Mapping Mendelian Factors Underlying Quantitative Traits Using RFLP Linkage Maps', *Genetics* **121**, 185–199.

Lee, S. T. M., Keshavmurthy, S., Fontana, S., Takuma, M., Chou, W.-H. and Chen, C. A. (2018), 'Transcriptomic response in Acropora muricata under acute temperature stress follows preconditioned seasonal temperature fluctuations', *BMC research notes* **11**(119), doi: 10.1186/s13104–018–3230–z.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and 1000 Genome Project Data Processing Subgroup (2009), 'The Sequence Alignment/Map format and SAMtools', *Bioinformatics* **25**(16), 2078–2079.

Li, M., Turner, D. J., Ning, Z., Yusa, K., Liang, Q., Eckert, S., Rad, L., Fitzgerald, T. W., Craig, N. L. and Bradley, A. (2011), 'Mobilization of giant piggyBac transposons in the mouse genome', *Nucleic Acids Research* **39**(22), e148.

Liefting, M., Hoffmann, A. A. and Ellers, J. (2009), 'Plasticity versus environmental canalization: population differences in thermal responses along a latitudinal gradient in Drosophila serrata', *Evolution; International Journal of Organic Evolution* **63**(8), 1954–1963.

Liu, J., Li, C., Yu, Z., Huang, P., Wu, H., Wei, C., Zhu, N., Shen, Y., Chen, Y., Zhang, B., Deng, W.-M. and Jiao, R. (2012), 'Efficient and specific modifications of the Drosophila genome by means of an easy TALEN strategy', *Journal of Genetics and Genomics* **39**(5), 209–215.

Love, M. I., Huber, W. and Anders, S. (2014), 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biology* **15**(12), doi: 10.1186/s13059–014–0550–8.

MacMillan, H. A., Andersen, J. L., Davies, S. and Overgaard, J. (2015a), 'The capacity to maintain ion and water homeostasis underlies interspecific variation in Drosophila cold tolerance', *Scientific Reports* **5**, 18607.

MacMillan, H. A., Andersen, J. L., Loeschcke, V. and Overgaard, J. (2015b), 'Sodium distribution predicts the chill tolerance of *Drosophila melanogaster* raised in different thermal conditions', *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* **308**(10), R823–R831.

MacMillan, H. A., Knee, J. M., Dennis, A. B., Udaka, H., Marshall, K. E., Merritt, T. J. S. and Sinclair, B. J. (2016), 'Cold acclimation wholly reorganizes the Drosophila melanogaster transcriptome and metabolome', *Scientific Reports* **6**, 28999.

Macmillan, H. A. and Sinclair, B. J. (2011), 'Mechanisms underlying insect chill-coma', *Journal of Insect Physiology* **57**(1), 12–20.

Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., Geer, R. C., He, J., Gwadz, M., Hurwitz, D. I., Lanczycki, C. J., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., Zheng, C. and Bryant, S. H. (2015), 'CDD: NCBI's conserved domain database', *Nucleic Acids Research* **43**(Database issue), D222–226.

Miles, C. and Wayne, M. (2008), 'Quantitative trait locus (QTL) analysis', *Nature Education* **1**(1), 208.

Morgan, T. H. (1910), 'Sex Limited Inheritance In Drosophila', *Science* **32**(812), 120–122.

Morin, J. P., Moreteau, B., Pétavy, G., Parkash, R. and David, J.R. (1997), 'Reaction norms of morphological traits in drosophila: adaptive shape changes in a stenotherm circumtropical species?', *Evolution* **51**(4), 1140–1148.

Morrow, G. and Tanguay, R. M. (2015), 'Drosophila melanogaster Hsp22: a mitochondrial small heat shock protein influencing the aging process', *Frontiers in Genetics* **6**, doi: 10.3389/fgene.2015.00103.

Mullins, R. D., Heuser, J. A. and Pollard, T. D. (1998), 'The interaction of Arp2/3 complex with actin: Nucleation, high affinity pointed end capping, and formation of branching networks of filaments', *Proceedings of the National Academy of Sciences* **95**(11), 6181–6186.

Mußgnug, S. (2018), 'Design and Construction of Vectors for CRISPR/Cas9- Mediated Knock-Out of Cold Tolerance Candidate Genes in Drosophila ananassae'. Master thesis, LMU München: Fakultät für Biologie.

Nielsen, M. M., Overgaard, J., Sørensen, J. G., Holmstrup, M., Justesen, J. and Loeschcke, V. (2005), 'Role of HSF activation for resistance to heat, cold and high-temperature knock-down', *Journal of Insect Physiology* **51**(12), 1320–1329.

Nilson, T. L., Sinclair, B. J. and Roberts, S. P. (2006), 'The effects of carbon dioxide anesthesia and anoxia on rapid cold-hardening and chill coma recovery in Drosophila melanogaster', *Journal of insect physiology* **52**(10), 1027–1033.

Norry, F. M., Scannapieco, A. C., Sambucetti, P., Bertoli, C. I. and Loeschcke, V. (2008), 'QTL for the thermotolerance effect of heat hardening, knockdown resistance to heat

and chill-coma recovery in an intercontinental set of recombinant inbred lines of Drosophila melanogaster', *Molecular Ecology* **17**(20), 4570–4581.

Oervar, B. L., Sangwan, V., Omann, F. and Dhindsa, R. S. (2000), 'Early steps in cold sensing by plant cells: the role of actin cytoskeleton and membrane fluidity', *The Plant Journal* **23**(6), 785–794.

Overgaard, J. and MacMillan, H. (2016), 'The Integrative Physiology of Insect Chill Tolerance', *Annual review of physiology* **79**, 187–208.

Paaby, A. B. and Rockman, M. V. (2013), 'The many faces of pleiotropy', *Trends in genetics* **29**(2), 66–73.

Paris, M., Kaplan, T., Li, X. Y., Villalta, J. E., Lott, S. E. and Eisen, M. B. (2013), 'Extensive Divergence of Transcription Factor Binding in Drosophila Embryos with Highly Conserved Gene Expression', *PLoS Genetics* **9**(9), e1003748.

Parsell, D. A. and Lindquist, S. (1993), 'The Function of Heat-Shock Proteins in Stress Tolerance: Degradation and Reactivation of Damaged Proteins', *Annual Review of Genetics* **27**(1), 437–496.

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S. and Hoekstra, H. E. (2012), 'Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species', *PLOS ONE* **7**(5), e37135.

Plantampa, C., Salort, K., P., G., Dumet, A., Mialdea, G., Mondy, N. and Voituron, B. (2016), 'All or nothing: Survival, reproduction and oxidative balance in Spotted Wing Drosophila (Drosophila suzukii) in response to cold', *Journal of Insect Physiology* **89**, 28–36.

Pollard, D. (2012), 'Design and Construction of Recombinant Inbred Lines', *Methods in molecular biology* **871**, 31–9.

Port, F., Chen, H.-M., Lee, T. and Bullock, S. L. (2014), 'Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila', *Proceedings of the National Academy of Sciences of the United States of America* **111**(29), E2967–E2976.

Poxleitner, M. (2010), 'Local adaptation to cold in Drosophila ananassae'. Bachelor thesis, LMU München: Fakultät für Biologie.

Quinn, P. J. (1985), 'A lipid-phase separation model of low-temperature damage to biological membranes', *Cryobiology* **22**(2), 128–146.

R Core Team (2018), 'R: a language and environment for statistical computing'. Available online at https://www.R-project.org/.

Ramnarine, T. J. S., Glaser-Schmitt, A., Catalán, A. and Parsch, J. (2019), 'Population Genetic and Functional Analysis of a cis-Regulatory Polymorphism in the Drosophila melanogaster Metallothionein A gene', *Genes* **10**(2), 147.

Ren, X., Sun, J., Housden, B. E., Hu, Y., Roesel, C., Lin, S., Liu, L.-P., Yang, Z., Mao, D., Sun, L., Wu, Q., Ji, J.-Y., Xi, J., Mohr, S. E., Xu, J., Perrimon, N. and Ni, J.-Q. (2013), 'Optimized gene editing technology for Drosophila melanogaster using germ line-specific Cas9', *Proceedings of the National Academy of Sciences of the United States of America* **110**(47), 19012–19017.

Rinehart, J. P., Li, A., Yocum, G. D., Robich, R. M., Hayward, S. A. L. and Denlinger, D. L. (2007), 'Up-regulation of heat shock proteins is essential for cold survival during insect diapause', *Proceedings of the National Academy of Sciences of the United States of America* **104**(27), 11130–11137.

Rinehart, J. P., Robich, R. M. and Denlinger, D. L. (2006), 'Enhanced Cold and Desiccation Tolerance in Diapausing Adults of Culex pipiens, and a Role for Hsp70 in Response to Cold Shock but Not as a Component of the Diapause Program', *Journal of Medical Entomology* **43**(4), 713–723.

Ring, R. A. (1982), 'Freezing-tolerant insects with low supercooling points', *Comparative Biochemistry and Physiology Part A: Physiology* **73**(4), 605–612.

Romero, I. G., Ruvinsky, I. and Gilad, Y. (2012), 'Comparative studies of gene expression and the evolution of gene regulation', *Nature reviews. Genetics* **13**(7), 505–516.

Ross, J., Jiang, H., Kanost, M. R. and Wang, Y. (2003), 'Serine proteases and their homologs in the Drosophila melanogaster genome: an initial analysis of sequence conservation and phylogenetic relationships', *Gene* **304**, 117–131.

Rubin, G. M. and Spradling, A. C. (1982), 'Genetic transformation of Drosophila with transposable element vectors', *Science* **218**(4570), 348–353.

Russo, C. A., Takezaki, N. and Nei, M. (1995), 'Molecular phylogeny and divergence times of drosophilid species', *Molecular Biology and Evolution* **12**(3), 391–404.

Ruttkay-Nedecky, B., Nejdl, L., Gumulec, J., Zitka, O., Masarik, M., Eckschlager, T., Stiborová, M., Adam, V. and Kizek, R. (2013), 'The Role of Metallothionein in Oxidative Stress', *International Journal of Molecular Sciences* **14**, 6044–6066.

Sedlazeck, F. J., Rescheneder, P. and von Haeseler, A. (2013), 'NextGenMap: fast and accurate read mapping in highly polymorphic genomes', *Bioinformatics* **29**(21), 2790–2791.

Sejerkilde, M., Sørensen, J. G. and Loeschcke, V. (2003), 'Effects of cold- and heat hardening on thermal resistance in Drosophila melanogaster', *Journal of Insect Physiology* **49**(8), 719–726.

Sen, S. and Churchill, G. A. (2001), 'A statistical framework for quantitative trait mapping.', *Genetics* **159**(1), 371–387.

Sformo, T., Walters, K., Jeannet, K., Wowk, B., Fahy, G. M., Barnes, B. M. and Duman, J. G. (2010), 'Deep supercooling, vitrification and limited survival to -100°}C in the Alaskan beetle Cucujus clavipes puniceus (Coleoptera: Cucujidae) larvae', *The Journal of Experimental Biology* **213**(3), 502–509.

Sinclair, B. J. (1999), 'Insect cold tolerance: How many kinds of frozen?', *European Journal of Entomology* **96**(2), 157–164.

Sinclair, B. J., Addo-Bediako, A. and Chown, S. L. (2003), 'Climatic variability and the evolution of insect freeze tolerance', *Biological Reviews of the Cambridge Philosophical Society* **78**(2), 181–195.

Sinclair, B. J., Gibbs, A. G. and Roberts, S. P. (2007), 'Gene transcription during exposure to, and recovery from, cold and desiccation stress in Drosophila melanogaster', *Insect Molecular Biology* **16**(4), 435–443.

Singh, B. N. (1983), 'An inversion within the subterminal inversion inDrosophila ananassae', *Experientia* **39**(1), 99–100.

Singh, B. N. (1984), 'High frequency of cosmopolitan inversions in natural populations of Drosophila ananassae from Kerala, South India', *Journal of Heredity* **75**(6), 506–507.

Singh, B. N. and Singh, P. (2010), 'Population Genetics of Drosophila ananassae: Evidence for Population Sub-Structuring at the Level of Inversion Polymorphism in Indian Natural Populations', *International Journal of Biology* **2**(1), 19–28.

Singh, B. N. and Yadav, J. P. (2015), 'Status of research on Drosophila ananassae at global level', *Journal of Genetics* **94**(4), 785–792.

Singh, P. and Singh, B. N. (2007), 'Population genetics of Drosophila ananassae: genetic differentiation among Indian natural populations at the level of inversion polymorphism', *Genetics Research* **89**(4), 191–199.

Sisodia, S. and Singh, B. N. (2010), 'Influence of developmental temperature on cold shock and chill coma recovery in Drosophila ananassae: Acclimation and latitudinal variations among Indian populations', *Journal of Thermal Biology* **35**(3), 117–124.

Sisodia, S. and Singh, B. N. (2012), 'Experimental Evidence for Nutrition Regulated Stress Resistance in Drosophila ananassae', *PLoS ONE* **7**(10), e46131.

Spradling, A. C. and Rubin, G. M. (1982), 'Transposition of cloned P elements into Drosophila germ line chromosomes', *Science* **218**(4570), 341–347.

Stephan, W. and Langley, C. H. (1989), 'Molecular Genetic Variation in the Centromeric Regionof the X Chromosome in Three Drosophila ananassae Populations. I. Contrasts Between the vermilion andforked Loci', *Genetics* **121**(1), 89–99.

Stephan, W. and Li, H. (2007), 'The recent demographic and adaptive history of Drosophila melanogaster', *Heredity* **98**(2), 65–68.

Stern, D. L., Crocker, J., Ding, Y., Frankel, N., Kappes, G., Kim, E., Kuzmickas, R., Lemire, A., Mast, J. D. and Picard, S. (2017), 'Genetic and Transgenic Reagents for Drosophila simulans, D. mauritiana, D. yakuba, D. santomea, and D. virilis', *G3: Genes, Genomes, Genetics* **7**(4), 1339–1347.

Storey, K. B. and Storey, J. M. (2012), 'Insect cold hardiness: metabolic, gene, and protein adaptation', *Canadian Journal of Zoology* **90**(4), 456–475.

Stětina, T., Koštál, V. and Korbelová, J. (2015), 'The Role of Inducible Hsp70, and Other Heat Shock Proteins, in Adaptive Complex of Cold Tolerance of the Fruit Fly (Drosophila melanogaster)', *PloS One* **10**(6), e0128976.

Supek, F., Bošnjak, M., Škunca, N. and Šmuc, T. (2011), 'REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms', *PLoS ONE* **6**(7), e21800.

Teets, N. and Denlinger, D. (2013), 'Physiological mechanisms of seasonal and rapid cold-hardening in insects', *Physiological Entomology* **38**, 105–116.

Teets, N. M., Peyton, J. T., Colinet, H., Renault, D., Kelley, J. L., Kawarasaki, Y., Lee, R. E. and Denlinger, D. L. (2012), 'Gene expression changes governing extreme dehydration tolerance in an Antarctic insect', *Proceedings of the National Academy of Sciences* **109**(50), 20744–20749.

The FlyBase Consortium (2009), 'FlyBase: enhancing Drosophila Gene Ontology annotations', *Nucleic Acids Research* **37**(Database issue), 7D555–D559.

Throckmorton et al. (1975), *The phylogeny, ecology, and geography of Drosophila*, Plenum Press, New York.

Tobari, Y. N. (1993), *Drosophila ananassae: genetical and biological aspects*, Japan Scientific Societies Press, Tokyo.

Udaka, H., Ueda, C. and Goto, S. G. (2010), 'Survival rate and expression of Heat-shock protein 70 and Frost genes after temperature stress in Drosophila melanogaster lines that are selected for recovery time from temperature coma', *Journal of Insect Physiology* **56**(12), 1889–1894.

Venken, K. T. J. and Bellen, H. J. (2007), 'Transgenesis upgrades for Drosophila melanogaster', *Development* **134**, 3571–3584.

Villa, J. D. and Rinderer, T. E. (1993), 'Cold room thermoregulation, store consumption, and survival of Africanized and European honey bees (Apis mellifera)', *Apidologie* **24**(4), 411–423.

von Heckel, K., Stephan, W. and Hutter, S. (2016), 'Canalization of gene expression is a major signature of regulatory cold adaptation in temperate Drosophila melanogaster', *BMC Genomics* **17**(1), 574.

Voris, H. (2000), 'Maps of Pleistocene sea levels in Southeast Asia: Shorelines, river systems and time durations', *Journal of Biogeography* **27**, 1153–1167.

Wang, C., Zeng, Z., Liu, Q., Zhang, R. and Ni, J. (2014), 'Se-methylselenocysteine inhibits apoptosis induced by clusterin knockdown in neuroblastoma N2a and SH-SY5y cell lines., Se-Methylselenocysteine Inhibits Apoptosis Induced by Clusterin Knockdown in Neuroblastoma N2a and SH-SY5y Cell Lines', *International journal of molecular sciences, International Journal of Molecular Sciences* **15**(11), 21331–21347.

Wang, Z., Gerstein, M. and Snyder, M. (2009), 'RNA-Seq: a revolutionary tool for transcriptomics', *Nature reviews. Genetics* **10**(1), 57–63.

Warren, R. J. and Chick, L. (2013), 'Upward ant distribution shift corresponds with minimum, not maximum, temperature tolerance', *Global Change Biology* **19**(7), 2082–2088.

Williams, C., Watanabe, M., Guarracino, M., Ferraro, M., Edison, A., Morgan, T., Boroujerdi, A. and Hahn, D. (2014), 'Cold adaptation shapes the robustness of metabolic networks in Drosophila melanogaster', *Evolution* **68**(12), 3505–3523.

Wu, J.-Y., Jin, C., Qu, H.-Y., Tao, S.-T., Xu, G.-H., Wu, J., Wu, H.-Q. and Zhang, S.-L. (2012), 'Low temperature inhibits pollen viability by alteration of actin cytoskeleton and regulation of pollen plasma membrane ion channels in Pyrus pyrifolia', *Environmental and Experimental Botany* **78**, 70–75.

Yi, S.-X. and Lee, R. E. (2011), 'Rapid cold-hardening blocks cold-induced apoptosis by inhibiting the activation of pro-caspases in the flesh fly Sarcophaga crassipalpis', *Apoptosis* **16**(3), 249–255.

Yi, S.-X., Moore, C. W. and Lee, R. E. (2007), 'Rapid cold-hardening protects Drosophila melanogaster from cold-induced apoptosis', *Apoptosis* **12**(7), 1183–1193.

Yuan, S., Yu, X., Topf, M., Dorstyn, L., Kumar, S., Ludtke, S. J. and Akey, C. W. (2011), 'Structure of the Drosophila apoptosome at 6.9Å resolution', *Structure* **19**(1), 128–140.

Acknowledgements

First of all, I would like to thank Sonja Grath for her excellent supervision and support during these last years. I could not have hoped for a better mentor. You gave me the opportunity to work independently, but were always there when I needed you - literally at any time! Your valuable feedback and comments on my presentations, abstracts, manuscripts and projects made me a better scientist.
I am grateful to John Parsch and Dirk Metzler who contributed with their ideas and feedback as constant members of my Thesis Advisory Committee significantly to the progress of this thesis. Further, I would like to thank Saad Arif for his support with the QTL mapping project and Wolfgang Enard, Stephan Hutter and Richard Merrill for their valuable comments during my Thesis Advisory Committee meetings.
I thank Karolina Worf for introducing me to the world of programming, Saurabh Pophaly and Dietfried Molter for sharing their expertise in MySQL and Linux.
About a dozen Bachelor and Master students have contributed to a greater or lesser extent with their thesis or research projects to the progress of my dissertation. While I am certainly grateful to all of them, I would like to highlight three persons: Selina Mußgnug, Eslam Katab and Timothy Ramnarine. Tim and Selina worked with me on the genome engineering project and Eslam carried out a lot of DNA extractions and PCRs. I enjoyed working with all of you and I wish you the best for your own PhDs!

I would like to thank Alistair McGregor, Sebastian Kittelman, Pedro Gaspar, Michaela Holzem and Franziska Franke for welcoming me in their group at Oxford brookes university and for sharing their knowledge about Crispr/Cas9.

I want to emphasize that I consider myself extremely lucky for having conducted my research in a department with a very friendly and supportive atmosphere. I attribute this in particular to Hilde, Gabi, Gisi, Ana and Ingrid. Special thanks go to Hilde for her

constant support in technical matters and fly husbandry.
I shared an office with the most awesome colleagues I could have wished for: Ann Kathrin Huylmans, Eliza Argyridou and Amanda Glaser-Schmitt. Ina, thank you for being such a nice host in Vienna! Liza, thank you for your endless words of support during this endeavor! Amanda, I very much enjoyed the two Evolution conferences that we attended together. I will always remember the bike trip through Groningen and the wine festival in Montpellier!

I thank my family and friends for their emotional support during the last years, especially Miri and Whoopie. You are the best companions!

Lastly, I thank my partner Marko and my parents Renate and Johann for their love and patience. This thesis is dedicated to the three of you, since it could not have been written without your support!