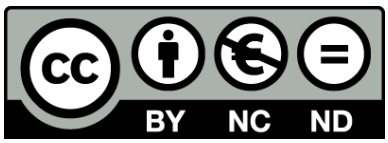
UNIVERSITAT DE BARCELONA

Adaptation in *Drosophila melanogaster* Natural Populations

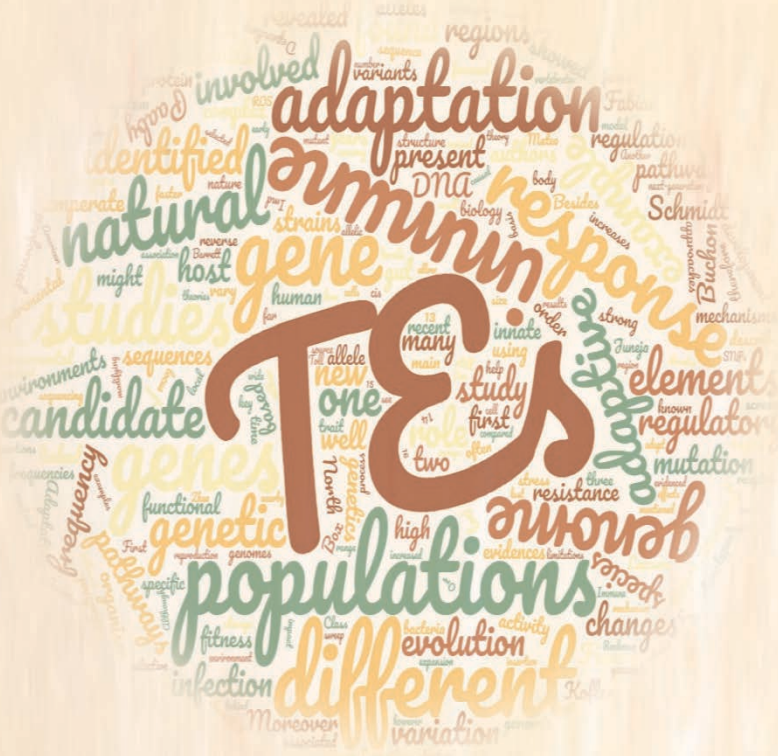
Fitness Effects and Evolutionary History of a Natural Insertion and Molecular Effects of Several Transposable Elements on Immune-Related Genes

Anna Ullastres i Coll



Adaptation in *Drosophila melanogaster* Natural Populations

Fitness Effects and Evolutionary History of a Natural Insertion and Molecular Effects of Several Transposable Elements on Immune-Related Genes



PhD Thesis

Anna Ullastres i Coll

2017

TESI DOCTORAL
UNIVERSITAT DE BARCELONA

FACULTAT DE BIOLOGIA
DEPARTAMENT DE GENÈTICA
PROGRAMA DE DOCTORAT EN GENÈTICA

Adaptation in *Drosophila melanogaster* Natural Populations: Fitness Effects and Evolutionary History of a Natural Insertion and Molecular Effects of Several Transposable Elements on Immune-Related Genes

"Adaptació en Poblacions Naturals de *Drosophila melanogaster*. Efectes en la *fitness* i Història Evolutiva d'una Inserció Natural i Efectes Moleculars dels Elements Mòbils en Gens Relacionats amb la Resposta Immune"

Memòria presentada per Anna Ullastres i Coll per a optar al grau de doctora per la Universitat de Barcelona

Aquest treball ha estat realitzat sota la supervisió de la Dra. Josefa González a l'Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra)

PhD candidate	**Tutor**	**Thesis advisor**
Anna Ullastres i Coll	Dr. Francesc Mestres Naval	Dra. Josefa González Pérez

Barcelona, Abril 2017
Institut de Biologia Evolutiva

INDEX

01

INTRODUCTION

1.1. Adaptation

Adaptation can be defined as the process that drives frequency changes in alleles involved in fitness increasing in a specific environment, thus increasing organism survival and/or reproduction (Haldane 1932; Dobzhansky 1951; Dobzhansky 1955; Orr 2005; Barker 2009; Orr 2009). Despite the numerous efforts in the evolutionary biology field to understand the process of adaptation, we still hold important unanswered questions about the mechanisms behind the process of adaptation. For example: What kind of mutations are involved in adaptation? How many mutations are needed to produce an adaptive phenotype? What are the traits under selection during this process? What is the role of epistasis and pleiotropy in adaptation?

1.1.1 The Study of Adaptation: A Historical Perspective

C. R. Darwin and A. R. Wallace introduced natural selection as a key driver in evolutionary adaptation to natural environments (1858). However, natural selection theory was rejected by most of the contemporary authors of the field, as well as during the following years, when other theories, such as Neo-Lamarckism and mutationist theories rediscovering Mendelian's laws, were better considered. It was not until 1930s that new theories of adaptation were elaborated considering natural selection as a key driver. R.A. Fisher, J.B.S. Haldane and S. Wright leaded the *modern synthesis*, which combined the idea of adaptation happening by mutations with the theory of natural selection (Fisher 1930; Wright 1931; Haldane 1932). These authors reconciled Darwin and Mendel's ideas by developing a mathematical theory of population genetics. Later on,

M. Kimura postulated the *neutral theory of molecular evolution* (1955), and proposed that molecular evolution is dominated by selectively neutral evolution, and that changes in allele frequencies are mainly because of genetic drift (Kimura 1991). Thus, he provided a foundation for discerning adaptive mutations from neutral mutations, and, therefore, for detecting the effects of natural selection on the DNA sequence. These authors, from C. R. Darwin to M. Kimura, established the essential evolutionary forces: natural selection, mutation, genetic drift, recombination, and gene flux. After that, other evolutionary biologists, such as T. Ohta and A. Orr, have been further elaborating on several mathematical models trying to predict how adaptation occurs in nature (Ohta 1973; Orr 2005).

The first experimental evidences that started to shed light on these theories were based on the identification of allozymes with electrophoretic studies. Allozymes are proteins resulting from allelic differences that can be distinguished by their electrophoretic

mobility. The initial studies in *Drosophila pseudobscura* (Lewontin and Hubby 1966) and in human (Harris 1966) revealed that populations harbored high levels of genetic variation. These levels of polymorphisms were higher than predicted by evolutionary biology theorists. A well-known example of natural selection acting on allozyme variants is the β-hemoglobin polymorphism in some African and Mediterranean human populations. This polymorphism is maintained by natural selection favoring heterozygote individuals, who are more resistant to malaria infections (Cavalli-Sforza and Bodmer 1971).

In the 1980s, new experimental approaches based on nucleotide polymorphisms allowed to find more evidences to understand how natural selection acts on genome evolution. The firsts experimental evidences from quantitative-trait *loci* (QTL), experimental evolution, and candidate genes further contributed to the understanding of the genetic basis of adaptation. These studies revealed some adaptations caused by large-effect alleles, such as mutations conferring insecticide resistance in dipterans (Guerrero et al. 1997), but failed detecting small-effect mutations that could potentially be involved in complex trait evolution (Rockman 2012).

More recently, next-generation sequencing has transformed our ability to identify alleles behind adaptation (Stapley et al. 2010). Whole-genome sequencing and genome-wide studies showed that adaptation can also occur through small-effect mutations (Collins and de Meaux 2009), and that it often results from small frequency changes in multiple alleles (Hancock et al. 2010). However, most of the alleles identified in genome-wide studies only contribute to a very small fraction of the studied phenotype (Rockman 2012). Moreover, only a small proportion of the candidate alleles are replicated comparing different studies, even within the same populations (Hong 2012).

Later on, several authors pointed out that researchers should integrate different areas of knowledge, such as population genomics, ecology and development, in order to uncover the genetic basis behind adaptation and phenotype evolution (Stern and Orgogozo 2008; Olson-Manning et al. 2012; Lowe et al. 2017).

1.1.2 Identifying the Genetic Basis of Adaptation

Three main approaches are followed to study adaptation in the evolutionary biology field: forward genetics, reverse genetics, and candidate gene (Barrett and Hoekstra 2011; Pardo-Diaz et al. 2015).

The **forward genetics** approach consists on looking for the genes controlling a phenotypic trait that is known to vary between environments or between individuals. Association mapping and QTL analysis are the two main methodologies followed. These methodologies have revealed many adaptive natural variants in a wide range of species

including plants, insects and vertebrates (Pardo-Diaz et al. 2015). An example of a forward genetics study identified the genetic basis of color-pattern differences in the beach mouse *Peromyscus polionotus* (Steiner et al. 2007). The authors performed a genome-wide linkage map using both microsatellite markers and SNPs in candidate genes for pigmentation. In this work, they linked a few large-effect mutations, both structural and regulatory, to the adaptive coat-color phenotype. In *Drosophila*, QTL studies have also allowed the discovery of genes playing crucial roles on different traits, such as diapause or body pigmentation (Schmidt et al. 2008; Mackay 2010; Bastide et al. 2016). Genome-wide association studies (GWAS) also revealed the genetic basis of other adaptive phenotypes in *Drosophila* such as desiccation resistance (Telonis-Scott et al. 2016), or oxidative stress resistance (Weber et al. 2012). GWAS analysis have also revealed some of the genes behind human adaptations, such as genes governing innate immunity (Deschamps et al. 2016), or the variants involved in high-altitude adaptation (Bigham and Lee 2014).

One of the limitations of the forward genetic approaches is that the identified locus could be in linkage disequilibrium with the actual causative adaptive mutation (Barrett and Hoekstra 2011). Another limitation is that forward genetics approaches are often biased towards the detection of large-effect *loci* and, therefore, they cannot detect cases of polygenic adaptation (Rockman 2012).

The **reverse genetics** approach aims to identify putatively adaptive alleles without a prior knowledge of the associated phenotype(s). It consists on the search of candidates based on genome-wide level signatures of selection, such as selective sweeps or allele frequency changes, either in the same population or among different populations. When a mutation is positively selected, its frequency in the genome increases rapidly. This is frequently accompanied by a decrease in the genetic diversity in the flanking regions, known as a *selective sweep* (Berry et al. 1991). Depending on the nature of each mutation and the signal they leave in the genome, we can distinguish between *hard* sweeps and *soft* sweeps (Figure 1.1). Hard sweeps occur when strong positive selection increases the frequency of one adaptive allele, thus leaving a strong signature in the genome (Figure 1.1A). The classic example of a hard sweep is when a new beneficial mutation appears in a population, and it increases its frequency starting from very low frequency in a short period of time. In a soft sweep, the adaptive mutation that is under selection is present in more than one allele; thus, it leaves a week signature on the pattern of genetic variation when they are selected (Figure 1.1B). This can happen when a mutation arise from standing variation, *i.e.* it was already present at some frequency in the population, and is segregating within several haplotypes. Thus, when it becomes advantageous, several haplotypes can be selected in the same population (Messer and Petrov 2013). Adaptive

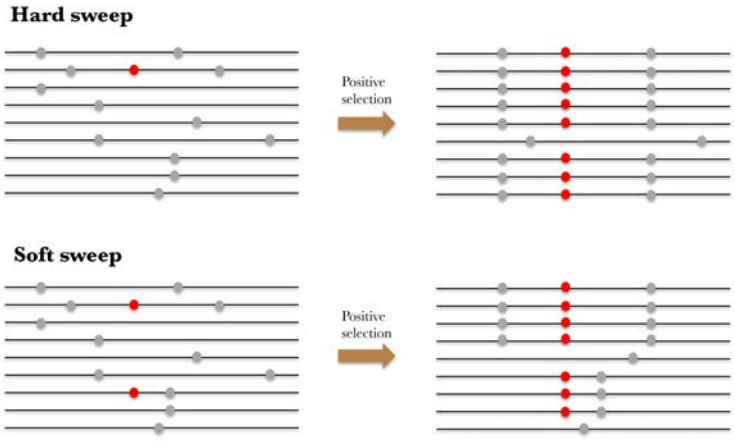


Figure 1.1. Genomic signatures of positive selection. **(A)** A new or very low frequent variant (in red) is positively selected in a population and increases its frequency, as well as the frequency of the linked neutral variants (in grey). **(B)** An advantageous variant (in red), which is linked to two different haplotypes, is positively selected. As a consequence, the two haplotypes are present at high frequency in the population.

mutations can emerge *de novo* or from standing variation. Moreover, selection can act on a single *locus* or on multiple *loci*. In some cases, adaptation in natural populations might imply subtle allele frequency changes at many *loci* controlling polygenic traits, and those changes are more complex to distinguish from genetic drift (Pritchard et al. 2010; Berg and Coop 2014). Therefore, each type of selection leaves a distinctive molecular signature in the genome. Hence, different statistic tests should be applied to detect selection depending on each particular case (Stephan 2016) (Box 1).

An example of the reverse genetics approach is the study of cichlid fishes from Lake Malawi and Lake Victoria that revealed the genetic variants implicated in visual pigment adaptation to the different waters (Hofmann et al. 2009). Also, a genome-wide screening of transposable element mutations frequencies in *Drosophila melanogaster* natural populations identified a total of 13 candidate mutations involved in adaptation of this species to out-of-Africa environments (González et al. 2008).

False positives are one of the limitations of the reverse genetics approach. Demographic factors can create patterns in the genome that are easily confounded with signals of selection. Moreover, tests looking for selection in the genome fail when selection is not strong. The use of several statistic approaches to analyze selection can help to overcome these limitations. A method usually followed to reduce false positives is the combination of population genetics with the association of environmental variables (González et al. 2010; Lowe et al. 2017).

The **candidate gene** approach is based on the knowledge of the gene adaptive function derived from a different species. Such studies are often the first step in unraveling the genetic determinants of complex human diseases (Tabor et al. 2002; Suh and Vijg 2005). Nowadays, *Drosophila* is still a key model organism for testing potential candidate genes

involved in human diseases such as the congenital heart disease (Zhu et al. 2017), or cancer (Sonoshita and Cagan 2016).

This methodology is often biased towards a few well-characterized genes with large effect (Rockman 2012). Although the knowledge of the candidate genes can help in identifying genes associated with particular phenotypes, the approach is poorly applied in evolutionary biology because, besides of the bias, it does not help in the identification of new *loci* involved in adaptation (Barrett and Hoekstra 2011; Pardo-Diaz et al. 2015).

1.1.3 Validating the Candidate Adaptive Mutations

Once the putatively adaptive *loci* are identified, it is fundamental to validate them in order to confidently claim that they are involved in adaptation (Barrett and Hoekstra 2011).

The methodologies followed to validate the role of the identified candidate mutations in adaptation can vary depending on the case under study. For example, the adaptive role of a candidate mutation affecting the protein-coding region could be traced with protein folding studies, as well as protein functional assays (Schmidt et al. 2008; Carnero-Montoro et al. 2012; Fernández-Sampedro et al. 2016). When the mutation is not interfering the protein-coding region, gene expression analysis can be performed in order to link the genetic change with a regulatory change of the nearby gene(s). Studies using microarrays and RNA-seq have produced many valuable catalogues of gene expression levels between populations (Zhao et al. 2016), or between different conditions (MacMillan et al. 2016). When studying a single strong candidate, other techniques, such as reverse-transcriptase quantitative PCR (qRT-qPCR) or *in situ* hybridization, can be used to test expression changes (Guio et al. 2014; Clemson et al. 2016). Both *cis* and *trans* regulatory elements can be involved in gene expression modifications. Thus, when observing gene expression differences between individuals, *trans*-regulatory changes could be the causal variant explaining such differences. Allele-specific expression (ASE) is a method that allows uncovering the *cis*-specific regulatory variation, as *trans*-regulation affects the expression of both alleles equally in diploid cells (Wittkopp et al. 2004). Thus, it is a strong technique to test whether a *cis*-regulatory mutation is modifying the expression of a gene.

Once the candidate variant has been linked to a transcript or protein modification, it is crucial to demonstrate that this modification has an impact on the organism fitness. This can be accomplished by performing functional assays with laboratory mutant strains generated with genome-editing techniques such as CRISPR/Cas9 (Bassett and Liu 2014). For example, Ding and collaborators demonstrated the causal *locus*, a calcium-activated potassium channel, and the causal mutation, a retroelement insertion, for courtship song differences in *Drosophila simulans* and *Drosophila mauritiana* by generating targeted deletion

Box 1. Methods for measuring selection in the genome

There are multiple statistics that can be applied in order to detect the distinctive molecular signatures that positive selection leaves in the genome (Stephan 2016). These statistics can be based on different parameters such as nucleotide diversity or linkage disequilibrium, among others (Casillas and Barbadilla 2017). Some examples of the tests applied to measure selection in the genome are: π, Tajima's D, CL, iHS, and Fst.

π

π measures the nucleotide diversity between two sequences. It is calculated as the mean number of nucleotide differences per site between two sequences (Jukes and Cantor 1969; Nei and Gojobori 1986; Nei 1987). A low nucleotide diversity of the candidate allele would indicate positive selection of that allele.

Tajima's D

Tajima's D is a neutrality test calculated as the ratio between the mean number of pairwise differences and the number of segregating sites (Tajima 1989). A ratio value of 0 would indicate that the candidate allele is segregating neutrally in the population. A negative value indicates an excess of low frequency polymorphisms in the population, which could be the consequence of an increase in the population size, because of a bottleneck or a selective sweep, and/or purifying selection. A positive value in Tajima's D test means low levels of both low and high frequency polymorphisms, indicating a decrease in the population size and/or balancing selection.

CL

The Composite Likelihood (CL) test scan the genome for regions with aberrant allele frequency distributions. CL is calculated by multiplying the marginal likelihoods for each site along the sequence (Nielsen et al. 2005). Higher CL values would indicate the presence of a selective sweep in the region analyzed. While most of the measures take as null hypothesis standard neutral models, CL test null hypothesis is derived from the background pattern of variation in the data itself (Nielsen et al. 2005). The use of the global observed frequency spectrum as the background makes CL more robust than other measures such as Tajima's D.

iHS

The integrated haplotype score (iHS) compares the frequency of derived alleles with the ancestral alleles, and measures the linkage disequilibrium (Voight et al. 2006). Values of 0 would indicate that there are no differences between the derived allele and the ancestral allele. Large negative values indicate linkage disequilibrium in the derived allele, while large positive values would indicate linkage disequilibrium in the ancestral allele. Linkage disequilibrium is found when the frequency of association of the different alleles is higher than expected if the loci were independent and associated randomly (Slatkin 2008). Different factors can influence linkage disequilibrium, including selection, recombination rate, mutation rate, and genetic drift, among others. Thus, large genomic regions with high linkage disequilibrium could be indicative of the presence of selective sweeps.

Fst

The Fixation index (Fst) is a measure to test population differentiation due to genetic structure. It calculates the average levels of gene flow based on allele frequencies (Hudson et al. 1992), often by using SNPs or microsatellites polymorphisms. Fst values range from 0 to 1, where a zero value would indicate no genetic subdivision between the populations considered. Fst is one of the main population genetics tests used in order to identify alleles involved in local adaptation.

with the CRISPR/Cas9 technique (Ding et al. 2016). Experiments performed with gene mutant strains, such as knockdowns or knockouts, can help to infer how perturbing the function of the genes associated to the candidate variants impacts on the phenotype (St Johnston 2013). However, in most cases, laboratory mutant strains tend to present extreme phenotypes rarely found in nature, and high pleiotropic effects (He 2016). Performing the assays using mutants generated in different genetic background should help to circumvent this problem (He 2016). Moreover, laboratory mutations might not be representative of the mutations segregating in the natural populations (Steiner et al. 2007; Kolaczkowski et al. 2011; Rose et al. 2011). For example, the mutations found to be involved in the color pattern in mice identified in natural populations differed from the candidate mutations identified in laboratory strains (Hoekstra et al. 2006; Steiner et al. 2007). Thus, experiments performed with natural populations, where the candidate mutation is present in its natural genomic context, might improve our understanding of phenotypic evolution (Gasch et al. 2016).

However, genotype-phenotype mapping in both mutant and natural population strains also present some caveats. First, the phenotypic effect of some mutations can only be observed under specific environmental conditions (Paaby and Schmidt 2008; Storz and Wheat 2010). Possible epistatic interactions, as well as other mutations in the genetic background tested could modify the observed phenotypes (Burnett et al. 2011; Huang et al. 2012; Chandler et al. 2013). Furthermore, one mutation can affect more than one phenotype, therefore, it can be beneficial for two different traits, or it can present tradeoffs (McGee et al. 2014). Genetic tradeoffs occur when one allele that is beneficial for one trait is deleterious for a different fitness component (Williams 1957; Edward and Chapman 2011). For example, it is well documented that an improvement in early reproduction has a physiological cost and shortens female lifespan in *Drosophila* (Partridge et al. 1999; Sgrò and Partridge 1999). Fecundity, measured as number of offspring produced per female, is also impaired in flies that are more resistant to stress such as cold or infection (Lazzaro et al. 2008; Marshall and Sinclair 2010). It was found that natural fly strains reared in the laboratory could adapt to multiple cold exposures, evidenced by a decreased mortality; however, their fecundity was significantly reduced (Marshall and Sinclair 2010). Thus, several backgrounds, as well as several phenotypes, should be analyzed to fully characterize the adaptive effects of the candidate mutations.

1.1.4 Adaptation in *Drosophila melanogaster*

Drosophila melanogaster has become an excellent model organism in functional genomic studies since the first works of T. H. Morgan in the beginning of the 20th century. Today

it has one of the best-annotated genomes (Ashburner and Bergman 2005), and there is a lot of information of the gene functions and pathways (Morgan 1911; Matthews et al. 2005; Jennings 2011; Gramates et al. 2017). Studies in *D. melanogaster* have played an important role in many areas of biology such as developmental biology, neurobiology or cell biology (Bellen et al. 2010; Jennings 2011). Moreover, there are plenty of resources to design experimental approaches to test candidate mutations: from online tools and molecular reagents, to fly stocks and genome engineering resources (Mohr et al. 2014). *D. melanogaster* is also an excellent model organism to study adaptation because of its recent demographic history (Figure 1.2).

1.1.4.1 *D. melanogaster* Natural Populations Vary Across Space and Time

D. melanogaster is a species original from subtropical Africa and just very recently, approximately 10,000-16,000 years ago, the population expanded to the Eurasian continent (David and Capy 1988; Li and Stephan 2006; Thornton and Andolfatto 2006) (Figure 1.2). Only between one hundred to a few hundred years ago, *D. melanogaster* colonized the American and Australian continents (Bock and Parsons 1981; Keller 2007). Nowadays *D. melanogaster* is a cosmopolitan species and its recent expansion suggests that the signatures of selection should still be detectable in its genome sequence (Przeworski 2002).

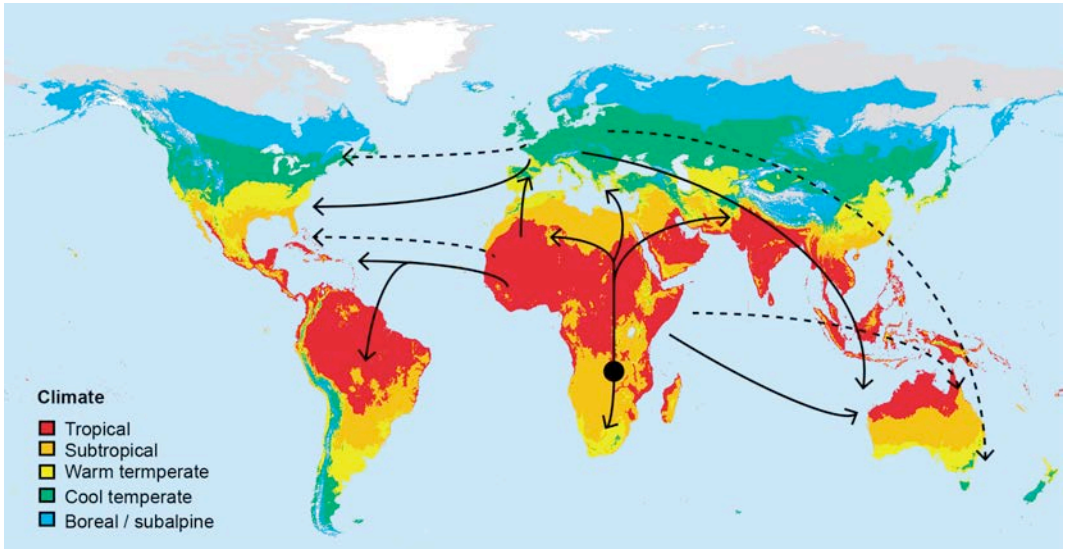


Figure 1.2. *Drosophila melanogaster* is present in almost all climatic regions. This specie is original from sub-tropical Africa, and recently has expanded to the rest of the continents (arrows). The present populations from different continents show evidences of admixture (depicted with dashed arrows).

Most of the relevant traits involved in *D. melanogaster* adaptation have been identified by comparing different geographic populations, either by using SNP-based GWAS, genome-

wide expression analyses, or by measuring life-history traits. Many of these studies analyzed natural populations from tropical and temperate climates of the east coasts from North America and Australia (Schmidt, Matzkin, et al. 2005; Kolaczkowski et al. 2011; Telonis-Scott et al. 2011; Fabian et al. 2012; Paaby et al. 2014; Reinhardt et al. 2014). Other studies compare African tropical populations and European temperate populations (Aguadé 2008; Klepsatel et al. 2014; Fabian et al. 2015; Bozicevic et al. 2016; Endler et al. 2016). However, secondary contacts among the populations from the different continents have been described that could hinder the study of allele variants showing latitudinal patterns (Caracristi and Schlötterer 2003; Duchen et al. 2013; Kao et al. 2015; Bergland et al. 2016). For example, there is admixture between temperate populations from the north of North America and Europe, and also between tropical populations from the south of North America and Africa. This could be interfering with the allelic variant frequencies of the populations resulting in latitudinal patterns (Bergland et al. 2016). The same was found in Australia, where secondary contacts with Europe and Africa would be also taking place in the populations in the extremes of the latitude (Bergland et al. 2016). While latitudinal clines could be explained by the above mentioned migration patterns, as well as by population bottlenecks, clinal variation as a consequence of the selective forces associated with the environment are still present (Przeworski 2002; Fabian et al. 2015; M. Kapun et al. 2016; Machado et al. 2016).

Besides geographical variation, *D. melanogaster* natural populations also harbor temporal variation. *D. melanogaster* inhabiting temperate environments expand their population size every spring, while they diminish their physiological and reproductive activity when environmental temperature drops and photoperiods shorten (Schmidt, Paaby, et al. 2005). Recently, it has been observed that phenotype and allele frequencies also vary seasonally in natural *D. melanogaster* populations inhabiting temperate environments (Bergland et al. 2014). Temperate populations are exposed to high levels of variation in temperature, humidity, and nutritional quality and quantity because of seasonal changes. Thus, rapid adaptation to environmental changes can be traced at the genetic level, as the allelic variants would change frequency in response to environmental conditions (Przeworski 2002; Bergland et al. 2014; Cogni et al. 2014; M. Kapun et al. 2016).

1.1.4.2 Traits Involved in *D. melanogaster* Adaptation

Some examples of classical life-history adaptive traits identified in *D. melanogaster* latitudinal population analyses are body size (Kennington et al. 2003; Paaby et al. 2010; Paaby et al. 2014; Fabian et al. 2015), female fecundity, lifespan, and developmental time (James and Partridge 1995; Schmidt, Matzkin, et al. 2005; Folguera et al. 2008; Paaby et al. 2010; Paaby et al. 2014; Fabian et al. 2015). Body size significantly increases with

latitude, as evidenced by the parallel clines found in North America and Australia (Paaby et al. 2010; Paaby et al. 2014; Fabian et al. 2015; Kapun et al. 2016). Temperature is suggested to be one of the main selective forces of body size, favoring larger body sizes in temperate climates, and smaller body sizes in tropical climates (Partridge et al. 1994). Developmental time (DT) is an especially relevant fitness trait for those organisms that occupy ephemeral habitats such as *D. melanogaster* (Chippindale et al. 1997). In nature, quick development favors *D. melanogaster* individuals for several reasons. First, larvae feed on rotting fruits that are ephemeral. Thus, quick development allows larvae to pupate before the food source is exhausted. Second, competition increases as more and more eggs are laid on a piece of fruit, also favoring individuals with faster DT (Nunney 1990). Third, breeding sites in nature can be destroyed by physical factors and predation, individuals that develop faster are thus more likely to escape microhabitat destruction. And fourth, faster DT accelerates the age of first breeding, which is relevant for the organism if most reproduction happens in expanding populations, such as *D. melanogaster* populations.

Besides life-history traits, other significant traits that have been associated with the adaptation to new environments in *D. melanogaster* are pigmentation (Telonis-Scott et al. 2011; Bastide et al. 2013; Endler et al. 2016), metabolism (Sezgin et al. 2004; Fabian et al. 2012; Lavington et al. 2014; Zhao et al. 2015; Bozicevic et al. 2016; Machado et al. 2016), circadian rhythm (Kyriacou et al. 2008; Kolaczkowski et al. 2011; Fabian et al. 2012; Zhao et al. 2015), olfaction (Aguadé 2008; Kolaczkowski et al. 2011; Reinhardt et al. 2014), and diapause (Schmidt et al. 2005; Schmidt and Paaby 2008; Zhao et al. 2015). Diapause incidence in temperate populations is an example of a key adaptation in *D. melanogaster* out-of-Africa expansion. Diapause is a period during which physiological activity is diminished. Undergoing diapause increases the probability of surviving of the populations inhabiting in temperate climates (Hand et al. 2016). Schmidt and collaborators identified an allelic variant of the gene *Couch potato* (*cpo*), which showed latitudinal differentiation in North American populations (Schmidt et al. 2008). The authors found an association of this *locus* with the arrest of ovarian development at low temperatures (Schmidt et al. 2008). Other studies in Australia and Europe have also linked *cpo* with diapause, however, they could not find association between the alleles identified by Schmith and colleagues and this trait (Lee et al. 2011; Zonato et al. 2016). These results manifest the complexity of mapping phenotypic adaptation to particular genomic variants. It is possible that other causal variants are playing a role in diapause in the different countries (Pegoraro et al. 2017).

Colonizing new environments also imply the exposure to new stressors that can be either abiotic, such as temperature, UV radiation, or precipitation; or biotic, such as new pathogens or species competition (Paaby et al. 2010; Kolaczkowski et al. 2011; Paaby et al. 2014; Bozicevic et al. 2016). There are some examples in the literature that have linked natural mutations in *Drosophila* to the resistance to abiotic stressors (Li et al. 2007). For example, insecticide resistance is improved by loss-of-function mutations leading to up-regulation of *P450* genes (Maitra et al. 2000). Xenobiotic substances can be naturally found in plants or can be synthetic compounds. A natural transposable element insertion, *FBti0019627*, was linked to both benzaldehyde and carbofuran resistance, which are natural and synthetic xenobiotic agents respectively (Mateo et al. 2014). This insertion modifies the 3'UTR structure of the gene *CG11699* and leads to increased expression levels. *CG11699* interacts with Aldh-III, the enzyme responsible for benzaldehyde metabolism. The authors also showed that increased *CG11699* expression lead to more Aldh-III enzymatic activity (Mateo et al. 2014).

Biotic factors like predators, resources competition or parasites, also affect the organism fitness. An important component of fitness in most organisms is thought to be immune defense against pathogens (Kolaczkowski et al. 2011; Levine et al. 2011; Fabian et al. 2012). Organisms have evolved a wide range of immune defense mechanisms to combat infection.

1.2. Immune Response in *Drosophila*

Most of our knowledge in innate immunity has been revealed through numerous studies in *Drosophila*. In fact, Jules A. Hoffmann received the Nobel Prize in 2011 for the discovery in *Drosophila* of the role of *Toll* gene in sensing pathogenic microorganisms and in activating the innate defense response (Lemaitre et al. 1996). The Nobel Prize was shared with Ralph Steinman and Bruce Beutler, the last one discovered *Toll*-like receptors in mammals in the light of J.A. Hoffmann results in *Drosophila* (Poltorak 1998).

The high conservation of the molecules and pathways involved in innate immune response, as well as in gut epithelium regeneration, and wound healing make possible the research with this model organism (Lemaitre and Hoffmann 2007; Buchon et al. 2014; Buchmann 2014; Bergman et al. 2016). Despite the lack of the adaptive immune response system, *Drosophila*, as the other invertebrates, has multiple innate immune response mechanisms to combat infection (Kounatidis and Ligoxygakis 2012). Thus, in order to survive infection, *Drosophila* strongly relies on both fast recognition and efficient killing of the pathogen, as well as on potent tissue regeneration systems. Innate immune response starts with the recognition of the pathogen by cell receptors, which activate the transcription of specific genes, and ends with the production of immune responsive genes,

such as antimicrobial peptides (AMPs), and reactive oxygen species (ROS) (Figure 1.3). Thus, the innate immune response is highly regulated at the transcriptional level.

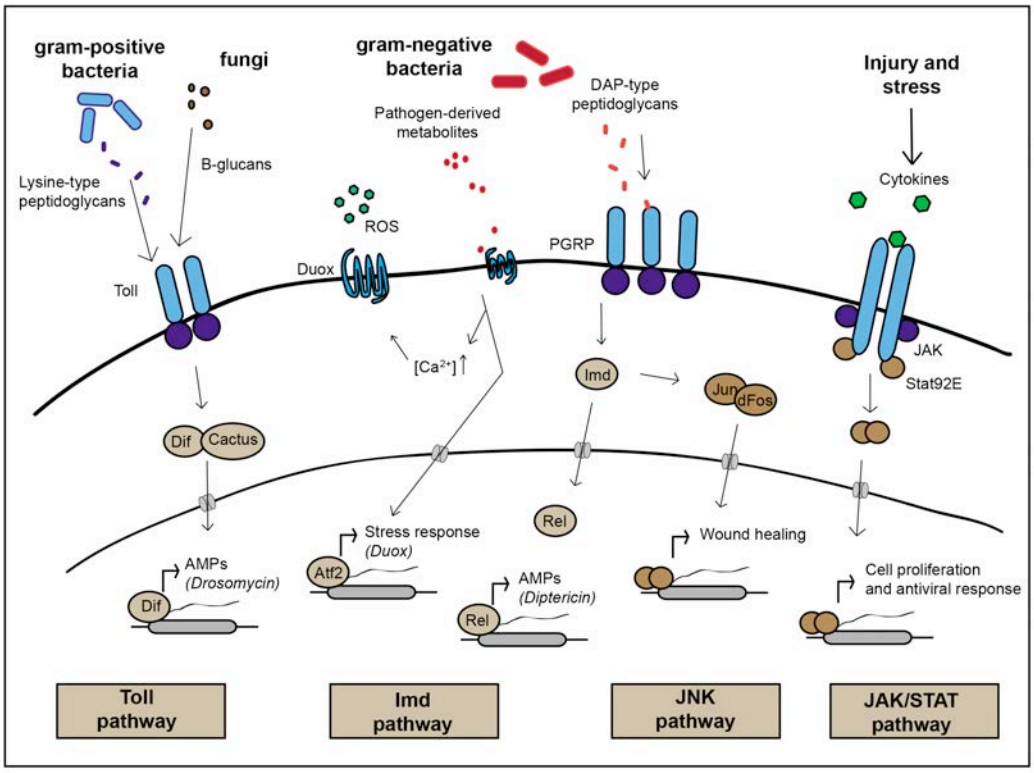


Figure 1.3. Summary of the innate immune response in *Drosophila*. Figure adapted from Buchon et al. 2014. Brief scheme of the main immune response pathways to combat different pathogens. The Toll pathway is mainly activated by gram-positive bacteria and fungi, and concludes with the activation of AMPs, such as *Drosomycin*, mainly by the transcription factor *Dorsal-related immunity factor* (*Dif*). The Imd pathway is mainly activated by gram-negative bacteria. The metabolites generated by these pathogens also activate ROS production. Imd pathway activates the transcription factor *Relish* (*Rel*), which activates AMPs production, such as *Diptericin*. Imd pathway activation also triggers the activation of Jun/dFos signaling, which is necessary for wound healing. Injury and stress activates JAK-STAT pathway and activates cell proliferation. This pathway has also been related to viral response.

Different pathways participate in the *Drosophila* innate immunity (Lemaitre and Hoffmann 2007; Buchon et al. 2014): some pathways show pathogen specificity, and others are general stress response pathways. Depending on the type of pathogen and on the infection route, the immune response uses distinct pathways, therefore, it has different genetic basis (Lemaitre and Hoffmann 2007; Teixeira 2012; Martins et al. 2013; Buchon et al. 2014). Depending on the pathogen, there are two main immune response pathways: the *Toll pathway*, which responds to gram-positive bacteria and fungi infections, and the *Imd pathway*, which responds to gram-negative bacteria (Figure 1.3). Depending on the

infection route, two main immune responses are well distinguished in *Drosophila*: the "systemic immune response", mainly occurring in the fat body, and the "local immune response", occurring in the epithelia such as the gut epithelia.
Besides Toll and Imd pathways, other pathways also participate in the innate immune response either by responding to pathogens or regenerating the damaged host tissue (Boutros et al. 2002) (Figure 1.3). For example, the *JAK/STAT pathway* is involved in cell proliferation and virus response (Myllymäki and Rämet 2014), and the *JNK pathway* is required for proper wound healing (Rämet et al. 2002). Finally, cellular processes such as phagocytosis or melanotic encapsulation also play a critical role in the innate immune response, for example defending against parasitoid eggs (Lemaitre and Hoffmann 2007).

Expression analysis studies after infection with different pathogens have shown that more than one pathway is activated (Boutros et al. 2002; Chakrabarti et al. 2012; Valanne 2014). Thus, a proper orchestration of the different set of pathways, rather than only one specific pathway, is probably responsible for an appropriate immune response in *Drosophila* (Teixeira 2012).

1.2.1 Immune Response in the Gut

The study of *Drosophila* immune response have traditionally been focused on the systemic response (Boman et al. 1972; Lemaitre et al. 1996; De Gregorio et al. 2001), however, several recent works are focusing on the study of local immune response in the gut (Vodovar et al. 2005; Buchon et al. 2009; Bou Sleiman et al. 2015; Capo et al. 2016). Oral infection is probably the most likely infection route happening in nature, and the gut epithelium is the first barrier that bacteria encounter in the organism (Bonfini et al. 2016; Capo et al. 2016). The gut immune response is still not completely understood and it is far more complex than the systemic immune response for several reasons. First, the *Drosophila* intestinal tract is a single tubule but it is anatomically composed by three different domains: the foregut, the midgut and the hindgut (Figure 1.4). At the same time, the midgut can be subdivided into five different histological and functional regions (Buchon et al. 2009; Buchon et al. 2013). Second, the gut is constantly in contact with bacteria composing the microbiota, therefore, the host has to differentiate between pathogenic bacteria and gut microbiota (Broderick et al. 2014; Bonfini et al. 2016). Thus, there must be a complex transcriptional regulatory toolkit in order to control the expression of immune responsive genes (Buchon et al. 2013).

The Imd pathway regulates the immune response in the whole intestinal tract, while the Toll pathway is only activated in the foregut and the hindgut (Buchon et al. 2013). Besides the main immune response triggered by the Imd pathway activation, ROS production

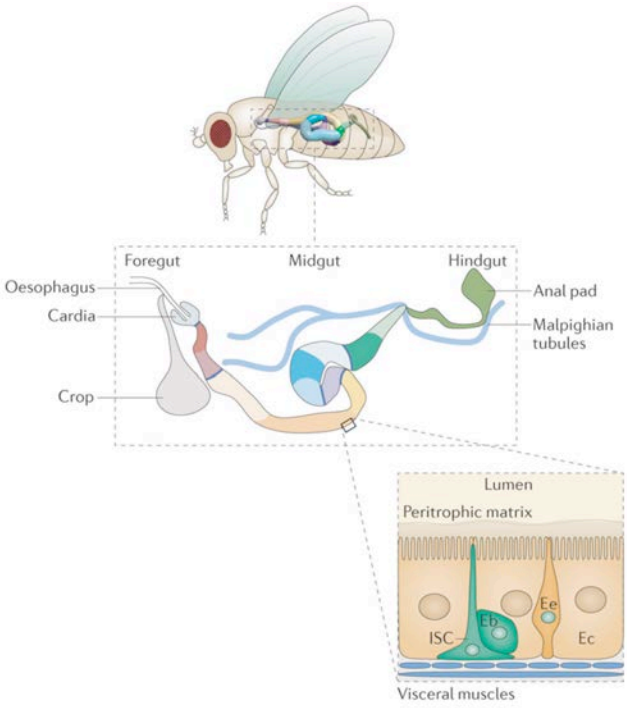


Figure 1.4. *D. melanogaster* gut is structured in different regions. Figure modified from Buchon et al. (2013). The gut is a tubular epithelium composed of a monolayer of different cell types: the enterocytes (Ec), large cells that absorb the nutrients from the lumen, and the secretory enteroendocrine cells (Ee). There are other cells in the gut that are in charge of the gut maintenance: the intestinal stem cells (ISC), and the progenitor cells enteroblasts (Eb). The gut tubule is surrounded by visceral muscles.

activated by the NADPH oxidase DUOX also plays a central role for combating gut local infection (Ha et al. 2005; Kim and Lee 2014). ROS products are secreted into the gut lumen with the aim of eliminating the ingested bacteria. However, this secretion also generates damage to the host cells, thus gut cells need to activate stress response pathways in order to proceed with ROS detoxification.

1.2.2 Natural Variation in the Innate Immune Response

Until recently, many studies focused on the characterization of the genes involved in immune response, mostly by using standard laboratory strains (De Gregorio et al. 2001; Irving et al. 2001; Ayres et al. 2008; Buchon et al. 2009). Several analysis of microarrays from infected laboratory flies revealed that infection triggered the expression of a wide range of genes that can be classified into three functional classes: recognition of the pathogen, signaling pathways, and effector molecules (De Gregorio et al. 2001; Irving et al. 2001; De Gregorio et al. 2002; Roxström-Lindquist et al. 2004; Vodovar et al. 2005; Buchon et al. 2009). Although some of the identified genes overlap in the different studies, there are many genes that are uniquely identified in one study (*e.g.* see Paparazzo et al. 2015). This might be because of the use of different fly strains, or different pathogens, or because the studies focused on different infection routes.

During the last years several studies have addressed genetic variation in immunity on natural populations (Bou Sleiman et al. 2015; Hotson and Schneider 2015; Paparazzo et

al. 2015; Early et al. 2016; Juneja et al. 2016; Howick and Lazzaro 2017). These last studies revealed that, despite of the essential role of immune response in fitness, there is a high genetic and phenotypic variation in the immune response among the strains from the same populations. Specifically, Bou Sleiman and colleagues (2015) revealed that flies with different genetic backgrounds derived from the same natural population harbored high variability in oral infection survival. The same natural population also showed high variation in resistance to and tolerance of infection with a different pathogen (Howick and Lazzaro 2017). These results reflect that immunocompetence is probably mediated by many different *loci* with individual small effects (Weinig et al. 2003; Bou Sleiman et al. 2015).

The extraordinary variability in immune response found by the studies mentioned above can be explained if we consider the evolutionary context of this trait. First, populations need to adapt to face new pathogens when colonizing new environments. Second, at the same time that populations adapt their immune system to overcome infections, the pathogens continuously adapt to circumvent the host immune system. And third, it has been described several interconnections between immunity and other fitness-related traits, like reproduction or metabolism, indicating that positive selection on other physiological traits can impair immune response (Short and Lazzaro 2013; Unckless and Lazzaro 2016).

1.2.3 Evolution of the Innate Immune System

As mentioned above, immunity is one of the traits that often arise when comparing different populations looking for signals of selection (Tinsley et al. 2006; Lazzaro et al. 2008; Fumagalli et al. 2011; Juneja et al. 2016). These evidences have shown that local adaptation is common in immune response not only in *Drosophila*, but also in human populations.

Studies looking for positive selection in *Drosophila* have been traditionally focused on the study of SNPs present in the immune genes (Sackton et al. 2007; Obbard et al. 2009; Early et al. 2016). These studies revealed several characteristics regarding the evolution of immune genes. First they showed that purifying selection act differently depending on the gene position in the network (Wertheim 2015). Thus, while central components of the molecular networks, such as TFs, are highly conserved, there is more diversification in the peripheries of the network (Sackton et al. 2010). These studies also revealed that the different immune pathways vary in the rate of adaptive evolution. A recent study identified a set of 595 genes involved in immune response, 361 of these genes had well-supported immune function (Early et al. 2016). Analyzing the 361 stronger candidate genes, they found that defense genes against RNA virus evolve faster compared to other

immune genes. This had already been observed in other works, where immune genes belonging to RNAi pathway and Imd pathway showed faster evolution rates (Obbard et al. 2009).

Instead of looking for variability in the immune gene sequences, Juneja and collaborators have focused on the study of the geographic variability of gene expression due to *cis* changes (Juneja et al. 2016). Selection on gene expression regulation is thought to be one of the major sources of adaptive evolution, and gene expression plasticity plays a central role when adapting to new environments (Sørensen et al. 2007; Levine et al. 2011). Gene regulation is achieved by *cis* acting elements, which are physically linked to the genes they control, or by *trans* acting elements, which can control many genes physically distant. The modification of *cis*-regulatory elements allows the fine-tuned regulation of gene expression, as it can have tissue specificity, or it can trigger expression at specific times. These characteristics allow *cis* modifications to affect fewer targets compared to *trans* alterations, and this is translated into less fitness costs (Prud'homme et al. 2007; Stern and Orgogozo 2008). Juneja and co-workers found that *cis*-regulatory variation contributed to latitudinal gene expression differences both in North America and Australia *D. melanogaster* natural populations (Juneja et al. 2016).

Transposable elements (TEs) are a source for *cis* regulatory elements that can influence genome regulation (Rebollo et al. 2012; Elbarbary et al. 2016). So far, some studies have directly linked TEs with immune response in a wide range of species from plants to humans (Magwire et al. 2011; Goic et al. 2013; Ali et al. 2014; Chuong et al. 2016; Wang et al. 2016). These examples evidence the impact of TEs on immunity adaptation by using different mechanisms: from gene network regulation (Chuong et al. 2016; Wang et al. 2016), to generating new transcripts (Aminetzach et al. 2005; Magwire et al. 2011), and participating in the creation of V(D)J recombination immune system in vertebrates (Agrawal et al. 1998). However, a systematic search for the role of TEs in immune response has never been performed.

1.3. Transposable Elements

1.3.1 Historical Perspective on the Role of TEs in the Genome

Transposable elements (TEs) are repetitive DNA sequences typically abundant in all the genomes. Barbara McClintock first described TEs in maize, for what she was awarded three decades later with the Nobel Prize in Physiology or Medicine. McClintock observed a changing color pattern in maize kernel, and associated that to the fact that some

chromosome regions had changed position. She first described TEs as "controlling elements" that jump from one site of the genome to another, in response to some change in the environment, thus modifying gene regulation (McClintock 1951; McClintock 1956). A few years later, Britten and Davidson hypothesized that TEs near functionally related genes could contribute to coordinate their expression (Britten and Davidson 1971). However, this idea was hushed during the following years, as the community moved to the view that TEs did not have any biological function and they were categorized as "junk DNA". During these years there were the first experimental evidences of TEs as functional regulatory elements (Samuelson et al. 1990), however, they were taken as sporadic events and did not change the idea of TEs being simply non-functional sequences that behave as DNA parasites (Hickey; Strobel et al. 1979; Doolittle and Sapienza 1980; Orgel and Crick 1980).

Nowadays, next-generation sequencing techniques have boosted the research on TEs. Today we know that TEs do not only constitute an important component of the genomes, but also that they are significant players in genomic functions (Warren et al. 2015; Elbarbary et al. 2016; Garcia-Pérez 2016). Over the last decade, evidences of TEs playing a role as genome regulators in different organisms are accumulating in the literature (*e.g.* Sorek et al. 2002; Leem et al. 2008; Mateo et al. 2014; Puig et al. 2015). The community is now integrating TEs as significant players in genome evolution (Biémont and Vieira 2006; Casacuberta and González 2013; Chuong et al. 2016). Specifically, population genetic studies in *Drosophila melanogaster* reveal that they might be standing strong in recent adaptation (González et al. 2008; González et al. 2010). In fact, it has been shown that they participate in different adaptations such as immune response (Magwire et al. 2011), xenobiotic stress resistance (Mateo et al. 2014), or oxidative stress (Guio et al. 2014).

1.3.2 Transposable Element Classification

Depending on their replication capability, TEs can be classified as autonomous or non-autonomous elements. Autonomous TEs contain ORFs and regulatory sequences that allow them to move from one position to another in the genome, *i.e.* transpose, while non-autonomous TEs depend on the enzymes encoded in the autonomous TEs to transpose. Those enzymes vary depending on the mechanism that the TE uses to transpose. Depending on the transposition mechanism, they are classified into two main categories: Class I elements, or retrotransposons, and Class II elements, or DNA transposons (see Box 2). Inside each category, TEs are further subdivided into orders, based on their structure, and into families, based on sequence similarities (Wicker et al. 2007).

Class I elements, traditionally known as "copy and paste" elements, use an RNA intermediate that is transcribed to dsDNA and then inserted in a different genome *locus*.

Box 2. TE classification in eukaryotes

Class I or retrotransposons

LTR retroelements, flanked by Long Terminal Repeats (LTR), produce target site duplication (TSD) of 4-6 bp upon insertion. They typically contain the ORFs GAG and POL. GAG encodes a structural protein for virus-like particles, and POL encodes for a reverse transcriptase (RT), RNase H (RH), aspartic proteinase (AP) and DDE integrase (INT). There are numerous families of LTR retrotransposons described and they are present in all species groups.

LTR order Ex: *Copia* family

TEs from the DIRS order differ in the mechanism of integration. They encode a tyrosine recombinase, are flanked either by inverted repeats or split direct repeats, and do not generate TSD when they transpose.

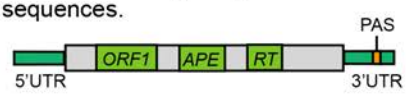
DIRS order Ex: *DIRS* family

Penelope-like elements (PLEs) have LTR-like sequences that can be found in direct or indirect orientation. These elements encode a RT and an endonuclease.

PLE order Ex: *Penelope* family

Long Interspersed Nuclear Elements (LINEs) do not contain LTRs. They encode at least a RT and a nuclease, and display a poly(A) tail at their 3' ends. They typically present truncated 5' ends, probably as a result of premature termination of reverse transcription (Eickbush et al. 2002). Thus, LINE elements usually lack their cis-regulatory sequences.

LINE order Ex: *L1* family



Finally, Short Interspersed Nuclear Elements (SINEs) are non-autonomous elements that originate from accidental retrotransposition of polymerase III (Pol III) transcripts (Kramerov and Vassetzky 2005). They can be expressed as they keep an internal Pol III promoter, however, they rely on LINE RT to transpose.

SINE order Ex: *Alu* family

Class II or DNA transposons

Subclass 1

The TE superfamilies from the TIR order are classified by their TIR sequences and the TSD size. They all encode a transposase, which mediates the transposition of the TE, with a DDE catalytic motif.

TIR order Ex: *Transib* superfamily

Subclass 2

Helitrons encode a Y2-type tyrosine recombinase, which trigger replication via a rolling-circle mechanism as it has an helicase domain and replication initiator activity (Kapitonov et al. 2001).

Helitron order Ex: *Helitron* superfamily

Maverick elements are flanked by long TIRs and encode up to 11 proteins. It is proposed that they are excised from a single strand and, after extrachromosomal replication, they integrate into a new site (Kapitonov et al. 2006).

Maverick order Ex: *Maverick* superfamily

Class I elements encode a reverse transcriptase and they are classified into five orders in eukaryotes: LTR retroelements, DIRS, PLEs, LINEs and SINEs (Box 2). Members of the LTR group are usually found as *solo*-LTRs, as after insertion they often undergo ectopic recombination between the LTRs. As a consequence, the coding regions of the element are removed, and only remains a chimeric copy of the flanking LTRs, which contain the *cis*-regulatory sequences.

Class II elements, traditionally known as "cut and paste" elements, transpose excising from one site of the genome without using an RNA intermediate. They are further classified into two subclasses depending on the number of DNA strand cuts they generate in the donor site (Box 2). Subclass 1 TEs generate double-strand cuts in the donor sequence, and contain TEs that belong to the Terminal Inverted Repeats (TIRs) order. TIR order includes TE families such as *P*-elements, *Merlin*, or *Transib,* among others. Subclass 2 TEs generate only one strand cut in the donor sequence when they transpose, following a process that involves replication. This subgroup comprises large TEs from the orders Helitron and Maverick.

1.3.3 Transposable Elements Abundance and Distribution

TEs are present in all eukaryotes and in almost all prokaryotes sequenced so far (Hua-Van et al. 2011). They usually represent a considerable fraction of the genome: from ~2% in *Pyrococcus furiosus*, to 56% in zebra fish, and 84% in maize (Filee et al. 2007; Schnable et al. 2009; Gao et al. 2016) (Figure 1.5). The TE content of the genome can vary a lot among the different species within the same group. For example, in *Drosophila*, TE content of the genome can vary from less than 5% in *D. busckii* up to 30% in *D. suzukii* (Sessegolo et al. 2016). Moreover, TE families are unequally represented in different species. For example, LINE *L1* elements are the most common family in human, while LTRs are the more abundant in *Drosophila* (Hua-Van et al. 2011; Sessegolo et al. 2016).

TEs are commonly distributed in heterochromatic regions, as well as pericentomeric and telomeric regions (Adams 2000). Those are regions with low gene content and, hence, the potential deleterious impact of the TEs is reduced. Moreover, they tend to accumulate on those regions, as they have a very low recombination rate and almost do not experience purifying selection (Betancourt et al. 2002; Campos et al. 2014; Blumenstiel et al. 2014). Nevertheless, some TEs present in heterochromatic regions have acquired essential roles in the genomes, both structural and functional. In *Drosophila*, TEs exclusively adopted the critical structural role of telomere structure and maintenance (Mason and Biessmann 1995). While in most eukaryotes telomeres are composed of simple repeating units, in *Drosophila* telomeres consist of tandem head-to-tail arrays of retrotransposons. In *Drosophila miranda*, TEs from the *Helitron* family are involved in male X chromosome dosage

compensation (Ellison and Bachtrog 2013). These TEs recruit male specific lethal (MSL) complex to achieve gene dosage compensation in this species.

TEs are also found in euchromatic regions close to genes, and even some of them have an impact on genome regulation. It is known that some TE families preferentially insert into specific sites such as 5' gene regulatory regions (Liao et al. 2000). Thus, we can find TEs that increased their frequency or became fixed as they acquired regulatory roles for the genes nearby and, therefore, were positively selected.

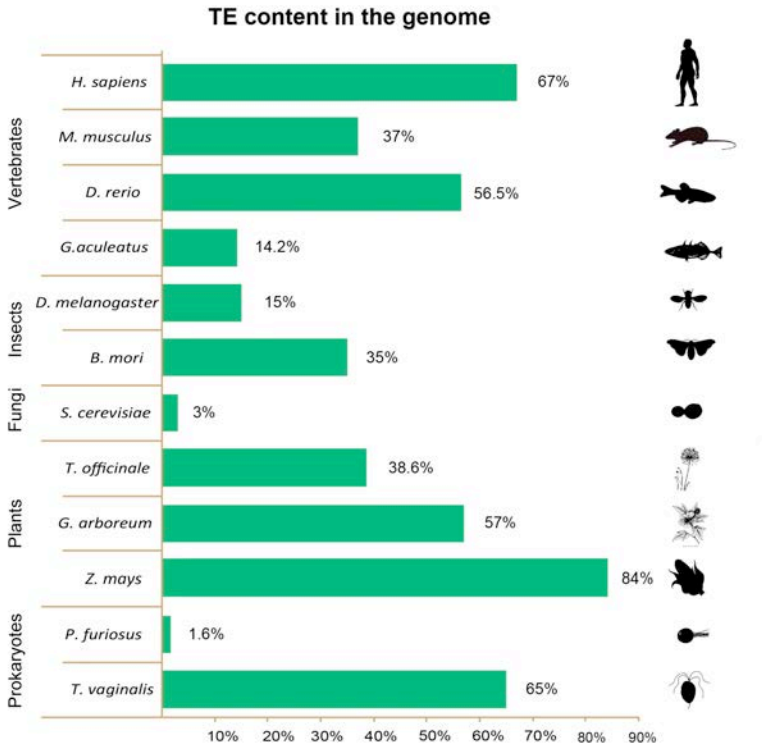


Figure 1.5. Transposable elements are an important component of the genome. The proportion of TEs in the genome vary among the different groups and also vary between closely related species.

1.3.4 Transposable Element Activity

It is assumed that most of the TEs in the human genome are inactive. Most of them have lost their ability to transpose as a result of losing part of their sequences or accumulating mutations over time. There are few TE families that remain active in the human genome, as it has been evidenced by several studies with LINE *L1* elements during early development (Garcia-Perez et al. 2007). Other studies in both plants and animals have also detected TE activity during early stages as well as during gametogenesis (Lisch 2012; Gerdes et al. 2016). Moreover, recent studies in humans have also detected LINE *L1* activity in tumoral cells (Tubio et al. 2014), or neurons (Evrony et al. 2012; Upton et al. 2015). In *Drosophila*, it is assumed that most of the full-length TEs are active, as it has been evidenced in the *gypsy* family (Kim et al. 1994; Leblanc et al. 2000). However, we lack experimental evidences for the activity of the other TE families.

TEs have long coexisted with the host genome and this has inevitably leaded to different types of interactions. As mentioned above, TEs jump in the genome disrupting and/or modifying its regulatory and structural landscape. Because TEs are a potential source for mutations, the host genome had to evolve mechanisms against TE expansion. There have been described several mechanisms to repress TE activity. Most of them are based on TE DNA alterations such as histone modifications, cytosine methylations, or nucleotide hypermutations. But also other mechanisms avoid TE expansion by inhibiting retrotransposition or through piRNA silencing cycles. An example is the APOBEC system, which edit C-to-U in DNA and hypermutates retrotransposon DNA, also interfering with reverse transcription. The APOBEC system has been described in different vertebrate species (Knisbacher and Levanon 2016). Another well-characterized mechanism for TE silencing is the piwi protein complex. This system is based on the production of piRNAs, which bind to TE sequences and block transposition. This mechanism was first reported in *Drosophila* fifteen years ago (Aravin et al. 2001), and nowadays it has been shown in many different organisms including vertebrates (Czech and Hannon 2016).

1.3.5 TEs Drive Genomic Variation

TEs are a significant source for generating genome variation in organisms. As DNA sequences, they can influence host genome in many different ways, such as changing gene regulation or genome structure (Warren et al. 2015; Chuong et al. 2016; Elbarbary et al. 2016). These changes can be both genetic and epigenetic, and can be exerted by the TE both directly and indirectly.

1.3.5.1 Regulatory Changes Induced by TEs

Several evidences show that some TEs have been co-opted by the host, as they acquired a regulatory function that confers adaptive changes (Box 3). TEs can model genomes either by influencing individual genes or by modulating gene networks.

Gene modeling by TEs can be achieved through multiple mechanisms: from generating new gene transcripts, to adding new regulatory elements, to altering the chromatin structure (see Box 3). For example, TEs can potentially act as enhancers or promoters for the nearby genes, as they carry regulatory sequences that are targeted by the host transcription machinery. A beautiful example is the role of *carb*-TE in the environmental adaptation of the peppered moth. This TE was found to up-regulate the *cortex* gene resulting in increased darker coloration, thus improving the fitness in polluted environments (van't Hof et al. 2016). Besides adding regulatory regions, TEs can also alter

chromatin structure by recruiting heterochromatin proteins and, thus, silencing the nearby genes (Sentmanat and Elgin 2012).

Besides modifying the regulation of individual genes, TEs can also regulate whole host pathways. The first studies on gene regulation evolution suggested that TEs might play a role in rewiring host regulatory networks. Recent findings evidence that TEs play a major role in gene network regulation, and show that TEs participate in critical physiological responses such as immune response in mammals (Chuong et al. 2016), sex chromosome dosage compensation in *Drosophila miranda* (Ellison and Bachtrog 2013), or early development in mammals (Gerdes et al. 2016).

1.3.5.2 Exaptation

Sometimes part of the TE sequences can be translated into functional proteins, thus generating innovations in the host protein repertoire. Different authors have described these events by using different terms such as *molecular domestication* (Miller et al. 1997), *co-opted events* (Sarkar et al. 2003), or *exaptation* (Brandt et al. 2005). Exaptation can lead to the domestication of an entire gene from the TE, such as the transposase, or to the generation of chimeric proteins, such as the fusion of the TE gene with a host gene.

A well-known example of TE exaptation is RAG1 and RAG2 antigen receptors which initiate the assemble of the gene segments that generate immunoglobulin and T cell receptors in vertebrates, known as the V(D)J recombination (Agrawal et al. 1998). A recent study provided the definitive evidence for the transposon exaptation of RAG antigen receptors (Huang et al. 2016). In this work, they found an evolutionary relative of the RAG transposon in lancelets, and propose that this TE was transmitted vertically through chordate and vertebrate evolution. Another well-known example is the primate protein *SETMAR*, which was found to be a fusion between a pre-existing SET histone methyltransferase gene and the TPase gene of *Hsmar1* transposon (Cordaux et al. 2006). The authors of this study showed that the DNA binding domain of this protein had evolved under continuous purifying selection. *In vitro* experiments demonstrated that the TPase region of *SETMAR* has retained a strong DNA-binding activity while it has lost its catalytic ability. Thus, TPase regions might be targeting the SET domain to different sites in the genome, and this might modify the chromatin and regulate gene expression of the genes in that region (Feschotte 2008).

There are other examples of TEs captured by host genomes that are translated as part of functional proteins. For example, an exon of *FASTKD3* in the bovine lineage (Almeida et al. 2008); or the centromere protein *CENP-B,* which has been independently

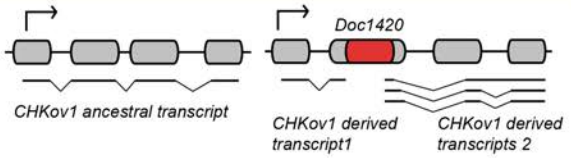
Box 3. Transposable elements generate genome changes through multiple mechanisms

TEs use a wide range of mechanisms to generate genome variation either as a consequence of the position where they insert, or because they add new regulatory sequences. In addition, they can alter genome structure by driving chromosomal rearrangements by ectopic recombination, as well as generate deletions when they jump.

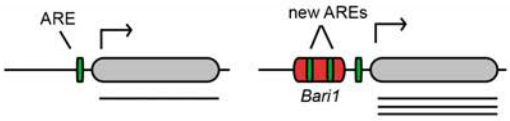
Depending on the position they insert in the genome TEs can modify gene regulation in different manners. The simplest example is when a TE **truncates genome sequences** (A) such as gene coding regions or gene promoters. *Doc1420* TE in *D. melanogaster* truncates *CHKov1* and generates a different transcript, thus driving adaptation to pesticide and virus infections (Aminetzach et al. 2005).

Nevertheless, TEs can generate more complex modifications by **adding new regulatory sequences** (B) such as promoters, enhancers or repressive elements, and hence, fine-tuning gene expression levels. For example, *Bari1* adds antioxidant response elements that function as an enhancer for the nearby gene (Guio et al. 2014). Some TE families insert preferentially to 5' regions of genes, and so they are more likely to modify gene expression regulation. This is the case of Tf1 retroelements in fission yeast (Leem et al. 2008) and P-elements in *Drosophila* (Liao et al. 2000).

(A) Truncate genes



(B) Add new regulatory sequences

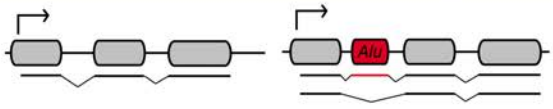


(C) Generate new transcripts

TEs can also trigger the **generation of new gene transcripts** (C) by contributing with a new TSS for the gene. It has been shown that TEs drive the expression of more than 150 genes in *Drosophila melanogaster* during development (Batut et al. 2012). One example is the *roo* element *FBti0019985*, which promotes transcription of the nearby gene *CG18446* in early embryos (Merenciano et al. 2016).

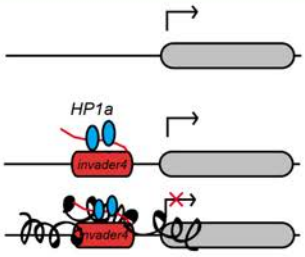
When TEs insert into introns they can potentially vary gene transcription by providing **alternative splicing sites** (D). Additionally, in some cases TEs can be incorporated in the transcript. It is estimated that 5% of human alternative spliced exons derive from Alu sequences (Sorek et al. 2002).

(D) Participate in alternative splicing



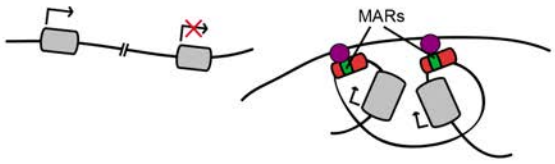
Furthermore, TEs can act as chromatin protein targets. Thus, they can trigger **heterochromatin formation** (E) that can be spread to the flanking genomic regions, so genes nearby can be silenced. For example, TEs from *1360* and *invader4* family, when transcribed, act as piRNA targets (Sentmanat et al. 2012). This is recognized by HP1a protein, which triggers the recruitment of heterochromatin proteins in the flanking regions, thus silencing the nearby genes.

(E) Contribute to heterochromatin formation



Box 3. Transposable elements generate genome changes through multiple mechanisms (continued)

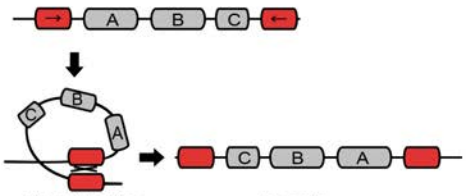
(F) Determine topological domains



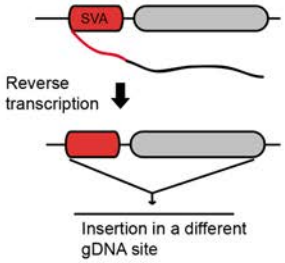
In addition, TEs can participate in the **determination of topological domains** (F) in the nucleus. It has been described that TEs can bind to matrix-attachment regions (MARs) and, hence, they could participate in the determination of chromatin loops in *D. melanogaster* nucleus (Mamillapalli et al. 2013). These structures might play a role in coordinating gene transcription by separating functional domains.

Finally, as a consequence of their high copy number and mobility, TEs also generate chromosomal structural mutations such as inversions, duplications, and deletions.

(G) Generate inversions



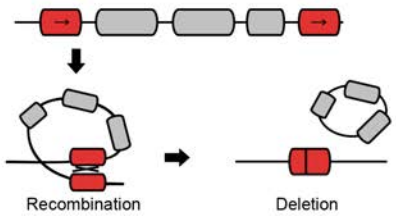
(H) Generate duplications



Due to the high sequence similarity between TEs from the same family, they act as substrate for ectopic recombination between two different TE copies, thus generating a chromosomal **inversion** (G) (Cáceres et al. 1999; Puig et al. 2015). SINE/VNTR/Alu (SVA) elements have been associated to gene **duplications** (H) in primates. For example, *AMAC* gene is found in three copies in the genome, as a consequence of the retrotransposition mediated by the TE (Xing et al. 2006).

In humans, SVA elements have been associated to large chromosome **deletions** (I) causing disease, probably by non-allelic homologous recombination (NAHR) (Vogt et al. 2014) TEs with similar sequences can undergo NAHR and generate deletions of the genomic regions between the two TE copies.

(I) Generate deletions



domesticated from a *pogo*-like transposase in several metazoan species (Casola et al. 2008; Mateo and González 2014).

1.3.6 Transposable Elements Are Effective Drivers of Adaptation in *D. melanogaster*

TEs have long been ignored as candidate mutations involved in key biological processes such as adaptation (González et al. 2010; Hoban et al. 2016). Because TEs are highly repetitive sequences and are found in many copies in the genome, both the identification and the annotation in the genome is a difficult task. Nowadays we are overcoming these challenges thanks to next-generation sequencing techniques, which facilitate the study of

TEs and allow considering them when studying complex processes such as genome evolution (Fiston-Lavier et al. 2015; Rahman et al. 2015; Kofler et al. 2016).
Unlike SNPs, which usually generate neutral mutations, the majority of TE-induced mutations are deleterious. Although a proportion of the TEs present at high frequencies could be neutral, we expect high frequency TEs to be enriched for adaptive mutations in this species (Barrón et al. 2011). This is especially true in *D.melanogaster*, where the efficiency of selection is high as it has a big population size and, hence, we would expect most TE insertions to be present at low population frequencies (Barrón et al. 2011; Kofler et al. 2012; Cridland et al. 2013; Barron et al. 2014; Blumenstiel et al. 2014). Moreover, genomic changes generated by TEs are more complex compared to other kind of mutations such as SNPs or InDels (see Box 3).

So far, there are few genome-wide studies surveying TE-induced adaptive mutations (González et al. 2008; González et al. 2010; Kofler et al. 2012; Blumenstiel et al. 2014). González and collaborators showed that TEs contributed significantly to *D. melanogaster* recent out-of-Africa adaptation. They screened North American natural populations by PCR, finding 18 polymorphic TEs likely involved in environmental adaptation (González et al. 2008; González et al. 2010). These authors detected signatures of selective sweep in the flanking regions for five of the TEs. Moreover, eight out of the 18 candidate TEs showed population differentiation. At the beginning of this thesis, only one out of the 18 TEs had been linked to its relevant adaptive phenotypes: insecticide resistance and virus resistance (Aminetzach et al. 2005; Magwire et al. 2011). Among these TEs, *FBti0019386* showed consistent population differentiation pattern in the two hemispheres (González et al. 2008; González et al. 2010). *FBti0019386* was found at higher frequency in temperate populations compared to tropical populations, suggesting a possible role in temperate environment adaptation. Besides the work from González and collaborators (2008, 2010), two other genome-wide studies screened *D. melanogaster* populations and found 20 new candidate adaptive TEs (Kofler et al. 2012; Blumenstiel et al. 2014). Kofler and colleagues (2012) analyzed *in silico* the genome of a European population and identified 13 fixed TEs showing genomic signatures of positive selection. In the third screening, Blumentiel and colleagues (2014) analyzed by PCRs 12 strains from a North American population, and 12 strains from an African population and identified a total of 9 candidate TEs.

So far, the number of TE insertions identified in the three genome-wide screenings, 38 candidate TEs, is probably underestimated due to some technical limitations. In two of the three studies, the screening methodology was based on PCRs (González et al. 2008,

Blumenstiel et al. 2014). This technique is limited to the possibility of designing primers to allow the TE identification, and also it is time-consuming. Nowadays the TE screening is facilitated by the development of new next-generation sequencing techniques, as well as the availability of software able to detect the TE insertions and calculate its frequencies (Fiston-Lavier et al. 2015; Kofler et al. 2016). Although Kofler et al. (2012) performed an *in silico* screening, they only considered as candidates the TEs fixed in one population. Another limitation for the two PCR-based screenings is that it was only focused on the identification of TEs present in the reference sequence, so it does not consider other adaptive insertions that might be segregating in natural populations. Indeed, in the genome-wide screening performed by Kofler et al. (2012), they already detected two candidate TEs not annotated in the reference genome showing genomic signatures of positive selection. Finally, in the three screenings, very few natural populations were used to identify the candidate TEs. González and collaborators already noticed that only half of the identified adaptive TE insertions were present in all the populations analyzed, indicating that local adaptation is common (González et al. 2008; González et al. 2010). Thus, sampling more populations should increase the number of identified adaptive TEs.

02

OBJECTIVES

2. OBJECTIVES

The objectives of this thesis are:

1. **To characterize the previously identified *FBti0019386* insertion.**

I will explore the adaptive phenotypes associated with *FBti0019386* insertion by performing phenotypic experiments with flies with and without the TE. I also will also analyze the molecular mechanisms underlying the phenotypes observed.

2. **To identify TEs candidate to be involved in *D. melanogaster* adaptation.**

I will identify the candidate TEs in several *D. melanogaster* natural populations from three different continents using *in silico* approaches. To detect a big dataset of candidate TEs for adaptation, I will consider both annotated and a subset of non-annotated TEs in the reference genome. Finally, I will analyze which phenotypes are more likely associated with the candidate adaptive TEs.

3. **To characterize several candidate adaptive TEs associated with a relevant phenotype.**

I will check whether the candidate TEs are associated with expression changes of the nearby genes. I will also identify the molecular mechanisms behind the expression changes.

03

RESULTS

3.1 RESULTS. CHAPTER 1

3.1.1 *FBti0019386* Flanking Regions Show Signatures of Positive Selection

We tested whether the region flanking *FBti0019386* showed signatures of positive selection (see Materials and Methods for a description of the different tests used). We found an extreme decrease of nucleotide diversity (π) in strains with *FBti0019386* insertion compared with strains without the insertion, which was accompanied by a decrease in Tajima's D statistic (Table 3.1.1, Annex Table S1.1, Figures S1A and S1B) (Tajima 1989; Hudson et al. 1992). The Composite Likelihood (CL) test, specifically designed to detect selective sweeps (Nielsen et al. 2005), was higher in flies with *FBti0019386* insertion compared with flies without the insertion, as expected if flies with the insertion show signatures of a selective sweep in the analyzed region (Table 3.1.1). We confirmed that values of π, Tajima's D, and CL were statistically different from neutral simulated scenarios in flies with *FBti0019386* insertion but not in flies without the insertion (Table 3.1.1 and Annex Table S1.2).

	Observed		Neutral simulations				Resampling of Strains	
			Mean (CI 95%)		p-value		Mean (CI 95%)	p-value
	P	A	P	A	P	A	P	A
π	0.43	4.51	3.92 (1.32, 7.81)	4.20 (1.33, 8.04)	0.001	> 0.05	3.35 (2.78, 3.87)	< 0.001
Tajima's D	-1.77	0.68	-0.11 (-1.46, 1.62)	-0.04 (-1.41, 1.64)	0.007	> 0.05	0.4 (-0.19, 1.02)	< 0.001
CL (log)	-5.95	-18.15	-18.69 (-29.67, -8.80)	-15.20 (-25.89, -6.82)	0.006	> 0.05	-12.18 (-15.23,-8.81)	< 0.001

Table 3.1.1: Summary of the analyses showing evidence of positive selection in the 1-Kb region around *FBti0019386* insertion.
NOTE: Neutral simulations were performed with MS program using the parameter theta = 4. For simulations with theta = 5, please see Annex Table S1.2. P, data set of strains with *FBti0019386* insertion; A, data set of strains without *FBti0019386* insertion.

To test whether the observed differences were due to the *FBti0019386* insertion, we estimated the three statistics in random samples of the strains (see Materials and Methods). None of the randomized data sets had lower π, lower Tajima's D, or higher CL value compared with the data set of strains with *FBti0019386* insertion (Table 3.1.1 and Annex Table S1.3). Finally, we performed the Composite Likelihood Ratio (CLR; Nielsen et al. 2005) test comparing strains with and without the *FBti0019386* insertion, and we found that it was significant: CLR = 24.40 p-value = 7.82×10^{-7}. Moreover, this CLR value is three times bigger than any of the CLR values calculated in a random sample of 1,000 1-kb-long regions from 3R chromosome, where *FBti0019386* is located (Annex Table S1.4). Note that estimates of π and Tajima's D in these 1,000 regions also showed that these two statistics did not significantly differ between strains with and without *FBti0019386* insertion (Annex Figure S1C and D).

Note that we checked whether polymorphisms other than TE were present in the flanking

regions analyzed. No other polymorphisms were found that could potentially confound the results of our tests of selection suggesting that the TE is the causative mutation.

Overall, we found evidence of positive selection in the region flanking *FBti0019386* insertion suggesting that *FBti0019386* is an adaptive insertion.

3.1.2 Exploring the Fitness Space of *FBti0019386*

To explore the phenotypic space of *FBti0019386* insertion, we investigated several traits related to the phenotypic effects of nearby genes: Fecundity and egg hatchability associated with *sra* mutant alleles. Related to egg hatchability, we also investigated egg hatching time, egg-to-adult viability, and DT. Additionally, we investigated cold stress, osmotic stress, and starvation stress as Bin1 mutants have been shown to play a role in stress resistance.

Because *FBti0019386* is located 242.4 kb away from the distal breakpoint of In(3R)Payne inversion and inversions are known to be under selection, we checked whether this inversion was present in any of the six strains used to perform the different phenotypic analyses (see Materials and Methods). We found that none of the strains used in our analyses carries In(3R)Payne inversion.

We also checked whether polymorphisms other than the *FBti0019386* insertion were present in the genomic region including *sra* and *Bin1* genes. We did not find any polymorphism linked to the *FBti0019386* that could potentially confound the results of the phenotypic assays performed.

3.1.2.1 *FBti0019386* Insertion Does Not Affect Fecundity or Egg Hatching

Laboratory mutant flies in which *sra* is underexpressed lay less eggs than wild-type flies and most of the eggs do not hatch (Horner et al. 2006). To check whether *FBti0019386* insertion has an effect on fecundity, we compared the number of eggs laid per female in outbred populations

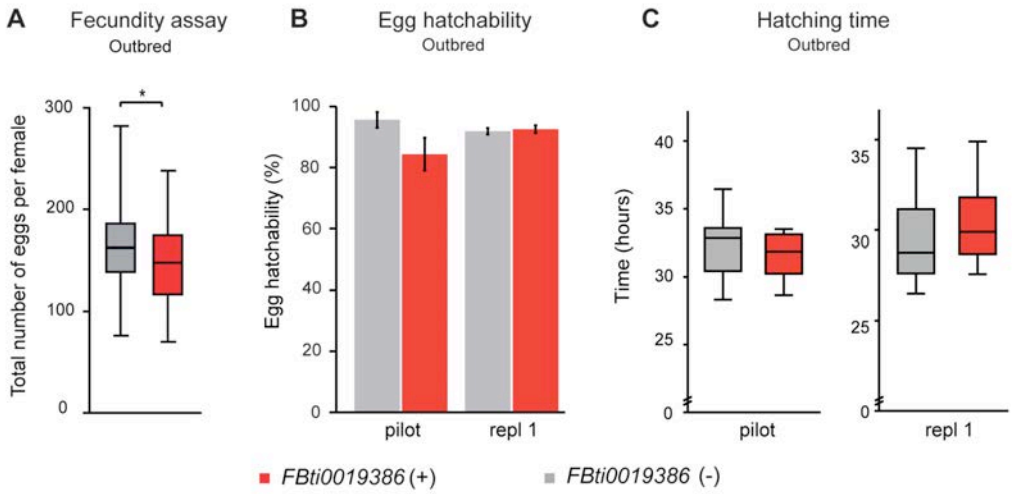


Figure 3.1.1 *FBti0019386* does not affect fecundity (A), egg hatchability (B), or hatching time (C) in outbred populations. (A) Average number of eggs laid by outbred females without *FBti0019386* insertion (*FBti0019386* (-)) and with *FBti0019386* insertion (*FBti0019386* (+)). (B) Percentage of hatched embryos. (C) Average hatching time. In all cases, error bars represent standard error of the mean (SEM).

with and without the insertion (see Materials and Methods). Our results showed that, on average, flies without the insertion laid slightly more eggs than flies with the insertion (t-test, p-value = 0.047) (Figure 3.1.1A). However, the size effect of the mutation was not significant (table 3.2). We also tested whether differences in fecundity were present early in life, as has been reported by Paaby et al. (2014). Although the mean number of eggs laid by flies with the insertion in the first 48 h of egg laying was bigger than the number laid by flies without the insertion (3.95 vs. 2.33 eggs), these differences were not statistically significant (t-test, p-value = 0.06) (Table 3.1.2).

Experiment	Strain	OR (CI)
Fecundity	Outbred	1.05 (0.67–1.64)
Hatching time in cold	Outbred pilot	7.07 (3.37–14.83)
	Outbred replica 1	2.21 (1.49–3.26)
DT	Outbred pilot	5.69 (2.72–11.94)
	Outbred replica 1	2.62 (1.88–3.66)
	Outbred replica 2	2.60 (1.94–5.88)
	Individual DGRP	1.95 (1.30–2.92)

Table 3.1.2: Odds ratios (OR) and confidence intervals (CI) for phenotypic experiments performed with embryos with and without *FBti0019386*.

We then checked whether outbred flies with and without *FBti0019386* differed in egg hatchability and/or hatching time. We first performed a pilot experiment using 150 embryos per strain and we found that flies with the insertion did not show significant differences compared with flies without the insertion in egg hatchability (t-test, p-value > 0.05) (Figure 3.1.1B) or hatching time (t-test, p-value > 0.05) (Figure 3.1.1C). Although differences were not significant, flies with the insertion showed a lower number of hatched eggs (Figure 3.1.1B) and a shorter hatching time (Figure 3.1.1C). We thus repeated the experiments using 500 embryos per strain and we found that flies with and without *FBti0019386* did not differ in egg hatchability (t-test, p-value > 0.05) (Figure 3.1.1B) or hatching time (t-test, p-value > 0.05) (Figure 3.1.1C).

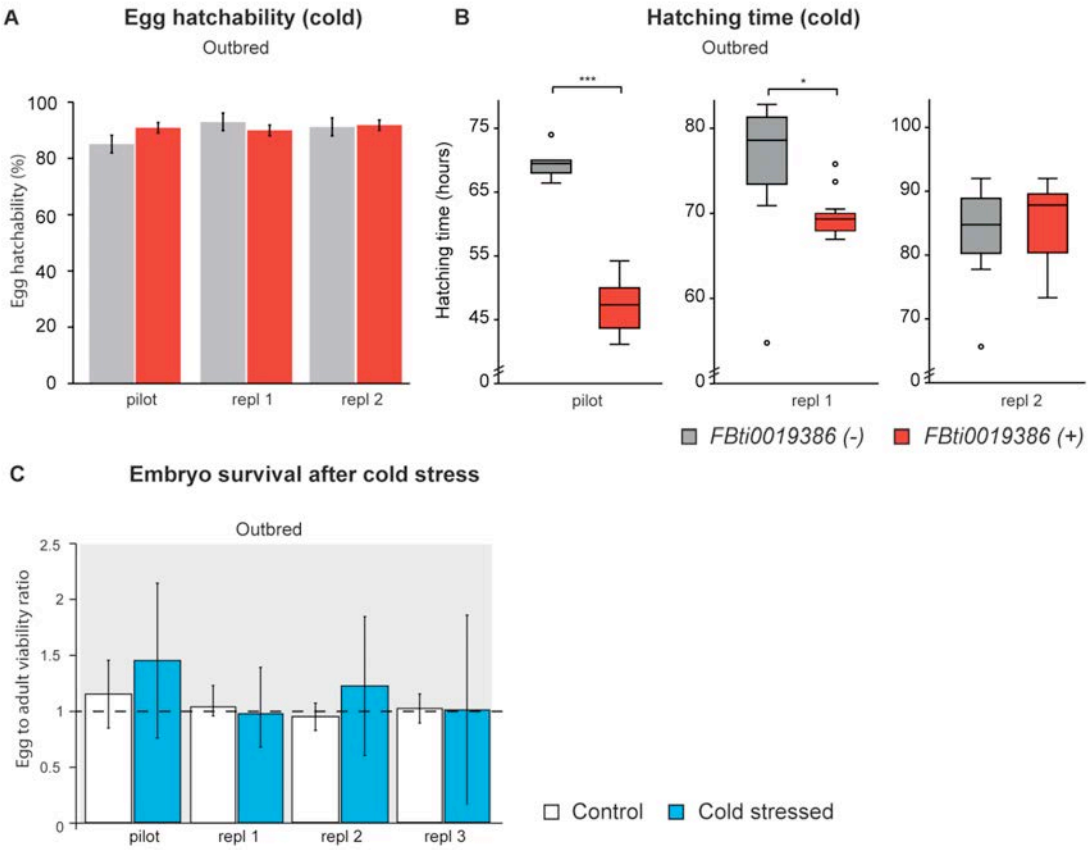
Overall, we did not find significant differences in fecundity, egg hatchability, or egg hatching time in flies with and without *FBti0019386* insertion. These results suggest that *FBti0019386* does not have a significant effect on these phenotypes.

3.1.2.2 *FBti0019386* Insertion Does Not Affect Egg Hatching or Egg- To-Adult Viability under Cold Stress Conditions

As mentioned above, *Bin1* plays a role in general environmental stress response in *Drosophila* (Costa et al. 2011). We thus screened several phenotypes in embryos under cold stress conditions: Egg hatching, egg hatching time, and egg-to-adult viability.

We performed egg hatchability and egg-hatching time assays in outbred populations under repeated cold stress exposure (see Materials and Methods). We did not detect differences in egg hatchability between flies with and without the insertion in any of the three replicas performed

(t-test, p-value > 0.05) (Figure 3.1.2A). However, flies with *FBti0019386* insertion from the pilot experiment and the first replica hatched significantly before flies without the element (t-test, p-value << 0.001 and p-value = 0.011, respectively) (table 3.1.2) whereas no differences were



observed in the second replica (t-test, p-value > 0.05) (Figure 3.1.2B).

Figure 3.1.2 *FBti0019386* does not affect embryo hatching or survival in cold stress conditions in outbred populations. (A) Percentage of embryos that hatched during cold-stress periods (see Materials and Methods). (B) Average egg hatching time. (C) Egg-to-adult survival after a single cold stress period during embryonic stage (cold stressed) and under control conditions (control). Bars represent the survival ratio between flies with *FBti0019386* and flies without *FBti0019386* and error bars represent SEM.

We further tested whether flies with and without *FBti0019386* differed in the egg-to-adult viability after exposing outbred flies to a single cold-stress period during early embryo stages. Our results showed that there are no differences in survival between flies with and without the insertion in control conditions or under cold-stress (two-way ANOVA [analysis of variance], p-value > 0.05, Figure 3.1.2C).

Overall, and although variability in hatching time was observed in some of the experiments performed, our results suggest that *FBti0019386* insertion does not affect cold-tolerance during the embryo stage.

3.1.2.3 *FBti0019386* Is Associated with Increased Sensitivity to Cold Stress in Adults

Because we could not find any significant difference between strains with and without *FBti0019386* in embryonic stage, we decided to test whether differences between the two strains were present in adult flies. We first tested whether adult flies with and without *FBti0019386* insertion differed in chill-coma recovery time (CCRT) and survival after cold stress. CCRT is used as a reliable measure of cold tolerance in *Drosophila* (Macdonald et al. 2004; Gibert et al. 2007). We observed that flies with the insertion showed significantly longer recovery time compared with flies without the insertion suggesting that they were more sensitive to cold stress (Mann–Whitney test, p-value << 0.001) (Figure 3.1.3A and Table 3.1.3). We replicated this result in flies with the same genetic background (Mann–Whitney test, p-value < 0.05) and in flies with two other genetic backgrounds: The introgressed strains generated in our laboratory (Mann–Whitney test, p-value << 0.001) and a couple of inbred strains from the DGRP (Drosophila Genetic Reference Panel) project (Mann–Whitney test, p-value << 0.001) (Figure 3.1.3A and Table 3.1.3) (see Materials and Methods).

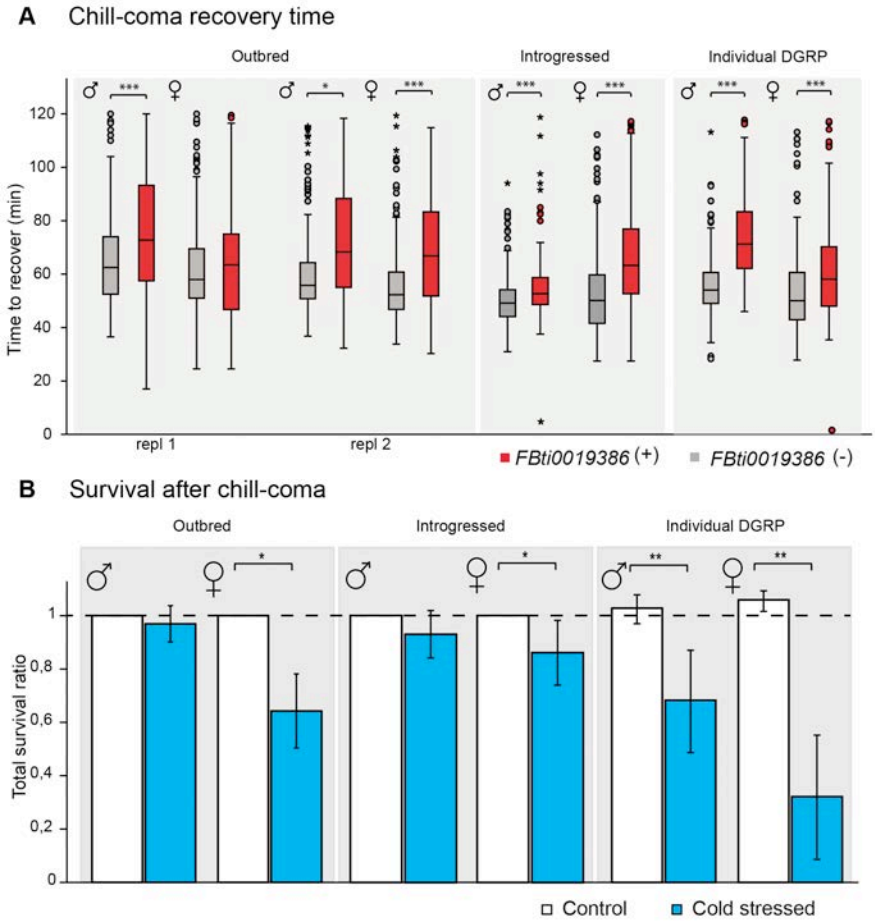


Figure 3.1.3 Flies with *FBti0019386* insertion are more sensitive to cold stress. (A) Average time to recover after chill coma in adult flies from outbred populations, introgressed strains, and inbred DGRP strains (RAL-857 and RAL-802). (B) Survival ratio between flies with *FBti0019386* insertion and flies without the insertion after chill coma exposure (cold stress) and in control conditions (control) in the three genetic backgrounds. Error bars represent SEM.

In accordance with this increased cold sensitivity, flies with the insertion also showed an increased mortality following chill-coma exposure, although these differences were not always significant (Figure 3.1.3B and Table 3.1.3). Finally, we also tested whether flies with *FBti0019386* insertion were more sensitive to osmotic stress and starvation stress. We found that outbred females with the insertion were more sensitive to high salt concentrations (Kaplan–Meyer, log rank p-value < 0.001) (Annex Figure S2A and Table 3.1.3), and outbred males with the insertion were more sensitive to starvation stress (Kaplan–Meyer, log rank p-value < 0.001) (Annex Figure S2B and Table 3.1.3).

Experiment	Strain	Males OR (CI)	Females OR (CI)
CCRT	Outbred replica 1	3.44 (2.31–5.18)	NA[a]
	Outbred replica 2	3.79 (2.54–5.67)	5.18 (3.43–7.82)
	Introgressed	2.44 (1.64–3.62)	4.16 (2.69–6.41)
	Individual DGRP	11.63 (6.79–19.93)	2.26 (1.54–3.33)
Survival after chill-coma	Outbred	NA	7.80 (3.27–18.60)
	Introgressed	NA	1.89 (0.99–3.62)
	Individual DGRP	9.94 (5.49–18)	6.88 (3.43–13.82)
Osmotic stress	Outbred	NA	1.61 (1.21–2.13)
Starvation stress	Outbred	1.52 (1.15–2.01)	NA

Table 3.1.3 Odds ratios (OR) and confidence intervals (CI) for phenotypic experiments performed with male and female flies with and without *FBti0019386*.
[a]NA (OR was estimated when differences between flies with and without *FBti0019386* were statistically significant).

Overall, longer CCRT and lower cold-stress survival in flies with *FBti0019386* insertion across backgrounds suggested that this mutation is negatively affecting adult cold-stress response. This high sensitivity to cold stress likely represents the cost of selection of this TE mutation. Furthermore, preliminary results are suggestive but not conclusive of a negative role of *FBti0019386* in general response to stress.

3.1.2.4 *FBti0019386* Insertion Is Associated with Shorter DT

During the course of the experiments, we noticed that flies with *FBti0019386* showed a shorter DT than flies without the insertion. Because DT is relevant to fitness in all organisms, and especially for those such as *D. melanogaster* that occupy ephemeral habitats (Chippindale et al. 1997), we tested this observation. We found that outbred flies (Mann–Whitney test, pilot experiment p-value = 0.006 and replica 1 and 2 p-value < 0.001) and inbred DGRP flies (t-test, p-value = 0.02) with the insertion developed faster compared with flies without the TE insertion (Figure 3.1.4 and Table 3.2). On average, flies with *FBti0019386* insertion developed 9.4–17.9 h before compared with flies without the insertion. However, we could not detect significant DT differences in the introgressed strains differing by the presence/absence of *FBti0019386* (t-test, p-value > 0.05) (Figure 3.1.4), suggesting that polymorphisms other than the TE influence DT in this background. Note that the effect size of the mutation on the other phenotypes studied also varies depending on the background being analyzed (Tables 3.1.2 and 3.1.3). This suggests

that polymorphisms other than *FBti0019386* play a role not only in DT but also in other phenotypes.

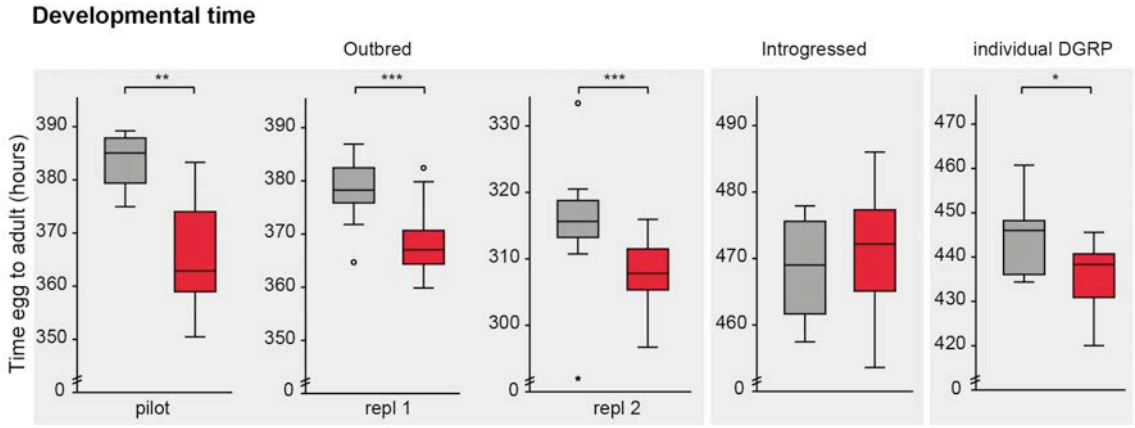


Figure 3.1.4 *FBti0019386* is associated with shorter DT. Average egg-to-adult DT in populations without *FBti0019386* insertion and with the insertion. Error bars represent SEM.

3.1.3 *FBti0019386* Frequency Showed Clinal Patterns in North America and Australia but No Correlation between Frequency and Latitude Is Found in Europe

Shorter DT and increased sensitivity to cold stress are not consistent with a role of *FBti0019386* in temperate adaptation (Gonzàlez et al. 2010). However, previous evidence for a role in temperate adaptation was based on the analysis of only two North American and five Australian populations (Gonzàlez et al. 2010). To further test these results, we estimated *FBti0019386* frequencies in additional populations from North America, Australia, Europe, and Africa (Annex Table S1.5) using T-lex2 pipeline (Fiston-Lavier et al. 2015). We found that *FBti0019386* insertion is present at 10% frequency in a Rwanda population confirming its low frequency in Africa (Annex Table S1.5). We confirmed that the TE is present at intermediate to high frequencies in 15 additional out-of-Africa populations (Figure 3.1.5 and Annex Table

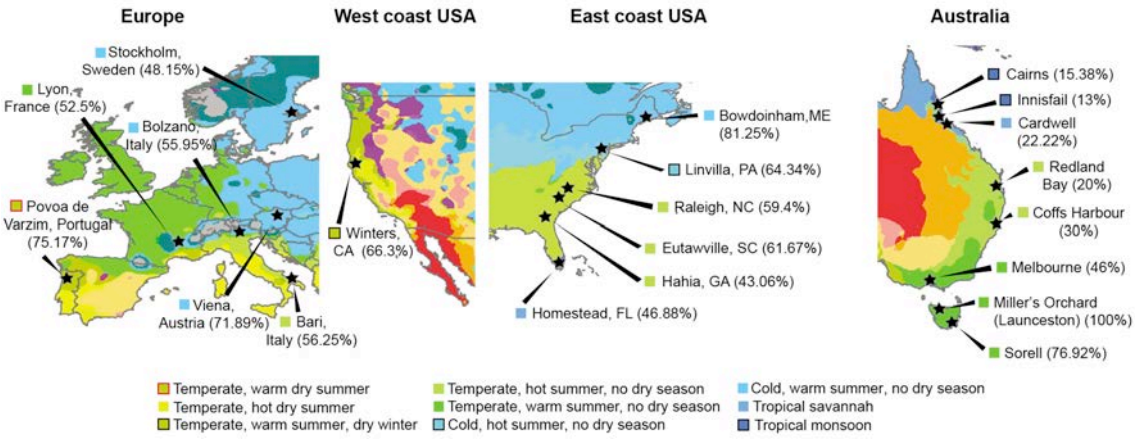


Figure 3.1.5 Climate map with *Drosophila melanogaster* population samples analyzed with T-lex2. The frequency of *FBti0019386* in each population is shown in brackets. Climate maps are modified from Peel et al. (2007).

S1.5). We also confirmed that the TE frequency varies clinally with latitude in North America and Australia (Pearson correlation p-value = 0.011 and p-value = 0.002, respectively; Annex Table S1.6). However, when we analyzed the *FBti0019386* frequency in six European populations we did not find any significant correlation between frequency and latitude (Pearson correlation p-value = 0.313; Annex Table S1.6).

Besides latitude, we also tested whether other geographical and climatic variables showed significant correlations with *FBti0019386* frequency. We found significant correlations between frequency and temperature-related variables in North America and between frequency and both temperature-related and precipitation-related variables in Australia (Annex Table S1.6). No significant correlation was found in Europe (Annex Table S1.6). Because most of the climatic variables are significantly correlated among them and with latitude (Annex Table S1.7), we performed a Principal Component Analysis (PCA) to disentangle the relationships between the variables. In North America, climate variables were grouped in two components, in Australia in three and in Europe in two (Annex Table S1.8). As expected based on the correlation analyses, only in North America and in Australia, some of the climatic variables grouped with latitude and frequency (Annex Figure S3A). In North America, the first component accounted for 46% of climatic variation (Annex Table S1.9) and explained 54% of the variation in *FBti0019386* frequency (Annex Figure S3B). In Australia, the first component accounted for 68% of climatic variation (Annex Table S1.9) and explained 86% of the frequency variation (Annex Figure S3B). Finally in Europe, the first principal component explained 54% of the climatic variation (Annex Table S1.9) but was not significantly correlated with *FBti0019386* frequency (Annex Figure S3B).
Overall, although we were able to confirm the clinal pattern of *FBti0019386* in North America and Australia, our results did not provide evidence for the presence of a clinal pattern in Europe. In Australia, the clinal pattern is well explained by the observed climatic variation, whereas in North America climatic variation did not fully explain the observed correlation between *FBti0019386* frequency and latitude, suggesting that other factors might be involved in the observed clinal pattern. As expected, none of the climatic variables significantly correlated with TE frequency in Europe.

3.1.4 *FBti0019386* Is Associated with Up-regulation of sra in Female Flies

To shed light on the molecular mechanism of *FBti0019386* insertion, we measured the expression of *sra* and *Bin1* in non-stress conditions in embryos and in non-stress and cold-stress conditions in female flies with and without FBti0019386 insertion.
We did not observe significant differences in *sra* or *Bin1* expression in embryos differing by the presence/absence of *FBti0019386* insertion (t-test, p-value > 0.05) (Figure 3.1.6A and B).

However, we observed that adult female flies with *FBti0019386* insertion showed an increase of *sra* expression compared with flies without the insertion both in control conditions and after cold-stress conditions, although results were only significant under non-stress conditions (t-test, p-value = 0.03) (Figure 3.1.6C). On the other hand, no significant differences in expression level between flies with and without *FBti0019386* were observed for *Bin1* (t-test, p-value > 0.05) (Figure 3.1.6D).

Interestingly, we observed a change in *sra* and *Bin1* expression after cold stress in flies with and without *FBti0019386* insertion: *sra* is up-regulated in cold stress conditions (t-test, p-value < 0.05 in both cases) (Figure 3.1.6C) whereas *Bin1* is down-regulated (t-test, p-value < 0.05 in both cases) (Figure 3.1.6D).

Overall, we did not observe any change in expression of *sra* and *Bin1* in embryos, in agreement with the lack of phenotypic consequences of *FBti0019386* in this developmental stage. However, we observed an up-regulation of *sra* in flies with *FBti0019386* insertion that was significant under non-stress conditions. Moreover, we showed that both *sra* and *Bin1* changed their expression in response to cold stress.

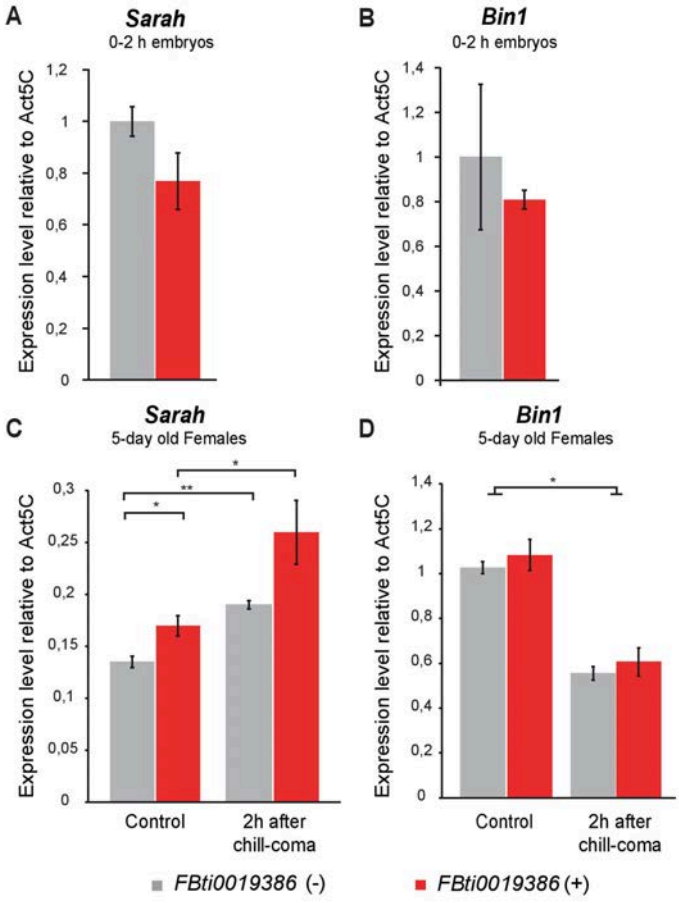


Figure 3.1.6 Flies with *FBit0019386* insertion showed sra upregulation. Real-time polymerase chain reaction quantification of sra and Bin1 transcript levels in outbred flies without *FBti0019386* insertion and with *FBti0019836* insertion. We represented the average expression level of sra (A and C) and Bin1 (B and D) relative to *Act5C* with SEM error bars for three biological replicates in 0–2h embryos and in 5-day-old females. Normalized expression measured 2h after chill-coma for *sra* and *Bin1* is depicted in (C) and (D), respectively.

3.1.5 *FBti0019386* Could Be Affecting gene Expression by Ectopically Assembling Heterochromatin

TEs from the *invader4* family contain sites with homology to PIWI interacting RNAs (piRNAs) that act as cis-acting targets for heterochromatin assembly by recruiting Heterochromatin Protein 1 a (HP1a) (Sentmanat and Elgin 2012). Specifically, these piRNA binding sites are located in the long terminal repeat (LTR) sequences. Because *FBti0019386* is a 347-bp solo-LTR, we hypothesized that it could be inducing the ectopic assembly of heterochromatin. We analyzed the 14.6-kb region containing *Bin1*, *sra*, and *FBti0019386* and found that both sense and antisense piRNAs bind specifically to *FBti0019386* (Figure 3.1.7A) (see Materials and Methods). Second, we tested whether there is evidence for the presence of HP1a binding to *FBti0019386* sequence. We found that HP1a specifically binds to *FBti0019386* sequence (Figure 3.1.7B) (see Materials and Methods). Thus, these results suggest that *FBti0019386* could be affecting gene expression by inducing the ectopic assembly of heterochromatin.

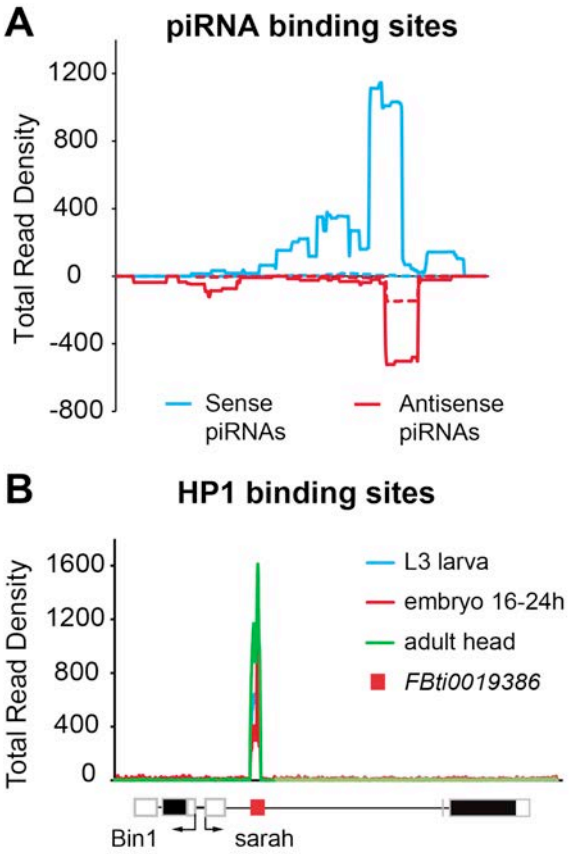


Figure 3.1.7 *FBti0019386* could bind piRNA and HP1a protein. (A) Mapping of piRNA sense and antisense RNA-seq reads against *FBti0019386* sequence. Data from Li et al. (2009) are depicted in dashed lines and data from Satyaki et al. (2014) are represented in continuous lines. (B) Mapping of reads coming from HP1a ChIP-Seq experimental data against the genome region containing *Bin1*, *FBti0019386*, and *sra*. Experimental data from L3 larva, 16–24h embryo, and adult heads are given.

3.2 RESULTS CHAPTER 2

3.2.1 Identifying Candidate Adaptive TEs

We performed a genome-wide screening of the *D. melanogaster* genome to identify TEs likely to be involved in adaptation. We looked for polymorphic TEs present at high frequencies in at least one of the out-of-Africa population analyzed (see below), and located in regions with high recombination rates (Comeron et al. 2012, Fiston-Lavier et al. 2010). We focused on polymorphic TEs so that it is possible to perform comparative functional experiments between flies with and without the candidate insertions. Besides, we focused on TEs located in high recombination regions because TEs present at high frequencies in regions with low recombination rates are more likely to be linked to an adaptive mutation rather than being the causal mutation (Hill and Robertson 1966, Smith and Haigh 1974, Charlesworth et al. 1993, Hudson and Kaplan 1995). Besides, purifying selection is low in these regions and thus slightly deleterious TEs could have reached high frequencies (Barrón et al. 2014, Castellano et al. 2015).

We analyzed not only TEs annotated in the reference genome, but also a subset of non-annotated TEs that were identified in DGRP strains by Rahman et al. (2015). To identify annotated candidate adaptive TEs, we estimated population frequencies of 815 TEs using T-lex2 (Fiston-Lavier et al. 2015) (see Material and Methods). We analyzed 280 *D. melanogaster* strains from four natural populations: two European populations, one from Bari (Italy) and one from Stockholm (Sweden), one North American population from North Carolina (DGRP), and one African population from Zambia (see Material and Methods). 577 out of the 815 TEs were polymorphic and 109 of the polymorphic TEs fulfilled our criteria and thus were considered as candidate adaptive TEs. 61 of the 109 TEs are present at low frequency in Africa and thus are likely to be involved in out-of-Africa adaptations, while 48 TEs are present at high frequencies in Africa and thus are likely to be involved in global adaptations (Annex Table S2.1).

To identify non-annotated candidate adaptive TEs, we analyzed a subset of 25 TEs previously identified in the DGRP strains using TIDAL (Rahman et al. 2015). This subset contains TEs located in high recombination regions and present at high frequencies according to Rahman et al. (2015). Because these TEs were not annotated in the reference genome but inferred using TIDAL, we first validated by PCR the presence of the TEs in DGRP strains. We were able to validate 20 out of 25 TEs (Annex Table S2.2). Thus, as previously reported, the majority of TIDAL predictions are likely to be real insertions (Rahman et al. 2015). We then estimated the TE frequencies based on a minimum of 7 strains per TE and considered as candidates 12 polymorphic TEs present at high frequencies and located in high recombination regions (Annex Table S2.1).

Thus overall, we identified 121 candidate adaptive TEs: 109 annotated TEs and 12 non-annotated TEs (Annex Table S2.1).

3.2.1.1 Candidate Adaptive TEs are Enriched for Truncated DNA Elements

We compared the genomic distribution, TE class identity, and TE length of the 109 annotated candidate adaptive TEs dataset vs the 577 annotated polymorphic TEs dataset. We found that most of the candidate adaptive TEs, are located inside genes or less than 1 kb from a gene (Table 3.2.1). Inside genes, most of the candidate adaptive TEs are located in the first intron suggesting that they might have some regulatory function (Cheng and Liang 2013; Park et al. 2014). However, there are no significant differences in the genomic distribution of candidate TEs compared with all polymorphic TEs (Table 3.2.1) (χ^2 test, p-value = 0.239).

				Inside Gene						
	Total TEs	**5'>1 kb**	**5'<1 kb**	**5' UTR**	**1st intron**	**Other intron**	**Exon**	**3' UTR**	**3'<1 kb**	**3'>1 kb**
Candidate adaptive TEs	**109**	8	10	3	33	25	3	5	14	8
Polymorphic TEs	**577**	54	42	12	181	141	10	24	46	67
Immune-related TEs	**16**	2	3	1	5	1	1	1	6	0

Table 3.2.1: Location of the candidate adaptive TEs, polymorphic TEs, and immune-related TEs annotated in the reference genome respect their nearby genes.

We found that the TE class identity of our candidate TEs differed from that of polymorphic TEs (χ^2 test, p-value << 0.001). While only 22% of the candidate TEs belongs to the LTR class, 52% of polymorphic TEs are LTRs (Table 3.2.2). On the other hand, 39% of the candidate adaptive TEs are DNA elements while only 17% of all polymorphic TEs belong to the DNA class (Table 3.2.2). Finally, we also checked whether TEs from the two datasets differed in

		TE class			TE length		
	Total TEs	**LTR**	**DNA**	**Non-LTR**	**Full-length**	**Truncated**	**% Canonical length**
Candidate adaptive TEs	**109**	24 (22%)	43 (39%)	42 (39%)	31 (28%)	78 (72%)	26
Polymorphic TEs	**577**	301 (52%)	99 (17%)	177 (31%)	295 (51%)	282 (49%)	35
Immune-related TEs	**16**	6 (32%)	5 (26%)	8 (42%)	4 (26%)	11 (74%)	17

Table 3.2.2: TE class and TE length of the candidate adaptive TEs, polymorphic TE, and immune-related TEs annotated in the reference genome. The percentage of TEs from each category is given between brackets. We considered as full-length TEs those TEs that conserve more than 95% of the canonical sequence. % Canonical length is the percentage of TE length conserved in the truncated TEs compared to the length of the canonical sequence.

length. We found that there are less full-length TEs in the candidate adaptive TE dataset compared to the polymorphic TEs (χ^2 test, p-value << 0.001). Moreover, truncated TEs from the candidate adaptive TEs are, on average, shorter than truncated TEs in the polymorphic TE dataset (Table 3.2.2).

Overall, although we found no differences in the genomic distribution of the candidate adaptive TEs compared to the polymorphic TEs, we found that candidate adaptive TEs have more DNA elements and less LTR elements, and that they are shorter in length.

3.2.1.2 Genes Located Nearby Candidate Adaptive TEs Are Enriched for Immune-Related Functions

We performed gene ontology (GO) analyses to check whether genes nearby candidate adaptive TEs were enriched for specific biological processes (see Material and Methods). DAVID annotation tool detected GO biological process information for a total of 85 genes associated with 74 TEs. We found two statistically significant enrichment clusters: the most significant cluster contains eight genes involved in response to biotic stimulus, and the second significant cluster contains 27 genes involved in transport and localization (Table 3.2.3). All the genes in the first cluster are related to immune response.

Cluster	Nº of genes	GO terms	p-value	Significant genes and associated TEs
Response to biotic stimulus ES = 1.5	8	GO:0043207 response to external biotic stimulus	0.0315	***pnr*** *(FBti0062242)*, ***cbx*** *(FBti0019985)*, ***Dif*** *(FBti0061506)*, ***Mef2*** *(FBti0018877)*, ***Dscam1*** *(FBti0061105)*, ***NUCB1*** *(FBti0020137)*, ***Tlk*** *(FBti0019564)*, ***AGO2*** *(FBti0020119)*
		GO:0009607 response to biotic stimulus	0.0315	
		GO:0051707 response to other organism	0.0315	
Transport and localization ES = 1.31	27	GO:0051234 establishment of localization	0.0215	***sgg*** *(FBti0019546)*, ***ken*** *(FBti0018868)*, ***nAChRalpha3*** *(FBti0019604)*, ***Cnx99A*** *(FBti0019453)*, ***TM4SF*** *(FBti0018868)*, ***GluClalpha*** *(FBti0019404)*, ***Kmn1*** *(FBti0019627)*, ***cindr*** *(FBti0020393)*, ***Vha16-3*** *(FBti0060715)*, ***Vha16-2*** *(FBti0060715)*, ***fab1*** *(FBti0019012)*, ***CG9413*** *(FBti0019056)*, ***MRP*** *(FBti0019158)*, ***Dscam1*** *(FBti0061105)*, ***CG8008*** *(FBti0018883)*, ***Indy*** *(FBti0020155)*, ***MFS9*** *(FBti0019410)*, ***lilli*** *(FBti0019112)*, ***CG30345*** *(FBti0018883)*, ***Frq2*** *(FBti0019079)*, ***AGO2*** *(FBti0020119)*, ***Sytbeta*** *(FBti0061417)*, ***Vps16A*** *(FBti0019344)*, ***Ppcs*** *(FBti0019400)*, ***Cngl*** *(FBti0019065)*, ***rdx*** *(FBti0019372)*, ***Bx*** *(FBti0019081)*
		GO:0006810 transport	0.0315	
		GO:1902578 single-organism localization	0.0453	

Table 3.2.3 Significant GO analysis results obtained with DAVID analyzing the genes associated with candidate TEs. ES: Enrichment Score. The p-value results from a modified Fisher's exact test (EASE score) (Huang et al. 2009).

Because not all genes nearby the candidate adaptive TEs have GO functional annotations, we further looked for additional functional information through literature searches (see Material and Methods). Taken together the information based on GO functional annotations and literature searches, we found functional information for the genes nearby 81 out of the 121 candidate adaptive TEs (Annex Table S2.3). 47 of these 81 TEs (58%) are associated with genes involved in stress response. Specifically, 19 TEs (23%) are associated with genes involved in immune response, 15 TEs (19%) are associated with genes involved in xenobiotic stress, and 14 TEs (17%) in oxidative stress (Annex Table S2.3). We also identified a considerable number of TEs associated with genes involved in cell signaling, behavior, or metabolism: 15 TEs (18%), 14 TEs (17%), and 12 TEs (15%), respectively (Annex Table S2.3).

Overall, we found that genes nearby candidate adaptive TEs are enriched for immune-related functions (Table 3.2.3). Additional functional information allowed us to identify a total of 19 candidate adaptive TEs located 21 nearby immune-related genes (Table 3.2.4). Because this is the most numerous subset, we decided to focus on TEs nearby immune-related genes for the rest of this work. Note that there are not significant differences between the dataset of candidate adaptive TEs and the 19 TEs located nearby immune-related genes in genomic distribution, class identity, or TE length (Table 3.2.1 and 3.2.2).

3.2.2 Functional Testing of Candidate Immune-Related Genes

The functional evidence for the majority of the 21 genes nearby the 19 candidate immune-related TEs comes from different types of experiments: transcriptional response to infection and/or survival experiments after infection (Table 3.2.4). The only exceptions are *TM4SF* and *Ken*, which are members of the JAK-STAT pathway that plays a role in *D. melanogaster* immune response (Myllymäke and Rämet 2014). For most of these genes, experimental evidence for their role in immune response was obtained infecting the flies with gram-negative bacteria (Table 3.2.4). Thus, to further confirm the role of these genes in immune response, we decided to perform survival experiments with laboratory mutant stocks (Table 3.2.5).

We used the gram-negative bacteria *Pseudomonas entomophila* a natural *D. melanogaster* pathogen (Vodovar et al. 2005). We focused on nine genes: six genes that did not have phenotypic evidence and three genes with phenotypic evidence obtained using a different pathogen (Table 3.2.5). We found that mutant strains of eight of these genes showed differences in survival after infection with *P. entomophila*: *NUCB1, CG2233,* and *Bin1* showed higher survival, *ken, CG8008, cbx* and *CG10943* mutants showed lower survival, and *TM4SF* mutants had higher survival in the firsts 30 hours and lower survival after that timepoint (Table 3.2.5). However, results were marginally significant for three genes: *Bin1*, *cbx,* and *CG10943*. Only for one gene, *CG15829,*

Table 3.2.4: Candidate TEs associated to immune-related genes. Gene functional envidence in immune response for the 16 annotated TEs (top) and the 3 non-annotated TEs (bottom).

TE	TE class	TE family	TE genomic position	TE length (bp)	TE distance to nearby gene(s)	TE position in the nearby gene	Gene immune-related evidences	Pathogen
FBti0018877	non-LTR	BS	2R: 9945496-9945626	131	0	first intron *Mef2*	**Survival and expression.** Adult *Mef2* mutant males are more sensitive to *E. cloacae* (gram-negative bacteria) and *M. marinum* (gram-positive bacteria) septic infection (Clark et al. 2013). Up-regulated after 4h of infection with *P. entomophila* (Bou Sleiman et al. 2015)	*E. cloacae* and *P.entomophila* (gram-negative bacteria) and *M. marinum* (gram-positive bacteria)
FBti0018883	LTR	Burdock	2R: 9151357-9157769	6413	136	3' *CG8008*	**Expression.** *CG8008* is induced by LPS (gram-negative bacteria) in an IKK-dependent manner in S2 cell cultures (Silverman et al. 2003). Up-regulated after *E.coli* (gram-negative bacteria) infection in S2 cells (Valanne et al. 2007).	LPS and *E. coli* (gram-negative bacteria)
FBti0019381	non-LTR	Juan	3R: 15132112-15135106	2995	180	5' *CG42788*	**Expression.** *CG42788* is down-regulated in response to gram-negative infection in virgin females (Short and Lazzaro 2013).	*P. rettgeri* (gram-negative bacteria)
FBti0019386	LTR	invader4	3R: 16189464-16189810	347	0	5'UTR *Bin1*	**Survival.** *Bin1* mutant larvaes are more sensitive to fungal *A. fumigatus* (fungi) infection (Costa et al. 2011).	*A. fumigatus* (fungi)
FBti0019457	DNA	pogo	3R: 29760415-29761560	1146	4434	5' *kay*	**Expression.** *kay* is a known component of the JNK pathway, which is essential for antimicrobial peptide release (Kleino et al. 2005; Kallio et al. 2005). *Kay* is up-regulated in *imd* and bsk mutant LPS-induced S2 cells, and down-regulated in *Rel* mutants (Kim et al. 2005). Kay is up-regulated in larvaes infected with gram-negative bacteria *P. entomophila* (Vodovar et al. 2005). Up-regulated after 4h of infection with *P. entomophila* (Bou Sleiman et al. 2015)	*P. entomophila* (gram-negative bacteria)
FBti0019602	non-LTR	Juan	X: 8031495-8035729	4249	12	3' *CG2233*	**Expression.** *CG2233* is down-regulated in *PEBP1* mutant L3 larvaes, which are more resistant to *M. luteus* (gram-positive bacteria) and *E. coli* (gram-negative bacteria) infection (Reumer et al. 2009).	*M. luteus* (gram-positive bacteria) and *E. coli* (gram-negative bacteria)
FBti0019985	LTR	roo	2R: 9871090-9871523	434	0	first intron *cbx*	**Survial.** *cbx* mutant flies are more sensitive to *S. aureus* (gram-positive bacteria) septic infection, but not to S. typhimurium (gram-negative bacteria) infection (Ayres et al. 2008).	*S. aureus* (gram-positive bacteria)
FBti0020046	non-LTR	Doc	3L: 6040416-6042720	2305	281	3' *Jon65Aiv*	**Expression.** *Jon65Aiv* is up-regulated after septic injury with mixed bacteria: *M. luteus* (gram-positive bacteria) and *E. coli* (gram-negative bacteria) (De Gregorio et al. 2012). Down-regulated after 4h of infection with *P. entomophila* (Bou Sleiman et al. 2015)	*M. luteus* (gram-positive bacteria) and *E. Coli* and *P. entomophila* (gram-negative bacteria)
FBti0020057	non-LTR	BS	3L: 7130011-7130136	126	338 / 739	3' *CG15829* / 5' *CG8628*	**Expression.** *CG15829* is up-regulated after infection by septic injury with mixed bacteria (gram-positive and gram-negative bacteria), and it is regulated by *Rel* (De Gregorio et al. 2002). Up-regulated after 4h of infection with *P. entomophila* (Bou Sleiman et al. 2015)// **Expression.** *CG8628* is up-regulated in microbiota assiciated flies vs germ free flies (Combe et al. 2014). Up-regulated after infection with several pathogens (gram-positive and gram-negative bacteria, fungi, protozoa) (Roxstrom-Lindquist et al. 2004). Down-regulated after 4h of infection with *P. entomophila* (Bou Sleiman et al. 2015)	mixed bacteria (gram-positive and gram-negative bacteria) // Different pathogens (gram positive and gram-negative bacteria, fungi, protozoa)

Table 3.2.4 (continued)

TE	TE class	TE family	TE genomic position	TE length (bp)	TE distance to nearby gene(s)	TE position in the nearby gene	Gene immune-related evidences	Pathogen
FBti0020119	DNA	S	3L: 15554974-15556705	1732	0	first intron *AGO2*	**Survival.** *AGO2* is involved in defense response to virus infections (Kemp et al. 2012), and interacts with Imd pathway proteins during gram-negative bacteria infection (Fukuyama et al. 2013).	RNA virus and *E. coli* (gram-negative bacteria)
FBti0020137	DNA	S	3L: 17799864-17801595	1732	0	first intron *NUCB1*	**Survival.** *NUCB1* mutants are more resistant to *V. cholerae* (gram-negative bacteria) oral infection (Berkey et al. 2009).	*V. cholerae* (gram-negative bacteria)
FBti0018868	LTR	297	2R: 23877783-23878196	414	1 / 340	5' *TM4SF* / 3' *ken*	**JAK-STAT.** *TM4SF* is a tetraspanin, which modulate immune-signaling in *Drosophila* (Levy and Shoham 2005). / **JAK-STAT.** *Ken* is a member of JAK-STAT pathway (Arbouzova et al. 2016). JAK-STAT pathway plays a role in immune response in *D. melanogaster* (Myllymäke and Rämet 2014).	Stress response and epithelium regeneration.
FBti0019564	LTR	mdg1	X: 3785867-3786055	189	0	intron *tlk*	**Phenotypic evidence.** *tlk* is involved in antimicrobial humoral response to gram-negative bacteria (Kleino et al. 2005). *tlk* knockdown, together with other five genes knocked-down, reduces phagocytosis of *E. coli* (gram-negative bacteria) and *S. aureus* (gram-positive bacteria) in S2 cells (Ulvila et al. 2011)	*E. coli* (gram-negative bacteria) and *S. aureus* (gram-positive bacteria)
FBti0061506	DNA	1360	2L: 17432071-17432118	48	0	first intron *Dif*	**Survival and expression.** *Dif* is the transcription factor involved in defense response to fungus and gram-positive bacteria and mediates Toll pathway activation (Rutschmann et al. 2000; Gobert et al. 2003; Brown et al. 2009; Christofi and Apidianakis 2013; Cornwell and Kirkpatrick 2001). Up-regulated in guts from *P. entomophila* infected flies (Bou-Sleiman et al. 2015).	gram-positive bacteria and fungi, *P. entomophila* (gram-negative bacteria)
FBti0061105	non-LTR	G5	2R: 7317828-7317878	51	46	3' *Dscam1*	**Expression.** *Dscam1* is involved in axon guidance and neuron development, detection of molecule of bacterial origin and phagocytosis (Watson et al. 2005)	*gram-negative bacteria (E. coli)*
FBti0062242	non-LTR	BS	3R: 16041234-16041335	102	0	3'UTR *pnr*	**Expression.** *pnr* is a modifier of the Toll pathway and *pnr* RNAi mutants show Imd pathway hyperactivation (Valanne et al. 2010)	*E. cloacae* (gram-negative bacteria) and *M. luteus* and *E. faecalis* (gram-positive bacteria)
tdn4	LINE	Jockey	2R: 18807871-18807898	800	479	3' CG15096	**Expression.** *CG15096* is down-regulated in Oregon R and *Rel*-mutant flies with microbiota compared to axenic flies (Broderick et al. 2014). It is down-regulated after *P. entomophila* infection (Bou-Sleiman et al. 2015).	flies with microbiota compared to axenic flies, *P. entomophila* (gram-negative bacteria)
tdn8	LTR	Gypsy	3L: 12863675-12863781	5,500	816	5' *CG10943*	**Expression.** *CG10943* is up-regulated in Oregon R and *Rel*-mutant flies with microbiota compared to axenic flies (Broderick et al. 2014). It is up-regulated 24h after infection with *O. muscaedomesticae* (protozoan) (Roxstrom-Lindquist et al. 2004). It is up-regulated after *P. entomophila* infection (Bou-Sleiman et al. 2015).	*O. muscaedomesticae* (protozoan), microbiota compared to axenic flies, *P. entomophila* (gram-negative bacteria)
tdn17	DNA	pogo	X: 21399382-21399471	1,000	2067	5' *lcs*	**Expression.** *lcs* is high up-regulated in young flies gut compared to old flies (Broderick et al. 2014). Involved in virus response, downregulated in males (Carpenter et al. 2009)	sigma virus (Rhabdoviridae)

mutant flies did not show differences in survival compared to a wild-type strain with a similar genetic background. Overall, we provide additional evidence for five of the six genes for which no phenotypic evidences were available and for the three genes that were previously tested with a different pathogen (Table 3.2.5).

Gene	Mutant type	Previous evidence	Survival experiment	p-value
NUCB1	PBac{PB} insertion	Survival (different pathogen)	Higher survival	0.006
CG2233	RNAi knockdown	Expression	Higher survival	0.0012
Bin1	Gal4/UAS overexpression	Survival (different pathogen)	Higher survival	0.044
ken	P{PZ} insertion	JAK-STAT	Lower survival	0.003
CG8008	Mi{MIC} insertion	Expression	Lower survival	0.031
TM4SF	RNAi knockdown	JAK-STAT	Lower survival*	0.00014
cbx	PBac{PB} insertion	Survival (different pathogen)	Lower survival	0.041
CG10943	Mi{MIC} insertion	Expression	Lower survival	0.045
CG15829	RNAi knockdown	Expression	No differences	0.136

Table 3.2.5: Mutant survival experiments results. Survival of mutant strains orally infected with *P. entomophila* compared to flies with a similar background. p-values obtained from log-rank survival test. OR (CI): odds ratio and confidence interval (95%) calculated when 50% of the flies from the sensitive background strain was dead. **TM4SF* RNAi knockdown flies had a higher survival compared to wild-type flies with a similar background in the first 30 hours of infection and, after that, they showed lower survival.

3.2.2.1 Immune-Related Candidate TEs are Associated with Gene Expression Changes

In order to explore whether the 19 candidate TEs were associated with expression changes of their nearby immune-related genes, we used allele-specific expression (ASE). ASE allows analyzing gene expression differences associated with *cis*-regulatory changes in the same genomic sample, thus, avoiding possible effects due to *trans*-regulatory changes (Wittkopp et al. 2004). For five of the 19 TEs we could not perform ASE, so overall we were able to analyze 16 genes located nearby 14 TEs (see Material and Methods) (Figure 3.2.1, Annex Table S2.4). We analyzed the expression in female fly guts both under non-infected conditions and 12 hours after infection with *P. entomophila*. We performed the analysis in flies with two different genetic backgrounds in order to detect possible background-dependent effects in gene expression changes (Figure 3.2.1, Annex Table S2.4).

In non-infected conditions, 10 out of the 16 genes showed statistically significant allele-specific expression differences in at least one of the two genetic backgrounds analyzed (Figure 3.2.1, Annex Table S2.4). For five genes, we found that the allele with the TE was up-regulated, and for five genes the allele with the TE was down-regulated.

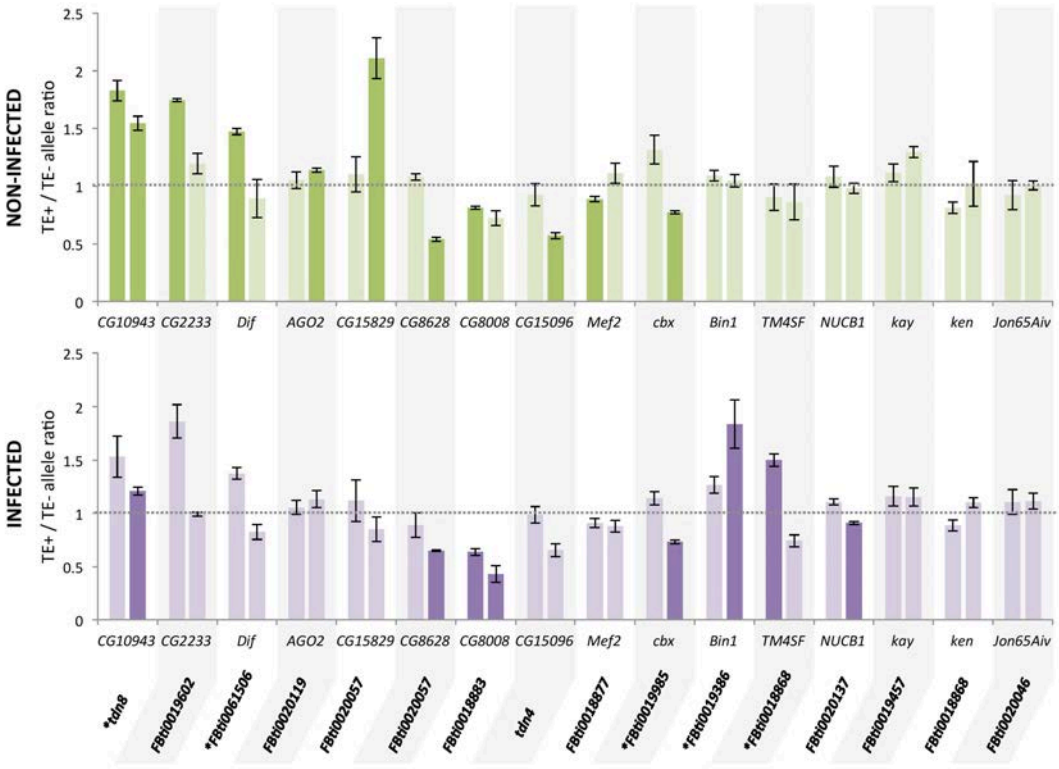


Figure 3.2.1. Allele-specific expression analysis. Results from female guts in non-infected conditions (in green) and in infected conditions (in purple). Each bar represents the average ratio of gene expression levels between the allele with the TE and the allele without the TE of the three replicas. Each gene has two bars representing each one of the two genetic backgrounds analyzed. Statistically significant differences are depicted as dark color (t-test p-values < 0.05, corrected for FDR). Error bars represent SEM. *TE further analyzed in this chapter.

In infected conditions, 7 out of the 16 genes showed statistically significant allele-specific expression differences in at least one of the two genetic backgrounds analyzed (Figure 3.2.1, Annex Table S2.4). For three genes, we found that the allele with the TE was up-regulated, and for four genes the allele with the TE was down-regulated.

Considering both non-infected and infected conditions, we found that the allele with the TE showed expression changes in the same direction in four genes (Figure 3.2.1, Annex Table S2.4). On the other hand, the allele with the TE was associated with expression changes only in non-

infected conditions for six genes and with changes only in infected conditions for three genes (Figure 3.2.1, Annex Table S2.4).

Our results also allowed checking whether differences in expression were background dependent. 17 out of the 32 expression analysis (16 genes in two conditions) gave the same result in the two backgrounds. In eight analyses, both backgrounds showed changes in expression in the same direction (up-regulated or down-regulated), although results were only statistically significant in one of the backgrounds. Finally, only seven analyses differed in the direction of the change of expression in the two backgrounds, although results were always statistically significant in only one of the backgrounds analyzed (Figure 3.2.1, Annex Table S2.4).

Overall, we found that most of the candidate immune-related TEs are associated with changes in expression of their nearby gene in at least one of the two conditions analyzed (Figure 3.2.1). While some expression changes are significant only in infected or only in non-infected conditions, a significant proportion of genes (31%) showed consistent changes in expression in both conditions. Finally, most of the analyses were either the same in the two backgrounds analyzed or consistent in the two backgrounds but statistically significant in only one of them (Figure 3.2.1, Annex Table S2.4).

3.2.2.2 Most of the TEs Are Likely to Be Responsible for the Expression Change in the Nearby Immune-Related gene

To further test whether the candidate adaptive TEs are responsible for the changes in expression of their nearby genes, we first checked whether there was any other polymorphism linked to the TE in the gene coding region or in the 1kb TE flanking regions. Only for the *AGO2* gene, we found two SNPs in the coding region that were linked to the TE insertion (Supplementary File 6). *AGO2* is a gene showing a fast rate of adaptive amino acid substitutions (Obbard et al. 2006; Obbard et al. 2009), and it is associated with recent selective sweep (Obbard et al. 2011). However, it is still not clear which is the genetic variant that is under positive selection (Obbard et al. 2011).

We then performed structural analysis and/or enhancer assays for a subset of TEs. We focused on five TEs: *FBti0019386* and *FBti0018868* that were associated with expression changes only in infected conditions, *FBti0061506* associated with expression changes only in non-infected conditions, and *tdn8* and *FBti0019985* associated with expression changes both in infected and non-infected conditions (Figure 3.2.1, Annex Table S2.4).

3.2.2.3 *FBti0019386* Provides a TSS to *Bin1* that Is Only Used in Infected Conditions

FBti0019386 is inserted in the 5'UTR region of *Bin1*, a gene required for the expression of immune and stress response genes (Costa et al. 2011) (Table 3.2.4). There is previous experimental evidence suggesting that *FBti0019386* adds a transcription start site (TSS) to *Bin1*: two of the three *Bin1* transcripts overlap 101 bp with *FBti0019386* (Figure 3.2.2A) (Batut et al. 2013). We thus performed RT-PCRs to detect whether flies homozygous for the presence and

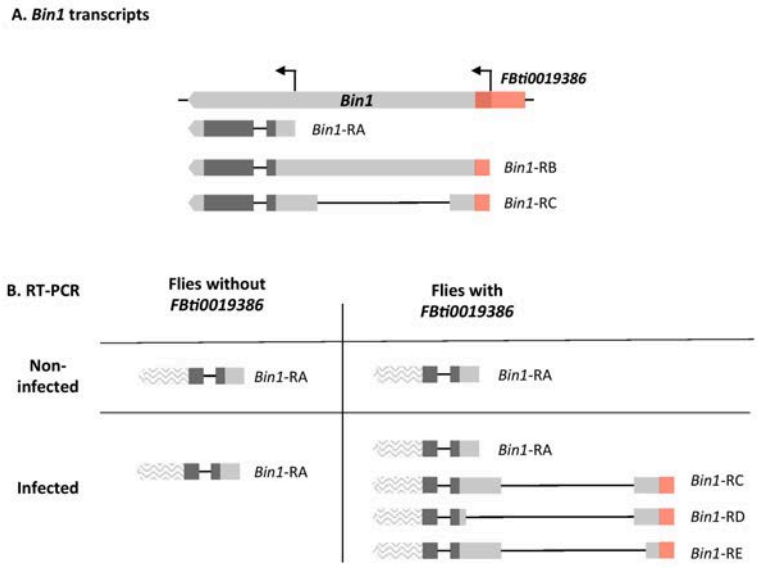


Figure 3.2.2. *FBti0019386* adds a new TSS to its nearby gene *Bin1*. Non-coding regions are depicted in grey, coding regions are depicted in black, and the TE is represented in red. (A) Transcripts annotated for *Bin1*. *FBti0019386* overlaps with two of the annotated transcripts. (B) Transcripts detected by RT-PCR in flies with and without *FBti0019386*, both in non-infected and infected conditions. Transcript regions wave-patterned are inferred from Flybase transcript annotation and were not sequenced in this work. *Bin1-RD* and *Bin1-RE* transcripts are, respectively, 318 bp and 172 bp shorter compared to *Bin1-RC* transcript.

for the absence of *FBti0019386* expressed different transcripts in non-infected and infected conditions. In non-infected conditions, we found that both flies with and without *FBti0019386* expressed only the short *Bin1-RA* transcript (Figure 3.2.2B). In infected conditions, we found differences between flies with and without *FBti0019386*. Flies without *FBti0019386* insertion only express *Bin1-RA*, while flies with *FBti0019386* express four different transcripts: *Bin1-RA*, and three transcripts starting in the TE: *Bin1-RC*, *Bin1-RD* and *Bin1-RE*. We confirmed these results by performing the experiments in a second genetic background (Figure 3.2.2B). Note that the later two transcripts were not described previously and differ in the size of the 5'UTR (Figure 3.2.2B).

Overall, we found that *FBti0019386* adds a TSS for *Bin1* that it is only used in infected conditions. While we were not able to detect *Bin1-RB* transcript, we found two previously not annotated transcripts (Figure 3.2.2B). Our results are in agreement with ASE results that showed that *FBti0019386* is associated with increased *Bin1* expression only in infected conditions in the two backgrounds analyzed (Figure 3.2.1).

3.2.2.4 *tdn8* Drives Expression of *CG10943* in Non-Infected and Infected Conditions

tdn8 is located 816 bp upstream of *CG10943*, a gene that is up-regulated in response to an immune challenge with different pathogens including *P. entomophila* (Figure 3.2.3A) (Broderick et al. 2014; Roxstrom-Lindquist et al. 2004; Bou-Sleiman et al. 2015). To test whether *tdn8* could be acting as an enhancer, we generated two reporter gene constructs with the *CG10943* upstream region in front of *lacZ* gene one including the *tdn8* insertion and another without the

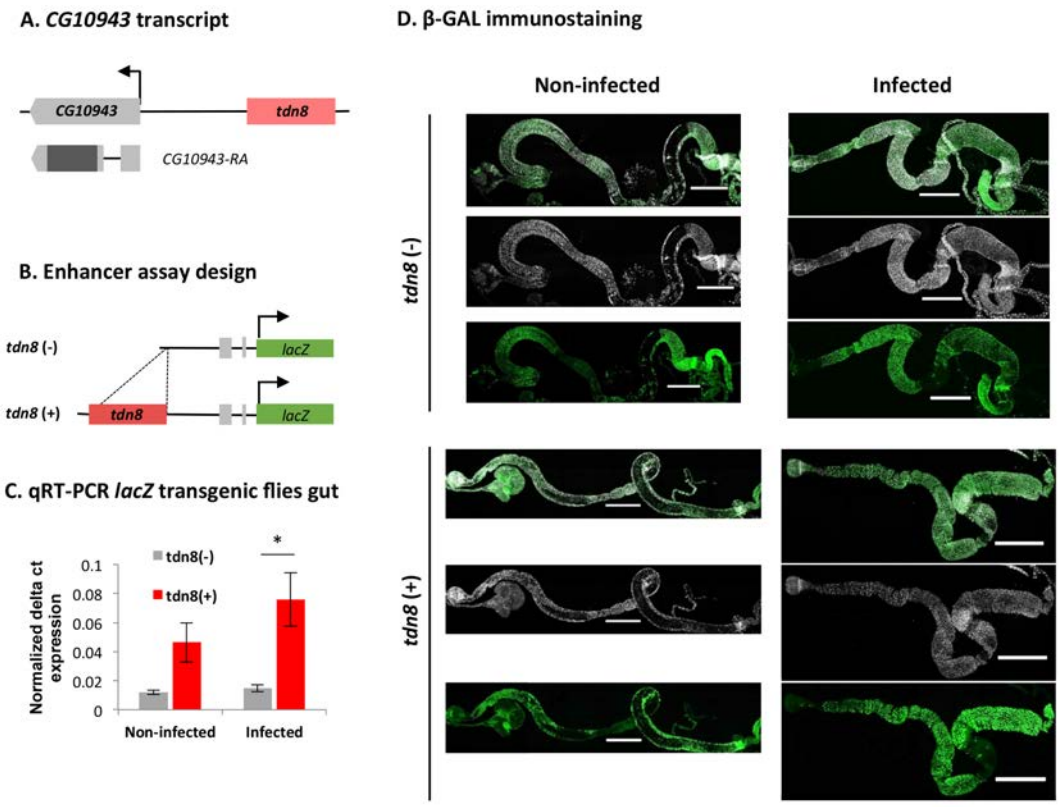


Figure 3.2.3. *tdn8* act as an enhancer regulatory sequence. (A) *tdn8* is located upstream the gene *CG10943*. (B) Vector construction without *tdn8* and with *tdn8* in the promoter region of the reporter gene *lacZ*. (C) Expression levels of the reporter gene *lacZ* in transgenic female guts without *tdn8* (in grey) and with *tdn8* (in red), both in non-infected and in infected conditions. (D) β-GAL immunostaining (in green), and DAPI staining (in grey) from female non-infected and infected guts. The scale bar represents 500 µm.

tdn8 insertion (Figure 3.2.3B). We found that transgenic strains with the upstream region of *CG10943* containing *tdn8* showed more expression than transgenic strains without the insertion, both in non-infected and in infected conditions, although these differences were only statistically significant in infected conditions (t-test, p-value = 0.095 and p-value = 0.046 respectively) (Figure 3.2.3C). We also checked whether the transgenic strains with and without *tdn8* differed in the localization of the *β-GAL* protein expression. We found no differences in non-infected or infected conditions (Figure 3.2.3D).

Overall, we found that *tdn8* is acting as an enhancer. These results are in agreement with our ASE results that showed that *tdn8* is associated with up-regulation of *CG10943* both in non-infected and infected conditions in the two genetic backgrounds analyzed (Figure 3.2.1).

3.2.2.5 *FBti0061506* Does Not Drive the Expression of a Reporter Gene

FBti0061506 is located in the 5'UTR intron of one of the four transcripts of the gene *Dif*, *Dif-RD*, and 3.8 kb upstream of the other three transcripts *Dif-RA*, *Dif-RB*¸ and *Dif-RC* (Figure 3.2.4A). All *Dif* transcripts are annotated as weakly supported, except *Dif-RB* that is strongly

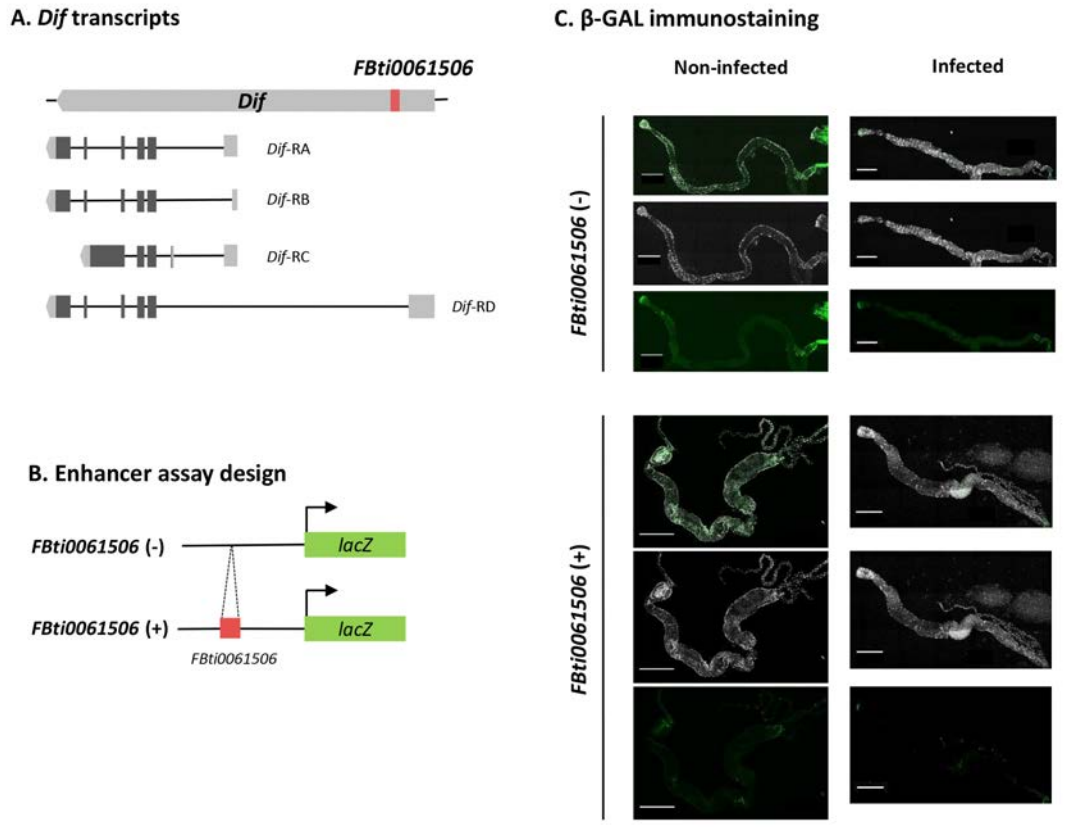


Figure 3.2.4. *FBti0061506* does not drive the expression of the reporter gene. (A) *FBti0061506* is located in the first intron of one of *Dif* transcripts, and upstream of the other transcripts. (B) Vectors construction for the enhancer assays with and without *FBti0061506* in front of the reporter gene *LacZ*. (C) β-GAL immunostaining (in green), and DAPI staining (in grey) from female non-infected and infected guts. The scale bar represents 500 µm.

supported (Gramates et al. 2017). Although *Dif* is a main transcription factor of the Toll-pathway, involved in gram-positive bacteria infection response (Gobert et al. 2003, Brown et al. 2009, Lemaitre and Hoffmann 2007), it was also found to be up-regulated in gut tissue after *P. entomophila* infection (Bou-Sleiman et al. 2015).

In order to study whether *FBti0061506* could act as an enhancer sequence, we generated two reporter gene constructs with part of the *Dif* intron were *FBti0061506* is inserted differing by the presence/absence of this insertion (see Material and Methods) (Figure 3.2.4B). None of the two gene constructs affected the expression of the reporter gene or the localization of the *β-GAL* protein (Figure 3.2.4C).

Overall, our results do not provide evidence for a role as an enhancer of *FBti0061506*. However, our ASE results showed that *FBti0061506* was associated with *Dif* up-regulation in non-infected conditions in one of two genetic backgrounds analyzed (Figure 3.2.1). It could be that the effect of *FBti0061506* is context depended. Therefore, a bigger genomic region with and without the insertion should be analyzed to discard an effect of *FBti0061506* on *Dif* expression. However, it might also be possible that the *Dif* expression change detected with ASE is due to a *cis*-mutation different from the *FBti0061506* insertion.

3.2.2.6 *FBti0018868* Adds a TSS Both in Infected and Non-Infected Conditions

FBti0018868 is annotated 1 bp upstream of one of the three *TM4SF* transcripts, and 310 bp upstream of the other two transcripts (Figure 3.2.5A). However, a previous genome-wide screening identified a new TSS for *TMSF* inside *FBti0018868* (Batut et al. 2013). We performed RT-PCR to check whether flies homozygous for the presence of *FBti0018868* expressed the transcript starting in the TE. We detected the presence of the transcript starting in *FBti0018868* in fly guts (Figure 3.2.5B). We further checked whether flies with the TE differed in the presence of this transcript in non-infected and infected conditions. However, we detected the presence of the transcript starting in the TE both in non-infected and in infected conditions.

We designed enhancer assays in order to test whether *FBti0018868* could affect the expression of its nearby gene *TM4SF* (Figure 3.2.5C). For that, we generated transgenic flies cloning the TE sequence in front of the reporter gene *lacZ*, and we checked *lacZ* expression both in non-infected and infected conditions. As a negative control, we generated transgenic strains with an empty vector carrying only a minimal promoter in front of *lacZ* gene (See Material and Methods). We did not detect *lacZ* expression by qRT-PCR in any of the transgenic strains in non-infected or in infected conditions. We also could not detect differences in *β-GAL* protein expression

localization comparing transgenic flies with *FBti0018868* with transgenic flies with the empty vector in any of the conditions (Figure 3.2.5D).

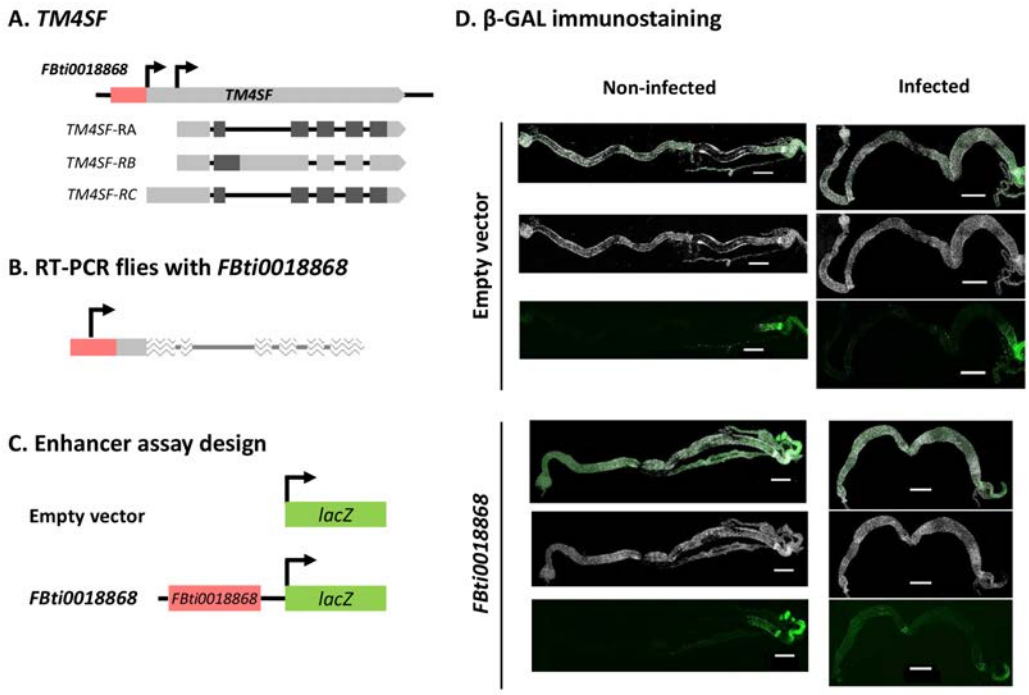


Figure 3.2.5. *FBti0018868* adds a new TSS to its nearby gene *TM4SF*. (A) *FBti0018868* is located upstream the gene *TM4SF* and it has three annotated transcripts described. (B) We detected a new *TM4SF* transcript overlapping with *FBti0018868* both in flies in non-infected and infected conditions. Transcript regions wave-patterned are inferred from Flybase transcript annotation and were not sequenced in this work. (C) Vector constructions with the empty vector as a negative control, and a vector carrying *FBti0018868* in front of the reporter gene *lacZ*. (D) β-GAL immunostaining (in green), and DAPI staining (in grey) of guts from transgenic strains guts with the empty vector and with *FBti0018868*. The scale bar represents 500 µm.

Overall, we found that *FBti0018868* adds a TSS for its nearby gene *TM4SF* that is used both in infected and non-infected conditions. On the other hand, we did not find evidences for *FBti0018868* affecting the expression of a reporter gene suggesting that the TE affects only the transcript structure. Thus, our current results do not explain the changes in *TM4SF* expression found only in infected conditions using ASE (Figure 3.2.1). However, transcript-specific qRT-PCRs could be performed to check whether flies with and without *FBti0018868* differed in the expression level of the different *TM4SF* transcripts in infected and non-infected conditions.

3.2.2.7 *FBti0019985* Drives the Expression of *cbx* Both in Non-Infected and Infected Conditions

FBti0019985 is located in the first 5'UTR intron of *cbx-RA* transcript, and 700 bp and 5.5 kb upstream of the other two annotated transcripts *cbx-RC* and *cbx-RB*, respectively (Figure 3.2.6A). We first checked whether the TE affects the expression of the different *cbx* transcripts by performing RT-PCR from non-infected guts of homozygous strains for the presence or the absence of the TE. We detected two of the three annotated transcripts, *cbx-RB* and *cbx-RC*, in both flies with and without *FBti0019985* (Figure 3.2.6A). Thus, we did not find evidences of *FBti0019985* affecting transcript choice or structure in non-infected conditions in the first background analyzed. *FBti0019985* could be acting as an upstream enhancer for *cbx-RB* and *cbx-RC* transcripts. Thus, we performed enhancer assays by generating transgenic strains with the TE sequence in front of the reporter gene *lacZ* (See Material and Methods, Figure 3.2.6B).

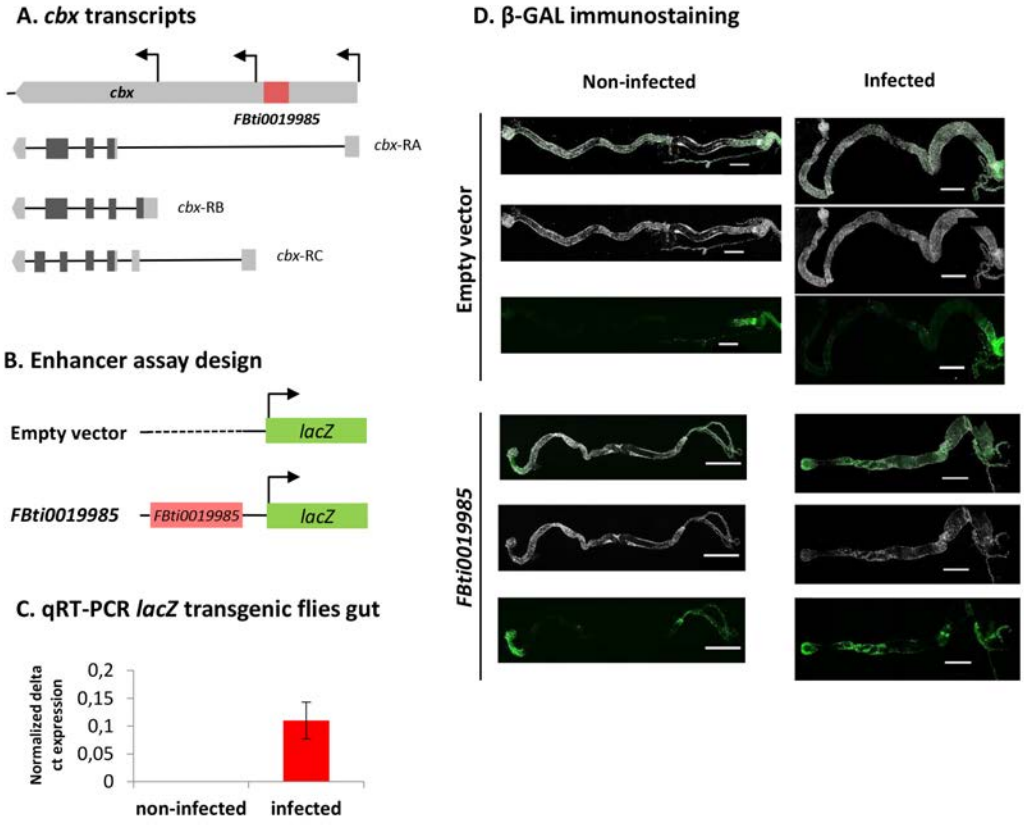


Figure 3.2.6 *FBti0019985* act as an enhancer regulatory sequence. (A) Transcripts annotated for the gene *cbx*, associated to *FBti0019985*. (B) Vector constructs for the enhancer assays. (C) Expression levels of the *lacZ* reporter gene both under non-infected and infected conditions. Empty vector showed no detectable expression levels in any of both conditions. (D) β-GAL immunostaining (in green), and DAPI staining (in grey) of guts from transgenic strains with the empty vector and with *FBti0018868*. Scale bar represents 500 µm.

As a negative control, we used the transgenic strains carrying the empty vector (see Material and Methods). We found that *FBti0019985* drives the expression of the reporter gene only in infected conditions (Figure 3.2.6C). We also checked the *β-GAL* protein expression localization in the guts by performing *β-GAL* immunostaining. In this case, we detected expression both in non-infected and in infected conditions localized in the anterior part of the gut (Figure 3.2.6D). The localization of the expression only in the anterior part of the gut could explain why we could not detect expression with the qRT-PCR of whole guts in non-infected conditions.

Overall, we showed that *FBti0019985* does not modify transcript structure under non-infected conditions. We also showed that the TE sequence act as an enhancer in the anterior part of the gut. The enhancer assays are in agreement with the ASE results for one of the ASE genetic backgrounds, although these results were only marginally significant before applying FDR correction (Annex Table S2.4). However, the observation that the enhancer effect of the TE is restricted to the anterior part of the gut, could explain why we cannot detect statistically significant differences analyzing whole gut expression. Further experiments restricted to the anterior part of the gut should be performed to test this hypothesis. Besides, we also found that *FBti0019985* was associated with *cbx* down-regulation. Our current results do not provide an explanation for the association between the presence of the insertion and *cbx* down-regulation.

04

DISCUSSION

4. DISCUSSION

In this thesis, we have identified and characterized the role of several candidate adaptive TEs in *Drosophila melanogaster* natural populations. To do that, we followed two different strategies: locus-specific and trait-specific (Figure 4). In the first chapter, we have characterized both at the molecular and phenotypic level a previously identified adaptive TE insertion (González et al. 2008; González et al. 2010). First, we provided more evidence supporting an adaptive role for this TE by elucidating its evolutionary history. Then, we have explored several ecologically relevant phenotypes associated with this insertion (Ullastres et al. 2015).

In the second chapter, we have studied the impact of several TE insertions in a highly conserved and ecologically relevant trait: the immune response. To do that, we first performed a new genome-wide screening in order to identify a bigger dataset of candidate TEs involved in adaptation. By increasing the number of populations and the number of TEs analyzed, we were able to increase the number of identified candidate TEs. Interestingly, we found that genes associated with those TEs are enriched for immune-related functions. We were also able to associate the candidate TEs with gene expression changes, and determine some of the molecular mechanisms behind these expression changes.

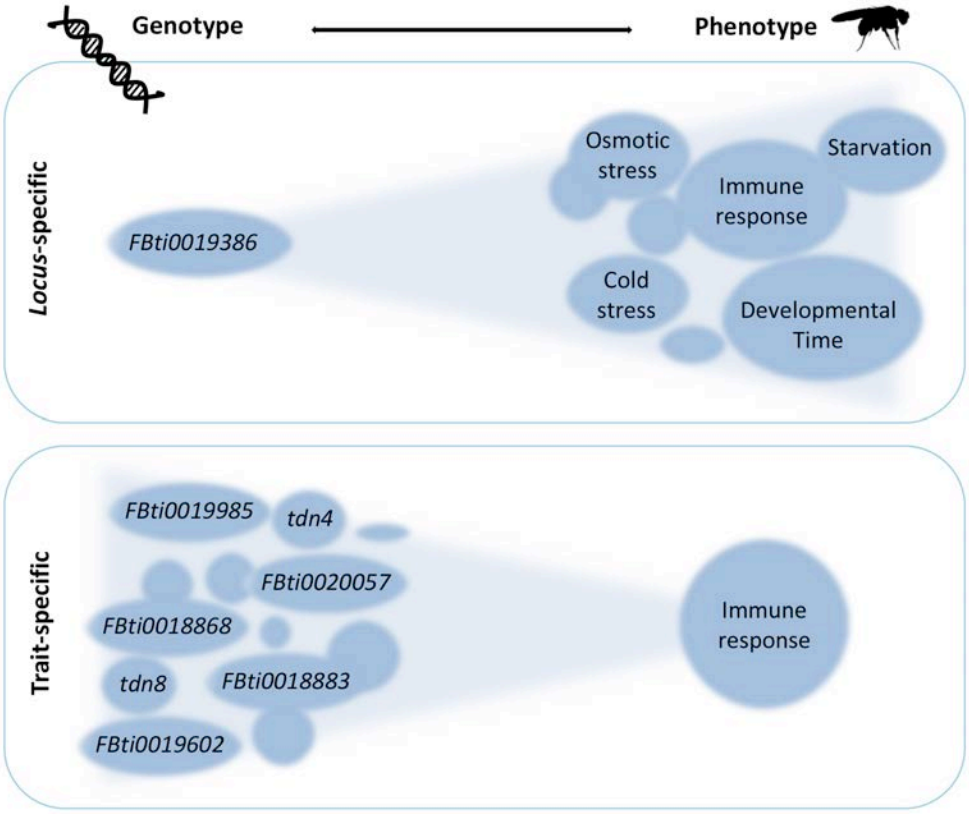


Figure 4. Strategies followed in this thesis for the characterization of the candidate TEs for adaptation.

4.1 Exploring the Phenotypic Space and the Evolutionary History of the Natural *FBti0019386* Insertion in *Drosophila melanogaster*

FBti0019386 was previously identified as a candidate insertion likely involved in adaptation to temperate environments (González et al. 2008, González et al. 2010). We have explored the phenotypes associated with *FBti0019386* in different developmental stages, embryo and adult, and in different environmental conditions, non-stress conditions and cold, osmotic, and starvation stress conditions. Overall, we found that *FBti0019386* mediates sensitivity to cold stress conditions and is associated with faster developmental time (DT) (Figures 3.1.3 and 3.1.4). These two phenotypic effects have plausible fitness consequences in nature that could explain why the mutation increased in frequency in natural populations but has not reached fixation. Increased sensitivity to cold stress conditions is likely to reduce fitness of the flies that carry *FBti0019386* insertion, and may represent the cost of selection of this mutation. On the other hand, faster DT is likely to increase the fitness of flies with *FBti0019386* insertion. Thus, it is plausible that *FBti0019386* increased in frequency in natural populations because of its positive effect on DT whereas it did not reach fixation because of its negative effect on cold-stress resistance. Our results emphasize the importance of exploring different phenotypes to fully characterize the effects of natural mutations, as have been suggested before (Mackay et al. 2010; Guio et al. 2014). Although our results provide a plausible explanation for the effect of *FBti0019386* insertion in natural populations, experiments under natural conditions are needed to unequivocally identify the effect of this insertion in nature.

4.1.1 *FBti0019386* Has Signatures of Positive Selection and It Is Not Involved in Temperate Climate Adaptation

By combining several tests that capture different signatures of selection at the DNA level, we demonstrate that *FBti0019386* shows signatures of positive selection suggesting that it is an adaptive mutation (Table 3.1.1, Annex Figure S1). However, our results also suggest that *FBti0019386* might not be involved in temperate adaptation as has been previously proposed (González et al. 2010). First, adaptation to temperate climates has been associated with increased stress resistance, increased DT and decreased fecundity (Stanley and Parsons 1981; Hoffmann et al. 2003; Schmidt et al. 2005; Folguera et al. 2008; Schmidt and Paaby 2008) but see also (James and Partridge 1995; James et al. 1997; Trotta et al. 2006). However, we found that *FBti0019386* is associated with increased sensitivity to cold stress (Figure 3.1.3), with shorter DT (Figure 3.1.4) and does not significantly affect fecundity (Figure 3.1.1). Thus, the phenotypic effects of *FBti0019386* are not consistent with a role of this insertion in temperate adaptation. Second, our global analyses of *FBti0019386* population frequency showed that *FBti0019386* frequency correlates with latitude and with climatic variables in North America and in Australia

but not in Europe (Figure 3.1.5, Table S1.6). We suggest that the clinal frequency patterns in North America and in Australia could be due to the dual colonization of these two continents by European and African populations rather than to the operation of spatially varying selection (Caracristi and Schlotterer 2003; Rouault et al. 2004; Duchen et al. 2013; Bergland et al. 2014). The lack of clinal frequency patterns in Europe would support this conclusion. However, it is also possible that phenotypic effects of *FBti0019386* not yet characterized could be consistent with a role of this natural mutation in temperate adaptation. Additionally, although there is evidence for the presence of clinal variation in European populations (David et al. 1985, 1986, 1998; Costa et al. 1992), other works have shown that clines are weaker in Europe compared to other continents (Oakeshott et al. 1983; Oakeshott et al. 1983). This could be partly due to differences in the latitudinal ranges spanned by populations analyzed in the different continents. In this work, the latitudinal range spanned by North American (25.82º to 45.06º) and Australian (-16.88º to -42.83º) populations is larger than the range spanned by European populations (41.13º to 59.33º). In any case, genome-wide scan studies that identify loci that are differentiated between populations should be taken as a first step towards the identification of loci that are subject to spatially varying selection (González et al. 2010; Kolaczkowski et al. 2011; Fabian et al. 2012; Reinhardt et al. 2014). Further functional validation should be gathered before concluding that the candidate loci are under spatially varying selection (Bergland et al. 2014).

4.1.2 *FBti0019386* Is Associated with *sra* Up-Regulation

Our results also shed light on the molecular processes that link genotype to phenotype variation. We found that *FBti0019386* is associated with up-regulation of *sra* (Figure 3.1.6C). As previously described for other elements from the *invader4* family, we showed that *FBti0019386* has piRNA binding sites (Figure 3.1.7A) (Sentmanat and Elgin 2012). We also showed that HP1a binds specifically to the *FBti0019386* sequence, further suggesting that *FBti0019386* could be inducing the ectopic assembly of heterochromatin (Figure 3.1.7B). These results highlight the potential role of TE remnants as silencing signals to be used by piRNAs to direct heterochromatin formation (Sentmanat et al. 2013). Although we observed an up-regulation of *sra* in adult females, we cannot discard that heterochromatin assembly induced by *FBti0019386* could be affecting gene expression in other developmental stages and/or specific tissues. In the case of *Bin1*, we did not find differences in expression associated with the presence of *FBti0019386* in adult females and embryos. Other developmental stages, tissues, or environmental conditions should be explored in order to discard a role of *FBti0019386* in *Bin1* expression regulation. Indeed, we later found that *FBti0019386* is associated with *Bin1* up-regulation after infection in adult females (see Results Chapter 2).

Although *sra* and *Bin1* have not been associated with DT, both genes play important roles during development and have been associated with a wide range of biological processes (Chang et al. 2003; Ejima et al. 2004; Horner et al. 2006; Takeo et al. 2006, 2010; Chang and Min 2009; Matyash et al. 2009; Costa et al. 2011; Nakai et al. 2011). A genome-wide screening looking for genes influencing DT in *D. melanogaster* has shown that the many candidate genes were involved in a wide range of biological processes such as cellular metabolic processes, organismal development, and response to stress (Mensch et al. 2008). More recently, developmental timing in insects has been associated with hormonal and circadian control (Di Cara and King-Jones 2013; Yadav et al. 2014). Interestingly, *sra* is regulated by Shaggy/GSK-3 β (*sgg*), a Ser–Thr kinase involved in the regulation of circadian rhythmicity (Martinek et al. 2001). On the other hand, both *Bin1* and *sra* are stress-response genes: *Bin1* is up-regulated in response to stress and *sra* is down-regulated (Figure 3.1.6). *Bin1* is a known key player in transcriptional response to environmental stress (Costa et al. 2011). Although there was no previous evidence for a direct role of *sra* in response to stress, *sra* could be affecting stress response through its role in the calcium pathway (Takeuchi et al. 2009; Teets et al. 2013; Davies et al. 2014). *sra* inhibits calcineurin, a highly conserved protein in eukaryotes that has the ability to sense calcium (Hogan et al. 2003). Although it is not deeply understood, calcium pathways play a role during general cell-stress response including cold stress response (Takeuchi et al. 2009; Teets et al. 2013; Davies et al. 2014). Note that many genes that affect complex traits in *Drosophila* had well-characterized roles in early development and were not previously annotated to affect adult quantitative traits (Mackay 2010). *FBti0019386* adds to the growing list of TE-induced adaptive mutations that have been linked to their fitness effects and their underlying molecular mechanisms in *Drosophila melanogaster* (Schmidt et al. 2010; Magwire et al. 2011; Guio et al. 2014; Mateo et al. 2014; Merenciano et al. 2016; Le Mahn et al. 2017).

4.2 Genome-Wide Screening for Candidate TEs Involved in Adaptation

In the second chapter of this thesis, we followed a trait-specific strategy: we looked for candidate TEs involved in adaptation to a specific trait. To do that, we first performed a genome-wide screening looking for candidate adaptive TE insertions in order to explore what traits are more likely to be under selection in *D. melanogaster* natural populations. A previous genome-wide screening in one African natural population and a few North American *D. melanogaster* natural populations, identified a total of 13 candidate TEs for out-of-Africa adaptation (González et al. 2008). This study was based on PCR estimations of the frequencies of a total of 902 TEs annotated in the reference genome. In a follow-up study, the same authors analyzed latitudinal frequencies in North American and Australian populations and ended up with a list of 18 candidate TEs for out-of-Africa adaptation (González et al. 2010). However, authors claimed that the screening was probably underestimating the number of candidate TEs, as other TEs

might be contributing to adaptation in other out-of-Africa populations, as well as in populations with different environments (González et al. 2008).

In this thesis, we performed a similar genome-wide screening by considering more natural populations and more TEs. Thus, while González and collaborators analyzed only one population from North America and one from Africa, we also added two European populations. Moreover, we analyzed a different North American population from North Carolina (Mackay et al. 2010), and a different African population from the ancestral range of *D. melangoaster*. We analyzed a total of 1,630 TEs annotated in the *D. melanogaster* reference genome (Annex Table S2.8), that is, 728 TEs more compared to the previous screening. Moreover, instead of performing PCRs to identify the TE frequencies in the populations analyzed, we run the software T-lex2 (Fiston-Lavier et al. 2015) with the NGS data available for the populations analyzed. Note that T-lex2 software identifies only TEs that are annotated in the reference genome (Fiston-Lavier et al. 2015). Thus, we cannot detect other insertions present in the genomes of the populations analyzed. *D. melanogaster* strains derived from natural populations have, on average, between 550 and 670 new TE insertions present in euchromatic regions compared to the reference genome (Rahman et al. 2015). We also included in our screening a small subset of TEs that are not annotated in the reference genome, which are present at high frequencies in the North American population from North Carolina (Rahman et al. 2015, Mackay et al. 2010). Overall, we were able to identify a total of 121 candidate TEs: 109 annotated TEs and 12 non-annotated TEs (Annex Table S2.3). From the set of 109 TEs, a total of 61 TEs were absent or present at very low frequencies (<10% frequency) in the African population from the *D. melanogaster* ancestral range. These TEs have likely increased its frequency in natural populations during the recent out-of-Africa expansion, suggesting a possible role in adaptation to out-of-Africa environments. Thus, we identified 3.5 times more candidate TEs for out-of-Africa adaptation compared to the previous work (González et al. 2010). Moreover, we have also identified 48 TEs that are high frequent in Africa as well as out-of-Africa populations, suggesting that they could be playing a role in global adaptation.

Two other genome-wide screenings previously performed in *D. melanogaster* populations revealed a total of 22 TEs associated with signatures of positive selection (Kofler et al. 2012; Blumenstiel et al. 2014). Kofler and collaborators considered as candidates TEs fixed in one European population, some of them not annotated in the reference genome. Thus, considering the three genome-wide screenings available before this work, a total of 38 TEs were identified as candidates for adaptation. With our screening, we detected all the 38 TEs but four: two TEs are not annotated in the reference genome, one TE is not present among the 1,630 TEs analyzed, and for another TE we could not obtain frequency estimations. As our screening is focused in

the identification of polymorphic TEs, although we detected the rest of candidate insertions, they were not considered because they are fixed (9 TEs) or absent (4 TEs) in the four populations analyzed. Moreover, two other polymorphic insertions were not considered because they are located in low recombination regions (Fiston-Lavier et al. 2010).

Note that our 121 candidate TEs contain most of the candidate TEs for adaptation identified so far in *D. melanogaster*, including six TEs that have been connected to ecologically relevant phenotypes (Table 4).

TE	TE class	TE family	Phenotype	Mechanism	References
FBti0019430	non-LTR	*Doc*	Insecticide and virus resistance	Modifies transcript structure	Aminetzach et al. 2005; Magwire et al. 2011
NA	LTR	*Accord*	Insecticide resistance	Enhancer	Chung et al. 2007; Daborn et al. 2002; Schmidt et al. 2010
FBti0018880	DNA	*Bari1*	Oxidative stress resistance	Adds antioxidant response elements	González et al. 2009; Guio et al. 2014; Guio et al. 2015
FBti0019627	DNA	*pogo*	Xenobiotic stress resistance	Modifies transcript structure	González et al. 2008; Mateo et al. 2014
FBti0020155	DNA	*1360*	Lifespan and fecundity	Up-regulates gene expression	Zhu et al. 2014
FBti0019386	LTR	*invader4*	Shorter developmental time	Probably adds regulatory regions	González et al. 2008; Ullastres et al. 2015
FBti0019985	LTR	*roo*	Cold stress resistance	Adds TSS	Merenciano et al. 2016
FBti0019170	non-LTR	*F*	Heavy metal stress response	Probably adds regulatory regions	Mahn Le et al. 2017
FBti0020123	non-LTR	*Doc*	Cold stress resistance	Probably adds regulatory regions	Falqués et al. (personal communication)

Table 4. TEs linked to fitness advantageous phenotypes. All of them provide evidences of selective sweeps associated with the TE and phenotypic assays. With our screening we were not able to detect *FBti0019430* (because it is fixed in the four populations analyzed), *Accord* (because it is not annotated in the reference genome), and *FBti0019170* (because could not estimate the frequencies).

Overall, by increasing the number of populations and annotated TEs analyzed, as well as considering a small subset of TEs that are not annotated in the reference genome, we were able to identify a bigger dataset of TEs likely involved in adaptation. Thus, we confirmed the predictions of González et al. (2008) that more TEs could be identified if more population were analyzed. Moreover, it is probable that besides the 121 candidate TEs identified in this thesis, there are still more TEs to be identified that would be found if more natural populations were analyzed. Furthermore, genome-wide studies focusing only in TEs annotated in the reference genome are underestimating the number of TE insertions that might be playing a role in

adaptation. In fact, one of the eight TEs with demonstrated advantageous fitness effects, *Accord*, is not annotated in the reference genome (Chung et al. 2007; Daborn et al. 2002; Schmidt et al. 2010), and two non-annotated TEs characterized in this thesis are likely playing a role in immune response (see below). Thus, new sequencing techniques that enable to obtain longer reads, such as PacBio, and thus to *de novo* annotate TEs, will help to uncover more candidate TEs present in the natural populations (Barrón et al. 2014; Disdero and Fileé 2017; Villanueva-Cañas et al. 2017).

4.2.1 TEs Are Likely Playing a Role in Stress Response

TEs have often been related to stress response in different organisms (Hua Van et al. 2011; Casacuberta and González 2013; Chuong et al. 2016; Ullastres et al. 2016). Previous genome-wide screenings looking for adaptive TEs were limited by the small number of candidates identified to determine what type of biological processes are being selected during adaptation (González et al. 2008; Kofler et al. 2012; Blumenstiel et al. 2014). From the total of 121 TEs, we found functional information for the genes nearby 81 TEs based on literature search. Our results show that 58% of these 81 TEs are associated with genes involved in stress response, including xenobiotic stress, oxidative stress, and immune response (Table 3.2.3 and Annex Table S2.3). We also found that TEs were associated with genes involved in behavior, metabolism, or circadian rhythm (Annex Table S2.3). All these traits have been previously associated with adaptation in studies looking for candidate SNP variants. For example, the comparison of African populations with North American and Caribbean populations detected candidate SNPs associated with genes involved in immune response, behavior, metabolism, circadian rhythm, stress response, development and morphogenesis (Yukilevich et al. 2010). Independent studies on clinal adaptation in the east coast of North America and Australia detected genes involved in olfaction and metabolism (Mackay et al. 2012; Machado et al. 2016; Kolaczkowski et al. 2011; Levine et al. 2011). Other studies of candidate SNPs along both North America and Australia both east and west coasts, showed that, besides olfaction and metabolism, immune response is also a significant trait (Fabian et al. 2012; Turner et al. 2008). Among the genes associated with the candidate SNPs for these traits in other publications, there are 135 genes that are also associated with our candidate TE dataset (Merenciano et al. personal communication). None of the identified SNP has been validated, therefore, it should be further studied in order to check whether they are involved in adaptation. Considering the nature of TE mutations, the candidate TEs identified in this work are the genetic variants more likely to affect gene regulation. Moreover, it is also probable that the SNPs found to be associated with these genes are in linkage disequilibrium with the candidate TEs. However, some of these SNPs might also be playing a role in adaptation, thus, different types of mutations in the same gene would be contributing to adaptation of a specific trait. Therefore, when screening for candidate mutations

for adaptation, integrating different types of mutation increases our power to identify what traits are more likely to be involved in adaptation. For example, we identified 21 genes associated only with TE insertions.

Further investigation analyzing more populations, as well as considering all the TE insertions present in the populations, will increase the number of candidate TEs for both local and global adaptations. A larger number of candidate adaptive TEs would also give a better picture of the traits and biological processes that are behind adaptation. This type of reverse genetics approaches will ultimately allow us to better understand what are the genetic basis underlying adaptation processes.

4.3 The Role of TEs in Immune Response

We found that our candidate TE dataset is enriched for genes involved in immune response (Table 3.2.3, Annex Table S2.3). As mentioned above, immune response is one of the traits that often arise when comparing different populations looking for signals of selection (Fabian et al. 2012; Kolaczkowski et al. 2011; Levine et al. 2011; Fumagalli et al. 2011, Tinsley et al. 2006; Lazzaro et al. 2008; Juneja et al. 2016). Recently, Juneja and collaborators found that *cis*-regulatory variation contribute to latitudinal differences of immune-related genes expression both in North America and Australia *D. melanogaster* natural populations. Although the authors detected several genes with allele expression changes, they did not identify the causal genetic variants for the gene expression variation observed (Juneja et al. 2016).

In this thesis, we have explored for the first time the possible genome-wide role of TEs in regulating the oral immune response to the gram-negative bacteria *Pseudomonas entomophila*, a natural *D. melanogaster* pathogen (Vodovar et al. 2005). It is known that local, but not systemic immunity, contributes to resistance against oral infection with *P. entomophila* (Liehl et al. 2006). In nature, bacteria are found at high concentrations in decaying fruits. The gut is the first barrier that pathogens encounter during infection, thus, adaptations improving the gut immune response are advantageous for the organism fitness (Buchon et al. 2014; Bonfini et al. 2016; Capo et al. 2016). The *D. melanogaster* gut is a compartmentalized tissue with rich gene expression diversity (Chintapalli et al. 2007; Buchon et al. 2013): more than half of the *D. melanogaster* annotated genes (62%) are expressed in the gut (Buchon et al. 2013). Moreover, a total of 460 transcription factors are expressed along the gut, 52 of them are expressed in a patterned manner, suggesting a high complex gene expression regulatory network (Buchon et al. 2013).

4.3.1 TEs Are Associated with Immune-Related Gene Expression Changes

We found that most of the TEs were associated with expression changes, both up-regulation and down-regulation, of the nearby genes in non-infected and/or infected conditions, and in at least one of the two genetic backgrounds analyzed (Figure 3.2.1). For that, we performed ASE analysis on heterozygous flies carrying one allele with the TE and the other allele without the TE. With this technique, we are able to detect the effect of *cis*-changes in the same genetic environment (Wittkopp et al. 2004). We found that four out of the 16 genes analyzed in our study showed TE-allele expression changes both in non-infected and infected conditions, while six genes showed TE-allele expression differences only in non-infected conditions and three only in infected conditions (Figure 3.2.1). Both Imd pathway and gut epithelium renewal are stimulated at a basal level by the gut microbiota (Buchon et al. 2009). Thus, gene expression regulation by the TEs not only can be regulating the response to infection in the gut but also might be playing a role in the gut-microbiota interactions. A transcriptome analysis comparing resistant and susceptible natural strains revealed that very few genes were expressed differently after *P. entomophila* oral infection (Bou Sleiman et al. 2015). Actually, resistant and susceptible strains differed in the basal intestinal transcriptome profile, *i.e.* in non-infected conditions. This suggests that gene expression variability in non-infected conditions would pre-dispose to enteric infection susceptibility (Bou Sleiman et al. 2015).

In the same study, Bou Sleiman and colleagues identified a total of 1,287 genes with expression differences comparing non-infected and infected flies, 4 hours after *P. entomophila* exposure (Bou Sleiman et al. 2015). This study included 14 out of the 16 genes analyzed in our study, nine of them showed expression differences after infection: six were up-regulated and three were down-regulated (Bou Sleiman et al. 2015). Note that with our ASE analysis we can only detect expression differences between the allele with the TE and the allele without the TE. Therefore, we cannot compare gene expression between non-infected and infected conditions. However, we found that five of our genes were associated with TE-allele expression changes in non-infected conditions in the same direction as Bou Sleiman and collaborators found they changed after infection. Thus, flies with the TE would have higher (or lower) gene expression levels before the infection happens, suggesting that these flies might be predisposed to a better response to infection.

While most of the TEs showed association with expression changes at basal levels, only three TEs were associated with expression changes only after infection. This suggests that these three TEs might be regulating gene expression specifically during infection. One of the TEs is *FBti0019386*, associated with *Bin1* allele up-regulation (Figure 3.2.1). This is likely due to the TSS signal present in *FBti0019386* (further discussed below). The other two TEs, *FBti0018868* and *FBti0020137*, associated with *TM4SF* and *NUCB1* respectively, showed gene expression

changes associated with the presence of the TE after infection: up-regulation in one background and down-regulation in the other background, however this was only significant in one of the backgrounds.

4.3.2 Background-Dependence in the Allele Specific Expression Changes

We found the same results in the two backgrounds in 17 out of the 32 expression analysis: *CG10943* showed up-regulation in non-infected conditions, *CG8008* showed down-regulation in infected conditions, and the other 15 genes did not show significant differences in any of the two backgrounds (Figure 3.2.1). However, for eight analyses, both backgrounds showed changes in expression in the same direction although results were only significant in one of the two backgrounds: five genes in non-infected conditions (*CG2233, AGO2, CG15829, CG8008,* and *CG15096*), and three genes in infected conditions (*CG10943, CG8628,* and *Bin1*) (Figure 3.2.1). Finally, only seven genes differed in the direction of the change of expression in the two backgrounds, although results were only statistically significant in only one of the genetic backgrounds analyzed: four genes in non-infected conditions (*Dif, CG8628, Mef2,* and *cbx*), and three genes in infected conditions (*cbx, TM4SF,* and *NUCB1*) (Figure 3.2.1).

The fact that we cannot detect statistically significant differences in the ASE analysis could be because we only used three biological replicates for each background and condition (Figure 3.2.1, Annex Table S2.4). Thus, by increasing the number of biological replicates we might gain statistical power to detect such allele expression differences in both backgrounds (Blainey et al. 2014). It might also be possible that other variants besides the TEs interfere in the gene expression regulation. Thus, epistatic interactions between the TE and other variants present in a genetic background may hinder the regulatory effect of the TE in that background (Chandler et al. 2013; Gasch et al. 2016). A detailed study in yeast on the quantitative trait nucleotides (QTN) interaction with both environment and genetic background found that, although the QTN effects were consistent in direction across backgrounds, the magnitude of their effect varied (Gerke et al. 2010). The analysis of the TEs 1kb flanking regions in the DGRP strains did not reveal the presence of other possible mutations in the regions conserved, considered as possible regulatory regions (Annex Table S2.5). Thus, we expect that the TE is the strongest candidate *cis* genetic variant explaining the gene expression changes found in the ASE. However, we cannot discard the presence of other variants in regions that are further from the 1kb flanking regions analyzed, which could be affecting the allele expression differences. Finally, although the expression changes were only significant in one of the genetic backgrounds analyzed, we found seven TEs to be associated with up-regulation in one background and down-regulation in the other background (Figure 3.2.1). Thus, it could be that a single TE insertion causes different effects depending on the genetic background (see below).

4.3.3 TEs Regulate Nearby Gene Expression by Adding Promoter and Enhancer Sequences

We explored the molecular mechanisms behind the expression changes detected in the ASE for five TEs: *FBti0019386* and *FBti0018868*, both up-regulated after infection, *FBti0061506*, up-regulated in non-infected conditions, *FBti0019985,* down-regulated both in non-infected and infected conditions, and *tdn8*, up-regulated in both conditions (Figure 3.2.1).

***FBti0019386* and *FBti0018868* Add a TSS to Their Nearby Genes.** In *D. melanogaster*, a genome-wide *in silico* study showed that TEs provide promoters that drive the expression of hundreds of annotated genes in different developmental stages (Batut et al. 2013). Two of these reported TEs are *FBti0019386* and *FBti0018868*, for which we confirmed the transcription initiation of their nearby genes, *Bin1* and *TM4SF* respectively, in the gut. Interestingly, *FBti0019386* only initiated transcription of *Bin1* in infected guts, which is consistent with the up-regulation of *Bin1* TE-allele found in the ASE analysis of infected guts. As seen in the first chapter, *FBti0019386* has signals of positive selection, and it is associated with a shorter developmental time and increased sensitivity to cold stress. Cold stress response has often been linked to immune response, as immune-related genes have been found up-regulated after cold exposure (Vermeulen et al. 2013; Zhang et al. 2011; MacMillan et al. 2016). Moreover, flies exposed to cold stress survive better to fungal infections (Marshall and Sinclair 2011; Le Bourg et al. 2009). Altogether these results suggest that *FBti0019386* might also be associated with an immune response phenotype. Previous evidences showed that *Bin1* mutant larvae are more sensitive to fungal infection (Costa et al. 2011). In this second chapter, we showed that adult mutant flies up-regulating *Bin1* gene, had a higher survival compared to wild-type flies to *P. entomophila* oral infection, although it was marginally statistically significant (Table 3.2.5, Annex Figures S4 and S5). Further phenotypic experiments using natural strains with and without *FBti0019386* should be performed in order to associate the TE with an increased survival after gram-negative infection. Moreover, it would be interesting to study the molecular and survival phenotypes of flies with and without *FBti0019386* after fungal infection.

We also detected that *FBti0018868* is modifying the transcript structure of the nearby gene *TM4SF* by adding a new TSS (Figure 3.2.5). However, we were not able to associate this change with the up-regulation of the allele in infected conditions. We also discarded that *FBti0018868* is playing a role as an enhancer in the gut (Figure 3.2.5). *TM4SF* RNAi mutants were more resistant to *P. entomophila* oral infection compared to the wild-type strain during the firsts 30 hours and, after that, they become more sensitive (Table 3.2.5, Annex Figures S4 and S5). Although *TM4SF* function is not known, this evidence showing susceptibility of the *TM4SF* mutant flies suggests that this gene is required for immune response to gram-negative bacteria.

In *Drosophila,* there are evidences for another TE involved in immune response by modifying the gene transcript structure (Magwire et al. 2011). In this case, the TE *FBti0019430* truncates the structure of *CHKov1* gene and generates two new transcripts, thus producing a shorter protein product (Aminetzach et al. 2005). Flies with *FBti0019430* insertion are more resistant to sigma virus infection compared to flies without the TE (Magwire et al. 2011).

***FBti0019985* and *tdn8* Act as Enhancer Elements.** Besides triggering structural changes, TEs can also add new regulatory regions able to enhance the expression of their nearby genes (Van't Hof et al. 2016; Chuong et al. 2016). We showed that at least two out of the four analyzed TEs that are associated with gene expression up-regulation, *FBti0019985* and *tdn8*, add enhancer regulatory sequences able to increase the expression of the nearby gene.

We found that *tdn8* is associated with *CG10943* up-regulation likely because it is adding enhancer regulatory sequences upstream the gene (Figure 3.2.3). This gene was found to be up-regulated in immune challenged flies in several studies (Table 3.2.4) (Broderick et al. 2014; Roxstrom-Lindquist et al. 2004; Bou Sleiman et al. 2015).

We also found that *FBti0019985* act as an enhancer sequence in the anterior part of the gut both in non-infected and infected conditions (Figure 3.2.6). These results would explain the marginally significant up-regulation of *cbx* associated with *FBti0019985* in the first genetic background analyzed in the ASE. Moreover, the regionalization of *FBti0019985* enhancer ability would explain why we did not detect statistically significance in the expression results using the whole gut (Figure 3.2.1, Annex Table S2.4).

Interestingly, *FBti0019985* is associated with *cbx* down-regulation in the second genetic background analyzed. Thus, it is probable that the same TE is causing gene expression changes in a background-dependent manner, and this could also be caused by different molecular mechanisms. More experiments should be performed with different genetic backgrounds to try to better understand the molecular mechanisms and molecular effects of *FBti0019985.*

cbx mutant flies are more sensitive to infection by injection with gram-positive bacteria but not gram-negative (Ayres et al. 2008). We also observed that *cbx* mutants do not show differences after oral infection with the gram-negative bacteria *P. entomophila* (Table 3.2.5, Annex Figure S4). *cbx* is an ubiquitin-conjugating enzyme that might have a role in crystal cell development (Milchanowski et al. 2004). Crystal cells compose 5% of *Drosophila* hemocyte and participate in immune response and wound healing through melanization, however, enzymes associated with melanization process and crystal cells are not expressed in the *Drosophila* gut (information from Flybase). Thus, more experiments should be performed in order to assess a possible role of *FBti0019985* in gram-positive bacteria response.

Besides locating in the first intron of *cbx* gene, *FBti0019985* overlaps with the 5'UTR of another gene, *CG18446* (Gramates et al. 2017). This TE adds a TSS to *CG18446* that is associated with embryo cold stress survival (Merenciano et al. 2016). Interestingly, this genomic region has been repeatedly reused as an insertion site for other TEs from the same family. The sequences of the different TEs barely vary (Merenciano et al. 2016), suggesting that probably all of them might be playing a role as an enhancer for *cbx* gene.

There are other examples in the literature showing how TEs regulate immune response by acting as enhancer regulatory sequences. For example, a study performed in human populations has linked the presence polymorphic TEs with immune gene expression differences (Wang et al. 2016). One of these insertions is associated with increased expression of *PAX5,* an important transcription factor for B cell differentiation. This suggests a possible impact of TEs on the whole regulatory network and, therefore, an important impact on the phenotype. A different study in mammalian cells, revealed that LTR retrotransposons nest interferon-γ-inducible enhancers in their sequences, which induced the expression of immune responsive genes to *Vaccinia virus* in cells (Chuong et al. 2016).

Finally, we were not able to link *FBti0061506*, associated with *Dif* up-regulation, with an enhancer role (Figure 3.2.4). With our experiments we cannot discard that *FBti0061506* can drive nearby gene expression, as it is known that some enhancer sequences might play their function only when placed into their genetic context (Spitz and Furlong 2012).

4.3.4 Other Molecular Mechanisms Could Underlie the Expression Changes

In our ASE analysis, there are eight other immune-related genes showing expression differences associated with the presence of TEs, most of them are associated with gene down-regulation (Figure 3.2.1). Thus, other mechanisms such as the addition of heterochromatin marks or the addition of downstream regulatory sequences should be explored. TEs are able to recruit heterochromatin proteins that participate in silencing the nearby genes (Sentmanat et al. 2012; Huisinga et al. 2016; Guio et al. personal communication). For example, in *Arabidopsis thaliana*, the TE COPIA-R7 mediates the regulation of *RPP7* gene, an immune regulatory protein that gives fungal resistance, by recruiting the histone mark H3K9me2 (Tsuchiya et al. 2013).

Another possible mechanism that could be explored is the modification of downstream regulatory sequences. For example, *FBti0019602* is located 12 bp downstream *CG2233*, likely modifying the gene polyadenilation signal (PAS). Downstream structural changes can modify the transcript expression levels, as seen for the TE *FBti0019627* associated with *CG11699* up-regulation (Mateo et al. 2014). This TE truncates *CG11699* PAS and generates a shorter

transcript that is associated with gene overexpression in flies with *FBti0019627* (Mateo et al. 2014).

Finally, it is also possible that the TEs are associated with other transcript changes that are not detectable with our ASE analysis. For example, differential expression of gene alternative transcripts associated with the presence of the TE. Moreover, it is also possible that some of the analyzed TEs show expression changes when infected with a different pathogen such as gram-positive bacteria or virus.

Overall, the examples presented in this thesis highlight the variety of mechanisms underlying adaptive mutations and point toward a significant role of TEs in response to stress (Casacuberta and González 2013). Although we found evidences supporting a role of TEs in immune-related gene expression regulation, more evidences are needed to conclude that these changes trigger phenotypic adaptation. Moreover, some TEs might exhibit their adaptive effect in a different timing, a different developmental stage, after infecting with a different pathogen, or even using a different infection route. Indeed, *FBti0019386* has been associated with expression changes in different developmental stages: females and gut, and under different conditions: non-stress, cold stress and infection. Moreover, we have also linked this insertion to several phenotypes: shorter developmental time, cold stress sensitivity, and probably immune response. Although we found evidences pointing to a possible role of TEs in immune response regulation, more experiments should be performed in order to link the identified TEs with a fitness effect in this trait.

05

CONCLUSIONS

5. CONCLUSIONS

From the results obtained in this thesis, we can conclude:

1. *FBti0019386* has genomic signatures of positive selection, thus reinforcing the previous identification of this TE as a candidate for adaptation.

2. Flies with *FBti0019386* have a shorter developmental time and are more sensitive to stress, which are likely to be the adaptive effect and the cost of selection of this mutation, respectively.

3. The observed phenotypic effects of *FBti0019386* are not consistent with a role of this TE in temperate adaptation as has been previously suggested. Indeed, our global analysis of the population frequency of *FBti0019386* show that climatic variables explain well the TE frequency patterns only in Australia. Thus, further functional validation should be gathered before concluding that a candidate *loci* is under spatially varying selection.

4. *FBti0019386* is associated with up-regulation of *sra* in adult females. There are no direct associations between *sra* and developmental time or stress. However, the role of *sra* as a calcium pathway regulator could be indirectly associated with both phenotypes.

5. Genome-wide screening of TE insertions, including natural populations from three different continents, as well as including both annotated and a subset of non-annotated TEs in the reference genome, allowed detecting more candidate insertions for adaptation compared to previous works.

6. Most of the candidate adaptive TEs are nearby genes associated with stress response. This suggests that TEs could be playing an important role in regulating genome response to environmental stressors in *Drosophila*.

7. An important part of the identified candidate TEs, 23%, are nearby immune-related genes. This suggests that TEs might be generating adaptive regulatory variation in immune response in *Drosophila*.

8. The immune-related genes nearby the candidate TEs show allele-specific expression changes in the gut associated with the TEs both in non-infected and infected conditions.

Thus, besides regulating the immune response, TEs might be also predisposing to a better response to oral infection.

9. Although we found background-dependent allele-specific expression changes, the analysis of the flanking regions show that the TEs identified are the strongest candidate *cis* genetic variants explaining the observed expression changes, as no other variants were detected.

10. We found that the candidate TEs were associated with both up-regulation and down-regulation of the nearby genes. This suggests that they might be modifying gene regulation by using different molecular mechanisms.

11. We found the mechanism underlying the expression changes for three out of the five TEs analyzed: *FBti0019386* adds a TSS to its nearby gene *Bin1*, and *FBti0019985* and *tdn8* act as enhancer elements. Therefore, most of the genes showing up-regulation associated with the candidate TEs add promoter or enhancer regulatory sequences.

06

MATERIAL AND METHODS

6. MATERIAL AND METHODS

6.1 CHAPTER 1

6.1.1 Sequence Analysis of the *FBti0019386* Flanking Regions

Single nucleotide polymorphism (SNP) data were downloaded from the DGRP2 webpage (https://www.hgsc.bcm.edu/arthropods/drosophila-genetic-reference-panel) in vcf format. Strains with (N = 65) and without (N = 38) *FBti0019386* insertion were filtered using vcftools v_0.1.10 (http://vcftools.sourceforge.net/).

We used three different statistics to detect positive selection: Nucleotide diversity (π), Tajima's D, and the CL of SNPs. Positive selection results in the elimination of standing genetic variation that is linked to the adaptive mutation. Thus, if *FBti0019386* has increased in frequency due to positive selection, we expect a decrease in π in flies with the insertion compared with flies without the insertion. π is calculated as the mean number of pairwise differences between two given sequences (Hudson et al. 1992). Tajima's D statistic is calculated as the ratio between the mean number of pairwise differences and the number of segregating sites (Tajima 1989). This ratio is expected to be 0 in a neutrally evolving population whereas negative values of Tajima's D can be taken as evidence of positive selection (Tajima 1989). Finally, CL test is calculated by multiplying the marginal likelihoods for each site along the studied sequences (Nielsen et al. 2005).

π, Tajima's D, and CL were calculated for the two sets of sequences, with and without the insertion, using the PopGenome package in R (Pfeifer et al. 2014). Sliding windows analyses were performed for 200-bp-size windows spanning 1 and 2-kb regions flanking the insertion. Differences between strains with and without the insertion were more drastic for the 1-kb region flanking the insertion; therefore, we focused our analysis in this region.

Simulations were performed using the MS program (Hudson 2002). Theta values were estimated using the 205 DGRP2 strains for the 2-kb region around *FBti0019386* (theta = 4.77/kb) and for the 3R chromosomal arm (theta = 4.5/kb). Thus, simulations were performed for theta values of 4/kb and 5/kb, which are frequently used as neutral values in *D. melanogaster*.

Ad hoc perl scripts were used for the resampling analyses. In total, 1,000 random samples of the 103 DGRP strains analyzed were obtained keeping the same proportion as in the original present and absent data sets (60%/40%, respectively) and a sample size of nearly 50% of the total data set.

We also computed CLR as 2*(log CL (present) _ log CL (absent)), for a 1-kb region around the TE insertion. Because demography could produce similar patterns as positive selection, we performed a random sampling of 1,000 1-kb-long regions from the 3R chromosome for the absent and present data sets and calculated π, Tajima's D, CL, and CLR tests in each one of them.

6.1.2 Fly Strains

Outbred Strains

We selected six inbred strains from the Drosophila Genetic Reference Panel (Mackay et al. 2012; Huang et al. 2014) homozygous for the presence of *FBti0019386* insertion (RAL-21, RAL-40, RAL-177, RAL-402, RAL-405, and RAL-857). We placed ten virgin females and ten males of each strain in a fly chamber to create an outbred population sharing the TE insertion. We also selected six inbred strains without the insertion (RAL-75, RAL-138, RAL-383, RAL-461, RAL-822, and RAL-908) and created an outbred strain following the same procedure explained above. Each outbred population was maintained by random mating (N=800 flies per generation) for at least ten generations before starting the experiments.

Introgressed Strains

We selected two DGRP strains: One homozygous for the presence of *FBti0019386* insertion (RAL-177) and one homozygous for the absence (RAL-802). We crossed RAL-177 virgin females with RAL-802 males and backcrossed the virgin females that carry *FBti0019386* insertion from the following generations with RAL-802 males for 12 generations. After that, we did brother–sister crosses until we obtained homozygous strains for the absence and homozygous strains for the presence of *FBti0019386*.

Individual DGRP Strains

We used a couple of individual DGRP strains differing by the presence/absence of *FBti0019386* insertion to perform our phenotypic assays. We used RAL-857 (homozygous for the presence of *FBti0019386* insertion) and RAL-802 (homozygous for the absence).

Presence/Absence of In(3R)Payne in the Analyzed Strains

To discard the effect of In(3R)Payne inversion on *FBti0019386* phenotypic effects, we genotyped the strains analyzed to detect the presence/absence of this inversion: The two outbred, the two introgressed, and the two individual DGRP strains. We used the primer sequences described in Matzkin et al. (2005). As a positive control, we used a strain that was previously genotyped in our laboratory and that carries the In(3R)Payne inversion.

6.1.3 Phenotypic Assays

All experiments were performed using outbred populations. Additionally, we used introgressed and individual DGRP strains to perform CCRT assay, survival after chill-coma, and DT assays.

Fecundity

In total, 40 virgin females from each strain were placed individually in vials with one male from the same strain. During 17 days flies were moved to new vials every 2 days and the number of eggs laid per female during that period was counted. Total fecundity, that is, average of the total number of eggs laid per female during the 17 days, and early fecundity, that is, average of the total number of eggs laid per female during the first 48 h of egg laying, was compared between flies with and without *FBti0019386*.

Egg Hatchability and Hatching Time

In total, 800 4- to 8-day-old flies were allowed to lay eggs for 3 h on apple juice-agar medium with fresh yeast. Embryos were separated in groups of 20 or 50 and placed into food vials. Vials were kept at room temperature (19–22 ºC) and checked during the following hours for hatched eggs (2–5 times per day). We analyzed the average time over the midpoint of each successive interval in order to estimate the hatching time. Two experiments were performed following this protocol: A first pilot experiment with 150 embryos per strain, and one replica with 500 embryos per strain.

Egg hatchability and egg hatching time were also analyzed under cold stress conditions. Embryos were placed at 1ºC overnight for 14 h and at 18ºC during the day, and this cycle was maintained until all the eggs had hatched. We performed a pilot experiment with 100 embryos per strain and additional experiments with 240 and 160 embryos per strain, respectively.

Cold Stress in Embryos

In total, 800 7- to 10-day-old flies were allowed to lay eggs for 3 h on apple juice-agar medium with fresh yeast. Embryos were collected following the methodology described in Schou (2013), and placed into food vials in groups of 50. When embryos were 3–6 h old, vials were placed at 1ºC for 14 h, and maintained at 18ºC until adult emergence. Simultaneously, control vials were always maintained at 18ºC and not cold-exposed to control for other variables affecting egg to adult survival. We performed a first pilot experiment using 280 embryos per strain and three biological replicas using 350 embryos per strain (replica 1) and 750 embryos per strain (replica 2 and replica 3, respectively). In all cases, we analyzed egg to adult survival after all the adults had emerged.

Chill-Coma Recovery Time

In total, 500 3- to 5-day-old flies were separated by sex and by strain and placed into five empty vials in groups of 50. We allowed flies to recover from CO_2 anesthesia for 1 h and then vials were put in ice and kept in a 4ºC chamber for 16 h as described in David et al. (1998). After the cold shock, adults were transferred to Petri dishes at room temperature (22–24ºC), and recovery time was monitored for successive intervals of 30 s during 2 h. We considered as recovered flies

those that were able to stand on their legs. As a control, we monitored survival of flies that were kept at room temperature: Three vials of 20 flies each, by sex and strain.

Survival after Chill-Coma

In total, 400 5- to 8-day-old flies were separated by sex and strain and placed into six food vials in groups of 20. We allowed flies to recover from CO_2 anesthesia for at least 2 days. After that, flies were changed to empty food vials and were put in ice, and kept in a 4ºC chamber for 16 h. When adults were recovered from chill-coma, we transferred them to food vials and we monitored mortality during the next 5 days. As a control, we monitored survival of flies that were kept at room temperature: Three vials of 20 flies each, by sex and strain.

Osmotic Stress

In total, 2,000 4- to 7-day-old flies were separated by sex and strain and placed in groups of 20 into 20 food vials containing 3% of NaCl, and into five vials with normal food as a control. Flies were maintained at room temperature (22–24 ºC) and dead flies were counted every 12–24 h until all the treated flies were dead.

Starvation Stress

In total, 2,000 3- to 4-day-old flies per strain were separated by sex and strain and placed in groups of 20 into 20 food vials containing only 1.5% agar, and into five vials with normal food as a control. Flies were maintained at room temperature (22–24 ºC) and dead flies were counted three times a day until all the treated flies were dead.

Developmental Time

In total, 800 7- to 10-day-old flies were allowed to lay eggs for 3 h. A total of 500 embryos per strain were collected and distributed in groups of 50 per food vial and were maintained at 18ºC. Vials were checked every 6–8 h for emerging adults until all flies had emerged. We estimated the average DT over the midpoint of each successive interval.

Statistical Analyses of the Phenotypic Assays

Analyses were performed with SPSS v21. We first tested whether data followed a normal distribution by performing Kolmogorov–Smirnov test. T-test was performed for normal data and Mann–Whitney test for non-normal data. Survival curves were compared with log-rank test. When the statistical test was significant, we estimated the size effect of the mutation by calculating the odds-ratio and its confidence interval.

6.1.4 *FBti0019386* Frequency Estimation for Natural Populations

To obtain *FBti0019386* frequency, we run T-lex2 (Fiston-Lavier et al. 2015) using *Drosophila* whole-genome sequences available from a total of 23 populations from North America, Australia, Europe, and Africa (Annex Table S1.5).

The accuracy of TE frequency estimates using *T-lex2* is affected by coverage. However, coverage for all samples was higher than 20x except for Lyon (France) and California (USA), which had 8x and 4.7x coverage respectively, suggesting that overall frequency estimates are accurate.

6.1.5 Correlation Analysis of *FBti0019386* Frequency with Geographic and Climate Variables

We analyzed whether the frequency of *FBti0019386* insertion correlated with different geographical and climate variables in North America, Australia, and Europe using Pearson product– moment correlations. We also performed a PCA to disentangle the relationships between the climatic variables using Statistica (v8.0, StatSoft, Inc. 2007). Climatic data were obtained from the weather stations adjacent to collection sites of each population, available in Peel et al. (2007). When necessary, data were transformed as described in Sokal and Rohlf (2012) (see pages 411–422).

6.1.6 mRNA Transcript Levels Analysis (quantitative reverse transcription polymerase chain reaction)

Total RNA was extracted from three biological samples of 40 adult females (5-day old) from outbred populations differing by the presence/absence of *FBti0019386* insertion using Trizol reagent and PureLink RNAMini kit (Ambion). RNA was treated on-column with DNase I (Trizol) and after RNA purification. Reverse transcription was carried out using 1 mg of total RNA, Anchored-oligo(dT) primer, and Transcription First Strand cDNA Synthesis Kit (Roche). The resulting cDNA was used for quantitative reverse transcription polymerase chain reaction with SYBR Green (BioRad) on an iQ5 Thermal cycler. sra total expression was measured using a pair of primers specific to a 124-bp cDNA amplicon spanning the 50- UTR/exon junction of the gene (5'-ACAACAACGGTGGAGAAGAGCCGT-3' and 5'-GGTGCATCGGCGGACGCA TTG-3'). For *Bin1*, we measured the 66-bp cDNA amplicon spanning the 50-UTR/exon junction using specific primers (5'-TGTCGTCCCGTAGAGCAGAA-3' and 5'-CA AGCAGATTGACCGCGAGA-3'). In both cases, we normalized the expression with Act5C (5'-GCGCCCTTA CTCTTTCACCA-3' and 5'-ATGTCACGGACGATTTCA CG-3'). Expression was measured in nonstress conditions and in cold-stress conditions: 16 h at 4ºC and 2h at room temperature to allow flies to recover.

We also analyzed the expression of both genes in 0–2 h embryos using the same procedure. We collected the embryos from population cages containing approximately 800 flies from outbred

populations differing by the presence/ absence of *FBti0019386* insertion. Briefly, 4 - to 8-day-old flies were allowed to lay eggs for 2 h on apple juice-agar medium with fresh yeast. Then, embryos were collected using a small brush and cleaned with water. Embryos were dechorionized by submerging them for 5 min in 50% bleach. After that, embryos were placed in a microcentrifuge tube, the excess of water was eliminated, and the samples were frozen at -80ºC until RNA extraction.

6.1.7 Detection of piRNA reads binding to *FBti0019386* sequence

We used small RNA sequencing data to check whether piRNAs reads mapped to *FBti0019386* sequence, following a methodology similar to that described in Sentmanat and Elgin (2012). Briefly, we obtained the small RNA reads from Oregon R ovaries (accession number SRP000458) (Li et al. 2009), and from wild type ovaries (accession number: SRX470700) (Satyaki et al. 2014). We aligned the reads by using BWA-MEM package version 0.7.5 a-r405 (Li 2013) to the 14.6-kb sequence obtained from Drosophila reference genome, containing *Bin1* and *sra* genes, and *FBti0019386* (release five chromosomal coordinates 3R: 12,010,721–12,025,306). Then, we used samtools and bamtools (Barnett et al. 2011) to index and filter by sense/antisense reads. Finally, we obtained the total read density using R (Rstudio v0.98.507).

6.1.8 Detection of HP1a Protein Binding in *FBti0019386* Sequence

We downloaded all available raw data from modEncode HP1a protein ChIP-Seq experiments: Embryos (ID 3391 and 3392), third instar larvae (ID 4936), and adult heads (ID 5592) (http://data.modencode.org). Then, we mapped the reads against the 14.6-kb region described above. We performed the alignments following the same methodology as for the piRNA reads analysis.

6. MATERIAL AND METHODS

6.1 CHAPTER 2

6.2.1 Fly Strains

DGRP strains. Raw reads from 141 DGRP strains were used to estimate the frequencies of TEs annotated in the *D. melanogaster* reference genome, (M.G. Barrón, personal communication) (Mackay et al. 2012). A subset of 37 DGRP strains were also used to analyze by PCR a subset of TEs not annotated in the reference genome (Rahman et al. 2015). Finally, DGRP strains were also used to perform allele specific expression analyses (ASE), transcription start site identification (TSS), and enhancer assays. The identity of the strains used for the different experiments can be found in Annex Table S2.5A.

African strains. A subset of 66 African strains collected in Siavonga (Zambia) with no evidence of cosmopolitan admixtures were used to estimate the frequency of TEs annotated in the reference genome (Lack et al. 2015) (Annex Table S2.5B).

European strains. Raw reads from 73 European strains: 57 from Stockholm (Sweden) and 16 from Bari (Italy) were used to estimate the frequencies of TEs annotated in the *D. melanogaster* reference genome, (M.G. Barrón, personal communication) (Ullastres et al. 2015) (Annex Table S2.5C). Additionally, one strain from Bari (CAS-49) was used for ASE and TSS experiments and one strain from Munich (MUN-8) was used for ASE experiments (Annex Table S2.5C).

Mutant strains. We used three RNAi mutant stocks from the VDRC stock center (Annex Table S2.5D). To generate the mutants, we crossed the stocks carrying the RNAi controlled by an *UAS* promoter with flies carrying a GAL4 driver to silence genes ubiquitously. We performed the experiments with F_1 flies that were obtained from each cross. Based on the phenotypic markers, we separated the RNAi mutant flies from the rest of the F_1 that do not have the gene expression altered. The flies with normal expression levels were used as a baseline of the experiment. To overcome the lethality of silencing *CG15829* during development, we used an *Act5c*-GAL4 strain regulated by the temperature sensitive repressor GAL80 (Annex Table S2.5D). For this mutant, we transferred flies from 25ºC to 29ºC 24h before performing the experiment.

We also used four mutant stocks generated with different transposable element insertions. In this case, we used strains with similar genetic backgrounds as a baseline for the experiments (Annex Table S2.5D).

6.2.2 Transposable Element Datasets

TEs annotated in the reference genome. There are 5,416 TEs annotated in the v6 of the *D. melanogaster* reference genome (Gramates et al. 2017). We did not consider the 2,234 TEs that belong to the INE-1 family, because this family has not been active for more than 3 million years. Thus INE-1 TEs are not likely to be involved in recent adaptation (Kapitonov and Jurka, 2003; Sackton et al., 2009; Singh and Petrov, 2004). We neither considered TEs that are flanked by simple repeats, nested TEs, or TEs that are part of segmental duplications because frequencies cannot be accurately estimated for these TEs using T-lex2 (Fiston-Lavier et al. 2015). Finally we discarded TEs present in genomic regions with a recombination rate = 0 according to Fiston-Lavier et al. (2010) or Comeron et al. (2012). Thus, we ended up with a dataset of 815 annotated TEs for which we estimated frequencies using T-lex2 (Fiston-Lavier et al. 2015). We considered high frequent TEs those present at a population frequency ≥ 10%.

TEs non-annotated in the reference genome. We also analyzed a subset of 25 TEs identified by Rahman et al. (2015) in DGRP strains that are not annotated in the reference genome (Annex Table S2.2). These 25 TEs are present in regions with recombination rate > 0 (Fiston-Lavier et al. 2010 and Comeron et al. 2012) and were inferred to be present in at least 15 DGRP strains out of the 177 strains analyzed by Rahman et al. (2015). We first confirmed the presence of these 25 TEs by PCR on a total of 37 DGRP strains (see below). For each TE, we sequenced at least one of the PCR products to confirm the presence and the family identity of the TE. We estimated the frequency of each TE based on the PCR results for a minimum of seven strains and considered as high frequent those present at a population frequency ≥ 10%.

6.2.3 Presence/Absence of TEs in the Analyzed Strains

We performed PCRs to confirm the *in silico* results obtained with T-lex2 (Fiston-Lavier et al. 2015) and TIDAL (Rahman et al. 2015). We designed specific primers for each analyzed TE, in order to confirm the presence and/or absence of that TE, using the online software Primer-BLAST (Ye et al. 2012) (Annex Table S2.7). Briefly, we designed a primer pair flanking the TE (FL and R primers), which produces a PCR product with different band sizes when the TE is present and when the TE is absent. For those TEs that are present in the reference genome, we also designed a primer inside the TE sequence (L primer) that, combined with the R primer, only amplifies when the TE is present (González et al. 2008). To perform the PCRs, genomic DNA was extracted from 10 females from each analyzed strain.

6.2.4 Functional Annotation of Genes Nearby Candidate Adaptive TEs

In order to search for enriched biological functions of the genes associated with the candidate TEs, we analyzed the Biological Process gene ontology (GO) terms using the DAVID functional annotation tool (Huang et al. 2009). We considered all the genes that were located less than 1kb from the candidate TEs. If the candidate TEs did not have any gene located in the 1kb flanking regions, then we considered only the closest gene. We compared the genes associated with the candidate TEs with the genes associated with all the polymorphic TEs. We run DAVID with the default parameters and using the statistical threshold with high stringency (Huang et al. 2009). We considered as significant gene functional clusters those above 1.3 Enrichment Score (ES), as recommended by Huang et al. (2009).

Additionally, we looked for functional information of the genes associated with the candidate adaptive TEs using Flybase (Gramates et al. 2017). We considered GO annotations based on experimental evidence and we also obtained functional information based on the publications cited in Flybase. Several lines of evidence were considered: genome-wide association studies in which SNPs in the analyzed genes were linked to a phenotypic trait, differential expression analyses, and phenotypic evidences based on the analyses of mutant stocks. We then classified all the TEs based on the gene functions associated with their nearby genes.

6.2.5 Bacterial Infection

We infected 5- to 7- day-old female flies with the gram-negative bacteria *Pseudomonas entomophila* (Vodovar et al. 2005). Flies were separated into food vials under CO_2 anesthesia two days before the bacteria exposure and were kept at 25ºC. The experiments were performed as described in Neyen et al. (2014). Briefly, flies were starved for two hours and then they were flipped to a food vial containing a filter paper soaked with 1.25% of sucrose and bacterial pellet. The bacterial preparation was adjusted to a final OD_{600} = 100, corresponding to 6.5 $x10^{10}$ colony forming units per ml (Vallet-Gely et al. 2010). Flies were kept at 29ºC and 70% humidity, which are the optimal infection conditions for *P. entomophila*. In parallel, we exposed non-infected flies to sterile LB with 1.25% sucrose.

6.2.6 Survival Experiments

We performed infection survival experiments with mutant flies, and we compared the mortality of the mutant flies to the mortality of flies with similar genetic backgrounds (Annex File S2.5D). Female flies were placed in groups of 10 per vial, and we performed the experiments with 5-12 vials (see Annex Figure S4), except for cn^1 considered as a wild-type background for which we

used 3 vials. As a control for each experiment, we exposed 3-4 vials containing 10 flies each to sterile LB with 1.25% sucrose.

Fly mortality was monitored at different time points until all the flies were dead. Survival curves were analyzed with log-rank test using SPSS v21 software. If the test was significant, we calculated the odds-ratio and its 95% confidence interval when 50% of the flies were dead.

6.2.7 RNA Extraction and cDNA Synthesis from Non-Infected and Infected Guts

We dissected 20-30 guts from both non-infected and orally infected 5- to 6-day-old females. Flies were infected with the gram-negative bacteria *P. entomophila* as mentioned above, and they were dissected after 12 hours of bacterial exposure. Samples were frozen in liquid nitrogen and stored at -80ºC until sample processing. RNA from gut tissue was extracted using Trizol reagent and PureLink RNA Mini kit (Ambion). We treated RNA on-column with DNase I (Thermo) during the RNA extraction, and we did an additional treatment after the RNA purification. We synthesized cDNA from a total of 500 ng – 1,000 ng of RNA using the Anchored-oligo (dT) primer and Transcription First Strand cDNA Synthesis kit (Roche).

6.2.8 Allele-Specific Expression Analysis (ASE)

For each TE analyzed, we first identified two strains homozygous for the presence of the TE and two strains homozygous for the absence of the TE according to T-lex2 or TIDAL (Fiston-Lavier et al. 2015; Rahman et al. 2015). We then looked for a synonymous SNP linked to the presence of the TE and located in the coding region of the nearby gene. Note that we only selected a SNP when it is present in the coding region of all the alternative transcripts described for that gene. To select the SNP, we downloaded the coding region of the nearby gene from the sequenced DGRP strains available in http://popdrowser.uab.cat/ (Ràmia et al. 2012). Once we identified a diagnostic SNP, we re-sequenced the strain to confirm the presence of the SNP and we performed a PCR to confirm the presence of the TE. We selected a synonymous SNP that is not linked to the TE in all the strains analyzed.

We also analyzed the coding region of the gene in order to discard the presence of nonsynonymous SNPs that could be linked to the TE (Annex Table S2.6A). Additionally, we analyzed the flanking regions of each TE in order to discard other variants that could be linked to the TE, or that could be potentially modifying the gene regulatory regions (Annex Table S2.6B). To do this, we used VISTA to define the conserved regions in the 1 kb TE flanking sequences between *D. melanogaster* and *D. yakuba*, which diverged approximately 11.6 Mya

(Junqueira et al. 2016). We then checked whether there is any SNP linked to the presence of the TE in the DGRP strains (Annex Table S2.6B).

We were not able to analyze five of the candidate TEs: for three TEs, *FBti0019381*, *FBti0061105* and *FBti0062242,* we could not identify homozygous strains for the presence or for the absence of the TE. For *FBti0019564*, we could not identify a diagnostic SNP. Finally, for *tdn17,* we could not design primers to validate the diagnostic SNP due to the presence of repetitive sequences in the nearby gene.

We then crossed a strain with the TE with a strain without the TE differing by the diagnostic SNP to obtain heterozygous flies in which allele-specific expression was measured (Annex Table S2.8). Note that for each TE two crosses were performed so that ASE was measured in two different genetic backgrounds.

ASE was measured in non-infected and infected conditions. We obtained cDNA samples from three biological replicates. We also extracted genomic DNA (gDNA) from 15-20 heterozygous females for each cross, which is needed to correct for any bias in PCR-amplification between alleles (Wittkopp et al. 2011). cDNA and gDNA samples were sent to an external company for primer design and pyrosequencing. We analyzed the pyrosequencing results as described in Wittkopp et al. (2011). Briefly, we calculated the ratios of the allele with the TE and the allele without the TE of the cDNA samples, and we normalized the values with the gDNA ratio. In order to perform the statistics, we transformed the ratios with log2 and we applied a two-tailed t-test in order to check whether there were allele expression differences between the alleles. We corrected the p-values for multiple testing using Benjamini-Hochberg's false discovery rate (FDR) (Benjamini and Hochberg 1995).

6.2.9 TSS Detection

To detect whether *FBti0019386* and *FBti0018868* are adding a Transcription Start Site (TSS) to their nearby gene, as suggested by Batut et al. (2013), we performed RT-PCR in gut tissue of non-infected and infected flies. For *FBti0019386,* associated with *Bin1* gene, we used the forward primer 5'- ATCTGAAGCTCGTTGGTGGG-3' and the reverse primer 5' ATGAGACTCCTGTTTCGCCG- 3' to detect *Bin1* transcript starting in the TE, and the same forward primer with the reverse primer 5' AAGAGCAAAGAGAAGCCGGAA-3' to detect *Bin1* short transcript. For *FBti0018868*, we used the forward primer located inside the TE sequence 5'-TCTTGGCGTTGTCCTTAGTCA -3' and the reverse primer 5'-CTGTCCCTTTCCCGCAATCA -3' to detect the *TM4SF* transcript starting in *FBti0018868*.

6.2.10 Enhancer Assays

We generated transgenic flies carrying the TE sequence in front of the reporter gene *LacZ* by using the *placZ.attB* vector (Bischoff et al. 2007, accession number: KC896840). In order to construct a clone with the correct orientation in the promoter region of *lacZ*, two cloning steps were necessary. We first had to introduce specific restriction sites into the flanking regions for each TE sequence. For that, we amplified the genomic regions containing the TE sequence by using a high fidelity Taq DNA polymerase (Expand High Fidelity PCR system from Sigma), and introduced the restriction sites with the primers used to amplify the region (Annex Table S2.9). After that, we cloned the PCR product into the vector pCR®4-TOPO® (Invitrogen). Finally, we digested both vectors and ligated the TE sequence into the *placZ.attB* and we sequenced the cloned insert to ensure that no polymerase errors were introduced in the PCR step. We purified the vector with the GeneEluteTM Plasmid Miniprep kit (Sigma), and prepared the injection mix at 300 ng/µl vector concentration diluted with injection buffer (5 mM KCl, 0.1 mM sodium phosphate, pH 6.8). The injection mix was sent to an external company to inject embryos from a strain that contain a stable integration site (Bloomington stock #24749). After microinjection, surviving flies were crossed in pairs and the offspring was screened for red eye color, which was diagnostic for stable mutants. We established three transgenic strains for each analyzed TE, which were considered as biological replicates in the expression experiments. As a negative control, we also established transgenic strains with the *placZ.attB* empty vector, in order to control for possible *LacZ* expression driven by the vector sequence.

For *FBti0018868* and *FBti0019985*, we designed primers flanking the TEs and cloned the PCR product in front of the reporter gene *lacZ* (Annex Table S2.9). For the TE *FBti0061506*, which spans only 48 bp, we constructed two different clones to generate two transgenic strains. One strain carries the TE and part of the flanking genomic region, and the other strain contains the same genomic region without the TE. Finally, for the TE *tdn8* we also produced two different clones to generate two transgenic strains. One strain carries the upstream region of *CG10943*, including the 5'UTR, with *tdn8*, and the other strain carries the same genomic region without *tdn8* (Annex Table S2.9).

6.2.11 qRT-PCR Expression Analysis

For the transgenic strains generated in the enhancer assays, we checked *lacZ* expression in non-infected and infected conditions. For most the mutant strains used in the infection survival experiments, we checked by qRT-PCR whether the expression of the mutated gene was affected in non-infected females (Annex Figure S5). We compared this expression to the expression measured in flies with a similar genetic background without the mutation. In all the cases, gene

expression was normalized with the housekeeping gene *Act5c*. The specific primers used for each gene can be found in Annex Table S2.10. We performed all RNA extractions and cDNA synthesis as mentioned above. We performed the qRT-PCR analysis with SYBR Green (BioRad) on an iQ5 Thermal cycler. Results were analyzed using the dCT method and following the recommendations of the MIQE guideline (Bustin et al. 2009).

6.2.12 Immunofluorescence Staining

We performed immunofluorescence gut staining to localize β-*GAL* expression in the transgenic flies from the enhancer assays, both in non-infected and infected conditions. Flies were dissected and gut tissue was fixed with 4% Formaldehyde. The tissue was then stained by using the primary antibody mouse anti-βGalactosidase (Hybridoma bank 40-1a), and the secondary antibody anti-mouse Alexa Fluor ® 555 (Sigma). Images were analyzed and captured using a Leica SP5 confocal microscope.

07

BIBLIOGRAPHY

7. BIBLIOGRAPHY

Adams MD. 2000. The Genome Sequence of *Drosophila melanogaster*. Science. 287:2185–2195.

Agrawal a, Eastman QM, Schatz DG. 1998. Implications of transposition mediated by V (D) J-recombination proteins RAG1 and RAG2 for origins of antigen-specific immunity. Nature 394:744–751.

Aguadé M. 2008. Nucleotide and Copy-Number Polymorphism at the Odorant Receptor Genes *Or22a* and Or22b in *Drosophila melanogaster*. Mol. Biol. Evol. 26:61–70.

Ali S, Laurie JD, Linning R, Cervantes-Chávez JA, Gaudet D, Bakkeren G. 2014. An Immunity-Triggering Effector from the Barley Smut Fungus Ustilago hordei Resides in an Ustilaginaceae-Specific Cluster Bearing Signs of Transposable Element-Assisted Evolution. PLoS Pathog. 10. (7):e1004223.

Almeida LM, Amaral MEJ, Silva IT, Silva WA, Riggs PK, Carareto CM. 2008. Report of a chimeric origin of transposable elements in a bovine-coding gene. Genet. Mol. Res. 7:107–116.

Aminetzach YT, Macpherson JM, Petrov DA. 2005. Pesticide resistance via transposition-mediated adaptive gene truncation in *Drosophila*. Science 309:764–767.

Aravin AA, Naumova NM, Tulin A V, Vagin V V, Rozovsky YM, Gvozdev VA. 2001. Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. Curr. Biol. 11:1017–1027.

Ashburner M, Bergman CM. 2005. *Drosophila melanogaster*: A case study of a model genomic sequence and its consequences. Genome Res. 15:1661–1667.

Ayres JS, Freitag N, Schneider DS. 2008. Identification of *Drosophila* mutants altering defense of and endurance to Listeria monocytogenes infection. Genetics 178:1807–1815.

Barker JSF. 2009. Defining fitness in natural and domesticated populations. Adapt. Fit. Anim. Popul.:3–14.

Barnett DW, Garrison EK, Quinlan AR, Strömberg MP, Marth GT. 2011. Bamtools: A C++ API and toolkit for analyzing and managing BAM files. Bioinformatics 27:1691–1692.

Barrett RDH, Hoekstra HE. 2011. Molecular spandrels: tests of adaptation at the genetic level. Nat. Rev. Genet. 12:767–780.

Barrón MG, Fiston-Lavier A-S, Petrov DA, González J. 2014. Population genomics of transposable elements in Drosophila. Annu. Rev. Genet. 48:561–581.

Bassett A, Liu JL. 2014. CRISPR/Cas9 mediated genome engineering in *Drosophila*. Methods 69:128–136.

Bastide H, Betancourt A, Nolte V, Tobler R, Stöbe P, Futschik A, Schlötterer C. 2013. A Genome-Wide, Fine-Scale Map of Natural Pigmentation Variation in *Drosophila melanogaster*. PLoS Genet. 9:e1003534.

Bastide H, Lange JD, Lack JB, Yassin A, Pool JE. 2016. A Variable Genetic Architecture of Melanic Evolution in *Drosophila melanogaster*. Genetics 204:1307–1319.

Batut P, Dobin A, Plessy C, Carninci P, Gingeras TR. 2013. High-fidelity promoter profiling reveals widespread alternative promoter usage and transposon-driven developmental gene expression. Genome Res. 23:169–180.

Bellen HJ, Levis RW, Liao G, He Y, Carlson JW, Tsang G, Evans-Holm M, Hiesinger PR, Schulze KL, Rubin GM, et al. 2004. The BDGP gene disruption project: Single transposon insertions associated with 40% of *Drosophila* genes. Genetics 167:761–781.

Bellen HJ, Tong C, Tsuda H. 2010. 100 years of *Drosophila* research and its impact on vertebrate neuroscience: a history lesson for the future. Nat. Rev. Neurosci. 11:514–522.

Benjamini Y, Hochberg Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful

Approach to Multiple Testing. J. R. Stat. Soc. Ser. B 57:289–300.

Berg JJ, Coop G. 2014. A Population Genetic Signal of Polygenic Adaptation. PLoS Genet. 10(8):e1004412.

Bergland AO, Behrman EL, O'Brien KR, Schmidt PS, Petrov DA. 2014. Genomic Evidence of Rapid and Stable Adaptive Oscillations over Seasonal Time Scales in *Drosophila*. PLoS Genet. 10 (11):e1004775.

Bergland AO, Tobler R, González J, Schmidt P, Petrov D. 2016. Secondary contact and local adaptation contribute to genome-wide patterns of clinal variation in *Drosophila melanogaster*. Mol. Ecol. 25:1157–1174.

Bergman P, Seyedoleslami Esfahani S, Engström Y. 2016. *Drosophila* as a Model for Human Diseases-Focus on Innate Immunity in Barrier Epithelia. Curr. Top. Dev. Biol.:29–81.

Berry AJ, Ajioka JW, Kreitman M. 1991. Lack of polymorphism on the drosophila fourth chromosome resulting from selection. Genetics 129:1111–1117.

Betancourt AJ, Presgraves DC. 2002. Linkage limits the power of natural selection in *Drosophila*. Proc. Natl. Acad. Sci. U. S. A. 99:13616–13620.

Biémont C, Vieira C. 2006. Genetics: junk DNA as an evolutionary force. Nature 443:521–524.

Bigham AW, Lee FS. 2014. Human high-altitude adaptation: Forward genetics meets the HIF pathway. Genes Dev. 28:2189–2204.

Bischof J, Maeda RK, Hediger M, Karch F, Basler K. 2007. An optimized transgenesis system for *Drosophila* using germ-line-specific phiC31 integrases. Proc. Natl. Acad. Sci. U. S. A. 104:3312–3317.

Blainey P, Krzywinski M, Altman N. 2014. Points of Significance: Replication. Nat. Methods 11:879–880.

Blumenstiel JP, Chen X, He M, Bergman CM. 2014. An age-of-allele test of neutrality for transposable element insertions. Genetics 196:523–538.

Bock I, Parsons P. 1981. Species of Australia and New Zealand. The Genetics and Biology of *Drosophila*, vol. 3a. (Ashburner M, Carson H, Thompson J, editors.). London: Academic press

Boman HG, Nilsson I, Rasmuson B. 1972. Inducible antibacterial defence system in *Drosophila*. Nature 237:232–235.

Bonfini A, Liu X, Buchon N. 2016. From pathogens to microbiota: How *Drosophila* intestinal stem cells react to gut microbes. Dev. Comp. Immunol. 64:22–38.

Bou Sleiman MS, Osman D, Massouras A, Hoffmann AA, Lemaitre B, Deplancke B. 2015. Genetic, molecular and physiological basis of variation in *Drosophila* gut immunocompetence. Nat Commun 6:7829.

Le Bourg E, Massou I, Gobert V. 2009. Cold stress increases resistance to fungal infection throughout life in *Drosophila melanogaster*. Biogerontology 10:613–625.

Boutros M, Agaisse H, Perrimon N. 2002. Sequential activation of signaling pathways during innate immune responses in Drosophila. Dev. Cell 3:711–722.

Bozicevic V, Hutter S, Stephan W, Wollstein A. 2016. Population genetic evidence for cold adaptation in European *Drosophila melanogaster* populations. Mol. Ecol. 25:1175–1191.

Brandt J, Schrauth S, Veith AM, Froschauer A, Haneke T, Schultheis C, Gessler M, Leimeister C, Volff JN. 2005. Transposable elements as a source of genetic innovation: Expression and evolution of a family of retrotransposon-derived neogenes in mammals. Gene 345:101–111.

Britten RJ, Davidson EH. 1971. Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty. Q. Rev. Biol. 46:111–138.

Broderick NA, Buchon N, Lemaitre B. 2014. Microbiota-Induced Changes in *Drosophila melanogaster* Host Gene Expression and Gut Morphology. MBio 5:e01117-14.

Brown AE, Baumbach J, Cook PE, Ligoxygakis P. 2009. Short-term starvation of immune deficient *Drosophila* improves survival to gram-negative bacterial infections. PLoS One 4:e4490.

Buchmann K. 2014. Evolution of innate immunity: Clues from invertebrates via fish to mammals. Front.

Immunol. 5:1–8.

Buchon N, Broderick NA, Lemaitre B. 2013. Gut homeostasis in a microbial world: insights from *Drosophila melanogaster*. Nat. Rev. Microbiol. 11:615–626.

Buchon N, Broderick NA, Poidevin M, Pradervand S, Lemaitre B. 2009. *Drosophila* Intestinal Response to Bacterial Infection: Activation of Host Defense and Stem Cell Proliferation. Cell Host Microbe 5:200–211.

Buchon N, Osman D, David FPA, Yu Fang H, Boquete JP, Deplancke B, Lemaitre B. 2013. Morphological and Molecular Characterization of Adult Midgut Compartmentalization in *Drosophila*. Cell Rep. 3:1725–1738.

Buchon N, Silverman N, Cherry S. 2014. Immunity in *Drosophila melanogaster* - from microbial recognition to whole-organism physiology. Nat. Rev. Immunol. 14:796–810.

Burnett C, Valentini S, Cabreiro F, Goss M, Somogyvári M, Piper MD, Hoddinott M, Sutphin GL, Leko V, McElwee JJ, et al. 2011. Absence of effects of *Sir2* overexpression on lifespan in *C. elegans* and *Drosophila*. Nature 477:482–485.

Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, et al. 2009. The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. Clin. Chem. 55:611–622.

Cáceres M, Ranz JM, Barbadilla A, Long M, Ruiz A. 1999. Generation of a Widespread *Drosophila* Inversion by a Transposable Element. Science 285:415–418.

Campo D, Lehmann K, Fjeldsted C, Souaiaia T, Kao J, Nuzhdin SV. 2013. Whole-genome sequencing of two North American *Drosophila melanogaster* populations reveals genetic differentiation and positive selection. Mol. Ecol. 22:5084–5097.

Campos JL, Halligan DL, Haddrill PR, Charlesworth B. 2014. The relation between recombination rate and patterns of molecular evolution and variation in *Drosophila melanogaster*. Mol. Biol. Evol. 31(4): 1010-1028

Capo F, Charroux B, Royet J. 2016. Bacteria sensing mechanisms in *Drosophila* gut: Local and systemic consequences. Dev. Comp. Immunol. 64:11–21.

Di Cara F, King-Jones K. 2013. How clocks and hormones act in concert to control the timing of insect development. Curr. Top. Dev. Biol. 105:1–36.

Caracristi G, Schlötterer C. 2003. Genetic differentiation between American and European *Drosophila melanogaster* populations could be attributed to admixture of African alleles. Mol. Biol. Evol. 20:792–799.

Carnero-Montoro E, Bonet L, Engelken J, Bielig T, Martínez-Florensa M, Lozano F, Bosch E. 2012. Evolutionary and functional evidence for positive selection at the human CD5 immune receptor gene. Mol. Biol. Evol. 29:811–823.

Casacuberta E, González J. 2013. The impact of transposable elements in environmental adaptation. Mol. Ecol. 22:1503–1517.

Casillas S, Barbadilla A. 2017. Molecular Population Genetics. Genetics 569:551–569.

Casola C, Hucks D, Feschotte C. 2008. Convergent domestication of pogo-like transposases into centromere-binding proteins in fission yeast and mammals. Mol. Biol. Evol. 25:29–41.

Castellano D, Coronado-Zamora M, Campos JL, Barbadilla A, Eyre-Walker A. 2016. Adaptive evolution is substantially impeded by hill-Robertson interference in drosophila. Mol. Biol. Evol. 33:442–455.

Cavalli-Sforza LL, Bodmer WF. 1971. The genetics of human populations. Dover Publications. San Francisco.

Chakrabarti S, Liehl P, Buchon N, Lemaitre B. 2012. Infection-induced host translational blockage inhibits immune responses and epithelial renewal in the *Drosophila* gut. Cell Host Microbe 12:60–70.

Chandler CH, Chari S, Dworkin I. 2013. Does your gene need a background check? How genetic

background impacts the analysis of mutations, genes, and evolution. Trends Genet.:1–9.

Chang KT, Min K-T. 2009. Upregulation of three *Drosophila* homologs of human chromosome 21 genes alters synaptic function: implications for Down syndrome. Proc. Natl. Acad. Sci. U. S. A. 106:17117–17122.

Chang KT, Shi Y-J, Min K-T. 2003. The *Drosophila* homolog of Down's syndrome critical region 1 gene regulates learning: implications for mental retardation. Proc. Natl. Acad. Sci. U. S. A. 100:15794–15799.

Charlesworth B, Morgan MT, Charlesworth D. 1993. The effect of deleterious mutations on neutral molecular variation. Genetics 134:1289–1303.

Chintapalli VR, Wang J, Dow J a T. 2007. Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. Nat. Genet. 39:715–720.

Chippindale A, Alipaz J, Chen H, Rose M. 1997. Experimental evolution of accelerated development in *Drosophila*. 1. Developmental speed and larval survival. Evolution (N. Y). 51:1536–1551.

Chung H, Bogwitz MR, McCart C, Andrianopoulos A, Ffrench-Constant RH, Batterham P, Daborn PJ. 2007. Cis-regulatory elements in the accord retrotransposon result in tissue-specific expression of the *Drosophila melanogaster* insecticide resistance gene Cyp6g1. Genetics 175:1071–1077.

Chuong EB, Elde NC, Feschotte C. 2016. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. Science. 351:1083–1087.

Chuong EB, Elde NC, Feschotte C. 2017. Regulatory activities of transposable elements: from conflicts to benefits. Nat. Rev. 18, 71-86.

Clemson AS, Sgrò CM, Telonis-Scott M. 2016. Thermal plasticity in *Drosophila melanogaster* populations from eastern Australia: Quantitative traits to transcripts. J. Evol. Biol.

Cogni R, Kuczynski C, Koury S, Lavington E, Behrman EL, O'Brien KR, Schmidt PS, Eanes WF. 2014. The intensity of selection acting on the *couch potato* gene-spatial-temporal variation in a diapause cline. Evolution (N. Y). 68:538–548.

Collins S, de Meaux J. 2009. Adaptation to different rates of environmental change in *Chlamydomonas*. Evolution 63:2952–2965.

Comeron JM, Ratnappan R, Bailin S. 2012. The Many Landscapes of Recombination in *Drosophila melanogaster*. PLoS Genet. 8:e1002905.

Cordaux R, Udit S, Batzer MA, Feschotte C. 2006. Birth of a chimeric primate gene by capture of the transposase gene from a mobile element. Proc Natl Acad Sci U S A 103:8101–8106.

Costa E, Beltran S, Espins ML. 2011. *Drosophila melanogaster* SAP18 protein is required for environmental stress responses. FEBS Lett. 585:275–280.

Costa R, Peixoto AA, Barbujani G, Kyriacou CP. 1992. A Latitudinal Cline in a *Drosophila* Clock Gene. Proc. R. Soc. B Biol. Sci. 250:43–49.

Cridland JM, Macdonald SJ, Long AD, Thornton KR. 2013. Abundance and distribution of transposable elements in two *Drosophila* QTL mapping resources. Mol. Biol. Evol. 30:2311–2327.

Czech B, Hannon GJ. 2016. One Loop to Rule Them All: The Ping-Pong Cycle and piRNA-Guided Silencing. Trends Biochem. Sci. 41:324–337.

Daborn PJ, Yen JL, Bogwitz MR, Le Goff G, Feil E, Jeffers S, Tijet N, Perry T, Heckel D, Batterham P, et al. 2002. A Single P450 Allele Associated with Insecticide Resistance in *Drosophila*. Science. 297:2253–2256.

David J, Capy P, Payant V, Tsakas S. 1985. Thoracic trident pigmentation in *Drosophila melanogaster*: Differentiation of geographical populations. Genet. Sel. Evol. 17:211.

David J, Merçot H, Capy P, McEvey S, Van Herrewege J. 1986. Alcohol tolerance and Adh gene frequencies in European and African populations of *Drosophila melanogaster.* Genet. Sel. Evol. 18:405–416.

David JR, Capy P. 1988. Genetic variation of *Drosophila melanogaster* natural populations. Trends Genet. 4:106–111.

Davies SA, Cabrero P, Overend G, Aitchison L, Sebastian S, Terhzaz S, Dow JAT. 2014. Cell signalling mechanisms for insect stress tolerance. J. Exp. Biol. 217:119–128.

Deschamps M, Laval G, Fagny M, Itan Y, Abel L, Casanova J-L, Patin E, Quintana-Murci L. 2016. Genomic Signatures of Selective Pressures and Introgression from Archaic Hominins at Human Innate Immunity Genes. Am. J. Hum. Genet. 98:5–21.

Dietzl G, Chen D, Schnorrer F, Su K-C, Barinova Y, Fellner M, Gasser B, Kinsey K, Oppel S, Scheiblauer S, et al. 2007. A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. Nature 448:151–156.

Ding Y, Berrocal A, Morita T, Longden KD, Stern DL. 2016. Natural courtship song variation caused by an intronic retroelement in an ion channel gene. Nature 536:329–332.

Dobzhansky T. 1951. Genetics and the Origin of Species. 2nd ed. New York: Columbia University Press

Dobzhansky T. 1955. A review of some fundamental concepts and problems of population genetics. Cold Spring Harb. Symp. Quant. Biol. 20:1–15.

Doolittle WF, Sapienza C. 1980. Selfish genes, the phenotype paradigm and genome evolution. Nature 284:601–603.

Duchen P, Živković D, Hutter S, Stephan W, Laurent S. 2013. Demographic inference reveals african and european admixture in the north american *Drosophila melanogaster* population. Genetics 193:291–301.

Early AM, Arguello JR, Cardoso-Moreira M, Gottipati S, Grenier JK, Clark AG. 2016. Survey of Global Genetic Diversity Within the *Drosophila* Immune System. Genetics. 205(1), 353-366.

Edward D a, Chapman T. 2011. Mechanisms underlying reproductive trade-offs: costs of reproduction. Mech. Life Hist. Evol. 423:137–152.

Eickbush TH, Furano A V. 2002. Fruit flies and humans respond differently to retrotransposons. Curr. Opin. Genet. Dev. 12:669–674.

Ejima A, Tsuda M, Takeo S, Ishii K, Matsuo T, Aigaki T. 2004. Expression level of sarah, a homolog of DSCR1, is critical for ovulation and female courtship behavior in *Drosophila melanogaster*. Genetics 168:2077–2087.

Elbarbary RA, Lucas BA, Maquat LE. 2016. Retrotransposons as regulators of gene expression. Science 351: 6274, aac7247.

Ellison CE, Bachtrog D. 2013. Dosage Compensation via Transposable Element Mediated Rewiring of a Regulatory Network. Science. 342:846–850.

Endler L, Betancourt AJ, Nolte V, Schlötterer C. 2016. Reconciling differences in pool-GWAS between populations: A case study of female abdominal pigmentation in *Drosophila melanogaster*. Genetics 202:843–855.

Evrony GD, Cai X, Lee E, Hills LB, Elhosary PC, Lehmann HS, Parker JJ, Atabay KD, Gilmore EC, Poduri A, et al. 2012. Single-Neuron Sequencing Analysis of L1 Retrotransposition and Somatic Mutation in the Human Brain. Cell 151:483–496.

Fabian DK, Kapun M, Nolte V, Kofler R, Schmidt PS, Schlötterer C, Flatt T. 2012. Genome-wide patterns of latitudinal differentiation among populations of *Drosophila melanogaster* from North America. Mol. Ecol. 21:4748–4769.

Fabian DK, Lack JB, Mathur V, Schlötterer C, Schmidt PS, Pool JE, Flatt T. 2015. Spatially varying selection shapes life history clines among populations of *Drosophila melanogaster* from sub-Saharan Africa. J. Evol. Biol. 28:826–840.

Fernández-Sampedro MA, Invergo BM, Ramon E, Bertranpetit J, Garriga P. 2016. Functional role of positively selected amino acid substitutions in mammalian rhodopsin evolution. Sci. Rep. 6:21570.

Feschotte C. 2008. Transposable elements and the evolution of regulatory networks. Nat. Rev. Genet. 9:397–405.

Filee J, Siguier P, Chandler M. 2007. Insertion Sequence Diversity in Archaea. Microbiol. Mol. Biol. Rev. 71:121–157.

Fisher RA. 1930. The Genetical Theory of Natural Selection. Oxford: Clarendon Press

Fiston-Lavier ASA-S, Barrón MG, Petrov DA, González J, Barron MG, Petrov DA, Gonzalez J. 2015. T-lex2: Genotyping, frequency estimation and re-annotation of transposable elements using single or pooled next-generation sequencing data. Nucleic Acids Res. 43:e22–e22.

Fiston-Lavier AS, Singh ND, Lipatov M, Petrov DA. 2010. *Drosophila melanogaster* recombination rate calculator. Gene 463:18–20.

Folguera G, Ceballos S, Spezzi L, Fananra JJ, Hasson E. 2008. Clinal variation in developmental time and viability, and the response to thermal treatments in two species of *Drosophila*. Biol. J. Linn. Soc. 95:233–245.

Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. 2004. VISTA: Computational tools for comparative genomics. Nucleic Acids Res. 32:W273-9.

Fumagalli M, Sironi M, Pozzoli U, Ferrer-Admettla A, Pattini L, Nielsen R, Nielsen R. 2011. Signatures of Environmental Genetic Adaptation Pinpoint Pathogens as the Main Selective Pressure through Human Evolution.Akey JM, editor. PLoS Genet. 7:e1002355.

Gao B, Shen D, Xue S, Chen C, Cui H, Song C. 2016. The contribution of transposable elements to size variations between four teleost genomes. Mob. DNA 7:4.

Garcia-Pérez JL. 2016. Transposons and Retrotransposons. Methods and Protocols. 1400:79–93.

Garcia-Perez JL, Doucet AJ, Bucheton A, Moran J V, Gilbert N. 2007. Distinct mechanisms for trans-mediated mobilization of cellular RNAs by the LINE-1 reverse transcriptase. Genome Res. 17:602–611.

Gasch AP, Payseur BA, Pool JE. 2016. The Power of Natural Variation for Model Organism Biology. Trends Genet. 32:147–154.

Gerdes P, Richardson SR, Mager DL, Faulkner GJ. 2016. Transposable elements in the mammalian embryo: pioneers surviving through stealth and service. Genome Biol. 17:100.

Gerke J, Lorenz K, Ramnarine S, Cohen B. 2010. Gene-environment interactions at nucleotide resolution. PLoS Genet. 6(9): e1001144.

Gibert P, Moreteau B, Pétavy G, Karan D, David JR. 2007. Chill-Coma Tolerance, a Major Climatic Adaptation Among *Drosophila* Species. Evolution (N. Y). 55:1063–1068.

Gobert V, Gottar M, Matskevich AA, Rutschmann S, Royet J, Belvin M, Hoffmann JA, Ferrandon D. 2003. Dual activation of the *Drosophila* Toll pathway by two pattern recognition receptors. Science. 302:2126–2130.

Goic B, Vodovar N, Mondotte JA, Monot C, Frangeul L, Blanc H, Gausson V, Vera-Otarola J, Cristofari G, Saleh M-C. 2013. RNA-mediated interference and reverse transcription control the persistence of RNA viruses in the insect model *Drosophila*. Nat. Immunol. 14:396–403.

González J, Karasov TL, Messer PW, Petrov DA. 2010. Genome-wide patterns of adaptation to temperate environments associated with transposable elements in *Drosophila*. PLoS Genet. 6:33–35.

González J, Lenkov K, Lipatov M, Macpherson JM, Petrov DA. 2008. High Rate of Recent Transposable Element Induced Adaptation in *Drosophila melanogaster*. PLoS Biol 6:e251.

González J, Macpherson JM, Petrov DA. 2009. A recent adaptive transposable element insertion near highly conserved developmental *loci* in *Drosophila melanogaster*. Mol. Biol. Evol. 26:1949–1961.

Gramates LS, Marygold SJ, Santos G dos, Urbano J-M, Antonazzo G, Matthews BB, Rey AJ, Tabone CJ, Crosby MA, Emmert DB, et al. 2017. FlyBase at 25: looking to the future. Nucleic Acids Res. 45:D663–D671.

David JR, Gibert P, Pla E, Petavy G, Karan D, Moreteau B. 1998. Cold stress tolerance in *Drosophila*: analysis of chill coma recovery in *D. melanogaster*. J. Therm. Biol. 23:291–299.

De Gregorio E, Spellman PT, Rubin GM, Lemaitre B. 2001. Genome-wide analysis of the *Drosophila* immune response by using oligonucleotide microarrays. Proc. Natl. Acad. Sci. U. S. A. 98:12590–12595.

De Gregorio E, Spellman PT, Tzou P, Rubin GM, Lemaitre B. 2002. The Toll and Imd pathways are the major regulators of the immune response in *Drosophila*. EMBO J. 21:2568–2579.

Disdero E, Filée J. 2017. LoRTE: Detecting transposon-induced genomic variants using low coverage PacBio long read sequences. Mob DNA. 8: 8:5.

Guerrero FD, Jamroz RC, Kammlah D, Kunz SE. 1997. Toxicological and molecular characterization of pyrethroid-resistant horn flies, *Haematobia irritans:* Identification of kdr and super-kdr point mutations. Insect Biochem. Mol. Biol. 27:745–755.

Guio L, Barrõn MG, González J. 2014. The transposable element *Bari-Jheh* mediates oxidative stress response in *Drosophila*. Mol. Ecol. 23:2020–2030.

Guio L, González J. 2015. The dominance effect of the adaptive transposable element insertion *Bari-Jheh* depends on the genetic background. Genome Biol. Evol. 7:1260–1266.

Ha E-M, Oh C-T, Bae YS, Lee W-J. 2005. A direct role for dual oxidase in *Drosophila* gut immunity. Science 310:847–850.

Haldane JBS. 1932. The causes of evolution. Cornell University Press. London.

Hancock AM, Witonsky DB, Ehler E, Alkorta-Aranburu G, Beall C, Gebremedhin A, Sukernik R, Utermann G, Pritchard JK, Coop G, et al. 2010. Colloquium paper: human adaptations to diet, subsistence, and ecoregion are due to subtle shifts in allele frequency. Proc. Natl. Acad. Sci. U. S. A. 107 Suppl:8924–8930.

Hand SC, Denlinger DL, Podrabsky JE, Roy R. 2016. Mechanisms of animal diapause: recent developments from nematodes, crustaceans, insects, and fish. Am. J. Physiol. - Regul. Integr. Comp. Physiol. 310:R1193–R1211.

Harris H. 1966. Enzyme polymorphisms in man. Proc. R. Soc. Lond. B. Biol. Sci. 164:298–310.

He X. 2016. The Biology Complicated by Genetic Analysis. Mol. Biol. Evol. 33:2177–2181.

Hickey DA. Selfish DNA: a sexually-transmitted nuclear parasite. Genetics 101:519–531.

Hill WG, Robertson A. 1966. The effect of linkage on limits to artificial selection. Genet. Res. 8:269–294.

Hoban S, Kelley JL, Lotterhos KE, Antolin MF, Bradburd G, Lowry DB, Poss ML, Reed LK, Storfer A, Whitlock MC. 2016. Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and Future Directions. Am. Nat. 188:379–397.

Hoekstra HE, Hirschmann RJ, Bundey RA, Insel PA, Crossland JP. 2006. A Single Amino Acid Mutation Contributes to Adaptive Beach Mouse Color Pattern. Science. 313:101–104.

Hoffmann AA, Sørensen JG, Loeschcke V. 2003. Adaptation of *Drosophila* to temperature extremes: Bringing together quantitative and molecular approaches. J. Therm. Biol. 28:175–216.

Hofmann CM, O'Quin KE, Justin Marshall N, Cronin TW, Seehausen O, Carleton KL. 2009. The eyes have it: Regulatory and structural changes both underlie cichlid visual pigment diversity. PLoS Biol. 7(12):e1000266.

Hogan PG, Chen L, Nardone J, Rao A. 2003. Transcriptional regulation by calcium, calcineurin, and NFAT. 15;17(18):2205–2232.

Hong H. 2012. Pitfall of genome-wide association studies: Sources of inconsistency in genotypes and their effects. J. Biomed. Sci. Eng. 5:557–573.

Horner VL, Czank A, Jang JK, Singh N, Williams BC, Puro J, Kubli E, Hanes SD, McKim KS, Wolfner MF, et al. 2006. The *Drosophila* Calcipressin *Sarah* Is Required for Several Aspects of Egg Activation. Curr. Biol. 16:1441–1446.

Hotson AG, Schneider DS. 2015. *Drosophila melanogaster* Natural Variation Affects Growth Dynamics of Infecting Listeria monocytogenes. G3 Genes|Genomes|Genetics 5:2593–2600.

Howick VM, Lazzaro BP. 2017. The Genetic Architecture of Defense as Resistance to and Tolerance of Bacterial Infection in *Drosophila melanogaster*. Mol. Ecol.:1533–1546.

Hua-Van A, Le Rouzic A, Boutin TS, Filée J, Capy P. 2011. The struggle for life of the genome's selfish architects. Biol. Direct 6:19.

Huang DW, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4:44–57.

Huang S, Tao X, Yuan S, Zhang Y, Li P, Beilinson H, Zhang Y, Yu W, Pontarotti P, Escriva H, et al. 2016. Discovery of an Active RAG Transposon Illuminates the Origins of V(D)J Recombination. Cell 166:102–114.

Huang W, Massouras A, Inoue Y, Peiffer J, Ràmia M, Tarone AM, Turlapati L, Zichner T, Zhu D, Lyman RF, et al. 2014. Natural variation in genome architecture among 205 *Drosophila melanogaster* Genetic Reference Panel lines. Genome Res. 24:1193–1208.

Huang W, Richards S, Carbone MA, Zhu D, Anholt RRH, Ayroles JF, Duncan L, Jordan KW, Lawrence F, Magwire MM, et al. 2012. Epistasis dominates the genetic architecture of *Drosophila* quantitative traits. Proc. Natl. Acad. Sci. 109:15553–15559.

Hudson RR. 2002. Generating samples under a Wright-Fisher neutral model of genetic variation. Bioinformatics 18:337–338.

Hudson RR, Kaplan NL. 1995. Deleterious background selection with recombination. Genetics 141:1605–1617.

Hudson RR, Slatkin M, Maddison WP. 1992. Estimation of levels of gene flow from DNA sequence data. Genetics 132:583–589.

Huisinga KL, Riddle NC, Leung W, Shimonovich S, McDaniel S, Figueroa-Clarevega A, Elgin SCR. 2016. Targeting of P-element reporters to heterochromatic domains by transposable element 1360 in *Drosophila melanogaster*. Genetics 202:565–582.

Irving P, Troxler L, Heuer TS, Belvin M, Kopczynski C, Reichart J-M, Hoffman JA, Hetru C, Reichhart JM, Hoffmann JA, et al. 2001. A genome-wide analysis of immune responses in *Drosophila*. Proc. Natl. Acad. Sci. U. S. A. 98:15119–15124.

James AC, Azevedo RB, Partridge L. 1997. Genetic and environmental responses to temperature of *Drosophila melanogaster* from a latitudinal cline. Genetics 146:881–890.

James AC, Partridge L. 1995. Thermal evolution of rate of larval development in *Drosophila melanogaster* in laboratory and field populations. J. Evol. Biol. 8:315–330.

Jennings BH. 2011. Drosophila-a versatile model in biology & medicine. Mater. Today 14:190–195.

Jukes TH, Cantor CR. 1969. Evolution of Protein Molecules. In: Mammalian Protein Metabolism: Volume III. p. 21–132.

Juneja P, Quinn A, Jiggins FM. 2016. Latitudinal clines in gene expression and cis -regulatory element variation in *Drosophila melanogaster*. BMC Genomics:1–11.

Junqueira A, Azeredo-Espin A, Paulo D, Marinho M, Tomsho LP, Drautz-Moses DI, Purbojati RW, Ratan A, Schuster SC. 2016. Large-scale mitogenomics enables insights into Schizophora (Diptera) radiation and population diversity. Sci. Rep. 6:21762.

Kao JY, Zubair A, Salomon MP, Nuzhdin SV., Campo D. 2015. Population genomic analysis uncovers African and European admixture in *Drosophila melanogaster* populations from the south-eastern United States and Caribbean Islands. Mol. Ecol. 24:1499–1509.

Kapitonov V V., Jurka J. 2001. Rolling-circle transposons in eukaryotes. Proc. Natl. Acad. Sci. U. S. A. 98:8714–8719.

Kapitonov V V, Jurka J. 2003. Molecular paleontology of transposable elements in the *Drosophila*

melanogaster genome. Proc. Natl. Acad. Sci. U. S. A. 100:6569–6574.

Kapitonov V V., Jurka J. 2006. Self-synthesizing DNA transposons in eukaryotes. Proc. Natl. Acad. Sci. 103:4540–4545.

Kapun M, Fabian DK, Goudet J, Flatt T. 2016. Genomic Evidence for Adaptive Inversion Clines in *Drosophila melanogaster*. Mol. Biol. Evol. 33:1317–1336.

Kapun M, Schmidt C, Durmaz E, Schmidt PS, Flatt T. 2016. Parallel effects of the inversion In(3R)Payne on body size across the North American and Australian clines in *Drosophila melanogaster*. J. Evol. Biol. 29:1059–1072.

Keller A. 2007. Drosophila melanogaster's history as a human commensal. Curr. Biol. 17:77–81.

Kennington WJ, Gockel J, Partridge L. 2003. Testing for asymmetrical gene flow in a *Drosophila melanogaster* body size cline. Genetics 165:667–673.

Kim A, Terzian C, Santamaria P, Pélisson A, Purd'homme N, Bucheton A. 1994. Retroviruses in invertebrates: the gypsy retrotransposon is apparently an infectious retrovirus of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. U. S. A. 91:1285–1289.

Kim S-H, Lee W-J. 2014. Role of DUOX in gut inflammation: lessons from *Drosophila* model of gut-microbiota interactions. Front. Cell. Infect. Microbiol. 3:116.

Kimura M. 1991. The neutral theory of molecular evolution: a review of recent evidence. Jpn. J. Genet. 66:367–386.

Klepsatel P, Gáliková M, Huber CD, Flatt T. 2014. Similarities and differences in altitudinal versus latitudinal variation for morphological traits in *Drosophila melanogaster*. Evolution (N. Y). 68:1385–1398.

Knisbacher BA, Levanon EY. 2016. DNA editing of LTR retrotransposons reveals the impact of APOBECs on vertebrate genomes. Mol. Biol. Evol. 33:554–567.

Kofler R, Betancourt AJ, Schlötterer C. 2012. Sequencing of pooled DNA samples (Pool-Seq) uncovers complex dynamics of transposable element insertions in *Drosophila melanogaster*. PLoS Genet. 8(1):e1002487.

Kofler R, Gómez-Sánchez D, Schlötterer C. 2016. PoPoolationTE2: Comparative Population Genomics of Transposable Elements Using Pool-Seq. Mol. Biol. Evol. 33:2759–2764.

Kolaczkowski B, Kern AD, Holloway AK, Begun DJ. 2011. Genomic differentiation between temperate and tropical Australian populations of *Drosophila melanogaster*. Genetics 187:245–260.

Kounatidis I, Ligoxygakis P. 2012. *Drosophila* as a model system to unravel the layers of innate immunity to infection. Open Biol. 2:120075.

Kramerov DA, Vassetzky NS. 2005. Short retroposons in eukaryotic genomes. Int. Rev. Cytol. 247:165–221.

Kyriacou CP, Peixoto AA, Sandrelli F, Costa R, Tauber E. 2008. Clines in clock genes: fine-tuning circadian rhythms to the environment. Trends Genet. 24:124–132.

Lack JB, Cardeno CM, Crepeau MW, Taylor W, Corbett-Detig RB, Stevens KA, Langley CH, Pool JE. 2015. The *Drosophila* genome nexus: A population genomic resource of 623 *Drosophila melanogaster* genomes, including 197 from a single ancestral range population. Genetics 199:1229–1241.

Lavington E, Cogni R, Kuczynski C, Koury S, Behrman EL, O'brien KR, Schmidt PS, Eanes WF. 2014. A small system-high-resolution study of metabolic adaptation in the central metabolic pathway to temperate climates in *Drosophila melanogaster*. Mol. Biol. Evol. 31:2032–2041.

Lazzaro BP, Flores HA, Lorigan JG, Yourth CP. 2008. Genotype-by-environment interactions and adaptation to local temperature affect immunity and fecundity in *Drosophila melanogaster*. PLoS Pathog. 4(3):e1000025.

Leblanc P, Desset S, Giorgi F, Taddei AR, Fausto AM, Mazzini M, Dastugue B, Vaury C. 2000. Life cycle of an endogenous retrovirus, ZAM, in *Drosophila melanogaster*. J. Virol. 74:10658–10669.

Lee SF, Sgrò CM, Shirriffs J, Wee CW, Rako L, van Heerwaarden B, Hoffmann AA. 2011. Polymorphism in the *couch potato* gene clines in eastern Australia but is not associated with ovarian dormancy in *Drosophila melanogaster*. Mol. Ecol. 20:2973–2984.

Leem Y-E, Ripmaster TL, Kelly FD, Ebina H, Heincelman ME, Zhang K, Grewal SIS, Hoffman CS, Levin HL. 2008. Retrotransposon Tf1 Is Targeted to Pol II Promoters by Transcription Activators. Mol. Cell 30:98–107.

Lemaitre B, Hoffmann J. 2007. The Host Defense of *Drosophila melanogaster*. Annu. Rev. Immunol. 25:697–743.

Lemaitre B, Nicolas E, Michaut L, Reichhart JM, Hoffmann JA. 1996. The dorsoventral regulatory gene cassette spätzle/Toll/cactus controls the potent antifungal response in *Drosophila* adults. Cell 86:973–983.

Levine MT, Eckert ML, Begun DJ. 2011. Whole-genome expression plasticity across tropical and temperate *Drosophila melanogaster* populations from eastern Australia. Mol. Biol. Evol. 28:249–256.

Lewontin RC, Hubby JL. 1966. A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. Genetics 54:595–609.

Li C, Vagin V V., Lee S, Xu J, Ma S, Xi H, Seitz H, Horwich MD, Syrzycka M, Honda BM, et al. 2009. Collapse of Germline piRNAs in the Absence of *Argonaute3* Reveals Somatic piRNAs in Flies. Cell 137:509–521.

Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv Prepr. arXiv 0:3.

Li H, Stephan W. 2006. Inferring the Demographic History and Rate of Adaptive Substitution in *Drosophila*. PLoS Genet. 2:e166.

Li X, Schuler MA, Berenbaum MR. 2007. Molecular Mechanisms of Metabolic Resistance to Synthetic and Natural Xenobiotics. Annu. Rev. Entomol. 52:231–253.

Liao GC, Rehm EJ, Rubin GM. 2000. Insertion site preferences of the P transposable element in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. U. S. A. 97:3347–3351.

Lisch D. 2012. Regulation of transposable elements in maize. Curr. Opin. Plant Biol. 15:511–516.

Lowe WH, Kovach RP, Allendorf FW. 2017. Population Genetics and Demography Unite Ecology and Evolution. Trends Ecol. Evol. 32(2): 141-152.

Macdonald SS, Rako L, Batterham P, Hoffmann AA. 2004. Dissecting chill coma recovery as a measure of cold resistance: Evidence for a biphasic response in *Drosophila melanogaster*. J. Insect Physiol. 50:695–700.

Machado HE, Bergland AO, O'Brien KR, Behrman EL, Schmidt PS, Petrov DA. 2016. Comparative population genomics of latitudinal variation in *Drosophila simulans* and *Drosophila melanogaster*. Mol. Ecol. 25:723–740.

Mackay TFC. 2010. Mutations and quantitative genetic variation: lessons from *Drosophila*. Philos. Trans. R. Soc. B-Biological Sci. 365:1229–1239.

Mackay TFC, Richards S, Stone EA, Barbadilla A, Ayroles JF, Zhu D, Casillas S, Han Y, Magwire MM, Cridland JM, et al. 2012. The *Drosophila melanogaster* Genetic Reference Panel. Nature 482:173–178.

MacMillan HA, Knee JM, Dennis AB, Udaka H, Marshall KE, Merritt TJS, Sinclair BJ. 2016. Cold acclimation wholly reorganizes the *Drosophila melanogaster* transcriptome and metabolome. Sci. Rep. 6:28999.

Magwire MM, Bayer F, Webster CL, Cao C, Jiggins FM. 2011. Successive increases in the resistance of *Drosophila* to viral infection through a transposon insertion followed by a duplication. PLoS Genet. 7(10):e1002337.

Maitra S, Dombrowski SM, Basu M, Raustol O, Waters LC, Ganguly R. 2000. Factors on the third

chromosome affect the level of *Cyp6a2* and *Cyp6a8* expression in *Drosophila melanogaster*. Gene 248:147–156.

Mamillapalli A, Pathak RU, Garapati HS, Mishra RK. 2013. Transposable element "*roo*" attaches to nuclear matrix of the *Drosophila melanogaster*. J Insect Sci 13:111.

Manh Le H, Guio L, Merenciano M, Rovira Q, Barrón MG, González J. 2017. Natural and laboratory mutations in *kuzbanian* are associated with zinc stress phenotypes in *Drosophila melanogaster*. Nat. Publ. Gr. 7:42663.

Marshall KE, Sinclair BJ. 2010. Repeated stress exposure results in a survival-reproduction trade-off in *Drosophila melanogaster*. Proc. R. Soc. B Biol. Sci. 277:963–969.

Martinek S, Inonog S, Manoukian AS, Young MW. 2001. A role for the segment polarity gene *shaggy/GSK-3* in the *Drosophila* circadian clock. Cell 105:769–779.

Martins NE, Faria VG, Teixeira L, Magalhães S, Sucena É. 2013. Host Adaptation Is Contingent upon the Infection Route Taken by Pathogens. PLoS Pathog. 9(9):e1003601.

Mason JM, Biessmann H. 1995. The unusual telomeres of *Drosophila*. Trends Genet. 11:58–62.

Mateo L, González J. 2014. Pogo-Like transposases have been repeatedly domesticated into CENP-B-Related Proteins. Genome Biol. Evol. 6:2008–2016.

Mateo L, Ullastres A, González J. 2014. A Transposable Element Insertion Confers Xenobiotic Resistance in *Drosophila*. PLoS Genet. 10(8):e1004560.

Matthews KA, Kaufman TC, Gelbart WM. 2005. Research resources for *Drosophila*: the expanding universe. Nat. Rev. Genet. 6:179–193.

Matyash A, Singh N, Hanes SD, Urlaub H, Jäckle H. 2009. *SAP18* promotes *Krüppel*-dependent transcriptional repression by enhancer-specific histone deacetylation. J. Biol. Chem. 284:3012–3020.

Matzkin LM, Merritt TJS, Zhu C-T, Eanes WF. 2005. The structure and population genetics of the breakpoints associated with the cosmopolitan chromosomal inversion In(3R)Payne in *Drosophila melanogaster*. Genetics 170:1143–1152.

McClintock B. 1951. Chromosome organization and genic expression. Cold Spring Harb. Symp. Quant. Biol. 16:13–47.

McClintock B. 1956. Controlling elements and the gene. Cold Spring Harb. Symp. Quant. Biol. 21:197–216.

McGee LW, Aitchison EW, Caudle SB, Morrison AJ, Zheng L, Yang W, Rokyta DR. 2014. Payoffs, Not Tradeoffs, in the Adaptation of a Virus to Ostensibly Conflicting Selective Pressures.Worobeg M, editor. PLoS Genet. 10:e1004611.

Mensch J, Lavagnino N, Carreira VP, Massaldi A, Hasson E, Fanara JJ. 2008. Identifying candidate genes affecting developmental time in *Drosophila melanogaster*: pervasive pleiotropy and gene-by-environment interaction. BMC Dev. Biol. 8:78.

Merenciano M, Ullastres A, de Cara MAR, Barrón MG, González J. 2016. Multiple Independent Retroelement Insertions in the Promoter of a Stress Response Gene Have Variable Molecular and Functional Effects in *Drosophila*. PLoS Genet. 12(8):e1006249.

Messer PW, Petrov DA. 2013. Population genomics of rapid adaptation by soft selective sweeps. Trends Ecol. Evol. 28:659–669.

Metaxakis A, Oehler S, Klinakis A, Savakis C. 2005. Minos as a genetic and genomic tool in *Drosophila melanogaster*. Genetics 171:571–581.

Milchanowski AB, Henkenius AL, Narayanan M, Hartenstein V, Banerjee U. 2004. Identification and characterization of genes involved in embryonic crystal cell formation during *Drosophila* hematopoiesis. Genetics 168:325–339.

Miller WJ, McDonald JF, Pinsker W. 1997. Molecular domestication of mobile elements. Genetica 100:261–270.

Mohr SE, Hu Y, Kim K, Housden BE, Perrimon N. 2014. Resources for functional genomics studies in *Drosophila melanogaster*. Genetics 197:1–18.

Morgan TH. 1911. The origin of nine wing mutations in *Drosophila*. Science. 33:496–499.

Myllymäki H, Rämet M. 2014. JAK/STAT Pathway in *Drosophila* Immunity. Scand. J. Immunol. 79:377–385.

Nakai Y, Horiuchi J, Tsuda M, Takeo S, Akahori S, Matsuo T, Kume K, Aigaki T. 2011. Calcineurin and its regulator *sra/DSCR1* are essential for sleep in *Drosophila*. J. Neurosci. 31:12759–12766.

Nei M. 1987. Molecular evolutionary genetics. Columbia University Press

Nei M, Gojobori T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3:418–426.

Neyen C, Bretscher AJ, Binggeli O, Lemaitre B. 2014. Methods to study *Drosophila* immunity. Methods 68:116–128.

Nielsen R, Williamson S, Kim Y, Hubisz MJ, Clark AG, Bustamante C. 2005. Genomic scans for selective sweeps using SNP data. Genome Res. 15:1566–1575.

Nunney L. 1990. *Drosophila* on oranges: colonization, competition, and coexistence. Ecology 71:1904–1915.

Oakeshott JG, Chambers GK, Gibson JB, Eanes WF, Willcocks DA. 1983. Geographic variation in *G6pd* and *Pgd* allele frequencies in *Drosophila melanogaster.* Heredity (Edinb). 50:67–72.

Oakeshott JG, Gibson JB, Willcocks DA, Chambers GK. 1983. Latitudinal variation in octanol dehydrogenase and acid phosphatase allele frequencies in *Drosophila melanogaster.* Theor. Appl. Genet. 65:191–196.

Obbard DJ, Jiggins FM, Bradshaw NJ, Little TJ. 2011. Recent and recurrent selective sweeps of the antiviral RNAi gene *argonaute-2* in three species of *Drosophila*. Mol. Biol. Evol. 28:1043–1056.

Obbard DJ, Jiggins FM, Halligan DL, Little TJ. 2006. Natural selection drives extremely rapid evolution in antiviral RNAi genes. Curr. Biol. 16:580–585.

Obbard DJ, Welch JJ, Kim KW, Jiggins FM. 2009. Quantifying adaptive evolution in the *Drosophila* immune system. PLoS Genet. 205(1):353-366.

Ohta T. 1973. Slightly deleterious mutant substitutions in evolution. Nature 246:96–98.

Olson-Manning CF, Wagner MR, Mitchell-Olds T. 2012. Adaptive evolution: evaluating empirical support for theoretical predictions. Nat. Rev. Genet. 13:867–877.

Orgel LE, Crick FH. 1980. Selfish DNA: the ultimate parasite. Nature 284:604–607.

Orr HA. 2005. The genetic theory of adaptation: a brief history. Nat. Rev. Genet. 6:119–127.

Orr HA. 2009. Fitness and its role in evolutionary genetics. Nat. Rev. Genet. 10:531–539.

Paaby AB, Bergland AO, Behrman EL, Schmidt PS. 2014. A highly pleiotropic amino acid polymorphism in the *Drosophila* insulin receptor contributes to life-history adaptation. Evolution (N. Y). 68:3395–3409.

Paaby AB, Blacket MJ, Hoffmann AA, Schmidt PS. 2010. Identification of a candidate adaptive polymorphism for *Drosophila* life history by parallel independent clines on two continents. Mol. Ecol. 19:760–774.

Paaby AB, Schmidt PS. 2008. Functional significance of allelic variation at *methuselah*, an aging gene in *Drosophila*. PLoS One 3.

Paparazzo F, Tellier A, Stephan W, Hutter S. 2015. Survival rate and transcriptional response upon infection with the generalist parasite *Beauveria bassiana* in a world-wide sample of *Drosophila melanogaster*. PLoS One 10:1–20.

Pardo-Diaz C, Salazar C, Jiggins CD. 2015. Towards the identification of the loci of adaptive evolution. Methods Ecol. Evol. 6:445–464.

Partridge L, Barrie B, Fowler K, French V. 1994. Evolution and Development of Body Size and Cell Size

in *Drosophila melanogaster* in Response to Temperature. Evolution (N. Y). 48:1269–1276.

Partridge L, Prowse N, Pignatelli P. 1999. Another set of responses and correlated responses to selection on age at reproduction in *Drosophila melanogaster.* Proc. R. Soc. B Biol. Sci. 266:255–261.

Peel MC, Finlayson BL, McMahon TA. 2007. Updated world map of the Köppen-Geiger climate classification. Hydrol. Earth Syst. Sci. 11:1633–1644.

Pegoraro M, Zonato V, Tyler ER, Fedele G, Kyriacou CP, Tauber E. 2017. Geographical analysis of diapause inducibility in European *Drosophila melanogaster* populations. J. Insect Physiol. 98:238–244.

Pfeifer B, Wittelsbürger U, Ramos-Onsins SE, Lercher MJ. 2014. PopGenome: An efficient swiss army knife for population genomic analyses in R. Mol. Biol. Evol. 31:1929–1936.

Poltorak A. 1998. Defective LPS Signaling in *C3H/HeJ* and *C57BL/10ScCr* Mice: Mutations in *Tlr4* Gene. Science. 282:2085–2088.

Pool JE, Corbett-Detig RB, Sugino RP, Stevens K a, Cardeno CM, Crepeau MW, Duchen P, Emerson JJ, Saelao P, Begun DJ, et al. 2012. Population Genomics of sub-saharan *Drosophila melanogaster*: African diversity and non-African admixture. PLoS Genet. 8:e1003080.

Pritchard JK, Pickrell JK, Coop G. 2010. The Genetics of Human Adaptation: Hard Sweeps, Soft Sweeps, and Polygenic Adaptation. Curr. Biol. 20(4):R208-15.

Prud'homme B, Gompel N, Carroll SB. 2007. Colloquium Papers: Emerging principles of regulatory evolution. Proc. Natl. Acad. Sci. USA 104 Suppl:8605–8612.

Przeworski M. 2002. The signature of positive selection at randomly chosen loci. Genetics 160:1179–1189.

Puig M, Castellano D, Pantano L, Giner-Delgado C, Izquierdo D, Gayà-Vidal M, Lucas-Lledó JI, Esko T, Terao C, Matsuda F, et al. 2015. Functional Impact and Evolution of a Novel Human Polymorphic Inversion That Disrupts a Gene and Creates a Fusion Transcript. PLoS Genet. 11:e1005495.

Rahman R, Chirn G, Kanodia A, Sytnikova YA, Brembs B, Bergman CM, Lau NC. 2015. Unique transposon landscapes are pervasive across *Drosophila melanogaster* genomes. Nucleic Acids Res. 43:10655–10672.

Rämet M, Lanot R, Zachary D, Manfruelli P. 2002. JNK Signaling Pathway Is Required for Efficient Wound Healing in *Drosophila*. Dev. Biol. 241:145–156.

Ràmia M, Librado P, Casillas S, Rozas J, Barbadilla A. 2012. PopDrowser: The population *Drosophila* browser. Bioinformatics 28:595–596.

Rebollo R, Romanish MT, Mager DL. 2012. Transposable Elements: An Abundant and Natural Source of Regulatory Sequences for Host Genes. Annu. Rev. Genet. 46:21–42.

Reinhardt JA, Kolaczkowski B, Jones CD, Begun DJ, Kern AD. 2014. Parallel geographic variation in *Drosophila melanogaster*. Genetics 197:361–373.

Rockman M V. 2012. The QTN program and the alleles that matter for evolution: All that's gold does not glitter. Evolution (N. Y). 66:1–17.

Rose MR, Mueller LD, Burke MK. 2011. New experiments for an undivided genetics. Genetics 188:1–10.

Rouault JD, Marican C, Wicker-Thomas C, Jallon JM. 2004. Relations between cuticular hydrocarbon (HC) polymorphism, resistance against desiccation and breeding temperature; a model for HC evolution in *D. melanogaster* and *D. simulans.* Genetica 120:195–212.

Roxström-Lindquist K, Terenius O, Faye I. 2004. Parasite-specific immune response in adult *Drosophila melanogaster*: a genomic study. EMBO Rep. 5:207–212.

Sackton TB, Kulathinal RJ, Bergman CM, Quinlan AR, Dopman EB, Carneiro M, Marth GT, Hartl DL, Clark AG. 2009. Population genomic inferences from sparse high-throughput sequencing of two populations of *Drosophila melanogaster.* Genome Biol. Evol. 1:449–465.

Sackton TB, Lazzaro BP, Clark AG. 2010. Genotype and gene expression associations with immune function in *Drosophila*. PLoS Genet. 6(1):e1000797.

Sackton TB, Lazzaro BP, Schlenke TA, Evans JD, Hultmark D, Clark AG. 2007. Dynamic evolution of the innate immune system in *Drosophila*. Nat. Genet. 39:1461–1468.

Samuelson LC, Wiebauer K, Snow CM, Meisler MH. 1990. Retroviral and pseudogene insertion sites reveal the lineage of human salivary and pancreatic amylase genes from a single gene during primate evolution. Mol. Cell. Biol. 10:2513–2520.

Sarkar A, Sim C, Hong YS, Hogan JR, Fraser MJ, Robertson HM, Collins FH. 2003. Molecular evolutionary analysis of the widespread piggyBac transposon family and related "domesticated" sequences. Mol. Genet. Genomics 270:173–180.

Satyaki PR V., Cuykendall TN, Wei KH-C, Brideau NJ, Kwak H, Aruna S, Ferree PM, Ji S, Barbash DA. 2014. The Hmr and Lhr Hybrid Incompatibility Genes Suppress a Broad Range of Heterochromatic Repeats. PLoS Genet. 10:e1004240.

Schmidt JM, Good RT, Appleton B, Sherrard J, Raymant GC, Bogwitz MR, Martin J, Daborn PJ, Goddard ME, Batterham P, et al. 2010. Copy number variation and transposable elements feature in recent, ongoing adaptation at the *Cyp6g1* locus. PLoS Genet. 6:1–11.

Schmidt PS, Matzkin L, Ippolito M, Eanes WF. 2005. Geographic variation in diapause incidence, life-history traits, and climatic adaptation in *Drosophila melanogaster*. Evolution 59:1721–1732.

Schmidt PS, Paaby AB. 2008. Reproductive diapause and life-history clines in North American populations of *Drosophila melanogaster.* Evolution (N. Y). 62:1204–1215.

Schmidt PS, Paaby AB, Heschel MS. 2005. Genetic variance for diapause expression and associated life histories in *Drosophila melanogaster.* Evolution (N. Y). 59:2616–2625.

Schmidt PS, Zhu C-T, Das J, Batavia M, Yang L, Eanes WF. 2008. An amino acid polymorphism in the *couch potato* gene forms the basis for climatic adaptation in *Drosophila melanogaster.* Proc. Natl. Acad. Sci. U. S. A. 105:16207–16211.

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, et al. 2009. The B73 Maize Genome: Complexity, Diversity, and Dynamics. Science. 326:1112–1115.

Schou MF. 2013. Fast egg collection method greatly improves randomness of egg sampling in *Drosophila melanogaster.* Fly (Austin). 7:44–46.

Sentmanat M, Wang SH, Elgin SCR. 2013. Targeting heterochromatin formation to transposable elements in *Drosophila*: potential roles of the piRNA system. Biochem. 78:562–571.

Sentmanat MF, Elgin SC. 2012. Ectopic assembly of heterochromatin in *Drosophila melanogaster* triggered by transposable elements. Proc. Natl. Acad. Sci. U. S. A. 109:14104–14109.

Sessegolo C, Burlet N, Haudry A, Biémont C, Vieira C, Tenaillon M, Hollister J, Gaut B, McClintock B, Mackay T, et al. 2016. Strong phylogenetic inertia on genome size and transposable element content among 26 species of flies. Biol. Lett. 12:521–524.

Sezgin E, Duvernell DD, Matzkin LM, Duan Y, Zhu CT, Verrelli BC, Eanes WF. 2004. Single-locus latitudinal clines and their relationship to temperate adaptation in metabolic genes and derived alleles in *Drosophila melanogaster.* Genetics 168:923–931.

Sgrò CM, Partridge L. 1999. A delayed wave of death from reproduction in *Drosophila*. Science 286:2521–2524.

Short SM, Lazzaro BP. 2013. Reproductive status alters transcriptomic response to infection in female *Drosophila melanogaster.* G3 3:827–840.

Singh ND, Petrov DA. 2004. Rapid Sequence Turnover at an Intergenic Locus in *Drosophila*. Mol. Biol. Evol. 21:670–680.

Slatkin M. 2008. Linkage disequilibrium--understanding the evolutionary past and mapping the medical future. Nat. Rev. Genet. 9:477–485.

Smith JM, Haigh J. 1974. The hitch-hiking effect of a favourable gene. Genet. Res. 23:23–35.

Sokal RR, Rohlf FJ. 2012. Biometry. The Principles and Practice of Statistics in Biological Research. 4th ed. New York: W.H. Freeman

Sonoshita M, Cagan RL. 2016. Modeling Human Cancers in *Drosophila*. Curr. Top. Dev. Biol. 121, 287-309.

Sorek R, Ast G, Graur D. 2002. Alu-containing exons are alternatively spliced. Genome Res. 12:1060–1067.

Sørensen JG, Nielsen MM, Loeschcke V. 2007. Gene expression profile analysis of *Drosophila melanogaster* selected for resistance to environmental stressors. J. Evol. Biol. 20:1624–1636.

Spitz F, Furlong EEM. 2012. Transcription factors: From enhancer binding to developmental control. Nat. Rev. Genet. 13:613–626.

Spradling AC, Stern D, Beaton A, Rhem EJ, Laverty T, Mozden N, Misra S, Rubin GM. 1999. The Berkeley *Drosophila* Genome Project Gene Disruption Project: Single. Screen.

St Johnston D. 2013. Using mutants, knockdowns, and transgenesis to investigate gene function in Drosophila. Wiley Interdiscip. Rev. Dev. Biol. 2:587–613.

Stanley S, Parsons P. 1981. The response of the cosmopolitan species, *Drosophila melanogaster,* to ecological gradients. Proc. Ecol. Soc. Aust. 11:121–130.

Stapley J, Reger J, Feulner PGD, Smadja C, Galindo J, Ekblom R, Bennison C, Ball AD, Beckerman AP, Slate J. 2010. Adaptation genomics: the next generation. Trends Ecol. Evol. 25:705–712.

Steiner CC, Weber JN, Hoekstra HE. 2007. Adaptive variation in beach mice produced by two interacting pigmentation genes. PLoS Biol. 5:1880–1889.

Stephan W. 2016. Signatures of positive selection: From selective sweeps at individual loci to subtle allele frequency changes in polygenic adaptation. Mol. Ecol. 25:79–88.

Stern DL, Orgogozo V. 2008. The loci of evolution: How predictable is genetic evolution? Evolution (N. Y). 62:2155–2177.

Storz JF, Wheat CW. 2010. Integrating evolutionary and functional approaches to infer adaptation at specific *loci*. Evolution (N. Y). 64:2489–2509.

Strobel E, Dunsmuir P, Rubin GM. 1979. Polymorphisms in the chromosomal locations of elements of the *412*, *copia* and *297* dispersed repeated gene families in *Drosophila*. Cell 17:429–439.

Suh Y, Vijg J. 2005. SNP discovery in associating genetic variation with human disease phenotypes. Mutat. Res. 573:41–53.

Tabor HK, Risch NJ, Myers RM. 2002. Candidate-gene approaches for studying complex genetic traits: Practical considerations. Nat. Rev. Genet. 3:391–397.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595.

Takeo S, Hawley RS, Aigaki T. 2010. Calcineurin and its regulation by *Sra/RCAN* is required for completion of meiosis in *Drosophila*. Dev. Biol. 344:957–967.

Takeo S, Tsuda M, Akahori S, Matsuo T, Aigaki T. 2006. The Calcineurin Regulator *Sra* Plays an Essential Role in Female Meiosis in *Drosophila*. Curr. Biol. 16:1435–1440.

Takeuchi K-I, Nakano Y, Kato U, Kaneda M, Aizu M, Awano W, Yonemura S, Kiyonaka S, Mori Y, Yamamoto D, et al. 2009. Changes in temperature preferences and energy homeostasis in dystroglycan mutants. Science 323:1740–1743.

Teets NM, Yi S-X, Lee RE, Denlinger DL. 2013. Calcium signaling mediates cold sensing in insect tissues. Proc. Natl. Acad. Sci. U. S. A. 110:9154–9159.

Teixeira L. 2012. Whole-genome expression profile analysis of *Drosophila melanogaster* immune responses. Brief. Funct. Genomics 11:375–386.

Telonis-Scott M, Hoffmann AA, Sgrò CM. 2011. The molecular genetics of clinal variation: A case study of ebony and thoracic trident pigmentation in *Drosophila melanogaster* from eastern Australia. Mol.

Ecol. 20:2100–2110.

Telonis-Scott M, Sgró CM, Hoffmann AA, Griffin PC. 2016. Cross-Study Comparison reveals common genomic, network, and functional signatures of desiccation resistance in *Drosophila melanogaster*. Mol. Biol. Evol. 33:1053–1067.

Thibault ST, Singer MA, Miyazaki WY, Milash B, Dompe NA, Singh CM, Buchholz R, Demsky M, Fawcett R, Francis-Lang HL, et al. 2004. A complementary transposon tool kit for *Drosophila melanogaster* using P and piggyBac. Nat. Genet. 36:283–287.

Thornton K, Andolfatto P. 2006. Approximate Bayesian inference reveals evidence for a recent, severe bottleneck in a Netherlands population of *Drosophila melanogaster*. Genetics 172:1607–1619.

Tinsley MC, Blanford S, Jiggins FM. 2006. Genetic variation in *Drosophila melanogaster* pathogen susceptibility. Parasitology 132:767–773.

Trotta V, Calboli FCF, Ziosi M, Guerra D, Pezzoli MC, David JR, Cavicchi S. 2006. Thermal plasticity in *Drosophila melanogaster*: a comparison of geographic populations. BMC Evol. Biol. 6:67.

Tsuchiya T, Eulgem T. 2013. An alternative polyadenylation mechanism coopted to the Arabidopsis RPP7 gene through intronic retrotransposon domestication. Proc. Natl. Acad. Sci. U. S. A. 110:E3535-43.

Tubio JMC, Li Y, Ju YS, Martincorena I, Cooke SL, Tojo M, Gundem G, Pipinikas CP, Zamora J, Raine K, et al. 2014. Mobile DNA in cancer. Extensive transduction of nonrepetitive DNA mediated by L1 retrotransposition in cancer genomes. Science 345:1251343.

Turner TL, Levine MT, Eckert ML, Begun DJ. 2008. Genomic analysis of adaptive differentiation in *Drosophila melanogaster*. Genetics 179:455–473.

Ullastres A, Merenciano M, Guio L, Gonz J. 2016. Transposable elements: a toolkit for stress and environmental adaptation in bacteria. In: de Bruijn FJ, editor. Stress and Environmental Regulation of Gene Expression and Adaptation in Bacteria. First edit. JohnWiley & Sons, Inc. p. 137–145.

Ullastres A, Petit N, González J. 2015. Exploring the phenotypic space and the evolutionary history of a natural mutation in *Drosophila melanogaster*. Mol. Biol. Evol. 32:1800–1814.

Unckless RL, Lazzaro BP. 2016. The potential for adaptive maintenance of diversity in insect antimicrobial peptides. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 371:20150291-.

Upton KR, Gerhardt DJ, Jesuadian JS, Richardson SR, Sánchez-Luque FJ, Bodea GO, Ewing AD, Salvador-Palomeque C, van der Knaap MS, Brennan PM, et al. 2015. Ubiquitous L1 Mosaicism in Hippocampal Neurons. Cell 161:228–239.

Valanne S. 2014. Functional genomic analysis of the *Drosophila* immune response. Dev. Comp. Immunol. 42:93–101.

Vallet-Gely I, Novikov A, Augusto L, Liehl P, Bolbach G, Péchy-Tarr M, Cosson P, Keel C, Caroff M, Lemaitre B. 2010. Association of hemolytic activity of *Pseudomonas entomophila*, a versatile soil bacterium, with cyclic lipopeptide production. Appl. Environ. Microbiol. 76:910–921.

van't Hof AE, Campagne P, Rigden DJ, Yung CJ, Lingley J, Quail MA, Hall N, Darby AC, Saccheri IJ. 2016. The industrial melanism mutation in British peppered moths is a transposable element. Nature 534:102–105.

Venken KJT, Schulze KL, Haelterman NA, Pan H, He Y, Evans-Holm M, Carlson JW, Levis RW, Spradling AC, Hoskins RA, et al. 2011. MiMIC: a highly versatile transposon insertion resource for engineering *Drosophila melanogaster* genes. Nat. Methods 8:737–743.

Vermeulen CJ, Sørensen P, Kirilova Gagalova K, Loeschcke V. 2013. Transcriptomic analysis of inbreeding depression in cold-sensitive *Drosophila melanogaster* shows upregulation of the immune response. J. Evol. Biol. 26:1890–1902.

Villanueva-Cañas JL, Rech G, de Cara MAR, González J. 2017. Beyond SNPs: how to detect selection on transposable element insertions. Methods Ecol. Evol. (in press).

Vodovar N, Vinals M, Liehl P, Basset A, Degrouard J, Spellman P, Boccard F, Lemaitre B. 2005. *Drosophila* host defense after oral infection by an entomopathogenic *Pseudomonas* species. Proc. Natl. Acad. Sci. U. S. A. 102:11414–11419.

Vogt J, Bengesser K, Claes KBM, Wimmer K, Mautner V-F, van Minkelen R, Legius E, Brems H, Upadhyaya M, Högel J, et al. 2014. SVA retrotransposon insertion-associated deletion represents a novel mutational mechanism underlying large genomic copy number changes with non-recurrent breakpoints. Genome Biol. 15:R80.

Voight BF, Kudaravalli S, Wen X, Pritchard JK. 2006. A map of recent positive selection in the human genome. PLoS Biol. 4:e72.

Wang L, Rishishwar L, Mariño-Ramírez L, Jordan IK. 2016. Human population-specific gene expression and transcriptional network modification with polymorphic transposable elements. Nucleic Acids Res. 45(5): 2318-2328.

Warren IA, Naville M, Chalopin D, Levin P, Berger CS, Galiana D, Volff JN. 2015. Evolutionary impact of transposable elements on genomic diversity and lineage-specific innovation in vertebrates. Chromosom. Res. 23:505–531.

Weber AL, Khan GF, Magwire MM, Tabor CL, Mackay TFC, Anholt RRH. 2012. Genome-Wide Association Analysis of Oxidative Stress Resistance in *Drosophila melanogaster*. Plos ONE. 7(6):e38722.

Weinig C, Stinchcombe JR, Schmitt J. 2003. QTL architecture of resistance and tolerance traits in Arabidopsis thaliana in natural environments. Mol. Ecol. 12:1153–1163.

Wertheim B. 2015. Genomic basis of evolutionary change: Evolving immunity. Front. Genet. 6:222.

Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, et al. 2007. A unified classification system for eukaryotic transposable elements. Nat. Rev. Genet. 8:973–982.

Williams GC. 1957. Pleiotropy, Natural Selection, and the Evolution of Senescence. Evolution (N. Y). 11:398.

Wittkopp PJ. 2011. Molecular Methods for Evolutionary Genetics. Mol. Methods Evol. Genet. 772:297–317.

Wittkopp PJ, Haerum BK, Clark AG. 2004. Evolutionary changes in cis and trans gene regulation. Nature 430:85–88.

Wright S. 1931. Evolution in Mendelian populations. Genetics. 16(2):97-159.

Xing J, Wang H, Belancio VP, Cordaux R, Deininger PL, Batzer M a. 2006. Emergence of primate genes by retrotransposon-mediated sequence transduction. Proc. Natl. Acad. Sci. U. S. A. 103:17608–17613.

Yadav P, Thandapani M, Sharma VK. 2014. Interaction of light regimes and circadian clocks modulate timing of pre-adult developmental events in *Drosophila*. BMC Dev. Biol. 14:19.

Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. 2012. Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics 13:134.

Yukilevich R, Turner TL, Aoki F, Nuzhdin S V, True JR. 2010. Patterns and processes of genome-wide divergence between North American and African *Drosophila melanogaster*. Genetics 186:219–239.

Zhang J, Marshall KE, Westwood JT, Clark MS, Sinclair BJ. 2011. Divergent transcriptomic responses to repeated and single cold exposures in *Drosophila melanogaster*. J. Exp. Biol. 214:4021–4029.

Zhao L, Wit J, Svetec N, Begun DJ. 2015. Parallel Gene Expression Differences between Low and High Latitude Populations of *Drosophila melanogaster* and *D. simulans*. PLoS Genet. 11:1–25.

Zhao X, Bergland AO, Behrman EL, Gregory BD, Petrov DA, Schmidt PS. 2016. Global transcriptional profiling of diapause and climatic adaptation in *Drosophila melanogaster*. Mol. Biol. Evol. 33:707–720.

Zhu CT, Chang C, Reenan RA, Helfand SL. 2014. *Indy* gene variation in natural populations confers fitness advantage and life span extension through transposon insertion. Aging (Albany. NY). 6:58–69.

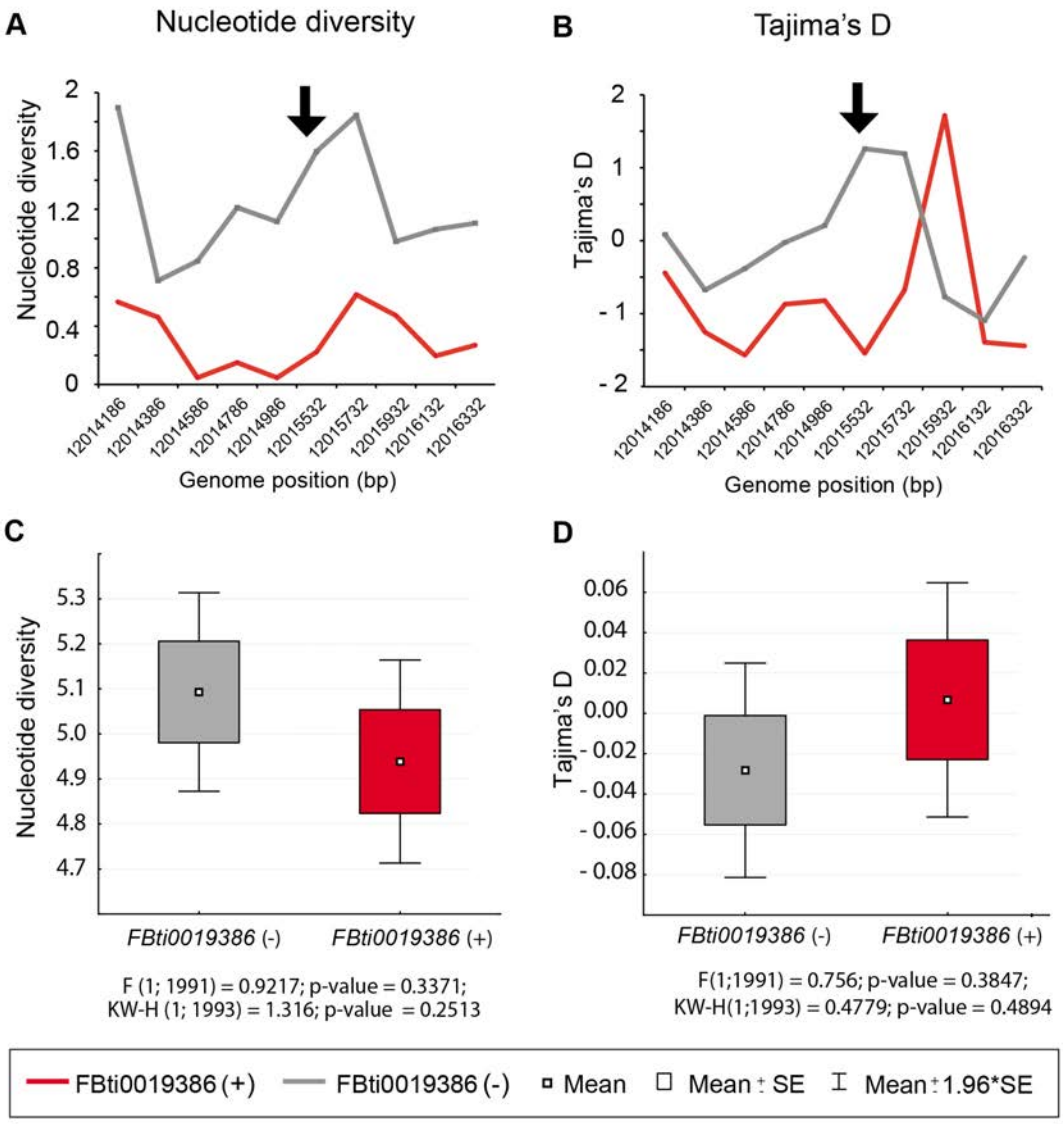
Zhu J, Fu Y, Nettleton M, Richman A, Han Z. 2017. High throughput in vivo functional validation of candidate congenital heart disease genes in *Drosophila*. Elife 6:e22617.

Zonato V, Fedele G, Kyriacou CP. 2016. An intronic polymorphism in *couch potato* is not distributed clinally in European *Drosophila melanogaster* populations nor does it affect diapause inducibility. PLoS One 11:1–14.

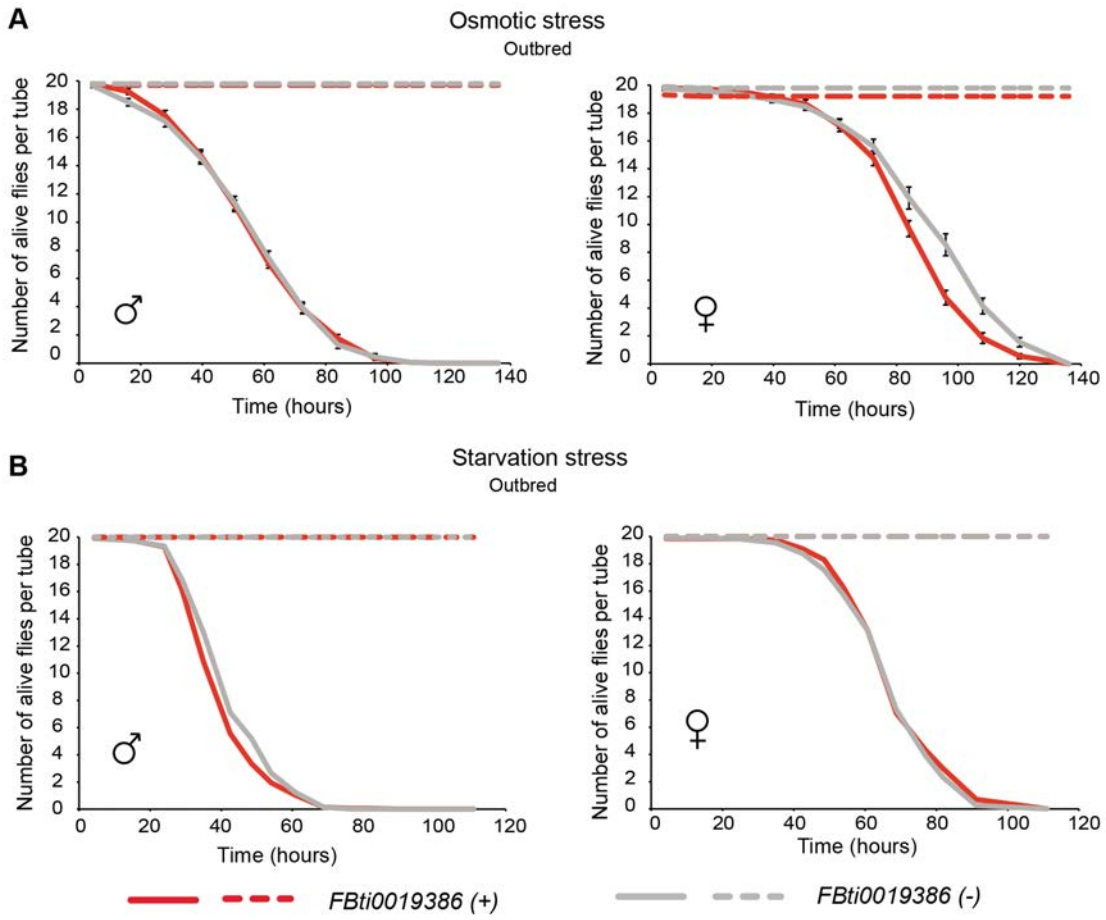
08

ANNEXES

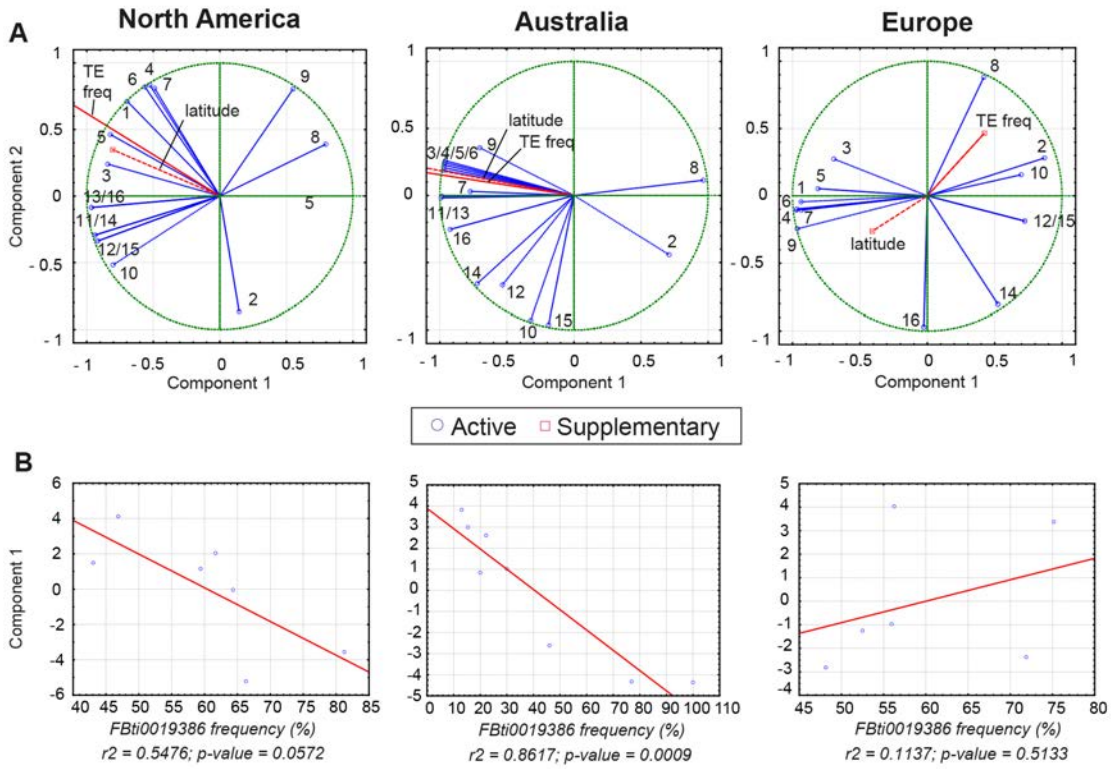
Annex Figure S1: Nucleotide diversity (A) and Tajima´s D (B) in the 2kb region around *FBti0019386* insertion. The arrow indicates the location of the TE. Box plot representation of nucleotide diversity (C) and Tajima´s D (D) estimated for 1,000 random 1kb regions.



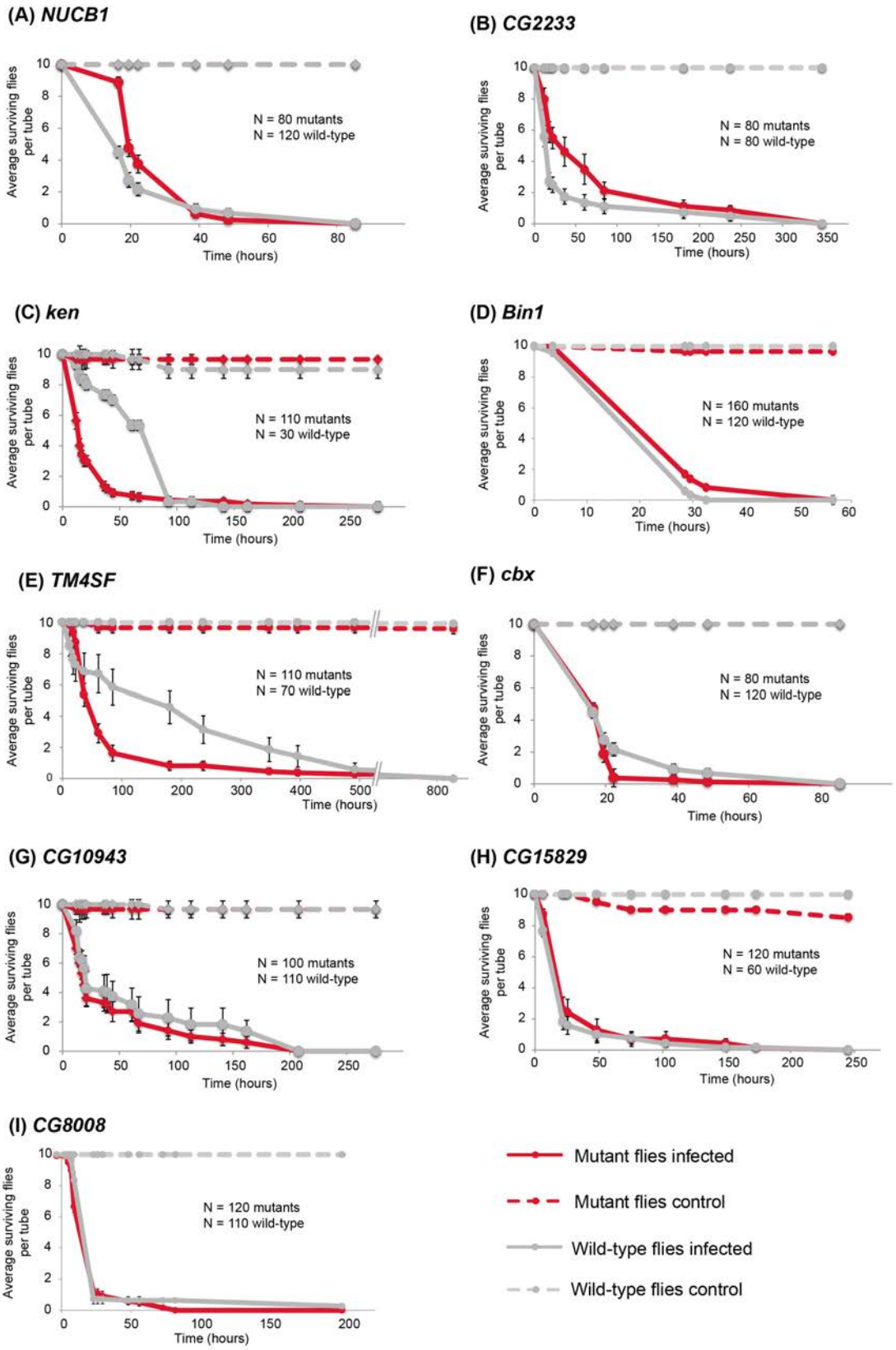
Annex Figure S2. Flies with *FBti0019386* are more sensitive to osmotic and starvation stress. (A) Females from outbred populations with the *FBti0019386* insertion (red) showed more mortality than females without *FBti0019386* insertion (gray). (B) Males with the *FBti0019386* insertion died more than males without the insertion. Survival under control conditions is represented as dashed lines.



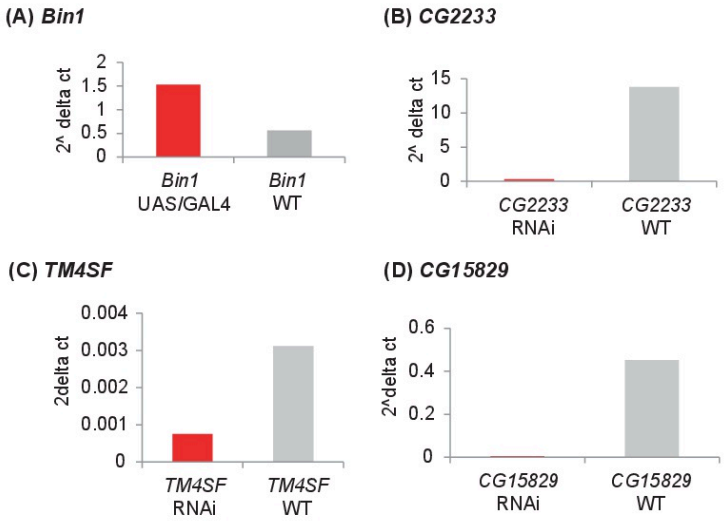
Annex Figure S3. Graphical representation of the Principal Component Analyses (PCA). (A) Grouping of the climatic variables (blue) in the PCA with latitude and TE frequency projected on the PCA. Variable are as follows: 1 AvMonTemp, 2 thermalAmp, 3 HotMont h, 4 ColdMonth, 5 summerSEASON, 6 winterSEASON, 7 monthabove10, 8 MAP, 9 Cv, 10 DryMonth, 11 summer_P,12 summer_DryM, 13 summer_wetM, 14 winter_P, 15 winter_DryM, 16 winter_wetM. (B) Correlation analyses between TE frequency and the first component of the PCA for the three continents.



Annex Figure S4. Mutant infection survival curves. N: total number of infected flies in each experiment. Because of the high sensitivity of *CG15829* mutant flies, this strain was infected with a total OD_{600} = 50. The rest of the mutant strains were infected with OD_{600}=100.



Annex Figure S5. Expression analysis for some of the mutant flies used for the infection experiments. Each bar represents the average ratio of gene expression relative to the housekeeping gene *Act5c*. Note that we only analyzed the expression of one biological replica.



Annex Table S1.1. Sliding windows analysis of the 2kb region flanking *FBti0019386* insertion (1kb on each side).

		Tajima's D		Nucleotide diversity	
Chr.start	**Chr.end**	***FBti0019386+***	***FBti0019386-***	***FBti0019386+***	***FBti0019386-***
12014186	12014385	-0.44	0.09	0.56	1.90
12014386	12014585	-1.25	-0.68	0.46	0.71
12014586	12014785	-1.57	-0.38	0.05	0.84
12014786	12014985	-0.87	-0.02	0.15	1.21
12014986	12015186	-0.82	0.21	0.05	1.12
12015532	12015731	-1.55	1.26	0.22	1.60
12015732	12015931	-0.67	1.19	0.62	1.85
12015932	12016131	1.72	-0.77	0.47	0.98
12016132	12016331	-1.39	-1.10	0.20	1.06
12016332	12016532	-1.44	-0.23	0.27	1.10

Annex Table S1.2. Results of simulations under a neutral model using MS program, with theta parameter=5 for flies with the element (*FBti0019386* +) and flies without the element (*FBti0019386* -) datasets in a region of 1 kb around the TE insertion.

				Neutral simulations					
		Observed	P-value	Valid N	Mean	Minimum	Maximum	Percentile 5	Percentile 95
	Tajima'sD	-1.7743102	0.0160	1000	-0.094790	-2.28808	2.783320	-1.43645	1.594586
FBti0019386 **+**	nucl.div	0.4332737	<0.001	1000	5.038808	0.773882	22.29469	1.708646	10.17030
	CL (log)	-5.95	<0.001	1000	-24.3104	-56.5700	-6.78000	-38.1700	-12.3400
	Tajima'sD	0.6775646	>0.05	1000	-0.096610	-2.14222	2.585194	-1.41546	1.590358
FBti0019386 **-**	nucl.div	4.5140351	>0.05	1000	4.953185	0.437544	16.16000	1.751053	10.19860
	CL (log)	-18.15	>0.05	1000	-19.6168	-43.7300	-3.51000	-32.7450	-8.95000

Annex Table S1.3. Statistics estimated in the 1000 datasets obtained by randomization of strains keeping the same proportion of the strains with and without the element.

dataset	CL	log(CL)	Tajima D	π
1	5.90E-12	-11.229313	0.8235081	3.797123
2	3.55E-15	-14.44944	-0.1670349	3.216816
3	1.94E-10	-9.7126312	0.21460617	2.785239
4	1.56E-13	-12.808249	0.27298718	3.352201
5	2.42E-08	-7.6157158	0.82448776	2.997203
6	1.07E-13	-12.972461	0.70823342	4.061495
7	8.45E-13	-12.073024	1.26670688	4.001695
8	1.56E-13	-12.808249	0.11532612	3.119774
9	5.76E-15	-14.239613	0.58744939	3.480829
10	8.45E-13	-12.073024	0.41540641	3.185876
11	1.43E-12	-11.84578	0.29438122	3.012121
12	5.41E-11	-10.266844	0.61799472	3.302458
13	3.55E-15	-14.44944	0.13584718	3.561608
14	2.66E-14	-13.574521	0.21921987	3.310996
15	2.66E-14	-13.574521	0.05058662	3.206061
16	4.13E-11	-10.384354	0.59262359	3.278314
17	5.71E-12	-11.24372	0.65259945	3.269474
18	3.85E-11	-10.414416	0.40631308	3.137401
19	1.83E-12	-11.738279	-0.4277619	2.732946
20	3.34E-17	-16.475968	0.06155533	3.601242
21	2.30E-14	-13.637553	0.64237548	3.634585
22	1.01E-12	-11.995585	0.71481474	3.988701
23	3.55E-15	-14.44944	0.22382834	3.686441
24	2.53E-15	-14.597012	0.6030543	4.023266
25	1.01E-14	-13.994952	0.26938517	3.604851
26	8.45E-13	-12.073024	0.47907708	3.246893
27	1.05E-12	-11.978945	0.60661451	3.576885
28	3.85E-11	-10.414416	0.85970978	3.5147
29	6.06E-09	-8.2177758	1.21843182	3.324942
30	3.55E-15	-14.44944	-0.0259769	3.332401
31	1.43E-12	-11.84578	0.82062707	3.553146
32	6.22E-13	-12.206189	0.67898357	3.629837
33	1.50E-15	-14.824256	0.24455025	3.685352
34	3.65E-10	-9.4375399	0.47367056	3.263158
35	4.09E-08	-7.388472	0.5119349	2.824026
36	2.14E-15	-14.669788	0.15736122	3.73705
37	3.79E-10	-9.4218958	0.36990539	2.604039
38	4.85E-11	-10.314691	1.00641645	3.402546
39	7.18E-12	-11.143995	0.83076601	3.299313
40	3.89E-14	-13.410309	0.39565522	3.342657
41	1.94E-10	-9.7126312	0.83120533	3.263403

42	4.57E-13	-12.340339	-0.0688167	3.039337
43	3.75E-16	-15.426316	0.2547472	3.629061
44	9.22E-14	-13.035493	0.92413619	3.878322
45	4.20E-12	-11.376885	0.36138798	3.327877
46	1.84E-09	-8.7357549	1.3499977	3.74504
47	1.58E-14	-13.801765	-0.195301	2.962302
48	1.15E-10	-9.9398751	0.44993913	2.875587
49	1.58E-14	-13.801765	0.084436	3.203313
50	4.85E-11	-10.314691	1.10396651	3.563842
51	7.18E-12	-11.143995	0.56291425	2.991304
52	1.94E-10	-9.7126312	0.85323799	3.266023
53	1.79E-12	-11.746055	1.18963955	3.565847
54	5.71E-12	-11.24372	0.43088248	3.180328
55	7.18E-12	-11.143995	0.79730766	3.288701
56	5.64E-17	-16.248724	-0.0795204	3.454785
57	1.54E-10	-9.8123561	0.44153628	3.232305
58	1.68E-11	-10.774825	0.13860677	3.143503
59	6.00E-15	-14.222196	0.47601221	4.000605
60	3.69E-13	-12.433433	0.69462182	3.645688
61	3.75E-16	-15.426316	0.35458249	3.835593
62	2.62E-13	-12.581005	-0.1273064	2.812587
63	2.26E-16	-15.646664	-0.1657208	3.310642
64	8.88E-16	-15.0515	-0.325746	3.08409
65	4.59E-10	-9.337815	0.57113254	3.048611
66	6.22E-13	-12.206189	0.20734012	3.213559
67	3.69E-13	-12.433433	0.77827037	3.79548
68	1.50E-15	-14.824256	0.15079809	3.511414
69	1.47E-12	-11.831373	0.24170571	3.206349
70	2.53E-15	-14.597012	0.17422824	3.581845
71	4.20E-12	-11.376885	0.56804048	3.558964
72	2.28E-11	-10.64166	1.02347141	3.688323
73	4.32E-08	-7.3645163	0.15564298	2.264972
74	6.00E-15	-14.222196	-0.0707918	3.281585
75	7.18E-12	-11.143995	0.35753312	2.840559
76	2.28E-11	-10.64166	0.01952139	2.750117
77	2.24E-10	-9.6491396	0.62208249	2.861449
78	4.20E-12	-11.376885	0.55893698	3.508159
79	3.38E-12	-11.470964	0.80273633	3.557062
80	3.55E-15	-14.44944	0.51307238	4.016384
81	3.74E-14	-13.426949	0.11223464	3.293155
82	1.94E-10	-9.7126312	0.99565894	3.376398
83	5.90E-12	-11.229313	0.21413921	3.121739
84	4.59E-10	-9.337815	1.00660914	3.437996
85	1.94E-10	-9.7126312	0.42160423	2.93284
86	2.28E-11	-10.64166	0.48063711	3.227922

87	3.75E-16	-15.426316	0.20782261	3.55528
88	5.64E-17	-16.248724	0.24214847	3.861072
89	2.87E-11	-10.541935	0.84739479	3.277855
90	2.28E-11	-10.64166	0.26138021	2.999008
91	2.14E-15	-14.669788	0.05936358	3.577856
92	2.11E-13	-12.675084	0.5223665	3.146316
93	8.45E-13	-12.073024	0.71791969	3.415851
94	3.85E-11	-10.414416	0.72111352	3.437996
95	1.94E-10	-9.7126312	0.51370991	3.015336
96	1.07E-13	-12.972461	0.19553048	3.303599
97	8.45E-13	-12.073024	0.29758708	2.964397
98	1.07E-13	-12.972461	0.06094805	3.125263
99	2.26E-16	-15.646664	-0.0440309	3.61948
100	1.07E-13	-12.972461	-0.2294247	2.992136
101	2.53E-15	-14.597012	0.40171985	3.774327
102	3.69E-13	-12.433433	0.67668383	3.627506
103	4.20E-12	-11.376885	0.07727958	2.934035
104	1.07E-13	-12.972461	-0.3358119	2.791142
105	2.53E-13	-12.597645	0.24278094	3.354069
106	2.62E-13	-12.581005	0.29808452	3.224759
107	1.94E-10	-9.7126312	0.81372005	3.214493
108	2.53E-15	-14.597012	0.00432936	3.305947
109	2.53E-15	-14.597012	-0.2492292	3.079254
110	3.95E-15	-14.403825	-0.1083576	3.015803
111	7.19E-13	-12.143158	0.45287336	3.682179
112	3.95E-15	-14.403825	-0.0549024	3.134849
113	1.50E-15	-14.824256	-0.0432031	3.334821
114	6.00E-15	-14.222196	0.27251523	3.670862
115	3.27E-10	-9.4853874	0.62870669	3.034429
116	4.26E-13	-12.370401	0.62678747	3.803775
117	2.40E-14	-13.620136	0.03958285	3.365217
118	9.22E-14	-13.035493	0.56403689	3.493854
119	6.06E-09	-8.2177758	0.86418173	3.119481
120	1.80E-11	-10.744762	0.66078628	3.591747
121	3.89E-14	-13.410309	0.32390323	3.26993
122	3.38E-12	-11.470964	0.31592484	3.051091
123	9.44E-11	-10.025193	-0.0054748	2.955357
124	5.41E-11	-10.266844	1.01929082	3.722718
125	6.31E-14	-13.199705	0.63496009	3.833566
126	3.75E-16	-15.426316	0.36218596	3.690115
127	1.65E-10	-9.7822937	1.06907513	3.750583
128	2.28E-11	-10.64166	0.82698329	3.519814
129	6.31E-14	-13.199705	-0.175894	2.88967
130	1.07E-13	-12.972461	-0.3870417	2.776838
131	2.24E-10	-9.6491396	1.00899944	3.167163

132	1.21E-11	-10.916751	0.45610353	2.881064
133	9.22E-14	-13.035493	0.35210534	3.226583
134	3.95E-15	-14.403825	0.02323482	3.17669
135	2.53E-13	-12.597645	0.24795225	3.397717
136	2.16E-10	-9.6647837	0.53256047	3.239161
137	1.05E-12	-11.978945	0.12558127	2.998519
138	6.09E-15	-14.2153	-0.3136064	3.220238
139	9.35E-15	-14.029009	0.43351162	3.576812
140	1.50E-13	-12.824889	0.03018131	3.184149
141	9.63E-12	-11.016476	0.47566164	3.223162
142	1.94E-10	-9.7126312	0.51016919	3.031073
143	4.20E-12	-11.376885	0.43612834	3.40377
144	2.49E-12	-11.604129	0.60286877	3.594395
145	8.97E-10	-9.0470796	0.99403892	3.171867
146	5.71E-12	-11.24372	0.2997674	3.096189
147	9.22E-14	-13.035493	0.83822475	3.698947
148	9.35E-15	-14.029009	0.64019759	3.77856
149	9.95E-12	-11.002069	-0.0628818	2.859526
150	1.94E-10	-9.7126312	0.53407516	2.981563
151	1.15E-10	-9.9398751	0.22279505	2.72028
152	1.35E-11	-10.868904	0.69735035	3.477919
153	3.95E-15	-14.403825	0.03363356	3.130673
154	3.75E-16	-15.426316	0.33270904	3.717296
155	1.50E-13	-12.824889	-0.107236	3.057044
156	4.20E-12	-11.376885	0.53193171	3.46137
157	1.05E-12	-11.978945	0.54496036	3.455901
158	2.66E-14	-13.574521	-0.3288654	2.798601
159	1.94E-10	-9.7126312	0.39139443	2.887897
160	3.79E-10	-9.4218958	0.71828691	2.959322
161	1.50E-15	-14.824256	0.3034645	3.728671
162	3.69E-13	-12.433433	-0.0162291	2.872066
163	1.56E-13	-12.808249	0.35422994	3.363277
164	3.69E-13	-12.433433	0.51585466	3.426501
165	5.71E-12	-11.24372	0.39426753	3.089552
166	3.81E-16	-15.41942	-0.0900951	3.511898
167	5.85E-16	-15.233129	0.63791859	3.90565
168	1.54E-10	-9.8123561	0.56111594	3.285218
169	1.01E-14	-13.994952	-0.0983386	3.25035
170	5.28E-14	-13.277144	0.74138113	3.385368
171	1.65E-10	-9.7822937	0.75152342	3.487044
172	8.45E-13	-12.073024	0.66803409	3.368298
173	2.53E-15	-14.597012	0.35050933	3.64036
174	9.02E-16	-15.044604	0.14416273	3.767857
175	6.31E-14	-13.199705	0.42196527	3.648863
176	1.01E-14	-13.994952	0.57954211	3.954617

177	2.11E-13	-12.675084	0.36824295	3.140678
178	1.43E-12	-11.84578	0.63699786	3.303106
179	1.50E-13	-12.824889	0.34395913	3.564781
180	5.71E-12	-11.24372	0.4686498	3.196925
181	5.68E-14	-13.24532	0.52347328	4.028249
182	3.34E-17	-16.475968	0.38194272	4.027972
183	2.87E-11	-10.541935	0.44146655	2.88323
184	1.43E-12	-11.84578	0.50373117	3.144021
185	2.49E-12	-11.604129	0.81046584	3.743196
186	1.21E-11	-10.916751	0.60141947	3.058275
187	2.14E-15	-14.669788	0.35191744	3.946584
188	1.15E-10	-9.9398751	1.05925894	3.503966
189	9.63E-12	-11.016476	0.41925924	3.079812
190	5.71E-12	-11.24372	0.68014557	3.344099
191	2.30E-14	-13.637553	0.40083671	3.292645
192	1.43E-12	-11.84578	0.24943234	2.987599
193	2.42E-08	-7.6157158	0.24179852	2.561017
194	9.02E-16	-15.044604	-0.4934054	2.941201
195	2.56E-09	-8.5925919	0.91266395	3.02543
196	5.99E-13	-12.222829	0.44743461	3.65377
197	5.71E-12	-11.24372	0.6239228	3.290683
198	2.53E-13	-12.597645	-0.1587673	3.045198
199	1.08E-08	-7.9665762	0.85341251	2.77381
200	1.58E-14	-13.801765	0.66849692	3.790078
201	4.04E-12	-11.393525	-0.1995824	3.074773
202	7.18E-12	-11.143995	1.11112948	3.447612
203	8.45E-13	-12.073024	1.27578088	3.875231
204	1.65E-10	-9.7822937	0.86799065	3.522567
205	1.70E-12	-11.768341	-0.2348516	2.843897
206	9.37E-17	-16.028376	-0.0654469	3.287646
207	2.62E-13	-12.581005	0.46694725	3.414918
208	9.63E-12	-11.016476	0.31077347	3.010097
209	1.58E-14	-13.801765	-0.3638617	2.870779
210	2.53E-15	-14.597012	-0.1004337	3.207039
211	3.55E-15	-14.44944	0.26172886	3.636962
212	4.20E-12	-11.376885	0.00722615	2.968254
213	5.71E-12	-11.24372	0.9235927	3.672881
214	7.19E-13	-12.143158	0.56097574	3.84755
215	3.75E-16	-15.426316	0.15742577	3.636419
216	6.50E-11	-10.187172	0.44823883	3.140913
217	1.65E-10	-9.7822937	0.72383798	3.336842
218	1.31E-09	-8.8833275	1.17889958	3.631073
219	4.20E-12	-11.376885	0.39243558	3.339394
220	4.85E-11	-10.314691	0.87350682	3.301166
221	2.49E-12	-11.604129	0.53366713	3.463126

222	8.88E-16	-15.0515	0.0810311	3.499207
223	5.68E-14	-13.24532	0.30109216	3.774576
224	9.63E-12	-11.016476	0.86098361	3.59175
225	5.55E-17	-16.25562	0.41480603	3.749352
226	3.81E-16	-15.41942	0.14416791	3.904962
227	2.87E-11	-10.541935	0.64006001	3.128503
228	5.71E-12	-11.24372	0.57119735	3.275991
229	1.56E-13	-12.808249	0.51024373	3.402973
230	3.38E-12	-11.470964	0.75117675	3.447552
231	2.14E-15	-14.669788	0.28314957	3.910023
232	4.20E-12	-11.376885	0.75525931	3.668323
233	2.66E-14	-13.574521	0.6922643	3.917163
234	1.05E-12	-11.978945	0.11180044	3.001174
235	5.07E-15	-14.294972	-0.1922504	3.299379
236	8.36E-18	-17.078028	-0.0853665	3.426501
237	1.80E-11	-10.744762	0.7395828	3.733474
238	1.52E-15	-14.81736	-0.1855476	3.569796
239	7.09E-12	-11.149641	0.77836066	3.710711
240	5.28E-14	-13.277144	0.73991401	3.475939
241	4.20E-12	-11.376885	0.8257689	3.739545
242	6.22E-13	-12.206189	0.67655093	3.58882
243	6.00E-15	-14.222196	0.26830095	3.623602
244	4.26E-13	-12.370401	0.4999244	3.83508
245	6.31E-14	-13.199705	-0.2220694	2.875362
246	1.43E-12	-11.84578	0.5894918	3.257971
247	2.87E-11	-10.541935	0.456822	2.896894
248	1.52E-15	-14.81736	0.31208951	3.921422
249	2.28E-11	-10.64166	1.13201639	3.830357
250	1.71E-14	-13.767708	0.27202865	3.692956
251	6.06E-09	-8.2177758	0.70933863	2.871222
252	1.01E-12	-11.995585	0.67315482	3.919619
253	1.84E-09	-8.7357549	0.04566916	2.633394
254	1.01E-14	-13.994952	-0.939998	2.315972
255	2.53E-15	-14.597012	0.00712918	3.46461
256	6.06E-09	-8.2177758	1.54595384	3.63247
257	4.20E-12	-11.376885	0.39179694	3.319579
258	6.22E-13	-12.206189	0.02605948	2.91471
259	8.45E-13	-12.073024	0.08623795	2.748612
260	3.27E-10	-9.4853874	0.73570822	3.214172
261	3.65E-10	-9.4375399	0.01433939	2.822142
262	5.64E-17	-16.248724	0.11821411	3.787571
263	6.22E-13	-12.206189	0.31977721	3.246708
264	1.50E-15	-14.824256	0.29449203	3.793103
265	9.44E-11	-10.025193	0.74935571	3.701166
266	2.24E-10	-9.6491396	0.40118583	2.660714

267	2.30E-14	-13.637553	0.28020811	3.225641
268	2.28E-11	-10.64166	0.45577579	3.165967
269	3.10E-09	-8.5085113	0.75339195	3.313068
270	4.26E-13	-12.370401	-0.1081257	3.123412
271	1.65E-10	-9.7822937	0.58102721	3.388961
272	6.22E-13	-12.206189	1.00474126	3.981151
273	3.95E-15	-14.403825	-0.0256262	3.144841
274	1.58E-14	-13.801765	0.61859644	3.794996
275	1.01E-14	-13.994952	0.0147266	3.378555
276	6.31E-14	-13.199705	0.0623769	3.17971
277	1.31E-09	-8.8833275	0.28403852	2.936327
278	4.09E-08	-7.388472	0.88251536	3.078794
279	7.18E-12	-11.143995	0.9941444	3.408858
280	3.38E-12	-11.470964	-0.0035508	2.746032
281	2.49E-12	-11.604129	0.4163529	3.363636
282	6.22E-13	-12.206189	1.05985539	4.082486
283	3.75E-16	-15.426316	0.01320781	3.421999
284	2.28E-11	-10.64166	0.60157759	3.385965
285	1.47E-12	-11.831373	0.23750576	3.145342
286	2.53E-15	-14.597012	-0.5007693	2.81498
287	5.71E-12	-11.24372	0.26330425	3.040113
288	1.31E-09	-8.8833275	0.82690308	3.315254
289	5.74E-08	-7.2408996	0.81671365	3.006944
290	1.94E-10	-9.7126312	0.41913134	2.863354
291	4.85E-11	-10.314691	0.71976685	3.199894
292	6.00E-15	-14.222196	0.16232613	3.59175
293	1.07E-13	-12.972461	0.22989823	3.464972
294	3.95E-15	-14.403825	0.13409691	3.27568
295	6.31E-14	-13.199705	-0.054479	3.09324
296	3.55E-15	-14.44944	0.64609748	3.990526
297	6.31E-14	-13.199705	-0.1230964	3.061343
298	3.65E-10	-9.4375399	-0.0851405	2.65035
299	4.20E-12	-11.376885	0.37541748	3.322145
300	1.50E-15	-14.824256	0.05074093	3.339504
301	1.42E-14	-13.84738	-0.0563244	3.276997
302	2.28E-11	-10.64166	0.99818061	3.79026
303	1.07E-13	-12.972461	0.27157184	3.443357
304	6.22E-13	-12.206189	0.75661262	3.633839
305	5.71E-12	-11.24372	0.49202607	3.219246
306	6.83E-14	-13.165648	0.2330452	3.626107
307	6.72E-11	-10.172765	0.87194728	3.89096
308	9.13E-11	-10.0396	0.31740975	3.113128
309	2.49E-12	-11.604129	0.74203563	3.782214
310	2.83E-11	-10.547581	0.15389276	3.205195
311	4.85E-11	-10.314691	0.56661436	3.101633

312	1.58E-14	-13.801765	0.40409948	3.545342
313	5.71E-12	-11.24372	0.88807224	3.617663
314	1.94E-10	-9.7126312	0.74764505	3.155694
315	9.95E-12	-11.002069	-0.0139003	3.058001
316	9.63E-12	-11.016476	0.8811111	3.590774
317	1.54E-10	-9.8123561	0.6209212	3.476621
318	2.66E-14	-13.574521	0.04186581	3.157764
319	4.46E-09	-8.3509298	0.09438273	2.787578
320	3.74E-14	-13.426949	-0.1625648	2.99752
321	9.02E-16	-15.044604	0.43164516	4.041408
322	1.27E-15	-14.897032	0.19309095	3.826389
323	1.15E-10	-9.9398751	1.05008342	3.630204
324	4.20E-12	-11.376885	0.42674565	3.484416
325	6.82E-12	-11.166281	0.23042133	3.465537
326	1.58E-14	-13.801765	0.4011069	3.674531
327	4.57E-13	-12.340339	0.542654	3.778953
328	1.94E-10	-9.7126312	0.37320034	2.854545
329	1.50E-13	-12.824889	0.756067	4.00899
330	1.94E-10	-9.7126312	0.78890385	3.225641
331	3.95E-15	-14.403825	0.28270574	3.396714
332	2.53E-13	-12.597645	0.4330483	3.735065
333	2.28E-11	-10.64166	1.02032977	3.72371
334	7.75E-10	-9.1105712	1.26928463	3.733212
335	2.16E-10	-9.6647837	0.39701371	3.2355
336	1.35E-11	-10.868904	1.03206866	3.776836
337	2.53E-13	-12.597645	-0.2535973	2.920677
338	1.43E-12	-11.84578	0.33546347	3.051282
339	2.87E-11	-10.541935	0.7588966	3.254237
340	1.05E-12	-11.978945	0.82949831	3.690175
341	3.38E-12	-11.470964	0.27156355	2.972783
342	2.87E-11	-10.541935	0.77821853	3.252247
343	9.22E-14	-13.035493	0.37964903	3.307287
344	2.24E-10	-9.6491396	1.15287136	3.322599
345	1.94E-10	-9.7126312	1.1676532	3.641863
346	1.01E-12	-11.995585	0.14727617	3.330853
347	9.59E-14	-13.018076	-0.119882	3.319419
348	6.00E-15	-14.222196	0.17247928	3.458246
349	1.51E-09	-8.8198357	0.86791048	3.049603
350	8.45E-13	-12.073024	0.95496457	3.68165
351	2.26E-16	-15.646664	0.22309971	3.793375
352	2.11E-13	-12.675084	0.25804909	2.854562
353	1.50E-13	-12.824889	-0.0585651	3.255656
354	1.58E-14	-13.801765	-0.2259705	2.909091
355	3.75E-16	-15.426316	0.39127738	3.783582
356	8.36E-18	-17.078028	0.25155603	3.896329

357	2.87E-11	-10.541935	0.50698572	2.957419
358	3.75E-16	-15.426316	0.12121345	3.499301
359	4.50E-12	-11.346823	0.57339658	3.543155
360	9.13E-11	-10.0396	0.34033027	2.989263
361	2.87E-11	-10.541935	0.75809165	3.198135
362	1.51E-09	-8.8198357	1.06425902	3.308176
363	9.22E-14	-13.035493	0.42003249	3.367366
364	3.69E-13	-12.433433	0.65448742	3.625496
365	6.22E-13	-12.206189	0.38139451	3.290683
366	8.18E-11	-10.087447	0.49592688	2.964103
367	6.31E-14	-13.199705	0.35760828	3.653247
368	6.22E-13	-12.206189	-0.0767652	2.845478
369	3.74E-14	-13.426949	0.12832371	3.37931
370	1.07E-13	-12.972461	-0.1082909	2.997101
371	1.77E-12	-11.751701	0.03439814	3.037288
372	3.89E-14	-13.410309	-0.0195263	2.885714
373	7.75E-10	-9.1105712	0.88511949	3.329365
374	2.66E-14	-13.574521	0.04399585	3.219742
375	8.45E-13	-12.073024	0.99881649	3.581419
376	3.59E-09	-8.4450195	0.55010797	2.753731
377	5.85E-16	-15.233129	0.61987586	3.839286
378	2.87E-11	-10.541935	0.68901512	3.10352
379	2.26E-16	-15.646664	0.15646518	3.860859
380	3.38E-12	-11.470964	0.53618125	3.261409
381	8.45E-13	-12.073024	0.34450336	3.139141
382	2.34E-15	-14.631069	0.72423059	3.99887
383	1.68E-11	-10.774825	0.61898335	3.610788
384	6.22E-13	-12.206189	-0.2668218	2.68998
385	4.13E-11	-10.384354	0.1569513	2.881119
386	9.59E-14	-13.018076	0.30090371	3.681299
387	1.58E-14	-13.801765	0.14523255	3.307692
388	1.58E-14	-13.801765	0.40837367	3.682396
389	1.58E-14	-13.801765	0.40870697	3.531299
390	8.45E-13	-12.073024	0.46995895	3.179487
391	9.22E-14	-13.035493	0.04514776	2.968832
392	2.92E-11	-10.534159	0.40670705	3.655932
393	7.18E-12	-11.143995	0.40854959	2.838811
394	2.66E-14	-13.574521	0.05095794	3.186567
395	4.20E-12	-11.376885	0.1909018	3.154762
396	3.69E-13	-12.433433	0.08927396	3.09322
397	1.05E-12	-11.978945	0.11532612	3.119774
398	2.40E-14	-13.620136	0.27428101	3.719196
399	9.35E-15	-14.029009	0.51370833	3.747753
400	2.14E-15	-14.669788	-0.0753922	3.639413
401	4.09E-08	-7.388472	0.31297351	2.657143

402	3.95E-15	-14.403825	0.46737602	3.653613
403	6.22E-13	-12.206189	0.27064842	3.324675
404	3.38E-12	-11.470964	0.82013624	3.513287
405	2.24E-10	-9.6491396	0.85698954	3.024242
406	9.63E-12	-11.016476	0.52099364	3.266526
407	3.27E-10	-9.4853874	0.36079087	2.878371
408	3.55E-15	-14.44944	-0.049001	3.306294
409	4.20E-12	-11.376885	0.44585222	3.479734
410	6.22E-13	-12.206189	-0.0185634	3.005445
411	5.71E-12	-11.24372	0.27021125	3.00744
412	3.38E-12	-11.470964	0.76786606	3.482639
413	1.34E-16	-15.873908	0.26980045	3.918155
414	1.58E-14	-13.801765	0.08973564	3.389238
415	1.94E-10	-9.7126312	0.91505963	3.39435
416	3.34E-17	-16.475968	0.22067169	3.835431
417	6.22E-13	-12.206189	0.68030971	3.611501
418	2.28E-11	-10.64166	0.66919826	3.351185
419	1.56E-13	-12.808249	0.29287827	3.219491
420	1.05E-12	-11.978945	0.10636888	2.995696
421	6.06E-09	-8.2177758	0.60299662	2.828869
422	1.50E-15	-14.824256	0.28076084	3.702877
423	1.58E-14	-13.801765	-0.0579013	3.110119
424	3.55E-15	-14.44944	0.11071239	3.668175
425	1.94E-10	-9.7126312	0.62864192	3.065847
426	9.22E-14	-13.035493	0.47520504	3.334035
427	1.47E-12	-11.831373	0.02217693	2.964103
428	5.71E-12	-11.24372	0.46673964	3.256503
429	6.32E-16	-15.199072	0.01209269	3.334161
430	3.65E-10	-9.4375399	0.78632096	3.610762
431	5.71E-12	-11.24372	0.53891844	3.245221
432	3.79E-10	-9.4218958	0.57709287	2.823903
433	4.59E-10	-9.337815	-0.221534	2.393224
434	6.22E-13	-12.206189	0.55603245	3.592257
435	7.71E-13	-12.113095	-0.018249	3.152778
436	6.22E-13	-12.206189	0.80178877	3.734416
437	1.21E-10	-9.9154588	0.69346155	3.732607
438	1.35E-11	-10.868904	0.83529755	3.527739
439	5.99E-13	-12.222829	0.52240599	3.805808
440	2.14E-15	-14.669788	0.02028866	3.696915
441	9.35E-15	-14.029009	0.27822087	3.471726
442	7.18E-12	-11.143995	0.45761992	2.965098
443	8.45E-13	-12.073024	0.62878283	3.27856
444	5.61E-09	-8.2512049	1.05996427	3.54507
445	1.58E-14	-13.801765	0.36642812	3.612429
446	1.56E-13	-12.808249	-0.3146827	2.587578

447	2.30E-14	-13.637553	0.77289816	3.66784
448	1.51E-09	-8.8198357	1.06376815	3.19627
449	1.47E-12	-11.831373	0.51345883	3.503437
450	1.50E-15	-14.824256	0.29449203	3.793103
451	3.10E-09	-8.5085113	0.38423127	2.918079
452	2.87E-11	-10.541935	0.98572997	3.384109
453	2.24E-10	-9.6491396	0.91077295	3.085317
454	9.63E-12	-11.016476	-0.1003864	2.672131
455	1.52E-15	-14.81736	0.0289425	3.735714
456	7.18E-12	-11.143995	0.25708151	2.704225
457	1.05E-12	-11.978945	0.78272829	3.777366
458	1.34E-16	-15.873908	0.37559204	3.974741
459	5.71E-12	-11.24372	0.71373253	3.493648
460	2.53E-15	-14.597012	0.32139867	3.79774
461	4.85E-11	-10.314691	0.14942014	2.689053
462	3.95E-15	-14.403825	0.34051091	3.496927
463	1.07E-13	-12.972461	0.54532425	3.830611
464	2.87E-11	-10.541935	0.95657233	3.341615
465	6.22E-13	-12.206189	0.2480017	3.277677
466	5.23E-09	-8.2812674	0.94435701	3.347234
467	2.87E-11	-10.541935	0.42527264	2.974592
468	1.54E-10	-9.8123561	0.59815911	3.283582
469	8.97E-10	-9.0470796	1.06314313	3.286364
470	9.59E-14	-13.018076	-0.0750554	3.321523
471	6.22E-13	-12.206189	0.67070449	3.621445
472	5.71E-12	-11.24372	1.22014569	3.857143
473	1.31E-09	-8.8833275	0.99314114	3.464407
474	1.01E-14	-13.994952	0.15907339	3.500207
475	1.56E-13	-12.808249	0.71926544	3.563104
476	3.79E-10	-9.4218958	0.62247668	2.799172
477	6.22E-13	-12.206189	0.5479843	3.477612
478	1.50E-15	-14.824256	-0.3768014	2.894824
479	1.71E-14	-13.767708	0.36451436	3.753292
480	1.58E-14	-13.801765	0.59279469	3.76734
481	6.00E-15	-14.222196	0.27673127	3.633126
482	1.01E-14	-13.994952	0.249957	3.582942
483	9.02E-16	-15.044604	-0.0693611	3.536753
484	1.62E-13	-12.790832	0.00875342	3.393849
485	8.45E-13	-12.073024	1.11309183	3.792541
486	2.53E-15	-14.597012	0.47155898	3.919643
487	2.34E-15	-14.631069	0.85163129	4.044776
488	6.83E-14	-13.165648	-0.0235556	3.357143
489	1.43E-12	-11.84578	1.02462105	3.748281
490	3.74E-14	-13.426949	0.32043285	3.437011
491	6.00E-15	-14.222196	0.45278479	3.898313

492	4.27E-15	-14.369768	0.10077637	3.41471
493	2.34E-15	-14.631069	0.74825097	3.893975
494	4.34E-14	-13.36283	0.04025519	3.386304
495	1.71E-14	-13.767708	0.44902233	3.969752
496	1.43E-12	-11.84578	0.51691009	3.224242
497	5.85E-16	-15.233129	0.74306484	3.94965
498	4.05E-14	-13.392892	-0.0924646	3.196792
499	6.22E-13	-12.206189	0.57980522	3.491097
500	2.66E-14	-13.574521	-0.3168989	2.75626
501	1.35E-11	-10.868904	0.5489591	3.219462
502	4.20E-12	-11.376885	0.71411136	3.707562
503	1.58E-14	-13.801765	0.54564313	3.759425
504	7.18E-12	-11.143995	0.20506162	2.688323
505	1.43E-12	-11.84578	0.77617762	3.490575
506	1.15E-10	-9.9398751	0.85904637	3.306052
507	9.35E-15	-14.029009	0.08637825	3.187011
508	1.51E-09	-8.8198357	0.84234265	2.996488
509	1.54E-10	-9.8123561	0.50419416	3.212121
510	4.95E-18	-17.305272	0.20967679	3.84623
511	2.87E-11	-10.541935	0.74132095	3.27987
512	1.07E-13	-12.972461	0.56416107	3.825989
513	6.31E-14	-13.199705	0.17282095	3.279343
514	1.15E-10	-9.9398751	0.38686845	3.005279
515	1.58E-14	-13.801765	0.52948807	3.679503
516	3.27E-10	-9.4853874	0.69414911	3.124232
517	1.58E-14	-13.801765	0.52561805	3.675362
518	5.76E-15	-14.239613	0.05795228	2.963975
519	3.89E-14	-13.410309	0.46104179	3.429067
520	5.71E-12	-11.24372	0.78221508	3.458736
521	4.27E-15	-14.369768	0.1030901	3.457419
522	1.52E-15	-14.81736	-0.0366247	3.484472
523	1.01E-12	-11.995585	-0.0313534	3.160233
524	2.87E-11	-10.541935	1.15849198	3.678546
525	3.85E-11	-10.414416	0.5173974	3.283616
526	3.85E-11	-10.414416	0.94335516	3.670545
527	9.35E-15	-14.029009	0.7238512	3.907814
528	3.61E-15	-14.442544	0.06838187	3.653613
529	9.59E-14	-13.018076	-0.4931441	2.845584
530	6.22E-13	-12.206189	-0.1928763	2.677897
531	6.31E-14	-13.199705	0.41068136	3.6367
532	1.01E-12	-11.995585	0.46788146	3.63345
533	1.07E-13	-12.972461	0.35320701	3.510536
534	6.60E-10	-9.1802336	0.84145577	3.44807
535	4.26E-13	-12.370401	0.33485653	3.628571
536	4.49E-14	-13.347278	-0.2136931	2.985876

537	1.58E-14	-13.801765	0.35426183	3.575886
538	6.22E-13	-12.206189	-0.510077	2.462189
539	1.15E-10	-9.9398751	0.66377533	3.131448
540	3.27E-10	-9.4853874	0.56594131	3.101028
541	2.87E-11	-10.541935	0.38676939	2.901639
542	3.74E-14	-13.426949	-0.0026933	3.129061
543	5.41E-11	-10.266844	0.55488227	3.279266
544	1.44E-08	-7.8429596	0.8874068	3.138961
545	6.22E-13	-12.206189	0.67250034	3.6036
546	3.38E-12	-11.470964	0.58512183	3.327869
547	1.14E-13	-12.942399	-0.1542274	2.966637
548	3.17E-16	-15.499092	0.48855345	4.131694
549	2.83E-11	-10.547581	0.30935546	3.364286
550	3.95E-15	-14.403825	0.40199913	3.650847
551	1.01E-14	-13.994952	0.07337155	3.467262
552	8.45E-13	-12.073024	0.80989267	3.485075
553	2.34E-15	-14.631069	0.2558778	3.386749
554	5.71E-12	-11.24372	0.20975079	2.9688
555	1.58E-14	-13.801765	0.19389144	3.339772
556	3.85E-11	-10.414416	0.36823146	3.082517
557	1.21E-10	-9.9154588	0.47547056	3.509982
558	5.71E-12	-11.24372	0.74488947	3.480698
559	2.53E-13	-12.597645	-0.2767667	2.854545
560	1.58E-14	-13.801765	0.69349599	4.017532
561	1.54E-10	-9.8123561	1.08943633	3.751097
562	1.58E-14	-13.801765	0.29218382	3.445127
563	9.02E-16	-15.044604	0.10747824	3.634977
564	1.50E-15	-14.824256	0.30992953	3.713287
565	9.22E-14	-13.035493	0.31760794	3.244513
566	4.20E-12	-11.376885	0.58150422	3.61827
567	1.50E-15	-14.824256	0.29610911	3.851153
568	1.01E-14	-13.994952	0.18044331	3.612374
569	1.71E-14	-13.767708	0.1777667	3.521325
570	1.56E-13	-12.808249	-0.1222874	2.85669
571	3.74E-14	-13.426949	0.03268839	3.251977
572	8.97E-10	-9.0470796	0.34463118	2.613591
573	4.04E-12	-11.393525	0.02291804	3.29026
574	1.51E-09	-8.8198357	0.70388817	2.852504
575	2.28E-11	-10.64166	0.71010972	3.32386
576	5.90E-12	-11.229313	0.50282325	3.451282
577	1.31E-09	-8.8833275	1.0264192	3.372455
578	8.45E-13	-12.073024	0.16618961	2.967937
579	1.58E-14	-13.801765	0.33539271	3.511888
580	1.34E-16	-15.873908	-0.0850184	3.426915
581	1.71E-14	-13.767708	0.11123537	3.466637

582	8.45E-13	-12.073024	0.69510314	3.498701
583	1.43E-12	-11.84578	0.74599513	3.481756
584	1.94E-10	-9.7126312	0.82811869	3.243635
585	9.22E-14	-13.035493	0.82740865	3.741201
586	2.49E-12	-11.604129	-0.3664313	2.672727
587	1.58E-14	-13.801765	-0.0200933	3.150794
588	3.81E-16	-15.41942	-0.0857924	3.626834
589	9.22E-14	-13.035493	0.25636298	3.164389
590	1.01E-14	-13.994952	-0.1274686	3.239087
591	1.31E-09	-8.8833275	0.98121674	3.474289
592	5.55E-17	-16.25562	0.18924395	3.514476
593	2.88E-14	-13.540465	0.19604205	3.584149
594	8.45E-13	-12.073024	-0.0331608	2.717758
595	1.56E-13	-12.808249	0.22167725	3.186012
596	3.74E-14	-13.426949	0.35140778	3.673655
597	1.02E-08	-7.9905321	0.27259842	2.553571
598	2.79E-10	-9.5550498	0.27726362	2.995804
599	2.53E-15	-14.597012	0.297184	3.698834
600	6.22E-13	-12.206189	0.33671826	3.323638
601	1.15E-10	-9.9398751	0.71703107	3.179067
602	1.58E-14	-13.801765	0.66510481	3.805164
603	4.20E-12	-11.376885	0.23881274	3.183683
604	7.18E-12	-11.143995	0.55180662	3.013986
605	9.22E-14	-13.035493	0.12682211	3.03354
606	2.28E-11	-10.64166	0.60323334	3.288411
607	3.79E-10	-9.4218958	1.03662171	3.225424
608	1.56E-13	-12.808249	0.41129347	3.358508
609	5.71E-12	-11.24372	1.26178478	3.975145
610	1.07E-13	-12.972461	0.33138022	3.574576
611	9.59E-14	-13.018076	-0.3090979	3.032738
612	7.18E-12	-11.143995	0.65751657	3.163277
613	2.30E-14	-13.637553	0.18357432	3.147321
614	5.85E-16	-15.233129	0.36508401	3.523266
615	1.02E-08	-7.9905321	0.97130617	3.119347
616	2.53E-15	-14.597012	-0.2413383	3.067603
617	1.01E-14	-13.994952	0.57634514	4.03869
618	1.50E-15	-14.824256	0.27489859	3.631056
619	1.21E-11	-10.916751	0.38421624	2.937689
620	3.27E-10	-9.4853874	0.7343821	3.128326
621	9.22E-14	-13.035493	0.23674673	3.266183
622	1.05E-12	-11.978945	0.71385542	3.626501
623	6.83E-14	-13.165648	0.03835276	3.500302
624	5.71E-12	-11.24372	0.91094879	3.683001
625	2.28E-11	-10.64166	0.79651517	3.472344
626	3.38E-12	-11.470964	0.80901766	3.542041

627	5.71E-12	-11.24372	0.28655603	3.004662
628	1.58E-14	-13.801765	0.608331	3.744523
629	2.14E-15	-14.669788	0.07574615	3.686012
630	1.73E-07	-6.7624563	1.67130767	3.405258
631	9.95E-12	-11.002069	0.59987241	3.54965
632	6.83E-14	-13.165648	0.38599156	3.799534
633	5.71E-12	-11.24372	0.88007215	3.589782
634	1.50E-15	-14.824256	0.69575519	4.128183
635	1.42E-14	-13.84738	0.42591118	3.801656
636	2.87E-11	-10.541935	0.65729922	3.108159
637	3.89E-14	-13.410309	-0.0875481	2.912994
638	1.31E-09	-8.8833275	0.92617079	3.366071
639	3.38E-12	-11.470964	0.17423069	2.880158
640	3.79E-10	-9.4218958	0.50254717	2.714223
641	2.53E-15	-14.597012	0.12314498	3.52381
642	3.75E-16	-15.426316	-0.0475284	3.161096
643	1.05E-12	-11.978945	0.000836	2.857193
644	3.38E-12	-11.470964	0.92566462	3.653622
645	5.71E-12	-11.24372	0.38203413	3.06087
646	1.56E-13	-12.808249	0.451309	3.399068
647	2.11E-13	-12.675084	0.20029075	2.888199
648	2.40E-14	-13.620136	-0.0066771	3.333187
649	1.51E-09	-8.8198357	0.95561565	3.106294
650	1.05E-12	-11.978945	0.77445598	3.70676
651	3.69E-13	-12.433433	0.93159302	3.975802
652	6.22E-13	-12.206189	-0.1362031	2.785338
653	1.31E-09	-8.8833275	0.39446954	2.841408
654	8.45E-13	-12.073024	0.48903285	3.197669
655	1.56E-13	-12.808249	-0.4681013	2.449517
656	8.45E-13	-12.073024	0.90367984	3.699351
657	1.07E-13	-12.972461	0.07314539	3.272871
658	9.22E-14	-13.035493	0.26620152	3.174327
659	2.87E-11	-10.541935	0.3478016	2.866737
660	2.53E-15	-14.597012	0.52464891	3.854035
661	2.34E-15	-14.631069	0.13940141	3.367232
662	2.53E-13	-12.597645	0.09626789	3.197574
663	1.71E-14	-13.767708	-0.0599676	3.362712
664	1.05E-12	-11.978945	0.79050857	3.723003
665	2.53E-13	-12.597645	0.48525326	3.717081
666	1.50E-13	-12.824889	-0.0519073	3.116567
667	9.22E-14	-13.035493	0.84821112	3.781387
668	9.95E-12	-11.002069	0.73923684	3.690909
669	5.41E-11	-10.266844	0.84899164	3.540793
670	1.01E-14	-13.994952	0.42178555	3.863095
671	3.75E-16	-15.426316	0.5098478	3.89648

672	1.58E-14	-13.801765	0.4300286	3.573085
673	1.94E-10	-9.7126312	0.73475865	3.319759
674	1.42E-14	-13.84738	0.17281441	3.628249
675	9.13E-11	-10.0396	0.28942483	3.108442
676	3.04E-11	-10.517519	0.21659885	3.124224
677	6.22E-13	-12.206189	0.63528845	3.585548
678	1.43E-12	-11.84578	0.64797877	3.330992
679	1.21E-10	-9.9154588	0.13971537	3.144633
680	5.71E-12	-11.24372	0.56817506	3.255048
681	6.06E-09	-8.2177758	1.26505023	3.435572
682	1.35E-11	-10.868904	1.09261905	3.813326
683	4.05E-14	-13.392892	-0.0222766	3.336597
684	3.74E-14	-13.426949	0.52921395	3.660016
685	9.95E-12	-11.002069	-0.335342	2.583845
686	9.37E-17	-16.028376	0.02930844	3.334116
687	1.07E-13	-12.972461	-0.258468	2.894345
688	4.85E-11	-10.314691	0.81372005	3.214493
689	2.53E-15	-14.597012	-0.0792087	3.251097
690	6.32E-16	-15.199072	-0.1283523	3.216317
691	1.07E-13	-12.972461	0.37898259	3.625989
692	1.51E-09	-8.8198357	0.88695096	3.082496
693	1.05E-12	-11.978945	-0.0858852	2.914689
694	1.47E-12	-11.831373	0.73411277	3.706349
695	2.27E-13	-12.64326	0.24583211	3.663194
696	8.45E-13	-12.073024	1.21840126	4.027254
697	6.22E-13	-12.206189	0.29407741	3.202484
698	2.92E-07	-6.5352125	0.9688049	2.948288
699	2.28E-11	-10.64166	0.9133184	3.621528
700	2.53E-13	-12.597645	0.13872021	3.300699
701	5.71E-12	-11.24372	0.82241068	3.534722
702	5.71E-12	-11.24372	0.52631127	3.251984
703	6.31E-14	-13.199705	0.45481178	3.661706
704	1.68E-11	-10.774825	0.4672875	3.395961
705	4.26E-13	-12.370401	0.47866695	3.606025
706	6.22E-13	-12.206189	0.28785469	3.273929
707	1.56E-13	-12.808249	0.53323327	3.462687
708	3.75E-16	-15.426316	0.22828587	3.691525
709	1.56E-13	-12.808249	0.5494697	3.460455
710	1.43E-12	-11.84578	-0.339209	2.462712
711	9.95E-12	-11.002069	0.54047515	3.553107
712	1.58E-14	-13.801765	0.48800902	3.697421
713	1.54E-10	-9.8123561	0.64127288	3.424077
714	6.00E-15	-14.222196	-0.0974386	3.230465
715	1.35E-11	-10.868904	0.51767666	3.263353
716	2.34E-15	-14.631069	0.15776543	3.245465

717	8.45E-13	-12.073024	0.65422548	3.393971
718	1.05E-12	-11.978945	0.1890879	3.173453
719	3.95E-15	-14.403825	-0.3379827	2.788811
720	4.50E-12	-11.346823	0.54128118	3.434272
721	1.58E-14	-13.801765	0.31359472	3.555367
722	1.94E-10	-9.7126312	0.63892279	3.187662
723	2.28E-11	-10.64166	1.05000123	3.732401
724	1.94E-10	-9.7126312	1.04730912	3.42236
725	9.22E-14	-13.035493	0.46188782	3.40979
726	1.07E-13	-12.972461	-0.1791683	3.023164
727	2.87E-11	-10.541935	0.59387718	3.068948
728	2.87E-11	-10.541935	0.86847707	3.314484
729	6.32E-16	-15.199072	-0.0602031	3.272608
730	8.97E-10	-9.0470796	0.09122181	2.490566
731	2.66E-14	-13.574521	0.24317933	3.434028
732	2.36E-11	-10.627253	-0.3161602	2.639881
733	2.87E-11	-10.541935	0.55424747	3.070621
734	1.47E-12	-11.831373	0.55058742	3.480246
735	1.51E-09	-8.8198357	1.03821694	3.331429
736	1.56E-13	-12.808249	-0.1269658	2.85193
737	1.35E-11	-10.868904	0.71184228	3.491833
738	6.22E-13	-12.206189	0.20143109	3.108903
739	1.50E-15	-14.824256	0.14670451	3.528205
740	4.05E-14	-13.392892	-0.4650692	2.776604
741	2.26E-16	-15.646664	0.18121456	3.682807
742	2.53E-15	-14.597012	0.11666555	3.472783
743	2.49E-12	-11.604129	0.32441702	3.355717
744	6.00E-15	-14.222196	-0.0341495	3.417423
745	6.06E-09	-8.2177758	0.73383192	2.972316
746	4.26E-13	-12.370401	-0.3949957	2.837013
747	3.95E-15	-14.403825	0.06989554	3.169405
748	1.94E-10	-9.7126312	0.75506999	3.271022
749	4.85E-11	-10.314691	0.50261292	3.024294
750	2.87E-11	-10.541935	0.71787913	3.162238
751	5.71E-12	-11.24372	0.25162453	3.072078
752	3.74E-14	-13.426949	0.3303221	3.527778
753	2.11E-13	-12.675084	1.20245316	3.858648
754	2.53E-15	-14.597012	-0.4081501	2.965537
755	2.30E-14	-13.637553	0.76794817	3.700176
756	9.37E-17	-16.028376	-0.0191896	3.362103
757	1.01E-14	-13.994952	0.35187103	3.718012
758	9.95E-12	-11.002069	0.80684929	3.759441
759	2.28E-11	-10.64166	-0.0284079	2.704429
760	5.71E-12	-11.24372	0.22004191	2.941259
761	2.40E-14	-13.620136	-0.0647883	3.436758

762	4.85E-11	-10.314691	0.6809725	3.146825
763	2.87E-11	-10.541935	1.04851541	3.556494
764	1.01E-14	-13.994952	0.204712	3.572432
765	6.22E-13	-12.206189	0.47487819	3.509377
766	1.50E-15	-14.824256	0.55276825	4.011905
767	2.53E-15	-14.597012	0.29542457	3.654244
768	6.22E-13	-12.206189	0.79850406	3.693662
769	8.88E-16	-15.0515	0.00264089	3.386905
770	3.04E-11	-10.517519	0.36946725	3.356954
771	3.34E-17	-16.475968	0.41903896	4.072261
772	6.22E-13	-12.206189	0.26789428	3.141799
773	2.49E-12	-11.604129	0.6699919	3.582195
774	6.60E-10	-9.1802336	0.66342846	3.311424
775	5.71E-12	-11.24372	0.61748634	3.251759
776	2.40E-14	-13.620136	-0.4346117	2.912216
777	1.51E-09	-8.8198357	0.68892427	2.813333
778	1.58E-14	-13.801765	0.62909732	3.849206
779	8.45E-13	-12.073024	0.88462217	3.55619
780	1.58E-14	-13.801765	0.22337916	3.371378
781	2.53E-13	-12.597645	0.19471055	3.34065
782	6.83E-14	-13.165648	0.12015638	3.476734
783	1.51E-09	-8.8198357	0.74211254	2.928671
784	3.95E-15	-14.403825	0.18273446	3.308489
785	6.31E-14	-13.199705	0.60201716	3.84294
786	1.94E-10	-9.7126312	-0.0952495	2.436364
787	5.71E-12	-11.24372	0.82427839	3.447612
788	9.59E-14	-13.018076	0.00457117	3.345917
789	1.50E-15	-14.824256	-0.4983838	2.796737
790	1.46E-09	-8.83548	0.71228248	3.429563
791	4.85E-11	-10.314691	0.67265129	3.139385
792	1.21E-11	-10.916751	0.43162783	2.890255
793	2.79E-10	-9.5550498	0.66584615	3.447671
794	1.70E-12	-11.768341	0.15375539	3.296752
795	2.42E-08	-7.6157158	0.62219215	2.828904
796	3.75E-16	-15.426316	0.28992962	3.713294
797	6.40E-09	-8.19382	1.65742506	3.335681
798	3.81E-16	-15.41942	0.20488491	3.793679
799	8.97E-10	-9.0470796	0.46686325	2.749153
800	1.58E-14	-13.801765	0.28974769	3.484127
801	7.75E-10	-9.1105712	0.88672359	3.368927
802	1.51E-09	-8.8198357	0.50485602	2.763617
803	3.38E-12	-11.470964	0.17192423	2.877963
804	3.27E-10	-9.4853874	0.99938148	3.37971
805	6.22E-13	-12.206189	0.40828338	3.33626
806	1.47E-12	-11.831373	0.48321986	3.542208

807	4.85E-11	-10.314691	0.64015685	3.060041
808	3.95E-15	-14.403825	0.72908375	3.979905
809	2.40E-14	-13.620136	0.08799924	3.507139
810	5.71E-12	-11.24372	0.60830756	3.29324
811	6.00E-15	-14.222196	0.19577323	3.629825
812	2.66E-14	-13.574521	0.18037624	3.325285
813	1.62E-11	-10.791465	0.49064991	3.722898
814	1.21E-11	-10.916751	0.45892543	3.00484
815	1.62E-13	-12.790832	0.32836168	3.831821
816	1.43E-12	-11.84578	0.8612185	3.552448
817	1.58E-14	-13.801765	0.65797266	3.816977
818	1.01E-12	-11.995585	0.14783748	3.31049
819	1.62E-13	-12.790832	-0.0029562	3.297731
820	5.71E-12	-11.24372	0.47093709	3.18042
821	5.55E-17	-16.25562	0.22531149	3.639881
822	2.49E-12	-11.604129	-0.1803575	2.777778
823	5.71E-12	-11.24372	0.63943285	3.400565
824	1.58E-14	-13.801765	0.61802802	3.815385
825	2.53E-13	-12.597645	-0.0558836	3.133792
826	1.58E-14	-13.801765	0.19081889	3.356643
827	9.35E-15	-14.029009	-0.0594022	3.031299
828	1.50E-15	-14.824256	-0.070792	3.459276
829	1.05E-12	-11.978945	0.38385894	3.350694
830	5.71E-12	-11.24372	0.69721777	3.415179
831	3.75E-16	-15.426316	0.53976022	3.973893
832	1.43E-12	-11.84578	0.01509211	2.681229
833	1.05E-12	-11.978945	0.52550902	3.537853
834	3.38E-12	-11.470964	0.71501241	3.377226
835	2.14E-15	-14.669788	0.05295092	3.736237
836	1.07E-13	-12.972461	0.13958853	3.301632
837	6.32E-16	-15.199072	0.2921771	3.764407
838	2.53E-13	-12.597645	0.51336803	3.724702
839	3.55E-15	-14.44944	-0.0366539	3.342262
840	3.75E-16	-15.426316	-0.1856166	3.195664
841	4.26E-13	-12.370401	-0.056518	3.111607
842	1.21E-10	-9.9154588	0.7006682	3.716384
843	1.50E-15	-14.824256	0.32514856	3.629825
844	7.75E-10	-9.1105712	0.22377796	2.705004
845	3.75E-16	-15.426316	-0.0848108	3.310418
846	1.51E-09	-8.8198357	0.20071902	2.562987
847	2.88E-12	-11.541098	0.4322192	3.637401
848	3.81E-16	-15.41942	0.03689419	3.593503
849	9.22E-14	-13.035493	0.51917691	3.509254
850	3.95E-15	-14.403825	-0.1942341	2.984664
851	2.11E-13	-12.675084	0.25799207	2.926839

852	1.80E-11	-10.744762	0.38359682	3.44026
853	1.56E-13	-12.808249	0.06164173	2.967702
854	5.71E-12	-11.24372	0.9501258	3.618525
855	1.07E-13	-12.972461	-0.0712724	3.095734
856	2.53E-15	-14.597012	0.11004671	3.508929
857	6.32E-16	-15.199072	-0.2227205	3.088674
858	4.61E-13	-12.336345	-0.1525851	3.257062
859	9.02E-16	-15.044604	0.36866617	4.012121
860	6.22E-13	-12.206189	0.4246262	3.435028
861	2.28E-11	-10.64166	0.61172432	3.374011
862	9.35E-15	-14.029009	0.18277878	3.369048
863	1.47E-12	-11.831373	0.24268271	3.207341
864	6.83E-14	-13.165648	0.03528391	3.401865
865	4.27E-15	-14.369768	0.14991001	3.489855
866	4.26E-13	-12.370401	0.46292376	3.608282
867	3.95E-15	-14.403825	0.30887488	3.574713
868	1.24E-08	-7.9064512	0.80039314	3.235897
869	7.18E-12	-11.143995	0.40158318	2.863477
870	2.66E-14	-13.574521	0.02599873	3.140787
871	2.28E-11	-10.64166	0.25740882	3.034463
872	1.79E-12	-11.746055	0.6583437	3.109091
873	8.88E-16	-15.0515	-0.0555446	3.343733
874	5.55E-17	-16.25562	0.2882173	3.735061
875	4.20E-12	-11.376885	-0.1814524	2.722153
876	1.94E-10	-9.7126312	0.3585776	2.841492
877	6.22E-13	-12.206189	0.2863861	3.17723
878	6.82E-12	-11.166281	0.67935426	3.903274
879	2.53E-15	-14.597012	0.13164428	3.430581
880	2.40E-14	-13.620136	0.33583173	3.742657
881	2.28E-11	-10.64166	0.55806783	3.195483
882	6.22E-13	-12.206189	0.40025173	3.309731
883	4.85E-11	-10.314691	0.19183226	2.785065
884	6.31E-14	-13.199705	-0.0235375	3.087785
885	1.58E-14	-13.801765	0.3961745	3.577156
886	2.53E-13	-12.597645	0.46259492	3.741077
887	3.69E-13	-12.433433	0.2785611	3.205004
888	2.56E-09	-8.5925919	0.70003248	2.909722
889	6.31E-14	-13.199705	-0.3353777	2.791608
890	6.82E-12	-11.166281	0.07242551	3.318814
891	6.31E-14	-13.199705	0.21436902	3.361721
892	6.31E-14	-13.199705	-0.208211	2.948413
893	6.60E-10	-9.1802336	0.86024035	3.551515
894	8.45E-13	-12.073024	0.62768968	3.368588
895	2.28E-11	-10.64166	1.15564158	3.873612
896	8.45E-13	-12.073024	0.45977116	3.151888

897	7.18E-12	-11.143995	1.14087857	3.522388
898	1.07E-13	-12.972461	-0.1834408	2.996298
899	9.22E-14	-13.035493	0.76775365	3.719814
900	8.97E-10	-9.0470796	0.87186523	2.977786
901	1.01E-14	-13.994952	0.16538519	3.619774
902	1.31E-09	-8.8833275	0.58545992	3.118572
903	1.70E-12	-11.768341	-0.0651319	3.146328
904	4.85E-11	-10.314691	1.09276863	3.431322
905	9.13E-11	-10.0396	0.57554507	3.24472
906	2.26E-16	-15.646664	0.04215982	3.645833
907	3.38E-12	-11.470964	0.88763056	3.638418
908	4.85E-11	-10.314691	0.62352822	3.113696
909	2.34E-15	-14.631069	0.39487203	3.575758
910	4.82E-12	-11.31681	-0.2881429	3.180294
911	2.14E-15	-14.669788	0.16025119	3.894156
912	4.85E-11	-10.314691	0.94977489	3.352063
913	2.87E-11	-10.541935	1.02103257	3.450893
914	8.45E-13	-12.073024	0.88532549	3.521909
915	2.66E-14	-13.574521	-0.5798073	2.529138
916	1.56E-13	-12.808249	0.62306603	3.637288
917	3.95E-15	-14.403825	0.0546439	3.190518
918	6.06E-09	-8.2177758	0.34935944	2.687662
919	5.20E-12	-11.283791	-0.0943588	3.012008
920	3.38E-12	-11.470964	0.86464202	3.555711
921	3.89E-14	-13.410309	0.10818733	3.051282
922	1.43E-12	-11.84578	0.05774693	2.804563
923	2.53E-15	-14.597012	0.12679625	3.444053
924	1.44E-14	-13.840484	0.19559119	3.967156
925	3.55E-15	-14.44944	0.09389218	3.490575
926	9.22E-14	-13.035493	0.47568628	3.423776
927	1.15E-10	-9.9398751	0.35755485	2.857639
928	2.11E-13	-12.675084	0.72838613	3.44494
929	1.01E-14	-13.994952	0.32677404	3.830006
930	6.31E-14	-13.199705	0.08869302	3.207867
931	7.18E-12	-11.143995	0.4925367	3.055844
932	1.04E-09	-8.9830523	1.08325829	3.764103
933	1.05E-12	-11.978945	0.42961491	3.440113
934	3.59E-09	-8.4450195	1.04853339	3.217345
935	9.35E-15	-14.029009	0.03820664	3.101053
936	7.18E-12	-11.143995	0.13720475	2.715668
937	1.62E-13	-12.790832	0.10786366	3.506448
938	1.80E-11	-10.744762	0.4914689	3.550649
939	1.07E-13	-12.972461	0.5301333	3.700176
940	9.95E-12	-11.002069	0.3799972	3.488688
941	2.49E-12	-11.604129	0.59365688	3.487089

942	2.11E-13	-12.675084	0.44784528	3.123395
943	2.28E-11	-10.64166	0.20903532	3.009074
944	3.89E-14	-13.410309	0.01089096	3.109351
945	3.85E-11	-10.414416	0.53928161	3.193662
946	3.89E-14	-13.410309	0.53158197	3.424491
947	1.05E-12	-11.978945	0.19725229	3.16121
948	7.75E-10	-9.1105712	1.15019245	3.648052
949	2.34E-15	-14.631069	0.15695475	3.32028
950	5.71E-12	-11.24372	0.74082613	3.519661
951	6.60E-10	-9.1802336	0.30373176	3.058699
952	6.06E-09	-8.2177758	1.20964444	3.370056
953	2.53E-13	-12.597645	-0.1009857	3.023705
954	3.95E-15	-14.403825	0.17632376	3.362103
955	1.14E-13	-12.942399	0.26531083	3.457837
956	3.74E-14	-13.426949	0.48524258	3.694444
957	5.23E-09	-8.2812674	0.94386447	3.346795
958	6.00E-15	-14.222196	-0.2174101	3.136905
959	1.35E-11	-10.868904	0.17546231	2.955932
960	1.43E-12	-11.84578	0.30339278	3.003073
961	1.05E-12	-11.978945	0.74009155	3.653002
962	1.07E-13	-12.972461	-0.1439515	2.923732
963	6.66E-15	-14.176581	0.04537907	3.22123
964	5.99E-13	-12.222829	0.08099507	3.281333
965	8.88E-16	-15.0515	0.09126434	3.423602
966	4.27E-15	-14.369768	0.06120844	3.389648
967	5.23E-09	-8.2812674	1.07663885	3.581818
968	1.07E-13	-12.972461	0.31121905	3.445963
969	9.22E-14	-13.035493	0.50209639	3.394757
970	1.56E-13	-12.808249	0.78896619	3.783712
971	1.05E-12	-11.978945	-0.1814643	2.83908
972	3.69E-13	-12.433433	0.65017286	3.642517
973	1.43E-12	-11.84578	0.22507768	3.094268
974	6.22E-13	-12.206189	0.80439189	3.73705
975	3.74E-14	-13.426949	0.19352705	3.283599
976	3.79E-10	-9.4218958	0.60359193	2.829365
977	3.89E-14	-13.410309	-0.0136991	2.927739
978	3.38E-12	-11.470964	0.9613555	3.731397
979	5.71E-12	-11.24372	0.87770277	3.549605
980	1.56E-13	-12.808249	0.08593002	3.010097
981	1.52E-15	-14.81736	0.16873696	3.773427
982	1.58E-14	-13.801765	-0.0009174	3.290706
983	1.01E-12	-11.995585	-0.3008867	2.84871
984	1.50E-13	-12.824889	-0.0137512	3.117208
985	4.85E-11	-10.314691	0.55807351	3.01958
986	7.18E-12	-11.143995	0.9483856	3.303221

987	1.44E-08	-7.8429596	0.4995482	2.813636
988	3.38E-12	-11.470964	0.74082613	3.519661
989	2.28E-11	-10.64166	0.26894935	2.987879
990	6.31E-14	-13.199705	0.03766266	3.134977
991	1.94E-10	-9.7126312	1.00149109	3.398156
992	3.38E-12	-11.470964	0.35881026	3.131638
993	5.71E-12	-11.24372	0.80059283	3.513889
994	8.88E-16	-15.0515	0.21714221	3.654151
995	2.40E-14	-13.620136	-0.1665888	3.17296
996	2.30E-14	-13.637553	0.59573287	3.545455
997	3.75E-16	-15.426316	0.04998475	3.488136
998	5.28E-14	-13.277144	0.55618143	3.243635
999	1.58E-14	-13.801765	0.08001386	3.280275
1000	6.00E-15	-14.222196	0.34170519	3.820904

Annex Table S1.4. Composite Likelihood (CL) and Ratio (CLR) estimated in 1000 regions of 1kb taken randomly from the 3R chromosome in the datasets with (*FBti0019386* +) and without (*FBti0019386* -) the element**.**

				CL		
chr.start	**chr.end**	**dataset**	**sample_id**	***FBti0019386* +**	**FBti0019386 -**	**CLR**
20078025	20079025	568	586	4.45E-05	5.43E-18	8.56
12015987	12016987	941	971	1.58E-07	5.77E-19	4.63
20154668	20155668	539	554	1.36E-03	1.94E-10	3.98
16592452	16593452	636	657	4.13E-08	2.53E-18	2.83
12030464	12031464	133	136	1.59E-20	1.38E-42	2.26
264845	265845	408	422	1.00E+00	1.48E-01	0.83
353411	354411	176	179	8.19E-02	2.31E-03	0.46
2951623	2952623	134	137	1.56E-02	1.31E-04	0.27
185647	186647	404	418	1.00E+00	1.00E+00	0.00
2723396	2724396	415	429	1.56E-02	3.20E-04	-0.12
21690092	21691092	710	732	2.50E-01	2.50E-01	-0.60
3749567	3750567	945	975	3.91E-03	1.31E-04	-0.93
134551	135551	679	701	3.70E-02	1.56E-02	-1.06
7837217	7838217	521	536	3.70E-02	1.56E-02	-1.06
516917	517917	580	598	3.20E-04	1.25E-06	-1.09
25563991	25564991	293	300	2.50E-01	1.00E+00	-1.20
10581903	10582903	380	393	2.14E-05	1.03E-08	-1.35
578232	579232	992	1022	1.56E-02	8.64E-03	-1.55
81752	82752	834	860	1.28E-03	6.10E-05	-1.57
142216	143216	124	127	5.49E-03	1.28E-03	-1.63
1984213	1985213	785	810	1.48E-01	1.00E+00	-1.66
13994239	13995239	553	570	1.68E-03	1.31E-04	-1.67
14576729	14577729	28	28	1.61E-06	1.54E-10	-1.77
17300978	17301978	10	10	2.46E-09	5.43E-16	-1.95
12128397	12129397	411	425	1.10E-08	1.90E-14	-2.20
27843708	27844708	394	408	3.70E-02	2.50E-01	-2.26
751957	752957	622	643	1.28E-03	3.43E-04	-2.32
27551611	27552611	127	130	1.24E-03	3.43E-04	-2.35
1294422	1295422	932	962	1.28E-03	4.12E-04	-2.40
313386	314386	306	314	5.49E-03	1.52E-02	-2.70
27482632	27483632	878	906	1.94E-05	2.38E-07	-2.80
27555017	27556017	895	924	8.07E-06	4.13E-08	-2.80
337231	338231	501	516	3.20E-04	8.57E-05	-2.92
5589861	5590861	382	395	3.20E-04	8.57E-05	-2.92
3741902	3742902	583	601	4.32E-08	1.70E-12	-2.96
4103830	4104830	361	372	1.12E-06	1.31E-09	-3.02
267400	268400	93	96	1.31E-04	1.94E-05	-3.05
472634	473634	688	710	1.28E-03	2.31E-03	-3.15
12124139	12125139	788	813	6.40E-09	6.31E-14	-3.19

20135081	20136081	357	368	3.17E-16	2.62E-28	-3.42
5161509	5162509	765	790	5.08E-11	7.81E-18	-3.48
1736399	1737399	11	11	3.11E-04	3.43E-04	-3.55
11313422	11314422	894	922	3.81E-06	5.54E-08	-3.58
1393207	1394207	358	369	5.79E-04	1.28E-03	-3.58
3719761	3720761	825	850	5.79E-04	1.28E-03	-3.58
6174906	6175906	82	84	5.79E-04	1.28E-03	-3.58
10768402	10769402	258	263	6.55E-06	1.65E-07	-3.59
3724019	3725019	933	963	3.11E-04	5.79E-04	-3.78
5967117	5968117	461	476	1.61E-06	2.42E-08	-3.97
9843571	9844571	609	630	1.61E-06	2.56E-08	-3.99
14514563	14515563	123	126	1.76E-19	3.19E-34	-4.01
19947731	19948731	63	64	6.18E-23	4.62E-41	-4.08
3735941	3736941	265	270	3.20E-04	1.28E-03	-4.10
19792741	19793741	759	784	1.94E-05	4.86E-06	-4.11
10707087	10708087	296	304	1.34E-06	2.94E-08	-4.21
11670240	11671240	224	228	1.56E-11	4.95E-18	-4.31
5876848	5877848	999	1029	1.28E-03	3.70E-02	-4.35
10759035	10760035	913	942	7.49E-12	1.82E-18	-4.51
12288497	12289497	959	989	8.57E-05	3.11E-04	-4.63
1009138	1010138	317	325	9.54E-07	4.13E-08	-4.66
8027973	8028973	920	949	6.06E-09	2.19E-12	-4.78
20139339	20140339	43	43	2.38E-07	3.59E-09	-4.80
8902560	8903560	733	758	9.54E-07	6.97E-08	-4.88
11736664	11737664	370	382	4.38E-08	2.33E-10	-5.08
17446600	17447600	110	113	6.82E-12	6.53E-18	-5.15
383217	384217	947	977	1.28E-03	2.50E-01	-5.18
4817465	4818465	188	191	6.61E-07	6.97E-08	-5.20
6685010	6686010	526	541	1.21E-06	2.38E-07	-5.21
27736407	27737407	255	259	1.21E-06	2.38E-07	-5.21
16906690	16907690	640	661	8.46E-14	1.21E-21	-5.23
16180	17180	151	154	7.77E-05	1.28E-03	-5.33
14412372	14413372	118	121	1.24E-12	6.82E-19	-5.64
8100359	8101359	665	687	2.24E-10	2.66E-14	-5.72
7322854	7323854	369	381	1.28E-06	1.18E-06	-5.86
3191773	3192773	773	798	5.79E-04	2.50E-01	-5.87
10477157	10478157	558	576	2.57E-15	5.84E-24	-5.94
402803	403803	163	166	3.28E-05	1.28E-03	-6.08
6410797	6411797	520	535	8.07E-06	7.77E-05	-6.08
4407849	4408849	245	249	3.69E-07	1.65E-07	-6.08
2709771	2710771	754	779	2.56E-08	1.51E-09	-6.36
709377	710377	633	654	1.80E-11	9.16E-16	-6.45
19636899	19637899	336	346	1.94E-05	1.28E-03	-6.53
2213292	2214292	218	222	4.86E-06	8.57E-05	-6.56
7758018	7759018	778	803	4.86E-06	8.57E-05	-6.56

5811275	5812275	996	1026	8.37E-08	2.56E-08	-6.56
892470	893470	899	928	2.38E-07	2.38E-07	-6.62
20062696	20063696	366	378	1.21E-11	6.78E-16	-6.66
3496643	3497643	222	226	2.38E-07	2.92E-07	-6.71
57908	58908	161	164	4.32E-08	1.08E-08	-6.76
9264488	9265488	956	986	2.72E-09	4.85E-11	-6.82
22194235	22195235	625	646	2.45E-12	3.99E-17	-6.82
19106356	19107356	205	209	2.24E-10	3.57E-13	-6.85
636992	637992	298	306	6.97E-08	4.13E-08	-6.93
1287609	1288609	231	235	1.65E-07	2.79E-07	-7.01
805607	806607	184	187	2.38E-07	6.61E-07	-7.07
12309787	12310787	676	698	1.15E-10	1.62E-13	-7.09
27773025	27774025	696	718	1.02E-08	1.31E-09	-7.10
5506404	5507404	74	76	4.86E-06	3.20E-04	-7.13
3830468	3831468	505	520	1.21E-06	2.14E-05	-7.16
4628411	4629411	56	57	2.38E-07	9.54E-07	-7.22
203530	204530	603	623	8.57E-05	1.48E-01	-7.30
3070846	3071846	413	427	6.97E-08	1.02E-07	-7.32
4897515	4898515	437	451	4.67E-06	5.79E-04	-7.42
1994432	1995432	412	426	2.70E-09	1.94E-10	-7.42
10717306	10718306	809	834	2.70E-09	1.94E-10	-7.42
5355672	5356672	101	104	2.38E-07	1.61E-06	-7.45
27410247	27411247	362	373	2.38E-07	1.61E-06	-7.45
677017	678017	842	868	1.21E-06	4.35E-05	-7.47
19812327	19813327	776	801	4.50E-12	7.20E-16	-7.55
17289907	17290907	349	360	1.52E-15	1.12E-22	-7.68
2972913	2973913	326	335	6.97E-08	2.38E-07	-7.69
13112840	13113840	794	819	1.10E-08	6.04E-09	-7.70
25773483	25774483	576	594	1.21E-06	8.57E-05	-7.76
8263013	8264013	419	433	2.62E-06	5.79E-04	-7.93
1472405	1473405	122	125	1.60E-09	2.24E-10	-7.94
4090204	4091204	216	220	1.08E-08	1.08E-08	-7.97
9482496	9483496	795	820	1.40E-09	1.94E-10	-7.99
16046580	16047580	931	961	2.38E-07	6.44E-06	-8.05
3010383	3011383	974	1004	7.75E-10	9.13E-11	-8.18
12978288	12979288	966	996	7.52E-12	8.78E-15	-8.19
27566940	27567940	574	592	7.18E-12	9.72E-15	-8.28
9857197	9858197	644	665	1.51E-09	4.59E-10	-8.30
20135933	20136933	130	133	3.02E-17	2.28E-25	-8.40
12469035	12470035	375	388	7.83E-10	1.54E-10	-8.40
591857	592857	140	143	6.06E-09	1.08E-08	-8.47
17098298	17099298	76	78	6.50E-11	1.25E-12	-8.47
20259414	20260414	810	835	6.97E-08	1.61E-06	-8.52
21565759	21566759	150	153	1.51E-09	8.66E-10	-8.58
8276639	8277639	954	984	4.13E-08	6.91E-07	-8.61

7919821	7920821	268	273	4.26E-13	9.37E-17	-8.71
18238582	18239582	88	91	6.31E-14	2.09E-18	-8.72
15933318	15934318	482	497	1.10E-08	6.97E-08	-8.76
3048705	3049705	263	268	1.21E-11	8.91E-14	-8.78
7464218	7465218	930	960	2.53E-15	4.22E-21	-8.82
6228556	6229556	917	946	3.95E-15	1.10E-20	-8.85
1305493	1306493	416	430	1.03E-08	1.02E-07	-8.98
8664966	8665966	643	664	6.91E-07	5.79E-04	-9.08
20023523	20024523	550	567	1.44E-08	2.94E-07	-9.15
8273232	8274232	305	313	2.38E-07	8.57E-05	-9.18
5576235	5577235	581	599	4.10E-07	3.43E-04	-9.31
21093976	21094976	911	940	9.24E-13	2.05E-15	-9.38
2423635	2424635	697	719	1.03E-08	2.79E-07	-9.42
16611187	16612187	165	168	8.18E-11	1.93E-11	-9.46
9491012	9492012	417	431	1.79E-12	9.72E-15	-9.48
3757231	3758231	614	635	3.59E-09	4.38E-08	-9.53
23406053	23407053	286	293	1.39E-07	7.77E-05	-9.60
5948382	5949382	995	1025	5.61E-11	1.35E-11	-9.63
1745767	1746767	656	677	3.57E-13	5.49E-16	-9.64
1802823	1803823	695	717	1.21E-11	7.14E-13	-9.69
27631661	27632661	209	213	1.10E-08	6.61E-07	-9.74
27499664	27500664	820	845	9.54E-07	5.12E-03	-9.75
20342870	20343870	832	857	3.52E-31	9.44E-52	-9.88
4570503	4571503	902	931	2.70E-09	5.96E-08	-9.91
16945011	16946011	922	951	4.50E-12	1.76E-13	-9.94
11713671	11714671	549	566	2.72E-09	8.37E-08	-10.05
1765353	1766353	442	456	1.21E-11	1.70E-12	-10.07
11723039	11724039	672	694	4.62E-14	3.18E-17	-10.17
4060398	4061398	838	864	1.03E-08	1.61E-06	-10.18
706822	707822	217	221	5.61E-11	4.85E-11	-10.19
1231404	1232404	20	20	5.65E-07	5.49E-03	-10.24
949526	950526	646	667	3.95E-15	2.70E-19	-10.24
3346763	3347763	612	633	3.79E-10	2.58E-09	-10.26
17489180	17490180	504	519	5.73E-18	6.74E-25	-10.31
12806266	12807266	322	331	1.71E-15	6.75E-20	-10.37
19089324	19090324	257	262	6.31E-14	9.66E-17	-10.38
1850513	1851513	935	965	7.18E-12	1.79E-12	-10.54
3781076	3782076	570	588	7.75E-10	2.09E-08	-10.54
3781076	3782076	624	645	7.75E-10	2.09E-08	-10.54
4241788	4242788	605	625	2.40E-14	2.03E-17	-10.55
27250147	27251147	800	825	3.79E-10	6.06E-09	-10.63
5927092	5928092	229	233	3.79E-10	6.40E-09	-10.65
5731225	5732225	314	322	1.50E-15	1.01E-19	-10.65
21174877	21175877	637	658	2.41E-12	2.62E-13	-10.66
11052834	11053834	535	550	1.03E-08	4.86E-06	-10.66

8290264	8291264	190	193	1.51E-09	1.39E-07	-10.78
19952840	19953840	516	531	1.02E-16	6.46E-22	-10.79
1148799	1149799	648	669	4.85E-11	1.54E-10	-10.82
14741938	14742938	837	863	1.51E-20	1.53E-29	-10.83
18875574	18876574	167	170	2.26E-16	3.45E-21	-10.83
2181783	2182783	427	441	5.28E-14	2.22E-16	-10.90
2047231	2048231	398	412	1.21E-11	1.21E-11	-10.92
15762999	15763999	494	509	1.56E-22	2.32E-33	-10.98
12448596	12449596	660	682	1.01E-14	1.02E-17	-11.00
13212476	13213476	112	115	1.79E-12	3.69E-13	-11.06
2617799	2618799	197	200	1.58E-14	3.43E-17	-11.14
10006226	10007226	493	508	2.62E-13	1.01E-14	-11.17
27853927	27854927	873	901	1.51E-09	3.59E-07	-11.19
13106878	13107878	943	973	3.48E-19	2.63E-26	-11.34
27765361	27766361	828	853	1.62E-13	5.76E-15	-11.34
6889393	6890393	573	591	2.42E-16	1.40E-20	-11.38
487111	488111	113	116	1.79E-12	7.71E-13	-11.38
18771680	18772680	106	109	4.00E-10	4.13E-08	-11.41
2812814	2813814	971	1001	1.21E-11	5.61E-11	-11.58
18672043	18673043	46	46	1.54E-10	1.08E-08	-11.66
755363	756363	454	468	4.13E-08	1.28E-03	-11.88
7574926	7575926	790	815	8.88E-16	6.37E-19	-11.91
1651240	1652240	691	713	6.06E-09	3.23E-05	-11.94
1387246	1388246	905	934	4.85E-11	2.56E-09	-12.04
5568571	5569571	908	937	5.61E-11	3.59E-09	-12.06
1875209	1876209	944	974	2.41E-12	7.09E-12	-12.09
8416300	8417300	745	770	1.18E-10	2.56E-08	-12.26
21073538	21074538	736	761	2.81E-18	1.66E-23	-12.32
17122143	17123143	812	837	6.97E-13	1.02E-12	-12.32
889915	890915	55	56	1.39E-17	5.18E-22	-12.43
2289083	2290083	206	210	9.87E-16	2.81E-18	-12.46
19013532	19014532	67	68	2.81E-18	3.19E-23	-12.61
11379846	11380846	340	350	1.38E-25	8.46E-38	-12.64
7895976	7896976	111	114	3.57E-13	6.22E-13	-12.69
23095221	23096221	774	799	1.40E-17	1.01E-21	-12.71
12836923	12837923	214	218	2.04E-20	2.15E-27	-12.72
6186828	6187828	758	783	3.79E-10	9.54E-07	-12.82
20344573	20345573	593	611	1.01E-14	9.02E-16	-12.95
16653766	16654766	331	341	1.44E-14	2.34E-15	-13.05
4526220	4527220	350	361	4.93E-11	3.20E-08	-13.12
20818911	20819911	793	818	1.75E-25	4.51E-37	-13.17
3356982	3357982	579	597	2.88E-10	1.28E-06	-13.19
10206350	10207350	242	246	6.31E-14	6.31E-14	-13.20
22040948	22041948	189	192	2.62E-13	1.09E-12	-13.20
837116	838116	243	247	2.72E-09	1.20E-04	-13.21

24863130	24864130	973	1003	7.20E-11	1.02E-07	-13.30
1819004	1820004	135	138	1.50E-13	5.48E-13	-13.39
964855	965855	682	704	3.89E-14	3.89E-14	-13.41
480298	481298	775	800	5.76E-15	8.88E-16	-13.43
12114772	12115772	181	184	2.81E-18	2.53E-22	-13.51
19483612	19484612	237	241	5.22E-19	9.98E-24	-13.56
7478696	7479696	22	22	2.11E-13	1.79E-12	-13.60
12477551	12478551	770	795	3.23E-13	4.68E-12	-13.65
16106192	16107192	613	634	4.64E-11	1.02E-07	-13.68
18760609	18761609	100	103	4.89E-22	1.37E-29	-13.76
12417088	12418088	396	410	2.85E-14	5.28E-14	-13.81
18404643	18405643	407	421	4.49E-13	1.35E-11	-13.83
4647146	4648146	597	616	1.21E-11	1.02E-08	-13.84
5238152	5239152	868	896	2.31E-14	3.74E-14	-13.85
13290823	13291823	783	808	4.95E-18	1.78E-21	-13.86
20188732	20189732	215	219	3.38E-12	8.97E-10	-13.89
18634573	18635573	424	438	4.05E-27	1.41E-39	-13.94
27633364	27634364	632	653	1.00E-10	9.54E-07	-13.98
811568	812568	928	957	1.56E-13	2.39E-12	-14.00
23408608	23409608	148	151	9.02E-16	8.30E-17	-14.01
3276081	3277081	228	232	1.01E-14	1.38E-14	-14.13
8125055	8126055	866	894	6.09E-15	6.00E-15	-14.21
5016738	5017738	727	751	1.59E-14	5.28E-14	-14.32
8219582	8220582	686	708	1.68E-11	6.97E-08	-14.39
24348767	24349767	247	251	1.62E-18	7.04E-22	-14.43
11173761	11174761	410	424	5.71E-12	1.08E-08	-14.52
8224692	8225692	397	411	1.42E-11	7.13E-08	-14.55
5459567	5460567	817	842	4.81E-10	8.57E-05	-14.57
1022764	1023764	620	641	8.97E-10	3.20E-04	-14.60
12606141	12607141	771	796	4.97E-20	9.96E-25	-14.61
19866829	19867829	365	377	1.02E-17	5.59E-20	-14.73
8797814	8798814	503	518	3.75E-16	7.59E-17	-14.73
8806330	8807330	668	690	1.58E-16	1.39E-17	-14.74
2016573	2017573	972	1002	4.34E-14	1.47E-12	-14.89
9484199	9485199	982	1012	3.50E-12	1.08E-08	-14.94
7821036	7822036	844	870	5.08E-18	2.36E-20	-14.96
1640169	1641169	569	587	1.43E-12	2.56E-09	-15.10
1855622	1856622	571	589	8.90E-17	1.12E-17	-15.15
22874658	22875658	934	964	1.84E-12	5.61E-09	-15.22
27268882	27269882	211	215	6.56E-14	7.18E-12	-15.22
12773054	12774054	846	872	6.32E-16	7.19E-16	-15.26
18785305	18786305	821	846	2.00E-18	8.14E-21	-15.31
21137407	21138407	262	267	8.28E-21	1.50E-25	-15.34
16789170	16790170	146	149	1.44E-15	5.76E-15	-15.44
2122171	2123171	48	48	1.77E-12	9.46E-09	-15.48

3060627	3061627	15	15	2.88E-14	2.88E-12	-15.54
20834240	20835240	288	295	3.25E-16	4.26E-16	-15.61
12155648	12156648	383	396	4.50E-12	1.02E-07	-15.70
6214930	6215930	129	132	6.78E-16	2.53E-15	-15.74
11845668	11846668	718	741	1.06E-23	6.93E-31	-15.79
7953033	7954033	307	315	1.07E-16	7.59E-17	-15.82
7871280	7872280	420	434	2.59E-17	4.95E-18	-15.87
25753897	25754897	647	668	3.34E-17	8.36E-18	-15.87
14270155	14271155	327	336	5.15E-20	2.00E-23	-15.88
12865026	12866026	769	794	1.58E-14	2.49E-12	-16.00
8674333	8675333	712	734	3.65E-17	1.39E-17	-16.02
19741645	19742645	156	159	8.91E-14	9.13E-11	-16.06
19807218	19808218	843	869	2.36E-16	6.52E-16	-16.07
15367860	15368860	533	548	1.51E-17	2.73E-18	-16.08
11513547	11514547	811	836	5.85E-18	5.22E-19	-16.18
16352302	16353302	302	310	1.01E-14	1.79E-12	-16.24
14558846	14559846	792	817	1.44E-15	3.74E-14	-16.26
19240056	19241056	120	123	1.99E-15	7.32E-14	-16.27
7254726	7255726	455	469	1.12E-12	2.56E-08	-16.31
3035080	3036080	373	385	5.76E-15	8.45E-13	-16.41
12685340	12686340	339	349	1.75E-24	7.83E-32	-16.41
4894109	4895109	489	504	7.52E-12	1.61E-06	-16.45
18415714	18416714	86	89	5.55E-17	9.52E-17	-16.49
6906425	6907425	402	416	1.52E-15	7.32E-14	-16.50
20457835	20458835	955	985	5.76E-15	1.05E-12	-16.50
17889429	17890429	557	575	1.31E-29	9.26E-42	-16.73
2218401	2219401	492	507	5.76E-15	1.79E-12	-16.73
27688718	27689718	354	365	3.65E-17	9.37E-17	-16.85
12274020	12275020	453	467	1.72E-17	2.09E-17	-16.85
9259378	9260378	980	1010	1.73E-17	2.56E-17	-16.93
2826439	2827439	431	445	2.22E-21	4.44E-25	-16.95
6080379	6081379	423	437	6.31E-14	3.79E-10	-16.98
8079921	8080921	483	498	3.52E-18	1.19E-18	-16.98
6581116	6582116	545	561	2.71E-15	7.05E-13	-16.98
15605454	15606454	714	736	9.02E-16	8.11E-14	-17.00
20042258	20043258	114	117	2.36E-16	6.66E-15	-17.08
20961127	20962127	169	172	1.76E-18	4.16E-19	-17.13
2434706	2435706	143	146	4.93E-11	3.43E-04	-17.15
2722545	2723545	807	832	9.35E-15	1.35E-11	-17.19
27746626	27747626	906	935	2.42E-16	9.72E-15	-17.22
13157122	13158122	699	721	4.02E-22	2.70E-26	-17.22
9106091	9107091	300	308	3.57E-13	2.56E-08	-17.30
7017132	7018132	631	652	5.85E-18	8.00E-18	-17.37
1696374	1697374	508	523	3.81E-16	3.74E-14	-17.41
4567097	4568097	602	622	1.46E-16	5.76E-15	-17.43

12891425	12892425	958	988	2.47E-16	2.40E-14	-17.60
12921231	12922231	171	174	2.19E-24	1.94E-30	-17.61
12492028	12493028	761	786	1.90E-17	1.46E-16	-17.61
4250304	4251304	462	477	3.14E-17	4.08E-16	-17.62
8968133	8969133	16	16	9.22E-14	3.59E-09	-17.63
5288396	5289396	337	347	3.19E-13	4.38E-08	-17.63
18870465	18871465	984	1014	1.81E-25	1.45E-32	-17.65
7704368	7705368	324	333	5.85E-16	1.56E-13	-17.66
1171792	1172792	204	208	1.01E-14	5.61E-11	-17.74
3843242	3844242	191	194	7.18E-12	3.28E-05	-17.80
10567426	10568426	50	50	1.87E-21	2.61E-24	-17.87
15958014	15959014	737	762	3.05E-20	7.18E-22	-17.89
23250211	23251211	983	1013	1.36E-16	1.42E-14	-17.89
2998461	2999461	829	854	1.02E-17	8.13E-17	-17.89
7505947	7506947	405	419	7.92E-17	6.00E-15	-17.98
19702472	19703472	144	147	9.65E-25	9.03E-31	-17.99
20837646	20838646	681	703	3.49E-21	1.23E-23	-18.00
5548133	5549133	269	274	1.07E-13	1.44E-08	-18.10
1587370	1588370	560	578	5.55E-17	4.57E-15	-18.17
12893128	12894128	225	229	3.60E-16	2.11E-13	-18.21
17444045	17445045	942	972	5.20E-15	4.60E-11	-18.23
558645	559645	748	773	3.28E-14	2.00E-09	-18.27
3836429	3837429	693	715	3.57E-13	2.38E-07	-18.27
11828636	11829636	487	502	1.71E-19	5.59E-20	-18.28
1333595	1334595	862	890	9.72E-15	2.24E-10	-18.38
4952017	4953017	986	1016	5.76E-15	9.46E-11	-18.46
17108517	17109517	236	240	1.67E-18	8.36E-18	-18.48
13367466	13368466	220	224	3.25E-16	3.57E-13	-18.53
2700403	2701403	421	435	3.61E-15	4.85E-11	-18.57
16248408	16249408	939	969	9.87E-16	4.50E-12	-18.66
4372933	4373933	723	746	6.47E-18	2.48E-16	-18.77
2371688	2372688	376	389	3.95E-17	9.72E-15	-18.79
11674498	11675498	904	933	1.72E-19	1.92E-19	-18.81
25414962	25415962	406	420	6.25E-22	2.95E-24	-18.88
19871087	19872087	192	195	1.20E-20	1.56E-21	-19.03
12641908	12642908	187	190	1.92E-19	4.04E-19	-19.04
5420394	5421394	967	997	1.21E-15	1.62E-11	-19.04
2272052	2273052	551	568	6.47E-13	4.67E-06	-19.05
16105340	16106340	351	362	5.35E-16	3.38E-12	-19.07
16364224	16365224	145	148	5.36E-15	3.79E-10	-19.12
14678920	14679920	356	367	5.85E-16	5.71E-12	-19.22
6528317	6529317	52	53	9.37E-17	1.56E-13	-19.25
12375359	12376359	629	650	3.58E-18	3.04E-16	-19.38
5559203	5560203	246	250	1.51E-18	7.92E-17	-19.54
6094856	6095856	364	375	6.91E-17	1.73E-13	-19.56

2822181	2823181	886	914	3.81E-16	7.18E-12	-19.69
27882029	27883029	888	916	1.79E-13	2.64E-06	-19.92
27755994	27756994	756	781	2.26E-16	4.50E-12	-19.95
1269726	1270726	730	754	4.59E-20	2.24E-19	-20.02
13498612	13499612	393	407	3.24E-25	1.16E-29	-20.04
10446500	10447500	541	557	4.56E-19	2.34E-17	-20.05
7034164	7035164	595	614	1.76E-18	3.86E-16	-20.10
13989130	13990130	715	737	2.96E-19	1.43E-17	-20.21
769840	770840	675	697	7.39E-16	9.13E-11	-20.22
27724485	27725485	530	545	3.18E-19	1.98E-17	-20.29
19876197	19877197	285	292	9.61E-30	1.82E-38	-20.29
11617441	11618441	915	944	8.18E-21	1.38E-20	-20.31
3775114	3776114	334	344	3.96E-21	4.17E-21	-20.42
3805772	3806772	60	61	2.34E-15	1.51E-09	-20.44
4831942	4832942	893	921	1.58E-16	7.18E-12	-20.46
13180115	13181115	379	392	2.50E-21	1.91E-21	-20.48
8205957	8206957	989	1019	9.52E-17	3.38E-12	-20.57
809865	810865	969	999	4.73E-24	8.61E-27	-20.58
7621763	7622763	663	685	1.83E-14	1.65E-07	-20.69
20053328	20054328	968	998	3.45E-21	5.90E-21	-20.69
26514370	26515370	865	893	7.72E-29	3.15E-36	-20.72
13015758	13016758	936	966	3.19E-23	6.13E-25	-20.78
4309915	4310915	180	183	9.82E-17	7.31E-12	-20.88
23845475	23846475	630	651	5.11E-22	2.16E-22	-20.92
4877077	4878077	994	1024	2.94E-18	9.72E-15	-21.05
5485966	5486966	44	44	8.73E-22	8.73E-22	-21.06
12693855	12694855	292	299	1.33E-19	2.03E-17	-21.06
18225808	18226808	25	25	8.76E-21	9.31E-20	-21.08
27803683	27804683	352	363	1.24E-18	2.34E-15	-21.18
7295603	7296603	395	409	3.98E-22	2.47E-22	-21.19
3369756	3370756	709	731	9.67E-16	1.51E-09	-21.21
2672301	2673301	924	953	1.15E-20	2.24E-19	-21.23
11484592	11485592	554	571	4.20E-21	3.93E-20	-21.35
7659233	7660233	610	631	1.36E-18	6.00E-15	-21.51
21609190	21610190	304	312	1.23E-27	4.89E-33	-21.51
22316865	22317865	54	55	3.17E-16	3.50E-10	-21.54
20229608	20230608	722	745	1.51E-26	8.32E-31	-21.56
4607122	4608122	497	512	1.51E-17	8.45E-13	-21.57
8703287	8704287	752	777	1.58E-16	1.15E-10	-21.66
9153780	9154780	753	778	2.94E-29	4.61E-36	-21.73
10361340	10362340	926	955	1.67E-18	1.55E-14	-21.75
3963317	3964317	261	266	5.85E-18	2.11E-13	-21.79
22396063	22397063	617	638	2.34E-21	3.95E-20	-21.86
2574367	2575367	108	111	5.22E-19	2.34E-15	-21.93
3052963	3053963	484	499	1.67E-18	2.66E-14	-21.98

8886380	8887380	703	725	2.07E-21	4.39E-20	-22.01
19323513	19324513	13	13	3.22E-24	1.09E-25	-22.02
27706601	27707601	542	558	1.79E-22	4.47E-22	-22.14
27706601	27707601	707	729	1.79E-22	4.47E-22	-22.14
23391576	23392576	607	627	5.33E-16	4.16E-09	-22.17
11082640	11083640	430	444	1.23E-17	2.31E-12	-22.18
11082640	11083640	517	532	1.23E-17	2.31E-12	-22.18
19234947	19235947	95	98	3.00E-21	2.21E-19	-22.39
15174548	15175548	634	655	3.90E-22	4.22E-21	-22.44
14638896	14639896	877	905	3.12E-27	3.05E-31	-22.50
20341167	20342167	653	674	1.91E-33	1.46E-43	-22.60
4033999	4034999	414	428	3.43E-17	4.85E-11	-22.61
19650525	19651525	577	595	3.35E-27	4.75E-31	-22.63
21792283	21793283	725	749	2.52E-25	3.06E-27	-22.68
13557372	13558372	279	286	3.42E-26	5.67E-29	-22.69
19446994	19447994	173	176	6.59E-23	2.36E-22	-22.74
2238839	2239839	198	201	3.39E-17	6.72E-11	-22.77
20852975	20853975	62	63	1.78E-21	1.92E-19	-22.78
18446371	18447371	853	879	5.18E-22	1.85E-20	-22.84
6509582	6510582	555	573	1.56E-27	1.81E-31	-22.87
11876326	11877326	626	647	6.00E-15	2.80E-06	-22.89
3333989	3334989	159	162	4.56E-19	1.83E-14	-22.95
10325573	10326573	174	177	1.05E-21	9.98E-20	-22.95
4179621	4180621	883	911	4.74E-18	2.41E-12	-23.03
13651898	13652898	75	77	1.09E-21	1.28E-19	-23.03
6523207	6524207	152	155	5.22E-19	3.89E-14	-23.15
18560485	18561485	856	883	1.03E-24	1.75E-25	-23.22
3362092	3363092	53	54	1.33E-19	3.71E-15	-23.32
15732342	15733342	102	105	7.76E-30	1.35E-35	-23.35
6045463	6046463	136	139	3.78E-16	3.22E-08	-23.35
6045463	6046463	977	1007	3.78E-16	3.22E-08	-23.35
8126758	8127758	89	92	4.81E-20	5.35E-16	-23.36
21925131	21926131	847	873	3.73E-19	3.89E-14	-23.45
16747442	16748442	443	457	1.62E-22	7.66E-21	-23.47
12215260	12216260	290	297	4.37E-25	6.23E-26	-23.51
5142774	5143774	871	899	4.41E-28	6.41E-32	-23.52
15818353	15819353	141	144	3.36E-29	4.18E-34	-23.57
25728349	25729349	319	327	6.13E-25	1.50E-25	-23.60
16707417	16708417	374	387	1.92E-25	1.51E-26	-23.61
20367566	20368566	253	257	8.94E-27	3.32E-29	-23.62
8844652	8845652	422	436	5.22E-19	1.14E-13	-23.62
20742268	20743268	949	979	2.52E-21	2.95E-18	-23.67
6962630	6963630	70	72	5.55E-17	1.60E-09	-23.72
20363309	20364309	79	81	2.43E-28	3.22E-32	-23.74
24173339	24174339	671	693	1.71E-19	1.59E-14	-23.74

27032139	27033139	172	175	8.36E-18	4.93E-11	-23.85
9941505	9942505	618	639	1.36E-25	1.35E-26	-23.86
15707646	15708646	294	302	1.60E-26	1.89E-28	-23.87
10095643	10096643	221	225	1.19E-22	1.29E-20	-23.96
1485179	1486179	652	673	1.41E-17	2.79E-10	-24.15
8489537	8490537	879	907	1.03E-19	1.58E-14	-24.17
11220598	11221598	99	102	4.16E-19	2.62E-13	-24.18
18100624	18101624	154	157	6.43E-24	6.63E-23	-24.21
2698700	2699700	284	291	1.66E-23	4.48E-22	-24.21
13470509	13471509	272	277	4.77E-22	3.77E-19	-24.22
13470509	13471509	360	371	4.77E-22	3.77E-19	-24.22
6957520	6958520	611	632	8.28E-21	1.34E-16	-24.29
4496414	4497414	979	1009	1.53E-15	4.86E-06	-24.32
19114021	19115021	731	755	1.10E-24	2.84E-24	-24.37
11637879	11638879	434	448	6.98E-28	1.32E-30	-24.43
24243170	24244170	85	87	2.51E-27	1.84E-29	-24.47
6506176	6507176	137	140	1.95E-28	1.14E-31	-24.48
27084086	27085086	251	255	2.31E-22	1.71E-19	-24.50
8090991	8091991	58	59	2.96E-25	3.24E-25	-24.57
18933483	18934483	155	158	2.77E-30	2.98E-35	-24.59
23218702	23219702	301	309	2.12E-18	1.80E-11	-24.60
15091943	15092943	985	1015	1.72E-34	1.26E-43	-24.63
23107995	23108995	529	544	4.49E-23	9.21E-21	-24.66
18150868	18151868	903	932	1.10E-24	6.98E-24	-24.76
23016875	23017875	687	709	3.66E-23	8.28E-21	-24.79
4515149	4516149	244	248	3.10E-19	7.19E-13	-24.87
3420851	3421851	182	185	2.56E-25	7.38E-25	-25.05
21616003	21617003	128	131	3.80E-20	2.10E-14	-25.16
11195902	11196902	698	720	1.12E-23	1.84E-21	-25.17
11406246	11407246	23	23	9.37E-21	1.30E-15	-25.17
15799618	15800618	21	21	8.94E-27	1.33E-27	-25.22
11393472	11394472	287	294	6.33E-29	6.72E-32	-25.23
10296619	10297619	104	107	3.21E-24	2.06E-22	-25.30
23079041	23080041	36	36	4.20E-25	4.15E-24	-25.37
9089059	9090059	14	14	1.56E-20	6.53E-15	-25.43
10618522	10619522	496	511	8.96E-27	2.24E-27	-25.45
11452232	11453232	199	202	1.03E-24	3.91E-23	-25.56
18285420	18286420	923	952	7.57E-23	2.24E-19	-25.59
19714394	19715394	515	530	1.12E-23	5.90E-21	-25.67
8729687	8730687	323	332	1.60E-19	1.79E-12	-25.85
11800534	11801534	818	843	2.12E-26	4.39E-26	-25.99
19670963	19671963	645	666	1.56E-22	2.55E-18	-26.02
25697692	25698692	29	29	5.90E-23	3.73E-19	-26.03
2799188	2800188	885	913	2.52E-24	7.76E-22	-26.09
11405394	11406394	289	296	6.13E-25	4.86E-23	-26.11

26365341	26366341	694	716	1.89E-23	5.15E-20	-26.16
16661431	16662431	24	24	1.07E-22	1.67E-18	-26.16
10156106	10157106	739	764	8.14E-30	1.12E-32	-26.23
10145887	10146887	321	330	1.85E-31	6.93E-36	-26.31
4080837	4081837	657	678	1.23E-19	3.38E-12	-26.35
3495792	3496792	705	727	1.05E-21	3.04E-16	-26.44
17455116	17456116	960	990	4.31E-28	6.12E-29	-26.52
16049986	16050986	273	278	8.23E-22	2.47E-16	-26.56
18656715	18657715	27	27	2.95E-30	8.04E-33	-26.97
12897386	12898386	869	897	6.47E-26	3.86E-24	-26.97
13482431	13483431	230	234	1.08E-35	1.11E-43	-26.98
24195480	24196480	998	1028	3.24E-25	1.11E-22	-27.03
17291610	17292610	311	319	2.27E-41	6.03E-55	-27.07
14300813	14301813	562	580	2.01E-25	4.86E-23	-27.08
815826	816826	264	269	8.09E-24	9.43E-20	-27.16
8620683	8621683	523	538	1.52E-20	3.57E-13	-27.19
20917696	20918696	757	782	2.78E-25	1.28E-22	-27.22
23779903	23780903	57	58	3.66E-26	2.29E-24	-27.23
3988013	3989013	852	878	2.31E-22	9.37E-17	-27.24
7326260	7327260	835	861	5.31E-21	6.56E-14	-27.37
9661330	9662330	532	547	2.96E-25	2.18E-22	-27.40
8362650	8363650	18	18	2.50E-20	1.70E-12	-27.44
19590061	19591061	378	391	5.30E-27	1.05E-25	-27.57
4803840	4804840	341	351	5.11E-22	9.93E-16	-27.58
4810652	4811652	465	480	5.60E-22	1.21E-15	-27.59
18765719	18766719	519	534	3.58E-26	6.52E-24	-27.71
20530221	20531221	436	450	2.15E-33	2.55E-38	-27.74
3097246	3098246	392	406	1.65E-23	1.51E-18	-27.74
3812585	3813585	31	31	2.02E-26	2.34E-24	-27.76
23648757	23649757	170	173	3.54E-28	8.80E-28	-27.85
16274807	16275807	728	752	2.71E-25	5.25E-22	-27.85
4691429	4692429	320	329	2.31E-26	3.86E-24	-27.86
21288991	21289991	963	993	4.75E-31	1.72E-33	-27.88
6969443	6970443	38	38	1.29E-22	1.34E-16	-27.90
7118471	7119471	559	577	2.26E-21	4.34E-14	-27.93
15110678	15111678	919	948	2.17E-23	4.05E-18	-27.93
5398252	5399252	186	189	1.20E-22	1.28E-16	-27.95
24150346	24151346	861	889	2.57E-31	6.33E-34	-27.98
17530056	17531056	929	958	2.78E-23	7.81E-18	-28.00
7137206	7138206	426	440	2.37E-25	6.28E-22	-28.05
8937476	8938476	975	1005	7.03E-19	5.61E-09	-28.06
7382465	7383465	962	992	1.36E-24	2.85E-20	-28.19
23252766	23253766	855	882	1.92E-25	6.17E-22	-28.22
8442700	8443700	248	252	3.36E-29	1.99E-29	-28.25
2239691	2240691	115	118	2.26E-21	9.22E-14	-28.26

9426290	9427290	565	583	3.40E-25	2.19E-21	-28.28
8739906	8740906	677	699	9.88E-27	2.19E-24	-28.35
19222173	19223173	851	877	2.09E-27	9.79E-26	-28.35
21527437	21528437	751	776	1.76E-21	7.32E-14	-28.37
8663262	8664262	784	809	2.40E-21	1.62E-13	-28.45
13268681	13269681	762	787	6.62E-23	1.29E-16	-28.47
6660314	6661314	720	743	1.65E-23	8.36E-18	-28.49
17867287	17868287	403	417	3.56E-27	4.04E-25	-28.51
17867287	17868287	606	626	3.56E-27	4.04E-25	-28.51
5124039	5125039	651	672	1.22E-28	5.01E-28	-28.53
20267078	20268078	863	891	7.03E-29	1.76E-28	-28.55
8699029	8700029	540	555	1.54E-22	8.88E-16	-28.57
14392785	14393785	548	565	4.43E-29	7.92E-29	-28.61
7634537	7635537	91	94	2.31E-22	2.34E-15	-28.64
14420036	14421036	7	7	5.63E-40	1.62E-50	-28.71
10237859	10238859	744	769	1.95E-31	3.07E-33	-28.91
16544762	16545762	45	45	3.84E-26	1.57E-22	-29.03
18643089	18644089	260	265	1.65E-23	3.34E-17	-29.09
6558974	6559974	543	559	6.03E-26	4.60E-22	-29.10
18349289	18350289	4	4	2.36E-25	7.63E-21	-29.14
11166948	11167948	511	526	5.52E-20	4.72E-10	-29.19
1757689	1758689	98	101	9.88E-25	1.76E-19	-29.25
8532117	8533117	880	908	3.09E-23	1.80E-16	-29.28
18115101	18116101	916	945	5.51E-28	6.04E-26	-29.30
5318202	5319202	444	458	1.20E-26	3.19E-23	-29.35
10777770	10778770	210	214	1.25E-29	3.60E-29	-29.36
2870722	2871722	282	289	3.02E-27	2.70E-24	-29.47
20613677	20614677	281	288	1.11E-26	4.54E-23	-29.57
14881600	14882600	742	767	1.90E-30	1.41E-30	-29.59
26522034	26523034	997	1027	1.34E-28	8.94E-27	-29.69
7036718	7037718	582	600	3.19E-23	5.85E-16	-29.76
2394681	2395681	599	618	7.54E-22	3.57E-13	-29.80
8554258	8555258	486	501	1.21E-31	9.73E-33	-29.82
14047889	14048889	346	356	8.28E-29	5.30E-27	-29.89
22706894	22707894	160	163	3.95E-32	1.24E-33	-29.90
23668344	23669344	119	122	1.69E-32	2.38E-34	-29.92
23454594	23455594	371	383	1.31E-31	2.97E-32	-30.24
5204940	5205940	952	982	3.62E-25	2.67E-19	-30.31
17819598	17820598	249	253	3.14E-27	2.07E-23	-30.32
19210251	19211251	803	828	1.34E-28	4.39E-26	-30.38
15305693	15306693	80	82	1.76E-32	8.00E-34	-30.41
25030894	25031894	766	791	1.56E-24	6.47E-18	-30.43
23463961	23464961	849	875	1.85E-25	1.01E-19	-30.47
25925067	25926067	621	642	2.61E-23	2.34E-15	-30.54
23303862	23304862	827	852	2.25E-30	1.84E-29	-30.56

21434613	21435613	740	765	2.37E-27	2.07E-23	-30.57
24310445	24311445	457	472	1.11E-26	5.60E-22	-30.66
6283058	6284058	467	482	1.02E-25	4.81E-20	-30.66
6822968	6823968	741	766	1.55E-31	1.21E-31	-30.71
12333631	12334631	274	279	1.76E-32	1.61E-33	-30.72
5651175	5652175	202	205	1.66E-23	2.05E-15	-30.87
5651175	5652175	267	272	1.66E-23	2.05E-15	-30.87
8109726	8110726	951	981	2.73E-32	6.38E-33	-30.93
1888834	1889834	252	256	8.75E-24	9.02E-16	-31.07
5489373	5490373	254	258	3.95E-24	2.47E-16	-31.20
25829689	25830689	450	464	2.73E-26	1.20E-20	-31.21
16845375	16846375	329	338	1.05E-24	1.98E-17	-31.25
23828443	23829443	433	447	1.87E-44	6.73E-57	-31.28
20888742	20889742	303	311	1.58E-30	5.09E-29	-31.31
25913145	25914145	92	95	9.79E-38	2.29E-43	-31.38
20710759	20711759	787	812	1.31E-29	5.30E-27	-31.49
17403169	17404169	673	695	3.43E-24	4.51E-16	-31.58
16702307	16703307	777	802	1.02E-30	4.28E-29	-31.61
15582461	15583461	669	691	1.31E-24	9.37E-17	-31.74
6072714	6073714	178	181	2.55E-20	3.79E-08	-31.76
21728413	21729413	600	619	4.91E-29	1.43E-25	-31.77
11420723	11421723	466	481	4.07E-31	1.02E-29	-31.79
9881893	9882893	12	12	1.92E-31	3.51E-30	-31.98
12747506	12748506	513	528	4.31E-38	1.85E-43	-32.00
7691594	7692594	764	789	1.04E-22	1.87E-12	-32.23
21808463	21809463	814	839	2.27E-28	1.12E-23	-32.34
22931715	22932715	524	539	2.25E-30	1.32E-27	-32.42
25638080	25639080	68	69	1.05E-34	3.16E-36	-32.46
3461728	3462728	32	32	1.34E-28	7.21E-24	-32.60
9047331	9048331	250	254	1.34E-28	9.15E-24	-32.70
13494354	13495354	623	644	7.22E-33	3.30E-32	-32.80
8874458	8875458	717	739	8.86E-31	5.76E-28	-32.87
18391869	18392869	226	230	3.25E-44	8.47E-55	-32.91
24976392	24977392	896	925	3.50E-35	1.25E-36	-33.01
16809608	16810608	49	49	8.85E-36	8.37E-38	-33.03
8544039	8545039	961	991	1.58E-23	2.93E-13	-33.07
22723075	22724075	78	80	6.57E-27	5.15E-20	-33.08
24110321	24111321	527	542	4.49E-33	2.69E-32	-33.13
6921753	6922753	77	79	2.33E-31	8.42E-29	-33.19
21282178	21283178	561	579	2.74E-31	1.24E-28	-33.22
21375853	21376853	615	636	5.18E-38	4.86E-42	-33.26
27061093	27062093	338	348	3.80E-26	2.94E-18	-33.31
2463660	2464660	510	525	8.84E-29	1.70E-23	-33.34
20885335	20886335	2	2	7.44E-35	2.12E-35	-33.58
10965972	10966972	854	880	5.76E-28	1.59E-21	-33.68

7949627	7950627	125	128	9.09E-34	4.07E-33	-33.69
7949627	7950627	276	282	9.09E-34	4.07E-33	-33.69
18171306	18172306	17	17	4.31E-31	9.73E-28	-33.72
18622651	18623651	399	413	1.26E-35	8.57E-37	-33.73
22442049	22443049	377	390	8.96E-27	4.37E-19	-33.74
2615244	2616244	946	976	2.73E-30	4.38E-26	-33.77
9407555	9408555	463	478	1.46E-25	1.34E-16	-33.80
4965642	4966642	937	967	7.32E-27	3.77E-19	-33.85
16844523	16845523	546	562	6.14E-26	2.86E-17	-33.88
19613054	19614054	59	60	1.88E-37	2.89E-40	-33.91
18098921	18099921	8	8	6.73E-33	4.31E-31	-33.98
25443917	25444917	683	705	6.93E-31	6.04E-27	-34.10
20571949	20572949	459	474	6.42E-38	5.80E-41	-34.15
12944224	12945224	528	543	5.65E-38	4.74E-41	-34.17
15864339	15865339	468	483	1.96E-28	6.64E-22	-34.24
26304026	26305026	239	243	9.09E-34	1.91E-32	-34.36
23857398	23858398	400	414	8.77E-27	2.09E-18	-34.43
24475655	24476655	456	471	1.08E-31	3.19E-28	-34.43
18393572	18394572	743	768	3.21E-30	3.24E-25	-34.50
9263636	9264636	586	604	6.10E-31	1.45E-26	-34.59
11780947	11781947	109	112	1.58E-25	1.27E-15	-34.71
11780947	11781947	892	920	1.58E-25	1.27E-15	-34.71
11419020	11420020	446	460	6.08E-29	1.95E-22	-34.72
17231998	17232998	721	744	4.21E-26	9.54E-17	-34.73
9793327	9794327	841	867	2.47E-32	3.53E-29	-34.76
19889822	19890822	772	797	2.08E-30	2.68E-25	-34.79
14273562	14274562	836	862	9.02E-32	5.21E-28	-34.81
3160264	3161264	438	452	3.10E-29	6.62E-23	-34.84
11258920	11259920	451	465	1.42E-28	1.59E-21	-34.90
23187193	23188193	1000	1030	3.31E-36	9.28E-37	-34.93
13906525	13907525	73	75	7.77E-37	5.74E-38	-34.98
24306187	24307187	666	688	6.38E-39	4.09E-42	-35.00
24244021	24245021	823	848	1.12E-34	1.28E-33	-35.01
24875052	24876052	333	343	3.43E-40	1.35E-44	-35.06
7961549	7962549	870	898	5.25E-33	4.20E-30	-35.18
15147297	15148297	41	41	1.82E-31	5.30E-27	-35.20
15147297	15148297	848	874	1.82E-31	5.30E-27	-35.20
9554029	9555029	19	19	2.62E-34	1.17E-32	-35.23
22780131	22781131	716	738	1.09E-29	2.09E-23	-35.24
18727397	18728397	445	459	1.32E-33	3.50E-31	-35.30
12731326	12732326	312	320	3.81E-29	3.03E-22	-35.32
9363272	9364272	564	582	1.62E-26	5.55E-17	-35.33
18936037	18937037	381	394	9.65E-36	2.69E-35	-35.46
14554588	14555588	594	612	1.12E-33	4.85E-31	-35.59
4968197	4969197	87	90	4.01E-27	8.36E-18	-35.72

24293414	24294414	670	692	1.69E-32	1.52E-28	-35.72
23940854	23941854	833	858	1.43E-34	2.31E-32	-36.05
24116282	24117282	898	927	2.78E-32	8.95E-28	-36.06
16895619	16896619	401	415	3.58E-32	1.95E-27	-36.18
14194364	14195364	449	463	4.51E-34	3.29E-31	-36.21
16331012	16332012	655	676	3.99E-33	3.36E-29	-36.32
5468934	5469934	5	5	1.89E-28	9.31E-20	-36.42
16187945	16188945	485	500	5.24E-29	9.02E-21	-36.52
18861097	18862097	105	108	2.07E-29	2.07E-21	-36.68
14431107	14432107	948	978	7.97E-29	3.27E-20	-36.71
10512924	10513924	389	403	7.99E-40	3.78E-42	-36.77
17684195	17685195	208	212	9.43E-48	5.59E-58	-36.80
22944489	22945489	39	39	4.80E-38	1.51E-38	-36.82
23538050	23539050	965	995	2.69E-29	5.72E-21	-36.90
26516925	26517925	343	353	3.03E-33	8.17E-29	-36.95
22395211	22396211	590	608	5.48E-28	3.08E-18	-37.01
17037835	17038835	240	244	1.36E-37	2.24E-37	-37.08
19142975	19143975	235	239	2.70E-40	8.77E-43	-37.08
20533627	20534627	858	885	2.16E-33	6.13E-29	-37.12
10227640	10228640	840	866	5.31E-36	3.83E-34	-37.13
17575191	17576191	506	521	7.76E-30	1.05E-21	-37.24
4099572	4100572	876	904	4.73E-35	4.52E-32	-37.31
18161087	18162087	193	196	1.07E-41	2.67E-45	-37.37
18161087	18162087	719	742	1.07E-41	2.67E-45	-37.37
18061451	18062451	755	780	1.93E-29	1.00E-20	-37.43
19269011	19270011	760	785	1.74E-35	8.40E-33	-37.44
10164622	10165622	750	775	4.10E-36	4.82E-34	-37.46
17110221	17111221	608	629	1.76E-30	1.80E-22	-37.76
9669846	9670846	556	574	4.22E-30	1.05E-21	-37.77
26944425	26945425	342	352	2.01E-33	3.14E-28	-37.89
26451352	26452352	801	826	6.04E-37	3.52E-35	-37.98
22707746	22708746	153	156	1.09E-36	1.39E-34	-38.07
5922834	5923834	907	936	1.82E-31	4.15E-24	-38.10
8603651	8604651	536	551	3.31E-30	1.39E-21	-38.10
23207632	23208632	798	823	1.95E-32	4.84E-26	-38.11
18809150	18810150	81	83	1.79E-43	4.46E-48	-38.14
21017332	21018332	94	97	1.96E-36	5.63E-34	-38.16
13736206	13737206	658	679	1.37E-38	3.12E-38	-38.22
13926963	13927963	912	941	3.90E-36	2.75E-33	-38.26
24842691	24843691	970	1000	2.27E-34	9.33E-30	-38.26
14049593	14050593	604	624	3.63E-39	3.35E-39	-38.41
4011006	4012006	316	324	1.50E-32	6.47E-26	-38.46
16854742	16855742	232	236	3.52E-35	5.14E-31	-38.62
9767779	9768779	391	405	1.66E-41	1.18E-43	-38.63
16207531	16208531	498	513	3.97E-28	6.91E-17	-38.64

25922512	25923512	887	915	6.27E-36	2.63E-32	-38.83
8482725	8483725	6	6	4.01E-34	1.24E-28	-38.89
13586326	13587326	147	150	1.16E-32	1.05E-25	-38.89
13567591	13568591	708	730	8.34E-37	5.63E-34	-38.91
16778951	16779951	227	231	2.05E-37	6.86E-35	-39.21
15973343	15974343	639	660	1.00E-34	1.73E-29	-39.24
17914125	17915125	724	747	6.73E-40	8.13E-40	-39.25
22616626	22617626	628	649	2.13E-33	8.61E-27	-39.28
15992930	15993930	51	51	1.04E-32	2.09E-25	-39.28
9219353	9220353	589	607	3.04E-32	2.80E-24	-39.48
9278113	9279113	701	723	1.20E-32	4.41E-25	-39.49
4713571	4714571	318	326	6.27E-36	1.66E-31	-39.63
18789563	18790563	734	759	6.24E-33	1.91E-25	-39.69
14584394	14585394	918	947	3.20E-38	5.29E-36	-39.71
10980449	10981449	310	318	6.80E-47	2.57E-53	-39.74
6247291	6248291	90	93	7.28E-39	3.01E-37	-39.75
15172845	15173845	34	34	1.02E-36	6.18E-33	-39.77
26448797	26449797	525	540	3.01E-38	5.78E-36	-39.81
18998204	18999204	667	689	1.45E-47	1.52E-54	-39.86
22172094	22173094	802	827	2.58E-34	5.76E-28	-39.94
18564743	18565743	432	446	1.43E-34	1.85E-28	-39.95
21817831	21818831	387	401	4.78E-32	2.07E-23	-39.96
23428194	23429194	418	432	1.83E-38	3.57E-36	-40.03
21369041	21370041	61	62	4.46E-47	2.81E-53	-40.15
26156701	26157701	166	169	2.98E-33	1.27E-25	-40.15
21246411	21247411	690	712	3.69E-36	2.43E-31	-40.25
18562188	18563188	359	370	2.26E-38	1.62E-35	-40.50
18565594	18566594	509	524	7.21E-37	2.69E-32	-40.71
13943995	13944995	241	245	6.97E-39	2.76E-36	-40.75
16061909	16062909	661	683	6.71E-24	2.89E-06	-40.81
18884942	18885942	439	453	2.20E-36	4.43E-31	-40.96
10575090	10576090	270	275	2.49E-37	6.73E-33	-41.03
6971997	6972997	805	830	9.81E-36	1.18E-29	-41.09
15997188	15998188	478	493	9.47E-45	1.25E-47	-41.14
19917925	19918925	116	119	2.08E-49	6.43E-57	-41.17
14288891	14289891	185	188	6.90E-37	8.44E-32	-41.25
14288891	14289891	275	281	6.90E-37	8.44E-32	-41.25
21584494	21585494	537	552	2.52E-44	1.31E-46	-41.32
22299833	22300833	940	970	6.29E-38	9.09E-34	-41.36
9830797	9831797	910	939	1.32E-38	5.94E-35	-41.53
16115559	16116559	704	726	7.75E-41	2.07E-39	-41.54
10512072	10513072	131	134	1.57E-42	9.46E-43	-41.58
26996372	26997372	779	804	1.44E-38	1.05E-34	-41.70
18572407	18573407	291	298	4.03E-34	9.21E-26	-41.75
14615903	14616903	330	339	7.59E-44	5.51E-45	-41.98

22139733	22140733	372	384	7.67E-35	5.77E-27	-41.99
21185948	21186948	441	455	1.68E-53	4.11E-64	-42.16
11486296	11487296	522	537	3.37E-41	1.74E-39	-42.18
14865419	14866419	107	110	6.62E-35	8.59E-27	-42.29
8418003	8419003	9	9	2.10E-35	9.73E-28	-42.34
22160171	22161171	822	847	3.16E-30	2.37E-17	-42.38
9078840	9079840	678	700	4.22E-32	4.25E-21	-42.38
16032103	16033103	726	750	1.84E-40	9.53E-38	-42.45
23386466	23387466	781	806	1.75E-42	9.29E-42	-42.48
24162268	24163268	702	724	3.33E-52	4.42E-61	-42.60
18086999	18087999	384	397	9.21E-32	4.59E-20	-42.73
15879668	15880668	490	505	1.79E-37	1.82E-31	-42.75
27018513	27019513	309	317	3.52E-35	8.73E-27	-42.85
16595006	16596006	566	584	3.25E-33	7.57E-23	-42.86
23282572	23283572	348	358	8.44E-42	6.39E-40	-42.95
26859265	26860265	388	402	1.13E-42	1.27E-41	-43.00
13789005	13790005	889	917	5.77E-40	3.34E-36	-43.00
24602542	24603542	746	771	5.92E-42	5.23E-40	-43.17
26492229	26493229	650	671	8.88E-44	1.22E-43	-43.19
7700962	7701962	42	42	2.75E-30	1.28E-16	-43.23
16867516	16868516	662	684	1.19E-45	2.58E-47	-43.26
14007865	14008865	475	490	8.04E-39	1.22E-33	-43.28
11285319	11286319	35	35	4.75E-38	5.09E-32	-43.35
6888541	6889541	797	822	2.22E-33	1.56E-22	-43.50
21264295	21265295	458	473	3.59E-47	6.12E-50	-43.68
12575484	12576484	481	496	3.60E-34	6.65E-24	-43.71
16639289	16640289	179	182	2.04E-42	5.23E-40	-44.10
26347458	26348458	914	943	8.76E-45	1.13E-44	-44.17
23372841	23373841	164	167	5.74E-42	6.23E-39	-44.28
11869513	11870513	909	938	1.96E-36	7.56E-28	-44.29
13261017	13262017	448	462	7.71E-43	1.22E-40	-44.31
19178742	19179742	534	549	1.62E-42	6.28E-40	-44.38
17646725	17647725	782	807	2.26E-62	1.25E-79	-44.39
25644893	25645893	830	855	3.76E-42	3.85E-39	-44.43
5709084	5710084	65	66	7.27E-40	1.83E-34	-44.54
8646231	8647231	390	404	5.47E-37	2.40E-28	-44.90
20709056	20710056	881	909	4.48E-38	1.77E-30	-44.95
23394131	23395131	313	321	5.35E-38	2.89E-30	-45.00
26154146	26155146	277	283	2.53E-43	8.19E-41	-45.11
25766671	25767671	689	711	3.86E-50	2.11E-54	-45.15
25429440	25430440	901	930	3.66E-49	1.91E-52	-45.15
15969085	15970085	233	237	3.33E-53	1.70E-60	-45.18
7654124	7655124	685	707	4.06E-40	2.53E-34	-45.19
10845897	10846897	584	602	1.56E-43	5.61E-41	-45.36
21122930	21123930	40	40	7.26E-51	1.58E-55	-45.48

23843772	23844772	900	929	1.02E-45	3.38E-45	-45.52
23326855	23327855	627	648	2.95E-36	4.21E-26	-45.68
22953857	22954857	749	774	3.50E-41	7.43E-36	-45.78
17018248	17019248	596	615	1.48E-42	1.41E-38	-45.81
18160236	18161236	882	910	5.23E-47	2.04E-47	-45.87
26390889	26391889	328	337	5.26E-37	2.15E-27	-45.89
5205792	5206792	616	637	5.24E-36	2.41E-25	-45.94
15831127	15832127	460	475	1.71E-39	2.69E-32	-45.96
15200096	15201096	488	503	4.55E-50	2.14E-53	-46.01
7129542	7130542	37	37	1.73E-31	3.25E-16	-46.04
25552069	25553069	711	733	7.62E-38	6.64E-29	-46.06
14033412	14034412	72	74	1.80E-43	4.00E-40	-46.09
17802566	17803566	259	264	8.38E-44	1.08E-40	-46.19
13375130	13376130	664	686	3.28E-42	2.30E-37	-46.33
26241860	26242860	207	211	2.81E-40	1.74E-33	-46.34
23719439	23720439	925	954	8.20E-36	1.74E-24	-46.41
16495370	16496370	649	670	8.73E-40	2.16E-32	-46.45
24723468	24724468	619	640	3.18E-40	3.63E-33	-46.55
13980614	13981614	332	342	1.03E-47	4.27E-48	-46.60
9782257	9783257	447	461	6.86E-49	2.21E-50	-46.67
26459868	26460868	732	756	6.45E-41	3.60E-34	-46.94
23424788	23425788	499	514	6.71E-42	5.78E-36	-47.11
21420988	21421988	544	560	1.13E-42	1.79E-37	-47.15
21316242	21317242	26	26	2.43E-46	1.68E-44	-47.45
7133800	7134800	177	180	6.89E-37	1.43E-25	-47.48
27188832	27189832	531	546	5.08E-41	8.05E-34	-47.49
14007013	14008013	355	366	1.98E-43	1.41E-38	-47.55
14380011	14381011	299	307	1.76E-46	1.89E-44	-47.78
14380011	14381011	680	702	1.76E-46	1.89E-44	-47.78
6573451	6574451	491	506	6.10E-42	5.02E-35	-48.13
22166984	22167984	981	1011	1.07E-36	1.55E-24	-48.13
12319154	12320154	452	466	3.20E-49	1.67E-49	-48.21
23246805	23247805	988	1018	2.50E-42	1.32E-35	-48.33
15215424	15216424	592	610	3.25E-44	2.38E-39	-48.35
16894767	16895767	864	892	1.21E-43	4.31E-38	-48.47
9619602	9620602	97	100	2.37E-41	1.72E-33	-48.49
23475884	23476884	212	216	1.38E-47	8.32E-46	-48.64
22729036	22730036	747	772	9.61E-41	4.17E-32	-48.65
2352101	2353101	162	165	7.83E-32	2.85E-14	-48.67
24613613	24614613	203	206	1.53E-44	1.92E-39	-48.91
27208419	27209419	874	902	2.48E-39	6.54E-29	-49.03
5041434	5042434	547	564	1.23E-38	3.00E-27	-49.30
10552949	10553949	700	722	1.47E-49	5.90E-49	-49.43
25549514	25550514	213	217	6.07E-43	1.10E-35	-49.48
25549514	25550514	839	865	6.07E-43	1.10E-35	-49.48

26382373	26383373	588	606	1.68E-42	8.94E-35	-49.50
8090140	8091140	767	792	2.09E-39	1.44E-28	-49.52
17794902	17795902	256	261	1.50E-53	8.05E-57	-49.56
20400779	20401779	659	680	6.34E-47	2.18E-43	-49.73
25833947	25834947	308	316	7.63E-55	3.91E-59	-49.83
16300355	16301355	789	814	6.70E-38	3.91E-25	-49.94
25813508	25814508	3	3	7.90E-45	7.28E-39	-50.07
26423249	26424249	976	1006	4.06E-40	1.99E-29	-50.08
18899419	18900419	867	895	1.21E-38	2.22E-26	-50.18
27097712	27098712	297	305	8.26E-44	1.16E-36	-50.23
14278671	14279671	738	763	8.83E-40	1.60E-28	-50.31
11515250	11516250	132	135	2.40E-43	1.58E-35	-50.44
9600015	9601015	654	675	8.35E-51	1.98E-50	-50.45
22430126	22431126	474	489	1.45E-32	7.32E-14	-50.54
23180381	23181381	367	379	9.98E-41	7.60E-30	-50.88
19767193	19768193	875	903	8.58E-52	7.90E-52	-51.03
16488557	16489557	138	141	4.06E-45	2.07E-38	-51.10
16980778	16981778	280	287	1.01E-51	2.17E-51	-51.33
8890638	8891638	538	553	1.21E-43	3.14E-35	-51.33
8102062	8103062	819	844	1.97E-45	1.20E-38	-51.49
21183393	21184393	587	605	1.51E-48	8.58E-45	-51.58
14718945	14719945	435	449	6.29E-46	3.93E-39	-52.00
22088637	22089637	713	735	2.46E-48	6.58E-44	-52.04
19076550	19077550	729	753	6.01E-46	4.19E-39	-52.06
24251685	24252685	353	364	3.96E-43	2.25E-33	-52.16
5693755	5694755	502	517	4.52E-43	3.74E-33	-52.26
22031581	22032581	857	884	3.76E-50	3.88E-47	-52.44
9392227	9393227	139	142	2.51E-45	3.05E-37	-52.68
17742955	17743955	978	1008	1.43E-51	1.33E-49	-52.81
26691501	26692501	845	871	9.76E-45	7.64E-36	-52.90
20520002	20521002	476	491	2.24E-50	6.19E-47	-53.09
15445355	15446355	469	484	1.44E-52	5.26E-51	-53.40
23800341	23801341	791	816	2.51E-39	2.80E-24	-53.65
22009439	22010439	763	788	4.83E-47	1.07E-39	-53.66
24443294	24444294	385	399	7.28E-49	5.05E-43	-53.98
20864897	20865897	121	124	1.36E-48	3.64E-42	-54.30
21256630	21257630	831	856	4.74E-44	5.04E-33	-54.35
11442013	11443013	796	821	1.62E-45	8.41E-36	-54.51
26628484	26629484	363	374	7.95E-43	2.28E-30	-54.56
23222109	23223109	479	494	1.17E-48	5.31E-42	-54.59
13874164	13875164	315	323	1.94E-50	1.69E-45	-54.65
13564184	13565184	335	345	5.07E-46	1.51E-36	-54.77
21402253	21403253	567	585	3.41E-45	7.27E-35	-54.80
13116246	13117246	223	227	3.75E-48	5.40E-40	-55.59
25297443	25298443	200	203	1.57E-46	1.23E-36	-55.70

27130924	27131924	735	760	1.41E-43	1.31E-30	-55.82
13430484	13431484	706	728	1.11E-45	1.32E-34	-56.03
10579348	10580348	464	479	1.34E-56	3.18E-56	-56.25
12681933	12682933	674	696	5.55E-49	1.31E-40	-56.63
23514205	23515205	884	912	3.69E-45	8.10E-33	-56.77
22448010	22449010	472	487	3.86E-46	9.30E-35	-56.79
13975504	13976504	96	99	4.55E-50	1.54E-42	-56.87
8357540	8358540	83	85	1.28E-48	2.89E-39	-57.25
10935314	10936314	103	106	6.92E-49	1.22E-39	-57.41
13307003	13308003	768	793	2.18E-56	1.60E-54	-57.53
21388627	21389627	815	840	1.58E-53	1.32E-48	-57.72
25137343	25138343	859	886	5.07E-48	1.48E-37	-57.76
9778850	9779850	512	527	3.16E-51	5.96E-44	-57.78
21724155	21725155	692	714	2.79E-46	1.19E-33	-58.19
9605125	9606125	518	533	2.65E-49	2.86E-39	-58.61
11417316	11418316	69	71	3.29E-49	5.90E-39	-58.74
26134559	26135559	806	831	7.12E-56	4.03E-52	-58.90
10833975	10834975	808	833	9.08E-51	6.91E-42	-58.92
13869906	13870906	194	197	1.57E-54	4.67E-49	-59.28
15641221	15642221	196	199	2.67E-44	1.36E-28	-59.28
23179529	23180529	168	171	1.24E-49	1.07E-38	-59.84
13867352	13868352	684	706	1.24E-55	1.11E-50	-59.86
25456691	25457691	440	454	3.60E-53	1.13E-45	-59.94
23377950	23378950	641	662	1.16E-53	1.95E-46	-60.16
14730016	14731016	344	354	5.08E-59	4.91E-57	-60.28
14730016	14731016	598	617	5.08E-59	4.91E-57	-60.28
19085066	19086066	278	285	1.91E-52	1.26E-43	-60.54
25725794	25726794	514	529	7.45E-56	3.64E-50	-60.82
22690714	22691714	938	968	2.37E-53	9.36E-45	-61.22
19211102	19212102	368	380	1.76E-61	7.56E-61	-61.39
24204848	24205848	816	841	4.54E-56	5.85E-50	-61.45
25718130	25719130	927	956	3.39E-51	6.39E-40	-61.75
23971511	23972511	149	152	7.04E-61	7.86E-59	-62.20
25477980	25478980	33	33	1.13E-52	2.17E-42	-62.23
26468384	26469384	575	593	9.37E-58	1.75E-52	-62.30
17563268	17564268	591	609	4.99E-66	5.24E-69	-62.32
9496973	9497973	826	851	5.58E-57	9.27E-51	-62.47
23627467	23628467	953	983	1.51E-55	1.41E-47	-62.79
21628777	21629777	950	980	2.20E-62	3.76E-61	-62.89
10418397	10419397	991	1021	6.30E-54	3.74E-44	-62.98
15454722	15455722	993	1023	5.04E-56	6.90E-48	-63.43
13691923	13692923	158	161	1.64E-60	8.81E-57	-63.52
22452268	22453268	66	67	3.85E-49	7.92E-34	-63.73
13306151	13307151	325	334	1.29E-50	1.27E-36	-63.88
10080314	10081314	850	876	6.25E-57	1.80E-48	-64.66

24083070	24084070	824	849	1.80E-54	2.56E-43	-64.90
10545285	10546285	480	495	3.60E-58	1.06E-50	-64.91
23436710	23437710	295	303	1.70E-58	2.50E-51	-64.94
19231540	19232540	1	1	3.89E-64	1.58E-62	-65.02
8396714	8397714	428	442	6.38E-56	5.01E-46	-65.09
13343621	13344621	921	950	6.13E-49	7.76E-32	-65.32
26658289	26659289	572	590	1.16E-50	2.93E-35	-65.34
14915663	14916663	813	838	1.74E-50	7.98E-35	-65.42
9964498	9965498	347	357	2.26E-53	1.67E-40	-65.52
25789664	25790664	477	492	1.58E-59	2.34E-52	-65.97
16704010	16705010	157	160	1.96E-58	6.87E-50	-66.25
8154009	8155009	890	918	3.95E-56	4.23E-45	-66.43
26815834	26816834	47	47	9.28E-58	1.57E-47	-67.26
13889493	13890493	175	178	2.93E-63	1.77E-58	-67.31
13889493	13890493	283	290	2.93E-63	1.77E-58	-67.31
10395404	10396404	585	603	1.25E-56	1.69E-44	-68.04
16018477	16019477	429	443	1.68E-66	2.10E-63	-68.87
24760938	24761938	495	510	2.74E-61	9.59E-53	-69.11
9902331	9903331	30	30	9.35E-59	3.13E-47	-69.55
10973636	10974636	201	204	2.19E-63	2.46E-56	-69.71
15447910	15448910	126	129	1.70E-69	3.92E-68	-70.13
12251027	12252027	238	242	7.01E-57	5.29E-42	-71.03
22769061	22770061	117	120	1.19E-57	1.68E-43	-71.07
16868368	16869368	500	515	5.06E-70	7.75E-68	-71.48
15611416	15612416	563	581	5.78E-60	2.14E-47	-71.81
23952776	23953776	552	569	4.99E-64	6.98E-55	-72.45
14286336	14287336	897	926	3.83E-72	6.95E-71	-72.68
9753302	9754302	891	919	2.60E-60	1.57E-46	-73.37
24800963	24801963	470	485	7.26E-62	2.02E-49	-73.58
25841611	25842611	780	805	8.03E-71	4.55E-67	-73.85
22901909	22902909	345	355	7.35E-62	2.28E-48	-74.62
26731526	26732526	872	900	2.18E-65	4.02E-55	-74.93
25992343	25993343	386	400	1.06E-67	1.16E-59	-75.01
25948912	25949912	84	86	5.92E-72	2.14E-67	-75.79
26387482	26388482	642	663	3.72E-68	1.98E-58	-77.15
11424129	11425129	638	659	5.45E-83	4.93E-83	-82.22
9970459	9971459	271	276	5.25E-70	1.77E-54	-84.81
14037670	14038670	507	522	3.86E-75	1.07E-63	-85.86
14119423	14120423	471	486	2.84E-70	9.41E-54	-86.07
23477587	23478587	987	1017	5.83E-73	2.52E-55	-89.87
26332129	26333129	990	1020	5.48E-75	1.13E-58	-90.58
14718094	14719094	195	198	2.25E-78	7.03E-64	-92.14
26394295	26395295	71	73	5.00E-79	2.11E-62	-94.93
26676173	26677173	601	621	1.04E-77	3.14E-59	-95.46
14197770	14198770	957	987	8.21E-97	7.50E-97	-96.05

22742661	22743661	409	423	1.50E-84	2.70E-72	-96.08
23656421	23657421	425	439	2.61E-82	4.71E-66	-97.84
25158633	25159633	473	488	2.85E-90	6.59E-81	-98.91
22445455	22446455	183	186	2.67E-69	1.12E-38	-99.19
14988049	14989049	142	145	4.00E-95	1.96E-89	-100.09
14618457	14619457	964	994	2.07E-97	1.11E-93	-100.41
25512044	25513044	578	596	1.80E-96	5.92E-88	-104.26
2646753	2647753	234	238	1.10E-100	2.72E-96	-104.35
26338942	26339942	219	223	6.98E-84	4.21E-61	-105.94
13442406	13443406	635	656	6.97E-103	7.63E-94	-111.20
9898925	9899925	266	271	6.15E-98	3.68E-77	-117.99
16130888	16131888	799	824	3.66E-175	9.74E-130	-219.86
1578003	1579003	804	829	NA	1.00E+00	NA
11642989	11643989	786	811	NA	NA	NA
192460	193460	860	887	NA	NA	NA
364482	365482	64	65	NA	NA	NA

Annex Table S1.5 Populations used to estimate the frequency of *FBti0019386.*

	Population	Köppen-Geiger climate classification / Latitude	Pool/ individual strain	Accession number	Reference
North America	Bowdoinham, ME	Cold climate, no dry season and warm summer / 45.5	50-100 pooled individuals	SRX661844-5	Bergland et al., 2014
	Linvilla, PA	Cold climate, no dry season and hot summer / 40	50-100 pooled individuals	SRX661837-43	Bergland et al., 2014
	Winters, CA	Temperate climate, dry winter and hot summer / 38.6	35 individual strains	SRP009033	Campo et al., 2013
	Raleigh, NC	Temperate climate, no dry season and hot summer / 35.5	141 individuals	SRX[a]	Huang et al., 2014
	Eutawville, SC	Temperate climate, no dry season and hot summer / 33	50-100 pooled individuals	SRX661835	Bergland et al., 2014
	Hahira, GA	Temperate climate, no dry season and hot summer / 30.9	50-100 pooled individuals	SRX661834	Bergland et al., 2014
	Homestead, FL	Tropical climate, monsoon / 25.5	50-100 pooled individuals	SRX661832-3	Bergland et al., 2014
Australia	Cairns	Tropical climate, monsoon / -16.88	Pool	SRR1177951	Reinhardt et al., 2014
	Innisfail	Tropical climate, monsoon / -17.52	-	-	González et al., 2010
	Cardwell	Tropical climate, savannah / - 18.25	Pool	SRR1177952	Reinhardt et al., 2014
	Redland Bay	Temperate climate, no dry season and hot summer / -27.48	-	-	González et al., 2010
	Coffs Harbour	Temperate climate, no dry season and hot summer / -30.32	-	-	González et al., 2010
	Melbourne	Temperate climate, no dry season and warm summer / -37.82	-	-	González et al., 2010
	Miller's Orchard, north Tasmania	Temperate climate, no dry season and warm summer / -41.53	Pool	SRR1177953	Reinhardt et al., 2014
	Sorell, south Tasmania	Temperate climate, no dry season and warm summer / -42.83	Pool	SRR1177955	Reinhardt et al., 2014
Europe	Stockholm, Sweden	Cold climate, no dry season and warm summer / 59.33	27 individual strains	-	This work
	Vienna, Austria	Cold climate, no dry season and warm summer / 48.25	Pool	ERR173232, ERR173238	Kofler et al., 2012
	Lyon, France	Temperate climate, no dry season and warm summer / 45.7	8 individual strains?	SRX058182-SRX058190	Pool et al., 2012
	Bolzano, Italy	Cold climate, no dry season and warm summer / 45.62	Pool	ERR173233, ERR173239	Kofler et al., 2012
	Povoa de Varzim, Portugal	Temperate climate, dry and warm summer / 41.23	Pool	SRR188217, SRR189066	Bastide et al., 2013
	Bari, Italy	Temperate climate, no dry season and hot summer / 41.13	16 individual strains	-	This work
Africa	Rwanda	Temperate climate, dry winter and warm summer / 2	22 individual strains	SRX058338-39, 41-57, 59, 62, 67, 69, 71	Pool et al., 2012

[a]Accession numbers for DGRP strains can be found in Supplemental Data File S1 in Huang et al 2014.

Annex Table S1.6. (**A**) Pearson product-moment correlations between the frequency of the TE and geographical and climatic variables for each one of the three continents analyzed. Significant correlations are in bold and highlighted in gray. Data for each one of the geographic and climatic variables is detailed in (**B**).

(A)

	USA		AUST		EUR	
%TE vs	r	p	r	p	r	p
Latitude	**0.87**	**p=0.011**	**0.91**	**p=0.002**	-0.50	p=0.313
Longitude	0.65	p=0.117	0.12	p=0.776	-0.68	p=0.139
Elevation	-0.27	p=0.556	**0.74**	**p=0.036**	0.18	p=0.735
AvMonTemp	**-0.75**	**p=0.050**	**-0.93**	**p=0.001**	0.37	p=0.470
ThermalAmp	0.2	p=0.664	0.55	p=0.160	-0.48	p=0.337
HotMonth	**-0.88**	**p=0.009**	**-0.95**	**p=0.000**	0.01	p=0.984
ColdMonth	-0.59	p=0.163	**-0.91**	**p=0.002**	0.38	p=0.458
SummerSEASON	**-0.87**	**p=0.012**	**-0.95**	**p=0.000**	0.29	p=0.573
WinterSEASON	-0.62	p=0.134	**-0.92**	**p=0.001**	0.40	p=0.434
Monthabove10	-0.56	p=0.185	**-0.94**	**p=0.000**	0.25	p=0.637
MAP	0.31	p=0.491	**0.84**	**p=0.009**	-0.66	p=0.156
Cv	0.16	p=0.734	-0.7	p=0.056	0.40	p=0.438
DryMonth	-0.48	p=0.278	-0.31	p=0.453	-0.12	p=0.817
Summer_P	-0.54	p=0.209	**-0.88**	**p=0.004**	0.15	p=0.773
Summer_DryM	-0.64	p=0.119	-0.58	p=0.135	-0.26	p=0.620
Summer_wetM	-0.63	p=0.127	**-0.85**	**p=0.007**	0.48	p=0.334
Winter_P	-0.54	p=0.209	-0.57	p=0.139	0.15	p=0.773
Winter_DryM	-0.64	p=0.119	-0.11	p=0.793	-0.26	p=0.620
Winter_wetM	-0.63	p=0.127	**-0.74**	**p=0.034**	0.48	p=0.334

(B) Climate data (from Peel et al., 2007)

LEGENDS:

Latitude	Latitude (decimal degrees)
Longitude	Longitude (decimal degrees)
Elevation	Elevation (m)
Temperature	
Ave. Mon. Temp.	The average annual temperature (Deg. C) (the average of the monthly averages)
Thermal Amplitude	HotMonth - ColdMonth
HotMonth	The hottest monthly average (Deg. C)
ColdMonth	The coldest monthly average (Deg. C)
SeasonA	Average of the monthly average temperatures (Deg. C) for Oct, Nov, Dec, Jan, Feb & Mar
SeasonB	Average of the monthly average temperatures (Deg. C) for Apr, May, Jun, Jul, Aug & Sep
MonthsAbove10	The number of monthly average temperatures greater than or equal to 10 Deg. C
Precipitation	
MAP	Mean annual precipitation for complete calendar years (mm)
Cv	Coefficient of variation of annual precipitation for complete calendar years
DryMonth	Monthly average precipitation (mm) for the driest month
ONDJFM_P*	Summation of the monthly average precipitation (mm) for Oct, Nov, Dec, Jan, Feb & Mar
ONDJFM_DryM*	Driest monthly average precipitation (mm) for the period Oct, Nov, Dec, Jan, Feb & Mar
ONDJFM_WetM*	Wettest monthly average precipitation (mm) for the period Oct, Nov, Dec, Jan, Feb & Mar
AMJJAS_P*	Summation of the monthly average precipitation (mm) for Apr, May, Jun, Jul, Aug & Sep
AMJJAS_DryM*	Driest monthly average precipitation (mm) for Apr, May, Jun, Jul, Aug & Sep
AMJJAS_WetM*	Wettest monthly average precipitation (mm) for Apr, May, Jun, Jul, Aug & Sep

*Oct, Nov, Dec, Jan, Feb & Mar were considered summer for AUST and winter for USA and EUR
*Apr, May, Jun, Jul, Aug & Sep were considered summer for USA and EUR and winter for AUST

Cont	pop	*FBti0019386* frequency	Latitude	Longitude	Elevation	AvMon Temp	Thermal Amplitude	Hot Month	Cold Month	SeasonA	SeasonB	Months Above10	MAP	Cv	Dry Month	ONDJFM _P	ONDJFM _DryM	ONDJFM _WetM	AMJJAS _P	AMJJAS _DryM	AMJJAS _WetM
USA	FL	46.88	25.5	-80.28	4	24.29	8.37	28.16	19.79	21.86	26.71	12	1475.09	0.244	50.06	499.15	50.06	206.81	987.58	81.84	225.32
USA	GA	43.06	30.9	-83.65	110	18.11	18.88	27.39	8.51	12.05	24.18	9	1140.95	0.166	57.46	572.03	57.46	127.13	575.78	78.10	125.10
USA	SC	61.67	33	-89.83	47	19.18	18.39	27.87	9.48	13.41	24.96	11	1133.07	0.182	70.56	544.61	70.56	110.60	587.73	83.22	118.23
USA	NC	59.40	35.5	-78.78	134	15.52	20.74	25.82	5.08	9.01	22.03	9	1149.69	0.157	68.48	503.57	68.48	99.16	646.30	83.32	132.44
USA	CA	66.30	38.6	-121.5	6	15.87	15.84	23.59	7.75	11.30	20.43	10	445.64	0.344	0.56	382.14	20.67	89.97	61.26	0.56	34.90
USA	PA	64.34	40	-75.2	9	12.49	24.77	24.88	0.11	5.06	19.93	7	1066.56	0.153	74.67	491.57	74.67	92.01	573.87	85.50	107.42
USA	ME	81.25	45.5	-122.6	12	11.85	15.69	19.79	4.10	7.30	16.41	7	1042.93	0.207	13.98	799.70	83.64	172.05	242.62	13.98	70.23
AUST	Cairns	15.38	-16.88	145.75	7	25.00	6.58	27.83	21.25	27.02	22.98	12	1990.58	0.288	25.78	1578.78	38.35	449.41	426.85	25.78	199.91
AUST	Innisfail	13	-17.52	146.02	4	22.97	7.78	26.46	18.68	25.29	20.64	12	3565.8	0.223	80.04	2262.50	80.04	673.34	1288.27	86.45	469.50
AUST	Cardwell	22.22	-18.25	146.02	7	23.74	8.03	27.16	19.13	26.19	21.30	12	2131.84	0.284	29.47	1674.83	52.02	464.08	452.90	29.47	212.15
AUST	Redland Bay	20	-27.48	153.03	38	20.24	9.78	24.62	14.84	23.24	17.25	12	1283.24	0.289	46.35	816.39	81.27	174.28	472.33	46.35	115.64
AUST	Coffs Harbour	30	-30.32	153.12	5	18.49	10.02	23.02	13.01	21.45	15.54	12	1697.47	0.282	60.71	1004.30	89.43	241.34	667.50	60.71	180.75
AUST	Melbourne	46	-37.82	144.97	113	15.49	10.37	20.52	10.15	18.54	12.44	12	656.86	0.192	47.11	333.97	47.11	67.80	324.01	48.71	59.21
AUST	Millers Orchard	100	-41.53	147.2	178	10.98	10.05	16.05	6.00	13.97	8.00	7	684.16	0.198	38.37	285.56	38.37	61.95	400.30	54.09	78.58
AUST	Sorell	76.92	-42.83	147.48	27	12.64	9.35	17.17	7.83	15.36	9.91	9	504.63	0.212	31.44	261.48	35.30	54.28	246.40	31.44	47.09
EUR	IT_bari	56.25	41.13	16.78	49	15.90	15.51	24.01	8.50	11.43	20.37	10	579.56	0.252	24.20	356.21	50.06	68.73	214.98	24.20	57.09
EUR	PORT	75.17	41.23	-8.68	77	14.23	11.18	19.93	8.75	10.93	17.53	9	1280.28	0.276	20.88	901.97	128.92	167.61	365.02	20.88	112.27
EUR	IT_bolzano	55.95	45.62	8.73	211	10.97	21.17	21.44	0.27	4.61	17.32	7	721.02	0.206	25.07	251.88	25.07	65.50	464.85	50.65	96.64
EUR	FR	52.5	45.7	4.7	201	11.08	17.75	20.28	2.53	5.74	16.42	7	735.43	0.198	37.35	314.96	37.35	83.25	424.73	56.47	77.54
EUR	AUST	71.89	48.25	16.37	209	9.47	21.06	19.61	-1.45	2.95	15.98	5	643.03	0.17	37.41	263.93	37.41	48.90	379.35	48.84	75.01
EUR	SW	48.15	59.33	18.05	52	6.04	21.14	17.49	-3.66	-0.18	12.27	4	555.76	0.171	28.57	244.64	28.57	51.01	308.92	33.21	72.95

Annex Table S1.7. Pearson's product-moment correlation coefficients among the different geographical and climatic variables in the three continents. Significant correlations are in **bold**.

USA

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
latitude (1)																		
longitude (2)	0.56																	
elevation(3)	-0.14	-0.31																
AvMonTemp(4)	**-0.94**	-0.32	-0.09															
thermalAmp(5)	0.39	-0.31	0.48	-0.69														
HotMonth(6)	**-0.95**	-0.71	0.29	**0.79**	-0.11													
ColdMonth(7)	**-0.79**	-0.09	-0.26	**0.95**	**-0.88**	0.57												
summerSEASON (8)	**-1.00**	-0.56	0.17	**0.93**	-0.37	**0.96**	**0.77**											
winterSEASON(9)	**-0.83**	-0.14	-0.24	**0.97**	**-0.84**	0.63	**1.00**	**0.81**										
monthabove10(10)	**-0.86**	-0.11	-0.11	**0.93**	-0.66	0.71	**0.89**	**0.85**	**0.91**									
MAP(11)	0.31	0.65	-0.33	-0.25	-0.01	-0.37	-0.17	-0.34	-0.17	0.02								
Cv(12)	0.02	0.67	-0.63	0.19	-0.58	-0.23	0.36	-0.07	0.34	0.39	**0.82**							
DryMonth(13)	-0.47	**-0.91**	0.50	0.20	0.46	0.67	-0.06	0.50	0.00	0.06	-0.72	**-0.82**						
summer_P(14)	-0.55	**-0.84**	0.33	0.43	0.03	0.62	0.28	0.57	0.30	0.17	**-0.94**	**-0.76**	**0.85**					
summer_DryM(15)	-0.64	**-0.95**	0.43	0.41	0.26	**0.78**	0.17	0.66	0.22	0.22	**-0.76**	-0.74	**0.97**	**0.92**				
summer_wetM(16)	-0.66	**-0.80**	0.23	0.59	-0.18	0.68	0.47	0.67	0.49	0.34	**-0.89**	-0.61	0.75	**0.97**	**0.87**			
winter_P(17)	-0.55	**-0.84**	0.33	0.43	0.03	0.62	0.28	0.57	0.30	0.17	**-0.94**	**-0.76**	**0.85**	**1.00**	**0.92**	**0.97**		
winter_DryM(18)	-0.64	**-0.95**	0.43	0.41	0.26	**0.78**	0.17	0.66	0.22	0.22	**-0.76**	-0.74	**0.97**	**0.92**	**1.00**	**0.87**	**0.92**	
winter_wetM(19)	-0.66	**-0.80**	0.23	0.59	-0.18	0.68	0.47	0.67	0.49	0.34	**-0.89**	-0.61	0.75	**0.97**	**0.87**	**1.00**	**0.97**	**0.87**

AUST

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
latitude (1)																		
longitude (2)	-0.03																	
elevation(3)	**0.77**	0.10																
AvMonTemp(4)	**-0.99**	0.06	**-0.78**															
thermalAmp(5)	**0.75**	-0.44	0.67	**-0.78**														
HotMonth(6)	**-0.98**	0.02	**-0.76**	**1.00**	**-0.73**													
ColdMonth(7)	**-0.98**	0.12	**-0.78**	**1.00**	**-0.83**	**0.99**												
summerSEASON (8)	**-0.99**	0.03	**-0.77**	**1.00**	**-0.74**	**1.00**	**0.99**											
winterSEASON(9)	**-0.98**	0.09	**-0.78**	**1.00**	**-0.81**	**0.99**	**1.00**	**1.00**										
monthabove10(10)	**-0.75**	-0.11	-0.61	**0.80**	-0.34	**0.84**	**0.76**	**0.82**	**0.78**									
MAP(11)	**0.95**	0.14	**0.77**	**-0.89**	0.60	**-0.90**	**-0.88**	**-0.90**	**-0.88**	-0.68								
Cv(12)	**-0.71**	-0.50	-0.62	**0.73**	-0.42	**0.74**	0.70	**0.74**	**0.72**	0.58	-0.70							
DryMonth(13)	-0.21	-0.23	-0.30	0.11	0.16	0.15	0.08	0.13	0.10	0.30	-0.37	-0.16						
summer_P(14)	**-0.98**	-0.01	**-0.86**	**0.95**	**-0.73**	**0.94**	**0.94**	**0.95**	**0.95**	0.70	**-0.97**	0.67	0.32					
summer_DryM(15)	-0.43	-0.69	-0.46	0.36	0.16	0.41	0.30	0.39	0.33	0.55	-0.60	0.42	**0.77**	0.51				
summer_wetM(16)	**-0.97**	0.04	**-0.86**	**0.94**	**-0.76**	**0.93**	**0.94**	**0.94**	**0.94**	0.67	**-0.96**	0.65	0.31	1.00	0.46			
winter_P(17)	-0.65	-0.15	-0.62	0.54	-0.32	0.55	0.52	0.54	0.53	0.42	**-0.79**	0.22	**0.83**	**0.75**	**0.74**	**0.75**		
winter_DryM(18)	-0.09	-0.15	-0.14	-0.03	0.20	0.00	-0.05	-0.02	-0.04	0.07	-0.29	-0.31	0.96	0.21	0.65	0.21	**0.80**	
winter_wetM(19)	**-0.89**	0.01	**-0.81**	**0.82**	-0.66	**0.81**	**0.82**	**0.81**	**0.82**	0.55	**-0.95**	0.49	0.51	**0.95**	0.56	**0.96**	**0.89**	0.45

EUR

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
latitude (1)																		
longitude (2)	0.16																	
elevation(3)	-0.19	0.51																
AvMonTemp(4)	**-0.92**	-0.29	-0.19															
thermalAmp(5)	0.65	0.61	0.41	-0.77														
HotMonth(6)	-0.77	0.11	-0.06	**0.84**	-0.32													
ColdMonth(7)	**-0.83**	-0.43	-0.35	**0.96**	**-0.92**	0.67												
summerSEASON(8)	**-0.93**	-0.10	-0.02	**0.95**	-0.54	**0.95**	**0.82**											
winterSEASON(9)	**-0.88**	-0.39	-0.28	**0.98**	**-0.87**	0.74	**0.99**	**0.88**										
monthabove10(10)	**-0.87**	-0.28	-0.27	**0.98**	-0.79	**0.82**	**0.97**	**0.91**	**0.98**									
MAP(11)	0.56	0.51	-0.21	-0.42	0.69	0.04	-0.53	-0.25	-0.50	-0.42								
Cv(12)	**-0.75**	-0.59	-0.41	**0.89**	**-0.92**	0.56	**0.96**	0.72	**0.94**	**0.92**	-0.62							
DryMonth(13)	0.28	0.67	0.60	-0.49	0.54	-0.32	-0.56	-0.36	-0.54	-0.58	0.34	-0.75						
summer_P(14)	0.02	0.26	**0.84**	-0.36	0.29	-0.40	-0.40	-0.31	-0.38	-0.37	-0.46	-0.33	0.33					
summer_DryM(15)	0.14	**0.82**	**0.89**	-0.45	0.68	-0.16	-0.60	-0.25	-0.54	-0.49	0.17	-0.69	0.77	0.71				
summer_wetM(16)	-0.18	-0.42	0.34	0.00	-0.28	-0.34	0.08	-0.12	0.07	0.03	**-0.85**	0.27	-0.32	0.72	0.02			
winter_P(17)	0.02	0.26	**0.84**	-0.36	0.29	-0.40	-0.40	-0.31	-0.38	-0.37	-0.46	-0.33	0.33	**1.00**	0.71	0.72		
winter_DryM(18)	0.14	**0.82**	**0.89**	-0.45	0.68	-0.16	-0.60	-0.25	-0.54	-0.49	0.17	-0.69	0.77	0.71	**1.00**	0.02	0.71	
winter_wetM(19)	-0.18	-0.42	0.34	0.00	-0.28	-0.34	0.08	-0.12	0.07	0.03	**-0.85**	0.27	-0.32	0.72	0.02	**1.00**	0.72	0.02

Annex Table S1.8. Climatic variables importance (given by its modeling power[a]) and contribution (given as correlation coefficients) to the principal components obtained in the three continents.

	Variable importance			Component 1			Component 2			Component 3
	USA	AUST	EUR	USA	AUST	EUR	USA	AUST	EUR	AUST
AvMonTemp	0.99	0.99	0.87	0.70	0.97	0.93	-0.71	-0.23	0.05	0.02
thermalAmp	0.78	0.96	0.84	-0.16	-0.71	-0.87	0.87	0.44	-0.29	0.52
HotMonth	0.77	0.98	0.55	0.84	0.97	0.69	-0.23	-0.19	-0.28	0.08
ColdMonth	0.97	0.99	0.95	0.54	0.96	0.97	-0.83	-0.26	0.12	-0.06
summerSEASON	0.89	0.99	0.66	0.83	0.97	0.81	-0.46	-0.21	-0.05	0.06
winterSEASON	0.99	0.99	0.94	0.57	0.97	0.96	-0.82	-0.24	0.11	-0.02
monthabove10	0.90	0.80	0.91	0.50	0.78	0.95	-0.81	-0.02	0.07	0.44
MAP	0.78	0.94	0.95	-0.79	-0.96	-0.41	-0.39	-0.11	-0.89	-0.01
Cv	0.95	0.83	0.99	-0.54	0.70	0.96	-0.81	-0.36	0.26	0.46
DryMonth	0.91	0.96	0.51	0.80	0.32	-0.70	0.53	0.93	-0.17	-0.01
summer_P	0.96	0.99	0.92	0.93	0.99	-0.54	0.30	0.02	0.79	-0.09
summer_DryM	0.95	0.97	0.56	0.91	0.53	-0.73	0.35	0.66	0.17	0.49
summer_wetM	0.94	0.99	0.95	0.96	0.98	0.01	0.09	0.01	0.97	-0.15
winter_P	0.96	0.99	0.92	0.93	0.72	-0.54	0.30	0.66	0.79	-0.19
winter_DryM	0.95	0.99	0.56	0.91	0.18	-0.73	0.35	0.96	0.17	-0.18
winter_wetM	0.94	0.98	0.95	0.96	0.92	0.01	0.09	0.26	0.97	-0.24

[a]A variable with a modeling power equal to one is completely relevant for building the PCA model. Variables with modeling power close to "number of components" divided by "number of variables" are regarded to be less significant.

Annex Table S1.9. Principal Component Analysis for climatic variables in the three continents. USA: North America; AUST: Australia; EUR: Europe.

	R^2X	R^2X (Cumul.)	Eigenvalues	Q^2
USA				
1	0.467734	0.467734	10.75788	0.076350
2	0.270011	0.737745	6.21025	0.293118
AUST				
1	0.684061	0.684061	10.94498	0.542123
2	0.207512	0.891573	3.32019	0.451575
3	0.067786	0.959359	1.08457	0.308800
EUR				
1	0.546608	0.546608	8.745728	0.217079
2	0.268152	0.814760	4.290434	0.277424

Annex Table S2.1 Candidate adaptive TE dataset. TE genomic positions from v6 *D. melanogaster* genome annotation. % length is the percentage length of the TE compared to the canonical sequence. ZB: TE frequency (%) in Zambia. IT: TE frequency (%) in Italy. SW: TE frequency (%) in Sweden. NC: TE frequency (%) in North Carolina (USA). RR1: recombination rate estimated in Comeron et al (2012). RR2: Recombination rate estimated in Fiston-Lavier et al (2010). The first 109 TEs correspond to the annotated TE dataset, and the last 12 TEs correspond to the non-annotated TE dataset. NA: no frequency data. ND: not determined.

TE	TEclass	TEfamily	chr	start	end	TE length	% length	ZB	IT	SW	NC	RR1	RR2
FBti0018862	LTR	17.6	2R	10948083	10955576	7494	100	4	46	25	18	3.15	2.25
FBti0018866	LTR	297	2R	15085034	15092025	6992	99.96	6	0	0	16	6.07	3.42
FBti0018877	non-LTR	BS	2R	9945496	9945626	131	2.55	0	0	0	23	0.76	1.84
FBti0018879	non-LTR	BS	2R	22465375	22465511	137	2.66	4	83	61	73	8.35	3.51
FBti0018883	LTR	Burdock	2R	9151357	9157769	6413	100	8	0	0	16	0.87	1.49
FBti0018884	LTR	Burdock	2R	16061783	16064358	2576	40.18	7	0	0	14	3.58	3.57
FBti0018916	non-LTR	F	2R	18111738	18115542	3805	80.82	7	0	0	11	3.15	3.76
FBti0018937	non-LTR	Rt1b	2R	12501751	12503829	2096	40.53	2	25	11	4	3.36	2.78
FBti0019008	non-LTR	Rt1a	2R	13905660	13907571	1912	37.43	0	19	0	1	3.58	3.16
FBti0019012	DNA	pogo	2R	17782416	17783563	1148	54.13	1	47	13	18	2.39	3.75
FBti0019056	DNA	pogo	X	14589730	14589915	186	8.77	1	100	96	63	1.90	3.56
FBti0019065	DNA	pogo	X	15421974	15423429	1456	68.65	4	78	56	57	2.35	3.32
FBti0019079	non-LTR	BS	X	18194124	18194597	474	9.22	5	0	0	23	1.27	2.26
FBti0019081	DNA	transib2	X	18555434	18556897	1464	51.48	8	56	39	76	1.18	2.08
FBti0019091	DNA	S	X	19710126	19711856	1731	99.71	2	0	0	10	3.62	1.48
FBti0019134	DNA	pogo	2L	7959095	7960450	1356	63.93	1	28	8	14	8.68	3.99
FBti0019158	non-LTR	BS	2L	12763895	12764039	145	2.82	2	NA	9	10	3.05	3.06
FBti0019164	non-LTR	X	2L	13036300	13036480	181	3.82	3	70	44	48	2.93	2.97
FBti0019165	non-LTR	BS	2L	13242015	13244341	2327	45.25	0	60	42	54	0.94	2.9
FBti0019176	LTR	copia	2L	14803935	14809079	5145	100	0	0	13	5	1.64	2.27
FBti0019177	non-LTR	jockey	2L	14890023	14890375	364	7.25	0	0	0	16	1.64	2.23
FBti0019279	DNA	1360	2R	10005906	10007010	1105	32.41	6	0	4	27	0.76	1.87
FBti0019354	LTR	17.6	3R	10804228	10811702	7475	100	3	9	0	20	0.98	1.05
FBti0019360	DNA	pogo	3R	12022938	12025059	2122	100	2	41	2	10	0.65	1.43
FBti0019381	non-LTR	Juan	3R	15132112	15135106	2995	70.70	0	19	3	3	4.03	2.23
FBti0019386	LTR	invader4	3R	16189464	16189810	347	11.18	3	56	45	65	1.31	2.44
FBti0019388	non-LTR	BS	3R	17294727	17295089	363	7.06	0	23	5	15	0.54	2.64
FBti0019389	non-LTR	F	3R	17346744	17348235	1492	31.69	0	13	2	21	0.54	2.65
FBti0019404	non-LTR	Rt1a	3R	19744302	19749476	5175	100	3	10	7	3	1.41	2.98
FBti0019410	non-LTR	BS	3R	20506641	20507386	746	14.51	6	43	10	49	2.18	3.05
FBti0019415	DNA	pogo	3R	22234737	22236000	1264	59.59	1	38	6	36	3.81	3.17
FBti0019453	non-LTR	jockey	3R	29319885	29320132	256	5.10	0	0	0	15	7.51	2.89
FBti0019457	DNA	pogo	3R	29760415	29761560	1146	54.03	2	6	8	25	1.63	2.83
FBti0019546	DNA	1360	X	2664442	2665527	1086	31.86	0	0	0	19	2.08	2.47
FBti0019602	non-LTR	Juan	X	8031495	8035729	4249	100	7	20	15	13	2.35	3.98
FBti0019604	non-LTR	BS	X	8364905	8365235	331	6.44	7	44	17	70	4.61	4.02
FBti0019624	DNA	hopper	X	11268618	11270052	1435	100	4	41	58	51	7.50	4.09
FBti0019627	DNA	pogo	X	11571507	11571692	186	8.77	5	78	51	81	2.35	4.07
FBti0019632	non-LTR	X	X	12302018	12303258	1241	26.18	2	81	56	67	2.08	3.99
FBti0019657	DNA	transib2	X	20504180	20505642	1463	51.44	4	44	13	50	2.80	1.04
FBti0019985	LTR	roo	2R	9871090	9871523	434	4.77	0	10	0	10	1.08	1.81
FBti0020036	non-LTR	Rt1a	3L	4773711	4774361	651	12.74	1	4	0	63	2.22	3.45
FBti0020046	non-LTR	Doc	3L	6040416	6042720	2305	48.78	1	61	11	19	0.90	3.45
FBti0020057	non-LTR	BS	3L	7130011	7130136	126	2.45	1	93	64	69	5.12	3.4
FBti0020089	non-LTR	X	3L	11105351	11106998	1648	34.77	0	21	11	35	4.22	2.87
FBti0020091	non-LTR	Rt1a	3L	11277515	11278450	936	18.32	3	96	79	72	4.08	2.83
FBti0020096	DNA	pogo	3L	11864607	11865846	1240	58.46	7	32	18	18	1.66	2.7
FBti0020110	non-LTR	Rt1b	3L	14721232	14723021	1790	34.62	1	0	5	10	0.48	1.9
FBti0020123	non-LTR	Doc	3L	16443902	16446186	2285	48.36	1	21	8	7	3.32	1.27
FBti0020125	non-LTR	BS	3L	16523336	16528459	5124	99.65	5	33	20	66	2.15	1.24
FBti0020137	DNA	S	3L	17799864	17801595	1732	99.77	1	25	2	11	0.28	0.71
FBti0020149	non-LTR	BS	3L	18514973	18520090	5118	99.53	2	97	93	85	1.32	0.38
FBti0020152	non-LTR	Doc	3L	18594004	18595923	1920	40.63	0	14	0	13	1.32	0.34
FBti0020155	DNA	1360	3L	18833837	18834940	1104	32.38	7	56	62	45	1.04	0.23
FBti0020323	DNA	1360	3R	8894928	8896023	1096	32.15	1	72	65	58	1.09	0.39
FBti0020390	DNA	hopper	3R	27856815	27858239	1425	99.30	0	19	7	2	4.35	3.05
FBti0020392	DNA	FB	3R	29079439	29080748	1310	100	5	12	24	0	1.63	2.92

Annex Table S2.1 (continued)

TE	TEclass	TEfamily	chr	start	end	TE length	% length	ZB	IT	SW	NC	RR1	RR2
FBti0020393	DNA	1360	3R	30818868	30820372	1505	44.15	2	19	0	22	0.65	2.68
FBti0060307	DNA	1360	2R	20088069	20088101	33	0.97	3	96	77	91	0.22	3.76
FBti0061303	DNA	1360	3L	14245616	14245658	43	1.26	2	7	11	6	2.28	2.05
FBti0063749	DNA	1360	3R	19106876	19106905	30	0.88	0	11	6	12	2.94	2.9
FBti0018867	LTR	297	2R	17798461	17798874	414	5.92	81	NA	NA	28	2.39	3.75
FBti0018868	LTR	297	2R	23877783	23878196	414	5.92	83	64	94	100	1.84	3.24
FBti0018880	DNA	Bari1	2R	18858291	18860019	1729	100	36	97	68	82	8.89	3.78
FBti0018936	non-LTR	Rt1b	2R	8490623	8492278	1656	32.02	10	43	48	17	0.22	1.17
FBti0018951	LTR	accord	2R	12862330	12867554	5225	70.57	46	15	17	29	1.52	2.89
FBti0018980	LTR	invader1	2R	6622200	6622615	416	10.32	31	75	69	96	0.54	0.17
FBti0019010	DNA	pogo	2R	11134165	11134350	186	8.77	82	NA	78	43	4.12	2.32
FBti0019055	LTR	opus	X	14551699	14559302	7604	100	85	NA	NA	61	1.90	3.57
FBti0019061	LTR	rover	X	15034147	15041616	7470	100	58	NA	0	24	1.72	3.44
FBti0019071	DNA	pogo	X	17057989	17058174	186	8.77	29	0	0	14	1.90	2.75
FBti0019082	non-LTR	Rt1b	X	18783882	18785788	1907	36.88	100	100	100	91	2.53	1.97
FBti0019088	LTR	Idefix	X	19437325	19444785	7461	100	69	NA	0	29	2.08	1.63
FBti0019112	DNA	pogo	2L	2933354	2935475	2122	100	10	65	29	56	0.94	3.53
FBti0019133	non-LTR	BS	2L	7579255	7579380	131	2.55	13	75	61	51	2.58	4.01
FBti0019144	non-LTR	Rt1b	2L	10138214	10143384	5171	100	36	33	25	60	9.50	3.73
FBti0019276	DNA	S	2R	6664234	6665968	1735	99.94	28	81	82	89	0.54	0.2
FBti0019344	non-LTR	Rt1a	3R	9278840	9284016	5177	100	12	13	29	28	0.98	0.53
FBti0019372	DNA	S	3R	14021702	14023463	1762	100	16	34	18	20	7.51	1.97
FBti0019378	non-LTR	BS	3R	15059337	15059465	129	2.51	82	68	63	58	3.92	2.21
FBti0019400	DNA	Bari1	3R	19137125	19138864	1740	100	80	97	100	99	2.94	2.91
FBti0019443	non-LTR	Rt1b	3R	27791698	27794772	3075	59.47	26	47	18	45	4.35	3.05
FBti0019552	LTR	opus	X	3178521	3186128	7608	100	72	0	0	10	3.44	2.69
FBti0019564	LTR	mdg1	X	3785867	3786055	189	2.53	41	70	80	45	2.71	2.92
FBti0019611	LTR	297	X	9798578	9805572	6995	100	92	NA	0	31	4.25	4.12
FBti0019612	LTR	297	X	10095218	10101092	5875	83.99	90	NA	NA	89	4.16	4.12
FBti0019613	DNA	1360	X	10101814	10102819	1006	29.51	11	NA	40	20	4.16	4.12
FBti0019623	LTR	297	X	11240552	11240965	414	5.92	89	NA	0	84	7.50	4.09
FBti0019677	DNA	hopper	X	21254758	21255285	528	36.79	59	100	100	97	0.72	0.57
FBti0019771	DNA	1360	2L	17335603	17336708	1106	32.44	14	71	90	63	1.17	0.95
FBti0019975	LTR	297	2R	7262242	7269237	6996	100	63	NA	0	12	0.22	0.53
FBti0019978	DNA	1360	2R	8566089	8566980	892	26.17	15	66	60	47	1.30	1.21
FBti0020041	LTR	Quasimodo	3L	5299207	5306585	7379	99.89	68	0	0	15	5.12	3.45
FBti0020086	LTR	17.6	3L	10060167	10067688	7522	100	20	43	25	37	1.04	3.06
FBti0020114	DNA	transib2	3L	14954422	14956032	1611	56.65	31	73	56	76	0.35	1.82
FBti0020119	DNA	S	3L	15554974	15556705	1732	99.77	13	100	89	55	1.25	1.61
FBti0020128	non-LTR	BS	3L	16730986	16731111	126	2.45	100	86	93	96	1.18	1.16
FBti0020146	DNA	S	3L	18189644	18190189	546	31.45	78	59	60	58	1.80	0.53
FBti0020151	non-LTR	Cr1a	3L	18590703	18591377	675	15.10	82	84	100	98	1.32	0.35
FBti0059782	LTR	297	X	18754392	18757999	3608	51.58	47	0	0	31	2.53	1.99
FBti0060443	DNA	1360	3R	9700026	9700057	32	0.94	17	84	56	67	0.44	0.68
FBti0060715	DNA	1360	3L	11474678	11474706	29	0.85	10	54	66	75	3.18	2.79
FBti0061105	non-LTR	G5	2R	7317828	7317878	51	1.05	89	100	100	100	0.76	0.56
FBti0061417	non-LTR	BS	3L	15056356	15056430	75	1.46	91	68	100	81	0.76	1.79
FBti0061428	DNA	H	2L	16858766	16859500	735	24.84	50	50	35	48	1.64	1.22
FBti0061506	DNA	1360	2L	17432071	17432118	48	1.41	97	NA	38	25	1.64	0.89
FBti0061529	non-LTR	BS	3R	12780325	12780388	64	1.24	92	100	100	97	0.44	1.64
FBti0062242	non-LTR	BS	3R	16041234	16041335	102	1.98	91	100	100	97	0.76	2.41
FBti0062309	DNA	1360	2R	10875686	10875723	38	1.11	86	100	100	95	4.34	2.22
tdn4	LINE	Jockey	2R	18807871	18807898	800	15.94	ND	ND	ND	88	3.69	3.76
tdn5	LINE	I	2L	8959897	8959967	300	5.55	ND	ND	ND	78	4.22	4.195
tdn6	LTR	Gypsy	3R	9192517	9192611	300	5.47	ND	ND	ND	63	1.31	0.395
tdn7	DNA	P	2R	15374746	15374797	500	42.52	ND	ND	ND	90	3.36	3.425
tdn8	LTR	Gypsy	3L	12863675	12863781	5,500	100	ND	ND	ND	63	3.39	2.58
tdn12	DNA	TcMar-Pogo	3L	14049977	14050077	1,500	68.18	ND	ND	ND	70	1.94	2.245
tdn13	DNA	TcMar-Pogo	3L	15035122	15035206	150	6.82	ND	ND	ND	86	0.76	1.93
tdn14	LTR	Gypsy	3R	18851003	18851056	250	5.21	ND	ND	ND	50	4.13	2.855
tdn15	DNA	TcMar-Pogo	3L	18815329	18815428	600	27.27	ND	ND	ND	80	1.04	0.35
tdn17	DNA	TcMar-Pogo	X	21399382	21399471	1,000	45.45	ND	ND	ND	75	1.09	0.79
tdn18	DNA	P	X	21204983	21205161	1,000	100	ND	ND	ND	60	0.90	0.915
tdn19	DNA	Transib	3R	14049333	14049435	1,500	53.57	ND	ND	ND	78	7.51	1.915

Annex Table S2.2: TEs not annotated in *D. melanogaster* v6 reference genome. TE length is inferred from the PCR product size. TE frequencies in NC population are calculated based on the obtained PCR results. TE+: number of strains homozygous for the presence of the TE. TE-: number of strains homozygous for the

							PCR screening				
TE	Chr.	Start	End	TE class	TE family	TE length (bp)	TE Frequency in NC	TE+ strains	TE- strains	PCR didn't work	Total strains checked (with results)
tdn1	3R	17673745	17673774	DNA	TcMar-Tc1	2,500	33	1	2	2	3
tdn2	3L	19339394	19339442	DNA	TcMar-Tc1	1,500	17	1	5	10	6
tdn3	X	2745133	2745242	DNA	FB	2,500	20	1	4	3	5
tdn4	2R	18807871	18807898	LINE	Jockey	800	88	14	2	0	16
tdn5	2L	8959897	8959967	LINE	I	300	78	7	2	2	9
tdn6	3R	9192517	9192611	LTR	Gypsy	300	63	5	3	2	8
tdn7	2R	15374746	15374797	DNA	P	500	90	9	1	1	10
tdn8	3L	12863675	12863781	LTR	Gypsy	5,500	63	5	3	0	8
tdn9	X	18193707	18193807	DNA	Transib	1,250	33	2	4	13	6
tdn10	X	19607809	19607911	DNA	TcMar-Tc1	1,000	25	1	3	7	4
tdn11	2R	11951802	11951878	DNA	TcMar-Pogo	1,000	67	4	2	0	6
tdn12	3L	14049977	14050077	DNA	TcMar-Pogo	1,500	70	7	3	1	10
tdn13	3L	15035122	15035206	DNA	TcMar-Pogo	150	86	6	1	3	7
tdn14	3R	18851003	18851056	LTR	Gypsy	250	27	3	8	1	11
tdn15	3L	18815329	18815428	DNA	TcMar-Pogo	600	80	4	1	4	5
tdn16	X	11221700	11221773	DNA	TcMar-Pogo	1,500	40	2	3	0	5
tdn17	X	21399382	21399471	DNA	TcMar-Pogo	1,000	75	9	3	0	12
tdn18	X	21204983	21205161	DNA	P	1,000	33	3	6	7	9
tdn19	3R	14049333	14049435	DNA	Transib	1,500	78	7	2	3	9
tdn20	3L	3798625	3798632	DNA	P	NA	0	0	4	0	4
tdn21	3R	26968419	26968433	DNA	TcMar-Tc1	NA	0	0	5	1	5
tdn22	X	10682928	10682947	DNA	TcMar-Pogo	NA	0	0	4	0	4
tdn23	X	8027461	8027490	LINE	Jockey	NA	0	0	3	3	3
tdn24	2R	14848980	14848986	LTR	Gypsy	NA	0	0	5	0	5
tdn25	X	21087630	21087639	LTR	Gypsy	NA	0	0	4	0	4

Annex Table S2.3 Gene functional information from candidate adaptive TE dataset. Genes nearby the TE are the total of genes located at less than 1 kb distance from each TE, or the closest nearby gene further than 1 kb. Funcional information of neaby genes contains information obtained from Flybase GO annotations as well as gene functional information retrieved from the literature. Functional annotation based on gene functional information found. The first 109 TEs correspond to the annotated TEs, and the last 12 TEs correspond to the non-annotated TEs.

TE	Gene(s) nearby the TE (Flybase ID)	Dist closest nearby gene (bp)	TE location from closest nearby gene	Nearby gene(s) names	Functional information nearby genes	Functional annotation
FBti0018862	*FBgn0266763*	988	3'	*CR45228*	no information	-
FBti0018866	*FBgn0265650*	0	exon	*CR44457*	no information	-
FBti0018877	*FBgn0011656*	0	first intron	*Mef2*	midgut development (Vining et al. 2005); antimicrobial humoral response, carbohydrate and lipid storage (Clark et al. 2013); development (Lovato et al. 2005; Furlong et al. 2001; Bour et al. 1995; Nguyen et al. 2002; Menon et al. 2005; Bryantsev et al. 2012; Brunetti et al. 2015); locomotor rhythm (Blanchard et al. 2010), regulation of gene expression (Elgar et al. 2008; Firdaus et al. 2015; Tanaka et al. 2008).	Immune response, metabolism, development and morphogenesis
FBti0018879	*FBgn0034731*	0	first intron	*CG10384*	no information	-
FBti0018883	*FBgn0050345/FBgn0033387*	0	3'UTR	*CG30345 / CG8008*	CG30345: no information. // CG8008: immune responsive (Valanne et al. 2007; Silverman et al. 2003), sensory perception of pain (Neely et al. 2010)	Immune response
FBti0018884	*FBgn0262446/FBgn0034071*	896	3'	*mir-137 / CG8405*	no information	-
FBti0018916	*FBgn0085225*	1244	3'	*CG34196*	no information	-
FBti0018937	*FBgn0045063/FBgn0050044*	0	first intron	*fdl / s-cup*	fdl: protein deglycosilation (Leonard et al. 2006; Rosenbaum et al. 2014), brain development (Boquet et al. 2000) // s-cup: no information	metabolism
FBti0019008	*FBgn0013733*	0	intron	*shot*	axonogenesis (Alves-Silva et al. 2012); microtubule cytoskeleton organization (Roper and Brown 2004; Subramanian et al. 2003); development and morphogenesis (Roper and Brown 2004; Lee and Kolodziej 2002; Reuter et al. 2003; Lee et al. 2003; Gao et al. 1999; Parrish et al. 2006; Bottenberg et al. 2009; Sanchez-Soriano et al. 2009)	development and morphogenesis
FBti0019012	*FBgn0262416/FBgn0028741/ FBgn0250851*	232	3'	*mir-31a / fab / CG33981*	mir-31a: segmentation (Leaman et al. 2005); muscle cell cellular homeostasis (Fulga et al. 2015). / fab: autophagic vacuole fusion (Rusten et al. 2007); endosome to lysosome transport and phosphatidylinositol phosphorylation (Rusten et al. 2006). / CG33981: no information.	signaling, other cell processes
FBti0019056	*FBgn0030574*	0	intron	*CG9413*	Hypoxia tolerance (Azad et al. 2012); carboplatin toxicity (King et al. 2014).	xenobiotic stress
FBti0019065	*FBgn0263257*	0	first intron	*Cngl*	oxidative stress (Weber et al. 2012), response to hypoxia (Dijkers and O'Farrell 2009; Vermehren-Schmaedick et al.. 2010). Possible role in the processing of visual and olfactory information in the neuron system (Miyazu et al. 2000).	oxidative stress, olfaction, photoreception.
FBti0019079	*FBgn0083228*	0	first intron	*Frq2*	Neuromuscular junction development, synaptic transmission (Romero-Pozuelo et al. 2007); regulation of neurotransmitter secretion (Dason et al. 2009).	development and morphogenesis
FBti0019081	*FBgn0265598*	0	first intron	*Bx*	chaeta development and positive regulation of transcription DNA-templated (Zenvirt et al. 2008); gastric emptying (Ren et al. 2014); imaginal disc-derived leg segmentation (Pueyo and Couso 2004); imaginal disc-derived wing morphogenesis (Milan et al. 1998); inter-male aggressive behavior (Edwards et al. 2009); locomotor rhythm and response to cocaine (Tsai et al. 2004); phagocytosis (Stroschein-Stevenson et al. 2006); reproductive process (Kairamkonda and Nongthomba 2014); oxidative stress (Weber et al. 2012).	Oxidative stress, xenobiotic stress, mating, development and morphogenesis
FBti0019091	*FBgn0040089*	0	first intron	*meso18E*	mesoderm development (Furlong et al. 2001)	development and morphogenesis

Annex Table S2.3 (continued)

TE	Gene(s) nearby the TE (Flybase ID)	Dist closest nearby gene (bp)	TE location from closest nearby gene	Nearby gene(s) names	Functional information nearby genes	Functional annotation
FBti0019134	*FBgn0085450*	0	first intron	*Snoo*	Neuron development (Takaesu et al. 2006); oxidative stress (Weber et al. 2012); negative regulation of decapentaplegic signaling pathway (Barrio et al. 2007); negative regulation of transforming growth factor beta receptor signaling pathway (Ramel et al. 2007)	Oxidative stress, development and morphogenesis, signaling
FBti0019158	*FBgn0032456*	0	intron	*MRP*	xenobiotic-transporting ATPase activity (Chahine and O'Donnell, 2009; 2010)	xenobiotic stress, membrane transport
FBti0019164	*FBgn0262160*	0	intron	*CG9932*	Wing disc and chaeta development (Bronstein et al. 2010); Starvation resistance and locomotor activity (Ayroles et al. 2009).	Other stress, development and morphogenesis
FBti0019165	*FBgn0266869/FBgn0032494/ FBgn0041720/FBgn0003935*	0	first intron	*CR45330 / CG5945 / snRNA:U2:34ABc / snRNA:U5:34A*	CR45330: no information / CG5945: Circadian clock and mating behavior (Kadener et al. 2006) / snRNA:U2:34ABc and snRNA:U5:34A: no information.	mating behavior, circadian rhythm
FBti0019176	*FBgn0250834*	642	3'	*CG33308*	no information	-
FBti0019177	*FBgn0266840*	4612	3'	*CR45302*	no information	-
FBti0019279	*FBgn0022382*	0	intron	*Pka-R2*	Behavioral response to cocaine, ethanol, circadian rhythm and locomotor rhythm (Park et al. 2000). Odour-guided behavior (Brown et al. 2013), axon guidance (Terman and Kolodkin 2004).	xenobiotic stress, behavior, olfaction
FBti0019354	*FBgn0037837/FBgn0037836*	0	3'UTR	*CG14693 / CG14692*	CG14693: Auditory perception (Senthilan 2012). CG14692: myosin light chain binding (Franke et al. 2006).	auditory perception
FBti0019360	*FBgn0051358*	2677	3'	*CG31358*	no information	-
FBti0019381	*FBgn0038290/FBgn0261859*	32	3'	*CG6912 / CG42788*	*CG42788:* response to infection (Short and Lazzaro 2013) / *CG6912:* no information	immune response
FBti0019386	*FBgn0024491/FBgn0086370*	0	first intron	*sra / Bin1*	sra: female meiotic division (Takeo et al. 2006, Horner et al. 2006, Takeo et al. 2012), long-term memory and olfactory behavior (Chang et al. 2003), courtship behavior (Ejima et al. 2004), egg activation (Horner et al. 2006) // Bin1: response ton environmental stress (Costa et al. 2011), chromatin silencing (Matyash et al. 2009).	immune response, learning or memory, courtship behavior, olfaction, development and morphogenesis, meiosis
FBti0019388	*FBgn0263501*	5909	3'	*CR43490*	no information	-
FBti0019389	*FBgn0038498*	0	first intron	*beat-IIa*	regulation of glucose metabolic process (Ugrankar et al. 2015); response to oxidative stress (Weber et al. 2012)	oxidative stress, metabolism
FBti0019404	*FBgn0024963/FBgn0263499*	0	first intron	*GluClalpha / CR43488*	GluClalpha: neuron projection morphogenesis (Sepp et al. 2008) / CR43488: no information	development and morphogenesis, membrane transport
FBti0019410	*FBgn0038799/FBgn0038798*	6	3'	*MFS9 / Or92a*	MFS9: Copper homeostasis and detoxification (Egli et al. 2006). // Or92a: olfactory receptor (Vogt et al. 2002).	xenobiotic stress, olfaction
FBti0019415	*FBgn0051163*	0	first intron	*SKIP*	Sensory perception of smell (Tunstall et al. 2012)	Sensory perception of smell
FBti0019453	*FBgn0015622*	241	5'	*Cnx99A*	rhodopsin biosynthetic process, regulation of calcium ion transport into cytosol, retina homeostasis (Rosenbaum et al. 2006); foraging behavior (Chen et al. 2012); and circadian rhythm (Ceriani et al. 2002).	Foraging behavior, photoreception, circadian rhythm

Annex Table S2.3 (continued)

TE	Gene(s) nearby the TE (Flybase ID)	Dist closest nearby gene (bp)	TE location from closest nearby gene	Nearby gene(s) names	Functional information nearby genes	Functional annotation
FBti0019457	*FBgn0266258 / FBgn0001297*	1375 /	5'	*CR44953 / kay*	CR44953: no information / kay: immune response (Kim et al. 2005; Kleino et al. 2005), pigmentation (Dembeck et al. 2015), locomotor rhythm (Ling et al. 2012), wound healing (Bosch et al. 2005; Ramet et al. 2002), development (Cerrato et al. 2006; Mathieu et al. 2007; Iyer et al. 2013; Grima et al. 2008; Jemc et al. 2012; Hyun et al. 2006)	Immune response, development and morphogenesis, pigmentation.
FBti0019546	*FBgn0003371*	0	first intron	*sgg*	response to anesthetic (Campbell et al. 2009); olfactory learning (Wolf et al. 2007); circadian rhythm (Martinek et al. 2001; Yuan et al. 2005; Wolf et al. 2007); chitin-based larval cuticle pattern formation (Kaplan et al. 2009); development and morphogenesis (Kaplan et al. 2011; Perrimon and Smouse 1989; Kanuka et al. 2005; Song and Xie 2003; Mohit et al. 2006); regulation of hemocyte differentiation (Zettervall et al. 2004); synapsis (Franco et al. 2004); female meiosis (Takeo et al. 2012; Song and Xie 2003; Jordan et al. 2006); signaling regulation (Franciscovich et al. 2008; Price and Kalderon 2002; Jia et al. 2002; Takeo et al. 2012); negative regulation of synaptic growth at neuromuscular junction (Franciscovich et al. 2008); protein catabolism and phosphorylation (Galletti et al. 2009; Price and Kalderon 2002; Price and Kalderon 2002 Jia et al. 2005)	Xenobiotic stress, learning or memory, behavior, olfaction, circadian rhythm, development and morphogenesis, signaling, meiosis
FBti0019602	*FBgn0029990*	12	3'	*CG2233*	Low larvae weight and high survival (Bochdanovits and de Jong 2004). Response to bacterial infection (Reumer et al. 2009).	immune response, metabolism
FBti0019604	*FBgn0015519*	0	intron	*nAChRalpha3*	insecticide resistance (Lansdell and Millar, 2000); acetylcholine-activated cation-selective channel activity (Schulz et al. 1998).	xenobiotic stress, membrane transport
FBti0019624	*FBgn0265595*	0	intron	*CG44422*	no information	-
FBti0019627	*FBgn0027259/FBgn0030311*	0	3'UTR	*Kmn1 / CG11699*	Kmn1: chromosome segregation (Przewloka et al. 2007; Venkei et al. 2011), neurogenesis (Neumüller et al. 2011), regulation of cell cycle (Clemente-Ruiz et al. 2014) / CG11699: Xenobiotic metabolism (Mateo et al. 2014)	xenobiotic stress, development and morphogenesis
FBti0019632	*FBgn0267001*	0	intron	*Ten-a*	synapsis (Kurusu et al. 2008; Hong et al. 2012; Mosca et al. 2012); immunolocalizes with adult olfactory receptor neurons (Hong et al. 2012), alcohol tolerance (Ghezzi et al. 2013).	xenobiotic stress, photoreception, development and morphogenesis
FBti0019657	*FBgn0031118*	0	intron	*RhoGAP19D*	imaginal disc-derived leg morphogenesis (Greenberg and Hatini 2011)	development and morphogenesis
FBti0019985	*FBgn0011241/FBgn0033458*	0	first intron	*cbx*	immune system (Ayres et al. 2008); spermatogenesis (Fabrizio et al. 1998; Castrillon et al. 1993)	immune response, spermatogenesis
FBti0020036	*FBgn0035574*	0	intron	*RhoGEF64C*	axon guidance (Bashaw et al. 2001); imaginal disc-derived leg morphogenesis (Greenberg and Hatini 2011); inter-male agressive behavior (Edwards et al. 2009); positive regulation of Rho protein signal transduction (Simoes et al. 2006); spiracle morphogenesis (Simoes et al. 2006)	behavior, development and morphogenesis, signaling
FBti0020046	*FBgn0250815*	281	3'	*Jon65Aiv*	serine-type endopeptidase activity (Ross et al. 2003); odor-guided behaviour (Anholt & Mackay 2001); mating-regulated (McGraw et al. 2004); immune response (De Gregorio et al. 2002; Short and Lazzaro 2013)	immune response, mating, behavior, olfaction, metabolism
FBti0020057	*FBgn0035743/FBgn0250836*	338	3'	*CG15829 / CG8628*	CG15829: immune response (Broderick et al. 2014; Rynes et al. 2012) // CG8628: IMD/NF-ŒʃB signaling (Combe et al. 2014), immune response (Roxstrom-Lindquist et al. 2004).	immune response, signaling
FBti0020089	*FBgn0052073*	10	3'	*CG32073*	no information	-
FBti0020091	*FBgn0265931*	2953	5'	*CR44720*	no information	-
FBti0020096	*FBgn0266100*	0	first intron	*CG44837*	Induced with phenobarbital (Sun et al. 2006)	xenobiotic stress
FBti0020110	*FBgn0259175*	0	first intron	*ome*	proteolysis (Chihara et al. 2005)	metabolism
FBti0020123	*FBgn0053158/FBgn0261799*	0	intron	*CG33158 / dsx-c73A*	CG33158: chill coma recovery (Ayroles et al. 2009). // dsx-c73A: constituent of chitin-based cuticle (Andrew and Baker 2008).	cold stress
FBti0020125	*FBgn0263131*	0	intron	*CG43373*	no information	-

Annex Table S2.3 (continued)

TE	Gene(s) nearby the TE (Flybase ID)	Dist closest nearby gene (bp)	TE location from closest nearby gene	Nearby gene(s) names	Functional information nearby genes	Functional annotation
FBti0020137	*FBgn0052190/FBgn0036754*	0	first intron	*NUCB1 / CG5589*	NUCB1: immune response (Berkey et al. 2009). / CG5589: no information	immune response
FBti0020149	*FBgn0052198*	6886	5'	*CG32198*	no information	-
FBti0020152	*FBgn0003683/FBgn0266938*	233	5'	*term / CR45389*	no information	-
FBti0020155	*FBgn0036816*	0	first intron	*Indy*	determination of adult lifespan (Wang et al. 2009; Rogina et al. 2000); regulation of sequestering of triglyceride (Wang et al. 2009); Fitness advantage by transposon insertion: increased fecundity and longevity through metabolic changes (Zhu et al. 2014)	fecundity, lifespan, transport
FBti0020323	*FBgn0262614*	0	first intron	*pyd*	development and morphogenesis (Jennings et al. 2007; Jung et al. 2006; Djiane et al. 2011; Seppa et al. 2008; Choi et al. 2011; Mummery-Widmer et al. 2009; Zhuang et al. 2009)	development and morphogenesis
FBti0020390	*FBgn0085382*	0	intron	*CG34353*	gravitaxis (Armstrong et al. 2006); oxidative stress (Weber et al. 2012); heavy metal stress (Zhou et al. 2016)	oxidative stress
FBti0020392	*FBgn0039633*	0	first intron	*CG11873*	response to endoplasmatic reticulum stress (Chow et al. 2013), oxidative stress (Weber et al. 2012)	oxidative stress, endoplasmatic reticulum stress
FBti0020393	*FBgn0027598*	0	intron	*cindr*	development and morphogenesis (Quinones et al. 2010; Johnson et al. 2008; Johnson and Cagan 2009; Mummery-Widmer et al. 2009); regulation of cytokinesis (Haglund et al. 2010).	development and morphogenesis
FBti0060307	*FBgn0265843*	1374	3'	*CR44632*	no information	-
FBti0061303	*FBgn0265754/FBgn0265753*	0	exon	*CR44561 / CR44560*	no information	-
FBti0063749	*FBgn0038679*	0	first intron	*CG6040*	no information	-
FBti0018867	*FBgn0028743*	0	intron	*Dhit*	positive regulation of GTPase activity (Lin et al. 2014)	signaling
FBti0018868	*FBgn0020372/FBgn0011236*	1	5'	*TM4SF / ken*	TM4SF: JAK/STAT pathway regulation - ken related (Arbouzova et al. 2006). // ken: imaginal disc-derived genitalia development (Lukacsovich et al. 2003); insemination (Castrillon et al. 1993); phagocytosis (Stroschein-Stevenson et al. 2006); regulation of JAK-STAT cascade (Arbouzova et al. 2006)	immune response, insemination, development and morphogenesis
FBti0018880	*FBgn0034405/FBgn0034406*	46	5'	*Jheh2 / Jheh3*	Jheh2: oxidative stress (Guio et al. 2014) // Jheh3: egg production (Terashima and Bownes 2005).	Oxidative stress, development and morphogenesis, egg production
FBti0018936	*FBgn0050361*	0	first intron	*mtt*	L-canavanine insecticide detection, feeding behavior (Mitri et al. 2009), G-protein coupled receptor activity (Mitri et al. 2004)	xenobiotic stress, behavior, signaling
FBti0018951	*FBgn0033777/FBgn0266633*	0	first intron	*CG17574 / CR45140*	no information	-
FBti0018980	*FBgn0086655*	408	3'	*jing*	axon guidance (Sun et al. 2006); development (Liu and Montell 2001; Sedaghat et al. 2002; Sedaghat and Sonnenfeld 2002; Carreira et al. 2011; Culi et al. 2006; Sonnenfeld et al. 2004); regulation of glucose metabolic process (Ugrankar et al. 2015); tissue regeneration (McClure and Schubiger 2008)	metabolism, development or morphogenesis, signaling
FBti0019010	*FBgn0033578*	1573	5'	*BBS4*	cilium assembly (Avidor-Reiss et al. 2004)	cillium assembly
FBti0019055	*FBgn0267077*	9702	5'	*CR45521*	no information	-
FBti0019061	*FBgn0030600/FBgn0052594*	0	intron	*hiw / be*	hiw: autophagy (Shen and Ganetzky 2009); BMP signaling pathway (McCabe et al. 2004); long-term memory (Huang et al. 2012); synapsis. / be: long-term memory (Zhao et al. 2009).	learning or memory, signaling, autophagy
FBti0019071	*FBgn0266354*	0	first intron	*CG45002*	no information	-
FBti0019082	*FBgn0030958/FBgn0030956/ FBgn0030959*	17	5'	*CR6900 / CG18259 / CG6961*	no information	-

Annex Table S2.3 (continued)

TE	Gene(s) nearby the TE (Flybase ID)	Dist closest nearby gene (bp)	TE location from closest nearby gene	Nearby gene(s) names	Functional information nearby genes	Functional annotation
FBti0019088	*FBgn0031016*	0	intron	*kek5*	Regulation of BMP signaling pathway (Evans et al. 2009)	development and morphogenesis, signaling
FBti0019112	*FBgn0041111*	0	first intron	*lilli*	compound eye photoreceptor development (Wittwer et al. 2001); learning or memory (Wang et al. 2008); olfactory behavior (Sambandan et al. 2006); regulation of cytoskeleton organization (Tang et al. 2001), development (Bejarano et al. 2008; Luschnig et al. 2004)	learning or memory, olfaction, photoreception, circadian rhythm, development and morphogenesis
FBti0019133	*FBgn0264895*	275	5'	*RapGAP1*	Intermale agressive behavior (Edwards et al. 2009); negative regulator of small GTPase mediated signal transduction (Chen et al. 1997), oxidative stress (Weber et al. 2012)	oxidative stress, behavior, signaling
FBti0019144	*FBgn0265002*	0	intron	*CG44153*	no information	-
FBti0019276	*FBgn0000054/FBgn0266621*	0	first intron	*Adf1 / CR45128*	Adf1: Dendrite morphogenesis and regulation of development (Timmerman et al. 2013 and Parrish et al. 2006); locomotion (Parrish et al. 2006); memory and synapse assembly (DeZazzo et al. 2000). / CR45128: no information	learning or memory, development and morphogenesis
FBti0019344	*FBgn0261241*	8616	3'	*Vps16A*	endosomal transport (Pulipparacharuvil et al. 2005; Kim et al. 2010); cellular response to starvation (Takáts et al. 2014)	metabolism
FBti0019372	*FBgn0264493*	0	first intron	*rdx*	morphogenesis and development (Kent et al. 2006; Mummery-Widmer et al. 2009); apoptotic process; apoptosis and positive regulation of JNK cascade (Liu et al. 2009), oxidative stress (Weber et al. 2012), protein ubiquitination and regulation of proteolysis (Zhang et al. 2006; Liu et al. 2009)	Oxidative stress, development and morphogenesis, metabolism, signaling
FBti0019378	*FBgn0038282*	0	intron	*dpr9*	behavioral response to ethanol (Kong et al. 2010; Nakamura et al. 2002)	xenobiotic stress, behavior
FBti0019400	*FBgn0038681/FBgn0261285*	0	3'UTR	*Cyp12a4 / Ppcs*	Cyp12a4: Response to insecticide (Bogwitz et al. 2005). / Ppcs: development (Bosveld et al. 2008)	oxidative stress, xenobiotic stress, development and morphogenesis
FBti0019443	*FBgn0085382*	0	intron	*CG34353*	no information	-
FBti0019552	*FBgn0000479*	0	first intron	*dnc*	dnc: learning and memory (Honjo and Furukubo-Tokunaga, 2005, 2009; Kamyshev et al. 2000); thermosensory behavior (Hong et al. 2008)	learning or memory, behavior
FBti0019564	*FBgn0086899*	0	intron	*tlk*	antimicrobial humoral response (Kleino et al. 2005); cell cycle and organization (Li et al. 2009; Kiger et al. 2003); protein phosphorylation (Carrera et al. 2003)	immune response, cell cycle
FBti0019611	*FBgn0052698*	0	first intron	*CG32698*	sensory perception of pain (Neely et al. 2010)	sensory perception of pain
FBti0019612	*FBgn0083940*	0	intron	*RhoU*	oxidative stress (Weber et al. 2012)	oxidative stress, signaling
FBti0019613	*FBgn0083940*	0	intron	*RhoU*	oxidative stress (Weber et al. 2012)	oxidative stress, signaling
FBti0019623	*FBgn0265595*	0	intron	*CG44422*	oxidative stress (Weber et al. 2012)	oxidative stress
FBti0019677	*FBgn0052521*	0	first intron	*CG32521*	no information	-
FBti0019771	*FBgn0267255*	14891	5'	*CG45691*	no information	-
FBti0019975	*FBgn0033154*	973	5'	*CG1850*	no information	-
FBti0019978	*FBgn0033302*	0	5'UTR	*Cyp6a14*	no information	-
FBti0020041	*FBgn0267305*	9628	3'	*CR45741*	no information	-
FBti0020086	*FBgn0040823*	0	intron	*dpr6*	salt aversion response, sensory perception of chemical stimulus (Nakamura et al. 2002).	xenobiotic stress
FBti0020114	*FBgn0267640*	3719	3'	*CR45978*	no information	-

Annex Table S2.3 (continued)

TE	Gene(s) nearby the TE (Flybase ID)	Dist closest nearby gene (bp)	TE location from closest nearby gene	Nearby gene(s) names	Functional information nearby genes	Functional annotation
FBti0020119	*FBgn0087035*	0	first intron	*AGO2*	defense response to virus (Wang et al. 2006; Mueller et al. 2010; Han et al. 2011; Zhang et al. 2015); response to gram-negative bacteria (Fukuyama et al. 2013); dosage compensation by hyperactivation of X chromosome (Menon and Meller; 2012); gene silencing by miRNA (Besnard-Guérin et al. 2015); heterochromatin organization involved in chromatin silencing (Fagegaltier et al. 2009); negative regulation of transposition; RNA-mediated (Berry et al. 2009); negative regulation of viral genome replication (Sabin et al. 2009); cellularization (Deshpande et al. 2005); pole cell formation (Deshpande et al. 2005); production of siRNA involved in RNA interference (Hammond et al. 2001; Okamura et al. 2008); salivary gland cell autophagic cell death (Gorski et al. 2003); segment polarity determination (Meyer et al. 2006); dsRNA transport (Saleh et al. 2006); RNA interference (Matranga et al. 2005; Rand et al. 2005; Ishizuka et al. 2002; Dorner et al. 2006;Rehwinkel et al. 2005; Meyer et al. 2006); siRNA loading onto RISC involved in RNA interference (Kim et al. 2007; Okamura et al. 2004); syncytial nuclear migration (Deshpande et al. 2005)	immune response, development and morphogenesis
FBti0020128	*FBgn0004556*	168	5'	*Dbp73D*	Neurogenesis (Neumüller et al. 2011)	development and morphogenesis
FBti0020146	*FBgn0003997*	4003	5'	*hid*	development and morphogenesis (Cullen and McCall 2004; Williams et al. 2006; Abbott and Lengyel, 1991; Guha and Kornberg, 2005; de la Cova et al. 2004); apoptosis (Sandu et al. 2010; Haining et al. 1999; Werz et al. 2005; Wang et al. 1999; Grether et al. 1995, Kurada and White, 1998; Moon et al. 2008; Yin and Thummel, 2004; Ribeiro et al. 2007; Tanaka-Matakatsu et al. 2009, Leulier et al. 2006; Zhou et al. 1997; Rodriguez Moncalvo and Campos, 2005; Leulier et al. 2006; Yin and Thummel, 2004; Jiang et al. 2000); cellular response to gamma radiation (Zhang et al. 2013); cellular response to starvation (Hou et al. 2008); circadian clock (Klarsfeld et al. 2004); positive regulation of cellular response to X-ray (Brodsky et al. 2004); endopeptidase activity (Wang et al. 1999; Yan et al. 2004); positive regulation of macroautophagy (Hou et al. 2008); response to red light (Klarsfeld et al. 2004); sex differentiation (DeFalco et al. 2003	starvation, response to radiation, circadian rhythm, metabolism, development and morphogenesis
FBti0020151	*FBgn0036791/FBgn0003683*	538	3'	*CG7271 / term*	no information	-
FBti0059782	*FBgn0030952*	6086	5'	*CG12609*	no information	-
FBti0060443	*FBgn0051352*	0	first intron	*Unc-115a*	axon guidance and photoreceptor cell axon guidance (Garcia et al. 2007)	photoreception
FBti0060715	*FBgn0267716/FBgn0028667/ FBgn0028668*	114	3'	*CR46049 / Vha16-2 / Vha16-3*	CR46049: no information / Vha16-2: phagocytosis (Stroschein-Stevenson et al. 2006), ATP transporter (Chintapalli et al. 2013) / Vha16-3: ATP transporter (Chintapalli et al. 2013)	phagocytosis
FBti0061105	*FBgn0033159*	46	3'	*Dscam1*	axon guidance and neuron development, detection of molecule of bacterial origin and phagocytosis (Watson et al. 2005)	immune response, development and morphogenesis
FBti0061417	*FBgn0261090/FBgn0036470*	0	first intron	*Sytbeta / EAChm*	Sytbeta: no information / EAChm: regulation of transcription (Nakagawa et al. 2015)	transcription regulation
FBti0061428	*FBgn0051809/FBgn0032615*	52	5'	*CG31809 / CG6012*	no information	-

Annex Table S2.3 (continued)

TE	Gene(s) nearby the TE (Flybase ID)	Dist closest nearby gene (bp)	TE location from closest nearby gene	Nearby gene(s) names	Functional information nearby genes	Functional annotation
FBti0061506	*FBgn0011274*	0	first intron	*Dif*	Immune response (Rutschmann et al. 2000; Gobert et al. 2003; Brown et al. 2009; Christofi and Apidianakis 2013; Cornwell and Kirkpatrick 2001; Bou-Sleiman et al. 2015); cellular response to DNA damage stimulus (Ravi et al. 2009); lamellocyte differentiation (Huang et al. 2005); peripheral nervous system neuron development (Ayyar et al. 2007); plasmatocyte differentiation (Huang et al. 2005); positive regulation of transcription (Park et al. 2003; Roxstrom-Lindquist et al. 2002; Brown et al. 2009); salivary gland histolysis (Lehmann et al. 2002).	immune response, development and morphogenesis
FBti0061529	*FBgn0038084*	0	first intron	*beat-Vc*	no information	-
FBti0062242	*FBgn0003117*	0	3'UTR	*pnr*	pigment metabolic process (Calleja et al. 2002); regulation of glucose metabolic process (Ugrankar et al. 2015), regulation of AMP biosynthesis process (Valanne et al. 2010); development (Calleja et al. 2000; Klinedinst and Bodmer 2003; Alvarez et al. 2003; Mandal et al. 2004; Reim and Frasch 2005; Hainaut et al. 2012; Stern et al. 2009; Qian et a l 2005; Han and Olson 2005; Qian and Bodmer 2009)	immune response, metabolism, development and morphogenesis
FBti0062309	*FBgn0263510/FBgn0050015/ FBgn0050016*	25	5'	*nclb / CG30015 / CG30016*	nclb: germ cell development (Casper et al. 2011) / CG30015 and CG30016: no information	germ cell development
tdn4	*FBgn0034394/FBgn0265661*	500	3'	*CG15096 / CR44468*	CG15096: starvation (Zinke et al. 2002); circadian behavior (Ceriani et al. 2002), mating response (Lawniczak et al. 2004); toxic challenge response (Stern et al. 2012), virus infection (Cordes 2013); immune response (Broderick et al. 2014) / CR44468: no information	xenobiotic stress, immune response, starvation, circadian rhythm, membrane transport
tdn5	*FBgn0032080/FBgn0032079*	0	3'UTR	*CG9525 / CG31886*	*CG9525:* multicellular organism reproduction (Ravi Ram and Wolfner 2007) / *CG31886:* no information.	reproduction
tdn6	*FBgn0003165*	0	intron	*pum*	behavioral response to ethanol (Berger et al. 2008), synapsis (Schweers et al. 2002; Mee et al. 2004; Menon et al. 2004), long-term memory (Dubnau et al. 2003, Akalal et al. 2011), morphogenesis (Ye et al. 2004; Gamberi et al. 2002), development (Kim et al. 2012; Asaoka-Taguchi et al. 1999), positive regulation of nuclear-transcribed mRNA poly(A) tail shortening (Weidmann et al. 2014)	learning or memory, morphogenesis and development, synapsis
tdn7	*FBgn0262964 / FBgn0262565*	50	exon -> esta a 50bp 5' CR43275 and 80 bp CR43105	*CR43275 / CR43105*	no information	-
tdn8	*FBgn0036320*	816	5'	*CG10943*	immune response (Broderick et al. 2014; Roxstrom-Lindquist et al. 2004; Bou-Sleiman et al. 2015)	immune response
tdn12	*FBgn0283709*	>100000	5'	*blue*	no information	-
tdn13	*FBgn0261090*	0	first intron	*Sytbeta*	no information	-
tdn14	*FBgn0051226/FBgn0267193*	350	3'	*CG31226 / CR45633*	no information	-
tdn15	*FBgn0036814*	0	first intron	*CG14073*	wing disc dorsal/ventral pattern formation (Bejarano et al. 2008).	development and morphogenesis
tdn17	*FBgn0028583*	2500	5'	*lcs*	upregulated in young flies compared to old flies guts (Broderick et al. 2014); virus response (Carpenter et al. 2009)	immune response
tdn18	*FBgn0052521/FBgn0031164*	300	3'	*CG32521 / CG1724*	no information	-
tdn19	*FBgn0011971*	>5000	5'	*tRNA:Ser-GCT-2-2*	no information	-

Annex Table S2.4 Allele ratios from ASE analysis, t-test p-values and false discovery rates for each gene and background analyzed. Significant values are depicted in bold. FDR: p-value corrected for multiple testing for FDR (Benjamini and Hochberg 1995).

			Non-infected			Infected		
Gene		**BG**	**Average ratio**	**p-value**	**FDR**	**Average ratio**	**p-value**	**FDR**
CG10943		I	**1.83**	**0.01**	**0.01**	1.53	0.03	0.02194
		II	**1.54**	**4.5 E-4**	**6.3 E-3**	**1.21**	**4.4 E-3**	**0.01**
CG2233		I	**1.74**	**2.8 E-7**	**7.8 E-4**	1.86	0.02	0.02
		II	1.19	0.08	0.03	0.99	0.58	0.05
Dif		I	**1.47**	**2.2 E-3**	**0.01**	1.37	0.02	0.02
		II	0.89	0.45	0.04	0.83	0.09	0.03
AGO2		I	1.05	0.57	0.04	1.05	0.47	0.04
		II	**1.14**	**1.1 E-3**	**0.01**	1.13	0.16	0.03
CG15829		I	1.10	0.68	0.05	1.12	0.54	0.04
		II	**2.11**	**1.1 E-3**	**8.6 E-3**	0.85	0.29	0.04
CG8628		I	1.08	0.09	0.03	0.89	0.39	0.04
		II	**0.54**	**6.7 E-05**	**3.1 E-3**	**0.65**	**2.4 E-6**	**1.56 E-3**
CG8008		I	**0.81**	**1.5 E-4**	**3.9 E-3**	**0.64**	**8.0 E-4**	**7.81 E-3**
		II	0.72	0.07	0.02	**0.43**	**8.3 E-3**	**0.01**
CG15096		I	0.92	0.42	0.04	0.99	0.81	0.04844
		II	**0.57**	**2.8 E-4**	**5.5 E-3**	0.65	9.3 E-3	0.02
Mef2		I	**0.89**	**6.9 E-3**	**0.01**	0.91	0.10	0.03
		II	1.11	0.27	0.04	0.88	0.10	0.03
cbx		I	1.32	0.06	0.02	1.14	0.07	0.02
		II	**0.77**	**6.2 E-5**	**2.3 E-3**	**0.73**	**2.3 E-4**	**4.69 E-3**
Bin1		I	1.09	0.26	0.03	1.27	0.06	0.02
		II	1.05	0.49	0.04	**1.84**	**7.1 E-3**	**0.01**
TM4SF		I	0.90	0.40	0.04	**1.50**	**4.8 E-4**	**7.03 E-3**
		II	0.86	0.45	0.04	0.74	0.02	0.02
NUCB1		I	1.08	0.47	0.04	1.11	0.02	0.02
		II	0.98	0.65	0.05	**0.91**	**4 E-3**	**0.01**
kay		I	1.11	0.22	0.03	1.16	0.15	0.03
		II	1.29	0.02	0.02	1.15	0.14	0.03
ken		I	0.81	0.03	0.02	0.89	0.09	0.03
		II	1.02	0.92	0.05	1.10	0.09	0.03
Jon65Aiv		I	0.92	0.58	0.05	1.11	0.43	0.04
		II	1.01	0.92	0.05	1.11	0.22	0.03

Annex Table S2.5 Strains used in the expriments.

A. DGRP Strains

Strain name	Tlex	TIDAL	ASE	TSS	enhancer assay
RAL-21	x	x	x		
RAL-26	x				
RAL-28	x				
RAL-38	x				
RAL-40	x	x			
RAL-42	x	x			
RAL-45	x				
RAL-49	x				
RAL-57	x				
RAL-59	x	x			
RAL-69	x				
RAL-73	x				
RAL-75	x	x	x		x
RAL-83	x				
RAL-85	x				
RAL-88	x	x			
RAL-91	x				
RAL-93	x				
RAL-101	x				
RAL-105	x				
RAL-109	x				
RAL-129	x				
RAL-136	x				
RAL-138	x				
RAL-142	x	x	x		
RAL-149	x				
RAL-158	x				
RAL-161	x				
RAL-176	x	x			
RAL-177	x	x			
RAL-181	x				
RAL-195	x	x			
RAL-208	x				
RAL-217	x				
RAL-227	x				
RAL-228	x				
RAL-229	x				
RAL-233	x				
RAL-235	x				
RAL-239	x				
RAL-256	x				
RAL-280	x				
RAL-287	x				
RAL-309	x				
RAL-310	x				
RAL-317	x				
RAL-318	x				
RAL-320	x				
RAL-321	x				
RAL-332	x	x			
RAL-338	x				
RAL-350	x	x			
RAL-352	x				
RAL-356	x				
RAL-357	x				
RAL-359	x				
RAL-365	x	x			
RAL-367	x				
RAL-370	x	x			
RAL-371	x	x			
RAL-373	x				
RAL-374	x				
RAL-375	x				
RAL-377	x				
RAL-380	x				
RAL-381	x				
RAL-383	x	x	x		
RAL-391	x				
RAL-392	x				

Strain name	Tlex	TIDAL	ASE	TSS	enhancer assay
RAL-399	x				
RAL-405	x	x	x		x
RAL-426	x				
RAL-439	x				
RAL-440	x				
RAL-441	x	x			
RAL-443	x				
RAL-461	x				
RAL-491	x	x			
RAL-492	x				
RAL-502	x	x	x		
RAL-508	x	x			
RAL-509	x				
RAL-513	x				
RAL-517	x		x		
RAL-531	x				
RAL-535	x				
RAL-555	x	x	x		
RAL-563	x				
RAL-589	x				
RAL-591	x				
RAL-595	x				
RAL-639	x	x			
RAL-642	x				
RAL-646	x				
RAL-703	x				
RAL-705	x				
RAL-707	x	x			
RAL-712	x				
RAL-714	x				
RAL-716	x	x	x		
RAL-721	x				
RAL-727	x				
RAL-730	x				
RAL-732	x				
RAL-737	x	x	x		x
RAL-738	x				
RAL-757	x	x	x		
RAL-761	x				
RAL-765	x				
RAL-776	x	x			
RAL-783	x	x	x	x	x
RAL-787	x	x			
RAL-790	x				
RAL-796	x				
RAL-799	x				
RAL-801	x		x	x	
RAL-802	x				
RAL-804	x				
RAL-805	x				
RAL-808	x				
RAL-810	x	x	x		x
RAL-812	x				
RAL-818	x				
RAL-820	x	x			
RAL-822	x				
RAL-832	x				
RAL-837	x				
RAL-852	x		x		
RAL-855	x	x	x		x
RAL-857	x	x			
RAL-859	x				
RAL-861	x				
RAL-879	x				
RAL-882	x				
RAL-887	x				
RAL-890	x				
RAL-892	x		x		
RAL-894	x	x			
RAL-907	x				
RAL-908	x				
RAL-911	x	x	x	x	

Annex Table S2.5 (continued)

B. African Strains

Strain name	Tlex	TIDAL	ASE	TSS	enhancer assay
ZI10	x				
ZI114N	x				
ZI117	x				
ZI161	x				
ZI184	x				
ZI194	x				
ZI206	x				
ZI207	x				
ZI210	x				
ZI213	x				
ZI214	x				
ZI219	x				
ZI228	x				
ZI230	x				
ZI232	x				
ZI235	x				
ZI237	x				
ZI239	x				
ZI250	x				
ZI252	x				
ZI253	x				
ZI255	x				
ZI264	x				
ZI265	x				
ZI27	x				
ZI271	x				
ZI292	x				
ZI296	x				
ZI303	x				
ZI311N	x				
ZI320	x				
ZI321	x				
ZI324	x				
ZI332	x				
ZI339	x				
ZI341	x				
ZI344	x				
ZI348	x				
ZI357N	x				
ZI364	x				
ZI365	x				
ZI378	x				
ZI379	x				
ZI386	x				
ZI384	x				
ZI398	x				
ZI400	x				
ZI402	x				
ZI418N	x				
ZI420	x				
ZI437	x				
ZI443	x				
ZI445	x				
ZI447	x				
ZI455N	x				
ZI456	x				
ZI457	x				
ZI460	x				
ZI476	x				
ZI477	x				
ZI486	x				
ZI517	x				
ZI76	x				
ZI85	x				
ZI90	x				
ZI99	x				

Annex Table S2.5 (continued)

C. European Strains

Strain name	Population	Tlex	TIDAL	ASE	TSS	enhancer assay
CAS-125	Bari, Italy	x				
CAS-127	Bari, Italy	x				
CAS-145	Bari, Italy	x				
CAS-148	Bari, Italy	x				
CAS-22	Bari, Italy	x				
CAS-33	Bari, Italy	x				
CAS-40	Bari, Italy	x				
CAS-42	Bari, Italy	x				
CAS-49	Bari, Italy	x		x	x	
CAS-50	Bari, Italy	x				
CAS-66	Bari, Italy	x				
CAS-68	Bari, Italy	x				
CAS-69	Bari, Italy	x				
CAS-72	Bari, Italy	x				
CAS-75	Bari, Italy	x				
CAS-52	Bari, Italy	x				
MUN-8	Munich, Germany			x		
STO-1	Stockholm, Sweden	x				
STO-10	Stockholm, Sweden	x				
STO-11	Stockholm, Sweden	x				
STO-12	Stockholm, Sweden	x				
STO-14	Stockholm, Sweden	x				
STO-15	Stockholm, Sweden	x				
STO-16	Stockholm, Sweden	x				
STO-17	Stockholm, Sweden	x				
STO-18	Stockholm, Sweden	x				
STO-19	Stockholm, Sweden	x				
STO-2	Stockholm, Sweden	x				
STO-20	Stockholm, Sweden	x				
STO-21	Stockholm, Sweden	x				
STO-22	Stockholm, Sweden	x				
STO-23	Stockholm, Sweden	x				
STO-24	Stockholm, Sweden	x				
STO-25	Stockholm, Sweden	x				
STO-26	Stockholm, Sweden	x				
STO-27	Stockholm, Sweden	x				
STO-29	Stockholm, Sweden	x				
STO-32	Stockholm, Sweden	x				
STO-33	Stockholm, Sweden	x				
STO-34	Stockholm, Sweden	x				
STO-35	Stockholm, Sweden	x				
STO-36	Stockholm, Sweden	x				
STO-38	Stockholm, Sweden	x				
STO-39	Stockholm, Sweden	x				
STO-4	Stockholm, Sweden	x				
STO-40	Stockholm, Sweden	x				
STO-41	Stockholm, Sweden	x				
STO-42	Stockholm, Sweden	x				
STO-44	Stockholm, Sweden	x				
STO-45	Stockholm, Sweden	x				
STO-46	Stockholm, Sweden	x				
STO-47	Stockholm, Sweden	x				
STO-48	Stockholm, Sweden	x				
STO-53	Stockholm, Sweden	x				
STO-6	Stockholm, Sweden	x				
STO-7	Stockholm, Sweden	x				
STO-8	Stockholm, Sweden	x				
STO-9	Stockholm, Sweden	x				
STO-63	Stockholm, Sweden	x				
STO-56	Stockholm, Sweden	x				
STO-50	Stockholm, Sweden	x				
STO-61	Stockholm, Sweden	x				
STO-60	Stockholm, Sweden	x				
STO-51	Stockholm, Sweden	x				
STO-62	Stockholm, Sweden	x				
STO-59	Stockholm, Sweden	x				
STO-53	Stockholm, Sweden	x				
STO-57	Stockholm, Sweden	x				
STO-58	Stockholm, Sweden	x				
STO-55	Stockholm, Sweden	x				
STO-52	Stockholm, Sweden	x				

Annex Table S2.5 (continued)

D. Mutant Strains

Gene	Stock number	Stock center	Genotype	References
CG2233	10089	VDRC	w1118 Mi{ET1}CG2233MB00881	Metaxakis et al. 2005; Bellen et al. 2011
TM4SF	8846	VDRC	w1118; P{GD3820}v8846	-
CG15829	104642	VDRC	P{KK111610}VIE-260B	Dietzl et al. 2007
Bin1	33574	BDSC	w1118; P{EP}Bin1G4692	Bellen et al. 2011
CG8008	25488	BDSC	w^{1118}; Mi{ET1}$CG8008^{MB06245}$	Metaxakis et al. 2005; Bellen et al. 2011
NUCB1	10581	BDSC	w1118; PBac{PB}NUCB1c01508	Bellen et al., 2004; Thibault et al., 2004
CG10943	56051	BDSC	y1 w*; Mi{MIC}MI08278	Venken et al. 2011
ken	11244	BDSC	cn1 P{PZ}ken02970/CyO; ry506	Spradling et al. (1999)
cbx	10067	BDSC	w1118; PBac{PB}cbxc00428	Bellen et al., 2004; Thibault et al., 2004
cn^1	263	BDSC	cn1	-
w^{1118}	-	From X. Franch-Marro lab	w^{1118}	-
Act5c-GAL4	4414	BDSC	y[1] w[*]; P{w[+mC]=Act5C-GAL4}25FO1/CyO, y[+]	-
Act5c-GAL4/ TubGAL80ts	-	From X. Franch-Marro lab	Ifi/CyO; (Act5c-GAL4 TubP-GAL80[ts])/SMTM	-

Annex Table S2.6: Analysis of the SNPs in the coding regions of the genes analyzed in the ASE (A), as well as the 1 kb TE flanking regions conserved between *D. melanogaster* and *D. yakuba* (B). *These SNPs are linked to the presence of the TE *FBti0020119*.

A: SNPs present in the gene-coding region. Only genes with missense amino acid changes are shown. The rest of the genes contained only synonymous SNPs.

Gene	SNP Ensembl ID	SNP Location	Alleles	Type	Amino Acid change	ASE strains
CG2233	ENSVDME05971278	X:8037060	T/G	Missense variant	Glu/Asp	RAL-892 (TE+ background II)
	ENSVDME05971279	X:8037116	T/A	Missense variant	Lys/Asp	
	ENSVDME05971280	X:8037118	T/C	Missense variant		
	ENSVDME05971281	X:8037129	T/G	Missense variant	Lys/Thr	
	ENSVDME05971282	X:8037132	T/G	Missense variant	Tyr/Ser	
	ENSVDME05971285	X:8037186	G/A	Missense variant	Thre/Ile	
	ENSVDME05971286	X:8037189	G/T	Missense variant	Thr/Asn	
	ENSVDME05971261	X:8036984	T/C	Synonymous variant	Ser/Gly	SNPs different in both TE+ and TE- strains
	ENSVDME05971241	X:8036599	T/C	Missense variant	Arg/Lys	
	ENSVDME05971241	X:8036599	T/C	Missense variant	Gly/Ser	
	ENSVDME05971236	X:8036549	C/T	Missense variant	Arg/Lys	
	ENSVDME05971235	X:8036513	A/T	Missense variant	Met/Leu	
	ENSVDME05971232	X:8036379	G/C	Missense variant	Leu/Phe	
	ENSVDME05971223	X:8035966	G/A	Missense variant	Ile/Thr	
Bin1	ENSVDME04579137	3R:16186278	T/C	Missense variant	Ala/Thr	RAL-801 (TE- background I) and RAL-21 (TE+ background II): Ala; RAL-75 (TE- background II) and RAL-911 (TE+ background I): Thr
CG8628	ENSVDME03237467	3L:8129688	G/T	Missense variant	Lys/Asp	RAL-405 (TE+ background II)
CG10943	ENSVDME03582113	3L:12862670	T/C	Missense variant	Gly/Lys	SNPs different in both TE+ and TE- strains
	ENSVDME03582101	3L:12862417	A/T	Missense variant	His/Phe	
	ENSVDME03582102	3L:12862418	A/G	Missense variant		
	ENSVDME03582093	3L:12862287	A/T	Missense variant	Glu/Asp	
	ENSVDME03582089	3L:12862173	C/A	Missense variant	Gln/Lys	RAL-75 (TE- bacground II)
	ENSVDME03582105	3L:12862456	G/T	Missense variant	Asn/Thr	RAL-405 (TE+ bacground II)
	ENSVDME03582104	3L:12862452	T/G	Missense variant	Asn/Lys	RAL-405 (TE+ bacground II)
NUCB1	ENSVDME03857267	3L:17798314	C/A	Missense variant	His/Gln	SNPs different in both TE+ and TE- strains
AGO2	ENSVDME03742113	3L:15556853	C/G	Missense variant	Arg/Gly	RAL-801 (TE+ background I)
	ENSVDME03742119	3L:15557202	T/A	Synonymous variant	Leu/Gln	RAL-757 (TE- background II)
	ENSVDME03742158	3L:15559281	G/A	Missense variant	Asn/Ser	SNPs different in both TE+ and TE- strains*
	ENSVDME03742167	3L:15559531	T/A	Missense variant	Asp/Gln	

B: SNPs present in the 1kb TE flanking regions.

Gene	TE	*D. melanogaster* coordinates	Sequence length	Conservation with *D. yakuba*	Location
AGO2	*FBti0020119*	3L: 15519466-15519481	16bp	75,00%	exon
		3L: 15519643-15519731	89bp	74,20%	exon
		3L: 15519832-15519924	93bp	75,30%	intron
		3L: 15522348-15522359	12bp	75,00%	exon
		3L: 15522360-15522406	47bp	87,20%	exon
		3L: 15522660-15522677	18bp	88,90%	exon
		3L: 15522696-15522706	11bp	72,70%	exon
		3L: 15522707-15522765	71bp	70,40%	exon
		3L: 15522914-15522936	23bp	87,00%	exon
Bin1	*FBti0019386*	3R: 12014432-12015173	771bp	83,80%	intron
		3R: 12015556-12015731	176bp	79,50%	intron
		3R: 12015734-12015881	150bp	79,30%	intron
		3R: 12016090-12016357	268bp	95,90%	intron
cbx	*FBti0019985*	2R: 6013225-6013399	175bp	89,10%	exon
Jon65Aiv	*FBti0020046*	3L: 6013478-6014031	554bp	87,00%	exon
		3L: 6013874-6014031	158bp	91,80%	exon
		3L: 6016723-6016839	117bp	91,50%	intergenic
		3L: 6016906-6017625	749bp	86,50%	intergenic
		3L: 7103232-7103366	153bp	73,20%	intergenic
CG15829 and *CG8628*	*FBti0020057*	3L: 7103731-7103846	116bp	87,10%	intergenic
		3L: 7104544-7104934	397bp	84,40%	intergenic
		3L: 7875736-7876156	440bp	82,50%	intergenic
CG2233	*FBti0019602*	X: 7876164-7876250	87bp	95,40%	intergenic
		X: 7876388-7876610	253bp	70,00%	intergenic
		X: 7881230-7881395	166bp	88,60%	exon
		X: 7881457-7881863	407bp	87,00%	exon
		X: 25585247-25585983	824bp	78,00%	intergenic
kay	*FBti0019457*	3R: 25587524-25588073	600bp	79,00%	intergenic
		3R: 5456773-5456891	121bp	91,70%	intron
Mef2	*FBti0018877*	2R: 5456894-5457198	305bp	88,90%	intron
		2R: 5457533-5457661	131bp	80,20%	intron
		2R: 5457677-5457784	108bp	100,00%	intron
		2R: 5457836-5458042	217bp	79,30%	intron
		2R: 5458164-5458264	111bp	74,80%	intron
		2R: 14321781-14322403	655bp	85,20%	intergenic
CG15096	*tdn4*	2R: 14322404-14322609	210bp	80,50%	intergenic
		2R: 14322673-14322794	123bp	81,30%	intergenic
		2R: 14322975-14323071	100bp	71,00%	intergenic
		2R: 14323201-14323251	53bp	81,10%	UTR
		2R: 14323256-14323359	104bp	96,20%	UTR
		2R: 14323360-14323386	27bp	92,60%	UTR
		2R: 14323390-14323454	66bp	87,90%	UTR
		2R: 14323475-14323638	164bp	97,00%	exon
		2R: 12836801-12836901	101bp	70,30%	intergenic
CG10943	*tdn8*	3L: 12837023-12837839	830bp	84,60%	intergenic
		3L: 12838159-12838507	349bp	81,10%	intergenic
		3L: 17421939-17422525	594bp	93,30%	intron
Dif	*FBti0061506*	2L: 17422539-17422637	100bp	70,00%	intron
		2L: 17422773-17422966	200bp	82,50%	intron
		2L: 17423088-17423194	113bp	75,20%	intron
		2L: 17423269-17423584	317bp	83,60%	intron
		2L: 17423585-17424119	562bp	78,80%	UTR
		2L: 4662006-4662188	183bp	88,00%	exon
CG8008	*FBti0018883*	2R: 4662249-4662384	136bp	93,40%	exon
		2R: 4662448-4662774	327bp	93,90%	exon
		2R: 4662834-4662990	157bp	91,10%	exon
		2R: 4663054-4663116	63bp	93,70%	exon
		2R: 4670137-4670410	274bp	84,30%	UTR
		2R: 4670425-4670482	58bp	89,70%	exon
		2R: 4670587-4670694	108bp	94,40%	exon
		2R: 19384369-19384377	9bp	100,00%	exon
TM4SF and *ken*	*FBti0018868*	X: 19384378-19384716	353bp	91,50%	UTR
		X: 19384665-19385042	380bp	88,40%	UTR
		X: 19384718-19384817	102bp	87,30%	UTR
		X: 19384818-19384819	2bp	100,00%	UTR
		X: 19384826-19385042	217bp	84,30%	UTR
		X: 19385231-19385339	109bp	73,40%	intergenic
		X: 19386059-19386491	433bp	91,20%	UTR
		X: 19386493-19386635	143bp	98,60%	exon
		X: 19386679-19386838	175bp	81,70%	UTR
		X: 17763880-17764757	878bp	92,40%	exon
NUCB1	*FBti0020137*	3L: 17764300-17764758	459bp	93,70%	exon
		3L: 17764759-17764811	53bp	96,20%	exon
		3L: 17766930-17767136	224bp	74,60%	intron
		3L: 17767152-17767393	242bp	90,90%	exon
		3L: 17767394-17767494	101bp	82,20%	UTR
		3L: 17767495-17767711	219bp	79,00%	intergenic

Annex Table S2.7. Primers used for the TE screening. To detect the presence/absence of the TE annotated in the reference sequence, two PCRs were performed: one PCR using the flanking primer (FL) and right primer (R), and the other PCR using the left primer (L) binding into the TE sequence, and the right primer (R). To detect the de novo TE insertions described in Rahman et al. 2015, we used only a flanking primer (F) and a reverse primer (R).

FBti0019386 FL	TTTGGAATCAATCACATCAACCC
FBti0019386 L	TTCGCATTCCAGAAATTCCTTCT
FBti0019386 R	CAATGTCCTGGGTGTAAGTCTCG
FBti0018883 FL	AGTGGTTGGCAGTACCATCG
FBti0018883 L	ATCAGACGCCAACCAGAGTG
FBti0018883 R	GCATAGCAAACACATCTCCGC
FBti0019985 FL	GGCATCATAAAACCGTTGAACAC
FBti0019985 L	AGTCCCTTAGTGGGAGACCACAG
FBti0019985 R	CGTAGGATCAGTGGGTGAAAATG
FBti0018868 FL	AGAGGAAGAGTGGGTGGTGTA
FBti0018868 L	GTCCAAACCAGCCACTTCCA
FBti0018868 R	TCTTGCGGATGCCTGTCTTT
FBti0061506 FL	TGCCATTCCAGTTCCCAGTC
FBti0061506 L	TGGCCGTTACGCATCTTGT
FBti0061506 R	TAGTGACCTGTTTTGCGGCT
FBti0020057 FL	AACAATAGGGTGGCGGATGT
FBti0020057 L	CAATAACAGTAACATAACAGCGCA
FBti0020057 R	GGAGATAGCCCCCGGATACA
FBti0019602 FL	ACGTTCACTGGACACCCATC
FBti0019602 L	AAATTGCTTGGAGCCCCGTT
FBti0019602 R	CAATCTGTGCCCCTCGATGT
FBti0020137 FL	TCGTTGTCGTGGTCCAGATG
FBti0020137 L	GCTTTAAGCACGTTTGATCAGC
FBti0020137 R	CCGACATTCGGGTGAGTAGG
FBti0019457 FL	CTTTGCTTTGTTCGGTGCGA
FBti0019457 L	TGGGTTTGCAGCAATTAAGGC
FBti0019457 R	CATTGCTCGAGTTCCCCGAT
FBti0018877 FL	TAGTTTCTCTGGGGGTGGCT
FBti0018877 L	CACATGATTAGTGAGAGGTTTGGT
FBti0018877 R	TTCCAGTTCAATAGGGCGCA
FBti0020119	GCTCCATAAACTTTCGAAATGCC
FBti0020119	AGCTAAAGCCAATGGGGAACATA
FBti0020119	TGTACCTGCTGTTTGCCTTGTTT
FBti0020046	TGGCTCGTGTTGAGTAAATGCTT
FBti0020046	ACCTATCTGGACTTATGGCTCCG
FBti0020046	GGCATCTAGGAAGGAGTCAGGAA

FBti0019381 FL	GGTGCGTGTCTCTGCTAAGT
FBti0019381 L	TAAATTGCTTGGAGCCCCGT
FBti0019381 R	ACATTTTGGATTGCTCCGGC
FBti0019564 FL	ATCCGCCGAAAATCTCCTCC
FBti0019564 R	GTTGGCAGCTAGGACGAACT
FBti0061105 FL	AGACCACCTTGACTGACTGAAC
FBti0061105 R	GGCATGCTGGGGATTCACTAT

Primers to detect de novo TEs (not annotated in the reference genome):

tdn1 F	TTGCAGATGGCCAAGAACTGC
tdn1 R	CAGGAAGAGCAAATGGCAGCA
tdn2 F	GTGAGTTTGTGGCAGGTGTG
tdn2 R	CGCTAACGAGGCGTGGTAAA
tdn3 F	TTCTTGCGGTTGCAAAACGA
tdn3 R	TCTGATAAGCGATTGGCGGT
tdn4 F	GTCTGCAATCTTTGCTGCGT
tdn4 R	ACTAATAGCAGGCCCCAACT
tdn5 F	ATTTCTTGACGCATCCCGCT
tdn5 R	AAAGCACTAGGTGCCATCCAT
tdn6 F	GGTTTCTGTGGTCTTGCCGT
tdn6 R	CGGTCTGCTGTCGCTCAAT
tdn7 F	CTTCCTTCTGCGACCGTAGT
tdn7 R	CTAACGCTTGTAGGCCAGGT
tdn8 FL	TTCGCTGGCGTCAGAAAATG
tdn8 R	TTCATTGGCCCCGGATATGG
tdn9 F	GAGGGGCCAACAACGACTAC
tdn9 R	TTGCTCCGCAATTTATGGGC
tdn10 F	GGATGGGATGGGATGGCTAC
tdn10 R	AACCAGAACAAGCGCAAACA
tdn11 F	GCAATTCATTCGGCAGCAAC
tdn11 R	AGCAGTCAGACACAAGTCGAA
tdn12 F	GTTGGCGATGTAAGTGCTGG
tdn12 R	TGCTTAAGATGCTGGAAGGCA
tdn13 F	ACTTGTTGCCTTGTGCGTTG
tdn13 R	AACAAAAAGTTGCTGGCGCA
tdn14 F	GGCACATCGCCTTGTTCATC
tdn14 R	GTAAGGGGTCGTGAGGGTC

tdn15 F	CATCAATACTAAGGTCGCTGCT
tdn15 R	TTCTTCGTCGTCTGTTGCCT
tdn16 F	TCGCTTTTGTATTTGCGGCT
tdn16 R	TGGAGAGGCCAACGAAAACA
tdn17 FL	ATTGGCCGTGGAGGTAAGTG
tdn17 R	ACCGGCATTCTCAATTGCAC
tdn18 FL	GGGTGGCTGGCTACTCAAA
tdn18 R	GCTCATGCGCGTTTTAATTGT
tdn19 F	CTCTTGCCACCCTCTTGACT
tdn19 R	AATTACGCGGTGCTGACATT
tdn20 F	ACAATCAACCAAATCCAAGAACG
tdn20 R	ATTTGATGAGCTTGTCGCAGC
tdn21 F	ATGCTCCGCTATGTGGCAAG
tdn21 R	AGGTGGCGAAGGTAGGAGAT
tdn22 F	CGTTTTCCCGCTTCAGCATT
tdn22 R	CGGGCAAATGTATCCACAGC
tdn23 F	ACATCCACACAACCACACCG
tdn23 R	TAAAATGGCCGCTCGCTGAA
tdn24 F	TCAGCGTTTGTTGGTTGTCG
tdn24 R	GCTACCGAGGTGAACACGAA
tdn25 F	ACATGTAGCTCCGGCCAATC
tdn25 R	TCTGGGTGGCTCAATTGGTG
tdn26 F	AGACTGCGATCTGGTTGTGT
tdn26 R	GAAGCCAACGGTCAAATGGT

Annex Table S2.8 Fly strains and SNPs used in the ASE crosses for each gene.
Females from the strains used in each cross are depicted in bold.

		1st background		2nd background			
Gene	**TE**	**TE+ strain**	**TE- strain**	**TE+ strain**	**TE- strain**	**SNP Ensembl ID**	**SNP**
AGO2	*FBti0020119*	RAL-801	**RAL-383**	**RAL-517**	RAL-757	ENSVDME03742174	C/T
NUCB1	*FBti0020137*	**RAL-383**	RAL-801	**RAL-852**	RAL-855	ENSVDME03857294	C/T
Jon65Aiv	*FBti0020046*	**RAL-911**	RAL-801	**RAL-517**	RAL-757	ENSVDME03101876	A/G
Bin1	*FBti0019386*	**RAL-911**	RAL-801	**RAL-21**	RAL-75	ENSVDME04579137	T/C
CG15829	*FBti0020057*	RAL-801	**RAL-911**	RAL-142	**RAL-757**	ENSVDME03169056	C/T
CG8628	*FBti0020057*	RAL-801	**RAL-911**	RAL-405	**RAL-75**	ENSVDME03169255	T/C
cbx	*FBti0019985*	**RAL-810**	RAL-855	**RAL-555**	RAL-757	ENSVDME01620861	T/G
CG2233	*FBti0019602*	**RAL-855**	RAL-502	RAL-892	**RAL-852**	ENSVDME05971216	A/C
kay	*FBti0019457*	**RAL-502**	RAL-801	**RAL-21**	RAL-75	ENSVDME05370341	T/C
Mef2	*FBti0018877*	**RAL-502**	RAL-801	RAL-142	**RAL-757**	ENSVDME01622694	A/G
CG15096	*tdn4*	RAL-810	**RAL-911**	**RAL-75**	RAL-405	ENSVDME02200235	A/T
CG10943	*tdn8*	RAL-716	**RAL-810**	RAL-405	**RAL-75**	ENSVDME01153579	T/G
Dif	*FBti0061506*	RAL-737	**RAL-855**	RAL-142	**RAL-757**	ENSVDME01153579	C/T
CG8008	*FBti0018883*	**RAL-75**	RAL-757	**RAL-852**	RAL-855	ENSVDME01586592	T/C
TM4SF	*FBti0018868*	**RAL-783**	CAS-49	**RAL-855**	MU-6	ENSVDME02541918	T/C
ken	*FBti0018868*	**RAL-783**	CAS-49	**RAL-852**	MU-6	ENSVDME02541755	A/G

Annex Table S2.9. Primers used for the amplification of the genomic regions analyzed in the enhancer assay. Primers include in their 5' region the restriction site for each enzyme used in the cloning process.

Primer ID	Primer sequence (5'-3')	Enzyme	*D. mel* v6 coordinates
tdn8_F	GGTACCTCGAAATCGTTGCAGTCACA	KpnI	3L: 12862729-12863886
tdn8_R	GCGGCCGCTTGGCCCCGGATATGGACTA	NotI	
18868_F	AAGCTTGCCCAGATGCCAACAAGTATATTTC	HindIII	2R: 23877727-23878275
18868_R	GGTACCTCGTATGGGGTGCTTAATTGAT	KpnI	
19985_F	AAGCTTCGACGTTTCTCTGCGGACTA	HindIII	2R: 9871040-9871567
19985_R	GGTACCACGAGAAGACAGCGTAGATCG	KpnI	
61506_F	GGTACCTTTTTGCGGTCCAGGAATGTG	KpnI	2L: 17431911-17432162
61506_R	AAGCTTGTAAACGACCTGTTTCCACCT	HindIII	

Annex Table S2.10. Primers used for qRT-PCR gene expression quantification.

Primer ID	Primer sequence (5'-3')	*D. mel* v6 coordinates
Act5c_RT_forward	GAGCAGTTGGAATCGGGTTTTAC	chr2R: 9873109-9873269
Act5c_RT_reverse	GTATGAATCGCAGTCCAGC	
lacZ_RT_forward	CCTGCTGATGAAGCAGAACAACT	-
lacZ_RT_reverse	CACCACATACAGGCCGTAGC	
CG15829_RT_forward	TGTCGAAGCCAACTAGCCAAC	chr3L: 7128923-7129109
CG15829_RT_reverse	GCCGTAGAACTCGAGGAACT	
CG2233_RT_forward	TCTCCTTTGCCAAGTACGCA	chrX: 8036600-8036794
CG2233_RT_reverse	GGGACAACTTAACGATATCGGACT	
Bin1_RT_forward	TGTCGTCCCGTAGAGCAGAA	chr3R: 16186595-16186733
Bin1_RT_reverse	CAAGCAGATTGACCGCGAGA	
TM4SF_RT_forward	GCAGCGAGGATAACGGGAAA	chr2R: 23878871-23879684
TM4SF_RT_reverse	AGTAGACCGAGTGACCCCAG	

Bibliography Annex Table 2.3

Abbott MK, Lengyel JA. 1991. Embryonic head involution and rotation of male terminalia require the *Drosophila* locus head involution defective. Genetics 129:783–789.

Akalal D-BG, Yu D, Davis RL. 2011. The long-term memory trace formed in the *Drosophila* α/β mushroom body neurons is abolished in long-term memory mutants. J. Neurosci. 31:5643–5647.

Alvarez AD, Shi W, Wilson B a, Skeath JB. 2003. pannier and pointedP2 act sequentially to regulate *Drosophila* heart development. Development 130:3015–3026.

Alves-Silva J, Sanchez-Soriano N, Beaven R, Klein M, Parkin J, Millard TH, Bellen HJ, Venken KJT, Ballestrem C, Kammerer RA, et al. 2012. Spectraplakins Promote Microtubule-Mediated Axonal Growth by Functioning As Structural Microtubule-Associated Proteins and EB1-Dependent +TIPs (Tip Interacting Proteins). J. Neurosci. 32:9143–9158.

Andrew DJ, Baker BS. 2008. Expression of the *Drosophila* secreted cuticle protein 73 (dsc73) requires *Shavenbaby*. Dev. Dyn. 237:1198–1206.

Anholt R, Mackay T. 2001. The Genetic Architecture of Odor-Guided Behavior in *Drosophila melanogaster*. Behav. Genet. 31:17–27.

Arbouzova NI, Bach EA, Zeidler MP. 2006. *Ken* & *Barbie* selectively regulates the expression of a subset of JAK/STAT pathway target genes. Curr. Biol. 16:80–88.

Armstrong JD, Texada MJ, Munjaal R, Baker DA, Beckingham KM. 2006. Gravitaxis in *Drosophila melanogaster*: A forward genetic screen. Genes, Brain Behav. 5:222–239.

Asaoka-Taguchi M, Yamada M, Nakamura a, Hanyu K, Kobayashi S. 1999. Maternal Pumilio acts together with Nanos in germline development in *Drosophila* embryos. Nat. Cell Biol. 1:431–437.

Avidor-Reiss T, Maer AM, Koundakjian E, Polyanovsky A, Keil T, Subramaniam S, Zuker CS. 2004. Decoding cilia function: Defining specialized genes required for compartmentalized cilia biogenesis. Cell 117:527–539.

Ayres JS, Freitag N, Schneider DS. 2008. Identification of drosophila mutants altering defense of and endurance to *Listeria monocytogenes* infection. Genetics 178:1807–1815.

Ayroles JF, Carbone MA, Stone EA, Jordan KW, Lyman RF, Magwire MM, Rollmann SM, Duncan LH, Lawrence F, Anholt RRH, et al. 2009. Systems genetics of complex traits in *Drosophila melanogaster*. Nat. Genet. 41:299–307.

Azad P, Zhou D, Zarndt R, Haddad GG. 2012. Identification of Genes Underlying Hypoxia Tolerance in *Drosophila* by a P-element Screen. G3: Genes|Genomes|Genetics 2:1169–1178.

Barrio R, López-Varea A, Casado M, de Celis JF. 2007. Characterization of dSnoN and its relationship to Decapentaplegic signaling in *Drosophila*. Dev. Biol. 306:66–81.

Bashaw GJ, Hu H, Nobes CD, Goodman CS. 2001. A novel Dbl family RhoGEF promotes Rho-dependent axon attraction to the central nervous system midline in *Drosophila* and overcomes Robo repulsion. J. Cell Biol. 155:1117–1122.

Bejarano F, Luque CM, Herranz H, Sorrosal G, Rafel N, Pham TT, Milán M. 2008. A gain-of-function suppressor screen for genes involved in dorsal-ventral boundary formation in the *Drosophila* wing. Genetics 178:307–323.

Berger KH, Kong EC, Dubnau J, Tully T, Moore MS, Heberlein U. 2008. Ethanol sensitivity and tolerance in long-term memory mutants of *Drosophila melanogaster*. Alcohol. Clin. Exp. Res. 32:895–908.

Berkey CD, Blow N, Watnick PI. 2009. Genetic analysis of *Drosophila melanogaster* susceptibility to intestinal *Vibrio cholerae* infection. Cell. Microbiol. 11:461–474.

Berry B, Deddouche S, Kirschner D, Imler JL, Antoniewski C. 2009. Viral suppressors of RNA silencing hinder exogenous and endogenous small RNA pathways in *Drosophila*. Preiss T, editor. PLoS One 4:e5866.

Besnard-Guérin C, Jacquier C, Pidoux J, Deddouche S, Antoniewsk C. 2015. The cricket paralysis virus suppressor inhibits microRNA silencing mediated by the drosophila argonaute-2 protein. PLoS One 10:e0120205.

Blanchard FJ, Collins B, Cyran SA, Hancock DH, Taylor M V, Blau J. 2010. The transcription factor *Mef2* is required for normal circadian behavior in *Drosophila*. J. Neurosci. 30:5855–5865.

Bochdanovits Z, de Jong G. 2004. Antagonistic pleiotropy for life-history traits at the gene expression level. Proc. Biol. Sci. 271 Suppl:S75–S78.

Bogwitz MR, Chung H, Magoc L, Rigby S, Wong W, O'Keefe M, McKenzie JA, Batterham P, Daborn PJ. 2005. Cyp12a4 confers lufenuron resistance in a natural population of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. U. S. A. 102:12807–12812.

Boquet I, Hitier R, Dumas M, Chaminade M, Préat T. 2000. Central brain postembryonic development in *Drosophila*: Implication of genes expressed at the interhemispheric junction. J. Neurobiol. 42:33–48.

Bosch M, Serras F, Martín-Blanco E, Baguñà J. 2005. JNK signaling pathway required for wound healing in regenerating *Drosophila* wing imaginal discs. Dev. Biol. 280:73–86.

Bosveld F, Rana A, Van Der Wouden PE, Lemstra W, Ritsema M, Kampinga HH, Sibon OCM. 2008. De novo CoA biosynthesis is required to maintain DNA integrity during development of the *Drosophila* nervous system. Hum. Mol. Genet. 17:2058–2069.

Bottenberg W, Sanchez-Soriano N, Alves-Silva J, Hahn I, Mende M, Prokop A. 2009. Context-specific requirements of functional domains of the Spectraplakin Short stop in vivo. Mech. Dev. 126:489–502.

Bou Sleiman MS, Osman D, Massouras A, Hoffmann AA, Lemaitre B, Deplancke B. 2015. Genetic, molecular and physiological basis of variation in *Drosophila* gut immunocompetence. Nat Commun 6:7829.

Bour BA, O'Brien MA, Lockwood WL, Goldstein ES, Bodmer R, Taghert PH, Abmayr SM, Nguyen HT. 1995. *Drosophila MEF2*, a transcription factor that is essential for myogenesis. Genes Dev. 9:730–741.

Broderick NA, Buchon N, Lemaitre B. 2014. Microbiota-Induced Changes in *Drosophila melanogaster* Host Gene Expression and Gut Morphology. MBio 5:e01117-14.

Bronstein R, Levkovitz L, Yosef N, Yanku M, Ruppin E, Sharan R, Westphal H, Oliver B, Segal D. 2010. Transcriptional regulation by CHIP/LDB complexes. PLoS Genet. 6(8):e1001063.

Brown AE, Baumbach J, Cook PE, Ligoxygakis P. 2009. Short-term starvation of immune deficient *Drosophila* improves survival to gram-negative bacterial infections. PLoS One 4:e4490.

Brown EB, Layne JE, Zhu C, Jegga AG, Rollmann SM. 2013. Genome-wide association mapping of natural variation in odour-guided behaviour in *Drosophila*. Genes, Brain Behav. 12:503–515.

Brunetti TM, Fremin BJ, Cripps RM. 2015. Identification of singles bar as a direct transcriptional target of *Drosophila Myocyte enhancer factor-2* and a regulator of adult myoblast fusion. Dev. Biol. 401:299–309.

Bryantsev AL, Baker PW, Lovato TL, Jaramillo MS, Cripps RM. 2012. Differential requirements for Myocyte Enhancer Factor-2 during adult myogenesis in *Drosophila*. Dev. Biol. 361:191–207.

Calleja M, Renaud O, Usui K, Pistillo D, Morata G, Simpson P. 2002. How to pattern an epithelium: Lessons from achaete-scute regulation on the notum of *Drosophila*. Gene 292:1–12.

Campbell JL, Gu Q, Guo D, Nash HA. 2009. Genetic effects in *Drosophila* on the potency of diverse general anesthetics: a distinctive pattern of altered sensitivity. J. Neurogenet. 23:412–421.

Carpenter J, Hutter S, Baines JF, Roller J, Saminadin-Peter SS, Parsch J, Jiggins FM. 2009. The transcriptional response of *Drosophila melanogaster* to infection with the sigma virus (Rhabdoviridae). PLoS One 4.

Carreira VP, Soto IM, Mensch J, Fanara JJ, Kerszberg M, Wolpert L, Celis J de, Blair S, Celis J de, Diaz-Benjumea F, et al. 2011. Genetic basis of wing morphogenesis in *Drosophila*: sexual dimorphism and non-allometric effects of shape variation. BMC Dev. Biol. 11:32.

Carrera P, Moshkin YM, Grönke S, Silljé HHW, Nigg EA, Jäckle H, Karch F. 2003. Tousled-like kinase functions with the chromatin assembly pathway regulating nuclear divisions. Genes Dev. 17:2578–2590.

Casper AL, Baxter K, Van Doren M. 2011. No Child Left Behind Encodes a Novel Chromatin Factor Required for Germline Stem Cell Maintenance in Males But Not Females. Development 138:3357–3366.

Castrillon DH, Gonczy P, Alexander S, Rawson R, Eberhart CG, Viswanathan S, DiNardo S, Wasserman SA. 1993. Toward a molecular genetic analysis of spermatogenesis in Drosophila melanogaster: Characterization of male-sterile mutants generated by single P element mutagenesis. Genetics 135:489–505.

Ceriani MF, Hogenesch JB, Yanovsky M, Panda S, Straume M, Kay SA. 2002. Genome-wide expression analysis in *Drosophila* reveals genes controlling circadian behavior. J. Neurosci. 22:9305–9319.

Cerrato A, Parisi M, Anna SS, Missirlis F, Guru S, Agarwal S, Sturgill D, Talbot T, Spiegel A, Collins F, et al. 2006. Genetic interactions between Drosophila melanogaster menin and Jun/Fos. Dev. Biol. 298:59–70.

Chahine S, O'Donnell MJ. 2009. Physiological and molecular characterization of methotrexate transport by Malpighian tubules of adult Drosophila melanogaster. J. Insect Physiol. 55:927–935.

Chang KT, Shi Y-J, Min K-T. 2003. The *Drosophila* homolog of Down's syndrome critical region 1 gene regulates learning: implications for mental retardation. Proc. Natl. Acad. Sci. U. S. A. 100:15794–15799.

Chen F, Barkett M, Ram KT, Quintanilla A, Hariharan IK. 1997. Biological characterization of *Drosophila Rapgap1*, a GTPase activating protein for Rap1. Proc. Natl. Acad. Sci. U. S. A. 94:12485–12490.

Chen S, Spletter M, Ni X, White KP, Luo L, Long M. 2012. Frequent Recent Origination of Brain Genes Shaped the Evolution of Foraging Behavior in *Drosophila*. Cell Rep. 1:118–132.

Chihara CJ, Song C, LaMonte G, Fetalvero K, Hinchman K, Phan H, Pineda M, Robinson K, Schneider GP. 2005. Identification and partial characterization of the enzyme of omega: one of five putative *DPP IV* genes in *Drosophila melanogaster*. J. Insect Sci. 5:26.

Chintapalli VR, Wang J, Herzyk P, Davies SA, Dow JAT. 2013. Data-mining the FlyAtlas online resource to identify core functional motifs across transporting epithelia. BMC Genomics 14:518.

Chow CY, Wolfner MF, Clark AG. 2013. Using natural variation in *Drosophila* to discover previously unknown endoplasmic reticulum stress genes. Proc. Natl. Acad. Sci. U. S. A. 110:9013–9018.

Christofi T, Apidianakis Y. 2013. *Drosophila* immune priming against *Pseudomonas aeruginosa* is short-lasting and depends on cellular and humoral immunity. F1000Research 2:76.

Clark RI, Tan SWS, Péan CB, Roostalu U, Vivancos V, Bronda K, Pilátová M, Fu J, Walker DW, Berdeaux R, et al. 2013. XMEF2 is an in vivo immune-metabolic switch. Cell 155.

Clemente-Ruiz M, Muzzopappa M, Milán M. 2014. Tumor suppressor roles of *CENP-E* and *Nsl1* in *Drosophila* epithelial tissues. Cell Cycle 13:1450–1455.

Combe BE, Defaye A, Bozonnet N, Puthier D, Royet J, Leulier FF. 2014. *Drosophila* microbiota modulates host metabolic gene expression via IMD/NF-KB signaling. PLoS One 9:e94729.

Cordes EJ, Licking-Murray KD, Carlson KA. 2013. Differential gene expression related to *Nora* virus infection of *Drosophila melanogaster*. Virus Res. 175:95–100.

Cornwell WD, Kirkpatrick RB. 2001. Cactus-independent nuclear translocation of *Drosophila RELISH*. J. Cell. Biochem. 82:22–37.

Costa E, Beltran S, Espins ML. 2011. *Drosophila melanogaster SAP18* protein is required for environmental stress responses. FEBS Lett. 585:275–280.

Culi J, Aroca P, Modolell J, Mann RS. 2006. jing is required for wing development and to establish the proximo-distal axis of the leg in Drosophila melanogaster. Genetics 173:255–266.

Cullen K, McCall K. 2004. Role of programmed cell death in patterning the *Drosophila* antennal arista. Dev. Biol. 275:82–92.

Dason JS, Romero-Pozuelo J, Marin L, Iyengar BG, Klose MK, Ferrús A, Atwood HL. 2009. Frequenin/NCS-1 and the Ca2+-channel alpha1-subunit co-regulate synaptic transmission and nerve-terminal growth. J. Cell Sci. 122:4109–4121.

DeFalco TJ, Verney G, Jenkins AB, McCaffery JM, Russell S, Van Doren M. 2003. Sex-specific apoptosis regulates sexual dimorphism in the *Drosophila* embryonic gonad. Dev. Cell 5:205–216.

Dembeck LM, Huang W, Magwire MM, Lawrence F, Lyman RF, Mackay TFC. 2015. Genetic Architecture of Abdominal Pigmentation in *Drosophila melanogaster*. PLOS Genet. 11:e1005163.

Deshpande G, Calhoun G, Schedl P. 2005. *Drosophila argonaute-2* is required early in embryogenesis for the assembly of centric/centromeric heterochromatin, nuclear division, nuclear migration, and germ-cell formation. Genes Dev. 19:1680–1685.

DeZazzo J, Sandstrom D, de Belle S, Velinzon K, Smith P, Grady L, DelVecchio M, Ramaswami M, Tully T. 2000. *nalyot*, a mutation of the *Drosophila myb-related Adf1* transcription factor, disrupts synapse formation and olfactory memory. Neuron 27:145–158.

Dijkers PF, O'Farrell PH. 2009. Dissection of a Hypoxia-induced, Nitric Oxide–mediated Signaling Cascade. Mol. Biol. Cell 20:4083–4090.

Djiane A, Shimizu H, Wilkin M, Mazleyrat S, Jennings MD, Avis J, Bray S, Baron M. 2011. Su(dx) E3 ubiquitin ligase-dependent and -independent functions of polychaetoid, the *Drosophila* ZO-1 homologue. J. Cell Biol. 192:189–200.

Dorner S, Lum L, Kim M, Paro R, Beachy PA, Green R. 2006. A genomewide screen for components of the RNAi pathway in *Drosophila* cultured cells. Proc. Natl. Acad. Sci. U. S. A. 103:11880–11885.

Dubnau J, Chiang AS, Grady L, Barditch J, Gossweiler S, McNeil J, Smith P, Buldoc F, Scott R, Certa U, et al. 2003. The *staufen/pumilio* pathway is involved in drosophila long-term memory. Curr. Biol. 13:286–296.

Edwards AC, Zwarts L, Yamamoto A, Callaerts P, Mackay TFC. 2009. Mutations in many genes affect aggressive behavior in *Drosophila melanogaster.* BMC Biol. 7:29.

Egli D, Yepiskoposyan H, Selvaraj A, Balamurugan K, Rajaram R, Simons A, Multhaup G, Mettler S, Vardanyan A, Georgiev O, et al. 2006. A family knockout of all four *Drosophila* metallothioneins reveals a central role in copper homeostasis and detoxification. Mol. Cell. Biol. 26:2286.

Ejima A, Tsuda M, Takeo S, Ishii K, Matsuo T, Aigaki T. 2004. Expression level of *sarah*, a homolog of DSCR1, is critical for ovulation and female courtship behavior in *Drosophila melanogaster*. Genetics 168:2077–2087.

Elgar SJ, Han J, Taylor M V. 2008. *mef2* activity levels differentially affect gene expression during Drosophila muscle development. Proc. Natl. Acad. Sci. U. S. A. 105:918–923.

Evans TA, Haridas H, Duffy JB. 2009. Kekkon5 is an extracellular regulator of BMP signaling. Dev. Biol. 326:36–46.

Fabrizio JJ, Hime G, Lemmon SK, Bazinet C. 1998. Genetic dissection of sperm individualization in *Drosophila melanogaster*. Development 125:1833–1843.

Fagegaltier D, Bougé A-L, Berry B, Poisot E, Sismeiro O, Coppée J-Y, Théodore L, Voinnet O, Antoniewski C. 2009. The endogenous siRNA pathway is involved in heterochromatin formation in *Drosophila*. Proc. Natl. Acad. Sci. U. S. A. 106:21258–21263.

Firdaus H, Mohan J, Naz S, Arathi P, Ramesh SR, Nongthomba U. 2015. A cis-regulatory mutation in troponin-i of *Drosophila* reveals the importance of proper stoichiometry of structural proteins during muscle assembly. Genetics 200:149–165.

Franciscovich AL, Vrailas Mortimer AD, Freeman AA, Gu J, Sanyal S. 2008. Overexpression screen in drosophila identifies neuronal roles of GSK-3B/shaggy as a regulator of AP-1-dependent developmental plasticity. Genetics 180:2057–2071.

Franco B, Bogdanik L, Bobinnec Y, Debec A, Bockaert J, Parmentier M-L, Grau Y. 2004. Shaggy, the Homolog of Glycogen Synthase Kinase 3, Controls Neuromuscular Junction Growth in *Drosophila*. J. Neurosci. 24:6573–6577.

Franke JD, Boury AL, Gerald NJ, Kiehart DP. 2006. Native nonmuscle myosin II stability and light chain binding in *Drosophila melanogaster*. Cell Motil. Cytoskeleton 63:604–622.

Fukuyama H, Verdier Y, Guan Y, Makino-Okamura C, Shilova V, Liu X, Maksoud E, Matsubayashi J, Haddad I, Spirohn K, et al. 2013. Landscape of protein-protein interactions in *Drosophila*

immune deficiency signaling during bacterial challenge. Proc Natl Acad Sci U S A 110:10717–10722.

Fulga TA, McNeill EM, Binari R, Yelick J, Blanche A, Booker M, Steinkraus BR, Schnall-Levin M, Zhao Y, DeLuca T, et al. 2015. A transgenic resource for conditional competitive inhibition of conserved *Drosophila* microRNAs. Nat Commun 6:7279.

Furlong EE, Andersen EC, Null B, White KP, Scott MP. 2001. Patterns of gene expression during *Drosophila* mesoderm development. Science 293:1629–1633.

Galletti M, Riccardo S, Parisi F, Lora C, Saqcena MK, Rivas L, Wong B, Serra A, Serras F, Grifoni D, et al. 2009. Identification of domains responsible for ubiquitin-dependent degradation of *dMyc* by *glycogen synthase kinase 3beta* and *casein kinase 1* kinases. Mol. Cell. Biol. 29:3424–3434.

Gamberi C, Peterson DS, He L, Gottlieb E. 2002. An anterior function for the *Drosophila* posterior determinant *Pumilio*. Development 129:2699–2710.

Gao F-BFB, Brenman JE, Jan LY, Jan YN. 1999. Genes regulating dendritic outgrowth, branching, and routing in *Drosophila*. 13:2549–2561.

Garcia MC, Abbasi M, Singh S, He Q. 2007. Role of *Drosophila* gene *dunc-115* in nervous system. Invertebr. Neurosci. 7:119–128.

Ghezzi A, Krishnan HR, Lew L, Prado FJ, Ong DS, Atkinson NS. 2013. Alcohol-Induced Histone Acetylation Reveals a Gene Network Involved in Alcohol Tolerance. PLoS Genet. 9(12):e1003986.

Gobert V, Gottar M, Matskevich AA, Rutschmann S, Royet J, Belvin M, Hoffmann JA, Ferrandon D. 2003. Dual activation of the *Drosophila* Toll pathway by two pattern recognition receptors. Science. 302:2126–2130.

Gorski SM, Chittaranjan S, Pleasance ED, Freeman JD, Anderson CL, Varhol RJ, Coughlin SM, Zuyderduyn SD, Jones SJM, Marra MA. 2003. A SAGE approach to discovery of genes involved in autophagic cell death. Curr. Biol. 13:358–363.

Greenberg L, Hatini V. 2011. Systematic expression and loss-of-function analysis defines spatially restricted requirements for *Drosophila RhoGEFs* and *RhoGAPs* in leg morphogenesis. Mech. Dev. 128:5–17.

De Gregorio E, Spellman PT, Tzou P, Rubin GM, Lemaitre B. 2002. The Toll and Imd pathways are the major regulators of the immune response in *Drosophila*. EMBO J. 21:2568–2579.

Grether ME, Abrams JM, Agapite J, White K, Steller H. 1995. The head involution defective gene of Drosophila melanogaster functions in programmed cell death. Genes Dev. 9:1694–1708.

Grima DP, Sullivan M, Zabolotskaya M V, Browne C, Seago J, Wan KC, Okada Y, Newbury SF. 2008. The 5'-3' exoribonuclease *pacman* is required for epithelial sheet sealing in *Drosophila* and genetically interacts with the phosphatase puckered. Biol. Cell 100:687–701.

Guha A, Kornberg TB. 2005. Tracheal branch repopulation precedes induction of the *Drosophila* dorsal air sac primordium. Dev. Biol. 287:192–200.

Guio L, Barrõn MG, González J. 2014. The transposable element Bari-Jheh mediates oxidative stress response in *Drosophila*. Mol. Ecol. 23:2020–2030.

Haglund K, Nezis IP, Lemus D, Grabbe C, Wesche J, Liestøl K, Dikic I, Palmer R, Stenmark H. 2010. *Cindr* Interacts with *Anillin* to Control Cytokinesis in *Drosophila melanogaster*. Curr. Biol. 20:944–950.

Hainaut M, Sagnier T, Berenger H, Pradel J, Graba Y, Miotto B. 2012. The MYST-containing protein chameau is required for proper sensory organ specification during drosophila thorax morphogenesis. PLoS One 7:e32882.

Haining WN, Carboy-Newcomb C, Wei CL, Steller H. 1999. The proapoptotic function of *Drosophila Hid* is conserved in mammalian cells. Proc. Natl. Acad. Sci. U. S. A. 96:4936–4941.

Hammond SM, Boettcher S, Caudy AA, Kobayashi R, Hannon GJ, Fire A, Hammond SM, Caudy AA, Hannon GJ, Hamilton AJ, et al. 2001. *Argonaute2*, a link between genetic and biochemical analyses of RNAi. Science 293:1146–1150.

Han Y-H, Luo Y-J, Wu Q, Jovel J, Wang X-H, Aliyari R, Han C, Li W-X, Ding S-W. 2011. RNA-based immunity terminates viral infection in adult *Drosophila* in the absence of viral suppression of RNA interference: characterization of viral small interfering RNA populations in wild-type and mutant flies. J. Virol. 85:13153–13163.

Han Z, Olson EN. 2005. Hand is a direct target of *Tinman* and GATA factors during *Drosophila* cardiogenesis and hematopoiesis. Development 132:3525–3536.

Hong W, Mosca TJ, Luo L. 2012. *Teneurins* instruct synaptic partner matching in an olfactory map. Nature 484:201–207.

Honjo K, Furukubo-Tokunaga K. 2005. Induction of cAMP response element-binding protein-dependent medium-term memory by appetitive gustatory reinforcement in *Drosophila* larvae. J. Neurosci. 25:7905–7913.

Honjo K, Furukubo-Tokunaga K. 2009. Distinctive neuronal networks and biochemical pathways for appetitive and aversive memory in *Drosophila* larvae. J. Neurosci. 29:852–862.

Horner VL, Czank A, Jang JK, Singh N, Williams BC, Puro J, Kubli E, Hanes SD, McKim KS, Wolfner MF, et al. 2006. The *Drosophila* Calcipressin *Sarah* Is Required for Several Aspects of Egg Activation. Curr. Biol. 16:1441–1446.

Hou YCC, Chittaranjan S, Barbosa SG, McCall K, Gorski SM. 2008. Effector caspase *Dcp-1* and IAP protein Bruce regulate starvation-induced autophagy during *Drosophila melanogaster* oogenesis. J. Cell Biol. 182:1127–1139.

Huang C, Zheng X, Zhao H, Li M, Wang P, Xie Z, Wang L, Zhong Y. 2012. A permissive role of mushroom body α/β core neurons in long-term memory consolidation in *Drosophila*. Curr. Biol. 22:1981–1989.

Huang L, Ohsako S, Tanda S. 2005. The lesswright mutation activates *Rel*-related proteins, leading to overproduction of larval hemocytes in *Drosophila melanogaster*. Dev. Biol. 280:407–420.

Hyun J, Bécam I, Yanicostas C, Bohmann D. 2006. Control of G2/M transition by *Drosophila Fos*. Mol. Cell. Biol. 26:8293–8302.

Ishizuka A, Siomi MC, Siomi H. 2002. A *Drosophila* fragile X protein interacts with components of RNAi and ribosomal proteins. Genes Dev. 16:2497–2508.

Iyer EPR, Iyer SC, Sullivan L, Wang D, Meduri R, Graybeal LL, Cox DN. 2013. Functional Genomic Analyses of Two Morphologically Distinct Classes of *Drosophila* Sensory Neurons: Post-Mitotic Roles of Transcription Factors in Dendritic Patterning. PLoS One 8.

Jemc JC, Milutinovich AB, Weyers JJ, Takeda Y, Van Doren M. 2012. Raw Functions through JNK signaling and cadherin-based adhesion to regulate *Drosophila* gonad morphogenesis. Dev. Biol. 367:114–125.

Jennings MD, Blankley RT, Baron M, Golovanov AP, Avis JM. 2007. Specificity and autoregulation of *notch* binding by tandem wwdomains in suppressor of *deltex*. J. Biol. Chem. 282:29032–29042.

Jia J, Amanai K, Wang G, Tang J, Wang B, Jiang J. 2002. *Shaggy/GSK3* antagonizes *Hedgehog* signalling by regulating *Cubitus interruptus*. Nature 416:548–552.

Jia J, Zhang L, Zhang Q, Tong C, Wang B, Hou F, Amanai K, Jiang J. 2005. Phosphorylation by double-time/CKIIepsilon and CKIalpha targets *Cubitus Interruptus* for Slimb/beta-TRCP-mediated proteolytic processing. Dev. Cell 9:819–830.

Jiang C, Lamblin A-FF, Steller H, Thummel CS. 2000. A Steroid-triggered Transcriptional Hierarchy Controls Salivary Gland Cell Death during *Drosophila* Metamorphosis. Mol. Cell 5:445–455.

Johnson RI, Cagan RL. 2009. A quantitative method to analyze *Drosophila* pupal eye patterning. PLoS One 4:e7008.

Johnson RI, Seppa MJ, Cagan RL. 2008. The *Drosophila* CD2AP/CIN85 orthologue *Cindr* regulates junctions and cytoskeleton dynamics during tissue patterning. J. Cell Biol. 180:1191–1204.

Jordan KC, Schaeffer V, Fischer KA, Gray EE, Ruohola-Baker H. 2006. Notch signaling through tramtrack bypasses the mitosis promoting activity of the JNK pathway in the mitotic-to-endocycle transition of *Drosophila* follicle cells. BMC Dev. Biol. 6:16.

Jung AC, Ribeiro C, Michaut L, Certa U, Affolter M. 2006. Polychaetoid/ZO-1 Is Required for Cell Specification and Rearrangement during *Drosophila* Tracheal Morphogenesis. Curr. Biol. 16:1224–1231.

Kadener S, Villella A, Kula E, Palm K, Pyza E, Botas J, Hall JC, Rosbash M. 2006. Neurotoxic protein expression reveals connections between the circadian clock and mating behavior in *Drosophila*. Proc. Natl. Acad. Sci. U. S. A. 103:13537–13542.

Kairamkonda S, Nongthomba U. 2014. Beadex function in the motor neurons is essential for female reproduction in *Drosophila melanogaster*. PLoS One 9(11):e113003.

Kamyshev NG, Iliadi KG, Bragina Y V, Savvateeva-Popova E V, Tokmacheva E V, Preat T. 2000. Identification of *Drosophila* mutant with memory defects after acquisition of conditioned reflex suppression of courtship. Neurosci. Behav. Physiol. 30:307–313.

Kanuka H, Kuranaga E, Takemoto K, Hiratou T, Okano H, Miura M. 2005. *Drosophila* caspase transduces Shaggy/GSK-3beta kinase activity in neural precursor development. EMBO J. 24:3793–3806.

Kaplan NA, Colosimo PF, Liu X, Tolwinski NS. 2011. Complex interactions between *GSK3* and *aPKC* in *Drosophila* embryonic epithelial morphogenesis. PLoS One 6:e18616.

Kaplan NA, Liu X, Tolwinski NS. 2009. Epithelial polarity: Interactions between junctions and apical-basal machinery. Genetics 183:897–904.

Kent D, Bush EW, Hooper JE. 2006. *Roadkill* attenuates *Hedgehog* responses through degradation of *Cubitus interruptus*. Development 133:2001–2010.

Kiger AA, Baum B, Jones S, Jones MR, Coulson A, Echeverri C, Perrimon N. 2003. A functional genomic analysis of cell morphology using RNA interference. J. Biol. 2:27.

Kim K, Lee YS, Carthew RW. 2007. Conversion of pre-RISC to holo-RISC by *Ago2* during assembly of RNAi complexes. RNA 13:22–29.

Kim S, Wairkar YP, Daniels RW, DiAntonio A. 2010. The novel endosomal membrane protein Ema interacts with the class C Vps-HOPS complex to promote endosomal maturation. J. Cell Biol. 188:717–734.

Kim SY, Kim JY, Malik S, Son W, Kwon KS, Kim C. 2012. Negative regulation of EGFR/MAPK pathway by *pumilio* in *Drosophila melanogaster*. PLoS One 7(4):e34016.

Kim T, Yoon JJH, Cho H, Lee W-B, Kim J, Song Y-H, Kim SN, Yoon JJH, Kim-Ha J, Kim Y-J. 2005. Downregulation of lipopolysaccharide response in *Drosophila* by negative crosstalk between the AP1 and NF-kappaB signaling modules. Nat. Immunol. 6:211–218.

King EG, Kislukhin G, Walters KN, Long AD. 2014. Using *Drosophila melanogaster* to identify chemotherapy toxicity Genes. Genetics 198:31–43.

Klarsfeld A, Malpel S, Michard-Vanhée C, Picot M, Chélot E, Rouyer F. 2004. Novel features of cryptochrome-mediated photoreception in the brain circadian clock of *Drosophila*. J. Neurosci. 24:1468–1477.

Kleino A, Valanne S, Ulvila J, Kallio J, Myllymäki H, Enwald H, Stöven S, Poidevin M, Ueda R, Hultmark D, et al. 2005a. Inhibitor of apoptosis 2 and TAK1-binding protein are components of the Drosophila Imd pathway. EMBO J. 24:3423–3434.

Kleino A, Valanne S, Ulvila J, Kallio J, Myllymäki H, Enwald H, Stöven S, Poidevin M, Ueda R, Hultmark D, et al. 2005b. Inhibitor of apoptosis 2 and TAK1-binding protein are components of the *Drosophila* Imd pathway. EMBO J. 24:3423–3434.

Klinedinst SL, Bodmer R. 2003. Gata factor Pannier is required to establish competence for heart progenitor formation. Development 130:3027–3038.

Kong EC, Allouche L, Chapot PA, Vranizan K, Moore MS, Heberlein U, Wolf FW. 2010. Ethanol-regulated genes that contribute to ethanol sensitivity and rapid tolerance in *Drosophila*. Alcohol. Clin. Exp. Res. 34:302–316.

Kurada P, White K. 1998. Ras promotes cell survival in *Drosophila* by downregulating hid expression. Cell 95:319–329.

Kurusu M, Cording A, Taniguchi M, Menon K, Suzuki E, Zinn K. 2008. A Screen of Cell-Surface Molecules Identifies Leucine-Rich Repeat Proteins as Key Mediators of Synaptic Target Selection. Neuron 59:972–985.

De La Cova C, Abril M, Bellosta P, Gallant P, Johnston LA. 2004. *Drosophila myc* regulates organ size by inducing cell competition. Cell 117:107–116.

Lansdell SJ, Millar NS. 2000. The influence of nicotinic receptor subunit composition upon agonist, alpha-bungarotoxin and insecticide (imidacloprid) binding affinity. Neuropharmacology 39:671–679.

Lawniczak MK, Begun DJ. 2004. A genome-wide analysis of courting and mating responses in *Drosophila melanogaster* females. Genome 47:900–910.

Leaman D, Po YC, Fak J, Yalcin A, Pearce M, Unnerstall U, Marks DS, Sander C, Tuschl T, Gaul U. 2005. Antisense-mediated depletion reveals essential and specific functions of microRNAs in *Drosophila* development. Cell 121:1097–1108.

Lee M. 2003. Distinct sites in E-cadherin regulate different steps in *Drosophila* tracheal tube fusion. Development 130:5989–5999.

Lee S, Kolodziej PA. 2002. The plakin Short Stop and the RhoA GTPase are required for E-cadherin-dependent apical surface remodeling during tracheal tube fusion. Development 129:1509–1520.

Lehmann M, Jiang C, Ip YT, Thummel CS. 2002. AP-1, but not NF-kappa B, is required for efficient steroid-triggered cell death in *Drosophila*. Cell Death Differ. 9:581–590.

Léonard R, Rendić D, Rabouille C, Wilson IBH, Préat T, Altmann F. 2006. The *Drosophila* fused lobes gene encodes an N-acetylglucosaminidase involved in N-glycan processing. J. Biol. Chem. 281:4867–4875.

Li H-H, Chiang C-S, Huang H-Y, Liaw G-J. 2009. mars and tousled-like kinase act in parallel to ensure chromosome fidelity in *Drosophila*. J. Biomed. Sci. 16:51.

Lin C, Koval A, Tishchenko S, Gabdulkhakov A, Tin U, Solis GP, Katanaev VL. 2014. Double Suppression of the Galpha Protein Activity by RGS Proteins. Mol. Cell 53:663–671.

Ling J, Dubruille R, Emery P. 2012. KAYAK-α modulates circadian transcriptional feedback loops in Drosophila pacemaker neurons. J. Neurosci. 32:16959–16970.

Liu J, Ghanim M, Xue L, Brown CD, Iossifov I, Angeletti C, Hua S, Nègre N, Ludwig M, Stricker T, et al. 2009. Analysis of *Drosophila* segmentation network identifies a JNK pathway factor overexpressed in kidney cancer. Science 323:1218–1222.

Liu Y, Montell DJ. 2001. Jing: a downstream target of slbo required for developmental control of border cell migration. Development 128:321–330.

Lovato TL, Benjamin AR, Cripps RM. 2005. Transcription of *Myocyte enhancer factor-2* in adult *Drosophila* myoblasts is induced by the steroid hormone ecdysone. Dev. Biol. 288:612–621.

Lukacsovich T, Yuge K, Awano W, Asztalos Z, Kondo S, Juni N, Yamamoto D. 2003. The *ken* and *barbie* gene encoding a putative transcription factor with a BTB domain and three zinc finger motifs functions in terminalia development of *Drosophila*. Arch. Insect Biochem. Physiol. 54:77–94.

Luschnig S, Moussian B, Krauss J, Desjeux I, Perkovic J, Nüsslein-Volhard C. 2004. An F1 genetic screen for maternal-effect mutations affecting embryonic pattern formation in *Drosophila melanogaster*. Genetics 167:325–342.

Mandal L, Banerjee U, Hartenstein V. 2004. Evidence for a fruit fly hemangioblast and similarities between lymph-gland hematopoiesis in fruit fly and mammal aorta-gonadal-mesonephros mesoderm. Nat. Genet. 36:1019–1023.

Martinek S, Inonog S, Manoukian AS, Young MW. 2001. A role for the segment polarity gene shaggy/GSK-3 in the Drosophila circadian clock. Cell 105:769–779.

Mateo L, Ullastres A, González J. 2014. A Transposable Element Insertion Confers Xenobiotic Resistance in *Drosophila*. PLoS Genet. 10.

Mathieu J, Sung HH, Pugieux C, Soetaert J, Rorth P. 2007. A sensitized PiggyBac-based screen for regulators of border cell migration in *Drosophila*. Genetics 176:1579–1590.

Matranga C, Tomari Y, Shin C, Bartel DP, Zamore PD. 2005. Passenger-strand cleavage facilitates assembly of siRNA into *Ago2*-containing RNAi enzyme complexes. Cell 123:607–620.

Matyash A, Singh N, Hanes SD, Urlaub H, Jäckle H. 2009. SAP18 promotes Krüppel-dependent transcriptional repression by enhancer-specific histone deacetylation. J. Biol. Chem. 284:3012–3020.

McCabe BD, Hom S, Aberle H, Fetter RD, Marques G, Haerry TE, Wan H, O'Connor MB, Goodman CS, Haghighi AP. 2004. Highwire regulates presynaptic BMP signaling essential for synaptic growth. Neuron 41:891–905.

McClure KD, Schubiger G. 2008. A screen for genes that function in leg disc regeneration in *Drosophila melanogaster*. Mech. Dev. 125:67–80.

McGraw LA, Gibson G, Clark AG, Wolfner MF. 2004. Genes regulated by mating, sperm, or seminal proteins in mated female *Drosophila melanogaster*. Curr. Biol. 14:1509–1514.

Mee CJ. 2004. Regulation of Neuronal Excitability through Pumilio-Dependent Control of a Sodium Channel Gene. J. Neurosci. 24:8695–8703.

Menon DU, Meller VH. 2012. A role for siRNA in X-chromosome dosage compensation in *Drosophila melanogaster*. Genetics 191:1023–1028.

Menon KP, Sanyal S, Habara Y, Sanchez R, Wharton RP, Ramaswami M, Zinn K. 2004. The translational repressor *Pumilio* regulates presynaptic morphology and controls postsynaptic accumulation of translation factor eIF-4E. Neuron 44:663–676.

Menon SD, Osman Z, Chenchill K, Chia W. 2005. A positive feedback loop between dumbfounded and rolling pebbles leads to myotube enlargement in *Drosophila*. J. Cell Biol. 169:909–920.

Meyer WJ, Schreiber S, Guo Y, Volkmann T, Welte MA, Müller HAJ. 2006. Overlapping functions of argonaute proteins in patterning and morphogenesis of *Drosophila* embryos. PLoS Genet. 2:1224–1239.

Milán M, Diaz-Benjumea FJ, Cohen SM. 1998. *Beadex* encodes an LMO protein that regulates Apterous LIM-homeodomain activity in *Drosophila* wing development: A model for LMO oncogene function. Genes Dev. 12:2912–2920.

Mitri C, Parmentier ML, Pin JP, Bockaert J, Grau Y. 2004. Divergent evolution in metabotropic glutamate receptors: A new receptor activated by an endogenous ligand different from glutamate in insects. J. Biol. Chem. 279:9313–9320.

Mitri C, Soustelle L, Framery B, Bockaert J, Parmentier ML, Grau Y. 2009. Plant insecticide L-canavanine repels *Drosophila* via the insect orphan GPCR DmX.Bellen HJ, editor. PLoS Biol. 7:e1000147.

Miyazu M, Tanimura T, Sokabe M. 2000. Molecular cloning and characterization of a putative cyclic nucleotide-gated channel from *Drosophila melanogaster*. Insect Mol Biol 9:283–292.

Mohit P, Makhijani K, Madhavi MB, Bharathi V, Lal A, Sirdesai G, Reddy VR, Ramesh P, Kannan R, Dhawan J, et al. 2006. Modulation of AP and DV signaling pathways by the homeotic gene Ultrabithorax during haltere development in *Drosophila*. Dev. Biol. 291:356–367.

Moon NS, Di Stefano L, Morris EJ, Patel R, White K, Dyson NJ. 2008. E2F and p53 induce apoptosis independently during *Drosophila* development but intersect in the context of DNA damage. PLoS Genet. 4:e1000153.

Mueller S, Gausson V, Vodovar N, Deddouche S, Troxler L, Perot J, Pfeffer S, Hoffmann JA, Saleh M-C, Imler J-L. 2010. RNAi-mediated immunity provides strong protection against the negative-strand RNA vesicular stomatitis virus in *Drosophila*. Proc. Natl. Acad. Sci. U. S. A. 107:19390–19395.

Mummery-Widmer JL, Yamazaki M, Stoeger T, Novatchkova M, Bhalerao S, Chen D, Dietzl G, Dickson BJ, Knoblich JA. 2009. Genome-wide analysis of Notch signalling in *Drosophila* by transgenic RNAi. Nature 458:987–992.

Nakagawa T, Ikehara T, Doiguchi M, Imamura Y, Higashi M, Yoneda M, Ito T. 2015. Enhancer of acetyltransferase chameau (EAChm) is a novel transcriptional co-activator. PLoS One 10:e0142305.

Nakamura M, Baldwin D, Hannaford S, Palka J, Montell C. 2002. *Defective proboscis extension response* (DPR), a member of the Ig superfamily required for the gustatory response to salt. J. Neurosci. 22:3463–3472.

Neely GG, Hess A, Costigan M, Keene AC, Goulas S, Langeslag M, Griffin RS, Belfer I, Dai F, Smith SB, et al. 2010. A Genome-wide *Drosophila* screen for heat nociception identifies α2δ3 as an evolutionarily conserved pain gene. Cell 143:628–638.

Neumüller RA, Richter C, Fischer A, Novatchkova M, Neumüller KG, Knoblich JA. 2011. Genome-wide analysis of self-renewal in *Drosophila* neural stem cells by transgenic RNAi. Cell Stem Cell 8:580–593.

Nguyen T, Wang J, Schulz RA. 2002. Mutations within the conserved *MADS* box of the *D-MEF2* muscle differentiation factor result in a loss of DNA binding ability and lethality in Drosophila. Differentiation 70:438–446.

Okamura K, Chung WJ, Ruby JG, Guo H, Bartel DP, Lai EC. 2008. The *Drosophila* hairpin RNA pathway generates endogenous short interfering RNAs. Nature 453:803–806.

Okamura K, Ishizuka A, Siomi H, Siomi MC. 2004. Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. Genes Dev. 18:1655–1666.

Park JM, Kim JM, Kim LK, Kim SN, Kim-Ha J, Kim JH, Kim YJ. 2003. Signal-induced transcriptional activation by *Dif* requires the dTRAP80 mediator module. Mol Cell Biol 23:1358–1367.

Park SK, Sedore SA, Cronmiller C, Hirsh J. 2000. Type II cAMP-dependent protein kinase-deficient *Drosophila* are viable but show developmental, circadian, and drug response phenotypes. J. Biol. Chem. 275:20588–20596.

Parrish JZ, Kim MD, Lily YJ, Yuh NJ. 2006. Genome-wide analyses identify transcription factors required for proper morphogenesis of *Drosophila* sensory neuron dendrites. Genes Dev. 20:820–835.

Perrimon N, Smouse D. 1989. Multiple functions of a *Drosophila* homeotic gene, zeste-white 3, during segmentation and neurogenesis. Dev. Biol. 135:287–305.

Price MA, Kalderon D. 2002. Proteolysis of the Hedgehog signaling effector Cubitus interruptus requires phosphorylation by Glycogen Synthase Kinase 3 and Casein Kinase 1. Cell 108:823–835.

Przewloka MR, Zhang W, Costa P, Archambault V, D'Avino PP, Lilley KS, Laue ED, McAinsh AD, Glover DM. 2007. Molecular analysis of core kinetochore composition and assembly in *Drosophila melanogaster*. PLoS One 2.

Pueyo JI, Galindo MI, Bishop S a, Couso JP. 2000. Proximal-distal leg development in *Drosophila* requires the apterous gene and the Lim1 homologue *dlim1*. Development 127:5391–5402.

Pulipparacharuvil S, Akbar MA, Ray S, Sevrioukov E a, Haberman AS, Rohrer J, Krämer H. 2005. *Drosophila* Vps16A is required for trafficking to lysosomes and biogenesis of pigment granules. J. Cell Sci. 118:3663–3673.

Qian L, Bodmer R. 2009. Partial loss of GATA factor Pannier impairs adult heart function in *Drosophila*. Hum. Mol. Genet. 18:3153–3163.

Qian L, Liu J, Bodmer R. 2005. Neuromancer Tbx20-related genes (H15/midline) promote cell fate specification and morphogenesis of the *Drosophila* heart. Dev. Biol. 279:509–524.

Ram KR, Wolfner MF. 2007. Seminal influences: *Drosophila Acps* and the molecular interplay between males and females during reproduction. Integr. Comp. Biol. 47:427–445.

Ramel MC, Emery CS, Foulger R, Goberdhan DCI, van den Heuvel M, Wilson C. 2007. *Drosophila SnoN* modulates growth and patterning by antagonizing TGF-beta signalling. Mech. Dev. 124:304–317.

Rämet M, Lanot R, Zachary D, Manfruelli P. 2002. JNK Signaling Pathway Is Required for Efficient Wound Healing in *Drosophila*. Dev. Biol. 241:145–156.

Rand TA, Petersen S, Du F, Wang X. 2005. *Argonaute2* cleaves the anti-guide strand of siRNA during RISC activation. Cell 123:621–629.

Ravi D, Wiles AM, Bhavani S, Ruan J, Leder P, Bishop AJR. 2009. A network of conserved damage survival pathways revealed by a genomic RNAi screen. PLoS Genet. 5:e1000527.

Rehwinkel J, Behm-Ansmant I, Gatfield D, Izaurralde E. 2005. A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing. RNA 11:1640–1647.

Reim I, Frasch M. 2005. The Dorsocross T-box genes are key components of the regulatory network controlling early cardiogenesis in *Drosophila*. Development 132:4911–4925.

Ren J, Zhu H, Chi C, Mehrmohamadi M, Deng K, Wu X, Xu T. 2014. Beadex affects gastric emptying in *Drosophila*. Cell Res. 24:636–639.

Reumer A, Bogaerts A, Van Loy T, Husson SJ, Temmerman L, Choi C, Clynen E, Hassan B, Schoofs L. 2009. Unraveling the protective effect of a Drosophila phosphatidylethanolamine-binding protein upon bacterial infection by means of proteomics. Dev. Comp. Immunol. 33:1186–1195.

Reuter JE, Nardine TM, Penton A, Billuart P, Scott EK, Usui T, Uemura T, Luo L. 2003. A mosaic genetic screen for genes necessary for Drosophila mushroom body neuronal morphogenesis. Development 130:1203–1213.

Rogina B, Reenan RA, Nilsen SP, Helfand SL. 2000. Extended life-span conferred by cotransporter gene mutations in *Drosophila*. Science 290:2137–2140.

Romero-Pozuelo J, Dason JS, Atwood HL, Ferrús A. 2007. Chronic and acute alterations in the functional levels of Frequenins 1 and 2 reveal their roles in synaptic transmission and axon terminal morphology. Eur. J. Neurosci. 26:2428–2443.

Röper K, Brown NH. 2004. A Spectraplakin Is Enriched on the Fusome and Organizes Microtubules during Oocyte Specification in *Drosophila*. Curr. Biol. 14:99–110.

Rosenbaum EE, Hardie RC, Colley NJ. 2006. Calnexin is essential for rhodopsin maturation, Ca2+ regulation, and photoreceptor cell survival. Neuron 49:229–241.

Rosenbaum EE, Vasiljevic E, Brehm KS, Colley NJ. 2014. Mutations in Four Glycosyl Hydrolases Reveal a Highly Coordinated Pathway for Rhodopsin Biosynthesis and N-Glycan Trimming in Drosophila melanogaster.Desplan C, editor. PLoS Genet. 10:e1004349.

Ross J, Jiang H, Kanost MR, 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.

Roxström-Lindquist K, Lindström-Dinnetz I, Olesen J, Engström Y, Faye I. 2002. An intron enhancer activates the immunoglobulin-related Hemolin gene in Hyalophora cecropia. Insect Mol. Biol. 11:505–515.

Roxström-Lindquist K, Terenius O, Faye I. 2004. Parasite-specific immune response in adult *Drosophila melanogaster*: a genomic study. EMBO Rep. 5:207–212.

Rusten TE, Rodahl LMW, Pattni K, Englund C, Samakovlis C, Dove S, Brech A, Stenmark H. 2006. Fab1 Phosphatidylinositol 3-Phosphate 5-Kinase Controls Trafficking but Not Silencing of Endocytosed Receptors. Mol. Biol. Cell 17:3989–4001.

Rusten TE, Vaccari T, Lindmo K, Rodahl LMW, Nezis IP, Sem-Jacobsen C, Wendler F, Vincent JP, Brech A, Bilder D, et al. 2007. ESCRTs and Fab1 Regulate Distinct Steps of Autophagy. Curr. Biol. 17:1817–1825.

Rutschmann S, Jung AC, Hetru C, Reichhart JM, Hoffmann JA, Ferrandon D. 2000. The *Rel* protein DIF mediates the antifungal but not the antibacterial host defense in *Drosophila*. Immunity 12:569–580.

Rynes J, Donohoe CD, Frommolt P, Brodesser S, Jindra M, Uhlirova M. 2012. Activating transcription factor 3 regulates immune and metabolic homeostasis. Mol Cell Biol 32:3949–3962.

Sabin LR, Zhou R, Gruber JJ, Lukinova N, Bambina S, Berman A, Lau CK, Thompson CB, Cherry S. 2009. *Ars2* Regulates Both miRNA- and siRNA- Dependent Silencing and Suppresses RNA Virus Infection in *Drosophila*. Cell 138:340–351.

Saleh M-C, van Rij RP, Hekele A, Gillis A, Foley E, O'Farrell PH, Andino R. 2006. The endocytic pathway mediates cell entry of dsRNA to induce RNAi silencing. Nat. Cell Biol. 8:793–802.

Sambandan D, Yamamoto A, Fanara JJ, Mackay TFC, Anholt RRH. 2006. Dynamic genetic interactions determine odor-guided behavior in *Drosophila melanogaster*. Genetics 174:1349–1363.

Sanchez-Soriano N, Travis M, Dajas-Bailador F, Gonçalves-Pimentel C, Whitmarsh AJ, Prokop A. 2009. Mouse ACF7 and drosophila short stop modulate filopodia formation and microtubule organisation during neuronal growth. J. Cell Sci. 122:2534–2542.

Sandu C, Ryoo HD, Steller H. 2010. *Drosophila* IAP antagonists form multimeric complexes to promote cell death. J. Cell Biol. 190:1039–1052.

Schulz R, Sawruk E, Mülhardt C, Bertrand S, Baumann a, Phannavong B, Betz H, Bertrand D, Gundelfinger ED, Schmitt B. 1998. D alpha3, a new functional alpha subunit of nicotinic acetylcholine receptors from *Drosophila*. J. Neurochem. 71:853–862.

Schweers BA, Walters KJ, Stern M. 2002. The *Drosophila melanogaster* translational repressor pumilio regulates neuronal excitability. Genetics 161:1177–1185.

Sedaghat Y, Miranda WF, Sonnenfeld MJ. 2002. The jing Zn-finger transcription factor is a mediator of cellular differentiation in the *Drosophila* CNS midline and trachea. Development 129:2591–2606.

Sedaghat Y, Sonnenfeld M. 2002. The *rjing* gene is required for embryonic brain development in *Drosophila*. Dev. Genes Evol. 212:277–287.

Senthilan PR, Piepenbrock D, Ovezmyradov G, Nadrowski B, Bechstedt S, Pauls S, Winkler M, Möbius W, Howard J, Göpfert MC. 2012. *Drosophila* auditory organ genes and genetic hearing defects. Cell 150:1042–1054.

Sepp KJ, Hong P, Lizarraga SB, Liu JS, Mejia LA, Walsh CA, Perrimon N. 2008. Identification of neural outgrowth genes using genome-wide RNAi. PLoS Genet. 4.

Seppa MJ, Johnson RI, Bao S, Cagan RL. 2008. Polychaetoid controls patterning by modulating adhesion in the *Drosophila* pupal retina. Dev. Biol. 318:1–16.

Shen W, Ganetzky B. 2009. Autophagy promotes synapse development in *Drosophila*. J. Cell Biol. 187:71–79.

Short SM, Lazzaro BP. 2013. Reproductive status alters transcriptomic response to infection in female *Drosophila melanogaster*. G3 3:827–840.

Silverman N, Zhou R, Erlich RL, Hunter M, Bernstein E, Schneider D, Maniatis T. 2003. Immune activation of NF-kappaB and JNK requires *Drosophila* TAK1. J. Biol. Chem. 278:48928–48934.

Simões S, Denholm B, Azevedo D, Sotillos S, Martin P, Skaer H, Hombría JC-G, Jacinto A. 2006. Compartmentalisation of *Rho* regulators directs cell invagination during tissue morphogenesis. Development 133:4257–4267.

Song X, Xie T. 2003. Wingless signaling regulates the maintenance of ovarian somatic stem cells in *Drosophila*. Development 130:3259–3268.

Stern MD, Aihara H, Roccaro GA, Cheung L, Zhang H, Negeri D, Nibu Y. 2009. *CtBP* is required for proper development of peripheral nervous system in *Drosophila*. Mech. Dev. 126:68–79.

Stern S, Fridmann-Sirkis Y, Braun E, Soen Y. 2012. Epigenetically Heritable Alteration of Fly Development in Response to Toxic Challenge. Cell Rep. 1:528–542.

Stroschein-Stevenson SL, Foley E, O'Farrell PH, Johnson AD. 2006. Identification of *Drosophila* gene products required for phagocytosis of *Candida albicans*. PLoS Biol. 4:0087–0099.

Sun W, Margam VM, Sun L, Buczkowski G, Bennett GW, Schemerhorn B, Muir WM, Pittendrigh BR. 2006. Genome-wide analysis of phenobarbital-inducible genes in *Drosophila melanogaster*. Insect Mol. Biol. 15:455–464.

Sun X, Morozova T, Sonnenfeld M. 2006. Glial and neuronal functions of the *Drosophila* homolog of the human SWI/SNF gene ATR-X (DATR-X) and the jing zinc-finger gene specify the lateral positioning of longitudinal glia and axons. Genetics 173:1397–1415.

Takaesu NT, Hyman-walsh C, Ye Y, Wisotzkey RG, Stinchfield MJ, Connor MBO, Wotton D, Newfeld SJ, Ahead PA. 2006. *dSno* facilitates baboon signaling in the *Drosophila* brain by switching the affinity of *Medea* away from *Mad* and toward *dSmad2*. Genetics. 174(3): 1299-1313.

Takeo S, Swanson SK, Nandanan K, Nakai Y, Aigaki T. 2012. *Shaggy / glycogen synthase kinase 3β* and phosphorylation of *Sarah* / regulator of *calcineurin* are essential for completion of *Drosophila* female meiosis. Proc. Natl. Acad. Sci. U. S. A. 109(17):6382-6389.

Takeo S, Tsuda M, Akahori S, Matsuo T, Aigaki T. 2006. The Calcineurin Regulator Sra Plays an Essential Role in Female Meiosis in *Drosophila*. Curr. Biol. 16:1435–1440.

Tanaka KKK, Bryantsev AL, Cripps RM. 2008. Myocyte Enhancer Factor 2 and Chorion Factor 2 Collaborate in Activation of the Myogenic Program in *Drosophila*. Mol. Cell. Biol. 28:1616–1629.

Tang AH, Neufeld TP, Rubin GM, Muller HA. 2001. Transcriptional regulation of cytoskeletal functions and segmentation by a novel maternal pair-rule gene, *lilliputian*. Development 128:801–813.

Terashima J, Bownes M. 2005. A microarray analysis of genes involved in relating egg production to nutritional intake in *Drosophila melanogaster*. Cell Death Differ. 12:429–440.

Terman JR, Kolodkin AL. 2004. Nervy links protein kinase a to plexin-mediated semaphorin repulsion. Science 303:1204–1207.

Timmerman C, Suppiah S, Gurudatta B V., Yang J, Banerjee C, Sandstrom DJ, Corces VG, Sanyal S. 2013. The *Drosophila* transcription factor *Adf-1* (nalyot) regulates dendrite growth by controlling *FasII* and *Staufen* expression downstream of CaMKII and neural activity. J. Neurosci. 33:11916–11931.

Tsai LTY, Bainton RJ, Blau J, Heberlein U. 2004. *Lmo* mutants reveal a novel role for circadian pacemaker neurons in cocaine-induced behaviors. PLoS Biol. 2(12):e408.

Tunstall NE, Herr A, de Bruyne M, Warr CG. 2012. A screen for genes expressed in the olfactory organs of *Drosophila melanogaster* identifies genes involved in olfactory behaviour. PLoS One 7(4):e35641.

Ugrankar R, Berglund E, Akdemir F, Tran C, Kim MS, Noh J, Schneider R, Ebert B, Graff JM. 2015. *Drosophila* glucome screening identifies *Ck1alpha* as a regulator of mammalian glucose metabolism. Nat. Commun. 6:7102.

Valanne S, Kleino A, Myllymäki H, Vuoristo J, Rämet M. 2007. *Iap2* is required for a sustained response in the *Drosophila* Imd pathway. Dev. Comp. Immunol. 31:991–1001.

Valanne S, Myllymäki H, Kallio J, Schmid MR, Kleino A, Murumägi A, Airaksinen L, Kotipelto T, Kaustio M, Ulvila J, et al. 2010. Genome-wide RNA interference in *Drosophila* cells identifies G protein-coupled receptor kinase 2 as a conserved regulator of NF-kappaB signaling. J. Immunol. 184:6188–6198.

Venkei Z, Przewloka MR, Glover DM. 2011. *Drosophila Mis12* complex acts as a single functional unit essential for anaphase chromosome movement and a robust spindle assembly checkpoint. Genetics 187:131–140.

Vermehren-Schmaedick A, Ainsley JA, Johnson WA, Davies SA, Morton DB. 2010. Behavioral responses to hypoxia in *Drosophila* larvae are mediated by atypical soluble guanylyl cyclases. Genetics 186:183–196.

Vining MS, Bradley PL, Comeaux CA, Andrew DJ. 2005. Organ positioning in *Drosophila* requires complex tissue-tissue interactions. Dev. Biol. 287:19–34.

Vogt RG, Rogers ME, Franco M, Sun M. 2002. A comparative study of odorant binding protein genes: differential expression of the PBP1-GOBP2 gene cluster in *Manduca sexta* (Lepidoptera) and the organization of OBP genes in *Drosophila melanogaster* (Diptera). J. Exp. Biol. 205:719–744.

Wang P-Y, Neretti N, Whitaker R, Hosier S, Chang C, Lu D, Rogina B, Helfand SL. 2009. Long-lived *Indy* and calorie restriction interact to extend life span. Proc. Natl. Acad. Sci. U. S. A. 106:9262–9267.

Wang SL, Hawkins CJ, Yoo SJ, Müller HAJ, Hay BA. 1999. The *Drosophila* caspase inhibitor *DIAP1* is essential for cell survival and is negatively regulated by *HID*. Cell 98:453–463.

Wang X-H, Aliyari R, Li W-X, Li H-W, Kim K, Carthew R, Atkinson P, Ding S-W. 2006. RNA interference directs innate immunity against viruses in adult *Drosophila*. Science 312:452–454.

Wang Z, Pan Y, Li W, Jiang H, Chatzimanolis L, Chang J, Gong Z, Liu L. 2008. Visual pattern memory requires foraging function in the central complex of *Drosophila*. Learn. Mem. 15:133–142.

Watson FL, Püttmann-Holgado R, Thomas F, Lamar DL, Hughes M, Kondo M, Rebel VI, Schmucker D. 2005. Extensive diversity of Ig-superfamily proteins in the immune system of insects. Science 309:1874–1878.

Weber AL, Khan GF, Magwire MM, Tabor CL, Mackay TF, Anholt RR. 2012. Genome-wide association analysis of oxidative stress resistance in *Drosophila melanogaster*. PLoS One 7:e34745.

Weidmann C a, Raynard N a, Blewett NH, Van Etten J, Goldstrohm AC. 2014. The RNA binding domain of *Pumilio* antagonizes poly-adenosine binding protein and accelerates deadenylation. RNA 20:1298–1319.

Werz C, Lee T V, Lee PL, Lackey M, Bolduc C, Stein DS, Bergmann A. 2005. Mis-specified cells die by an active gene-directed process, and inhibition of this death results in cell fate transformation in *Drosophila*. Development 132:5343–5352.

Williams DW, Kondo S, Krzyzanowska A, Hiromi Y, Truman JW. 2006. Local caspase activity directs engulfment of dendrites during pruning. Nat. Neurosci. 9:1234–1236.

Wittwer F, van der Straten A, Keleman K, Dickson BJ, Hafen E. 2001. *Lilliputian*: an AF4/FMR2-related protein that controls cell identity and cell growth. Development 128:791–800.

Wolf FW, Eddison M, Lee S, Cho W, Heberlein U. 2007. *GSK-3/Shaggy* regulates olfactory habituation in *Drosophila*. Proc. Natl. Acad. Sci. U. S. A. 104:4653–4657.

Yan N, Wu J-W, Chai J, Li W, Shi Y. 2004. Molecular mechanisms of DrICE inhibition by DIAP1 and removal of inhibition by Reaper, Hid and Grim. Nat. Struct. Mol. Biol. 11:420–428.

Ye B, Petritsch C, Clark IE, Gavis ER, Jan LY, Jan YN. 2004. nanos and pumilio are essential for dendrite morphogenesis in *Drosophila* peripheral neurons. Curr. Biol. 14:314–321.

Yin VP, Thummel CS. 2004. A balance between the diap1 death inhibitor and reaper and hid death inducers controls steroid-triggered cell death in *Drosophila*. Proc. Natl. Acad. Sci. U. S. A. 101:8022–8027.

Yuan Q, Lin F, Zheng X, Sehgal A. 2005. Serotonin modulates circadian entrainment in *Drosophila*. Neuron 47:115–127.

Zenvirt S, Nevo-Caspi Y, Rencus-Lazar S, Segal D. 2008. Drosophila LIM-only is a positive regulator of transcription during thoracic bristle development. Genetics 179:1989–1999.

Zettervall C-J, Anderl I, Williams MJ, Palmer R, Kurucz E, Ando I, Hultmark D. 2004. A directed screen for genes involved in *Drosophila* blood cell activation. Proc. Natl. Acad. Sci. U. S. A. 101:14192–14197.

Zhang C, Hong Z, Ma W, Ma D, Qian Y, Xie W, Tie F, Fang M. 2013. *Drosophila* UTX coordinates with p53 to regulate ku80 expression in response to DNA damage.Müller R, editor. PLoS One 8:e78652.

Zhang Q, Zhang L, Gao X, Qi S, Chang Z, Wu Q. 2015. DIP1 plays an antiviral role against DCV infection in *Drosophila melanogaster*. Biochem. Biophys. Res. Commun. 460:222–226.

Zhang Q, Zhang L, Wang B, Ou CY, Chien CT, Jiang J. 2006. A Hedgehog-Induced BTB Protein Modulates Hedgehog Signaling by Degrading Ci/Gli Transcription Factor. Dev. Cell 10:719–729.

Zhao H, Zheng X, Yuan X, Wang L, Wang X, Zhong Y, Xie Z, Tully T. 2009. *ben* Functions with scamp during synaptic transmission and long-term memory formation in *Drosophila*. J. Neurosci. 29:414–424.

Zhou S, Morozova T V., Hussain YN, Luoma SE, McCoy L, Yamamoto A, MacKay TFC, Anholt RRH. 2016. The genetic basis for variation in sensitivity to lead toxicity in *Drosophila melanogaster*. Environ. Health Perspect. 124:1062–1070.

Zhu CT, Chang C, Reenan RA, Helfand SL. 2014. *Indy* gene variation in natural populations confers fitness advantage and life span extension through transposon insertion. Aging (Albany. NY). 6:58–69.

Zhuang S, Shao H, Guo F, Trimble R, Pearce E, Abmayr SM. 2009. *Sns* and *Kirre*, the *Drosophila* orthologs of *Nephrin* and *Neph1*, direct adhesion, fusion and formation of a slit diaphragm-like structure in insect nephrocytes. Development 136:2335–2344.

Zinke I, Schütz CS, Katzenberger JD, Bauer M, Pankratz MJ. 2002. Nutrient control of gene expression in *Drosophila*: Microarray analysis of starvation and sugar-dependent response. EMBO J. 21:6162–6173.

Exploring the Phenotypic Space and the Evolutionary History of a Natural Mutation in *Drosophila melanogaster*

Anna Ullastres,[1] Natalia Petit,[1] and Josefa González*[,1]
[1]Institute of Evolutionary Biology, CSIC-Universitat Pompeu Fabra, Barcelona, Spain
***Corresponding author:** E-mail: josefa.gonzalez@ibe.upf-csic.es.
Associate editor: John H. McDonald

Abstract

A major challenge of modern Biology is elucidating the functional consequences of natural mutations. Although we have a good understanding of the effects of laboratory-induced mutations on the molecular- and organismal-level phenotypes, the study of natural mutations has lagged behind. In this work, we explore the phenotypic space and the evolutionary history of a previously identified adaptive transposable element insertion. We first combined several tests that capture different signatures of selection to show that there is evidence of positive selection in the regions flanking *FBti0019386* insertion. We then explored several phenotypes related to known phenotypic effects of nearby genes, and having plausible connections to fitness variation in nature. We found that flies with *FBti0019386* insertion had a shorter developmental time and were more sensitive to stress, which are likely to be the adaptive effect and the cost of selection of this mutation, respectively. Interestingly, these phenotypic effects are not consistent with a role of *FBti0019386* in temperate adaptation as has been previously suggested. Indeed, a global analysis of the population frequency of *FBti0019386* showed that climatic variables explain well the *FBti0019386* frequency patterns only in Australia. Finally, although *FBti0019386* insertion could be inducing the formation of heterochromatin by recruiting HP1a (Heterochromatin Protein 1a) protein, the insertion is associated with upregulation of *sra* in adult females. Overall, our integrative approach allowed us to shed light on the evolutionary history, the relevant fitness effects, and the likely molecular mechanisms of an adaptive mutation and highlights the complexity of natural genetic variants.

Key words: **transposable elements, selective sweep, gene regulation, fitness, adaptation.**

Article

Introduction

Understanding the functional consequences of naturally occurring mutations remains a largely open question in Biology. Most of our knowledge on the effect of mutations comes from the analyses of laboratory-induced mutations. However, it is not clear whether laboratory mutations are representative of mutations that arise and persist in natural populations (Kolaczkowski et al. 2011; Rose et al. 2011). First, most laboratory mutations studied are loss-of-function mutations that are most likely rare in natural populations and/or their effects are masked by the presence of buffering mechanisms (Landry and Rifkin 2012). Additionally, laboratory-induced mutations tend to be highly pleiotropic and it is difficult to infer which of the phenotypes might be targets of selection in nature (Kolaczkowski et al. 2011).

The recent explosion in the number of studies aimed at identifying natural adaptive mutations in several organisms allows us to study the effect of natural genetic variants at an unprecedented scale (González et al. 2008; Turner et al. 2010; Jones et al. 2012; Huang et al. 2014; Tobler et al. 2014). These studies are revealing that mapping genotype to phenotype is even more complex than previously thought due to the prevalence of gene-by-environment interactions, gene-by-gene interactions, and pleiotropy (Rockman 2012; Lehner 2013; Mackay 2014). First, being able to map a putatively adaptive mutation to its relevant phenotypic effect depends partly on finding the particular environmental conditions in which the mutation is adaptive (Paaby and Schmidt 2008; Storz and Wheat 2010). Thus, taking into account environmental information of the populations where putative adaptive mutations are identified should help mapping the mutation to its relevant phenotype. Second, epistatic interactions also affect the phenotypic outcome of mutations. The phenotypic effect of mutations could be enhanced or suppressed depending on the background being analyzed (Huang et al. 2012). Additionally, several backgrounds should be analyzed to discard the effect of other mutations and reliably attribute the identified phenotypic effect to the candidate mutation (Burnett et al. 2011). Third, many genes are linked to several traits (Paaby and Rockman 2013). In some cases, mutations can have antagonistic effects, that is, beneficial effects in a trait/environment and deleterious effects on a different trait/environment. Pleiotropic mutations can also have beneficial effects on two different traits (McGee et al. 2014). Tradeoffs are prevalent when selection acts on a single trait, whereas payoffs arise when multiple traits are selected for simultaneously (McGee et al. 2014). Thus, if we want to fully characterize the effects of a given natural mutation, several phenotypes need to be studied (Mackay 2010; Guio et al. 2014).

Finally, a comprehensive understanding of adaptation goes beyond identifying fitness consequences of adaptive mutations. Pinpointing the molecular mechanisms underlying adaptation is needed to provide conclusive support for the

Table 1. Summary of the Analyses Showing Evidence of Positive Selection in the 1-kb Region around *FBti0019386* Insertion.

	Observed		Neutral Simulations				Resampling of Strains	
			Mean (CI 95%)		*P* value		Mean (CI 95%)	*P* Value
	P	*A*	*P*	*A*	*P*	*A*		
π	0.43	4.51	3.92 (1.32, 7.81)	4.20 (1.33, 8.04)	0.001	> 0.05	3.35 (2.78, 3.87)	<0.001
Tajima's *D*	−1.77	0.68	−0.11 (−1.46, 1.62)	−0.04 (−1.41, 1.64)	0.007	> 0.05	0.4 (−0.19, 1.02)	<0.001
CL (log)	−5.95	−18.15	−18.69 (−29.67, −8.80)	−15.20 (−25.89, −6.82)	0.006	> 0.05	−12.18 (−15.23, −8.81)	<0.001

NOTE.—Neutral simulations were performed with MS program using the parameter theta = 4. For simulations with theta = 5, please see supplementary table S2, Supplementary Material online. P, data set of strains with *FBti0019386* insertion; A, data set of strains without *FBti0019386* insertion.

adaptive role of the mutation (Storz and Wheat 2010). Additionally, elucidating the evolutionary history of adaptive variation for fitness traits allows to start answering long-standing questions on the genetic basis of adaptation (Orr 2005).

In this work, we focused on characterizing the functional effects, the molecular mechanism, and the evolutionary history of a natural transposable element (TE)-induced mutation in *Drosophila melanogaster*: *FBti0019386* belonging to the *invader4* retrotransposon family (González et al. 2008, 2010; St Pierre et al. 2014). *FBti0019386* has been identified as a candidate adaptive TE insertion based on its population dynamics (González et al. 2008). González et al. (2010) further reported that *FBti0019386* shows parallel clinal frequency patterns in North America and Australia suggesting that it is involved in adaptation to temperate environments. *FBti0019386* is inserted in the 5′-untranslated region (UTR) intron of *sarah* (*sra*) and 2.5 kb upstream of *Bicoid-interacting protein 1* (*Bin1*) in the 3 R chromosomal arm (St Pierre et al. 2014). *sra* laboratory mutants affect several biological processes, such as egg activation, female meiosis, and long-term memory among others (Ejima et al. 2001, 2004; Chang et al. 2003; Horner et al. 2006; Takeo et al. 2006; Sakai and Aigaki 2010; Nakai et al. 2011). In most cases, these phenotypes are the result of the deregulation of *calcineurin*, which is inhibited by *sra* (Takeo et al. 2006; Sakai and Aigaki 2010; Nakai et al. 2011). Laboratory-induced mutations affecting *Bin1*, a highly conserved transcriptional corepressor, play a role during environmental stress response in *Arabidopsis* (Song and Galbraith 2006) and in *Drosophila* (Costa et al. 2011). Thus, to identify the phenotypic consequences of *FBti0019386* mutation, we explored several candidate phenotypes previously associated with *sra* and *Bin1* mutants in different developmental stages, in different environmental conditions, and in flies with different genetic backgrounds.

Our results showed that *FBti0019386* increased in frequency in out-of-Africa populations due to positive selection and is associated with shorter developmental time (DT) and increased sensitivity to cold-stress. These two phenotypic effects together with the lack of correlation between *FBti0019386* frequency and latitude in European populations raised doubts about the role of *FBti0019386* in temperate adaptation. Finally, although *FBti0019386* insertion could be inducing pi-RNA mediated heterochromatin assembly, the insertion is associated with upregulation of *sra* in adult females.

Results

FBti0019386 Flanking Regions Show Signatures of Positive Selection

We tested whether the region flanking *FBti0019386* showed signatures of positive selection (see Materials and Methods for a description of the different tests used). We found an extreme decrease of nucleotide diversity (π) in strains with *FBti0019386* insertion compared with strains without the insertion, which was accompanied by a decrease in Tajima's *D* statistic (table 1, supplementary fig. S1*A* and *B* and table S1, Supplementary Material online) (Hudson et al. 1992; Tajima 1989). The Composite Likelihood (CL) test, specifically designed to detect selective sweeps (Nielsen et al. 2005), was higher in flies with *FBti0019386* insertion compared with flies without the insertion, as expected if flies with the insertion show signatures of a selective sweep in the analyzed region (table 1). We confirmed that values of π, Tajima's *D*, and CL were statistically different from neutral simulated scenarios in flies with *FBti0019386* insertion but not in flies without the insertion (table 1 and supplementary table S2, Supplementary Material online).

To test whether the observed differences were due to the *FBti0019386* insertion, we estimated the three statistics in random samples of the strains (see Materials and Methods). None of the randomized data sets had lower π, lower Tajima's *D*, or higher CL value compared with the data set of strains with *FBti0019386* insertion (table 1 and supplementary table S3, Supplementary Material online). Finally, we performed the Composite Likelihood Ratio (CLR; Nielsen et al. 2005) test comparing strains with and without the *FBti0019386* insertion, and we found that it was significant: CLR = 24.40 *P* value = 7.82×10^{-7}. Moreover, this CLR value is three times bigger than any of the CLR values calculated in a random sample of 1,000 1-kb-long regions from 3 R chromosome, where *FBti0019386* is located (supplementary table S4, Supplementary Material online). Note that estimates of π and Tajima's *D* in these 1,000 regions also showed that these two statistics did not significantly differ between strains with and without *FBti0019386* insertion (supplementary fig. S1C and D, Supplementary Material online).

Note that we checked whether polymorphisms other than TE were present in the flanking regions analyzed. No other polymorphisms were found that could potentially confound the results of our tests of selection suggesting that the TE is the causative mutation.

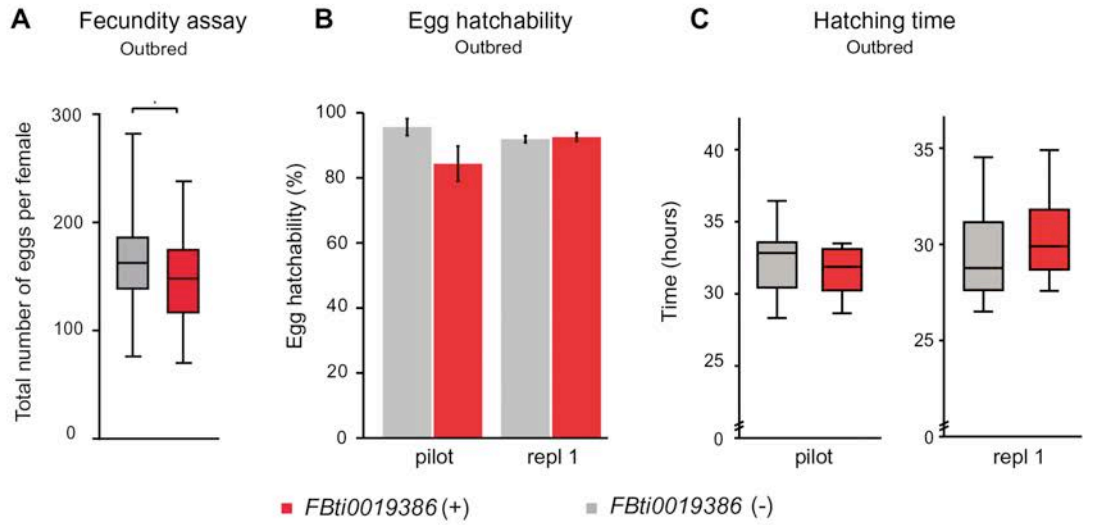


FIG. 1. *FBti0019386* does not affect fecundity (*A*), egg hatchability (*B*), or hatching time (*C*) in outbred populations. (*A*) Average number of eggs laid by outbred females without *FBti0019386* insertion (*FBti0019386* (−)) and with *FBti0019386* insertion (*FBti0019386* (+)). (*B*) Percentage of hatched embryos. (*C*) Average hatching time. In all cases, error bars represent standard error of the mean (SEM).

Table 2. Odds Ratios (OR) and Confidence Intervals (CI) for Phenotypic Experiments Performed with Embryos with and without *FBti0019386*.

Experiment	Strain	OR (CI)
Fecundity	Outbred	1.05 (0.67–1.64)
Hatching time in cold	Outbred pilot	7.07 (3.37–14.83)
	Outbred replica 1	2.21 (1.49–3.26)
DT	Outbred pilot	5.69 (2.72–11.94)
	Outbred replica 1	2.62 (1.88–3.66)
	Outbred replica 2	2.60 (1.94–5.88)
	Individual DGRP	1.95 (1.30–2.92)

Overall, we found evidence of positive selection in the region flanking *FBti0019386* insertion suggesting that *FBti0019386* is an adaptive insertion.

Exploring the Fitness Space of *FBti0019386*

To explore the phenotypic space of *FBti0019386* insertion, we investigated several traits related to the phenotypic effects of nearby genes: Fecundity and egg hatchability associated with *sra* mutant alleles. Related to egg hatchability, we also investigated egg hatching time, egg-to-adult viability, and DT. Additionally, we investigated cold stress, osmotic stress, and starvation stress as *Bin1* mutants have been shown to play a role in stress resistance.

Because *FBti0019386* is located 242.4 kb away from the distal breakpoint of *In(3R)Payne* inversion and inversions are known to be under selection, we checked whether this inversion was present in any of the six strains used to perform the different phenotypic analyses (see Materials and Methods). We found that none of the strains used in our analyses carries *In(3R)Payne* inversion.

We also checked whether polymorphisms other than the *FBti0019386* insertion were present in the genomic region including *sra* and *Bin1* genes. We did not find any polymorphism linked to the *FBti0019386* that could potentially confound the results of the phenotypic assays performed.

FBti0019386 Insertion Does Not Affect Fecundity or Egg Hatching

Laboratory mutant flies in which *sra* is underexpressed lay less eggs than wild-type flies and most of the eggs do not hatch (Horner et al. 2006). To check whether *FBti0019386* insertion has an effect on fecundity, we compared the number of eggs laid per female in outbred populations with and without the insertion (see Materials and Methods). Our results showed that, on average, flies without the insertion laid slightly more eggs than flies with the insertion (*t*-test, $P = 0.047$) (fig. 1*A*). However, the size effect of the mutation was not significant (table 2). We also tested whether differences in fecundity were present early in life, as has been reported by Paaby et al. (2014). Although the mean number of eggs laid by flies with the insertion in the first 48 h of egg laying was bigger than the number laid by flies without the insertion (3.95 vs. 2.33 eggs), these differences were not statistically significant (*t*-test, $P = 0.06$).

We then checked whether outbred flies with and without *FBti0019386* differed in egg hatchability and/or hatching time. We first performed a pilot experiment using 150 embryos per strain and we found that flies with the insertion did not show significant differences compared with flies without the insertion in egg hatchability (*t*-test, $P > 0.05$) (fig. 1*B*) or hatching time (*t*-test, $P > 0.05$) (fig. 1*C*). Although differences were not significant, flies with the insertion showed a lower number of hatched eggs (fig. 1*B*) and a shorter hatching time (fig. 1*C*). We thus repeated the experiments using 500 embryos per strain and we found that flies with and without *FBti0019386* did not differ in egg hatchability (*t*-test,

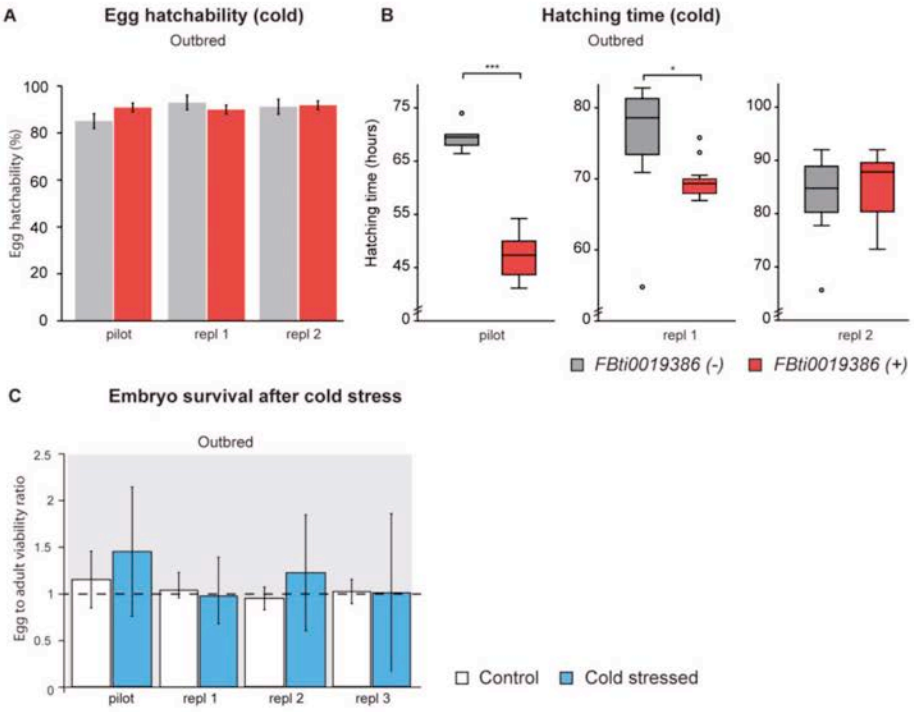


FIG. 2. *FBti0019386* does not affect embryo hatching or survival in cold stress conditions in outbred populations. (*A*) Percentage of embryos that hatched during cold-stress periods (see Materials and Methods). (*B*) Average egg hatching time. (*C*) Egg-to-adult survival after a single cold stress period during embryonic stage (cold stressed) and under control conditions (control). Bars represent the survival ratio between flies with *FBti0019386* and flies without *FBti0019386* and error bars represent SEM.

$P > 0.05$) (fig. 1*B*) or hatching time (*t*-test, $P > 0.05$) (fig. 1*C*).

Overall, we did not find significant differences in fecundity, egg hatchability, or egg hatching time in flies with and without *FBti0019386* insertion. These results suggest that *FBti0019386* does not have a significant effect on these phenotypes.

FBti0019386 Insertion Does Not Affect Egg Hatching or Egg-To-Adult Viability under Cold Stress Conditions

As mentioned above, *Bin1* plays a role in general environmental stress response in *Drosophila* (Costa et al. 2011). We thus screened several phenotypes in embryos under cold stress conditions: Egg hatching, egg hatching time, and egg-to-adult viability.

We performed egg hatchability and egg-hatching time assays in outbred populations under repeated cold stress exposure (see Materials and Methods). We did not detect differences in egg hatchability between flies with and without the insertion in any of the three replicas performed (*t*-test, $P > 0.05$) (fig. 2*A*). However, flies with *FBti0019386* insertion from the pilot experiment and the first replica hatched significantly before flies without the element (*t*-test, $P << 0.001$ and $P = 0.011$, respectively) (table 2) whereas no differences were observed in the second replica (*t*-test, $P > 0.05$) (fig. 2*B*).

We further tested whether flies with and without *FBti0019386* differed in the egg-to-adult viability after exposing outbred flies to a single cold-stress period during early embryo stages. Our results showed that there are no differences in survival between flies with and without the insertion in control conditions or under cold-stress (two-way ANOVA [analysis of variance], $P > 0.05$, fig. 2*C*).

Overall, and although variability in hatching time was observed in some of the experiments performed, our results suggest that *FBti0019386* insertion does not affect cold-tolerance during the embryo stage.

FBti0019386 Is Associated with Increased Sensitivity to Cold Stress in Adults

Because we could not find any significant difference between strains with and without *FBti0019386* in embryonic stage, we decided to test whether differences between the two strains were present in adult flies. We first tested whether adult flies with and without *FBti0019386* insertion differed in chill-coma recovery time (CCRT) and survival after cold stress. CCRT is used as a reliable measure of cold tolerance in *Drosophila* (Gibert et al. 2001; Macdonald et al. 2004). We observed that flies with the insertion showed significantly longer recovery time compared with flies without the insertion suggesting that they were more sensitive to cold stress (Mann–Whitney test, $P << 0.001$) (fig. 3*A* and table 3). We

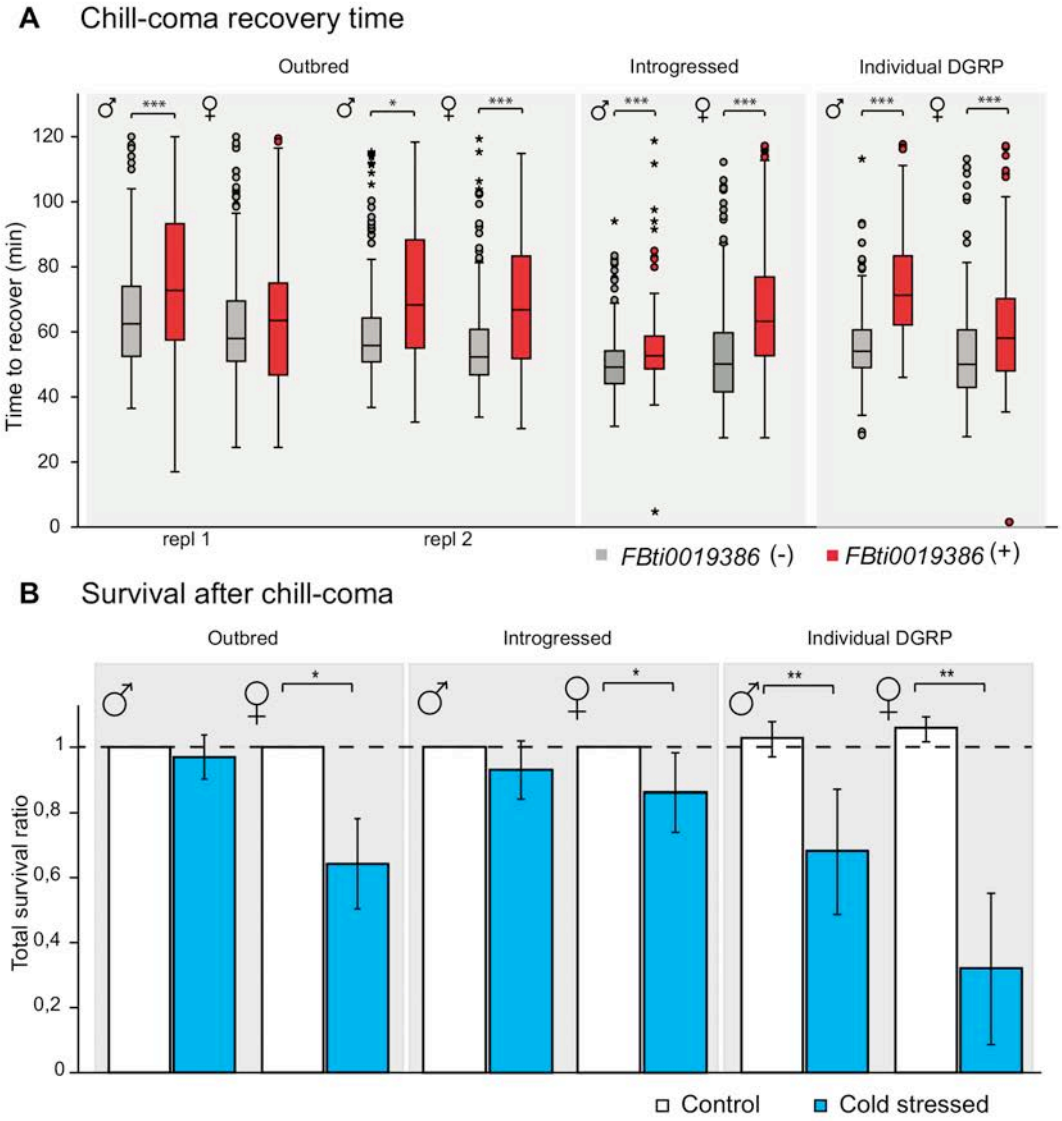


FIG. 3. Flies with *FBti0019386* insertion are more sensitive to cold stress. (*A*) Average time to recover after chill coma in adult flies from outbred populations, introgressed strains, and inbred DGRP strains (RAL-857 and RAL-802). (*B*) Survival ratio between flies with *FBti0019386* insertion and flies without the insertion after chill coma exposure (cold stress) and in control conditions (control) in the three genetic backgrounds. Error bars represent SEM.

replicated this result in flies with the same genetic background (Mann–Whitney test, $P < 0.05$) and in flies with two other genetic backgrounds: The introgressed strains generated in our laboratory (Mann–Whitney test, $P << 0.001$) and a couple of inbred strains from the DGRP (Drosophila Genetic Reference Panel) project (Mann–Whitney test, $P << 0.001$) (fig. 3*A* and table 3) (see Materials and Methods).

In accordance with this increased cold sensitivity, flies with the insertion also showed an increased mortality following chill-coma exposure, although these differences were not always significant (fig. 3*B* and table 3).

Finally, we also tested whether flies with *FBti0019386* insertion were more sensitive to osmotic stress and starvation stress. We found that outbred females with the insertion were more sensitive to high salt concentrations (Kaplan–Meyer, log rank $P < 0.001$) (supplementary fig. S2*A*, Supplementary Material online, and table 3), and outbred males with the insertion were more sensitive to starvation stress (Kaplan–Meyer, log rank $P < 0.001$)

(supplementary fig. S2*B*, Supplementary Material online, and table 3).

Overall, longer CCRT and lower cold-stress survival in flies with *FBti0019386* insertion across backgrounds suggested that this mutation is negatively affecting adult cold-stress response. This high sensitivity to cold stress likely represents the cost of selection of this TE mutation. Furthermore, preliminary results are suggestive but not conclusive of a negative role of *FBti0019386* in general response to stress.

FBti0019386 Insertion Is Associated with Shorter DT

During the course of the experiments, we noticed that flies with *FBti0019386* showed a shorter DT than flies without the insertion. Because DT is relevant to fitness in all organisms, and especially for those such as *D. melanogaster* that occupy ephemeral habitats (Chippindale et al. 1997), we tested this observation. We found that outbred flies (Mann–Whitney test, pilot experiment $P = 0.006$ and replica 1 and 2 $P < 0.001$) and inbred DGRP flies (*t*-test, $P = 0.02$) with the insertion developed faster compared with flies without the TE insertion (fig. 4 and table 2). On average, flies with *FBti0019386* insertion developed 9.4–17.9 h before compared with flies without the insertion. However, we could not detect significant DT differences in the introgressed strains differing by the presence/absence of *FBti0019386* (*t*-test, $P > 0.05$) (fig. 4), suggesting that polymorphisms other than the TE influence DT in this background. Note that the effect size of the mutation on the other phenotypes studied also varies depending on the background being analyzed (tables 2 and 3). This suggests that polymorphisms other than *FBti0019386* play a role not only in DT but also in other phenotypes.

Table 3. Odds Ratios (OR) and Confidence Intervals (CI) for Phenotypic Experiments Performed with Male and Female Adult Flies with and without *FBti0019386*.

Experiment	Strain	Males OR (CI)	Females OR (CI)
CCRT	Outbred replica 1	3.44 (2.31–5.18)	N/A[a]
	Outbred replica 2	3.79 (2.54–5.67)	5.18 (3.43–7.82)
	Introgressed	2.44 (1.64–3.62)	4.16 (2.69–6.41)
	Individual DGRP	11.63 (6.79–19.93)	2.26 (1.54–3.33)
Survival after chill-coma	Outbred	N/A	7.80 (3.27–18.60)
	Introgressed	N/A	1.89 (0.99–3.62)
	Individual DGRP	9.94 (5.49–18)	6.88 (3.43–13.82)
Osmotic stress	Outbred	N/A	1.61 (1.21–2.13)
Starvation stress	Outbred	1.52 (1.15–2.01)	N/A

[a]N/A (OR was estimated when differences between flies with and without *FBti0019386* were statistically significant).

FBti0019386 Frequency Showed Clinal Patterns in North America and Australia but No Correlation between Frequency and Latitude Is Found in Europe

Shorter DT and increased sensitivity to cold stress are not consistent with a role of *FBti0019386* in temperate adaptation (González et al. 2010). However, previous evidence for a role in temperate adaptation was based on the analysis of only two North American and five Australian populations (González et al. 2010). To further test these results, we estimated *FBti0019386* frequencies in additional populations from North America, Australia, Europe, and Africa (supplementary table S5, Supplementary Material online) using *T-lex2* pipeline (Fiston-Lavier et al. 2014). We found that *FBti0019386* insertion is present at 10% frequency in a Rwanda population confirming its low frequency in Africa (supplementary table S5, Supplementary Material online). We confirmed that the TE is present at intermediate to high frequencies in 15 additional out-of-Africa populations (fig. 5 and supplementary table S5, Supplementary Material online). We also confirmed that the TE frequency varies clinally with latitude in North America and Australia (Pearson correlation $P = 0.011$ and $P = 0.002$, respectively; supplementary table S6, Supplementary Material online). However, when we analyzed the *FBti0019386* frequency in six European populations we did not find any significant correlation between frequency and latitude (Pearson correlation $P = 0.313$; supplementary table S6, Supplementary Material online).

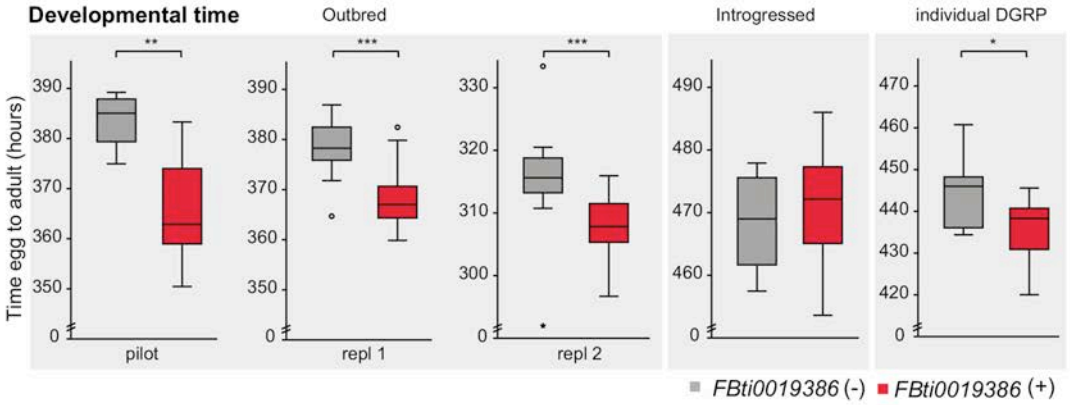


FIG. 4. *FBti0019386* is associated with shorter DT. Average egg-to-adult DT in populations without *FBti0019386* insertion and with the insertion. Error bars represent SEM.

Besides latitude, we also tested whether other geographical and climatic variables showed significant correlations with *FBti0019386* frequency. We found significant correlations between frequency and temperature-related variables in North America and between frequency and both temperature-related and precipitation-related variables in Australia (supplementary table S6, Supplementary Material online). No significant correlation was found in Europe (supplementary table S6, Supplementary Material online). Because most of the climatic variables are significantly correlated among them and with latitude (supplementary table S7, Supplementary Material online), we performed a Principal Component Analysis (PCA) to disentangle the relationships between the variables. In North America, climate variables were grouped in two components, in Australia in three and in Europe in two (supplementary table S8, Supplementary Material online). As expected based on the correlation analyses, only in North America and in Australia, some of the climatic variables grouped with latitude and frequency (supplementary fig. S3*A*, Supplementary Material online). In North America, the first component accounted for 46% of climatic variation (supplementary table S9, Supplementary Material online) and explained 54% of the variation in *FBti0019386* frequency (supplementary fig. S3*B*, Supplementary Material online). In Australia, the first component accounted for 68% of climatic variation (supplementary table S9, Supplementary Material online) and explained 86% of the frequency variation (supplementary fig. S3*B*, Supplementary Material online). Finally in Europe, the first principal component explained 54% of the climatic variation (supplementary table S9, Supplementary Material online) but was not significantly correlated with *FBti0019386* frequency (supplementary fig. S3*B*, Supplementary Material online).

Overall, although we were able to confirm the clinal pattern of *FBti0019386* in North America and Australia, our results did not provide evidence for the presence of a clinal pattern in Europe. In Australia, the clinal pattern is well explained by the observed climatic variation, whereas in North America climatic variation did not fully explain the observed correlation between *FBti0019386* frequency and latitude, suggesting that other factors might be involved in the observed clinal pattern. As expected, none of the climatic variables significantly correlated with TE frequency in Europe.

FBti0019386 Is Associated with Upregulation of sra in Female Flies

To shed light on the molecular mechanism of *FBti0019386* insertion, we measured the expression of *sra* and *Bin1* in nonstress conditions in embryos and in nonstress and cold-stress conditions in female flies with and without *FBti0019386* insertion.

We did not observe significant differences in *sra* or *Bin1* expression in embryos differing by the presence/absence of *FBti0019386* insertion (*t*-test, $P > 0.05$) (fig. 6*A* and *B*). However, we observed that adult female flies with *FBti0019386* insertion showed an increase of *sra* expression compared with flies without the insertion both in control conditions and after cold-stress conditions, although results were only significant under control conditions (*t*-test, $P = 0.03$) (fig. 6*C*). On the other hand, no significant differences in expression level between flies with and without *FBti0019386* were observed for *Bin1* (*t*-test, $P > 0.05$) (fig. 6*D*).

Interestingly, we observed a change in *sra* and *Bin1* expression after cold stress in flies with and without *FBti0019386* insertion: *sra* is upregulated in cold stress conditions (*t*-test, $P < 0.05$ in both cases) (fig. 6*C*) whereas *Bin1* is downregulated (*t*-test, $P < 0.05$ in both cases) (fig. 6*D*).

Overall, we did not observe any change in expression of *sra* and *Bin1* in embryos, in agreement with the lack of phenotypic consequences of *FBti0019386* in this developmental stage. However, we observed a upregulation of *sra* in flies with *FBti0019386* insertion that was significant under nonstress conditions. Moreover, we showed that both *sra* and *Bin1* changed their expression in response to cold stress.

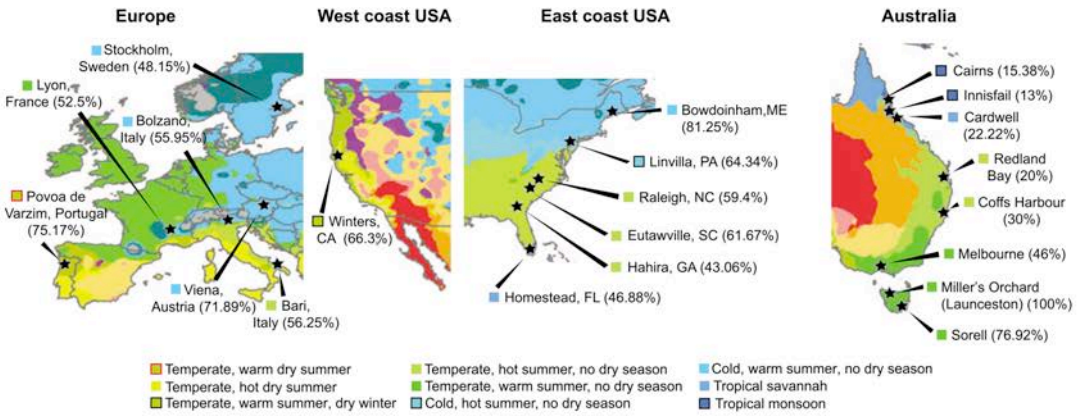


FIG. 5. Climate map with *Drosophila melanogaster* population samples analyzed with T-lex2. The frequency of *FBti0019386* in each population is shown in brackets. Climate maps are modified from Peel et al. (2007).

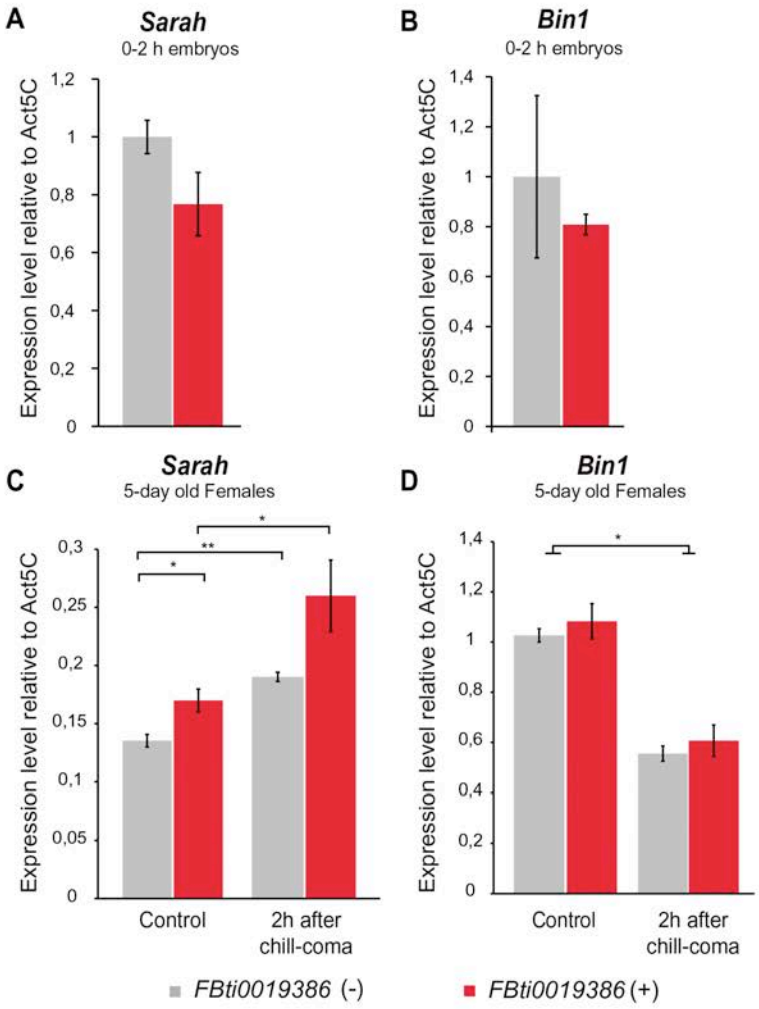


Fig. 6. Flies with *FBit0019386* insertion showed *sra* upregulation. Real-time polymerase chain reaction quantification of *sra* and *Bin1* transcript levels in outbred flies without *FBti0019386* insertion and with *FBti0019836* insertion. We represented the average expression level of *sra* (*A* and *C*) and *Bin1* (*B* and *D*) relative to *Act5C* with SEM error bars for three biological replicates in 0–2 h embryos and in 5-day-old females. Normalized expression measured 2 h after chill-coma for *sra* and *Bin1* is depicted in (*C*) and (*D*), respectively.

FBti0019386 Could Be Affecting gene Expression by Ectopically Assembling Heterochromatin

TEs from the *invader4* family contain sites with homology to PIWI interacting RNAs (piRNAs) that act as *cis*-acting targets for heterochromatin assembly by recruiting Heterochromatin Protein 1a (HP1a) (Sentmanat and Elgin 2012). Specifically, these piRNA binding sites are located in the long terminal repeat (LTR) sequences. Because *FBti0019386* is a 347-bp solo-LTR, we hypothesized that it could be inducing the ectopic assembly of heterochromatin. We analyzed the 14.6-kb region containing *Bin1*, *sra*, and *FBti0019386* and found that both sense and antisense piRNAs bind specifically to *FBti0019386* (fig. 7*A*) (see Materials and Methods). Second, we tested whether there is evidence for the presence of HP1a binding to *FBti0019386* sequence. We found that HP1a specifically binds to *FBti0019386* sequence (fig. 7*B*) (see Materials and Methods). Thus, these results suggest that *FBti0019386* could be affecting gene expression by inducing the ectopic assembly of heterochromatin.

Discussion

In this work, we explored the plausible phenotypic space of the putatively adaptive *FBti0019386* insertion in different developmental stages, embryo and adult, and in different environmental conditions, nonstress conditions and cold, osmotic, and starvation stress conditions. Overall, we found

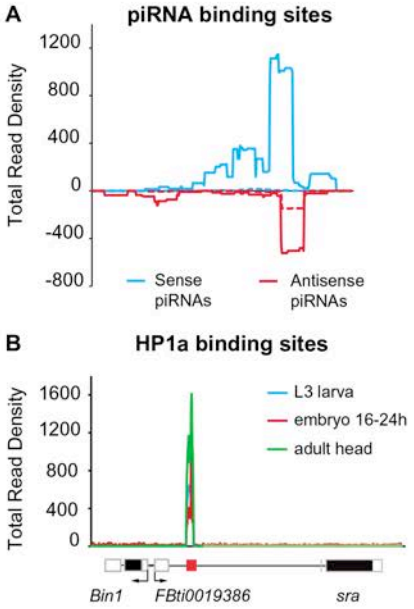


FIG. 7. *FBti0019386* could bind piRNA and HP1a protein. (*A*) Mapping of piRNA sense and antisense RNA-seq reads against *FBti0019386* sequence. Data from Li et al. (2009) are depicted in dashed lines and data from Satyaki et al. (2014) are represented in continuous lines. (*B*) Mapping of reads coming from HP1a ChIP-Seq experimental data against the genome region containing *Bin1*, *FBti0019386*, and *sra*. Experimental data from L3 larva, 16–24 h embryo, and adult heads are given.

that *FBti0019386* mediates sensitivity to cold stress conditions (fig. 3) and is associated with faster DT (fig. 4). These two phenotypic effects have plausible fitness consequences in nature that could explain why the mutation increased in frequency in natural populations but has not reached fixation. Increased sensitivity to cold stress conditions is likely to reduce fitness of the flies that carry *FBti0019386* insertion, and may represent the cost of selection of this mutation. On the other hand, faster DT is likely to increase the fitness of flies with *FBti0019386* insertion. In nature, quick development favors *D. melanogaster* individuals for several reasons. First, larvae feed on rotting fruits that are ephemeral. Thus, quick development allows larvae to pupate before the food source is exhausted. Second, competition increases as more and more eggs are laid on a piece of fruit, also favoring individuals with faster DT (Nunney 1990). Third, breeding sites in nature can be destroyed by physical factors and predation, individuals that develop faster are thus more likely to escape microhabitat destruction. And fourth, faster DT accelerates the age of first breeding, which is relevant for the organism if most reproduction happens in expanding populations. This is the case for *D. melanogaster* populations that expand their population size every spring. Thus, it is plausible that *FBti0019386* increased in frequency in natural populations because of its positive effect on DT whereas it did not reach fixation because of its negative effect on cold-stress resistance. Our results emphasize the importance of exploring different phenotypes to fully characterize the effects of natural mutations, as have been suggested before (Mackay 2010; Guio et al. 2014). Although our results provide a plausible explanation for the effect of *FBti0019386* insertion in natural populations, experiments under natural conditions are needed to unequivocally identify the effect of this insertion in nature.

By combining several tests that capture different signatures of selection at the DNA level, we demonstrate that *FBti0019386* shows signatures of positive selection suggesting that it is an adaptive mutation (table 1). However, our results are not entirely consistent with a role of *FBti0019386* in temperate adaptation as has been previously proposed (González et al. 2010). First, adaptation to temperate climates has been associated with increased stress resistance, increased DT, and decreased fecundity (Stanley and Parsons 1981; Hoffmann et al. 2003; Schmidt et al. 2005; Folguera et al. 2008; Schmidt and Paaby 2008; but see also James and Partridge 1995; James et al. 1997; Trotta et al. 2006). However, we found that *FBti0019386* is associated with increased sensitivity to cold stress (fig. 3), with shorter DT (fig. 4) and does not significantly affect fecundity (fig. 1). Thus, the phenotypic effects of *FBti0019386* are not consistent with a role of this insertion in temperate adaptation. Second, our global analyses of *FBti0019386* population frequency showed that *FBti0019386* frequency correlates with latitude and with climatic variables in North America and in Australia but not in Europe (fig. 5 and supplementary table S6, Supplementary Material online). We suggest that the clinal frequency patterns in North America and in Australia could be due to the dual colonization of these two continents by European and African populations rather than to the operation of spatially varying selection (Caracristi and Schlotterer 2003; Rouault et al. 2004; Duchen et al. 2013; Bergland et al. 2014). The lack of clinal frequency patterns in Europe would support this conclusion. However, it is also possible that phenotypic effects of *FBti0019386* not yet characterized could be consistent with a role of this natural mutation in temperate adaptation. Additionally, although there is evidence for the presence of clinal variation in European populations (David et al. 1985, 1986, 1989; Costa et al. 1992), other works have shown that clines are weaker in Europe compared with other continents (Oakeshott, Chambers, et al. 1983; Oakeshott, Gibson, et al. 1983). This could be partly due to differences in the latitudinal ranges spanned by populations analyzed in the different continents. In this work, the latitudinal range spanned by North American (25.82°–45.06°) and Australian (−16.88° to −42.83°) populations is larger than the range spanned by European populations (41.13°–59.33°). In any case, genome-wide scan studies that identify loci that are differentiated between populations should be taken as a first step toward the identification of loci that are subject to spatially varying selection (González et al. 2010; Kolaczkowski et al. 2011; Fabian et al. 2012; Reinhardt et al. 2014). Further functional validation should be gathered before concluding that the candidate loci are under spatially varying selection (Bergland et al. 2014).

Our results also shed light on the molecular processes that lead from genotype to phenotype. We found that *FBti0019386* is associated with upregulation of *sra* (fig. 6*C*). As previously described for other elements from the *invader4* family, we showed that *FBti0019386* has piRNA binding sites (fig. 7*A*) (Sentmanat and Elgin 2012). We also showed that HP1a binds specifically to the *FBti0019386* sequence, further suggesting that *FBti0019386* could be inducing the ectopic assembly of heterochromatin (fig. 7*B*). These results highlight the potential role of TE remnants as silencing signals to be used by piRNAs to direct heterochromatin formation (Sentmanat et al. 2013). Although we observed an upregulation of *sra* in adult females, we can not discard that heterochromatin assembly induced by *FBti0019386* could be affecting gene expression in other developmental stages and/or specific tissues.

A recent update of FlyBase, the database of *Drosophila* genes and genomes, annotated two new *Bin1* transcripts that have their transcription start site inside *FBti0019386* (St Pierre et al. 2014). As a consequence, these two new transcripts would only be produced in strains with the insertion, and could contribute to differences in the level of *Bin1* expression in flies with and without the insertion. Although we did not find differences in *Bin1* expression, we cannot discard that differences in the level of expression of *Bin1* are present in developmental stages, tissues, or environmental conditions that we have not investigated.

Although *sra* and *Bin1* have not been associated with DT, both genes play important roles during development and have been associated with a wide range of processes (Chang et al. 2003; Ejima et al. 2004; Horner et al. 2006; Takeo et al. 2006, 2010; Chang and Min 2009; Matyash et al. 2009; Costa et al. 2011; Nakai et al. 2011). A genome-wide screening looking for genes influencing DT in *D. melanogaster* has shown that the many candidate genes were involved in a wide range of biological processes such as cellular metabolic processes, organismal development, and response to stress (Mensch et al. 2008). More recently, developmental timing in insects has been associated with hormonal and circadian control (Di Cara and King-Jones 2013; Yadav et al. 2014). Interestingly, *sra* is regulated by *Shaggy/GSK-3β* (*sgg*), a Ser–Thr kinase involved in the regulation of circadian rhythmicity (Martinek et al. 2001). On the other hand, both *Bin1* and *sra* are stress-response genes: *Bin1* is upregulated in response to stress and *sra* is downregulated (fig. 6). *Bin1* is a known key player in transcriptional response to environmental stress (Costa et al. 2011). Although there was no previous evidence for a direct role of *sra* in response to stress, *sra* could be affecting stress response through its role in the calcium pathway (Takeuchi et al. 2009; Teets et al. 2013; Davies et al. 2014). *sra* inhibits *calcineurin*, a highly conserved protein in eukaryotes that has the ability to sense calcium (Hogan et al. 2003). Although it is not deeply understood, calcium pathways play a role during general cell-stress response including cold stress response (Takeuchi et al. 2009; Teets et al. 2013; Davies et al. 2014). Note that many genes that affect complex traits in *Drosophila* had well-characterized roles in early development and were not previously annotated to affect adult quantitative traits (Mackay 2010).

FBti0019386 adds to the growing list of TE-induced adaptive mutations that have been linked to their fitness effects and their underlying molecular mechanisms (Schmidt et al. 2010; Magwire et al. 2011; Guio et al. 2014; Mateo et al. 2014; Sun et al. 2014). Overall, these examples highlight the variety of mechanisms underlying adaptive mutations and point toward a significant role of TEs in response to stress (Casacuberta and González 2013). However, the number of characterized mutations is still too small to obtain an overall picture of adaptation. In depth, characterization of a representative set of adaptive mutations in natural populations will allow us to start answering long-standing questions in the field such as which traits are more relevant for adaptation? What is the effect-size distribution of adaptive mutations? and What evolutionary processes underlie adaptive evolution?

Materials and Methods

Sequence Analysis of the *FBti0019386* Flanking Regions

Single nucleotide polymorphism (SNP) data were downloaded from the DGRP2 webpage (https://www.hgsc.bcm.edu/arthropods/drosophila-genetic-reference-panel) in vcf format. Strains with ($N = 65$) and without ($N = 38$) *FBti0019386* insertion were filtered using vcftools v_0.1.10 (http://vcftools.sourceforge.net/).

We used three different statistics to detect positive selection: Nucleotide diversity (π), Tajima's *D*, and the CL of SNPs. Positive selection results in the elimination of standing genetic variation that is linked to the adaptive mutation. Thus, if *FBti0019386* has increased in frequency due to positive selection, we expect a decrease in π in flies with the insertion compared with flies without the insertion. π is calculated as the mean number of pairwise differences between two given sequences (Hudson et al. 1992). Tajima's *D* statistic is calculated as the ratio between the mean number of pairwise differences and the number of segregating sites (Tajima 1989). This ratio is expected to be 0 in a neutrally evolving population whereas negative values of Tajima's *D* can be taken as evidence of positive selection (Tajima 1989). Finally, CL test is calculated by multiplying the marginal likelihoods for each site along the studied sequences (Nielsen et al. 2005).

π, Tajima's *D*, and CL were calculated for the two sets of sequences, with and without the insertion, using the PopGenome package in R (Pfeifer et al. 2014). Sliding windows analyses were performed for 200-bp-size windows spanning 1 and 2-kb regions flanking the insertion. Differences between strains with and without the insertion were more drastic for the 1-kb region flanking the insertion; therefore, we focused our analysis in this region.

Simulations were performed using the MS program (Hudson 2002). Theta values were estimated using the 205 DGRP2 strains for the 2-kb region around *FBti0019386* (theta = 4.77/kb) and for the 3R chromosomal arm (theta = 4.5/kb). Thus, simulations were performed for theta

values of 4/kb and 5/kb, which are frequently used as neutral values in *D. melanogaster*.

Ad hoc perl scripts were used for the resampling analyses. In total, 1,000 random samples of the 103 DGRP strains analyzed were obtained keeping the same proportion as in the original present and absent data sets (60%/40%, respectively) and a sample size of nearly 50% of the total data set.

We also computed CLR as 2*(log CL (present) − log CL (absent)), for a 1-kb region around the TE insertion. Because demography could produce similar patterns as positive selection, we performed a random sampling of 1,000 1-kb-long regions from the 3 R chromosome for the absent and present data sets and calculated π, Tajima's *D*, CL, and CLR tests in each one of them.

Fly Strains

Outbred Strains

We selected six inbred strains from the Drosophila Genetic Reference Panel (Mackay et al. 2012; Huang et al. 2014) homozygous for the presence of *FBti0019386* insertion (RAL-21, RAL-40, RAL-177, RAL-402, RAL-405, and RAL-857). We placed ten virgin females and ten males of each strain in a fly chamber to create an outbred population sharing the TE insertion. We also selected six inbred strains without the insertion (RAL-75, RAL-138, RAL-383, RAL-461, RAL-822, and RAL-908) and created an outbred strain following the same procedure explained above. Each outbred population was maintained by random mating (N ≈ 800 flies per generation) for at least ten generations before starting the experiments.

Introgressed Strains

We selected two DGRP strains: One homozygous for the presence of *FBti0019386* insertion (RAL-177) and one homozygous for the absence (RAL-802). We crossed RAL-177 virgin females with RAL-802 males and backcrossed the virgin females that carry *FBti0019386* insertion from the following generations with RAL-802 males for 12 generations. After that, we did brother–sister crosses until we obtained homozygous strains for the absence and homozygous strains for the presence of *FBti0019386*.

Individual DGRP Strains

We used a couple of individual DGRP strains differing by the presence/absence of *FBti0019386* insertion to perform our phenotypic assays. We used RAL-857 (homozygous for the presence of *FBti0019386* insertion) and RAL-802 (homozygous for the absence).

Presence/Absence of *In(3 R)Payne* in the Analyzed Strains

To discard the effect of *In(3 R)Payne* inversion on *FBti0019386* phenotypic effects, we genotyped the strains analyzed to detect the presence/absence of this inversion: The two outbred, the two introgressed, and the two individual DGRP strains. We used the primer sequences described in Matzkin et al. (2005). As a positive control, we used a strain that was previously genotyped in our laboratory and that carries the *In(3 R)Payne* inversion.

Phenotypic Assays

All experiments were performed using outbred populations. Additionally, we used introgressed and individual DGRP strains to perform CCRT assay, survival after chill-coma, and DT assays.

Fecundity

In total, 40 virgin females from each strain were placed individually in vials with one male from the same strain. During 17 days flies were moved to new vials every 2 days and the number of eggs laid per female during that period was counted. Total fecundity, that is, average of the total number of eggs laid per female during the 17 days, and early fecundity, that is, average of the total number of eggs laid per female during the first 48 h of egg laying, was compared between flies with and without *FBti0019386*.

Egg Hatchability and Hatching Time

In total, 800 4 - to 8-day-old flies were allowed to lay eggs for 3 h on apple juice-agar medium with fresh yeast. Embryos were separated in groups of 20 or 50 and placed into food vials. Vials were kept at room temperature (19–22°C) and checked during the following hours for hatched eggs (2–5 times per day). We analyzed the average time over the midpoint of each successive interval in order to estimate the hatching time. Two experiments were performed following this protocol: A first pilot experiment with 150 embryos per strain, and one replica with 500 embryos per strain.

Egg hatchability and egg hatching time were also analyzed under cold stress conditions. Embryos were placed at 1°C overnight for 14 h and at 18°C during the day, and this cycle was maintained until all the eggs had hatched. We performed a pilot experiment with 100 embryos per strain and additional experiments with 240 and 160 embryos per strain, respectively.

Cold Stress in Embryos

In total, 800 7 - to 10-day-old flies were allowed to lay eggs for 3 h on apple juice-agar medium with fresh yeast. Embryos were collected following the methodology described in Schou (2013), and placed into food vials in groups of 50. When embryos were 3–6 h old, vials were placed at 1°C for 14 h, and maintained at 18°C until adult emergence. Simultaneously, control vials were always maintained at 18°C and not cold-exposed to control for other variables affecting egg to adult survival. We performed a first pilot experiment using 280 embryos per strain and three biological replicas using 350 embryos per strain (replica 1) and 750 embryos per strain (replica 2 and replica 3, respectively). In all cases, we analyzed egg to adult survival after all the adults had emerged.

Chill-Coma Recovery Time

In total, 500 3 - to 5-day-old flies were separated by sex and by strain and placed into five empty vials in groups of 50. We allowed flies to recover from CO_2 anesthesia for 1 h and then vials were put in ice and kept in a 4°C chamber for 16 h as described in David et al. (1998). After the cold shock, adults were transferred to Petri dishes at room temperature (22–24°C), and recovery time was monitored for successive

Detection of piRNA Reads Binding to *FBti0019386* Sequence

We used small RNA sequencing data to check whether piRNAs reads mapped to *FBti0019386* sequence, following a methodology similar to that described in Sentmanat and Elgin (2012). Briefly, we obtained the small RNA reads from Oregon R ovaries (accession number SRP000458) (Li et al. 2009), and from wild type ovaries (accession number: SRX470700) (Satyaki et al. 2014). We aligned the reads by using BWA-MEM package version 0.7.5 a-r405 (Li 2013) to the 14.6-kb sequence obtained from *Drosophila* reference genome, containing *Bin1* and *sra* genes, and *FBti0019386* (release five chromosomal coordinates 3 R: 12,010,721–12,025,306). Then, we used samtools and bamtools (Barnett et al. 2011) to index and filter by sense/antisense reads. Finally, we obtained the total read density using R (Rstudio v0.98.507).

Detection of HP1a Protein Binding in *FBti0019386* Sequence

We downloaded all available raw data from modEncode HP1a protein ChIP-Seq experiments: Embryos (ID 3391 and 3392), third instar larvae (ID 4936), and adult heads (ID 5592) (http://data.modencode.org). Then, we mapped the reads against the 14.6-kb region described above. We performed the alignments following the same methodology as for the piRNA reads analysis.

Supplementary Material

Supplementary tables S1–S9 and figures S1–S3 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

The authors thank Maite G. Barrón, Lain Guio, and Miriam Merenciano for comments on the manuscript. A.U. is an FPI fellow (BES-2012-052999) and J.G. is a *Ramón y Cajal* fellow (RYC-2010-07306). This work was supported by grants from the European Comission (Marie Curie CIG PCIG-2011-293860) and from the Spanish Government (Fundamental Research Projects Grant BFU-2011-24397) to J.G.

References

Barnett DW, Garrison EK, Quinlan AR, Stromberg MP, Marth GT. 2011. BamTools: a C++ API and toolkit for analyzing and managing BAM files. *Bioinformatics* 27:1691–1692.

Bergland AO, Tobler R, González J, Schmidt PS, Petrov DA. 2014. Secondary contact and local adaptation contribute to genome-wide patterns of clinal variation in *Drosophila melanogaster*. *bioRxiv* doi: 10.1101/009084.

Burnett C, Valentini S, Cabreiro F, Goss M, Somogyvari M, Piper MD, Hoddinott M, Sutphin GL, Leko V, McElwee JJ, et al. 2011. Absence of effects of Sir2 overexpression on lifespan in *C. elegans* and *Drosophila*. *Nature* 477:482–485.

Caracristi G, Schlotterer C. 2003. Genetic differentiation between American and European *Drosophila melanogaster* populations could be attributed to admixture of African alleles. *Mol Biol Evol.* 20:792–799.

Casacuberta E, González J. 2013. The impact of transposable elements in environmental adaptation. *Mol Ecol.* 22:1503–1517.

Chang KT, Min KT. 2009. Upregulation of three *Drosophila* homologs of human chromosome 21 genes alters synaptic function: implications for Down syndrome. *Proc Natl Acad Sci U S A.* 106:17117–17122.

Chang KT, Shi YJ, Min KT. 2003. The *Drosophila* homolog of Down's syndrome critical region 1 gene regulates learning: implications for mental retardation. *Proc Natl Acad Sci U S A.* 100: 15794–15799.

Chippindale AK, Alipaz JA, Chen HW, Rose RM. 1997. Experimental evolution of accelerated development in *Drosophila*. 1. Developmental speed and larval survival. *Evolution* 51:1536–1551.

Costa E, Beltran S, Espinas ML. 2011. *Drosophila melanogaster* SAP18 protein is required for environmental stress responses. *FEBS Lett.* 585: 275–280.

Costa R, Peixoto AA, Barbujani G, Kyriacou CP. 1992. A latitudinal cline in a *Drosophila* clock gene. *Proc Biol Sci.* 250:43–49.

David J, Capy P, Payant V, Tsakas S. 1985. Thoracic trident pigmentation in *Drosophila melanogaster*: differentiation of geographical populations. *Genet Sel Evol.* 17:211–224.

David J, Mercot H, Capy P, McEvey S, Van Herrewege J. 1986. Alcohol tolerance and Adh gene frequencies in European and African populations of *Drosophila melanogaster*. *Genet Sel Evol.* 18:405–416.

David JR, Alonso-Moraga A, Borai F, Capy P, Mercot H, McEvey SF, Munoz-Serrano A, Tsakas S. 1989. Latitudinal variation of Adh gene frequencies in *Drosophila melanogaster*: a Mediterranean instability. *Heredity (Edinb)* 62(Pt 1):11–16.

David RJ, Gibert P, Pla E, Petavy G, Karan D, Moreteau B. 1998. Cold stress tolerance in Drosophila: analysis of chill coma recovery in *D. melanogaster*. *J Therm Biol.* 23(5):291–299.

Davies SA, Cabrero P, Overend G, Aitchison L, Sebastian S, Terhzaz S, Dow JA. 2014. Cell signalling mechanisms for insect stress tolerance. *J Exp Biol.* 217:119–128.

Di Cara F, King-Jones K. 2013. How clocks and hormones act in concert to control the timing of insect development. *Curr Top Dev Biol.* 105: 1–36.

Duchen P, Zivkovic D, Hutter S, Stephan W, Laurent S. 2013. Demographic inference reveals African and European admixture in the North American *Drosophila melanogaster* population. *Genetics* 193:291–301.

Ejima A, Nakayama S, Aigaki T. 2001. Phenotypic association of spontaneous ovulation and sexual receptivity in virgin females of *Drosophila melanogaster* mutants. *Behav Genet.* 31:437–444.

Ejima A, Tsuda M, Takeo S, Ishii K, Matsuo T, Aigaki T. 2004. Expression level of sarah, a homolog of DSCR1, is critical for ovulation and female courtship behavior in *Drosophila melanogaster*. *Genetics* 168:2077–2087.

Fabian DK, Kapun M, Nolte V, Kofler R, Schmidt PS, Schlotterer C, Flatt T. 2012. Genome-wide patterns of latitudinal differentiation among populations of *Drosophila melanogaster* from North America. *Mol Ecol.* 21:4748–4769.

Fiston-Lavier AS, Barrón M, Petrov DA, González J. 2014. T-lex2: genotyping, frequency estimation and re-annotation of transposable elements using single or pooled next-generation sequencing data. *Nucleic Acids Res.* 43(4):e22.

Folguera G, Ceballos S, Spezzi L, Fanara JJ, Hasson E. 2008. Clinal variation in developmental time and viability, and the response to thermal treatments in two species of *Drosophila*. *Biol J Linn Soc Lond.* 95: 233–245.

Gibert P, Moreteau B, Petavy G, Karan D, David JR. 2001. Chill-coma tolerance, a major climatic adaptation among *Drosophila* species. *Evolution* 55:1063–1068.

González J, Karasov TL, Messer PW, Petrov DA. 2010. Genome-wide patterns of adaptation to temperate environments associated with transposable elements in *Drosophila*. *PLoS Genet.* 6:e1000905.

González J, Lenkov K, Lipatov M, Macpherson JM, Petrov DA. 2008. High rate of recent transposable element-induced adaptation in *Drosophila melanogaster*. *PLoS Biol.* 6:e251.

intervals of 30 s during 2 h. We considered as recovered flies those that were able to stand on their legs. As a control, we monitored survival of flies that were kept at room temperature: Three vials of 20 flies each, by sex and strain.

Survival after Chill-Coma

In total, 400 5 - to 8-day-old flies were separated by sex and strain and placed into six food vials in groups of 20. We allowed flies to recover from CO_2 anesthesia for at least 2 days. After that, flies were changed to empty food vials and were put in ice, and kept in a 4°C chamber for 16 h. When adults were recovered from chill-coma, we transferred them to food vials and we monitored mortality during the next 5 days. As a control, we monitored survival of flies that were kept at room temperature: Three vials of 20 flies each, by sex and strain.

Osmotic Stress

In total, 2,000 4 - to 7-day-old flies were separated by sex and strain and placed in groups of 20 into 20 food vials containing 3% of NaCl, and into five vials with normal food as a control. Flies were maintained at room temperature (22–24°C) and dead flies were counted every 12–24 h until all the treated flies were dead.

Starvation Stress

In total, 2,000 3 - to 4-day-old flies per strain were separated by sex and strain and placed in groups of 20 into 20 food vials containing only 1.5% agar, and into five vials with normal food as a control. Flies were maintained at room temperature (22–24°C) and dead flies were counted three times a day until all the treated flies were dead.

Developmental Time

In total, 800 7 - to 10-day-old flies were allowed to lay eggs for 3 h. A total of 500 embryos per strain were collected and distributed in groups of 50 per food vial and were maintained at 18°C. Vials were checked every 6–8 h for emerging adults until all flies had emerged. We estimated the average DT over the midpoint of each successive interval.

Statistical Analyses of the Phenotypic Assays

Analyses were performed with SPSS v21. We first tested whether data followed a normal distribution by performing Kolmogorov–Smirnov test. *t*-Test was performed for normal data and Mann–Whitney test for nonnormal data. Survival curves were compared with log-rank test. When the statistical test was significant, we estimated the size effect of the mutation by calculating the odds-ratio and its confidence interval.

FBti0019386 Frequency Estimation for Natural Populations

To obtain *FBti0019386* frequency, we run *T-lex2* (Fiston-Lavier et al. 2014) using *Drosophila* whole-genome sequences available from a total of 23 populations from North America, Australia, Europe, and Africa (supplementary table S5, Supplementary Material online).

The accuracy of TE frequency estimates using *T-lex2* is affected by coverage. However, coverage for all samples was higher than 20× except for Lyon (France) and California (USA), which had 8× and 4.7× coverage respectively, suggesting that overall frequency estimates are accurate.

Correlation Analysis of *FBti0019386* Frequency with Geographic and Climate Variables

We analyzed whether the frequency of *FBti0019386* insertion correlated with different geographical and climate variables in North America, Australia, and Europe using Pearson product–moment correlations. We also performed a PCA to disentangle the relationships between the climatic variables using Statistica (v8.0, StatSoft, Inc. 2007). Climatic data were obtained from the weather stations adjacent to collection sites of each population, available in Peel et al. (2007). When necessary, data were transformed as described in Sokal and Rohlf (2012) (see pages 411–422).

mRNA Transcript Levels Analysis (quantitative reverse transcription polymerase chain reaction)

Total RNA was extracted from three biological samples of 40 adult females (5-day old) from outbred populations differing by the presence/absence of *FBti0019386* insertion using Trizol reagent and PureLink RNA Mini kit (Ambion). RNA was treated on-column with DNase I (Trizol) and after RNA purification. Reverse transcription was carried out using 1 µg of total RNA, Anchored-oligo(dT) primer, and Transcription First Strand cDNA Synthesis Kit (Roche). The resulting cDNA was used for quantitative reverse transcription polymerase chain reaction with SYBR Green (BioRad) on an iQ5 Thermal cycler. *sra* total expression was measured using a pair of primers specific to a 124-bp cDNA amplicon spanning the 5′-UTR/exon junction of the gene (5′-ACAACAACGGTGG AGAAGAGCCGT-3′ and 5′-GGTGCATCGGCGGACGCA TTG-3′). For *Bin1*, we measured the 66-bp cDNA amplicon spanning the 5′-UTR/exon junction using specific primers (5′-TGTCGTCCCGTAGAGCAGAA-3′ and 5′-CA AGCAGATTGACCGCGAGA-3′). In both cases, we normalized the expression with *Act5C* (5′-GCGCCCTTA CTCTTTCACCA-3′ and 5′-ATGTCACGGACGATTTCA CG-3′). Expression was measured in nonstress conditions and in cold-stress conditions: 16 h at 4°C and 2 h at room temperature to allow flies to recover.

We also analyzed the expression of both genes in 0–2 h embryos using the same procedure. We collected the embryos from population cages containing approximately 800 flies from outbred populations differing by the presence/absence of *FBti0019386* insertion. Briefly, 4 - to 8-day-old flies were allowed to lay eggs for 2 h on apple juice-agar medium with fresh yeast. Then, embryos were collected using a small brush and cleaned with water. Embryos were dechorionized by submerging them for 5 min in 50% bleach. After that, embryos were placed in a microcentrifuge tube, the excess of water was eliminated, and the samples were froze at −80°C until RNA extraction.

Guio L, Barrón MG, González J. 2014. The transposable element Bari-Jheh mediates oxidative stress response in *Drosophila*. *Mol Ecol.* 23: 2020–2030.
Hill T, Lewicki P. 2007. STATISTICS: Methods and Applications. Tulsa (OK): StatSoft.
Hoffmann AA, Sorensen JG, Loeschcke V. 2003. Adaptation of *Drosophila* to temperature extremes: bringing together quantitative and molecular approaches. *J Therm Biol.* 28:175–216.
Hogan PG, Chen L, Nardone J, Rao A. 2003. Transcriptional regulation by calcium, calcineurin, and NFAT. *Genes Dev.* 17:2205–2232.
Horner VL, Czank A, Jang JK, Singh N, Williams BC, et al. 2006. The *Drosophila* calcipressin sarah is required for several aspects of egg activation. *Curr Biol.* 16:1441–1446.
Huang W, Massouras A, Inoue Y, Peiffer J, Ramia M, et al. 2014. Natural variation in genome architecture among 205 *Drosophila melanogaster* Genetic Reference Panel lines. *Genome Res.* 24:1193–1208.
Huang W, Richards S, Carbone MA, Zhu D, Anholt RR, et al. 2012. Epistasis dominates the genetic architecture of *Drosophila* quantitative traits. *Proc Natl Acad Sci U S A.* 109:15553–15559.
Hudson RR. 2002. Generating samples under a Wright-Fisher neutral model of genetic variation. *Bioinformatics* 18:337–338.
Hudson RR, Slatkin M, Maddison WP. 1992. Estimation of levels of gene flow from DNA sequence data. *Genetics* 132:583–589.
James AC, Azevedo RB, Partridge L. 1997. Genetic and environmental responses to temperature of *Drosophila melanogaster* from a latitudinal cline. *Genetics* 146:881–890.
James AC, Partridge L. 1995. Thermal evolution of rate of larval development in *Drosophila melanogaster* in laboratory and field populations. *J Evol Biol.* 8:315–330.
Jones FC, Grabherr MG, Chan YF, Russell P, Mauceli E, et al. 2012. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484:55–61.
Kolaczkowski B, Kern AD, Holloway AK, Begun DJ. 2011. Genomic differentiation between temperate and tropical Australian populations of *Drosophila melanogaster*. *Genetics* 187:245–260.
Landry CR, Rifkin SA. 2012. The genotype-phenotype maps of systems biology and quantitative genetics: distinct and complementary. *Adv Exp Med Biol.* 751:371–398.
Lehner B. 2013. Genotype to phenotype: lessons from model organisms for human genetics. *Nat Rev Genet.* 14:168–178.
Li C, Vagin VV, Lee S, Xu J, Ma S, et al. 2009. Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. *Cell* 137:509–521.
Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. bioRxiv doi: 1303.3997v2..
Macdonald SS, Rako L, Batterham P, Hoffmann AA. 2004. Dissecting chill coma recovery as a measure of cold resistance: evidence for a biphasic response in *Drosophila melanogaster*. *J Insect Physiol.* 50: 695–700.
Mackay TF. 2010. Mutations and quantitative genetic variation: lessons from *Drosophila*. *Philos Trans R Soc Lond B Biol Sci.* 365:1229–1239.
Mackay TF. 2014. Epistasis and quantitative traits: using model organisms to study gene-gene interactions. *Nat Rev Genet.* 15:22–33.
Mackay TF, Richards S, Stone EA, Barbadilla A, Ayroles JF, et al. 2012. The *Drosophila melanogaster* Genetic Reference Panel. *Nature* 482: 173–178.
Magwire MM, Bayer F, Webster CL, Cao C, Jiggins FM. 2011. Successive increases in the resistance of *Drosophila* to viral infection through a transposon insertion followed by a Duplication. *PLoS Genet.* 7: e1002337.
Martinek S, Inonog S, Manoukian AS, Young MW. 2001. A role for the segment polarity gene shaggy/GSK-3 in the *Drosophila* circadian clock. *Cell* 105:769–779.
Mateo L, Ullastres A, González J. 2014. A transposable element insertion confers xenobiotic resistance in *Drosophila*. *PLoS Genet.* 10:e1004560.
Matyash A, Singh N, Hanes SD, Urlaub H, Jackle H. 2009. SAP18 promotes Kruppel-dependent transcriptional repression by enhancer-specific histone deacetylation. *J Biol Chem.* 284:3012–3020.
Matzkin LM, Merritt TJ, Zhu CT, Eanes WF. 2005. The structure and population genetics of the breakpoints associated with the cosmopolitan chromosomal inversion In(3R)Payne in *Drosophila melanogaster*. *Genetics* 170:1143–1152.
McGee LW, Aitchison EW, Caudle SB, Morrison AJ, Zheng L, Yang W, Rokyta DR. 2014. Payoffs, not tradeoffs, in the adaptation of a virus to ostensibly conflicting selective pressures. *PLoS Genet.* 10:e1004611.
Mensch J, Lavagnino N, Carreira VP, Massaldi A, Hasson E, Fanara JJ. 2008. Identifying candidate genes affecting developmental time in *Drosophila melanogaster*: pervasive pleiotropy and gene-by-environment interaction. *BMC Dev Biol.* 8:78.
Nakai Y, Horiuchi J, Tsuda M, Takeo S, Akahori S, Matsuo T, Kume K, Aigaki T. 2011. Calcineurin and its regulator sra/DSCR1 are essential for sleep in *Drosophila*. *J Neurosci.* 31:12759–12766.
Nielsen R, Williamson S, Kim Y, Hubisz MJ, Clark AG, Bustamante C. 2005. Genomic scans for selective sweeps using SNP data. *Genome Res.* 15:1566–1575.
Nunney L. 1990. *Drosophila* on oranges: colonization, competition and coexistance. *Ecology* 71:1904–1915.
Oakeshott JG, Chambers GK, Gibson JB, Eanes WF, Willcocks DA. 1983. Geographic variation in G6pd and Pgd allele frequencies in *Drosophila melanogaster*. *Heredity (Edinb)* 50(Pt 1):67–72.
Oakeshott JG, Gibson JB, Willcocks DA, Chambers GK. 1983. Latitudinal variation in octanol dehydrogenase and acid phosphatase allele frequencies in *Drosophila melanogaster*. *Theor Appl Genet.* 65:191–196.
Orr HA. 2005. The genetic theory of adaptation: a brief history. *Nat Rev Genet.* 6:119–127.
Paaby AB, Bergland AO, Behrman EL, Schmidt PS. 2014. A highly pleiotropic amino acid polymorphism in the *Drosophila* insulin receptor contributes to life-history adaptation. *Evolution* 68:3395–3409.
Paaby AB, Rockman MV. 2013. The many faces of pleiotropy. *Trends Genet.* 29:66–73.
Paaby AB, Schmidt PS. 2008. Functional significance of allelic variation at methuselah, an aging gene in *Drosophila*. *PLoS One* 3:e1987.
Peel MC, Finlayson BL, McMahon TA. 2007. Updated world map of the Köppen-Geiger climate classification. *Hydrol Earth Syst Sci.* 11: 1633–1644.
Pfeifer B, Wittelsburger U, Ramos-Onsins SE, Lercher MJ. 2014. PopGenome: an efficient Swiss army knife for population genomic analyses in R. *Mol Biol Evol.* 31:1929–1936.
Reinhardt JA, Kolaczkowski B, Jones CD, Begun DJ, Kern AD. 2014. Parallel geographic variation in *Drosophila melanogaster*. *Genetics* 197:361–373.
Rockman MV. 2012. The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution* 66:1–17.
Rose MR, Mueller LD, Burke MK. 2011. New experiments for an undivided genetics. *Genetics* 188:1–10.
Rouault JD, Marican C, Wicker-Thomas C, Jallon JM. 2004. Relations between cuticular hydrocarbon (HC) polymorphism, resistance against desiccation and breeding temperature; a model for HC evolution in *D. melanogaster* and D. simulans. *Genetica* 120:195–212.
Sakai T, Aigaki T. 2010. The *Drosophila* calcineurin regulator, Sarah, is involved in male courtship. *Neuroreport* 21:985–988.
Satyaki PR, Cuykendall TN, Wei KH, Brideau NJ, Kwak H, Aruna S, Ferree PM, Ji S, Barbash DA. 2014. The Hmr and Lhr hybrid incompatibility genes suppress a broad range of heterochromatic repeats. *PLoS Genet.* 10:e1004240.
Schmidt JM, Good RT, Appleton B, Sherrard J, Raymant GC, et al. 2010. Copy number variation and transposable elements feature in recent, ongoing adaptation at the Cyp6g1 locus. *PLoS Genet.* 6:e1000998.
Schmidt PS, Paaby AB. 2008. Reproductive diapause and life-history clines in North American populations of *Drosophila melanogaster*. *Evolution* 62:1204–1215.
Schmidt PS, Paaby AB, Heschel MS. 2005. Genetic variance for diapause expression and associated life histories in *Drosophila melanogaster*. *Evolution* 59:2616–2625.
Schou MF. 2013. Fast egg collection method greatly improves randomness of egg sampling in *Drosophila melanogaster*. *Fly (Austin)* 7: 44–46.

Sentmanat M, Wang SH, Elgin SC. 2013. Targeting heterochromatin formation to transposable elements in *Drosophila*: potential roles of the piRNA system. *Biochemistry (Mosc)* 78:562–571.

Sentmanat MF, Elgin SC. 2012. Ectopic assembly of heterochromatin in *Drosophila melanogaster* triggered by transposable elements. *Proc Natl Acad Sci U S A.* 109:14104–14109.

Sokal RR, Rohlf FJ. 2012. Biometry: the principles and practice of statistics in biological research, 4th ed. New York: W. H. Freeman and Co.

Song CP, Galbraith DW. 2006. AtSAP18, an orthologue of human SAP18, is involved in the regulation of salt stress and mediates transcriptional repression in *Arabidopsis*. *Plant Mol Biol.* 60:241–257.

St Pierre SE, Ponting L, Stefancsik R, McQuilton P, FlyBase C. 2014. FlyBase 102—advanced approaches to interrogating FlyBase. *Nucleic Acids Res.* 42:D780–D788.

Stanley SM, Parsons PA. 1981. The response of the cosmopolitan species *Drosophila melanogaster* to ecological gradients. *Proc Ecol Soc Aust.* 11:121–130.

Storz JF, Wheat CW. 2010. Integrating evolutionary and functional approaches to infer adaptation at specific loci. *Evolution* 64: 2489–2509.

Sun W, Shen YH, Han MJ, Cao YF, Zhang Z. 2014. An adaptive transposable element insertion in the regulatory region of the EO gene in the domesticated silkworm, *Bombyx mori*. *Mol Biol Evol.* 31(12):3302–3313.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595.

Takeo S, Hawley RS, Aigaki T. 2010. Calcineurin and its regulation by Sra/RCAN is required for completion of meiosis in *Drosophila*. *Dev Biol.* 344:957–967.

Takeo S, Tsuda M, Akahori S, Matsuo T, Aigaki T. 2006. The calcineurin regulator sra plays an essential role in female meiosis in *Drosophila*. *Curr Biol.* 16:1435–1440.

Takeuchi K, Nakano Y, Kato U, Kaneda M, Aizu M, et al. 2009. Changes in temperature preferences and energy homeostasis in dystroglycan mutants. *Science* 323:1740–1743.

Teets NM, Yi SX, Lee RE Jr, Denlinger DL. 2013. Calcium signaling mediates cold sensing in insect tissues. *Proc Natl Acad Sci U S A.* 110: 9154–9159.

Tobler R, Franssen SU, Kofler R, Orozco-Terwengel P, Nolte V, Hermisson J, Schlotterer C. 2014. Massive habitat-specific genomic response in *D. melanogaster* populations during experimental evolution in hot and cold environments. *Mol Biol Evol.* 31: 364–375.

Trotta V, Calboli FC, Ziosi M, Guerra D, Pezzoli MC, David JR, Cavicchi S. 2006. Thermal plasticity in *Drosophila melanogaster*: a comparison of geographic populations. *BMC Evol Biol.* 6:67.

Turner TL, Bourne EC, Von Wettberg EJ, Hu TT, Nuzhdin SV. 2010. Population resequencing reveals local adaptation of *Arabidopsis lyrata* to serpentine soils. *Nat Genet.* 42:260–263.

Yadav P, Thandapani M, Sharma VK. 2014. Interaction of light regimes and circadian clocks modulate timing of pre-adult developmental events in *Drosophila*. *BMC Dev Biol.* 14:19.