Transposable Elements and the Evolution of Insects

Clément Gilbert, Jean Peccoud, Richard Cordaux

Transposable elements and the evolution of insects

Clément Gilbert[1*], Jean Peccoud[2], Richard Cordaux[2]

[1] Laboratoire Evolution, Génomes, Comportement, Écologie, Unité Mixte de Recherche 9191 Centre National de la Recherche Scientifique and Unité Mixte de Recherche 247 Institut de Recherche pour le Développement, Université Paris-Saclay, Gif-sur-Yvette, France

[2] Laboratoire Ecologie et Biologie des Interactions, Equipe Ecologie Evolution Symbiose, Unité Mixte de Recherche 7267 Centre National de la Recherche Scientifique, Université de Poitiers, Poitiers, France

*Corresponding author : clement.gilbert@egce.cnrs-gif.fr

Abstract

Insects are major contributors to our understanding of the interaction between transposable elements (TEs) and their hosts, owing to seminal discoveries, as well as to the growing number of sequenced insect genomes, population genomics and functional studies. Insect TE landscapes are highly variable both within and across insect orders, although phylogenetic relatedness appears to correlate with similarity in insect TE content. This correlation is unlikely to be solely due to inheritance of TEs from shared ancestors and may partly reflect preferential horizontal transfer of TEs between closely related species. The influence of insect traits on TE landscapes remains however unclear. Recent findings indicate that in addition to being involved in insect adaptations and aging, TEs are seemingly at the cornerstone of insect antiviral immunity. Thus, TEs are emerging as essential insect symbionts, which may have deleterious or beneficial consequences on their hosts depending on context.

Introduction

Transposable elements (TEs) are pieces of DNA able to move from one genomic locus to another, often duplicating themselves in the process (13). The large diversity of TE types is classified in two classes: retrotransposons (Class-I TEs), which move through a replicative process involving reverse transcription of the TE mRNA, and DNA transposons (Class-II TEs), which mobilize through replicative or non-replicative transposition of DNA copies (24, 144). TEs are found in the genomes of virtually all organisms, often in very large amounts. The structure of TEs, their mobility and repeated nature have profoundly shaped the evolution of their hosts (22, 39, 63, 123).

TEs were discovered in the 1940s by Barbara McClintock, who showed that some genetic elements could move in the maize genome. This activity would induce chromosome breaks and sometimes modify the expression of neighboring genes with visible changes in maize kernel color (92). These initial studies have been mostly ignored for over three decades until they were finally rewarded by the Nobel prize of Physiology or Medicine (in 1983). TEs then became increasingly considered as important drivers of evolution, notably owing to major discoveries made in insects, mainly in *Drosophila* flies (8). Examples include the hybrid dysgenesis phenomenon, whereby F1 hybrids between males harboring *P* elements (class-II TEs) and females lacking them showed several defects such as gonadal atrophy and sterility (17, 70, 72). The molecular underpinnings of hybrid dysgenesis were only deciphered in 2008 through the discovery of the PIWI-associated RNAs (piRNAs)-mediated control of TEs (15). In the meantime, molecular cloning and extensive studies of the *P* element led to the development of this TE as one of the most widely used germline transformation tools in *Drosophila* (87, 125). Other TE-based transformation systems were developed for non-*Drosophila* species, including the one using *Piggybac*, a TE from the cabbage looper moth (*Trichoplusia ni*) that was discovered because of its ability to jump into a baculovirus (41, 56).

Surveys of the distribution of *P* elements in wild and laboratory strains of various *Drosophila* species also revealed that horizontal transfer (HT) of genetic material (here TEs) was not restricted to bacteria but that it could occur in eukaryotes (27, 45). Other landmark TE discoveries made in *Drosophila* include the so-called *P* neogenes derived from stationary *P* element copies, which led to the realization that TE protein domains can be co-opted to fulfill cellular functions (95, 102, 108). Perhaps even more surprising, the finding that the classical telomeres had been replaced in *Drosophila* flies with *HetA/TART/TAHRE* retrotransposons, which only transpose at chromosome ends, revealed that *bona fide* TEs could be beneficial to their host (9, 101). Finally, seminal studies mapped the genetic basis of insecticide resistance in the tobacco budworm (*Heliothis virescens*) and

61   *D. melanogaster* to TE-driven mutations, which contributed to set TEs as powerful facilitators of

62   genome evolution and host adaptation (26, 43, 98, 116).

63   These case-studies along with more recent multi-species, genome-wide surveys allow painting a

64   general picture of the composition, evolution and potential roles of TEs in insects, which is the topic

65   of this review. We first expose emerging macroevolutionary trends that underlie insect TE genomic

66   landscapes. We then discuss how the latest major discoveries on the impact of TEs in insect

67   adaptation, antiviral immunity and aging are continuously changing the way we view these genomic

68   symbionts.

69

1 - Evolution of TEs within and across insect orders

1 - 1 Diversity and abundance of insect TEs

72   The first comprehensive characterizations of TE abundance, diversity, age and genome distribution in

73   an insect were made possible thanks to analyses of the *D. melanogaster* genome (5, 7, 66, 69). These

74   studies illustrated how estimates of TE abundance and diversity heavily depend on the completeness

75   and quality of assembled genomes as well as on the methods used to annotate TEs. TE content

76   ranged from 2% of the *D. melanogaster* genome in early studies of euchromatic regions relying only

77   on similarity-based approaches, to about 15 % when heterochromatic regions were included and

78   combined de *novo* annotation pipelines were developed. The *D. melanogaster* TE landscape is

79   dominated by LTR-retrotransposons, which make up 10% of the genome, with non-LTR

80   retrotransposons and DNA transposons altogether occupying ~5% of the genome. The TE distribution

81   along *D. melanogaster* chromosomes is non-random, with a much lower density in euchromatic

82   regions where TEs generally occur as single full-length copies, than in pericentromeric

83   heterochromatin where TEs are arranged in dense islands of nested and often rearranged copies.

84   Bergman et al. (7) proposed that expression of such TE islands would produce transcripts made of

85   chimeric TEs which could be involved in suppressing the activity of multiple TE families. The

86   hypothesis was later confirmed with the discovery of dozens of pericentromeric TE-rich regions

87   called piRNA clusters acting as TE traps, the expression of which initiates piRNA-mediated TE

88   suppression (15, 150). Early studies based on between-copy genetic distances within TE families

89   found that most euchromatic TE copies resulted from very recent transposition events (14, 69). Their

90   recent origin is in agreement with subsequent studies showing that >80% of polymorphic TE copies

91   segregate at low (<0.2) frequency in *D. melanogaster* populations (75).

92　The analysis of other genomes quickly revealed that characteristics of fly TEs were not necessarily

93　representative of those found in other insects and that TE landscapes were highly dynamic in this

94　group (90). For example, although the overall content of classified TEs in the mosquito *Anopheles*

95　*gambiae* (about 13%) was close to that of *D. melanogaster*, LTR retrotransposons occupy about the

96　same amount of genomic space as non-LTR retrotransposons and DNA transposons in this species

97　(61, 96). Genome sequencing of another mosquito (*Aedes aegypti*), a Coleoptera (*Tribolium*

98　*castaneum*) and a Lepidoptera (*Bombyx mori*) further revealed that TE content varies substantially in

99　insects (from 6% in *T. castaneum* to 48% in *A. aegypti*) and that both DNA transposons and non-LTR

100　retrotransposons can be the dominant TEs in other species (97, 100, 139). The decrease in

101　sequencing costs was then accompanied by a rapid augmentation of the number of genome papers

102　with detailed sections on TE annotation (e.g. 19, 21, 23, 42, 44). One study reported that the

103　antarctic midge (*Belgica antarctica*) had both the smallest genome and the lowest TE content among

104　arthropods (71). It is now common to see papers reporting comparative analyses of multiple insect

105　genomes with efforts dedicated to surveying the evolution of TE contents across genomes (e.g. 1, 6,

106　15, 21, 28, 37, 45), as well as studies specifically dedicated to comparisons of TE contents at scales

107　ranging from insect genus to whole arthropods (55, 105, 110, 114, 148). We ourselves performed an

108　automated TE annotation on 195 publicly available insect genomes (103) (Figure 1).

109　A recent study also reported automated *de novo* TE-annotation and comparison of 73 arthropod

110　genomes (including 62 insects) (105). An interesting outcome of this analysis, as well as ours (Figure

111　1), is that the nature of a large fraction of the repetitive content of most insect genomes (11% in *D.*

112　*simulans* up to 93% in *Ephemera danica*; mean 56%, (105)) cannot be assessed using automated

113　procedures. Thus, interpretations on variation in TE content in insects must be taken with caution.

114　Bearing this limitation in mind, broad-scale TE landscape analyses reveal a high diversity of TEs in

115　insects, with most known TE superfamilies present in most insect species (105, 148). They also unveil

116　some degree of homogeneity in TE composition among insect groups spanning diverse taxonomic

117　levels (105). For example, in six tsetse fly species (*Glossina* spp), the genome proportion occupied by

118　DNA transposons (12.8-14.7%), LTR (0.4-0.8%) and non-LTR (5.8-2.9%) retrotransposons has

119　remained stable over 25 million years (2). In three Hemipteran species having diverged >200 million

120　years ago (*Cimex lectularius*, *Halyomorpha halys* and *Oncopeltus fasciatus*), genome proportions of

121　TE classes differ between species but non-LTR retrotransposons (7-14%) are always in higher

122　amounts than DNA transposons (1.3-4%) and LTR retrotransposons (0.6-1.4%). The situation is similar

123　in four lepidopteran species which diverged >100 million years ago (*Helicoverpa armigera*, *Bombyx*

124　*mori*, *Danaus plexippus*, *Heliconius melpomene*), in which non-LTR retrotransposons (4.2-20.4%) are

125　also always in higher amount than DNA transposons (0.9-3.1%) and LTR retrotransposons (0.3-0.7%).

126   While homogeneity in TE composition exists in some cases, pronounced differences have been found

127   in other cases. In hemipterans for example, SINEs and DNA transposons of the pea aphid

128   (*Acyrtosiphon pisum*) occupy the largest genome fraction (7.1% and 8.6%, respectively) while they

129   are either present in low amount (mean=1.2% and 2.3%, respectively) and/or occupy the lowest

130   genome fraction in the other six surveyed species (105). In coleopterans, DNA transposons occupy a

131   much larger genome fraction than retrotransposons in *T. castaneum* and *Anoplophora glabripennis*

132   while they are in minority compared to retrotransposons in *Agrilus plannipenis* and *Leptinotarsa*

133   *decemlineata* (105). Large differences in TE landscapes also exist at the genus level. Among 16

134   *Heliconius* butterfly species, TE composition may be similar at the level of TE class (DNA transposons

135   versus retrotransposons), but may differ drastically within each TE Class (110). For example, LTR

136   retrotransposons are almost absent in *H. doris* and *H. burneyi* but they make up 2 to 6% of the

137   genome of the other 14 species. Similarly, eight *Heliconius* species have a genome proportion of

138   SINEs <1% while SINEs make up 2 to 6% of the genome of the remaining eight species.

139   **1 – 2 The influence of vertical and horizontal transmission on insect TE landscapes**

140   The TE content of a genome is an equilibrium between the rate of TE acquisition, the replication

141   dynamics of TEs within the genome and the rate of TE loss by degradation/deletion. TEs are acquired

142   from ancestors (vertical inheritance) and potentially from other organisms (HT). TE replication

143   dynamics reflect rates of TE insertions (TE activity), which may be suppressed by host factors, as well

144   as the fate of these insertions under the influence of natural selection and genetic drift (73, 79, 115,

145   135). While the relative importance of the factors and processes shaping TE landscapes is not well

146   understood, insects are one of the groups in which they have been the most intensely studied so far

147   (see also (129) for a study in nematodes).

148   Multiple large-scale studies indicate a positive correlation between phylogenetic relatedness and

149   similarity of TE landscapes in insects (see above and refs. (105, 148)), although this correlation is not

150   formally tested. Variation in genome size also shows a strong phylogenetic signal and is positively

151   correlated to TE content both at the scale of arthropods (102, 146) and within insect orders: in

152   *Drosophila* species (40, 122), wood-white butterflies (130) and fireflies (84). This observation is

153   somewhat expected given that TEs, like any other genome component, are inherited from parents to

154   descendants. Yet in several pairs of closely related species showing relatively similar TE landscapes,

155   such as *D. melanogaster* and *D. simulans*, most TE copies have originated from transposition events

156   that are much more recent than the species split (14, 105). Their recent origin is further supported by

157   the fact that TE insertions in *D. melanogaster* are generally polymorphic within the species (75). As a

158   result, relatively few TE copies may be inherited over millions of years in a context of ongoing TE

159   activity, and species would be expected to diverge quickly with respect to their TE content.

160   On the other hand, a systematic survey of HT of TEs (HTT) between *D. melanogaster*, *D. simulans* and

161   *D. yakuba*, which diverged 11 million years ago, revealed that at least one third of the TE families in

162   these three species have been acquired through HTT, with an overall HTT rate of 0.035 events per TE

163   family per million years (4). It was also found that the number of HTT was higher between closely

164   related species than between distantly related ones, with twice as many events between *D.*

165   *melanogaster* and *D. simulans* as between any of these species and *D. yakuba* (4). A similar study

166   extended to 195 insect genomes inferred no fewer than 2248 HTT events between species having

167   diverged >40 million years ago (103). In addition to showing that HTT is not infrequent and has had a

168   strong impact on genome evolution in insects, with on average 2% of insect genome content deriving

169   from horizontally transferred TEs (and up to 25%), this study confirmed the role of phylogenetic

170   relatedness in favoring HTT. Preferential HTT between related species may thus explain the

171   phylogenetic signal in insect TE landscape variation in conjunction to inheritance from shared

172   ancestors.

173   In addition to phylogeny, geography was found to partly explain the distribution of HTT along the

174   insect tree. More HTT was detected between species originating from the same biogeographic realm

175   than between species from different ones (103). The observation that species more likely to be in

176   (direct or indirect) contact are also more likely to exchange TEs supports a role for ecological

177   relationships in shaping HTT. This is in line with a series of studies reporting HTT between animal

178   hosts and their parasites (47, 78, 128, 137, 141), but not with a study that did not detect HTT

179   between *Drosophila* species and two of their parasitoid wasps (99). Given the relatively high rate of

180   HTT in insects (103), it is likely that future studies implementing appropriate taxon sampling will

181   further delineate the types of ecological relationships, including host-parasite relationships, that

182   affect HTT patterns in insects (135). Interestingly, if was recently found that lepidopterans were

183   significantly more prone to HTT than other insect orders (112). The higher susceptibility of

184   lepidopterans to HTT may be explained by a number of different factors, one of which being

185   exposure to large double-stranded DNA viruses belonging to the Baculoviridae family. Most known

186   baculoviruses indeed infect lepidopterans and a large number of TEs are able to jump from infected

187   moth species to the genome of several baculoviruses, suggesting a role for these viruses as vectors of

188   HTT (46).

1 - 3 Are insect TE landscapes shaped by ecological and life history traits?

The replication dynamics of TEs may depend on certain life history traits of their hosts, as these traits may condition the impact of TE insertions on fitness and how these insertions evolve in host populations. One candidate trait is reproductive mode, since reduced recombination rates in asexual taxa may lead to the accumulation of TE insertions, as seen in non-recombinant sex chromosomes (18). However, asexual arthropods (including asexual wasps) do not have more TEs than their sexual counterparts (6, 77, 119). By contrast, social complexity is negatively correlated with TE abundance and TE diversity in bees, which may be explained by increased recombination rates in highly social insects and/or the pronounced hygienic behavior in these insects, leading to a lower exposure to parasites that may act as vectors of HTT (68). Yet, in social ants displaying hygienic behavior, TE content and diversity are similar to non-solitary insects (42). Furthermore, in termites, the socially more complex *Macrotermes natalensis* has a larger genome (1.31 Gb) and higher TE content (about 30%) than *Zootermopsis nevadensis* (562 Mb and about 16%, respectively) (76). Thus, whether or not TE landscapes are shaped by levels of sociality in bees, the influence of this factor may not hold for other social insects.

The ecological niche has been proposed as a potential factor shaping insect TE landscapes. For example, the low TE content of the antarctic midge *B. antartica* (0.12%) is thought to result from the paucity of contacts this extreme specialist has with other organisms in the species-poor polar environment, combined to strong selective pressures due to pronounced physiological and ecological constraints (71). In apparent agreement with a link between ecological niche and TE content, another extreme specialist, the fig wasp *Ceratosolen solmsi* also has one of the lowest TE contents of all insects (6.4%) (149). Yet, when comparing TE landscapes between *Drosophila* species in a phylogenetic context, no significant difference was found between specialists and generalists (40). The level of ecological specialization of *Drosophila* species, which is not as extreme as in *B. antartica* and *C. solmsi*, may not be high enough to impact TE content. Thus, reevaluating TE content in a controlled phylogenetic framework including related specialist and generalist species is required to properly test the influence of ecological niche on this variable. More generally, how host physiology and ecology govern TE replication dynamics deserves further studies. Such works may help to assess how the similarity in TE contents among related species, despite sustained TE activity, reflects some form of stabilizing/parallel evolution under similar constraints.

2 – The fate and impact of TEs in insect genomes

The development of next-generation sequencing technologies and analytical tools have enabled studies of the population genomics of insect TEs, many of which have focussed on *D. melanogaster* (3, 12). These studies have shown that a large fraction of the TEs inserted in the *D. melanogaster* reference genome are not fixed in the species (106). In addition, thousands of non-reference TE insertions segregate in *D. melanogaster* populations (25, 74, 80, 83, 109).

The majority of polymorphic TE insertions are present at lower frequencies than expected for neutral variants, suggesting that purifying selection (deleterious effects) plays a major role in TE population dynamics in *D. melanogaster* (10). A similar pattern of TE dynamics has been found in *D. simulans* (75). The deleterious effect of TEs can in part be explained by insertional mutagenesis upon integration in the genome. Consistently, more than half of spontaneous mutant phenotypes in *D. melanogaster* are caused by TE insertions disrupting gene function (33) and very few TE insertions are found in protein-coding genes in natural populations (74, 106). At the post-insertional level, an important cause for TE deleteriousness is ectopic recombination between non allelic, homologous TE insertions, which can generate deleterious genomic rearrangements (74, 106). It will be interesting to assess to what extent TE insertions are deleterious for insects other than *Drosophila*.

2 - 1 Insect adaptation

Even though most TE insertions are deleterious or neutral, some insertions are expected to be beneficial to their carriers. This is the case of several TE-mediated mutations underlying resistance to insecticides in the pink bollworm *Pectinophora gossypiella*, a major pest of cotton, and in *D. melanogaster* (116). In *P. gossypiella*, several independent TE insertions disrupting the *PgCad1* gene confer resistance to the widely used *Bacillus thuringiensis* Cry1Ac toxin (37, 138). In *D. melanogaster*, *cis*-regulatory sequences of an *Accord* LTR retrotransposon inserted in the 5' regulatory region of the *Cyp6g1* gene are responsible for *Cyp6g1* upregulation and increased resistance to DDT (21, 26). Furthermore, an adaptive allelic series in *Cyp6g1* involving successive mutations (including *P* and *HMS-Beagle* TE insertions) was later found to be associated with increasing DDT resistance levels (120). Similarly, resistance to organophosphate insecticide is caused by a *Doc1420* LINE retrotransposon that truncates the *CHKov1* gene and produces altered transcripts (1). Interestingly, the new variant may have initially been selected as a defense against viral infections, thereby pre-adapting flies to insectides (86). Yet another mechanism underlies TE-mediated increased resistance to cyclodiene dieldrin: *Rdl* gene duplication through ectopic recombination between *Roo* LTR retrotransposons (113). DNA transposons can also be involved in xenobiotic resistance, as exampified by the *Pogo* TE affecting polyadenylation signal choice in the *CG11699* gene and

253   conferring resistance to carbofuran and benzaldehyde (89). These examples serve to highlight the

254   diversity of TE families and molecular mechanisms through which TEs can be adaptive to their hosts,

255   which is all the more spectacular with respect to insecticide resistance in *D. melanogaster*, when

256   considering that this species is not a pest and has not been targeted by insecticides.

257   The contribution of TEs to *D. melanogaster* adaptation is not restricted to insecticide resistance, as

258   genome-wide studies of TE insertion polymorphisms have identified dozens of candidate adaptive TE

259   insertions (50, 51, 74, 80, 111). Population frequency distributions of TE insertions suggest that some

260   may play a role in adaptation to temperate climates (*D. melanogaster* has originated in the

261   afrotropics), often by induction of regulatory changes in nearby genes (50, 51). Functional gene

262   categories such as stress response, behavior and development appear to be particularly prone to TE-

263   mediated adaptation (111). In some cases, ecologically-relevant fitness benefits have been clearly

264   connected to specific TE insertions, such as a *Bari1* DNA transposon mediating oxidative stress

265   response (52, 54), *P*, *Jockey* and *Roo* elements involved in thermotolerance (81, 93) and an *invader4*

266   retrotransposon inducing faster development time (133). Interestingly, negative fitness effects of

267   such TE insertions have sometimes been identified, which may represent costs of selection and may

268   explain why the TE insertions have not reached fixation (52, 133).

269   Adaptive TE insertions have also been found in various insects other than *D. melanogaster*. For

270   example, in the peppered moth *Biston betularia*, industrial melanism is caused by a *carb-TE* DNA

271   transposon in the first intron of the *cortex* gene which affects gene expression (60). This mutation

272   underlies the replacement of a pale form by a black form driven by the interplay between bird

273   predation and coal pollution during the Industrial Revolution in the UK. Wing color polymorphism is

274   also associated with a TE insertion in the *Colias crocea* butterfly, confering developmental and

275   reproductive advantages (147). In Chinese populations of the migratory locust (*Locusta migratoria*),

276   an *Lm1* SINE insertion in the third exon of the *Hsp90* gene produces an alternative splicing form

277   associated with faster development and higher developmental synchrony (20). Remarkably, only

278   heterozygotes are present in natural populations, suggesting there is an heterozygote advantage that

279   is maintained by balancing selection. It has also been proposed that TEs may be powerful agents of

280   rapid adaptation in invasive species (126), an hypothesis that has received empirical support from

281   several insects such as the ant *Cardiocondyla obscurior* (121) and the Asian tiger mosquito *Aedes*

282   *albopictus* (53).

283   While the aforementioned examples of adaptive TEs consider individual copies, the repetitive nature

284   of TEs and existence of multiple homologous sequences scattered in genomes offer the perspective

285   for TEs to collectively impact cellular processes. For example, *G2/Jockey-3* retrotransposons appear

286 to contribute to the structure and function of centromeres in *D. melanogaster* and *D. simulans* (19).

287 TEs may also rewire regulatory networks by supplying *cis* regulatory elements at chromosomal scale,

288 as shown with the recruitment of *Helitron* elements to regulate dosage compensation in *D. miranda*

289 sex chromosomes (34, 35). There is however no formal demonstration that TEs have actually

290 improved these cellular processes from their ancestral state. The involvement of TEs in these

291 processes could represent neutral or even slightly deleterious mutational changes.

292 **2 – 2 Antiviral immunity**

293 In addition to their role in adaptation of insect populations, recent studies are setting TEs as an

294 important component of insect antiviral immunity. In insects, viruses are mainly restricted by RNA

295 interference (RNAi) (134). This pathway relies on the recognition and cleavage of intracellular long

296 viral double-stranded (ds) RNAs by the endoribonuclease Dicer2, yielding 21-nucleotide small

297 interfering RNAs (siRNAs). One strand of siRNAs is then used by the endoribonuclease Argonaute 2 to

298 target and cleave complementary long single-stranded viral RNAs, resulting in viral restriction. While

299 this pathway confers immunity to infected insect cells, it has long remained unclear whether such

300 immunity could spread systemically. In plants and nematodes, systemic RNAi spread relies on

301 amplification of dsRNA molecules by RNA-dependent RNA-polymerases (RdRp) (59, 124), which are

302 absent in insects (82). Yet, Goic et al. (48, 49) showed that antiviral RNAi can be amplified through a

303 hitherto unknown mechanism in *Drosophila* flies and *Aedes* mosquitoes, which relies on the

304 formation of viral DNA (vDNA). Remarkably, vDNAs are produced through reverse transcription of

305 viral RNAs by insect retrotransposon proteins, instead of an RdRp. This process primarily takes place

306 in hemocytes, in which viral dsRNAs released by infected cells accumulate (107, 131). vDNAs are then

307 transcribed and they trigger the formation of secondary viral siRNAs which are loaded into exosome-

308 like vesicles. They are then secreted in the extracellular environment and spread over the entire

309 insect body, conferring systemic antiviral immunity (131). vDNAs are generated as both circular and

310 linear molecules, some of which may end up being integrated into the host genome (132). Isolation

311 and sequencing of episomal circular DNA from infected mosquito cells and *Drosophila* flies revealed

312 that many circular vDNAs (cvDNAs) correspond to virus-LTR retrotransposon chimeras. Strikingly, the

313 amount of circular DNA containing virus-LTR retrotransposon chimeras produced by mosquito cells is

314 typically much higher than that containing chimeras of non-virus host sequence and

315 retrotransposons (132). Furthermore, the viral portion of virus-LTR retrotransposon chimeras is

316 enriched for defective viral genomes (107). Together, these observations suggest that viral defective

317 RNAs are recognized and preferentially reverse-transcribed over non-viral RNAs by LTR

318 retrotransposon-encoded reverse transcriptases, through a process involving copy-choice

319 recombination and seemingly depending on Dicer2 (107, 131, 132). Importantly, inhibition of reverse

320 transcriptase activity in flies and mosquitoes leads to accelerated death of infected individuals (48,

321 49). Thus, in the context of viral infections, the presence of LTR retrotransposons capable of

322 producing active reverse transcriptases is crucial for fly and mosquito survival. In principle, as long as

323 sources of reverse transcriptase activity are available throughout the evolutionary history of insects,

324 this antiviral protection system does not require co-option of a specific TE copy that would become

325 stationary and evolve under purifying selection like a regular gene. Thus, while in conflict with their

326 hosts because of the overall deleterious consequences of transposition, insect LTR retrotransposons

327 may become beneficial when host individuals incur viral infections.

328 Even more remarkably, insect retrotransposons have recently been shown to catalyse vDNA

329 integration into piRNA clusters of an *Aedes aegypti* cell line, leading to formation of endogenous viral

330 elements (EVE) (132, 143). Such EVEs are co-transcribed with other TE remnants from piRNA clusters,

331 leading to the production of anti-sense viral piRNAs that provide antiviral immunity. Thus, in addition

332 to their role in the amplification and systemic spread of antiviral immunity, insect retrotransposons

333 and their piRNA-mediated host surveillance system have the potential to confer transgenerational

334 antiviral immunity to their hosts (132, 143). The finding that LTR retrotransposons may be essential

335 to insect antiviral immunity opens new perspectives on the way we conceive insect/TE relationships.

336 Much like other TEs such as *TBE* DNA transposons in the ciliate *Oxytricha trifillax* (136) or the

337 telomeric non-LTR retrotransposons in *Drosophila* flies (118), and much like cellular endosymbionts

338 such as *Wolbachia* and *Cardinium* (44, 65, 142), the position of insect LTR retrotransposons on the

339 continuum of host/symbiont interactions may vary in unexpected ways between conflict and

340 cooperation, depending on context (23).

341 **2 – 3 Aging**

342 Aging is typically associated with the progressive alteration of genome regulation, notably due to the

343 loss of repressive heterochromatin marks (145). Several studies have shown that this dysregulation is

344 associated with an increase of TE expression in various eukaryotic organisms including yeasts (91),

345 nematodes (30), mice (29) and human cell lines (28). Thus, it has been proposed that expression and

346 transposition of TEs that are silenced in young individuals may strongly affect cell and genome

347 integrity in old individuals, triggering or reinforcing aging symptoms (31, 145). A series of elegant

348 experiments performed in *D. melanogaster* seemingly support this model. In particular, it was shown

349 that Dicer2-mutant individuals had more DNA double-strand breaks, presumably caused by activated

350 TEs, than controls in the nuclei of fat body cells (146). This is consistent with the known role of Dicer2

351 in the formation of repressive heterochromatin around repeated DNA (including TEs) via the

352 recruitment of the Su(var)3-9 histone methyl-transferase (38, 104). Importantly, decreasing TE

353 activity in these Dicer2 mutants through the addition of a reverse-transcriptase inhibitor led to

354 reduced transposition and increased lifespan. Thus, TE activity rather than the mere deterioration of

355 chromatin structure could be directly implicated in aging in *D. melanogaster* (146). In apparent

356 agreement, the lower lifespan of male versus female *D. melanogaster* flies has been attributed to a

357 higher amount of heterochromatin in males, which possess a highly repeat- and heterochromatin-

358 rich Y chromosome (16, 88).

359 The finding that TEs largely contribute to aging in *D. melanogaster* has recently been complemented

360 by studies in social insects. These insects are ideal models in aging research because individuals from

361 a given colony sharing an identical genotype show large differences in lifespan, depending on the

362 caste they belong to. In the termite *Macrotermes bellicosus* for example, reproductive castes (kings

363 and queens) can live for as long as 20 years whereas the lifespan of major and minor workers never

364 exceeds a few months. Elsner et al. (36) have shown that, while overall TE and gene expression levels

365 in heads were stable through the lifetime of kings and queens, several hundreds of TE copies were

366 expressed at higher levels in old versus young major workers. The higher TE expression in old

367 workers was associated with down-regulation of four genes involved in the piRNA pathway,

368 suggesting that aging in major workers may be due to a decrease in TE suppression. Since these

369 genes were expressed at equal levels in old and young reproductive castes, it was proposed that the

370 longer lifespan of kings and queens evolved thanks to cooption of the piRNA pathway (36) to

371 suppress TEs in their somatic tissues (here heads). Though interesting, this hypothesis does not

372 explain why piRNA genes are expressed (and TEs suppressed) in heads of young major workers (85).

373 In fact, somatic expression of the piRNA pathway has been inferred to be ancestral to all arthropods,

374 with only few species (including *D. melanogaster*), having lost this pathway in the soma (82). Thus,

375 rather than cooption of the piRNA pathway in heads of *M. bellicosus* reproductives, observations

376 may be more consistent with a secondary loss of this pathway in heads of old major workers.

377

3 – Conclusions and perspectives

379 Major breakthroughs in our understanding of eukaryote genome regulation and evolution have been

380 made through the study of insect TEs. Our knowledge of TE landscapes and their impact on insect

381 genomes and phenotypes has rapidly increased over the last ten years thanks to the availability of an

382 increasing number of genome sequences. Current large-scale surveys of insect TE landscapes must be

383 interpreted cautiously as TEs have been thoroughly mined only in a handful of insect genomes and

384 new TE families are still being discovered even in the best annotated genomes (148). Yet, global

385 trends emerge, suggesting that while TE landscapes can be highly variable both within and across

386 insect orders, phylogenetic relatedness generally correlates with similarity in TE content overall (84,

387 105, 122, 130, 148), even in cases where most extant TEs are not inherited from shared ancestors.

388 Vertical inheritance, horizontal transfer and the ecology/physiology of hosts are likely involved in

389 shaping insect TE landscapes, and it will be interesting to assess their relative contribution along the

390 insect phylogeny.

391 Because of their high diversity and the wide spectrum of molecular mechanisms through which they

392 can generate evolutionary innovations (including insertion, deletion, rearrangement, dispersion of

393 cis-regulatory elements and TE protein sequences), TEs appear to be involved in a variety of adaptive

394 events in many insect species. Thus, insects may be good models to quantify the respective

395 contribution of TE-mediated genomic change versus other mutations in adaptation.

396 The finding that retrotransposons are at the heart of antiviral response in some insects emerges as

397 one of the most important discoveries made in invertebrate immunity in the last few years (94). Yet,

398 the range of viruses that are targeted by TE-mediated small RNA responses in the wild remains to be

399 defined and it is unclear to what extent retrotransposons play a role in suppressing viruses beyond *D.*

400 *melanogaster* and *Aedes* mosquitoes. Whether and how retrotransposons involved in this response

401 vary among insects is also unknown. In addition, it will be interesting to further decipher the

402 mechanisms by which insect retrotransposons selectively recognize and reverse-transcribe viral RNAs

403 over cellular RNAs.

404 Monitoring and manipulating insect TE expression in old versus young insects uncovered a likely

405 causal role for TE activity in aging, in agreement with the retrotransposon theory of aging (31, 145).

406 In social insects, it will be interesting to assess whether TE activity in old workers is associated with a

407 deterioration of heterochromatin structure, as observed in *D. melanogaster*. Furthermore, a general

408 prediction of the retrotransposon theory of aging is that individuals with higher heterochromatin

409 loads should be shorter lived than individuals with low amounts of heterochromatin. While this

410 hypothesis seems to hold in *D. melanogaster* (16), testing it could also take advantage of species in

411 which B chromosomes, known to be largely heterochromatic, are present at substantial frequencies

412 (57). More generally, these results call for a detailed evaluation of the impact of somatic

413 transposition on insect biology.

414

415 **Figure legend**

416 Figure 1. Proportions of DNA transposons, LTR, LINE and SINE retrotransposons in insect genomes.

417 TEs were automatically annotated in 195 insect species (103). Because of space constraints, only one

species per clade that are ≤ 45 My old is shown here (to the exception of the pair *D. melanogaster*

and *D. simulans*). The "Unclassified" category of TEs represents elements that were not classified

(and categorized as "unknown") by the procedure. These exclude satellites and other repeat

elements that are unlikely to be TEs. The design of the figure is inspired from figure 1 of ref. (105).

The insect phylogeny on the left is that used in ref. (103). The data used to build the figure is

available upon request.

References

1. Aminetzach YT. 2005. Pesticide Resistance via Transposition-Mediated Adaptive Gene Truncation in Drosophila. *Science*. 309(5735):764–67

2. Attardo GM, Abd-Alla AMM, Acosta-Serrano A, Allen JE, Bateta R, et al. 2019. Comparative genomic analysis of six Glossina genomes, vectors of African trypanosomes. *Genome Biol.* 20(1):187

3. Barrón MG, Fiston-Lavier A-S, Petrov DA, González J. 2014. Population Genomics of Transposable Elements in *Drosophila*. *Annual Review of Genetics*. 48(1):561–81

4. Bartolome C, Bello X, Maside X. 2009. Widespread evidence for horizontal transfer of transposable elements across Drosophila genomes. *Genome Biol*. 10:R22

5. Bartolomé C, Maside X, Charlesworth B. 2002. On the abundance and distribution of transposable elements in the genome of Drosophila melanogaster. *Mol. Biol. Evol.* 19(6):926–37

6. Bast J, Schaefer I, Schwander T, Maraun M, Scheu S, Kraaijeveld K. 2016. No Accumulation of Transposable Elements in Asexual Arthropods. *Mol. Biol. Evol.* 33(3):697–706

7. Bergman CM, Quesneville H, Anxolabéhère D, Ashburner M. 2006. Recurrent insertion and duplication generate networks of transposable element sequences in the Drosophila melanogaster genome. *Genome Biol.* 7(11):R112

8. Biémont C. 2010. A Brief History of the Status of Transposable Elements: From Junk DNA to Major Players in Evolution: Figure 1.—. *Genetics*. 186(4):1085–93

9. Biessmann H, Valgeirsdottir K, Lofsky A, Chin C, Ginther B, et al. 1992. HeT-A, a transposable element specifically involved in "healing" broken chromosome ends in Drosophila melanogaster. *Mol. Cell. Biol.* 12(9):3910–18

10. Blumenstiel JP, Chen X, He M, Bergman CM. 2014. An age-of-allele test of neutrality for transposable element insertions. *Genetics*. 196(2):523–38

11. Bonasio R, Zhang G, Ye C, Mutti NS, Fang X, et al. 2010. Genomic comparison of the ants Camponotus floridanus and Harpegnathos saltator. *Science*. 329(5995):1068–71

451   12. Bourgeois Y, Boissinot S. 2019. On the Population Dynamics of Junk: A Review on the Population
452   Genomics of Transposable Elements. *Genes*. 10(6):419

453   13. Bourque G, Burns KH, Gehring M, Gorbunova V, Seluanov A, et al. 2018. Ten things you should
454   know about transposable elements. *Genome Biology*. 19(1):

455   14. Bowen NJ, McDonald JF. 2001. Drosophila euchromatic LTR retrotransposons are much younger
456   than the host species in which they reside. *Genome Res.* 11(9):1527–40

457   15. Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, et al. 2007. Discrete small RNA-generating loci
458   as master regulators of transposon activity in Drosophila. *Cell*. 128(6):1089–1103

459   16. Brown EJ, Nguyen AH, Bachtrog D. 2019. The Y chromosome contributes to sex-specific aging in
460   Drosophila. *bioRxiv*

461   17. Bucheton A, Paro R, Sang HM, Pelisson A, Finnegan DJ. 1984. The molecular basis of I-R hybrid
462   Dysgenesis in drosophila melanogaster: Identification, cloning, and properties of the I factor.
463   *Cell*. 38(1):153–63

464   18. Chalopin D, Volff J-N, Galiana D, Anderson JL, Schartl M. 2015. Transposable elements and early
465   evolution of sex chromosomes in fish. *Chromosome Research*. 23(3):545–60

466   19. Chang C-H, Chavan A, Palladino J, Wei X, Martins NMC, et al. 2019. Islands of retroelements are
467   major components of Drosophila centromeres. *PLOS Biology*. 17(5):e3000241

468   20. Chen B, Zhang B, Xu L, Li Q, Jiang F, et al. 2017. Transposable Element-Mediated Balancing
469   Selection at *Hsp90* Underlies Embryo Developmental Variation. *Molecular Biology and Evolution*,
470   p. msx062

471   21. Chung H, Bogwitz MR, McCart C, Andrianopoulos A, ffrench-Constant RH, et al. 2007. *Cis* -
472   Regulatory Elements in the *Accord* Retrotransposon Result in Tissue-Specific Expression of the
473   *Drosophila melanogaster* Insecticide Resistance Gene *Cyp6g1*. *Genetics*. 175(3):1071–77

474   22. Cordaux R, Batzer MA. 2009. The impact of retrotransposons on human genome evolution. *Nat
475   Rev Genet*. 10:691–703

476   23. Cosby RL, Chang N-C, Feschotte C. 2019. Host-transposon interactions: conflict, cooperation, and
477   cooption. *Genes Dev.* 33(17–18):1098–1116

478   24. Craig, Chandler, Gellert, Lambowitz, Rice, Sandmeyer, eds. 2015. *Mobile DNA III*. American
479   Society of Microbiology

480   25. Cridland JM, Macdonald SJ, Long AD, Thornton KR. 2013. Abundance and Distribution of
481   Transposable Elements in Two Drosophila QTL Mapping Resources. *Molecular Biology and
482   Evolution*. 30(10):2311–27

483   26. Daborn PJ, Yen JL, Bogwitz MR, Le Goff G, Feil E, et al. 2002. A single p450 allele associated with
484   insecticide resistance in Drosophila. *Science*. 297(5590):2253–56

485 27. Daniels SB, Peterson KR, Strausbaugh LD, Kidwell MG, Chovnick A. 1990. Evidence for horizontal
486   transmission of the P transposable element between Drosophila species. *Genetics*. 124:339–55

487 28. De Cecco M, Criscione SW, Peckham EJ, Hillenmeyer S, Hamm EA, et al. 2013. Genomes of
488   replicatively senescent cells undergo global epigenetic changes leading to gene silencing and
489   activation of transposable elements. *Aging Cell*. 12(2):247–56

490 29. De Cecco M, Criscione SW, Peterson AL, Neretti N, Sedivy JM, Kreiling JA. 2013. Transposable
491   elements become active and mobile in the genomes of aging mammalian somatic tissues. *Aging*
492   *(Albany NY)*. 5(12):867–83

493 30. Dennis S, Sheth U, Feldman JL, English KA, Priess JR. 2012. C. elegans Germ Cells Show
494   Temperature and Age-Dependent Expression of Cer1, a Gypsy/Ty3-Related Retrotransposon.
495   *PLoS Pathogens*. 8(3):e1002591

496 31. Driver CJ, McKechnie SW. 1992. Transposable elements as a factor in the aging of Drosophila
497   melanogaster. *Ann. N. Y. Acad. Sci.* 673:83–91

498 32. Drosophila 12 Genomes Consortium, Clark AG, Eisen MB, Smith DR, Bergman CM, et al. 2007.
499   Evolution of genes and genomes on the Drosophila phylogeny. *Nature*. 450(7167):203–18

500 33. Eickbush T. 2002. Fruit flies and humans respond differently to retrotransposons. *Current*
501   *Opinion in Genetics & Development*. 12(6):669–74

502 34. Ellison CE, Bachtrog D. 2013. Dosage Compensation via Transposable Element Mediated
503   Rewiring of a Regulatory Network. *Science*. 342(6160):846–50

504 35. Ellison CE, Bachtrog D. 2015. Non-allelic gene conversion enables rapid evolutionary change at
505   multiple regulatory sites encoded by transposable elements. *eLife*. 4:

506 36. Elsner D, Meusemann K, Korb J. 2018. Longevity and transposon defense, the case of termite
507   reproductives. *Proceedings of the National Academy of Sciences*. 115(21):5504–9

508 37. Fabrick JA, Mathew LG, Tabashnik BE, Li X. 2011. Insertion of an intact CR1 retrotransposon in a
509   cadherin gene linked with Bt resistance in the pink bollworm, Pectinophora gossypiella: CR1
510   retrotransposon and Bt resistance. *Insect Molecular Biology*. 20(5):651–65

511 38. Fagegaltier D, Bougé A-L, Berry B, Poisot É, Sismeiro O, et al. 2009. The endogenous siRNA
512   pathway is involved in heterochromatin formation in Drosophila. *Proceedings of the National*
513   *Academy of Sciences*. 106(50):21258–63

514 39. Feschotte C, Pritham EJ. 2007. DNA transposons and the evolution of eukaryotic genomes. *Annu*
515   *Rev Genet*. 41:331–68

516 40. Fonseca PM, Moura RD, Wallau GL, Loreto ELS. 2019. The mobilome of Drosophila incompta, a
517   flower-breeding species: comparison of transposable element landscapes among generalist and
518   specialist flies. *Chromosome Res.* 27(3):203–19

519 41. Fraser MJ, Smith GE, Summers MD. 1983. Acquisition of Host Cell DNA Sequences by

520   Baculoviruses: Relationship Between Host DNA Insertions and FP Mutants of Autographa

521   californica and Galleria mellonella Nuclear Polyhedrosis Viruses. *Journal of Virology*. 47:287–300

522 42. Gadau J, Helmkampf M, Nygaard S, Roux J, Simola DF, et al. 2012. The genomic impact of 100

523   million years of social evolution in seven ant species. *Trends Genet.* 28(1):14–21

524 43. Gahan LJ, Gould F, Heckel DG. 2001. Identification of a gene associated with Bt resistance in

525   Heliothis virescens. *Science*. 293(5531):857–60

526 44. Gavotte L, Mercer DR, Stoeckle JJ, Dobson SL. 2010. Costs and benefits of Wolbachia infection in

527   immature Aedes albopictus depend upon sex and competition level. *Journal of Invertebrate*

528   *Pathology*. 105(3):341–46

529 45. Gilbert C, Feschotte C. 2018. Horizontal acquisition of transposable elements and viral

530   sequences: patterns and consequences. *Current Opinion in Genetics & Development*. 49:15–24

531 46. Gilbert C, Peccoud J, Chateigner A, Moumen B, Cordaux R, Herniou EA. 2016. Continuous Influx

532   of Genetic Material from Host to Virus Populations. *PLoS Genetics*. 12:e1005838

533 47. Gilbert C, Schaack S, Pace JK, Brindley PJ, Feschotte C. 2010. A role for host-parasite interactions

534   in the horizontal transfer of transposons across phyla. *Nature*. 464:1347–50

535 48. Goic B, Stapleford KA, Frangeul L, Doucet AJ, Gausson V, et al. 2016. Virus-derived DNA drives

536   mosquito vector tolerance to arboviral infection. *Nat Commun*. 7:12410

537 49. Goic B, Vodovar N, Mondotte JA, Monot C, Frangeul L, et al. 2013. RNA-mediated interference

538   and reverse transcription control the persistence of RNA viruses in the insect model Drosophila.

539   *Nature Immunology*. 14:396–403

540 50. González J, Karasov TL, Messer PW, Petrov DA. 2010. Genome-Wide Patterns of Adaptation to

541   Temperate Environments Associated with Transposable Elements in Drosophila. *PLoS Genetics*.

542   6(4):e1000905

543 51. González J, Lenkov K, Lipatov M, Macpherson JM, Petrov DA. 2008. High Rate of Recent

544   Transposable Element–Induced Adaptation in Drosophila melanogaster. *PLoS Biology*. 6(10):e251

545 52. Gonzalez J, Macpherson JM, Petrov DA. 2009. A Recent Adaptive Transposable Element Insertion

546   Near Highly Conserved Developmental Loci in Drosophila melanogaster. *Molecular Biology and*

547   *Evolution*. 26(9):1949–61

548 53. Goubert C, Henri H, Minard G, Valiente Moro C, Mavingui P, et al. 2017. High-throughput

549   sequencing of transposable element insertions suggests adaptive evolution of the invasive Asian

550   tiger mosquito towards temperate environments. *Molecular Ecology*. 26(15):3968–81

551 54. Guio L, Barrón MG, González J. 2014. The transposable element *Bari-Jheh* mediates oxidative

552   stress response in Drosophila. *Molecular Ecology*. 23(8):2020–30

553   55. Han M-J, Zhou Q-Z, Zhang H-H, Tong X, Lu C, et al. 2016. iMITEdb: the genome-wide landscape of
554   miniature inverted-repeat transposable elements in insects. *Database*. 2016:baw148

555   56. Handler AM, McCombs SD, Fraser MJ, Saul SH. 1998. The lepidopteran transposon vector,
556   piggyBac, mediates germ-line transformation in the Mediterranean fruit fly. *Proceedings of the*
557   *National Academy of Sciences*. 95(13):7520–25

558   57. Hanlon S, Hawley R. 2018. B Chromosomes in the Drosophila Genus. *Genes*. 9(10):470

559   58. Harrison MC, Jongepier E, Robertson HM, Arning N, Bitard-Feildel T, et al. 2018. Hemimetabolous
560   genomes reveal molecular basis of termite eusociality. *Nat Ecol Evol*. 2(3):557–66

561   59. Himber C, Dunoyer P, Moissiard G, Ritzenthaler C, Voinnet O. 2003. Transitivity-dependent and -
562   independent cell-to-cell movement of RNA silencing. *EMBO J.* 22(17):4523–33

563   60. Hof AE van't, Campagne P, Rigden DJ, Yung CJ, Lingley J, et al. 2016. The industrial melanism
564   mutation in British peppered moths is a transposable element. *Nature*. 534(7605):102–5

565   61. Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, et al. 2002. The genome sequence
566   of the malaria mosquito Anopheles gambiae. *Science*. 298(5591):129–49

567   62. Honeybee Genome Sequencing Consortium. 2006. Insights into social insects from the genome
568   of the honeybee Apis mellifera. *Nature*. 443(7114):931–49

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

571   64. International Aphid Genomics Consortium. 2010. Genome sequence of the pea aphid
572   Acyrthosiphon pisum. *PLoS Biol.* 8(2):e1000313

573   65. Joshi D, McFadden MJ, Bevins D, Zhang F, Xi Z. 2014. Wolbachia strain wAlbB confers both fitness
574   costs and benefit on Anopheles stephensi. *Parasites & Vectors*. 7(1):336

575   66. Kaminker JS, Bergman CM, Kronmiller B, Carlson J, Svirskas R, et al. 2002. The transposable
576   elements of the Drosophila melanogaster euchromatin: a genomics perspective. *Genome Biol*.
577   3(12):research0084.1

578   67. Kanost MR, Arrese EL, Cao X, Chen YR, Chellapilla S, et al. 2016. Multifaceted biological insights
579   from a draft genome sequence of the tobacco hornworm moth, Manduca sexta. *Insect Biochem*
580   *Mol Biol*. 76:118–47

581   68. Kapheim KM, Pan H, Li C, Salzberg SL, Puiu D, et al. 2015. Social evolution. Genomic signatures of
582   evolutionary transitions from solitary to group living. *Science*. 348(6239):1139–43

583   69. Kapitonov VV, Jurka J. 2003. Molecular paleontology of transposable elements in the Drosophila
584   melanogaster genome. *Proc. Natl. Acad. Sci. U.S.A.* 100(11):6569–74

585   70. Kelleher ES. 2016. Reexamining the *P* -Element Invasion of *Drosophila melanogaster* Through the
586   Lens of piRNA Silencing. *Genetics*. 203(4):1513–31

587   71. Kelley JL, Peyton JT, Fiston-Lavier A-S, Teets NM, Yee M-C, et al. 2014. Compact genome of the

588   Antarctic midge is likely an adaptation to an extreme environment. *Nat Commun*. 5:4611

589   72. Kidwell MG, Kidwell JF. 1975. Cytoplasm–chromosome interactions in Drosophila melanogaster.

590   *Nature*. 253(5494):755–56

591   73. Kofler R. 2019. Dynamics of Transposable Element Invasions with piRNA Clusters. *Molecular*

592   *Biology and Evolution*. 36(7):1457–72

593   74. Kofler R, Betancourt AJ, Schlötterer C. 2012. Sequencing of Pooled DNA Samples (Pool-Seq)

594   Uncovers Complex Dynamics of Transposable Element Insertions in Drosophila melanogaster.

595   *PLoS Genetics*. 8(1):e1002487

596   75. Kofler R, Nolte V, Schlötterer C. 2015. Tempo and Mode of Transposable Element Activity in

597   Drosophila. *PLOS Genetics*. 11(7):e1005406

598   76. Korb J, Poulsen M, Hu H, Li C, Boomsma JJ, et al. 2015. A genomic comparison of two termites

599   with different social complexity. *Frontiers in Genetics*. 6:

600   77. Kraaijeveld K, Zwanenburg B, Hubert B, Vieira C, De Pater S, et al. 2012. Transposon proliferation

601   in an asexual parasitoid. *Mol. Ecol.* 21(16):3898–3906

602   78. Kuraku S, Qiu H, Meyer A. 2012. Horizontal transfers of Tc1 elements between teleost fishes and

603   their vertebrate parasites, lampreys. *Genome Biol Evol*. 4:929–36

604   79. Le Rouzic A, Dupas S, Capy P. 2007. Genome ecosystem and transposable elements species.

605   *Gene*. 390(1–2):214–20

606   80. Lerat E, Goubert C, Guirao-Rico S, Merenciano M, Dufour A, et al. 2019. Population-specific

607   dynamics and selection patterns of transposable element insertions in European natural

608   populations. *Molecular Ecology*. 28(6):1506–22

609   81. Lerman DN, Feder ME. 2005. Naturally occurring transposable elements disrupt hsp70 promoter

610   function in Drosophila melanogaster. *Mol. Biol. Evol.* 22(3):776–83

611   82. Lewis SH, Quarles KA, Yang Y, Tanguy M, Frézal L, et al. 2018. Pan-arthropod analysis reveals

612   somatic piRNAs as an ancestral defence against transposable elements. *Nature Ecology &*

613   *Evolution*. 2(1):174–81

614   83. Linheiro RS, Bergman CM. 2012. Whole Genome Resequencing Reveals Natural Target Site

615   Preferences of Transposable Elements in Drosophila melanogaster. *PLoS ONE*. 7(2):e30008

616   84. Lower SS, Johnston JS, Stanger-Hall KF, Hjelmen CE, Hanrahan SJ, et al. 2017. Genome Size in

617   North American Fireflies: Substantial Variation Likely Driven by Neutral Processes. *Genome*

618   *Biology and Evolution*. 9(6):1499–1512

619   85. Lucas ER, Keller L. 2018. New explanation for the longevity of social insect reproductives:

620   Transposable element activity. *Proc. Natl. Acad. Sci. U.S.A.* 115(21):5317–18

621 86. Magwire MM, Bayer F, Webster CL, Cao C, Jiggins FM. 2011. Successive Increases in the

622   Resistance of Drosophila to Viral Infection through a Transposon Insertion Followed by a

623   Duplication. *PLoS Genetics*. 7(10):e1002337

624 87. Majumdar S, Rio DC. 2015. P Transposable Elements in Drosophila and other Eukaryotic

625   Organisms. *Microbiol Spectr*. 3(2):MDNA3-0004–2014

626 88. Marais GAB, Gaillard J-M, Vieira C, Plotton I, Sanlaville D, et al. 2018. Sex gap in aging and

627   longevity: can sex chromosomes play a role? *Biol Sex Differ*. 9(1):33

628 89. Mateo L, Ullastres A, González J. 2014. A Transposable Element Insertion Confers Xenobiotic

629   Resistance in Drosophila. *PLoS Genetics*. 10(8):e1004560

630 90. Maumus F, Fiston-Lavier AS, Quesneville H. 2015. Impact of transposable elements on insect

631   genomes and biology. *Current Opinion in Insect Science*. 7:30–36

632 91. Maxwell PH, Burhans WC, Curcio MJ. 2011. Retrotransposition is associated with genome

633   instability during chronological aging. *Proceedings of the National Academy of Sciences*.

634   108(51):20376–81

635 92. McClintock B. 1950. The origin and behavior of mutable loci in maize. *Proceedings of the

636   National Academy of Sciences*. 36(6):344–55

637 93. Merenciano M, Ullastres A, de Cara MAR, Barrón MG, González J. 2016. Multiple Independent

638   Retroelement Insertions in the Promoter of a Stress Response Gene Have Variable Molecular and

639   Functional Effects in Drosophila. *PLOS Genetics*. 12(8):e1006249

640 94. Miesen P, Joosten J, van Rij RP. 2016. PIWIs Go Viral: Arbovirus-Derived piRNAs in Vector

641   Mosquitoes. *PLOS Pathogens*. 12(12):e1006017

642 95. Miller WJ, McDonald JF, Nouaud D, Anxolabéhère D. 1999. Molecular domestication--more than

643   a sporadic episode in evolution. *Genetica*. 107(1–3):197–207

644 96. Neafsey DE, Waterhouse RM, Abai MR, Aganezov SS, Alekseyev MA, et al. 2015. Mosquito

645   genomics. Highly evolvable malaria vectors: the genomes of 16 Anopheles mosquitoes. *Science*.

646   347(6217):1258522

647 97. Nene V, Wortman JR, Lawson D, Haas B, Kodira C, et al. 2007. Genome sequence of Aedes

648   aegypti, a major arbovirus vector. *Science*. 316(5832):1718–23

649 98. Oliver KR, Greene WK. 2012. Transposable elements and viruses as factors in adaptation and

650   evolution: an expansion and strengthening of the TE-Thrust hypothesis. *Ecology and Evolution*.

651   2(11):2912–33

652 99. Ortiz MF, Wallau GL, Graichen DA, Loreto EL. 2015. An evaluation of the ecological relationship

653   between Drosophila species and their parasitoid wasps as an opportunity for horizontal

654   transposon transfer. *Molecular Genetics and Genomics*. 290:67–78

655   100.   Osanai-Futahashi M, Suetsugu Y, Mita K, Fujiwara H. 2008. Genome-wide screening and

656   characterization of transposable elements and their distribution analysis in the silkworm,

657   Bombyx mori. *Insect Biochem. Mol. Biol.* 38(12):1046–57

658   101.   Pardue M-L, DeBaryshe PG. 2011. Retrotransposons that maintain chromosome ends. *Proc.*

659   *Natl. Acad. Sci. U.S.A.* 108(51):20317–24

660   102.   Paricio N, Pèrez-Alonso M, Martinez-Sebastián maria J, de Frutos R. 1991. P sequences of

661   *Drosophilla Subobscura* lack exon 3 and may encode a 66 kd repressor-like protein. *Nucleic Acids*

662   *Research*. 19(24):6713–18

663   103.   Peccoud J, Loiseau V, Cordaux R, Gilbert C. 2017. Massive horizontal transfer of transposable

664   elements in insects. *Proc. Natl. Acad. Sci. U.S.A.* 114(18):4721–26

665   104.   Peng JC, Karpen GH. 2007. H3K9 methylation and RNA interference regulate nucleolar

666   organization and repeated DNA stability. *Nat. Cell Biol.* 9(1):25–35

667   105.   Petersen M, Armisén D, Gibbs RA, Hering L, Khila A, et al. 2019. Diversity and evolution of the

668   transposable element repertoire in arthropods with particular reference to insects. *BMC Evol.*

669   *Biol.* 19(1):11

670   106.   Petrov DA, Fiston-Lavier A-S, Lipatov M, Lenkov K, Gonzalez J. 2011. Population Genomics of

671   Transposable Elements in Drosophila melanogaster. *Molecular Biology and Evolution*.

672   28(5):1633–44

673   107.   Poirier EZ, Goic B, Tomé-Poderti L, Frangeul L, Boussier J, et al. 2018. Dicer-2-Dependent

674   Generation of Viral DNA from Defective Genomes of RNA Viruses Modulates Antiviral Immunity

675   in Insects. *Cell Host & Microbe*. 23(3):353-365.e8

676   108.   Quesneville H, Nouaud D, Anxolabehere D. 2005. Recurrent Recruitment of the THAP DNA-

677   Binding Domain and Molecular Domestication of the P-Transposable Element. *Molecular Biology*

678   *and Evolution*. 22(3):741–46

679   109.   Rahman R, Chirn G, Kanodia A, Sytnikova YA, Brembs B, et al. 2015. Unique transposon

680   landscapes are pervasive across *Drosophila melanogaster* genomes. *Nucleic Acids Research*.

681   43(22):10655–72

682   110.   Ray DA, Grimshaw JR, Halsey MK, Korstian JM, Osmanski AB, et al. 2019. Simultaneous TE

683   Analysis of 19 Heliconiine Butterflies Yields Novel Insights into Rapid TE-Based Genome

684   Diversification and Multiple SINE Births and Deaths. *Genome Biol Evol*. 11(8):2162–77

685   111.   Rech GE, Bogaerts-Márquez M, Barrón MG, Merenciano M, Villanueva-Cañas JL, et al. 2019.

686   Stress response, behavior, and development are shaped by transposable element-induced

687   mutations in Drosophila. *PLOS Genetics*. 15(2):e1007900

688   112.   Reiss D, Mialdea G, Miele V, de Vienne DM, Peccoud J, et al. 2019. Global survey of mobile

689   DNA horizontal transfer in arthropods reveals Lepidoptera as a prime hotspot. *PLOS Genetics*.

690   15(2):e1007965

691   113.   Remnant EJ, Good RT, Schmidt JM, Lumb C, Robin C, et al. 2013. Gene duplication in the

692   major insecticide target site, *Rdl* , in *Drosophila melanogaster*. *Proceedings of the National*

693   *Academy of Sciences*. 110(36):14705–10

694   114.   Rius N, Guillen Y, Delprat A, Kapusta A, Feschotte C, Ruiz A. 2016. Exploration of the

695   Drosophila buzzatii transposable element content suggests underestimation of repeats in

696   Drosophila genomes. *BMC Genomics*. 17:344

697   115.   Robillard É, Le Rouzic A, Zhang Z, Capy P, Hua-Van A. 2016. Experimental evolution reveals

698   hyperparasitic interactions among transposable elements. *Proceedings of the National Academy*

699   *of Sciences*. 113(51):14763–68

700   116.   Rostant WG, Wedell N, Hosken DJ. 2012. Transposable Elements and Insecticide Resistance.

701   In *Advances in Genetics*. 78:169–201. Elsevier

702   117.   Sadd BM, Barribeau SM, Bloch G, de Graaf DC, Dearden P, et al. 2015. The genomes of two

703   key bumblebee species with primitive eusocial organization. *Genome Biol.* 16:76

704   118.   Saint-Leandre B, Nguyen SC, Levine MT. 2019. Diversification and collapse of a telomere

705   elongation mechanism. *Genome Res.* 29(6):920–31

706   119.   Schaack S, Pritham EJ, Wolf A, Lynch M. 2010. DNA transposon dynamics in populations of

707   Daphnia pulex with and without sex. *Proc. Biol. Sci.* 277(1692):2381–87

708   120.   Schmidt JM, Good RT, Appleton B, Sherrard J, Raymant GC, et al. 2010. Copy Number

709   Variation and Transposable Elements Feature in Recent, Ongoing Adaptation at the Cyp6g1

710   Locus. *PLoS Genetics*. 6(6):e1000998

711   121.   Schrader L, Kim JW, Ence D, Zimin A, Klein A, et al. 2014. Transposable element islands

712   facilitate adaptation to novel environments in an invasive species. *Nature Communications*. 5(1):

713   122.   Sessegolo C, Burlet N, Haudry A. 2016. Strong phylogenetic inertia on genome size and

714   transposable element content among 26 species of flies. *Biology Letters*. 12(8):20160407

715   123.   Siguier P, Gourbeyre E, Chandler M. 2014. Bacterial insertion sequences: their genomic

716   impact and diversity. *FEMS Microbiology Reviews*. 38(5):865–91

717   124.   Sijen T, Fleenor J, Simmer F, Thijssen KL, Parrish S, et al. 2001. On the Role of RNA

718   Amplification in dsRNA-Triggered Gene Silencing. *Cell*. 107(4):465–76

719   125.   Spradling A, Rubin G. 1982. Transposition of cloned P elements into Drosophila germ line

720   chromosomes. *Science*. 218(4570):341–47

721   126.   Stapley J, Santure AW, Dennis SR. 2015. Transposable elements as agents of rapid adaptation

722   may explain the genetic paradox of invasive species. *Molecular Ecology*. 24(9):2241–52

723   127.   Suen G, Teiling C, Li L, Holt C, Abouheif E, et al. 2011. The genome sequence of the leaf-cutter

724   ant Atta cephalotes reveals insights into its obligate symbiotic lifestyle. *PLoS Genet.*

725   7(2):e1002007

726   128.   Suh A, Witt CC, Menger J, Sadanandan KR, Podsiadlowski L, et al. 2016. Ancient horizontal

727   transfers of retrotransposons between birds and ancestors of human pathogenic nematodes.

728   *Nat Commun*. 7:11396

729   129.   Szitenberg A, Cha S, Opperman CH, Bird DM, Blaxter ML, Lunt DH. 2016. Genetic Drift, Not

730   Life History or RNAi, Determine Long-Term Evolution of Transposable Elements. *Genome Biol*

731   *Evol*. 8(9):2964–78

732   130.   Talla V, Suh A, Kalsoom F, Dincă V, Vila R, et al. 2017. Rapid Increase in Genome Size as a

733   Consequence of Transposable Element Hyperactivity in Wood-White (Leptidea) Butterflies.

734   *Genome Biology and Evolution*. 9(10):2491–2505

735   131.   Tassetto M, Kunitomi M, Andino R. 2017. Circulating Immune Cells Mediate a Systemic RNAi-

736   Based Adaptive Antiviral Response in Drosophila. *Cell*. 169(2):314-325.e13

737   132.   Tassetto M, Kunitomi M, Whitfield ZJ, Dolan PT, Sánchez-Vargas I, et al. 2019. Control of RNA

738   viruses in mosquito cells through the acquisition of vDNA and endogenous viral elements. *eLife*.

739   8:

740   133.   Ullastres A, Petit N, González J. 2015. Exploring the Phenotypic Space and the Evolutionary

741   History of a Natural Mutation in *Drosophila melanogaster*. *Molecular Biology and Evolution*.

742   32(7):1800–1814

743   134.   van Rij RP, Saleh M-C, Berry B, Foo C, Houk A, et al. 2006. The RNA silencing endonuclease

744   Argonaute 2 mediates specific antiviral immunity in Drosophila melanogaster. *Genes &*

745   *Development*. 20(21):2985–95

746   135.   Venner S, Miele V, Terzian C, Biémont C, Daubin V, et al. 2017. Ecological networks to unravel

747   the routes to horizontal transposon transfers. *PLOS Biology*. 15(2):e2001536

748   136.   Vogt A, Goldman AD, Mochizuki K, Landweber LF. 2013. Transposon domestication versus

749   mutualism in ciliate genome rearrangements. *PLoS Genet.* 9(8):e1003659

750   137.   Walsh AM, Kortschak RD, Gardner MG, Bertozzi T, Adelson DL. 2013. Widespread horizontal

751   transfer of retrotransposons. *Proc. Natl. Acad. Sci. U.S.A.* 110(3):1012–16

752   138.   Wang L, Wang J, Ma Y, Wan P, Liu K, et al. 2019. Transposon insertion causes cadherin mis-

753   splicing and confers resistance to Bt cotton in pink bollworm from China. *Scientific Reports*. 9(1):

754   139.   Wang S, Lorenzen MD, Beeman RW, Brown SJ. 2008. Analysis of repetitive DNA distribution

755   patterns in the Tribolium castaneum genome. *Genome Biol.* 9(3):R61

756   140.   Wang X, Fang X, Yang P, Jiang X, Jiang F, et al. 2014. The locust genome provides insight into

757   swarm formation and long-distance flight. *Nat Commun*. 5:2957

758   141.   Wang X, Liu X. 2016. Close ecological relationship among species facilitated horizontal
759   transfer of retrotransposons. *BMC Evol Biol*. 16:201

760   142.   White JA, Kelly SE, Cockburn SN, Perlman SJ, Hunter MS. 2011. Endosymbiont costs and
761   benefits in a parasitoid infected with both Wolbachia and Cardinium. *Heredity*. 106(4):585–91

762   143.   Whitfield ZJ, Dolan PT, Kunitomi M, Tassetto M, Seetin MG, et al. 2017. The Diversity,
763   Structure, and Function of Heritable Adaptive Immunity Sequences in the Aedes aegypti
764   Genome. *Current Biology*. 27(22):3511-3519.e7

765   144.   Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, et al. 2007. A unified classification
766   system for eukaryotic transposable elements. *Nat. Rev. Genet.* 8(12):973–82

767   145.   Wood JG, Helfand SL. 2013. Chromatin structure and transposable elements in organismal
768   aging. *Front Genet*. 4:274

769   146.   Wood JG, Jones BC, Jiang N, Chang C, Hosier S, et al. 2016. Chromatin-modifying genetic
770   interventions suppress age-associated transposable element activation and extend life span in
771   Drosophila. *Proc. Natl. Acad. Sci. U.S.A.* 113(40):11277–82

772   147.   Woronik A, Tunström K, Perry MW, Neethiraj R, Stefanescu C, et al. 2019. A transposable
773   element insertion is associated with an alternative life history strategy. *Nature Communications*.
774   10(1):

775   148.   Wu C, Lu J. 2019. Diversification of Transposable Elements in Arthropods and Its Impact on
776   Genome Evolution. *Genes (Basel)*. 10(5):

777   149.   Xiao J-H, Yue Z, Jia L-Y, Yang X-H, Niu L-H, et al. 2013. Obligate mutualism within a host drives
778   the extreme specialization of a fig wasp genome. *Genome Biol.* 14(12):R141

779   150.   Zanni V, Eymery A, Coiffet M, Zytnicki M, Luyten I, et al. 2013. Distribution, evolution, and
780   diversity of retrotransposons at the flamenco locus reflect the regulatory properties of piRNA
781   clusters. *Proc. Natl. Acad. Sci. U.S.A.* 110(49):19842–47

782

783   **Figure 1**

784

785