Transposable element landscape in *Drosophila* populations selected for longevity

Daniel K. Fabian[1,2]*, Handan Melike Dönertaş[1], Matías Fuentealba[1,2], Linda Partridge[2,3] and Janet M. Thornton[1]

[1] European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, UK

[2] Institute of Healthy Ageing, Department of Genetics, Evolution and Environment, University College London, London, UK

[3] Max Planck Institute for Biology of Ageing, Cologne, Germany

* Corresponding author

E-mail: daniel.fabian@ebi.ac.uk

ABSTRACT

Transposable elements (TEs) inflict numerous negative effects on health and fitness as they replicate by integrating into new regions of the host genome. Even though organisms employ powerful mechanisms to demobilize TEs, transposons gradually lose repression during aging. The rising TE activity causes genomic instability and was implicated in age-dependent neurodegenerative diseases, inflammation and the determination of lifespan. It is therefore conceivable that long-lived individuals have improved TE silencing mechanisms resulting in reduced TE expression relative to their shorter-lived counterparts and fewer genomic insertions. Here, we test this hypothesis by performing the first genome-wide analysis of TE insertions and expression in populations of *Drosophila melanogaster* selected for longevity through late-life reproduction for 50-170 generations from four independent studies. Contrary to our expectation, TE families were generally more abundant in long-lived populations compared to non-selected controls. Although simulations showed that this was not expected under neutrality, we found little evidence for selection driving TE abundance differences. Additional RNA-seq analysis revealed a tendency for reducing TE expression in selected populations, which might be more important for lifespan than regulating genomic insertions. We further find limited evidence of parallel selection on genes related to TE regulation and transposition. However, telomeric TEs were genomically and transcriptionally more abundant in long-lived flies, suggesting improved telomere maintenance as a promising TE-mediated mechanism for prolonging lifespan. Our results provide a novel viewpoint indicating that reproduction at old age increases the opportunity of TEs to be passed on to the next generation with little impact on longevity.

INTRODUCTION

Aging, also known as senescence, is an evolutionary conserved process described as the progressive loss of physiological homeostasis starting from maturity with disease promotion, decline in phenotypic function, and increased chance of mortality over time as a consequence (Fabian and Flatt 2011; Flatt and Heyland 2011; López-Otín et al. 2013). At the molecular level, studies of loss-of-function mutations in model organisms such as yeast, *Caenorhabditis elegans*, *Drosophila melanogaster*, and mice have successfully identified key pathways underlying aging and longevity including the conserved insulin/insulin-like growth factor signaling (IIS) and target of rapamycin (TOR) nutrient-sensing network (Piper et al. 2008; Fontana et al. 2010; Gems and Partridge 2013; Pan and Finkel 2017). More recently, sequencing of whole genomes, transcriptomes, and epigenomes corroborated that aging has a complex genetic basis involving many genes and is accompanied by changes across a broad range of interconnected molecular functions (López-Otín et al. 2013).

While there has been a predominant focus on understanding the links between genes and phenotypes correlated with aging, the role of transposable elements (TEs) in senescence and longevity has received less attention even though their discovery by Barbara McClintock goes back more than half a century ago (McClintock 1950). TEs, or transposons, are selfish genetic elements that replicate and move within genomes of their hosts. In eukaryotes, TEs typically constitute a considerable portion of the genome, with estimates around ~3% in yeast, ~20% in *D. melanogaster*, ~70% in humans and ~85% in maize (Quesneville et al. 2005; Schnable et al. 2009; de Koning et al. 2011; Carr et al. 2012). To date, several thousand TE families broadly classified into DNA-transposons and retrotransposons multiplying via RNA intermediates have been identified and are known to vary hugely in their transpositional mobility (Jurka et al. 2011; Deniz et al. 2019). For example, only a small fraction of L1 retrotransposons are responsible for most of the transposition events in the human genome, while the vast majority of L1s and other TE families have been inactivated by the accumulation of structural and point mutations over evolutionary time scales (Brouha et al. 2003).

In spite of the substantial evidence implicating TEs in adaptive evolution and diseases, the majority of transposons residing in the genome are likely to be neutral or only slightly deleterious for host fitness (Barrón et al. 2014; Arkhipova 2018). Yet, their exact physiological functions and the extent to which particular TE insertions or whole TE classes contribute to host fitness is still under debate (Brunet and Doolittle 2015). In general, TE mobility causes genomic instability through insertional mutagenesis, which can directly affect coding sequences of genes or modify their transcription. Typically, TE insertions into or close to genes

100 impose negative consequences on health and have been associated with ~100 diseases in
101 humans, including cystic fibrosis, haemophilia and cancer (Hancks and Kazazian 2012). It is
102 not just through the insertion of TEs that their presence may be deleterious, but also by
103 causing detrimental chromosomal rearrangements resulting from ectopic recombination
104 between TE families with similar sequences in different genomic locations (Montgomery et al.
105 1987; Charlesworth et al. 1992; Petrov et al. 2011). Additionally, TE expression and translation
106 also allow the formation of toxic TE products that, for example, contribute to autoimmune
107 diseases, while TE activity and replication of an increased genomic TE content might indirectly
108 impose metabolic costs to the host (Kaneko et al. 2011; Barrón et al. 2014; Volkman and
109 Stetson 2014; Bogu et al. 2019). On the other hand, there is mounting experimental evidence
110 for positive selection on segregating TE insertions from multiple taxa confirming beneficial
111 phenotypic properties including insecticide and virus resistance in *Drosophila* (Daborn et al.
112 2002; Magwire et al. 2011; Kuhn et al. 2014; Li et al. 2018; Rech et al. 2019).
113
114 A common feature of TEs observed in various organisms including yeast, *D. melanogaster*,
115 *C. elegans*, mice, and humans, is the age-associated increase in transposition and
116 expression, which usually coincides with weakening of the host TE silencing machinery and
117 loss of genomic stability (Maxwell et al. 2011; Dennis et al. 2012; Solyom et al. 2012; De
118 Cecco et al. 2013; Li et al. 2013; Gorbunova et al. 2014; Chen et al. 2016; Bogu et al. 2019;
119 De Cecco et al. 2019). TEs have further been implicated in age-related neurodegenerative
120 diseases (e.g. Krug et al. 2017; Prudencio et al. 2017; Guo et al. 2018) and might promote
121 chronic inflammation observed during aging (Chen et al. 2014; De Cecco et al. 2019) further
122 supporting the involvement of TEs in senescence and longevity as proposed by the emerging
123 'transposable element theory of aging' (Kirkwood 1989; Sedivy et al. 2013). The age-related
124 change in TE activity detected in many tissues has mainly been attributed to chromatin
125 remodeling and the decline in repressive heterochromatin structure which is commonly rich in
126 transposable elements (Dimitri and Junakovic 1999; Wood and Helfand 2013; Chen et al.
127 2016; Wood et al. 2016). TEs that are not suppressed by chromatin structure are the target of
128 post-transcriptional silencing by the host RNA-interference (RNAi) machinery, mostly the piwi-
129 interacting RNA (piRNA) pathway, which is in turn also necessary for heterochromatin
130 formation and stability (Lippman and Martienssen 2004; Martienssen and Moazed 2015).
131 Indeed, research has identified longevity-promoting effects of several genes involved in the
132 RNAi machinery and heterochromatin formation (Mori et al. 2012; Wood and Helfand 2013;
133 Wood et al. 2016). Interestingly, it is possible that age-related misexpression of TEs is
134 exclusive to the soma due to efficient post-transcriptional TE silencing mediated by the piRNA
135 machinery in the germline (Sturm et al. 2015; Elsner et al. 2018; Erwin and Blumenstiel 2019).
136 Considering current evidence, it seems natural that longevity can be achieved through

137 impeding TE activity and controlling the genomic content of TEs. However, whether variation
138 in aging and lifespan within species is also mediated by transposons and their role in the
139 evolution of senescence is largely unknown.
140
141 Here, we analyze genomes of *D. melanogaster* populations experimentally selected for
142 increased lifespan through postponed reproduction from four independent studies to
143 understand the role of TEs in the evolution and genomic basis of late-life performance and
144 aging. The invertebrate *D. melanogaster* is an excellent model in this respect as it exhibits
145 abundant genetic and phenotypic variation in fecundity and traits related to aging that can be
146 selected for. In the present experiments, replicate populations derived from nature were
147 subjected to a late-life breeding scheme in which only flies surviving and fertile at old age
148 contributed to the subsequent generations, while control individuals reproduced earlier in life.
149 When the genomes were sequenced, the selection process had continued for over 30 years
150 with ~170 and ~150 generations of selection for Carnes et al. 2015 (Carnes2015) and Fabian
151 et al. 2018 (Fabian2018), and for 58 and 50 generations for Hoedjes et al. 2019 (Hoedjes2019)
152 and Remolina et al. 2012 (Remolina2012) enabling us to quantify differences in TE content of
153 long- and short-term evolutionary responses. Selection for postponed senescence has
154 resulted in phenotypic divergence of multiple fitness traits, most notably an ~8% to ~74%
155 increase in lifespan and improved old age fecundity at the cost of reduced early reproduction
156 (Luckinbill et al. 1984; Rose 1984; Remolina et al. 2012; Carnes et al. 2015; Fabian et al.
157 2018; Hoedjes et al. 2019; May et al. 2019). At the genome level, analysis of genetic
158 differentiation has revealed a significant sharing in candidate genes across the four studies
159 indicating parallel evolution (Hoedjes et al. 2019), but at the same time exposed multiple novel
160 targets of selection. For instance, three of the studies report genetic and/or transcriptomic
161 divergence in immunity genes, and it has recently been confirmed that these molecular
162 changes reflect differences in traits related to pathogen resistance (Fabian et al. 2018). Thus,
163 despite variations in the experimental designs, numerous evolutionary repeatable phenotypic
164 and genetic adaptations have been observed, but the importance of TEs in these studies has
165 remained unexplored. Therefore, our main objective was to investigate for the first time
166 whether TE abundance in the genome, and host genes related to TE regulation, had
167 undergone similar parallel changes. Using RNA-seq data from Carnes et al. (2015), we further
168 test if males and females of selected populations evolved to suppress TE transcription to
169 mitigate potentially negative effects on longevity.
170
171
172
173

174 **RESULTS**

175 **Selection for postponed reproduction affects genomic abundance of TE families**

176 To analyze if selection for longevity affected TE copy number, we used DeviaTE (Weilguny
177 and Kofler 2019) on whole genome pool-sequences of a total of 24 late-breeding, long-lived
178 selection (S) and 22 early-breeding control (C) populations from four studies (see **Table S1**
179 for details on experimental designs) (Remolina et al. 2012; Carnes et al. 2015; Fabian et al.
180 2018; Hoedjes et al. 2019). DeviaTE is an assembly-free tool that estimates genomic
181 abundance of 179 TE families by contrasting the sequencing depth of TEs and five single-
182 copy genes taking internal deletions within TEs into account (**Fig. S1** and **Fig. S2**).

183 After employing coverage and mapping quality filters (**Fig. S3**), we screened for differences in
184 abundance between control and selection regimes of 110 to 115 TE families dependent on
185 the study, using three complementary approaches that vary in stringency (see overview in
186 **Fig. S1** and Materials and Methods, summary statistics in **Table S2**). In brief, we (1) analyzed
187 studies independently, (2) fit models combining all studies using proportions of TE family
188 abundance relative to the total genomic TE content, and (3) tested if copy number differences
189 are driven by TE expansions specific to particular populations by investigating if changes in
190 TE abundance are consistent across all replicates within regime and study. For all methods,
191 we found more TE families with higher copy numbers in selected populations relative to
192 controls than vice versa, with the exception of the high protein/sugar larval diet regime in
193 Hoedjes2019 (**Table 1,** see **Supplementary Results**, for breeding regime differences within
194 each diet also see **Table S3** and **Fig. S4**). We further obtained qualitatively similar results
195 when we only considered the last 200 bp at the 3'-ends of the TE families, which are thought
196 to harbor less deletions and truncations (**Table S4**), and when we analyzed sequence
197 abundance using sums of normalized coverage values across the TE family consensus (**Table**
198 **S5**).

199 For the downstream analysis, we describe TE families varying between regimes as defined
200 by approach #1 (**Fig. 1A, Table 1**). In this approach, between 46% and 77% of all TEs had a
201 significantly larger number of genomic insertions in the selected populations relative to
202 controls after Bonferroni correction for multiple testing (from here on referred to as S>C TEs).
203 In contrast, only 12% - 31% of TEs showed the opposite pattern and had more insertions in
204 the controls (from here on referred to as C>S TEs).

205 To explore if the dynamics of TE copy number change are similar among studies, we first
206 contrasted $\log_2$ fold changes in abundance between S>C and C>S TEs. S>C TEs had a
207 significantly larger magnitude of change than C>S TEs in the two short-term evolution studies
208 of Hoedjes2019 and Remolina2012, while the opposite pattern was observed for Fabian2018
209 and no difference for Carnes2015 (**Fig. 1B**; t-tests, all $P < 0.05$ except Carnes2015: P =

210 0.466). Moreover, studies differed significantly in the size of $\log_2$ FC values in the order of
211 Carnes2015 > Fabian2018 > Hoedjes2019 = Remolina2012 (**Fig. 1C**; ANOVA with Study
212 term, Tukey HSD, P < 0.001 for all pairwise comparisons except Hoedjes2019-Remolina2012,
213 P = 0.924), seemingly scaling with the length of selection (Carnes2015: 170; Fabian2018:
214 ~146; Hoedjes2019: 58, Remolina2012: 50 generations).
215 We next asked if changes in TE abundance are driven by certain TE subclasses (Long
216 Terminal Repeat, LTR; Non-Long Terminal Repeat, non-LTR; Terminal Inverted Repeat, TIR)
217 or class (RNA, DNA) and tested S>C and C>S TEs for enrichment of these types using two-
218 sided Fisher's exact tests. We only detected a significant underrepresentation of TIRs and
219 DNA-class TEs (i.e. overrepresentation of RNA-class) in the C>S group of TEs of Carnes2015
220 and Hoedjes2019 (Carnes2015, TIRs: P = 0.044; DNA/RNA class: P = 0.024; and
221 Hoedjes2019, TIRs: P = 0.013; DNA/RNA class: P = 0.008), while there was no enrichment in
222 Fabian2018 and Remolina2015.
223 Despite many individual TEs having a higher genomic abundance in the selected populations,
224 the whole genomic TE content was not significantly different between the regimes, but varied
225 among studies (**Fig. 1D**, and **Table S6**). This was at least partly driven by the fact that although
226 C>S TEs were fewer in number than S>C TEs, they showed a significantly higher difference
227 in insertion counts in two studies (**Fig. S5,** t-test using δInsertion values; Carnes2015 P = 0.04;
228 Fabian2018 P = 0.005). The non-significant difference in overall genomic TE load could
229 therefore be a result of a large number of S>C TEs with small differences that are balanced
230 by fewer C>S TEs with large differences. We further analyzed the whole genomic abundance
231 of individual subclasses of TEs and identified a significantly higher TIR content in selected
232 populations compared to controls (**Fig. S6**, ANOVA, both Regime and Regime x Study factors,
233 P < 0.001), but this effect was strongly influenced by Carnes2015 (Tukey HSD, Regime x
234 Study factor testing for C vs S within studies, Carnes2015: P < 0.0001; other studies: P >
235 0.85). We also detected that selected populations had a larger LTR retrotransposon load than
236 controls (ANOVA, Regime factor, P = 0.026), whereas non-LTR content did not differ
237 significantly. Finally, we note that studies in general varied significantly in total TE content and
238 subclass-specific loads (ANOVA, Study factor, P < 0.0001 in all models).
239 In summary, our results demonstrate that selection for postponed reproduction leads to
240 evolutionary repeatable increases in copy number of many TE families relative to early bred
241 controls, but without affecting the overall genomic TE load.
242
243 **TE families varying in genomic abundance differ in evolutionary age and activity**
244 We next tested if differences in TE activity explain the changes in abundance between control
245 and selected populations. In *Drosophila*, most TE families are considered to be active (Guio
246 and González 2019), and it has been shown that the average population frequency of TE

insertions within a family serves as a good proxy for recent activity and age of TE invasion (Kofler et al. 2012; Kofler et al. 2015).

We first determined the exact genomic location and frequency of TE insertions using PoPoolationTE2 (Kofler et al. 2016) and calculated average population frequency across all insertion sites for each TE family. As expected, the number of detected TE insertions which could be mapped to genomic locations partially scaled with coverage (see Materials and Methods): across all populations within a study, we found 13,018 TE insertions in Hoedjes2019, 8,402 in Fabian2018, and 4,502 in Remolina2012, which is in the range recently identified in natural populations (i.e. 4,277 - 11,649 TE insertions in Lerat et al. 2019). The least number of TE insertion locations was found for Carnes2015 for which we detected an unusually small number of 567 TE insertions, likely reflecting a large number of false negatives due to low sequencing depth. For each TE family, we then averaged frequencies across all of its detected genomic positions to estimate the mean frequency at which a TE is segregating in a population (Kofler et al. 2015). Studies varied in the minimum average TE family frequency in the order of Carnes2015 > Remolina2012 > Fabian2018 > Hoedjes2019, which is likely a further effect of dissimilar sequencing depths and other experimental factors (average frequency ranges of Hoedjes2019: 0.01 - 0.9; Fabian2018: 0.02 - 1; Remolina2012: 0.04 - 0.84; Carnes2015: 0.19 – 0.9). Therefore, the TE frequencies of Carnes2015 need to be interpreted with care, considering the likely insufficient amount of data.

To get unbiased average TE frequency estimates independent of coverage fluctuations across studies, we also obtained average frequencies from a single natural South African (SA) population (Kofler et al. 2015; Kofler 2019). The SA population had a higher sequencing depth than all studies here (i.e. 381x) and thus presumably a more accurate estimate of TE frequencies. Notably, this population was not subjected to any selection or control treatment and was only maintained 8 generations in the lab before sequencing. Average genome-wide TE frequencies of control and selected populations of Fabian2018, Hoedjes2019 and Remolina2012, but not Carnes2015, were significantly correlated with the South African TE frequencies (**Fig. S7**; Spearman's ρ, Fabian2018: 0.65; Hoedjes2019: 0.61; Remolina2012: 0.58, all three $P < 0.0001$; Carnes2015: 0.1, $P = 0.403$), demonstrating that the SA population can function as an appropriate reference here.

In accordance with previous reports, we confirmed that the TE content of all populations consists of a large number of low frequency and fewer high frequency TE families (**Fig. S8**, Spearman's ρ between TE abundance and average frequency of SA population, $\rho = -0.4$ to -0.54, all $P < 0.0001$; similar when frequencies of experimental evolution studies were used, see **Fig. S9**) (Petrov et al. 2011; Kofler et al. 2015).

282 We then examined the data of the SA population and found that C>S TEs had a significantly
283 lower frequency than S>C TEs in all four studies (**Fig. 2**, t-tests between C>S and S>C
284 frequencies, $P < 0.05$ for all four studies). As there were more S>C than C>S TEs, we also
285 contrasted the average frequencies of the top 10 C>S and S>C TEs with the biggest changes
286 in genomic abundance defined by $\log_2$ FC values (**Fig. 1A**). We only detected a significantly
287 higher frequency in top 10 S>C relative to C>S TEs for Carnes2015 (t-test, $P = 0.03$), but not
288 in the other three studies. Considering the relationship between insertion age, frequency and
289 activity of TE families (Kofler et al. 2015), the lower frequency of C>S TEs suggests that they
290 are evolutionary younger and potentially more active than S>C TEs.
291
292 **Genetic drift is not driving differences in TE abundance**
293 A major challenge in experimental evolution studies is to differentiate selection from the
294 confounding genomic signals of genetic drift, which might be amplified by small effective
295 population sizes (N_e) or varying generation times spent in the lab between control and selected
296 populations. We therefore calculated genome-wide nucleotide diversity π and Watterson's θ
297 across 100kb windows as a proxy for N_e. With the exception of Fabian2018, where π was
298 equal between regimes (ANOVA, Regime factor, $P = 0.179$), we found that both estimators
299 were significantly higher in selected relative to control populations (**Table S7**; ANOVA, Regime
300 factor, all $P < 0.0001$). Even though a generally reduced N_e in controls should lead to the loss
301 of low and fixation of high frequency TEs under neutrality, we observed the opposite pattern
302 in our analysis above (**Fig. 2**).
303 To further formally test if the increased abundance of many TEs is driven by selection on
304 preexisting TE insertions or genetic drift alone, we performed population genetic simulations
305 using the correlated average TE frequencies from the natural South African population (Kofler
306 et al. 2015) as a starting point (see **Fig. S7** and results above). We simulated TE frequency
307 change in selected and control populations 5,000 times given the reported consensus
308 population sizes as N_e, generation times and number of replicates. We then asked how often
309 the same or a higher relative proportion of S>C to C>S TEs as in our observations is obtained
310 (**Table 1**). While the results from Carnes2015, Hoedjes2019, and Remolina2012 were
311 significantly different from the expected proportions, the TE abundance differences of
312 Fabian2018 could be caused by genetic drift alone (**Fig. S10**). Testing different ranges of the
313 reported population sizes and assuming that only 50% and 25% of flies in the selected
314 populations were able to breed at old age resulted in qualitatively similar results (not shown).
315 We also quantified expected proportions of TEs consistently varying in frequency across
316 simulated replicates: while there were generally more TEs consistently higher in abundance
317 in selected populations (**Table 1**, approach #3), all our simulations resulted in more TE families

318 with a consistently higher frequency in controls. The increased genomic abundance of many
319 TEs in selected populations is therefore unlikely to be solely caused by genetic drift.
320
321 **Limited evidence for selection on TE abundance and insertion frequencies**
322 Considering the deviation from neutrality, we next asked if the parallel patterns in TE
323 abundance are caused by the same or different TEs, which could indicate selection acting on
324 genomic copy number of certain TEs. Among the 103 common TE families, we identified 14
325 S>C and 2 C>S TEs shared across all four studies (**Fig. 3A**, **Table S8**). Despite the seemingly
326 large number of shared S>C TEs, only the overlap between Remolina2012 and Hoedjes2019
327 was significant (P = 0.025). Yet, we found that the most common telomeric TE *HeT-A*
328 (Casacuberta 2017) was on average more abundant in selected populations in all four studies
329 (**Fig. S1,** also identified by approach #2, see **Table S2**), suggesting that long-lived populations
330 might have evolved longer telomeres to avoid attrition, which is considered to be a key
331 conserved mechanism of aging (López-Otín et al. 2013). In contrast to S>C TEs, C>S TEs
332 showed significant overlaps across all four studies, two triple set comparisons, and between
333 Remolina2012 and Hoedjes2019 (**Fig. 3A, Table S2**). Potentially, a high genomic abundance
334 of *G-element* and *G2* found in the control populations of all studies is detrimental for longevity
335 and late-reproduction (**Fig. 3B**). However, we did not observe any significant Spearman's
336 correlation coefficients in pairwise comparisons of $\log_2$ FC values between studies except for
337 Hoedjes2019-Remolina2012 (ρ = 0.28, P = 0.004), showing that TE families generally lack
338 parallel changes in abundance.
339 Genomic TE abundance in selected populations might also be increased because selection
340 acted on a large number of segregating TE insertions resulting in frequency divergence
341 between control and selected populations. We therefore screened all identified TE insertion
342 sites for significant frequency differences between regimes in each study by performing
343 ANOVAs on arcsine square root transformed frequencies **(Table S9**). After correcting for
344 multiple testing, we detected significant frequency differences for 38 TE insertions in
345 Fabian2018 and 100 in Hoedjes2019 (**Fig. 3C** and **Fig. 3D**). At the gene level, the significant
346 TEs defined 29 and 98 genes in Fabian2018 and Hoedjes2019, respectively, and none were
347 shared between the two studies. However, in Carnes2015 and Remolina2012 insertions did
348 not show significant frequency differentiation even at a less stringent cut-off (FDR < 0.05).
349 We further tested if TE families varying in genomic abundance also differ significantly in
350 frequency between the regimes (**Table S10**, see **Table S11** for statistics on each TE family).
351 There was little evidence for parallel patterns in all studies except from Carnes2015
352 (Carnes2015: 27 TE families significant for abundance and frequency; other studies: 0 to 3).

353 Thus, although differences in TE abundance are unlikely to be driven by neutral evolution
354 alone, we only found limited evidence for parallel evolution of TE copy numbers and sparse
355 TE frequency differentiation.
356
357 **Sex, age, and selection regime affect TE expression**
358 To test whether the increased genomic abundance of TE families in selected flies is explained
359 by a higher transcriptional activity we analyzed RNA-seq data from whole flies of Carnes2015
360 (**Fig. 4** and **Table S12,** see **Table S13** for the complete statistical analysis). We first fit a model
361 with Sex, Age, and Regime to every TE family and each gene on the major chromosomal arms
362 **(Fig. S11)**. In line with sex differences in gene expression observed by Carnes et al. (2015),
363 ~92% of TE families had a significant sex term of which most had a higher expression in males
364 than females.
365 We therefore decided to test the effects of Regime, Age, and the Regime x Age interaction in
366 the sexes separately (**Fig. 4A**, **Table S12**). We detected 41 (~34% of total) and 27 TEs (~22%)
367 significantly different between regimes in males and females, respectively, with the majority
368 being upregulated in controls (**Fig. 4B**). Among these, 19 TEs significant in both sexes also
369 had the same directionality of expression change: 10 LTR-class TEs and 6 non-LTRs were
370 higher expressed in controls, whereas 3 non-LTR TEs (*TART-A*, *TART-B, and TAHRE*) were
371 upregulated in selected populations **(Table S14)**. Interestingly, *TART-A*, *TART-B,* and *TAHRE*
372 provide the enzymatic machinery for telomeric maintenance (Casacuberta 2017), again
373 suggesting that reduced telomere attrition evolved in response to selection, paralleling the
374 genome-based analysis. In general, regime affected TE expression in males and females
375 similarly, as indicated by a significant correlation of $\log_2$ fold change values between sexes
376 (**Fig. 4B**, Pearson's $r = 0.73$, $P < 0.0001$). We further asked if the magnitude of $\log_2$ fold change
377 varies between TEs more expressed in controls or selected populations, and did not find any
378 significant difference (**Fig. S12**, t-test, females: $P = 0.86$; males: $P = 0.95$).
379 Supporting the notion that TEs become derepressed during aging, the effect of age on TE
380 expression in males was general as 107 of the 108 significant TEs (i.e. ~88% of all included
381 TE families) had a higher expression in older flies. Less pronounced differences were found
382 in females where 8% of all TEs – all of which were retrotransposons – increased and 4% of
383 TEs decreased expression with age (**Fig. 4A** and **Fig. 4C**). Moreover, consistent with a recent
384 study (Chen et al. 2016), the TEs upregulated in older females had on average a significantly
385 higher $\log_2$ fold change relative to the downregulated TEs (**Fig. S12**, t-test, $P = 0.018$). We
386 further found 13 TEs with a significant age factor in both sexes (**Fig. 4C**, **Table S15**), of which
387 *copia, Burdock, R1* and *R2* are already known to increase expression with age (Li et al. 2013;
388 Chen et al. 2016).

389 No TE families showed a significant Regime x Age term in males, but the interaction was
390 significant for 28 TEs (~23% of total) in females (**Fig. 4A**). Interestingly, most of these TEs
391 were defined by a higher expression in young controls compared with selected flies of the
392 same age (see **Fig. S13** for example). Selected populations subsequently increased while
393 controls decreased expression, meeting at a similar expression level at old age. This is
394 comparable with recent studies which suggested that age-dependent changes in TE
395 expression differ between genotypes (Erwin and Blumenstiel 2019; Everett et al. 2020).
396 We next investigated if differential expression of TEs is specific or similar to the overall
397 transcriptomic changes by comparing proportions of TEs and genes up- or downregulated or
398 unchanged within levels of sex, regime, and age (**Fig. S14**). Distributions generally varied
399 significantly (χ^2 tests, P < 0.001 for all, except age factor in females: P = 0.129), demonstrating
400 that these factors have different effects on TE and gene expression.
401 To further examine if the selected populations might have evolved to maintain a young TE
402 expression profile, we compared differences between regimes to those that occurred with age
403 (**Fig. 4D and 4E**). The correlation of $\log_2$ FC values between regime and age was positive for
404 TEs in females (Pearson's correlation, females: r = 0.21, P = 0.021; males: r = -0.01, P =
405 0.875), and varied from the one for genes (1000 bootstrap replicates resampling 100 genes:
406 mean Pearson's correlation, females: r = -0.12, 95% CI: -0.13 to -0.11; males: r = 0.09, 95%
407 CI: 0.08 to 0.1). Thus, expression of TEs between selected and control populations only
408 mirrors the changes between young and old flies in females.
409 In summary, our results suggest that selected populations of Carnes2015 evolved to reduce
410 TE expression, but differences across sex and age were overall more dominant than variation
411 between regimes.
412
413 **Differences in TE abundance do not match TE expression patterns**
414 We also asked if the change in genomic TE abundance parallels the expression differences
415 between selected and control populations. Notably, as the genomic TE abundance measures
416 came from DNA pools of female flies, we did not do this comparison in males. We first
417 confirmed that TE expression scaled robustly with the number of genomic insertions in each
418 age-regime combination (Spearman's ρ = 0.72; P < 0.0001; **Fig. S15** and **Table S16**). Next,
419 we investigated if there were parallel changes in 23 TEs significantly varying between regimes
420 in expression and genomic abundance. We found that a majority of 13 TE families had non-
421 parallel changes (**Table S17**). Indeed, $\log_2$ FC expression and $\log_2$ FC insertions between
422 regimes were not significantly correlated (**Fig. 4F**, Spearman's ρ = 0.14, P = 0.149), indicating
423 that differences in TE abundance poorly predict differential expression between control and
424 selected populations. As expected, correcting RNA-seq read counts for TE copy number to

425 examine if average expression per TE insertion varies between regimes yielded qualitatively
426 similar results compared to analyzing overall TE expression (**Table S18**). However, the
427 tendency of TE families to be more highly expressed in controls was substantially larger (63
428 TEs more, 3 TEs less expressed in controls), further emphasizing that selection for late-
429 reproduction leads to a reduction in TE expression.
430

431 **Little study-wide sharing in candidate genes involved in regulation of TE activity**

432 We next hypothesized that if TE expression and transposition are predominantly detrimental
433 for lifespan and aging, as proposed by many studies, experimental evolution for longevity
434 would have likely resulted in selection on host alleles that influence TE activity. To test this,
435 we screened 96 chromatin-structure, piRNA, and transposition-associated genes known to be
436 involved in TE regulation and silencing for clear-cut genetic and expression differentiation
437 possibly driven by selection (**Table S19**). Of these, 3 to 10 genes were implicated under
438 selection across the four studies, and only *E2f1* (FBgn0011766) and *Hsp83* (FBgn0001233)
439 were shared between two datasets (**Fig. 5A**). Moreover, the four studies did not report any
440 significant enrichment of GO terms related to transposon silencing and chromatin structure.
441 Using the available RNA-seq data from whole flies of both sexes in Carnes2015 and
442 microarray data from female heads and abdomens in Remolina2012, we then asked if TE
443 regulation genes are differentially expressed (**Fig. 5B**). We found that the 42 TE regulation
444 genes significant for regime tended to be upregulated in controls in Carnes2015, but only two
445 genes differed in Remolina2012. Interestingly, similar to TE expression patterns in
446 Carnes2015 (**Fig. 4A**), TE regulation genes showed a clear tendency for upregulation with
447 age in males but to a lesser degree in females (**Fig. 5B**). Comparable patterns were detected
448 in Remolina2012, where the age effect was stronger in abdominal compared to head tissue.
449 Thus, the boosted expression of TE regulation genes at older ages appears to be common
450 and might be a response to increased TE transcription in old flies.
451 Taken together, the small number of genetically differentiated TE regulation genes, lack of
452 TE-associated GO enrichment, and overall missing parallel patterns suggest that improving
453 TE repression was either specific to studies and/or not a prime target of selection.
454

455 **DISCUSSION**

456 Are transposable elements conferring an adaptive advantage as shown for many traits
457 (Daborn et al. 2002; Magwire et al. 2011; Kuhn et al. 2014; Li et al. 2018; Rech et al. 2019) or
458 should they be purged and repressed during the evolution of longevity due to their widespread
459 negative effects on fitness (Chen et al. 2014; Krug et al. 2017; Prudencio et al. 2017; Guo et
460 al. 2018; De Cecco et al. 2019)? In this report, we attempt to answer this controversial question

461 by employing four independent data sets to present the first characterization of the genome-
462 wide TE content and expression in *D. melanogaster* populations that were experimentally
463 selected for late-life reproduction and longevity.
464
465 **Does longevity-selection lead to changes in TE abundance?**
466 Variation in TE copy number has been associated with some geographic and climatic factors
467 (Kalendar et al. 2000; Kreiner and Wright 2018; Lerat et al. 2019) in natural populations of
468 plants and *Drosophila* and was shown to change during experimental evolution in different
469 temperatures (Kofler et al. 2018). Our analysis revealed a repeatable trend showing that
470 many, but not all, TE families have an increased number of genomic insertions in late-
471 breeding, long-lived populations, which indicates that reproductive age, with some
472 dependency on developmental diet, is another factor influencing divergence in TE abundance
473 **(Fig. 1A** and **Table 1)**. Interestingly, we found a significant difference in the magnitude of TE
474 abundance change between studies that roughly scaled with the number of generations under
475 selection (**Fig. 1C**). While parallel changes in TE characteristics within populations of the same
476 selection regime have been reported by similar experiments (Graves et al. 2017; Kofler et al.
477 2018), it is striking that we observed this pattern in data created by four independent studies.
478 Despite a lot of TE families being more abundant in long-lived populations, our analysis shows
479 no significant difference in the total genomic TE content between control and selected
480 populations (**Fig. 1D**), possibly because there were a few TEs with large increases in copy
481 number in controls in contrast to many TEs with small increases in abundance in selected
482 populations (**Fig. S5**). Changes in the overall genomic TE load are therefore likely not
483 essential to evolve longevity or fecundity at old age in *Drosophila*. These findings are in
484 contrast to recent work in several killifish species, which reported that TE expansion can cause
485 an increased genome size with possible negative effects on lifespan (Cui et al. 2019).
486 However, our analyses focused exclusively on the genomic TE load and as such we cannot
487 exclude a difference in genome size between control and selected populations, which may be
488 caused by other factors such as non-repetitive InDels or repetitive DNA unrelated to TEs.
489
490 **Are TEs adaptive during the evolution of aging?**
491 The genomic content of TEs evolves through various factors, including replicative
492 transposition, selection, genetic drift, and the TE defense machineries of the host
493 (Charlesworth and Charlesworth 1983; Kofler 2019). By performing population genetic
494 simulations that consider only genetic drift, we were able to exclude that population size and
495 generations spent in the lab *per se* cause an increased abundance of TE families in selected
496 populations (**Fig. S10**). Even though it is known that the majority of TE insertions are neutral

497 to fitness (Arkhipova 2018), our findings suggest that factors other than genetic drift influenced
498 TEs.
499 From a selective point of view, increasing many TE families might be beneficial for longevity,
500 while fewer families could affect lifespan negatively. Under this scenario, selection would lead
501 to parallel increases or decreases of the same TE families across studies. However, when we
502 screened for parallel patterns in abundance change, we found only two TEs (*G-element* and
503 *G2*) that had decreased copy numbers in selected flies and were significantly shared across
504 all studies (**Fig. 3A,B)**. Both elements are *jockey*-like non-LTR TEs, of which *G2* is highly
505 enriched in centromeric regions of the genome (Chang et al. 2019). Thus, changing
506 centromeric structure by altering its TE content could be one mechanism modulating aging,
507 although experimental evidence for this is still missing. In contrast to this, we did not find any
508 significant overlap between all four studies among TEs with an increased abundance in the
509 late-breeding populations. Unless many TE families had non-repeatable effects on longevity,
510 the small amount of significant sharing suggests that abundance of most TEs is neutral.
511 Another possibility is that TE abundance is altered through selection affecting TE insertions at
512 a genome-wide scale, resulting in a large number of insertions significantly varying in
513 frequency between control and selected populations. We found only a minor fraction of TE
514 insertions in Fabian2018 and Hoedjes2019, but not in the other two studies, with significantly
515 different frequencies between the regimes that are in or close to <100 genes (**Fig. 3C,D** and
516 **Table S9**). A small fraction of TE insertions with a higher frequency in selected populations
517 were found in two of the studies. Taken together with the fact that there were very few
518 differences in frequency of TE families, we propose that standing genetic variation presented
519 by TEs plays a role in the evolution of aging, but it is unlikely to be a major driver of TE
520 abundance differentiation. However, as we identified genomic locations of TEs only using
521 PoPoolationTE2, which has been shown to have a low rate of false positives, we might miss
522 insertions that would otherwise have been found by comparable software (Kofler et al. 2016;
523 Nelson et al. 2017; Lerat et al. 2019).
524 Yet, we found that telomere maintenance, a key hallmark of aging known to be associated
525 with mortality, diseases and the rate of senescence in several organisms might be improved
526 in the late-breeding populations (Canela et al. 2007; López-Otín et al. 2013; Dantzer and
527 Fletcher 2015; Foley et al. 2018; Whittemore et al. 2019). Among the three TE families
528 constituting and maintaining *D. melanogaster* telomeres (Casacuberta 2017), *HeT-A* showed
529 parallel increases in copy number in long-lived flies although the difference was less clear in
530 two studies **(Fig. 3A** and **Fig. S1)**. Simultaneously, the few TEs transcriptionally upregulated
531 in long-lived populations of Carnes2015 were almost exclusively telomeric elements (**Fig. 4B**).
532 Despite similarities, the fundamental differences in telomeres between species make
533 generalizations difficult (Mason et al. 2008). Moreover, previous studies in *D. melanogaster*

534 and *C. elegans* failed to establish a connection between telomeres and lifespan, but telomere
535 length might affect other traits such as fecundity (Raices et al. 2005; Walter et al. 2007). Also,
536 in several species the rate of telomere shortening rather than the initial length itself was a
537 better predictor for lifespan (Whittemore et al. 2019). Another complication yet to be addressed
538 is if these patterns are caused by 'intergenerational plasticity' of telomere length, determined
539 by paternal age at reproduction as observed in several mammals including humans
540 (Eisenberg et al. 2012; Eisenberg and Kuzawa 2018). Thus, the exact impact of telomere
541 length on evolutionary fitness and aging remains to be poorly understood.
542
543 **Is TE expression detrimental for longevity?**
544 At the transcriptional level, age-dependent misregulation of TEs, thought to be resulting from
545 a gradual decline in heterochromatin maintenance, has been proposed to be harmful for
546 lifespan in *Drosophila* (Li et al. 2013; Chen et al. 2016; Wood et al. 2016; Brown and Bachtrog
547 2017; Guo et al. 2018), mice (De Cecco et al. 2019), and humans (Bogu et al. 2019). Further
548 supporting the notion that expression of many TEs is detrimental, our RNA-seq analysis
549 indicates that long-lived populations evolved to downregulate TE families, and this effect was
550 even more apparent after we corrected for genomic copy numbers (**Fig. 4A,B** and **Table S18**).
551 Considering the missing association between genomic abundance and TE transcription (**Fig.
552 4F**), this further suggests that lowering expression of TEs might be more important than
553 purging them from the genome during the evolution of longevity.
554 Overall, however, TE expression appeared to be more strongly influenced by sex and age
555 compared to selection regime. Interestingly, the trend of TEs being less expressed in late-
556 breeding populations and upregulated with age was more pronounced in male flies, which
557 further had generally higher levels of TE expression relative to females (**Fig. 4** and **Fig. S11A**).
558 These findings are consistent with recent work showing that males suffer more from TE
559 derepression during aging due to their entirely repetitive, heterochromatin-rich *Y* chromosome
560 (Brown and Bachtrog 2017). However, if the divergent TE expression patterns between sexes
561 are caused by differences in tissue compositions and whether this disparity explains sexual
562 dimorphism in lifespan is yet to be confirmed. DNA-sequencing of male flies in the four
563 experimental evolution studies would be necessary to determine if selection for postponed
564 senescence had similarly strong effects on TE copy number of the *Y* chromosome.
565
566 **Did selection lead to differentiation in genes related to regulation of TE activity?**
567 We also hypothesized that potential detrimental effects of TEs on longevity should be reflected
568 by selection on genes related to TE regulation and transposition (**Fig. 5**). Although parallel
569 genetic changes have been reported among the four studies (Fabian et al. 2018; Hoedjes et
570 al. 2019), genetically and transcriptionally differentiated TE regulation genes were generally

571 not shared between studies. Together with the missing functional enrichment associated with
572 TE regulation, we hypothesize that improvement of chromatin structure/heterochromatin
573 maintenance, piRNA-mediated silencing and modulators of transposition are not prime targets
574 of selection during the evolution of longevity. This, however, does not preclude that other
575 means of TE protection have evolved. It is becoming increasingly evident that TE expression
576 acts as a causative agent of inflammation and immune activation in mammals (Kassiotis and
577 Stoye 2016; De Cecco et al. 2019). Interestingly, Carnes2015, Fabian2018, and
578 Remolina2012 all found significant divergence in innate immunity genes, whereas Fabian et
579 al. (2018) demonstrated an improved survival upon infection and alleviated
580 immunosenescence in the long-lived populations. Rather than reducing TE copy number and
581 expression, selection might preferentially act on immunity genes to reduce TE-mediated
582 inflammation and increase tolerance to TEs with extended lifespan as a consequence. It
583 remains to be explored to what degree innate immune pathways other than the RNAi
584 machinery contribute to TE regulation in *D. melanogaster*.
585
586 **Is reproduction at old age associated with an increased TE content?**
587 Our findings suggest that neither genetic drift nor pervasive selection on TEs or genes related
588 to TE regulation are predominant drivers of the differences in TE abundance. The most
589 parsimonious explanation for our results therefore is that postponed reproduction increases
590 the chance for many TEs to be inserted into the germline and passed on to the next
591 generation. In particular, TE families of high frequency which are putatively low in
592 transpositional activity might need the prolonged chronological time offered by late-life
593 reproduction to achieve a successful genomic insertion (**Fig. 2**). Over many generations, flies
594 breeding at old age would have accumulated more TEs in the genome than populations
595 reproducing early in life. Supporting this hypothesis, it has been demonstrated that most TE
596 families had a higher rate of insertions in the ovaries of older relative to young *P-element*
597 induced dysgenic hybrids, even though at the same time fertility was restored and improved
598 with age (Khurana et al. 2011). However, if this applies to non-dysgenic fruit flies and whether
599 it can result in a larger number of TEs over multiple generations has to our knowledge not yet
600 been observed. Thus, TE accumulation in late-breeding populations is comparable to the
601 regularly observed positive correlation between parental age and number of *de novo*
602 mutations in offspring (Goldmann et al. 2019; Sasani et al. 2019). In line with this, genome-
603 wide measures of nucleotide diversity were also repeatably larger in late-breeding populations
604 across four experiments (**Table S7**). Although, we have not ruled out that greater nucleotide
605 diversity was driven by genetic drift or balancing selection as proposed by one study (Michalak
606 et al. 2017).

607 Opposing our hypothesis, two recent studies in termites (Elsner et al. 2018) and *D.*
608 *melanogaster* (Erwin and Blumenstiel 2019) suggest that the germline is protected from TE
609 invasions through increased transcription of the piRNA machinery. Indeed, our expression
610 analysis confirms that many genes associated with transcriptional and post-transcriptional TE
611 silencing tend to be upregulated with age. Despite this, many TE families had a higher copy
612 number in populations reproducing late in life. It therefore remains to be determined whether
613 this age-dependent upregulation of TE regulation genes really equates to reduced insertional
614 activity, since potential and realized TE repression might not necessarily match. The
615 observation that these genes also tended to be more expressed in controls relative to selected
616 flies in Carnes2015 further poses the question whether there is a trade-off between TE
617 silencing in the germline and lifespan, which could be another mechanism explaining the rising
618 TE abundance in the genomes of long-lived flies.
619
620 Altogether, our work presents a novel viewpoint on the poorly understood role of TEs in aging
621 and longevity that is largely, but not exclusively, neutral. However, the caveat remains that we
622 are unable to rule out that survival of selected populations would be further extended if they
623 had a reduced TE content and expression. In-depth studies tracking piRNA production in the
624 germline together with direct measures of TE transposition rates throughout life or measuring
625 longevity upon knockdown and overexpression of TEs would be crucial experiments to obtain
626 a more complete picture.
627

628 MATERIALS AND METHODS

629 Datasets

630 We utilized genomic data from four independent studies performing laboratory selection for
631 postponed reproduction on wild-derived replicate populations by only allowing flies of relatively
632 old age to contribute to subsequent generations, whereas controls reproduced early in life
633 (Remolina et al. 2012; Carnes et al. 2015; Fabian et al. 2018; Hoedjes et al. 2019) (**Table S1**).
634 The experimental designs of the studies were overall comparable, but notable differences
635 include the mode of selection, maintenance of controls, variable source populations, number
636 of replicate populations and generations at the time of sequencing. Moreover, Hoedjes2019
637 performed the selection for postponed senescence on three varying larval diets ranging from
638 low to high sugar/protein content. The genomic analysis was based on available raw fastq
639 files from whole-genome pool-sequencing of 100 to 250 females. RNA-seq data from Carnes
640 et al. (2015) consisted of raw fastq files from pools of 50 flies. The study included
641 transcriptomes of all selected and control populations, for which both sexes at two ages 3-5
642 days (young) and 26-35 days of age (old) have been sequenced in replicates. Microarray

643 expression data from Remolina et al. (2012) are derived from heads and abdomens from
644 females at the age of 1, 5, 15, 30, and 50 days of age from the three control and selected
645 populations. See methods in the publications of each study for details on experimental design
646 and **Table S1** for a summary. For simplicity, we refer to Carnes et al. (2015) as Carnes2015,
647 Fabian et al. (2018) as Fabian2018, Hoedjes et al. (2019) as Hoedjes2019, and Remolina et
648 al. (2012) as Remolina2012 throughout this report. All statistics were done in *R* using in-built
649 functions unless otherwise stated. More details on the bioinformatic pipeline are available in
650 **Table S20**.
651
652 **Genome-wide TE abundance**
653 To quantify the number of genomic insertions for each TE family in selected and control
654 populations we used DeviaTE (Weilguny and Kofler 2019) (**Table S20**). In brief, DeviaTE
655 maps raw reads to an incorporated library of 179 TE family consensus sequences (Sackton
656 et al. 2009; Bergman et al. 2018) and normalizes the obtained coverage values by the average
657 depth of the same five single-copy genes (**Fig. S2**). The distribution of normalized values
658 reflects fluctuations in insertion number estimates within a TE family, where averaging over
659 all consensus positions of a TE family gives the mean abundance per haploid genome (see
660 Weilguny and Kofler 2019 for details). We restricted our downstream analysis to TE families
661 that had a study-average of >=0.5 insertions for at least 80% of the consensus positions within
662 a TE family sequence (**Fig. S3**). Thus, we excluded all families with very low abundance and
663 potentially wrongly mapped reads, and TE families of which only small fractions of the whole
664 consensus sequence were covered.
665 We then investigated if TE families vary in genomic abundance between control and selected
666 populations using three different approaches (see **Fig. S1** for a comprehensive description).
667 In our least conservative approach #1, we analyzed studies by fitting *Regime* (control,
668 selected) and *Population[Regime]* (replicate populations nested within regime) to normalized
669 coverage values of consensus sequence positions within a TE family. For Hoedjes2019, we
670 used a different model and included *Regime*, *Diet* (low, medium, high protein/sugar larval diet
671 regime), and the *Regime x Diet* interaction. To correct for multiple testing, we applied a
672 Bonferroni cut-off at α = 0.01. We further used *SuperExactTest* (Wang et al. 2015) to analyze
673 if the overlap of TEs with a significantly higher genomic abundance in selected ("S>C") or
674 control populations ("C>S") between postponed senescence studies is expected by chance.
675 The normalized coverage values were averaged to obtain a single insertion estimate per TE
676 family and population, and these values used for all the remaining analyses.
677 For approach #2, we arcsine square root transformed proportions of TE family copy number
678 relative to the total genomic TE content within a population and analyzed all studies together
679 rather than independently by fitting *Study* (four levels: Carnes2015, Fabian2018,

680 Hoedjes2019, Remolina2012), *Regime* and the *Study x Regime* interaction as factors. TE
681 families with an FDR < 0.05 were considered significant.
682 Finally, our approach #3 is the most conservative as we only considered TE families that
683 showed a consistent increase or decrease in copy number (i.e. average of insertion estimates
684 across all consensus positions) within all selected relative to all control populations in each
685 study and within diet regimes of Hoedjes2019.
686 To analyze differences in the total genomic and subclass-specific (LTR, non-LTR, TIR) TE
687 content, we summed up all TE insertion estimates within a population and fit models with
688 *Study*, *Regime* and the *Study x Regime* interaction.
689
690 **Genomic TE locations and activity/age of TE families**
691 We first masked the *D. melanogaster* reference (v.6.27) for TEs present in the DeviaTE library
692 using RepeatMasker (Smit et al. 1996) (**Table S20**). We then trimmed reads with cutadapt
693 (Martin 2011) and mapped them using bwa bwasw (Li and Durbin 2009). PoPoolationTE2 was
694 then employed to obtain the exact genomic positions and population frequency of TE
695 insertions on chromosomes *X*, *2, 3,* and *4* of each study using the joint analysis mode, which
696 finds insertions by combining all samples rather to considering them separately (Kofler et al.
697 2016). Importantly, while TE abundance is quantified by the total number of reads mapping to
698 a TE relative to single-copy genes (Weilguny and Kofler 2019), identifying the exact genomic
699 location of insertions requires mates of a read-pair to map discordantly to the reference
700 genome and TE sequence, and strongly depends on the sequencing depth and number of
701 populations (Cridland et al. 2013; Kofler et al. 2016; Lerat et al. 2019). For each TE family, we
702 calculated the average population frequency across all of its detected genomic locations within
703 a population as a proxy for active or recent transposition events and evolutionary age (Kofler
704 et al. 2015). We used Spearman's correlation analysis to compare average frequency values
705 of each study to average frequencies from a natural South African (SA) population sequenced
706 to a high genomic coverage (Kofler et al. 2015), and to correlate TE abundance with average
707 frequency. We employed t-tests to analyze if average population frequency from the SA
708 population varies between TE families more abundant in selected or control populations, and
709 also performed this analysis using only the top10 TEs with the largest $\log_2$ FC values of
710 abundance change.
711
712 **Genome-wide nucleotide diversity and genetic drift simulations**
713 We mapped trimmed paired-end reads against the repeat-masked reference genome, the TE
714 library from DeviaTE (Weilguny and Kofler 2019), *Wolbachia pipientis* (NC_002978.6), and
715 two common gut bacteria *Acetobacter pasteurianus* (AP011121.1), and *Lactobacillus brevis*
716 (CP000416.1) using bwa *mem* (Li and Durbin 2009), and removed duplicates using

717 PicardTools (**Table S20**). We then filtered and created pileup files using samtools *mpileup* (Li
718 et al. 2009). To calculate nucleotide diversity π and Watterson's θ across non-overlapping
719 100kb windows we used Popoolation (Kofler et al. 2011) and then fitted ANOVA models
720 including the factors *Chromosome* (*X, 2L, 2R, 3L, 3R, 4*), *Diet, Regime*, and the *Diet x Regime*
721 interaction for Hoedjes2019, and *Population[Regime]*, *Chromosome*, and *Regime* for all other
722 studies. Average coverage across major chromosomal arms was 162x, 101x, 41x, and 23x
723 for Fabian2018, Hoedjes2019, Remolina2012, and Carnes2015, respectively. We detected
724 reads mapping to the genome of the intracellular bacterium *Wolbachia* in all populations.
725 To test if TE family abundance differences can be caused by genetic drift alone, we compared
726 proportions of S>C and C>S TEs from 5,000 simulations of TE frequency change to observed
727 proportions from approach #1 and #3. See Supplementary Methods for more details.
728
729 **TE frequency differences**
730 To identify genomic TE insertion sites putatively involved in lifespan and aging, we analyzed
731 differences in arcsine square root transformed insertion frequencies between selected and
732 control populations fitting models with *Regime* for Carnes2015, Fabian2018 and
733 Remolina2012, and with factors *Diet*, *Regime*, and *Diet x Regime* for Hoedjes2019. Bonferroni
734 correction at $\alpha = 0.05$ was used to correct for multiple testing. Functional annotations were
735 supplemented using SnpEff (v.4.0e, Cingolani et al. 2012) considering TE insertions within
736 1000 bp of the 5' and 3' UTR as upstream or downstream of a gene.
737 We further analyzed if each TE family varies in frequency between regimes by fitting the
738 factors of *Diet*, *Regime, and Diet x Regime* for Hoedjes2019, or *Regime* and
739 *Population[Regime]* for all other studies on arcsine square root transformed insertion site
740 frequencies. FDR values were obtained by using "p.adjust" in *R* and TE families considered
741 significant at FDR < 0.05.
742
743 **RNA-seq analysis**
744 RNA-seq data from Carnes et al. (2015) consisted of two replicates of young and old males
745 and females from all control and selected populations (**Table S1**). Raw reads were filtered
746 using cutadapt (Martin 2011) and mapped to the repeat-masked reference genome, the TE
747 library from DeviaTE, *Wolbachia pipientis*, *Acetobacter pasteurianus*, and *Lactobacillus brevis*
748 (see above) using STAR (Dobin et al. 2013) (**Table S20**). Read counts were obtained using
749 featureCounts (Liao et al. 2013). We next pre-filtered read count data by excluding all genes
750 and TEs that did not have a sum of 400 counts across all 80 samples (i.e. on average 5 counts
751 per sample). Five TE families that are not known to occur in *D. melanogaster* passed this filter
752 and were excluded. For simplicity, the analysis was performed on average read counts from

753 two replicates, as all replicates were highly significantly correlated (Pearson's r ranging from
754 0.95 to 1, significant after Bonferroni correction). To analyze differential expression, we fit
755 models using read counts of genes and TEs with DESEq2 in *R* (Love et al. 2014). First, a
756 model testing the main effects of *Regime* (selected vs control), *Sex* (male vs female), and *Age*
757 (young vs old) was fit. As the sex term was significant for most TEs, we decided to analyze
758 males and females separately and fitted models with *Regime* and *Age* to analyze the main
759 effects. To examine the interaction, we also fitted models including *Regime x Age*. We
760 obtained $\log_2$ fold change values for each factor and the library-size normalized read counts
761 from DEseq2 for further analysis. To investigate average expression per TE insertion, we
762 divided read counts of TEs from females by the number of genomic insertions observed in
763 each population, assuming that genes and 13 TEs that did not pass our filters in the genomic
764 analysis have a single copy in the genome.
765

766 **Evolution of TE regulation genes**

767 The list of genes involved in TE regulation consisted of piRNA pathway genes also analyzed
768 in Erwin and Blumenstiel 2019 and Elsner et al. 2018, and genes involved in heterochromatic
769 and chromatin structure from Lee and Karpen 2017. We further added 7 genes involved in
770 these functions, and genes annotated to "regulation of transposition" (GO:0010528) and
771 "transposition" (GO:0032196) according to FlyBase so that we ended up with a total of 96
772 genes (**Table S19**). We then screened the published genomic candidate gene lists from
773 Carnes2015, Fabian2018, Hoedjes2019 and Remolina2012 for these genes. We also
774 compared TE regulation genes to differentially expressed genes from the RNA-seq analysis
775 of Carnes2015 (see above). We further obtained normalized microarray expression data from
776 Remolina2012 of female flies at 1, 5, 15, 30, and 50 days of age (**Table S1**). Notably, the
777 expression data were created from flies at 40 generations of selection compared to 50
778 generations in the genomic analysis. We fit a mixed effects model similar to the one used in
779 their original publication with *Age*, *Regime*, and *Age x Regime* as fixed and *replication within*
780 *population-age combination* as random effect. The two available tissues (heads and
781 abdomens) were analyzed separately. A gene was considered to be differentially expressed
782 if it had an FDR < 0.05 unless otherwise stated.
783

784 **DATA ACCESSIBILITY**

785 Accession numbers to the raw genomic and transcriptomic data can be found in **Table S1** and
786 in the original studies (Remolina et al. 2012; Carnes et al. 2015; Fabian et al. 2018; Hoedjes
787 et al. 2019). RNA-seq data were obtained directly from the authors (Wen Huang and Trudy
788 Mackay, active download links in supplementary code on GitHub). Scripts to all analyses and

raw output files are available at: https://github.com/FabianDK/LongeviTE. Additional raw output and edited files used to analyze TE abundance and nucleotide diversity, results for the microarray analysis, and boxplots showing the number of insertions for significant TE families are available on Dryad (DOI: 10.5061/dryad.s7h44j13r).

ACKNOWLEDGEMENTS

We are grateful to Alexis Braun, three anonymous reviewers and the associate editors for comments on the manuscript. We also thank Robert Kofler, Andrea Betancourt, Frank Jiggins and Lukas Weilguny for helpful discussions. This work was supported by the Wellcome Trust (WT098565/Z/12/Z to J.M.T. and L.P.), EMBL (H.M.D. and J.M.T.), and Comisión Nacional de Investigación Científica y Tecnológica – Government of Chile (CONICYT scholarship to M.F.). We are further grateful for financial support from the Society for Molecular Biology & Evolution enabling us to present this work at the annual meeting (SMBE 2019, Carer travel award & registration award to D.F.).

REFERENCES

Arkhipova IR. 2018. Neutral theory, transposable elements, and eukaryotic genome evolution. *Mol Biol Evol*. 35:1332–1337.

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

Bergman CM, Han S, Benos T, Bayraktaroglu L, Ashburner M, de Grey A, Chillemi J, Reese M, Lewis S, Guochun L, et al. 2018. Drosophila transposable element consensus sequences - v10.1. https://github.com/cbergman/transposons

Bogu GK, Reverter F, Marti-Renom MA, Snyder MP, Guigó R. 2019. Atlas of transcriptionally active transposable elements in human adult tissues. *unpublished data*. www.biorxiv.org/content/10.1101/714212v1

Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran J V., Kazazian HH. 2003. Hot L1s account for the bulk of retrotransposition in the human population. *Proc Natl Acad Sci U S A*. 100:5280–5285.

Brown EJ, Bachtrog D. 2017. The Y chromosome contributes to sex-specific aging in Drosophila. *unpublished data*. www.biorxiv.org/content/10.1101/156042v1

Brunet TDP, Doolittle WF. 2015. Multilevel selection theory and the evolutionary functions of transposable elements. *Genome Biol Evol*. 7:2445–2457.

Canela A, Vera E, Klatt P, Blasco MA. 2007. High-throughput telomere length quantification by FISH and its application to human population studies. *Proc Natl Acad Sci U S A*. 104:5300–5305.

825 Carnes MU, Campbell T, Huang W, Butler DG, Carbone MA, Duncan LH, Harbajan S V.,
826 King EM, Peterson KR, Weitzel A, et al. 2015. The genomic basis of postponed
827 senescence in Drosophila melanogaster. *PLoS One*. 10:e0138569.
828 Carr M, Bensasson D, Bergman CM. 2012. Evolutionary genomics of transposable elements
829 in Saccharomyces cerevisiae. *PLoS One*. 7:e50978.
830 Casacuberta E. 2017. Drosophila: Retrotransposons making up telomeres. *Viruses*. 9:E192.
831 De Cecco M, Criscione SW, Peckham EJ, Hillenmeyer S, Hamm EA, Manivannan J,
832 Peterson AL, Kreiling JA, Neretti N, Sedivy JM. 2013. Genomes of replicatively
833 senescent cells undergo global epigenetic changes leading to gene silencing and
834 activation of transposable elements. *Aging Cell*. 12:247–256.
835 De Cecco M, Ito T, Petrashen AP, Elias AE, Skvir NJ, Criscione SW, Caligiana A, Brocculi
836 G, Adney EM, Boeke JD, et al. 2019. L1 drives IFN in senescent cells and promotes
837 age-associated inflammation. *Nature*. 566:73–78.
838 Chang CH, Chavan A, Palladino J, Wei X, Martins NMC, Santinello B, Chen CC, Erceg J,
839 Beliveau BJ, Wu CT, et al. 2019. Islands of retroelements are major components of
840 Drosophila centromeres. *PLoS Biol*. 17:e3000241.
841 Charlesworth B, Charlesworth D. 1983. The population dynamics of transposable elements.
842 *Genet Res*. 42:1–27.
843 Charlesworth B, Lapid A, Canada D. 1992. The distribution of transposable elements within
844 and between chromosomes in a population of Drosophila melanogaster. II. Inferences
845 on the nature of selection against elements. *Genet Res*. 60:115–130.
846 Chen H, Zheng X, Xiao D, Zheng Y. 2016. Age-associated de-repression of
847 retrotransposons in the Drosophila fat body, its potential cause and consequence.
848 *Aging Cell*. 15:542–552.
849 Chen H, Zheng X, Zheng Y. 2014. Age-associated loss of lamin-B leads to systemic
850 inflammation and gut hyperplasia. *Cell*. 159:829–843.
851 Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM.
852 2012. A program for annotating and predicting the effects of single nucleotide
853 polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118;
854 iso-2; iso-3. *Fly (Austin)*. 6:80–92.
855 Cridland JM, Macdonald SJ, Long AD, Thornton KR. 2013. Abundance and distribution of
856 transposable elements in two drosophila QTL mapping resources. *Mol Biol Evol*.
857 30:2311–2327.
858 Cui R, Medeiros T, Willemsen D, Leonardo L Iasi, Collier G, Graef M, Reichard M,
859 Valenzano D. 2019. Relaxed selection limits lifespan by increasing mutation load. *Cell*.
860 178:385-399.e20.
861 Daborn PJ, Yen JL, Bogwitz MR, Le Goff G, Feil E, Jeffers S, Tijet N, Perry T, Heckel D,

862 Batterham P, et al. 2002. A single P450 allele associated with insecticide resistance in
863 Drosophila. *Science*. 297:2253–2256.
864 Dantzer B, Fletcher QE. 2015. Telomeres shorten more slowly in slow-aging wild animals
865 than in fast-aging ones. *Exp Gerontol*. 71:38–47.
866 Deniz Ö, Frost JM, Branco MR. 2019. Regulation of transposable elements by DNA
867 modifications. *Nat Rev Genet*. 20:417–431.
868 Dennis S, Sheth U, Feldman JL, English KA, Priess JR. 2012. C. elegans germ cells show
869 temperature and age-dependent expression of Cer1, a Gypsy/Ty3-related
870 retrotransposon. *PLoS Pathog*. 8:e1002591.
871 Dimitri P, Junakovic N. 1999. Revising the selfish DNA hypothesis: New evidence on
872 accumulation of transposable elements in heterochromatin. *Trends Genet*. 15:123–124.
873 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M,
874 Gingeras TR. 2013. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*. 29:15–
875 21.
876 Eisenberg DTA, Hayes MG, Kuzawa CW. 2012. Delayed paternal age of reproduction in
877 humans is associated with longer telomeres across two generations of descendants.
878 *Proc Natl Acad Sci U S A*. 109:10251–10256.
879 Eisenberg DTA, Kuzawa CW. 2018. The paternal age at conception effect on offspring
880 telomere length: Mechanistic, comparative and adaptive perspectives. *Philos Trans R*
881 *Soc B Biol Sci*. 373:pii: 20160442.
882 Elsner D, Meusemann K, Korb J. 2018. Longevity and transposon defense, the case of
883 termite reproductives. *Proc Natl Acad Sci U S A*. 115:5504–5509.
884 Erwin AA, Blumenstiel JP. 2019. Aging in the Drosophila ovary: contrasting changes in the
885 expression of the piRNA machinery and mitochondria but no global release of
886 transposable elements. *BMC Genomics*. 20:305.
887 Everett LJ, Huang W, Zhou S, Carbone MA, Lyman RF, Arya GH, Geisz MS, Ma J,
888 Morgante F, Armour G, et al. 2020. Gene expression networks in the Drosophila
889 genetic reference panel. *Genome Res*. 30:485–496.
890 Fabian D, Flatt T. 2011. The evolution of aging. *Nat Educ Knowl*. 2:9.
891 Fabian DK, Garschall K, Klepsatel P, Santos-Matos G, Sucena É, Kapun M, Lemaitre B,
892 Schlötterer C, Arking R, Flatt T. 2018. Evolution of longevity improves immunity in
893 Drosophila. *Evol Lett*. 2:567–579.
894 Flatt T, Heyland A. 2011. Mechanisms of life history evolution: The genetics and physiology
895 of life history traits and trade-offs. *Oxford Univ Press Oxford*.
896 Foley NM, Hughes GM, Huang Z, Clarke M, Jebb D, Whelan C V., Petit EJ, Touzalin F,
897 Farcy O, Jones G, et al. 2018. Growing old, yet staying young: The role of telomeres in
898 bats' exceptional longevity. *Sci Adv*. 4:eaao0926.

899 Fontana L, Partridge L, Longo VD. 2010. Extending healthy life span-from yeast to humans.
900 *Science*. 328:321–326.
901 Gems D, Partridge L. 2013. Genetics of longevity in model organisms: Debates and
902 paradigm shifts. *Annu Rev Physiol*. 75:621–644.
903 Goldmann J, Veltman J, Gilissen C. 2019. De novo mutations reflect development and aging
904 of the human germline. *Trends Genet*. 35:828–839.
905 Gorbunova V, Boeke JD, Helfand SL, Sedivy JM. 2014. Sleeping dogs of the genome.
906 *Science*. 346:1187–118.
907 Graves JL, Hertweck KL, Phillips MA, Han M V., Cabral LG, Barter TT, Greer LF, Burke MK,
908 Mueller LD, Rose MR, et al. 2017. Genomics of parallel experimental evolution in
909 Drosophila. *Mol Biol Evol*. 34:831–842.
910 Guio L, González J. 2019. New insights on the evolution of genome content: Population
911 dynamics of transposable elements in flies and humans. In: Anisimova M. (eds)
912 Evolutionary Genomics. Methods in Molecular Biology. p. 505–530.
913 Guo C, Jeong HH, Hsieh YC, Klein HU, Bennett DA, De Jager PL, Liu Z, Shulman JM. 2018.
914 Tau activates transposable elements in Alzheimer's disease. *Cell Rep*. 23:2874–2880.
915 Hancks DC, Kazazian HH. 2012. Active human retrotransposons: Variation and disease.
916 *Curr Opin Genet Dev*. 22:191–203.
917 Hoedjes KM, van den Heuvel J, Kapun M, Keller L, Flatt T, Zwaan BJ. 2019. Distinct
918 genomic signals of lifespan and life history evolution in response to postponed
919 reproduction and larval diet in Drosophila. *Evol Lett*. 3:598–609.
920 Jurka J, Bao W, Kojima KK. 2011. Families of transposable elements, population structure
921 and the origin of species. *Biol Direct*. 6:44.
922 Kalendar R, Tanskanen J, Immonen S, Nevo E, Schulman AH. 2000. Genome evolution of
923 wild barley (Hordeum spontaneum) by BARE-1 retrotransposon dynamics in response
924 to sharp microclimatic divergence. *Proc Natl Acad Sci U S A*. 97:6603–6607.
925 Kaneko H, Dridi S, Tarallo V, Gelfand BD, Fowler BJ, Cho WG, Kleinman ME, Ponicsan SL,
926 Hauswirth WW, Chiodo VA, et al. 2011. DICER1 deficit induces Alu RNA toxicity in age-
927 related macular degeneration. *Nature*. 471:325–330.
928 Kassiotis G, Stoye JP. 2016. Immune responses to endogenous retroelements: Taking the
929 bad with the good. *Nat Rev Immunol*. 16:207–219.
930 Khurana JS, Wang J, Xu J, Koppetsch BS, Thomson TC, Nowosielska A, Li C, Zamore PD,
931 Weng Z, Theurkauf WE. 2011. Adaptation to P element transposon invasion in
932 Drosophila melanogaster. *Cell*. 147:1551–1563.
933 Kirkwood TBL. 1989. DNA, mutations and aging. *Mutat Res*. 219:1–7.
934 Kofler R. 2019. Dynamics of Transposable Element Invasions with piRNA Clusters. *Mol Biol*
935 *Evol*. 36:1457–1472.

936 Kofler R, Betancourt AJ, Schlötterer C. 2012. Sequencing of pooled DNA samples (Pool-
937 Seq) uncovers complex dynamics of transposable element insertions in Drosophila
938 melanogaster. *PLoS Genet*. 8:e1002487.
939 Kofler R, Gómez-Sánchez D, Schlötterer C. 2016. PoPoolationTE2: Comparative population
940 genomics of transposable elements using pool-seq. *Mol Biol Evol*. 33:2759–2764.
941 Kofler R, Nolte V, Schlötterer C. 2015. Tempo and mode of transposable element activity in
942 Drosophila. *PLoS Genet*. 11:e1005406.
943 Kofler R, Orozco-terWengel P, de Maio N, Pandey RV, Nolte V, Futschik A, Kosiol C,
944 Schlötterer C. 2011. Popoolation: a toolbox for population genetic analysis of next
945 generation sequencing data from pooled individuals. *PLoS One*. 6:e15925.
946 Kofler R, Senti KA, Nolte V, Tobler R, Schlötterer C. 2018. Molecular dissection of a natural
947 transposable element invasion. *Genome Res*. 28:824–835.
948 de Koning APJ, Gu W, Castoe TA, Batzer MA, Pollock DD. 2011. Repetitive elements may
949 comprise over two-thirds of the human genome. *PLoS Genet*. 7:e1002384.
950 Kreiner JM, Wright SI. 2018. A less selfish view of genome size evolution in maize. *PLoS*
951 *Genet*. 14:e1007249.
952 Krug L, Chatterjee N, Borges-Monroy R, Hearn S, Liao WW, Morrill K, Prazak L, Rozhkov N,
953 Theodorou D, Hammell M, et al. 2017. Retrotransposon activation contributes to
954 neurodegeneration in a Drosophila TDP-43 model of ALS. *PLoS Genet*. 13:e1006635.
955 Kuhn A, Ong YM, Cheng C-Y, Wong TY, Quake SR, Burkholder WF. 2014. Linkage
956 disequilibrium and signatures of positive selection around LINE-1 retrotransposons in
957 the human genome. *Proc Natl Acad Sci U S A*. 111:8131–8136.
958 Lee YCG, Karpen GH. 2017. Pervasive epigenetic effects of Drosophila euchromatic
959 transposable elements impact their evolution. *Elife*. 6:pii: e25762.
960 Lerat E, Goubert C, Guirao-Rico S, Merenciano M, Dufour AB, Vieira C, González J. 2019.
961 Population-specific dynamics and selection patterns of transposable element insertions
962 in European natural populations. *Mol Ecol*. 28:1506–1522.
963 Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler
964 transform. *Bioinformatics*. 25:1754–1760.
965 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin
966 R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 25:2078–
967 2079.
968 Li W, Prazak L, Chatterjee N, Grüninger S, Krug L, Theodorou D, Dubnau J. 2013.
969 Activation of transposable elements during aging and neuronal decline in Drosophila.
970 *Nat Neurosci*. 16:529–531.
971 Li ZW, Hou XH, Chen JF, Xu YC, Wu Q, Gonzalez J, Guo YL. 2018. Transposable elements
972 contribute to the adaptation of Arabidopsis thaliana. *Genome Biol Evol*. 10:2140–2150.

973 Liao Y, Smyth GK, Shi W. 2013. FeatureCounts: an efficient general-purpose read
974 summarization program. *Bioinformatics*. 30:923–930.
975 Lippman Z, Martienssen R. 2004. The role of RNA interference in heterochromatic silencing.
976 *Nature*. 431:364–370.
977 López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. 2013. The hallmarks of
978 aging. *Cell*. 153:1194–1217.
979 Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for
980 RNA-seq data with DESeq2. *Genome Biol*. 15:550.
981 Luckinbill LR, Arking R, Clare MJ. 1984. Selection for delayed senescence in Drosophila
982 melanogaster. *Evolution*. 38:996–1003.
983 Magwire MM, Bayer F, Webster CL, Cao C, Jiggins FM. 2011. Successive increases in the
984 resistance of Drosophila to viral infection through a transposon insertion followed by a
985 duplication. *PLoS Genet*. 7:e1002337.
986 Martienssen R, Moazed D. 2015. RNAi and heterochromatin assembly. *Cold Spring Harb*
987 *Perspect Biol*. 7:a019323.
988 Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing
989 reads. *EMBnet.journal*. 17:10–12.
990 Mason JM, Frydrychova RC, Biessmann H. 2008. Drosophila telomeres: An exception
991 providing new insights. *BioEssays*. 30:25–37.
992 Maxwell PH, Burhans WC, Curcio MJ. 2011. Retrotransposition is associated with genome
993 instability during chronological aging. *Proc Natl Acad Sci U S A*. 108:20376–20381.
994 May CM, van den Heuvel J, Doroszuk A, Hoedjes KM, Flatt T, Zwaan BJ. 2019. Adaptation
995 to developmental diet influences the response to selection on age at reproduction in the
996 fruit fly. *J Evol Biol*. 32:425–437.
997 McClintock B. 1950. The origin and behavior of mutable loci in maize. *Proc Natl Acad Sci U*
998 *S A*. 36:344–355.
999 Michalak P, Kang L, Sarup PM, Schou MF, Loeschcke V. 2017. Nucleotide diversity inflation
1000 as a genome-wide response to experimental lifespan extension in Drosophila
1001 melanogaster. *BMC Genomics*. 18:84.
1002 Montgomery E, Charlesworth B, Langley CH. 1987. A test for the role of natural selection in
1003 the stabilization of transposable element copy number in a population of Drosophila
1004 melanogaster. *Genet Res*. 49:31–41.
1005 Mori MA, Raghavan P, Thomou T, Boucher J, Robida-Stubbs S, MacOtela Y, Russell SJ,
1006 Kirkland JL, Blackwell TK, Kahn CR. 2012. Role of microRNA processing in adipose
1007 tissue in stress defense and longevity. *Cell Metab*. 16:336–3347.
1008 Nelson MG, Linheiro RS, Bergman CM. 2017. McClintock: An integrated pipeline for
1009 detecting transposable element insertions in whole-genome shotgun sequencing data.

1010 *G3 Genes, Genomes, Genet*. 7:2763–2778.
1011 Pan H, Finkel T. 2017. Key proteins and pathways that regulate lifespan. *J Biol Chem*.
1012 292:6452–6460.
1013 Petrov DA, Fiston-Lavier AS, Lipatov M, Lenkov K, González J. 2011. Population genomics
1014 of transposable elements in Drosophila melanogaster. *Mol Biol Evol*. 28:1633–1644.
1015 Piper MDW, Selman C, McElwee JJ, Partridge L. 2008. Separating cause from effect: How
1016 does insulin/IGF signalling control lifespan in worms, flies and mice? *J Intern Med*.
1017 263:179–191.
1018 Prudencio M, Gonzales PK, Cook CN, Gendron TF, Daughrity LM, Song Y, Ebbert MTW,
1019 van Blitterswijk M, Zhang YJ, Jansen-West K, et al. 2017. Repetitive element
1020 transcripts are elevated in the brain of C9orf72 ALS/FTLD patients. *Hum Mol Genet*.
1021 26:3421–3431.
1022 Quesneville H, Bergman CM, Andrieu O, Autard D, Nouaud D, Ashburner M, Anxolabehere
1023 D. 2005. Combined evidence annotation of transposable elements in genome
1024 sequences. *PLoS Comput Biol*. 1:166–175.
1025 Raices M, Maruyama H, Dillin A, Kariseder J. 2005. Uncoupling of longevity and telomere
1026 length in C. elegans. *PLoS Genet*. 1:e30.
1027 Rech GE, Bogaerts-Márquez M, Barrón MG, Merenciano M, Villanueva-Cañas JL, Horváth
1028 V, Fiston-Lavier AS, Luyten I, Venkataram S, Quesneville H, et al. 2019. Stress
1029 response, behavior, and development are shaped by transposable element-induced
1030 mutations in Drosophila. *PLoS Genet*. 15:e1007900.
1031 Remolina S, Chang P, Leips J. 2012. Genomic basis of aging and life history evolution in
1032 Drosophila melanogaster. *Evolution*. 66:3390–3403.
1033 Rose M. 1984. Laboratory evolution of postponed senescence in Drosophila melanogaster.
1034 *Evolution*. 38:1004–1010.
1035 Sackton TB, Kulathinal RJ, Bergman CM, Quinlan AR, Dopman EB, Carneiro M, Marth GT,
1036 Hartl DL, Clark AG. 2009. Population Genomic Inferences from Sparse High-
1037 Throughput Sequencing of Two Populations of Drosophila melanogaster. *Genome Biol*
1038 *Evol*. 1:449–465.
1039 Sasani TA, Pedersen BS, Gao Z, Baird L, Przeworski M, Jorde LB, Quinlan AR. 2019.
1040 Large, three-generation human families reveal post-zygotic mosaicism and variability in
1041 germline mutation accumulation. *Elife*. 8:pii: e46922.
1042 Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L,
1043 Graves TA, et al. 2009. The B73 maize genome: Complexity, diversity, and dynamics.
1044 *Science*. 326:1112–1115.
1045 Sedivy JM, Kreiling JA, Neretti N, Cecco M De, Criscione SW, Hofmann JW, Zhao X, Ito T,
1046 Peterson AL. 2013. Death by transposition - the enemy within? *BioEssays*. 35:1035–

1043.
Smit A, Hubley R, Green P. 1996. RepeatMasker Open-3.0. www.repeatmasker.org
Solyom S, Ewing AD, Rahrmann EP, Doucet T, Nelson HH, Burns MB, Harris RS, Sigmon DF, Casella A, Erlanger B, et al. 2012. Extensive somatic L1 retrotransposition in colorectal tumors. *Genome Res*. 22:2328–2338.
Sturm Á, Ivics Z, Vellai T. 2015. The mechanism of ageing: Primary role of transposable elements in genome disintegration. *Cell Mol Life Sci*. 72:1839–1847.
Volkman HE, Stetson DB. 2014. The enemy within: Endogenous retroelements and autoimmune disease. *Nat Immunol*. 15:415–422.
Walter MF, Biessmann MR, Benitez C, Török T, Mason JM, Biessmann H. 2007. Effects of telomere length in Drosophila melanogaster on life span, fecundity, and fertility. *Chromosoma*. 116:41–51.
Wang M, Zhao Y, Zhang B. 2015. Efficient test and visualization of multi-set intersections. *Sci Rep*. 5:16923.
Weilguny L, Kofler R. 2019. DeviaTE: Assembly-free analysis and visualization of mobile genetic element composition. *Mol Ecol Resour*. 19:1346–1354.
Whittemore K, Vera E, Martínez-Nevado E, Sanpera C, Blasco MA. 2019. Telomere shortening rate predicts species life span. *Proc Natl Acad Sci U S A*. 116:15122–15127.
Wood JG, Helfand SL. 2013. Chromatin structure and transposable elements in organismal aging. *Front Genet*. 4:274.
Wood JG, Jones BC, Jiang N, Chang C, Hosier S, Wickremesinghe P, Garcia M, Hartnett DA, Burhenn L, Neretti N, et al. 2016. Chromatin-modifying genetic interventions suppress age-associated transposable element activation and extend life span in Drosophila. *Proc Natl Acad Sci U S A*. 113:11277–11282.

1084 **FIGURE LEGENDS**

1085 **Figure 1. Dynamics of TE copy number change between breeding regimes.** (A) Log_2 fold
1086 change in average genomic insertions of the late-breeding selected populations ("S") relative
1087 to early-breeding controls ("C"). The dashed line indicates no difference between regimes. >0
1088 denote TE families with a larger abundance in selected populations ("S>C"), while <0 TEs with
1089 more insertions in controls ("C>S"). Number of TE families in these two categories are given
1090 in the center at the top and bottom of each plot. TE subclasses are given in different colors.
1091 Selected flies had more genomic insertions than controls for most TE families (also see **Table**
1092 **1**). (B) Difference in the magnitude of absolute log_2 fold change between C>S and S>C TE
1093 groups. Significant difference between TE groups was determined using t-tests for each study.
1094 (C) Magnitude of absolute log_2 fold change between studies, analyzed using ANOVA with
1095 Study as single term ($F_{3,358}$ = 106.5, P < 2e-16) and pairwise Tukey post-hoc tests. * P < 0.05;
1096 ** P < 0.01; *** P < 0.001; ns, not significant. (D) Total number of genomic TE insertions. We
1097 used ANOVA to test the effects of Study, Regime and the Study x Regime interaction. See
1098 **Table S6** for a summary of the statistical analysis.
1099

1100 **Figure 2. Differences in average TE frequency.** Average TE frequency from the South
1101 African population separated into C>S (blue) and S>C TEs (red) are shown on the Y-axis. We
1102 investigated differences considering all C>S and S>C TEs ("All") or only the top 10 TEs with
1103 the biggest differences in log_2 FC of insertions ("Top 10"). t-tests were used to assess
1104 statistical significance. ns, not significant; * P < 0.05; ** P < 0.01; *** P < 0.001.
1105

1106 **Figure 3. Selection on TE abundance and insertions.** (A) Overlap of TE families with
1107 significant abundance differences among studies. S>C and C>S denote TEs with a higher
1108 abundance in selected or control populations, respectively. Red bars indicate a significant
1109 overlap at P < 0.05 (also see **Table S8**). (B) Boxplots of the number of genomic insertions
1110 relative to the total genomic content of the 2 significantly shared C>S TEs. (C) Genome-wide
1111 differentiation in TE insertion frequency between selected and control populations in
1112 Fabian2018 and (D) Hoedjes2019. Every point indicates the $-log_{10}$ P-value of a TE insertion
1113 across chromosomal arms (alternating black and grey color). The solid orange line
1114 corresponds to the Bonferroni cut-off at α = 0.05 (Fabian2018: P < 5.9 x 10^{-6}; Hoedjes2019:
1115 P < 3.8 x 10^{-6}). Red and blue points denote TE insertions with a significantly higher frequency
1116 in selected or control populations, respectively. More details including exact positions,
1117 frequency and annotation of candidate TE insertions can be found in **Table S9**.
1118

1119 **Figure 4. Multiple factors influence TE expression.** (A) Proportions of differentially
1120 expressed TEs at adjusted P < 0.05 and directionality relative to 123 TEs with detectable
1121 expression for factors from statistical models on pre-filtered read counts in DESeq2 (also see
1122 **Table S12**). "Sex" refers to the results of the model including Sex (M, males; F, females), Age
1123 (young; old), and Regime (C, control; S, selected). "Regime", "Age" and "RxA" (i.e. Regime x
1124 Age interaction) refer to results from model fits with males and females separately analyzed.
1125 The absolute number of TEs for factor levels are given above or below bars. (B) Log_2 fold
1126 change of regime (selected vs control) and (C) age (young vs old) for males and females.
1127 Colors designate TEs significant only in males (blue), or females (red), or shared between
1128 both sexes (orange). Not significant TEs are in grey. (D and E) Log_2 fold changes across
1129 regime against age differences in males and females. Colors designate TEs significant only
1130 for regime (blue), or age (red), or for both factors (orange). Not significant TEs are in grey. (F)
1131 Relationship of log_2 fold changes in TE expression and genomic abundance between regimes
1132 in females. (B to E): r, Pearson's correlation coefficient; (F): ρ, Spearman's correlation
1133 coefficient; * P < 0.05; *** P < 0.0001; ns, not significant.
1134
1135 **Figure 5. Number of genetically and transcriptionally differentiated genes involved in**
1136 **regulation of TE activity.** (A) Counts of genetically differentiated (G.D.) TE regulation genes
1137 reported in the four experimental evolution studies. (B) Number of TE regulation genes
1138 differentially expressed (D.E.) between regimes (C, control; S, selected) and ages (young; old)
1139 in the RNA-seq data of Carnes2015 (whole female flies) and microarray data of Remolina2012
1140 (female heads and abdomens). Also see **Table S19** for information on all 96 TE regulation
1141 genes.
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155

1156 **TABLES**

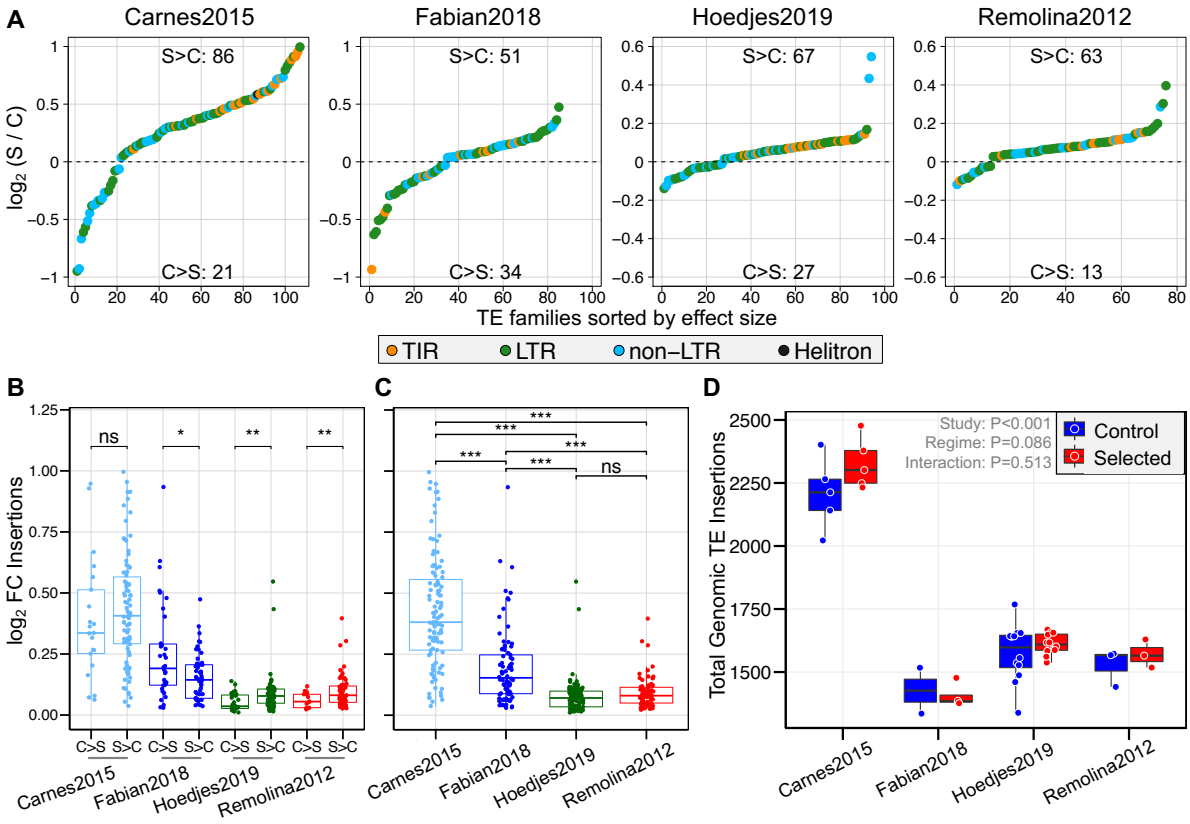
1157 **Table 1.** Number of detected TE families (N) and percentage of families more abundant in
1158 selected (S>C) or control regimes (C>S) or not different (n.s.) using three different approaches
1159 (also see Fig. S1 and Table S2).

Approach	Study	N	N (sign.)[a]	S>C	C>S	n.s.
#1 For Hoedjes2019: ~Regime+Diet+Regime x Diet For other studies: ~Regime+Pop[Regime]	Carnes2015	112	107	77%	19%	4%
	Fabian2018	110	85	46%	31%	23%
	Hoedjes2019	115	94	58%	24%	18%
	Remolina2012	110	76	57%	12%	31%
#2 ~Study+Regime+Study x Regime	Studies Combined	103	Regime: 41 Study: 101 Study x Regime: 65	33%	7%	60%
#3 Consistent differences between all S and C populations	Carnes2015	112		43%	2%	55%
	Fabian2018	110		14%	7%	79%
	Hoedjes2019: Low[b]	115		37%	2%	61%
	Hoedjes2019: Medium[b]	115		3%	0%	97%
	Hoedjes2019: High[b]	115		3%	29%	69%
	Remolina2012	110		3%	0%	97%

1160 [a]Significant after Bonferroni correction at α=0.01 and FDR<0.05 in approach #1 and #2,
1161 respectively. [b]Three larval diet conditions; low had 0.25x less and high had 2.5x more sugar
1162 and protein compared to medium diet.

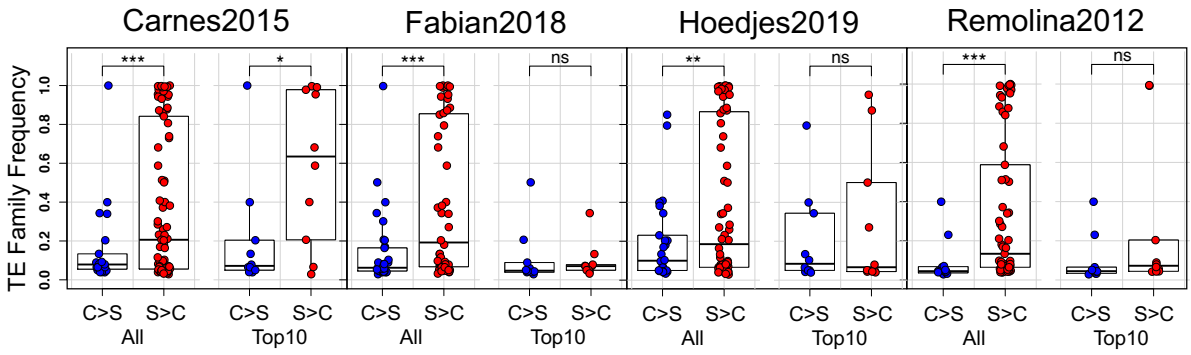
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181

1182 **FIGURE 1**



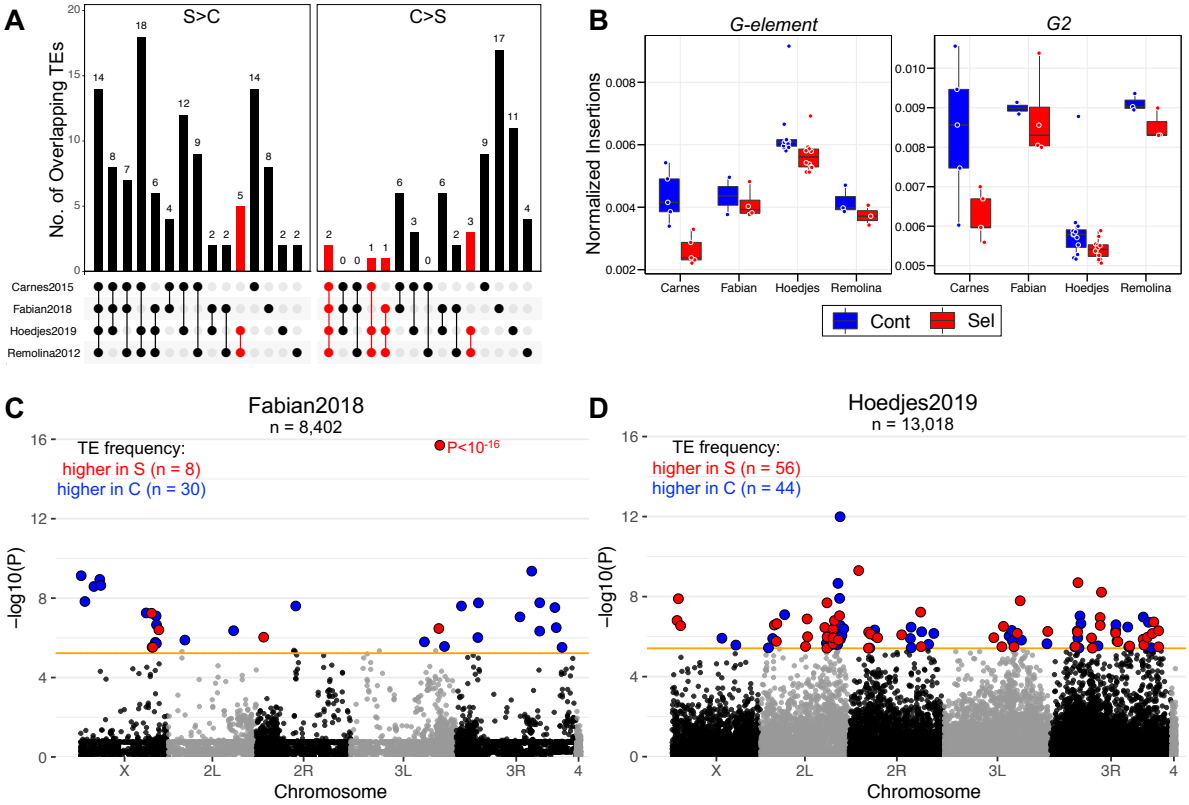
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202

1203 **FIGURE 2**



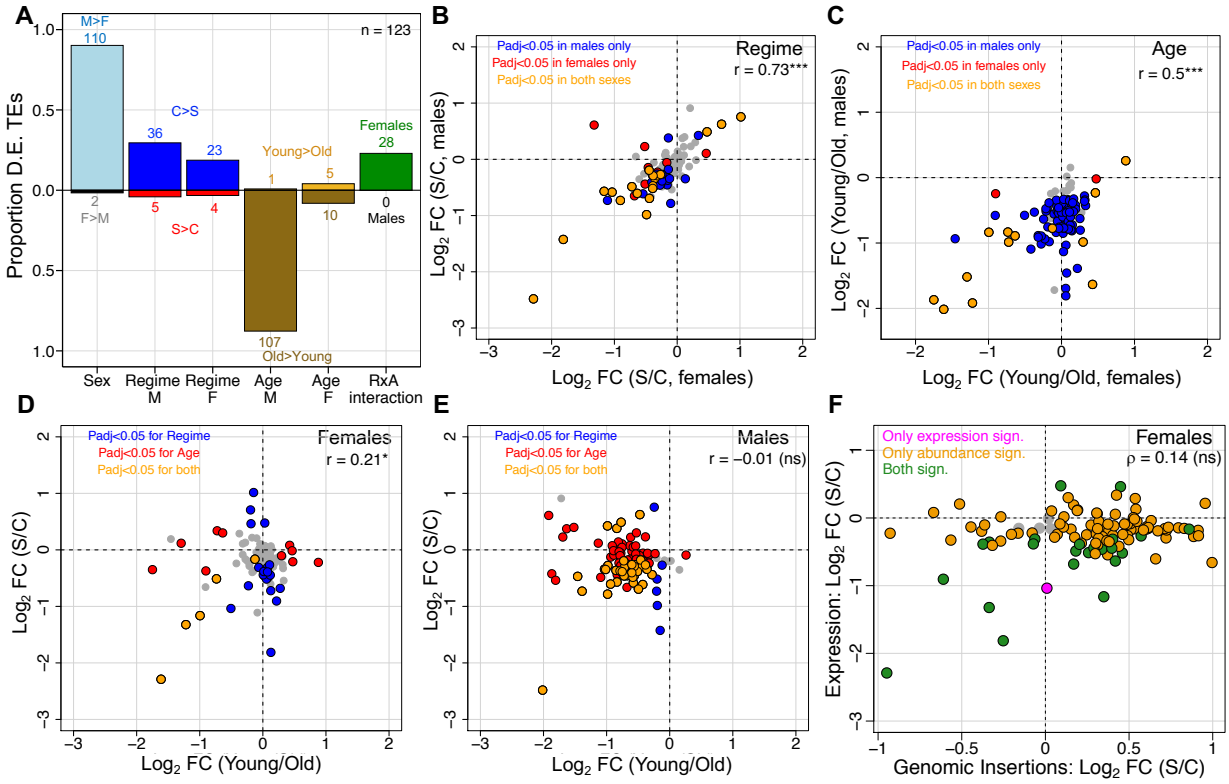
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232

1233 **FIGURE 3**



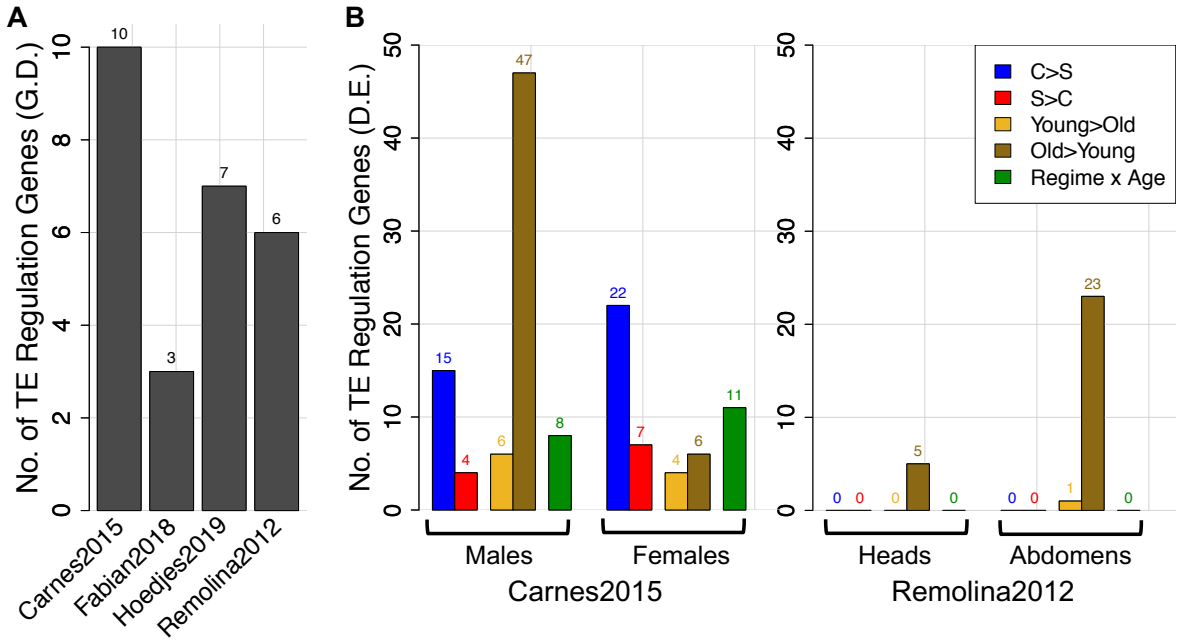
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253

1254 **FIGURE 4**



1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274

1275 **FIGURE 5**



1276