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 utilize powerful mechanisms to demobilize TEs, transposons become increasingly derepressed during aging. The rising TE activity causes genomic instability and was suggested to be involved in age-dependent neurodegenerative diseases, inflammation and the determination of lifespan. It is therefore conceivable that long-lived individuals have improved TE silencing mechanisms and consequently fewer genomic insertions and a reduced TE expression relative to their shorter-lived counterparts. Here, we test this hypothesis by performing the first analysis of genome-wide insertions and expression of TEs in populations of *Drosophila melanogaster* selected for longevity through late-life reproduction for 50-170 generations from four independent studies. Surprisingly, we found that TE families were generally more abundant in long-lived populations compared to non-selected controls. Despite simulations showed that this was not expected under neutrality, we found little evidence for selection driving TE abundance differences. Additional RNA-seq analysis revealed only few differentially expressed TEs whereas reducing TE expression might be more important 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 prolonging lifespan. Our results provide a novel viewpoint proposing 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 identifying 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 to 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 they 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 instance, 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 through the accumulation of structural and point mutations over evolutionary time scales (Brouha et al. 2003).

For a long time, transposons were unfairly deemed as ‘junk DNA’ that do not have any significant impact on organismal function, but a series of studies proofing TE-mediated adaptive evolution and their role in diseases initiated a paradigm shift in thinking. Nevertheless, the extent to which active transposition and immobile TEs residing in the genome contribute to host fitness are still controversial. 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 impose negative consequences on health and have been associated with ~100 diseases in humans, including cystic fibrosis, haemophilia and cancer (Hancks and Kazazian 2012). Moreover, TE expression and translation allow the formation of toxic TE products that, for example, contribute to autoimmune diseases, whereas TE activity and replication of an increased genomic TE content might also impose metabolic costs to the host (Kaneko et al. 2011; Barrón et al. 2014; Volkman and Stetson 2014; Bogu et al. 2019).

On the other hand, there is mounting experimental evidence for positive selection on segregating TE insertions from multiple taxa confirming beneficial phenotypic properties including insecticide and virus resistance in *Drosophila* (Daborn et al. 2002; Magwire et al. 2011; Kuhn et al. 2014; Li et al. 2018; Rech et al. 2019). Still, recent arguments have emphasized the notion that most TE insertions are likely to be slightly deleterious or neutral with little effects on fitness, and that adaptive TEs are negligibly rare (Arkhipova 2018).

An evolutionary conserved characteristic of TEs is that their rate of transposition and expression is not constant throughout life, but rather changes during aging as demonstrated in various organisms including yeast, *D. melanogaster*, *C. elegans*, mice, and humans (Maxwell et al. 2011; Dennis et al. 2012; Solyom et al. 2012; De Cecco et al. 2013; Li et al. 2013; Chen et al. 2016; Bogu et al. 2019; De Cecco et al. 2019). TEs have further been implicated in age-related neurodegenerative diseases (e.g. Krug et al. 2017; Prudencio et al. 2017; Guo et al. 2018) and might promote chronic inflammation observed during aging (Chen et al. 2014; De Cecco et al. 2019) further supporting the involvement of TEs in senescence and longevity as proposed by the emerging ‘transposable element theory of aging’. The age-related change in TE activity detected in many tissues has mainly been attributed to chromatin remodeling and the decline in repressive heterochromatin structure which is

126 commonly rich in transposable elements (Dimitri and Junakovic 1999; Wood and
127 Helfand 2013; Chen et al. 2016; Wood et al. 2016). TEs that are not suppressed by
128 chromatin structure are target of post-transcriptional silencing by the host RNA-
129 interference (RNAi) machinery, mostly the piwi-interacting RNA (piRNA) pathway, that
130 is in turn also necessary for heterochromatin formation and stability (Lippman and
131 Martienssen 2004; Martienssen and Moazed 2015). Indeed, research has identified
132 longevity-promoting effects of several genes involved in the RNAi machinery and
133 heterochromatin formation (Mori et al. 2012; Wood and Helfand 2013; Wood et al.
134 2016). Interestingly, it is possible that age-related misexpression of TEs is exclusive
135 to the soma due to efficient post-transcriptional TE silencing mediated by the piRNA
136 machinery in the germline (Sturm et al. 2015; Elsner et al. 2018; Erwin and Blumenstiel
137 2019). Considering current evidence, it seems natural that longevity can be achieved
138 through impeding TE activity and controlling the genomic content of TEs. However,
139 whether variation in aging and lifespan within species is mediated by transposons and
140 their role in the evolution of senescence is largely unknown.
141
142 Here, we analyze genomes of *D. melanogaster* populations experimentally selected
143 for increased lifespan through postponed reproduction from four independent studies
144 to understand the role of TEs in the evolution and genomic basis of late-life
145 performance and aging. The invertebrate *D. melanogaster* is an excellent model in
146 this respect as it exhibits abundant genetic and phenotypic variation in fecundity and
147 traits related to aging that can be selected for. In the present experiments, replicate
148 populations derived from nature were subjected to a late-life breeding scheme in which
149 only flies surviving and fertile at old age contributed to the subsequent generations,
150 while control individuals reproduced earlier in life. At the point of whole genome
151 sequencing, this process had continued for over 30 years with ~170 and ~150
152 generations of selection for Carnes et al. 2015 (Carnes2015) and Fabian et al. 2018
153 (Fabian2018), and for 58 and 50 generations for Hoedjes et al. 2019 (Hoedjes2019)
154 and Remolina et al. 2012 (Remolina2012) enabling us to quantify differences in TE
155 content of long- and short-term evolutionary responses. Selection for postponed
156 senescence has resulted in phenotypic divergence of multiple fitness traits, most
157 notably an ~8% to ~74% increase in lifespan and improved old age fecundity at the
158 cost of reduced early reproduction (Luckinbill et al. 1984; Rose 1984; Remolina et al.
159 2012; Carnes et al. 2015; Fabian et al. 2018; Hoedjes et al. 2019; May et al. 2019). At

160 the genome level, analysis of genetic differentiation has revealed a significant sharing
161 in candidate genes across the four studies indicating parallel evolution (Hoedjes et al.
162 2019), but at the same time exposed multiple novel targets of selection. For instance,
163 three of the studies report genetic and/or transcriptomic divergence in immunity genes,
164 and it has recently been confirmed that these molecular changes reflect differences in
165 traits related to pathogen resistance (Fabian et al. 2018). Thus, despite variations in
166 the experimental designs, numerous evolutionary repeatable phenotypic and genetic
167 adaptations have been observed, but the importance of TEs in these studies has
168 remained unexplored. Our main objective therefore was to investigate for the first time
169 whether TE abundance in the genome, and host genes related to TE regulation, had
170 undergone similar parallel changes. Using RNA-seq data from Carnes et al. (2015),
171 we further test if males and females of selected populations evolved to suppress TE
172 transcription to mitigate potentially negative effects on longevity.
173

174 # RESULTS

175 **Genomic TE abundance**

176 Our initial goal was to test if transposable elements differ in the number of genomic
177 insertions between late-breeding, long-lived selection (S) and early-breeding control
178 (C) populations of four studies performing experimental evolution of postponed
179 senescence (**Table S1**). As a first step, we excluded all TEs that did not pass our
180 coverage and mapping quality filters and therefore included between 110 to 115 TEs
181 dependent on the study in our downstream analysis. We then used three different
182 approaches to test differences in TE abundance across control and selection regimes
183 (see overview in **Fig. S1**).
184 As sequencing depth varies throughout the positions within a TE, resulting in a
185 distribution of insertion counts rather than a single point-estimate, we first analyzed
186 the differences between regimes using all insertion value estimates within individual
187 TEs separately for each study (**Fig. 1** and **Table 1**). We found that between 46% and
188 77% of all TEs had a significantly larger number of genomic insertions in the selected
189 populations relative to controls after Bonferroni correction for multiple testing (from
190 here on referred to as S>C TEs). In contrast, only 12% - 31% of TEs showed the
191 opposite pattern and had more insertions in the controls (from here on referred to as
192 C>S TEs). As expected, the magnitude of difference in TE insertions between regimes

193 was correlated with the genomic copy number of TE families (Pearson's r,
194 Carnes2015: 0.81, Fabian2018: 0.52, Hoedjes2019: 0.77, Remolina2012: 0.6, all P <
195 0.0001). To correct for this, we also obtained $\log_2$ fold change (FC) values by dividing
196 average insertions of selected by control populations for each TE (lower panels in **Fig.**
197 **1**). S>C TEs had significantly larger $\log_2$ FC values in the two short-term evolution
198 studies of Hoedjes2019 and Remolina2012, while the opposite pattern and no
199 difference was observed for Fabian2018 and Carnes2015, respectively (**Fig. S2A**; t-
200 tests between absolute $\log_2$ FC values of S>C and C>S; Carnes2015: P ~ 0.47;
201 Fabian2018: P ~ 0.01; Hoedjes2019: P ~ 0.001, Remolina2012: P ~ 0.005).
202 This demonstrates that the dynamics of TE change between control and selected
203 populations varies among studies. Indeed, we find that studies also differed
204 significantly in the size of $\log_2$ FC values in the order of Carnes2015 > Fabian2018 >
205 Hoedjes2019 = Remolina2012 (**Fig. S2A**; ANOVA with Study term, Tukey post hoc
206 test; P < 0.001 for all pairwise comparisons except Hoedjes2019-Remolina2012,
207 which was not significant), seemingly scaling with the length of selection (Carnes2015:
208 170; Fabian2018: ~146; Hoedjes2019: 58, Remolina2012: 50 generations).
209 As a second, complementary approach to test for TE abundance differences between
210 selection and control regimes, we also fitted linear models using average number of
211 insertions per TE and population and analyzed the four studies together. Because the
212 total number of TEs was significantly different between studies (ANOVA P < 0.0001
213 of Study factor; **Fig. S3**), we normalized the insertion counts of each individual TE by
214 the total insertions in the corresponding population and analyzed arcsine square root
215 transformed proportions (**Fig. S1B** and **Fig. S1C**). At an FDR of <0.05, we identified
216 41 TEs with a significant regime factor of which 34 were more and 7 less abundant in
217 selected populations, confirming our findings of individual studies (**Table 1**).
218 Comparable tendencies were obtained when we analyzed the raw insertion counts or
219 standardized z-scores of insertions and changing the significance threshold did not
220 alter the fact that S>C TEs were more frequent than C>S (not shown).
221 In our third - and most conservative - analysis, we identified TEs with consistent
222 differences in genomic TE insertions between all control and all selected populations
223 (**Table S2**). Interestingly, the number of TEs were generally smaller compared to our
224 statistical approach indicating that TEs within control or selected populations do not
225 always behave in a parallel way. The insertion tendency in favor of S>C was again
226 apparent in three studies. In Hoedjes2019, however, we observed the same pattern

227 for populations selected on larval diet with low and medium sugar/protein
228 concentration, but the opposite for the high sugar/protein diet. A qualitatively similar
229 results was obtained applying the statistical approach from above (**Fig. S4** and **Table**
230 **S3**). Yet, the overall increase in TE insertions in late-breeding populations was
231 apparent in our model which corrects for the effect of diet (**Fig. 1** and **Table 1**).
232 We further tested if our results are influenced by the chosen mapping quality cut-off of
233 20 and redid the analyses without applying a threshold. The number of genomic
234 insertions per TE were highly correlated between both mapping approaches
235 (Pearson's r range = 0.94 – 1 in all populations and studies, P < 0.0001), and changing
236 the cut-off resulted in similar proportions of S>C and C>S TEs (not shown).
237
238 To analyze if changes in TE abundance are driven by certain TE subclasses (Long
239 Terminal Repeat, LTR; Non-Long Terminal Repeat, non-LTR; Terminal Inverted
240 Repeat, TIR) or class (RNA, DNA) we tested S>C and C>S TEs for enrichment of
241 these types using two-sided Fisher's exact tests. We only detected a significant
242 underrepresentation of TIRs and DNA-class TEs (i.e. overrepresentation of RNA-
243 class) in the C>S group of TEs of Carnes2015 and Hoedjes2019 (Carnes2015, TIRs:
244 P ~ 0.044; DNA/RNA class: P ~ 0.024; and Hoedjes2019, TIRs: P ~ 0.013; DNA/RNA
245 class: P ~ 0.008), while significant TEs of Fabian2018 and Remolina2015 had no
246 enrichment in class or subclass.
247 Despite the bulk of individual TEs had a higher genomic abundance in the selected
248 populations, the whole genomic TE content was not significantly different between the
249 regimes (ANOVA Regime factor: P ~ 0.086; **Fig. S3**). This was at least partly driven
250 by the fact that relative to S>C TEs, the fewer C>S TEs showed a significantly higher
251 magnitude of insertion differences in two studies (**Fig. S2B,** t-test on S>C vs C>S TEs
252 δInsertion values; Carnes2015 P ~ 0.04; Fabian2018 P ~ 0.005). The non-significant
253 difference in overall genomic TE load could therefore be a result of a large number of
254 S>C TEs with small differences that are balanced by fewer C>S TEs with large
255 differences. We further analyzed the whole genomic content of individual subclasses
256 of TEs and identified that selected populations had on average a significantly higher
257 number of TIRs than controls (**Fig. S3**, ANOVA, both Regime and Regime x Study
258 factors, FDR < 0.001). Tukey tests analyzing differences between selected and control
259 populations indicated that this effect is strongly influenced by Carnes2015 (Tukey

260 HSD, Regime x Study factor testing for C vs S within studies, Carnes2015: P < 0.0001;
261 Hoedjes2019: P ~ 0.85; Fabian2018 and Remolina2012: P ~ 1). We also detected that
262 selected populations had a larger LTR retrotransposon load than controls (ANOVA,
263 Regime factor, P ~ 0.026), whereas non-LTR content did not differ significantly.
264 Finally, we note that studies in general varied significantly in total TE content and
265 subclass-specific loads (ANOVA, Study factor, P < 0.0001 in all models).
266 In summary, our results demonstrate that selection for postponed reproduction leads
267 to evolutionary repeatable increases in copy number of many TEs relative to early
268 bred controls, without affecting the overall genomic TE load.
269
270 **Approximate TE mobility**
271 We next tested if the change in abundance of TEs between control and selected
272 populations can be explained by differences in average population frequency of TE
273 insertions which serves as a proxy for recent mobility and age of TE invasion (Kofler
274 et al. 2015).
275 We first determined the exact genomic location and frequency of TE insertions and
276 calculated average population frequency across all insertion sites for each TE family.
277 Importantly, while TE abundance (as above) is quantified by the total number of reads
278 mapping to a TE relative to single-copy genes (Weilguny and Kofler 2019), identifying
279 the exact genomic location of insertions requires mates of a read-pair to map
280 discordantly to the chromosome and TE sequence, and therefore strongly depends
281 on the sequencing depth and number of populations (Cridland et al. 2013; Kofler et al.
282 2016; Lerat et al. 2019). The genome-wide average coverage ~162, ~101, ~41, and
283 ~23 for Fabian2018, Hoedjes2019, Remolina2012, and Carnes2015, respectively. As
284 expected, the number of detected TE insertions partially scaled with coverage: across
285 all populations within a study, we found 13,018 TE insertions in Hoedjes2019, 8,402
286 in Fabian2018, and 4,502 in Remolina2012, which is in the range recently identified in
287 natural populations (i.e. 4,277 - 11,649 TE insertions in Lerat et al. 2019). The least
288 number of insertions was found for Carnes2015 for which we identified an unusually
289 small number of 567 TE insertions.
290 For each TE family, we then averaged frequencies across all of its detected genomic
291 positions to estimate the mean frequency at which a TE is segregating in a population
292 (Kofler et al. 2015). Studies varied in the minimum average TE frequency in the order
293 of Carnes2015 > Remolina2012 > Fabian2018 > Hoedjes2019, which is likely a further

294 effect of dissimilar sequencing depths and other experimental factors (average
295 frequency ranges of Hoedjes2019: 0.01 - 0.9; Fabian2018: 0.02 - 1; Remolina2012:
296 0.04 - 0.84; Carnes2015: 0.19 – 0.9). The TE frequencies of Carnes2015 therefore
297 need wo be interpreted with care, considering the likely insufficient amount of data.
298
299 To get unbiased average TE frequency estimates independent of coverage
300 fluctuations across studies, we also obtained average frequencies from a single
301 natural South African (SA) population (Kofler et al. 2015; Kofler 2019). The SA
302 population had a higher sequencing depth than all studies here (Fabian2018: ~2x,
303 Hoedjes2019: 3.2x; Remolina2012: ~7.8x, Carnes2015: ~13.9x times more) and thus
304 presumably a more accurate estimate of TE frequencies. Notably, this population was
305 not subjected to any selection or control treatment and was only maintained 8
306 generations in the lab before sequencing. Average genome-wide TE frequencies of
307 control and selected populations of Fabian2018, Hoedjes2019 and Remolina2012, but
308 not Carnes2015, were significantly correlated with the South African TE frequencies
309 (**Fig. S5**; Spearman's r, Fabian2018: 0.65; Hoedjes2019: 0.61; Remolina2012: 0.58,
310 all three $P < 0.0001$; Carnes2015: 0.1, $P \sim 0.4$), demonstrating that the SA population
311 can function as an appropriate reference here.
312 Previous studies reported that low frequency TEs are more abundant in genomes of
313 *D. melanogaster* (Petrov et al. 2011; Kofler et al. 2015). We confirmed that for all
314 populations, TE abundance was similarly negatively correlated with average TE
315 frequency of the SA population (Spearman's r range, Carnes2015: -0.49 to -0.54;
316 Fabian2018: -0.4 to -0.5; Hoedjes2019: -0.42 to -0.45; Remolina2012: -0.5 to -0.51;
317 all $P < 0.0001$). Moreover, we obtained comparable results when we performed this
318 analysis using the average TE frequencies of Fabian2018, Hoedjes2019 and
319 Remolina2012 (Spearman's r averaged across all populations, Fabian2018: -0.33;
320 Hoedjes2019: -0.37; Remolina2012: -0.4, all populations $P < 0.001$; Carnes2015: -
321 0.29 to 0.28, two populations significant at $P < 0.05$). The transposable element
322 content of all populations can therefore be characterized by a large number of TEs
323 with low frequency, whereas TEs with a high population frequency are less abundant
324 in the genomes.
325 We then employed the data of the SA population and asked if C>S and S>C TEs vary
326 in average frequency to identify TEs that have presumably been mobile in the two
327 regimes (**Fig. 2**). C>S TEs had a significantly lower frequency than S>C TEs in all four

328 studies (t-tests between C>S and S>C frequencies, $P < 0.05$ for all three studies). As
329 there were more S>C than C>S TEs, we also contrasted the average frequencies of
330 the top 10 C>S and S>C TEs with the biggest changes in genomic abundance defined
331 by $\log_2$ FC values (lower panel of **Fig. 1**). We only detected a significantly higher
332 frequency in top 10 S>C relative to C>S TEs for Carnes2015 (t-test, $P \sim 0.03$), but not
333 in the other two studies. Except from the marginal difference in Carnes2015, this
334 suggests that TEs with the biggest abundance changes between regimes have similar
335 levels of mobility.
336 In summary, our analysis proposes that C>S TEs are potentially mobile, low frequency
337 TEs expanding to new sites in the genome, whereas S>C TEs span from low to high
338 frequency, possibly reflecting TEs with a broad range of transposition rates.
339
340 **Selection on TE abundance and insertions**
341 A major challenge in experimental evolution studies is to differentiate selection from
342 the confounding genomic signals of genetic drift, which might be amplified by small
343 effective population sizes (N_e) or varying generation times spent in the lab between
344 control and selected populations. We therefore calculated genome-wide nucleotide
345 diversity π and Watterson's θ across 100kb windows as a proxy for N_e. With the
346 exception of Fabian2018, where π was equal between regimes (ANOVA, Regime
347 factor, $P \sim 0.18$), we found that both estimators were significantly higher in selected
348 relative to control populations (**Table S4**; ANOVA, Regime factor, all $P < 0.0001$). Even
349 though a generally reduced N_e in controls should lead to the loss of low and fixation of
350 high frequency TEs under neutrality, we observed the opposite pattern in our analysis
351 above (**Fig. 2**).
352 To further formally test if the increased abundance of many TEs is driven by genetic
353 drift alone, we performed population genetic simulations using the correlated average
354 TE frequencies from the natural South African population (Kofler et al. 2015) as a
355 starting point (see **Fig. S5** and results above). We simulated TE frequency change in
356 selected and control populations 5,000 times given the reported consensus population
357 sizes as N_e, generation times and number of replicates. We then asked how often the
358 same or a higher relative proportion of S>C to C>S TEs as in our observations is
359 obtained (**Table 1**). While the results from Carnes2015, Hoedjes2019, and
360 Remolina2012 were significantly different from the expected proportions, the TE

361 abundance differences of Fabian2018 could be caused by genetic drift alone (**Fig.**
362 **S6**). Testing different ranges of the reported population sizes and assuming that only
363 50% and 25% of flies in the selected populations were able to breed at old age resulted
364 in qualitatively similar results (not shown). We also quantified expected proportions of
365 TEs consistently varying in frequency across simulated replicates: while there were
366 generally more TEs consistently higher in abundance in selected populations (**Table**
367 **S2**), all our simulations resulted in more TEs with a consistently higher frequency in
368 controls. The increased genomic abundance of many TEs in selected populations is
369 therefore unlikely to be solely caused by genetic drift.
370
371 Considering the deviation from neutrality, we next asked if the parallel patterns in TE
372 abundance are caused by the same or different TEs, which could indicate selection
373 acting on genomic copy number of certain TEs. We considered the TEs significantly
374 varying in abundance from the first linear model in which we analyzed each study
375 separately (**Fig. 1** and **Table 1**), and created study-overlaps of S>C and C>S TEs.
376 Among the 103 common TEs, we identified a sharing of 14 S>C and 2 C>S TEs across
377 all four studies (**Fig. 3A**). Even though this seemingly large number of shared S>C
378 TEs, only the overlap between Remolina2012 and Hoedjes2019 was significant ($P <$
379 0.05). Yet, we found that the most common telomeric TE *HeT-A* (Casacuberta 2017)
380 was significantly more abundant in selected populations in all four studies and in our
381 combined model (**Fig. S1**), suggesting that long-lived populations might have evolved
382 longer telomeres to avoid attrition, which is considered to be a key conserved
383 mechanism of aging (López-Otín et al. 2013). In contrast to S>C TEs, we detected
384 several significant overlaps for the C>S group of TEs. Potentially, a high genomic
385 abundance of several TEs, most importantly *G-element* and *G2* found in all four
386 studies, is detrimental for longevity and late-reproduction (**Fig. 3B**). Despite some
387 significant overlaps, we did not observe any significant Spearman's correlation
388 coefficients in pairwise comparisons of $\log_2$ FC values between studies except for
389 Hoedjes2019-Remolina2012 ($r = 0.28$, $P \sim 0.004$).
390
391 Genomic TE abundance in selected populations might also be increased because
392 selection acted on a large number of TEs segregating in the base populations resulting
393 in frequency divergence between control and selected populations. We therefore
394 screened all identified TE insertion sites across the genome for significant frequency

395 differences between regimes in each study by performing ANOVAs on arcsine square
396 root transformed frequencies and Bonferroni correction. We did not find any TE
397 insertions significantly varying in frequency in Carnes2015 and Remolina2012, and
398 using a less stringent cut-off at FDR < 0.05 did not change this result. However, we
399 detected 38 significant TE insertions out of 8,402 in Fabian2018 and 100 out of 13,018
400 in Hoedjes2019 (**Fig. 3C** and **Fig. 3D, Table S5**). While there were more TEs with a
401 significantly higher frequency in control populations in Fabian2018 (8 higher vs 30
402 lower frequency in S), the opposite pattern was observed in Hoedjes2019 (56 higher
403 vs 44 lower frequency in S).
404 At the gene level, the significant TEs defined 29 and 98 genes in Fabian2018 and
405 Hoedjes2019, respectively, and none were shared between the two studies. In
406 accordance with the notion that most TEs are neutral (Arkhipova 2018), only a
407 negligible fraction of insertions might be beneficial for longevity and postponed
408 reproduction.
409 To further investigate if frequency changes explain variation in abundance, we tested
410 if each TE family varying in abundance also differs significantly in frequency between
411 the regimes, rather than analyzing individual insertion sites (**Table S6**). We fit the
412 same model as for the abundance analysis but using arcsine square root transformed
413 TE frequencies (**Figure S1A**). At an FDR < 0.05, we found that in Carnes2015, 27 TE
414 families had a higher frequency and abundance in selected populations, while 7
415 showed a mismatch in direction of frequency and abundance change. In Fabian2018
416 and Hoedjes2019, we only detected 1 and 3 TE families with significant frequency
417 variation all of which matched the differences in abundance, whereas none were
418 significant in Remolina2012.
419 Thus, despite differences in TE abundance are likely not driven by neutral evolution
420 alone, we only found limited evidence for parallel evolution of TE copy numbers and
421 sparse TE frequency differentiation.
422
423 **Differential TE expression**
424 To test whether the increased genomic abundance of TEs in selected flies is explained
425 by a higher transcriptional activity we analyzed RNA-seq data from whole flies of
426 Carnes2015 (**Fig. 4** and **Table S7**). We first fit a model with *Sex*, *Age*, *Regime* and all
427 interactions to every TE and each gene on the major chromosomal arms. In line with
428 abundant sexual dimorphism of gene expression observed by Carnes et al. (2015),

429 we identified that ~93% of TEs had a significant main factor of sex or interaction
430 including it, whereby all 109 TEs significant for the sex term had a higher expression
431 in males than females **(Fig. 4A** and **Fig. S7)**. To exclude that this is caused by a
432 technical artifact, we asked if a similar male-bias is observed for the genes on the
433 major chromosomal arms. We found that 53% of all 13,255 included genes had a
434 significantly higher expression in males while 37% were biased towards females.
435 Moreover, the proportions of genes with male-, female- or no bias was significantly
436 different from the proportions observed for TEs (χ^2 test on counts of TEs vs genes, P
437 < 0.0001). Because we do not find a similarly strong male-biased expression for genes
438 or a higher number of read counts in males (not shown), we conclude that TEs are
439 generally higher expressed in males than females in the present populations.
440
441 We therefore decided to test the effects of *Regime*, *Age*, and the *Regime x Age*
442 interaction in the sexes separately (**Fig. 4**, **Table S7**). We detected several TEs
443 differentially expressed between selected and control regimes. In males, 19 TEs
444 (~16% of total) had a significant regime term of which a majority of 17 TEs were
445 upregulated in controls (**Fig. 4B**). Among these, there were 10 LTRs, 6 non-LTRs, and
446 1 DNA-class foldback TE, whereas the 2 upregulated in selected populations were
447 both non-LTR TEs (*TART-A*, and *TART-B*). Less TEs changed across regimes in
448 females (10 TEs, ~8%) compared to males. However, females showed a similar
449 tendency towards higher expression in controls: 6 TEs had a higher and 4 TEs a lower
450 expression in controls relative to selected flies. The 6 TEs with a higher expression in
451 controls consist of 1 non-LTR (*R1A1-element*) and 5 LTR TEs (*copia, flea, blood,*
452 *mdg1,* and *rooA*), while the 4 more in selected were 3 non-LTR (*TART-A*, *TART-B,*
453 *TAHRE*) and 1 LTR (*1731*). We also observed that the 6 TEs significant in both sexes
454 also had the same directionality of expression change: 4 LTR-class TEs (*copia, flea,*
455 *blood,* and *mdg1*) were higher expressed in controls, whereas 2 non-LTR TEs (*TART-*
456 *A*, and *TART-B*) were upregulated in selected populations. Interestingly, *TART-A*,
457 *TART-B,* and *TAHRE* provide the enzymatic machinery for telomeric maintenance
458 (Casacuberta 2017), again suggesting that reduced telomere attrition evolved in
459 response to selection paralleling the genome-based analysis. In general, regime
460 affected TE expression in males and females similarly as indicated by a significant
461 correlation of $\log_2$ fold change values between sexes (**Fig. 4B**, Pearson's r = 0.7, P <

462 0.0001). We further asked if the magnitude of log_2 fold change varies between TEs
463 more expressed in controls or selected populations, and did not find any significant
464 difference (**Fig. S8**, t-test, females: P ~ 0.12; males: P ~ 0.49).
465
466 To investigate if derepression of TEs occurs with age, we compared the proportion of
467 TEs up- or downregulated or unchanged with age to those of genes. For both male
468 and females, we found a significant difference in the distributions of TEs and genes
469 with age (χ^2 test on counts of TEs vs genes, for males: P < 0.0001; for females: P ~
470 0.01), demonstrating that TEs show generally different age-related expression
471 changes than genes. Supporting the notion that TEs become derepressed during
472 aging, the effect of age on TE expression in males was general as all 89 significant
473 TEs (i.e. ~73% of all included TEs) had a higher expression in older flies (**Fig. 4A** and
474 **Fig. 4C**), whereas ~26% of all genes increased and also ~26% decreased expression
475 with age. In contrast to this, the effect of age was less pronounced in females in which
476 only 35 TEs (~28%) had a significant Age factor. Surprisingly, less TEs than genes
477 were upregulated with age in females (7% TEs and 15% genes), while 21% of TEs
478 and 14% of genes decreased expression with age. Still, the TEs upregulated in older
479 females had on average a significantly higher log_2 fold change of 1.3, relative to 0.53
480 for the downregulated TEs (**Fig. S8**, t-test: P < 0.01), consistent with the results of a
481 recent study (Chen et al. 2016).
482 Moreover, 27 TEs had a significant age term in both male and female models, but only
483 9 of them had the same sign and increased in expression with age (**Fig. 4A** and **Fig.**
484 **4C**). Among these were 6 LTR (*copia, gypsy, mdg1, Burdock, rooA, springer*) and 3
485 non-LTR transposons (*jockey, R1A1- and R2-elements*). This confirms previous
486 findings that showed *copia, gypsy, Burdock, R1* and *R2* upregulation with increasing
487 age (Li et al. 2013; Chen et al. 2016).
488 The 16 remaining shared TEs showed downregulation with age in females, but
489 upregulation in males, further highlighting that age has generally different effects on
490 TE expression in both sexes. In females, the 35 age-related TEs were also signficantly
491 enriched for retrotransposons whereas DNA-class TEs were underrepresented
492 (Fisher's exact test, P ~ 0.02). Male flies in contrast, did not show any enrichment.
493 Thus, males exhibit a general upregulation of TEs regardless of their class during

494 aging, whereas females show a more complicated pattern with mainly RNA-class TEs
495 affected by age.
496
497 No TE families showed a significant *Regime x Age* term in males, but the interaction
498 was significant for 28 TEs (~23%) in females (**Fig. 4A**). These consisted of 13 non-
499 LTR, 13 LTR, and 2 TIRs. Interestingly, most of these TEs were defined by a higher
500 expression in young controls compared with selected flies of the same age (see **Fig.**
501 **S9** for example). Selected populations then increased while controls decreased
502 expression, meeting at a similar expression level at old age. This is comparable with
503 recent studies which suggested that age-dependent changes in TE expression differ
504 between genotypes (Erwin and Blumenstiel 2019; Everett et al. 2019).
505
506 To examine if the selected populations might have evolved to maintain a young TE
507 expression profile, we compared differences between regimes to those that occurred
508 with age (**Fig. 4D and 4E**). In both males and females $\log_2$ FC values of all TEs were
509 significantly correlated between regime and age (Pearson's correlation, females: $r =$
510 0.55, $P < 0.0001$; males: $r = 0.19$, $P < 0.05$). Performing the same analysis only
511 considering genes also resulted in significant positive correlations (Pearson's
512 correlation, females: $r = 0.37$, males: $r = 0.41$, both $P < 0.0001$), and the coefficients
513 between males and females were more similar. Thus, variation in expression of genes
514 and TEs between selected and control populations mirrors the changes between
515 young and old flies.
516
517 In summary, our results propose that selected populations of Carnes2015 evolved to
518 reduce TE expression particularly in males and to a smaller magnitude in females, but
519 expression of most TEs was at the levels observed in controls. In agreement with
520 sexual dimorphism (Brown and Bachtrog 2017) and age-related deregulation of TEs
521 (Li et al. 2013; Chen et al. 2016), the effects of sex and age on TE expression were
522 more dominant compared to regime.
523
524 **Link between genomic and expression differentiation**
525 We next asked if the change in genomic TE abundance reflects the expression
526 differences between selected and control populations. On average, TE expression
527 was positively correlated to the number of genomic insertions in the genome (**Fig. 4F**,

528 Spearman's r ~ 0.74; P < 0.0001). To analyze if this association depends on the
529 treatment, we separated our samples into different levels of sex, age, and regime. The
530 number of genomic TE insertions predicted TE expression significantly across all
531 conditions (**Table S8**). Interestingly, the correlation coefficients were similar across
532 different sample groups, demonstrating that the relationship between expression and
533 genomic abundance of insertions is not dependent on regime, sex or age. Thus, in
534 Carnes2015, TE copy number is a robust predictor of TE expression regardless of
535 treatment.
536
537 Next, we investigated if there are parallel changes in expression and genomic
538 abundance of TEs significantly varying between regimes (**Table S9**). In males, 14 TEs
539 were significant for regime in the genomic and expression analysis all of which were
540 higher expressed in controls. Of these, 9 TEs were upregulated and had significantly
541 more genomic insertions in controls (*blood, copia*, *F-element*, *Doc2-element, Doc3-*
542 *element, G2, Porto1, ZAM, 17.6*), while 5 TEs (*Circe, G5, mdg1, Tom1, Quasimodo*)
543 had more copy numbers in selected populations. In females, 3 TEs (*blood*, *copia*,
544 *R1A1-element*) were higher expressed and had a higher genomic abundance in
545 controls, while 2 TEs (*1731*, *TAHRE*) were significantly more expressed and more
546 abundant in selected populations. The remaining two TEs signficant in the expression
547 and insertion analysis (*mdg1, rooA*) showed different signs and had despite a larger
548 number of genomic insertions in selected populations a higher expression in controls.
549 This suggests that the absence of a strong link between expression and number of
550 genomic insertions at the individual TE level. Indeed, $\log_2$ FC expression and $\log_2$ FC
551 insertions between regimes was only correlated in males (Pearson's r = 0.31, P <
552 0.001), and marginally not significant in females (r = 0.18, P ~ 0.05) (**Fig. S10**).
553 Because our results suggests some positive relationship between TE expression and
554 genomic abundance, we re-did the differential expression analysis from above using
555 RNA-seq read counts corrected for the number of TE insertions in each sample (**Table**
556 **S10**). The tendency of TEs to be higher expressed in controls was substantially larger
557 compared to the analysis where we did not correct for insertion numbers (males: 55
558 TEs more, 2 TEs less expressed in controls; females: 16 TEs more, 3 TEs less
559 expressed in controls).
560 Our results further emphasize that selection for late-reproduction leads to a reduction
561 in TE expression in whole flies. Notably, we assumed that males have the same copy

562 number as females for the analyses here. As the genomic TE abundance measures
563 came from DNA pools of female flies, copy number could differ from males due to the
564 repetitive nature of the *Y* chromosome.
565
566 **Genes involved in TE regulation**
567 We next hypothesized that if the effects of TEs on lifespan and aging are
568 predominantly negative, as proposed by many studies, experimental selection for
569 longevity would have likely resulted in clear-cut genetic and expression differentiation
570 in 96 known chromatin-structure, piRNA, and transposition-associated genes likely
571 involved in TE regulation and silencing (**Table 2**, a complete list of genes can be found
572 in **Table S11**). Of these, 10 genes were implicated under selection in Carnes2015, 3
573 in Fabian2018, 7 in Hoedjes2019, and 6 in Remolina2012. *E2f1* (FBgn0011766) and
574 *Hsp83* (FBgn0001233) were the only genes occurring in more than one study. Thus,
575 the significant candidate gene sharing reported by Fabian et al. (2018) and Hoedjes
576 et al. (2019) does generally not include genes regulating TEs. In line with this, the four
577 studies did not report any significant enrichment of GO terms related to transposon
578 silencing and chromatin structure.
579 Using the available RNA-seq data from whole flies of both sexes in Carnes2015 and
580 microarray data from female heads and abdomens in Remolina2012, we then asked
581 if TE regulation genes are differentially expressed (**Table S11**). In Carnes2015, we
582 found that the 30 TE regulation genes with a significant regime term tended to be
583 upregulated in controls relative to selected populations (males: 3 more and 8 less
584 expressed; females: 6 more and 14 genes less expressed relative to selected
585 populations). In Remolina2012, we detected generally little differences in transcript
586 abundance between regimes at FDR < 0.05 (only 5 out of 13,995 genes significant in
587 head tissue, none significant for abdomen). We therefore analyzed genes significant
588 at P < 0.01 (491 genes for heads, 8 for abdomen) for the regime term, and only
589 detected 2 TE regulation genes of which one was more and the other lower expressed
590 in selected flies.
591 Similar to the TE expression patterns in Carnes2015 (**Fig. 4A**), the effect of age was
592 stronger than selection regime in both studies and significant TE regulation genes
593 showed a clear tendency for upregulation with age. In males of Carnes2015, 41 TE
594 regulation genes increased and 6 decreased expression with age. The effect was less
595 pronounced in females in which 19 genes increased transcription and 8 decreased

596 with age. Likewise, in Remolina2012, all 5 TE regulation genes significant for age in
597 heads ramped up transcription with age. The age effect was even stronger in
598 abdomens, where 24 genes increased with age, while only one showed the opposite
599 pattern.
600 Taken together, the small number of genetically differentiated TE regulation genes,
601 lack of TE-associated GO enrichment, and overall missing sharing between studies
602 suggests that improving TE repression was either specific to studies and/or not a prime
603 target of selection. Indeed, experimental selection has mainly affected expression of
604 TE regulation genes in the Carnes2015 long-term evolution study but not
605 Remolina2012, suggesting that altering expression of these genes is not a general
606 necessity for increasing lifespan. The boosted expression of TE regulation genes at
607 older ages appears to be common and might be a response to increased TE
608 transcription in old flies.
609
610 **DISCUSSION**
611 Are transposable elements conferring an adaptive advantage as shown for many traits
612 (Daborn et al. 2002; Magwire et al. 2011; Kuhn et al. 2014; Li et al. 2018; Rech et al.
613 2019) or should they be purged and repressed during the evolution of longevity due to
614 their widespread negative effects on fitness (Chen et al. 2014; Krug et al. 2017;
615 Prudencio et al. 2017; Guo et al. 2018; De Cecco et al. 2019)? In this report, we
616 attempt to answer this controversial question by employing four independent data sets
617 to present the first characterization of the genome-wide TE content and expression in
618 *D. melanogaster* populations experimentally selected for late-life reproduction and
619 longevity.
620
621 **Does longevity-selection lead to changes in TE abundance?**
622 Variation in TE copy number has been associated with some geographic and climatic
623 factors (Kalendar et al. 2000; Kreiner and Wright 2018; Lerat et al. 2019) in natural
624 populations of plants and *Drosophila* and was shown to change during experimental
625 evolution in different temperatures (Kofler et al. 2018). Our analysis revealed
626 repeatable predominance of TE families with increased genomic insertions in late-
627 breeding, long-lived populations and indicates that reproductive age, with some
628 dependency on developmental diet, is another factor influencing divergence in TE
629 abundance **(Fig. 1** and **Table 1)**. Interestingly, we find that the degree of TE

630 abundance change was significantly different between studies and roughly scaled with
631 the number of generations under selection (**Fig. S2A**). While parallel changes in TE
632 characteristics within populations of the same selection regime have been reported by
633 similar experiments (Graves et al. 2017; Kofler et al. 2018), it is striking that we
634 observed this pattern in data created by four independent studies. Despite most TE
635 families are more abundant in long-lived populations, our analysis shows no significant
636 difference in the total genomic TE content between control and selected populations
637 (**Fig. S3**), which was partly driven by a few TEs with large increases in copy number
638 in controls that oppose many TEs with small increases in abundance in selected
639 populations (**Fig. S2B**). Changes in the overall genomic TE load are therefore likely
640 not essential to evolve longevity or fecundity at old age in *Drosophila*. These findings
641 are in contrast to recent work in several killifish species, which reported that TE
642 expansion can cause an increased genome size with possible negative effects on
643 lifespan (Cui et al. 2019). However, our analyses focused exclusively on the genomic
644 TE load, so that we cannot exclude a difference in genome size between control and
645 selected populations, as might be caused by other factors such as non-repetitive
646 InDels or repetitive DNA unrelated to TEs.
647
648 **Are TEs adaptive during the evolution of aging?**
649 The genomic content of TEs evolves through various factors including replicative
650 transposition, selection, genetic drift, and the TE defense machineries of the host
651 (Charlesworth and Charlesworth 1983; Kofler 2019). By performing population genetic
652 simulations considering just genetic drift, we were able to exclude that population size
653 and generations spent in the lab *per se* cause an increased abundance of TE families
654 in selected populations (**Fig. S6**). Even though it is known that the majority of TE
655 insertions are neutral to fitness (Arkhipova 2018), this proposes that factors other than
656 genetic drift influenced TEs in our experiments.
657 From a selective point of view, increasing many TE families might be beneficial for
658 longevity, while the fewer TEs less abundant in selected flies could have a negative
659 effect. Under this scenario, selection would lead to parallel increases or decreases of
660 the same TE families across studies. However, when we screened for parallel patterns
661 in abundance change, we found only two TEs (*G-element* and *G2*) that had a
662 decreased copy number in selected flies and were significantly shared across all
663 studies. Both elements are *jockey*-like non-LTR TEs, of which *G2* is highly enriched in

664 centromeric regions of the genome (Chang et al. 2019). Thus, changing centromeric
665 structure by altering its TE content could be one mechanism modulating aging, but
666 experimental evidence for this is still missing. In contrast to this, we did not find any
667 significant overlap between all four studies among TEs with an increased abundance
668 in the late-breeding populations. Unless many TE families had non-repeatable effects
669 on longevity, the small amount of significant sharing suggests that abundance of most
670 TEs is neutral.
671 Another possibility to alter TE abundance is through selection affecting TE insertions
672 that were already present in the base population at a genome-wide scale, resulting in
673 a large number of TE insertions significantly varying in frequency between control and
674 selected populations. We found only a minor fraction of TE insertions in Fabian2018
675 and Hoedjes2019, but not in the other two studies, with significantly different
676 frequencies between the regimes that are in or close to <100 genes (**Fig. 3** and **Table**
677 **S5**). The small fraction of TE insertions with a higher frequency in selected populations
678 found in only two studies together with little differences in frequency of TE families
679 suggests that standing genetic variation presented by TEs plays a role in the evolution
680 of aging, but it is unlikely to be a major driver of TE abundance differentiation.
681 Yet, we found a signal indicating potentially improved telomere maintenance – a key
682 hallmark of aging and lifespan (López-Otín et al. 2013) – as an exciting mechanism
683 possibly adaptive during the evolution of aging. Telomere shortening coincides with
684 aging and has been associated to mortality, diseases and the rate of senescence in
685 several organisms (Canela et al. 2007; López-Otín et al. 2013; Dantzer and Fletcher
686 2015; Foley et al. 2018; Whittemore et al. 2019). In *D. melanogaster*, three TE families
687 exclusively found in telomeres comprise most of their DNA-content, and are further
688 essential for their enzymatic maintenance (Casacuberta 2017). In all four studies, we
689 found that long-lived populations had on average an increased genomic abundance
690 of *HeT-A*, which is thought to be the most common telomeric TE, although the
691 difference was less clear in two studies after correcting for TE content **(Fig. 3A** and
692 **Fig. S1)**. Moreover, the very few TEs transcriptionally upregulated in males and
693 females of long-lived populations in Carnes2015 were almost exclusively telomeric
694 elements (**Fig. 4B**). Our study therefore provides the first indication of a genomic and
695 transcriptional response of telomeric TEs to selection for postponed reproduction and
696 longevity. Previous studies, however, failed to establish a connection between
697 telomere length and lifespan. For instance, telomere length was not associated with

698 survival in *D. melanogaster* and *C. elegans*, but might affect other traits such as
699 fecundity (Raices et al. 2005; Walter et al. 2007). Moreover, in several species the
700 rate of telomere shortening rather than the initial length itself was a better predictor for
701 lifespan (Whittemore et al. 2019). Another complication yet to be addressed in
702 *Drosophila* concerns the phenomenon of 'intergenerational plasticity' of telomere
703 length observed in several mammals including humans, where increased paternal age
704 at reproduction causes longer telomeres in offspring for at least two generations
705 (Eisenberg et al. 2012; Eisenberg and Kuzawa 2018). The exact impact of telomere
706 length on evolutionary fitness and aging therefore remains to be poorly understood.
707
708 **Is TE expression detrimental for longevity?**
709 At the transcriptional level, age-dependent TE misregulation thought to be resulting
710 from a gradual decline in heterochromatin maintenance has been proposed to be
711 harmful for lifespan in *Drosophila* (Chen et al. 2016; Wood et al. 2016; Brown and
712 Bachtrog 2017), mice (De Cecco et al. 2019), and humans (Bogu et al. 2019). Our
713 RNA-seq analysis in whole flies did not identify a TE-wide decrease in expression in
714 selected populations of Carnes2015 (**Fig. 4** and **Table S7**). In fact, TE expression
715 appeared to be more strongly influenced by sex and age compared to selection
716 regime. Nevertheless, TEs with a significant regime effect tended to be downregulated
717 in selected populations, and this effect was even more apparent after we corrected for
718 genomic copies. Our results are in line with the notion that transcription of several TEs
719 is detrimental and reducing their expression could be beneficial for longevity (Li et al.
720 2013; Wood et al. 2016; Guo et al. 2018). Together with the weak association between
721 genomic abundance and expression (**Fig. S10**), this further proposes that lowering
722 expression of TEs might be more important than purging TEs from the genome in the
723 evolution of longevity.
724 Interestingly, the tendency towards lower TE expression in late-breeding populations
725 as well as the age-related increase in expression with age was more pronounced in
726 male flies. Almost all TEs had a higher expression in males relative to females (**Fig. 4**
727 and **Fig. S7**). These findings are consistent with recent work showing that males suffer
728 more from TE derepression during aging due to their entirely repetitive,
729 heterochromatin-rich *Y* chromosome (Brown and Bachtrog 2017). If a systemically
730 increased TE expression in males contributes to sexual dimorphism in lifespan is yet
731 to be confirmed. DNA-sequencing of male flies in the four experimental evolution

732 studies would be necessary to determine if selection for postponed senescence had
733 similarly strong effects on TE copy number of the *Y* chromosome.
734
735 **Did selection lead to differentiation in genes related to TE regulation?**
736 We also hypothesized that potential detrimental effects of TEs on longevity should be
737 reflected by selection on genes related to TE regulation and transposition (**Table 2**
738 and **Table S11**). Across the four studies, 3 to 10 out of 96 TE regulation genes have
739 been reported as genetically differentiated candidates, of which only *E2f1* and *Hsp83*
740 were shared between two studies. There was further little transcriptional differentiation
741 between control and selected populations of Carnes2015 and Remolina2012 and no
742 overlaps of these genes have been observed. The absence of a strong, shared signal
743 of selection on these genes and missing GO enrichment associated to regulation of
744 TEs allows us to hypothesize that improvement of chromatin
745 structure/heterochromatin maintenance, piRNA-mediated silencing and modulators of
746 transposition are generally not prime targets of selection during the evolution of
747 longevity. This, however, does not preclude that other means of TE protection have
748 evolved. It is becoming increasingly evident that TE expression acts as a causative
749 agent of inflammation and immune activation in mammals (Kassiotis and Stoye 2016;
750 De Cecco et al. 2019). Interestingly, Carnes2015, Fabian2018, and Remolina2012 all
751 found significant divergence in innate immunity genes, whereas Fabian et al. (2018)
752 demonstrated an improved survival upon infection and alleviated immunosenescence
753 in the long-lived populations. Rather than reducing TE copy number and expression,
754 selection might preferentially act on immunity genes to reduce TE-mediated
755 inflammation and increase tolerance to TEs with improved lifespan as a consequence.
756 It remains yet to be explored to what degree innate immune pathways other than the
757 RNAi machinery contribute to TE regulation in *D. melanogaster*.
758
759 **Is reproduction at old age associated with an increased TE content?**
760 Our findings suggest that neither genetic drift nor pervasive selection on TEs or genes
761 related to TE regulation are predominant drivers of the differences in TE abundance.
762 The most parsimonious explanation for our results therefore is that postponed
763 reproduction increases the chance of many TEs to be inserted into the germline and
764 passed on to the next generation. TEs of low transpositional activity in particular might
765 need a prolonged chronological time offered by late-life reproduction to achieve a

successful genomic insertion (**Figure 2**). Over many generations, flies breeding at old age would have accumulated more TEs in the genome than populations reproducing early in life. Supporting this hypothesis, it has been demonstrated that most TE families had a higher rate of insertions in the ovaries of older relative to young *P-element* induced dysgenic hybrids, even though at the same time fertility was restored and improved with age (Khurana et al. 2011). However, if this applies to non-dysgenic fruit flies and whether it can result in a larger number of TEs over multiple generations has to our knowledge not yet been observed. Thus, TE accumulation in late-breeding populations could be similar to the regularly observed positive correlation between parental age and number of *de novo* mutations in offspring (Goldmann et al. 2019; Sasani et al. 2019). In line with this, genome-wide measures of nucleotide diversity were also repeatably larger in late-breeding populations across four experiments (**Table S4**), although we have not ruled out that this was driven by genetic drift or balancing selection as proposed by one study (Michalak et al. 2017).

Opposing our hypothesis, two recent studies in termites (Elsner et al. 2018) and *D. melanogaster* (Erwin and Blumenstiel 2019) suggest that the germline is protected from TE invasions through increased transcription of the piRNA machinery. Indeed, our expression analysis confirms that many genes associated with transcriptional and post-transcriptional TE silencing tend to be upregulated with age. Despite this, TE families had a generally higher copy number in populations reproducing late in life. It therefore remains to be determined whether this age-dependent upregulation of TE regulation genes really equates to reduced insertional activity, since potential and realized TE repression might not necessarily match. The observation that these genes also tended to be more expressed in controls relative to selected flies in Carnes2015 further poses the question whether there is a trade-off between TE silencing in the germline and lifespan, which could be another mechanism explaining the rising TE abundance in the genomes of long-lived flies.

Altogether, our work presents a novel viewpoint on the poorly understood role of TEs in aging and longevity that is largely, but not exclusively, neutral. However, the caveat remains that we are unable to rule out that survival of selected populations would be further extended if they had a reduced TE content and expression. In-depth studies tracking piRNA production in the germline together with direct measures of TE transposition rates throughout life or measuring longevity upon knockdown and

800 overexpression of TEs would be crucial experiments to obtain a more complete
801 picture.
802

803 MATERIALS & METHODS

804 Datasets

805 We utilized genomic data from four independent studies performing laboratory
806 selection for postponed reproduction on wild-derived replicate populations by only
807 allowing flies of relatively old age to contribute to subsequent generations, whereas
808 controls reproduced early in life (Remolina et al. 2012; Carnes et al. 2015; Fabian et
809 al. 2018; Hoedjes et al. 2019) (**Table S1**). The experimental designs of the studies
810 were overall comparable, but notable differences include the mode of selection,
811 maintenance of controls, variable source populations, number of replicate populations
812 and generations at sequencing. Moreover, Hoedjes2019 performed the selection for
813 postponed senescence on three varying larval diets ranging from low to high
814 sugar/protein content. The genomic analysis was based on available raw fastq files
815 from whole-genome pool-sequencing of 100 to 250 females. RNA-seq data from
816 Carnes et al. (2015) consisted of raw fastq files from pools of 50 flies. The study
817 included transcriptomes of all selected and control populations, for which both sexes
818 at two ages 3-5 days (young) and 26-35 days of age (old) have been sequenced in
819 replicates. Microarray expression data from Remolina et al. (2012) are derived from
820 heads and abdomens from females at the age of 1, 5, 15, 30, and 50 days of age from
821 the three control and selected populations. See methods in the publications of each
822 study for details on experimental design and **Table S1** for a summary. For simplicity,
823 we refer to Carnes et al. (2015) as Carnes2015, Fabian et al. (2018) as Fabian2018,
824 Hoedjes et al. (2019) as Hoedjes2019, and Remolina et al. (2012) as Remolina2012
825 throughout this report.
826

827 Genome-wide TE abundance

828 To quantify the number of genomic insertions for each TE family in selected and
829 control populations we used the assembly-free tool deviaTE (Weilguny and Kofler
830 2019). In brief, deviaTE maps raw reads to a library of known TE sequences and
831 normalizes the sequencing depth of each position within a TE to the average depth of
832 multiple single-copy genes. The obtained distribution of normalized values reflects

833 coverage fluctuations within a TE, where averaging over all positions of a TE gives the
834 mean abundance per haploid genome. We used raw fastq files as input and ran
835 deviaTE with options: --read_type phred+33 (phred+64 for Fabian2018), --
836 min_read_len 50, --quality_threshold 18, --hq_threshold 20, --min_alignment_len 30,
837 --threads 20, --families ALL, --single_copy_genes
838 Dmel_rpl32,Dmel_RpII140,Dmel_Act5C,Dmel_piwi,Dmel_p53. TE abundance was
839 estimated considering reads with a mapping quality of >=20 by adding the "hq_cov"
840 and "phys_cov" column from the deviaTE output. Due to sequence similarity within
841 and between TEs, we further repeated most analyses using a less conservative
842 mapping approach without filtering for mapping quality by using values from the "cov"
843 column instead of "hq_cov". We restricted our downstream analysis to TEs that had a
844 study-average of >=0.5 insertions for at least 80% of the positions within a TE
845 sequence, thereby excluding all TEs with very low abundance and potentially wrongly
846 mapped reads.
847 We then investigated if the number of genomic abundance per TE varies between
848 control and selected populations using three different approaches (**Fig. S1**). First, we
849 tested each study (and diet regime for Hoedjes2019) independently and fit a model for
850 every detected TE as: *NormalizedCoverage ~ Regime + Population[Regime]*, where
851 *NormalizedCoverage* are the normalized coverage values reflecting insertion
852 estimates of each position within a TE, allowing us to take the variance of insertion
853 abundance within a TE into account. *Regime* is the effect of selection regime (levels:
854 selected and control), and *Population[Regime]* is the effect of the different replicate
855 populations nested within Regime. To obtain the overall effect of breeding regime in
856 Hoedjes2019, we fit: *NormalizedCoverage ~ Diet + Regime + Diet x Regime*, where
857 *Diet* corresponds to the three larval diets the flies were raised on (low, medium, high).
858 To correct for multiple testing and to define significant differences in TE abundance,
859 we applied a Bonferroni cut-off at α = 0.01.
860 To visualize differences in genomic abundance of TEs shown in **Fig. 1**, we first
861 calculated the average number of insertions per TE for each breeding regime. We then
862 took the absolute difference per TE by subtracting control from selected insertion
863 values (δInsertions, upper panel in **Fig. 1**). We used Pearson's correlation to show
864 that δInsertions from each TE were expectedly correlated to its copy number in the
865 genomes. To correct for this, we divided the selected by control number of insertions

866 per TE and took the log_2 of these values to obtain the log_2 fold change (log_2 (S / C),
867 lower panel in **Fig. 1**). We further used *SuperExactTest* (Wang et al. 2015) to analyze
868 if the overlap of TEs with a significantly higher genomic abundance in selected ("S>C")
869 or control populations ("C>S") between postponed senescence studies is expected by
870 chance.
871 In our second approach, we combined all four studies and only considered a single
872 average insertion value per TE instead of taking all positions into account. We then
873 summed up all TE insertion values within a population to obtain the total genomic TE
874 content. Because the total insertion TE content was significantly different between
875 studies (see **Fig. S3** and methods below), we normalized the insertion counts of each
876 TE by the total insertions in the corresponding population, and fit a combined model
877 as: asin(sqrt(*NormalizedInsertions*) ~ *Study* + *Regime* + *Regime*Study*, where
878 asin(sqrt(*NormalizedInsertions*) is the arcsine square root transformed proportion of
879 insertions per TE relative to the total genome-wide TE content in each population. The
880 *Study* term had four levels (Carnes2015, Fabian2018, Hoedjes2019, Remolina2012),
881 whereas *Regime* is as described above. As not all positions within a TE are considered
882 in this model, statistical power is lower than in the previous approach above. We
883 therefore performed a less stringent multiple testing correction by employing an FDR
884 cut-off of 0.05 using the "p.adjust" function with method "fdr" in *R*.
885 Third, we identified all TEs that showed a consistent increase or decrease in the
886 genomic abundance within all selected populations relative to all control populations
887 of a given study and within diet regimes of Hoedjes2019.
888 To analyze if the total genomic TE content varies between control and selected
889 populations, we calculated the total TE load across the whole genome, and TE
890 subclasses of Long Terminal Repeat (LTR), Non-Long Terminal Repeat (non-LTR),
891 and Terminal Inverted Repeat (TIR) by summing up all insertion values of each TE
892 family per population. We then used these sums as dependent variables in a statistical
893 model with *Study*, *Regime* and the *Study x Regime* interaction as factors.
894 All statistics were done in *R* using in-built functions unless otherwise stated. We did
895 not consider the effect of diet in the models combining all studies.
896
897 **Genomic TE locations and approximate mobility**
898 To identify the exact genomic positions and population frequency of TE insertions, we
899 used the program PoPoolationTE2 (Kofler et al. 2016). We first quality trimmed paired-

900 end raw reads using cutadapt (v.2.0) (Martin 2011) with options: --minimum-length 50
901 -q 18 and also --quality-base=64 for Fabian2018. Next, we masked the *D.*
902 *melanogaster* reference (v.6.27) for TEs present in the deviaTE library using
903 RepeatMasker (v.4.0.9) (Smit et al. 1996) with the RMBlast (v.2.9.0) search engine
904 and set options as recommended in the PopoolationTE2 manual (-gccalc -s -cutoff
905 200 -no_is -nolow -norna -gff -u -pa 18). Trimmed reads were then mapped against
906 the masked reference genome using bwa bwasw (v.0.7.17) (Li and Durbin 2009). We
907 restored paired-end information as outlined in the PopoolationTE2 pipeline (function
908 *se2pe*) and generated a ppileup file (function *ppileup*) applying a mapping quality cut-
909 off of 15. TE insertions were then identified using the functions *identifySignatures* (--
910 signature-window fix500, --min-valley fix150; and due to differences in the number of
911 replicate populations we chose --min-count 3 for Remolina2012 and Fabian2018, --
912 min-count 5 for Carnes2015, or --min-count 12 for Hoedjes2019). Finally, we applied
913 the *frequency* and *pairupSignatures* (--min-distance -200 --max-distance 300)
914 functions and filtered identified TEs using *filterSignatures* (--max-otherte-count 2 --
915 max-structvar-count 2) to create the final set of TE insertions. We only included sites
916 covered in all populations within a study and restricted our analysis to chromosomes
917 *X*, *2, 3,* and *4*. For each TE family, we calculated the average population frequency
918 across all of its identified genomic locations within a population as a proxy for active
919 or recent transposition events (Kofler et al. 2015). We used Spearman's correlation
920 analysis to compare average frequency values of each study to average frequencies
921 from a natural South African (SA) population sequenced to a high genomic coverage
922 (Kofler et al. 2015), and to correlated TE abundance with average frequency. We
923 employed t-tests to analyze if average population frequency from the SA population
924 varies between TEs more abundant in selected or controls, and also performed this
925 analysis using only the top10 TEs with the largest $\log_2$ FC values of abundance
926 change.
927
928 **Genome-wide nucleotide diversity**
929 We mapped trimmed paired-end reads against a fasta file including the repeat-masked
930 v6.27 reference genome (see above), TE library from deviaTE (Weilguny and Kofler
931 2019), *Wolbachia pipientis* (NC_002978.6), and two common gut bacteria *Acetobacter*
932 *pasteurianus* (AP011121.1), and *Lactobacillus brevis* (CP000416.1) using bwa *mem*
933 (v.0.7.17) (Li and Durbin 2009), and then created sorted bam files with the samtools

934 (v.1.9) (Li et al. 2009) functions *view* (-Sb) followed by *sort*. Duplicates were removed
935 using PicardTools (v.2.18.27) function *MarkDuplicates* with
936 REMOVE_DUPLICATES=true. For Remolina2012, we merged files from the same
937 population sequenced on multiple lanes using samtools *merge*. We then filtered and
938 created pileup files using samtools *mpileup* with option -q 20 --rf 0x2 --ff 0x4 --ff 0x8 -
939 f (and -6 for Fabian2018). Average coverage per study was determined only using
940 major chromosomal arms. We detected reads mapping to the genome of the
941 intracellular bacterium *Wolbachia* for all populations. To calculate nucleotide diversity
942 π and Watterson's θ across non-overlapping 100kb windows we used *Variance-*
943 *sliding.pl* implemented in PoPoolation (Kofler et al. 2011) with parameters --min-count
944 2 --min-coverage 5 --window-size 100000 --step-size 100000 --pool-size 100 --
945 measure pi (or theta) --fastq-type sanger --min-covered-fraction 0.6. We set the
946 maximum coverage threshold (--max-coverage) to two times the average genome
947 coverage of each population. The range of cut-offs used in each study was 20-74 in
948 Carnes2015, 286-355 in Fabian2018, 67-107 in Remolina2012 and 166-228 in
949 Hoedjes2019. We then analyzed variation in π and θ using ANOVA models including
950 the factors *Chromosome* (*X, 2L, 2R, 3L, 3R, 4*), *Diet* (low, medium, high), *Regime*
951 (control, selected) and the *Diet x Regime* interaction for Hoedjes2019, and
952 *Population[Regime]*, *Chromosome*, and *Regime* for all other studies.
953
954 **Genetic drift simulations**
955 As the analyzed studies did not include data from the ancestral populations, we used
956 average TE frequencies from the South African population (Kofler et al. 2015) as a
957 starting point in our simulations. To simulate frequency change of each TE, we set
958 population sizes (N), generation times and number of replicate populations as
959 mentioned in the original publications (**Table S1**). As some studies reported a range
960 of population sizes, we performed simulations using the lower and upper limits. We
961 used the *rbinom* function in *R* and drew from a size of 2N considering the diploid
962 genome and given the average frequency of a TE as probability to be drawn. The
963 number of successful draws was divided by 2N to obtain the new TE frequency and
964 used as input for the next draw. We repeated this process until the generation times
965 at sequencing of control and selected populations were reached. Using the simulated
966 TE frequencies in the last generation, we calculated average TE frequency across all

967 simulated replicates within breeding regime. Next, we obtained the proportion of TEs
968 with a higher (S>C) and lower frequency (C>S) in selected populations and calculated
969 the $\log_2$ relative proportion (i.e. $\log_2$ of S>C proportion divided by C>S proportion). As
970 an additional approach, we also acquired the $\log_2$ relative proportion by using the
971 number of TEs with consistent differences in frequency between all the populations of
972 the regimes instead of averaging. We performed 5,000 simulations to get a distribution
973 of relative proportions expected under genetic drift. P-values were calculated by
974 dividing the number of simulations that resulted in a larger or equal proportion as
975 observed in the actual analysis by the total number of simulations.
976
977 **TE frequency differences**
978 We further asked if certain genomic TE insertions are putatively involved in lifespan
979 and aging as might be indicated by a consistent TE frequency differences at any
980 genomic site between selected and control populations. We used the identified TE
981 insertions and frequency estimates from PopoolationTE2 (see above). Significant
982 difference in frequency between selected and control populations were determined by
983 t-tests analyzing differences in arcsine square root transformed frequencies between
984 regimes for Carnes2015, Fabian2018, and Remolina2012. For Hoedjes2019, we used
985 a linear model with *Diet*, *Regime*, and *Diet x Regime* as factors. Bonferroni correction
986 at $\alpha = 0.05$ was used to correct for multiple testing. Functional annotations were
987 supplemented using SnpEff (v.4.0e, Cingolani et al. 2012) and the *D. melanogaster*
988 reference v6.27 considering TE insertions within 1000 bp of the 5' and 3' UTR as
989 upstream or downstream of a gene.
990 We analyzed if each TE family varies in frequency between regimes by fitting the
991 factors of *Diet*, *Regime, and Diet x Regime* for Hoedjes2019, or *Regime* and
992 *Population[Regime]* for all other studies on arcsine square root transformed insertion
993 site frequencies. FDR values were obtained by using "p.adjust" in *R* and TE families
994 considered significant at FDR<0.05.
995
996 **RNA-seq analysis**
997 RNA-seq data from Carnes et al. (2015) consisted of two replicates of young and old
998 males and females from all control and selected populations (**Table S1**). Raw reads
999 were filtered to remove remaining adapter sequences and to exclude low quality

1000 positions using cutadapt (v.2.0) (Martin 2011) with options: -a AGATCGGAAGAGC --
1001 minimum-length 75 -q 20. Filtered reads were then mapped to the repeat-masked
1002 reference genome (see above), the TE library from deviaTE, *Wolbachia pipientis*,
1003 *Acetobacter pasteurianus*, and *Lactobacillus brevis* (see above) using STAR (Dobin
1004 et al. 2013) with options --alignIntronMin 23 --alignIntronMax 268107 –-
1005 outFilterMultimapNmax 10 --outReadsUnmapped FastX --outSAMstrandField
1006 intronMotif --outSAMtype BAM SortedByCoordinate. Read counts were obtained
1007 using the command-line version of featureCounts (Liao et al. 2013) with options -t
1008 exon -g gene_id --extraAttributes gene_symbol. We next pre-filtered the read count
1009 data by excluding all genes and TEs that did not have a sum of 400 counts across all
1010 80 samples (i.e. on average 5 counts per sample). Five TEs that are not known to
1011 occur in *D. melanogaster* passed this filter and were excluded. For simplicity, the
1012 analysis was performed on average read counts from two replicates, as all replicates
1013 were highly significantly correlated (Pearson's r ranging from 0.95 to 1, significant after
1014 Bonferroni correction). To analyze differential expression, we fit two models using read
1015 counts of genes and TEs with DESEq2 in *R* (Love et al. 2014). First, a full model with
1016 *Regime* (selected vs control), *Sex* (male vs female), *Age* (young vs old) and all
1017 interactions were fit. As the sex term was significant for most TEs, we decided to
1018 analyze males and females separately and by fitting models with *Regime*, *Age* and
1019 the *Regime x Age* interaction. We obtained $\log_2$ fold change values for each factor and
1020 the normalized read counts from DEseq2 for further analysis. To correct for the effect
1021 of copy number on expression, we divided read counts of TEs by the number of
1022 insertions observed in each population, assuming that genes and 13 TEs that did not
1023 pass out filters in the genomic analysis have a single copy in the genome.
1024
1025 **Evolution of TE regulation genes**
1026 The list of genes involved in TE regulation consisted of piRNA pathway genes also
1027 analyzed in Erwin and Blumenstiel 2019 and Elsner et al. 2018, and genes involved
1028 in heterochromatic and chromatin structure from Lee and Karpen 2017. We further
1029 added 7 additional genes involved in these functions, and genes annotated to
1030 "regulation of transposition" (GO:0010528) and "transposition" (GO:0032196)
1031 according to FlyBase so that we ended up with a total of 96 genes (**Table S11**). We
1032 then screened the published genomic candidate gene lists from Carnes2015,
1033 Fabian2018, Hoedjes2019 and Remolina2012 for these genes. We also compared TE

regulation genes to differentially expressed genes from the RNA-seq analysis of Carnes2015 (see above). We further obtained normalized microarray expression data from Remolina2012 of female flies at 1, 5, 15, 30, and 50 days of age (**Table S1**). Notably, the expression data were created from flies at 40 generations of selection compared to 50 generations in the genomic analysis. We fit a mixed effects model similar to the one used in the original publication with *Age*, *Regime*, and *Age x Regime* as fixed and *replication within population-age combination* as random effect. The two available tissues (heads and abdomens) were analyzed separately. A gene was considered to be differentially expressed if it had an FDR < 0.05 unless otherwise stated.

DATA ACCESSIBILITY

All data used in this study have been previously published. Accession numbers to the raw genomic and transcriptomic data can be found in **Table S1** and in the original studies (Remolina et al. 2012; Carnes et al. 2015; Fabian et al. 2018; Hoedjes et al. 2019). RNA-seq data were obtained directly from the authors. Additional raw data and results files on TE abundance, TE positions and frequency differentiation, nucleotide diversity, RNA-seq results for genes and TEs, and microarray results are available at: *to be made available upon publication*.

ACKNOWLEDGEMENTS

We thank Robert Kofler, Andrea Betancourt, Frank Jiggins, Lukas Weilguny and Suse Franssen for helpful comments and 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.

1068 Bogu GK, Reverter F, Marti-Renom MA, Snyder MP, Guigó R. 2019. Atlas of
1069 transcriptionally active transposable elements in human adult tissues.
1070 *unpublished data*. www.biorxiv.org/content/10.1101/714212v1
1071 Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran J V., Kazazian
1072 HH. 2003. Hot L1s account for the bulk of retrotransposition in the human
1073 population. *Proc Natl Acad Sci U S A*. 100:5280–5285.
1074 Brown EJ, Bachtrog D. 2017. The Y chromosome contributes to sex-specific aging in
1075 Drosophila. *unpublished data*. www.biorxiv.org/content/10.1101/156042v1
1076 Canela A, Vera E, Klatt P, Blasco MA. 2007. High-throughput telomere length
1077 quantification by FISH and its application to human population studies. *Proc Natl*
1078 *Acad Sci U S A*. 104:5300–5305.
1079 Carnes MU, Campbell T, Huang W, Butler DG, Carbone MA, Duncan LH, Harbajan
1080 S V., King EM, Peterson KR, Weitzel A, et al. 2015. The genomic basis of
1081 postponed senescence in Drosophila melanogaster. *PLoS One*. 10:e0138569.
1082 Carr M, Bensasson D, Bergman CM. 2012. Evolutionary genomics of transposable
1083 elements in Saccharomyces cerevisiae. *PLoS One*. 7:e50978.
1084 Casacuberta E. 2017. Drosophila: Retrotransposons making up telomeres. *Viruses*.
1085 9:E192.
1086 De Cecco M, Criscione SW, Peckham EJ, Hillenmeyer S, Hamm EA, Manivannan J,
1087 Peterson AL, Kreiling JA, Neretti N, Sedivy JM. 2013. Genomes of replicatively
1088 senescent cells undergo global epigenetic changes leading to gene silencing
1089 and activation of transposable elements. *Aging Cell*. 12:247–256.
1090 De Cecco M, Ito T, Petrashen AP, Elias AE, Skvir NJ, Criscione SW, Caligiana A,
1091 Brocculi G, Adney EM, Boeke JD, et al. 2019. L1 drives IFN in senescent cells
1092 and promotes age-associated inflammation. *Nature*. 566:73–78.
1093 Chang CH, Chavan A, Palladino J, Wei X, Martins NMC, Santinello B, Chen CC,
1094 Erceg J, Beliveau BJ, Wu CT, et al. 2019. Islands of retroelements are major
1095 components of Drosophila centromeres. *PLoS Biol*. 17:e3000241.
1096 Charlesworth B, Charlesworth D. 1983. The population dynamics of transposable
1097 elements. *Genet Res*. 42:1–27.
1098 Chen H, Zheng X, Xiao D, Zheng Y. 2016. Age-associated de-repression of
1099 retrotransposons in the Drosophila fat body, its potential cause and
1100 consequence. *Aging Cell*. 15:542–552.
1101 Chen H, Zheng X, Zheng Y. 2014. Age-associated loss of lamin-B leads to systemic

1102 inflammation and gut hyperplasia. *Cell*. 159:829–843.
1103 Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden
1104 DM. 2012. A program for annotating and predicting the effects of single
1105 nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila
1106 melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)*. 6:80–92.
1107 Cridland JM, Macdonald SJ, Long AD, Thornton KR. 2013. Abundance and
1108 distribution of transposable elements in two drosophila QTL mapping resources.
1109 *Mol Biol Evol*. 30:2311–2327.
1110 Cui R, Medeiros T, Willemsen D, Leonardo L Iasi, Collier G, Graef M, Reichard M,
1111 Valenzano D. 2019. Relaxed selection limits lifespan by increasing mutation
1112 load. *Cell*. 178:385-399.e20.
1113 Daborn PJ, Yen JL, Bogwitz MR, Le Goff G, Feil E, Jeffers S, Tijet N, Perry T,
1114 Heckel D, Batterham P, et al. 2002. A single P450 allele associated with
1115 insecticide resistance in Drosophila. *Science*. 297:2253–2256.
1116 Dantzer B, Fletcher QE. 2015. Telomeres shorten more slowly in slow-aging wild
1117 animals than in fast-aging ones. *Exp Gerontol*. 71:38–47.
1118 Deniz Ö, Frost JM, Branco MR. 2019. Regulation of transposable elements by DNA
1119 modifications. *Nat Rev Genet*. 20:417–431.
1120 Dennis S, Sheth U, Feldman JL, English KA, Priess JR. 2012. C. elegans germ cells
1121 show temperature and age-dependent expression of Cer1, a Gypsy/Ty3-related
1122 retrotransposon. *PLoS Pathog*. 8:e1002591.
1123 Dimitri P, Junakovic N. 1999. Revising the selfish DNA hypothesis: New evidence on
1124 accumulation of transposable elements in heterochromatin. *Trends Genet*.
1125 15:123–124.
1126 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson
1127 M, Gingeras TR. 2013. STAR: Ultrafast universal RNA-seq aligner.
1128 *Bioinformatics*. 29:15–21.
1129 Eisenberg DTA, Hayes MG, Kuzawa CW. 2012. Delayed paternal age of
1130 reproduction in humans is associated with longer telomeres across two
1131 generations of descendants. *Proc Natl Acad Sci U S A*. 109:10251–10256.
1132 Eisenberg DTA, Kuzawa CW. 2018. The paternal age at conception effect on
1133 offspring telomere length: Mechanistic, comparative and adaptive perspectives.
1134 *Philos Trans R Soc B Biol Sci*. 373:pii: 20160442.
1135 Elsner D, Meusemann K, Korb J. 2018. Longevity and transposon defense, the case

1136 of termite reproductives. *Proc Natl Acad Sci U S A*. 115:5504–5509.
1137 Erwin AA, Blumenstiel JP. 2019. Aging in the Drosophila ovary: contrasting changes
1138 in the expression of the piRNA machinery and mitochondria but no global
1139 release of transposable elements. *BMC Genomics*. 20:305.
1140 Everett LJ, Huang W, Zhou S, Carbone MA, Lyman RF, Arya GH, Geisz MS, Ma J,
1141 Morgante F, Armour G, et al. 2019. Gene expression networks in the Drosophila
1142 genetic reference panel. *unpublished data*.
1143 www.biorxiv.org/content/10.1101/816579v1
1144 Fabian D, Flatt T. 2011. The evolution of aging. *Nat Educ Knowl*. 2:9.
1145 Fabian DK, Garschall K, Klepsatel P, Santos-Matos G, Sucena É, Kapun M,
1146 Lemaitre B, Schlötterer C, Arking R, Flatt T. 2018. Evolution of longevity
1147 improves immunity in Drosophila. *Evol Lett*. 2:567–579.
1148 Flatt T, Heyland A. 2011. Mechanisms of life history evolution: The genetics and
1149 physiology of life history traits and trade-offs. *Oxford Univ Press Oxford*.
1150 Foley NM, Hughes GM, Huang Z, Clarke M, Jebb D, Whelan C V., Petit EJ, Touzalin
1151 F, Farcy O, Jones G, et al. 2018. Growing old, yet staying young: The role of
1152 telomeres in bats' exceptional longevity. *Sci Adv*. 4:eaao0926.
1153 Fontana L, Partridge L, Longo VD. 2010. Extending healthy life span-from yeast to
1154 humans. *Science*. 328:321–326.
1155 Gems D, Partridge L. 2013. Genetics of longevity in model organisms: Debates and
1156 paradigm shifts. *Annu Rev Physiol*. 75:621–644.
1157 Goldmann J, Veltman J, Gilissen C. 2019. De novo mutations reflect development
1158 and aging of the human germline. *Trends Genet*. 35:828–839.
1159 Graves JL, Hertweck KL, Phillips MA, Han M V., Cabral LG, Barter TT, Greer LF,
1160 Burke MK, Mueller LD, Rose MR, et al. 2017. Genomics of parallel experimental
1161 evolution in Drosophila. *Mol Biol Evol*. 34:831–842.
1162 Guo C, Jeong HH, Hsieh YC, Klein HU, Bennett DA, De Jager PL, Liu Z, Shulman
1163 JM. 2018. Tau activates transposable elements in Alzheimer's disease. *Cell*
1164 *Rep*. 23:2874–2880.
1165 Hancks DC, Kazazian HH. 2012. Active human retrotransposons: Variation and
1166 disease. *Curr Opin Genet Dev*. 22:191–203.
1167 Hoedjes KM, van den Heuvel J, Kapun M, Keller L, Flatt T, Zwaan BJ. 2019. Distinct
1168 genomic signals of lifespan and life history evolution in response to postponed
1169 reproduction and larval diet in Drosophila. *Evol Lett*.

1170 Jurka J, Bao W, Kojima KK. 2011. Families of transposable elements, population
1171 structure and the origin of species. *Biol Direct*. 6:44.
1172 Kalendar R, Tanskanen J, Immonen S, Nevo E, Schulman AH. 2000. Genome
1173 evolution of wild barley (Hordeum spontaneum) by BARE-1 retrotransposon
1174 dynamics in response to sharp microclimatic divergence. *Proc Natl Acad Sci U*
1175 *S A*. 97:6603–6607.
1176 Kaneko H, Dridi S, Tarallo V, Gelfand BD, Fowler BJ, Cho WG, Kleinman ME,
1177 Ponicsan SL, Hauswirth WW, Chiodo VA, et al. 2011. DICER1 deficit induces
1178 Alu RNA toxicity in age-related macular degeneration. *Nature*. 471:325–330.
1179 Kassiotis G, Stoye JP. 2016. Immune responses to endogenous retroelements:
1180 Taking the bad with the good. *Nat Rev Immunol*. 16:207–219.
1181 Khurana JS, Wang J, Xu J, Koppetsch BS, Thomson TC, Nowosielska A, Li C,
1182 Zamore PD, Weng Z, Theurkauf WE. 2011. Adaptation to P element transposon
1183 invasion in Drosophila melanogaster. *Cell*. 147:1551–1563.
1184 Kofler R. 2019. Dynamics of Transposable Element Invasions with piRNA Clusters.
1185 *Mol Biol Evol*. 36:1457–1472.
1186 Kofler R, Gómez-Sánchez D, Schlötterer C. 2016. PoPoolationTE2: Comparative
1187 population genomics of transposable elements using pool-seq. *Mol Biol Evol*.
1188 33:2759–2764.
1189 Kofler R, Nolte V, Schlötterer C. 2015. Tempo and mode of transposable element
1190 activity in Drosophila. *PLoS Genet*. 11:e1005406.
1191 Kofler R, Orozco-terWengel P, de Maio N, Pandey RV, Nolte V, Futschik A, Kosiol C,
1192 Schlötterer C. 2011. Popoolation: a toolbox for population genetic analysis of
1193 next generation sequencing data from pooled individuals. *PLoS One*. 6:e15925.
1194 Kofler R, Senti KA, Nolte V, Tobler R, Schlötterer C. 2018. Molecular dissection of a
1195 natural transposable element invasion. *Genome Res*. 28:824–835.
1196 de Koning APJ, Gu W, Castoe TA, Batzer MA, Pollock DD. 2011. Repetitive
1197 elements may comprise over two-thirds of the human genome. *PLoS Genet*.
1198 7:e1002384.
1199 Kreiner JM, Wright SI. 2018. A less selfish view of genome size evolution in maize.
1200 *PLoS Genet*. 14:e1007249.
1201 Krug L, Chatterjee N, Borges-Monroy R, Hearn S, Liao WW, Morrill K, Prazak L,
1202 Rozhkov N, Theodorou D, Hammell M, et al. 2017. Retrotransposon activation
1203 contributes to neurodegeneration in a Drosophila TDP-43 model of ALS. *PLoS*

1204 *Genet*. 13:e1006635.
1205 Kuhn A, Ong YM, Cheng C-Y, Wong TY, Quake SR, Burkholder WF. 2014. Linkage
1206 disequilibrium and signatures of positive selection around LINE-1
1207 retrotransposons in the human genome. *Proc Natl Acad Sci U S A*. 111:8131–
1208 8136.
1209 Lee YCG, Karpen GH. 2017. Pervasive epigenetic effects of Drosophila euchromatic
1210 transposable elements impact their evolution. *Elife*. 6:pii: e25762.
1211 Lerat E, Goubert C, Guirao-Rico S, Merenciano M, Dufour AB, Vieira C, González J.
1212 2019. Population-specific dynamics and selection patterns of transposable
1213 element insertions in European natural populations. *Mol Ecol*. 28:1506–1522.
1214 Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler
1215 transform. *Bioinformatics*. 25:1754–1760.
1216 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G,
1217 Durbin R. 2009. The Sequence Alignment/Map format and SAMtools.
1218 *Bioinformatics*. 25:2078–2079.
1219 Li W, Prazak L, Chatterjee N, Grüninger S, Krug L, Theodorou D, Dubnau J. 2013.
1220 Activation of transposable elements during aging and neuronal decline in
1221 Drosophila. *Nat Neurosci*. 16:529–531.
1222 Li ZW, Hou XH, Chen JF, Xu YC, Wu Q, Gonzalez J, Guo YL. 2018. Transposable
1223 elements contribute to the adaptation of Arabidopsis thaliana. *Genome Biol*
1224 *Evol*. 10:2140–2150.
1225 Liao Y, Smyth GK, Shi W. 2013. FeatureCounts: an efficient general-purpose read
1226 summarization program. *Bioinformatics*. 30:923–930.
1227 Lippman Z, Martienssen R. 2004. The role of RNA interference in heterochromatic
1228 silencing. *Nature*. 431:364–370.
1229 López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. 2013. The hallmarks
1230 of aging. *Cell*. 153:1194–1217.
1231 Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and
1232 dispersion for RNA-seq data with DESeq2. *Genome Biol*. 15:550.
1233 Luckinbill LR, Arking R, Clare MJ. 1984. Selection for delayed senescence in
1234 Drosophila melanogaster. *Evolution*. 38:996–1003.
1235 Magwire MM, Bayer F, Webster CL, Cao C, Jiggins FM. 2011. Successive increases
1236 in the resistance of Drosophila to viral infection through a transposon insertion
1237 followed by a duplication. *PLoS Genet*. 7:e1002337.

1238 Martienssen R, Moazed D. 2015. RNAi and heterochromatin assembly. *Cold Spring*
1239 *Harb Perspect Biol*. 7:a019323.
1240 Martin M. 2011. Cutadapt removes adapter sequences from high-throughput
1241 sequencing reads. *EMBnet.journal*. 17:10–12.
1242 Maxwell PH, Burhans WC, Curcio MJ. 2011. Retrotransposition is associated with
1243 genome instability during chronological aging. *Proc Natl Acad Sci U S A*.
1244 108:20376–20381.
1245 May CM, van den Heuvel J, Doroszuk A, Hoedjes KM, Flatt T, Zwaan BJ. 2019.
1246 Adaptation to developmental diet influences the response to selection on age at
1247 reproduction in the fruit fly. *J Evol Biol*. 32:425–437.
1248 McClintock B. 1950. The origin and behavior of mutable loci in maize. *Proc Natl*
1249 *Acad Sci U S A*. 36:344–355.
1250 Michalak P, Kang L, Sarup PM, Schou MF, Loeschcke V. 2017. Nucleotide diversity
1251 inflation as a genome-wide response to experimental lifespan extension in
1252 Drosophila melanogaster. *BMC Genomics*. 18:84.
1253 Mori MA, Raghavan P, Thomou T, Boucher J, Robida-Stubbs S, MacOtela Y,
1254 Russell SJ, Kirkland JL, Blackwell TK, Kahn CR. 2012. Role of microRNA
1255 processing in adipose tissue in stress defense and longevity. *Cell Metab*.
1256 16:336–3347.
1257 Pan H, Finkel T. 2017. Key proteins and pathways that regulate lifespan. *J Biol*
1258 *Chem*. 292:6452–6460.
1259 Petrov DA, Fiston-Lavier AS, Lipatov M, Lenkov K, González J. 2011. Population
1260 genomics of transposable elements in Drosophila melanogaster. *Mol Biol Evol*.
1261 28:1633–1644.
1262 Piper MDW, Selman C, McElwee JJ, Partridge L. 2008. Separating cause from
1263 effect: How does insulin/IGF signalling control lifespan in worms, flies and mice?
1264 *J Intern Med*. 263:179–191.
1265 Prudencio M, Gonzales PK, Cook CN, Gendron TF, Daughrity LM, Song Y, Ebbert
1266 MTW, van Blitterswijk M, Zhang YJ, Jansen-West K, et al. 2017. Repetitive
1267 element transcripts are elevated in the brain of C9orf72 ALS/FTLD patients.
1268 *Hum Mol Genet*. 26:3421–3431.
1269 Quesneville H, Bergman CM, Andrieu O, Autard D, Nouaud D, Ashburner M,
1270 Anxolabehere D. 2005. Combined evidence annotation of transposable
1271 elements in genome sequences. *PLoS Comput Biol*. 1:166–175.

1272 Raices M, Maruyama H, Dillin A, Kariseder J. 2005. Uncoupling of longevity and
1273 telomere length in C. elegans. *PLoS Genet*. 1:e30.
1274 Rech GE, Bogaerts-Márquez M, Barrón MG, Merenciano M, Villanueva-Cañas JL,
1275 Horváth V, Fiston-Lavier AS, Luyten I, Venkataram S, Quesneville H, et al.
1276 2019. Stress response, behavior, and development are shaped by transposable
1277 element-induced mutations in Drosophila. *PLoS Genet*. 15:e1007900.
1278 Remolina S, Chang P, Leips J. 2012. Genomic basis of aging and life history
1279 evolution in Drosophila melanogaster. *Evolution*. 66:3390–3403.
1280 Rose M. 1984. Laboratory evolution of postponed senescence in Drosophila
1281 melanogaster. *Evolution*. 38:1004–1010.
1282 Sasani TA, Pedersen BS, Gao Z, Baird L, Przeworski M, Jorde LB, Quinlan AR.
1283 2019. Large, three-generation human families reveal post-zygotic mosaicism
1284 and variability in germline mutation accumulation. *Elife*. 8:pii: e46922.
1285 Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J,
1286 Fulton L, Graves TA, et al. 2009. The B73 maize genome: Complexity, diversity,
1287 and dynamics. *Science*. 326:1112–1115.
1288 Smit A, Hubley R, Green P. 1996. RepeatMasker Open-3.0. www.repeatmasker.org
1289 Solyom S, Ewing AD, Rahrmann EP, Doucet T, Nelson HH, Burns MB, Harris RS,
1290 Sigmon DF, Casella A, Erlanger B, et al. 2012. Extensive somatic L1
1291 retrotransposition in colorectal tumors. *Genome Res*. 22:2328–2338.
1292 Sturm Á, Ivics Z, Vellai T. 2015. The mechanism of ageing: Primary role of
1293 transposable elements in genome disintegration. *Cell Mol Life Sci*. 72:1839–
1294 1847.
1295 Volkman HE, Stetson DB. 2014. The enemy within: Endogenous retroelements and
1296 autoimmune disease. *Nat Immunol*. 15:415–422.
1297 Walter MF, Biessmann MR, Benitez C, Török T, Mason JM, Biessmann H. 2007.
1298 Effects of telomere length in Drosophila melanogaster on life span, fecundity,
1299 and fertility. *Chromosoma*. 116:41–51.
1300 Wang M, Zhao Y, Zhang B. 2015. Efficient test and visualization of multi-set
1301 intersections. *Sci Rep*. 5:16923.
1302 Weilguny L, Kofler R. 2019. DeviaTE: Assembly-free analysis and visualization of
1303 mobile genetic element composition. *Mol Ecol Resour*. 19:1346–1354.
1304 Whittemore K, Vera E, Martínez-Nevado E, Sanpera C, Blasco MA. 2019. Telomere
1305 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.

1340 **FIGURE LEGENDS**

1341 **Figure 1. Difference in TE abundance between regimes.** Upper row: δinsertions (Y-
1342 axis) were obtained by subtracting the average genomic insertions of the controls from
1343 the selected populations. Bottom row: $\log_2$ fold change between selected and control
1344 populations. The dashed line indicates no difference between regimes. >0 denote TEs
1345 with a larger abundance in selected populations, while <0 TEs with more insertions in
1346 controls. TE subclasses are given in different colors. Selected flies had more genomic
1347 insertions than controls for most TE families (also see Table 1).

1348

1349 **Figure 2. Differences in average TE frequency.** Average TE frequency from the
1350 South African population separated into C>S (blue) and S>C TEs (red) are shown on
1351 the Y-axis. Differences are shown considering all C>S and S>C TEs ("All") or only the
1352 top 10 TEs with the biggest differences in $\log_2$ FC of insertions ("Top 10"). t-tests were
1353 used to assess statistical significance. ns, not significant; * P < 0.05; *** P < 0.001.

1354

1355 **Figure 3. Selection on TE abundance and insertions.** (A) Sharing of TEs with
1356 significant differences in abundance between regimes of each study. S>C and C>S
1357 denote TEs with a higher abundance in selected or control populations, respectively.
1358 Red bars indicate a significant overlap at P < 0.05. (B) Boxplots of the number of
1359 genomic insertions relative to the total genomic content of the 2 significantly shared
1360 C>S TEs. (C) Genome-wide differentiation in TE insertion frequency between selected
1361 and control populations in Fabian2018 and (D) Hoedjes2019. Every point indicates
1362 the $-\log_{10}$ P-value of a TE insertion across chromosomal arms (alternating black and
1363 grey color). The solid orange line corresponds to the Bonferroni cut-off at $\alpha = 0.05$.
1364 Red points denote TEs with a significantly higher frequency in selected relative to
1365 control populations. More details including exact positions, frequency and annotation
1366 of candidate TE insertions can be found in **Table S5**.

1367

1368 **Figure 4. Multiple factors influence TE expression.** (A) Proportions of differentially
1369 expressed TEs at adjusted P < 0.05 and directionality relative to 123 TEs with
1370 detectable expression for factors from statistical models on pre-filtered read counts in
1371 DESeq2 (also see **Table S7**). "Sex" refers to the results of the full model including Sex
1372 (M, males; F, females), Age (young; old), Regime (C, control; S, selected), and all

interactions. “Regime”, “Age” and “RxA” (i.e. *Regime x Age* interaction) refer to results from model fits with males and females separately analyzed. The absolute number of TEs for factor levels are given above or below bars. (B) Log_2 fold change of regime (selected vs control) and (C) age (young vs old) for males and females. Colors designate TEs significant only in males (blue), or females (red), or shared between both sexes (orange). Not significant TEs are in grey. (D and E) Log_2 fold changes across regime against age differences in males and females. Colors designate TEs significant only for regime (blue), or age (red), or for both factors (orange). Not significant TEs are in grey. (F) Normalized expression counts were significantly correlated to the number of insertions in the genome. Expression values for each TE were averaged across all conditions (sex, age, regime). Separating samples into different conditions resulted in similar significant correlation coefficients (**Table S8**). Colors correspond to TEs with a significantly higher genomic abundance in selected (red) or control populations (blue). (B to E): r, Pearson’s correlation coefficient; (F): r, Spearman’s correlation coefficient; * $P < 0.05$, *** $P < 0.0001$.

1407 **TABLES**

1408

1409 **Table 1.** Number of detected and significant TE families with percentage of TEs more
1410 abundant in selected (S>C) or controls (C>S) or not different between regimes (n.s.)
1411 from models analyzing studies independently or combined.

1412

Study	N	N (sign.)[a]	S>C	C>S	n.s.
Carnes2015	112	107	77%	19%	4%
Fabian2018	110	85	46%	31%	33%
Hoedjes2019	115	94	58%	24%	18%
Remolina2012	110	76	57%	12%	31%
Combined Model	103	Regime: 41 Study: 101 Study x Regime: 65	33%	7%	60%

1413 [a]Significant after Bonferroni correction at α=0.01 in single studies or FDR<0.05 in combined
1414 model.

1415

1416

1417

1418

1419

1420

1421

1422

1423

1424

1425

1426

1427

1428

1429

1430

1431

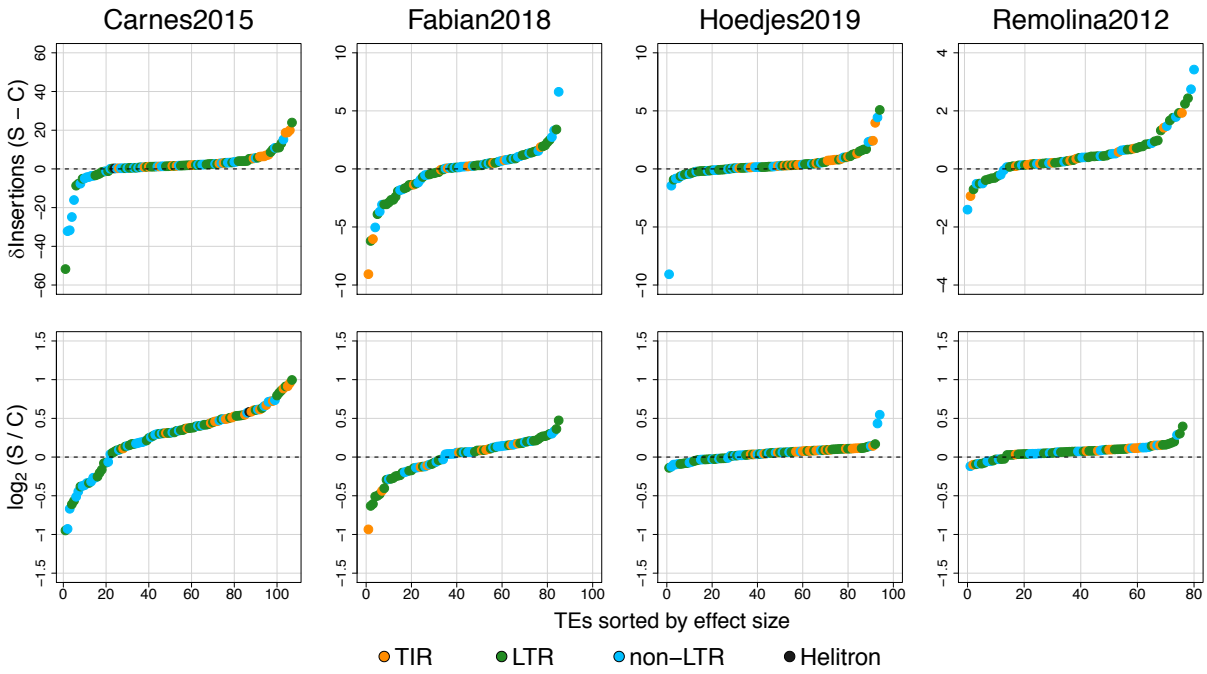
1432 **Table 2.** TE regulation genes identified by each study as genetically differentiated
1433 candidate genes for lifespan combined with differential expression in Carnes2015 and
1434 Remolina2012 (see **Table S11** for a full summary of all 96 genes).
1435

		Genetic Differentiation				Differential expression		
Gene	**Type**	**Carnes 2015**	**Fabian 2018**	**Hoedjes 2019**	**Remolina 2012**	**Direction of Expression[a]**	**Sex**	**Study**
E2f1	epigenetic	✓	✓	-	-	Old>Young	M	Carnes
Hsp83	transposition	✓	-	-	✓	Old>Young/RxA	M	Carnes
csul	piRNA	✓	-	-	-	Old>Young	M	Carnes
fs(1)Yb	piRNA	✓	-	-	-	-	-	-
G9a	epigenetic	✓	-	-	-	Old>Young	M	Carnes
nub	epigenetic	✓	-	-	-	Old>Young	M	Carnes
piwi	epigenetic & piRNA	✓	-	-	-	S>C	F	Carnes
Su(var)2-HP2	epigenetic	✓	-	-	-	Old>Young	M	Carnes
tej	piRNA	✓	-	-	-	Old>Young/RxA	F	Carnes
wapl	epigenetic	✓	-	-	-	Old>Young	F	Remolina
tj	piRNA	-	✓	-	-	Young>Old/RxA	F	Carnes
vas	piRNA	-	✓	-	-	Old>Young	F	Remolina
del	piRNA	-	-	✓	-	Old>Young	F	Remolina
Hen1	piRNA	-	-	✓	-	C>S/Old>Young/RxA	F	Carnes
kis	epigenetic	-	-	✓	-	Old>Young Young>Old/RxA	M F	Carnes
mino	piRNA	-	-	✓	-	C>S Old>Young/RxA	M&F M	Carnes
Nbr	piRNA	-	-	✓[b]	-	Old>Young	M&F	Carnes
qin	piRNA	-	-	✓	-	C>S Old>Young	M F	Carnes Remolina
Su(var)3-7	epigenetic	-	-	✓	-	Old>Young	F	Remolina
AGO3	piRNA	-	-	-	✓	Old>Young RxA	M&F F	Carnes
armi	piRNA	-	-	-	✓	-	-	-
CG8108	epigenetic	-	-	-	✓	Old>Young	F	Remolina
moon	piRNA	-	-	-	✓	-	-	-
nos	piRNA	-	-	-	✓	C>S Old>Young	M F	Carnes Remolina

1436 [a]Padj.<0.05 for *Age*, *Regime* and interaction of both (RxA), where “>” indicates larger. [b]*P-element*
1437 frequency divergence found in the present study.
1438
1439
1440
1441
1442

1443 **FIGURE 1**

1444



1445

1446

1447

1448

1449

1450

1451

1452

1453

1454

1455

1456

1457

1458

1459

1460

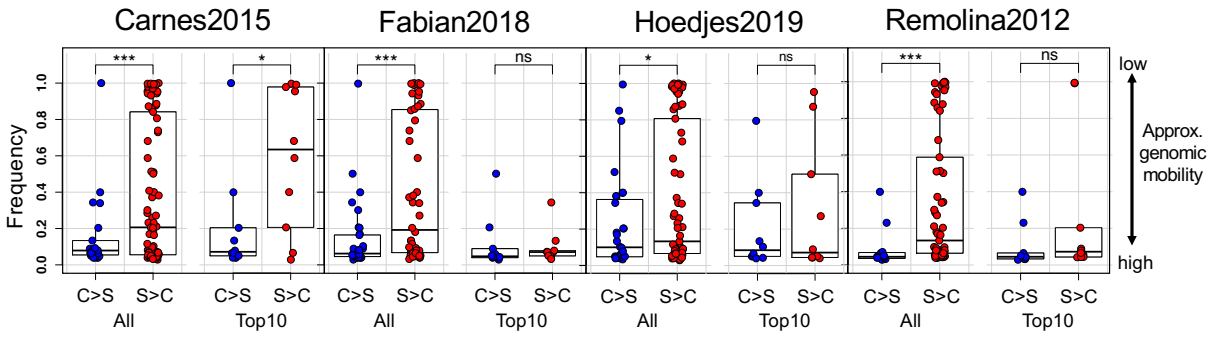
1461

1462

1463

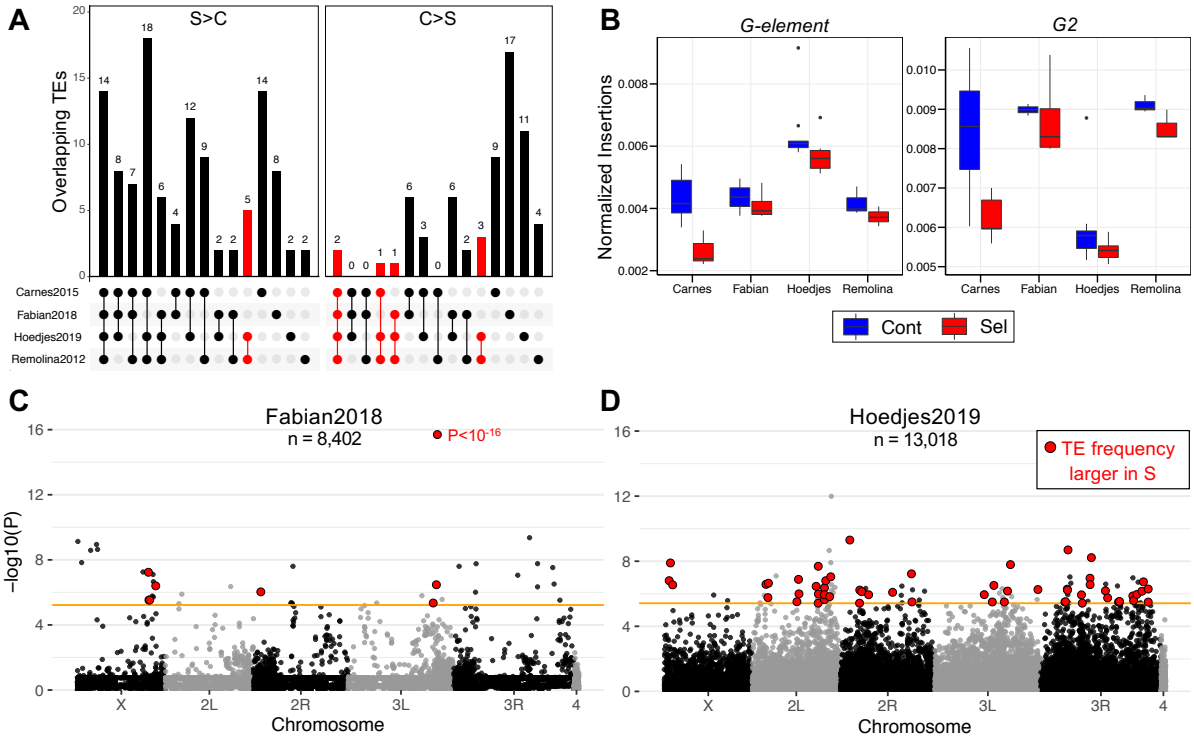
1464

FIGURE 2


Carnes2015
Fabian2018
Hoedjes2019
Remolina2012
Frequency
0.0
0.2
0.4
0.6
0.8
1.0
***
*
***
ns
*
ns
***
ns
C>S
S>C
All
Top10
low
high
Approx. genomic mobility

1493 **FIGURE 3**

1494



1495

1496

1497

1498

1499

1500

1501

1502

1503

1504

1505

1506

1507

1508

1509

1510

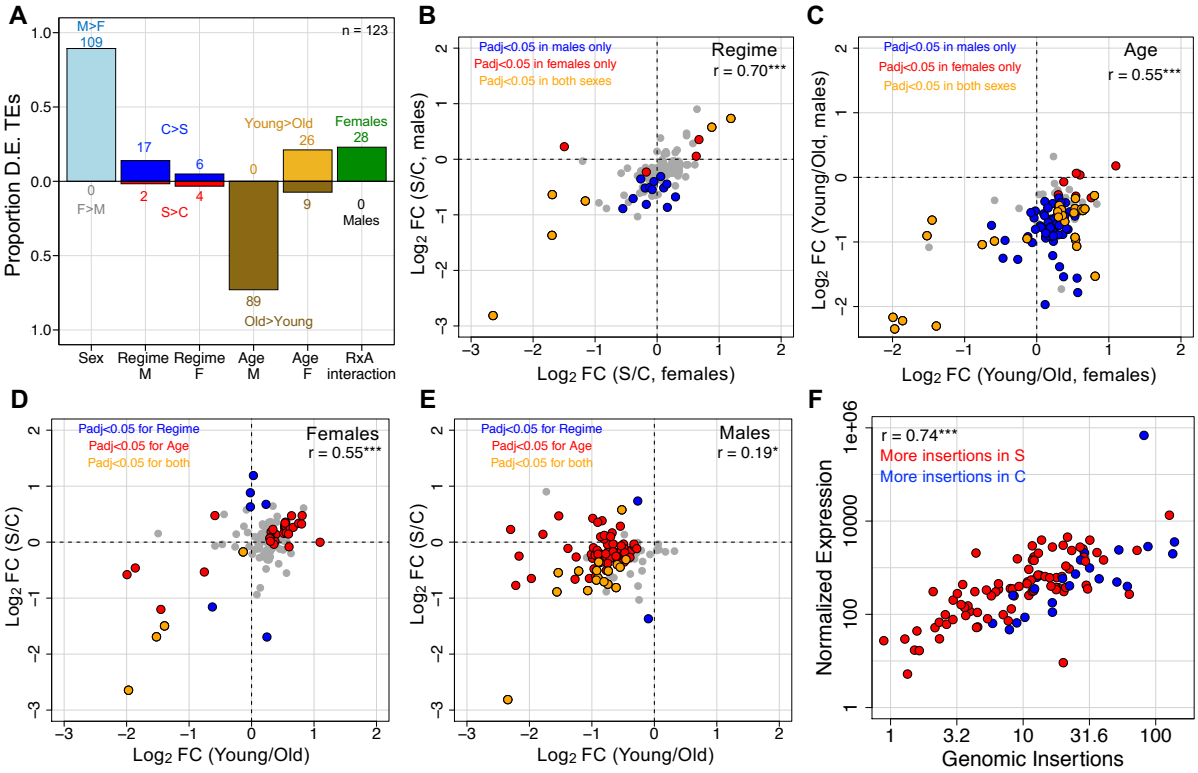
1511

1512

1513

1514 **FIGURE 4**

1515



1516