Long-term experimental evolution reveals purifying selection on piRNA-mediated control of transposable element expression

Ulfar Bergthorsson[*1], Sheeba Singh[*2,3], Anke Konrad[1], Tony Belicard[2,3], Toni Beltran[2,3], Vaishali Katju[^1], Peter Sarkies[^2,3]

^Joint corresponding
*co-first authors
[1] Department of Veterinary Integrative Biosciences, Texas A&M University, College Station, TX 77845, USA
[2] MRC London Institute of Medical Sciences, Du Cane Road, London W12 0NN UK
[3] Institute of Clinical Sciences, Imperial College London, Du Cane Road, London W12 0NN UK
^Vaishali Katju vkatju@cvm.tamu.edu
^ Peter Sarkies psarkies@imperial.ac.uk

Abstract

Transposable elements (TEs) are a key component of eukaryotic genomes. TEs can copy themselves independently of the host genome, thus are often considered to be selfish genomic elements. However, TE dynamics within genomes have contributed to adaptive evolution leading to speculation that natural selection preserves TE expression. Here we used experimental evolution of *C. elegans* to study the dynamics of TE expression over 400 generations in population sizes maintained at 1, 10 and 100 individuals. We demonstrate increased TE expression relative to the ancestral population, with the largest increases corresponding to smaller population sizes. Using high-throughput small RNA sequencing we show that the transcriptional activation of TEs within active regions of the genome is associated with loss of piRNA-mediated silencing, whilst desilenced TEs in repressed chromatin domains retain small RNAs. Additionally, we find that the sequence context of the surrounding region influences the propensity of TEs to lose silencing through failure of small RNA-mediated silencing. Together, our results show that natural selection in *C. elegans* is responsible for maintaining low levels of TE expression, and provide new insights into the epigenomic features responsible. [183 words]

Introduction

Transposable elements (TEs) are almost ubiquitous across eukaryotic genomes(Pritham, 2009). Their ability to replicate independently of the host genome, coupled with the existence of multiple copies liable to ectopic recombination means they present a potential threat to genome stability. Moreover, TEs pose a threat to genome function as new integrations can disrupt

47 genes or gene regulatory elements. As a result, organisms have evolved
48 sophisticated control strategies, which protect the genome from TE proliferation.
49
50 Across eukaryotes short (20-33 nucleotides) small RNAs are key to TE control.
51 Within animals, Piwi-interacting small RNAs (piRNAs) are paramount in the TE
52 defence armoury(Siomi et al., 2011). piRNAs are produced from defined
53 genomic loci named piRNA clusters and after processing, associate with the Piwi
54 subfamily of argonaute proteins(Brennecke et al., 2007). They recognise TEs
55 through sense-antisense base pairing and target TEs for transcriptional and
56 post-transcriptional silencing(Siomi et al., 2011). In many model organisms,
57 piRNAs are essential for fertility through their role in controlling TE proliferation
58 in the germline(Weick and Miska, 2014)
59
60 The nematode *Caenorhabditis elegans* is a well-established model for small-RNA
61 mediated silencing. piRNAs in *C. elegans* are unusual in that the two piRNA
62 clusters on Chromosome IV are composed of individual RNA polymerase II (RNA
63 pol II) transcription loci where each piRNA has its own upstream motif(Batista et
64 al., 2008; Das et al., 2008; Ruby et al., 2006; Wang and Reinke, 2008). piRNA
65 clusters are located within H3K27me3-rich chromatin, which, together with cis-
66 acting RNA pol II pausing sequences downstream of the piRNA, enforce
67 production of ~28 nucleotide piRNA precursors (Beltran et al., 2019[ref to be
68 inserted after pub]). piRNA precursors are further trimmed to result in mature
69 21 nucleotide piRNAs with a Uracil as the first nucleotide (21U-RNAs), which
70 associate with the *C. elegans* Piwi protein PRG-1(Batista et al., 2008; Das et al.,
71 2008; Wang and Reinke, 2008). Downstream of PRG-1, piRNA silencing relies on
72 a nematode-specific class of secondary small RNAs known as 22G-RNAs(Das et
73 al., 2008). 22G-RNA synthesis is carried out by RNA-dependent RNA
74 polymerases using the target RNA as a template, following initiation by piRNA
75 target recognition(Pak and Fire, 2007). 22G-RNAs bind to Argonaute proteins
76 and lead to transcriptional and post-transcriptional silencing of target
77 RNAs(Yigit et al., 2006). Additionally 22G-RNAs can be transmitted
78 transgenerationally(Buckley et al., 2012) and as a result piRNA-initiated
79 silencing can persist for many generations even after piRNAs themselves are

removed by mutating PRG-1(Ashe et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012). Consequently whilst removal of piRNAs alone has mild effects on TE expression, combining mutations of PRG-1 with mutations disrupting the 22G-RNA biogenesis machinery leads to reactivation of several TEs(de Albuquerque et al., 2015; Phillips et al., 2015)

Despite the threat posed by TEs to genome integrity, many events in adaptive evolution have been linked to TE insertions. Across a variety of species, individual examples have been documented whereby TE insertions can rewire gene expression leading to differences both within and between species that can be positively selected for. On a more widespread level, up to 60% of human-specific enhancers may be TE-derived(Rebollo et al., 2011) and TE insertions have been proposed to substantially rewire the human immune cell transcriptome(Imbeault et al., 2017). Such observations have led to the speculation that TEs are retained within eukaryotic genomes because of their potential contribution to "evolvability"(Fablet and Vieira, 2011). However, this idea is still unproven, and a strong null hypothesis is that TEs are generally detrimental and beneficial TE insertions appear overrepresented due to the effects of natural selection in weeding out deleterious insertions(Simonti et al., 2017)

In the context of the uncertainty surrounding the contribution of TE mobility and expression to organismal fitness, it is important to understand the extent to which the balance between TE expression and TE regulation is under selection. One way to study this is to use a mutation accumulation (MA) framework in which replicate lines descended from a single common ancestor are propagated under a regime of drastic population bottlenecks for several hundred generations(Halligan and Keightley, 2009; Katju and Bergthorsson, 2019). The maintenance of these lines at a minimal population size attenuates the efficacy of selection, thereby enabling the accumulation of a large, unbiased sample of spontaneous mutations under conditions of genetic drift which can subsequently be identified and their fitness effects investigated. Previous studies in single-celled organisms and multicellular model organisms have used the MA approach

113 to investigate the effect of mutations on protein-coding gene expression
114 divergence(Denver et al., 2005; Landry et al., 2007; Rifkin et al., 2005). Here, we
115 extend this approach to investigate the effect of mutations on TE expression
116 divergence.
117
118 We created spontaneous MA lines of *C. elegans* that were descended from a
119 single worm ancestor and propagated for ~400 generations under three
120 population size treatments of N = 1, 10 and 100 individuals per generation(Katju
121 et al., 2015). The varying population size treatment in the experiment permitted
122 a manipulation of the strength of selection, with the N =1 lines evolving under
123 close to neutral conditions (minimal selection) and an incremental increase in
124 the strength of selection with increasing population size. We employed this
125 framework to investigate how TE expression evolves under conditions of near
126 neutrality and under the influence of increasing selection intensity. We show
127 that overall TE expression increases in MA lines with the smallest population
128 size. We further show that expression increase results in part from failure of
129 piRNA-mediated silencing. Intriguingly, differences in the responses of different
130 TEs to reduced piRNA-mediated silencing depend on the chromatin environment
131 of the TE loci, such that TEs in repressed chromatin domains largely remain
132 silent due to epigenetic memory imparted by 22G-RNAs, whilst in active
133 chromatin domains, increased TE expression is much more likely to occur.
134 Together our results demonstrate for the first time that robust TE control is
135 under selection in animals. Importantly further, our results provide new insight
136 into how the chromatin environment interacts with piRNA-mediated silencing to
137 control TE expression.
138
139 **Results**
140
141 *Relaxed selection leads to increased TE expression*
142
143 In order to assess the effect of selection on TE expression, we generated
144 mutation accumulation (MA) lines and propagated them by randomly selecting N
145 individuals at each generation, where N was either 1, 10 or 100(Katju et al.,

146 2015). Henceforth we refer to the three conditions as *N*.1, *N*.10 or *N*.100 (Figure
147 1A). After 409 generations, we isolated RNA and performed RNA sequencing to
148 investigate TE expression. All three conditions showed an increase in total TE
149 expression relative to the pre-MA ancestral control; moreover, *N*.1 had higher
150 total TE expression than *N*.10 or *N*.100. Across increasing population sizes, we
151 observed a monotonic decrease in total TE expression (Jonckheere test for
152 ordered medians; henceforth Jh, $p = 0.019$; Figure 1B). Similarly, linear
153 regression analysis showed a significant negative correlation between increasing
154 population size and TE expression (Linear regression; henceforth LR $p = 0.023$
155 Figure 1B). Protein-coding gene expression diverges during MA in a variety of
156 model organisms including *C. elegans*(Denver et al., 2005; Rifkin et al., 2005;
157 Landry et al., 2007; Hodgins-Davis et al., 2015). To test whether this effect also
158 occurred for TEs in our dataset, we estimated the variation in the expression of
159 each individual TE and each individual gene. To control for potential changes in
160 the mean expression, which can affect noise, we calculated the Fano factor
161 [var(x)/mean(x)] within *N*.1, *N*.10 and *N*.100 lines separately. Fano factors for
162 TEs and genes were higher in *N*.1 lines than in *N*.10 or *N*.100 lines (TEs:
163 Wilcoxon paired test $p = 0.015$ and 0.004 for *N*.10 and *N*.100; genes: Wilcoxon
164 paired test $p < 1e^{-16}$ for both *N*.10 and *N*.100; Supplemental Figure 1A and B). To
165 control for the possibility that the larger number of *N*.1 lines might lead to higher
166 variance, we calculated TE fano factors from 1,000 subsets of five *N*.1 lines and
167 all 252 subsets of five *N*.10 lines and compared these to the five *N*.100 lines. This
168 showed the same trend as the full dataset (Supplemental Figure 1C). To further
169 investigate the variation in expression we calculated the total variance in the
170 change in expression of all TEs or all genes between each line and the pre-MA
171 ancestral control. Variance in the differences in both TE and gene expression
172 increased with smaller population sizes (Jh $p = 0.008$ and 0.009, respectively)
173 (Supplemental Figure 1 D,E). However, importantly, there was no correlation in
174 overall variance between TEs and genes in the same line (Supplemental Figure
175 1F), showing that TE expression and gene expression diverge independently.
176
177 In order to understand loss of repression of TE expression in more detail,
178 we classified TEs into different families using RepeatMasker annotations and

179 compared the median TE expression for each TE family across different lines. TE
180 expression was compared between *N*.1 and *N*>1 and tested for monotonic
181 median increase with decreasing population size. As expected, the majority of
182 TEs that showed a significant difference between *N*.1 and combined *N*.10 and
183 *N*.100 lines had increased expression in *N*.1 compared to the other lines (Figure
184 1C and Supplemental Figure 2). However, individual TEs displayed different
185 patterns of expression change. Some TEs, notably the DNA transposon Turmoil2,
186 showed more consistent increases across the *N*.1 lines (Figure 1D). Indeed, the
187 majority of the total effect on TE expression seen in Figure 1B could be
188 attributed to one TE family, the Turmoil2 TEs, which showed a large expression
189 increase across the majority of the *N*.1 lines (Figure 1C,D). Contrastingly some
190 TEs, notably the Mariner family DNA transposon Tc1, showed a burst-like
191 pattern of expression where a large increase in expression was observed in a few
192 *N*.1 lines whilst Tc1 remained low in expression in the remainder (Figure 1E).
193
194 We next investigated the timecourse of TE desilencing during
195 propagation of the MA lines. We performed gene expression analysis by RNAseq
196 on 11 *N*.1 population size lines at 25 and 100 generations. Median total TE
197 expression showed a highly significant increase with increasing numbers of
198 generations (Jh p = 1.1e-9; Figure 1F). Linear regression analysis confirmed a
199 positive relationship between increased numbers of generations and increased
200 TE expression (LR p = 6.7e-8; Figure 1F). Different TEs showed different kinetics
201 of desilencing. Turmoil2 showed a positive relationship between increased
202 numbers of generations and increased expression (LR p = 3.9e-3) and a
203 monotonic increase in median expression (Jh p = 2.55e-3; Figure 1G).
204 Contrastingly, Tc1 desilencing did not show a positive linear relationship
205 between increased generations and increased expression (LR p = 0.13; Figure
206 1H) though there was a significant increase in median expression (Jh p = 4.22e-
207 6).
208
209 We further investigated whether the expression of TEs in individual MA
210 lines correlated with the expression of other TEs. The majority of TEs showed
211 little correlation with the expression of other TEs (Figure 2). The TEs with the

212 most statistically significant increases in expression, Turmoil2 and non-LTR
213 retrotransposons of the LINE2 family (Figure 1C and Supplemental Figure 2)
214 clustered together (Figure 2) suggesting coregulation of these TE families
215 despite their different mechanisms of replication.
216
217 *Expression of TEs is weakly associated with increased copy-number*
218
219 TEs are capable of replicating independently of the host genome and thus
220 their copy-number might change across MA lines. We sequenced the genomes of
221 the MA lines after 400 generations and mapped the reads to consensus TE
222 sequences thereby obtaining estimates of copy-number variation (CNV) for each
223 TE family. Median TE copy-number increased with decreasing population size
224 (Jh p = 1e-7; LR p=1e-3 Figure 3A). Within the *N*.1 lines, there was a weak
225 significant positive correlation between increased copy-number and increased
226 expression. In contrast, we found no significant correlation between copy-
227 number and expression in the *N*.10 or *N*.100 lines (Figure 3B). Moreover,
228 increased expression of specific TEs in individual MA lines was often
229 unaccompanied by increased copy-number (Figure 3C,D). We conclude that
230 increased TE copy number is not the primary cause of increased TE expression.
231
232 *Alterations in small RNA levels are associated with TE expression changes*
233
234 Since changes in TE copy number are unlikely to explain altered TE
235 expression, we investigated whether changes in regulation of TE expression
236 could explain the loss of silencing observed during mutation accumulation. In *C.*
237 *elegans*, piRNAs and 22G-RNAs are important small RNA classes involved in TE
238 silencing(de Albuquerque et al., 2015; Bagijn et al., 2012; Das et al., 2008; Phillips
239 et al., 2015). To test whether piRNAs are important in the loss of silencing of TEs
240 in *N*.1 lines, we remapped recently published cross-linking immunoprecipitation
241 (CLIP) data(Shen et al., 2018) to identify TE transcripts that are bound by piRNA-
242 Piwi complexes *in vivo*. Approximately 25% of TE transcripts with RNAseq
243 reads were targeted by piRNAs. TEs targeted by piRNAs showed a statistically
244 significant increase in total expression in *N*.1 lines (Jh p = 0.02; LR p = 0.033;

245 Figure 4A) whilst TEs that were not targeted by piRNAs were not significantly
246 altered (Jh p = 0.47, LR p = 0.27; Figure 4B).
247
248 We tested whether defective piRNA-mediated silencing might account for
249 increased expression of TEs targeted by piRNAs. piRNAs in *C. elegans* are
250 expressed from individual promoters as a result of RNA polymerase II
251 transcription(Billi et al., 2013; Cecere et al., 2012; Gu et al., 2012). We
252 considered two potential mechanisms that might give rise to altered piRNAs.
253 One possibility is that mutations in the piRNA sequences might occur in
254 individual lines, which might affect their ability to recognise transposable
255 elements and thus interfere with TE silencing. There was a trend for *N*.1 lines to
256 have more mutations in piRNAs than *N*.10 or *N*.100 lines (Supplemental Figure
257 3). However, the trend was not significant (Jh p>0.05). Moreover, we identified
258 on average only 1.1 piRNA sequences with mutations across the *N*.1 lines. Thus
259 we conclude that the increase in TE expression is unlikely to be related to
260 mutations in specific piRNAs.
261
262 We next considered the expression of individual piRNA loci. We
263 identified piRNA loci with significantly altered median expression between *N*.1
264 lines and *N*.10 and *N*.100 lines combined using the Wilcoxon unpaired test. A
265 small percentage of piRNA loci showed significant differences but overall there
266 was no trend for these loci to show reduced expression in the *N*.1 lines; indeed
267 these loci were more likely to have higher expression in the *N*.1 lines (Figure 4C).
268 There was also no change in the expression of piRNAs identified by CLIP(Shen et
269 al., 2018) to be directly binding to TEs (Figure 4D). Thus changes in piRNA
270 expression are unlikely to explain the changes we observed in TE expression.
271
272 22G-RNAs act downstream of piRNAs to bring about target silencing(Das et al.,
273 2008). Surprisingly, although 22G-RNAs silence TEs, we found that the total
274 levels of 22G-RNAs mapping to TEs were increased in lines with smaller
275 population sizes (Jh p=0.05, LR p=0.068; Figure 4E). To examine this in more
276 detail, we analysed 22G-RNAs at individual TEs. The majority of TEs showing
277 significantly increased 22G-RNAs in *N*.1 lines were non-LTR transposons (Figure

278 4F). Notably, these TEs were not significantly increased in transcript levels in
279 the *N*.1 lines (Figure 1C; Supplemental Figure 2). Thus, the global increase in
280 22G-RNA levels does not relate directly to the increased transcript levels of TEs
281 in *N*.1 lines.
282
283 To compare directly how increased expression of TEs relates to 22G-RNA
284 mediated silencing, for each TE we compared 22G-RNA levels in lines with
285 increased expression to 22G-RNA levels in lines without increased expression.
286 The only TE showing alterations in both expression and 22G-RNA levels was
287 Turmoil2. Turmoil2 showed decreased small RNA levels in lines showing
288 increased Turmoil2 expression (Figure 4G). Thus, the increased transcript levels
289 of Turmoil2 elements is associated with reduced 22G-RNAs. We conclude that
290 the increased transcript levels of some TEs may be caused by reduced small
291 RNA-mediated silencing, but that some increases in TE transcript levels occur
292 independently of 22G-RNA changes.
293
294
295 *Chromatin environment determines relationship between 22G-RNA levels and TE*
296 *expression*
297
298 We investigated what determines the different changes in 22G-RNAs at
299 different TE types. 22G-RNAs interact with chromatin modifying factors to
300 control expression of TEs(McMurchy et al., 2017). piRNA-mediated silencing has
301 been directly linked to the generation of H3K9me2/3 marked nucleosomes
302 ("classical heterochromatin") and this has been proposed to be important for
303 transcriptional silencing induced by piRNAs(Ashe et al., 2012; McMurchy et al.,
304 2017; Shirayama et al., 2012; Woodhouse et al., 2018). Additionally, it is
305 becoming clear that a large proportion of the autosomal DNA in *C. elegans* can be
306 divided into active domains, containing H3K36me3 and germline expressed
307 genes and regulated domains, containing H3K27me3 marked nucleosomes and
308 silent genes(Evans et al., 2016; Liu et al., 2011). These domains are largely
309 stable through development, including in adult germline(Evans et al., 2016;
310 Rechtsteiner et al., 2010). We examined the influence of these three types of

311 chromatin on the response of TEs to mutation accumulation. TEs in regulated
312 domains and in classical heterochromatin showed no significant overall increase
313 in expression (Figure 5A,B). The rare examples of N.1 lines showing significantly
314 higher expression of TEs in classical heterochromatin corresponded to lines in
315 which Tc1 reactivation occurred, consistent with enrichment of Tc1 elements
316 within these regions(McMurchy et al., 2017). Contrastingly, TEs in active
317 domains had significantly higher expression in N.1 than *N*>1 lines (Wilcoxon
318 unpaired test, $p = 0.02$). There was also a trend towards decreased median
319 expression with increasing population size although this was on the border of
320 significance (Jh, $p = 0.056$, LR $p=0.064$; Figure 5C).
321
322 We next examined 22G-RNAs mapping to TEs across different chromatin
323 domains. TEs in repressed domains showed significantly increased levels of
324 22G-RNAs in *N*.1 lines relative to *N*.10 and *N*.100. However, in active domains
325 there was no significant change in 22G-RNA levels (Figure 5D,E,F). Thus
326 increased 22G-RNAs occur predominantly in TEs within repressed chromatin.
327
328
329 *AT-rich sequences in TEs reduce 22G-RNA generation*
330
331 A recent study has demonstrated that silencing of both transgenes and
332 endogenous genes by 22G-RNAs is inhibited by a high content of periodic repeats
333 of AT-rich sequences, known as PATCs(Frøkjær-Jensen et al., 2016). We tested
334 how PATC density within TEs influenced their expression under reduced
335 selection. High PATC density corresponded to reactivation of TEs (Jh $p = 0.018$,
336 LR p=0.02) whereas TEs with low PATC density did not show an increase in
337 expression (Figure 6A). Contrastingly, only TEs with low PATC density showed
338 significantly increased 22G-RNAs in *N*.1 lines relative to *N*.10 and *N*.100 (Jh $p =$
339 0.018, LR p=0.011; Figure 6B). Importantly this effect was specific to PATC
340 sequences as GC-content alone had no significant effect on either TE expression
341 or small RNA generation (Supplemental Figure 4A,B). We conclude that low
342 PATC density is required for 22G-RNA generation, which may be required to
343 restrain TE activation. We tested whether the chromatin environment

344 modulated the effect of PATC sequences on TE reactivation in MA lines.
345 Importantly, PATC content was similar in TEs across active, classic
346 heterochromatin and regulated domains (Supplemental Figure 4C). 22G-RNAs
347 were significantly increased in low PATC regions within regulated domain (Jh
348 p=0.018, LR =0.11) and classical heterochromatin domains (Jh p=0.055, LR
349 p=0.035) but not in active domains. In contrast there were no significant
350 changes in 22G-RNA levels in high PATC regions within these domains (Figure
351 6C,D). Thus generation of increased 22G-RNAs against TEs in MA lines is
352 promoted by both low PATC content and a repressive chromatin environment.
353
354
355 *Epigenetic memory of piRNA silencing correlates to compensatory 22G-RNA*
356 *biogenesis*
357
358 On the basis of these results we hypothesised that reactivation of TEs in
359 repressed chromatin regions may be restrained by a response whereby
360 increased 22G-RNAs are generated from TEs with low PATC content. In contrast,
361 loss of 22G-RNAs from TEs with higher PATC content can lead to their
362 reactivation in *N*.1 lines. The compensatory increase in 22G-RNAs is reminiscent
363 of recent results showing that silencing initiated by piRNAs can become
364 independent of the piRNA pathway due to the ability of 22G-RNAs to be
365 transmitted across generations(Ashe et al., 2012; Luteijn et al., 2012; Shirayama
366 et al., 2012). As a result, some genes and TEs show unchanged or even increased
367 levels of 22G-RNAs when piRNAs are removed whilst a functional 22G-
368 biogenesis machinery is still present(de Albuquerque et al., 2015; Phillips et al.,
369 2015). These 22G-RNAs can only be removed by mutating the 22G-RNA
370 biogenesis machinery. However, if the 22G-RNA biogenesis machinery is
371 subsequently reintroduced into a background lacking piRNAs, 22G-RNAs cannot
372 be made indicating that piRNAs are required to initiate their silencing(de
373 Albuquerque et al., 2015; Phillips et al., 2015). We used previously published
374 small RNA sequencing data from these studies to test whether epigenetic
375 memory of piRNA silencing might occur only in regulated domains and not active
376 domains. TEs in active domains showed significantly reduced 22G-RNAs when

377 *prg-1* was deleted in the presence of 22G-RNA biogenesis factors (Wilcoxon
378 Paired Test, p = 6e-5) whilst TEs in repressed chromatin did not show an overall
379 reduction (Wilcoxon Paired Test, p = 0.84) (Figure 7A). Importantly however,
380 TEs in all chromatin types showed a significant reduction in 22G-RNAs when the
381 22G-RNA biogenesis machinery was reintroduced in the absence of piRNAs
382 (Figure 7B). This suggests that the changes we observed in 22G-RNAs in MA are
383 similar to changes in 22G-RNAs that occur as a result of loss of piRNAs in the
384 context of functional 22G-RNA biogenesis factors.
385
386 **Discussion**
387
388 Our analysis of how the interplay between TE expression and TE silencing
389 factors changes over 400 generations at small population sizes provides the first
390 clear view of how TE expression diverges under reduced selection in animals.
391 Additionally, closer analysis of how TE control mechanisms are affected in the
392 MA lines offers new insight into the fundamental mechanisms of TE silencing in
393 *C. elegans*, underlining the ability of evolutionary studies to derive fundamental
394 molecular insights. Here we discuss each of these aspects of our work in turn.
395
396 *The effect of selection on TE expression*
397
398 Whilst TEs pose a threat to genome integrity, expression of TEs has been
399 suggested to have positive roles in gene expression networks; indeed the ability
400 of transposable element polymorphisms to rewire expression of some genes has
401 led to speculation that eukaryotic genomes may even avoid complete silencing of
402 TEs in order to facilitate this activity(Jacques et al., 2013). Here, we showed that,
403 at least in *C. elegans*, TE expression drifts to higher, rather than lower expression
404 under conditions of reduced selection. This demonstrates that, in general, low
405 expression of TEs is under purifying selection. Of all the TEs showing changes in
406 expression in MA lines, the only one showing a trend towards decreased
407 expression in *N*.1 lines than *N*.10 or *N*.100 was a non-LTR retrotransposon,
408 Vingi-1 (Figure 1B); however, the expression of this element was actually lower
409 in the starting population than in the *N*.10 or *N*.100 lines thus the significance of

410 this observation is unclear. Whilst it is possible that increased expression of TEs
411 may become beneficial under certain conditions, such as fluctuating
412 environmental conditions or pathogen infection, our observations strongly
413 suggest that the expression of most TEs is largely detrimental and TE expression
414 is under purifying selection. Thus our research would support the proposal that
415 beneficial events of TEs within genomes represent the exception- "making the
416 best of a bad job"- rather than the rule. However, an important possibility that
417 remains to be tested is whether TE activation is beneficial under fluctuating
418 environmental conditions as opposed to the stable environment of laboratory
419 maintained MA lines.
420
421 Phenotypic analyses of previous *C. elegans* MA experiments suggest that
422 the decline in fitness in *N* = 1 lines results primarily from a few mutations of
423 large effect(Estes et al., 2004; Halligan et al., 2003; Keightley and Cabellero,
424 1997). Similar results have been obtained in experimental evolution studies in
425 *Drosophila melanogaster*(Ávila and García-Dorado, 2002) and bacteria(Dillon
426 and Cooper, 2016; Heilbron et al., 2014). In our study, purifying selection at
427 larger population sizes (*N*.10 and *N*.100) would eliminate such large-effect
428 mutations and indeed, our *N*.10 and N. 100 lines exhibited no evidence of fitness
429 reduction over the course of successive bottlenecking for 409 generations(Katju
430 et al., 2015). Analysis of gene expression data from MA experiments from *C.*
431 *elegans*, *D. melanogaster* and *S. cerevisiae* concluded that large effect mutations
432 are also responsible for changes in protein-coding gene expression(Hodgins-
433 Davis et al., 2015). . However, our results for TE expression do not seem to fit
434 with this model because overall TE expression, which is largely driven by
435 Turmoil2 and non-LTR elements (see Figure 1), increases gradually with time
436 across the *N*.1 lines and is also increased, although less so, in *N*.10 and *N*.100
437 lines. The expression of Tc1 is an exception to the overall trend as it is not
438 affected in *N*.10 or *N*.100 lines but a small number of *N*.1 lines show markedly
439 increased Tc1 expression. Thus Tc1 reactivation may be dominated by a few
440 mutations with large effect. This difference might be related to the different
441 mechanism of silencing of Tc1 compared to Turmoil2 elements as discussed
442 further below.

443

444 Previous theories on the effect of natural selection on TE polymorphisms
445 have focussed on the copy-number of TEs as the major potential fitness
446 penalty(Bartolomé et al., 2002; Charlesworth and Langley, 1986; Lee and
447 Langley, 2010, 2012; Pasyukova et al., 2004). However, we observed that not all
448 expression increases of TEs are linked to increasing copy-number; indeed many
449 lines with very high expression of specific TEs display no evidence of increased
450 copy-number at all. This suggests that many TEs replicate inefficiently in *C.*
451 *elegans* such that even very large increases in expression levels do not
452 automatically result in increased copy-numbers. As a corollary of this point, TE
453 expression may be detrimental without directly posing a threat to genome
454 integrity, potentially through effects on endogenous gene expression networks
455 or through toxicity of repetitive RNA within the cell(Simon et al., 2014).
456

457 *Weakened piRNA silencing is responsible for increased expression of TEs under*
458 *relaxed selection.*
459

460 Our investigations of the molecular mechanisms behind reduced silencing
461 of TEs in MA lines strongly suggest that defective piRNA silencing is a major
462 culprit. Only piRNA-targeted TEs show increased expression in MA lines, and,
463 whilst piRNAs themselves do not seem to change significantly in MA lines, the
464 levels of 22G-RNAs that act as effectors of piRNA silencing are perturbed at TEs
465 showing increased expression. Why is piRNA silencing so vulnerable to mutation
466 accumulation? TE silencing and TE activation in organisms are likely in a
467 precarious equilibrium due to a constant evolutionary arms race between TEs
468 and their host genome. As a result, many mutations could converge on the
469 piRNA pathway to throw TE silencing out of balance.
470

471 *New insights into the role of chromatin in the piRNA pathway*
472

473 Our analysis of how mutation accumulation affects TE silencing provides
474 novel insights into how the chromatin environment of TEs affects piRNA-
475 mediated silencing in *C. elegans* (Figure 7C). Importantly, TEs in repressive

476 chromatin are much less prone to reactivation than those in active chromatin
477 regions. Mechanistically this is because 22G-RNAs in repressive chromatin
478 regions are stable or even increased, whilst 22G-RNAs mapping to TEs in active
479 regions are reduced in MA lines with increased expression. We show that this
480 also applies in a previously described experimental model where piRNAs are
481 removed without removing the 22G-RNA pathway(Phillips et al., 2015). 22G-
482 RNAs mapping to TEs within active domains are lost whereas those within
483 repressed regions are maintained. This result may also explain why reactivation
484 of Tc1 elements occurs less frequently than Turmoil2 elements, because Tc1
485 elements are predominantly located in repressed domains and are therefore
486 silenced more robustly.
487
488 What is the mechanism whereby silencing memory is supported in
489 repressed chromatin regions? The simplest possibility is that silencing and
490 generation of 22G-RNAs are directly promoted by repressive chromatin
491 modifications. In line with this possibility a mutually reinforcing loop between
492 H3K9 methylation and small RNAs is well documented in fission yeast(Bühler,
493 2009), and H3K9 methylation factors contribute to silencing of transgenes in *C.*
494 *elegans*(Ashe et al., 2012; McMurchy et al., 2017; Shirayama et al., 2012;
495 Woodhouse et al., 2018) although the situation is more complicated for
496 endogenous genes(Ni et al., 2014). However, our observations hold equally well
497 for H3K27me3-repressed chromatin, which has not been directly linked to 22G-
498 RNA silencing. We propose therefore that the nuclear small RNA pathway
499 responds differently depending on whether surrounding genes are active or
500 repressed to detect and quell aberrant gene activation. This model will be of
501 interest for further mechanistic investigation of small-RNA mediated silencing in
502 *C. elegans*.
503
504
505
506 *Figure Legends*
507

508 **Figure 1** Increased expression of transposable elements in mutation
509 accumulation lines. **A** Diagram of mutation accumulation (MA) experimental
510 design. **B** Overall transposable element expression in MA lines as a function of
511 population size. **C** Volcano plot for individual TE genes coloured according to
512 family. The p value is a Wilcoxon unpaired test comparing the median in *N*.1
513 lines to *N*.10 and *N*.100. **D** Expression changes in Turmoil2 elements. The box
514 shows interquartile range with a line at median and the whiskers extend to the
515 furthest point <=1.5 times the interquartile range from the median. **E** Expression
516 changes in Tc1 elements. Boxplot as in D. **F** Total TE expression after different
517 numbers of generations of mutation accumulation in lines with a population size
518 of 1. **G** Turmoil2 expression after different numbers of generations of mutation
519 accumulation with a population size of 1. **H** Tc1 expression after different
520 numbers of generations of mutation accumulation with a population size of 1.
521
522 **Figure 2** Heatmap showing the significance in correlation in expression between
523 different TEs across all the MA lines. The colour intensity shows the significance
524 of the Spearman's rank correlation coefficient with blue showing a negative
525 correlation and orange a positive correlation.
526
527 **Figure 3** Copy-number increases are weakly associated with increased TE
528 expression in MA lines. **A** Copy-number of TEs are increased with weaker
529 selection. Data shows mean of TE copy-number across all TEs normalized to the
530 mean across all lines. **B** Correlation between total change in TE copy-number
531 across all TEs and total TE expression. **C&D** Representative examples (Turmoil2
532 and non-LTR transposons) showing how copy-number changes in high or
533 expression lines. High expression lines are defined as being >1 standard
534 deviation away from the mean expression across all lines.
535
536 **Figure 4** Perturbed 22G-RNAs are associated with changes in TE expression in
537 MA lines. **A&B** Comparison of change in TE expression for piRNA targeted and
538 non-piRNA targeted TEs. **C** Total piRNA levels in MA lines across different
539 population sizes. **D** Levels of piRNAs shown to directly bind to TEs by CLIP-seq,
540 across different population size MA lines. **E** Levels of 22G-RNAs mapping to all

541 TEs in MA lines at different population sizes. **F** Volcano plot showing the
542 gradient of change in 22G levels as estimated by a linear model against the p-
543 value for the linear model on the y axis. G 22G-RNAs mapping to Turmoil2 in
544 lines showing increased Turmoil2 expression compared to lines where Turmoil2
545 expression was not increased.
546
547 **Figure 5** Chromatin environment controls alterations in piRNA mediated
548 silencing in MA lines. **A,B,C** Expression changes in TEs in Regulated
549 (H3K27me3), Classic Heterochromatin (H3K9me2) and Active (H3K36me3)
550 chromatin domains. *P*-values are Wilcoxon unpaired test. **D,E,F** Changes in TE-
551 mapping 22Gs in Regulated, Classic Heterochromatin and Active domains.
552
553 **Figure 6 PATC content influences 22G-RNA generation in MA lines**
554 **A** Expression changes in TEs in 4 equal size bins of 21 TEs with decreasing PATC
555 content. **B** 22G-RNAs across the bins used in A. **C** and **D** Stratification of bins
556 from A and B into active (H3K36me3) and regulated (H3K27me3) chromatin
557 domains. High PATC is the top bin and low PATC is the lowest bin.
558
559
560 **Figure 7 Chromatin and PATC together influence TE control in MA lines**
561 **A** Alterations in 22G-RNAs in the absence of *prg-1* whilst 22G-RNA biogenesis
562 pathways are still active for TEs divided into Regulated, Classic Heterochromatin
563 and Active domains. **B** Alterations in 22G-RNAs in the absence of both *prg-1* and
564 22G-RNA biogenesis pathways for TEs from Regulated, Classic Heterochromatin
565 and Active domains. Boxplots in A and B as in Figure 1D.
566 **C** Model for how chromatin environment contributes to TE desilencing in MA
567 lines. Left- repressed chromatin domains, taking H3K27me3 enriched domains
568 as an example. Right, Active chromatin domains.
569
570
571 Supplemental Figure 1
572 **A** Fano factor of individual TE transcript levels across all lines of the indicated
573 population size; B Fano factor in individual protein-coding gene transcript levels

574 across all lines of the indicated population size. C Variance of TE transcript
575 levels in 1000 samples of 5 N.1 lines, 253 samples of 5 N.10 lines [the maximum]
576 and all 5 N.100 lines. D Total variance in transcript level differences between
577 each TE and its corresponding value in the starting population across lines of
578 different population size. D Total variance of transcript level differences
579 between each protein-coding gene and the starting population in lines of the
580 indicated population size. Boxplots for A-D are as in Figure 1D. E Variance in TE
581 transcript changes relative to starting population compared to the variance in
582 protein-coding gene transcript changes in the same line.
583 Supplemental Figure 2
584 Volcano plot showing the logarithm (base 10) p-value of a linear model relating
585 population size to expression of TEs on the y axis to the gradient of the linear
586 model on the x axis.
587 Supplemental Figure 3
588 Total piRNA reads showing one base pair mismatch to the reference sequence
589 divided by the total number of mismatched loci, in MA lines across different
590 population sizes
591 Supplemental Figure 4
592 A GC content does not affect expression of TEs. Bins with high to low GC content
593 left to right. B GC content does not affect 22G-RNA levels. Bins with high to low
594 GC content left to right. C No clear difference in the proportion of TEs from
595 different chromatin domains within bins of different PATC content. D
596 Expression of TEs from regulated, heterochromatic and active chromatin regions
597 in the top PATC bin. E Expression of TEs from regulated, heterochromatic and
598 active chromatin regions in the lowest PATC bin.
599 Additional Files
600 R code and input files required to generate the plots in the figures will released
601 along with final publication and are available upon reasonable request to
602 psarkies@imperial.ac.uk
603
604
605
606

607 **Methods**

608

609 **RNA Library Preparation, Sequencing, and Analysis of Transcript**
610 **Abundance.**
611
612 The library preparation and RNA-sequencing procedures have previously
613 been described in detail (Konrad et al., 2018). Briefly, we isolated one, two, and
614 three individuals each from populations of sizes N = 1, 10, and 100, respectively.
615 These 55 worms, as well as one individual from the ancestral population were
616 each sequestered to NGM plates seeded with OP50, where they were allowed to
617 self-fertilize and reproduce at 20ºC. Three offspring worms at the L4 larval stage
618 were isolated from each of the F_1 populations to serve as biological replicates.
619 These 168 individual worm samples were allowed to reproduce for three
620 generations to yield enough tissue for RNA extraction. A standard bleaching
621 protocol was used to collect gravid eggs from adults in order to generate
622 synchronized populations of L1 larvae. Total RNA was isolated from L1 larvae
623 via the Qiagen RNeasy Mini Kit. The Nanodrop 2000, Qubit 3.0 Fluorometer, and
624 an Agilent RNA Analyzer were used to evaluate the quality of the RNA samples,
625 the and Illumina TruSeq RNA library Prep Kit v2 was used with standard
626 procedures to prepare the RNA sequencing libraries for each sample at the Texas
627 A&M University Genomics and Bioinformatics Services Center. The RNA was
628 fragmented and Illumina adapters were annealed for amplification. Size selected
629 cDNA fragments were isolated via a Qiagen Gel Extraction Kit. Finally,
630 sequencing was performed on the Illumina HiSeq 4000 platform with default
631 quality filters.
632 Demultiplexing and prefiltering of the sequencing reads was performed
633 based on default Illumina QC protocols. Reads containing abnormally short
634 insert lengths were removed, and adapters were discarted from the reads.
635 The raw RNA-sequencing reads in fastq format were aligned to the
636 protein-coding transcriptome of *C. elegans* (Wormbase reference N2 genome
637 version WS247) using TopHat(Trapnell et al., 2009) via the "very sensitive"
638 bowtie2 algorithm with a maximum of one mismatch in the anchor region for
639 each spliced alignment and a minimum and maximum intron length of 20 and
640 3,000 bp, respectively. Cufflinks(Trapnell et al., 2010) with default settings and
641 gene annotations from the N2 genome version WS247 was used to estimate the
642 relative transcript abundance for each protein-coding gene. All following
643 analyses were focused on FPKM values calculated on the per gene level. The
644 relative transcript abundances (FPKM) from the three biological replicates for
645 each original sample were averaged to get mean relative transcript abundance
646 for each gene in that sample.
647

648

649 **Small RNA sequencing**

650 MA lines were synchronised using hypochlorite treatment and embryos were

651 isolated after 12 hours. RNA was extracted using trizol and small RNA libraries

652 were prepared using the Illumina Small RNA sequencing kit as described
653 previously(Sarkies et al., 2015).
654 Small RNAs were aligned to a genome built using bowtie from a fasta file
655 containing all piRNAs, miRNAs, ncRNAs and genes including TEs, extracted from
656 Wormbase (WS264; ce11), requiring perfect mapping. Reads mapping to the
657 sense strand of ncRNAs and miRNAs were extracted using bedtools intersect –c –
658 S. We used DEseq using ncRNAs and miRNAs to extract size factors. 22G-RNAs
659 mapping to TEs and genes were extracted using a custom Perl script and the
660 number of 22G-RNAs mapping antisense to each gene and TE was then counted
661 using bedtools intersect –c -s. The 22G-RNAs were then normalized to the size
662 factors from ncRNAs and miRNAs combined.
663
664 TE copy-number analysis
665 Alignments to determe TE copy number changes in MA lines.
666 DNA sequencing was aligned to a genome built using bowtie2-build from TE
667 consensus sequences extracted from repbase combined with all coding
668 sequences. Bowtie2 was used to map PE reads to this genome and the readcount
669 mapping to each cds or TE was obtained using bedtools intersect –c.
670
671 Data integration
672
673 Computational analysis of TE expression and small RNA analysis
674 TEs from WS264 were annotated using Repeatmasker(Smit et al., 2017). All data
675 analysis was conducted using the R environment for statistical analysis
676 (www.Rproject.com). Details of the individual analyses are documented in the R
677 markdown file accompanying this manuscript (additional file 1), with the raw
678 data tables required to run these programs in as a zipped file (additional file 2).
679 Previously published datasets containing Chromatin domain annotations from
680 Early Embryo ChiP-Seq were taken from(Evans et al., 2016) updated to WS264
681 using liftover (https://genome.ucsc.edu/). Small RNA sequencing data from
682 reactivation of small RNA pathways in the presence or absence of piRNAs was
683 taken from (Phillips et al., 2015)and aligned to the *C. elegans* genome as
684 described above. The average PATC score for each TE was calculated by taking

685 the average PATC score across the element from per-base sliding window
686 genome-wide PATC scores from (Frøkjær-Jensen et al., 2016).
687
688 REFERENCES

689 de Albuquerque, B.F.M., Placentino, M., and Ketting, R.F. (2015). Maternal piRNAs
690 Are Essential for Germline Development following De Novo Establishment of
691 Endo-siRNAs in Caenorhabditis elegans. Dev. Cell *34*, 448–456.

692 Ashe, A., Sapetschnig, A., Weick, E.M., Mitchell, J., Bagijn, M.P., Cording, A.C.,
693 Doebley, A.L., Goldstein, L.D., Lehrbach, N.J., Le Pen, J., et al. (2012). PiRNAs can
694 trigger a multigenerational epigenetic memory in the germline of C. elegans. Cell
695 *150*, 88–99.

696 Ávila, V., and García-Dorado, A. (2002). The effects of spontaneous mutation on
697 competitive fitness in Drosophila melanogaster. J. Evol. Biol. *15*, 561–566.

698 Bagijn, M.P., Goldstein, L.D., Sapetschnig, A., Weick, E.M., Bouasker, S., Lehrbach,
699 N.J., Simard, M.J., and Miska, E.A. (2012). Function, targets, and evolution of
700 Caenorhabditis elegans piRNAs. Science (80-.). *337*, 574–578.

701 Bartolomé, C., Maside, X., and Charlesworth, B. (2002). On the Abundance and
702 Distribution of Transposable Elements in the Genome of Drosophila
703 melanogaster. Mol. Biol. Evol. *19*, 926–937.

704 Batista, P.J., Ruby, J.G., Claycomb, J.M., Chiang, R., Fahlgren, N., Kasschau, K.D.,
705 Chaves, D.A., Gu, W., Vasale, J.J., Duan, S., et al. (2008). PRG-1 and 21U-RNAs
706 Interact to Form the piRNA Complex Required for Fertility in C. elegans. Mol. Cell
707 *31*, 67–78.

708 Billi, A.C., Freeberg, M.A., Day, A.M., Chun, S.Y., Khivansara, V., and Kim, J.K.
709 (2013). A Conserved Upstream Motif Orchestrates Autonomous, Germline-
710 Enriched Expression of Caenorhabditis elegans piRNAs. PLoS Genet. *9*.

711 Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and
712 Hannon, G.J. (2007). Discrete Small RNA-Generating Loci as Master Regulators of
713 Transposon Activity in Drosophila. Cell *128*, 1089–1103.

714 Buckley, B.A., Burkhart, K.B., Gu, S.G., Spracklin, G., Kershner, A., Fritz, H., Kimble,
715 J., Fire, A., and Kennedy, S. (2012). A nuclear Argonaute promotes
716 multigenerational epigenetic inheritance and germline immortality. Nature *489*,
717 447–451.

718 Bühler, M. (2009). RNA turnover and chromatin-dependent gene silencing.
719 Chromosoma *118*, 141–151.

720 Cecere, G., Zheng, G.X.Y., Mansisidor, A.R., Klymko, K.E., and Grishok, A. (2012).
721 Promoters Recognized by Forkhead Proteins Exist for Individual 21U-RNAs. Mol.
722 Cell *47*, 734–745.

723 Charlesworth, B., and Langley, C.H. (1986). The evolution of self-regulated

724 transposition of transposable elements. Genetics *112*, 359–383.

725 Das, P.P., Bagijn, M.P., Goldstein, L.D., Woolford, J.R., Lehrbach, N.J., Sapetschnig,
726 A., Buhecha, H.R., Gilchrist, M.J., Howe, K.L., Stark, R., et al. (2008). Piwi and
727 piRNAs Act Upstream of an Endogenous siRNA Pathway to Suppress Tc3
728 Transposon Mobility in the Caenorhabditis elegans Germline. Mol. Cell *31*, 79–90.

729 Denver, D.R., Morris, K., Streelman, J.T., Kim, S.K., Lynch, M., and Thomas, W.K.
730 (2005). The transcriptional consequences of mutation and natural selection in
731 Caenorhabditis elegans. Nat. Genet. *37*, 544–548.

732 Dillon, M.M., and Cooper, V.S. (2016). The fitness effects of spontaneous
733 mutations nearly unseen by selection in a bacterium with multiple
734 chromosomes. Genetics *204*, 1225–1238.

735 Estes, S., Phillips, P.C., Denver, D.R., Thomas, W.K., and Lynch, M. (2004).
736 Mutation Accumulation in Populations of Varying Size: The Distribution of
737 Mutational Effects for Fitness Correlates in Caenorhabditis elegans. Genetics *166*,
738 1269–1279.

739 Evans, K.J., Huang, N., Stempor, P., Chesney, M.A., Down, T.A., and Ahringer, J.
740 (2016). Stable Caenorhabditis elegans chromatin domains separate broadly
741 expressed and developmentally regulated genes. Proc. Natl. Acad. Sci. *113*,
742 E7020--E7029.

743 Fablet, M., and Vieira, C. (2011). Evolvability, epigenetics and transposable
744 elements. Biomol. Concepts *2*.

745 Frøkjær-Jensen, C., Jain, N., Hansen, L., Davis, M.W., Li, Y., Zhao, D., Rebora, K.,
746 Millet, J.R.M., Liu, X., Kim, S.K., et al. (2016). An Abundant Class of Non-coding
747 DNA Can Prevent Stochastic Gene Silencing in the C. elegans Germline. Cell *166*,
748 343–357.

749 Gu, W., Lee, H.-C., Chaves, D., Youngman, E.M., Pazour, G.J., Conte, D., and Mello,
750 C.C. (2012). CapSeq and CIP-TAP identify Pol II start sites and reveal capped
751 small RNAs as C. elegans piRNA precursors. Cell *151*, 1488–1500.

752 Halligan, D.L., and Keightley, P.D. (2009). Spontaneous Mutation Accumulation
753 Studies in Evolutionary Genetics. Annu. Rev. Ecol. Evol. Syst. *40*, 151–172.

754 Halligan, D., Peters, A., and Keightley, P. (2003). Estimating numbers of EMS-
755 induced mutations affecting life history traits in Caenorhabditis elegans in
756 crosses between inbred sublines. Genet. Res. *82*, 191–205.

757 Heilbron, K., Toll-Riera, M., Kojadinovic, M., and MacLean, R.C. (2014). Fitness is
758 strongly influenced by rare mutations of large effect in a microbial mutation
759 accumulation experiment. Genetics *197*, 981–990.

760 Hodgins-Davis, A., Rice, D.P., and Townsend, J.P. (2015). Gene Expression Evolves
761 under a House-of-Cards Model of Stabilizing Selection. Mol. Biol. Evol. *32*, 2130–
762 2140.

763 Imbeault, M., Helleboid, P.-Y., and Trono, D. (2017). KRAB zinc-finger proteins
764 contribute to the evolution of gene regulatory networks. Nature *543*, 550–554.

765 Jacques, P.-É., Jeyakani, J., and Bourque, G. (2013). The Majority of Primate-
766 Specific Regulatory Sequences Are Derived from Transposable Elements. PLoS
767 Genet. *9*, e1003504.

768 Katju, V., and Bergthorsson, U. (2019). Old Trade, New Tricks: Insights into the
769 Spontaneous Mutation Process from the Partnering of Classical Mutation
770 Accumulation Experiments with High-Throughput Genomic Approaches.
771 Genome Biol. Evol. *11*, 136–165.

772 Katju, V., Packard, L.B., Bu, L., Keightley, P.D., and Bergthorsson, U. (2015).
773 Fitness decline in spontaneous mutation accumulation lines of *Caenorhabditis*
774 *elegans* with varying effective population sizes. Evolution (N. Y). *69*, 104–116.

775 Katju, V., Packard, L.B., and Keightley, P.D. (2018). Fitness decline under osmotic
776 stress in *Caenorhabditis elegans* populations subjected to spontaneous mutation
777 accumulation at varying population sizes. Evolution (N. Y). *72*, 1000–1008.

778 Keightley, P.D., and Cabellero, A. (1997). Genomic mutation rates for lifetime
779 reproductive output and lifespan in Caenorhabditis elegans. Proc. Natl. Acad. Sci.
780 USA *94*, 3823–3827.

781 Konrad, A., Flibotte, S., Taylor, J., Waterston, R.H., Moerman, D.G., Bergthorsson,
782 U., and Katju, V. (2018). Mutational and transcriptional landscape of spontaneous
783 gene duplications and deletions in *Caenorhabditis elegans*. Proc. Natl. Acad. Sci.
784 *115*, 7386–7391.

785 Landry, C.R., Lemos, B., Rifkin, S.A., Dickinson, W.J., and Hartl, D.L. (2007). Genetic
786 properties influencing the evolvability of gene expression. Science (80-.). *317*,
787 118–121.

788 Lee, Y.C.G., and Langley, C.H. (2010). Transposable elements in natural
789 populations of Drosophila melanogaster. Philos. Trans. R. Soc. B Biol. Sci. *365*,
790 1219–1228.

791 Lee, Y.C.G., and Langley, C.H. (2012). Long-term and short-term evolutionary
792 impacts of transposable elements on Drosophila. Genetics *192*, 1411–1432.

793 Liu, T., Rechtsteiner, A., Egelhofer, T.A., Vielle, A., Latorre, I., Cheung, M.-S., Ercan,
794 S., Ikegami, K., Jensen, M., Kolasinska-Zwierz, P., et al. (2011). Broad
795 chromosomal domains of histone modification patterns in C. elegans. Genome
796 Res. *21*, 227–236.

797 Luteijn, M.J., van Bergeijk, P., Kaaij, L.J.T., Almeida, M.V., Roovers, E.F., Berezikov,
798 E., and Ketting, R.F. (2012). Extremely stable Piwi-induced gene silencing in
799 Caenorhabditis elegans. EMBO J. *31*, 3422–3430.

800 McMurchy, A.N., Stempor, P., Gaarenstroom, T., Wysolmerski, B., Dong, Y.,
801 Aussianikava, D., Appert, A., Huang, N., Kolasinska-Zwierz, P., Sapetschnig, A., et

802 al. (2017). A team of heterochromatin factors collaborates with small RNA
803 pathways to combat repetitive elements and germline stress. Elife *6*.

804 Ni, J.Z., Chen, E., and Gu, S.G. (2014). Complex coding of endogenous siRNA,
805 transcriptional silencing and H3K9 methylation on native targets of germline
806 nuclear RNAi in C. elegans. BMC Genomics *15*, 1157.

807 Pak, J., and Fire, A. (2007). Distinct populations of primary and secondary
808 effectors during RNAi in C. elegans. Science (80-.). *315*, 241–244.

809 Pasyukova, E.G., Nuzhdin, S. V., Morozova, T. V., and Mackay, T.F.C. (2004).
810 Accumulation of transposable elements in the genome of Drosophila
811 melanogaster is associated with a decrease in fitness. J. Hered. *95*, 284–290.

812 Phillips, C.M., Brown, K.C., Montgomery, B.E., Ruvkun, G., and Montgomery, T.A.
813 (2015). PiRNAs and piRNA-Dependent siRNAs Protect Conserved and Essential
814 C. elegans Genes from Misrouting into the RNAi Pathway. Dev. Cell *34*, 457–465.

815 Pritham, E.J. (2009). Transposable Elements and Factors Influencing their
816 Success in Eukaryotes. J. Hered. *100*, 648–655.

817 Rebollo, R., Romanish, M.T., and Mager, D.L. (2011). Transposable Elements: An
818 Abundant and Natural Source of Regulatory Sequences for Host Genes. Annu.
819 Rev. Genet. *46*, 120913153128008.

820 Rechtsteiner, A., Ercan, S., Takasaki, T., Phippen, T.M., Egelhofer, T.A., Wang, W.,
821 Kimura, H., Lieb, J.D., and Strome, S. (2010). The Histone H3K36
822 Methyltransferase MES-4 Acts Epigenetically to Transmit the Memory of
823 Germline Gene Expression to Progeny. PLOS Genet. *6*, e1001091.

824 Rifkin, S.A., Houle, D., Kim, J., and White, K.P. (2005). A mutation accumulation
825 assay reveals a broad capacity for rapid evolution of gene expression. Nature
826 *438*, 220–223.

827 Ruby, J.G., Jan, C., Player, C., Axtell, M.J., Lee, W., Nusbaum, C., Ge, H., and Bartel,
828 D.P. (2006). Large-Scale Sequencing Reveals 21U-RNAs and Additional
829 MicroRNAs and Endogenous siRNAs in C. elegans. Cell *127*, 1193–1207.

830 Sarkies, P., Selkirk, M.E., Jones, J.T., Blok, V., Boothby, T., Goldstein, B., Hanelt, B.,
831 Ardila-Garcia, A., Fast, N.M., Schiffer, P.M., et al. (2015). Ancient and Novel Small
832 RNA Pathways Compensate for the Loss of piRNAs in Multiple Independent
833 Nematode Lineages. PLoS Biol. *13*.

834 Shen, E.-Z., Chen, H., Ozturk, A.R., Tu, S., Shirayama, M., Tang, W., Ding, Y.-H., Dai,
835 S.-Y., Weng, Z., and Mello, C.C. (2018). Identification of piRNA Binding Sites
836 Reveals the Argonaute Regulatory Landscape of the C. elegans Germline. Cell *172*,
837 937–951.e18.

838 Shirayama, M., Seth, M., Lee, H.-C., Gu, W., Ishidate, T., Conte, D., and Mello, C.C.
839 (2012). piRNAs initiate an epigenetic memory of nonself RNA in the C. elegans
840 germline. Cell *150*, 65–77.

841 Simon, M., Sarkies, P., Ikegami, K., Doebley, A.L., Goldstein, L.D., Mitchell, J.,
842 Sakaguchi, A., Miska, E.A., and Ahmed, S. (2014). Reduced Insulin/IGF-1 Signaling
843 Restores Germ Cell Immortality to Caenorhabditis elegans Piwi Mutants. Cell
844 Rep. *7*, 762–773.

845 Simonti, C.N., Pavličev, M., and Capra, J.A. (2017). Transposable Element
846 Exaptation into Regulatory Regions Is Rare, Influenced by Evolutionary Age, and
847 Subject to Pleiotropic Constraints. Mol. Biol. Evol. *34*, 2856–2869.

848 Siomi, M.C., Sato, K., Pezic, D., and Aravin, A.A. (2011). PIWI-interacting small
849 RNAs: the vanguard of genome defence. Nat. Rev. Mol. Cell Biol. *12*, 246.

850 Smit, A., Hubley, R., and Green, P. (2017). RepeatMasker Open-4.0.6 2013-2015 .
851 Http://Www.Repeatmasker.Org.

852 Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice
853 junctions with RNA-Seq. Bioinformatics *25*, 1105–1111.

854 Trapnell, C., Williams, B. a, Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J.,
855 Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). How Cufflinks works. Nat.
856 Biotechnol. *28*, 511–515.

857 Wang, G., and Reinke, V. (2008). A C. elegans Piwi, PRG-1, Regulates 21U-RNAs
858 during Spermatogenesis. Curr. Biol. *18*, 861–867.

859 Weick, E., and Miska, E. a (2014). piRNAs : from biogenesis to function. Genes
860 Dev. 1–41.

861 Woodhouse, R.M., Buchmann, G., Hoe, M., Harney, D.J., Low, J.K.K., Larance, M.,
862 Boag, P.R., and Ashe, A. (2018). Chromatin Modifiers SET-25 and SET-32 Are
863 Required for Establishment but Not Long-Term Maintenance of
864 Transgenerational Epigenetic Inheritance. Cell Rep. *25*, 2259–2272.e5.

865 Yigit, E., Batista, P.J., Bei, Y., Pang, K.M., Chen, C.-C.G., Tolia, N.H., Joshua-Tor, L.,
866 Mitani, S., Simard, M.J., and Mello, C.C. (2006). Analysis of the C. elegans
867 Argonaute family reveals that distinct Argonautes act sequentially during RNAi.
868 Cell *127*, 747–757.

869

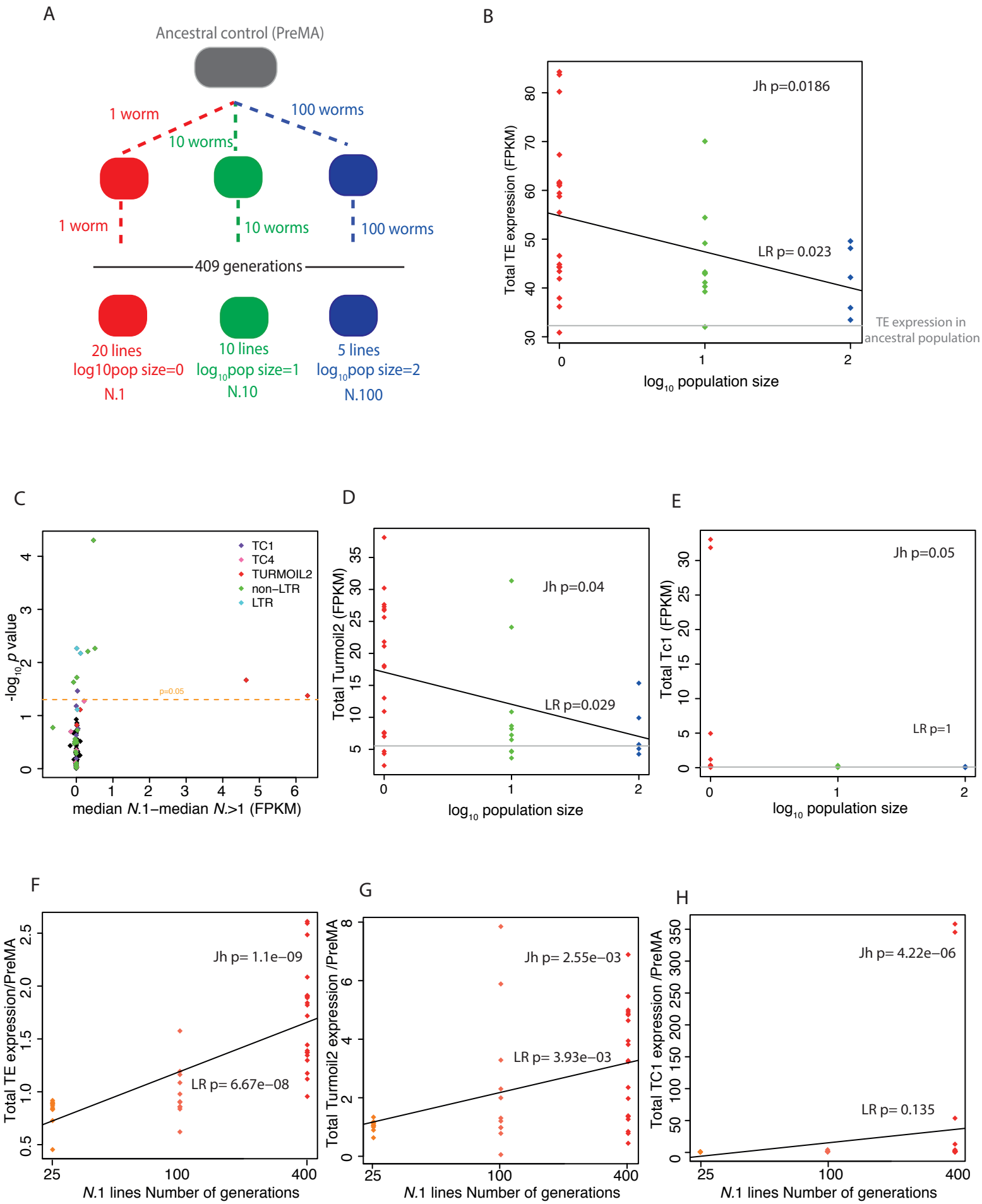

A
Ancestral control (PreMA)
1 worm
10 worms
100 worms
1 worm
10 worms
100 worms
409 generations
20 lines
log10pop size=0
N.1
10 lines
log10pop size=1
N.10
5 lines
log10pop size=2
N.100
B
Jh p=0.0186
LR p= 0.023
Total TE expression (FPKM)
TE expression in ancestral population
log10 population size
C
TC1
TC4
TURMOIL2
non−LTR
LTR
p=0.05
-log10 p value
median N.1−median N.>1 (FPKM)
D
Jh p=0.04
LR p=0.029
Total Turmoil2 (FPKM)
log10 population size
E
Jh p=0.05
LR p=1
Total Tc1 (FPKM)
log10 population size
F
Jh p= 1.1e−09
LR p= 6.67e−08
Total TE expression/PreMA
N.1 lines Number of generations
G
Jh p= 2.55e−03
LR p= 3.93e−03
Total Turmoil2 expression /PreMA
N.1 lines Number of generations
H
Jh p= 4.22e−06
LR p= 0.135
Total TC1 expression /PreMA
N.1 lines Number of generations

Figure 1

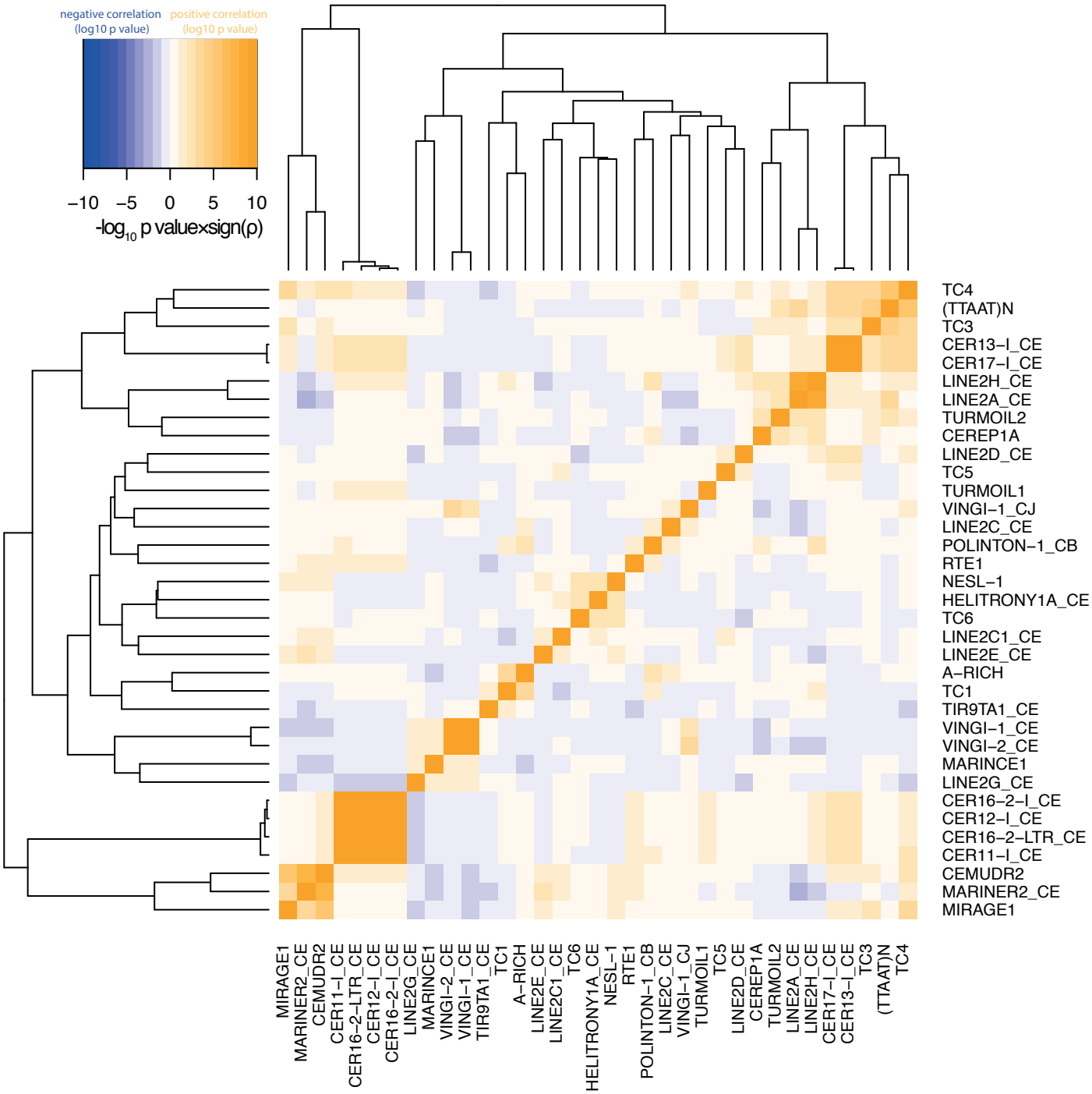

negative correlation
(log10 p value)
positive correlation
(log10 p value)
−10
−5
0
5
10
-log10 p value×sign(ρ)
TC4
(TTAAT)N
TC3
CER13−I_CE
CER17−I_CE
LINE2H_CE
LINE2A_CE
TURMOIL2
CEREP1A
LINE2D_CE
TC5
TURMOIL1
VINGI−1_CJ
LINE2C_CE
POLINTON−1_CB
RTE1
NESL−1
HELITRONY1A_CE
TC6
LINE2C1_CE
LINE2E_CE
A−RICH
TC1
TIR9TA1_CE
VINGI−1_CE
VINGI−2_CE
MARINCE1
LINE2G_CE
CER16−2−I_CE
CER12−I_CE
CER16−2−LTR_CE
CER11−I_CE
CEMUDR2
MARINER2_CE
MIRAGE1
MIRAGE1
MARINER2_CE
CEMUDR2
CER11−I_CE
CER16−2−LTR_CE
CER12−I_CE
CER16−2−I_CE
LINE2G_CE
MARINCE1
VINGI−2_CE
VINGI−1_CE
TIR9TA1_CE
TC1
A−RICH
LINE2E_CE
LINE2C1_CE
TC6
HELITRONY1A_CE
NESL−1
RTE1
POLINTON−1_CB
LINE2C_CE
VINGI−1_CJ
TURMOIL1
TC5
LINE2D_CE
CEREP1A
TURMOIL2
LINE2A_CE
LINE2H_CE
CER17−I_CE
CER13−I_CE
TC3
(TTAAT)N
TC4

Figure 2

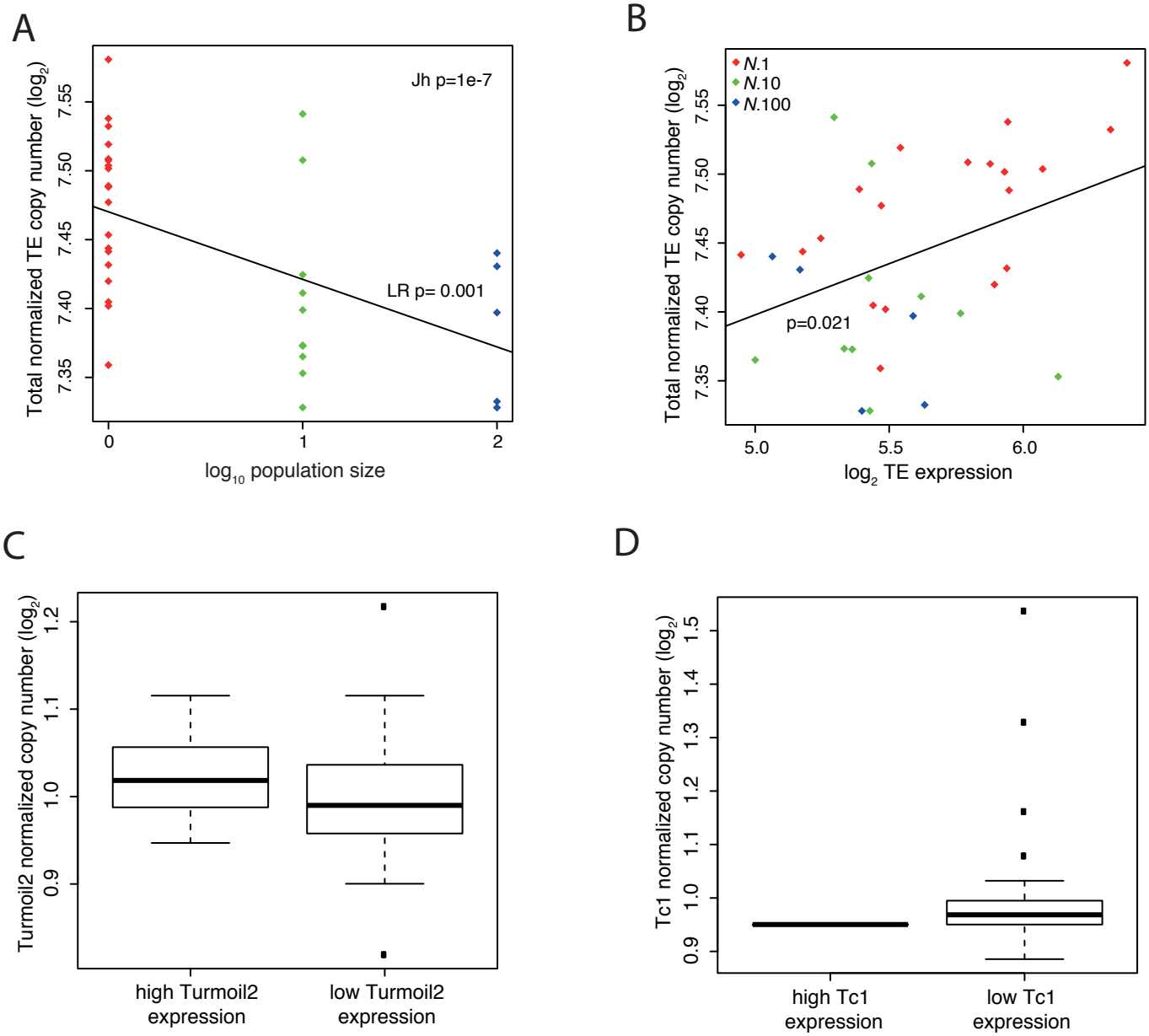

A
Total normalized TE copy number (log2)
7.35
7.40
7.45
7.50
7.55
Jh p=1e-7
LR p= 0.001
0
1
2
log10 population size
B
N.1
N.10
N.100
Total normalized TE copy number (log2)
7.35
7.40
7.45
7.50
7.55
p=0.021
5.0
5.5
6.0
log2 TE expression
C
Turmoil2 normalized copy number (log2)
0.9
1.0
1.1
1.2
high Turmoil2 expression
low Turmoil2 expression
D
Tc1 normalized copy number (log2)
0.9
1.0
1.1
1.2
1.3
1.4
1.5
high Tc1 expression
low Tc1 expression

Figure 3

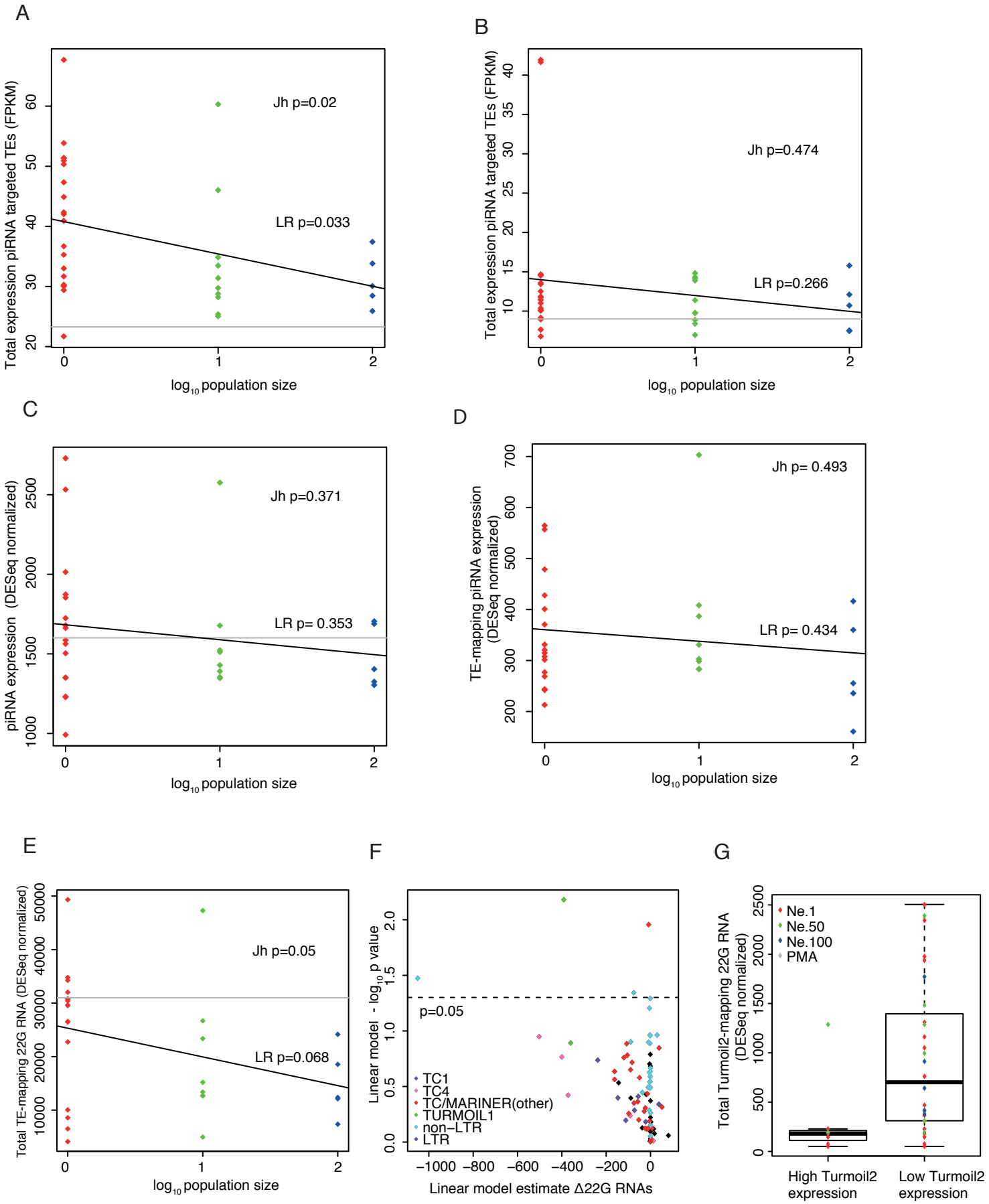

A
Total expression piRNA targeted TEs (FPKM)
Jh p=0.02
LR p=0.033
log10 population size
B
Total expression piRNA targeted TEs (FPKM)
Jh p=0.474
LR p=0.266
log10 population size
C
piRNA expression (DESeq normalized)
Jh p=0.371
LR p= 0.353
log10 population size
D
TE-mapping piRNA expression (DESeq normalized)
Jh p= 0.493
LR p= 0.434
log10 population size
E
Total TE-mapping 22G RNA (DESeq normalized)
Jh p=0.05
LR p=0.068
log10 population size
F
Linear model - log10 p value
p=0.05
TC1
TC4
TC/MARINER(other)
TURMOIL1
non−LTR
LTR
Linear model estimate Δ22G RNAs
G
Total Turmoil2-mapping 22G RNA (DESeq normalized)
Ne.1
Ne.50
Ne.100
PMA
High Turmoil2 expression
Low Turmoil2 expression

Figure 4

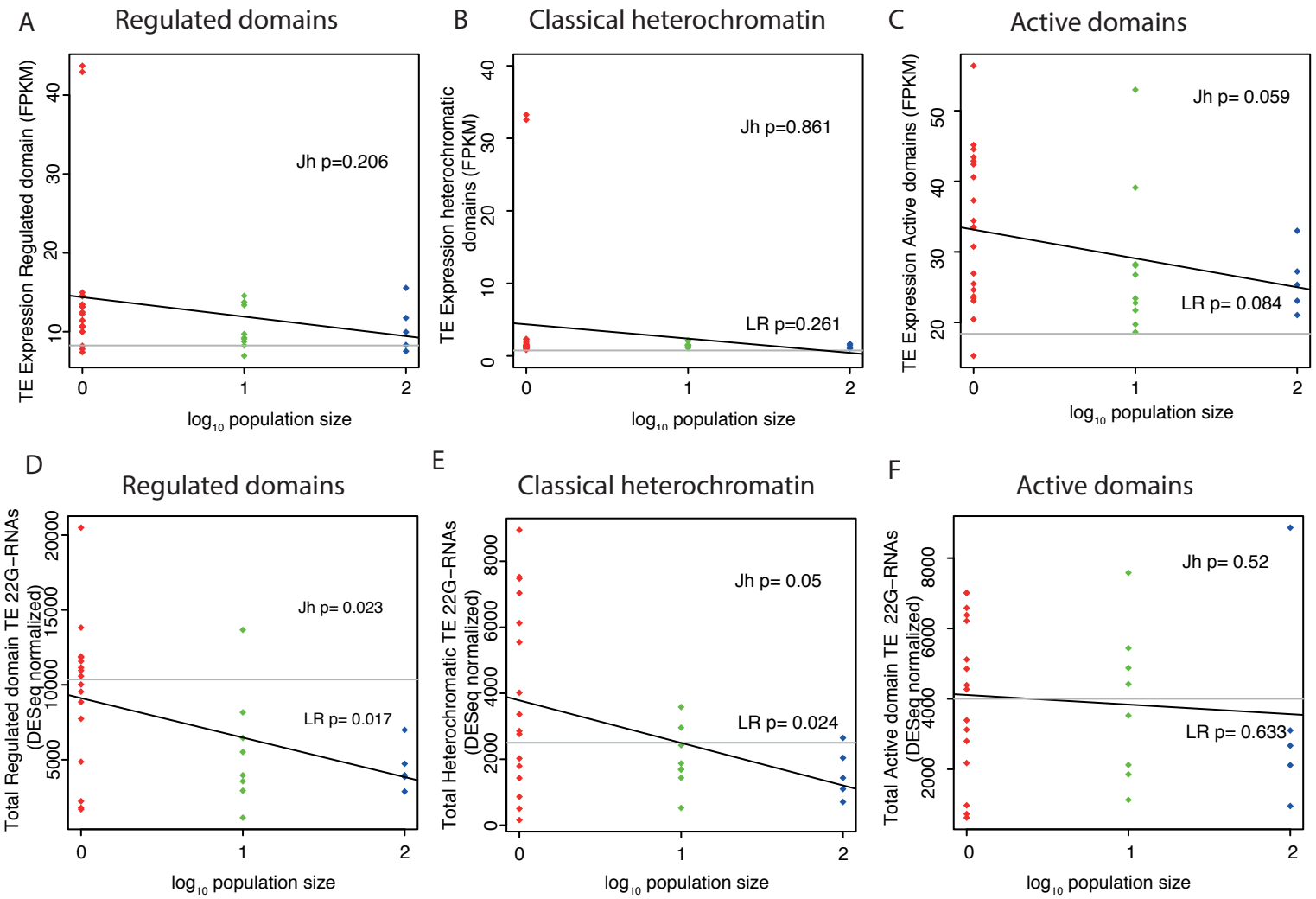

A
Regulated domains
TE Expression Regulated domain (FPKM)
Jh p=0.206
log10 population size
B
Classical heterochromatin
TE Expression heterochromatic domains (FPKM)
Jh p=0.861
LR p=0.261
log10 population size
C
Active domains
TE Expression Active domains (FPKM)
Jh p= 0.059
LR p= 0.084
log10 population size
D
Regulated domains
Total Regulated domain TE 22G−RNAs (DESeq normalized)
Jh p= 0.023
LR p= 0.017
log10 population size
E
Classical heterochromatin
Total Heterochromatic TE 22G−RNAs (DESeq normalized)
Jh p= 0.05
LR p= 0.024
log10 population size
F
Active domains
Total Active domain TE 22G−RNAs (DESeq normalized)
Jh p= 0.52
LR p= 0.633
log10 population size

Figure 5

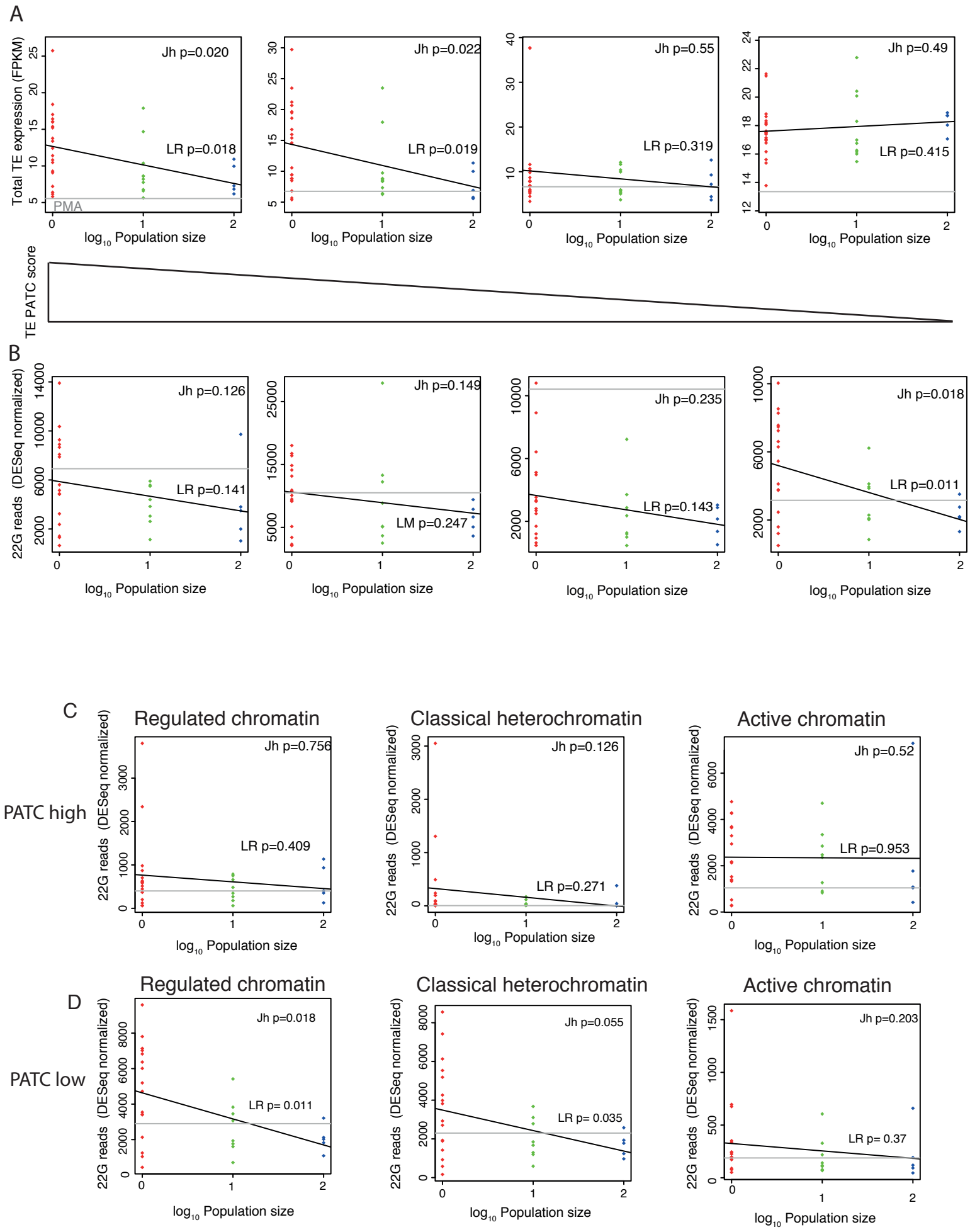
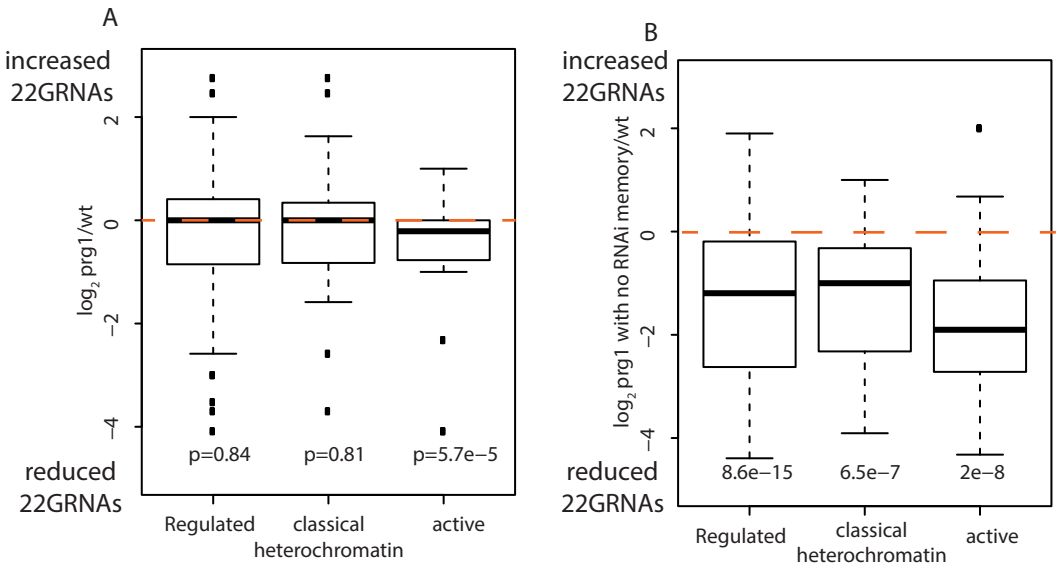

A
Total TE expression (FPKM)
Jh p=0.020
LR p=0.018
PMA
Jh p=0.022
LR p=0.019
Jh p=0.55
LR p=0.319
Jh p=0.49
LR p=0.415
log10 Population size
TE PATC score
B
22G reads (DESeq normalized)
Jh p=0.126
LR p=0.141
Jh p=0.149
LM p=0.247
Jh p=0.235
LR p=0.143
Jh p=0.018
LR p=0.011
C
PATC high
Regulated chromatin
Jh p=0.756
LR p=0.409
Classical heterochromatin
Jh p=0.126
LR p=0.271
Active chromatin
Jh p=0.52
LR p=0.953
D
PATC low
Regulated chromatin
Jh p=0.018
LR p= 0.011
Classical heterochromatin
Jh p=0.055
LR p= 0.035
Active chromatin
Jh p=0.203
LR p= 0.37

Figure 6


A
increased 22GRNAs
reduced 22GRNAs
log2 prg1/wt
2
0
−2
−4
p=0.84
p=0.81
p=5.7e−5
Regulated
classical heterochromatin
active
B
increased 22GRNAs
reduced 22GRNAs
log2 prg1 with no RNAi memory/wt
2
0
−2
−4
8.6e−15
6.5e−7
2e−8
Regulated
classical heterochromatin
active

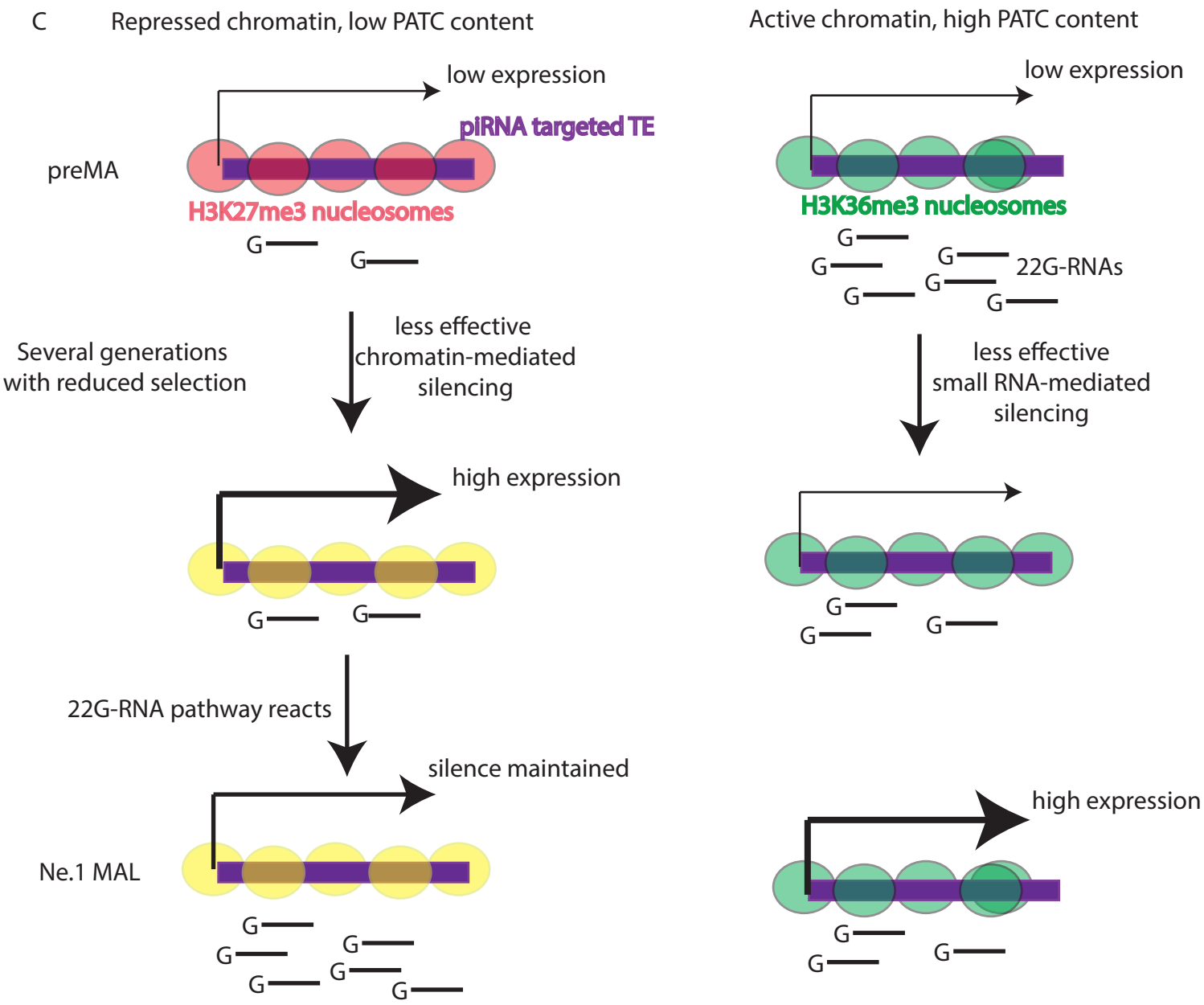
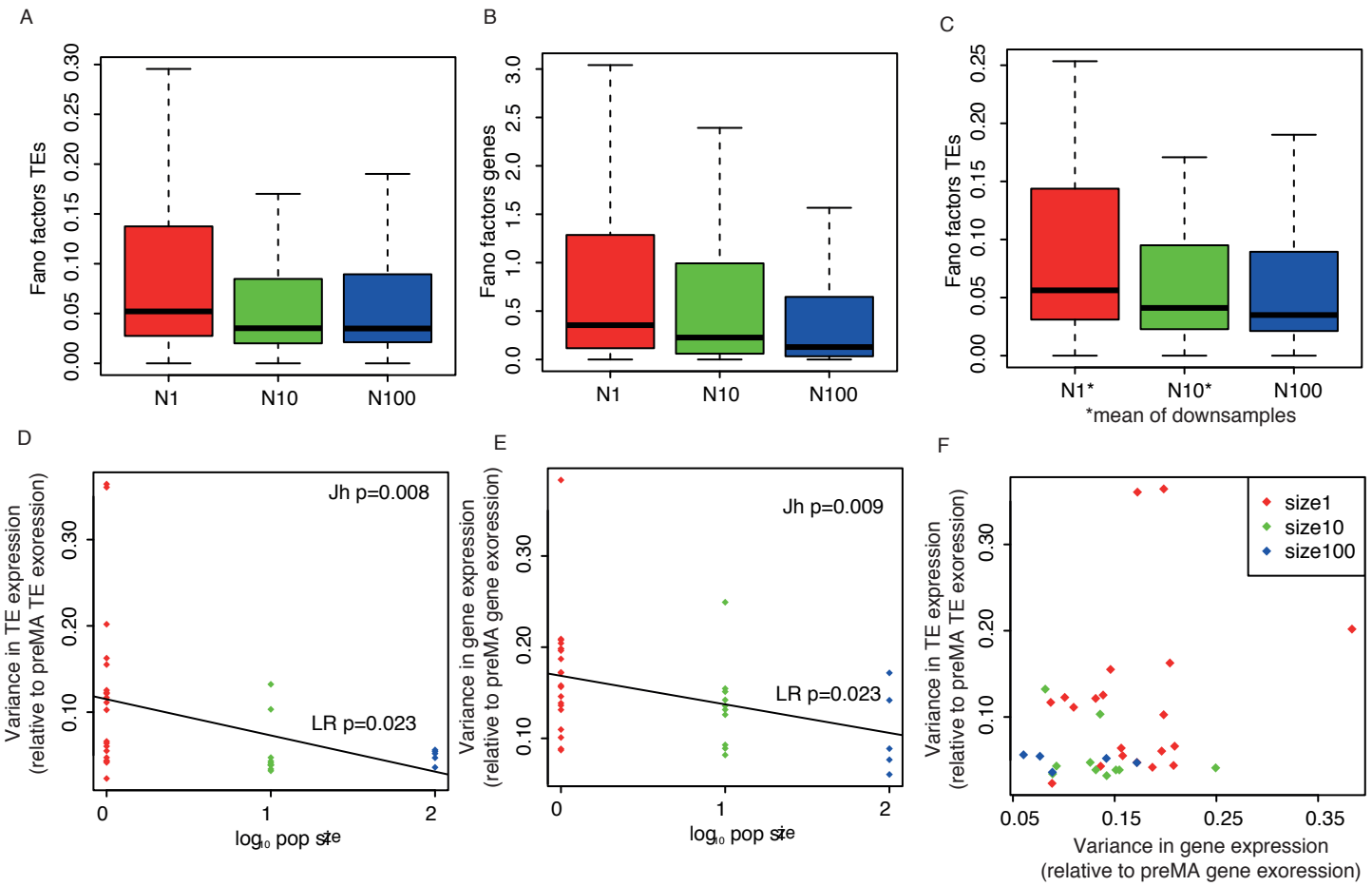
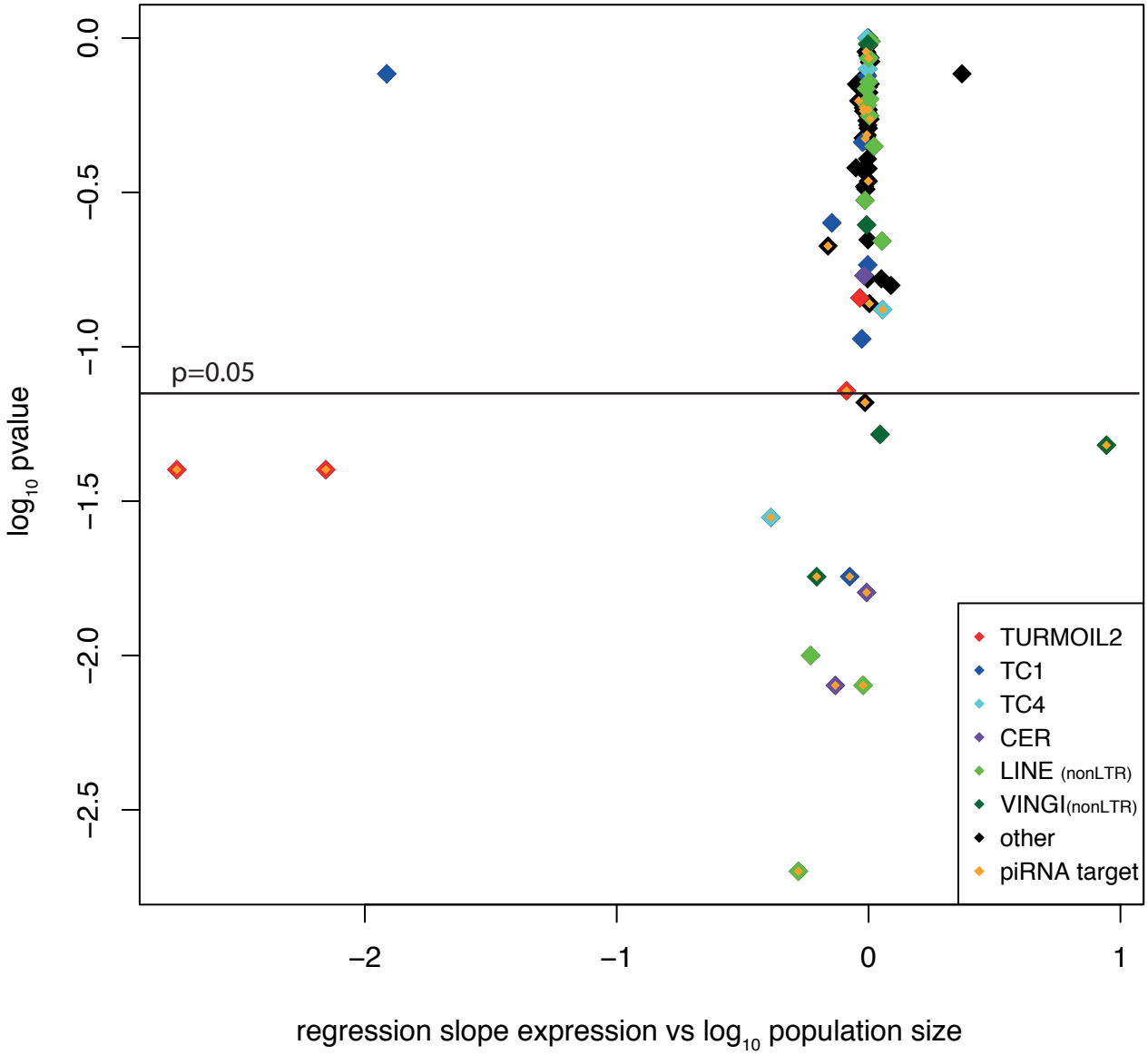

C
Repressed chromatin, low PATC content
Active chromatin, high PATC content
low expression
piRNA targeted TE
preMA
H3K27me3 nucleosomes
G
low expression
H3K36me3 nucleosomes
22G-RNAs
Several generations with reduced selection
less effective chromatin-mediated silencing
less effective small RNA-mediated silencing
high expression
22G-RNA pathway reacts
silence maintained
Ne.1 MAL
high expression

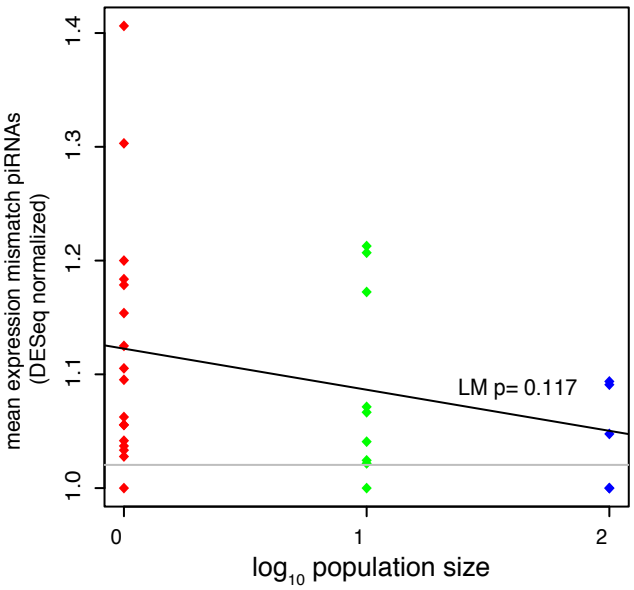
Figure 7



Supplemental Figure 1



Supplemental Figure 2



Supplemental Figure 3

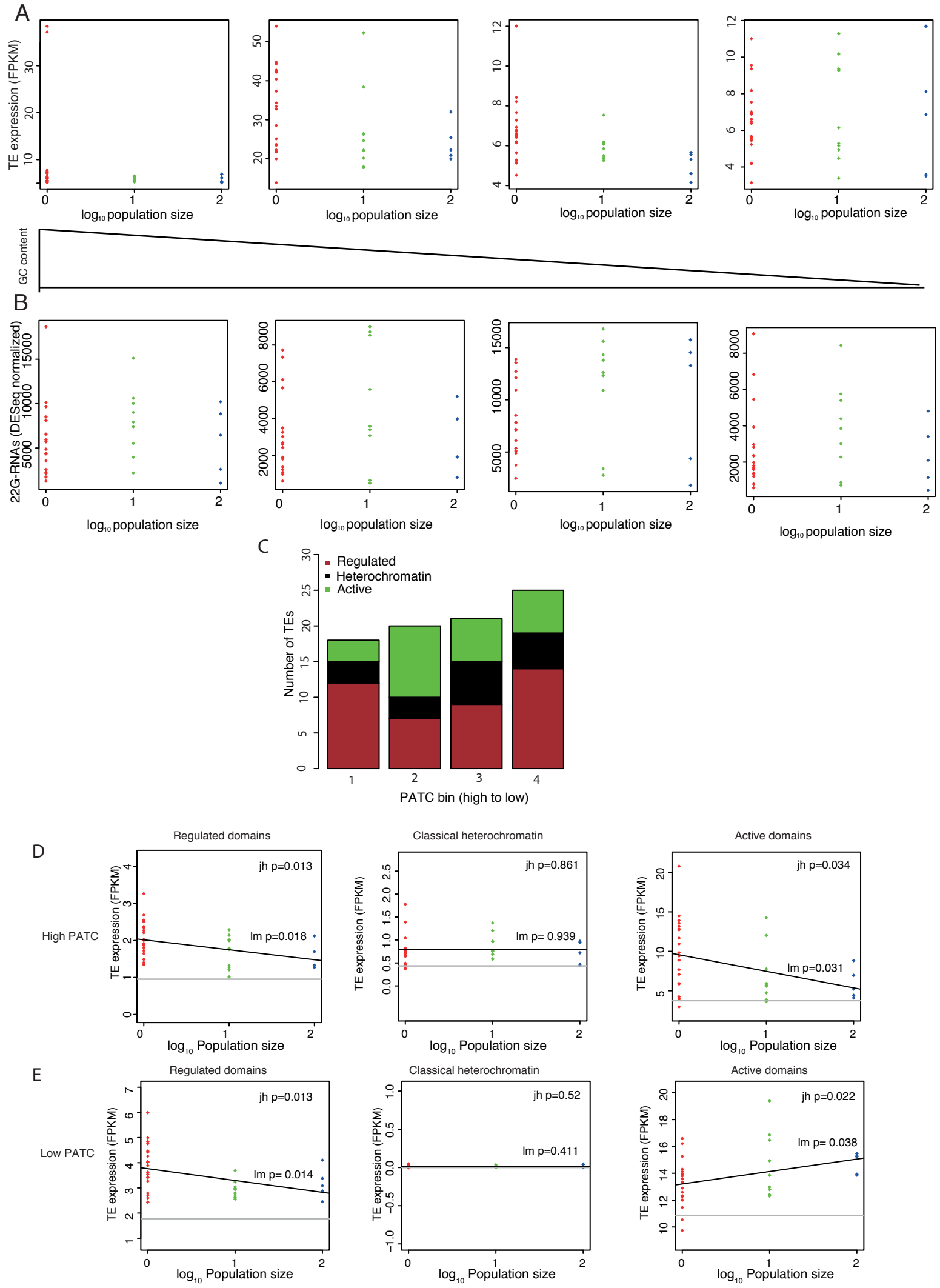

A
TE expression (FPKM)
log10 population size
GC content
B
22G-RNAs (DESeq normalized)
log10 population size
C
Regulated
Heterochromatin
Active
Number of TEs
PATC bin (high to low)
D
High PATC
Regulated domains
Classical heterochromatin
Active domains
jh p=0.013
lm p=0.018
jh p=0.861
lm p= 0.939
jh p=0.034
lm p=0.031
TE expression (FPKM)
log10 Population size
E
Low PATC
Regulated domains
Classical heterochromatin
Active domains
jh p=0.013
lm p= 0.014
jh p=0.52
lm p=0.411
jh p=0.022
lm p= 0.038
TE expression (FPKM)
log10 Population size

Figure S4