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

genes or gene regulatory elements. As a result, organisms have evolved sophisticated control strategies, which protect the genome from TE proliferation.

Across eukaryotes short (20-33 nucleotides) small RNAs are key to TE control. Within animals, Piwi-interacting small RNAs (piRNAs) are paramount in the TE defence armoury(Siomi et al., 2011). piRNAs are produced from defined genomic loci named piRNA clusters and after processing, associate with the Piwi subfamily of argonaute proteins(Brennecke et al., 2007). They recognise TEs through sense-antisense base pairing and target TEs for transcriptional and post-transcriptional silencing(Siomi et al., 2011). In many model organisms, piRNAs are essential for fertility through their role in controlling TE proliferation in the germline(Weick and Miska, 2014)

The nematode *Caenorhabditis elegans* is a well-established model for small-RNA mediated silencing. piRNAs in *C. elegans* are unusual in that the two piRNA clusters on Chromosome IV are composed of individual RNA polymerase II (RNA pol II) transcription loci where each piRNA has its own upstream motif(Batista et al., 2008; Das et al., 2008; Ruby et al., 2006; Wang and Reinke, 2008). piRNA clusters are located within H3K27me3-rich chromatin, which, together with cis-acting RNA pol II pausing sequences downstream of the piRNA, enforce production of ~28 nucleotide piRNA precursors (Beltran et al., 2019[ref to be inserted after pub]). piRNA precursors are further trimmed to result in mature 21 nucleotide piRNAs with a Uracil as the first nucleotide (21U-RNAs), which associate with the *C. elegans* Piwi protein PRG-1(Batista et al., 2008; Das et al., 2008; Wang and Reinke, 2008). Downstream of PRG-1, piRNA silencing relies on a nematode-specific class of secondary small RNAs known as 22G-RNAs(Das et al., 2008). 22G-RNA synthesis is carried out by RNA-dependent RNA polymerases using the target RNA as a template, following initiation by piRNA target recognition(Pak and Fire, 2007). 22G-RNAs bind to Argonaute proteins and lead to transcriptional and post-transcriptional silencing of target RNAs(Yigit et al., 2006). Additionally 22G-RNAs can be transmitted transgenerationally(Buckley et al., 2012) and as a result piRNA-initiated 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

to investigate the effect of mutations on protein-coding gene expression divergence(Denver et al., 2005; Landry et al., 2007; Rifkin et al., 2005). Here, we extend this approach to investigate the effect of mutations on TE expression divergence.

We created spontaneous MA lines of *C. elegans* that were descended from a single worm ancestor and propagated for ~400 generations under three population size treatments of N = 1, 10 and 100 individuals per generation(Katju et al., 2015). The varying population size treatment in the experiment permitted a manipulation of the strength of selection, with the N =1 lines evolving under close to neutral conditions (minimal selection) and an incremental increase in the strength of selection with increasing population size. We employed this framework to investigate how TE expression evolves under conditions of near neutrality and under the influence of increasing selection intensity. We show that overall TE expression increases in MA lines with the smallest population size. We further show that expression increase results in part from failure of piRNA-mediated silencing. Intriguingly, differences in the responses of different TEs to reduced piRNA-mediated silencing depend on the chromatin environment of the TE loci, such that TEs in repressed chromatin domains largely remain silent due to epigenetic memory imparted by 22G-RNAs, whilst in active chromatin domains, increased TE expression is much more likely to occur. Together our results demonstrate for the first time that robust TE control is under selection in animals. Importantly further, our results provide new insight into how the chromatin environment interacts with piRNA-mediated silencing to control TE expression.

Results

Relaxed selection leads to increased TE expression

In order to assess the effect of selection on TE expression, we generated mutation accumulation (MA) lines and propagated them by randomly selecting N individuals at each generation, where N was either 1, 10 or 100(Katju et al.,

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

In order to understand loss of repression of TE expression in more detail, we classified TEs into different families using RepeatMasker annotations and

compared the median TE expression for each TE family across different lines. TE expression was compared between *N*.1 and *N*>1 and tested for monotonic median increase with decreasing population size. As expected, the majority of TEs that showed a significant difference between *N*.1 and combined *N*.10 and *N*.100 lines had increased expression in *N*.1 compared to the other lines (Figure 1C and Supplemental Figure 2). However, individual TEs displayed different patterns of expression change. Some TEs, notably the DNA transposon Turmoil2, showed more consistent increases across the *N*.1 lines (Figure 1D). Indeed, the majority of the total effect on TE expression seen in Figure 1B could be attributed to one TE family, the Turmoil2 TEs, which showed a large expression increase across the majority of the *N*.1 lines (Figure 1C,D). Contrastingly some TEs, notably the Mariner family DNA transposon Tc1, showed a burst-like pattern of expression where a large increase in expression was observed in a few *N*.1 lines whilst Tc1 remained low in expression in the remainder (Figure 1E).

We next investigated the timecourse of TE desilencing during propagation of the MA lines. We performed gene expression analysis by RNAseq on 11 *N*.1 population size lines at 25 and 100 generations. Median total TE expression showed a highly significant increase with increasing numbers of generations (Jh p = 1.1e-9; Figure 1F). Linear regression analysis confirmed a positive relationship between increased numbers of generations and increased TE expression (LR p = 6.7e-8; Figure 1F). Different TEs showed different kinetics of desilencing. Turmoil2 showed a positive relationship between increased numbers of generations and increased expression (LR p = 3.9e-3) and a monotonic increase in median expression (Jh p = 2.55e-3; Figure 1G). Contrastingly, Tc1 desilencing did not show a positive linear relationship between increased generations and increased expression (LR p = 0.13; Figure 1H) though there was a significant increase in median expression (Jh p = 4.22e-6).

We further investigated whether the expression of TEs in individual MA lines correlated with the expression of other TEs. The majority of TEs showed little correlation with the expression of other TEs (Figure 2). The TEs with the

most statistically significant increases in expression, Turmoil2 and non-LTR retrotransposons of the LINE2 family (Figure 1C and Supplemental Figure 2) clustered together (Figure 2) suggesting coregulation of these TE families despite their different mechanisms of replication.

Expression of TEs is weakly associated with increased copy-number

TEs are capable of replicating independently of the host genome and thus their copy-number might change across MA lines. We sequenced the genomes of the MA lines after 400 generations and mapped the reads to consensus TE sequences thereby obtaining estimates of copy-number variation (CNV) for each TE family. Median TE copy-number increased with decreasing population size (Jh p = 1e-7; LR p=1e-3 Figure 3A). Within the *N*.1 lines, there was a weak significant positive correlation between increased copy-number and increased expression. In contrast, we found no significant correlation between copy-number and expression in the *N*.10 or *N*.100 lines (Figure 3B). Moreover, increased expression of specific TEs in individual MA lines was often unaccompanied by increased copy-number (Figure 3C,D). We conclude that increased TE copy number is not the primary cause of increased TE expression.

Alterations in small RNA levels are associated with TE expression changes

Since changes in TE copy number are unlikely to explain altered TE expression, we investigated whether changes in regulation of TE expression could explain the loss of silencing observed during mutation accumulation. In *C. elegans*, piRNAs and 22G-RNAs are important small RNA classes involved in TE silencing(de Albuquerque et al., 2015; Bagijn et al., 2012; Das et al., 2008; Phillips et al., 2015). To test whether piRNAs are important in the loss of silencing of TEs in *N*.1 lines, we remapped recently published cross-linking immunoprecipitation (CLIP) data(Shen et al., 2018) to identify TE transcripts that are bound by piRNA-Piwi complexes *in vivo*. Approximately 25% of TE transcripts with RNAseq reads were targeted by piRNAs. TEs targeted by piRNAs showed a statistically significant increase in total expression in *N*.1 lines (Jh p = 0.02; LR p = 0.033;

Figure 4A) whilst TEs that were not targeted by piRNAs were not significantly altered (Jh $p = 0.47$, LR $p = 0.27$; Figure 4B).

We tested whether defective piRNA-mediated silencing might account for increased expression of TEs targeted by piRNAs. piRNAs in *C. elegans* are expressed from individual promoters as a result of RNA polymerase II transcription(Billi et al., 2013; Cecere et al., 2012; Gu et al., 2012). We considered two potential mechanisms that might give rise to altered piRNAs. One possibility is that mutations in the piRNA sequences might occur in individual lines, which might affect their ability to recognise transposable elements and thus interfere with TE silencing. There was a trend for *N*.1 lines to have more mutations in piRNAs than *N*.10 or *N*.100 lines (Supplemental Figure 3). However, the trend was not significant (Jh p>0.05). Moreover, we identified on average only 1.1 piRNA sequences with mutations across the *N*.1 lines. Thus we conclude that the increase in TE expression is unlikely to be related to mutations in specific piRNAs.

We next considered the expression of individual piRNA loci. We identified piRNA loci with significantly altered median expression between *N*.1 lines and *N*.10 and *N*.100 lines combined using the Wilcoxon unpaired test. A small percentage of piRNA loci showed significant differences but overall there was no trend for these loci to show reduced expression in the *N*.1 lines; indeed these loci were more likely to have higher expression in the *N*.1 lines (Figure 4C). There was also no change in the expression of piRNAs identified by CLIP(Shen et al., 2018) to be directly binding to TEs (Figure 4D). Thus changes in piRNA expression are unlikely to explain the changes we observed in TE expression.

22G-RNAs act downstream of piRNAs to bring about target silencing(Das et al., 2008). Surprisingly, although 22G-RNAs silence TEs, we found that the total levels of 22G-RNAs mapping to TEs were increased in lines with smaller population sizes (Jh p=0.05, LR p=0.068; Figure 4E). To examine this in more detail, we analysed 22G-RNAs at individual TEs. The majority of TEs showing significantly increased 22G-RNAs in *N*.1 lines were non-LTR transposons (Figure

4F). Notably, these TEs were not significantly increased in transcript levels in the *N*.1 lines (Figure 1C; Supplemental Figure 2). Thus, the global increase in 22G-RNA levels does not relate directly to the increased transcript levels of TEs in *N*.1 lines.

To compare directly how increased expression of TEs relates to 22G-RNA mediated silencing, for each TE we compared 22G-RNA levels in lines with increased expression to 22G-RNA levels in lines without increased expression. The only TE showing alterations in both expression and 22G-RNA levels was Turmoil2. Turmoil2 showed decreased small RNA levels in lines showing increased Turmoil2 expression (Figure 4G). Thus, the increased transcript levels of Turmoil2 elements is associated with reduced 22G-RNAs. We conclude that the increased transcript levels of some TEs may be caused by reduced small RNA-mediated silencing, but that some increases in TE transcript levels occur independently of 22G-RNA changes.

Chromatin environment determines relationship between 22G-RNA levels and TE expression

We investigated what determines the different changes in 22G-RNAs at different TE types. 22G-RNAs interact with chromatin modifying factors to control expression of TEs(McMurchy et al., 2017). piRNA-mediated silencing has been directly linked to the generation of H3K9me2/3 marked nucleosomes ("classical heterochromatin") and this has been proposed to be important for transcriptional silencing induced by piRNAs(Ashe et al., 2012; McMurchy et al., 2017; Shirayama et al., 2012; Woodhouse et al., 2018). Additionally, it is becoming clear that a large proportion of the autosomal DNA in *C. elegans* can be divided into active domains, containing H3K36me3 and germline expressed genes and regulated domains, containing H3K27me3 marked nucleosomes and silent genes(Evans et al., 2016; Liu et al., 2011). These domains are largely stable through development, including in adult germline(Evans et al., 2016; Rechtsteiner et al., 2010). We examined the influence of these three types of

chromatin on the response of TEs to mutation accumulation. TEs in regulated domains and in classical heterochromatin showed no significant overall increase in expression (Figure 5A,B). The rare examples of N.1 lines showing significantly higher expression of TEs in classical heterochromatin corresponded to lines in which Tc1 reactivation occurred, consistent with enrichment of Tc1 elements within these regions(McMurchy et al., 2017). Contrastingly, TEs in active domains had significantly higher expression in N.1 than *N*>1 lines (Wilcoxon unpaired test, $p = 0.02$). There was also a trend towards decreased median expression with increasing population size although this was on the border of significance (Jh, $p = 0.056$, LR $p=0.064$; Figure 5C).

We next examined 22G-RNAs mapping to TEs across different chromatin domains. TEs in repressed domains showed significantly increased levels of 22G-RNAs in *N*.1 lines relative to *N*.10 and *N*.100. However, in active domains there was no significant change in 22G-RNA levels (Figure 5D,E,F). Thus increased 22G-RNAs occur predominantly in TEs within repressed chromatin.

AT-rich sequences in TEs reduce 22G-RNA generation

A recent study has demonstrated that silencing of both transgenes and endogenous genes by 22G-RNAs is inhibited by a high content of periodic repeats of AT-rich sequences, known as PATCs(Frøkjær-Jensen et al., 2016). We tested how PATC density within TEs influenced their expression under reduced selection. High PATC density corresponded to reactivation of TEs (Jh $p = 0.018$, LR p=0.02) whereas TEs with low PATC density did not show an increase in expression (Figure 6A). Contrastingly, only TEs with low PATC density showed significantly increased 22G-RNAs in *N*.1 lines relative to *N*.10 and *N*.100 (Jh $p = 0.018$, LR p=0.011; Figure 6B). Importantly this effect was specific to PATC sequences as GC-content alone had no significant effect on either TE expression or small RNA generation (Supplemental Figure 4A,B). We conclude that low PATC density is required for 22G-RNA generation, which may be required to 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

this observation is unclear. Whilst it is possible that increased expression of TEs may become beneficial under certain conditions, such as fluctuating environmental conditions or pathogen infection, our observations strongly suggest that the expression of most TEs is largely detrimental and TE expression is under purifying selection. Thus our research would support the proposal that beneficial events of TEs within genomes represent the exception- "making the best of a bad job"- rather than the rule. However, an important possibility that remains to be tested is whether TE activation is beneficial under fluctuating environmental conditions as opposed to the stable environment of laboratory maintained MA lines.

Phenotypic analyses of previous *C. elegans* MA experiments suggest that the decline in fitness in $N = 1$ lines results primarily from a few mutations of large effect(Estes et al., 2004; Halligan et al., 2003; Keightley and Cabellero, 1997). Similar results have been obtained in experimental evolution studies in *Drosophila melanogaster*(Ávila and García-Dorado, 2002) and bacteria(Dillon and Cooper, 2016; Heilbron et al., 2014). In our study, purifying selection at larger population sizes (*N*.10 and *N*.100) would eliminate such large-effect mutations and indeed, our *N*.10 and N. 100 lines exhibited no evidence of fitness reduction over the course of successive bottlenecking for 409 generations(Katju et al., 2015). Analysis of gene expression data from MA experiments from *C. elegans*, *D. melanogaster* and *S. cerevisiae* concluded that large effect mutations are also responsible for changes in protein-coding gene expression(Hodgins-Davis et al., 2015). . However, our results for TE expression do not seem to fit with this model because overall TE expression, which is largely driven by Turmoil2 and non-LTR elements (see Figure 1), increases gradually with time across the *N*.1 lines and is also increased, although less so, in *N*.10 and *N*.100 lines. The expression of Tc1 is an exception to the overall trend as it is not affected in *N*.10 or *N*.100 lines but a small number of *N*.1 lines show markedly increased Tc1 expression. Thus Tc1 reactivation may be dominated by a few mutations with large effect. This difference might be related to the different mechanism of silencing of Tc1 compared to Turmoil2 elements as discussed further below.

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

Weakened piRNA silencing is responsible for increased expression of TEs under relaxed selection.

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

New insights into the role of chromatin in the piRNA pathway

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

chromatin are much less prone to reactivation than those in active chromatin regions. Mechanistically this is because 22G-RNAs in repressive chromatin regions are stable or even increased, whilst 22G-RNAs mapping to TEs in active regions are reduced in MA lines with increased expression. We show that this also applies in a previously described experimental model where piRNAs are removed without removing the 22G-RNA pathway(Phillips et al., 2015). 22G-RNAs mapping to TEs within active domains are lost whereas those within repressed regions are maintained. This result may also explain why reactivation of Tc1 elements occurs less frequently than Turmoil2 elements, because Tc1 elements are predominantly located in repressed domains and are therefore silenced more robustly.

What is the mechanism whereby silencing memory is supported in repressed chromatin regions? The simplest possibility is that silencing and generation of 22G-RNAs are directly promoted by repressive chromatin modifications. In line with this possibility a mutually reinforcing loop between H3K9 methylation and small RNAs is well documented in fission yeast(Bühler, 2009), and H3K9 methylation factors contribute to silencing of transgenes in *C. elegans*(Ashe et al., 2012; McMurchy et al., 2017; Shirayama et al., 2012; Woodhouse et al., 2018) although the situation is more complicated for endogenous genes(Ni et al., 2014). However, our observations hold equally well for H3K27me3-repressed chromatin, which has not been directly linked to 22G-RNA silencing. We propose therefore that the nuclear small RNA pathway responds differently depending on whether surrounding genes are active or repressed to detect and quell aberrant gene activation. This model will be of interest for further mechanistic investigation of small-RNA mediated silencing in *C. elegans*.

Figure Legends

Figure 1 Increased expression of transposable elements in mutation accumulation lines. **A** Diagram of mutation accumulation (MA) experimental design. **B** Overall transposable element expression in MA lines as a function of population size. **C** Volcano plot for individual TE genes coloured according to family. The p value is a Wilcoxon unpaired test comparing the median in *N*.1 lines to *N*.10 and *N*.100. **D** Expression changes in Turmoil2 elements. The box shows interquartile range with a line at median and the whiskers extend to the furthest point <=1.5 times the interquartile range from the median. **E** Expression changes in Tc1 elements. Boxplot as in D. **F** Total TE expression after different numbers of generations of mutation accumulation in lines with a population size of 1. **G** Turmoil2 expression after different numbers of generations of mutation accumulation with a population size of 1. **H** Tc1 expression after different numbers of generations of mutation accumulation with a population size of 1.

Figure 2 Heatmap showing the significance in correlation in expression between different TEs across all the MA lines. The colour intensity shows the significance of the Spearman's rank correlation coefficient with blue showing a negative correlation and orange a positive correlation.

Figure 3 Copy-number increases are weakly associated with increased TE expression in MA lines. **A** Copy-number of TEs are increased with weaker selection. Data shows mean of TE copy-number across all TEs normalized to the mean across all lines. **B** Correlation between total change in TE copy-number across all TEs and total TE expression. **C&D** Representative examples (Turmoil2 and non-LTR transposons) showing how copy-number changes in high or expression lines. High expression lines are defined as being >1 standard deviation away from the mean expression across all lines.

Figure 4 Perturbed 22G-RNAs are associated with changes in TE expression in MA lines. **A&B** Comparison of change in TE expression for piRNA targeted and non-piRNA targeted TEs. **C** Total piRNA levels in MA lines across different population sizes. **D** Levels of piRNAs shown to directly bind to TEs by CLIP-seq, across different population size MA lines. **E** Levels of 22G-RNAs mapping to all

TEs in MA lines at different population sizes. **F** Volcano plot showing the gradient of change in 22G levels as estimated by a linear model against the p-value for the linear model on the y axis. G 22G-RNAs mapping to Turmoil2 in lines showing increased Turmoil2 expression compared to lines where Turmoil2 expression was not increased.

Figure 5 Chromatin environment controls alterations in piRNA mediated silencing in MA lines. **A,B,C** Expression changes in TEs in Regulated (H3K27me3), Classic Heterochromatin (H3K9me2) and Active (H3K36me3) chromatin domains. *P*-values are Wilcoxon unpaired test. **D,E,F** Changes in TE-mapping 22Gs in Regulated, Classic Heterochromatin and Active domains.

Figure 6 PATC content influences 22G-RNA generation in MA lines

A Expression changes in TEs in 4 equal size bins of 21 TEs with decreasing PATC content. **B** 22G-RNAs across the bins used in A. **C** and **D** Stratification of bins from A and B into active (H3K36me3) and regulated (H3K27me3) chromatin domains. High PATC is the top bin and low PATC is the lowest bin.

Figure 7 Chromatin and PATC together influence TE control in MA lines

A Alterations in 22G-RNAs in the absence of *prg-1* whilst 22G-RNA biogenesis pathways are still active for TEs divided into Regulated, Classic Heterochromatin and Active domains. **B** Alterations in 22G-RNAs in the absence of both *prg-1* and 22G-RNA biogenesis pathways for TEs from Regulated, Classic Heterochromatin and Active domains. Boxplots in A and B as in Figure 1D.

C Model for how chromatin environment contributes to TE desilencing in MA lines. Left- repressed chromatin domains, taking H3K27me3 enriched domains as an example. Right, Active chromatin domains.

Supplemental Figure 1

A Fano factor of individual TE transcript levels across all lines of the indicated 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

Methods

RNA Library Preparation, Sequencing, and Analysis of Transcript Abundance.

The library preparation and RNA-sequencing procedures have previously been described in detail (Konrad et al., 2018). Briefly, we isolated one, two, and three individuals each from populations of sizes N = 1, 10, and 100, respectively. These 55 worms, as well as one individual from the ancestral population were each sequestered to NGM plates seeded with OP50, where they were allowed to self-fertilize and reproduce at 20ºC. Three offspring worms at the L4 larval stage were isolated from each of the F_1 populations to serve as biological replicates. These 168 individual worm samples were allowed to reproduce for three generations to yield enough tissue for RNA extraction. A standard bleaching protocol was used to collect gravid eggs from adults in order to generate synchronized populations of L1 larvae. Total RNA was isolated from L1 larvae via the Qiagen RNeasy Mini Kit. The Nanodrop 2000, Qubit 3.0 Fluorometer, and an Agilent RNA Analyzer were used to evaluate the quality of the RNA samples, the and Illumina TruSeq RNA library Prep Kit v2 was used with standard procedures to prepare the RNA sequencing libraries for each sample at the Texas A&M University Genomics and Bioinformatics Services Center. The RNA was fragmented and Illumina adapters were annealed for amplification. Size selected cDNA fragments were isolated via a Qiagen Gel Extraction Kit. Finally, sequencing was performed on the Illumina HiSeq 4000 platform with default quality filters.

Demultiplexing and prefiltering of the sequencing reads was performed based on default Illumina QC protocols. Reads containing abnormally short insert lengths were removed, and adapters were discarted from the reads.

The raw RNA-sequencing reads in fastq format were aligned to the protein-coding transcriptome of *C. elegans* (Wormbase reference N2 genome version WS247) using TopHat(Trapnell et al., 2009) via the “very sensitive” bowtie2 algorithm with a maximum of one mismatch in the anchor region for each spliced alignment and a minimum and maximum intron length of 20 and 3,000 bp, respectively. Cufflinks(Trapnell et al., 2010) with default settings and gene annotations from the N2 genome version WS247 was used to estimate the relative transcript abundance for each protein-coding gene. All following analyses were focused on FPKM values calculated on the per gene level. The relative transcript abundances (FPKM) from the three biological replicates for each original sample were averaged to get mean relative transcript abundance for each gene in that sample.

Small RNA sequencing

MA lines were synchronised using hypochlorite treatment and embryos were isolated after 12 hours. RNA was extracted using trizol and small RNA libraries

were prepared using the Illumina Small RNA sequencing kit as described previously(Sarkies et al., 2015).

Small RNAs were aligned to a genome built using bowtie from a fasta file containing all piRNAs, miRNAs, ncRNAs and genes including TEs, extracted from Wormbase (WS264; ce11), requiring perfect mapping. Reads mapping to the sense strand of ncRNAs and miRNAs were extracted using bedtools intersect –c –S. We used DEseq using ncRNAs and miRNAs to extract size factors. 22G-RNAs mapping to TEs and genes were extracted using a custom Perl script and the number of 22G-RNAs mapping antisense to each gene and TE was then counted using bedtools intersect –c -s. The 22G-RNAs were then normalized to the size factors from ncRNAs and miRNAs combined.

TE copy-number analysis

Alignments to determe TE copy number changes in MA lines.

DNA sequencing was aligned to a genome built using bowtie2-build from TE consensus sequences extracted from repbase combined with all coding sequences. Bowtie2 was used to map PE reads to this genome and the readcount mapping to each cds or TE was obtained using bedtools intersect –c.

Data integration

Computational analysis of TE expression and small RNA analysis

TEs from WS264 were annotated using Repeatmasker(Smit et al., 2017). All data analysis was conducted using the R environment for statistical analysis (www.Rproject.com). Details of the individual analyses are documented in the R markdown file accompanying this manuscript (additional file 1), with the raw data tables required to run these programs in as a zipped file (additional file 2). Previously published datasets containing Chromatin domain annotations from Early Embryo ChiP-Seq were taken from(Evans et al., 2016) updated to WS264 using liftover (https://genome.ucsc.edu/). Small RNA sequencing data from reactivation of small RNA pathways in the presence or absence of piRNAs was taken from (Phillips et al., 2015)and aligned to the *C. elegans* genome as described above. The average PATC score for each TE was calculated by taking

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

REFERENCES

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

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

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

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

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

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

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

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

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

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

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

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

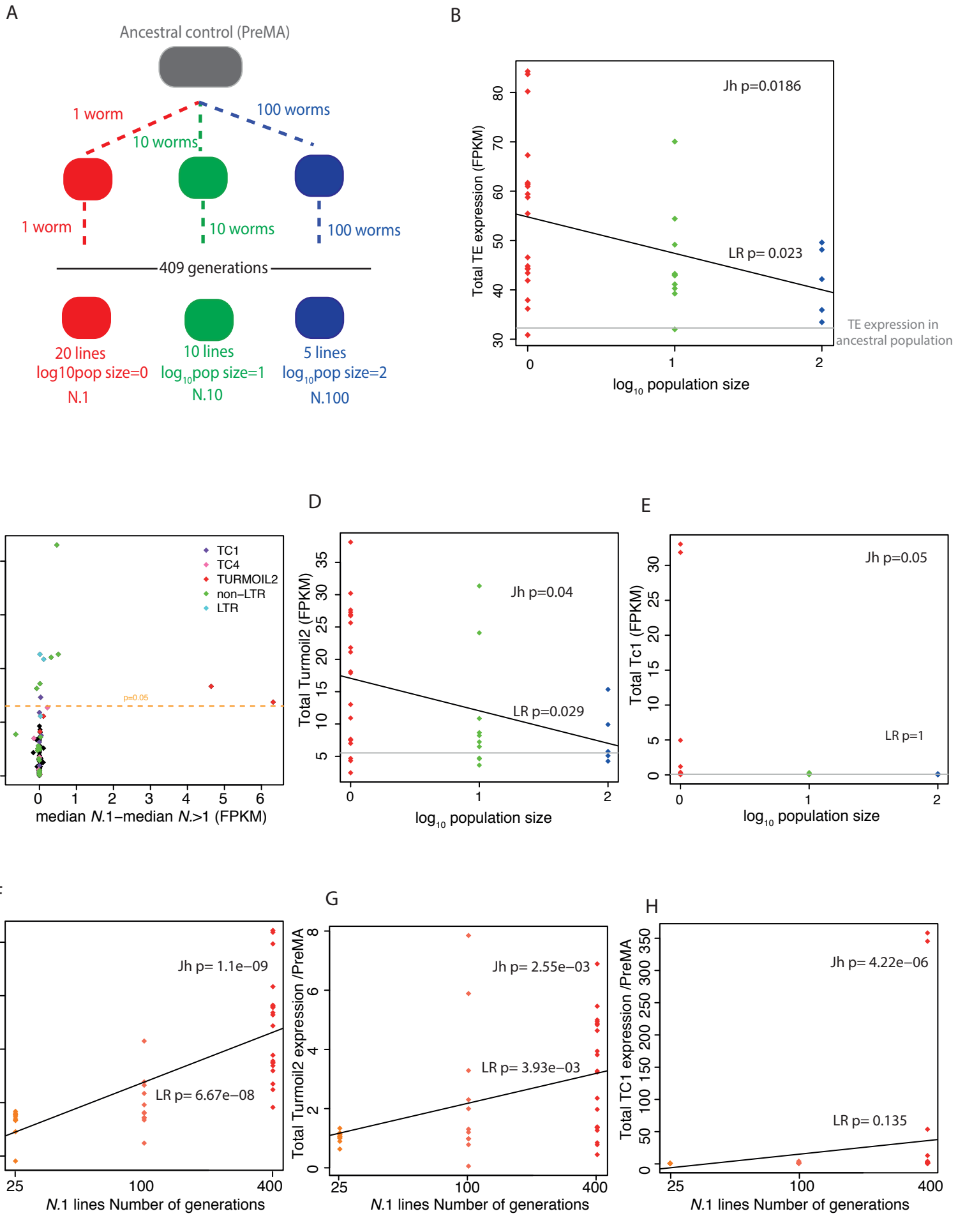


Figure 1

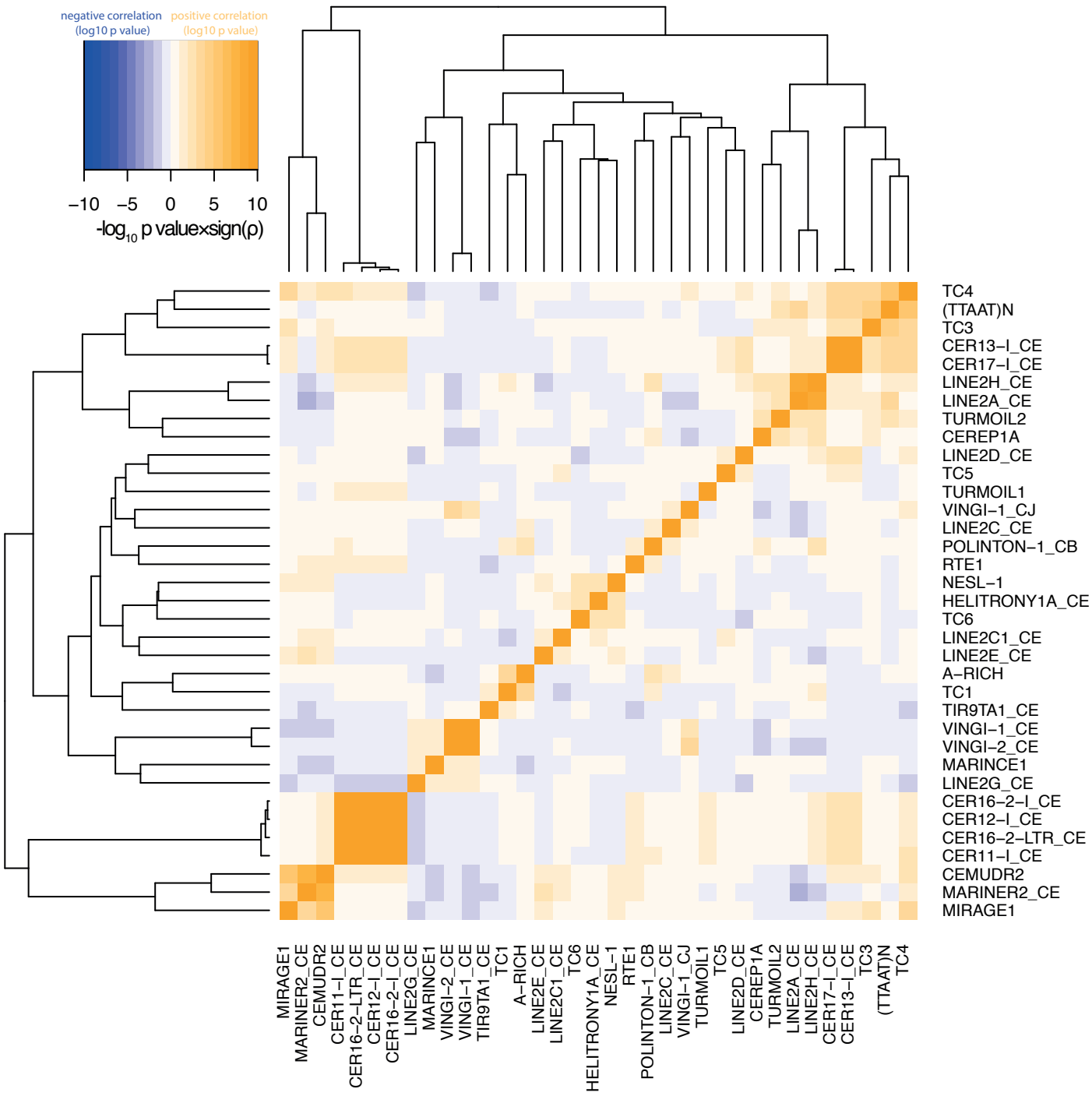


Figure 2

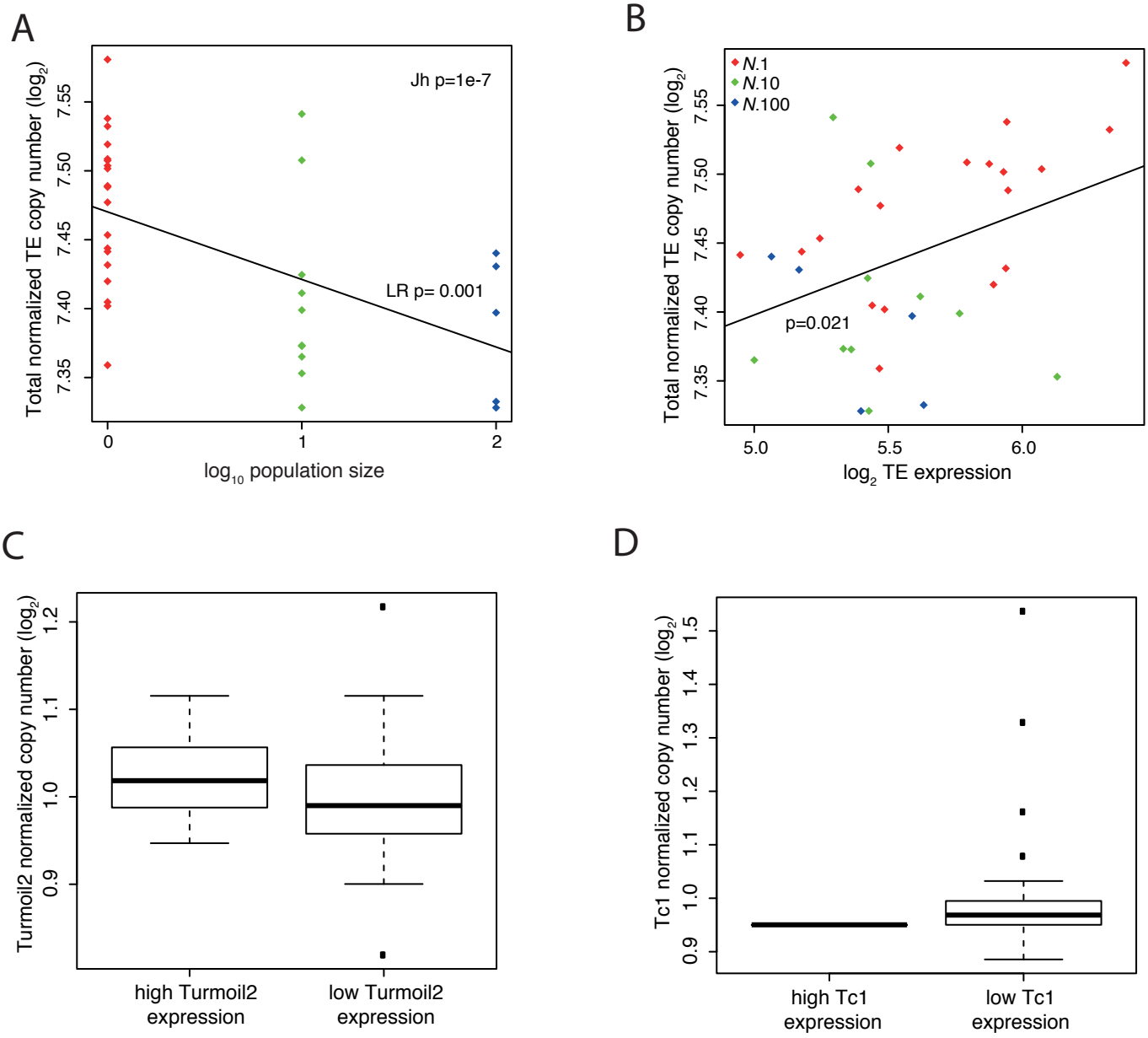


Figure 3

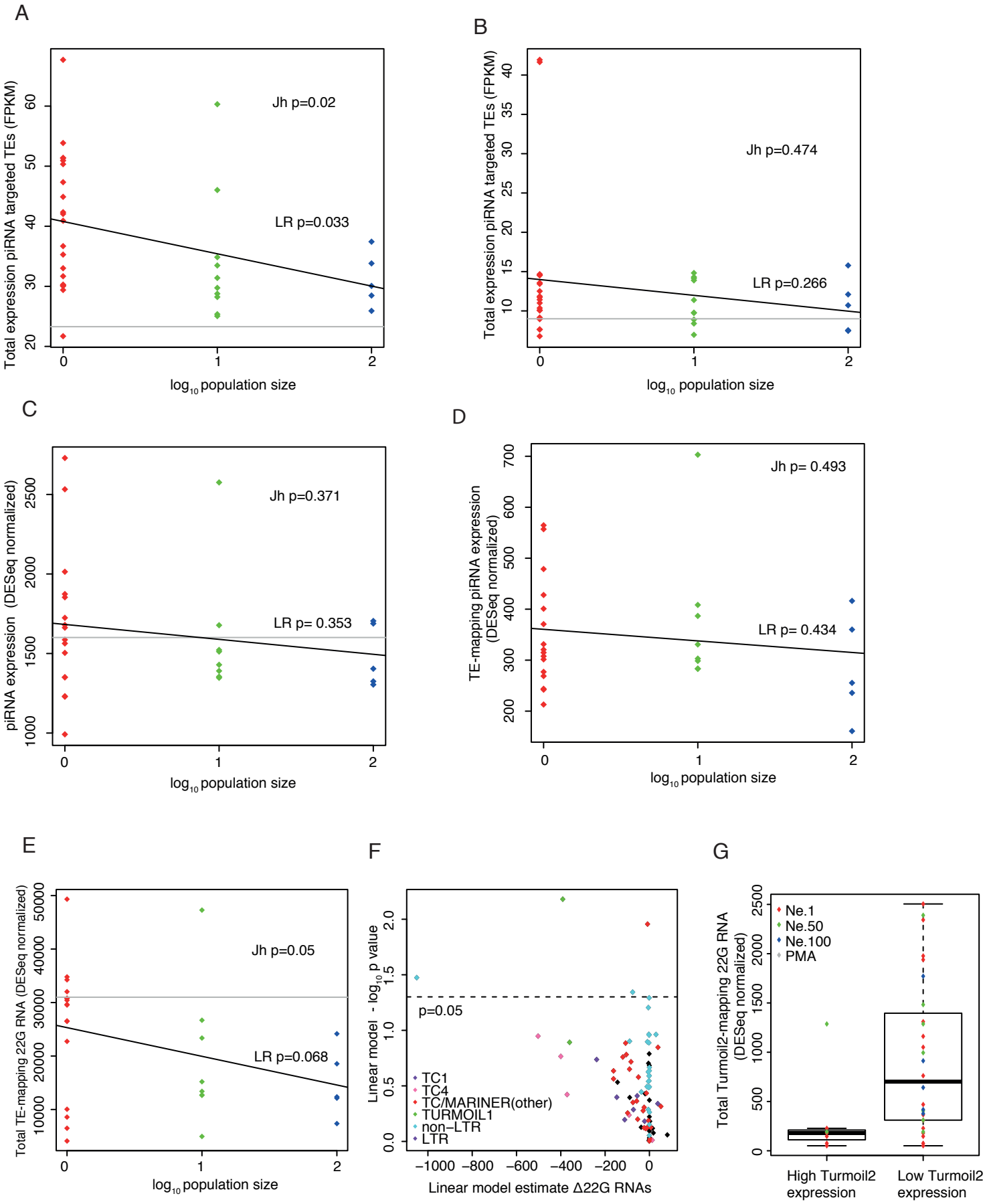


Figure 4

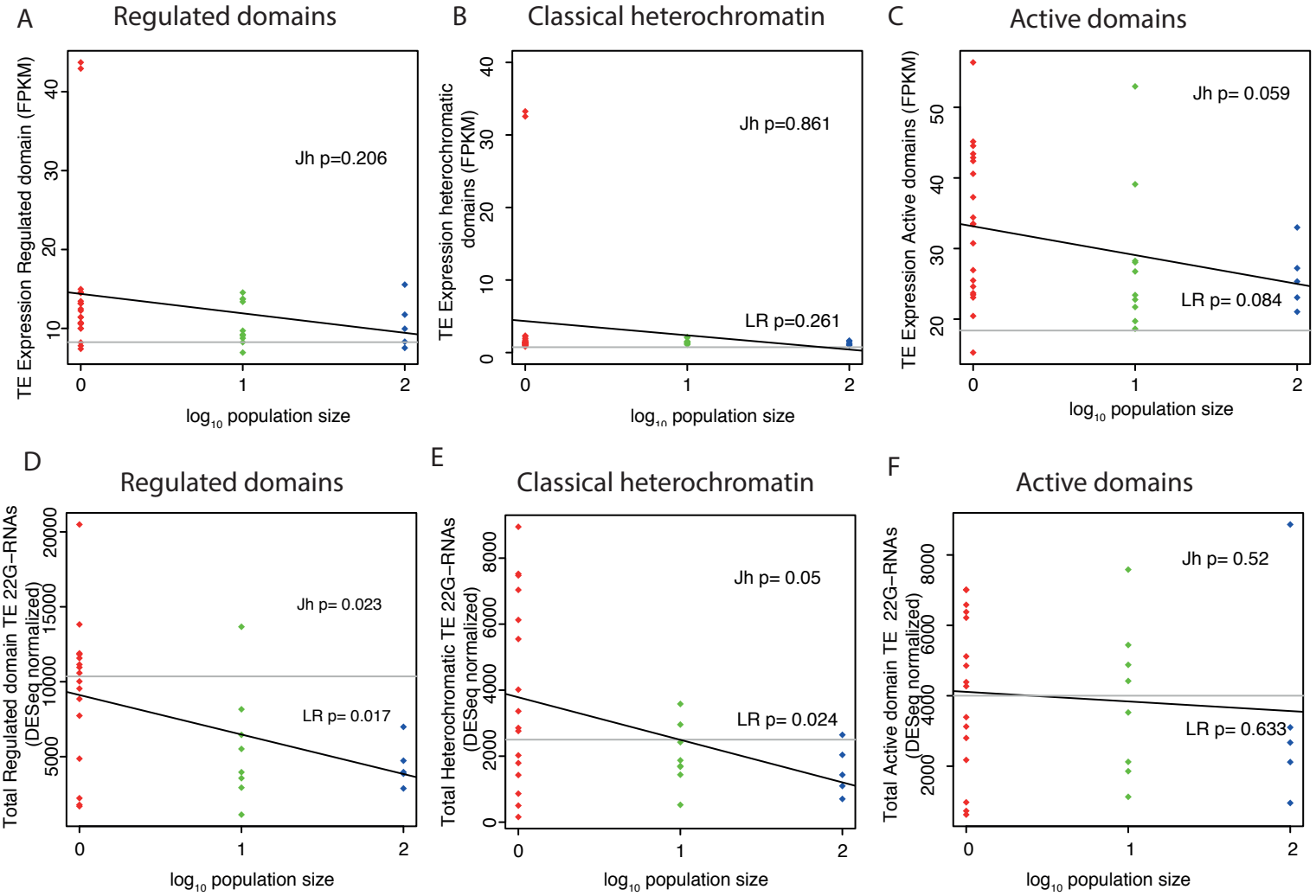


Figure 5

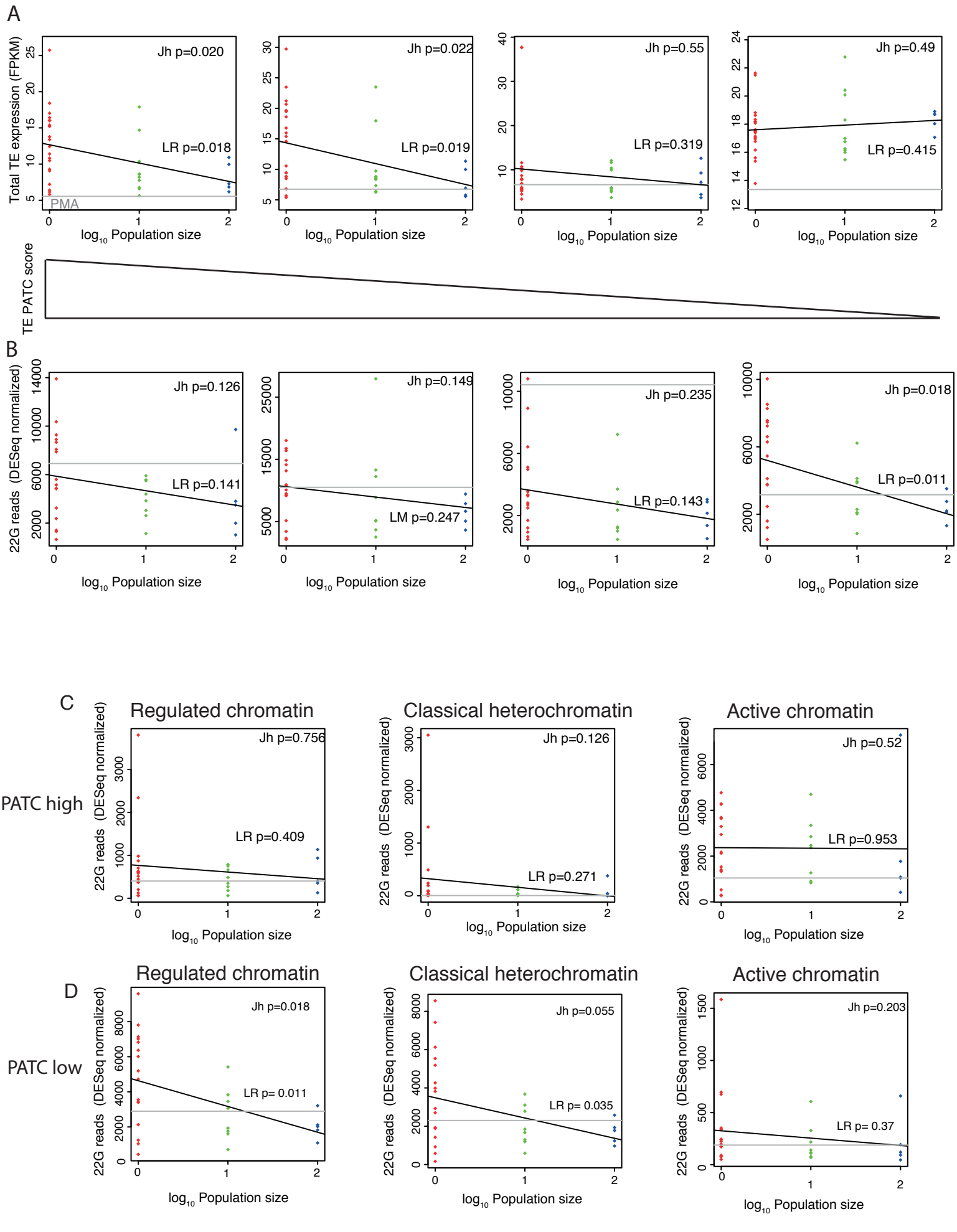
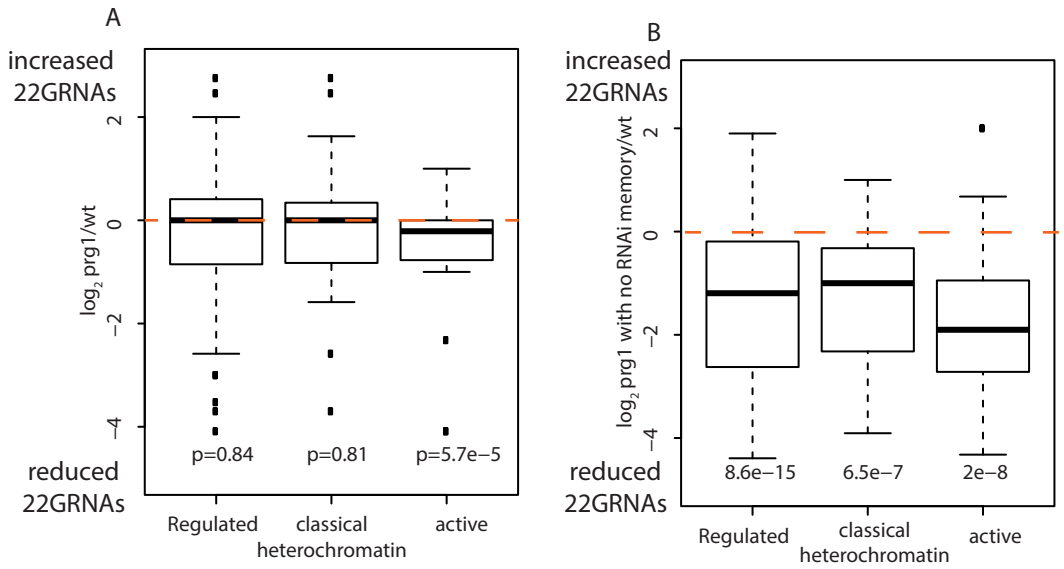


Figure 6



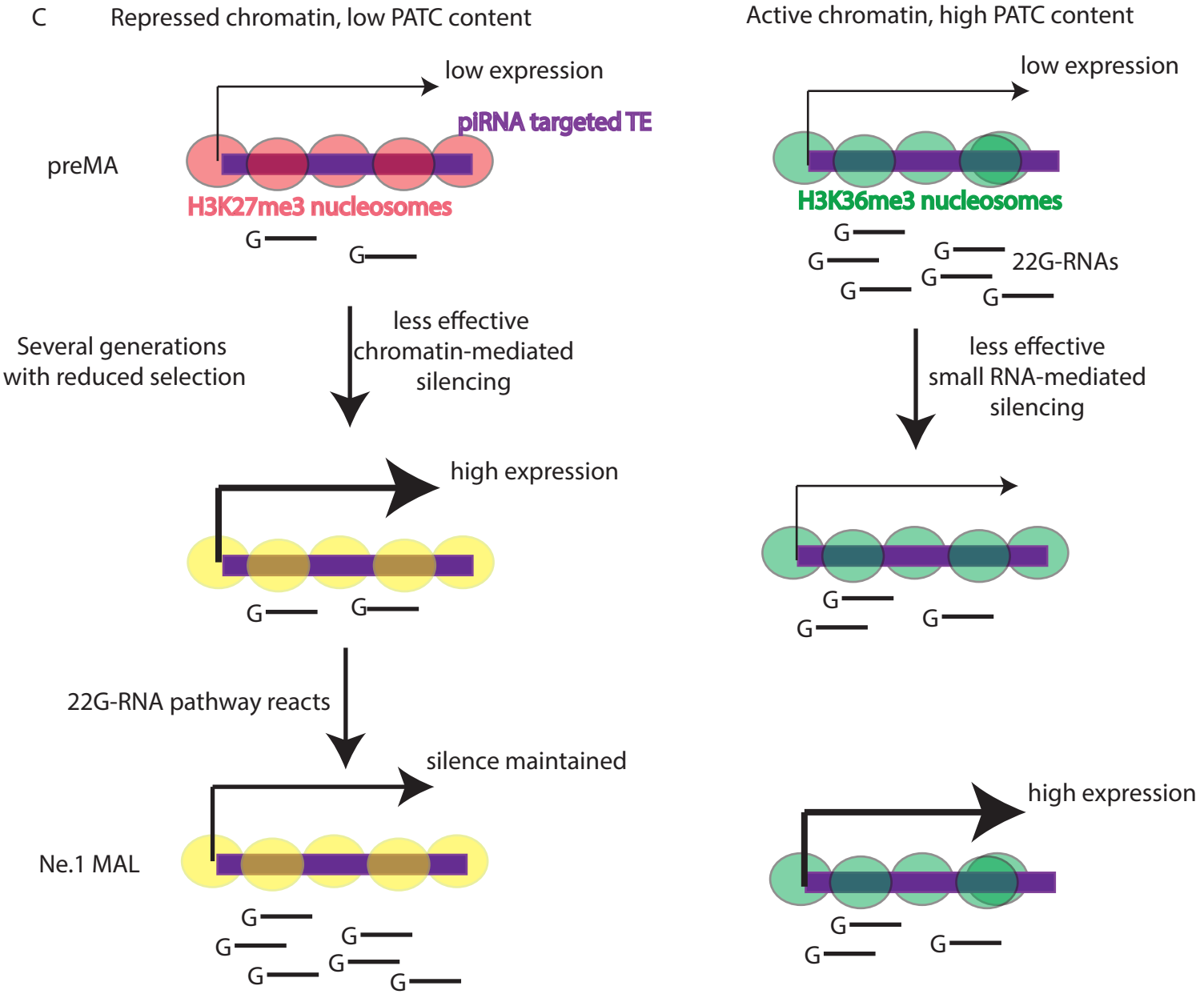
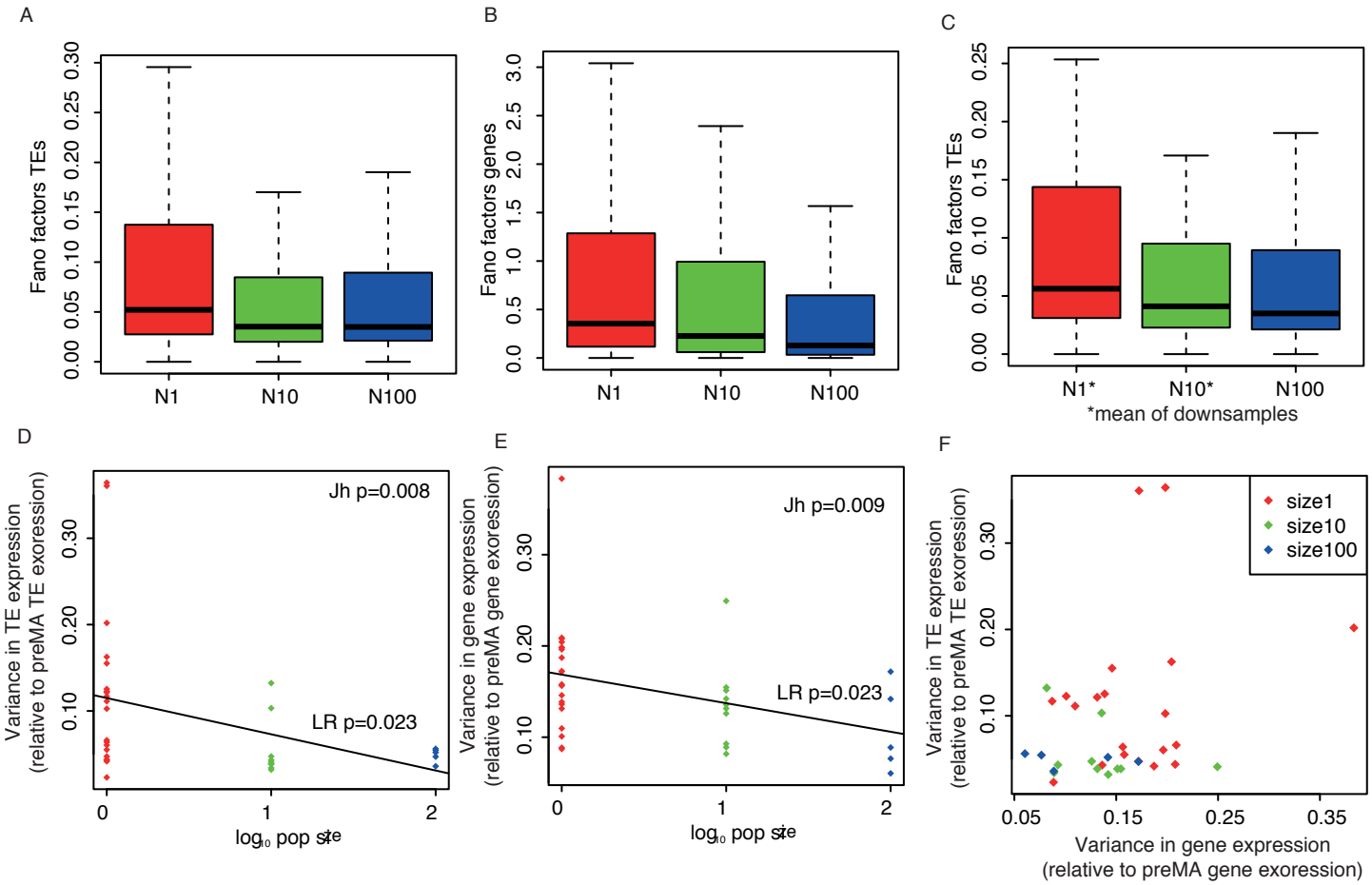
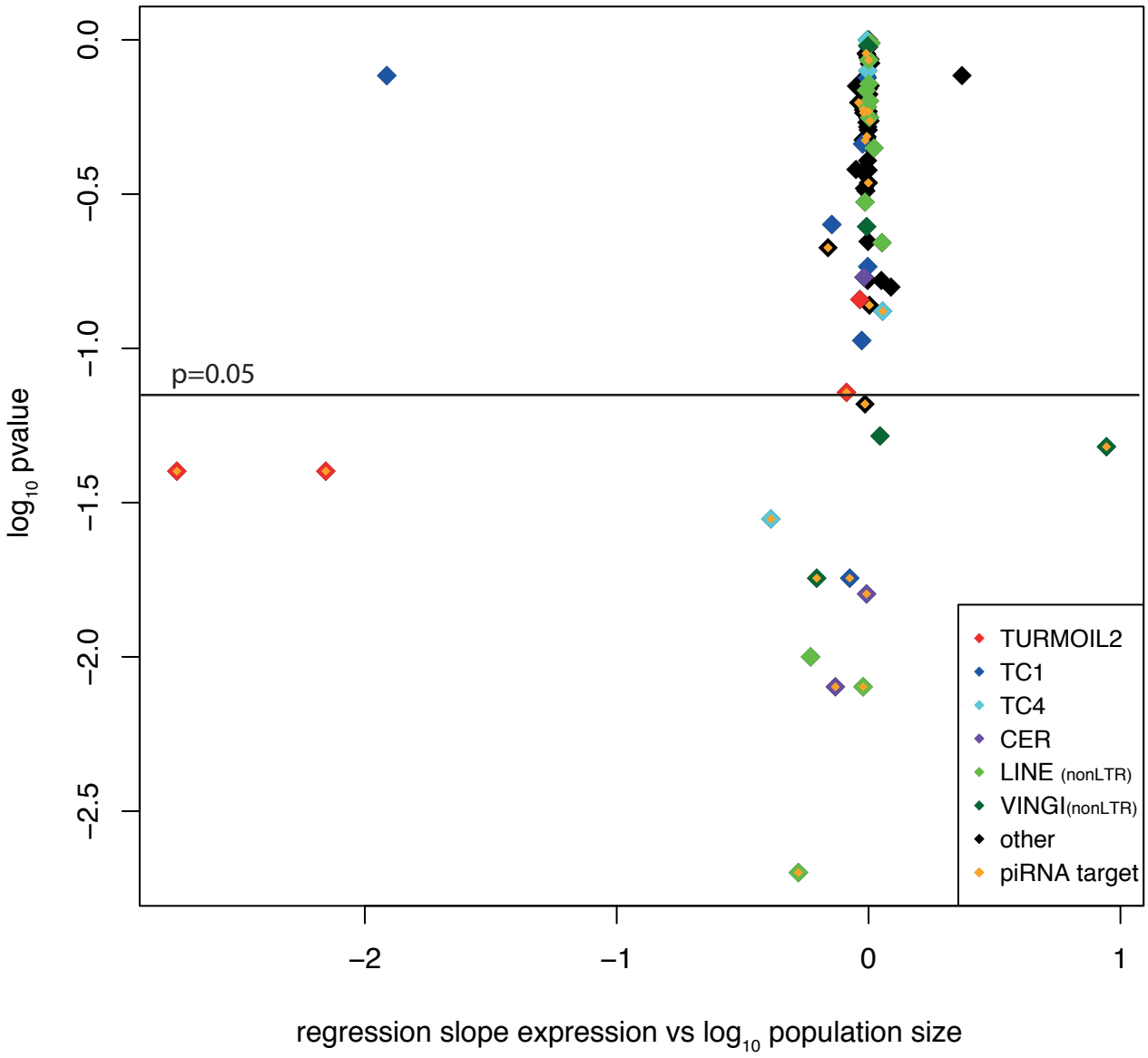


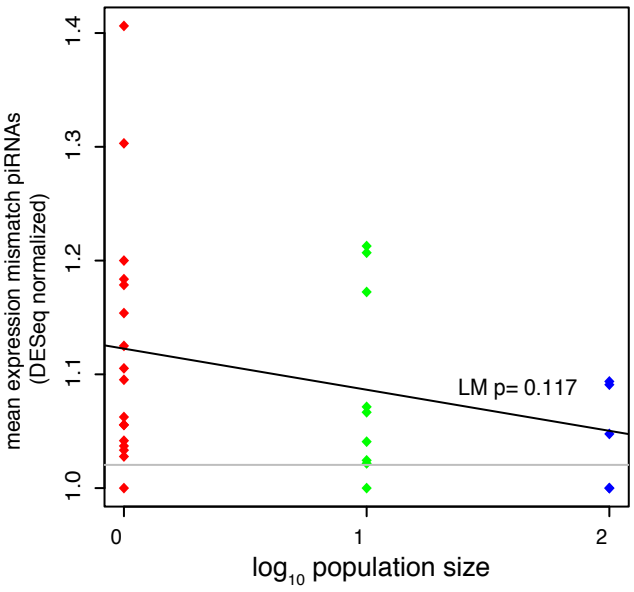
Figure 7



Supplemental Figure 1



Supplemental Figure 2



Supplemental Figure 3

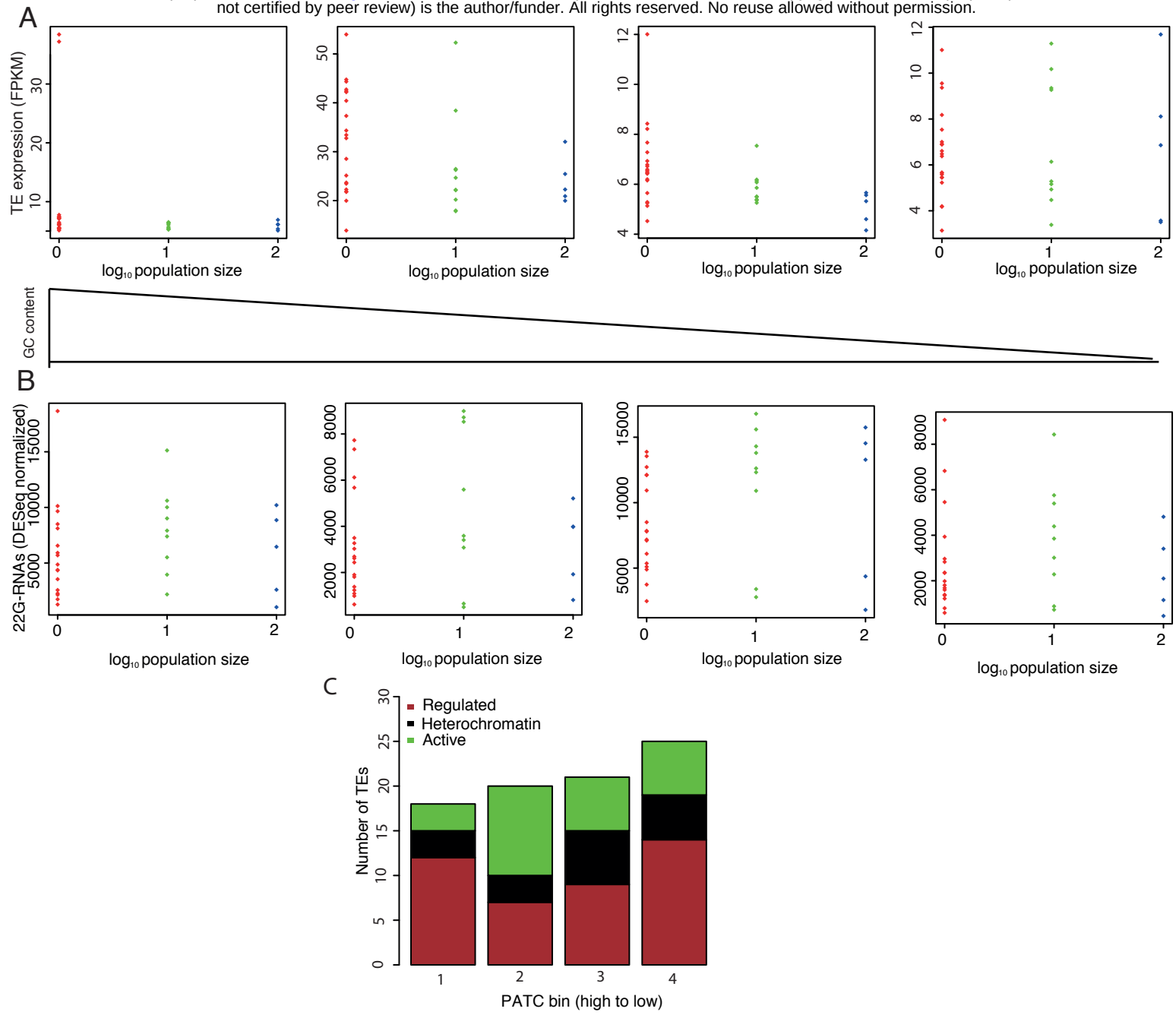


Figure S4