Segregating Complete Tf2 Elements Are Largely Neutral in Fission Yeast

Yan Wang, Qin Wang, Zhiwei Wu, and Guan-Zhu Han (iD)*

College of Life Sciences, Nanjing Normal University, Nanjing, China

*Corresponding author: E-mail: guanzhu@njnu.edu.cn.

Abstract

Transposable elements (TEs) comprise a large proportion of the eukaryote genomes. Yet it remains poorly understood how TEs influence the fitness of the hosts carrying them. Here, we empirically test the impact of TEs on the host fitness in the fission yeast *Schizosaccharomyces pombe*. We find that two families of TEs (Tf1 and Tf2 elements), both of which belong to long terminal repeat retrotransposons, are highly polymorphic among individual *S. pombe* strains. Only 13 complete Tf2 elements are identified in *S. pombe* laboratory strain 972. These 13 Tf2 elements integrated into host genomes in very recent time and are segregating within the *S. pombe* population. Through knocking out each of the 13 Tf2 elements in *S. pombe* strain 972, we find Tf2 knockout does not affect the host fitness, and Tf2 elements do not alter the expression of nearby genes. Challenged by diverse forms of stress, the Tf2 knockout strains do not exhibit different growth rates from wild-type strain. Together, we conclude that segregating complete Tf2 elements insertions are largely neutral to host fitness in the fission yeast. Our study provides genome-wide empirical support for the selfish nature of TEs in fission yeast.

Key words: transposable elements, fitness, selfish DNA, yeast.

Significance

Few genome-wide empirical tests have been performed to explore the effects of transposable elements (TEs) on the host fitness. We knocked out each of the 13 complete TEs in *Schizosaccharomyces pombe* strain 972, and found that these segregating TEs neither have significant effect on the host fitness nor alter the expression of nearby genes. Our study provides empirical support for neutral evolution and selfish nature of TEs in fission yeast.

Introduction

Transposable elements (TEs) comprise a substantial fraction of the genomes of eukaryotes (Wicker et al. 2007; Wells and Feschotte 2020). TEs are typically classified into two classes based on their transposition mechanisms, namely class I (retrotransposons) and class II (DNA transposons) (Wicker et al. 2007; Kapitonov and Jurka 2008; Wells and Feschotte 2020). Class I TEs replicate through RNA intermediates and reverse transcription, whereas most of class II TEs mobilize through excising themselves and moving to new genomic locations (Wicker et al. 2007; Wells and Feschotte 2020). Both classes can be further divided into distinct lower-ranking groups (subclasses, superfamilies, or families) based on mechanistic features, genome organizations, and phylogenetic analyses of hallmark genes (Wicker et al. 2007; Kapitonov and Jurka 2008).

Like most new mutations, the majority of new TE insertions have been thought to be deleterious to the host fitness (Lynch 2007; Arkhipova 2018; Cosby et al. 2019; Payer and Burns 2019). When a TE inserts into an essential cellular gene and disrupts its function, it could be a lethal mutation (Kazazian et al. 1988; Arkhipova 2018; Payer and Burns 2019). Indeed, some human genetic diseases are known to be caused by new TE insertions (Kazazian et al. 1988; Payer

and Burns 2019). These TE insertions that are selectively disadvantageous will be rapidly removed from the host population by purifying selection. However, nearly all the eukaryote genomes (with a few exceptions, such as the human malaria parasite *Plasmodium falciparum*) are known to harbor TEs (Gardner et al. 2002; Wells and Feschotte 2020). Many nonmutually exclusive hypotheses have been proposed to account for the function, maintenance, and evolution of TEs within host genomes (Charlesworth 1987; Finnegan 1989; Sundaram and Wang 2018). 1) TEs are colloquially referred to as "junk" DNA that conveys little selective advantage to the organism (Doolittle and Sapienza 1980); 2) TEs might induce mutations and facilitate genomic rearrangements which contribute the evolutionary potential of the host population in the long run. However, "evolution is not anticipatory; structures do not evolve because they might later prove useful" (Doolittle and Sapienza 1980); 3) TEs are selfish genomic parasites with two properties: amplifying by forming additional copies within the genome and making no specific contribution to the phenotype, that is, the selfish DNA hypothesis (Doolittle and Sapienza 1980; Orgel and Crick 1980); 4) The mutations caused by TEs might be beneficial, and fuel adaptation to changing environmental conditions (Daborn et al. 2002; Blanc and Adams 2003; Arkhipova 2018; Esnault et al. 2019). For example, the mutation giving rise to industrial melanism in the peppered moth *Biston betularia* is a TE insertion into the first intron of the gene cortex that is estimated to occur around 1819 (Van't Hof et al. 2016); and 5) TEs have long been postulated to have a fundamental role in gene regulation in eukaryotes (McClintock 1956; Britten and Davidson 1969). Donating the same transcription factor binding sites to nearby genes where they integrated into the genome, TEs might contribute to the origin, evolution, and rewiring of gene regulatory networks (Chuong et al. 2016, 2017; Lynch 2016; Nishihara 2019; Ullastres et al. 2021). In general, hosts can repurpose TEs to perform cellular function, a process known as co-option or exaptation (Brosius and Gould 1992; Cosby et al. 2019; Wang and Han 2020, 2021). Albeit with these hypotheses and models, it remains obscure and contentious about the role of TEs in the host genome. Vanishing few genome-wide empirical tests has been performed to directly interrogate the effect of TEs on the host fitness.

The fission yeast *Schizosaccharomyces pombe* represents an important model for studying the biology of eukaryotes (Rhind et al. 2011). For TEs, two closely related families of long terminal repeat (LTR) retrotransposons, designated Tf1 and Tf2, have been identified within diverse *S. pombe* strains (Levin et al. 1990; Hoff et al. 1998; Bowen et al. 2003; Esnault and Levin 2015). The genome of *S. pombe* laboratory strain 972 contains only 13 complete Tf2 retrotransposons, and 35 solo-LTRs generated by homologous recombination between the two LTRs flanking an LTR retrotransposon (Wood et al. 2002; Bowen et al. 2003). The chromosome locations and the copy numbers

of Tf1 and Tf2 are different among different *S. pombe* isolates (Levin et al. 1990). Although new Tf1 element insertions might provide a potential path to cope with environmental stress (Feng et al. 2013; Esnault et al. 2019), it remains unclear the effect of segregating TEs on the host fitness.

In this study, we empirically tested the effect of complete TEs on host fitness using *S. pombe* as the model. We found that complete TEs were polymorphic among *S. pombe* isolates. Through knocking out each of the 13 complete Tf2 elements in *S. pombe* laboratory strain 972, we found that these segregating complete Tf2 elements neither have a significant effect on the host fitness nor alter the expression of nearby genes. Our results might provide novel insights into the relationship between TEs and their hosts.

Results

TEs Are Polymorphic within the *S. pombe* Population

First, we analyzed the distribution and diversity of TEs within the *S. pombe* population. As previous studies (Levin et al. 1990; Hoff et al. 1998; Bowen et al. 2003; Esnault and Levin 2015), we identified two major families of LTR retrotransposons, namely Tf1 and Tf2 retrotransposons, and found that Tf1 and Tf2 elements are present within all the 19 *S. pombe* strains sampled globally (fig. 1A; supplementary tables S1–S3, Supplementary Material online). The copy numbers of complete Tf1 and Tf2 elements are different among *S. pombe* strains, ranging from 1 (in strain JB939) to 64 (in strain JB1110). Complete Tf2 elements are present in all the *S. pombe* strains studied, whereas complete Tf1 elements are not identified in nearly half (8/19) of *S. pombe* strains. The absence of Tf1 elements in some *S. pombe* strains is possibly due to internal region removal resulting from recombination between LTRs after insertions into host genomes, as a large number of Tf1 solo-LTRs were identified in all these *S. pombe* strains (fig. 1A). It should be noted that these numbers are rough estimates, because TEs represent obstacles in genome assembly. We built phylogenetic networks for all the complete Tf1 and Tf2 elements and found a complex network structure, indicating that recombination might have taken place within and between Tf1 and Tf2 elements (fig. 1B). These results suggest that Tf1 and Tf2 elements are highly polymorphic among individual fission yeast strains.

Consistent with previous studies (Bowen et al. 2003), we identified a total of 13 complete Tf2 elements within the genome of *S. pombe* laboratory strain 972 (fig. 1C). The 13 Tf2 elements of strain 972 are closely related and distributed at a tree-like region (fig. 1B). Indeed, no recombination signal was detected within these Tf2 elements with the exception of Tf2–11. Tf2–11 is derived from recombination between Tf2 and Tf1 with Tf1 contributing to part of 5′-LTR (fig. 1C; supplementary fig. S1, Supplementary Material online) (Bowen et al. 2003). Tf2–7 and Tf2–8 are tandem elements which share a
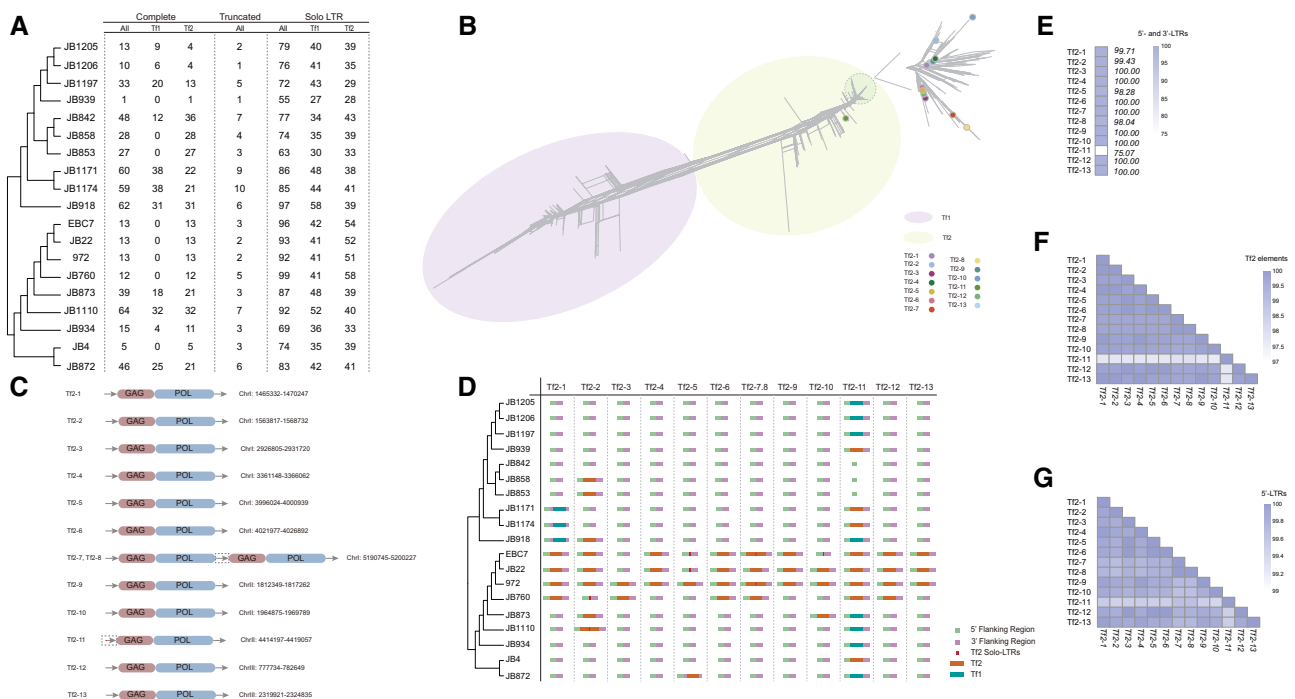
FIG. 1.—The evolution and distribution of Tf1 and Tf2 elements in *S. pombe*. (A) The distribution and copy numbers of Tf1 and Tf2 elements in 19 *S. pombe* strains sampled globally. Phylogenetic relationship among 19 *S. pombe* strains was reconstructed based on 30 randomly selected genes. The copy numbers of complete Tf elements, truncated Tf elements, solo LTRs for each strain are shown in the corresponding rows. (B) Phylogenetic networks for all the full length Tf1 and Tf2 elements in 19 *S. pombe* strains. The 13 Tf2 elements in strain 972 are indicated with filled circles with different colors and enlarged. (C) The structure and position of 13 Tf2 elements in strain 972. The pink and blue boxes represent GAG and POL proteins, and the arrows on both sides are 5'- and 3'-LTRs. The small red fragment within the 5' LTR of Tf2–11 indicates that it is derived from the LTR of Tf1. (D) The syntenies of 13 Tf2 elements in 19 *S. pombe* isolates. 5' flanking and 3' flanking regions are 1,000 bp sequences flanking Tf2 elements. The Tf1 and Tf2 elements are highlighted in blue and orange, respectively. (E–G) The pairwise nucleotide identity among 13 Tf2 elements. (E) The pairwise identity of 5'- and 3'- LTRs for each Tf2 element. (F) The pairwise identity between complete Tf2 elements. (G) The pairwise identity between complete Tf2 elements without 5'-LTR.

common internal LTR (fig. 1C). To further investigate the evolutionary history of the 13 Tf2 elements, we analyzed the syntenies among 19 *S. pombe* strains for each of these 13 Tf2 elements (fig. 1D). We found these 13 Tf2 elements are segregating within the *S. pombe* populations. In general, the integration of these Tf2 elements occurred among closely related strains. However, the integration pattern is occasionally inconsistent with strain phylogeny, which might be explained by: 1) Multiple independent integration into similar genomic positions, given Tf1 and Tf2 have integration preferences (Bowen et al. 2003). For Tf2–1, Tf2–2, Tf2–5, and Tf2–10, we found orthologous Tf2 elements do not cluster together, indicative of multiple independent integration (supplementary fig. S2, Supplementary Material online). 2) Introgression involving regions with TEs occurred among *S. pombe* strains. Nevertheless, our results show that the 13 complete Tf2 elements are segregating within the *S. pombe* population.

The pairwise identity between 5'- and 3'-LTRs for each complete Tf2 element is greater than 98%, except for Tf2–11 (only 75.07%) (fig. 1E), which is due to recombination between Tf2 and Tf1 in 5'-LTR (fig. 1C). Although the pairwise identity between complete Tf2 elements is greater than 97%

(fig. 1F), the pairwise identity between Tf2 elements (excluding 5'-LTR of each of Tf2 elements) is greater than 99% (fig. 1G). These lines of evidence suggest that Tf2 elements have been very recently proliferating during the evolutionary course leading to strain 972.

Tf2 Knockout Has No Effect on *S. pombe* Growth

To explore the effect of segregating complete Tf2 elements on the growth of *S. pombe*, we knocked out each of the 13 Tf2 elements in strain 972 using "deletion cassette" through homologous recombination with a KanMX marker gene flanked by two homologous fragments (fig. 2A). We generated a total of 12 Tf2 knockout (ΔTf2) strains, because Tf2–7 and Tf2–8 are in tandem with each other (ΔTf2–7 represent the strain without Tf2–7 and Tf2–8) (fig. 1B). To exclude the effect of the marker gene on the growth of *S. pombe*, we inserted a KanMX in the gene-poor region (Chromosome I (ChrI): 2,973,561-2,973,583), generating a knockin strain E1 (fig. 2B). The spot assay, a semi-quantitative method commonly used to analyze yeast growth phenotypes, shows that there was no significant difference between wild type (WT)
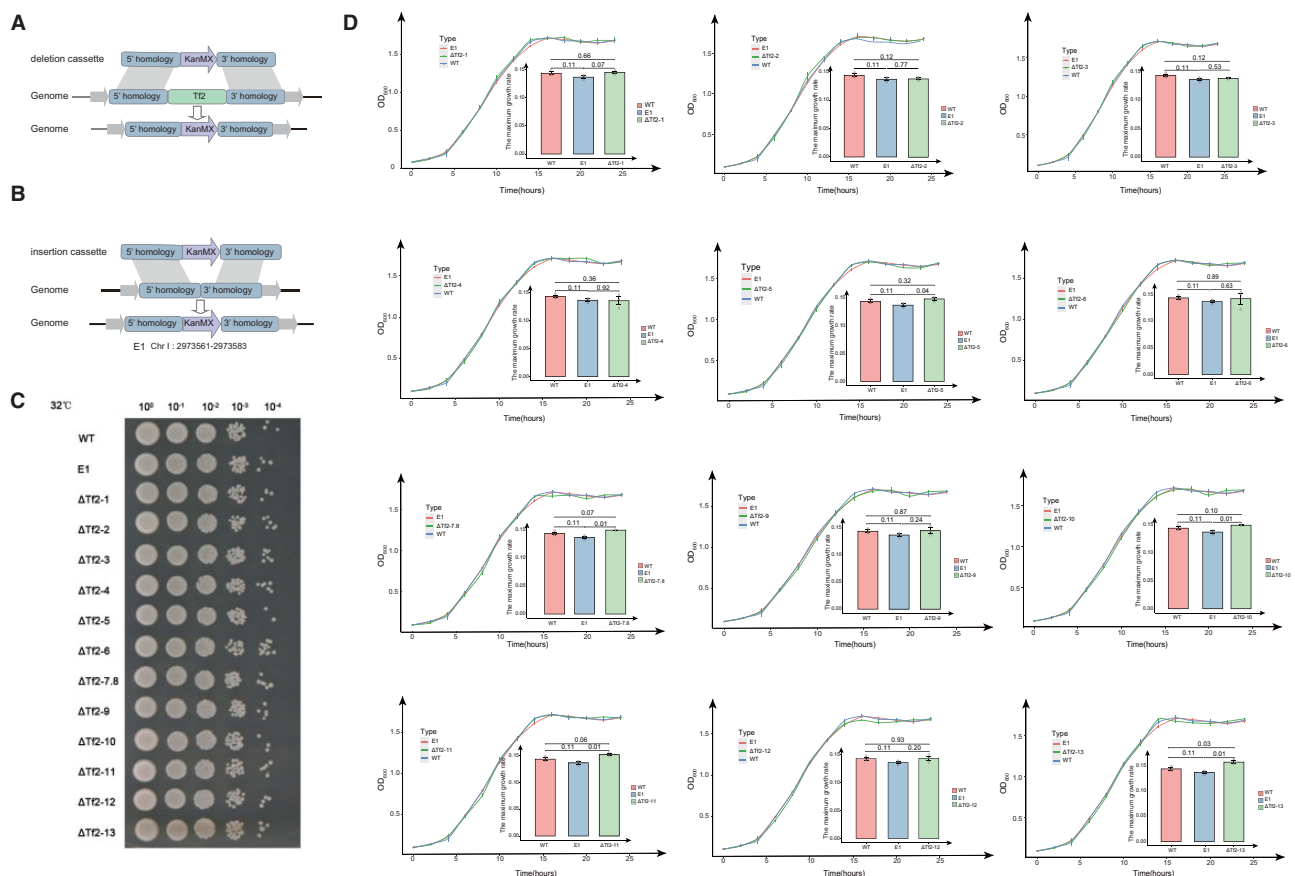
Fig. 2.—Tf2 knockout and the effects of Tf2 on growth of *S. pombe*. (*A* and *B*) The schematic diagram of a homologous recombination gene "deletion cassette" (*A*) and 'insertion cassete' (*B*). Both the 5'homology and the 3'homology are 500 bp sequences flanking Tf2. The position of E1 is insertion in a gene poor region. (*C*) The spot assay of Tf knockout, E1, and WT strains cultured at 32°C. (*D*) The growth curves and maximum growth rates of WT, E1, and Tf2 knockout isolates. The growth curves of WT and Tf2 knockout strains were obtained by kinetic experiments, and each of the 13 Tf2 knockout strains was compared with the WT strain. Three biological replicates were performed. The average OD_{600} is shown as solid line with the error bar representing standard deviation (SD). The filled circles with different colors are the data of the replicates. The average of maximum growth rates for all the three replicates is shown with the error bar representing SD.

and knockout strains (fig. 2*C*). Furthermore, there was no significant difference in growth curves was observed among WT, E1, and ΔTf2 strains (fig. 2*D*). All the strains reached the logarithmic growth phase at ∼6 h, and reached the stationary phase at ∼14 h (fig. 2*D*). Moreover, we estimated the maximum growth rates for WT, E1, and ΔTf2 strains, and found that the maximum growth rates for all the experimental strains were ∼0.14 (fig. 2*D*). No significant difference in the maximum growth rates was observed between WT and knockout strains (*P* > 0.01), although weak difference was observed for some comparisons between E1 and knockout strains possibly due to stochastic errors. Taken together, these results imply that segregating complete Tf2 elements might have limited effect on the growth of *S. pombe*.

Tf2 Knockout Has No Effect on *S. pombe* Fitness

To determine the effect of TEs on yeast fitness, we calculated the relative fitness of each knockout strain and to exclude the

effect of marker gene, the E1 was also calculated. We found that knocking out Tf2 elements had limited effect on yeast fitness. The relative fitness of ΔTf2–1 (1.0710 ± 0.2463), ΔTf2–2 (1.0181 ± 0.0518), ΔTf2–3 (1.0031 ± 0.1681), ΔTf2–4 (1.0031 ± 0.1681), ΔTf2–5 (0.9831 ± 0.0462), ΔTf2–6 (1.0140 ± 0.0794), ΔTf2–7 (0.9666 ± 0.1234), ΔTf2–9 (0.9964 ± 0.2128), ΔTf2–10(1.0033 ± 0.0556), ΔTf2–11 (0.9345 ± 0.1434), ΔTf2–12(1.0017 ± 0.1247), and ΔTf2–13(1.0029 ± 0.0063) are all very close to 1 (table 1, fig. 2*D*). Therefore, Tf2s might have only very small, if not any, effect on the fitness of yeast.

Next, we used competition experiment to measure the relative fitness of yeast strains. We expressed a green fluorescent protein (GFP) under the control of the ADH1 promoter in WT and ΔTf2 strains, and GFP was consistently expressed throughout the competition experiment (fig. 3*A*). We mixed WT with GFP (WT-GFP) with ΔTf2 in roughly equal ratio (note: not exactly 50% due to stochastic errors), making WT-GFP and ΔTf2 coexist and compete in a 7-day simulated

Table 1

The Relative Fitness of ΔTf2s Relative to WT Estimated from Kinetics Experiments

Strain	Fitness ± SE Relative to WT
E1	1.0146 ± 0.1119
ΔTf2–1	1.0710 ± 0.2463
ΔTf2–2	1.0181 ± 0.0518
ΔTf2–3	1.0031 ± 0.1681
ΔTf2–4	1.0134 ± 0.2424
ΔTf2–5	0.9831 ± 0.0462
ΔTf2–6	1.0140 ± 0.0794
ΔTf2–7	0.9666 ± 0.1234
ΔTf2–9	0.9964 ± 0.2128
ΔTf2–10	1.0033 ± 0.0556
ΔTf2–11	0.9345 ± 0.1434
ΔTf2–12	1.0017 ± 0.1247
ΔTf2–13	1.0029 ± 0.0063

competition environment (fig. 3B). Similar to the competition between WT-GFP and WT strains (fig. 3C), the proportion of WT-GFP population fluctuated around 50% without a directional trend for the competition between WT-GFP and ΔTf2 strains (fig. 3D). No obvious decrease or increase in fitness was observed for either ΔTf2 strains (fig. 3D). Moreover, we also expressed a GFP in each ΔTf2 strain (ΔTf2-GFP), and

performed the competition experiment between WT and ΔTf2-GFP strains. Once again, no directional trend was observed for any ΔTf2-GFP stain (fig. 3E). These results suggest that the segregating complete Tf2 elements are largely neutral for the fitness of *S. pombe*.

Tf2 Knockout Has No Effect on the Expression of Nearby Genes

TE integration might affect the expression of nearby host genes in *S. pombe* (Bowen et al. 2003). To explore whether Tf2 element insertions affect the expression of nearby genes, we mapped their positions in the genome (fig. 4) and used real-time quantitative fluorescent PCR to measure the transcriptional level of nearby genes for WT and ΔTf2 strains (fig. 4). Except for Tf2–11, the distance between the Tf2 elements and their adjacent genes was within less than 1,000 bp (fig. 4). For Tf2–3 (fig. 4C), Tf2–4 (fig. 4D), Tf2–6 (fig. 4F), and Tf2–12 (fig. 4K), they are less than 200 bp away from their nearby genes, but their knockouts did not affect the expression of nearby genes ($P = 0.57$ for Tf2–3, $P = 0.45$ for Tf2–4, $P = 0.89$ for Tf2–6, and $P = 0.07$ for Tf2–12). For Tf2–3, it partially overlaps with the 5′-UTR of *amt3*, but no significant difference in nearby gene expression was observed
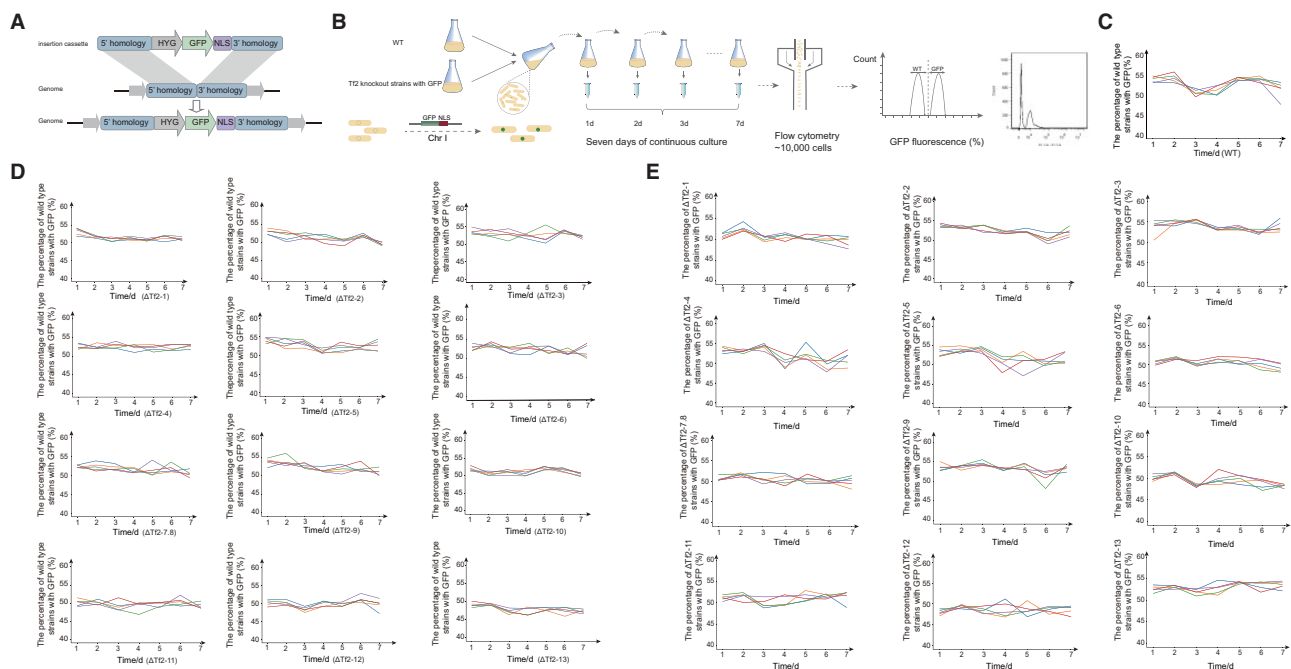
FIG. 3.—Competition experiment between WT and Tf2 knockout strains. (A) The schematic diagram of a "insertion cassette." Both the 5′ homology and the 3′ homology are 500 bp sequences. HYG: Hygromycin B. (B) The model of competition experiment. In the competition experiment, the stains with and without GFP were equally mixed and continuously cultured for 7 days. We sampled daily and detected the percentage of strains with GFP throughout the fermentation process by flow cytometry. The dotted circle represents the nucleus, and the green circle indicates the nucleus with GFP expression. (C) The percentage of the WT-GFP strain cells during competition between WT and WT-GFP strains. (D) The percentage of the WT-GFP strain cells during competition between WT and Tf2 knockout strains. (E) The percentage of Tf2 knockout strain cells during competition between WT and Tf2 knockout strains. The lines in different colors represent five replicate populations. The starting percentage is about but not exactly 50% due to stochastic errors.
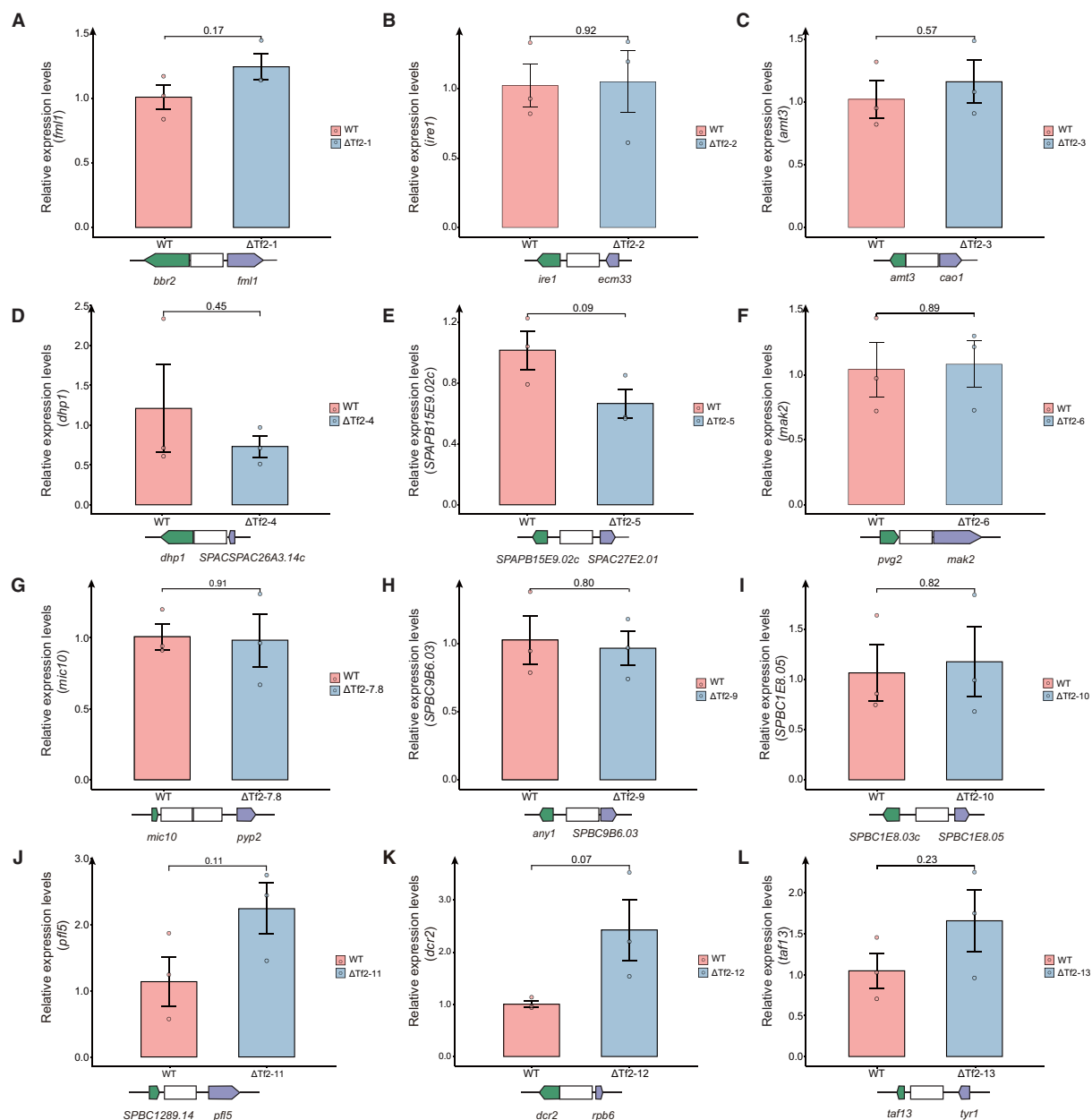
Fig. 4.—The effects of Tf2 elements on the expression of nearby genes. (A–L) Compression of the expression levels of nearby genes between WT and Tf2 knockout strains. The nearby genes of 13 Tf2 elements are shown with white boxes representing Tf2 elements, and green and purple boxes represent the nearby genes. All experiments were performed for three biological replicates which are represented by colored filled circles, and the error bar represents SD. Statistically, significance was determined by Student's t-test ($^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$).

between WT and ΔTf2–3 strains ($P = 0.57$) (fig. 4C). For Tf2–1 (fig. 4A), Tf2–2 (fig. 4B), Tf2–10 (fig. 4I), Tf2–13 (fig. 4L), they are ~1,000 bp away from their nearby genes, and no significant difference in nearby gene expression was observed between WT and ΔTf2 ($P = 0.17$ for Tf2–1, $P = 0.92$ for Tf2–2, $P = 0.82$ for Tf2–10, and $P = 0.23$ for Tf2–13). In summary, our results suggest that the Tf2 elements do not affect the expression of nearby genes.

Tf2 Knockout Has No Effect on *S. pombe* Stress Response

To explore the effect of Tf2 elements on stress response of *S. pombe*, we used the spot assay to observe the growth of yeast challenged by four different stresses (hyperthermal, nutritional, oxidative, and osmotic stresses). The growth of WT and ΔTf2 strains is similar under the normal culture condition of 32 °C (fig. 2C). When challenged by hyperthermal stress (cultured in 37 °C), the growth was slower for both WT and
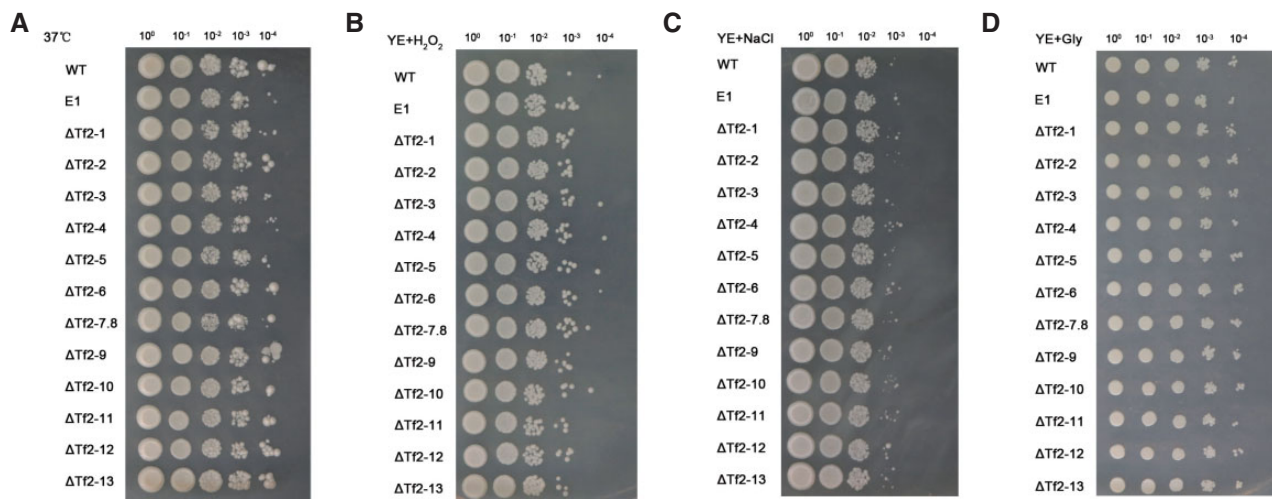
FIG. 5.—Spot assay of Tf2 knockout strains. WT and Tf2 knockout strains were cultured in liquid medium until logarithmic growth phase, and the stock suspensions obtained by 10-fold serial dilution were spotted onto the rich medium (*A*), rich media containing 5 mM H_2O_2 (YE + H_2O_2) (*B*), rich media containing 0.15 M NaCl (YE + NaCl) (*C*), and rich media containing 3% glycerol (YE + Gly) (*D*). Plate A was cultured at 37 °C for 5 days. Plates B–D were cultured at 32 °C for 3–7 days.

ΔTf2 strains, but no significant difference in growth was observed between WT and ΔTf2 strains (fig. 5*A*). Under oxidative stress (cultured by adding hydrogen peroxide), no obvious difference in growth was observed between WT and ΔTf2 strains (fig. 5*B*). Under osmotic stress (cultured by adding 0.15 M NaCl) and nutritional stress (cultured by adding 3% glycerol), WT and ΔTf2 strains were similar in growth. Taken together, our results suggest that the complete Tf2 elements might have no obvious effect on stress responses of *S. pombe*.

Discussion

In this study, we empirically tested the effect of segregating TEs on the fitness of the host using *S. pombe* as the model system. Two families of LTR retrotransposons, namely Tf1 and Tf2 elements, were widely present in *S. pombe* strains sampled globally. Consistent with previous studies (Bowen et al. 2003), we identified 13 Tf2 elements, a family of LTR-retrotransposon, within the genome of *S. pombe* strain 972. These Tf2 elements are segregating within the *S. pombe* population, and appear to be resulted from recent integration. Recombination occurred among and between Tf1 and Tf2 elements, potentially facilitating chromosomal rearrangement or structural variation in *S. pombe* (Fawcett et al. 2014; Jeffares et al. 2017; Tusso et al. 2019).

To empirically test the effect of TEs on the host fitness, we knocked out each of Tf2 elements within *S. pombe* laboratory strain 972. We found that Tf2 knockout has no obvious effect on the growth of *S. pombe*. For each of the knockout strains, the relative fitness to WT is close to 1. Although the relative fitness for ΔTf2–1 and ΔTf2–11 is ~1.07 (±0.25) and ~0.93 (±0.14), the standard error appears to be pretty large and no

directional trend was observed for either strain in competition experiment. Through competition experiment, we found Tf2 knockout did not affect the fitness of *S. pombe*. Moreover, no significant effect of Tf2 elements on the expression of nearby genes was observed, although the possibility of Tf2 elements acting as long-distance enhancers cannot formally excluded. Taken together, our results suggest segregating complete Tf2 elements are largely neutral to the fitness of *S. pombe*.

TE insertions have been thought to impact the expression of host genes, providing a potential means to cope with environmental stress (McClintock 1984; Negi et al. 2016; Horváth et al. 2017; Lanciano and Mirouze 2018; Esnault et al. 2019). In this study, we challenged WT and Tf2 knockout strains with hyperthermal, oxidative, osmotic, and nutritional stresses. Under these stresses, no difference in growth was observed between WT and Tf2 knockout strains, suggesting that Tf2 elements might play limited role in response to these stresses. Given that environmental stresses might be far more complex in nature, it is possible that these Tf2 elements help adapt to other stresses. However, these Tf2 insertions appear to have limited effect on the expression of nearby genes, making the possibility less likely.

Our study suggests that segregating complete Tf2 elements are largely neutral in *S. pombe*. But previous studies found, when subjecting TE-tagged *S. pombe* cells to stresses, novel TE integration promotes the expression of stress response genes, proving a path to cope with stresses (Feng et al. 2013; Esnault et al. 2019). This paradox can be readily explained by the neutral theory of molecular evolution (Arkhipova 2018). Most new TE insertions are deleterious, which will be rapidly removed from the host population (Arkhipova 2018). A majority of the remaining new TE

insertions are selectively neutral or slightly deleterious (Arkhipova 2018), and their fates are governed by genetic drift, accounting for most of TE insertions polymorphism within the host population. Consistently, our study suggests segregating Tf2 elements are largely neutral to the host fitness. Only a minor fraction of novel insertions might be beneficial, which might be fixed by positive selection (Arkhipova 2018), as exemplified by those novel TE insertions upregulating stress response genes (Feng et al. 2013; Esnault et al. 2019). However, when stresses are alleviated or changed, these insertions might not be beneficial anymore. Our study provides empirical support for the selfish nature of Tf2 elements in fission yeasts: they amplified very recently within the genome, and making no obvious contribution to the host phenotype (Doolittle and Sapienza 1980; Orgel and Crick 1980). The possibility that these complete TE elements and their derivatives gain new biological functions during the evolutionary course cannot be formally excluded.

What are the ultimate fates of these segregating complete Tf2 elements? Based on our results, these segregating complete Tf2 elements will mainly evolve under genetic drift, accumulating disruptive mutations (including internal region removal by recombination between LTRs) and degrading sometime in the future. Or the segregating complete Tf2 elements might provide raw materials for co-option by hosts, which might be rare but not impossible. Nevertheless, our study provides empirical support for a neutral evolution model and selfish nature of TEs and might have general implications in understanding the relationship between the TEs and their hosts.

Materials and Methods

Population Genomics Analysis

We used similarity search with diverse TE proteins to identify TEs within in 19 *S. pombe* strains sampled globally (supplementary table S1, Supplementary Material online). In previous studies (Levin et al. 1990), we identified two families of LTR retrotransposons, namely Tf1 and Tf2 elements. Then, we used the BLASTn algorithm to identify Tf1 and Tf2 elements within 19 *S. pombe* strains with Tf1 and Tf2 elements as queries and an *e* cutoff value of 10^{-5}. The putative proteins and domains were annotated using Conserved Domain Search (Marchler-Bauer and Bryant 2004). LTRs were annotated using LTRfinder (Xu and Wang 2007). Given Tf1 and Tf2 elements exhibit high sequence divergence in Gag, we distinguished Tf1 and Tf2 elements based on phylogenetic analysis of Gag proteins.

To reconstruct the phylogenetic relationship among 19 *S. pombe* strains sampled globally, we randomly selected 30 genes (supplementary table S4, Supplementary Material online) and concatenated these genes by Phylosuite version 1.2.1 (Zhang et al. 2020). Sequences were aligned using MAFFT (Katoh and Standley 2013). Phylogenetic analysis was performed using a maximum likelihood algorithm implemented in IQ-TREE (Nguyen et al. 2015). We generated split networks of complete Tf1 and Tf2 elements through neighborhood network analysis implemented Splittree4 (Huson and Bryant 2006). RDP5 (Martin et al. 2020) was used to examine recombination in Tf1 and Tf2 elements. Sequence identity among Tf2–11, Tf2 (accession No. L10324), and Tf1 (accession No. M38526) was examined using Simplot (Bowen et al. 2003) with a window size of 100 nt and a step size of 10 nt. Pairwise identity was calculated by BioAider version 1.334 (Zhou et al. 2020).

Strains, Media, and Culture Conditions

The strains used in this study were from American type culture collection. The *S. pombe* strains were grown using standard methods and media. The rich media used were Yeast Extract (YE) media. All cells were cultured at 32 °C and cell growth was monitored by OD_{600}.

Tf2 Knockout

All the strains used in this study were derived from *S. pombe* laboratory strain 972. We knocked out TEs with "deletion cassette" through homologous recombination with a KanMX flanked by two homologous fragments (Moreno et al. 1991; García-Ríos et al. 2014). Strains were generated by transformation with a lithium acetate-based method (Moreno et al. 1991; García-Ríos et al. 2014). We used the same method to express a GFP with the nuclear localization signal (NLS) controlled by ADH1 promoter in each ΔTf2 strain.

Spot Assay

The yeast strains were cultured in YE liquid medium at 32 °C with shaking at 200 rpm overnight. The strains incubated overnight were transferred to a new YE liquid, and grow to OD_{600} value of ~3. The cells were collected by centrifugation, and OD_{600} was adjusted to 3. Then, the strains were continuously diluted by a gradient of 10 times to 10^{-5}, generating five stock suspensions. The five stock suspensions were used to perform the spot assay. We plated 2.5 μl of each stock suspension on the surface of YE, YE + 0.15M NaCl, YE + 3% glycerol, YE + 5mM H_2O_2 solid medium. The plates were cultured at 32 °C before observing the growth of colonies. At the same time, we cultured the YE plates at 37 °C to observe the growth of colonies.

Growth Curve Analysis and Fitness Measurement

All cells were cultured at 32 °C with a starting OD600 value of 0.2. We measured the number of cells as well as the maximum growth rates of WT and knockout strains after 24 h of incubation following inoculation of the seed solution.

We took samples at 0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24 h, and measured OD_{600} in 96-well plates to characterize their growth curves. Three replicates were performed for each sample. We used the growth rate of the logarithmic growth period in the growth curve as the maximum growth rate of each knockout strain, which is the ratio of the difference between the end of logarithmic growth and the beginning of logarithmic growth to the time difference. The relative fitness of each TE knockout strain was calculated by calculating the doubling time of each knockout strain from the growth curve and comparing with the doubling time of WT strain (Gui et al. 2021).

Competition Experiment

We tested the fitness of WT and knockout strains by competition experiments (García-Ríos et al. 2014). The initial seed solution for the competition experiment was a mixed culture consisting of two strains (WT-GFP and ΔTf2 strains, WT-GFP and WT strains, or WT and ΔTf2-GFP strains) with equal ratio (note: not exactly 50% due to stochastic errors). Before mixing the solution, two strains were cultured in YE medium overnight with OD_{600} value of 3 to ensure that they have the same growth status. We passaged for seven consecutive days with five replicates for each mixed sample. We sampled daily and detected the percentage of strains with GFP throughout the fermentation process by flow cytometry (BD Accuri C6). A total of 10,000 cells from the samples were measured with the FL1-fluorescein isothiocyanate sensor, and the number of fluorescent cells was measured.

RNA Extraction and RT-qPCR

We tested whether the Tf2 element insertions affect the expression of nearby genes. The nearby genes are genes with complete Tf2 elements (except for Tf2–7.8) located upstream of their 5' UTRs. Tf2–7.8 is physically far (2,139 bp upstream) from *pyp2* gene, and we measured the expression change of *mic10* gene. Yeast samples were grown overnight in 10 ml YE medium. The cultures were diluted to OD_{600} value of 0.2, and then were grown OD_{600} value of 0.8. RNA was extracted by Yeast RNA Kit (Omega). RNA was reverse transcribed into cDNA using HiScript III RT SuperMix for qPCR (+gDNA wiper) (Vazyme). Quantitative Real-Time RT-PCR (qRT-PCR) was carried out using intercalating dye ChamQTM SYBR qPCR Master Mix (High ROX Premixed) (Vazyme) with each primer set (supplementary table S5, Supplementary Material online). The qPCR reactions were performed in 10 μl volume. We extracted RNA three times. Data analysis was performed by StepOne. The C_T values were normalized against actin mRNA levels from the same preparation to give the ΔC_T values, and the relative changes in gene expression were estimated using the $2^{-\Delta\Delta Ct}$ method.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Data Availability

The data underlying this article are available in the article and in its online supplementary material.

Literature Cited

Arkhipova IR. 2018. Neutral theory, transposable elements, and eukaryotic genome evolution. Mol Biol Evol. 35(6):1332–1337.

Blanc VM, Adams J. 2003. Evolution in *Saccharomyces cerevisiae*: identification of mutations increasing fitness in laboratory populations. Genetics 165(3):975–983.

Bowen NJ, Jordan IK, Epstein JA, Wood V, Levin HL. 2003. Retrotransposons and their recognition of pol II promoters: a comprehensive survey of the transposable elements from the complete genome sequence of *Schizosaccharomyces pombe*. Genome Res. 13(9):1984–1997.

Britten RJ, Davidson EH. 1969. Gene regulation for higher cells: a theory. Science 165(3891):349–357.

Brosius J, Gould SJ. 1992. On "genomenclature": a comprehensive (and respectful) taxonomy for pseudogenes and other "junk DNA". Proc Natl Acad Sci USA. 89(22):10706–10710.

Charlesworth B. 1987. The population biology of transposable elements. Trends Ecol Evol. 2(1):21–23.

Chuong EB, Elde NC, Feschotte C. 2017. Regulatory activities of transposable elements: from conflicts to benefits. Nat Rev Genet. 18(2):71–86.

Chuong EB, Elde NC, Feschotte C. 2016. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. Science 351(6277):1083–1087.

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

Doolittle WF, Sapienza C. 1980. Selfish genes, the phenotype paradigm and genome evolution. Nature 284(5757):601–603.

Feng G, Leem YE, Levin HL. 2013. Transposon integration enhances expression of stress response genes. Nucleic Acids Res. 41(2):775–789.

Esnault C, Levin HL. 2015. The long terminal repeat retrotransposons Tf1 and Tf2 of *Schizosaccharomyces pombe*. Microbiol Spectr. 3(4):10.1128/microbiolspec.

Esnault C, Lee M, Ham C, Levin HL. 2019. Transposable element insertions in fission yeast drive adaptation to environmental stress. Genome Res. 29(1):85–95.

Daborn PJ, et al. 2002. A single p450 allele associated with insecticide resistance in Drosophila. Science 297(5590):2253–2256.

Fawcett JA, et al. 2014. Population genomics of the fission yeast *Schizosaccharomyces pombe*. PLoS One 9(8):e104241.

Finnegan DJ. 1989. Eukaryotic transposable elements and genome evolution. Trends Genet. 5(4):103–107.

Gardner MJ, et al. 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature 419(6906):498–511.

García-Ríos E, Gutiérrez A, Salvadó Z, Arroyo-López FN, Guillamon JM. 2014. The fitness advantage of commercial wine yeasts in relation to the nitrogen concentration, temperature, and ethanol content under microvinification conditions. Appl Environ Microbiol. 80(2):704–713.

Gui Q, et al. 2021. Transcriptome analysis in yeast reveals the externality of position effects. Mol Biol Evol. 38(8):3294–3307.

Hoff EF, Levin HL, Boeke JD. 1998. *Schizosaccharomyces pombe* retrotransposon Tf2 mobilizes primarily through homologous cDNA recombination. Mol Cell Biol. 18(11):6839–6852.

Horváth V, Merenciano M, González J. 2017. Revisiting the relationship between transposable elements and the eukaryotic stress response. Trends Genet. 33(11):832–841.

Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. Mol Biol Evol. 23(2):254–267.

Jeffares DC, et al. 2017. Transient structural variations have strong effects on quantitative traits and reproductive isolation in fission yeast. Nat Commun. 8:14061.

Kapitonov VV, Jurka J. 2008. A universal classification of eukaryotic transposable elements implemented in Repbase. Nat Rev Genet. 9(5):411–412.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780.

Kazazian HH, Jr, et al. 1988. Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. Nature 332(6160):164–166.

Lanciano S, Mirouze M. 2018. Transposable elements: all mobile, all different, some stress responsive, some adaptive? Curr Opin Genet Dev. 49:106–114.

Levin HL, Weaver DC, Boeke JD. 1990. Two related families of retrotransposons from *Schizosaccharomyces pombe*. Mol Cell Biol. 10(12):6791–6798.

Lynch M. 2007. The origins of genome architecture. Sunderland (MA): Sinauer.

Lynch VJ. 2016. A copy-and-paste gene regulatory network. Science 351(6277):1029–1030.

Marchler-Bauer A, Bryant SH. 2004. CD-Search: protein domain annotations on the fly. Nucleic Acids Res. 32:W327–W331.

Martin DP, et al. 2020. RDP5: a computer program for analyzing recombination in, and removing signals of recombination from, nucleotide sequence datasets. Virus Evol. 27(1):veaa087.

McClintock B. 1956. Controlling elements and the gene. Cold Spring Harb Symp Quant Biol. 21:197–216.

McClintock B. 1984. The significance of responses of the genome to challenge. Science 226(4676):792–801.

Moreno S, Klar A, Nurse P. 1991. Molecular genetic analysis of fission yeast *Schizosaccharomyces pombe*. Methods Enzymol. 194:795–823.

Negi P, Rai AN, Suprasanna P. 2016. Moving through the stressed genome: emerging regulatory roles for transposons in plant stress response. Front Plant Sci. 7:1448.

Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 32(1):268–274.

Nishihara H. 2019. Retrotransposons spread potential cis-regulatory elements during mammary gland evolution. Nucleic Acids Res. 47(22):11551–11562.

Orgel LE, Crick FH. 1980. Selfish DNA: the ultimate parasite. Nature 284(5757):604–607.

Payer LM, Burns KH. 2019. Transposable elements in human genetic disease. Nat Rev Genet. 20(12):760–772.

Rhind N, et al. 2011. Comparative functional genomics of the fission yeasts. Science 332(6032):930–936.

Sundaram V, Wang T. 2018. Transposable element mediated innovation in gene regulatory landscapes of cells: re-visiting the gene-battery model. Bioessays 40(1):1700155.

Tusso S, et al. 2019. Ancestral admixture is the main determinant of global biodiversity in fission yeast. Mol Biol Evol. 36(9):1975–1989.

Ullastres A, Merenciano M, González J. 2021. Regulatory regions in natural transposable element insertions drive interindividual differences in response to immune challenges in Drosophila. Genome Biol. 22(1):265.

Van't Hof AE, et al. 2016. The industrial melanism mutation in British peppered moths is a transposable element. Nature 534(7605):102–105.

Wang J, Han GZ. 2020. Frequent retroviral gene co-option during the evolution of vertebrates. Mol Biol Evol. 37(11):3232–3242.

Wang J, Han GZ. 2021. Unearthing LTR retrotransposon gag genes co-opted in the deep evolution of eukaryotes. Mol Biol Evol. 38(8):3267–3278.

Wells JN, Feschotte C. 2020. A field guide to eukaryotic transposable elements. Annu Rev Genet. 54:539–561.

Wicker T, et al. 2007. A unified classification system for eukaryotic transposable elements. Nat Rev Genet. 8(12):973–982.

Wood V, et al. 2002. The genome sequence of *Schizosaccharomyces pombe*. Nature 415(6874):871–880.

Xu Z, Wang H. 2007. LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. Nucleic Acids Res. 35:W265–W268.

Zhang D, et al. 2020. PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Mol Ecol Resour. 20(1):348–355.

Zhou ZJ, Qiu Y, Pu Y, Huang X, Ge XY. 2020. BioAider: an efficient tool for viral genome analysis and its application in tracing SARS-CoV-2 transmission. Sustain Cities Soc. 63:102466.

Associate editor: Josefa Gonzalez