# One Cell, Two Gears: Extensive Somatic Genome Plasticity Accompanies High Germline Genome Stability in *Paramecium*

Francesco Catania [1,2,*], Rebecca Rothering [1], and Valerio Vitali [1]

[1]Institute for Evolution and Biodiversity, University of Münster, Germany
[2]Institute of Environmental Radioactivity, Fukushima University, Japan

*Corresponding author: E-mail: francesco.catania@uni-muenster.de.

**Accepted:** 22 November 2021

## Abstract

Mutation accumulation (MA) experiments are conventionally employed to study spontaneous germline mutations. However, MA experiments can also shed light on somatic genome plasticity in a habitual and genetic drift-maximizing environment. Here, we revisit an MA experiment that uncovered extraordinary germline genome stability in *Paramecium tetraurelia*, a single-celled eukaryote with nuclear dimorphism. Our re-examination of isogenic *P. tetraurelia* MA lines propagated in nutrient-rich medium for >40 sexual cycles reveals that their polyploid somatic genome accrued hundreds of intervening DNA segments (IESs), which are normally eliminated during germline-soma differentiation. These IESs frequently occupy a fraction of the somatic DNA copies of a given locus, producing IES excision/retention polymorphisms, and preferentially fall into a class of epigenetically controlled sequences. Relative to control lines, retained IESs are flanked by stronger *cis*-acting signals and interrupt an excess of highly expressed coding exons. These findings suggest that *P. tetraurelia*'s elevated germline DNA replication fidelity is associated with pervasive somatic genome plasticity. They show that MA regimes are powerful tools for investigating the role that developmental plasticity, somatic mutations, and epimutations have in ecology and evolution.

**Key words:** mutation accumulation, germline-soma differentiation, programmed DNA elimination, developmental plasticity, epigenetics.

## Significance

Mutation accumulation (MA) experiments are conventionally used to study de novo germline mutations. Here, however, we use an MA experiment to investigate developmental variation. We find that genetically identical *Paramecium* lines, which are independently cultured for tens of sexual generations under conditions that minimize the power of selection, accrue hundreds of developmental variants in their somatic nucleus. Two models, one based on random (epigenetic) drift and one of evolutionary adaptation that does not require positive Darwinian selection, explain our observations equally well. This study shows that functional specialization of nuclei in *Paramecium* is accompanied by vastly different levels of genome stability.

## Introduction

Mutations are critical for life. They fuel evolutionary change (Loewe and Hill 2010), affect disease susceptibility (Pritchard 2001; Veltman and Brunner 2012; Poduri et al. 2013), and contribute to ecosystem functioning (Reusch et al. 2005; Whitham et al. 2006). Yet, several important questions about mutation rates and the significance of mutations still warrant further attention. Four of these questions are considered below.

First, how stable are somatic (nuclear) genomes? Although much is known about the variation in germline mutation rates across species (Lynch et al. 2016 and references therein),

© The Author(s) 2021. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

somatic mutation rates are considerably less investigated. Previous observations indicate that somatic genomes are less stable than germline genomes in a few animal species (Lynch 2010; Milholland et al. 2017; García-Nieto et al. 2019; Yizhak et al. 2019), an instability that may be cell division independent (Krašovec et al. 2014, 2017; Abascal et al. 2021). Whether this disparity holds for most animals and, more generally, for eukaryotes with nuclear dimorphism such as microbial ciliates remains unknown.

Second, what are the implications (if any) of somatic genome instability? Although somatic mutations are often regarded as a hallmark of aging (Martincorena and Campbell 2015) and are expected to be largely lost upon sexual reproduction, a rapid somatic mutation rate could be an enduring source of variation in organisms where transgenerational somatic inheritance is at play.

Third, how do external changes affect the emergence of mutations? In multicellular species such as *Drosophila*, increased environmental temperature activates transposon mobilization in germ cells (Cappucci et al. 2019). In bacteria, yeast, and cancer cells, adverse conditions raise the frequency of mutations, via a mechanism known as stress-induced mutagenesis (Bjedov et al. 2003; Foster 2007; Heidenreich 2007; Shor et al. 2013; Eisen 2014; Fitzgerald et al. 2017; Russo et al. 2019). Furthermore, mutation rate increases as the degree of cell–cell interaction decreases (Krašovec et al. 2014). Together, these observations align with the hypothesis that stress-induced mutagenesis may be a conserved program that operates in unicellular organisms and in cells that have "lost tissue-imposed homeostasis" within a multicellular system (Russo et al. 2019). The validity of this hypothesis requires further investigation.

Fourth, and last, what environmental conditions affect genome stability? Mutations rates may increase in response to a number of adverse, presumably stressful conditions, such as high temperatures (Belfield et al. 2021). However, little is known about whether exposure to new environmental conditions that an experimenter may perceive as benign (e.g., nutrient enrichment) can also induce mutations.

The free-living single-celled *Paramecium tetraurelia* is an ideal system to gain insights into these questions. In nature, *P. tetraurelia* often lives in ephemeral aquatic environments (e.g., temporary natural ponds). The occurrence of genetically similar con-specific strains in geographically distant and distinct locations (Catania et al. 2009; Johri et al. 2017) suggests that this ciliate may move frequently between water bodies, possibly via migrating birds that transport ciliates in water drops (Sonneborn 1975; Foissner et al. 2008). Given its low levels of intra-specific genetic diversity (and its elevated germline genome stability, see below), it is unclear how *P. tetraurelia* manages to colonize different environments with presumably varying physical properties. Differently put, where does the variation through which *P. tetraurelia* copes with new environments come from? A close look into this ciliate's biology might help uncover an answer to this question.

*P. tetraurelia* contains a germline and a somatic nucleus within a single cell. At each sexual event, a new somatic genome regenerates from the zygotic genome through a process that can yield somatic genome variation (Duret et al. 2008; Catania et al. 2013). More specifically, during the process of germ-soma differentiation, *P. tetraurelia*'s newly developing somatic genome experiences a number of changes that fall under the name of programmed DNA elimination (PDE) (Chalker et al. 2013; Betermier and Duharcourt 2014). These changes include genome amplification via endoreplication (from $2n$ to $\sim 860n$, Woodard et al. 1961), chromosome fragmentation, transposon elimination, and de novo telomere addition. Moreover, ~45,000 intervening DNA segments, which are known as internal eliminated sequences (IESs) and occupy inter- and intragenic positions (~80% in exons, a nearly random distribution of IESs with respect to genes), are eliminated at a genome-wide level (Arnaiz et al. 2012; Vitali et al. 2019). IES elimination is not foolproof, however. At hundreds of loci, IESs can be retained in one or more copies of the polyploid somatic genome (Duret et al. 2008; Catania et al. 2013). Besides, when *P. tetraurelia* is exposed to environmental (temperature) changes during vegetative life and/or development, a considerable number of IESs that are excised at standard culture conditions are retained in the somatic nucleus, often to a non-trivial extent (e.g., >10% of the somatic DNA copies) (Vitali et al. 2019; Hagen et al. 2020). Retained IESs may be considered developmental variants in that they are present in the germline genome and their incorporation into the somatic genome does not require de novo mutations. However, they may also be regarded as de novo somatic mutations because they are novel somatic insertions. Either way, somatic IESs may alter gene expression levels, and give rise to phenotypic effects that are detrimental and thus selected against (Arnaiz et al. 2012; Ferro et al. 2015; Vitali et al. 2019). Importantly, retained IESs may also be passed down to sexual offspring via conserved RNA-mediated epigenetic mechanisms (Duharcourt et al. 1995, 1998, 2009), the same mechanisms that contribute to the trans-generational transmission of parental phenotypic responses in animals, fungi, and plants (Duempelmann et al. 2020) and that regulate cellular genes in the absence of germline mutations enabling the inheritance of mating types in *Paramecium* (Singh et al. 2014). More specifically, retained IESs can be inherited beyond the vegetative stage through a form of non-Mendelian inheritance termed homology-dependent inheritance, which allows cells to inherit retained IESs across sexual generations despite the replacement of the somatic nucleus (Meyer and Garnier 2002). In sum, current observations suggest that *per*-locus alternative DNA splicing ($IES^+$ and $IES^-$) variants in *P. tetraurelia* can arise during somatic development, especially upon exposure to a new environment. Furthermore, some of those variants that are under

epigenetic control may persist and accumulate across sexual generations. By greatly increasing variability, it is possible that these alternative DNA splicing variants may help *P. tetraurelia* cope with new environmental conditions.

To begin to assess the plausibility of this hypothesis, we revisited a published mutation accumulation (MA) study where multiple *P. tetraurelia* lines were propagated from an isogenic state for ~3,300 asexual generations at room temperature and in a nutrient-rich and antibiotic-free culture medium (Sung et al. 2012) (supplementary fig. S1, Supplementary Material online). Importantly, these culture conditions were new to the MA lines at the beginning of the experiment. For one, the MA lines originated from a progenitor *P. tetraurelia* stock (d4-2) that was normally kept at <18 °C (rather than at room temperature), constantly at dark (rather than partly at light), and fed moderately once every few weeks (rather than copiously, daily). For another, the progenitor *P. tetraurelia* d4-2 stock was kept at low/moderate cell density—the isogenic MA lines originated from one of these cells—whereas the MA lines were systematically propagated via daily single-cell transfers to maximize the efficacy of random genetic drift. In the MA study, cell density was only allowed to increase once during the vegetative stage to trigger starvation-induced self-fertilization. The ~3,300 asexual generations in the MA experiment were intercalated with >40 episodes of self-fertilization, each of which is expected to give rise to sexual offspring that are completely homozygous and genetically identical to its parent (Sonneborn 1957). Based on these experimental conditions, which the MA lines experienced for ~4 years, the authors reported no more than 29 putative germline mutations across seven MA lines (Sung et al. 2012).

Here, we examined the somatic genomes of these MA lines (with a particular focus on those MA lines with the highest levels of genome-wide sequencing coverage). We find that these somatic genomes possess considerably more retained IESs compared with con-specific stocks and *P. tetraurelia* lines that were cultured at various environmental temperatures. A relative excess of these retained IESs is under epigenetic control and occur in coding sequences. Further, in each of the independently evolved MA lines, retained IESs preferentially disrupt genes that are highly expressed in the parental stock. These observations suggest that the MA experimental regime promotes substantial, potentially heritable, and nonrandom developmental/somatic genome plasticity in *Paramecium* lines with otherwise virtually uniform germline DNA sequences.

## Results

### An Excess of IESs Accrues in the Somatic Nucleus of *P. tetraurelia* MA Lines

In a typical deep-sequencing study, IES retention in *P. tetraurelia* is limited to a few hundred somatic loci and mainly involves ≤5% of the ~860 copies *per* locus when the ciliate is cultured under standard conditions (supplementary table S1, Supplementary Material online). Because exposure to a new cultivation environment can significantly perturb PDE in *P. tetraurelia*, generating potentially heritable DNA variation (Vitali et al. 2019; Hagen et al. 2020), we asked whether the long-term exposure of *P. tetraurelia* cells to the MA experimental regime is coupled with an enhanced accrual of IESs in the somatic genome. After filtering out somatic loci with a trivial fraction of IES-retaining mapping reads (arbitrarily set to ≤5% or with IES retention Score [IRS] ≤ 0.05), we found that the surveyed MA lines (the five lines with adequate median IES coverage, supplementary table S1, Supplementary Material online) exhibit between 2 and 4 times more IES-containing loci than the three control *P. tetraurelia* stocks including d4-2, the stock used for the MA experiment (supplementary table S1, Supplementary Material online and fig. 1*A*). The increase in IES-containing loci is relative to the overall number of analyzable IES loci (i.e., loci covered by >20 reads) in each sample and thus accounts for differences in genome coverage. Four additional and independent somatic genomes of *P. tetraurelia* lines cultured at standard temperatures (25 °C/27 °C, AR, EV, ND7 [stock 51] and 25 °C F1 [stock d12]; supplementary table S1, Supplementary Material online), display an amount of retained IESs (with IRS > 0.05) that is comparable to that of the three control stocks (fig. 1*A*). In contrast, increased counts of retained IESs were observed for F1 lines that originate from the same parental somatic genome, that is, Ctrl d12, but were exposed to 18 °C or 32 °C during macronuclear development, or 40 °C (daily, for 30 s) during vegetative life and cultured at 25 °C during development (Vitali et al. 2019; Hagen et al. 2020) (these lines are henceforth referred to as "environmental lines"; fig. 1*A* and supplementary fig. S1 and table S1, Supplementary Material online). The relative excess of retained IESs in the MA lines and environmental lines raises the possibility that the MA lines' parental culture acquired many retained IESs soon upon exposure to the MA experimental conditions. Unfortunately, this hypothesis cannot be tested because the parental culture is no longer available. Alternatively, or additionally, the independently evolved MA lines may have accrued retained IESs during the long-term (~4-year) experiment.

### IES Retention Levels across MA Lines and Con-Specific Stocks

A matching number of samples in MA, control, and environmental lines makes comparisons between these lines easier, therefore for the rest of our analyses we decided to focus on the three MA lines with the highest median coverage (MA25, MA70, MA30) (supplementary table S1, Supplementary Material online; Sung et al. 2012). Similarly, the set of control lines was reduced to one control *per P. tetraurelia* strain (Ctrl d4-2, Ctrl 51, and Ctrl d12).

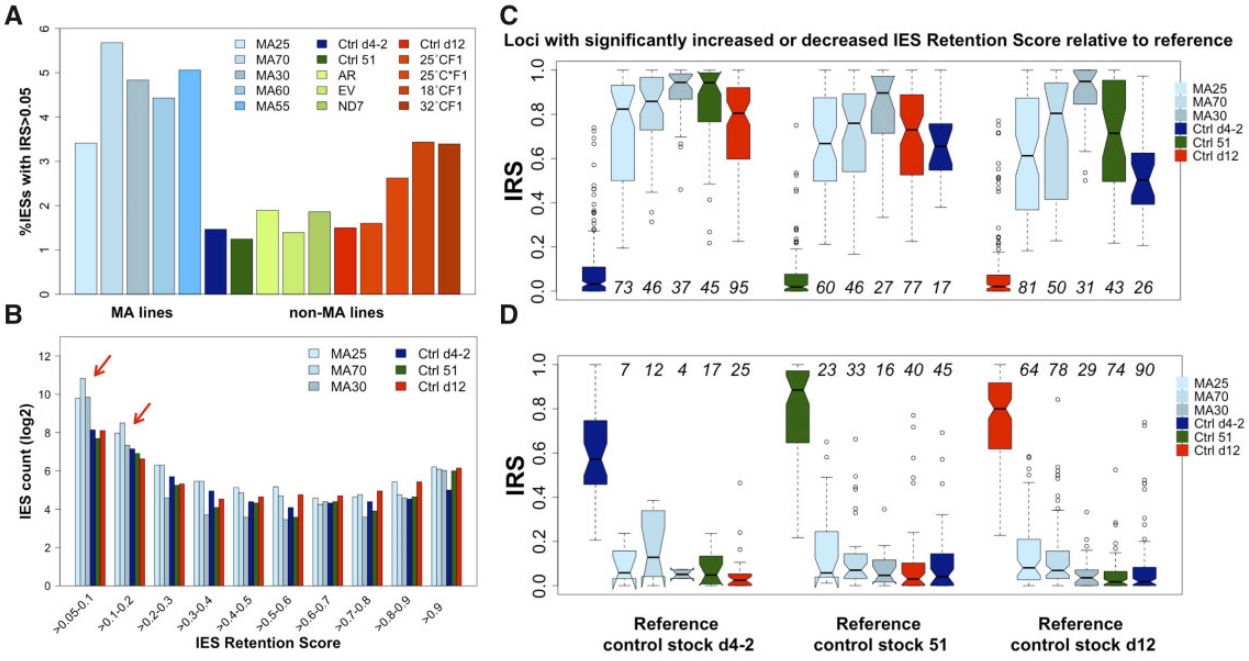


**Fig. 1.**—IES retention profile in MA, control (Ctrl), and environmental (Env) lines. (*A*) MA lines contain a higher percentage of retained IESs with IES retention score (IRS) > 0.05 compared with control *P. tetraurelia* stocks d4-2, 51, and d12 cultured at standard conditions. Ctrl d4-2 is the same stock used as progenitor of the MA lines. Stock 51 includes four independent experiments: Ctrl 51, AR, EV, and ND7 (more details in Materials and Methods). Stock d12 includes four co-experimental lines derived from isogenic parental cells (Ctrl d12) and exposed to 1) 25 °C during vegetative life and self-fertilization (25 °CF1); 2) 25 °C intercalated with 40 °C per 30 s daily during vegetative life and 25 °C during self-fertilization (25 °C * F1); 3) 25 °C during vegetative life and 18 °C during self-fertilization (18 °CF1); and 4) 25 °C during vegetative life and 32 °C during self-fertilization (32 °CF1) (more details in Materials and Methods). (*B*) MA lines display a relatively higher count (log2) of retained IESs with $0.05 < \text{IRS} \leq 0.2$. (*C*) At tens of loci (counts in italics), the IRSs in the MA lines are significantly elevated or (*D*) significantly reduced compared with reference control stocks. Statistical significance was estimated through a binomial test after taking the upper and lower bound of the 95% confidence interval around the retention score of the reference stock. For all the analyses, only retained IESs with >20× sequence coverage were considered.

When we consider the MA lines, we note that the enrichment of retained IESs is largely confined to IESs with $0.05 < \text{IRS} \leq 0.2$ ($P < 2.2\text{e-}16$, proportion test; fig. 1*B* and supplementary table S1, Supplementary Material online). Furthermore, we detected differences in the magnitude of IES retention at tens of loci. More specifically, IES retention can either increase (fig. 1*C*) or decrease (fig. 1*D*) significantly in the MA lines compared with the control stocks as well as among the control stocks (fig. 1*C* and *D*). These findings suggest that the somatic nucleus of *P. tetraurelia* undergoes recurrent and pronounced seesaw IES retention dynamics, extending previous across-species observations (Catania et al. 2013).

## Recurrence of Developmental IES-Related Changes in MA Lines and Control Stocks

We next observed that numerous retained IESs co-occur in the MA lines (fig. 2). The degree of this overlap is greater for the MA lines than for both control stocks and environmental lines when we focus on IESs with IRS > 0.2. More explicitly, an excess of three-way shared IESs with IRS > 0.2 is detected for the MA lines (25.3%) compared with control stocks (18.5%) and environmental lines (18.8%) (Bonferroni corrected $P \leq 0.01$, proportion test). Instead, the MA lines exhibit a less marked intersection of retained IESs relative to the control stocks when we consider the range $0.05 < \text{IRS} \leq 0.2$. About 4.6% and 2.7% of IESs with $0.05 < \text{IRS} \leq 0.2$ are shared among the three control stocks and the three environmental lines, respectively, versus 2.5% in the three MA lines (MA lines vs control lines: $P = 0.001$; MA lines vs environmental lines: $P = 0.55$; Proportion test). Thus, if retained IESs occurred in the MA lines' parental cells (i.e., before the MA lines were allowed to evolve independently), then only a subset of IESs with $0.05 < \text{IRS} \leq 0.2$ were inherited/recurred across >40 sexual cycles. Alternatively, or in addition, retained IESs with $0.05 < \text{IRS} \leq 0.2$ may have accrued in parallel in the MA lines' somatic nucleus during the MA experiment. Either way, our observations suggest that *P. tetraurelia* contains a subset of IESs that are particularly prone to incomplete excision.

## An Excess of Retained IESs in *Paramecium* Falls into a Class of Epigenetically Controlled IESs

Although the excision of all *P. tetraurelia* IESs relies on the domesticated piggyBac transposase PiggyMac (Baudry et al. 2009; Bischerour et al. 2018), a subset of IESs also rely on

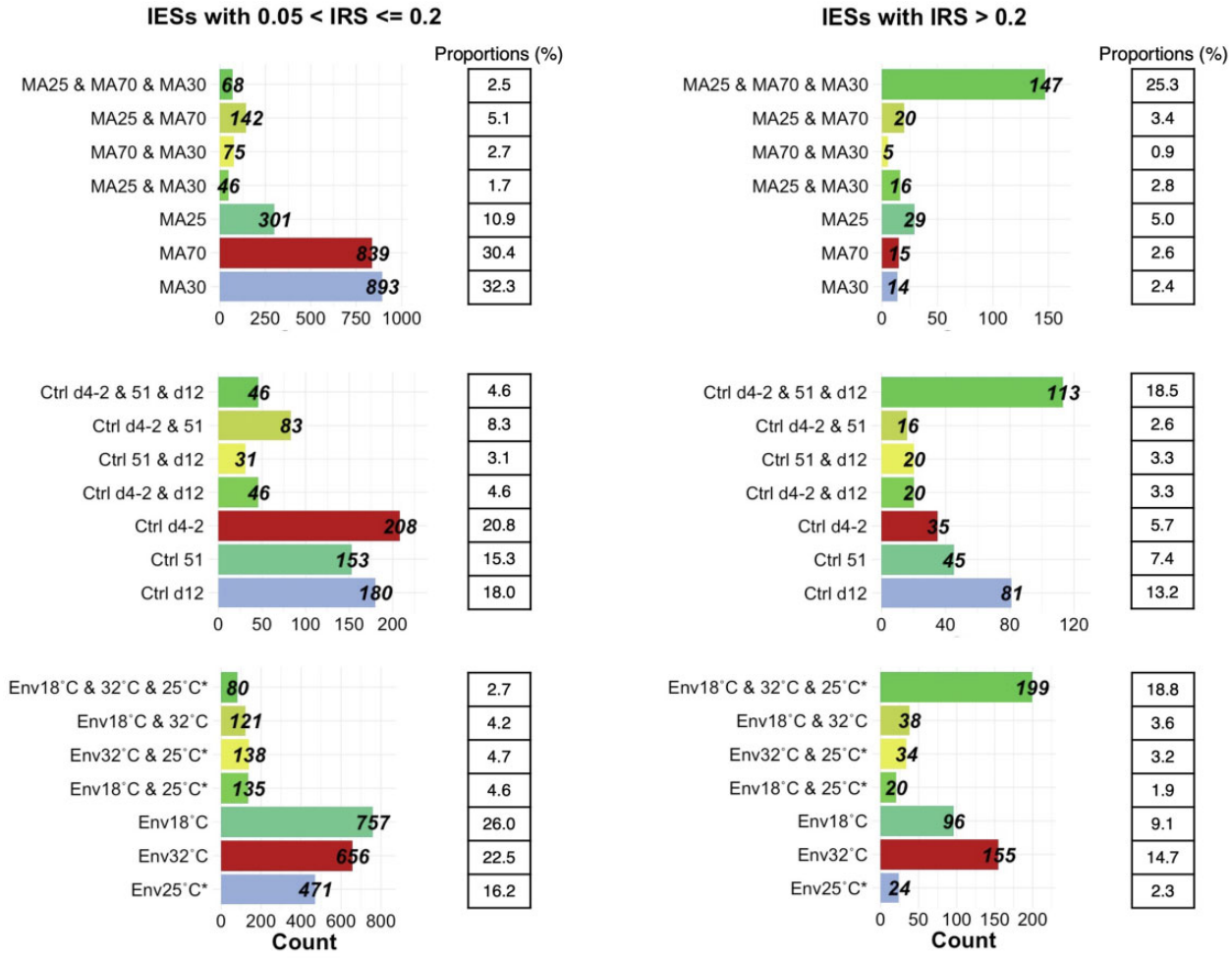


**Fig. 2.**—Unique and shared retained IESs in MA, Ctrl, and Env lines. Retained IESs are grouped according to their IRS, that is, $0.05 < \text{IRS} \leq 0.2$ and $\text{IRS} > 0.2$. Counts and corresponding proportions (%) relative to the total number of IESs *per* IRS interval and group are shown. Only IESs supported by >20 reads simultaneously in all lines of each group were considered.

epigenetic mechanisms. More specifically, the excision of a few thousands IESs (~6% of the total) further requires the intervention of piRNA-like small RNAs, known as scnRNAs, that are generated in the germline nucleus during meiosis via the Dicer-like proteins Dcl2 and Dcl3 (Lepere et al. 2008, 2009; Duharcourt et al. 2009; Lhuillier-Akakpo et al. 2014; Sandoval et al. 2014; Maliszewska-Olejniczak et al. 2015). After silencing of *Dcl2* and *Dcl3*, epigenetically controlled IESs (epi-IESs hereinafter) show a level of retention that is significantly higher than the nonsilenced control. Furthermore, IESs that are retained in the somatic nucleus and whose excision is epigenetically controlled may be passed down to the next sexual generation(s) (Duharcourt et al. 1995, 1998).

We found that a significant excess of retained IESs in the cultures under study falls into a class of IESs that are under the control of scnRNAs in stock 51 (Lhuillier-Akakpo et al. 2014). The excess of these epi-IESs can be detected regardless of the level of IES retention. However, whereas the MA lines and the control lines exhibit comparable fractions of epi-IESs when we consider the set of IESs with $\text{IRS} > 0.2$ ($22.5 \pm 1.0\%$ [$n = 193$] vs $19.5 \pm 5.3\%$ [$n = 143$], $P = 0.27$), the MA lines exhibit a relative deficit of epi-IESs in the set of IESs with $0.05 < \text{IRS} \leq 0.2$ ($12.4 \pm 2.6\%$ [$n = 548$] vs $18.2 \pm 4.4\%$ [$n = 202$], $P = 2.1\text{e-}06$; proportion test after summing counts across lines *per* set). The environmental lines, on the other hand, exhibit fractions of epi-IESs that are statistically indistinguishable from those estimated for the MA lines ($14 \pm 1.6\%$ [$n = 420$] and $21 \pm 1.1\%$ [$n = 237$] in the set of IESs with $0.05 < \text{IRS} \leq 0.2$ ($P = 0.07$) and $\text{IRS} > 0.2$ ($P = 0.48$), respectively, proportion test).

Finally, we examined the putative epigenetic regulation of the MA lines' IESs with significantly higher or lower IRS values compared with the counterpart in the parental stock (fig. 1C and *D*). We found that the control stock 51 and the MA lines exhibit comparable fractions of epi-IESs in the sets of IESs with

significantly reduced IRS values (10/17 [58.8%] vs overall 12/23 [56.0 ± 16.9%], respectively; $P = 0.733$, proportion test). In contrast, the fraction of epi-IESs in the MA lines' set of IESs with significantly increased IRS values (overall 52/156 [32.8 ± 2.6%]) is three times higher than that recorded for the control stock 51 (5/45 [~11%]) ($P = 0.031$, proportion test).

Collectively, these findings suggest that a subset of the retained IESs in the MA lines may be under the control of scnRNAs. This epigenetic control may be more pronounced for IESs with particularly elevated retention scores.

## Retained IESs in MA Lines Are Associated with Strong *Cis*-Acting Regulatory Signals and Unique Domains-Encoding Genes

We previously described a quality measure for *cis*-acting IES recognition/excision signals called $C_{in}$-score. Shaped by natural selection, this score corresponds to the degree of complementary base pairing between IESs' subterminal sites (Ferro et al. 2015). We found that IESs with $0.05 < IRS \leq 0.2$ exhibit stronger *cis*-acting signals in the MA lines compared with the counterpart in control stocks (respective $C_{in}$-scores: 0.598 [$n = 3{,}819$] vs 0.556 [$n = 864$], $P = 0.00019$, Wilcoxon rank sum test). In contrast, the *cis*-acting signals of IESs with IRS > 0.2 are comparably weak in the MA lines and the control stocks (respective $C_{in}$-scores: 0.518 [$n = 431$] vs 0.520 [$n = 410$], $P = 0.897$, Wilcoxon rank sum test).

Next, we examined the relationship between retained IESs and repetitive sequences, such as the extremely abundant Tetratricopeptide repeat domain, within *P. tetraurelia* genes. After filtering out retained IESs in genes without annotated protein domains (see Materials and Methods), we found that a similar average proportion of retained IESs with IRS > 0.2 map to genes encoding repeat domains in the MA lines and the control stocks (39 ± 9.6% vs 42 ± 7.2%, respectively; $P = 0.707$, proportion test). In contrast, IESs with $0.05 < IRS \leq 0.2$ in the MA lines map considerably less often to repeat domains-encoding genes compared with the control lines (25 ± 7.6% vs 37 ± 4.5%, respectively; $P = 1.835e\text{-}10$, proportion test) (supplementary table S2, Supplementary Material online). When we focus on retained IESs that fall within protein domains, a comparable proportion of IESs with IRS > 0.2 interrupt repeat protein domains in the MA lines and the control stocks (49% and 48%, respectively, $P = 1$, proportion test). Instead, the IESs with $0.05 < IRS \leq 0.2$ in the MA lines interrupt repeat protein domains considerably less (23%) than the control stocks (38%) ($P = 1.913e\text{-}05$, proportion test; supplementary table S2, Supplementary Material online).

Overall, these findings suggest that the IESs that are preferentially retained in the MA lines (i.e., IESs with $0.05 < IRS \leq 0.2$) are subject to substantial selective pressure in the environment to which control stocks are normally exposed.

## Retained IESs in the MA Lines Map Preferentially within Coding Exons and May Impact Gene Expression Levels

We investigated the genomic distribution of the retained IESs. We found that the IESs with $0.05 < IRS \leq 0.2$ in the MA lines reside most frequently within coding exons (MA25, $n = 628$ [58%]; MA70, $n = 1399$ [67%]; MA30, $n = 732$ [71%]), an unexpected finding. Because exon-mapping IESs have, on average, stronger *cis*-acting recognition/excision signals compared with intergenic IESs (Ferro et al. 2015), intergenic IESs should be at least as or more likely to undergo incomplete excision. Indeed, retained IESs with comparable IRSs in the control stocks are distributed evenly across intergenic and exonic regions, a significantly different configuration compared with that recorded for the MA lines ($P < 2.2e\text{-}16$; proportion test after summing across lines *per* set *per* genomic position; fig. 3*A*). As for retained IESs with IRS > 0.2, they reside preferentially in intergenic regions to a similar extent in the MA lines and in the control stocks ($P > 0.05$; proportion test after summing across lines *per* set *per* genomic position; fig. 3*B*). Notably, many retained IESs with $0.05 < IRS \leq 0.2$ in the MA lines introduce premature termination codons (PTCs) in the nascent transcript: 77% of the ORFs acquire PTCs upon these IESs' retention. These transcripts may be targeted for degradation by the cellular surveillance systems.

Next, we found that genes with reduced expression levels preferentially accumulate retained IESs, an unsurprising observation (Arnaiz et al. 2012; Ferro et al. 2015; Vitali et al. 2019). More specifically, an excess of exon-mapping retained IESs in the MA lines or in the control stocks maps to genes that are weakly expressed during the vegetative life of the control stocks 51 and/or d12 (fig. 3*C* and *D*). Surprising, however, is that retained IESs with $0.05 < IRS \leq 0.2$ (but not with IRS > 0.2) in the MA lines map to highly expressed genes significantly more often than in the control stocks ($P < 0.0001$, proportion test after summing across lines *per* set *per* gene expression category). Thus, loci that are highly expressed during the vegetative life of *Paramecium* may preferentially experience reduced allele dosage in the MA lines.

## IESs with Divergent Levels of Retention between the MA Lines and the Parental Stock

Finally, we considered those IESs that show divergent retention levels between the MA lines and the parental stock d4-2. We found two IESs that are under-represented in the somatic genome of all the MA lines relative to the parental stock (supplementary fig. S2, Supplementary Material online). The only intragenic IES maps to a *P. tetraurelia* gene with currently unknown function (*PTET.51.1.G1120131*). This single-copy gene is predicted to contain a trans-membrane helix that spans the differentially retained 45 nt-long exon-mapping IES. The pronounced IES retention in the parental stock is predicted to perturb the topology of the encoded trans-membrane protein (supplementary fig. S3, Supplementary

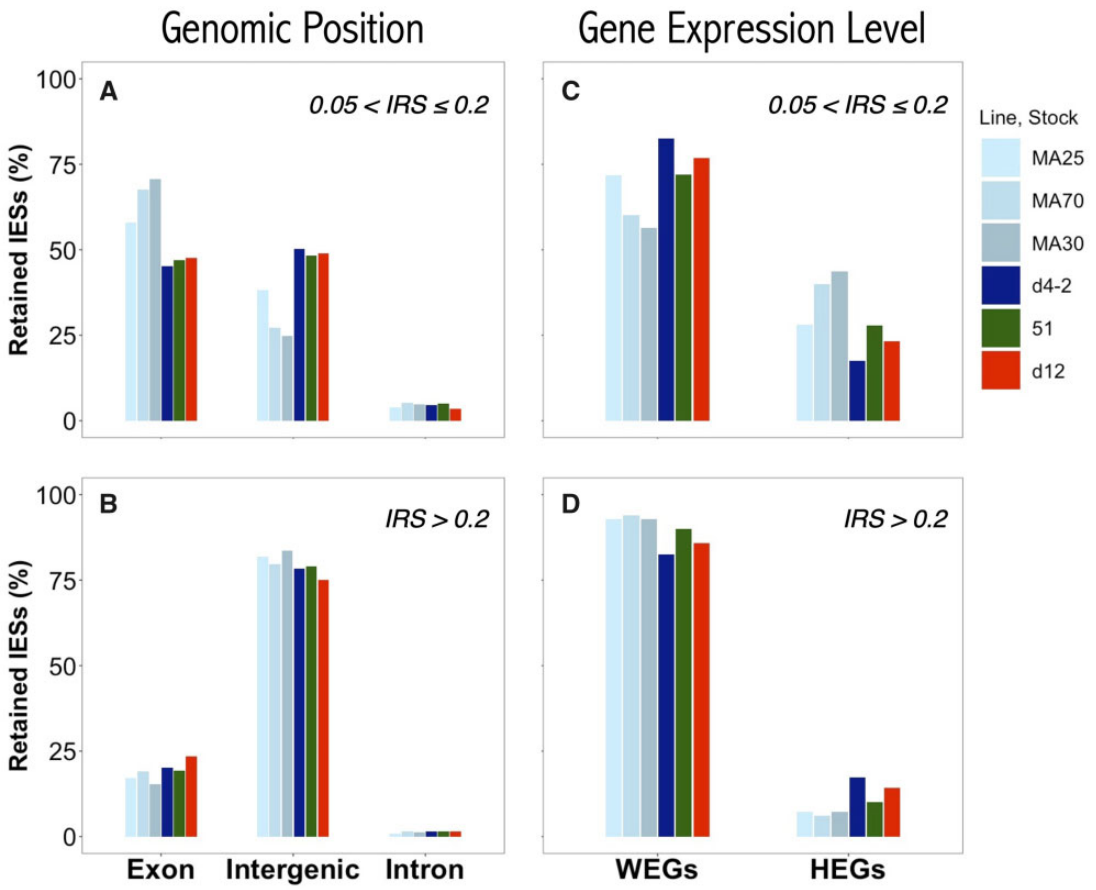


**Fig. 3.**—Positional profile of retained IESs in MA lines and Ctrl stocks. (*A*) Genomic distribution of retained IESs with IRS ranging between >0.05 and ≤0.2 and (*B*) >0.2 in MA lines and Ctrl stocks. (*C*) In the MA lines, highly expressed genes (HEGs) accumulate a higher fraction of incompletely excised IESs (IRS > 0.05 and ≤0.2) compared with the control stocks d4-2, 51, and d12. (*C* and *D*) Overall, an excess of retained exon-mapping IESs (regardless of their IRS) maps to weakly expressed genes (WEGs) at standard culture conditions (i.e., these genes' log2-transformed vegetative expression values fall in the first quartile of an underlying distribution of gene expression levels) in stock 51 and stock d12. The transcriptome of the control stock 51 (Arnaiz et al. 2017) was used as reference for the MA lines, stock d4-2 and stock 51, whereas the transcriptome of the control stock d12 (Vitali et al. 2019) was used as reference for stock d12. Only IESs supported by >20 reads were considered.

Material online). It follows that *PTET.51.1.G1120131* in the MA lines might generate relatively larger amounts of a transmembrane protein whose function remains to be characterized. The other (intergenic) IES flanks a gene (*PTET.51.1.G1590064*) also with unknown function, which has a predicted trans-membrane domain (supplementary table S3, Supplementary Material online).

Twenty-one IESs are fully or largely absent from the somatic DNA of the parental stock d4-2 but have high levels of somatic retention in all the MA lines (supplementary fig. S4, Supplementary Material online). Thirteen IESs are intergenic and most often flank gene 3′ ends with unknown implications (supplementary table S4, Supplementary Material online). Three maps to noncoding genes with unknown function. The remaining five IESs interrupt five genes, three with a putative description (supplementary fig. S4, Supplementary Material online): 1) *PTET.51.1.G0280259* is single-copy and encodes a WD repeat-containing protein that is involved in ribosome biogenesis in yeast; 2) *PTET.51.1.G0310293* encodes a guanine nucleotide exchange factor that is involved in signaling pathways such as cell proliferation; and last, 3) *PTET.51.1.G0480149* is a *CYClin* PHO80-like gene that is highly conserved from unicellular to multicellular species. *PTET.51.1.G0480149* is interrupted in the MA lines but neither in the parental stock nor in the other control stocks (see IESPGM.PTET51.1.48.253167 in supplementary fig. S4, Supplementary Material online). The cyclin PHO80-like domain (IPR013922) breaks down upon the retention of the 76 nt-long IES (Arnaiz and Sperling 2011). In other eukaryotes, the gene *PHO80* is a key effector of the so-called PHO pathway—by which phosphate availability is sensed—and a positive regulator of the insulin-signaling pathway. This gene has three presumably functional ohnologs in *P. tetraurelia* (stock 51). Thus, the disruption of *PTET.51.1.G0480149* in the MA lines might contribute to reducing—rather than fully abolishing—*CYClin*' s activity,

possibly blunting nutrient signaling sensors and slowing down cell growth.

## Discussion

In an experimental setting where the power of random genetic drift is maximized (Katju and Bergthorsson 2019) and new germline DNA variation is negligible (Sung et al. 2012), we found that replicate MA lines of *P. tetraurelia* accrue a multitude of developmental/somatic variants over the course of ~4-year exposure to the same culture conditions (fig. 1*A* and supplementary table S1, Supplementary Material online). An excess of these molecular variants corresponds to somatically retained IESs with $0.05 < \text{IRS} \leq 0.2$ (fig. 1*B*), a retention interval that implies that IESs preferentially interrupt a modest fraction of somatic DNA copies at a given locus. A similar pattern of increased IES retention was previously recorded for F1 lines (here referred to as environmental lines) that originate from *P. tetraurelia* cells exposed to presumably stressful thermal changes during vegetative life or self-fertilization (Vitali et al. 2019; Hagen et al. 2020). Thus, as in the case of the environmental lines, the MA lines may have experienced elevated levels of IES retention in response to environmental stress. But what is the origin of this stress?

As the MA lines' progenitor stock (*P. tetraurelia* d4-2) was kept at conditions that differ substantially from the MA conditions (see Introduction), it is possible that the impact of the MA lines' parental cells with the new MA experimental conditions generated sufficient stress to promote incomplete IES excision. That noted, *Paramecium* stock cells are often, if not routinely, exposed to standard, MA-like environmental conditions across laboratories (albeit not for as many sexual generations as in the case of the MA lines), and never have to our knowledge particularly increased levels of IES retention been detected/reported (e.g., see control stocks in fig. 1*B*). Therefore, it is also possible that stressful circumstances, which affect the process of IES excision, may have occurred during the MA study.

And when and how might these retained IESs have accrued in the MA lines? The *Paramecium* cells that Sung et al. (2012) used to start the MA study are no longer available, as the MA lines are no longer living, so we can only speculate about the dynamics that unfolded during the MA experiment. That noted, retained IESs might have occurred in the MA lines' progenitor cells, in the independently evolving MA lines, or both. To begin, it is unlikely that most of the IES retention occurred from the time of the most recent new macronucleus formation: there is no evidence of increased IES retention under experimental conditions that overlap with those of the MA lines (Vitali et al. 2019; Hagen et al. 2020). Furthermore, the significantly reduced intersection of retained IESs with $0.05 < \text{IRS} \leq 0.2$ in the MA lines compared with the control stocks (fig. 2) may be interpreted as indicating that these IESs were first retained in the MA lines' progenitor cells and later partially "lost" (e.g., reverted to complete excision) in the independently evolving MA lines. However, the narrow overlap among retained IESs with $0.05 < \text{IRS} \leq 0.2$ in the MA lines may also result from parallel IES retentions occurred during the MA study. Either way, retained IESs may have been trans-generationally inherited due to heritable small RNA-mediated epimutations, consistent with the excess of scnRNA-controlled IESs detected among the MA lines' retained IESs. The alternative explanation that retained IESs in the MA lines are transmitted subsequent to genetic mutations is less plausible. Given the low germline mutation rate in *Paramecium* (Sung et al. 2012; Long et al. 2018), it is unlikely that *bona fide* germline mutations in *cis* (i.e., in the IES itself) or in *trans* (i.e., in genes encoding factors affecting IES excision) can explain the hundreds of retained IESs reported here.

Regardless of when exactly IESs populated the MA lines' somatic genome, additional observations suggest that retained IESs in the MA lines accumulated nonrandomly with regard to location and gene expression levels. More specifically, IESs with $0.05 < \text{IRS} \leq 0.2$ in each of the MA lines (but not in the control stocks) preferentially reside in coding exons and genes that are highly expressed at standard culture conditions (fig. 3). If we postulate that IESs with $0.05 < \text{IRS} \leq 0.2$ were already present in the MA lines' progenitor and that this initially shared set of retained IESs narrowed down in the evolving MA lines (fig. 2), then the peculiar positional distribution of remaining retained IESs in the MA lines suggest that these IESs were maintained nonrandomly. Alternatively, if we postulate that retained IESs in the MA lines accrued in parallel over the course of the experiment, then the peculiar distribution of these retained IESs suggests once again nonrandom dynamics. This begs the question: what dynamics gave rise to this nonrandom pattern?

It is difficult to explain our observations by leveraging biases in chromosome segregation. Amitosis might have favored a stochastic increase (and/or decrease) of IES retention levels during asexual divisions, although somatic assortment in *P. tetraurelia* is expected to proceed at a very slow pace (Vitali et al. 2021). Instead, two models provide plausible explanations for our observations. In the erroneous excision model, the excess of IESs with $0.05 < \text{IRS} \leq 0.2$ observed in MA lines corresponds to retained IESs that would be counterselected in normal cultivation conditions—because their retention, even at a moderate level, impacts fitness—but persist in the MA experiment where selection is much less efficient. These "fitness-impacting" IESs are expected to accumulate more epimutations (and hence to have a higher IRS) in MA lines compared with control stocks.

In the regulated retention model, IESs with $0.05 < \text{IRS} \leq 0.2$ in the MA lines serve a regulatory purpose. They may help downregulate (and/or stabilize the downregulation of) pathways that underlie disfavored phenotypes in the MA environment. They may also stimulate/stabilize the expression of

favorable alternative pathways/phenotypes. For example, under the MA experimental regime—where the supply of nutrients is rich and constant and population density is minimized—IES-coupled developmental changes might both slow cell growth and increase stress resistance. This speculative regulated retention model links previous findings in *Paramecium*, such as the trade-off between cell growth and stress resistance (Thind et al. 2020) and the increase in IES retention upon exposure to stressful cultivation regimes (Vitali et al. 2019; Hagen et al. 2020). It predicts that during evolutionary times some IESs were co-opted to help maintain cellular homeostasis, and is consistent with a previously proposed non-Darwinian model of adaptive evolution (Hughes 2012; Catania et al. 2021). The erroneous excision and the regulated retention models are not mutually exclusive: some retained IESs might correspond to errors, whereas others might perform a regulatory function. Moreover, the accumulation of incorrectly excised IESs could itself modulate biological functions.

In conclusion, we leveraged a classical MA experiment to investigate the source, the amount, and the possible consequences of somatic variation in the single-celled *Paramecium*. Our observations align with previous findings in species such as humans and mice where somatic genomes are less stable than germline genomes. They also indicate that a sufficiently lengthy exposure of genetically narrow/homogeneous cell populations to a new environment can induce potentially lasting somatic variation, which may result from small RNA-mediated epimutations. At present, two models—one non-adaptive and one adaptive—can be leveraged to explain our results. Further work is needed to evaluate them.

## Materials and Methods

### *Paramecium* Stocks

Details concerning the culture conditions, the macronuclear DNA isolation and the whole-genome sequencing of the MA *Paramecium tetraurelia* lines were previously reported (Sung et al. 2012). MA lines' DNA sequences are deposited in the National Center for Biotechnology Information Short Read Archive (Sung et al. 2012). Three *P. tetraurelia* MA lines (MA25 [run SRR652989], MA70, and MA30 [run SRR652988]) were selected for this study based on the high analyzable percentage of the total genome (Sung et al. 2012) and their relatively higher median coverage of IESs (supplementary table S1, Supplementary Material online). The unevolved *P. tetraurelia* stocks d4-2, 51, and d12 were used as control. Stock d4-2 (the same stock as the MA lines) and stock d12 are derivatives of stock 51 (Sonneborn 1974; Rudman et al. 1991).

Reads used to quantify IES retention in the control stocks were obtained from Arnaiz et al. (2012): (stock d4-2; ERR138952 [Ctrl d4-2]) (Arnaiz et al. 2012; Lhuillier-Akakpo et al. 2014; Swart et al. 2017), (stock 51, cultured at 27 °C; ERR138450 [AR], ERR501376 [Ctrl 51], ERR1212640 [ND7], ERR1212635 [EV]), and (Vitali et al. 2019) (stock d12, F0 line cultured at 25 °C; ERR2807204 [Ctrl d12]).

Last, data relative to the "environmental lines" was published in Vitali et al. (2019): (stock d12, F1 lines cultured at 25 °C, but whose macronuclear genome developed at 18 °C; ERR2807205 or 32 °C; ERR2807207) and (Hagen et al. 2020) (stock d12, F1 line cultured and developed at 25 °C with F0 exposed to daily heat shocks of 30 s at 40 °C during vegetative life; ERR4179861). Somatic genomes for these lines were obtained from cells that had undergone more than 10 (and in any case less than 30) fissions postautogamy to avoid carryover of parental somatic DNA.

### IES Data Sets

For each of the above-mentioned samples, bowtie2 (with the option –very-sensitive-local) was used to map reads against the reference somatic genome of strain 51 (MAC) (Arnaiz et al. 2017), as well as a pseudo-germline version of it, in which the full reference set of 44,928 known IESs (Arnaiz et al. 2012) was inserted (MAC + IES). IES retention scores (IRSs), that is the *per*-locus ratio between IES-containing reads and the total number of mapping reads, were calculated with ParTIES (Denby Wilkes et al. 2016) using the IES score method and allowing for a single mismatch. IES loci with significantly different retention levels in an MA line compared with a control stock or between control stocks were designated as described in Vitali et al. (2019) using a 95% confidence interval around the IRS value of the control stock. IES loci supported by $\leq$20 sequence reads were excluded. Only IESs with a size larger than 25 nucleotides are considered for this study.

### Identification of IESs Disrupting Protein Domains

We use the results of a previous scan of *Paramecium* proteins for known domains ((Arnaiz et al. 2017), Supplementary Material online). Interpro domains and their relative positions within genes were extracted and checked against IES coordinates. When an overlap was detected the IES was tagged as domain disrupting.

## Supplementary Material

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

## Acknowledgments

We thank the editor and two anonymous reviewers for their careful evaluations of this manuscript and their constructive comments and suggestions. This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—RTG2220—project number 281125614.

## Author Contributions

F.C. conceived the study and carried out computational analysis, supervised, wrote the manuscript, and secured funding. R.R. performed the data acquisition, carried out computational analysis, and wrote the manuscript. V.V. performed the data acquisition and wrote the manuscript. All authors have read and approved the manuscript.

## Data Availability

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

## Literature Cited


Abascal F, et al. 2021. Somatic mutation landscapes at single-molecule resolution. Nature 593(7859):405–410.

Arnaiz O, et al. 2012. The *Paramecium* germline genome provides a niche for intragenic parasitic DNA: evolutionary dynamics of internal eliminated sequences. PLoS Genet. 8(10):e1002984.

Arnaiz O, et al. 2017. Improved methods and resources for paramecium genomics: transcription units, gene annotation and gene expression. BMC Genomics. 18(1):483.

Arnaiz O, Sperling L. 2011. ParameciumDB in 2011: new tools and new data for functional and comparative genomics of the model ciliate *Paramecium tetraurelia*. Nucleic Acids Res. 39(Database issue):D632–D636.

Baudry C, et al. 2009. PiggyMac, a domesticated piggyBac transposase involved in programmed genome rearrangements in the ciliate *Paramecium tetraurelia*. Genes Dev. 23(21):2478–2483.

Belfield EJ, et al. 2021. Thermal stress accelerates *Arabidopsis thaliana* mutation rate. Genome Res. 31(1):40–50.

Betermier M, Duharcourt S. 2014. Programmed rearrangement in ciliates: *paramecium*. Microbiol Spectr. 2(6):1–20. doi:10.1128/microbiolspec.MDNA3-0035-2014.

Bischerour J, et al. 2018. Six domesticated PiggyBac transposases together carry out programmed DNA elimination in *Paramecium*. Elife 7:e37927.

Bjedov I, et al. 2003. Stress-induced mutagenesis in bacteria. Science 300(5624):1404–1409.

Cappucci U, et al. 2019. The Hsp70 chaperone is a major player in stress-induced transposable element activation. Proc Natl Acad Sci U S A. 116(36):17943–17950.

Catania F, McGrath CL, Doak TG, Lynch M. 2013. Spliced DNA sequences in the *Paramecium* germline: their properties and evolutionary potential. Genome Biol Evol. 5(6):1200–1211.

Catania F, Ujvari B, Roche B, Capp J-P, Thomas F. 2021. Bridging tumorigenesis and therapy resistance with a non-darwinian and non-lamarckian mechanism of adaptive evolution. Front Oncol. 11:732081.

Catania F, Wurmser F, Potekhin AA, Przybos E, Lynch M. 2009. Genetic diversity in the *Paramecium aurelia* species complex. Mol Biol Evol. 26(2):421–431.

Chalker DL, Meyer E, Mochizuki K. 2013. Epigenetics of Ciliates. Cold Spring Harb Perspect Biol. 5(12):a017764.

Denby Wilkes C, Arnaiz O, Sperling L. 2016. ParTIES: a toolbox for *Paramecium* interspersed DNA elimination studies. Bioinformatics 32(4):599–601.

Duempelmann L, Skribbe M, Bühler M. 2020. Small RNAs in the transgenerational inheritance of epigenetic information. Trends Genet. 36(3):203–214.

Duharcourt S, Butler A, Meyer E. 1995. Epigenetic self-regulation of developmental excision of an internal eliminated sequence on *Paramecium tetraurelia*. Genes Dev. 9(16):2065–2077.

Duharcourt S, Keller AM, Meyer E. 1998. Homology-dependent maternal inhibition of developmental excision of internal eliminated sequences in *Paramecium tetraurelia*. Mol Cell Biol. 18(12):7075–7085.

Duharcourt S, Lepere G, Meyer E. 2009. Developmental genome rearrangements in ciliates: a natural genomic subtraction mediated by non-coding transcripts. Trends Genet. 25(8):344–350.

Duret L, et al. 2008. Analysis of sequence variability in the macronuclear DNA of *Paramecium tetraurelia*: a somatic view of the germline. Genome Res. 18(4):585–596.

Eisen HN. 2014. Affinity enhancement of antibodies: how low-affinity antibodies produced early in immune responses are followed by high-affinity antibodies later and in memory B-cell responses. Cancer Immunol Res. 2(5):381–392.

Ferro D, Lepennetier G, Catania F. 2015. Cis-acting signals modulate the efficiency of programmed DNA elimination in *Paramecium tetraurelia*. Nucleic Acids Res. 43(17):8157–8168.

Fitzgerald DM, Hastings PJ, Rosenberg SM. 2017. Stress-induced mutagenesis: implications in cancer and drug resistance. Annu Rev Cancer Biol. 1:119–140.

Foissner W, Chao A, Katz LA. 2008. Diversity and geographic distribution of ciliates (Protista: Ciliophora). Biodivers Conserv. 17(2):345–363.

Foster PL. 2007. Stress-induced mutagenesis in bacteria. Crit Rev Biochem Mol Biol. 42(5):373–397.

García-Nieto PE, Morrison AJ, Fraser HB. 2019. The somatic mutation landscape of the human body. Genome Biol. 20(1):298.

Hagen R, Vitali V, Catania F. 2020. Cross-generational effects and non-random developmental response to temperature variation in *paramecium*. Front Cell Dev Biol. 8:584219.

Heidenreich E. 2007. Adaptive mutation in *Saccharomyces cerevisiae*. Crit Rev Biochem Mol Biol. 42(4):285–311.

Hughes AL. 2012. Evolution of adaptive phenotypic traits without positive Darwinian selection. Heredity 108(4):347–353.

Johri P, et al. 2017. Population genomics of *paramecium* species. Mol Biol Evol. 34(5):1194–1216.

Katju V, Bergthorsson U. 2019. Old trade, new tricks: insights into the spontaneous mutation process from the partnering of classical mutation accumulation experiments with high-throughput genomic approaches. Genome Biol Evol. 11(1):136–165.

Krašovec R, et al. 2014. Mutation rate plasticity in rifampicin resistance depends on *Escherichia coli* cell-cell interactions. Nat Commun. 5(1):1–8.

Krašovec R, et al. 2017. Spontaneous mutation rate is a plastic trait associated with population density across domains of life. PLoS Biol. 15(8):e2002731.

Lepere G, Betermier M, Meyer E, Duharcourt S. 2008. Maternal noncoding transcripts antagonize the targeting of DNA elimination by scanRNAs in *Paramecium tetraurelia*. Genes Dev. 22(11):1501–1512.

Lepere G, et al. 2009. Silencing-associated and meiosis-specific small RNA pathways in *Paramecium tetraurelia*. Nucleic Acids Res. 37:903–915.

Lhuillier-Akakpo M, et al. 2014. Local effect of enhancer of zeste-like reveals cooperation of epigenetic and cis-acting determinants for zygotic genome rearrangements. PLoS Genet. 10(9):e1004665.

Loewe L, Hill WG. 2010. The population genetics of mutations: good, bad and indifferent. Philos Trans R Soc Lond B Biol Sci. 365(1544):1153–1167.

Long HA, Doak TG, Lynch M. 2018. Limited mutation-rate variation within the *paramecium aurelia* species complex. G3 8(7):2523–2526.

Lynch M. 2010. Evolution of the mutation rate. Trends Genet. 26(8):345–352.

Lynch M, et al. 2016. Genetic drift, selection and the evolution of the mutation rate. Nat Rev Genet. 17(11):704–714.

Maliszewska-Olejniczak K, et al. 2015. TFIIS-dependent non-coding transcription regulates developmental genome rearrangements. PLoS Genet. 11(7):e1005383.
Martincorena I, Campbell PJ. 2015. Somatic mutation in cancer and normal cells. Science 349(6255):1483–1489.
Meyer E, Garnier O. 2002. Non-Mendelian inheritance and homology-dependent effects in ciliates. Adv Genet. 46:305–337.
Milholland B, et al. 2017. Differences between germline and somatic mutation rates in humans and mice. Nat Commun. 8:15183.
Poduri A, Evrony GD, Cai X, Walsh CA. 2013. Somatic mutation, genomic variation, and neurological disease. Science 341(6141):1237758.
Pritchard JK. 2001. Are rare variants responsible for susceptibility to complex diseases? Am J Hum Genet. 69(1):124–137.
Reusch TBH, Ehlers A, Hämmerli A, Worm B. 2005. Ecosystem recovery after climatic extremes enhanced by genotypic diversity. Proc Natl Acad Sci U S A. 102(8):2826–2831.
Rudman B, Preer LB, Polisky B, Preer JR. Jr 1991. Mutants affecting processing of DNA in macronuclear development in *Paramecium*. Genetics 129(1):47–56.
Russo M, et al. 2019. Adaptive mutability of colorectal cancers in response to targeted therapies. Science 366(6472):1473–1480.
Sandoval PY, Swart EC, Arambasic M, Nowacki M. 2014. Functional diversification of Dicer-like proteins and small RNAs required for genome sculpting. Dev Cell. 28(2):174–188.
Shor E, Fox CA, Broach JR. 2013. The yeast environmental stress response regulates mutagenesis induced by proteotoxic stress. PLoS Genet. 9(8):e1003680.
Singh DP, et al. 2014. Genome-defence small RNAs exapted for epigenetic mating-type inheritance. Nature 509(7501):447–452.
Sonneborn TM. 1957. Breeding systems, reproductive methods, and species problems in protozoa. In: Mayr E, trans, editor. The species problem. Washington (DC): AAAS. p. 155–324.
Sonneborn TM. 1974. *Paramecium aurelia*. In: King RC, editor. Handbook of genetics. Vol. 2 New York and London: Plenum Press. p. 469–594.
Sonneborn TM. 1975. The *paramecium-aurelia* complex of 14 sibling species. Trans Am Microscopical Soc. 94(2):155–178.
Sung W, et al. 2012. Extraordinary genome stability in the ciliate *Paramecium tetraurelia*. Proc Natl Acad Sci U S A. 109(47):19339–19344.
Swart EC, et al. 2017. Identification and analysis of functional associations among natural eukaryotic genome editing components [version 1; referees: 1 approved]. F1000Research. 6:1374.
Thind AS, Vitali V, Guarracino MR, Catania F. 2020. What's genetic variation got to do with it? Starvation-induced self-fertilization enhances survival in *Paramecium*. Genome Biol Evol. 12(5):626–638.
Veltman JA, Brunner HG. 2012. De novo mutations in human genetic disease. Nat Rev Genet. 13(8):565–575.
Vitali V, Hagen R, Catania F. 2019. Environmentally induced plasticity of programmed DNA elimination boosts somatic variability in *Paramecium tetraurelia*. Genome Res. 29(10):1693–1704.
Vitali V, Rothering R, Catania F. 2021. Fifty generations of amitosis: tracing asymmetric allele segregation in polyploid cells with single-cell DNA sequencing. Microorganisms 9(9):1979.
Whitham TG, et al. 2006. A framework for community and ecosystem genetics: from genes to ecosystems. Nat Rev Genet. 7(7):510–523.
Woodard J, Gelber B, Swift H. 1961. Nucleoprotein changes during the mitotic cycle in *Paramecium aurelia*. Exp Cell Res. 23:258–264.
Yizhak K, et al. 2019. RNA sequence analysis reveals macroscopic somatic clonal expansion across normal tissues. Science 364(6444):eaaw0726.

**Associate editor:** Rebecca Zufall