Published in final edited form as:

J Evol Biol. 2016 June ; 29(6): 1284–1292. doi:10.1111/jeb.12868.

Hidden genetic variation in the germline genome of *Tetrahymena thermophila*

Kristen L. Dimond and Rebecca A. Zufall

Department of Biology and Biochemistry, University of Houston, Houston, TX 77204

Abstract

Genome architecture varies greatly among eukaryotes. This diversity may profoundly affect the origin and maintenance of genetic variation within a population. Ciliates are microbial eukaryotes with unusual genome features, such as separation of germline and somatic genomes within a single cell and a mitotic division. These features have previously been proposed to increase the rate of molecular evolution in these species. Here, we assessed the fitness effects of genetic variation in the two genomes of natural isolates of the ciliate *Tetrahymena thermophila*. We find more extensive genetic variation in fitness in the transcriptionally-silent germline genome than in the expressed somatic genome. Surprisingly, this variation is not primarily deleterious, but has both beneficial and deleterious effects. We conclude that *Tetrahymena* genome architecture allows for the maintenance of genetic variation that would otherwise be eliminated by selection. We consider the impacts of reproductive strategies and the mechanism of sex determination on the structure of this variation.

Keywords

population genetics; sex determination; inbreeding depression; ciliate

Introduction

Evolution is dependent on variation among individuals within a population (Darwin, 1859). The amount of genetic variation in a population is determined by many factors such as mutation rate, effective population size, and strength of selection. Genomic structure and content can also affect the amount of genetic variation in a population by influencing the types of mutations that are likely to occur and how efficiently genetic variation is sorted by selection (Lynch *et al.*, 2006, Lynch, 2010).

Corresponding author: KLD, kldimond@uh.edu, 713-743-8136, 4800 Calhoun, Department of Biology and Biochemistry, University of Houston, Houston, TX 77204-5001.

Availability of supporting data: The data set supporting the results of this article is deposited in the Dryad repository [Identifier and hyperlink to be updated following deposition].

Competing interests: The authors declare that they have no competing interests.

Authors' contributions: KD and RZ conceived of, designed, and coordinated the study. KD carried out the experiments. KD and RZ wrote the manuscript.

Genome architecture varies widely across the tree of life. We broadly define genome architecture to include how genomes are contained in a cell and how information is structured within those genomes (McGrath & Katz, 2004, Parfrey *et al.*, 2008, Oliverio & Katz, 2014). For example, among eukaryotes genome size varies by over 4 orders of magnitude (Gregory, 2005a, Gregory, 2015). Similarly, ploidy varies both among species and across life history stages within a species (Parfrey *et al.*, 2008). Some species have unusual genome structures such as subtelomeric rDNA genes, hypervariable karyotypes, or dual genomes (reviewed in McGrath & Katz, 2004). Other species such as bdelloid rotifers rebuild their genomes following massive desiccation-induced double-strand breaks (Hespeels *et al.*, 2014). How this diversity of genome architecture affects the origin and maintenance of genetic variation in a population remain unclear.

Ciliates are microbial eukaryotes that are known for having unconventional genome architecture (Prescott, 1994, Katz, 2001, Zufall *et al.*, 2006) and various features of this architecture have the potential to affect the amount of genetic variation in a population (Zufall *et al.*, 2006, Morgens *et al.*, 2014). Ciliates have two different types of nuclei in a single cell (Fig. 1A). The micronucleus contains a genome that performs germline functions—it undergoes meiosis and is transmitted between cells during conjugation—but no transcription occurs from this genome during most of the life cycle. The macronucleus contains the somatic genome. This genome is highly polyploid and is the site of all transcription during vegetative growth (Prescott, 1994). The dual nature of ciliate genomes allows mutations to accumulate in the germline (micronuclear) genome during many rounds of asexual division without being exposed to selection (Long *et al.*, 2013). This is thought to be responsible for the increased rates of evolution observed in ciliate proteins (Riley & Katz, 2001, Zufall *et al.*, 2006).

Furthermore, some ciliates have unusual mechanisms of nuclear division that also affect genetic variation. Prior to cell division, the micronucleus divides by mitosis, but the macronucleus divides by a process called amitosis. During amitosis, homologous chromosomes are randomly segregated among daughter cells. This results in a phenomenon called phenotypic assortment, where one allele can become fixed in a cell lineage while the other is lost due to selection or drift. This leads to homozygosity in the macronuclear genome regardless of any heterozygosity in the micronuclear genome (Fig. 1A, Martindale *et al.*, 1982, Doerder *et al.*, 1992, Simon & Nanney, 1984). Homozygosity in the macronucleus exposes recessive deleterious mutations, which allows them to be eliminated by selection from the macronuclear genome. However, those recessive deleterious mutations will still be maintained in the micronuclear genome.

Mating system features are highly diverse in ciliates, specifically, variation in the number of mating types (or sexes) and the mechanism of sex determination (Phadke & Zufall, 2009). Although microbial eukaryotes vary greatly in frequency of sexual vs. asexual reproduction and rates of selfing, inbreeding, and outcrossing, little is known about the natural history of reproductive strategies in ciliates (Sonneborn, 1957, Dini & Nyberg, 1993). These features almost certainly have important consequences on the extent and distribution of genetic variation in populations.

To assess the effects of nuclear dualism and phenotypic assortment on the maintenance of genetic variation, we compared the amount of variation in fitness in the micronuclear and macronuclear genomes of natural isolates of *T. thermophila*. We first measured the fitness of 14 natural isolates after >500 generations of asexual reproduction, to evaluate the variation in macronuclear genomes. We then measured the fitness of those isolates after a form of inbreeding that produces whole genome homozygotes and after outcrossing to determine the extent and type of variation present in the micronuclear genomes.

Materials & Methods

T. thermophila Strains

Tetrahymena thermophila wild strains used in this study were isolated from natural populations collected from 4 geographic regions across the known distribution of the species in the eastern United States: New England (NE), Western Pennsylvania (WPA), Eastern Pennsylvania (EPA), and South (SO). We previously demonstrated that these regions are genetically differentiated subpopulations (Zufall *et al.*, 2013). Wild strains have been growing asexually in nutrient-rich media for long enough that we confidently assume the somatic genome has become largely homozygous due to phenotypic assortment (>500 generations; Doerder *et al.*, 1992).

Star strains, which have inactive germline nuclei, were used in inbred crosses (see below). Two star strains, A*III and B*VII were obtained from the *Tetrahymena* Stock Center (Cornell University).

For all fitness assays, we used the reference strain LA011-7 (Zufall *et al.*, 2013) as a plate control. This strain was chosen because it produced the lowest variance both among replicates in a single assay plate for 2 independent assay plates, in a test of 6 possible reference strains.

Inbred Crosses

We generated inbred and outbred progeny to examine the effects of alleles present in the germline genomes of wild strains. Inbred progeny were created using a technique called genomic exclusion (GE; Fig. 1B; Allen, 1963). The progeny of GE crosses are whole genome homozygotes in both nuclear genomes.

GE involves two rounds of crosses. In round 1 (R1), the strain of interest is crossed with a star strain. All strains were grown separately in 6 ml of growth medium (SSP; Lynn & Doerder, 2012) plus 1% Penicillin, Streptomycin, Amphotericin-B (PSA; Amresco, Solon, OH) and incubated at 30°C until reaching mid-late log phase, $\sim 2 \times 10^5$ cells/ml (1-2 days). Cultures were then prepared for starvation by centrifugation at 3000rpm for 3 min. The supernatant was removed and 6ml of 10mM Tris-HCl (pH 7.5) was added to the pellet. Cultures were incubated at 30°C for 2.5 – 3 days to induce sexual competence. Equal volumes (0.5 – 1ml) of a wild strain and star strain were then mixed in one well of a multi-well plate and incubated at 30°C. After 6 – 8 hours we checked plates microscopically for the presence of mating pairs.

For each R1 cross, we isolated 24-48 mating pairs into 800ul of SSP with 1% PSA on a 48 well plate. Plates were incubated for 2-3 days at 30°C. After R1, progeny are homozygous in the germline genome, but the somatic genome remains intact (Fig. 1B). Each R1 mating pair is formed from independent meiotic products and therefore different R1 pairs contain different combinations of alleles from the parental germline genome (Fig. 1B); we call strains derived from these independent R1 pairs “GE lines.”

R1 progeny were then re-mated with each other in GE Round 2 (R2; Fig. 1B), following a similar procedure as R1. For each R1 pair that successfully produced progeny, we inoculated 24-48 new cultures into 6ml SSP with 1% PSA. These were incubated and starved as in R1. Once cells reached sexual competence, the progeny of a single R1 pair were mated among themselves (re-mating), i.e. there was no need to mix cultures for mating. After 6 – 8 hours we checked plates for the presence of mating pairs and isolated 4 - 6 replicate mating pairs per GE line in 800ul of SSP with 1% PSA. R2 progeny are whole genome homozygotes in both germline and somatic nuclei (Fig. 1B; Allen, 1963).

Outbred crosses

Outbred crosses were performed among various wild strains. Mating procedures were similar to inbred crosses. Starved cells of wild strains with complementary mating types were mixed and incubated for 6 – 8 hours, after which we checked plates for the presence of mating pairs and picked pairs for fitness assays (see below). Fitness assays were performed immediately after mating in order to preserve as much heterozygosity in the macronucleus as possible, prior to phenotypic assortment.

Fitness assays

Fitness of each strain was estimated by the maximum population growth rate ($r_{\max}$). Single cells were isolated into 180ul of SSP with 1% PSA in a 96-well plate and pre-incubated at 30°C for 24 hours. Plates were then loaded into a 30°C microplate reader and OD650 was measured every 5 minutes for 72 hours. From these data, $r_{\max}$ was estimated as in Wang *et al.*, (2012) and Long *et al.*, (2013).

For fitness assays on the inbred strains, we scored 2 - 13 GE lines per inbred cross, with 1 – 31 replicates per GE line (Additional file 1). Every assay plate also included 12-24 replicates of the reference strain and 12-24 replicates of the wild strain from which the GE lines were derived. Wells were randomized to control for possible position effects for each GE line, wild strain, and reference strain on the 96-well plate.

For fitness assays on the outbred strains, 60 outbred mating pairs from each cross were isolated into separate wells. Each plate also included 12 replicates of each wild strain that was a parent of that outbred cross and 12 replicates of the reference strain.

Data Analysis

All data were analyzed in JMP (version 8.0.2). Values of $r_{\max}$ were natural-log transformed to increase normality and homogeneity of variances. Data were analyzed using both

arithmetic and geometric means; no significant differences were found between these so all results are shown based on arithmetic means only.

Results

Variation in fitness among asexually dividing wild isolates

Despite over 500 generations of asexual division under laboratory conditions, we found significant differences in fitness among 14 wild isolates (Table 1, Fig. 2A; ANOVA, $F_{13} = 14.90$, $P < 0.001$). There was no detectable effect of geographic region of origin on fitness (ANOVA, $F_3 = 4.83$, $P = 0.825$), although the two most fit isolates (19625-2 & 18901), which are significantly more fit than the remaining isolates (Fig. 2A; Tukey HSD, $Q_{13} = 3.40$, $P < 0.0001$), are both from Western PA. This period of asexual division is long enough that most heterozygosity that may have been present in the macronucleus following any previous mating has since been lost due to phenotypic assortment. In addition, this time in growth under laboratory conditions has also likely resulted in adaptation to the conditions in which we commonly grow cells in lab.

Mating reveals hidden variation in the micronuclear genome

To examine variation present in the micronuclear genome, 2 – 13 independent genomic exclusion lines were created for 13 of the wild isolates (Additional file 1). Each GE line is homozygous across the genome and different GE lines can carry alternate alleles at loci that are heterozygous in the micronucleus of the wild isolate (Fig. 1). The GE progeny are highly variable among lines (Fig. 2B; ANOVA, $F_{60} = 8.00$, $P < 0.0001$) and there is a significant interaction between strain and source, i.e. whether we are measuring fitness of the wild isolate or its GE progeny (Fig. 2; ANOVA, $F_{12} = 4.10$, $P < 0.0001$). This means that the fitness of GE progeny cannot be predicted from the fitness of the wild isolate parent. In addition, for several of the isolates, the independent GE lines differ from one another (Fig. 2B, Fig. 3). Both of these results suggest that there are alleles hidden in the micronuclei that are not present in the macronuclei of the wild isolates.

For a subset of the isolates, we made outbred crosses among wild isolates. Four out of eight crosses produced offspring whose fitness differed significantly from the mean of the parents (Table 2). This result is consistent with the presence of alleles in the micronucleus that are not present in the parental macronucleus. Only two of the outbred crosses produced offspring whose fitness differed significantly from the mean fitness of the GE lines produced from their parents (Table 2). If the difference between outbred progeny and their parents is indeed due to unexpressed variation in the micronuclear genome, then we would expect less of a difference between outbred progeny and inbred progeny because both types of progeny express parental micronuclear alleles. Though the difference between these two effects, i.e. outbred progeny vs. parents and outbred vs. inbred progeny, is not significantly different ($\chi^2 = 2$, $d.f. = 1$, $P = 0.1573$), these results may nonetheless support the presence of hidden variation in the micronuclear genome of the wild isolates. An alternative explanation for these results, which we cannot rule out, is the presence of widespread epistasis.

Micronuclear variation has both beneficial and deleterious effects

Among the GE lines, we found no consistent direction of fitness effects relative to either the wild isolates (Table 1) or the outbred progeny (Table 2). Of the GE progeny that differ significantly from the wild isolates, 4 have on average higher fitness than the wild isolates and 3 have lower fitness (Table 1). When we compare the outbred progeny to their parents, i.e. the wild isolates, we find that 2 of them have significantly higher fitness than the mean of their parents and 2 have significantly lower fitness; the remaining 4 are not significantly different from the mean of their parents (Table 2). There is no detectable relationship between these fitness differences and whether the parents were from the same or different geographic regions (Table 2).

If we consider these results in the context of inbreeding depression, we can compare the fitness of the inbred GE progeny to the outbred progeny and calculate an overall coefficient of inbreeding depression across all lines: the difference between the means of the inbred and outbred progeny fitness, $\delta_{i-o} = 0.0049$, is not significantly different from zero ($t_1 = 1.96$, $P = 0.8093$). Only 2 out of 8 crosses show a significant difference in fitness between the outbred progeny and the mean of the inbred progeny (Table 2). The overall difference between the inbred progeny and wild isolates is likewise not significantly different from zero ($\delta_{i-w} = 0.0339$; $t_1 = 2.064$ $P = 0.066$), nor is the difference between the mean fitness of the outbred progeny and the mean of their parental wild isolates ($\delta_{w-o} = 0.0182$, $t_1 = 1.9632$, $P = 0.449$).

Discussion

The dual nature of ciliate genomes has long been known, however, the evolutionary effects of this unusual genome architecture have yet to be fully understood. Previous work has demonstrated the effects of ciliate genome architecture on rates of protein evolution (Zufall *et al.*, 2006), gene family evolution (Gao *et al.*, 2014), and mutation rate (Sung *et al.*, 2012, H. Long, personal communication). Here, we found that *Tetrahymena* genome architecture allows for the maintenance of genetic variation for fitness in the micronuclear genome despite low variation present in the macronuclear genome.

We measured fitness effects of alleles present in the macronuclear genomes of wild isolates and, by two different mating schemes, measured the fitness effects of alleles present in the micronuclear genomes of those lines. Two lines of evidence support our conclusion of hidden variation in the germline genome. First, we find that the fitness of neither the inbred nor outbred progeny can be predicted based on the fitness of their wild parents. This indicates that there are different alleles affecting fitness in the two genomes. Second, we find that GE lines from a single wild isolate often differ from each other in fitness, revealing the presence of heterozygosity in the micronuclear genome. Both of these results match our expectation that the micronuclear genome will tend to accumulate mutations that are not acted on by selection during asexual division (Zufall *et al.*, 2006, Long *et al.*, 2013) and will maintain allelic diversity for long periods of time (Morgens *et al.*, 2014).

Surprisingly, most of the variation harbored in the micronuclear genome is not deleterious. Given that most mutations are deleterious (Fisher, 1930), if micronuclear mutations are

rarely exposed to selection, we might expect a build up of deleterious alleles in the micronuclear genome. A previous study on *T. thermophila* did indeed find inbreeding depression in viability in this species (Nanney, 1956), likely the result of the expression of recessive deleterious alleles in the micronuclear genome. However, we found an overall coefficient of inbreeding depression on growth rate that was not different from zero and neither inbred nor outbred progeny had consistently lower (or higher) fitness than the wild strains.

One possible explanation for our observed lack of inbreeding depression, or rather lack of deleterious alleles in the micronuclear genome, comes from the mechanism of sex determination and reproductive strategies in this species. *T. thermophila* has seven self-incompatible mating types, or sexes. Normally, such self-incompatibility would promote outcrossing, and allow the accumulation of recessive deleterious alleles in a population (Lande & Schemske, 1985, Charlesworth & Charlesworth, 1987). However, in addition to unusual genome architecture, ciliates also have unconventional mechanisms of sex determination (Phadke & Zufall, 2009). Mating type in *T. thermophila* is determined by a mechanism that we have called probabilistic sex determination (PSD; Paixao *et al.*, 2011). Unlike organisms with exact sex determination, the alleles at the mating type locus (*mat*) in *T. thermophila* specify the probability that one of the seven sexes will be expressed (Nanney & Caughey, 1953, Orias, 1981). A normal mating reaction in this species produces four genetically identical individuals (Nanney & Caughey, 1953, Orias, 1963), however due to PSD these four progeny can express different mating types. This means that despite the fact that mating types are self-incompatible, genetically identical individuals can nonetheless mate with each other (equivalent to selfing). It is likely that PSD, like selfing, may be favored in order to ensure the presence of compatible mates in newly-founded or sparse populations where mates are scarce (reviewed in Goodwillie *et al.*, 2005), which may frequently be the case in *T. thermophila* (Zufall *et al.*, 2013). A side effect of PSD then is that in newly-founded or sparse populations, mating can occur between genetically identical siblings, subsequently resulting in the expression and purging of deleterious recessive alleles from both genomes.

Previously, we demonstrated that *T. thermophila* populations are subdivided and experience little gene flow between geographic regions (Zufall *et al.*, 2013). If these populations are experiencing different selective conditions, we might expect them to become locally adapted. In the presence of local adaptation, progeny of crosses between populations or geographic regions may display “outbreeding depression” (Baker, 1955, Lynch, 1991, reviewed in Goodwillie *et al.*, 2005). However, our results show no correlation between geographic distance between parents and progeny fitness (geographic distances were ranked by region, where parents from the same pond had a rank of 1 followed by WPA-EPA, WPA-NE, and SO-EPA (Table 2); Spearman's $\rho = 0.012$, $d.f. = 6$, $P = 0.90$). Given our small sample size and the fact that we measure fitness under only a single condition, it is possible that we lack power to detect such an effect. Alternatively, it may be that populations become locally adapted in the macronucleus through phenotypic assortment, but that micronuclear variation is not exposed to selection frequently enough to remove alleles that are transferred by rare migration events. Knowledge of the geographic structure of alleles present in the

micronucleus, and further insight into the frequency with which *T. thermophila* undergo sexual reproduction in natural populations (Doerder *et al.*, 1995) would help resolve this.

Overall, our results indicate that *T. thermophila* genome architecture leads to increased genetic variation in fitness hidden in the micronuclear genome. It remains unclear why most of this variation is not deleterious and what role reproductive strategies and gene flow play in determining the extent and type of micronuclear variation. Future investigations into the molecular nature of micronuclear vs. macronuclear genetic variation and further work on the reproductive biology and ecology of this species will help to elucidate these issues.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to Paul Doerder for providing strains and logistical support, Joe West, Hongan Long, and Heather Votaw for helpful discussion and comments. JL Martinez, Stephanie Pham, Jennifer Guerrero, Kyle Long, Laura Potter-Bosch, Tuan Phan, and Isaac Norris helped carry out crosses. Special thanks to: Michael Dimond, Lynne Dimond, and Jillian VanZelfden for providing continuous support. This research was supported in part by the National Institutes of General Medical Science of the National Institutes of Health (award no. R01GM101352) and the Texas Higher Education Coordinating Board (Norman Hackerman Advanced Research Program, 003652-0102-2009).

References

- Allen SL. Genomic exclusion in *Tetrahymena*: Genetic basis. *J Protozool.* 1963; 10(4):413–420. [PubMed: 14074437]
- Baker HG. Self-incompatibility and establishment after ‘long-distance’ dispersal. *Evolution.* 1955; 9:347–348.
- Charlesworth D, Charlesworth B. Inbreeding depression and its evolutionary consequences. *Annu Rev Ecol Syst.* 1987; 18:237–268.
- Darwin, C. *On the Origin of Species.* John Murray; London: 1859.
- Dini, F.; Nyberg, D. Sex in ciliates. In: Jones, JG., editor. *Advances in microbial ecology.* Vol. 13. Plenum Press; New York NY: 1993. p. 85-154.
- Doerder FP, Deak JC, Lief JH. Rate of phenotypic assortment in *Tetrahymena thermophila*. *Dev Genet.* 1992; 13:126–132. [PubMed: 1499154]
- Doerder FP, Gates MA, Eberhardt FP, Arslanyolu M. High frequency of sex and equal frequencies of mating types in natural populations of the ciliate *Tetrahymena thermophila*. *Proc Natl Acad Sci USA.* 1995; 92:8715–8718. [PubMed: 7568003]
- Fisher, RA. *The genetical theory of natural selection.* Oxford University Press; Oxford, UK: 1930.
- Gao F, Song W, Katz LA. Genome structure drives patterns of gene family evolution in ciliates, a case study using *Chilodonella uncinata*. *Evolution.* 2014; 68(8):2287–2295. [PubMed: 24749903]
- Goodwillie C, Kalisz S, Eckert CG. The evolutionary enigma of mixed mating systems in plants: occurrence, theoretical explanations, and empirical evidence. *Annu Rev Ecol Evol Syst.* 2005; 36:47–79.
- Gregory, TR. Genome size evolution in animals. In: Gregory, TR., editor. *The Evolution of the Genome.* Elsevier; San Diego: 2005a. p. 3-87.
- Gregory, TR. Animal Genome Size Database. 2015. See <http://www.genomesize.com>
- Hespeels B, Knapen M, Hanot-Mambres D, Heuskin AC, Pineux F, Lucas S, et al. Gateway to genetic exchange? DNA double-strand breaks in the bdelloid rotifer *Adineta vaga* submitted to desiccation. *J Evol Biol.* 2014; 27(7):1334–1345. [PubMed: 25105197]

- Katz LA. Evolution of nuclear dualism in ciliates: a reanalysis in light of recent molecular data. *Int J Syst Evol Microbiol*. 2001; 51:1587–1592. [PubMed: 11491362]
- Lande R, Schemske DW. The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. *Evolution*. 1985; 39:24–40.
- Long H, Paixao T, Azevedo RBR, Zufall RA. Accumulation of spontaneous mutations in the ciliate *Tetrahymena thermophila*. *Genetics*. 2013; 195:527–540. [PubMed: 23934880]
- Lynch M. The genetic interpretation of inbreeding depression and outbreeding depression. *Evolution*. 1991; 45(3):622–629.
- Lynch M, Koskella B, Schaack S. Mutation pressure and the evolution of organelle genomic architecture. *Science*. 2006; 311:1727–1730. [PubMed: 16556832]
- Lynch M. Evolution of the mutation rate. *Trends Genet*. 2010; 26:345–352. [PubMed: 20594608]
- Lynn, DH.; Doerder, FP. The life and times of *Tetrahymena*. In: *Methods in Cell Biology*. In: Collins, K., editor. *Tetrahymena thermophila*. Vol. 109. Academic Press; San Diego: 2012. p. 12–25.
- Martindale DW, Allis CD, Bruns PJ. Conjugation in *Tetrahymena thermophila*. A temporal analysis of cytological stages. *Exp Cell Res*. 1982; 140:227–236. [PubMed: 7106201]
- McGrath CL, Katz LA. Genome diversity in microbial eukaryotes. *Trends Ecol Evol*. 2004; 19:32–38. [PubMed: 16701223]
- Morgens DW, Strutz TC, Cavalcanti ARO. Novel population genetics in ciliates due to life cycle and nuclear dimorphism. *Mol Biol Evol*. 2014; 31(8):2084–2093. [PubMed: 24784136]
- Muller HJ. The relation of recombination to mutational advance. *Mutat Res*. 1964; 106:2–9. [PubMed: 14195748]
- Nanney DL. Inbreeding degeneration in *Tetrahymena*. *Genetics*. 1956; 42(2):137–146. [PubMed: 17247685]
- Nanney DL, Caughey PA. Mating type determination in *Tetrahymena pyriformis*. *Proc Natl Acad Sci USA*. 1953; 39(10):1057–1063. [PubMed: 16589373]
- Oliverio AM, Katz LA. The dynamic nature of genomes across the tree of life. *Gen Biol Evol*. 2014; 6(3):482–488.
- Orias E. Mating type determination in variety 8, *Tetrahymena pyriformis*. *Genetics*. 1963; 48(11):1509–1518. [PubMed: 14083046]
- Orias E. Probable somatic DNA rearrangements in mating type determination in *Tetrahymena thermophila* – A review and model. *Dev Genet*. 1981; 2:185–202.
- Paixao T, Phadke SS, Azevedo RBR, Zufall RA. Sex ratio evolution under probabilistic sex determination. *Evolution*. 2011; 65(7):2050–2060. [PubMed: 21729059]
- Parfrey LW, Lahr DJG, Katz LA. The dynamic nature of eukaryotic genomes. *Mol Biol Evol*. 2008; 25:787–794. [PubMed: 18258610]
- Phadke SS, Zufall RA. Rapid diversification of mating system in ciliates. *Biol J Linn Soc*. 2009; 98:187–197.
- Prescott DM. The DNA of ciliated protozoa. *Microbiol Rev*. 1994; 58:233–267. [PubMed: 8078435]
- Riley JL, Katz LA. Widespread distribution of extensive chromosomal fragmentation in ciliates. *Mol Biol Evol*. 2001; 18(7):1372–1377. [PubMed: 11420375]
- Simon EM, Nanney DL. Karyonidal inheritance of mating types in *Tetrahymena melaccensis*. *Dev Genet*. 1984; 5:43–58.
- Sonneborn, TM. Breeding systems, reproductive methods, and the species problems in protozoa. In: Mayr, E., editor. *The species problem*. AAAS; Washington DC: 1957. p. 155–324.
- Sung W, Tucker AE, Doak TG, Choi E, Thomas WK, Lynch M. Extraordinary genome stability in the ciliate *Paramecium tetraurelia*. *Proc Natl Acad Sci USA*. 2012; 109:19339–19344. [PubMed: 23129619]
- Wang Y, Diaz AC, Stoebel DM, Cooper TF. Genetic background affects epistatic interactions between two beneficial mutations. *Biol Lett*. 2012; 9:201–208?.
- Zufall RA, McGrath CL, Muse SV, Katz LA. Genome architecture drives protein evolution in ciliates. *Mol Biol Evol*. 2006; 23(9):1681–1687. [PubMed: 16760419]
- Zufall RA, Dimond KL, Doerder FP. Restricted distribution and limited gene flow in the model ciliate *Tetrahymena thermophila*. *Mol Ecol*. 2013; 22(4):1081–1091. [PubMed: 23094694]

Figure 1.

(A) Sexual conjugation produces progeny that are heterozygous in the diploid micronucleus (MIC) and the polyploid macronucleus (MAC). During asexual division, amitosis results in unequal segregation of homologous chromosomes and loss of heterozygosity in the MAC, termed phenotypic assortment. Cells that have undergone phenotypic assortment may be heterozygous in the MIC despite the lack of diversity in the MAC. (B) Genomic exclusion is a procedure that produces whole genome homozygotes in both the MIC and MAC. In round 1 (R1), cells are mated with a star strain. This produces cells that are homozygous in the MIC, but retain the parental MAC. When R1 cells are remated in round 2 (R2), the progeny are homozygous in both genomes. Independent R1 mating pairs can produce progeny that carry different parental alleles.

Figure 2. Relative fitness of (A) wild strains and (B) inbred progeny. Y-axis shows the mean of $\ln(r_{\max})$ for each line relative to the reference strain. Error bars are standard error. Letters indicate results of Tukey HSD test; values not connected by the same letter are significantly different from each other at $P = 0.05$. In (B), * indicates lines where inbred progeny fitness differs significantly from their parental wild strain (Table 1), and + indicates a strain where GE lines from a single wild strain differed significantly from each other (Table 1, Fig. 3).

Figure 3.

Relative fitness of GE lines for strains (A) 18473-2, (B) 21157-1, (C) 20438-1, (D) 20452-5, and (E) FL032-2. Y-axis and letters are the same as in Fig. 2. Strains 21157-1 and FL032-2 show a significant effect of GE line in an ANOVA (Fig. 2), but a Tukey HSD test cannot distinguish among GE lines for these strains.

Table 1

Relative fitness of wild strains and inbred progeny

Strain	Region	Fitness of wild strain (SE) ^a	N _w	Fitness of inbred strains (SE) ^a	N _i	$\delta_{(i-w)}$	Inbred vs. wild (p-value)	Effect of GE line (p-value)
18473-2	WPA	-0.0751 (0.0419)	19	-0.1259 (0.0359)	26	-0.0608	0.2766	0.0008 ⁺
19624-4	WPA	-0.0557 (0.0346)	30	-0.0699 (0.0515)	28	-0.0821	0.1912	0.1750
19625-2	WPA	0.6327 (0.0603)	13	0.3768 (0.0463)	25	-0.2358	0.0048 [*]	0.8995
21157-1	WPA	-0.0098 (0.0591)	5	-0.0637 (0.0260)	32	-0.0829	0.2082	0.0490 ⁺
18901-1	WPA	0.5147 (0.0605)	12	na	na	na	na	na
19179-2	NE	-0.0493 (0.0673)	8	-0.0448 (0.0359)	31	0.0071	0.926	0.9542
20429-1	NE	-0.1139 (0.0433)	9	0.2846 (0.0238)	34	0.4263	<0.0001 [*]	0.062
20430-1	NE	0.1323 (0.0540)	7	-0.1342 (0.0154)	87	-0.2636	<0.0001 [*]	0.1061
20430-2	NE	-0.0204 (0.0426)	17	0.0931 (0.0328)	28	0.1157	0.0405 [*]	0.9778
20438-1	NE	0.1999 (0.0545)	14	-0.0120 (0.0370)	34	-0.1570	0.0216 [*]	<0.0001 ⁺
20452-5	NE	-0.0190 (0.0370)	15	0.1355 (0.0257)	32	0.1637	0.0008 [*]	0.0006 ⁺
20453-1	NE	-0.1891 (0.0634)	7	-0.2280 (0.0538)	18	-0.0517	0.5406	0.8229
19147-1	EPA	0.0175 (0.0390)	31	-0.2359 (0.0312)	79	-0.2600	<0.0001 [*]	0.2495
FL032-2	SO	0.0101 (0.0500)	18	0.0436 (0.0397)	29	-0.0456	0.4794	0.0394 ⁺

^a Fitness values are the mean of the $\ln(r_{\max})$ relative to the reference strain and SE is standard error.

N_w and N_i are the sample sizes for the wild strains and inbred progeny, respectively.

$\delta_{(i-w)}$ is the difference between the fitness of the inbred and wild progeny. Negative values indicate lower fitness of inbreds.

^{*} indicates a significant effect of strain type on fitness in ANOVA.

⁺ indicates a significant difference among GE lines in ANOVA.

Table 2

Relative fitness of outbred progeny

		Wild vs. outbred		Inbred vs. outbred	
Parent A (region)	Parent B (region)	Fitness of outbred progeny (SE)	N ₀	δ	p-value
18473-2 (WPA)	19624-4 (WPA)	-0.1221 (0.0351)	29	0.0589	0.1629
18473-2 (WPA)	19147-1 (EPA)	-0.2485 (0.0382)	28	0.2308	< 0.0001*
19624-4 (WPA)	19147-1 (EPA)	0.0244 (0.0394)	28	-0.0429	0.3686
19624-4 (WPA)	20453-1 (NE)	-0.2381 (0.0540)	12	0.1696	0.0078*
21157-1 (WPA)	20452-5 (NE)	0.0098 (0.0506)	13	-0.0265	0.6845
20430-2 (NE)	20452-5 (NE)	0.0733 (0.0265)	37	-0.0930	0.0182*
FL032-2 (SO)	19147-1 (EPA)	-0.0986 (0.0611)	15	0.1134	0.1065
19624-4 (WPA)	20438-1 (NE)	0.1502 (0.0403)	33	-0.1246	0.0209*
					0.1884
					0.0002*

N₀ indicates the number of progeny scored.

δ indicates the difference in fitness between the mean of the parental wild strains ('wild vs. outbred') or mean of the inbred progeny from those parental strains ('inbred vs. outbred') and the outbred progeny. Positive values indicate lower fitness of outbred strains.

* indicates a significant difference between the wild strain or inbred progeny compared to the outbred progeny using a T-test.