# Spontaneous mutation accumulation and reproductive strategies in free-living ciliates

---

A Dissertation

Presented to

the Faculty of the Department of Biology and Biochemistry

University of Houston

---

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

By Hongan Long

December 2012

# Spontaneous mutation accumulation and reproductive strategies in free-living ciliates

**Hongan Long**

APPROVED BY:

**Rebecca A. Zufall, Chair**
Department of Biology and Biochemistry

**Ricardo B. R. Azevedo**
Department of Biology and Biochemistry

**Tim F. Cooper**
Department of Biology and Biochemistry

**Dan Graur**
Department of Biology and Biochemistry

**Michael H. Kohn**
Department of Ecology and Evolutionary Biology,
Rice University

**Mark A. Smith**
Dean, College of Natural Sciences and Mathematics

To my most wonderful advisor, Becky Zufall

# Acknowledgements

Being a great biologist is my dream ever since childhood and getting a Biology PhD degree is a necessary step towards this dream, but I never imagined that I would have quit two Biology PhD programs and eventually attained this goal under the supervision of my dream advisor in my current 3rd Biology PhD program.

Yes, Becky is my dream advisor, especially because she is super nice to us students. Under her supervision, I got complete academic freedom, all possible financial and mental support for my research within her capacity, as well as her concerns with my family in Houston and great patience for my hasty personality. My harshness and rudeness to her occasional mistakes are in sharp contrast to her forgiveness and politeness to my lack of organization and endless mistakes. In another word, I was literally spoiled during my time being a student of hers. Ricardo, Tim, Dan, and Michael are four other professors on my committee. Through collaborating with Ricardo and bugging him for technical/analytical details, I really feel educated and inspired by a top evolutionary biologist. I still remember the moments when he was driven crazy by my messy raw data and he almost punched his own face because of my unclear expression when we had a discussion, but with his forgiveness and my soon realizing what was going on, I just hope to have the opportunity to collaborate with him again in the future. Tim provided essential equipment and technical support and deep insights for finishing the major part of the project. I earned my lowest grade in his Microbial Population Biology course; however, that was one of the courses I learned the most on microbiology. Speaking of courses, Dan is my favorite course instructor and I took two of his courses. His broad

knowledge and funny teaching style made classes like thought-provoking TV shows, and students in his classes rarely got sleepy (even if they did, his yelling on them won't let that come true). He always provides strong supporting reference letters for me. Whenever I had requests to bug him, he always tried his best to help me and I could feel consistently his kind heart under his sometimes strict-looking face. Michael is the one I bugged least, mostly because I am a slacker in going to his office at Rice University; however, he is very forgiving and always gave very straight-forward advice as well as encouragement during the committee meetings.

I also appreciate other faculty/staff members in the Department of Biology and Biochemistry, University of Houston, Tony Frankino, Diane Wiernasz, Blaine Cole, Anne Delcour, Stuart Dryer, Elizabeth Ostrowski, Yonia Pulido, and Amanda Paul for their consistent help and encouragement throughout my time in Houston.

Rolf Lohaus and Tiago Paixão from Ricardo's lab, Liang Qu, Yabin Cheng, and Steve Pennings' lab members helped me a lot in sample collection or data analysis and are greatly appreciated. Peter Keightley at University of Edinburgh helped me in providing programs for mutation accumulation data analysis. I really appreciate his kind help and would continue reading his great papers on evolutionary genetics.

Drs. Larry Rapp, Tom Vida, and Richard Knapp are my favorite coordinators during my four years' working as a teaching assistant. Their teaching trainings, the professional coordination or personal help made my part of the job much easier than it was supposed to be.

I would also like to thank my friends in Houston and China for having fun with me or helping me through my PhD time: Xiangan Liu picked me up from the airport the first day I arrived in the US and now I am still renting his condo with a price much lower than the market one; Sujal Phadke, a previous Zufall lab member, helped me a lot in grant proposal writing, personal life, as well as being a close friend; Kevin J. Spring, also a previous Zufall lab member and close friend, saved me from my first car accident. He, his wife, his sister, and their friends literally provided all the baby stuff for our first baby, and greatly relieved my financial pressure; I also thank Kristen Dimond, my current lab mate, for always talking in the lab and keeping me from getting aphasia. Other friends in the Department of Biology and Biochemistry, and the University of Houston include: Xiaopeng Shen, Drew Russey, Lara Appleby, Gayani Rajapaksa, Haosi Chen, Ruijuan Zhu, Nicholas Price, Xin Wang, Chen Liu, Yichen Zheng, Sneha Koneru, Wenfu Li, Maithili Gupte, Hongyu Guo, Yinhua Wang, Daniel Jeffcoat, Stephanie Pham, Robert Hudson, Jun Wang, Fen Peng, Brian Mahon, Bingjun Zhang, Gang Liu, Lian Lin, Xiao Xu, Li Yao, Ming Xiao, Jing Xu, Weiguo Zhang, Zhen Zheng, etc., I really appreciate their making my time in U of H so memorable and wonderful and helping me out of any difficult situations in all kinds of ways. I am grateful to the Tenney's family, Mike, Sarah, and their kids John, Colin, and Beth, for their Houstonians' hospitality and friendship. For those friends I did not mention here, you are all in my heart (saying learned from my undergraduate advisor Dr. Fan).

My family in the US and in China are strong and consistent supporters for my pursuit for academics. My parents in China showed extreme forgiveness and openness to

my repeatedly quitting PhD programs and provided financial/mental support anytime I needed. They came to the US to take care of my first baby, though not long, even with their own health issues and without any knowledge of English. My wife Ping has been with me for more than 12 years and even indirectly sacrificed her PhD program in China for me. She basically takes over all the hard work of taking care of our daughter and supports me for anything I want (except too many beers or wines). With her efforts, my little daughter Anna is healthily and happily growing up day by day. I feel so lucky to have such wonderful and beautiful wife and daughter. My parents-in-law are also thoughtful and really nice to me, and never complained about their eldest daughter never going back home ever since 2009 and never getting a chance to see their first grand-child in person. My elder brother and sister in-law in China are taking care of our parents in our hometown, and this greatly reduces my worries. I feel so fortunate to have such a harmonious and supportive big family and hope I would be able to do something in return in the future.

The funding for my PhD research from Sigma Xi-GIAR, Houston Coastal Center, and University of Houston-GEAR are greatly appreciated. I am also grateful to the Chip-Perinatal, Medicaid, WIC, and Food Stamp programs from the Texas government to help me raise my small family here in the US. I also really thank U of H alumni for the University of Houston Asian Alumni Johnson Tsui Memorial Scholarship and the Friends of NSM Graduate Fellowship.

# Spontaneous mutation accumulation and reproductive strategies in free-living ciliates

---

An Abstract of a Dissertation

Presented to

the Faculty of the Department of Biology and Biochemistry

University of Houston

---

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

By Hongan Long

December 2012

## Abstract

Knowledge of the frequency and fitness effects of mutations is essential for understanding a diversity of issues in evolution; thus, many efforts have gone into elucidating the mutational process. However, our understanding of mutation is far from complete, largely due to the fact that mutations are rare and frequently eliminated by natural selection. This renders studies of the mutational process inherently difficult. The ciliate *Tetrahymena thermophila* provides a unique opportunity to overcome these difficulties by allowing the accumulation of mutations over many generations in the absence of selection on the germline genome. Estimates of the rate and fitness effect of mutations, the first from the eukaryotic supergroup Chromalveolata, are within the range of those of previously studied eukaryotes. Mutations are partially recessive on average and the rate of lethal mutations is substantially lower than the deleterious mutation rate.

Germline mutation accumulation in *T. thermophila* used the reproductive idiosyncrasy that the germline genome only gets expressed during sexual reproductions; thus, germline mutations could be hidden from selection during asexual transfers. There might be other useful reproduction strategies for evolutionary genetics, so as a start, another marine free-living ciliate *Glauconema trihymene* was explored and shown to have diverse asexual reproductive strategies. Selfing was also observed with peculiar macronuclear events in *G. trihymene*.

# Table of Contents

# Chapter 1 Introduction

## 1.1 What are mutations and mutational parameters?

The term "mutation" was coined by the Dutch botanist and geneticist Hugo de Vries in the late 19th century. It referred to the sudden origin of new morphological forms of the plant evening primroses (*Oenothera lamarckiana*) (DE VRIES 1905), though those new forms could also be from other genetic processes (EMERSON 1935), *e.g.*, meiotic recombination. The current view on mutation is DNA sequence change (RNA sequence change in the case of RNA based organisms). Such change varies from a nucleotide to a whole genome (reviewed in GRAUR in press). Mutations causing harmful effects are called deleterious mutations, those increasing fitness of the individuals carrying them are called beneficial mutations, while others with no effects are called neutral mutations (reviewed in GRAUR in press).

Major mutational parameters involved in this study are: genomic mutation rate, mutational effect, and dominance coefficient. The genomic mutation rate refers to the frequency of mutations occurring in one genome and is essential in evolutionary models and predictions. For example, the observation that the deleterious mutation rate is less than 1 per genome per generation in several organisms suggests that selection against deleterious mutations is not sufficient to explain the widespread maintenance of sexual reproduction (HALLIGAN and KEIGHTLEY 2009; KONDRASHOV 1988). There are two units for the mutation rate in this study: mutational events with fitness effects per haploid

genome per generation from the Bateman-Mukai estimates (assuming equal and additive mutational effects) and nucleotide changes per nucleotide site per generation from genome sequencing. Mutational effect is the consequence of a spontaneous mutation and it is usually deleterious (harmful) (EYRE-WALKER and KEIGHTLEY 2007; FISHER 1930; SILANDER *et al.* 2007). The distribution of fitness effects is one important evolutionary topic (see section 1.2 for details) (reviewed by EYRE-WALKER and KEIGHTLEY 2007). Dominance coefficient or degree of dominance describes the effect of a mutation on a heterozygote carrier. It is measured as the fitness difference between the non-mutated homozygote and the heterozygote genotypes, relative to the fitness difference between the non-mutated homozygote and the homozygote mutant genotypes. Current research shows that mutations are generally incompletely recessive (reviewed in CHARLESWORTH 2009; HALLIGAN and KEIGHTLEY 2009).

## 1.2 Why study mutations and use *T. thermophila* as the experimental subject?

Mutations create genetic variation and cause many human diseases like cancers, and can lead to senescence (HALLIGAN and KEIGHTLEY 2009; LYNCH 2010; LYNCH *et al.* 1999). They are crucial in understanding evolutionary topics like the maintenance of sex (OTTO and LENORMAND 2002), genetic variation at the molecular level (CHARLESWORTH *et al.* 1995) and persistence of small populations (LANDE 1994). Intensive research has been conducted on mutations; however, our understanding of mutation is far from complete largely due to the fact that mutations are inherently difficult to study as they are rare and frequently eliminated by natural selection. Among eukaryotes, mutational parameters have only been estimated in Opisthokonta (which includes animals and fungi). The first

human mutation rate (for hemophilia) was measured early by J.B.S. Haldane (HALDANE 1935; HALDANE 1946; HALDANE 1949); however, the mutation rates of the majority of eukaryotic groups are still unexamined (reviewed in HALLIGAN and KEIGHTLEY 2009).

As the only ciliate that could be cultured in chemically defined medium and easily frozen in liquid nitrogen, and which has the most diverse genetic tools, the ciliate *T. thermophila* also has one transcriptionally active macronucleus (containing the somatic genome) and one inert micronucleus (containing the germline genome) like most ciliates. It is not unusual for *T. thermophila* to become incapable of producing viable progeny from sexual reproduction after thousands of generations of asexual reproductions (NANNEY 1974). This was assumed by NANNEY (1974) to be related to the accumulation of deleterious mutations in the germline micronucleus. The dimorphic nuclear apparatus of *T. thermophila* thus provides a unique model to study the germline mutation accumulation in the absence of selection.

### 1.3 How to study mutations?

There are many ways to study mutation rate (reviewed in LYNCH 2010); however, the major way to study genomic mutation rate, fitness effects of mutations, as well as dominance coefficient is through mutation accumulation (MA) experiments (EYRE-WALKER and KEIGHTLEY 2007; HALLIGAN and KEIGHTLEY 2009).

MA experiments were pioneered by Hermann Joseph Muller using *Drosophila* (MULLER 1927; MULLER 1928). MUKAI (1964) estimated genomic mutational parameters using such experiments, which were recently conducted in other model organisms like

*Arabidopsis* (SCHOEN 2005; SHAW et al. 2000), *Caenorhabditis elegans* (KEIGHTLEY and CABALLERO 1997; VASSILIEVA et al. 2000), *Escherichia coli* (KIBOTA and LYNCH 1996), *Saccharomyces cerevisiae* (ZEYL and DEVISSER 2001), and Φ6 virus (BURCH et al. 2007).

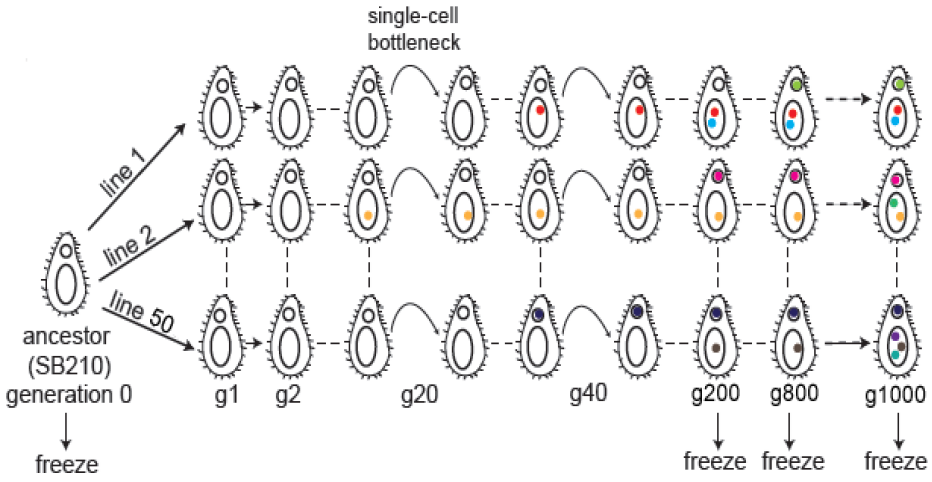


**Figure 1.1 Experimental protocol for mutation accumulation transfers.** Starting from a single *T. thermophila* SB210 cell, 50 independent lines were grown for 1000 generations. Every 20 generations (~3 days), a single cell was transferred from early log-phase culture and every ~200 generations the lines were cryopreserved. Colored dots indicate the independent accumulation of mutations in the germline and somatic genomes. Picture courtesy of Rebecca A. Zufall.

In a typical MA study, mutation accumulation lines originating from the same ancestor line are cultured over many generations through inbreeding and population bottle-necking. Selection is greatly reduced in these experiments, so that drift dominates over selection (reviewed in HALLIGAN and KEIGHTLEY 2009). Because most mutations with fitness effects are deleterious, the fitness of cell lines in an MA experiment would decrease compared with the ancestor fitness; different MA lines usually accumulate different mutations, so the among-line variance of fitness increases. The fitness decrease ($\Delta M$) and among-line variance increase ($\Delta V$) thus become signs of MA and these values

can be used to calculate the Bateman-Mukai estimates of mutational parameters (BATEMAN 1959; MUKAI 1964) (see details in section 2.2.3, chapter 2).

Fitness, either relative or absolute, is usually determined with competition assays or fitness assays on fitness metrics, usually the growth rate in microbes. Different fitness metrics can result in large variation in mutational estimates (VASSILIEVA *et al.* 2000). The mutational effect and genomic mutation rate could then be estimated based on the fitness change $\Delta M$ and among-line variance change $\Delta V$ between the ancestor and the final evolved lines, using the Bateman-Mukai method (BATEMAN 1959; MUKAI 1964). There are also recent methods based on maximum likelihood analyses (KEIGHTLEY 2000; SHAW *et al.* 2002), which could even include fitness from lines at multiple generations. The dominance coefficient could be estimated by fitness assays on parental lines and on progeny from the crossing of the ancestor and the final evolved individuals in the same lineage. Long-term storage, *e.g.*, cryopreservation, is thus needed for ideal ancestor controls, which are feasible in microbes and some multicellular organisms, while not easy in *Drosophila*.

MA experiments are not the only experimental tools used especially for the genomic mutation rate estimation. With the advent of the next-generation genome sequencing, the genomic mutation rate has been shown to be underestimated in MA experimental studies, in which only mutations with fitness effects above a certain level could be detected; thus, estimating genomic mutation rate can also be done using genome-sequencing (LYNCH *et al.* 2008). However, distribution of fitness

effects/dominance coefficient estimations require fitness assays on MA to be conducted, and these parameters are far from being understood (HALLIGAN and KEIGHTLEY 2009).

### 1.4 Why study mutations on morphological traits?

Morphological species identification has been based on morphological traits for centuries; however, the variation of those traits within one population has rarely been empirically tested. This is one reason why the morphological species diversity of ciliates was challenged (CARON 2009; FINLAY and CLARKE 1999). In addition, phylogenetic trees based on morphological traits and those based on marker genes were not consistent with each other (WIENS 2004), studying the mutational effects on morphological traits might give insight on such inconsistency. Morphological traits have also been shown to be useful in estimating mutational parameters (AZEVEDO *et al.* 2002).

### 1.5 Why study reproductive strategies of ciliates?

As shown by the genetic tools used in the above mutation accumulation experiments, the unique nuclear structure and reproductive strategies helped overcome difficulties posed in other model organisms, *e.g.*, mutations were silent in the germline genome during asexual transfers and got expressed by a special sexual reproduction process. Nonetheless, *T. thermophila* is known to be a ciliate with relatively simple life history/reproductive strategies; thus, exploring the life history traits like reproductive strategies of other non-model ciliates might provide more perspective into ciliate evolution or genetics, despite that they are rarely studied (LYNN 2008).

## 1.6 The study systems, *Tetrahymena thermophila* and *Glauconema trihymene*

The unusual nuclear architecture, diverse genetic tools, and short generation time of the ciliated unicellular eukaryote *Tetrahymena thermophila* (Chromalveolata) makes it particularly suitable for MA experiments. *T. thermophila*, like most ciliates, maintains two types of nuclear genomes: a transcriptionally active somatic genome in the macronucleus, and a sequestered, largely transcriptionally inert germline genome in the micronucleus. Because the germline genome can remain silent throughout many rounds of asexual reproduction, mutations that occur in that genome are hidden from selection until they get expressed after conjugation. Thus, all mutations, regardless of fitness effect, have an equal probability of fixing in the germline genome. Genetic tools like genomic exclusion also make this ciliate particularly suited for the purpose of mutation accumulation (BRUNS and CASSIDY-HANLEY 1999). Accumulation of deleterious mutations in the germline genome of lab cultures of *Tetrahymena* was first observed over 50 years ago (NANNEY 1959). Remarkably, strains with deteriorated, or even absent, germline genomes are nonetheless capable of vigorous asexual reproduction (ALLEN *et al.* 1984; NANNEY 1974).

Free-living ciliates are known to exhibit diversity in modes of reproduction (LYNN 2008) and these reproduction modes provide unique tools for genetic studies, like the type of conjugation (genomic exclusion, see section 2.4 in chapter 2 for details) used in the MA studies of *T. thermophila*. *Glauconema trihymene* is a free-living cosmopolitan scuticociliate, which belongs to the class Oligohymenophorea including *Tetrahymena*. It is widely distributed in warm coastal waters, while extremely understudied (LONG and

ZUFALL 2010). Several reproductive strategies that have been previously unknown in the free-living ciliate *G. trihymene* were discovered and might provide potential tools for evolutionary genetic studies of ciliates.

## 1.7 Dissertation outline

This dissertation is aimed at conquering the difficulties unavoidable in most other mutation accumulation systems using the idiosyncratic nuclear genomes and reproductive strategies of the ciliate *T. thermophila*, finding morphological traits variation after long-term mutation accumulation, as well as exploring the peculiar reproductive strategies of another marine ciliate *G. trihymene*. I addressed the following questions: (1) What are the genomic mutation rate, fitness effects, and dominance coefficient in the germline genome without being affected by selection in *T. thermophila* (Chapter 2)? Are the mutational parameters of the somatic and the germline genomes identical within one species (Chapter 2)? (2) Do morphological traits, which are used in ciliate taxonomy, change after germline mutations get expressed after long term asexual divisions (Chapter 3)? (3) What are the reproductive strategies in the non-model ciliate *G. trihymene* (Chapter 4)?

In chapter 2, I conducted one large MA experiment on *T. thermophila*, using single-cell transfers, fitness assays, and genetic tools by manipulating its unique sexual reproductive strategies, as well as individual-based simulations (simulations done by collaborators Tiago Paixão (TP) and Ricardo B. R. Azevedo (RBRA)) to understand the mutational dynamics in the somatic and the germline genomes. I show that mutations were successfully accumulated in the germline genome and mutational estimates are

within the range of other studied organisms. The somatic genome had very different mutational dynamics, which needs further study, from those of the germline genome.

In chapter 3, morphology of cells bearing germline mutations and somatic mutations are compared with that of the ancestors. Four morphological traits: numbers of somatic kineties (No. SK) and post-oral kineties (No. PK), macronucleus (Ma), and cell sizes, as well as the fitness were observed and analyzed. My results showed No. PK had no variability (thus good for ciliates' taxonomy), but not the macronucleus and cell size, which showed a drop from the ancestor. The cell size drop also showed positive correlation with that of the fitness, which is known to be affected by mutations, indicating that cell size is not an ideal trait for morphological species identification/definition.

In chapter 4, reproductive strategies of another free-living but non-model ciliate *G. trihymene* were explored. This species was shown to be the one known with the most diverse asexual reproduction modes among all free-living ciliates: fission, reproductive cysts, and asymmetric division. The selfing in this species also showed peculiar nuclear events.

## 1.8 References


ALLEN, S., P. ERVIN, N. MCLARSEN, and R. BRAND, 1984 The 5S ribosomal RNA gene clusters in *Tetreahymena thermophila*: strain differences, chromosomal localization, and loss during micronuclear ageing. Mol Gen Genet **197:** 244–253.

AZEVEDO, R. B. R., P. D. KEIGHTLEY, C. LAUREN-MAATTA, L. L. VASSILIEVA, M. LYNCH *et al.*, 2002 Spontaneous mutational variation for body size in *Caenorhabditis elegans*. Genetics **162:** 755–765.

BATEMAN, A. J., 1959 The viability of near-normal irradiated chromosomes. Int J Radiat Biol **1:** 170–180.

BRUNS, P., and D. CASSIDY-HANLEY, 1999 Methods for genetic analysis. Methods Cell Biol **62:** 229–240.

BURCH, C. L., C. GUYADER, D. SAMAROV, and H. SHEN, 2007 Experimental estimate of the abundance and effects of nearly neutral mutations in the RNA virus Φ6. Genetics **176:** 467–476.

CARON, D., 2009 Protistan biogeography: why all the fuss? J Eukaryot Microbiol **56:** 105–112.

CHARLESWORTH, B., 2009 Effective population size and patterns of molecular evolution and variation. Nat Rev Genet **10:** 195–205.

CHARLESWORTH, D., B. CHARLESWORTH, and M. T. MORGAN, 1995 The pattern of neutral molecular variation under the background selection model. Genetics **141:** 1619–1632.

DE VRIES, H., 1905 *Species and varieties: their origin by mutation.* Open Court Publishing Co., Chicago.

EMERSON, S. H., 1935 The genetic nature of de Vries' mutations in *Oenothera lamarckiana*. Amer Nat **69:** 545–559.

EYRE-WALKER, A., and P. D. KEIGHTLEY, 2007 The distribution of fitness effects of new mutations. Nat Rev Genet **8:** 610–618.

FINLAY, B. J., and K. J. CLARKE, 1999 Ubiquitous dispersal of microbial species. Nature **400:** 828.

FISHER, R., 1930 *The genetical theory of natural selection*. Clarendon Press, Oxford.

GRAUR, D., *Fundamentals of molecular evolution*. In press.

HALDANE, J. B. S., 1935 The rate of spontaneous mutation of a human gene. J Genet **31:** 317–326.

HALDANE, J. B. S., 1946 The mutation rate of the gene for haemophilia, and its segregation ratios in males and females. Ann Hum Genet **13:** 262–271.

HALDANE, J. B. S., 1949 The rate of mutation of human genes. Hereditas **35:** 262–273.

HALLIGAN, D. L., and P. D. KEIGHTLEY, 2009 Spontaneous mutation accumulation studies in evolutionary genetics. Annu Rev Ecol Evol Syst **40:** 151–172.

JOSEPH, S., and D. HALL, 2004 Spontaneous mutations in diploid *Saccharomyces cerevisiae*: more beneficial than expected. Genetics **168:** 1817–1825.

KEIGHTLEY, P. D., and T. M. BATAILLON, 2000 Multigeneration maximum-likelihood analysis applied to mutation-accumulation experiments in *Caenorhabditis elegans*. Genetics **154:** 1193–1201.

KEIGHTLEY, P. D., and A. CABALLERO, 1997 Genomic mutation rates for lifetime reproductive output and lifespan in *Caenorhabditis elegans*. Proc Natl Acad Sci USA **94:** 3823–3827.

KIBOTA, T., and M. LYNCH, 1996 Estimate of the genomic mutation rate deleterious to overall fitness in *E. coli*. Nature **381:** 694–696.

KONDRASHOV, A. S., 1988 Deleterious mutations and the evolution of sexual reproduction. Nature **336:** 435–440.

LANDE, R., 1994 Risk of population extinction from fixation of new deleterious mutations. Evolution **48:** 1460–1469.

LONG, H., and R. ZUFALL, 2010 Diverse modes of reproduction in the marine free-living ciliate *Glauconema trihymene*. BMC Microbiol **10:** 108.

LYNCH, M., 2010 Evolution of the mutation rate. Tr Genet **26:** 345–352.

LYNCH, M., J. BLANCHARD, D. HOULE, T. KIBOTA, S. SCHULTZ *et al.*, 1999 Perspective: spontaneous deleterious mutation. Evolution **53:** 645–663.

LYNCH, M., W. SUNG, K. MORRIS, N. COFFEY, and C. R. LANDRY *et al.*, 2008 A genome-wide view of the spectrum of spontaneous mutations in yeast. Proc Natl Acad Sci USA **105:** 9272–9277.

LYNN, D. H., 2008 *The ciliated protozoa. Characterization, classification and guide to the literature*. Springer, New York.

MUKAI, T., 1964 The genetic structure of natural populations of *Drosophila melanogaster*. I. Spontaneous mutation rate of polygenes controlling viability. Genetics **50:** 1–19.

MULLER, H. J., 1927 Artificial transmutation of the gene. Science **66:** 84–87.

MULLER, H. J., 1928 The measurement of gene mutation rate in *Drosophila*, its high variability, and its dependence upon temperature. Genetics **13:** 279–357.

NANNEY, D., 1959 Vegetative mutants and clonal senility in *Tetrahymena*. J Eukaryot Microbiol **6:** 171–177.

NANNEY, D., 1974 Aging and long-term temporal regulation in ciliated protozoa. A critical review. Mech Ageing Dev **3:** 81–105.

OTTO, S. P., and T. LENORMAND, 2002 Resolving the paradox of sex and recombination. Nat Rev Genet **3:** 252–261.

SCHOEN, D. J., 2005 Deleterious mutation in related species of the plant genus *Amsinckia* with contrasting mating systems. Evolution **59:** 2370–2377.

SHAW, F. H., C. J. GEYER, and R. G. SHAW, 2002 A comprehensive model of mutations affecting fitness and inferences for *Arabidopsis thaliana*. Evolution **56:** 453–463.

SHAW, R. G., D. L. BYERS, and E. DARMO, 2000 Spontaneous mutational effects on reproductive traits of *Arabidopsis thaliana*. Genetics **155:** 369–378.

SILANDER, O. K., O. TENAILLON, and L. CHAO, 2007 Understanding the evolutionary fate of finite populations: the dynamics of mutational effects. PLoS Biol **5:** e94.

VASSILIEVA, L. L., A. M. HOOK, and M. LYNCH, 2000 The fitness effects of spontaneous mutations in *Caenorhabditis elegans*. Evolution **54:** 1234–1246.

WIENS, J. J., 2004 The role of morphological data in phylogeny reconstruction. Syst Biol **53:** 653–661.

ZEYL, C., and J. A. G. M. DEVISSER, 2001 Estimates of the rate and distribution of fitness effects of spontaneous mutation in *Saccharomyces cerevisiae*. Genetics **157:** 53–61.

# Chapter 2 Mutation accumulation in *T. thermophila*

Due to its central role in both evolutionary change and human disease, mutation has been the focus of intensive research. The probability that a spontaneous mutation will occur and its impact on fitness are parameters of central importance for making predictions in evolutionary biology; however, estimates of these parameters are difficult to make and exist only for a few taxonomically-restricted eukaryotic species. Here we introduce a new experimental system for elucidating these and other mutational parameters, the microbial eukaryote *T. thermophila*. The nuclear structure of this species allows for the accumulation of mutations of all fitness effects, including strongly deleterious and lethal mutations, which are usually eliminated by selection in similar experiments. Despite the unusual genome structure and the wide evolutionary distance from previously studied species, the estimated mutational parameters in *T. thermophila* do not differ substantially from those of other species.

## 2.1 Introduction

Mutations are the ultimate source of all variation responsible for evolutionary change. Thus, knowledge of the rates and fitness consequences of spontaneous mutations are essential to understanding evolution (CHARLESWORTH 1996). These parameters have been estimated in a handful of systems and have provided many important insights into the evolutionary process (reviewed in HALLIGAN and KEIGHTLEY 2009; LYNCH 2010; LYNCH *et al.* 1999). For example, the observation that the deleterious mutation rate is less

than 1 per genome per generation in several organisms suggests that selection against deleterious mutations is not sufficient to explain the widespread maintenance of sexual reproduction (HALLIGAN and KEIGHTLEY 2009; KONDRASHOV 1988); however, among eukaryotes, mutational parameters have only been estimated in Opisthokonta (which includes animals and fungi) and Archaeplastida (comprising red and green algae and land plants), leaving the majority of eukaryotic diversity unexamined.

Mutation accumulation (MA) is the principal experimental tool available to study mutation rates and fitness effects. Typically, MA experiments consist of allowing the accumulation of spontaneous mutations in replicate populations over many generations. Small populations are used in order to allow new mutations to drift to fixation regardless of their impact on fitness. The efficacy of an MA experiment in estimating mutational parameters is constrained by the biology of the organism used. Organisms with long generation times accumulate mutations slowly (*e.g.*, one MA study on *Arabidopsis thaliana* (SHAW *et al.* 2000), which has a generation time of ~10 weeks, only lasted 17 generations), leading to imprecise estimates; however, in organisms with short generation times, mutations with stronger deleterious effects are more likely to be eliminated by selection—especially if populations are allowed to expand between transfers (KIBOTA and LYNCH 1996)—, leading to biased estimates.

The unusual nuclear architecture and short generation time of the ciliated unicellular eukaryote *T. thermophila* (Chromalveolata) makes it particularly suitable for MA experiments. *T. thermophila*, like most ciliates, maintains two types of nuclear genomes: a transcriptionally-active somatic genome in the macronucleus, and a

sequestered, largely transcriptionally-inert germline genome in the micronucleus. Because the germline genome can remain silent throughout many rounds of asexual reproduction, mutations that occur in that genome are hidden from selection until they get expressed after conjugation; thus, all mutations, regardless of fitness effects, have an equal probability of fixing in the germline genome. Accumulation of deleterious mutations in the germline genome of lab cultures of *Tetrahymena* was first observed over 50 years ago (NANNEY 1959). Remarkably, strains with deteriorated, or even absent, germline genomes are nonetheless capable of vigorous asexual reproduction (ALLEN *et al.* 1984; NANNEY 1974).

An MA experiment was conducted to determine the mutation rate and fitness effects of mutations in the germline and somatic genomes of *T. thermophila*. Mutations in the germline genome caused a substantial decline in mean fitness and an increase in the variance in fitness among MA lines. In contrast to a previous study (Brito *et al.* 2010), no evidence was found for the accumulation of deleterious mutations in the somatic genome over 1,000 generations. The deleterious mutation rate and fitness effects of mutations in the germline genome are consistent with estimates of these parameters from MA experiments in other eukaryotes. The results of backcrossing and genomic exclusion experiments suggest that most germline mutations are mildly recessive when expressed in the somatic genome, and that lethal mutations are rare.

## 2.2 Results

### *2.2.1 The germline genome accumulated deleterious mutations*

To measure the fitness consequences of mutation accumulation in the germline genome, The maximum population growth rate ($r_{max}$) of 25 MA lines after genomic exclusion (GE) to express germline mutations at different time points during evolution was measured (Supplemental Information Figure S1; see S2 for results on lag time). After 1000 generations of MA, the mean $r_{max}$ of 19 MA lines was 40.6% lower than that of the ancestral line (95% credible interval (CI): 31.6%, 48.4%). When additional data from generations 200 and 800 were taken into account, we (all data analyses were in collaboration with RBRA & TP) estimated that mean ln($r_{max}$) declined by $\Delta M$ = –0.0509% per generation (95% CI: –0.0649%, –0.0368%; Figure 2.1A).

In parallel, the among-line variance in fitness also changed over time. The best fitting linear mixed model assumed that the among-line variance components differed between generations (Table 2.1). The among-line variance component in ln($r_{max}$) increased at a rate of $\Delta V$ = 0.00552% per generation (95% CI: 0.00227%, 0.01279%; Figure 2.1C).

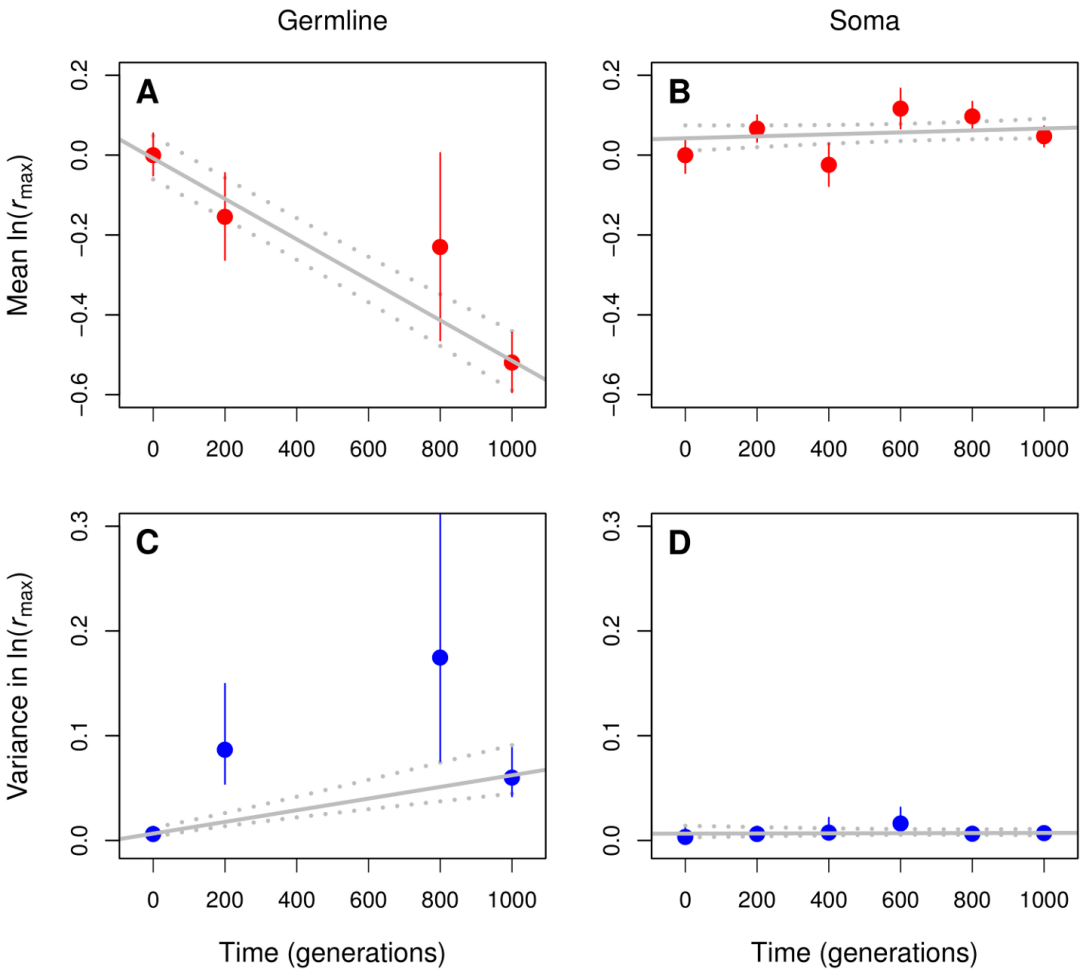


**Figure 2.1 Accumulation of spontaneous mutations in *T. thermophila* over 1000 generations.** Gradual accumulation of deleterious mutations in the germline genome causes mean fitness to decrease (A) and among-line variance in fitness to increase (C). In contrast, neither the mean nor the among-line variance in somatic fitness changes significantly over the course of the experiment (B, D). Our fitness measure is the maximum population growth rate ($r_{max}$) relative to the ancestral population. Means and among-line variance components were estimated through linear mixed models. Error bars are 68.3% credible intervals (CIs), based on the posterior distribution of the mixed model. In (C), the upper CI for $t$ = 800 generations (0.508) is truncated to improve visualization (the wide CI is caused by an outlier). Solid gray lines are weighted least-squares regression fits where the mean or among-line variance at each time was weighted by the inverse of the sampling variance calculated from the posterior distribution; the dotted lines show the 68.3% CIs on the fits calculated from the posterior distribution.

**Table 2.1. DIC (Deviance Information Criterion) analysis of among-line variance components in the MA experiment.** Three linear mixed models were fit by making different assumptions on the among-line variance components at different generations: (1) equal to zero, (2) equal to each other, and (3) different from each other. Generation was treated as a fixed effect and MA line and Plate were treated as random effects. The germline and somatic fitness data were analyzed separately. Values are the DIC corresponding to each model. Lower values of DIC indicate a better fit (SPIEGELHALTER *et al.* 2002). A difference in the DIC of five is considered substantial, and a difference of 10 rules out the model with the larger DIC (BARNETT *et al.* 2010). Values in bold indicate the simplest adequate model.

| Model | Germline | Soma |
|---|---|---|
| $V_0 = V_{200} = \ldots = V_{1000} = 0$ | 92.32 | –246.59 |
| $V_0 = V_{200} = \ldots = V_{1000} > 0$ | –78.07 | **–294.06** |
| $V_0 \neq V_{200} \neq \ldots \neq V_{1000} > 0$ | **–178.45** | –294.42 |

### *2.2.2 The somatic genome did not accumulate deleterious mutations*

To measure the fitness consequences of mutation accumulation in the somatic genome, the $r_{max}$ of 42 MA lines (without GE) was measured at different times during evolution (see Supplemental Information Figure S2 for results on lag time). Mean $\ln(r_{max})$ *increased* slightly by $\Delta M = 0.00240\%$ per generation, and the 68.3% CI overlapped with zero (–0.00117%, 0.00595%; Figure 2.1B). Similarly, although there was evidence that the among-line variance component was greater than zero, there was little support for the hypothesis that it changed over time (Table 2.1; Figure 2.1D).

### *2.2.3 Estimating the deleterious mutation rate and the average mutational effect*

The expected number of new germline deleterious mutations per MA line per generation is $2UN_e$, where $U$ is the deleterious mutation rate per haploid germline genome per

generation, and $N_e$ is the effective population size. Since the germline nucleus divides mitotically and these mutations are neutral (hidden from selection), the probability of fixation of a diploid genotype containing a new mutation in a heterozygous state is $1/N_e$; thus, the rate of deleterious germline mutation accumulation per line per generation is $2U$. After $t$ generations, the expected number of deleterious germline mutations per MA line is $k = 2Ut$. All these mutations are expected to be in a heterozygous (or heteroallelic) state.

To assay the fitness of the germline of an MA line, a GE line was constructed. If an MA line has $k$ mutations in a heterozygous state, a GE line derived from it will have a subset of, on average, $k/2$ mutations in a homozygous state; different GE lines will have different combinations of the original $k$ mutations. We (in collaboration with RBRA) assume that the fitness of a genotype containing a deleterious mutation in the homozygous state is $1 - s$ and that all mutations have equal effects; then, following the approach of Bateman (1959) and Mukai (1964), the expected decline per generation in the mean fitness of MA lines after GE relative to the ancestral line is given by $\Delta M = -Us$; the expected increase per generation in the among-line variance in fitness is given by $\Delta V = Us^2$ (Table 2.2), which lead to the Bateman-Mukai estimators for the mutational parameters in the experiments shown in Table 2.2 (MA column). Applying these estimators to the $r_{max}$ data we got a deleterious mutation rate per haploid germline genome per generation of $U = 0.00470$ (95% CI: 0.00154, 0.01249) and a deleterious effect of a mutation of $s = 0.109$ (95% CI: 0.0437, 0.272).

Individual-based simulations by TP and RBRA confirmed that these estimates were consistent with the germline fitness data (Figures 2.2A, C, S4; see Figure S3 for results on lag time). Although the MA lines were propagated by single-cell bottlenecks, in the simulations, an effective population size $N_e = 10$ was assumed to take population expansion between transfers into account, obtained by taking the harmonic mean of successive doublings from population size of 1 until the final population size is reached. This value is consistent with cell counts at transfer during the first half of the experiment (WAHL and GERRISH 2001).

Assuming $N_e > 1$ has no effect on the trends in $\Delta M$ and $\Delta V$ for germline fitness. However, it does have an effect on the accumulation of deleterious mutations in the soma: it allows purging of deleterious mutations by selection during clonal expansion. Simulations (by TP and RBRA) under $U = 0.0045$ and $s = 0.11$ lead to the following somatic trends: $\Delta M = -0.0133\%$ and $\Delta V = 0.000942\%$. Our somatic fitness data were consistent with the expectation for the among-line variance ($\Delta V = 0.0000160\%$; 95% CI: –0.00187%, 0.00127%), but not the mean ($\Delta M = 0.00240\%$; 95% CI: –0.00486%, 0.00962%) (Figure 2.2B, 2.2D). Introducing a fixed proportion of beneficial mutations in addition to the deleterious ones, was not sufficient to resolve this discrepancy (Figure 2.3). Further research would reveal whether the two nuclear genomes in *T. thermophila* have different mutational parameters.

**Table 2.2 Expected values of experimental and mutational parameters for the three types of experiments used in this study**. MA, mutation accumulation experiment followed by a single genomic exclusion per line; GE, multiple genomic exclusions per mutation accumulation line; BX, multiple genomic exclusions per mutation accumulation line followed by backcrossing to the ancestral line. From these experiments we measured the following two quantities directly: $\Delta M$, change in the mean fitness of the lines per generation; $\Delta V$, change in the among-line variance per generation. (In the MA experiments these quantities were measured across MA lines; in the GE and BX experiments they were measured among replicate genomic exclusions, within a single MA line.) From these quantities we estimated the mutational parameters: $U$, deleterious mutation rate per haploid germline genome per generation; $s$, deleterious effect of a mutation when in a homozygous state; $h$, dominance coefficient of a mutation. (The fitness of a genotype containing a mutation in the homozygous state is $1 - s$; that of a genotype containing the same mutation in the heterozygous state is $1 - hs$; $\Delta M_{BX} = 1 - M_{BX}$, $\Delta M_{GE} = 1 - M_{GE}$.) The expressions assume that all mutations have equal effects, following the approach pioneered by Bateman and Mukai (BATEMAN 1959; MUKAI 1964). The estimators for the MA experiment were validated by individual-based simulations (Figure S4).

| Parameters | MA | GE | BX |
|---|---|---|---|
| $\Delta M$ | $-Us$ | $-Us$ | $-Uhs$ |
| $\Delta V$ | $Us^2$ | $Us^2 / 2$ | $U(hs)^2 / 2$ |
| $U$ | $(\Delta M)^2 / \Delta V$ | $(\Delta M)^2 / (2\Delta V)$ | $(\Delta M)^2 / (2\Delta V)$ |
| $S$ | $-\Delta V / \Delta M$ | $-2\Delta V / \Delta M$ | $-2\Delta V / (h\Delta M)$ |
| $H$ | | $\Delta M_{BX} / \Delta M_{GE}$ | |

### *2.2.4 Dominance*

Four MA lines were taken (Figure S1) and multiple (16–27) viable GE lines were generated from each. Then each GE line was backcrossed (BX) against the ancestor and 7–12 BX lines were generated for each MA line (BX was unsuccessful for some GE lines). As explained above, a GE line derived from an MA line that has accumulated $k$ mutations for $t$ generations, is expected to have $k/2$ deleterious mutations in a homozygous state. A BX line derived from this MA line is expected to have the same

deleterious mutations but in a *heterozygous* state. Therefore, the mean and variance in fitness across the GE or BX lines generated from a given MA line can also be used to estimate mutational parameters (Table 2.2, GE and BX columns). Furthermore, by combining the data from both types of lines, we (through collaboration with RBRA & TP) can estimate the dominance coefficient of a mutation ($h$), such that the fitness of a genotype containing a mutation in the heterozygous state is $1 - hs$.

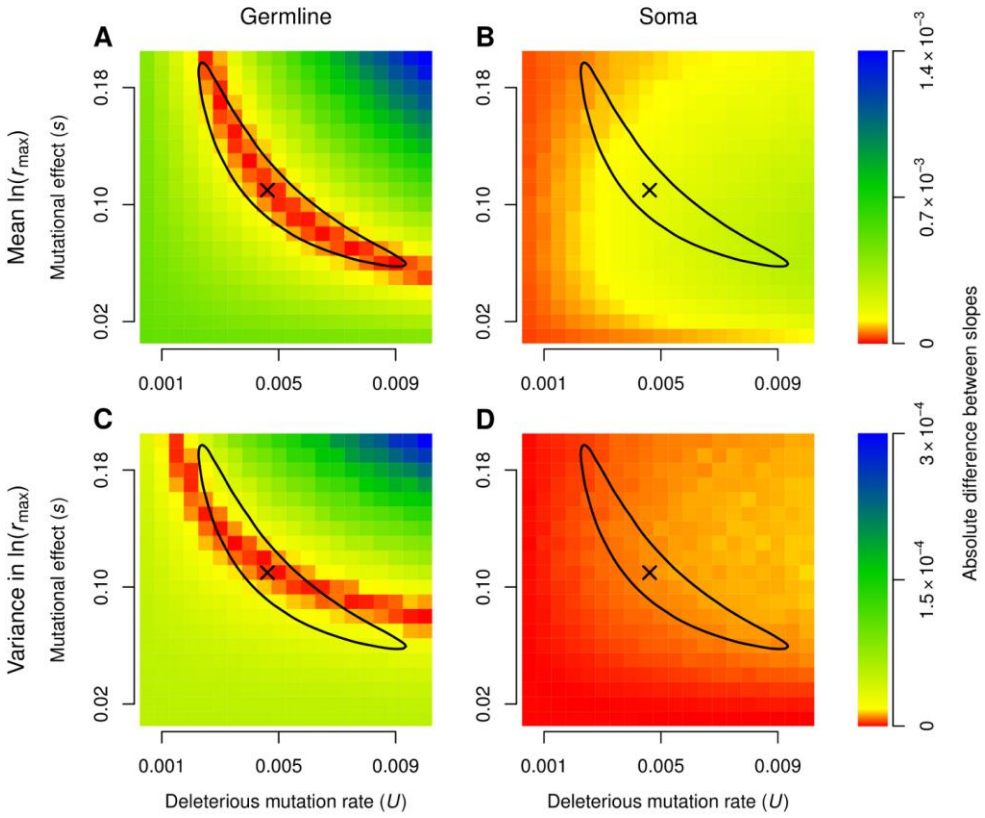


**Figure 2.2 Comparison between experimental and simulated data.** The slopes of the regressions shown in Figure 2.1 are compared to those obtained from individual-based simulations of mutation accumulation under a broad range of values of the deleterious mutation rate per haploid germline genome per generation ($U$) and the deleterious effect of a mutation when in a homozygous state ($s$). Each simulated MA line was allowed to evolve for 1000 generations with a constant population size of $N = 10$ (see Figure S4 for more details). Colors indicate absolute differences between the slopes, such that low values (red) denote better fits and higher values (blue) worse fits. The cross marks the estimate of the mutational parameters based on the estimators in Table 2.2 (MA); the black line marks the 68.3% credible region based on the posterior distribution of the mixed model described in Figure 2.1.

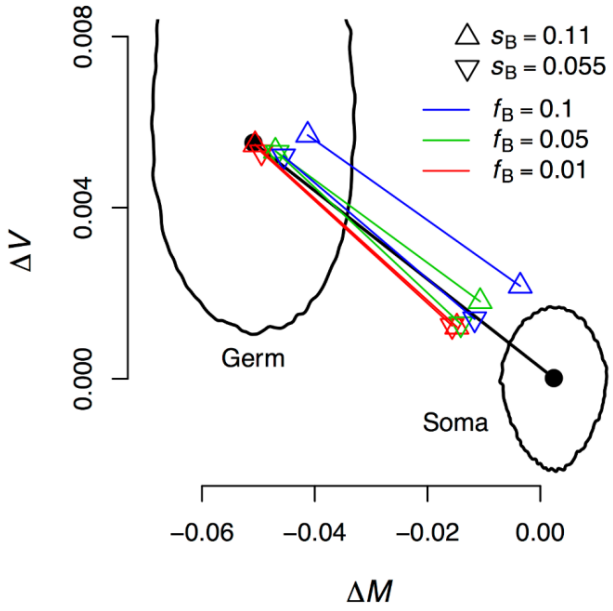


**Figure 2.3 Adding beneficial mutations does not resolve the discrepancy between the germline and somatic responses**. The solid circles and black lines mark the median and 95% credible regions for the responses in the germline (left) and somatic (right) mean ($\Delta M$) and variance ($\Delta V$) in fitness estimated from the MA experiment (the confidence region for the germline is truncated to improve visualization). The open triangles show values of $\Delta M$ and $\Delta V$ for germline and soma from simulations assuming that beneficial mutations occur with different proportions ($f_B$) and effects ($s_B$). Simulations were conducted as described in Figure S4. In all simulations, we assumed that deleterious mutations occurred at rate $U = 0.0045$ and had an effect of $s = 0.11$. The proportion of beneficial mutations is defined as: $f_B = U_B / (U_B + U)$, where $U_B$ is the beneficial mutation rate. Increasing both $f_B$ and $s_B$ substantially fails to generate somatic $\Delta M$ and $\Delta V$ consistent with those observed in our experiment.

Applying these estimators to the GE data we (in collaboration with RBRA) got a deleterious mutation rate per haploid germline genome per generation of $U = 0.0205$ (95% CI: 0.00583, 0.101) and a deleterious effect of a mutation of $s = 0.0198$ (95% CI: 0.00423, 0.0558); similarly, we got the following estimates from the BX data $U = 0.00155$ (95% CI: 0.0000359, 0.0134), $h = 0.257$ (95% CI: –0.0221, 0.622), and $s = 0.258$ (95% CI: 0.0277, 1). Differences among MA lines were small (Table 3) so they were ignored. The estimate of the dominance coefficient suggests that most of the mutations in these lines

are incompletely recessive (Figure 2.4). The lower credibility of the BX estimates result from the fact that they depend on *h*, which is itself estimated with uncertainty.

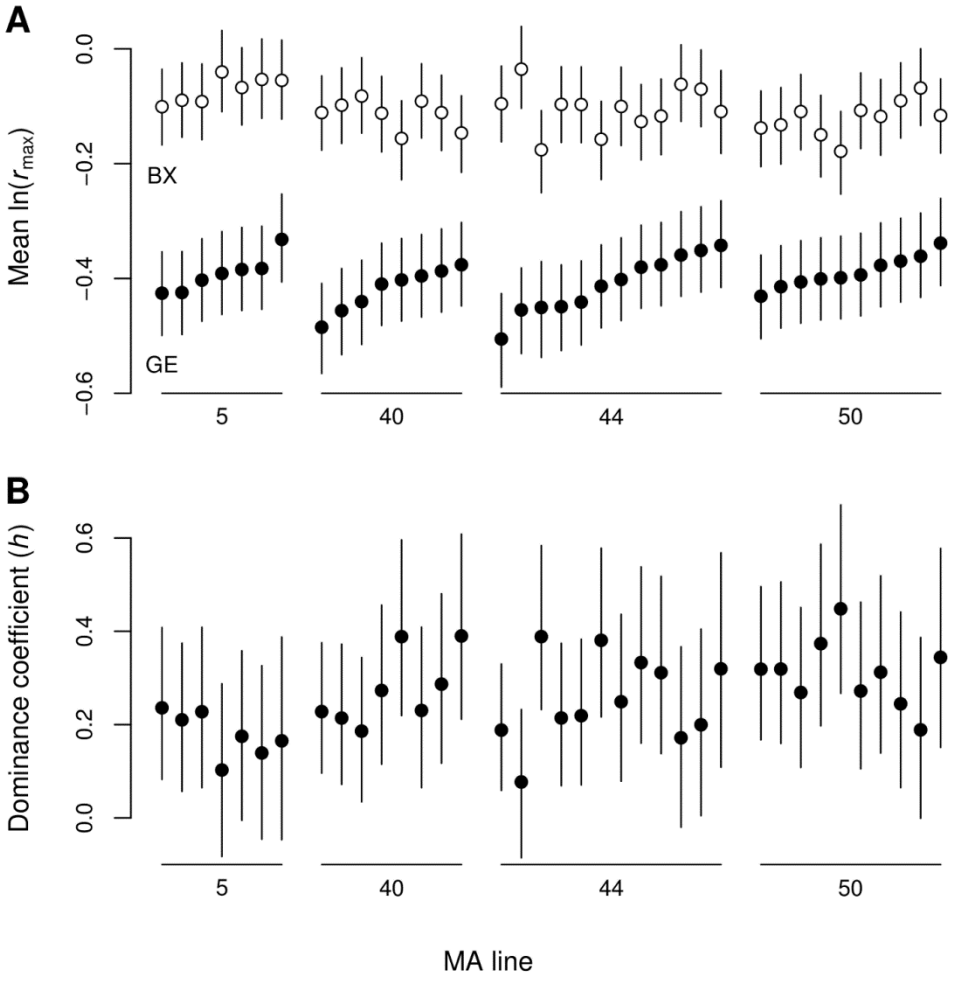


**Figure 2.4 Most mutations are incompletely recessive.** (A) Results of multiple genomic exclusions per mutation accumulation line (GE) followed by backcrossing to the ancestral line (BX). Line means were estimated from a linear mixed model; error bars are 68.3% CIs. (B) Estimates of the dominance coefficient based on the data in A (see Table 2.2).

### *2.2.5 Mutations with effects on viability*

The previous estimates assume that a GE generates an unbiased sample of the mutations present in an MA line; however, this will not be the case if some of those mutations are lethal. If an MA line contains $k_L$ unlinked lethal mutations, then a GE line derived from it will only be free of lethal mutations with probability $1/2^{k_L}$. To test this possibility, multiple, independent GEs were conducted and the viability of each GE germline genome was estimated.

**Table 2.3 DIC analysis of variation among MA lines in the GE and BX experiments.** Three linear mixed models were fit making different assumptions about the fixed effect MA line (4, 40, 44 or 50, see Figure 2.4). Treatment (GE or BX) was also treated as a fixed effect and Replicate line and Plate were treated as random effects. Values are the Deviance Information Criterion (DIC) corresponding to each model. Values in bold indicate the simplest adequate model. See legend of Table 2.1 for more details.

| Model | DIC |
|---|---|
| Treatment | **–246.97** |
| Treatment + MA line | –248.64 |
| Treatment + MA line + Treatment:MA line | –248.47 |

During the first round of GE, cells pair and the MA strain donates a haploid pronucleus to the "star" strain, but a new somatic genome does not develop. During the second round of GE, germline mutations become expressed in the soma. For each of the four MA lines, a large number of independent first round GE crosses (144-288 mating pairs were isolated per MA line) were conducted. 91–181 independent surviving first round GE pairs per MA line were obtained, each containing a different germline genome. 16 to 30 first round independent GE mating pair cultures later succeeded in the second

round GE crossing. Then each independent mating pair culture was used to set up 48 independent second round GE crosses. Each of these 48 replicates contains identical germline nuclei, so if none of them survived, it would be strong evidence that the GE had "picked up" a lethal mutation.

Only 3 out of 84 independent GE from the four MA lines displayed 0/48 = 0% viability in GE round II, all in the same MA line (44), which had 30 independent GE trials. If that MA line contained a lethal mutation and only that mutation could influence GE line viability; then the probability that only 3/30 or fewer independent GE lines show 0% viability is $P = 4.2 \times 10^{-6}$, suggesting that this MA line does not carry any lethal mutations; the equivalent probabilities for each of the other three MA lines are all $P < 1.5 \times 10^{-5}$. A lethal mutation rate of $U_L = 0.000374$ per haploid genome per generation would imply a probability of $P = 0.05$ of not finding a lethal mutation in four MA lines after $t = 1000$ generations.

Although the four MA lines do not show any evidence for lethal mutations, their average viability is 0.204 (95% CI: 0.174, 0.238), whereas that of the ancestor is 0.845 (95% CI: 0.775, 0.897). These results suggest that the MA lines are carrying mutations that reduce GE line viability. In collaboration with RBRA, we estimated the number of mutations $k$ per line and their deleterious effect $s$ using approximate Bayesian computation (ABC). We estimated that the original MA lines contain $k \approx 6$–12 mutations with effects $s \approx 0.2$–0.6 (Figure 2.5). Based on the estimate of $U$ in the MA experiment, we would have expected to find approximately the same number of mutations, but of

smaller effect (Figure 2.5). These results suggest that the estimates of mutational parameters based on "viable" GE lines could be biased towards smaller mutational effects, although the discrepancy might also reflect differences between the two traits.

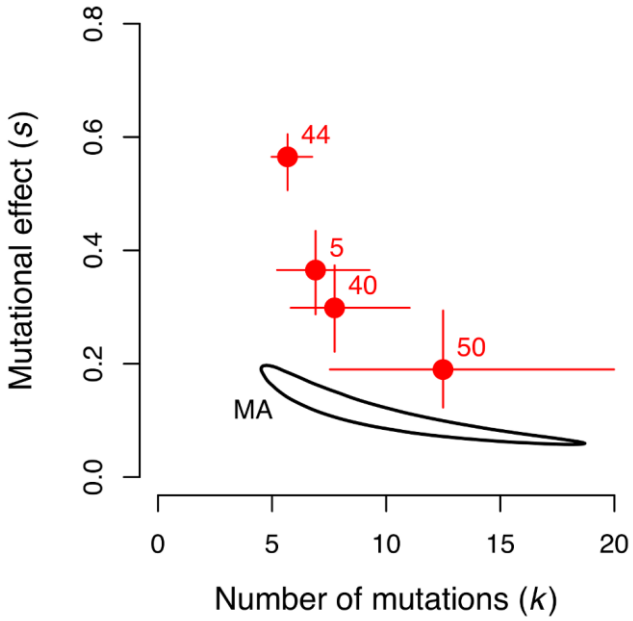


**Figure 2.5 Number of mutations and their effects on viability in four MA lines.** Solid circles are estimates and 68.3% CIs from an ABC analysis of the GE viability data. Numbers indicate the MA line used. The black line marks the 68.3% credible region shown in Figure 2.2 (obtained by setting $k = 2Ut$).

## 2.3 Discussion

Despite the unusual genetic features that make *T. thermophila* a useful system in which to study spontaneous mutation, the deleterious mutation rate ($U = 0.0047$ per haploid germline genome per generation) and average fitness effect ($s = 0.11$) do not differ substantially from those estimated for other eukaryotes (Figure 2.6; HALLIGAN and KEIGHTLEY 2009). For example, $U$ and $s$ are similar between *T. thermophila* and *C. elegans* ($U = 0.0053$, $s = 0.16$, median of estimates based on six fitness metrics

VASSILIEVA *et al.* 2000). The haploid genome sizes (~104 Mb and ~100 Mb for *T. thermophila* and *C. elegans* respectively) and number of genes (~27,000 and ~20,000) in these two species are also remarkably similar (EISEN *et al.* 2006; HILLER *et al.* 2005). Likewise, our estimates of dominance coefficients of new mutations ($h = 0.257$) are consistent with previous estimates from other species (HALLIGAN and KEIGHTLEY 2009), *e.g.*, $h = 0.12$, average of five traits, in *D. melanogaster* (HOULE *et al.* 1997). However, there is also large variation across species and experiments in these estimates (HALLIGAN and KEIGHTLEY 2009).

The estimates of mutational parameters could be biased if GE progeny were lost due to lethal mutations; however, the observation that four MA lines failed to accumulate any germline lethal mutations indicates that lethal mutations are unlikely to bias the results. It also suggests that the lethal mutation rate in *T. thermophila* is less than 10% of the deleterious mutation rate. Previous studies in yeast estimate the lethal mutation rate to be ~0.00007 (HALL and JOSEPH 2010) and 0.00031 (WLOCH *et al.* 2001), which make up 12–30% of total mutations with fitness effects. There are few estimates of lethal mutation rate because in most experiments, unlike in *T. thermophila*, mutations are exposed to selection throughout the course of MA (but see SIMMONS and CROW 1977). The tests for lethal mutations also allowed us to consider an additional fitness component: viability. Estimates of mutation rate and fitness effects based on viability correspond reasonably well with those based on growth rate (Figure 2.5).

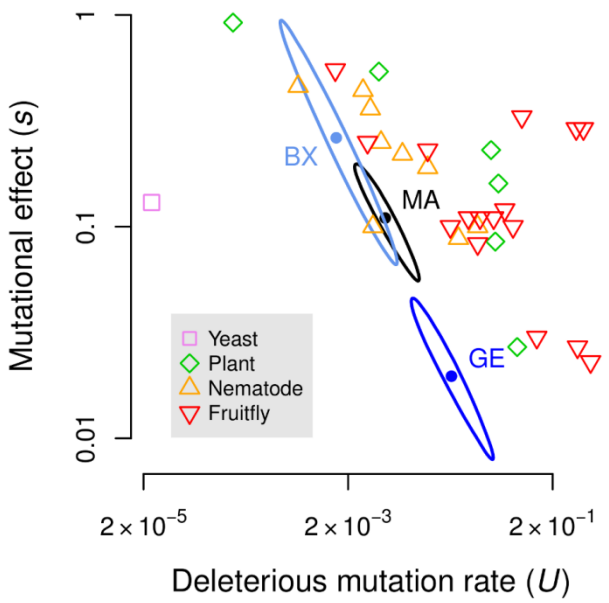


**Figure 2.6 Mutational parameters of *Tetrahymena thermophila* in the context of those of other eukaryotes.** Solid circles and elliptical shapes indicate the estimates and 68.3% credible regions, respectively, based on the MA, GE and BX experiments (see Table 2.2 for details). The data from other eukaryotes were taken from Table 1 of (HALLIGAN 2009). Note that both axes are log-transformed.

In addition to examining different fitness components, we (in collaboration with RBRA) used multiple approaches to estimate mutational parameters. The resulting estimates were not always entirely consistent with each other (Figure 2.6), highlighting the need for statistical methods that can combine the different types of experiments (Table 2.2) to generate more robust estimates. One possible reason for the variation among estimates is that the assumption that different mutations do not interact epistatically may not be met in this system. Crosses between BX lines derived from different MA lines could be used to test for directional epistasis among mutations (DEVISSER *et al.* 1997; DEVISSER *et al.* 1996; WEST *et al.* 1998).

In an earlier MA study in *T. thermophila*, Brito *et al.* (BRITO *et al.* 2010) used the rate of clonal extinction as a measure of fitness. They found that ~1.25 MA lines went extinct per bottleneck and interpreted this as evidence for the rapid accumulation of deleterious mutations in the somatic genome. They speculated that their lines may have experienced gains and/or losses of chromosomes during amitosis, a phenomenon that, although common in ciliates (BLACKBURN and KARRER 1986), has never been observed in *T. thermophila* as a result of a copy number control mechanism (LARSON *et al.* 1991). In contrast, no evidence was found that MA lines accumulated any deleterious somatic mutations over the course of 1,000 generations. A high rate of clonal extinction, or rather transfer failure, was observed during our experiment, but it did not change throughout the experiment. The use of multiple single-cell transfer cultures at each bottleneck event completely suppressed line loss during the experiment. Transfer failure might be caused by the dependency of *T. thermophila* survival, growth and proliferation on cell-cell interactions (reviewed in CHRISTENSEN *et al.* 1995).

The comparison of mutational parameters from somatic and germline genomes suggests these genomes may experience different rates and average fitness effects of mutations. This result is consistent with previous estimates of germline and somatic mutation parameters from multicellular organisms (reviewed in LYNCH 2010). While it is unclear what results in these different rates in multicellular organisms, it is possible that in *T. thermophila* the different parameters may reflect different polymerase or DNA repair machinery operating in the different genomes. Resolving the molecular basis of mutation in the two genomes may help answer this question.

## 2.4 Materials and Methods

*Strains and media*

*T. thermophila* strains SB210 and B*VII were acquired from the *Tetrahymena* Stock Center (Cornell University). SB210, the strain that was used for MA, is highly inbred and a heterokaryon (EISEN *et al.* 2006). B*VII has a dysfunctional germline nucleus and carries mating type VII. SSP medium was used during MA: 2% proteose peptone (EMD Chemicals, New York, USA), 0.2% glucose, 0.1% yeast extract (BD, Spark, Maryland, USA), and 0.003% Fe-EDTA (Acros Organics, New Jersey, USA) (GOROVSKY *et al.* 1975). Tris buffer (10mM Tris-HCl pH 7.5) was used to starve cells in preparation for conjugation (BRUNS and BRUSSARD 1974). 2% proteose peptone was used for mating pair re-feeding during conjugation (BRUNS and CASSIDY-HANLEY 1999).

*Mutation accumulation*

Fifty mutation accumulation (MA) lines were established from single-cell isolates of strain SB210. Each line was grown to log-phase in 3ml of fresh SSP medium in a borosilicate glass test tube (Corning, New York, USA) on a shaker at 200rpm and 30˚C. When cultures reached log phase, usually 3–4 days from the last transfer, MA lines were bottlenecked by single cell transfer. 4–8 replicates were prepared for each line in order to prevent the accidental loss of MA lines. The transfer success rate for each MA line was defined as the ratio of successfully established replicate cultures out of the first four replicates. The average transfer success rate was 51% (s.d. = 29%) and did not change

significantly over the course of the experiment (one-way ANOVA for generations 0, 500 and 1,000: $F_{2,144} = 0.89$, $P = 0.4$).

Prior to each transfer, the optical density at 650nm ($OD_{650}$) of the log phase culture was measured on a Versamax™ microplate reader (Molecular Devices, Sunnyvale, CA, USA). The $OD_{650}$ was used to calculate the number of cell divisions between two transfers using a standard curve plotted based on five cell densities (1X, 3X, 9X, 27X, and 81X; 3 replicates each) of the ancestor SB210 culture. Cell size was monitored to assure consistency of these measurements (Supplementary Text). When an MA line took longer than seven days to reach a measureable $OD_{650}$ (0.01 in our system), cell density was determined by C-Chip haemocytometer (Incyto, SKC, Korea). Single-cell transfers were repeated until each line had undergone ~1,000 asexual generations. This took 256 days for the slowest growing lines.

Rarely, none of the replicate single-cell transfers survived. In that case, the MA line was reinitiated using a starvation backup culture. The backups were created by adding 200µl of the log-phase culture to 800µl Tris buffer and centrifuging (1100g for 3 min). The supernatant was removed and replaced with 500µl Tris-HCl pH 7.5; this culture was kept at 30˚C and 200 rpm. To reinitiate a line, starved cells were washed four times in 100µl drops of Tris supplemented with 1X Penicillin, Streptomycin, and Amphotericin B (PSA; final concentration; Amresco, Solon, OH, USA). Single cells were then transferred to the growth medium in 16 replicates. As a result of these procedures only one cell line was lost due to fungal contamination.

Cells from each MA line were cryopreserved in liquid $N_2$ (BRUNS *et al.* 1999) every 200 generations. 3ml of log-phase cultures were centrifuged (1100g, 3 min), the supernatant removed and 100µl of the remaining cells were transferred to 500µl Tris buffer, with at least two replicates. These starvation cultures were incubated at 30˚C and 200rpm. After 2 days, cells were centrifuged and supernatant removed, as above, and moved to cryogenic vials with DMSO (Sigma-Aldrich; final concentration 8%). Cells were then frozen using "Mr. Frosty" –1˚C cryogenic freezing containers (Nalgene) and stored in liquid $N_2$.

*Somatic fitness assays*

After completion of 1,000 generations of mutation accumulation, frozen MA lines were thawed in a 42˚C water bath and transferred to pre-warmed SSP medium, following (BRUNS *et al.* 1999). Some cultures could not be recovered from frozen stocks, due to cells either having entered the late-log phase or being at too low density when frozen. Successfully thawed cultures from all time points were analyzed together. Fitness changes associated with mutations that accumulated in the somatic genome were assayed directly from the thawed cultures. 10 cells from each culture were counted and inoculated into 180µl SSP supplemented with 1X PSA in a well on a 96 well plate. After 14 hours of pre-incubation at 30˚C, the cultures were loaded onto a microplate reader at 30˚C where $OD_{650}$ was read every 5 minutes for 60 hours (Figure S5). Location of cultures on a plate was randomized in each of 3–5 of replicates in different blocks. Maximum population growth rate ($r_{max}$) and lag time were calculated using the methods modified from (WANG *et al.* 2012). The fitness effects of mutations were determined in the somatic genome in 4

lines from the initial culture, 16 lines after 200 generations of MA, 6 lines after 400 generations, 10 after 600 generations, 13 after 800 generations, and 38 after 1000 generations.

*Germline fitness assays*

In order to measure the fitness effects of mutations accumulated in the germline genome over the course of MA, those mutations were expressed by allowing the cells to undergo conjugation and development. Conjugation was performed using genomic exclusion (GE) cross procedures (BRUNS and CASSIDY-HANLEY 1999) (Fig. 2.7). GE involves conjugation with the mating partner B*VII with a dysfunctional germline nucleus and, after two rounds of crosses, results in genetically identical progeny that are homozygous in both somatic and germline genomes for ~1/2 of all mutations that have accumulated in the germline during MA. Mitochondria, however, are parentally inherited, so half of the progeny contain mitochondria from B*VII and half from the MA lines. These two mitochondrial types were not distinguished in the experiments; thus, fitness assays contained approximately one half of each.

In round I crosses of GE, cells from MA lines and B*VII were cultured in 5.5ml SSP until they reached a density of $\sim 2 \times 10^5$ / ml at which time they were centrifuged and starved for 1–3 days (24–36 hours' starvation time was optimal for matings) in Tris buffer. 1ml starved MA culture was mixed with 1ml starved B*VII and incubated in a well on a 6-well plate in a wet chamber at 30˚C. After 6–8 hours, mating pairs were re-fed with 2% proteose peptone (final concentration 1%). 2 hours after re-feeding, 16–48

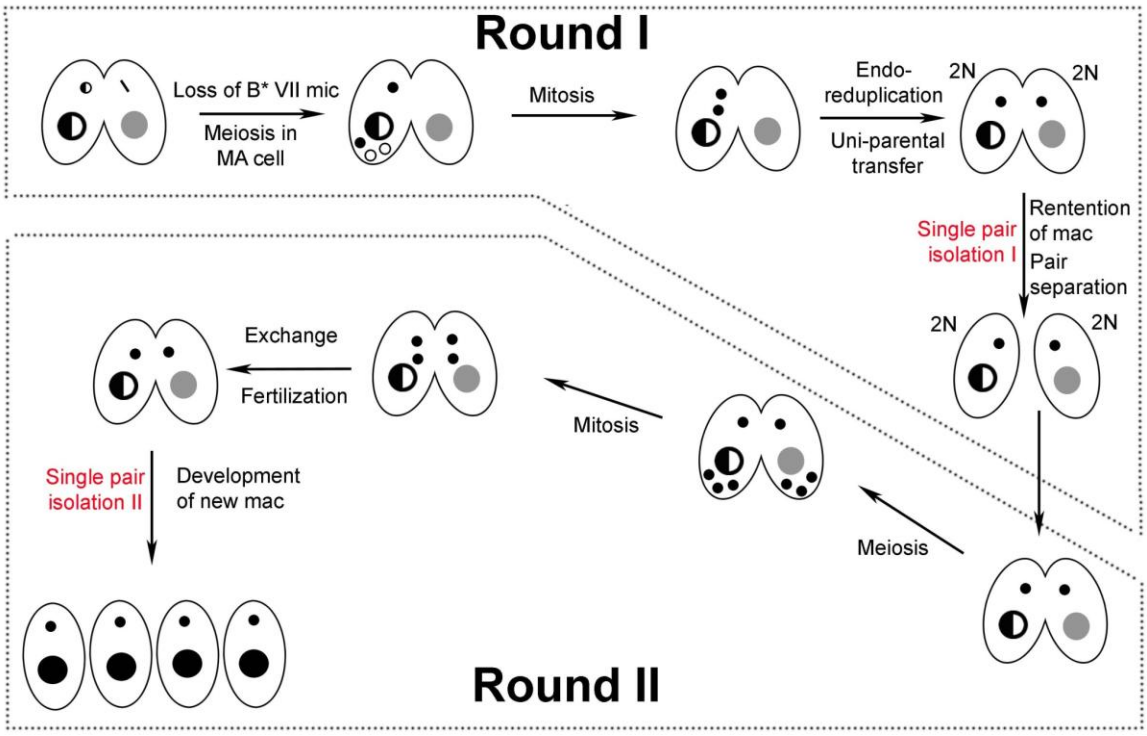


**Figure 2.7 Nuclear events during the two rounds of mating in genomic exclusion.** Modified from (BRUNS and CASSIDY-HANLEY 1999)

(for round I) single mating pairs were isolated and inoculated into 800µl SSP in wells of a 48-well plate. All single pair isolation was carried out using a 0.1–2µl pipette. Round I offspring viability was defined as the fraction of mating pairs producing viable progeny after 72 hours. Viability was additionally confirmed after one week. After round I, the germline nucleus of progeny is the homozygous product of the diploidization of the pronucleus from the MA line; the somatic nucleus is retained from the parent.

The round II cross of GE was performed by culturing and starvation (as above) of the successful round I cultures. 2ml of starved cultures were allowed to mate and 24–96 single mating pairs were isolated as above. Round II offspring viability was determined

as for round I. Progeny cells from round II were tested with 2-Deoxy-D-galactose (2-dgal; Sigma-Aldrich; COLE and BRUNS 1992). The allele conferring resistance to 2-dgal is carried only in the germline genome of SB210, in homozygous state; thus, only cells that have successfully undergone both rounds of crosses will be resistant to the drug. Drug effects were checked 96 hours after 2-dgal was added to the culture. Progeny cells with resistance to 2-dgal were then used in fitness assays as described in the previous section. Fitness effects of germline mutations were assessed for one GE line per MA line (except for the ancestral lines, also see next section). GE crosses were repeated independently at least twice in the cases where there was no mating, no viable progeny, or no progeny with drug resistance. GE crosses were successful for 25 MA lines at one or more of generations $t$ = 200, 800 or 1000 (Figure S1).

*Multiple genomic exclusions and backcrosses*

Multiple independent GE crosses were conducted on four lines at generation $t$ = 1000 (Figure S1, "*" labeled lines), following the above procedures, with 48 mating pairs isolated in each round of the GE crossings. Viability of mating pairs in round I and round II was also recorded.

Fitness of the multiple GE lines was determined by assaying progeny cells from the successful multiple GE crossings as above, with only one cell inoculated into one well of a 96-well plate, instead of 10. These multiple GE lines were then backcrossed (BX) with an ancestral (control) GE line (4-0-AI1). The fitness of 4-0-AI1 was also assayed in each assay of GE lines before and after BX. 7–12 GE lines resulted in

successful BX for each of the four MA lines at generation 1000. Fitness after BX was determined as before BX, except that mating pairs were inoculated into the wells on a 96-well plate and then incubated less than 48 hours at 30˚C to avoid phenotypic assortment, which would cause loss of heterozygosity during amitosis of the somatic nucleus (DOERDER *et al.* 1992).

*Statistical analyses*

**Mutation accumulation.** The fitness was defined as the maximum population growth rate ($r_{max}$) divided by that of the ancestral strain. Values were natural log-transformed to increase normality and homogeneity of variances. We (in collaboration with RBRA) used a Bayesian approach, fitting linear mixed models separately for the germline and somatic fitness data using Markov Chain Monte Carlo with the MCMCglmm 2.15 package (HADFIELD 2010) in R 2.13.0 (R DEVELOPMENT CORE TEAM 2011). Generation was treated as a fixed effect and MA line and Plate were treated as random effects. Separate among-MA line variance components were fit for each Generation using the 'idh' variance function. The default uninformative proper priors was used—normally distributed with mean $\mu = 0$ and large variance $\sigma^2 = 10^8$ for the fixed effect, and inverse Wishart distributed with parameters $V = 1$ and $\nu = 0.002$ for the random effects. We allowed the Markov chain to run for a burn-in period of $10^6$ iterations after which we ran $1.1 \times 10^7$ iterations and sampled from the posterior distribution every 50 iterations, resulting in $2 \times 10^5$ stored values. The autocorrelation function between consecutive parameter values of the Markov chain at successive iterations indicates that sampling

every ~20 iterations is sufficient to achieve independence. Applying the method of (GELMAN and RUBIN 1992) (implemented in R through the coda 0.14-4 package PLUMMER *et al.* 2006) to three parallel Markov chains suggests that convergence was achieved within ~$2 \times 10^5$ iterations.

The linear mixed model directly estimates the among-line variance component, $V_t$, at each generation $t$. We (in collaboration with RBRA) used the estimates of the fixed effects to calculate the mean fitness of the lines, $M_t$, at each generation. To estimate the change in mean fitness ($\Delta M$) and the change in the among-line variance in fitness ($\Delta V$) we used weighted least-squares regression of $M_t$ and $V_t$, respectively, against $t$, where each value is weighted by the inverse of its sampling variance calculated from the posterior distribution. To estimate the germline haploid deleterious mutation rate ($U$) and the deleterious effect of a mutation ($s$) we used the formulas listed under MA in Table 2.2 for the germline fitness data. Unless otherwise stated, we reported the posterior median and 68.3% credible intervals (corresponding to ±1 standard deviation for a normal distribution) for $M_t$, $V_t$, $\Delta M$, $\Delta V$, $U$, and $s$.

We (in collaboration with RBRA) evaluated the estimates of $M_t$, $\Delta M$, and $\Delta V$, by the extent to which their posterior distribution overlaps with zero. This approach is not appropriate for the estimates of $V_t$ because their prior specification implies that they must be positive; instead we compared different models using the Deviance Information Criterion, DIC (BARNETT *et al.* 2010; SPIEGELHALTER *et al.* 2002; WILSON *et al.* 2010).

**Multiple genomic exclusions and backcrosses.** A linear mixed model to the GE and BX experiments was fitted as described above for the MA experiment. Treatment (GE or BX) and MA line were treated as crossed fixed effects and Replicate line and Plate were treated as random effects. Separate among-replicate line variance components were fit for each treatment using the 'idh' variance function. The same priors, number of iterations, sampling frequency, and burn-in period were used as described for the MA experiment. The convergence and autocorrelation patterns for the Markov chains were similar to those described for the MA experiment.

The linear mixed model directly estimates the among-line variance components, $V_{1000}$, at generation $t = 1000$. We (in collaboration with RBRA & TP) used the estimates of the fixed effects to calculate the mean fitness of the lines $M_{1000}$ in each treatment. In each treatment, we estimated the change in mean fitness and the change in the among-line variance in fitness as $\Delta M = (M_{1000} - M_0) / 1000$ and $\Delta V = V_{1000} / 1000$, respectively. To estimate the mutational parameters $U$, $s$ and $h$ we used the formulas listed under GE and BX in Table 2.2. Unless otherwise stated, we reported the posterior median and 68.3% credible intervals for $U$, $s$ and $h$.

**Viability.** We (in collaboration with RBRA & TP) estimated mutational parameters for the viability data using Approximate Bayesian Computation (ABC) (BEAUMONT *et al.* 2002). The ABC modeling process involved simulating our viability data under a simple model of GE. We assume that an MA line carries $k$ mutations that are capable of influencing viability in a heterozygous state. GE samples each of these mutations with probability $P = 0.5$ and makes them homozygous. The viability of a GE genotype with

$k_{GE}$ mutations is $w = w_0(1–s)^{k_{GE}}$, where $s$ is the effect of a mutation and $w_0$ is the viability of a genotype with no mutations. We assume that mutations act multiplicatively, which is appropriate if $s$ can be large. We then simulate an experiment involving $n$ second round GE crosses as a binomial process with probability of "success" (*i.e.*, survival) $w$.

For each MA line tested ($t$ = 0 ancestor, and $t$ = 1000 MA lines 5, 40, 44 and 50) $4 \times 10^6$ simulated data sets were generated by RBRA and TP involving $N$ independent GEs (*e.g.*, $n$ = 48 and $N$ = 20 for MA line 40). Values for the number of mutations were drawn at random from a uniform prior of $k \sim [0, 30]$, and those for the mutational effect were drawn at random from a uniform prior of $s \sim [0, 1]$, independently from $k$. Values for the logit of baseline viability, $\ln[w_0 / (1 – w_0)]$, were drawn from a normal distribution with mean 1.6977 (corresponding to $w_0 = 0.845$) and s.d. = 0.2364, estimated from the GE viability data for the ancestor using a generalized linear model with logit link and quasibinomial errors.

RBRA and TP compared the results from the simulated data to the empirical data using the following 4 summary statistics on the proportion of viable second round GE crosses: mean, standard deviation, skewness and kurtosis. From a set of $4 \times 10^6$ simulations, they retained the $10^4$ simulations (*i.e.*, 0.25% of the pooled set) with the lowest Euclidian distance to the summary statistics using the R package abc 1.4 (CSILLÉRY *et al.* 2012). To estimate the posterior distribution of the parameters $k$ and $s$ from the accepted simulations, they corrected for the discrepancy between the accepted and the observed summary statistics using the local linear regression method with log-transformed parameters (BEAUMONT *et al.* 2002). Unless otherwise stated, the posterior

median and 68.3% credible intervals of $k$ and $s$ were reported. All the estimates using posterior predictive checks were validated.

*Individual-based simulations*

This part was done by collaborators TP and RBRA. Individuals consist of a diploid germline genome and a 45-ploid somatic genome each with $L = 100$ loci and reproduce asexually. Each allele at each locus in both genomes can irreversibly mutate to a deleterious form. Each somatic locus $i$ contributes equally and additively to fitness, in proportion to the number of mutated copies at that locus $k_i$. Thus, the fitness of a genotype is given by:

$$w = 1 - \sum_{i}^{L} \frac{k_i}{45} s$$

where $s$ is the deleterious effect of a mutation present in $k = 45$ copies. The germline genome does not contribute to fitness. There is concern that the assumption of additive mutational effects in this simulation might not be perfectly met in the mutation accumulation process if there is some dominance between mutant and wild-type alleles. That is to say, with respect to one locus, one individual might have high fitness even if half of its macronuclear alleles are mutants (mutants' effects are masked by the other wild-type alleles). In such a case, a much lower rate of fitness decline than the model used here might be expected. TP and RBRA are now including the dominance effects into new versions of simulations, and the results would be shown in one coming article based on this chapter.

Asexual reproduction proceeds by generating an individual with a germline genome identical to its progenitor (mitosis); the somatic genome of a daughter cell is obtained by duplicating and sampling without replacement each of the progenitor's somatic loci until 45-ploidy is reached (amitosis); thus, mutations at intermediate copy number ($0 < k < 45$) in the somatic genome will increase or decrease stochastically in copy number within individuals from generation to generation, until they ultimately fix or disappear, as occurs in phenotypic assortment (DOERDER *et al.* 1992). This is not true in the germline genome, where all the descendants of an individual carrying a germline mutation will also contain that mutation. At reproduction, each allele at each locus in each genome mutates at a rate $\mu$; the germline haploid deleterious mutation rate is $U = L\mu$.

Evolution follows a Wright-Fisher scheme (EWENS 2004) of reproduction-mutation-selection of a population of constant size $N$. In order to advance the population one generation, individuals are sampled randomly, copied, mutated, and allowed to "live" with a probability proportional to their somatic fitness until the population size $N$ is reached. In the simulations reported here we set $N = 10$, the estimated effective population size in the MA experiment (WAHL and GERRISH 2001).

In order to "assay" the germline fitness of an individual, a single GE for each MA line was simulated. A homozygous version of the individual was generated by randomly picking one of the copies at every locus and copying it to the other locus. This fully homozygous germline genome was then allowed to generate a new 45-ploid somatic genome and fitness was assayed as described above.

## 2.5 Supplementary Information

*Supplementary Text*

**Cell size did not change throughout the transfers** To ensure that cell size differences did not bias our estimates of number of cell divisions between transfers, we measured the cell sizes of 4 MA lines at generations 0 and 1,000. Micrographs were taken of cells from each culture in late log-phase. Length ($L$) and width ($W$) of 100 cells per culture were measured using calibrated scale bars. To calculate the volume of a cell we assume that it is a prolate spheroid (HELLUNG-LARSEN and ANDERSON 1989): $V = (\pi/6) \cdot L \cdot W^2$. Cell size did not change significantly between the beginning and end of the experiment in the 4 MA lines (Welch's two sample $t$-test on line means: $t$=0.31, df=5.7, $p$=0.76); thus, the use of $OD_{650}$ to determine number of generations elapsed between transfers was consistent throughout the course of the experiment.

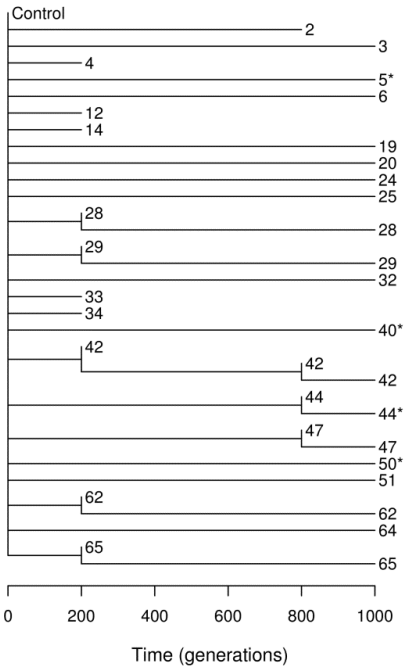


**Figure S1 Evolutionary history of the MA lines assayed for germline fitness.** Lines marked with an asterisk were also assayed in the GE and BX experiments (see Table 2.2 for details).

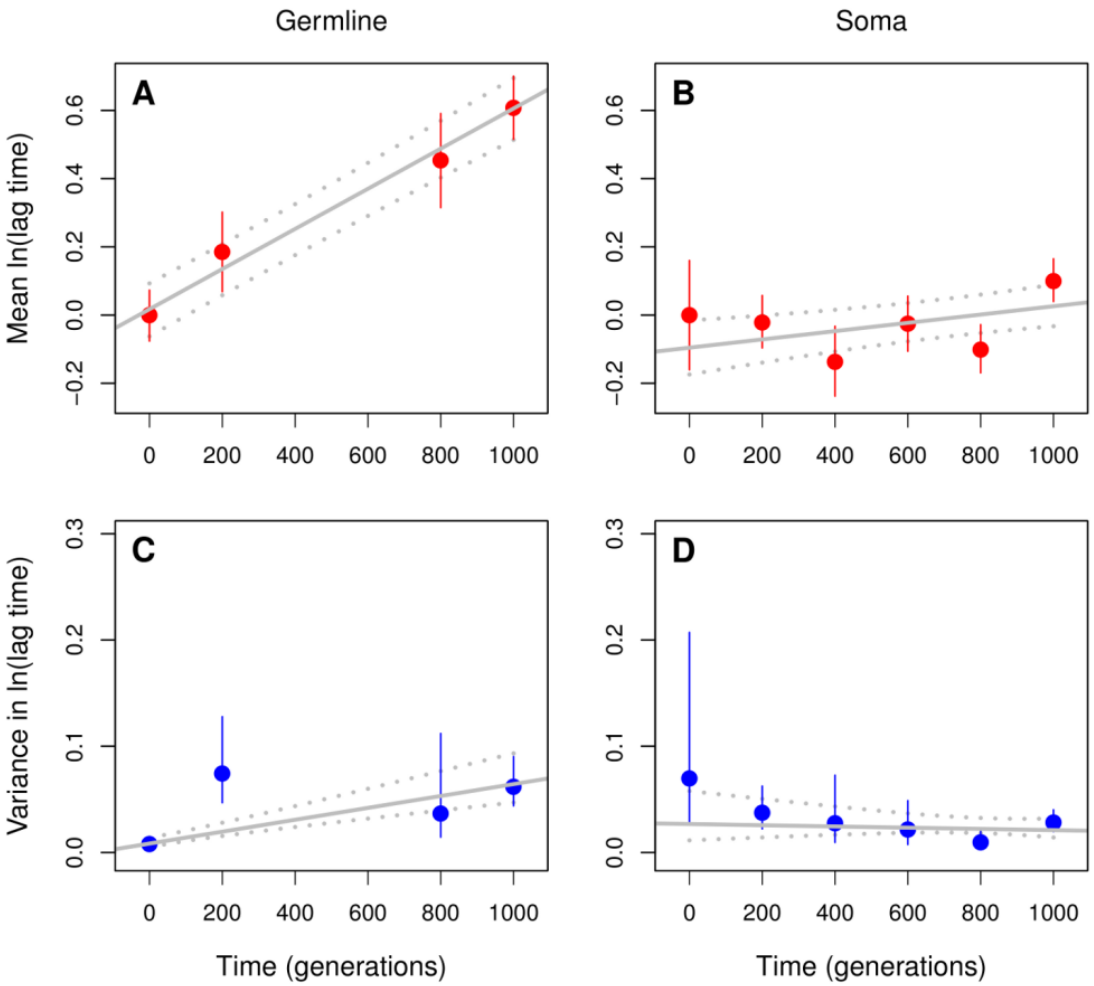


**Figure S2** **Results for lag time.** See the legend of Figure 2.1 for more details.

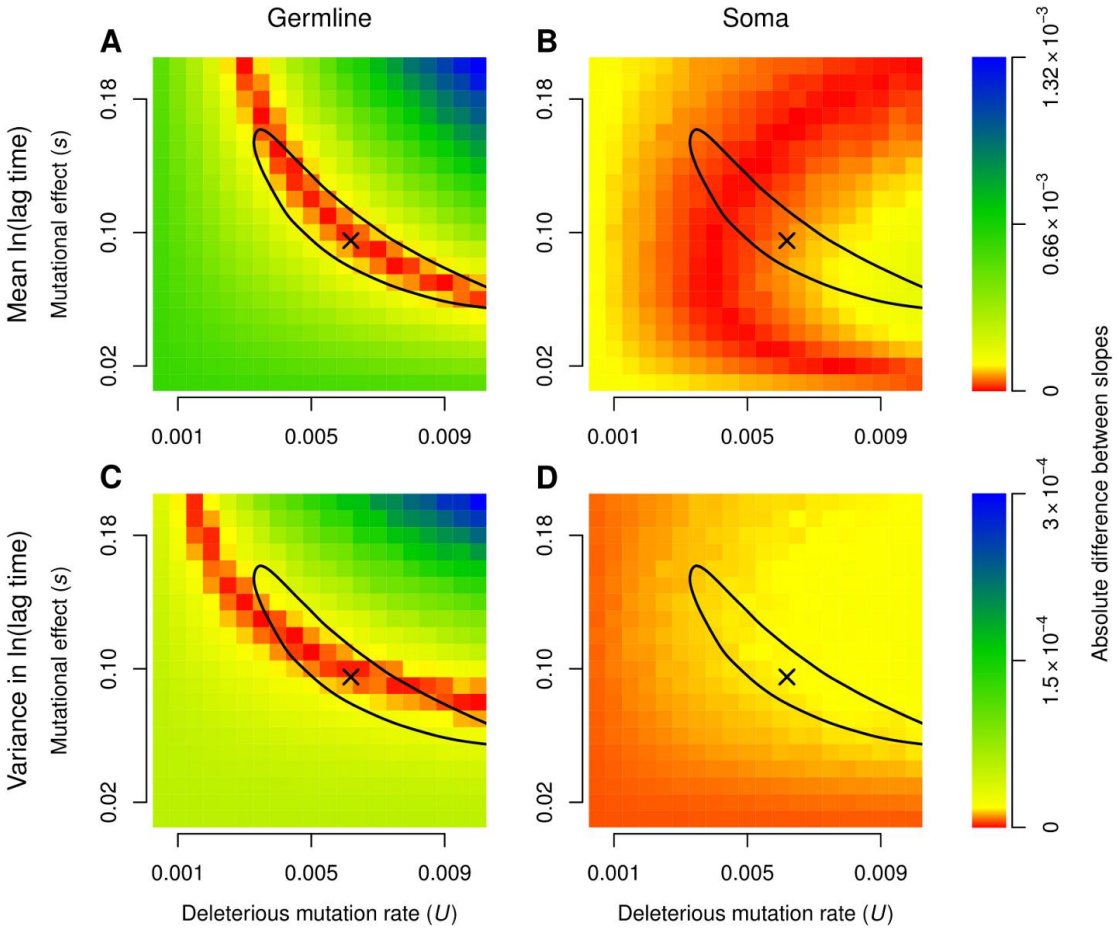


**Figure S3 Comparison between the slopes of the regressions shown in Figure S2 and those obtained from individual-based simulations of mutation accumulation under a broad range of mutational parameters.** See the legend of Figure 2.2 for more details.

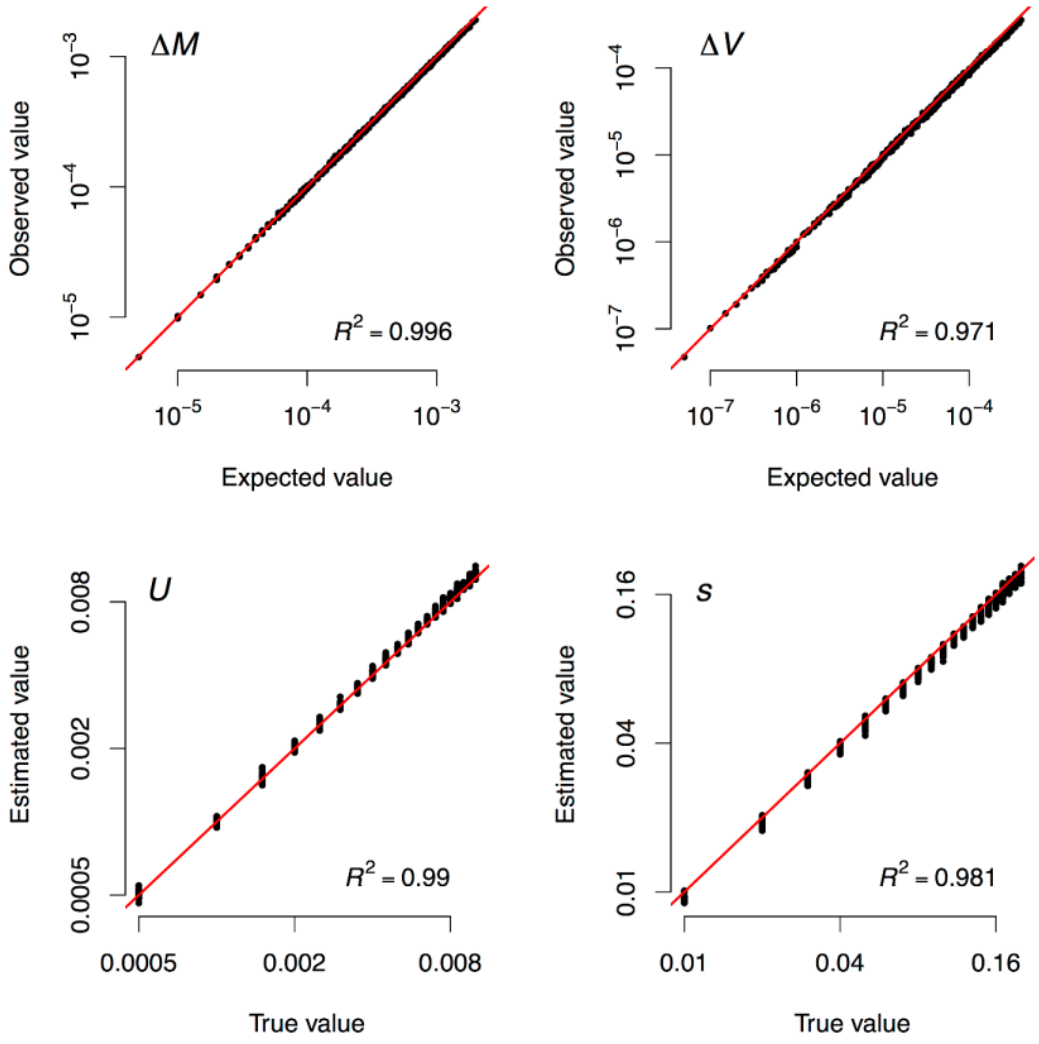


**Figure S4 Our estimators of mutational parameters in the MA experiment (Table 2.2) predict the results of individual-based simulations.** Each panel summarizes the results of the 400 individual-based simulations for different values of the deleterious mutation rate per haploid germline genome per generation ($U$) and the deleterious effect of a mutation when in a homozygous state ($s$) used to generate Figure 2.2. In each simulation, 1000 MA lines were evolved for 1000 generations. Estimates of germline fitness are based on a single genomic exclusion of a random individual sampled from each MA line. (Top) The change in mean fitness ($\Delta M$) and the change in among-line variance in fitness ($\Delta V$) were estimated by linear regression based on time-series data from every 10 generations (these estimates are used in Figure 2.2) and compared to their expected values obtained using the formulas in Table 2.2 (MA column). (Bottom) The mutational parameters $U$ and $s$ estimated from the observed values of $\Delta M$ and $\Delta V$ using the formulas in Table 2.2 (MA column) are compared to the actual values used in the simulations.

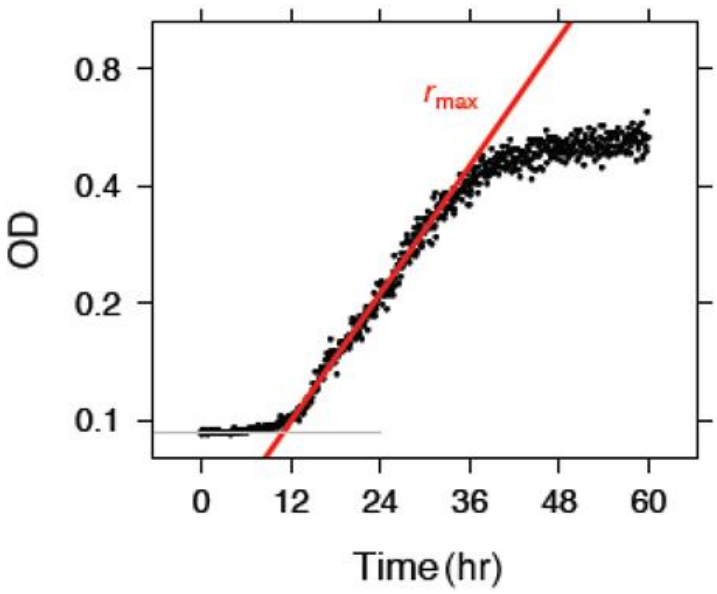


**Figure S5 Fitness assay example using one MA line started with 10 cells.** Division of cells from one cell line was monitored by measuring OD650 every 5 minutes for continuously 60 hours. Maximum population growth rate ($r_{max}$) was estimated by the slope of the curve in log phase (red line).

## 2.6 References


Allen, S., P. Ervin, N. McLarsen, and R. Brand, 1984 The 5S ribosomal RNA gene clusters in *Tetreahymena thermophila*: strain differences, chromosomal localization, and loss during micronuclear ageing. Mol Gen Genet 197: 244–253.

BARNETT, A. G., N. KOPER, A. J. DOBSON, F. SCHMIEGELOW, and M. MANSEAU, 2010 Using information criteria to select the correct variance-covariance structure for longitudinal data in ecology. Meth Ecol Evol **1:** 15–24.

BATEMAN, A., 1959 The viability of near-normal irradiated chromosomes. Int J Radiat Biol **1:** 170–180.

BEAUMONT, M. A., W. ZHANG, and D. J. BALDING, 2002 Approximate Bayesian computation in population genetics. Genetics **162:** 2025–2035.

BLACKBURN, E. H., and K. M. KARRER, 1986 Genomic reorganization in ciliated protozoans. Annu Rev Genet **20:** 501–521.

BRITO, P., E. GUILHERME, H. SOARES, and I. GORDO, 2010 Mutation accumulation in *Tetrahymena*. BMC Evol Biol **10:** 354.

BRUNS, P., and T. BRUSSARD, 1974 Positive selection for mating with functional heterokaryons in *Tetrahymena pyriformis*. Genetics **78:** 831–841.

BRUNS, P., and D. CASSIDY-HANLEY, 1999 Methods for genetic analysis. Methods Cell Biol **62:** 229–240.

BRUNS, P., H. SMITH, and D. CASSIDY-HANLEY, 1999 Long-term storage. Methods Cell Biol **62:** 213–218.

CHARLESWORTH, B., 1996 The good fairy godmother of evolutionary genetics. Curr Biol **6:** 220.

CHRISTENSEN, S., D. WHEATLEY, M. RASMUSSEN, and L. RASMUSSEN, 1995 Mechanisms controlling death, survival and proliferation in a model unicellular eukaryote *Tetrahymena thermophila*. Cell Death Differ **2:** 301–308.

COLE, E., and P. BRUNS, 1992 Uniparental cytogamy: a novel method for bringing micronuclear mutations of *Tetrahymena* into homozygous macronuclear expression with precocious sexual maturity. Genetics **132:** 1017–1031.

CSILLÉRY, K., O. FRANÇOIS, and M. G. B. BLUM, 2012 abc: an R package for approximate Bayesian computation (ABC). Method Ecol Evol **3:** 475–479.

DEVISSER, J., R. HOEKSTRA, and H. VAN DEN ENDE, 1997 An experimental test for synergistic epistasis and its application in *Chlamydomonas*. Genetics **145:** 815–819.

DEVISSER, J., R. HOEKSTRA, and H. VAN DEN ENDE, 1996 The effect of sex and deleterious mutations on fitness in *Chlamydomonas*. Proc R Soc B **263:** 193–200.

DOERDER, F. P., J. DEAK, and J. LIEF, 1992 Rate of phenotypic assortment in *Tetrahymena thermophila*. Dev Genet **13:** 126–132.

EISEN, J., R. COYNE, M. WU, D. WU, M. THIAGARAJAN *et al.*, 2006 Macronuclear genome sequence of the ciliate *Tetrahymena thermophila*, a model eukaryote. PLoS Biol **4:** 1620–1642.

EWENS, W. J., 2004 *Mathematical population genetics*. Springer, New York.

GELMAN, A., and D. B. RUBIN, 1992 Inference from iterative simulation using multiple sequences. Stat Sci **7:** 457–511.

GOROVSKY, M., M. YAO, J. KEEVERT, and G. PLEGER, 1975 Isolation of micro- and macronuclei of *Tetrahymena pyriformis*. Methods Cell Biol **9:** 311–327.

HADFIELD, J. D., 2010 MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. J Stat Softw **22:** 1–22.

HALL, D. W., and S. B. JOSEPH, 2010 A high frequency of beneficial mutations across multiple fitness components in *Saccharomyces cerevisiae*. Genetics **185:** 1397–1409.

HALLIGAN, D. L., and P. D. KEIGHTLEY, 2009 Spontaneous mutation accumulation studies in evolutionary genetics. Annu Rev Ecol Evol Syst **40:** 151–172.

HELLUNG-LARSEN, P., and A. P. ANDERSON, 1989 Cell volume and dry weight of cultured *Tetrahymena*. J Cell Sci **92:** 319–324.

HILLER, L. W., A. COULSON, J. I. MURRAY, Z. BAO, J. E. SULSTON *et al.*, 2005 Genomics in *C. elegans*: so many genes, such a little worm. Genome Res **15:** 1651–1660.

HOULE, D., K. A. HUGHES, S. ASSIMACOPOULOS, and B. CHARLESWORTH, 1997 The effects of spontaneous mutation on quantitative traits. II. Dominance of mutations with effects on life-history traits. Genet Res **70:** 27–34.

KIBOTA, T., and M. LYNCH, 1996 Estimate of the genomic mutation rate deleterious to overall fitness in *E. coli*. Nature **381:** 694–696.

KONDRASHOV, A., 1988 Deleterious mutations and the evolution of sexual reproduction. Nature **336:** 435–440.

LARSON, D., A. UMTHUN, and W. SHAIU, 1991 Copy number control in the *Tetrahymena* macronuclear genome. J Eukaryot Microbiol **38:** 258–263.

LYNCH, M., 2010 Evolution of the mutation rate. Tr Genet **26:** 345–352.

LYNCH, M., J. BLANCHARD, D. HOULE, T. KIBOTA, S. SCHULTZ *et al.*, 1999 Perspective: spontaneous deleterious mutation. Evolution **53:** 645–663.

MUKAI, T., 1964 The genetic structure of natural populations of *Drosophila melanogaster*. I. Spontaneous mutation rate of polygenes controlling viability. Genetics **50:** 1–19.

NANNEY, D., 1959 Vegetative mutants and clonal senility in *Tetrahymena*. J Eukaryot Microbiol **6:** 171–177.

NANNEY, D., 1974 Aging and long-term temporal regulation in ciliated protozoa. A critical review. Mech Ageing Dev **3:** 81–105.

PLUMMER, M., N. BEST, K. COWLES, and K. VINES, 2006 CODA: convergence diagnosis and output analysis for MCMC. R News **6:** 7–11.

R DEVELOPMENT CORE TEAM, 2011 R: A language and environment for statistical computing., pp. in *R Foundation for statistical computing*, Vienna, Austria.

SHAW, R. G., D. L. BYERS, and E. DARMO, 2000 Spontaneous mutational effects on reproductive traits of *Arabidopsis thaliana*. Genetics **155:** 369–378.

SIMMONS, M. J., and J. CROW, 1977 Mutations affecting fitness in *Drosophila* populations. Annu Rev Genet **11:** 49–78.

SPIEGELHALTER, D. J., N. G. BEST, B. P. CARLIN, and A. VAN DER LINDE, 2002 Bayesian measures of model complexity and fit. J R Statist Soc B **64:** 583–639.

VASSILIEVA, L. L., A. M. HOOK, and M. LYNCH, 2000 The fitness effects of spontaneous mutations in *Caenorhabditis elegans*. Evolution **54:** 1234–1246.

WAHL, L., and P. GERRISH, 2001 The probability that beneficial mutations are lost in populations with periodic bottlenecks. Evolution **55:** 2606–2610.

WANG, Y., A. C. DÍAZ, D. M. STOEBEL, and T. F. COOPER, 2012 Genetic background affects epistatic interactions between two beneficial mutations. Biol Lett **In Press**.

WEST, S. A., PETERS A. D., and B. N.H., 1998 Testing for epistasis between deleterious mutations. Genetics **149:** 435–444.

WILSON, A. J., D. RÉALE, M. N. CLEMENTS, M. M. MORRISSEY, E. POSTMA *et al.*, 2010 An ecologist's guide to the animal model. J Anim Ecol **79:** 13–26.

WLOCH, D. M., K. SZAFRANIEC, R. H. BORTS, and R. KORONA, 2001 Direct estimate of the mutation rate and distribution of fitness effects in the yeast *Saccharomyces cerevisiae*. Genetics **159:** 441–452.

# Chapter 3 Mutations on morphological traits of *T. thermophila*

Morphological traits are widely and popularly used in ciliates taxonomy; however, whether these traits are stable enough to be used for distinguishing species relies on taxonomists judgment, without being quantitatively tested. The study here shows how to test the stability of morphological traits to mutations using the *T. thermophila* mutation accumulation system we recently developed (collaborators: TP, RBRA, RAZ). The number of post-oral kineties never changed either in cell lines with or without mutations accumulated, so it should be given more weight as a species-distinguishing trait, supporting what taxonomists have long observed and been doing on the genus *Tetrahymena*; however, cell size of *T. thermophila* decreased as mutations accumulated in the germline genome for 1000 generations, consistent with findings in other unicellular or multicellular organisms. This indicates that cell size should be treated cautiously, although it is widely used in species distinguishment or definition. It was also shown here that cell size and fitness (maximum population growth rate as the fitness metric) both decreased after long-term mutation accumulation.

## 3.1 Introduction

Mutations can result in changes in a variety of life history or morphological traits, for example, spontaneous mutations could decrease longevity, survival rate, and body size of *C. elegans* over dozens to hundreds of generations (AZEVEDO *et al.* 2002; KEIGHTLEY and OHNISHI 1998; VASSILIEVA and LYNCH 1999). In addition, change in a morphological

trait may also be coupled with fitness change, *e.g.*, larger individuals having shorter generation time is not uncommon in unicellular organisms (MOLENAAR *et al.* 2009; MONGOLD and LENSKI 1996); however, this has never been reported in ciliates.

As for ciliates, mutations affecting morphological traits are especially important for taxonomy, which is mostly based on morphological traits in this group (LYNN 2008). Dozens of new morphological species are being reported every year, due to the wide application of silver-staining methods and improved microscopy. Morphological traits, like the number of post-oral kineties (No. PK, ciliary rows posterior to the oral region, also called post-oral meridians, POM) (Fig. 3.1), which is used as a species distinguishing trait in the genus *Tetrahymena* (CORLISS 1973), have never been tested with evolutionary genetic techniques to determine whether they are really stable to environmental or genetic influences to provide precise morphological species identification. Resolving such uncertainty in morphological traits might contribute to our understanding of the true protist species diversity which is under debate between ecologists and taxonomists (CARON 2009; FENCHEL and FINLAY 2004; FOISSNER *et al.* 2008; WEISSE and RAMMER 2006).

Four morphological traits, which are widely used in ciliate taxonomy and morphometric studies, were measured in this study (Fig. 3.1) (AGATHA and TSAI 2008; FAN *et al.* 2011; FOISSNER and STOECK 2011; JUNG *et al.* 2012; MCMANUS *et al.* 2010; TOKIWA *et al.* 2010; TSAI *et al.* 2010). Two types of cell lines were analyzed using mutation accumulation and morphological techniques. The two types of cell lines were: mutation accumulation (MA) lines, the cell lines experienced approximately 1000

generations' mutation accumulation; genomic exclusion (GE) lines, the cell lines bear homozygous mutations expressed from the germline genome of MA lines and are described in Chapter 2 (Table 3.1). I assessed whether the four morphological traits are stable to mutations and thus suitable for precise morphological taxonomy.

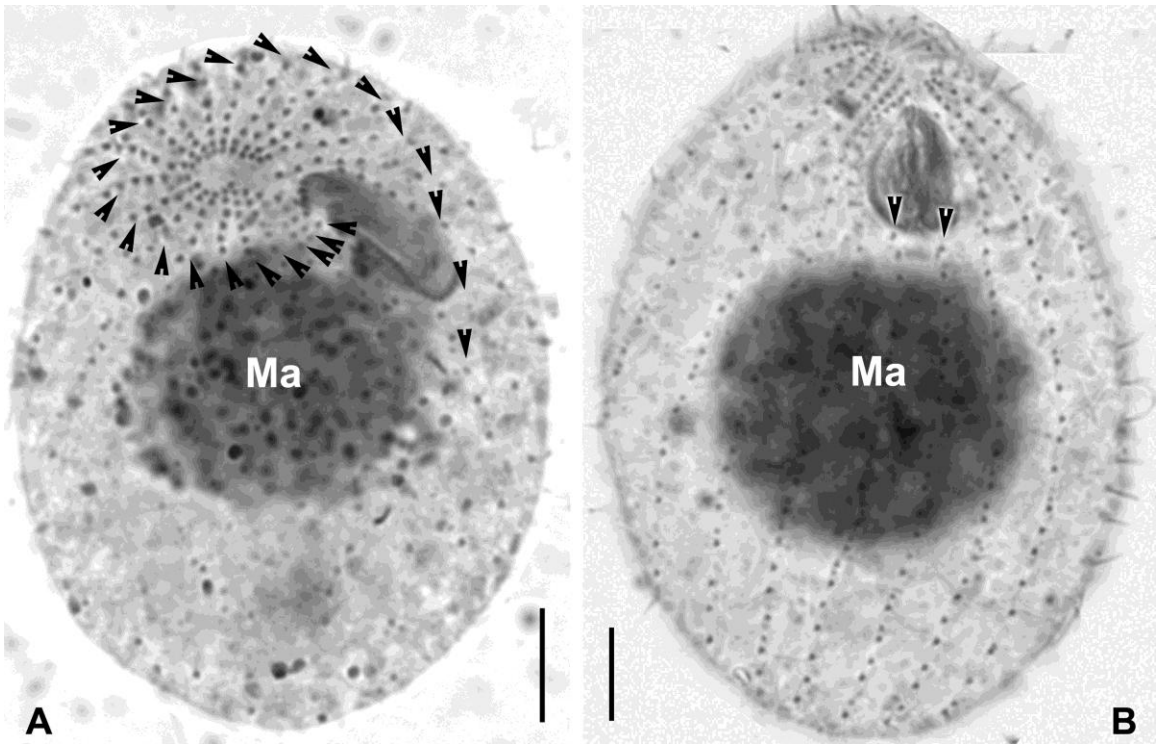


**Figure 3.1 Protargol-stained *T. thermophila* cells (A from the GE line 4-0-A, B from GE line 51-1000-A).** Arrowheads in **A** show all the somatic kineties (SK, including the two post-oral kineties) of the cell, arrowheads in **B** show the post-oral kineties (PK). The dark regions in the center in both pictures are the macronucleus (Ma). Both scale bars are 4.5 μm.

## 3.2 Results

The patterns of the four morphological traits (No. SK, No. PK, macronuclear size, and cell size) and fitness showing changes from the ancestor to the evolved MA and GE lines were as shown in Fig. 3.2A&B. In the evolved GE lines, macronuclear (Ma) size, cell size, and fitness of GE lines all decreased from those of the ancestral GE lines (decrease of 26%, 23%, and 26% separately for standardized macronuclear size, cell size, and fitness) (Fig. 3.2B, Ma Size, Cell Size, Fitness; Table 3.1). Though only marginal p-values from t-tests on cell size, fitness, and macronuclear size of ancestral and evolved GE lines were detected (Table 3.2), for the purpose of this discussion I will consider the differences as if they are significant. Since only 6 evolved GE lines were used, even the difference in fitness was not significant here, but fitness was significant in Chapter 2 using a much larger sample N=19). In contrast, the number of somatic kineties (No. SK) of the evolved lines showed an increase from the ancestral GE lines (12% for standardized No. SK) (Fig. 3.2B, No. SK; Tables 3.1&3.2; p=0.07). The among-line variance of Ma size, cell size, fitness, and No. SK in the evolved GE lines all increased. The number of post-oral kineties remained constantly two in all the ancestral and the evolved MA lines and GE lines.

**Table 3.1 Morphological traits (before standardization) and fitness of the cell lines in this study.** Cell line names with two hyphens indicate GE lines, with one hyphen refer to MA lines, 0 and 1000 after the first hyphen indicate generation numbers. No. SK: number of somatic kineties; No. PK: number of post-oral kineties; Ma: macronuclear size, both cell and macronuclear sizes are in $\mu m^3$; Cell: cell size; Sd: standard deviation; n: number of specimens observed, n* means the number of replicates in the fitness assay.

| Lines | No. SK(Sd)(n) | No. PK(Sd)(n) | Ma (Sd)(n) | Cell (Sd)(n) | Fitness(Sd)(n*) |
|---|---|---|---|---|---|
| 4-0 | 19.45(1.17)(20) | 2(0)(24) | 633(345)(110) | 4519(1136)(60) | 0.93(0.04)(3) |
| 27-0 | 19.88(1.05)(26) | 2(0)(32) | 1194(458)(98) | 8179(2461)(60) | 0.99(0.14)(2) |
| 60-0 | 18.96(0.79)(24) | 2(0)(23) | 844(324)(97) | 5256(1951)(60) | 1.08(0.17)(6) |
| Averages | 19.43 | 2 | 890 | 5984.81 | 1 |
| 5-1000 | 19.27(0.82)(15) | 2(0)(16) | 980(384)(72) | 6340(1897)(60) | 0.75(0.11)(6) |
| 32-1000 | 20.21(1.57)(28) | 2(0)(28) | 591(243)(108) | 6769(1985)(94) | 1.16(0.17)(4) |
| 40-1000 | 18.68(1.63)(25) | 2(0)(25) | 801(330)(133) | 5445(2196)(60) | 0.84(0.07)(2) |
| 50-1000 | 17.5(0.80)(12) | 2(0)(12) | 955(460)(95) | 5091(1343)(72) | 1.03(0.33)(4) |
| 51-1000 | 18.43(1.92)(40) | 2(0)(40) | 687(250)(105) | 6285(1727)(80) | 0.98(0.15)(5) |
| 62-1000 | 18.62(1.9)(34) | 2(0)(34) | 604(220)(114) | 6553(1389)(86) | 0.99(0.21)(3) |
| Averages | 18.79 | 2 | 770 | 6080 | 0.96 |
| 4-0-A | 18.22(1.27)(27) | 2(0)(41) | 814(429)(88) | 5900(1694)(60) | 0.85(0.09)(3) |
| 27-0-X | 17.77(1.04)(22) | 2(0)(13) | 1233(754)(97) | 7714(1675)(60) | 0.93(0.03)(4) |
| 60-0-A | 17(0.83)(10) | 2(0)(10) | 791(383)(102) | 6470(1511)(60) | 1.22(0.29)(6) |
| Averages | 17.66 | 2 | 946 | 6695 | 1 |
| 5-1000-3U1 | 18.09(0.88)(11) | 2(0)(10) | 577(249)(78) | 4617(1048)(60) | 0.69(0.05)(7) |
| 32-1000-A | 21.54(1.47)(24) | 2(0)(24) | 709(352)(110) | 5376(1694)(85) | 0.68(0.16)(3) |
| 40-1000-AM1 | 16.5(0.63)(14) | 2(0)(10) | 755(336)(49) | 3849(1025)(60) | 0.62(0.09)(6) |
| 50-1000-3W3 | 19.17(1.1)(12) | 2(0)(16) | 568(233)(71) | 4693(1098)(72) | 0.64(0.05)(6) |
| 51-1000-A | 22.1(2.11)(39) | 2(0)(39) | 694(240)(117) | 6755(2046)(95) | 0.83(0.13)(6) |
| 62-1000-B | 21.73(1.87)(15) | 2(0)(15) | 854(366)(113) | 5584(1296)(84) | 0.94(0.12)(8) |
| Averages | 19.86 | 2 | 693 | 5146 | 0.73 |

In contrast, in the MA lines, none of the mean values of the morphological traits or fitness changed significantly, not even close to the cutoff p-value (Fig. 3.2A; Table 3.2).

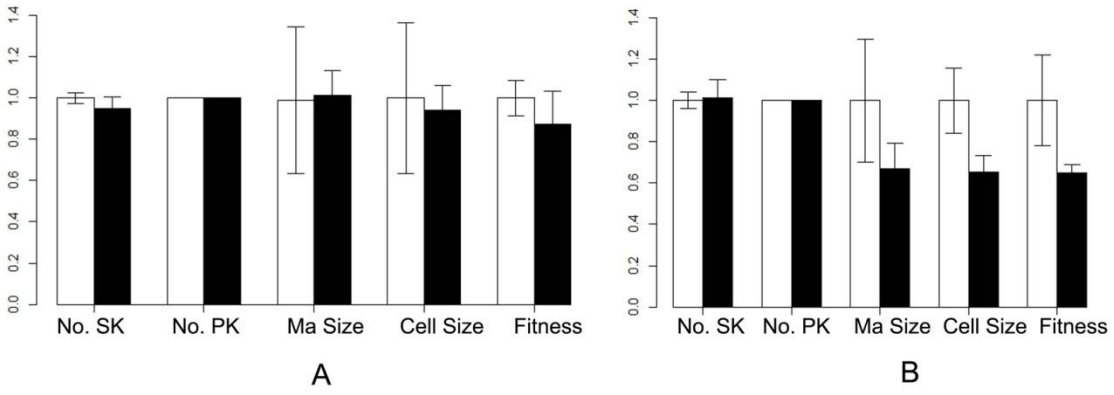


**Figure 3.2 Morphological traits and fitness from ancestor (blank bars, generation 0) and evolved (filled bars, generation 1000) MA lines (A) and GE lines (B)**. Note each trait or fitness was standardized by dividing the observed values with the ancestral mean. Error bars are 95% confidence intervals.

Linear regression analysis on the cell size and the fitness of GE lines showed they were significantly and positively correlated (Fig. 3.3D; $r=0.47$, $p<0.05$; Table 3.2). This is an indication of cell size and fitness both being affected by accumulated mutations, but whether there is mutational pleiotropy needs further study. None of the other morphological traits of GE lines (Fig. 3.3A-C) nor all the morphological traits of the MA lines (Fig. 3.4) had a significant correlation with fitness, though the macronuclear size regression with fitness did show very similar trend with that of the cell size vs. fitness regression (Figs. 3.3C).

**Table 3.2 p-values from two sided Student t-tests on the ancestral and the evolved cell lines.** GE, genomic exclusion lines with germline mutations in the MA lines expressed; MA, mutation accumulation lines without germline mutations expressed; Ma, macronuclear.

| Cell lines | No. SK | No. PK | Ma size | Cell size | Fitness |
|---|---|---|---|---|---|
| MA | 0.21 | N/A | 0.55 | 0.94 | 0.60 |
| GE | 0.07 | N/A | 0.23 | 0.08 | 0.12 |

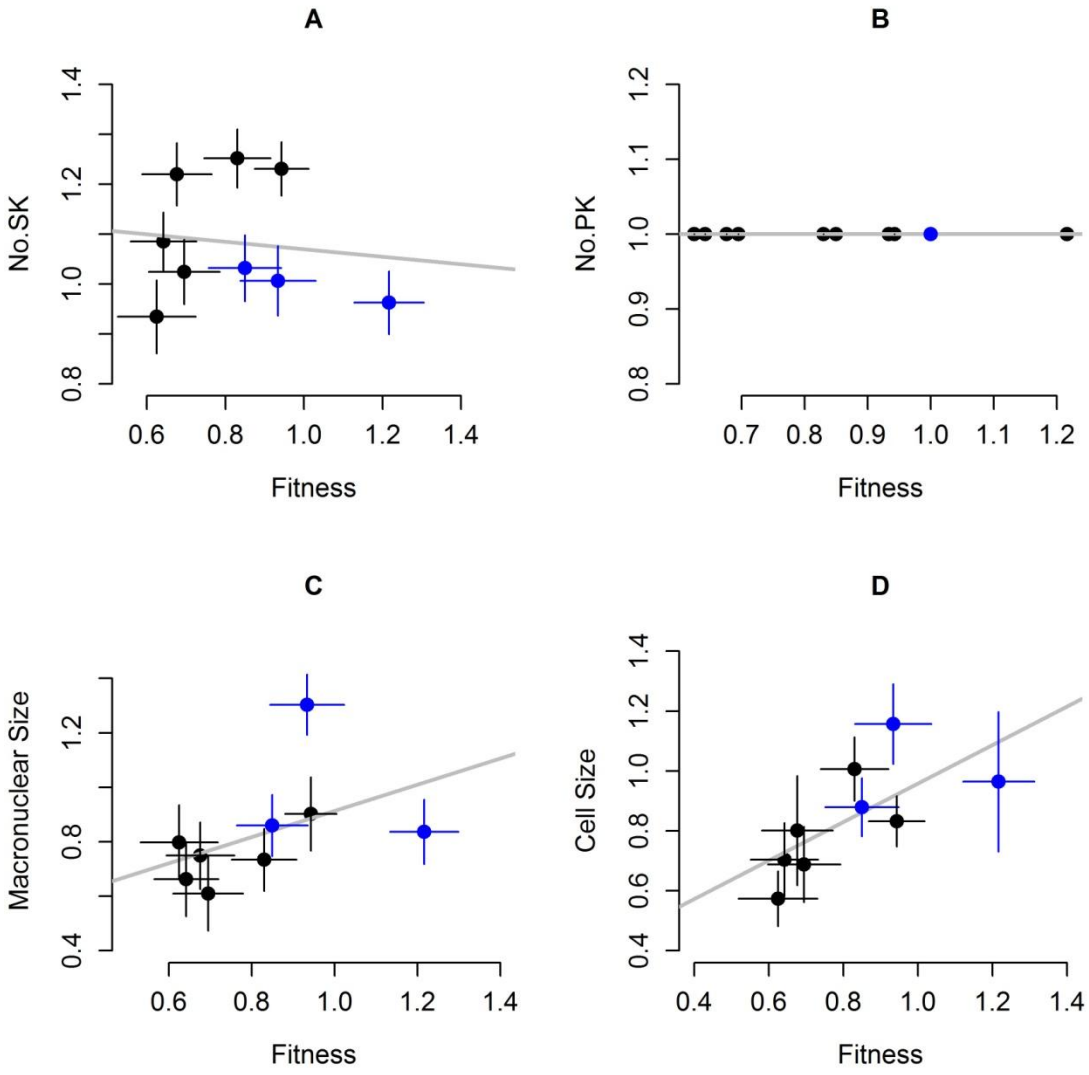


**Figure 3.3 Linear regressions between morphological traits and fitness of GE lines.** Blue dots show the ancestor and black dots show the evolved lines. No. SK: number of somatic kineties; No. PK: number of post-oral kineties. Error bars are 95% confidence intervals.

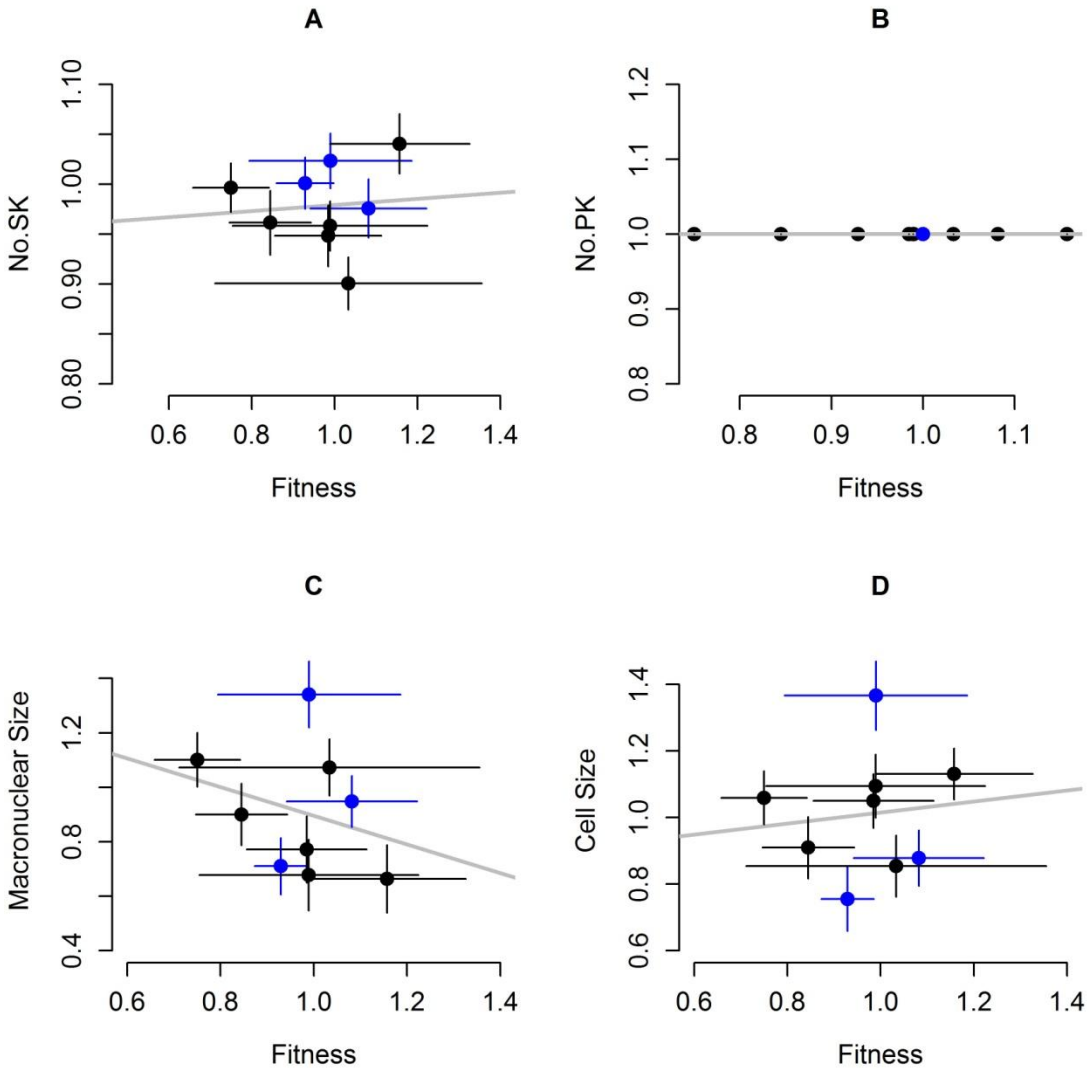


**Figure 3.4 Linear regressions between morphological traits and fitness of MA lines.** Blue dot shows the ancestor and black dots show the evolved lines. No. SK: number of somatic kineties; No. PK: number of post-oral kineties. Error bars are 95% confidence intervals. Note the three ancestral MA lines had identical genotypes and thus combined into one blue dot here.

## 3.3 Discussion

The cell size and fitness in the GE lines are positively correlated (Fig. 3.3D). In fact, such positive correlation between growth rate and cell size is not unprecedented, but rather

widespread (MONGOLD and LENSKI 1996; STEARNS 1992). Mongold and Lenski proposed that larger cells have more reserves for growth and that this is the cause of the correlation between growth rate and cell size; however, since the cell size was measured on starved cells here, the observed size differences are not likely indicative of different reserves from the log phase before starvation; thus, not supporting this hypothesis (MONGOLD and LENSKI 1996).

In addition, cell size decreased only when mutations were shielded from selection (cell size decreased in GE lines, but not in MA lines, see Fig. 3.2A,B, Cell Size). This suggests that mutations that decrease cell size are selected against in MA lines, highly similar to what happened to fitness of MA lines (See Section 2.2.1 in chapter 2). In nature, the selection against cell size decrease might also occur due to the avoidance of size-selective predators (*e.g.*, copepods, amoebae), which are extremely abundant in the freshwater habitats of *T. thermophila* and known to induce morphological change in ciliates (KUSCH 1993).

Finally, whether such correlation was a reflection of mutational pleiotropy on both fitness and cell size remains uncertain, and it is already known that many life history, behavioral and morphological traits are positively correlated (KEITHLEY and OHNISHI 1998). One cause of such uncertainty was the uncertain and probably different number of mutations in each GE line involved in this study. Multiple genomic exclusions on each MA line and further assays on fitness or cell size could be used to find out an accurate number of mutations affecting fitness or cell size in each MA line. If the numbers of mutations causing fitness or cell size to change are similar, this would be one sign of

mutational pleiotropy, *viz.* one mutation could have multiple effects on both fitness and cell size. Since mutational effects are also affected by other complicating factors like dominance coefficient and epistasis of new mutations, sorting out mutational pleiotropy effects would need taking these factors into account, as in (CABALLERO and KEIGHTLEY 1994).

**Table 3.3 Three morphological traits of 10 *Tetrahymena* species.** Length and width in μm; [+] normal dimensions of *T. thermophila*; [a] trophont life stage data; [b] free-living phase data; [c] cell length; [d] microstome life stage data.

| Species | Length×Width | No. SK | No. PK | References |
|---|---|---|---|---|
| *T. thermophila*[+] | 50×20 | 18–21 | 2 | (NANNEY and SIMON 1999) |
| *T. setifera* | 40×25 | 22–26 | 1–3 | (CORLISS 1973) |
| *T. chironomi* | 40×23 | 23–28 | 2 | (CORLISS 1973) |
| *T. rostrata*[a] | 20–80[c] | 27–35 | 1–4 | (CORLISS 1973) |
| *T. limacis*[b] | 40–45[c] | 24–32 | 1–4 | (CORLISS 1973) |
| *T. corlissi* | 47×31 | 25–31 | 2 | (CORLISS 1973) |
| *T. stegomyiae* | 60–100[c] | 25–30 | 2 | (CORLISS 1973) |
| *T. patula*[d] | 45×28 | 32–41 | 3–5 | (CORLISS 1973) |
| *T. vorax*[d] | 31–115[c] | 18–23 | 2 | (CORLISS 1973) |
| *T. paravorax*[d] | 70–90[c] | 22–30 | 2 | (CORLISS 1973) |

Similar to the number of post-oral kineties, cell size is also one of the most widely recorded traits in ciliate morphological taxonomy; however, in contrast to the limited variability of the number of post-oral kineties even across the whole *Tetrahymena* genus (Table 3.3), cell size being affected by spontaneous mutations accumulated in the germline genome in this research raises question on whether this trait should be used for

distinguishing or defining different species, especially in sexual species like *T. thermophila*. Morphological traits changing before and after sex in sexual ciliates, which experienced long-term asexual reproduction, would cause confusion in morphological species identification. Though many species in the *Tetrahymena* genus did take cell size as one of the species' features (Table 3.3), my results indicate that the cell size should be treated cautiously, especially in defining species. For example, the cell length×width (in μm) of the GE lines' cells decreased from the ancestral 35.34×18.86 μm to the evolved (at generation 1000) 34.80×16.58 μm, further deviated from previously reported *T. thermophila* size (in length×width; Table 3.3). Mutational variability of more quantitative morphological traits should also be tested following this evolutionary genetic case study, to provide solid data support for current ciliate morphological taxonomy. Morphological traits, like the number of post-oral kineties which are robust to the test, should be put more weight in species identification. Fortunately, morphological taxonomists have long noticed the stability of such traits and did give them more weight in species distinguishment, for example, the 3–5 (mostly 4) post-oral kineties is the species-unique trait of *Tetrahymena patula* (Table 3.3).

## 3.4 Materials and Methods

*Cell lines and media*

See section 2.4-*Strains and media*, chapter 2 for details. Nine mutation accumulation (MA) lines (three ancestral lines and six evolved lines about 1000 cell divisions in time apart from the ancestor) and another nine genomic exclusion (GE) lines derived from

these MA lines (one GE line for each MA line, see "genomic exclusion" in section 2.4, chapter 2 for detailed procedures) in this research were from a previous mutation accumulation experiment and frozen in liquid nitrogen (section 2.4, chapter 2; Table 3.1; the cell lines were evenly sampled from all the MA lines in section 2.4, chapter 2, with reference to the fitness of the GE lines).

*Morphological traits*

All cell lines were thawed from liquid nitrogen following (BRUNS *et al.* 1999): cryovials from a liquid nitrogen tank were quickly transferred to 42℃ water bath, thawed cells were then transferred to 4ml SSP medium supplemented with 1X Penicillin, Streptomycin, and Amphotericin B (PSA; final concentration; Amresco, Solon, OH, USA) in wells of a 6-well plate, pre-warmed in a 30℃ incubator. For each line, the plate with thawed cells was incubated in the 30℃ incubator for 48 hours; then, 100μl culture from each well was transferred to 12.5ml SSP supplemented with 1X PSA in a 125ml glass flask. Future morphological analysis should use more batch cultures for each thawed line to disentangle the mutational and the environmental variance of morphological traits. Another 24 hours later, 100μl cells in log phase from each flask were stained with protargol (WILBERT 1975): cells were fixed with Bouins' fluid/saturated $HgCl_2$ mixture (volume ratio 1:1) for 1 min; then fixatives were removed and fixed cells were rinsed with DI water at least six times, after this, cells were bleached with sodium hypochlorite solution (NaClO) 1:600 diluted from the stock solution (CAS Number 7681-52-9, Sigma-Aldrich Co. LLC.), this procedure should be finished within 2 min to avoid over-bleaching. Bleached cells were then immerged with 1% protargol solution (protargol S™

from Polysciences Inc., Warrington, PA, Cat No.: 01070) in an embryo dish and heated at 60 ℃ for 1 hour before silver developing and fixing. Numbers of somatic kineties and post-oral kineties of protargol stained specimens were recorded using a compound Olympus microscope. The length and width of macronuclei were measured by first taking photomicrographs using digital camera connected to an Olympus inverted microscope; then the photomicrographs were analyzed for macronuclear length and width in Image J (1.46r, NIH, USA). Scale bars in the photomicrograph program were calibrated with a stage micrometer. Macronuclear size was then calculated based on the length and width measurements, by assuming a prolate spheroid shape of the macronucleus, using the formula from (HELLUNG-LARSEN and ANDERSON 1989):

$$V = (\pi/6) \times L \times W^2$$

where $V$ is the macronuclear volume (size), $L$ is the macronuclear length, and $W$ is the macronuclear width.

To get the size of synchronized cells, cells in log-phase from the above thawing cultures in one 125ml flask were also centrifuged in a 50ml conical tube. Supernatant in the conical tube was removed and cells were re-suspended with 12.5ml Tris buffer and transferred to a clean 125ml flask and starved for 24 hours. Starved cells were then transferred to a 6-well plate and photomicrographs were taken on them on an Olympus inverted microscope. Cell length and width were acquired from the photomicrographs analyzed in Image J (1.46r, NIH, USA). Similar to macronuclear size, cell size was then calculated based on the length and width measurements, by assuming a prolate spheroid-shape of the cell, using the above formula from (HELLUNG-LARSEN and ANDERSON 1989).

Morphological traits for each line were listed in Table 3.1. Each trait was standardized by dividing the corresponding mean ancestor values.

*Plotting and statistical analyses*

Plots, linear regression and statistical tests were all done in R (2.15.0) (IHAKA and GENTLEMAN 1996).

## 3.5 References


AGATHA, S., and S. F. TSAI, 2008 Redescription of the Tintinnid *Stenosemella pacifica* Kofoid and Campbell, 1929 (Ciliophora, Spirotricha) based on live observation, protargol impregnation, and scanning electron microscopy. J Eukaryot Microbiol **55:** 75–85.

AZEVEDO, R. B. R., P. D. KEIGHTLEY, C. LAUREN-MAATTA, L. L. VASSILIEVA, M. LYNCH *et al.*, 2002 Spontaneous mutational variation for body size in *Caenorhabditis elegans*. Genetics **162:** 755–765.

BRUNS, P., H. SMITH, and D. CASSIDY-HANLEY, 1999 Long-term storage. Methods Cell Biol **62:** 213–218.

CABALLERO, A., and P. D. KEIGHTLEY, 1994 A pleiotropic nonadditive model of variation in quantitative traits. Genetics **138:** 883–900.

CARON, D., 2009 Protistan biogeography: why all the fuss? J Eukaryot Microbiol **56:** 105–112.

CORLISS, J., 1973 History, taxonomy, ecology, and evolution of species of *Tetrahymena* in *Biology of Tetrahymena*, edited by A. M. ELLIOTT. Dowden, Hutchinson & Ross Inc., Stroudsburg, Pennsylvania.

FAN, X. P., X. F. LIN, K. A. S. AL-RASHEID, S. A. AL-FARRAJ, A. WARREN *et al.*, 2011 The diversity of scuticociliates (Protozoa, Ciliophora): a report on eight marine forms found in coastal waters of China, with a description of one new species. Acta Protozool **50:** 219–234.

FENCHEL, T., and B. J. FINLAY, 2004 The ubiquity of small species: patterns of local and global diversity. Bioscience **54:** 777–784.

FOISSNER, W., A. CHAO, and L. A. KATZ, 2008 Diversity and geographic distribution of ciliates (Protista: Ciliophora). Biodivers Conserv **17:** 345–363.

FOISSNER, W., and T. STOECK, 2011 *Cotterillia bromelicola* nov. gen., nov. spec., a gonostomatid ciliate (Ciliophora, Hypotricha) from tank bromeliads (Bromeliaceae) with de novo originating dorsal kineties. Euro J Protistol **47:** 29–50.

HELLUNG-LARSEN, P., and A. P. ANDERSON, 1989 Cell volume and dry weight of cultured *Tetrahymena*. J Cell Sci **92:** 319–324.

IHAKA, R., and R. GENTLEMAN, 1996 R: A language for data analysis and graphics. J Comput Graph Stat **9:** 491–508.

JUNG, J. H., K. M. PARK, and G. S. MIN, 2012 Morphology, morphogenesis, and molecular phylogeny of a new brackish water ciliate, *Pseudourostyla cristatoides* n. sp., from Songjiho lagoon on the coast of East Sea, South Korea. Zootaxa **3334:** 42–54.

KEIGHTLEY, P. D., and O. OHNISHI, 1998 EMS-induced polygenic mutation rates for nine quantitative characters in *Drosophila melanogaster*. Genetics **148:** 753–766.

KUSCH, J., 1993 Behavioural and morphological changes in ciliates induced by the predator *Amoeba proteus*. Oecologia **96:** 354–359.

LYNN, D. H., 2008 *The ciliated protozoa. Characterization, classification and guide to the literature*. Springer, New York.

MCMANUS, G. B., D. P. XU, B. A. COSTAS, and L. A. KATZ, 2010 Genetic identities of cryptic species in the *Strombidium stylifer*/*apolatum*/*oculatum* cluster, including a description of *Strombidium rassoulzadegani* n. sp. J Eukaryot Microbiol **57:** 369–378.

MOLENAAR, D., R. VAN BERLO, D. DE RIDDER, and B. TEUSINK, 2009 Shifts in growth strategies reflect tradeoffs in cellular economics. Mol Syst Biol **5:** 323.

MONGOLD, J. A., and R. E. LENSKI, 1996 Experimental rejection of a nonadapaptive explanation for increased cell size in *Escherichia coli*. J Bacteriol **178:** 5333–5334.

NANNEY, D. L., and E. M. SIMON, 1999 Laboratory and evolutionary history of *Tetrahymena thermophila* in *Tetrahymena thermophila*, edited by D. ASAI and J. FORNEY. Academic Press, San Diego, California.

STEARNS, S. C., 1992 *The evolution of life histories.* Oxford University Press, New York.

TOKIWA, T., D. MODRY, A. ITO, K. POMAJBIKOVA, K. J. PETRZELKOVA *et al.*, 2010 A new entodiniomorphid ciliate, *Troglocorys cava* n. g., n. sp., from the wild eastern chimpanzee (Pan troglodytes schweinfurthii) from Uganda. J Eukaryot Microbiol **57:** 115–120.

TSAI, S. F., J. Y. CHEN, and K. P. CHIANG, 2010 *Spirotontonia taiwanica* n. sp. (Ciliophora: Oligotrichida) from the coastal waters of northeastern Taiwan: morphology and nuclear small subunit rDNA sequences. J Eukaryot Microbiol **57:** 429–434.

VASSILIEVA, L. L., and M. LYNCH, 1999 The rate of spontaneous mutation for life-history traits in *Caenorhabditis elegans*. Genetics **151:** 119–129.

WEISSE, T., and S. RAMMER, 2006 Pronounced ecophysiological clonal differences of two common freshwater ciliates, *Coleps spetai* (Prostomatida) and *Rimostrombidium lacustris* (Oligotrichida), challenge the morphospecies concept. J Plankton Res **28:** 55–63.

WILBERT, N., 1975 Eine verbesserte Technik der Protargolimprägnation für Ciliaten. Mikrokosmos **64:** 171–179.

# Chapter 4 Reproductive strategies in the marine free-living ciliate *Glauconema trihymene*

The previous chapters 2 & 3 used genetic tools (*e.g.*, genomic exclusion) developed from ciliates' sexual reproduction (conjugation). I also explored whether other types of reproduction could be used for evolutionary genetic studies in another ciliate *G. trihymene*. Most free-living ciliates reproduce by equal fission or budding during vegetative growth. In certain ciliates, reproduction occurs inside the cyst wall, *viz*. reproductive cysts, but more complex reproductive strategies have generally been thought to be confined to parasitic or symbiotic species, *e.g.*, *Radiophrya* spp.

## 4.1 Introduction

Ciliates are a diverse group of unicellular eukaryotes characterized by two kinds of nuclei in each cell: a germline micronucleus and a somatic macronucleus. Free-living ciliates are known to exhibit diversity in modes of reproduction (FOISSNER 1996; FOISSNER *et al.* 2008; LYNN 2008). Most of these reproductive modes include equal fission or budding. In certain ciliates, including *Tetrahymena patula* and *Colpoda inflata*, reproduction can also occur inside the cyst wall, *viz*., reproductive cysts (CORLISS 2001; LYNN 2008).

Symbiotic ciliates like the astomatid ciliates, *e.g.*, *Radiophrya* spp., and certain apostomatid ciliates, *e.g.*, *Polyspira* spp., reproduce by forming cell chains, also called catenoid colonies, which are usually brought about by repeated asymmetric division without separation of the resulting filial products (CHATTON and LWOFF 1935; LYNN

2008). Some *Tetrahymena*, such as temperature-sensitive cytokinesis-arrest mutants of *T. thermophila*- strain cdaC, and *T. pyriformis* also showed similar cell chains at high temperature (FRANKEL 1964; FRANKEL 1977), and similar morphotypes were also recently reported in the non-reproductive artificial lethal mutants of *T. thermophila* (THAZHATH *et al.* 2002). Certain free-living flagellates are also known to conduct repeated nuclear replication without cytokinesis and then releasing daughter cells, when prey cells are abundant (CHANTANGSI and LEANDER 2010). However, no free-living ciliates have been reported to form cell chains in response to food (bacteria) concentration.

During early and late phases of equal fission, most ciliates share certain features, such as common positioning of the macronucleus and micronucleus, synchronization of macronuclear amitosis and fission furrow, and a specific and well defined dividing size (ADL and BERGER 1996; COHEN and BEISSON 1980; LYNN and TUCKER 1976). It is generally assumed that if food density meets requirements of both cell development and division, the daughter cells will be identical, so after division, the two daughter cells could not be differentiated from each other (FENCHEL 1990; JAWORSKA *et al.* 1996; ORIAS 1976).

However, ciliates from the same single cell isolate were reported to have high diversity in physiological states, such as cell size and volume, growth rate, feeding, and digestion (DOLAN and COATS 2008; HATZIS *et al.* 1994; LYNN 1975; WEISSE and RAMMER 2006), and certain ciliates even develop highly unique physiological strategies to maximally adapt to their habitats. For example, after feeding on the cryptomonad

*Geminigera cryophila*, the mixotrophic red-tide-causing ciliate *Myrionecta rubra* retains the prey organelles, which continue to function in the ciliate for up to 30 days (JOHNSON *et al.* 2007; TAYLOR *et al.* 1969). Comprehensive analysis of physiological state changes of ciliates usually requires monitoring of individuals for a relatively long period and therefore is rarely conducted (DOLAN and COATS 2008). Most ciliates are currently un-culturable or swim too fast for microscopic observation, further hindering such analyses.

In this study, I described a series of reproductive strategies that have been previously unknown in free-living ciliates. These types of reproduction occurred in all newly established cultures of *Glauconema trihymene*, a free-living scuticociliate belonging to the class Oligohymenophorea, which also includes *Tetrahymena* and *Paramecium*. The division processes and the relationship between persistence time of asymmetric divisions and bacteria concentrations are described, and an updated life cycle and phylogenetic position of *G. trihymene* are presented.

## 4.2 Results

### *4.2.1 Natural history of G. trihymene*

The *G. trihymene* isolate described here, collected in Hong Kong, is free-living and bacterivorous. It has a polyphenic life cycle that includes the following three previously described stages (MA *et al.* 2006; THOMPSON 1966): trophont, reniform, the feeding, and division stage, mostly 35 ×20 μm *in vivo* (Fig. 4.1A, B); tomite, the dispersion, and

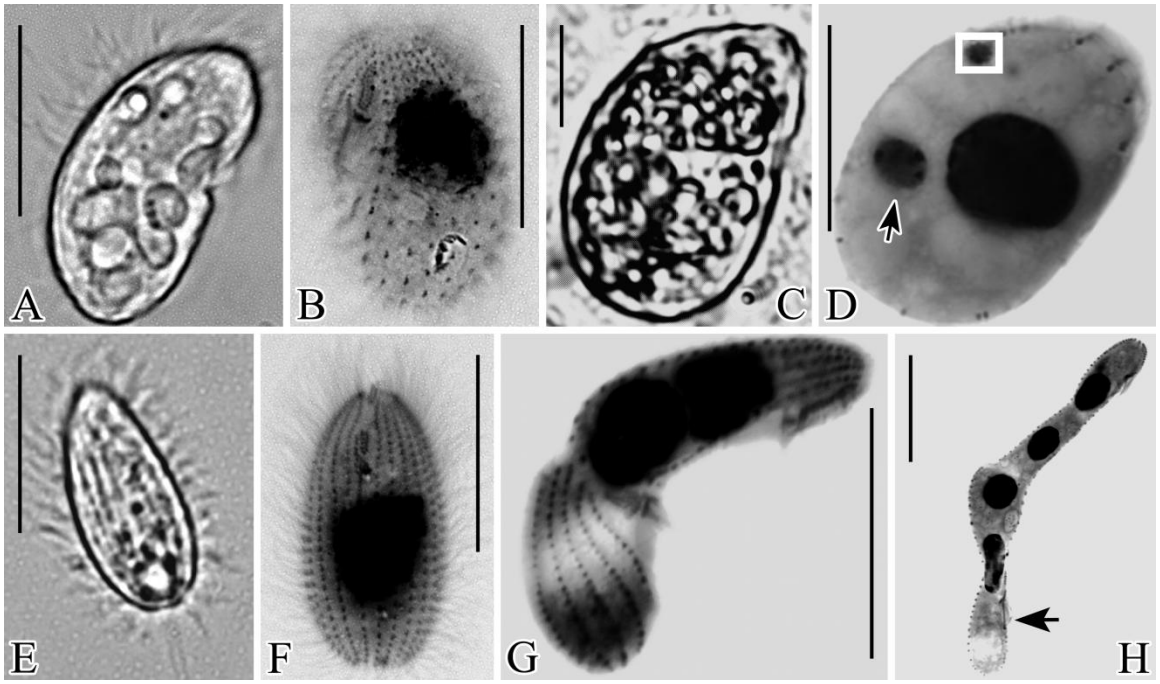


**Figure 4.1** ***G. trihymene*** **morphotypes.** A, C, E were from living cells; B, D, F- H were from protargol impregnated specimens. A, B. Lateral and ventral view of trophonts. C. A well-fed trophont. D. One probable asymmetric divider. Arrow marks the smaller macronucleus. The white square frame marks the micronucleus from a different plane of focus. The smaller macronucleus differs from the micronucleus by having many nucleoli. E, F. Ventral view of tomites. G. One asymmetric divider with two displaced macronuclei. H. One long asymmetric divider, probably releasing one trophont (arrow). Scale bars: A-H: 25 μm.

fast-swimming stage in response to starvation, with a spindle-shaped cell, mostly 30×15 μm *in vivo* (Fig. 4.1E, F); resting cyst, mostly rounded, dormant stage during trophic depletion, about 20 μm in diameter. Like other free-living ciliates, *G. trihymene* has a transcriptionally-active macronucleus and a germline micronucleus. The infraciliature and buccal apparatus are the same as in previous reports; however, I found the life cycle was much more complicated and included two reproductive modes new to scuticociliates, asymmetric division and reproductive cysts.

### *4.2.2 Processes of asymmetric division in young cultures*

Many slowly moving, well-fed trophonts (Fig. 4.1C) appeared within 24 hours after inoculation with tomites in cultures of wheat grain medium. In all of the cultures, a trophont underwent a cell division, but cytokinesis was arrested prior to completion, creating a unit consisting of two cells, now called "subcells" because of their failure to separate. Typically, each of the two connected subcells later underwent a second transverse division, resulting in a chain of four subcells, each with a macronucleus, an oral apparatus, and a contractile vacuole (Figs. 4.1H; 4.2A). I define these chains of subcells as asymmetric dividers. Asymmetric dividers vary in sizes from 30×15 μm to 180×30 μm *in vivo*, have diverse shapes consisting of chains of 2–4 subcells (Figs. 4.1G, H; 4.2A, J, O) and give rise to two filial cells that could be morphologically differentiated from each other after each division.

Similar asymmetric dividers were also repeatedly found in different cultures, though the sizes varied with media type. Up to four macronuclei were found in the cytoplasm of each asymmetric divider (Fig. 4.1H). Most un-disturbed asymmetric dividers attached to the bottom of Petri dishes, moved very slowly or stayed immobile and had two or more rounded contractile vacuoles, pulsating with different frequencies (arrows in Fig. 4.2C). The number of asymmetric dividers in the cultures increased with time from appearance of the first asymmetric divider.

Several asymmetric dividers were continuously followed on inverted microscopes. Two typical division processes of asymmetric dividers in young cultures (the 3rd or 4th day after inoculation) are described in detail (Fig. 4.2A–M):

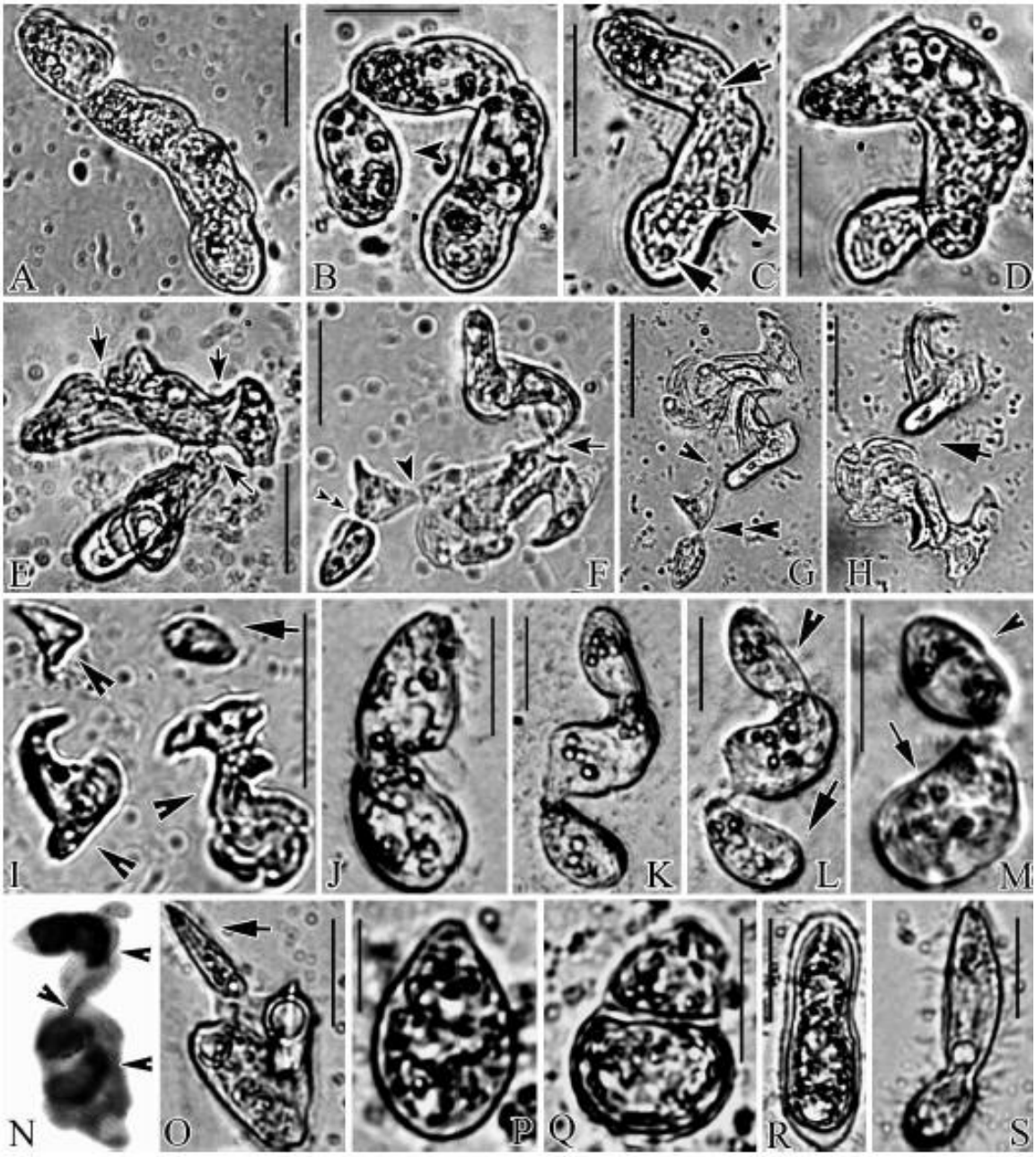

**Figure 4.2 Division processes of two *G. trihymene* asymmetric dividers in young cultures (A–I, J–M), other asymmetric dividers in young (N) and old cultures (O, S), and reproductive cysts (P–R).** A. One four-subcell asymmetric divider. B. The first asymmetric division. Arrowhead marks the trophont to be released. C-E. The new asymmetric divider gradually became highly deformed and many cleavage furrows appeared (arrows in E). Note the three contractile vacuoles in C (arrows). F. The arrowhead, double-arrowheads and arrow show the sites of the second, third and fourth cleavage furrows respectively. G. The second asymmetric division is completed at the arrowhead. The double arrowheads show the furrow that will shortly be broken in the third asymmetric division. H. The trophont resulting from the completion of the third asymmetric division has swum out of the field of view. The fourth asymmetric division has just been completed near the arrow, at a site corresponding to the furrow indicated by the arrow in F. I. Three new asymmetric dividers (arrowheads) and one trophont (arrow) were present by the end of the fourth asymmetric division. J. One two-subcell asymmetric divider. K, L. After elongation, the first asymmetric division produced one trophont (arrow in L) and one asymmetric divider (arrowhead in L). M. The second asymmetric division, producing one trophont (arrowhead) and another asymmetric divider (arrow). N. Arrowheads mark oral apparatuses (after protargol). O. One asymmetric divider releasing a tomite (arrow). P, Q. The division process of reproductive cysts. R. Another asymmetric divider forming a cyst wall. S. An asymmetric divider resembling a dividing tomite. Scale bars: A–H: 50 μm; I: 100 μm; J–M, O–S: 25 μm.).

The first division of one long asymmetric divider (Fig. 4.2A) occurred about two hours after it was found. During this first division, the cell's most anterior part was released (the anterior and posterior ends were judged from the moving direction and posterior position of the contractile vacuoles) as a trophont and quickly swam away (Fig. 4.2B, arrowhead). The larger posterior part became a new asymmetric divider (Fig. 4.2C), which then deformed so much that no clear body axis could be determined (Fig. 4.2D, E). The division types (transverse or longitudinal) were thus not easily categorized and many cleavage furrows appeared (Fig. 4.2E, arrows). The second asymmetric division occurred through disjuncture or fission at the most mature cleavage furrow (Fig. 4.2F, G, arrowheads). Then after about three minutes, the other two furrows broke (Fig. 4.2F-H, double-arrowheads, arrows). And finally three new asymmetric dividers, which were also slowly moving or immobile and continued dividing highly unequally (Fig. 4.2I, arrowheads), and one trophont (Fig. 4.2I, arrow) were produced. The entire process described above occurred over the course of 22 hours.

The most common asymmetric dividers in young cultures had two subcells (Fig. 4.2J), which divided over the course of 6 hours. The division process (Fig. 4.2K–M) was similar to the one described above in that the first division yielded one active trophont (Fig. 4.2L, arrow) and one new asymmetric divider (Fig. 4.2L, arrowhead). After that, however, the newly formed asymmetric divider divided into one trophont (Fig. 4.2M, arrowhead) and one new asymmetric divider (Fig. 4.2M, arrow), which became deformed and continued dividing highly unequally. During each division, the asymmetric dividers either produced one trophont and one new asymmetric divider (as shown in Fig. 4.2B, L, M) or two new asymmetric dividers (Fig. 4.2G, H).

### *4.2.3 Asymmetric dividers and reproductive cysts in old cultures*

When bacteria were depleted, most trophonts transformed into tomites and the cultures were termed "old". In the soil extract medium with various bacteria concentrations, this usually occurred between 141 and 175 hours after inoculation (Table 4.1). In old cultures, asymmetric division continued, but produced tomites instead of trophonts (Fig. 4.2O, arrow). Small asymmetric dividers producing tomites sometimes looked like dividing tomites (Fig. 4.2S). Some asymmetric dividers were also found to die and were observed with a large central vacuole. Reproductive cysts were also found: some asymmetric dividers developed transparent cyst walls and continued to divide unequally one or two times inside the cyst walls (Fig. 4.2P–R).

**Table 4.1 Average first appearance time of tomites in three different concentrations of bacteria in soil extract medium (four replicates for each concentration).**

| Bacterial concentrations of cultures | Tomite first appearance time (hours after inoculation) |
|---|---|
| 0.01X | 141.5 |
| 0.1X | 168.1 |
| 1X | 174.9 |

### *4.2.4 Is asymmetric division a cultural artifact?*

Actively dividing asymmetric dividers were found in all wheat grain medium cultures and cultures with bacterial suspensions in the soil extract medium, as well as cultures started with single cells as inocula. Even though the seawater for cultures was changed twice (natural seawater from coastal areas of Galveston TX, USA), asymmetric dividers were found in all cultures under study. Asymmetric dividers also showed up in early cultures of another seven *G. trihymene* isolates collected from coastal areas of Galveston, TX, USA (Table 4.2). The regularity with which asymmetric dividers appear and their consistent response to bacterial concentrations (Table 4.1) suggest that these asymmetric dividers are not cultural artifacts. The phylogenetic relationship based on the ITS marker between isolates was also shown in Fig. 4.3.

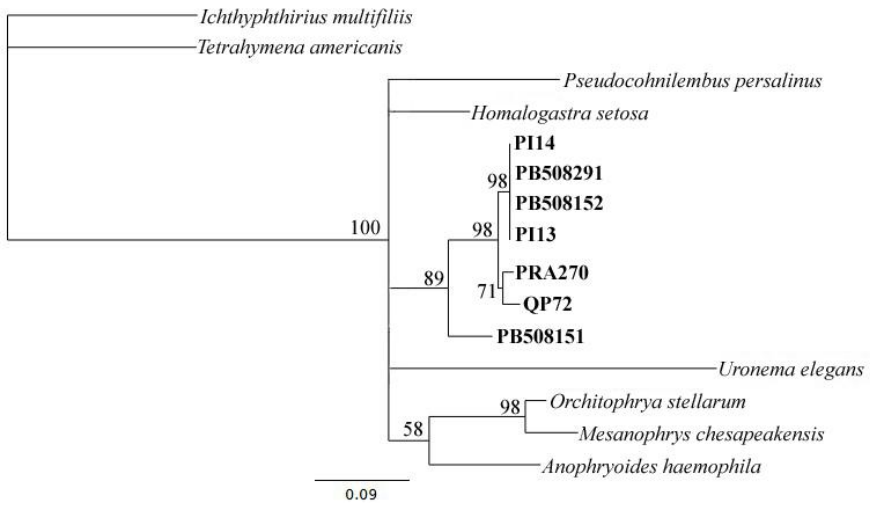


**Figure 4.3 Phylogenetic relationship of seven *G. trihymene* isolates (in bold) and eight other oligohymenophorean ciliates.** Bootstrap values (in %) are shown above selected nodes (with bootstrap support >50%). Scale bar: 9 substitutions per 100 nucleotide positions. PI14: isolate PI108294; PI13: isolate PI108293.

### *4.2.5 Somatic and nuclear characteristics of asymmetric dividers after protargol impregnation*

Some asymmetric dividers had similar body shape to trophonts, except having two highly unequal macronuclei (Fig. 4.1D). Macronuclear divisions could also happen several times before the completion of cytokinesis, producing up to 4 macronuclei in the same cytoplasm (Fig. 4.1H). The positioning of macronuclei was highly variable even if the cleavage furrows were clearly formed (Figs 4.1G, H; 4.2N). Usually more than two buccal apparatuses were present in bigger asymmetric dividers (Fig. 4.2N, arrowheads).

### *4.2.6 Relationship between asymmetric dividers and food abundance*

All asymmetric dividers first appeared on the 3$^{rd}$ to 4$^{th}$ day (51-93 hours) (Fig. 4.4, hollow bars) after inoculation of tomites into three bacterial concentrations. The earliest

asymmetric dividers appeared in the cultures with the highest bacterial concentration ($P<0.05$, One-way ANOVA; Fig. 4.4, hollow bar B), on average 54 hours after inoculation. There was no significant difference between the time of first appearance of asymmetric dividers in the other cultures ($P>0.05$, One-way ANOVA; Fig. 4.4, hollow bars A).

After the first asymmetric dividers appeared in each culture, they were checked every 12 hours until no asymmetric dividers remained. The time interval between first appearance of asymmetric dividers and the time when no asymmetric divider could be found was recorded for each culture (Fig. 4.4, filled bars). The time during which no asymmetric divider could be found was probably the stationary phase, when cells had run out of food so that they could not divide at all. This time interval, reflecting the total time of asymmetric divisions in each culture, was found to increase with bacterial concentration (Fig. 4.4, filled bars, a-c; One-way ANOVA, $P<0.05$).

### *4.2.7 Phylogenetic position of* G. trihymene

Maximum likelihood, maximum parsimony and Bayesian trees, inferred from 18S SSU rDNA sequences, all show that *G. trihymene* (Hong Kong isolate) groups with typical scuticociliates, like *Anophryoides haemophila* and *Miamiensis avidus* (Fig. 4.5). The Hong Kong isolate shares 81.2% DNA pair-wise identity with a previously submitted *G. trihymene* sequence [GenBank Accession No.: AY169274].

**Table 4.2 Collection efforts of *G. trihymene,* based on previous and current studies.**

| Isolates | Collection sites & Time | Temperature & Habitats | Data Sources |
|---|---|---|---|
| *G. trihymene* | Cedar Island VA, US, - | - | (THOMPSON 1966) |
| *U. tortum** | Mie Prefecture, Japan, 09/1995 | -, sea lettuce | (PEREZ-UZ and GUINEA 2001) |
| *G. trihymene* | Qingdao, China, 09/2000 | -, coastal water | (MA *et al.* 2006) |
| PRA-270 | Hong Kong, 08/20/2007 | 26 ℃, Rinsing/crab | (LONG and ZUFALL 2010) |
| PB508151 | Port Bolivar, TX US, 08/15/2009 | 31 ℃, sea lettuce | (LONG and ZUFALL 2010) |
| PB508152 | Port Bolivar, TX US, 08/15/2009 | 31 ℃, sea lettuce | (LONG and ZUFALL 2010) |
| PB508291 | Port Bolivar, TX US, 08/29/2009 | 31.2 ℃, sea lettuce | Current |
| PB508292 | Port Bolivar, TX US, 08/29/2009 | 31.2 ℃, sea lettuce | Current |
| PB508293 | Port Bolivar, TX US, 08/29/2009 | 31.2 ℃, sea lettuce | (LONG and ZUFALL 2010) |
| PB508294 | Port Bolivar, TX US, 08/29/2009 | 31.2 ℃, sea lettuce | Current |
| PB508295 | Port Bolivar, TX US, 08/29/2009 | 31.2 ℃, sea lettuce | Current |
| PI108291 | Pelican Island, TX US, 08/29/2009 | 31 ℃, sea lettuce | Current |
| PI108292 | Pelican Island, TX US, 08/29/2009 | 31 ℃, sea lettuce | Current |
| PI108293 | Pelican Island, TX US, 08/29/2009 | 31 ℃, sea lettuce | (LONG and ZUFALL 2010) |
| PI108294 | Pelican Island, TX US, 08/29/2009 | 31 ℃, sea lettuce | (LONG and ZUFALL 2010) |
| PI408291 | Pelican Island, TX US, 08/29/2009 | 31 ℃, sea lettuce | Current |
| PI408292 | Pelican Island, TX US, 08/29/2009 | 31 ℃, sea lettuce | Current |
| PI408293 | Pelican Island, TX US, 08/29/2009 | 31 ℃, sea lettuce | Current |
| PI508291 | Pelican Island, TX US, 08/29/2009 | 30.8 ℃, sea lettuce | Current |
| PI508291 | Pelican Island, TX US, 08/29/2009 | 30.8 ℃, sea lettuce | Current |
| PI608291 | Pelican Island, TX US, 08/29/2009 | 30.8 ℃, sea lettuce | (LONG and ZUFALL 2010) |
| QP72 | Free Port, TX US, 10/24/2009 | 23 ℃, sea lettuce | Current |
| QP76 | Free Port, TX US, 10/24/2009 | 23 ℃, sea lettuce | (LONG and ZUFALL 2010) |
| † | Vancouver, BC Canada, 09/15/2009 | 15 ℃, sea lettuce | Current |
| † | Mt Pearl, NF Canada, 9/29/2009 | 13 ℃, sea lettuce | Current |

*Synonym of *G. trihymene*

† Failed collection endeavors

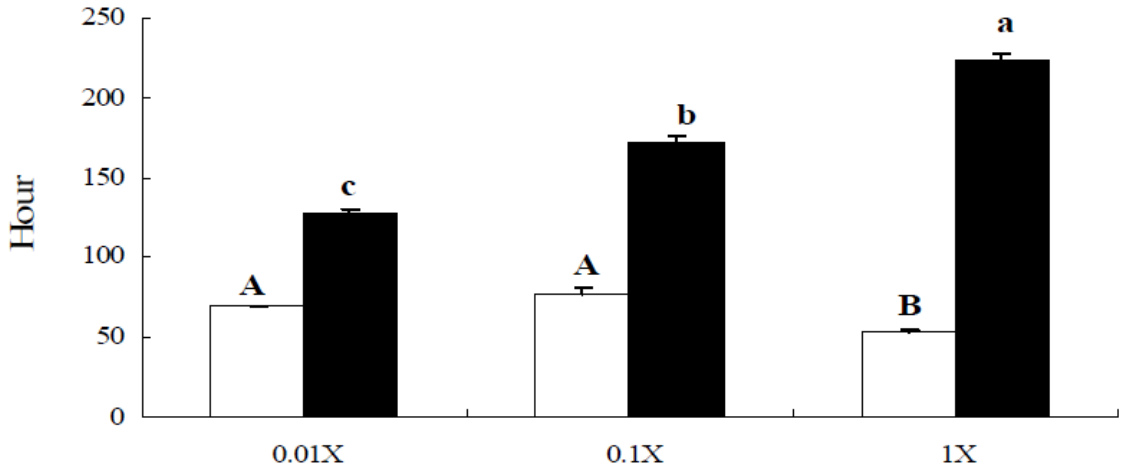


**Figure 4.4** **First appearance time and duration of persistence of asymmetric divisions.** The time of appearance of the first asymmetric divider in the newly inoculated cultures (hollow bars) and the duration of persistence of asymmetric divisions after the appearance of the first asymmetric divider (filled bars) were noted for cells maintained in the Erd-Schreiber soil extract cultures with one of three different bacterial concentrations. Appearance time of first asymmetric dividers and persistence time of asymmetric divisions were analyzed independently. Error bars: standard error. Levels not connected by the same letter are significantly different ($P < 0.05$).

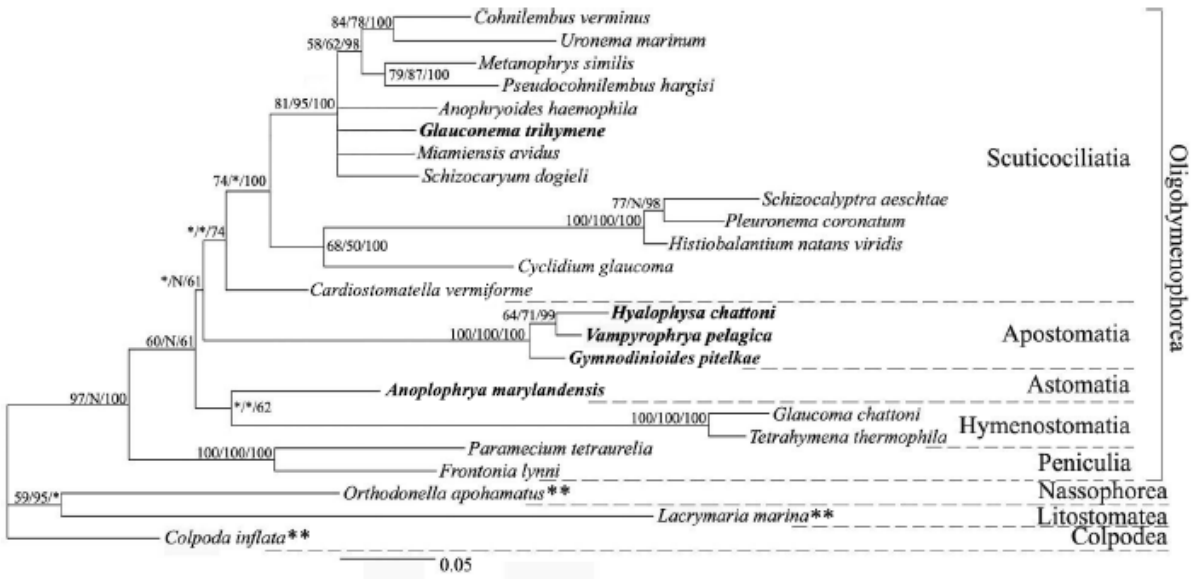


**Figure 4.5** **Phylogenetic position of *G. trihymene*.** Maximum likelihood tree topology and branch lengths, rooted with species marked with **. Support for clades is indicated by ML bootstrap/MP bootstrap/MB posterior probabilities. N indicates that this clade was not found in the given analysis and asterisks indicate clades with support of less than 50%. Nodes with <50% support in all methods are shown as a polytomy. Scale bar: 5 substitutions per 100 nucleotide positions.

## 4.3 Discussion

*The updated life cycle of* G. trihymene *during vegetative growth*

The life cycle during vegetative growth of *G. trihymene* is generalized in Fig. 4.6, based on previous and current studies (MA *et al.* 2006) (THOMPSON 1966). The life cycle has multiple stages, as is typical in polyphenic ciliates. These life stages could be highly diverse and complex, depending on the total number of asymmetric divider morphotypes and food concentration. For simplification and clarity, most intermediate asymmetric dividers are not shown in Fig. 4.6.

Some free-living ciliates, for example, *Tetrahymena pyriformis*, produce maximal progeny cells by shifting their physiological states during starvation (CAMERON 1973). Similarly, *G. trihymene* produces progeny cells by combining three reproductive modes: asymmetric division, reproductive cysts and equal fission. In addition, this is the first report of reproductive cysts in scuticociliates, though they are not uncommonly found in certain ciliate genera, like *Colpoda* and *Tetrahymena* (CORLISS 2001). If each morphotype of asymmetric dividers could be deemed as one life stage, which could probably be the case as many similar or continuous morphotypes were repeatedly found in cultures with different "age" or media (Fig. 4.1G, H); then the updated life cycle of *G. trihymene* might rival most known life cycles of free-living ciliates in complexity (Fig. 4.6). *G. trihymene* thus provides a special opportunity for studying ciliate polyphenism.

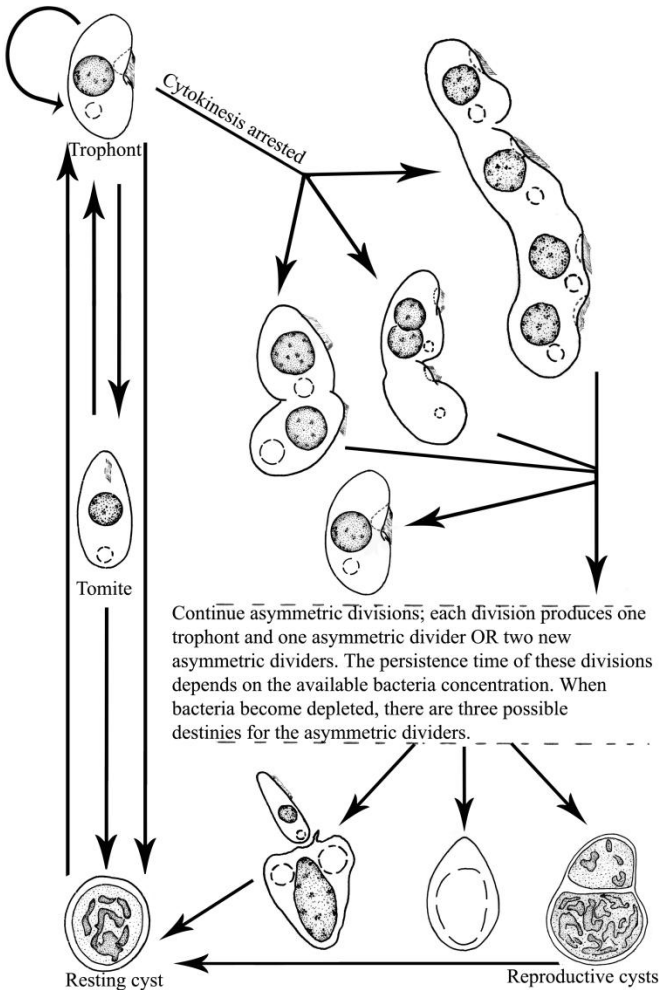


**Figure 4.6 Updated life cycle of *G. trihymene* in vegetative growth.** This is generalized from continuous microscopy and observation of specimens after protargol impregnation. Note the first asymmetric dividers (probably more than three morphotypes) with different sizes and shapes in early cultures developed through the arrest of cytokinesis in some trophonts. Drawings are not strictly to scale. Information on micronuclei is not available.

Although *G. trihymene* was first discovered early in 1966, it was believed to reproduce only by equal fission, during vegetative growth (MA *et al.* 2006; THOMPSON 1966). One reason for the persistence of this narrow view of *G. trihymene* reproduction is that, to date, few studies have been conducted on *G. trihymene* and they have mainly focused on morphology or systematics rather than reproduction dynamics (MA *et al.* 2006)

(THOMPSON 1966). Secondly, some of the reproduction forms appear only under particular food conditions, for example, in the Hong Kong isolate, asymmetric dividers appeared on the 3rd or 4th day after inoculation, when bacterial supply was high and disappeared soon after the appearance of tomites. The disappearance of asymmetric dividers was probably associated with the transition from exponential culture growth to the stationary phase. Third, the relative immobility and irregular body shapes of most asymmetric dividers (Figs 4.1G, H; 4.2E, N), could cause them to be mistaken as cultural artifacts or debris. Lastly, some asymmetric dividers are easily mistaken as conjugating cells or equal binary dividers, if observed on low magnifications (<100×) (Fig. 4.2J). Thus, it is no wonder that these usually large, irregularly shaped asymmetric dividers were unreported until this study.

The class Oligohymenophorea, to which all scuticociliates and the well-known *Tetrahymena* and *Paramecium* belong, contains highly diverse species (LYNN 1996), but only a few model species, such as *Tetrahymena thermophila* and *Paramecium tetraurelia*, are under intensive biological study. Most members of Oligohymenophorea, especially the marine species, are limited to taxonomic and systematic studies or are un-described (FOISSNER *et al.* 2008; YI *et al.* 2009). I predict that as life histories of more species are closely examined, much more diversity in reproductive strategies will be discovered among free-living protists.

*Proposed ecological roles of various life cycle stages*

The high feeding efficiency, slow movement, and arrested cytokinesis observed in *G. trihymene* asymmetric dividers might be advantageous. Based on the results of our culturing experiments, I conclude that asymmetric dividers are innate physiological states of *G. trihymene*, which can be induced to occur in bacteria-sufficient media. Cells with asymmetric divisions may ingest more food than those without-most asymmetric dividers had many oral apparatuses with oral membranes beating quickly. They may be able to consume as many bacteria as several trophonts in the same period of time (Fig. 4.2N, arrowheads). In addition, the relative immobility of these asymmetric dividers may minimize their energy consumption (FENCHEL 1987). The arrested cytokinesis should also save energy for asymmetric dividers, compared with equal dividers.

I propose the following ecological scenario that comes about as *G. trihymene* with a capacity for asymmetric divisions explores its surrounding environment. Suppose one *G. trihymene* trophont finds a food patch with plenty of bacteria, but also with many other bacteria-feeding protists. To avoid being a loser in this resource exploitation competition, for 2-3 days *G. trihymene* vigorously feeds on bacteria and divides equally. While plenty of bacteria remain, some trophonts asymmetrically divide, producing trophonts and more asymmetric dividers. When the food patch is nearly exhausted, most trophonts transform into tomites, and the asymmetric dividers instead of producing trophonts, produce tomites. After most of the bacteria are consumed, most tomites become resting cysts. Asymmetric dividers secrete a cyst wall and continue dividing inside, producing reproductive cysts, which ultimately become resting cysts. Some tomites transformed from trophonts or

released by asymmetric dividers swim rapidly to seek more food patches, transforming back into trophonts when they find new food patches and repeating the above processes. The quickly dispersing tomites, the tolerating resting cysts, and the diverse reproductive strategy may enable *G. trihymene* to identify and dominate enough food patches and survive in the coastal water community.

*Frequent collection of 'rare' species*

Along with all previous collection information, it is clear that *G. trihymene* has a world-wide distribution and preferably inhabits surfaces of sea lettuces (*Ulva* spp.) in warm seawater (≥23 ºC) (Table 3.2). Prior to this study, I suspected that *G. trihymene* was a rare species, given that there had only been a few reports of it in the past 44 years (LONG and ZUFALL 2010; MA *et al.* 2006; PEREZ-UZ and GUINEA 2001; THOMPSON 1966); however, my frequent collection success suggests that this "rare species" may be far more abundant than would have been predicted based on traditional sampling protocols. Sampling methods relying on artificial enrichment substrates, like glass slides in frames and PFU (polyurethane foam unit), neglect most habitat information (YONGUE and CAIRNS 1973). Likewise, environmental sampling for microbial metagenomic studies often relies on samples from the water column, and does not include all possible habitat niches, such as plant or animal surfaces. My sampling results suggest that in order to understand microbial distribution and abundance at a locale, specific habitat information is crucial. As more and more habitat information is gathered, these habitats will eventually become another stock center for ciliates, especially those that cannot currently

be cultured or cryopreserved. Further, many "rare species" may become common species in certain habitats, such as *G. trihymene*.

*Phylogenetic position of* G. trihymene*, and asymmetric division*

*G. trihymene* groups with typical scuticociliates with high bootstrap support and posterior probability, though the precise relationships within the clades remain unresolved (Fig. 4.5). In addition, *G. trihymene* has high SSU rDNA pair-wise identity with *Anophryoides haemophila* (96%), the scuticociliate causing the "Bumper car disease" of American lobsters and *Miamiensis avidus* (96%), a polyphenic, parasitic ciliate, which causes diseases in fish (CAWTHORN *et al.* 1996; GOMEZ-SALADIN and SMALL 1993). My result supports the monophyly of scuticociliatia, despite what was found in earlier studies utilizing a previously reported *G. trihymene* SSU rDNA sequence [GenBank Accession No.: AY169274] (DUNTHORN *et al.* 2008; UTZ and EIZIRIK 2007) which I believe to be erroneous. In line with my interpretation, the most recent study on morphology and morphogenesis of *G. trihymene* (performed by the same group that submitted the previous *G. trihymene* SSU rDNA sequence) showed that it is indeed a typical scuticociliate (MA *et al.* 2006).

Asymmetric divisions, similar to those in *G. trihymene*, occur in certain apostomic and many astomic ciliates (see phylogenetic position in Fig. 4.5), though the details of division had never been studied using continuous microscopy (CHATTON and LWOFF 1935). Such asymmetric dividers were called catenoid colonies in these host-dependent ciliates. Asymmetric dividers were so named in the present study to emphasize

the difference between the two filial cells. As in the asymmetric division of *G. trihymene* in Fig. 4.2A, long cell chains in the parasitic and commensal astome and apostome ciliates are formed by repeated incomplete divisions without separation of the resulting filial products, after which some subcells are fully or partially pinched off. These subcells require subsequent metamorphosis to regain the form typical of the normal trophont stage of the life cycle (CHATTON and LWOFF 1935; LYNN 2008).

The results of the phylogenetic analysis suggest that complex life cycles including asymmetric division are either 1) an ancestral feature of these three groups that has been modified, lost, or not yet discovered in other free-living species, or 2) a convergent trait that has arisen multiple times independently in these closely related taxa.

I found no obvious morphological differences between any of the isolates during asexual growth and division; all isolates have similar lifecycles, as described in (LONG and ZUFALL 2010); however, unlike in other isolates, conjugation occurred frequently in repeated single-cell isolations of the PB508151 lineage, which has the greatest genetic distance from the other isolates (DNA pairwise identity at ITS locus from 85.8% to 87.1%; Table 4.2). Such selfing was observed frequently in 2-3 day old cultures (Fig. 4.7). No selfing was observed in cultures of other isolates, nor could it be induced by starvation (a common cue for conjugation in ciliates). There has been no previous report of selfing of *G. trihymene*.

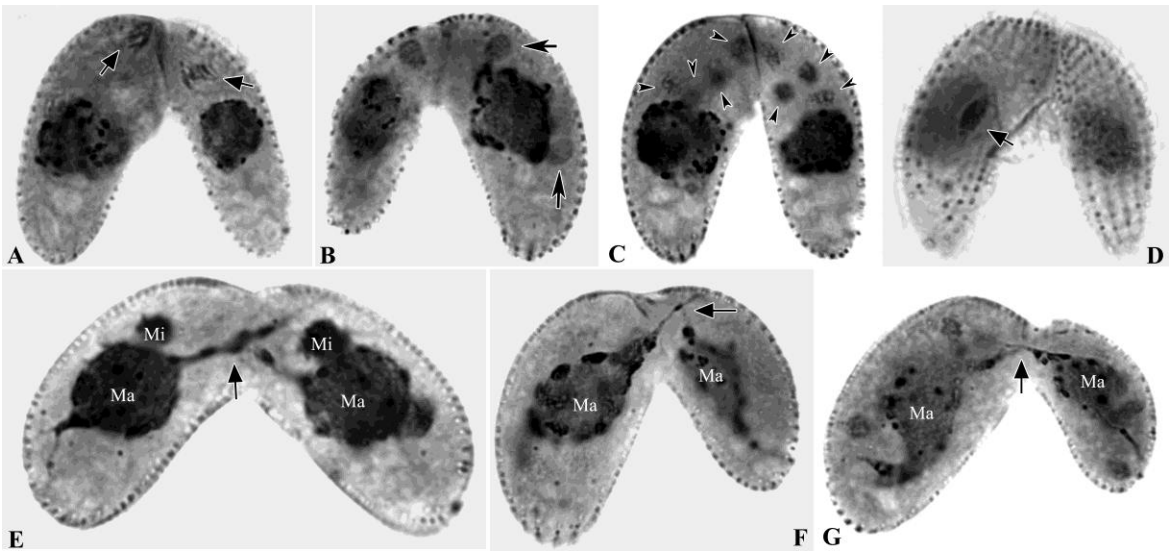


**Figure 4.7 Selfing pairs after protargol staining from *G. trihymene* isolate PB508151. (A)** Condensation of micronuclear chromatin (arrows). **(B)** Completion of meiosis I, arrows mark the micronuclei. **(C)** Completion of meiosis II, arrowheads indicate the pronuclei. **(D)** Mitosis of the micronucleus, arrow marks the condensed chromosomes in metaphase. **(E-G)** Macronuclear contact between two conjugants, arrows mark the contacting sites. Ma-macronucleus; Mi-micronucleus.

*Nuclear events during selfing*

Nuclear events were observed over the course of conjugation in this self-compatible strain using protargol staining. The only selfing strain PB508151 diverged from all other isolates with high bootstrap support (Fig. 4.3). Some typical nuclear activities were observed in the selfing pairs of PB508151, including condensation of chromosomes (Fig. 4.7A, arrows) and meiosis I (Fig. 4.7B), II (Fig. 4.7C) and prezygotic mitosis (Fig. 4.7D) of the micronucleus; however, I also observed previously unreported nuclear processes: macronuclear streaming and direct contact between macronuclei of the conjugants (Fig. 4.7E-G, arrows). The macronucleus from each conjugant elongated and then came in direct contact with the elongated macronucleus of the other conjugant. This

usually occurred around the connected region of the two conjugants (Fig. 4.7E-G). Macronuclear contact lasted until the two conjugants separated. Because I was unable to induce conjugation between isolates, I cannot determine whether macronuclear contact is a general property of conjugation in *G. trihymeme*, or is unique to selfing. Whether there is any macronuclear genetic exchange between the two exconjugants or selfing could be involved in sympatric speciation (Fig. 4.3) needs further exploration.

*Asymmetric division: one clue to multicellularity?*

The colonial flagellate hypothesis, claiming that flagellated protists living as colonies evolved into the first animals, has inspired extensive productive exploration on the origin of multicellularity (CARR *et al.* 2008; HYMAN 1940; KING *et al.* 2008; ROKAS 2008). The asymmetric division of *G. trihymene* serves as an alternative mechanism through which protists may have led to a multicellular form: a multicellular form could arise by a ciliate with one single-macronucleus and micronucleus subdividing itself as a result of growth followed by arrested cytokinesis. It should be noted, however, that such asymmetric division does not result in different developmental fates akin to truly multicellular ciliate species, such as *Zoothamnium alternans* (FAURÉ-FREMIET 1930; SUMMERS 1938).

As is shown in this study, asymmetric dividers produce new asymmetric dividers and trophonts by successive asymmetric divisions, in favorable conditions, and the more available food, the longer the asymmetric divisions persisted (Fig. 4.4, filled bars). If asymmetric dividers lived in consistently bacteria-rich environments for a long time, they might retain the multicellular form, but lose the ability to produce trophonts or tomites.

Bacteria-rich environments were common in the ancient ocean, which had very different chemistry from that of today's (CROWE *et al.* 2008; ZERKLE *et al.* 2005). Thus, it is possible that some multicellular organisms, which have not yet been discovered or have since gone extinct, originated from certain asymmetric dividers of ciliates.

To test whether the asymmetric dividers are adaptive, I tried to find out if they are advantageous in fitness over the normal binary dividers. The fitness metric I measured was doubling time, which was the time for one single asymmetric divider or one normal binary divider to produce one progeny cell, using direct time-lapse microscopy; however, the asymmetric dividers budded off not only asymmetric dividers, but also normal cells, which later underwent normal binary division; thus, such preliminary fitness measurement is not accurate for finding whether fitness advantage of asymmetric dividers and needs improvement. If future accurate fitness measurement shows there is no fitness difference between asymmetric and normal dividers, considering the low frequency of the asymmetric dividers in newly collected isolates, the possibility that asymmetric dividers are just one transient physiological state could not be excluded.

## 4.4 Materials and Methods

*Sampling and identifying* G. trihymene

*G. trihymene* was isolated with a fine pipette from a seawater rinse of a newly dead crab (species unknown) collected from a sand beach near the pier of Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong (22° 20' N; 114° 17' E) on August 20, 2007. The salinity was about 33‰, temperature 26 ℃, and pH 8.1. The

cultures used in this study were derived from a single *G. trihymene* cell of the Hong Kong isolate. Seven other isolates were collected from Texas coastal areas (Table 4.2). The salinity was about 33‰ and temperature ranged from 23 to 31 ℃. Trophonts and tomites of *G. trihymene* were observed *in vivo* first using a stereomicroscope and then an epi-fluorescence microscope at 100-1000X. The nuclear apparatuses and infraciliature were revealed by the protargol impregnation method (WILBERT 1975). The protargol S™ was manufactured by Polysciences Inc., Warrington, PA (Cat No.: 01070). Drawings were based on free-hand sketches. One subculture of the isolate in this study was deposited in ATCC (American Type Culture Collection; Reg. No.: PRA-270).

*Monitoring individual asymmetric dividers with continuous microscopy*

For continuous microscopy of *G. trihymene* reproductions, 50 cultures were established in wheat grain medium (100×15mm plastic Petri dishes each with 3 autoclaved wheat grains in 30ml autoclaved seawater, 0.2g/grain, and with about 50 tomites in 100μl stock culture medium as inoculum). The salinity was about 31‰, pH 8.0. All cultures were maintained at room temperature, about 23 ℃. Most asymmetric dividers, which were first observed under a stereomicroscope, were immobile or slowly moving on bottoms of Petri dishes, and their position was marked on the Petri dish bottom. The asymmetric dividers were then observed and followed under an inverted microscope (100–400X; Olympus IX71). To minimize disturbance to asymmetric dividers during continuous multi-day observation, low light intensity and low magnification were used. Asymmetric dividers from 3–7 day-old cultures were continuously isolated with fine pipettes and impregnated

with protargol, in order to check the nuclei and infraciliature characters during asymmetric divisions.

*Effect of bacterial concentration on asymmetric division*

The Erd-Schreiber soil extract medium added with bacterial suspension has recently been shown to be efficient for culturing *G. trihymene* (PEREZ-UZ and GUINEA 2001; TOMPKINS *et al.* 1995) (I believe *Urocryptum tortum* in PEREZ-UZ and GUINEA 2001 is a junior synonym of *G. trihymene*, because of their great similarity in living morphology, infraciliature, and habitat, as well as the life cycle characteristics). To prepare bacterial suspension, 10μl stock culture medium without cells was inoculated into 3ml autoclaved seawater LB medium in test tubes (seawater LB recipe: 12.5g LB broth in 500ml autoclaved filtered natural seawater) and cultured at 30 ℃, 200rpm, overnight, to maximal growth. The bacteria were harvested by centrifugation at 7378g in 1.5ml eppendorf tubes with a microcentrifuge and the supernatant was removed. Then 1ml sterile Erd-Schreiber soil extract medium was added to wash the bacteria pellets, at 7378g. This washing procedure was repeated twice. Each pellet was finally re-suspended with 1ml soil extract medium and combined in a sterile 50ml polypropylene conical tube (BD Flacon$^{TM}$).

Bacterial suspensions of 3ml, 0.3ml, and 0.03ml were added separately into 3 Petri dishes with sterile soil extract medium to reach a final volume of 30ml (marked as 1X, 0.1X and 0.01X for each concentration, respectively). It should be noted that the Erd-Schreiber soil extract medium was not a rich medium supporting growth of a large number of bacteria. Four replicates were prepared for each concentration. After each

culture was inoculated with about 50 tomites in 100μl stock culture medium, all 12 cultures were placed on a rocking platform at 3rpm. Each culture was checked every 12 hours for asymmetric dividers, until 50 hours after the inoculation (preliminary experiments showed that the earliest appearance of asymmetric dividers occurred at $50^{th}$ hour after inoculation with tomites). After the $50^{th}$ hour, all cultures were checked for appearance of asymmetric dividers every two hours until they were first observed in each culture. The first appearance time of asymmetric dividers and tomites was recorded for each culture. Subsequently, all cultures were checked for the presence of asymmetric dividers every 12 hours, until all of them disappeared from each culture. The disappearance time of asymmetric dividers for each culture was also recorded.

*Amplifying, cloning, and sequencing of SSU rDNA*

Cells from the stock culture were harvested in one 1.5ml eppendorf tube with a micro-centrifuge, at 1844g. Supernatant was removed and the pellet was re-suspended with 20μl autoclaved seawater. The cell suspension was directly used as DNA template for amplifying the SSU rDNA. Universal eukaryotic primers for SSU rRNA were used: forward 5'-AACCTGGTTGATCCTGCCAGT-3', reverse 5'-TGATCCTTCTGCAGGTTCACCTAC-3' (MEDLIN *et al.* 1988). PCR programs were performed using the iProof™ High-Fidelity PCR kit (Bio-Rad, CA): 1 cycle (98 ℃, 2min); 30 cycles (98 ℃, 10s; 70 ℃, 30s; 72 ℃, 50s); 1 cycle (72 ℃, 7min). The PCR products were then purified with the QIAquick gel extraction kit (QIAGEN Sciences, MD) and cloned with the Zero Blunt TOPO kit (Invitrogen, CA). The plasmid DNA was isolated from transformant colonies using the QIAprep spin miniprep kit (Qiagen, CA) and four

clones were sequenced with the BigDye terminator kit (Applied Biosystems, CA) on an automated ABI 3130 XL sequencer in the Department of Microbiology and Molecular Genetics, University of Texas Health Sciences Center at Houston.

*SSU rDNA sequence availability and phylogenetic tree reconstruction*

The SSU rDNA sequence of *G. trihymene* was deposited in GenBank [GenBank: GQ214552]. The accession numbers of the additional SSU rDNA sequences used in this study were as follows: *Anophryoides haemophila* [GenBank:U51554], *Anoplophrya marylandensis* [GenBank:AY547546], *Cardiostomatella vermiforme* [GenBank:AY881632], *Cohnilembus verminus* [GenBank:Z22878], *Colpoda inflata* [GenBank:M97908], *Cyclidium glaucoma* [GenBank:EU032356], *Entorhipidium pilatum* [GenBank:AY541689], *Gymnodinioides pitelkae* [GenBank:EU503534], *Histiobalantium natans viridis* [GenBank:AB450957], *Hyalophysa chattoni* [GenBank:EU503536], *Metanophrys similes* [GenBank:AY314803], *Miamiensis avidus* [GenBank:AY550080], *Pleuronema coronatum* [GenBank:AY103188], *Pseudocohnilembus hargisi* [GenBank:AY833087], *Schizocalyptra aeschtae* [GenBank:DQ777744], *Schizocaryum dogieli* [GenBank:AF527756], *Uronema marinum* [GenBank:AY551905], *Vampyrophrya pelagica* [GenBank:EU503539].

*Amplifying, cloning, and sequencing of ITS sequence*

Cells from each isolate were directly used as DNA templates for amplifying the ITS1-5.8S rDNA-ITS2 locus, using the degenerative primers designed by (SNOEYENBOS-WEST *et al.* 2004). PCR was performed using the iProof™ High-Fidelity PCR kit (Bio-Rad,

CA): 1 cycle (98 ℃, 2 min); 36 cycles (98 ℃, 10s; 63 ℃, 30s; 72 ℃, 50s); 1 cycle (72 ℃, 7min). The PCR products were then run on 2% low melting agarose gel for about 6 hours to separate the target products from contaminating bacterial products (usually about 20bp longer than the target products). Target PCR products were purified with the QIAquick gel extraction kit (QIAGEN Sciences, MD) and cloned with the Zero Blunt TOPO kit (Invitrogen, CA). The plasmid DNA was isolated from transformant colonies using the QIAprep spin miniprep kit (Qiagen, CA) and clones were then sequenced on an automated sequencer (ABI 3730xl) at Eton Bioscience Inc., Houston Texas, USA.

The ITS1-5.8S rDNA-ITS2 sequences of 7 *G. trihymene* isolates were deposited in GenBank (see accession numbers in Table 4.1). The GenBank accession numbers of the additional sequences used in this study were as follows: *Anophryoides haemophila* (AF107779), *Homalogastra setosa* (EF158845), *Ichthyphthirius multifiliis* (DQ270015), *Mesanophrys chesapeakensis* (AF107778), *Orchitophrya stellarum* (AF107776), *Pseudocohnilembus persalinus* (EU262622), *Tetrahymena americanis* (AY833381), *Uronema elegans* (AY513760).

Selfing in the isolate PB508151 usually occurred 2–3 days after cultures were established in the wheat grain medium. Nuclear apparatus activities were revealed by continuously isolating and staining selfing pairs with protargol after the first selfing pair was observed. Photomicrographs were then taken on a compound microscope at 600X.

Sequences were aligned in ClustalW (LARKIN *et al.* 2007) (executed as a plug-in in Geneious Pro 4.0.4 (DRUMMOND *et al.* 2008)) and adjusted by hand. 1707 nucleotides

(positions) were used in the analysis. Maximum likelihood (ML) and parsimony (MP) phylogenetic analyses were performed in PAUP* (SWOFFORD 2003) and Bayesian analyses (MB) in Mr. Bayes (RONQUIST and HUELSENBECK 2003) (both executed in Geneious Pro 4.0.4) using the best fit model as determined by ModelTest (POSADA and CRANDALL 1998) (GTR+I+G). Support was determined based on 100 bootstrap replicates (ML and MP) or the posterior probability after one million generations, with an initial 10% burn-in (MB).

*Statistical analysis*

One-way ANOVA analysis (Tukey HSD Test, $\alpha$=0.05, JMP 7 software package) was conducted to check the differences among first appearance time and persistence time of asymmetric dividers in cultures with three different concentrations of bacterial suspension (data was log-transformed into normal distribution).

## 4.5 References:


ADL, S. M., and J. D. BERGER, 1996 Commitment to division in ciliate cell cycles. J Eukaryot Microbiol **43:** 77–86.

CAMERON, I. L., 1973 Growth characteristics of *Tetrahymena*. pp. 199-226 in *Biology of Tetrahymena.*, edited by A. ELLIOTT. Dowden, Hutchinson & Ross Inc, Stroudsburg, Pennsylvania.

CARR, M., B. S. C. LEADBEATER, R. HASSAN, M. NELSON, and S. L. BALDAUF, 2008 Molecular phylogeny of choanoflagellates, the sister group of Metazoa. Proc Natl Acad Sci USA **105:** 16641–16646.

CAWTHORN, R. J., D. H. LYNN, B. DESPRES, R. MACMILLAN, R. MALONEY *et al.*, 1996 Description of *Anophryoides haemophila* n. sp. (Scuticociliatida: Orchitophryidae), a pathogen of American lobsters *Homarus americanus*. Dis Aquat Org **24:** 143–148.

CHANTANGSI, C., and B. S. LEANDER, 2010 Ultrastructure, life cycle and molecular phylogenetic position of a novel marine sand-dwelling Cercozoan: *Clautriavia biflagellata* n. sp. Protist **161:** 133–147.

CHATTON, É., and A. LWOFF, 1935 Les ciliés Apostomes. Aperçu historique et général; étude mongraphique des genres et des espèces. Arch Zool Exp Gén **77:** 1–453.

COHEN, J., and J. BEISSON, 1980 Genetic analysis of the relationship between the cell surface and the nuclei in *Paramecium tetraurelia*. Genetics **95:** 797–818.

CORLISS, J. O., 2001 Protozoan cysts and spores., pp. in *Encyclopedia of life sciences.* John Wiley & Sons, Ltd., Chichester, West Sussex.

CROWE, S. A., C. JONES, S. KATSEV, C. MAGEN, A. H. O'NEILL *et al.*, 2008 Photoferrotrophs thrive in an Archean Ocean analogue. Proc Natl Acad Sci USA **105:** 15938–15943.

DOLAN, J. R., and D. W. COATS, 2008 Physiological diversity in widely distributed microzooplankton: digestion in the ciliate *Euplotes vannus*., pp. 207–220 in *Microbial ecology research trends*, edited by T. VANDIJK. Nova Science Publishers, New York.

DRUMMOND, A. J., B. ASHTON, M. CHEUNG, J. HELED, M. KEARSE *et al.*, 2008 Geneious v4.0.

DUNTHORN, M., W. FOISSNER, and L. A. KATZ, 2008 Molecular phylogenetic analysis of class Colpodea (phylum Ciliophora) using broad taxon sampling. Mol Phylogenet Evol **46:** 316–327.

FAURÉ-FREMIET, E., 1930 Growth and differentiation of the colonies of *Zoothamnium alternans* (Clap. and Lachm.). Biol Bull **58:** 28–51.

FENCHEL, T., 1987 *Ecology of protozoa: the biology of free-living phagotrophic protists.* Science Tech Publishers, Madison, Wisconsin.

FENCHEL, T., 1990 Adaptive significance of polymorphic life cycles in protozoa: responses to starvation and refeeding in two species of marine ciliates. J Exp Mar Biol Ecol **136:** 159–177.

FOISSNER, W., 1996 Ontogenesis in ciliated protozoa, with emphasis on stomatogenesis., pp. 95-177 in *Ciliates, cells as organisms.*, edited by K. HAUSMANN, BRADBURY PC. Gustav Fischer Press, Stuttgart, Germany.

FOISSNER, W., A. CHAO, and L. A. KATZ, 2008 Diversity and geographic distribution of ciliates (Protista: Ciliophora). Biodivers Conserv **17:** 345–363.

FRANKEL, J., 1964 Morphogenesis and division in chains of *Tetrahymena pyriformis* GL. J Protozool **11:** 514–526.

FRANKEL, J., 1977 Mutations affecting cell division in *Tetrahymena pyriformis*, syngen 1. 2. Phenotypes of single and double homozygotes. Dev Biol **58:** 255–275.

GOMEZ-SALADIN, E., and E. B. SMALL, 1993 Prey-induced transformation of *Miamiensis avidus* strain Ma/2 by a soluble factor. J Eukaryot Microbiol **40:** 550–556.

HATZIS, C., F. SRIENC, and A. G. FREDRICKSON, 1994 Feeding heterogeneity in ciliate populations: effects of culture age and nutritional state. Biotechnol Bioeng **43:** 371–380.

HYMAN, L. H., 1940 *The invertebrates: protozoa through Ctenophora.* McGraw-Hill, New York.

JAWORSKA, J. S., T. G. HALLAM, and T. W. SCHULTZ, 1996 A community model of ciliate *Tetrahymena* and bacteria *E. coli*: part I. Individual-based models of *Tetrahymena* and *E. coli* populations. B Math Biol **58:** 247–264.

JOHNSON, M. D., D. OLDACH, C. F. DELWICHE, and D. K. STOECKER, 2007 Retention of transcriptionally active cryptophyte nuclei by the ciliate *Myrionecta rubra*. Nature **445:** 426–428.

KING, N., M. J. WESTBROOK, S. L. YOUNG, A. KUO, M. ABEDIN, *et al.*, 2008 The genome of the choanoflagellate *Monosiga brevicollis* and the origins of metazoan multicellularity. Nature **451:** 783–788.

LARKIN, M. A., G. BLACKSHIELDS, N. P. BROWN, R. CHENNA, P. A. MCGETTIGAN *et al.*, 2007 Clustal W and Clustal X version 2.0. Bioinformatics **23:** 2947-2948.

LONG, H., and R. ZUFALL, 2010 Diverse modes of reproduction in the marine free-living ciliate *Glauconema trihymene*. BMC Microbiol **10:** 108.

LYNN, D. H., 1975 The life cycle of the histophagous ciliate *Tetrahymena corlissi* Thompson, 1955. J Protozool **22:** 188–195.

LYNN, D. H., 1996 Systematics of ciliates., pp. 51–72 in *Ciliates, cells as organisms*, edited by K. HAUSMANN, BRADBURY PC. Gustav Fischer Press, Stuttgart, Germany.

LYNN, D. H., 2008 *The ciliated protozoa. Characterization, classification and guide to the literature*. Springer, New York.

LYNN, D. H., and J. B. TUCKER, 1976 Cell size and proportional distance assessment during determination of organelle position in the cortex of the ciliate *Tetrahymena*. J Cell Sci **21:** 35–46.

MA, H., W. SONG, A. WARREN, D. ROBERTS, J. GONG *et al.*, 2006 Redescription of the marine scuticociliate *Glauconema trihymene* Thompson, 1966 (Protozoa: Ciliophora): life cycle and stomatogenesis. Zootaxa **1296:** 1–17.

MEDLIN, L., H. J. ELWOOD, S. STICKEL, and M. L. SOGIN, 1988 The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. Gene **71:** 491–499.

ORIAS, E., 1976 Derivation of ciliate architecture from a simple flagellate: an evolutionary model. Am Microsc Soc **95:** 415–429.

PEREZ-UZ, B., and A. GUINEA, 2001 Morphology and infraciliature of a marine scuticociliate with a polymorphic life cycle: *Urocryptum tortum* n. gen., n. comb. J Eukaryot Microbiol **48:** 338–347.

POSADA, D., and K. CRANDALL, 1998 MODELTEST: testing the model of DNA substitution. Bioinformatics **14:** 817–818.

ROKAS, A., 2008 The origins of multicellularity and the early history of the genetic toolkit for animal development. Annu Rev Genet **42:** 235–251.

RONQUIST, F., and J. HUELSENBECK, 2003 MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19:** 1572–1574.

SNOEYENBOS-WEST, O., J. COLE, A. CAMPBELL, D. COATS and L. A. KATZ, 2004 Molecular phylogeny of Phyllopharyngean ciliates and their group I introns. Journal of Eukaryotic Microbiology **51:** 441-450.

SUMMERS, F. M., 1938 Some aspects of normal development in the colonial ciliate *Zoothamnium alternans*. Biol Bull **74:** 117–129.

SWOFFORD, D., 2003 *PAUP*. Phylogenetic analysis using parsimony (*and other methods).* Sinauer Associates, Sunderland, MA.

TAYLOR, F. J. R., D. J. BLACKBOURN, and J. BLACKBOURN, 1969 Ultrastructure of the chloroplasts and associated structures within the marine ciliate *Mesodinium rubrum* (Lohmann). Nature **224:** 819–821.

THAZHATH, R., C. LIU, and J. GAERTIG, 2002 Polyglycylation domain of β-tubulin maintains axonemal architecture and affects cytokinesis in *Tetrahymena*. Nat Cell Biol **4:** 256–259.

THOMPSON, J. C., 1966 *Glauconema trihymene* n. g., n. sp., a hymenostome ciliate from the Virginia coast. J Protozool **13:** 393–395.

TOMPKINS, J., M. M. DEVILLE, J. G. DAY, and M. F. TURNER, 1995 *Culture collection of algae and protozoa. Catalogue of strains.* The Culture Collection of Algae and Protozoa, Cumbria, UK.

UTZ, L. R. P., and E. EIZIRIK, 2007 Molecular phylogenetics of subclass Peritrichia (Ciliophora: Oligohymenophorea) based on expanded analysis of 18S rRNA sequences. J Eukaryot Microbiol **54:** 303–305.

WEISSE, T., and S. RAMMER, 2006 Pronounced ecophysiological clonal differences of two common freshwater ciliates, *Coleps spetai* (Prostomatida) and *Rimostrombidium lacustris* (Oligotrichida), challenge the morphospecies concept. J Plankton Res **28:** 55–63.

WILBERT, N., 1975 Eine verbesserte Technik der Protargolimprägnation für Ciliaten. Mikrokosmos **64:** 171–179.

YI, Z., W. SONG, J. GONG, A. WARREN, K. A. S. AL-RASHEID *et al.*, 2009 Phylogeny of six oligohymenophoreans (Protozoa, Ciliophora) inferred from small subunit rRNA gene sequences. Zool Scr **38:** 323–331.

YONGUE, J. W. H., and J. J. CAIRNS, 1973 Long term exposure of artificial substrates to colonization by protozoans. J Elisha Mitchell Sci Soc **81:** 115–119.

ZERKLE, A. L., C. H. HOUSE, and S. L. BRANTLEY, 2005 Biogeochemical signatures through time as inferred from whole microbial genomes. Am J Sci **305:** 467–502.

# Chapter 5 Discussion

"*The capacity to blunder slightly is the real marvel of DNA. Without this special attribute, we would still be anaerobic bacteria and there would be no music*" - Lewis Thomas 1979, *The Medusa and the Snail*

In previous chapters, I have addressed how to accumulate and estimate mutational parameters using the somatic and germline genomes of *T. thermophila*, by measuring the changes in fitness, how morphological traits change due to mutations, as well as the diverse modes of reproduction strategies in the marine free-living ciliate *G. trihymene*. In this chapter, I will summarize the previous chapters and discuss other potential uses of the currently established mutation accumulation system in exploring other evolutionary genetic topics, the application of mutation accumulation techniques in current morphological taxonomy, the need to explore life history strategies of free-living ciliates, as well as some potential problems with the studies in each chapter.

## 5.1 Findings and potential problems of the mutation accumulation system

### *5.1.1 Findings in mutation accumulation study using T. thermophila*

The dimorphic nuclear system is the unique feature of ciliates among eukaryotes, which provides opportunity for exploring evolutionary topics in an idiosyncratic way that could not be easily realized in most other model organisms. For example, selection against

deleterious mutations is unavoidable in counterpart unicellular organisms with only one nuclear genome or it takes too long to accumulate mutations in multicellular organisms. In chapter 2, I showed that the mutation accumulation system using the model ciliate *T. thermophila* worked successfully. This species is the only ciliate growing in sterile well-defined medium and has the most diverse genetic tools. The diverse fitness metrics, like maximum population growth rate, lag-time, viability, as well as the statistical and genetic tools used, all demonstrate the power of this system in studying broad aspects of evolutionary genetics, even lethal mutation rate and dominance coefficient of new mutations, which are important but rarely available. In the future, this system could also be used to estimate directional epistasis, *viz*. the interaction between new mutations, by manipulating the germline genomes through genomic exclusion and backcrossing (DEVISSER *et al.* 1997; WEST *et al.* 1998).

### *5.1.2 Potential problems in the mutation accumulation system using T. thermophila*

There are actually also some potential problems with the mutation accumulation system currently established. First of all, as most mutation accumulation experiments, there are some limitations: the sensitivity of the fitness assay may cause small effects mutations undetected (KEIGHTLEY and EYRE-WALKER 1999); thus, causing the mutational estimates biased to large effects mutations; using competition assays might improve on this. Also, the true distribution of fitness effects (DFE) in mutation accumulation experiments is always difficult to infer from mutation accumulation experiments, because the fitness effects and the genomic mutation rate have usually un-

avoidable sampling covariance. Disentangling the DFE needs usually very high sample size and complicated algorithms (EYRE-WALKER and KEIGHTLEY 2007).

Second, epigenetic influence of soma mutations on germline genome expressed into the somatic genomes in the genomic exclusion progeny. In *T. thermophila*, an epigenetic process is regulating sequence elimination during development from the zygotic nucleus to a new somatic nucleus. Small RNAs "compare" the parental somatic genome to the developing one and signal the elimination of any sequences that are not present in the parent (reviewed in CHALKER 2008; DUHARCOURT *et al.* 2009). Thus, though the germline genome is not affected by somatic mutations, how the germline genome gets expressed after development is indeed affected. Particularly, deletion mutations in a somatic genome may result in subsequent elimination of that deleted sequence from the new somatic genome after conjugation. The experimental design for both measuring germline fitness and sequencing the germline genome (see below), would be confounded by any such events. Although we will not be able to correct such effects in the fitness assays, they will be revealed in the sequence data as a shared deletion in somatic and germline genomes from a single MA line. In any such case, the germline deletion could be detected as developmental by sequencing developmentally eliminated sequences from the germline.

Third, preliminary SNP analysis (by Reed Cartwright) on germline mutations of four MA lines at generation 1000 and the ancestor lines after next-generation whole-genome sequencing found 5, 92, 19, and 13 point-mutation candidates (32 point mutations per MA line on average). Considering that the point mutation candidates need

further Sanger sequencing confirmation, the genomic mutation rate 3.08e-10 point mutation candidates per generation is probably an over-estimate, and the true genomic mutation rate may be similar to the estimate from another model ciliate *Paramecium tetraurelia* (SUNG *et al.* 2012b). The point-mutation rate per site per generation in both ciliates are lower than most other eukaryotic organisms, this is probably due to *T. thermophila* and *P. tetraurelia* having similarly large effective population size and high fidelity DNA polymerases (SUNG *et al.* 2012a; SUNG *et al.* 2012b).

Finally, there is a potential complication of genome sequencing when the somatic genome is sequenced in the future. Somatic genome sequences might be contaminated by germline sequences during next-generation whole genome deep sequencing. Because the somatic genome has many more copies (usually 45×) than the germline genome (2×), most sequence data recovered will be from the somatic genome; however, some contaminating sequences from the germline genome are expected (EISEN *et al.* 2006). Any sequences that do not map to the reference genome are likely from contaminating, developmentally eliminated sequences. In sequencing the somatic genome, mapped sequences with apparent heterozygosity may be due to either incomplete phenotypic assortment or germline contamination. These possibilities will be distinguished by comparison with the sequenced germline genome from that line. This will not be a problem when sequencing the germline genome, since after GE the germline and soma will be identical, with the exception of developmentally eliminated sequences.

## 5.2 Insights and improvements of mutations on ciliates morphology

### *5.2.1 Insights from mutations influencing morphology*

Four morphological traits of ciliates, *viz*. numbers of somatic and post-oral kineties, macronucleus, and cell sizes were used in chapter 3 for exploring morphological variation from germline mutations after long term asexual reproduction in *T. thermophila*. Cell size change from cells bearing germline mutations was also positively correlated with fitness change. In addition, morphological traits' change due to mutations after 1000 generations established a case study of testing the genetic variability of morphological traits used for species identification, using mutation accumulation techniques. Consistent with previous observations by taxonomists, the number of post-oral kineties, which has been given more weight in species identification (CORLISS 1973), never changed in cell lines with or without mutations accumulated, but the cell size did decrease in cell lines with germline mutations expressed. In the future, such tests on morphological traits could be further explored by including sexual reproductions and different populations.

### *5.2.2 Improving the current research*

There are also some potential problems with this part using morphological traits to estimate mutational parameters: first, the sample size was small. Due to time limitation (I was still working on the experiment several weeks before the thesis submission deadline), only six MA lines and six GE lines derived from them were evenly sub-sampled with reference to the 19 GE lines' fitness from the 19 MA lines with successful genomic

exclusion. A large sample size including more MA and GE lines would give more credibility to the mutation accumulation tests on morphological traits variability. Besides, checking finer structures or applying automatic assays could improve the morphological traits measurement. The four morphological traits chosen in this study are common ones used in morphological taxonomy; however, there might be better resolution for detecting mutations if finer morphological structures were used. For example, the number of kinetosomes in a certain somatic kinety might provide more information than the number of somatic kineties on a whole cell. All the morphological traits here were measured by observing living or stained cells on microscopes. At least cell size was known to be measured automatically by flow cytometry (TZUR *et al.* 2011). Using automatic measurement would definitely decrease the systematic errors and increase within line replicates and be more accurate.

## 5.3 Discovery of diverse asexual reproductions and future research on the ciliate *G. trihymene*

### *5.3.1 Discovery of the diverse asexual reproduction modes in G. trihymene*

The last chapter focused on reproductive strategies, using a different ciliate species *G. trihymene*. This topic is not directly related with mutation accumulation, but could be potentially used for evolutionary genetic studies on marine ciliates. For example, the reproductive cysts might be a life stage conveniently used as equivalence with liquid nitrogen frozen ancestor cells, since most marine ciliates could not be frozen in liquid nitrogen. Also, this part of research showed that a ciliate species previously thought to be

"rare" is not actually rare but rather due to that the original habitats and seasonality were unknown. In the future, this ciliate with diverse reproductive strategies could be a model for studying marine ciliates speciation, as indicated in chapter 4 that frequent selfing might have caused sympatric speciation by expressing germline mutations into a new somatic genome soon after mutations' appearance; then, the individuals with the expressed mutations soon got selected against or for. Other lineages without selfing could diverge from this selfing lineage, for example, accumulating mutations in the germline genome due to long term no sex, which would even slightly change conservative genes since germline mutations can accumulate for many generations without selection. Exploration on mating systems and population structures of *G. trihymene* would be a start on this sympatric speciation topic.

### *5.3.2 Future research on the reproduction strategies of G. trihymene*

There are also some potential problems with this part exploring reproductive strategies: sexual reproductive strategies should have been investigated more thoroughly. Except for selfing, all reproductive strategies explored in this study were asexual and mostly equivalent to fission. Though mating tests were tried using the limited number of strains collected mostly from Texas, no successful conjugation except selfing was found. More strains should have been collected from a broader geographical range. Besides, starvation was used to induce mating, which is regular operation on model freshwater ciliates, but since the ocean environment is relatively homogeneous and disturbing compared with the heterogeneous and static freshwater habitats, there might be alternative triggers for

mating other than starvation, for example, higher salinity. Furthermore, species identification could be more explored, in addition to morphology and genetic markers. All the collected strains were identified by morphology and the ITS gene marker, which were mainstream identifying methodology in marine ciliates taxonomy; however, mating test must be done to make sure they are the same biology species in order to establish it as an evolutionary genetic studying subject or to explore speciation. Finally, genetics of *G. trihymene* was never explored. Including chapter 4 study, the genetics of *G. trihymene* remains blank. Even the nuclear activities during conjugation were derived from other studied model ciliates, though such derivation seems to be reasonable as most studied ciliates fit the nuclear developmental patterns from model ciliates. Such blank genetic background is actually not bad news, since most people working on marine ciliates are obsessed with morphology and phylogeny with conservative markers, exploring genetics of this marine ciliate would teach us a lesson on more conceptually grounded thinking and moving on from the years when the microscope was shortly invented.

## 5.4 References


CHALKER, D., 2008 Dynamic nuclear reorganization during genome remodeling of *Tetrahymena*. Biochem Biophys Acta **1783:** 2130–2136.

CORLISS, J., 1973 History, taxonomy, ecology, and evolution of species of *Tetrahymena* in *Biology of Tetrahymena*, edited by A. M. ELLIOTT. Dowden, Hutchinson & Ross Inc., Stroudsburg, Pennsylvania.

DEVISSER, J., R. HOEKSTRA, and H. VAN DEN ENDE, 1997 An experimental test for synergistic epistasis and its application in *Chlamydomonas*. Genetics **145:** 815–819.

DUHARCOURT, S., G. LEPÈRE, and E. MEYER, 2009 Developmental genome rearrangements in ciliates: a natural genomic subtraction mediated by non-coding transcripts. Tr Genet **25:** 344–350.

EISEN, J., R. COYNE, M. WU, D. WU, M. THIAGARAJAN *et al.*, 2006 Macronuclear genome sequence of the ciliate *Tetrahymena thermophila*, a model eukaryote. PLoS Biol **4:** 1620–1642.

EYRE-WALKER, A., and P. D. KEIGHTLEY, 2007 The distribution of fitness effects of new mutations. Nat Rev Genet **8:** 610–618.

KEIGHTLEY, P. D., and A. EYRE-WALKER, 1999 Terumi Mukai and the riddle of deleterious mutation rates. Genetics **153:** 515–523.

TZUR, A., J. K. MOORE, P. JORGENSEN, H. M. SHAPIRO, and M. W. KIRSCHNER, 2011 Optimizing optical flow cytometry for cell volume-based sorting and analysis. PLoS One **6:** e16053.

WEST, S. A., A. D. PETERS, and N. H. BARTON, 1998 Testing for epistasis between deleterious mutations. Genetics **149:** 435–444.