2010

Dynamic Genomes of Eukaryotes and the Maintenance of Genomic Integrity

Laura Wegener Parfrey
University of Massachusetts Amherst

Laura A. Katz
Smith College, lkatz@smith.edu

Dynamic Genomes of Eukaryotes and the Maintenance of Genomic Integrity

Eukaryotes specify a genome to be inherited stably, enabling dynamic rearrangements and amplifications of other genomic elements

Laura Wegener Parfrey and Laura A. Katz

Many biologists assume that eukaryotic genomes are transmitted stably between generations with only minor variations. Yet, this presumed constancy is at odds with data indicating that eukaryotic genomes are dynamic, varying extensively in content among many different lineages. Thus, rather than being constant, genomes vary considerably within individuals during their lifetimes.

Despite the dynamic nature of eukaryotic genomes, it appears likely that eukaryotes specify the portion of their genomes that will be inherited, either by placing it in a separate germline nucleus, by marking it through epigenetics, or both. Epigenetic mechanisms such as methylation, histone acetylation, and genome scanning through RNAi may differentiate between the inherited and somatic genetic material within a single nucleus. Such a distinction between somatic and inherited genomes is key to enabling dynamic variation in genomes within life cycles and among individuals within populations, while also maintaining integrity of the genome between generations.

Much of our traditional view of genomes and their inheritance is based on insights from plants and animals, species whose cells carry a single nucleus and a relatively small number of chromosomes. Ploidy variation—changes in the number of whole genome complements in a nucleus—in these organisms is generally limited to fluctuations between haploid (one copy of each chromosome) and diploid (two copies) within life cycles. According to this view, genome content is maintained at a constant ploidy level during cell propagation by mitosis and cytokinesis, resulting in a population or a multicellular organism that is composed of genetically identical cells.

We focus on two genome modifications in which nuclear DNA content changes during the life cycle: cyclic polyploidy and differential amplification. During cyclic polyploidy, the number of whole genome copies increases, in some lineages by a factor of 1,000, and then decreases prior to reproduction such that each daughter cell inherits either one or two genome complements. In contrast, during differential amplification, the copy number for only portions of the

Summary

- The presumed constancy of eukaryotic genomes between generations is at odds with the observations of their being dynamic and varying extensively within a generation in some eukaryotic lineages.
- It appears likely that eukaryotes specify the portion of their genome to be inherited, either placing it in a separate germline nucleus or marking it through epigenetics.
- Both cyclic polyploidy and differential amplification elevate gene copy numbers and likely provide advantages to the organism as more templates become available for transcription, thus increasing metabolic efficiency.
- Microbial eukaryotes display diverse mechanisms for reducing ploidy levels and thus maintaining the fidelity of the inherited genome.

Laura Wegener Parfrey is a Ph.D. candidate in the Program in Organismic and Evolutionary Biology, University of Massachusetts, Amherst, and Laura A. Katz is Elsie Damon Simonds Professor in the Department of Biological Sciences, Smith College, Northampton, Mass., and a faculty member in the Program in Organismic and Evolutionary Biology, University of Massachusetts, Amherst.

genome increases, and these portions range in size from whole chromosomes to single genes. Both cyclic polyploidy and differential amplification elevate gene copy numbers and likely provide advantages to the organism as more templates become available for transcription, thus increasing metabolic efficiency.

Inherited Genome Concept Could Explain How Integrity Is Maintained

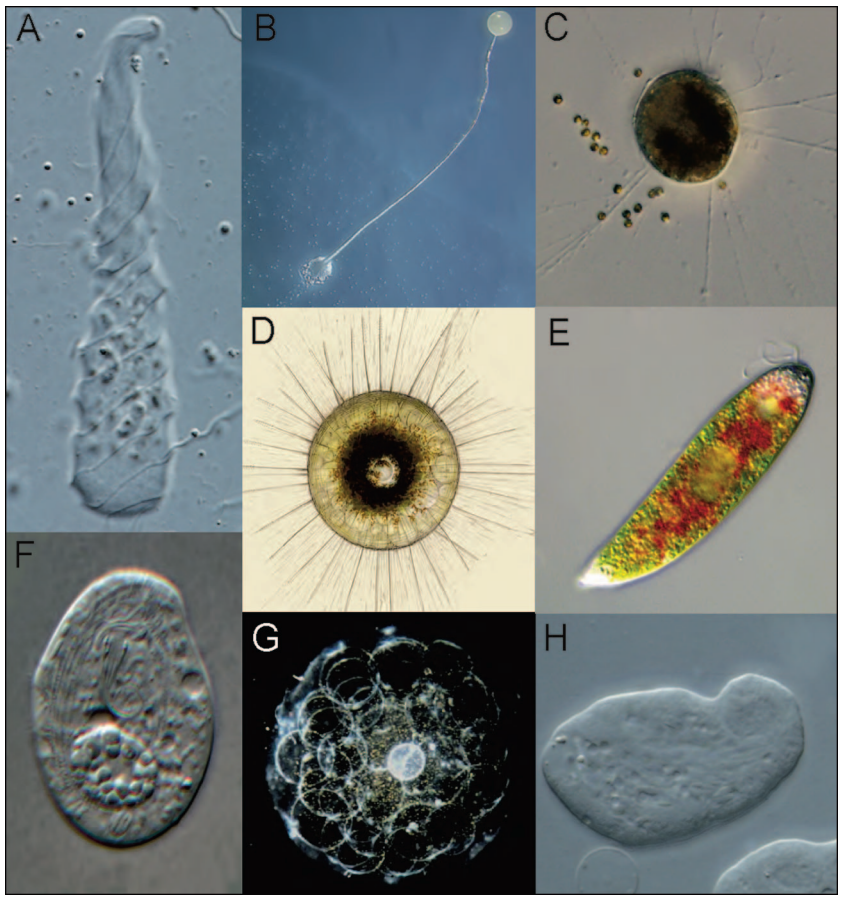
Microbial eukaryotes (Fig. 1) have many means for amplifying their genomes while still maintaining hereditary integrity. We use a phylogenetic framework to highlight examples of genomic features across the eukaryotic tree of life along with diverse mechanisms for resetting the inherited genome (Fig. 2 and 3). The broad distribution of these features suggests that a shared mechanism for distinguishing between inherited and somatic genomes enables the evolution of dynamic genomes among diverse eukaryotes.

To bridge the gulf between the traditional and dynamic views of genomes, we propose the inherited genome concept for eukaryotes. This concept involves organisms marking their inherited genome so that it may be transmitted to the next generation with its integrity intact.

Eukaryotic genomes transmit information both genetically, through DNA via its nucleotide sequence, and epigenetically through chemical changes to that DNA or its associated histone proteins. There is increasing recognition that epigenetic processes play important roles in regulating activities within the nucleus, such as heterochromatin formation and transcription levels. We suspect that epigenetic mechanisms also play a role in marking inherited genomes. Once the inherited genome is marked, the somatic genome becomes free to vary in copy number and composition.

We use somatic here to refer to the genome that is used for body cells as opposed to the germ line or inherited material. This distinction between inherited and somatic genomes is unmistakable in animals where the inherited genome is sequestered within germline cells early during embryonic development, while the somatic genome varies among cell types. Indeed, within humans and other mammals, genomes become polyploid in specialized cells such as megakaryocytes and liver cells, while they are dramatically rearranged in immune cells (Fig. 3B). Further, many other eukaryotes have mechanisms for marking their inherited genome without necessarily sequestering them in specialized germ cells, as described below.

FIGURE 1



Some microbial eukaryotes with dynamic genomes. Approximate size of organisms is given for reference. (A) *Pyrsonympha* sp. (oxymonad)–95 μm; (B) *Dictyostelium discoideum* (acellular slime mold)–fruiting body 2.5 mm, amoeba 10 μm; (C) *Allogromia laticollaris* strain CSH (Foraminifera)–75 μm; (D) *Aulacantha scolymantha* (Phaeodarea)–200 μm; (E) *Euglena sanguinea* (Euglenid)–120 μm; (F) *Chilodonella uncinata* (ciliate)–25 μm; (G) *Thalassophysa* sp. (Polycystina)–6 mm (including bubble capsule); (H) *Entamoeba* sp.–25 μm.

Katz: Ciliates That Do Not Heed Mendel Might Shed Light on Other Biological Principles

"The organisms I study forgot to read the introductory biology textbook," says Laura Katz of Smith College, whose research focuses on the evolution of microbial eukaryotes. "Nobody told them about Mendel's rules, or that Lamarckism is out of favor. The organisms we study also have dynamic genomes that vary in content and structure throughout life cycles."

Katz, a professor in biological sciences, studies eukaryotic evolution through phylogenetic reconstruction, community sampling, and analyses of genome evolution. "We still know very little about the diversity of our microbial relatives," she says. Students in her lab do research in three areas: reconstructing the eukaryotic tree of life, with particular focus on ciliates and amoebae; determining the principles of genome evolution through studies of ciliates and foraminifera; and characterizing the diversity of ciliates in near-shore environments.

"Working on eukaryotic microbes, I can begin to imagine what it was like to be off exploring on the *Beagle*," she says, referring to the ship that carried Charles Darwin to the Galapagos Islands among other far-flung places while he studied biological diversity. "For example, our work on marine ciliates is contributing to the picture of a 'rare biosphere'—the surprising observation that most microbes in the environment are rare." She is confident the work will add to other insights from microorganisms that have helped to transform biology. "For example, telomeres and self-splicing RNA were originally discovered in ciliates, one of the groups we focus on in particular," she says. "My belief is that work on eukaryotic microbes will shed light on processes found across the tree of life, and perhaps even within humans."

Because Smith College is predominantly an undergraduate institution, Katz teaches more than do others like her at larger universities. "Smith is unusual, in that we provide an elite education to women of promise, independent as to whether they are women of privilege," she says. "When I last looked, Smith was second [after Berea College] among liberal arts colleges in the nation for attracting students on Pell grants." Nonetheless, she maintains an active research lab, and she is aggressive in seeking grants to support that work. "Recently, I had two technicians, two postdocs, four Ph.D. students, a master's student, and an uncountable number of undergraduates," she says.

In addition to teaching two biology courses each semester, Katz also is deeply involved in a program for underrepresented students, including those who are first in their families to attend col

Processed Genomes in Ciliates

Genetic material in ciliates (Fig. 1F) is divided between a silenced germline micronucleus and a transcriptionally active macronucleus, each of which is found within the same cell. The micronucleus behaves like a canonical nucleus. It has a relatively small number of large chromosomes that divide by mitosis during asexual reproduction and by meiosis leading up to sexual reproduction.

Ciliates exchange meiotic products of the micronucleus during sexual reproduction, and these fuse to form a zygotic nucleus. The zygotic nucleus then divides to produce a new micronucleus and, after a period of genome rearrangement, a macronucleus. During macronuclear development the large micronuclear chromosomes are fragmented, micronuclear-limited sequences are excised, and the remaining material is amplified to yield as many as 25 million gene-sized chromosomes in some lineages. This processing results in numerous small chromosomes with most noncoding sequences removed (Fig. 3A).

With the inherited genome sequestered in the micronucleus, ciliates dramatically modify their macronuclear genome, which may offer selective benefits. For example, we find that elevated rates of protein evolution correlate with the extent of chromosomal processing. Thus, ciliates with extensively processed chromosomes have more divergent proteins than those lineages that undergo limited processing.

The micronuclear genome is not seen by selec-

lege. "Among my most humbling experiences was watching my first student, whose father was a truck driver and whose mother was unemployed, win a prize for best student presentation at a regional meeting," Katz says. "This was particularly impressive, as she was competing against graduate students."

In 2007, working with Kate Queeney—a colleague in the chemistry department—Katz launched the program called Achieving Excellence in Mathematics, Engineering, and Sciences (AEMES). Its goal is to enable underrepresented and first-generation college students succeed in science, math, and engineering early during their time at Smith. AEMES targets students through the admission process, then provides them with a science-focused program and an adaptive learning strategies course to help build skills and to support group study. The idea is "to coordinate existing efforts and to create new opportunities for diverse students," she says. AEMES also matches students with faculty and peers for additional support, and encourages students to engage in research early in their college careers. "In three years, 60 students have participated," she says.

Katz, who "moved around a lot as a child," lived mostly in the Northeast, but also spent time in Chicago and Germany. Her father is a university-associated cardiologist, her mother, a classicist. Her brother studies Chinese history and lives in Taiwan, one sister sells echocardiogram equipment for General Electric and lives near Boston, and her second sister is an events organizer in London.

"I always loved doing science," Katz says. "I started in basic science research as a high school student, and I was particularly encouraged by my dad. I did not recognize my love of microbes until the end of my graduate career." She earned her A.B. in the history of science from Harvard College in 1989, and her Ph.D. in ecology and evolutionary biology from Cornell University in 1996.

Katz was a postdoctoral fellowat Princeton University, where she fell under the influence of John Tyler Bonner, her postdoctoral advisor and a developmental biologist whose chief research interest is in cellular slime molds. Bonner, now 89, professor emeritus, has been at Princeton since 1947. "He loves to learn and he loves to teach," she says of Bonner. "As a postdoctoral fellow at Princeton, I used to have lunch with John, and we'd talk about everything— life, science. John is truly an inspiring person. I still rely heavily on John for advice about my research, and about academia in general."

Katz, who is married to an immigration lawyer and has two children, 12 and 10, says she loves what she does. "I do feel lucky that I get paid to look down the microscope at pond scum," she says.

Marlene Cimons
Marlene Cimons lives and writes in Bethesda, Md.

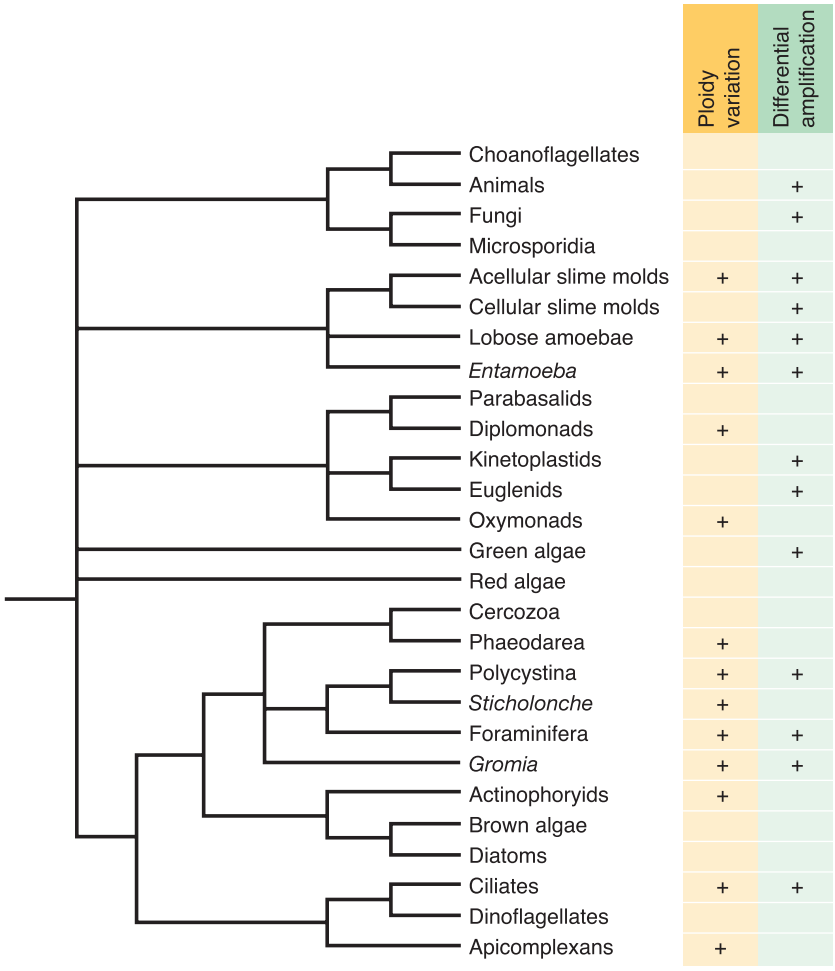
tion because it is transcriptionally inactive, perhaps allowing ciliates to explore protein space and accumulate potentially beneficial compensatory mutations. Finally, processing of the macronuclear genome confers the advantages of polyploidy without the necessity of returning to the inherited genome.

Foraminifera Undergo Complex Nuclear Events

Foraminifera (Fig. 1C) undergo complex nuclear events, sometimes greatly modifying their genomes during vegetative growth but then passing only their inherited genome to the next generation. The life cycle of Foraminifera alternates between uninucleate and multinucleate adults that biologists traditionally assumed to be haploid and diploid, respectively. However, events that occur before reproduction in the uninucleate phase suggest that both polyploidy and differential amplification occur during development.

Consider the dynamics within the nucleus during the uninucleate stage. The single nucleus, which arises from meiosis, is haploid. According to classic accounts of Foraminifera, this nucleus remains haploid while the cell and nucleus grow tremendously—in some species, the nucleus can expand to 400 μm in diameter. Shortly before reproduction, the nucleus enlarges, while DNA and other material within it condense into granules. These granules are subsequently degraded in a process referred to as *Zerfall*.

FIGURE 2



Distribution of dynamic genomic features found throughout the eukaryotic tree of life. + indicates that cyclic polyploidy or differential amplification are reported in at least one taxon within the lineage.

The ability of Foraminifera to degrade portions of their nuclear DNA prior to reproduction indicates that DNA is amplified during the growth of the nucleus and the amplified DNA can be distinguished from the inherited genome. *Zerfall*, aptly described as nuclear cleansing, appears to facilitate a return to the inherited genome before reproduction.

Following *Zerfall*, the single nucleus divides rapidly by mitosis into hundreds or sometimes thousands of daughter nuclei, demonstrating that this single nucleus must be polyploid. There is no detectable DNA synthesis during the proliferation of daughter nuclei in the only species in which it was measured. The details of this proliferation vary from one species to another, but all indicate polyploidy of the parental nucleus. Thus, it is likely that multiple whole genome copies are produced by polyploidization of the parental nucleus during growth, and that these genome copies segregate into daughter nuclei during reproduction. Other loci are also amplified extrachromosomally in eukaryotes. Foraminifera extensively modify their genome content and have developed at least one mechanism, *Zerfall*, to return to the inherited genome.

Phylogenetic Diversity of Dynamic Genomes

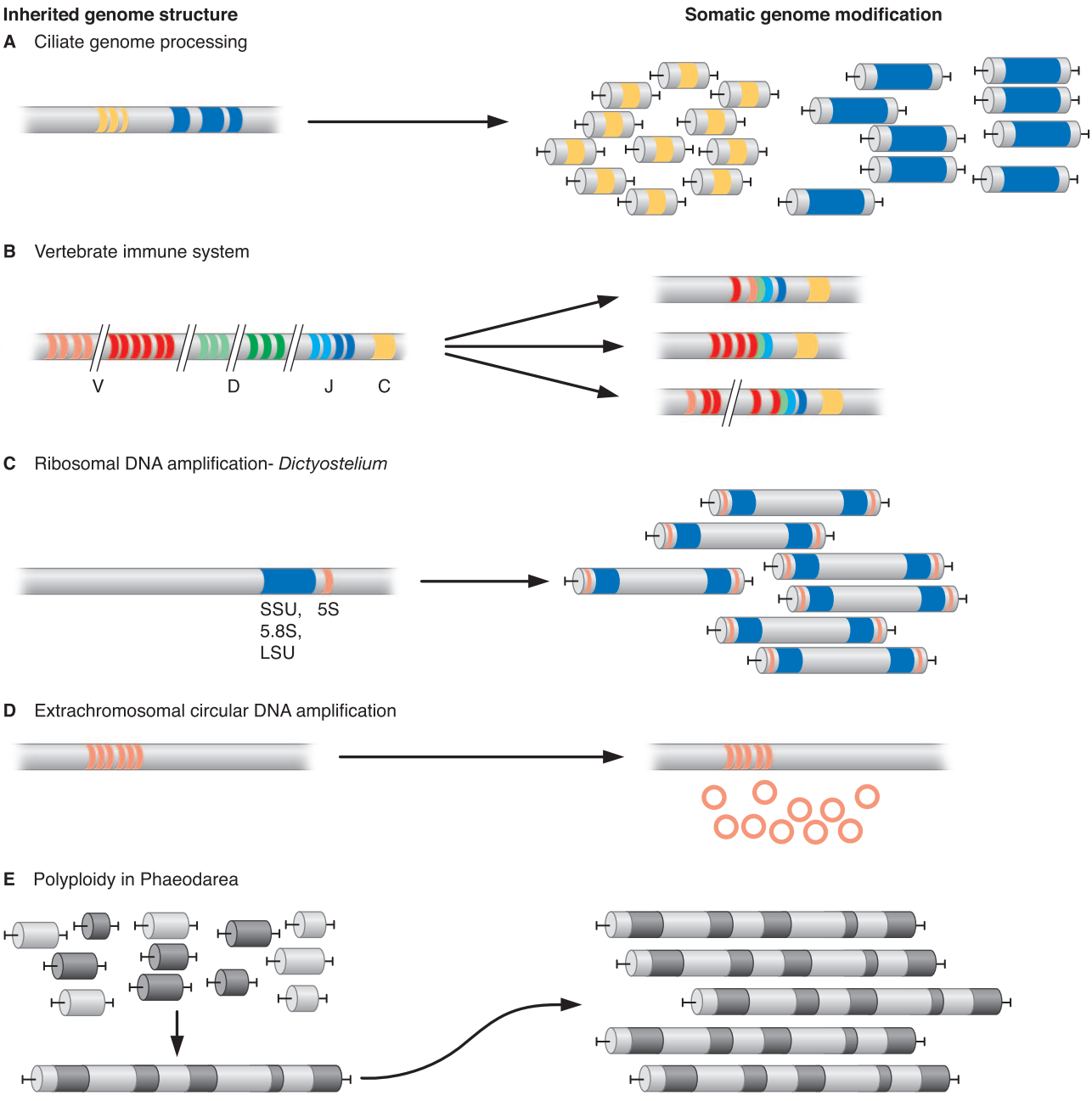
Beyond the specific life cycles that are the focus of work in our lab, differentiation between inherited and somatic genomes is apparent in other lineages (Fig. 1). Even though data are limited, these lineages appear to be widespread across the eukaryotic tree of life (Fig. 2). The number of amplified loci ranges from the whole genome in ciliates to ribosomal DNA in many other lineages. This broad distribution indicates that the ability to recognize the inherited genome selectively is widespread among eukaryotes.

Ribosomal DNA is the most commonly amplified locus, including among fungi, animals, *Entamoeba*, slime molds, ciliates, and *Euglena*. In many cases, amplification of the chromosomal rDNA locus leads to extrachromosomal copies of rDNA. These may be circular, as in animals and *Euglena* (Fig. 1E), or linear as in ciliates and the slime mold *Dictyostelium* (Fig. 1B and 3C).

Extrachromosomal copies are presumably excluded from the inherited genome because they lack chromosome features (such as kinetochores) necessary for making it into daughter cells. In contrast, *Entamoeba histolytica* (Fig. 1H), the causative agent of amoebiasis (or dysentery), contains no chromosomal copy of rDNA. Instead, rDNA is located on up to 200 circular plasmid-like molecules that are maintained at variable copy numbers during the life cycle of *Entamoeba* and that are inherited through an unknown mechanism.

Other loci are also amplified in eukaryotes. For instance, Sarit Cohen and colleagues at Tel

FIGURE 3



Selected examples of genomic processes that lead to somatic genome variation. In all cases, the extent of amplification is highly variable. (A) Genome-wide rearrangements in ciliates involve chromosome breakage, elimination of DNA, and amplification of processed chromosomes. Colored areas represent coding regions and white regions are non-coding; "T" bars at the end of chromosomes represent telomeres that are added *de novo*. (B) Targeted rearrangements are exemplified here by V(D)J processing in vertebrate immune systems. In this case, a single locus is processed to produce a diversity of antibody genes by joining various V, D, and J regions faded ends indicate that the chromosome continues. (C) Extrachromosomal amplification of rDNA in *Dictyostelium* produces linear chromosomes with the rDNA locus arranged as palindromic repeats. (D) Amplification of a tandemly repeated gene leading to extrachromosomal circular DNA. (E) Illustration of polyploidization in Phaeodarea. Individual chromosomes are joined end to end into a composite chromosome, which is amplified many times.

Aviv University in Israel recently demonstrated that extrachromosomal DNA is generated from multiple loci, including loci for ribosomal DNA, in several animal and plant taxa. The unifying feature of the amplified loci is that they are arranged as tandem repeats on chromosomes (Fig. 3D). Extrachromosomal circles of DNA are generated by DNA repair mechanisms, and can replicate independently of the chromosomes. This mechanism for differential amplification may be universal throughout eukaryotes.

Dynamic processes that resemble *Zerfall* are also found in Polycystinea (a lineage formerly called radiolaria) and *Gromia*, both of which are large amoebae closely related to Foraminifera. These lineages discard and degrade DNA before reproduction, again suggesting that a portion of the genome was amplified before that phase. As in Foraminifera, the identity of the degraded DNA is unknown, but ribosomal DNA and other highly expressed loci are good candidates.

Cyclic Polyploidy Arose Several Times

Polyploidy arose numerous times in eukaryotes (Fig. 2). Terminal polyploidy is common in differentiated tissues of multicellular organisms such as plants and animals. In these cases, polyploidy is not reduced and these cells eventually die, while the sequestered germline cells give rise to subsequent generations. In contrast, if polyploid single-cells failed to transmit their genome, they would be evolutionary dead ends. Hence, microbial eukaryotes have numerous mechanisms to reduce ploidy levels and to maintain the fidelity of their inherited genome. Cyclic polyploidy is described in members of the radiolaria, Foraminifera, oxymonads, and Apicomplexa.

Radiolaria are a polyphyletic assemblage of large marine amoebae whose ornate shells were drawn by Haeckel. They fall into three clades on the basis of molecular analysis and differences in shell composition: the Phaeodarea, Polycystina, and Acantharea. All become polyploid during their life cycle and then segregate these genome copies into gametic nuclei. The resulting biflagellate cells are described as spores or swarmers, but are likely gametes that will not complete their life cycle in culture. The details of gametogenesis differ among members of these groups, which diverged at least 550 million years ago. In each case, however, the result is a return to the inherited genome during reproduction.

Phaeodarea (Fig. 1D) have a single polyploid nucleus with roughly 1,000 composite chromosomes (Fig. 3E). Prior to reproduction, these composite chromosomes condense and are segregated along microtubules into secondary nuclei. Subsequently, the composite chromosomes break down into 10–12 smaller chromosomes, presumably the inherited genome complement, and divide mitotically to give rise to gametic nuclei.

In Polycystina (Fig. 1G), gametogenesis begins with nucleoli and other peripheral nuclear material pinching off from the nucleus and being degraded in a *Zerfall*-like process. Then intranuclear microtubule organizing centers (MTOCs) form at intervals along the nuclear envelope, and chromosomes condense onto these MTOCs. Each MTOC and its associated chromosomes give rise to a secondary nucleus, which divides mitotically to produce gametes.

The oxymonad *Pyrsonympha* (Fig. 1A) is a symbiont of termites and is very large for a flagellate, up to 200 μm in length. *Pyrsonympha* becomes polyploid as it grows. Upon reproduction, whole chromosomes sets are segregated into numerous daughter nuclei, thereby returning to the inherited genome. The homologous chromosomes are connected by synaptonemal complexes that aid in distributing them to daughter nuclei. The presence of synaptonemal complexes, which are commonly found during meiosis, is intriguing as they may be evidence of meiosis occurring within the polyploid nucleus.

Apicomplexa are a diverse lineage of intracellular parasites and include the causative agents of human diseases such as malaria. Apicomplexa have complex life cycles that often include cyclic polyploidy. For example, during growth the genome and microtubule organizing centers in *Sarcocystis neuorna*, which infects horses, replicate six times without nuclear division, leading to a nucleus with 32 genome copies and MTOCs. Subsequently, *Sarcocystis* segregates these copies of the inherited genome into daughter nuclei.

Since dynamic genome processes are widespread, eukaryotes must have evolved a mechanism to mark their inherited genome. Such a distinction between somatic and inherited genomes may be key in enabling the variation observed within life cycles while also maintaining integrity of the genome between generations.

SUGGESTED READING

Cerutti, H., and J. A. Casas-Mollano. 2006. On the origin and functions of RNA-mediated silencing: from protists to man. Curr. Genet. **50:**81–99.

Cohen, S., and D. Segal. 2009. Extrachromosomal circular DNA in eukaryotes: possible involvement in the plasticity of tandem repeats. Cytogenetic Genome Res.**124:**327–338.

Goldstein, S. T. 1999. Foraminifera: a biological overview, p. 37–56. *in* B. K. Sen Gupta (ed.), Modern Foraminifera. Kluwer, Dordrecht.

Kloc, M., and B. Zagrodzinska. 2001. Chromatin elimination—an oddity or a common mechanism in differentiation and development? Differentiation **68:**84–91.

Kondrashov, A. S. 1997. Evolutionary genetics of life cycles. Annu. Rev. Ecol. Systematics **28:**391–435.

McGrath, C. L., and L. A. Katz. 2004. Genome diversity in microbial eukaryotes. Trends Ecol. Evol. **19:**32–38.

Parfrey, L. W., D. J. G. Lahr, and L. A. Katz. 2008. The dynamic nature of eukaryotic genomes. Mol. Biol. Evol. **25:**787–794.

Raikov, I. B. 1982. The protozoan nucleus: morphology and evolution. Springer-Verlag, Wien.

Zufall, R. A., C. L. McGrath, S. V. Muse, and L. A. Katz. 2006. Genome architecture drives protein evolution in ciliates. Mol. Biol. Evol. **23:**1681–1687.

Zufall, R. A., T. Robinson, and L. A. Katz. 2005. Evolution of developmentally regulated genome rearrangements in eukaryotes. J. Exp. Zool. Part B-Mol. Dev. Evol. **304B:**448–455.