3-12-2021

Opinion: Genetic Conflict With Mobile Elements Drives Eukaryotic Genome Evolution, and Perhaps Also Eukaryogenesis

Adena B. Collens  
*Smith College*

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

Follow this and additional works at: [https://scholarworks.smith.edu/bio_facpubs](https://scholarworks.smith.edu/bio_facpubs)

Part of the Biology Commons

**Recommended Citation**
[https://scholarworks.smith.edu/bio_facpubs/76](https://scholarworks.smith.edu/bio_facpubs/76)

This Article has been accepted for inclusion in Biological Sciences: Faculty Publications by an authorized administrator of Smith ScholarWorks. For more information, please contact scholarworks@smith.edu
Symposium Article

Opinion: Genetic Conflict With Mobile Elements Drives Eukaryotic Genome Evolution, and Perhaps Also Eukaryogenesis

Adena B. Collens and Laura A. Katz

From the Department of Biological Sciences, Smith College, Northampton, MA 01063 (Collens and Katz); Program in Organismic and Evolutionary Biology, University of Massachusetts, Amherst, MA 01003 (Katz).

Address correspondence to Laura A. Katz at the address above, or e-mail: lkatz@smith.edu.

Received June 8, 2020; First decision July 31, 2020; Accepted December 17, 2020.

Abstract

Through analyses of diverse microeukaryotes, we have previously argued that eukaryotic genomes are dynamic systems that rely on epigenetic mechanisms to distinguish germline (i.e., DNA to be inherited) from soma (i.e., DNA that undergoes polyploidization, genome rearrangement, etc.), even in the context of a single nucleus. Here, we extend these arguments by including two well-documented observations: (1) eukaryotic genomes interact frequently with mobile genetic elements (MGEs) like viruses and transposable elements (TEs), creating genetic conflict, and (2) epigenetic mechanisms regulate MGEs. Synthesis of these ideas leads to the hypothesis that genetic conflict with MGEs contributed to the evolution of a dynamic eukaryotic genome in the last eukaryotic common ancestor (LECA), and may have contributed to eukaryogenesis (i.e., may have been a driver in the evolution of FECA, the first eukaryotic common ancestor). Sex (i.e., meiosis) may have evolved within the context of the development of germline–soma distinctions in LECA, as this process resets the germline genome by regulating/eliminating somatic (i.e., polyploid, rearranged) genetic material. Our synthesis of these ideas expands on hypotheses of the origin of eukaryotes by integrating the roles of MGEs and epigenetics.

Subject area: Population structure and phylogeography

Key words: epigenetics, eukaryotic diversity, LECA, meiosis, transposable elements, viruses

Overview

Based on observations of dynamic genomes (i.e., cyclical polyploidy, genome rearrangements) in diverse eukaryotic lineages, we have previously argued that last eukaryotic common ancestor (LECA) used epigenetic mechanisms to distinguish germline from somatic DNA, even in the context of a single nucleus (McGrath and Katz 2004; Zufall et al. 2005; Parfrey et al. 2008; Parfrey and Katz 2010; Maurer-Alcala and Katz 2015; Weiner et al. 2020). As discussed in this series of papers from our lab, examples of such germline/soma distinctions include: sequestered germline nuclei in animals, ciliates, and some foraminifera; cyclical polyploidization throughout life cycles of apicomplexans such as Plasmodium (the causative agent of malaria); generation of extrachromosomal DNA, including amplification of ribosomal RNA loci in many eukaryotes; developmentally regulated genome rearrangements, for example, in trypanosomes and immune cells of vertebrates (i.e., V(D)J recombination); and even the mis-regulation of DNA through polyploidization in cancer cells (Erenpreisa et al. 2017). Despite this long list, examples of genome dynamics in diverse lineages of eukaryotic microbes are still limited as the bulk of their life cycle data come from a small number...
of model lineages (e.g., Tetrahymena, Plasmodium). However, promising recent evidence of chromatin extraction and deoxyribonucleic acid in *Amoeba proteus* (Goodkov et al. 2020) suggests that more examples of such dynamics are on the horizon.

We have also argued that germline/soma distinctions in eukaryotes are regulated by epigenetic tools including histone modification, DNA methylation, and scanning by small nonprotein-coding RNAs (Zufall et al. 2005; Parfrey et al. 2008; Parfrey and Katz 2010; Maurer-Alcala and Katz 2015; Weiner et al. 2020). Here, we extend this hypothesis by combining it with two observations: (1) the widespread occurrence of mobile genetic elements (MGEs) (e.g., transposable elements [TEs], viruses) and (2) data on the epigenetic regulation of MGEs within eukaryotes. Synthesis of these observations leads to the hypothesis that genetic conflict has shaped the evolution of eukaryotic genomes and, as others have also argued (e.g., Aravind et al. 2012; Koonin 2017; Massey and Mishra 2018; Havird et al. 2019), the evolution of eukaryotes themselves.

### MGEs Are Widespread

The function and abundance of MGEs such as viruses and TEs has been extensively reviewed, and we provide only a few highlights here. TEs are present in genomes across the tree of life (e.g., Kidwell and Lisch 2001; Suzuki and Bird 2008; Keinovskiy et al. 2012; Campbell et al. 2017) and can constitute more than half the genome of many eukaryotic lineages (e.g., Kazazian 2004; Fedoroff 2012; Song and Schaack 2018). Viruses are the most abundant biological entities on Earth (e.g., Edwards and Rohwer 2005; Koonin 2017), and, like TEs, they are able to integrate into eukaryotic genomes (Chalker and Yao 2011; Koonin 2017; Song and Schaack 2018).

Though early studies characterized MGEs as “parasitic” and/or “selfish” because of the harm they can cause to host genomes, it is now clear that MGEs also generate novel genetic variation that can be the source of adaptation (e.g., Fedoroff 2012; Koonin and Krupovic 2018). Some of the damage TEs can cause includes mutations, DNA breaks, and rearrangement of chromosomes as they move through host genomes (e.g., Kazazian 2004; Fedoroff 2012; Parhad and Theurkauf 2019). Similarly, rapid evolution and replication of viruses create an “arms race” with the host genomes that have evolved to eliminate them (e.g., Bruscella et al. 2017; Koonin and Krupovic 2018). Consequently, replication and mobilization of MGEs is a substantial source of genetic variation in eukaryotes, and these abilities allow MGEs to both resist elimination and create an immediate and lasting impact on host evolution (e.g., Kidwell and Lisch 2001; Schaack et al. 2010; Campbell et al. 2017; Koonin and Krupovic 2018).

### MGEs Are Regulated by Epigenetics

Epigenetic mechanisms are key to eukaryotic responses to MGEs (e.g., Levine et al. 2016; Campbell et al. 2017; Song and Schaack 2018; Parhad and Theurkauf 2019). In many cases, epigenetic responses protect the host’s germline by limiting TE mobilization (Chung et al. 2008; Suzuki and Bird 2008; Parhad and Theurkauf 2019). *Drosophila* exemplify this through expansion of the *HP1D* gene family, which silences TEs in the female germline (Levine et al. 2016). While under epigenetic regulation, TEs display a spectrum of fitness effects within host genomes from parasitism to mutualism (Kidwell and Lisch 2001; Vogt et al. 2013; Cosby et al. 2019). This relationship can also change over time as the epigenetic systems that regulate them may evolve such that transposons ultimately become domesticated (e.g., neutral or used for host function, Kidwell and Lisch 2001; Vogt et al. 2013; Piegut et al. 2013; Cosby et al. 2019; Doyle and Coate 2019).

Epigenetic mechanisms can also regulate viruses within eukaryotic genomes. Endogenous retroviruses, like TEs, occur at various levels of mobility and can be epigenetically regulated via processes like histone methylation (Manghera and Douville 2013; Collins et al. 2015; Meyer et al. 2017). Viruses have also been observed to regulate their replication cycles through epigenetic mechanisms of their own (Woellmer and Hammerschmidt 2013; Balakrishnan and Milavetz 2017; Bruscella et al. 2017). The human Epstein–Barr herpesvirus represents one such intimate relationship, as the latent virus is restrained by Polycomb proteins, but in the lytic replication stage, when Polycomb repression is erased, the virus escapes from the methylation network of the host (Woellmer and Hammerschmidt 2013). This type of multilayered epigenetic relationship reflects the complexity of interactions between viral replication systems and eukaryotic hosts.

### Genetic Conflict Is Foundational to Eukaryotic Genome Evolution, and Perhaps Eukaryogenesis

The widespread occurrence and epigenetic regulation of MGEs engender the hypothesis that genetic conflict between host and MGEs led to the evolution of a dynamic eukaryotic genome that distinguishes germline and soma (Figure 1). Genetic conflict, the competitive relationship between MGEs and host genomes, has been well described as a driving force of evolutionary change (e.g., Hurst et al. 1996; Werren 2011; McLaughlin and Malik 2017; Massey and Mishra 2018; Song and Schaack 2018). Hurst et al. (1996) argued for a “gene’s-eye view” of such conflict to describe the strategies that MGEs and hosts deploy in the struggle over inheritance and proliferation. Nearly two decades later, Song and Schaack (2018) provide an extensive review on the nature of genetic conflict between hosts and MGEs and the possible mechanisms of resolution. In light of this conflict, we and others (e.g., Aravind et al. 2012; Fedoroff et al. 2012; Koonin 2017) propose that epigenetic mechanisms resulting from interactions with MGEs were likely fundamental to eukaryotic evolution. Indeed, the genetic mechanisms that underlie epigenetic regulation (i.e., the epigenetic toolkit) clearly predate the evolution of eukaryotes (e.g., Oliverio and Katz 2014; Weiner et al. 2020), though the specific machinery may have been replaced and/or elaborated on over time (Maurer-Alcala and Katz 2015). Here, we expand on these ideas by linking them explicitly to the origin of germline–soma distinctions during eukaryogenesis.

Consistent with the idea that genetic conflict between host and MGEs specifically led to distinction of germline and somatic genome material are observations on the differential epigenetic regulation of MGEs in extant eukaryotic lineages. For example, flowering plant pollen possesses the ability to epigenetically regulate and de-regulate transcription of TEs in a cyclical manner (Slotkin et al. 2009). In animals like *Drosophila*, TEs are silenced in the germline through female-specific RNA silencing mechanisms (Levine et al. 2016), while a different set of small interfering RNAs regulate TEs in the soma (Chung et al. 2008). In the nematode *Caenorhabditis elegans*, piRNA epigenetic silencing networks suppress TE mobility in the germline, and this silencing can be inherited across more than 20 generations (Ashe et al. 2012). In ciliates, epigenetic mechanisms
including small RNAs and transposases co-opted from transposons are used to shape somatic genomes following conjugation (e.g., Chalker and Yao 2011; Bracht et al. 2013; Maurer-Alcala and Nowacki 2019). The observation of differential epigenetic regulation of MGEs between germline and somatic nuclei in diverse extant eukaryotes raises the possibility that such a mechanism was present in LECA and perhaps even FECA (the first eukaryotic common ancestor).

A special case of conflict at the origin of eukaryotes stems from the acquisition of mitochondria, an event extensively reviewed in the literature (though there remain debates on the timing and physiology of the events; e.g., Pittis and Gabaldon 2016; Lopez-Garcia et al. 2017; Martin 2017; Gabaldon 2018; Lopez-Garcia and Moreira 2019; Wein et al. 2019). At the time of the acquisition of mitochondria, the chimeric cell had to navigate two distinct genomes in a shared cytoplasm. Certainly, there is evidence of conflict between mitochondria and nuclei of extant organisms; for example, in humans, nucleocytoplasmic conflict can lead to disease (e.g., Cummins 2001; Havird et al. 2019) and there are data indicating epigenetic interactions between mitochondria and nuclei (Harvey 2019). Hence, it is possible that conflict from a single but significant “mobile” event, the acquisition of an alphaproteobacterial symbiont in FECA, contributed to the invasion/expansion of MGEs (Krupovic and Koonin 2015) and ultimately the evolution of eukaryotic genome structures.

We suggest that eukaryogenesis resulted in the evolution of a genome that distinguishes germline from soma, which was fueled by genetic conflict between MGEs and hosts (Figure 1). Our hypothesis does not specify the timing of events between FECA and LECA, nor do we address the origin of the eukaryotic cytoskeleton, the synapomorphy of eukaryotes that allowed for the evolution of diverse morphologies and life histories. Instead, we suggest that germline–soma distinctions evolved as a response to genetic conflict with MGEs and contributed to the second major epoch of evolution, the origin of eukaryotes with meiotic sex, as described in Bonner (2019). Under such a scenario, the nucleus may have evolved to “protect” the genome from viruses (e.g., Bell 2009; Aravind et al. 2012; Forterre and Gaia 2016; Poole and Hendrickson 2019) or may have resulted from selection to separate transcription from translation, allowing for the excision of mobile elements (Martin and Koonin 2006; Brunk and Martin 2019). It may also be the case that the nuclear envelope is just a byproduct of events at the time (i.e., resulting from the chaos of the acquisition of mitochondria with its genome [including its own MGEs], or some other autogenous event).

Sex (i.e., meiosis and syngamy) is argued to be ancestral in eukaryotes based on the widespread distribution of meiotic genes coupled with other evidence (i.e., cell fusion, cryptic sexual cycles) in lineages previously thought to be asexual (Lahr et al. 2011; Tekle et al. 2017; Hofstatter et al. 2018) but see Maciver (2019). Kondrashov (1994, 1997) argued that meiosis evolved as a means to regulate cycles of ploidy, which are part of what we refer to as somatic genome content (i.e., cyclical polyploidization, along with the generation of extrachromosomal DNA and developmental regulated rearrangements, all represented by the thin

Figure 1. Genetic conflict during eukaryogenesis resulted in epigenetically regulated germline–soma distinctions in eukaryotes. This figure depicts the players at the origin of eukaryotes, namely the diversity of viruses and the presence of TEs integrated within both bacteria, including the ancestor of mitochondria, and archaea, including the likely host cell of FECA (top panel). Conflict among these genomes and mobile genetic elements (MGEs; middle panel) resulted in eukaryotes that distinguish germline (i.e., marked for inheritance, capable of meiosis to reset genome, represented by the condensed chromosomes in LECA) and somatic (e.g., cyclical polyploidy, extrachromosomal DNA, developmentally regulated genome rearrangements, DNA elimination, represented by the thinner lines within the nucleus of LECA) material (bottom panel). The inset under the somatic functions in LECA represents three somatic chromosomes generated from a single germline region in the ciliate Chilodonella uncinata (redrawn from Gao et al. 2015). Additional details and references can be found in the text.
lines within the nucleus of LECA in Figure 1). In fact, Kondrashov (1994) suggested that sex may have evolved as a means for “ordered genetic reduction,” which would be required in novel eukaryotic lineages with complex genome dynamics (e.g., McGrath and Katz 2004; Zufall et al. 2003; Parfrey et al. 2008; Parfrey and Katz 2010; Maurer-Alcala and Katz 2015; Goodkov et al. 2020; Weiner et al. 2020). Despite open questions (e.g., on the timing of events, the origin of the nuclear envelope and cytoskeleton), we believe consideration of our hypothesis—that genetic conflict between host and MGEs at the time of the origin of eukaryotes led to dynamic genomes in which germline–soma distinctions are regulated by epigenetics and reset through meiosis—provides an important expansion on models of eukaryogenesis.

### Funding

L.A.K. is supported by grants a National Institute of Health award (R15HG010409) and three awards from the National Science Foundation (OCE-1924570, DEB-1651908, and DEB-1541511).

### Acknowledgments

We thank Maria Orive (University of Kansas) for organizing the 2019 American Genetics Association symposium that inspired this manuscript and several members of the Katzlab for comments on earlier versions. We are grateful to Jailene C. Gonzalez for creating the illustration (Figure 1).

### References


