Cell

Leading Edge
Essay

The Frustrated Gene: Origins of Eukaryotic Gene Expression

Hiten D. Madhani[1,*]
[1]Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94158, USA
*Correspondence: hitenmadhani@gmail.com
http://dx.doi.org/10.1016/j.cell.2013.10.003

Eukaryotic gene expression is frustrated by a series of steps that are generally not observed in prokaryotes and are therefore not essential for the basic chemistry of transcription and translation. Their evolution may have been driven by the need to defend against parasitic nucleic acids.

Introduction

The goal of this piece is to consider why gene expression in eukaryotes is the way that it is. Students of molecular biology learn that many key elements of eukaryotic gene expression are generally absent from eubacteria. Because eukaryotic features such as chromatin, pre-mRNA processing, and small RNAs offer many opportunities for regulatory control, it might be tempting to think that these attributes evolved to drive the evolution of complex, multicellular organisms. However, the ubiquity of these gene expression elements and available phylogenetic data argue that the core elements of eukaryotic gene expression were established within the ancient unicellular progenitor of modern eukaryotes. In addition, the widespread abundance of prokaryotes throughout the biosphere means that none of the eukaryotic "embellishments" are required for the operation of the central dogma of molecular biology per se. What, then, drove their initial evolution?

The Frustrated Gene: A Metaphor

As a frame for thinking about this question, consider how one might view the eukaryotic gene expression apparatus as a human organization. It would seem to be a highly bureaucratic one, replete with unnecessary impediments that reduce its ostensible output. It would be as if overeager managers implemented bureaucratic roadblocks to each phase: chromatin obstructs transcription; introns and the requirement for a cap structure on pre-mRNA stymie translation; and continuous degradation of RNAs lacking caps or polyadenylate (polyA) tails diminishes overall output. For the individual gene (as for many an office worker), this would be a frustrating environment in which to work. As detailed below, the hoops through which a gene is forced to jump between transcription and translation may have evolved as part of a cellular defensive strategy rather than a desire for efficiency. Through this metaphorical lens, eukaryotic evolution can be seen as the consequences of what bureaucrats term "enterprise risk management," wherein a focus on potential hazards drives the management of the organization.

The Existential Threat of Parasitic Nucleic Acids

To understand the origins of complex eukaryotic gene expression mechanisms, it is helpful to consider that the evolution of the earliest life forms likely coincided with or was quickly followed by the evolution of the first selfish nucleic acid parasites (Dawkins, 1976). Whatever their form, these could have extinguished early life (via genome damage or the exhaustion of cellular resources) were it not for the rapid evolution of host mechanisms limiting their negative impacts.

Here, parasitic nucleic acids will be defined as those that, at the very least, utilize host ribosomes in order to synthesize proteins required for their own replication, thereby resulting in fitness costs for the host. An understanding of the impact of parasitic nucleic acids on host genomes, as well as features of specific parasites, will be important for the argument developed below. I will focus on the most ubiquitous parasitic DNAs: transposable elements and viruses.

Eukaryotes are targeted by three types of transposable elements: cut-and-paste DNA transposons, non-LTR (long terminal repeat) retrotransposons, and LTR retrotransposons (Craig et al., 2002). Each has a distinct replicative program, but they share the common goal of increasing their copy number relative to the remainder of the genome. Integration into new sites in chromatin is a requirement for success, as is the ability to use host RNA polymerase II (RNA pol II) and host ribosomes. Because of sexual reproduction, transposable elements can proliferate within populations despite a negative impact on host fitness (Hickey, 1982). Sex facilitates population spread beyond a single maternal or paternal lineage. The ability of mobile genetic elements to spread through a population via sex partly accounts for the widespread evolution of antitransposon defense mechanisms, as well as for the focused action of these systems in the germline. Eukaryotes are also infected by three broad classes of exogenous parasitic nucleic acids: DNA viruses, RNA viruses, and retroviruses. Viruses use diverse strategies for replication and spread, among which the retroviruses notably require entry into the nucleus and integration of double-stranded DNA into chromatin. Like transposable elements, retroviruses use host RNA Pol II to express their genes. Viruses that use their own RNA polymerase face many challenges to gene expression, as discussed below.

Thus, a conflict is set up in which the host, which can be from any of the three domains of life, needs to develop strategies to sense and silence parasitic nucleic acids, whereas the latter need to replicate

despite these host strategies. The existence of dedicated antiviral immunity mechanisms among all domains of life speaks to the importance of this conflict.

Below, I will describe a series of examples aimed at developing the underlying thesis that the threat posed by parasitic nucleic acids drove the evolution of contemporary eukaryotic gene expression.

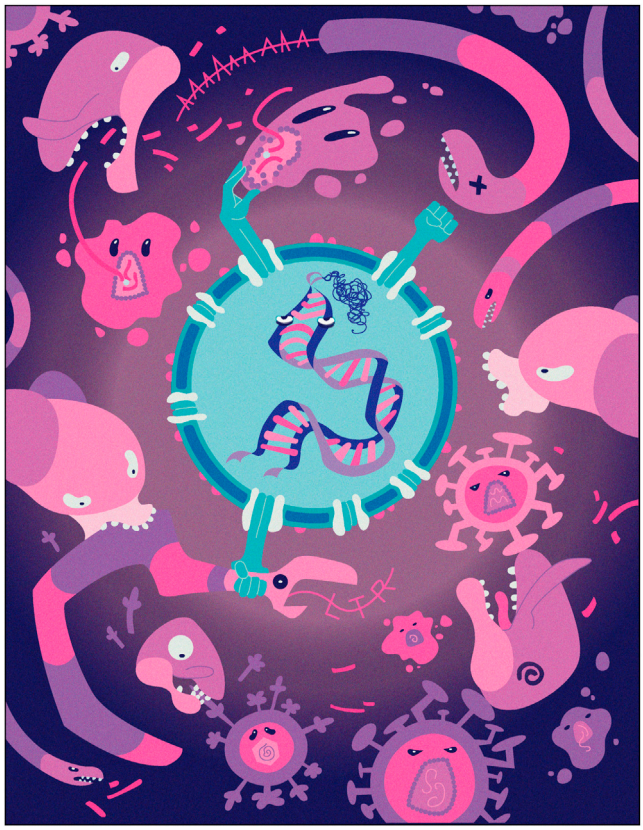
Chromatin: DNA Access Restriction

Histones and chromatin are found in nearly all eukaryotes. It is widely thought that chromatin evolved to allow for the extraordinary DNA condensation required for mitosis and for gene regulation. However, high levels of DNA condensation and elaborate gene regulation also occur in eubacteria that lack chromatin. Could another function of chromatin have driven its initial evolution? Genome-wide surveys of retroviral integration reveal a predilection for integration sites within DNase I-sensitive regions, a marker for accessible chromatin (e.g., Roth et al., 2011), leading me to suggest that chromatin may have evolved to protect genomes from mobile elements and retroviruses, both of which require integration for replication.

Importantly, many studies also show that retroelements have additional integration preferences, commonly including a preference for integration into DNA sites exposed on the outer face of nucleosomes (e.g., Roth et al., 2011). Viewed in the context of the host-pathogen conflict, these preferences can be seen as a way for the mobile elements to "turn the tables" and to adapt to a host block on its spread (the nucleosome).

This model offers a potential explanation for why the control of chromatin at promoters and enhancers is such a prominent aspect of eukaryotic gene regulation. The goal for cells would be to sequester as much of the genome as possible to limit the spread of transposons and the integration of retroviruses. This operational framework would also explain why gene bodies (i.e., open reading frames) are targets for both histone modifications promoting nucleosome re-deposition after transcription and DNA methylation (Smolle et al., 2013; Zilberman et al., 2007). Shielding gene bodies from integration of a transposon or retrovirus likely helps host and parasite alike by avoiding insertion into essential sequences. Although it has been suggested that chromatin modifications such as H3 lysine 36 methylation in gene bodies serve to prevent the use of cryptic promoters (Smolle et al., 2013), the hypothesis put forth here posits that this may not be the ancestral or critical current function for this modification. Rather, this and other chromatin modifications may have evolved to suppress the integration of parasitic elements into essential genes by maintaining nucleosome density to as high a degree as possible. Archaea also possess histones, a nucleosome-like structure, and nucleosome-free regions, indicating that chromatin evolved prior to the existence of eukaryotes (Ammar et al., 2012). As with eukaryotes, it may have evolved to protect archaeal genomes from transposon integration.

Dinoflagellates are unusual among eukaryotes, as they lack histones and instead pack genomic DNA in a paracrystalline array. They also have very large genomes with a high content of repeat sequences. A recent study shows that dinoflagellates acquired a nonhistone DNA-packaging protein (perhaps from an algal virus) around the time when their histones were lost and their genome massively increased (Gornik et al., 2012). Further study of these organisms may allow for reconstruction of the evolutionary events occurring after loss of canonical chromatin and assessment of whether the chromatin loss may have triggered the increase in genome size via proliferation of transposable elements.



Conflict in the Eukaryotic Cell
Artistic interpretation of the themes of this Essay. The proposed defensive role of attributes of the eukaryotic cell is fancifully rendered. Illustration by Teny Issakhanian.

Coactivator Relics of Ancient Conflicts

The number of factors required for the relevant RNA polymerase to initiate transcription on mRNA-coding genes is vastly larger in eukaryotes than in prokaryotes. The general transcription factors (TFIIA, B, D, E, F, G, H) required for basal transcription and the Mediator complex required for activated transcription are highly conserved across eukaryotes. The complexity of Mediator, whose mass exceeds that of RNA Pol II itself, is remarkable. The orthodox view is that this complexity offers opportunities for regulation. However, phylogenetic studies suggest that a 17-subunit Mediator complex existed in the protoeukaryote some 1–2 billion years ago, likely prior to the

evolution of significant eukaryotic complexity (Bourbon, 2008). Why did such a complex transcriptional coactivator evolve at such an early stage?

A resolution to this conundrum comes out of the fact that transposons and retroviruses utilize host transcriptional activators along with Mediator to express their genes. In principle, this dependence may have created a situation in which the protoeukaryote was under pressure to alter its gene expression requirements via changes or elaborations of Mediator to block the binding of activators used by transposon and retroviruses, helping to avoid lethal parasitism by selfish nucleic acids. Parasites would then have been subject to selection for adaptations, permitting continued gene expression in the host (namely the acquisition of binding sites of other host activators). Such an arms race could have driven the observed complexity in eukaryotic Pol II initiation machinery and in Mediator—in particular at a time near the birth of the eukaryotic lineage before complex eukaryotes arose.

In support of this hypothesis, recent work suggests that the evolution of paralogs of subunits of the coactivator TFIID (composed of the TATA-box-binding protein TBP and associated TAFs) may have been driven by genetic conflict. For example, testis-specific TAFs (tTAFs) in *Drosophila* have evolved very rapidly, and there is evidence for positive selection and coevolution among the tTAFs (Li et al., 2009). Although it remains unclear what drove this fast pace of evolution, conflicts with parasitic nucleic acids are a possibility.

An extant role of the transcriptional machinery in genome defense comes from plants. Plants contain two RNA polymerases missing from other eukaryotes called Pol IV and Pol V (Pikaard et al., 2013). These enzymes are paralogs of Pol II but do not function to synthesize mRNA. Rather, they suppress transposable elements. Pol IV is recruited to many of its targets by a protein that recognizes repressive histone methylation on H3 lysine 9 (Law et al., 2013), where it synthesizes uncapped and nonpolyadenylated transcripts that are then used as templates for an RNA-dependent RNA polymerase. Small RNAs produced from the ensuing double-stranded RNA were used to target repressive histone and DNA methylation factors to transposons transcribed by the other plant-specific RNA polymerase, Pol V (presumably by base-pairing between the small RNAs and the nascent transcript). Much remains to be learned about how Pol IV and Pol V function to recognize and silence transposons, but their established function in genome defense makes it plausible that core components of the gene expression machinery can evolve because of the need to suppress parasitic elements.

Antiviral Identification Cards for Messenger RNA

Eukaryotic premessenger RNAs bear a distinctive nucleotide structure at their 5′ ends termed the "cap." Through a series of posttranscriptional modification reactions, enzymes install a 5′-5′-linked 7-methylguanine nucleotide on the free end, along with a cluster of 2′-O-methyl groups. Eubacteria and archaea lack this structure. (Shuman, 2002). Caps are required to prevent RNA degradation, to stimulate splicing and RNA export, and to recruit ribosomes. However, caps are also an impediment to viruses, which must conform to or circumvent this eukaryote-specific requirement for gene expression. According to this framework, caps evolved as another roadblock to viral gene expression.

RNA viruses have met this challenge in a variety of ways (Walsh et al., 2013). Some, such as hepatitis C virus, solve the problem by replicating in the cytoplasm and evolving internal ribosome entry sites (IRESs) to permit RNA translation independent of a cap. Viruses that have evolved IRESs often go a step further to shut down all cap-dependent host translation (Ehrenfeld, 1982). Others, such as influenza A virus, steal caps from host mRNAs. Several other cytoplasmic viruses, such as flaviviruses and poxviruses, encode their own capping enzymes. Retroviruses integrate into the host genome and use RNA Pol II and host-capping enzymes. Thus, all currently successful selfish DNAs have evolved complex mechanisms for either meeting or bypassing the cap requirement.

In the case of cap bypass, eukaryotic cells have evolved an additional related means of antiviral defense. Viral RNAs containing 5′ triphosphates instead of caps are recognized in the cytosol by receptors such as RIG-I proteins, which then trigger an antiviral response by the innate immune system (Loo and Gale, 2011). The 2′-*O*-methylation of riboses near the 5′ ends of eukaryotic mRNAs serve as additional signals to distinguish normal capped cellular pre-mRNAs from viral RNAs (Daffis et al., 2010). Interestingly, the evolution of the eukaryotic cap methylation machinery appears to have involved horizontal gene transfer between eukaryotes and viruses, supporting the possibility of conflict between host and viruses with respect to these sugar modifications of mRNA (Werner et al., 2011).

Similar to 5′ cap structures, 3′ polyA tails are required for mRNA stability in eukaryotic cells. Consistent with the idea that polyadenylation may have evolved to allow cells to distinguish normal cellular RNAs from those of viral parasites, all successful viruses either encode polyA at the end of their transcripts or utilize mechanisms for protection of transcripts against degradation by eukaryotic 3′-5′ exonucleases. Further support for this notion comes from studies of genetic suppressors of yeast killer virus. *Saccharomyces cerevisiae SKI* genes are bona fide antiviral genes required to suppress the replication of a toxin-encoding cytoplasmic RNA virus. *SKI* gene products include ones promoting the 5′-3′ degradation of uncapped transcripts, as well as 3′-5′ degradation of deadenylated RNAs (Araki et al., 2001).

The Spliceosome: A Transposon-Derived Transposon Censor

Introns were perhaps the biggest surprise uncovered when eukaryotic genes were characterized at the molecular level. Similarities in the chemistry of splicing between spliceosomal pre-mRNAs and the RNA-only self-splicing group II introns have long supported the hypothesis of a common origin (van der Veen et al., 1986). This idea has gained further support as similarities between them in both primary and secondary structures have been recognized (Madhani and Guthrie, 1992; Yu et al., 1995). A recent crystal structure of the spliceosomal protein Prp8 exhibits similarities to group II intron-encoded protein maturases (Galej et al., 2013). Thus, there is strong

evidence that the eukaryotic spliceosome arose from the invasion of mobile ancestors of modern-day group II introns (likely from the α-proteobacterial ancestor of mitochondria) and their associated protein maturases (Rogozin et al., 2012).

Some have argued that there is no need to invoke selection for the maintenance of introns: their rates of loss are slow, they have minimal negative impact on fitness as they are removed by splicing, and there are mechanisms for intron gain (Rogozin et al., 2012). Some introns have been modified for alternative RNA splicing or for encoding snoRNAs and miRNAs. However, it remains an open question whether these examples are sufficient to account for the maintenance of introns. Indeed, it has been suggested that something may be missing from the intron puzzle, as there are no known examples of a free-living eukaryote that has lost all of its introns and the spliceosome (Rogozin et al., 2012). Although there are examples of genomes containing introns in as few as 2% of genes, why are there no intron-free eukaryotes?

Our recent discovery of a function for introns and the spliceosome in transposon recognition offers an explanation. We found that, in the intron-rich yeast *Cryptococcus neoformans*, siRNAs suppress the movement of transposons. Several lines of evidence suggest that, in this organism, nonoptimal splicing signals result in the stalling of transposon-related mRNA precursors on the spliceosome (Dumesic et al., 2013). Stalling is required for targeting of the splicing intermediates to an RNA-dependent RNA polymerase complex and, ultimately, the production of complementary suppressive siRNAs. These data indicate that the spliceosome and RNA splicing serve to monitor the genome (Dumesic et al., 2013). Thus, beyond their better-known role as an impediment to gene expression, introns play an additional role as a mechanism for discriminating foreign elements, which are then targeted for silencing by small RNAs. Additional circumstantial evidence for a role for stalled spliceosomes in triggering small RNA production exists in protists, fungi, plants, and animals. One example comes from studies of the *Drosophila* germline genome defense system in which primary transcripts from specific chromosomal clusters containing a history of transposon insertions are processed into protective piRNAs. The mechanism by which these transcripts (and not others) are targeted for piRNA biogenesis is not understood. However, the recent identification of a mutation in an RNA splicing and export factor that produces a defect in piRNA biogenesis is intriguing (Zhang et al., 2012).

Gateway to the Genome

Eukaryotic cells segregate their genomes in the nucleus, away from the cytoplasm using the nuclear envelope. The evolution of the nucleus is a topic of much debate. Regardless of its origins, the nuclear envelope, like chromatin, offers a shield for the genome. Retroviruses and retrotransposons, which require genome integration for replication, must navigate this barrier. Indeed, it has been suggested that the nucleus evolved in response to a singular retroelement invasion, namely that of the group II ribozyme presumed to be the ancestor of modern-day introns and the spliceosome (Koonin, 2006). Other classes of parasitic nucleic acids arguably may have represented a greater threat than group II introns, as they do not have the ability to mitigate their impact by precisely splicing themselves out of RNA transcribed from their insertion site. DNA viruses that require host RNA Pol II, spliceosomes, and polyadenylation factors (e.g., adenovirus) have also evolved elaborate mechanisms to gain entry into the protectorate established by the nuclear envelope (Greber et al., 1993). RNA viruses that replicate only in the cytoplasm avoid having to invade the nucleus but face the challenge of producing mRNAs that can be translated and are resistant to mRNA degradation.

Another consequence of nucleocytoplasmic separation is the requirement for mRNAs to be exported to the cytoplasm for translation. In many eukaryotes, this requires the deposition of the exon-junction complex (EJC) via RNA splicing (Le Hir and Séraphin, 2008). Retrotranposons and the full-length genomic transcripts of retroviruses necessarily lack introns, which force them to evolve mechanisms for export that are independent of this mechanism. For example, HIV-1 encodes a dedicated export factor Rev to export unspliced genomic RNA (Cullen, 2003). Thus, the combination of a splicing-coupled EJC deposition required for RNA export and nucleocytoplasmic segregation can be viewed as antiretroviral strategy.

Degrade First and Ask Questions Later

Retroviruses and retrotransposons produce all of the key viral enzymes—protease, reverse transcriptase, RNase H, and integrase—via unusual mechanisms involving either stop codon readthrough or ribosome frameshifting. These measures are required due to the presence of a premature termination codon in the genomic RNA encoding the gag-pol polyprotein. If the stop codon is utilized, only the gag polyporotein is produced, but if it is bypassed, gag-pol is synthesized and viral enzymes are produced that fuel viral replication. This feature of retroelements may have driven the evolution of a eukaryotic ribosome-associated mechanism for the detection of premature termination codons: namely, the nonsense-mediated decay (NMD) pathway that destroys mRNAs whose coding segments are not translated. In *S. cerevisiae*, for instance, ribosomal frameshifting signals in endogenous genes have been shown to destabilize mRNAs, in part, via NMD (Belew et al., 2011). Moreover, genomic RNAs of the HTLV-I retrovirus have recently been identified as NMD substrates. Strikingly, the viral RNA-binding protein Rex blocks NMD, presumably to defend against this antiviral mechanism (Nakano et al., 2013). Although NMD is widely thought to have evolved to deal with introns and inefficient splicing (Koonin, 2006), it is also plausible that detection of premature stop codons in retroelement transcripts played a role in its appearance and/or maintenance. Similarly, other mRNA decay pathways, the no-go and nonstop decay pathways (Parker, 2012), detect ribosome stalling and may have evolved to detect features of transposon and viral RNAs.

Genome Defense Roots of RNA Regulators

Although the evolution of many of the mechanisms described above has been ascribed to different pressures, there is agreement that one class of players in modern-day gene expression, small

RNAs, initially evolved because of the need for genome defense (Cerutti and Casas-Mollano, 2006; Shabalina and Koonin, 2008). These noncoding RNAs and the PIWI-PAZ domain proteins that mediate their action are the focus of much recent work (Joshua-Tor and Hannon, 2011). miRNAs, for example, play diverse cellular regulatory roles through both RNA transcript stability and translational control. Related RNAi systems, in turn, function to protect eukaryotic cells from RNA viruses through detection of the double-strand RNA intermediates of viral replication and subsequent triggering of a small RNA defensive response. Plants, animals, and fungi all use RNA silencing to defend against RNA viruses (Ding and Voinnet, 2007), and the shared requirement in both miRNA and RNAi pathways for the Dicer endonuclease and Argonaute proteins suggests that miRNA systems evolved from more ancient RNAi systems. Thus, it has been widely proposed (Cerutti and Casas-Mollano, 2006; Shabalina and Koonin, 2008) that genome defense against selfish nucleic acids (endogenous siRNA pathways) may have indirectly driven the evolution of a current feature of eukaryotic gene expression (miRNAs). The *Drosophila* piRNA/Piwi system offers another example of this concept, as it appears to be a system evolved to defend the genome against transposable elements in the germline yet has additional developmental functions (Peng and Lin, 2013). The arguments that I make in this essay extend the concept to essentially all eukaryotic gene expression elaborations.

Prokaryotes: Less Defended but Also Less Frustrated?

The notion that parasitic nucleic acids drove the evolution of the extant features of eukaryotic gene expression begs the question of how prokaryotes defend themselves without access to the defense capacity afforded by the additional steps of gene expression that are characteristic of eukaryotes. Part of the answer may be that natural selection is likely to be considerably stronger in prokaryotes because their population sizes are larger (Lynch and Conery, 2003). Thus, individuals harboring a parasitic nucleic acid may be removed by selection before the population fixation of the deleterious element is achieved. The high ploidy of replicons and the lack of obligate sex may also mitigate the impact of transposons. However, in many instances, prokaryotes do benefit from one or more genome defense system. Restriction-modification systems (which use DNA methylation to distinguish self from nonself) and CRISPR-CAS systems (which use a cache of small RNAs to recognize and cleave parasitic DNAs) are important arbiters of virus and transposon resistance (Makarova et al., 2013). These systems, though widespread, are far from universal, which again may be related to the power of selection in prokaryotic populations. In terms of the metaphor introduced above, prokaryotic genomes would appear to be less defended but also less frustrated in their expression than their eukaryotic counterparts.

Beyond the Frustrated Gene

Eukaryotic gene expression may thus be seen as a complex defense bureaucracy that employs genome access restriction (chromatin), transcript inspectors (the spliceosome and the ribosome), and security gates (the nuclear pore). Other aspects of eukaryotic cell biology may also have evolved to combat external threats. For example, it has been suggested that the evolution of apoptosis was driven by the need to rid a eukaryotic ancestor of intracellular bacteria: a prokaryotic pattern was detected—namely, bacterial cytochrome *c*—triggering a self-destruct sequence in the infected cell and thereby protecting other members of the population (James and Green, 2002). In this scenario, apoptosis preceded the appearance of a stable mitochondrial symbiont in which cytochrome *c* release became a regulated event. The evolution of the unfolded protein response may have evolved as an alarm system for viral infection, detecting ER stress produced by the massive load of viral glycoproteins, followed by triggered destruction of ER-associated (including viral) mRNAs and translation shut-down (Hollien, 2013). Many other eukaryotic features (autophagy, lysosomes, and endocytosis to name a few) may have evolved for similar reasons. Understanding the eukaryotic cell in this way requires a mindset in which parasitism and defense are seen as central drivers of biological innovation.

ACKNOWLEDGMENTS

I thank Matt Daugherty, Victor Hansen-Smith, Suzanne Noble, Joy Noble, Alok Srivastava, Jordan Burke, Phillip Dumesic, and Leslie Spector for comments on drafts; Steve Jacobsen for clarifying the plant literature; Harmit Malik, Patrick O'Farrell, and Olivier Voinnet for stimulating conversations; and the three anonymous reviewers for their criticism. I apologize to those whose work could not be cited due to citation limits. This work was supported by a grant from the NIH (GM071801).

REFERENCES

Ammar, R., Torti, D., Tsui, K., Gebbia, M., Durbic, T., Bader, G.D., Giaever, G., and Nislow, C. (2012). Chromatin is an ancient innovation conserved between Archaea and Eukarya. Elife *1*, e00078.

Araki, Y., Takahashi, S., Kobayashi, T., Kajiho, H., Hoshino, S., and Katada, T. (2001). Ski7p G protein interacts with the exosome and the Ski complex for 3′-to-5′ mRNA decay in yeast. EMBO J. *20*, 4684–4693.

Belew, A.T., Advani, V.M., and Dinman, J.D. (2011). Endogenous ribosomal frameshift signals operate as mRNA destabilizing elements through at least two molecular pathways in yeast. Nucleic Acids Res. *39*, 2799–2808.

Bourbon, H.M. (2008). Comparative genomics supports a deep evolutionary origin for the large, four-module transcriptional mediator complex. Nucleic Acids Res. *36*, 3993–4008.

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

Craig, N., Craigie, R., Gellert, M., and Lambowitz, A. (2002). Mobile Genetic Elements II (Washington, DC: ASM Press).

Cullen, B.R. (2003). Nuclear mRNA export: insights from virology. Trends Biochem. Sci. *28*, 419–424.

Daffis, S., Szretter, K.J., Schriewer, J., Li, J., Youn, S., Errett, J., Lin, T.Y., Schneller, S., Zust, R., Dong, H., et al. (2010). 2′-O methylation of the viral mRNA cap evades host restriction by IFIT family members. Nature *468*, 452–456.

Dawkins, R. (1976). The Selfish Gene (Oxford, UK: Oxford University Press).

Ding, S.W., and Voinnet, O. (2007). Antiviral immunity directed by small RNAs. Cell *130*, 413–426.

Dumesic, P.A., Natarajan, P., Chen, C., Drinnenberg, I.A., Schiller, B.J., Thompson, J., Moresco, J.J., Yates, J.R., 3rd, Bartel, D.P., and Madhani, H.D. (2013). Stalled spliceosomes are a signal for RNAi-mediated genome defense. Cell *152*, 957–968.

Ehrenfeld, E. (1982). Poliovirus-induced inhibition of host-cell protein synthesis. Cell *28*, 435–436.

Galej, W.P., Oubridge, C., Newman, A.J., and Nagai, K. (2013). Crystal structure of Prp8 reveals active site cavity of the spliceosome. Nature *493*, 638–643.

Gornik, S.G., Ford, K.L., Mulhern, T.D., Bacic, A., McFadden, G.I., and Waller, R.F. (2012). Loss of nucleosomal DNA condensation coincides with appearance of a novel nuclear protein in dinoflagellates. Curr. Biol. *22*, 2303–2312.

Greber, U.F., Willetts, M., Webster, P., and Helenius, A. (1993). Stepwise dismantling of adenovirus 2 during entry into cells. Cell *75*, 477–486.

Hickey, D.A. (1982). Selfish DNA: a sexually-transmitted nuclear parasite. Genetics *101*, 519–531.

Hollien, J. (2013). Evolution of the unfolded protein response. Biochim. Biophys. Acta *1833*, 2458–2463.

James, E.R., and Green, D.R. (2002). Infection and the origins of apoptosis. Cell Death Differ. *9*, 355–357.

Joshua-Tor, L., and Hannon, G.J. (2011). Ancestral roles of small RNAs: an Ago-centric perspective. Cold Spring Harb. Perspect. Biol. *3*, a003772.

Koonin, E.V. (2006). The origin of introns and their role in eukaryogenesis: a compromise solution to the introns-early versus introns-late debate? Biol. Direct *1*, 22.

Law, J.A., Du, J., Hale, C.J., Feng, S., Krajewski, K., Palanca, A.M., Strahl, B.D., Patel, D.J., and Jacobsen, S.E. (2013). Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. Nature *498*, 385–389.

Le Hir, H., and Séraphin, B. (2008). EJCs at the heart of translational control. Cell *133*, 213–216.

Li, V.C., Davis, J.C., Lenkov, K., Bolival, B., Fuller, M.T., and Petrov, D.A. (2009). Molecular evolution of the testis TAFs of Drosophila. Mol. Biol. Evol. *26*, 1103–1116.

Loo, Y.M., and Gale, M., Jr. (2011). Immune signaling by RIG-I-like receptors. Immunity *34*, 680–692.

Lynch, M., and Conery, J.S. (2003). The origins of genome complexity. Science *302*, 1401–1404.

Madhani, H.D., and Guthrie, C. (1992). A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. Cell *71*, 803–817.

Makarova, K.S., Wolf, Y.I., and Koonin, E.V. (2013). Comparative genomics of defense systems in archaea and bacteria. Nucleic Acids Res. *41*, 4360–4377.

Nakano, K., Ando, T., Yamagishi, M., Yokoyama, K., Ishida, T., Ohsugi, T., Tanaka, Y., Brighty, D.W., and Watanabe, T. (2013). Viral interference with host mRNA surveillance, the nonsense-mediated mRNA decay (NMD) pathway, through a new function of HTLV-1 Rex: implications for retroviral replication. Microbes Infect. *15*, 491–505.

Parker, R. (2012). RNA degradation in Saccharomyces cerevisae. Genetics *191*, 671–702.

Peng, J.C., and Lin, H. (2013). Beyond transposons: the epigenetic and somatic functions of the Piwi-piRNA mechanism. Curr. Opin. Cell Biol. *25*, 190–194.

Pikaard, C.S., Haag, J.R., Pontes, O.M., Blevins, T., and Cocklin, R. (2013). A transcription fork model for Pol IV and Pol V-dependent RNA-directed DNA methylation. Cold Spring Harb. Symp. Quant. Biol. *77*, 205–212.

Rogozin, I.B., Carmel, L., Csuros, M., and Koonin, E.V. (2012). Origin and evolution of spliceosomal introns. Biol. Direct *7*, 11.

Roth, S.L., Malani, N., and Bushman, F.D. (2011). Gammaretroviral integration into nucleosomal target DNA in vivo. J. Virol. *85*, 7393–7401.

Shabalina, S.A., and Koonin, E.V. (2008). Origins and evolution of eukaryotic RNA interference. Trends Ecol. Evol. *23*, 578–587.

Shuman, S. (2002). What messenger RNA capping tells us about eukaryotic evolution. Nat. Rev. Mol. Cell Biol. *3*, 619–625.

Smolle, M., Workman, J.L., and Venkatesh, S. (2013). reSETting chromatin during transcription elongation. Epigenetics *8*, 10–15.

van der Veen, R., Arnberg, A.C., van der Horst, G., Bonen, L., Tabak, H.F., and Grivell, L.A. (1986). Excised group II introns in yeast mitochondria are lariats and can be formed by self-splicing in vitro. Cell *44*, 225–234.

Walsh, D., Mathews, M.B., and Mohr, I. (2013). Tinkering with translation: protein synthesis in virus-infected cells. Cold Spring Harb. Perspect. Biol. *5*, a012351.

Werner, M., Purta, E., Kaminska, K.H., Cymerman, I.A., Campbell, D.A., Mittra, B., Zamudio, J.R., Sturm, N.R., Jaworski, J., and Bujnicki, J.M. (2011). 2′-O-ribose methylation of cap2 in human: function and evolution in a horizontally mobile family. Nucleic Acids Res. *39*, 4756–4768.

Yu, Y.T., Maroney, P.A., Darzynkiwicz, E., and Nilsen, T.W. (1995). U6 snRNA function in nuclear pre-mRNA splicing: a phosphorothioate interference analysis of the U6 phosphate backbone. RNA *1*, 46–54.

Zhang, F., Wang, J., Xu, J., Zhang, Z., Koppetsch, B.S., Schultz, N., Vreven, T., Meignin, C., Davis, I., Zamore, P.D., et al. (2012). UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery. Cell *151*, 871–884.

Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T., and Henikoff, S. (2007). Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. Nat. Genet. *39*, 61–69.