United we stand: Big roles for small RNA gene clusters

Brice Felden, Luc Paillard

United we stand: Big roles for small RNA gene clusters

Bacterial and eukaryotic immunity RNA clusters share unexpected strong structural and functional resemblances in fighting DNA intruders.

Brice Felden[a,b] and Luc Paillard[b,c]

[a] Inserm U835 Biochimie Pharmaceutique, University of Rennes 1, Rennes, France. [b] Biosit, University of Rennes 1/Université Européenne de Bretagne, Rennes, France. [c] Centre National de la Recherche Scientifique UMR 6290, Institut de Génétique et Développement de Rennes, Rennes, France.

Corresponding authors: brice.felden@univ-rennes1.fr, luc.paillard@univ-rennes1.fr

Regulatory small RNAs (sRNAs) are essential for protecting living cells against potentially harmful DNA challenges. They are found in all domains of life, and the parallels between their functions and mechanistic properties in eukaryotes and prokaryotes are striking. Their shared key role is to specify the targets of their associated effectors. During the interference phase, the sRNAs guide ribonucleoproteic complexes to the proper nucleic acid targets by base-pairing. This guidance capacity is surprisingly conserved in prokaryotes and eukaryotes (Figure 1), so it may be a leftover from an ancestral RNA world where the RNA was responsible both for the interaction specificity and the biochemical fate of the targets.

The sRNAs of each domain differ in their partners and biological consequences. In prokaryotes, CRISPR RNAs (clustered regularly interspaced short palindromic repeat RNAs, crRNAs) guide Cas proteins to foreign DNA to trigger endonucleolytic cleavage, thus preventing bacteriophage infections and plasmid invasions (1). In eukaryotes, PIWI-interacting RNAs (piRNAs) guide complexes that include PIWI, a member of the Argonaute (Ago) superfamily of proteins, to transposon DNA. This induces gene silencing through epigenetic modifications, and hence represses transposition. In eukaryotes too, short, double-stranded RNAs (small interfering RNAs, siRNAs) associate into a complex also containing Ago, and target it to a complementary RNA to trigger its degradation. If the siRNA comes from a double-stranded viral RNA, this mechanism contributes to host antiviral defenses (2, 3). Ago proteins are the central hubs of sRNA-mediated silencing devices in eukaryotes, but recent evidences also highlight the involvement of prokaryotic Ago in protecting bacterial genomes against foreign, and possibly invasive, genomic elements such as plasmids (4, 5). The molecular mechanisms that control sRNA association with the effector complex differ between crRNAs, piRNAs, and siRNAs. The prokaryotic crRNAs interact by pairing with another RNA, tracrRNA, which tethers them to Cas proteins. There is no known eukaryotic equivalent for tracrRNA, and how a transcribed RNA is recognized as a piRNA to be processed and linked to PIWI is currently not well understood. Meanwhile, a double-strand conformation triggers siRNA recognition by Ago. Thanks to this limited structural requirement, the siRNA pathway is used in reverse genetics to silence gene expression by RNA interference.

Dissimilarities aside, eukaryotic and prokaryotic sRNAs do both act as guide RNAs. Another resemblance is their shared genomic organization as clusters. Gene clusters are associations of genes expressing similar macromolecules with shared functions. Most investigations have focused on protein gene clusters (*e.g.* the vertebrate *hox* genes), but sRNAs are also frequently produced from clusters. piRNA cluster sequences are poorly conserved throughout evolution, despite the strong prevalence of the clustered organization itself, and they share extensive sequence conservation with non-functional transposon fragments. This suggests that they may have formed

by aggregating captured transposons sequentially (3), thereby expanding cluster complexity. It is unknown whether this appropriation is due to the random insertion of active transposons into a preexisting piRNA cluster, or if there is a direct capture mechanism. siRNAs are exogenous (viral infections or experimentally introduced nucleic acid sequences), and there are no documented siRNA gene clusters. However, siRNAs substantially resemble another class of eukaryotic sRNAs, the microRNAs (miRNAs). miRNAs differ form siRNAs by their origins, as they are all genome-encoded. Cellular miRNAs have antiviral functions in mammalian cells (6), but their major function is to regulate endogenous gene expression. Most often in animals, they imperfectly pair with their target RNAs, and the primary outcome of miRNA-mediated regulations is translational repression. Interestingly, miRNA genes are clustered. miRNA gene clusters may have come about due to the co-transcriptional recruitment of miRNA processing enzymes, which facilitates the emergence of new sRNAs near existing ones (2). In addition, miRNA genes, within a cluster, share sequence conservation, suggesting that these clusters have expanded by local duplications. miRNA gene clusters may be subjected to duplications and to the subsequent appearance of new miRNA genes as a consequence of genomic imprinting. Indeed, miRNA clusters are enriched in imprinted regions (7). However, the driving forces for eukaryotic miRNA cluster formation remain essentially elusive.

The molecular events that direct prokaryotic crRNA cluster dynamics are far better understood (Figure 1). During adaptation, bacteria and archaea acquire new sequences from foreign DNA, and once integrated these spacers serve as templates for crRNA synthesis. Three recent studies explained foreign DNA acquisition preference and how spacer invaders are captured. Spacer acquisition is replication-dependent, and replication forks are more frequent on multicopy plasmids than on the chromosome. In addition, Chi sites, highly repeated octamer sequences on the bacterial chromosome, negatively impact spacer acquisition. Together, these observations explain how non-self DNA is preferred over the self for spacer acquisition (8). Two structural studies revealed how the Cas1-Cas2 complex acts as an integrase (9, 10). This conserved complex binds a protospacer sequence to catalyze spacer acquisition. These recent works showed how protospacer length is predetermined, selected, processed into a spacer and integrated into the bacterial genome. These studies also showed that the Cas1-Cas2 complex has a similar structure to the eukaryotic transposases used by transposons to integrate into chromosomes.

The comparison of eukaryotic and prokaryotic sRNA clusters raises many questions. Are similar mechanisms at play in both domains? Can the growing knowledge gleaned from bacteria help us to understand eukaryotic cluster dynamics? What selection pressure induced such a peculiar genomic organization throughout evolution? Bringing together genes involved in similar functions is often seen to ease their co-regulation. Furthermore, piRNA clusters have specific epigenetic profiles (11) that may help to tag the transcribed RNAs as being piRNAs. Assembling sRNA genes into functional immunity clusters is probably a way to improve the efficiency of defenses against DNA intruders in cells from all domains of life.

References

1. B. Wiedenheft *et al.*, *Nature* **482**, 331 (2012).
2. M.J. Axtell *et al.*, *Genome Biol* **12**, 221 (2011).
3. B. Czech, G.J. Hannon. *Trends Biochem Sci.* **41**, 324 (2016).
4. I. Olovnikov *et al.*, *Mol Cell* **51**, 594 (2013).
5. D.C. Swarts *et al.*, *Nature* **507**, 258 (2014).
6. I.M. Pedersen *et al.*, *Nature* **449**, 919 (2007).
7. M. Girardot *et al.*, *Epigenetics* **7**, 1341 (2012).

8. A. Levy *et al.*, *Nature* **520**, 505 (2015).
9. J.K. Nunez *et al.*, *Nature* **527**, 535 (2015).
10. J. Wang *et al.*, *Cell* **163**, 840 (2015).
11. Le Thomas *et al.*, *Cell Rep* **25**, 1617 (2014).

Figure 1. **Comparison showing the close resemblance of prokaryotic and eukaryotic immunity cluster RNA.** Molecular defense against DNA intruders in the two domains was subdivided into nine distinct steps. During the initial adaptation phase, invasive DNAs (phages, plasmids, or transposons) are inserted into specialized genomic clusters. During the following interference phase, RNAs transcribed from these clusters associate into ribonucleoproteic complexes (yellow ovals) to target the invasive phage DNA for destruction in prokaryotes, or the transposon DNA for transcriptional gene silencing (TGS) in eukaryotes. sRNAs transcribed from clustered miRNA genes, as well as viral siRNAs, regulate post-transcriptionally gene expression by translational inhibition and/or RNA degradation (post-transcriptional gene silencing, PTGS).

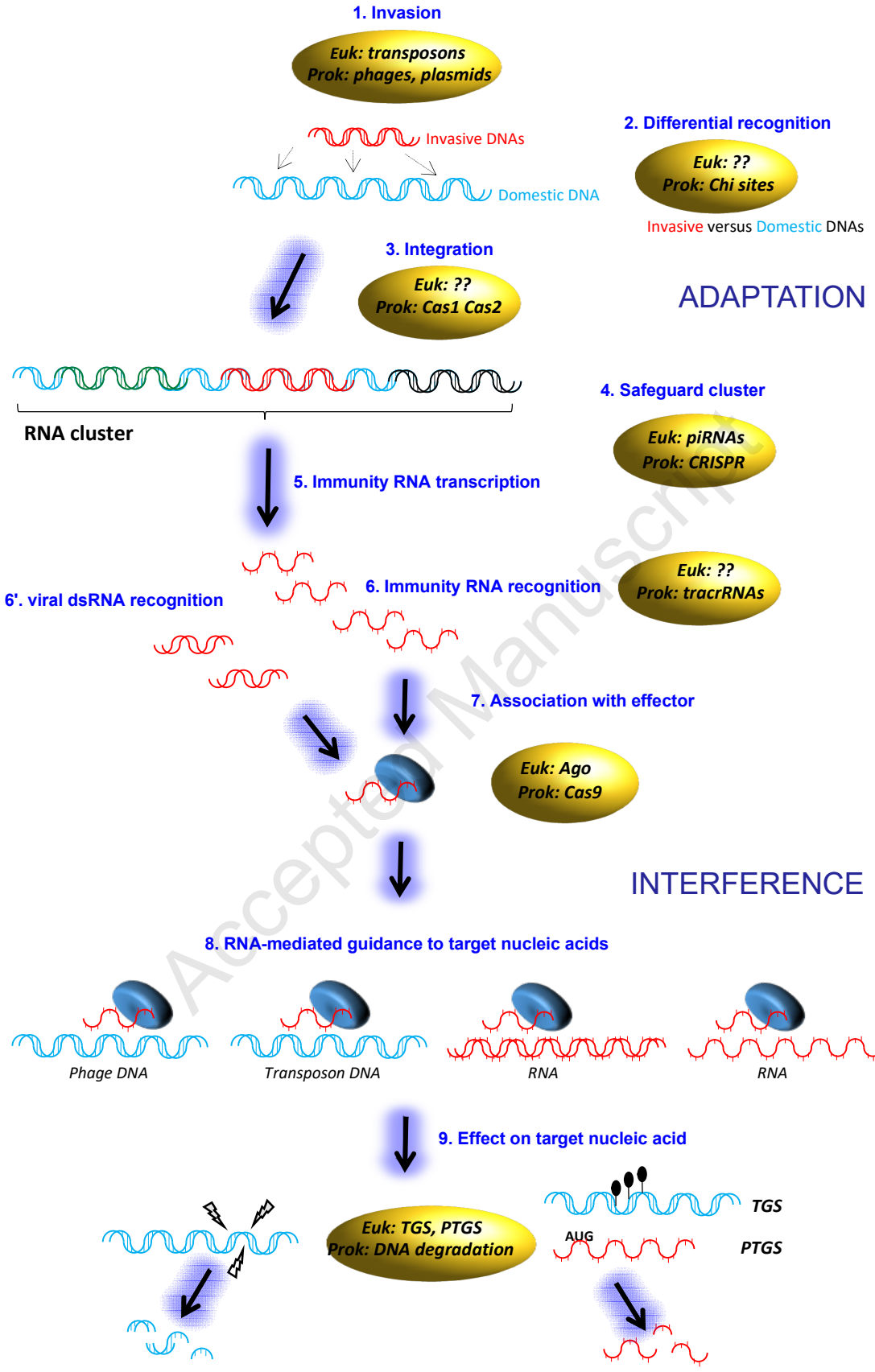

Figure 1