in the tentacles and peduncle, respectively, and by position likely represent the CB circuit (likely innervated by Ec2 and Ec4 sensory neurons in tentacles and hypostome). The ec3 ganglion cell clusters uniquely span the basal disk ectoderm and therefore likely represent the RP1 cluster active in the basal disk. Finally, they have proposed that the endodermal en1 ganglion cell neurons make up the RP2 circuit, likely innervated by en2 sensory cells [2].

Figure 1D shows the distinct transcription factor signatures of neuronal cell types. Some important lessons can be learned. First, all neurons (and related gland cells) specifically express transcription factors of the class A basic helix–loop–helix family (*E12/E47*, *AscA*, *Delilah*, *Ptf1b*, and *Neurogenin*), supporting the ancestral and conserved role of this family in neuronal specification. Second, neuron types representing the three subnetworks CB, RP1, and RP2, also differ by transcription factor signature, expressing *Zic* (ec1,2, and 4), *GATA* (ec3), *otp* (en2 and 3), or *Lmx* (en1), (red, violet, and gray in Figure 1D). Three of these factors have also been identified to specifically demarcate neuronal subgroups in sea anemone [5], opening up the exciting possibility that some subnetworks already existed in cnidarian ancestors. More comparative work in these and other cnidarian species will be needed to trace the evolution of nerve nets in molecular and cellular detail across the cnidarian phylum.

[1]Developmental Biology Unit, European Molecular Biology Laboratory, Meyerhofstrasse 1, 69012 Heidelberg, Germany

[2]Centre for Organismal Studies, University of Heidelberg, Im Neuenheimer Feld 230, 69120 Heidelberg, Germany

*Correspondence:
arendt@embl.de

https://doi.org/10.1016/j.tig.2019.09.003

References

1. Bosch, T.C. *et al.* (2010) The Hydra polyp: nothing but an active stem cell community. *Dev. Growth Differ.* 52, 15–25
2. Siebert, S. *et al.* (2019) Stem cell differentiation trajectories in *Hydra* resolved at single-cell resolution. *Science.* Published online July 26, 2019. https://doi.org/10.1126/science.aav9314.
3. Marioni, J.C. and Arendt, D. (2017) How single-cell genomics is changing evolutionary and developmental biology. *Annu. Rev. Cell Dev. Biol.* 33, 537–553
4. Marlow, H. *et al.* (2013) Ectopic activation of the canonical Wnt signaling pathway affects ectodermal patterning along the primary axis during larval development in the anthozoan *Nematostella vectensis*. *Dev. Biol.* 380, 324–334
5. Sebe-Pedros, A. *et al.* (2018) Cnidarian cell type diversity and regulation revealed by whole-organism single-cell RNA-Seq. *Cell* 173, 1520–1534
6. Arendt, D. *et al.* (2015) From nerve net to nerve ring, nerve cord and brain – evolution of the nervous system. *Nat. Rev. Neurosci.* 17, 61–72
7. Leclere, L. *et al.* (2019) The genome of the jellyfish *Clytia hemisphaerica* and the evolution of the cnidarian life-cycle. *Nat. Ecol. Evol.* 3, 801–810
8. Rankin, S.A. *et al.* (2011) A gene regulatory network controlling hhex transcription in the anterior endoderm of the organizer. *Dev. Biol.* 351, 297–310
9. Steinmetz, P.R.H. *et al.* (2017) Gut-like ectodermal tissue in a sea anemone challenges germ layer homology. *Nat. Ecol. Evol.* 1, 1535–1542
10. Dupre, C. and Yuste, R. (2017) Non-overlapping neural networks in *Hydra vulgaris*. *Curr. Biol.* 27, 1085–1097

Forum

A Genetic Instruction Code Based on DNA Conformation

Alan Herbert[1,*]

Flipons are sequences capable of forming either right- or left-handed DNA under physiological conditions, forming a class of dissipative structures that trade metabolic energy for information by cycling DNA between different chromatin states. Flipons enhance the diversity of transcriptomes, increasing entropy while enabling the evolution of features both new and unexpected.

The Unexpected Answer to a Question Nobody Was Asking

DNA comes in many different forms [1]. The Watson and Crick B-DNA conformer is a right-handed helix, while, in Z-DNA, the bases are flipped over, causing the phosphate backbone to zag left [2] (Figure 1A). Z-DNA was an accidental discovery found in the first-ever synthetic DNA crystal. Z-DNA forms dynamically under physiological conditions when certain enzymes untwist right-handed DNA while working to make RNA. The *in vivo* relevance of Z-DNA was an unresolved question until recently [3]. One early clue to its function was the nonrandom distribution, within the human genome, of sequences capable of adopting either right- or left-handed DNA conformations, referred to here as 'flipons'. Flipons are enriched in the promoters and 5′ untranslated regions of genes [4] and are associated with active histone marks [5]. By setting chromatin state, they provide instructions for compiling genetic information into RNA, a process dependent upon the consumption of free energy. As such, flipons are a class of dissipative structures, nonequilibrium states initially described by Prigogine [6] that, in the case of nucleic acids, convert the thermodynamic entropy of DNA structures (ΔE) into the information entropy of transcripts (ΔI) (Figure 1). Here, I describe how flipons work, how they are encoded, and how they accelerate evolution through the dissipative structures they form.

Experimental Approaches to the Z-DNA Question

One approach to understanding the biological role of Z-DNA has been to

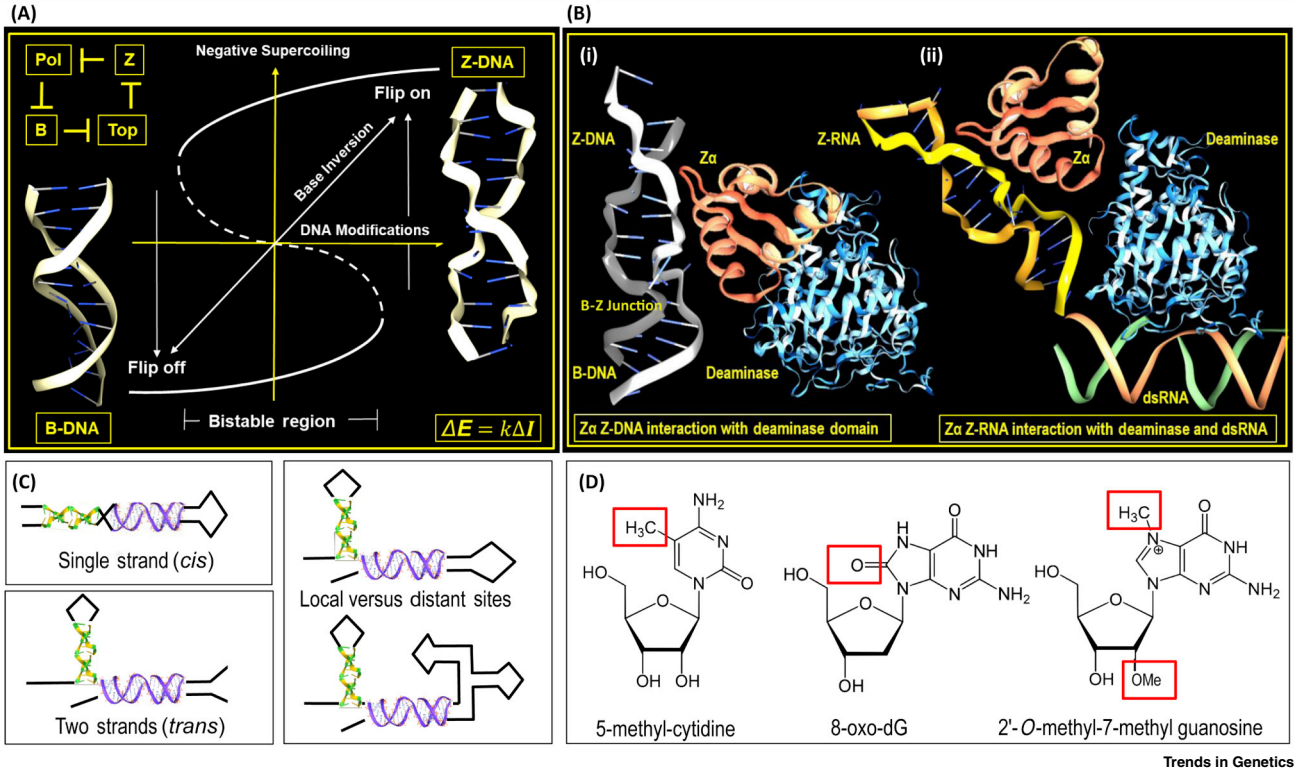

Figure 1. Flipons Integrate Many Environmental Inputs.

(A) Alternative DNA conformations. Negative DNA supercoiling (NSC) and nucleotide modifications (NM) create conditions where the alternative right- and left-handed flipon conformers are energetically accessible (the bistable region). NSC relaxation is slower than generation, releasing free energy for epigenetically modifying the chromatin state. The thermodynamic (*E*) and Shannon information (*I*, base *e*) entropies quantify possible microstates (k = the Boltzmann constant, Δ = change). The cycle (upper left) is intransitive, maximizing entropy by making all outcomes possible. RNA polymerase (Pol) disrupts B-DNA ('B'), Topoisomerase (Top) disrupts Z-DNA ('Z'), 'B' deactivates Top, and 'Z' pauses Pol (blunt-ended arrows indicate inhibition). [Protein Data Bank (PDB): 4OCB ('Z') and 355D ('B')]. (B) Adenosine deaminase RNA-specific (ADAR) model. (i) Zα bound to Z-DNA. The abrupt transition at the 'B–Z' junction everts a base pair. (ii) Zα engages Z-RNA while the deaminase operates on the double-stranded (ds)RNA strand colored green. [PDB: 5ZU1 (Zα) and 5ED2 (ADARB1 deaminase domain)]. (C) Conformation and chromatin states: Z-RNA (green) and dsRNA (purple) formation is affected by chromatin state, involving a single RNA strand with base-pairing either between close-by or distant sequences, or two different RNAs, coding or noncoding. RNA backbones are in black. (D) Nucleotide modifications. Adducts (boxed in red) to bases favor Z-DNA formation, including 5-methylcytosine (5mC), 5-formylcytosine (5fC), 5-carboxylcytosine (5cC), 8-oxo-, 8-nitro-, and 7-methyl-guanine. Deoxyribose adducts such as 2′-*O*-methylated ribosides (2′OmR), adopt the C3′-endo conformation needed to kink the Z-DNA backbone [11]. 5-hydroxymethylcytosine, an intermediate between 5mC and 5fC, does not stabilize Z-formation, making this adduct a key regulator of flipon conformation [12].

isolate proteins that bind specifically to the left-handed Z-DNA conformation, and then study their function. This work led to the discovery of the Zα domain family [2], which binds tightly to both Z-DNA and Z-RNA. X-ray studies revealed that the binding was specific for the Z-conformation without any sequence preference. The co-crystals of Zα with Z-DNA and with Z-RNA allowed identification of key protein residues essential to these interactions [7].

The Zα domain is present in an editing enzyme called adenosine deaminase RNA specific (ADAR). ADAR edits double-stranded RNAs (dsRNA) that usually form when an RNA transcript base-pairs with itself [8]. The enzyme changes adenosine to inosine, which is then read out as guanosine, altering both the information content of the RNA and its downstream processing, allowing the production of many different RNAs from a single gene. Early studies suggested that ADAR was involved in antiviral interferon responses [8]. However, most edited dsRNAs in a cell originate from repetitive, noncoding inverted repeats (IRE) formed by palindromic copies of *Alu* elements that arose early during the evolution of the human genome through a process of copy and paste. IREs fold back on themselves to form dsRNA editing substrates (dRES). They frequently contain sequences capable of forming

Feature	Codon	Flipon
Encoding	Linear sequence	3D Structure
Type	Static, codes words, encodes message	Dynamic, instructive, compiles different messages by altering editing splicing and the engagement of other regulators (e.g., miRNA)
Sequence complexity	High	Low
Advantages	Precise, tunes protein properties	Flexible, adjusts chromatin state
Size	Four bases (adenine, thymine, cytosine, guanosine)	Multiple conformations and chromatin states (B-DNA, Z-DNA, Z-RNA, DNA:RNA hybrids, etc.)
Readout	Depends on base composition (different combinations)	Depends on local chromosomal energetics (transcription, competing conformations, base modifications, oxidative stress and/or mutagens)
Epigenetic	Expand size [e.g., 5-methyl-cytidine (5-mC), 5-hydroxymethylcytosine (5-hC)]	Alter energetics of state transitions (5mC, 5fC, 5cC, 8-oxo, 8-nitro, 7-methyl purines, 2′OmR) (Figure 1D)
Evolution	Base mutation and DNA editing sequence deletion/duplication (replication, repair, recombination)	Base variations altering flip energetics; flipon insertion in new locations (transposon, recombination, repair)
Machinery	Sequence-specific (transcription factors, etc.)	Conformation-specific (ADAR, ZBP1, topoisomerases, etc.)
Thermodynamics	Equilibrium	Non-Equilibrium
Limitations	Multiple sequence motifs in large genomes impact targeting accuracy	Genomic instability associated with replication, transposition, flipon boundaries (B-Z junctions)

Table 1. Ways in Which Codons and Flipons Differ in Their Encoding of Genetic Information

double-stranded Z-RNA (Figure 1B) [2]. Much of this biology remains untapped. It is routine to exclude *Alu* sequences from analyses [e.g., protocols for analysis of d(CpG) islands routinely exclude *Alu* sequences and RNA-seq pipelines discard edited *Alu* as unmatched reads]. Experimental manipulations can inadvertently induce Z-DNA, for example, by extracting proteins, such as histones, that reduce negative supercoiling and produce misleading results. Additional tools for studying Z-DNA *in vivo* based on two photon infrared and Raman spectroscopy of living cells [5], *in situ* Z-specific chemical modification [9], and variants of Zα family members [2] engineered for super-resolution imaging would help overcome current limitations.

Mendelian genetics has proven the most robust approach so far for studies demonstrating an *in vivo* role for Z-DNA [3], especially those involving studies of families with Aicardi–Goutières syndrome (AGS) (OMIM: 615010) and Bilateral Striatal Necrosis/Dystonia (BSD) [3], both type I interferonopathies. Nonsynonymous, non-Z binding Zα variants have exaggerated dsRNA-induced interferon responses due to impaired dsRNA editing, unambiguously confirming a biological role for the left-handed nucleic acid conformation. Recent studies extended these findings by demonstrating that ~40% of tumors depend on their editing by ADAR for survival, with inhibition of interferon production preventing antitumor immune responses [10]. In these scenarios, Z-DNA targets ADAR to actively transcribed genes. The Z-DNA forming sequences also form Z-RNA. Whether it is Z-DNA, Z-RNA, or a Z-DNA:RNA hybrid that targets the enzyme is still an open question because the Zα domain binds all things 'Z'. Binding to the hybrid potentially facilitates the handoff of ADAR from Z-DNA to Z-RNA while pausing RNA polymerases long enough for the dRES to finish folding. It could also allow time to complete RNA modifications that alter the efficiency of editing, including those that favor Z-RNA formation. Z-RNA has the potential to latch ADAR to dRES and increase editing efficiency [2]. The biology of ADAR in the type I interferon response provides a model for how flipons alter the readout of genetic information.

Coding with Flipons

Flipping DNA to the left-handed conformation can alter ADAR localization, histone placement, transcription factor binding, splice site selection, and processing of noncoding RNAs [4]. These outcomes all involve a switch from one chromatin state to another. In the case of ADAR, the use of chromatin-dependent RNA-folding pathways provides a flexible mechanism for changing the readout from genes through the creation of different dRES from a single transcript (Figure 1C). The localization of ADAR to Z-DNA by interferon augments this process by enhanced editing of substrates such as signal recognition protein 9 (SRP9) and cathepsin S (CTSS) [2].

ADAR has roles independent of editing, during which it partners with proteins such as DROSHA and DICER1, targeting them to miRNAs and other noncoding RNAs, often in a tissue-specific manner [8]. Localization of different classes of catalytic machinery to active genes is also possible through their binding to flipons, usually by capturing them in the 'B'-conformation. Many of the resulting epigenetic nucleotide modifications (NMs) are known to impact flip energetics (Figure 1D and Table 1). By pushing flipons either left or right, NMs are able to lock in a particular chromatin state. Flip energetics also change when oxidative stresses and mutagens produce DNA and RNA adducts, especially on the C5 position of cytidine and the C8 position of guanosine. Here, flipons act as damage sensors, enabling a direct and rapid transcriptional response to cellular stress through changes in chromatin state. In all these cases, flipon conformation instructs on how to compile response-specific transcripts.

Generation of Diversity

The ease with which flipons form, along with their location, is subject to selection just like any other genetic variation. Flipons create phenotypic diversity by increasing transcriptome entropy. The genomes that emerge encode information by sequence and instructions by conformation (Table 1). Regions where flipon and codon sequences overlap have lower entropy (i.e., they have a fixed information content) and likely become hotspots for spawning species-specific phenotypic variability. Germline retrotransposition, recombination, and repair enhance transcriptome diversity by spreading flipon sequences to other parts of the genome. The noncoding IREs targeted by ADAR exemplify how this process works. During insertion into active genes, they bring flipons along for the ride [2]. While active transposons threaten genomic stability, the instruction sets they carry enhance the creation and capture of novel genetic programs.

Concluding Remarks: Entropy, Flipons, and Evolution

Codons enable the mapping of nucleotide sequence to protein sequence. Altering their usage is only one way to diversify phenotypes. Nature has discovered other strategies to create novelty through editing and splicing of RNA, resulting in multiple transcripts from the same reading frame. Flipons provide a novel innovation for changing the transcriptome by dynamically switching chromatin states to change how messages are compiled. The sequences encoding flipons often overlap those of codons and other regulatory elements. Each encodes a different set of information, is subject to natural selection, and causes Mendelian diseases in its own way. Flipons trade free energy for the extra possibilities that novel transcripts provide. The increased entropy enhances the reworking of existing adaptations and speeds the evolution of traits both new and unexpected.

Acknowledgment

Many thanks to the reviewer who highlighted the importance of Ilya Prigogine's work on dissipative structures.

[1]Discovery, InsideOutBio, Inc, 42 8th Street, Unit 3412, Charlestown, MA 02129, USA

*Correspondence:
alan.herbert@insideoutbio.com

https://doi.org/10.1016/j.tig.2019.09.007

References

1. Shing, P. and Carter, M. (2011) DNA structure: alphabet soup for the cellular soul. In *DNA Replication: Current Advances* (Seligmann, H. ed), pp. 3–28, IntechOpen
2. Herbert, A. (2019) Z-DNA and Z-RNA in human disease. *Commun. Biol.* 2, 7
3. Herbert, A. (2019) Mendelian disease caused by variants affecting recognition of Z-DNA and Z-RNA by the Zα domain of the double-stranded RNA editing enzyme ADAR. *Eur. J. Hum. Genet.* Published online July 18, 2019. https://doi.org/10.1038/s41431-019-0458-6.
4. Shin, S.I. *et al.* (2016) Z-DNA-forming sites identified by ChIP-Seq are associated with actively transcribed regions in the human genome. *DNA Res* 23, 477–486
5. Zhang, F. *et al.* (2016) Histone acetylation induced transformation of B-DNA to Z-DNA in cells probed through FT-IR spectroscopy. *Anal. Chem.* 88, 4179–4182
6. Prigogine, I. (1978) Time, structure, and fluctuations. *Science* 201, 777–785
7. Athanasiadis, A. (2012) Zalpha-domains: at the intersection between RNA editing and innate immunity. *Semin. Cell Dev. Biol.* 23, 275–280
8. Nishikura, K. (2016) A-to-I editing of coding and non-coding RNAs by ADARs. *Nat. Rev. Mol. Cell Biol.* 17, 83–96
9. Xu, Y. *et al.* (2007) Photochemical determination of different DNA structures. *Nat. Protoc.* 2, 78–87
10. Herbert, A. (2019) ADAR and immune silencing in cancer. *Trends Cancer* 5, 272–282
11. Balasubramaniyam, T. *et al.* (2018) 2′-O-Methyl-8-methylguanosine as a Z-form RNA stabilizer for structural and functional study of Z-RNA. *Molecules* 23, 2572
12. Wang, S. *et al.* (2014) Systematic investigations of different cytosine modifications on CpG dinucleotide sequences: the effects on the B-Z transition. *J. Am. Chem. Soc.* 136, 56–59

Forum

Compensatory Evolution of Gene Expression

Sarah A. Signor[1,2,*]
and Sergey V. Nuzhdin[3]

Determining the contribution of *cis* and *trans* components to differences in gene expression is a powerful approach for understanding gene regulatory evolution. Specifically, differences in gene expression that are due to linked polymorphisms (*cis*, allele-specific and local to the affected gene), or differences due to diffusible products that do not need to be linked with the affected gene (*trans*, affecting both alleles equally in diploids). Decomposing the evolution of gene expression into its