Review

Triplexes Color the Chromaverse by Modulating Nucleosome Phasing and Anchoring Chromatin Condensates

Alan Herbert

Discovery, InsideOutBio, Charlestown, MA 02129, USA; alan.herbert@insideoutbio.com

Abstract: Genomic sequences that form three-stranded triplexes (TPXs) under physiological conditions (called T-flipons) play an important role in defining DNA nucleosome-free regions (NFRs). Within these NFRs, other flipon types can cycle conformations to actuate gene expression. The transcripts read from the NFR form condensates that engage proteins and small RNAs. The helicases bound then trigger RNA polymerase release by dissociating the 7SK ribonucleoprotein. The TPXs formed usually incorporate RNA as the third strand. TPXs made only from DNA arise mostly during DNA replication. Many small RNA types (sRNAs) and long noncoding (lncRNA) can direct TPX formation. TPXs made with circular RNAs have greater stability and specificity than those formed with linear RNAs. LncRNAs can affect local gene expression through TPX formation and transcriptional interference. The condensates seeded by lncRNAs are updated by feedback loops involving proteins and noncoding RNAs from the genes they regulate. Some lncRNAs also target distant loci in a sequence-specific manner. Overall, lncRNAs can rapidly evolve by adding or subtracting sequence motifs that modify the condensates they nucleate. LncRNAs show less sequence conservation than protein-coding sequences. TPXs formed by lncRNAs and sRNAs help place nucleosomes to restrict endogenous retroelement (ERE) expression. The silencing of EREs starts early in embryogenesis and is essential for bootstrapping development. Once the system is set, EREs play a different role, with a notable enrichment of Short Interspersed Nuclear Repeats (SINEs) in Enhancer–Promoter condensates. The highly programmable TPX-dependent processes create a chromaverse capable of many complexities.

Keywords: Triplex; G-quadruplex; Z-DNA; flipons; condensates; gene expression; nucleosome phasing; endogenous retroelements; embryonic development; helicase; microRNA; lncRNA; anti-sense RNA; argonaute; enhancer; promoter

1. Introduction

Genetic elements, called flipons, adopt alternative structures to produce different biological outcomes. They encode information by their shape rather than by their sequence. By switching states from one conformation to another, flipons change the readout of RNA from the genome. The transition from one flipon structure to another occurs without strand breakage or any alteration to the nucleic acid sequence.

Flipons are named for the alternative conformation they encode, e.g., Z-flipons form Z-DNA and Z-RNA, G-flipons form DNA and RNA G-quadruplexes (GQs), and T-flipons encode triplexes (TPXs). The formation of these alternative conformations inside cells is now well supported for both Z- and G-flipons. Here, the focus is on three-stranded TPX (Figure 1). In nondividing cells, the extra strand is usually composed of RNA, rather than DNA. The third strand can be encoded in regions quite distant from where the TPX forms. This design is quite variable and helps integrate the readout of genetic information in cells

by establishing open regions of chromatin that are depleted of nucleosomes (historically called nucleosome-free regions (NFRs)). The flipons and transcription binding sites within NFRs enable the promoter–enhancer contacts that modulate gene expression. At these locations, an RNA cloud composed of promoter and enhancer transcripts engages RNA and protein effectors that set the chromatin state. The process is dynamic with both DNA and RNA conformations reset as an RNA polymerase is released from the promoter. This cycle allows an update of the promoter based on the current cellular condition.
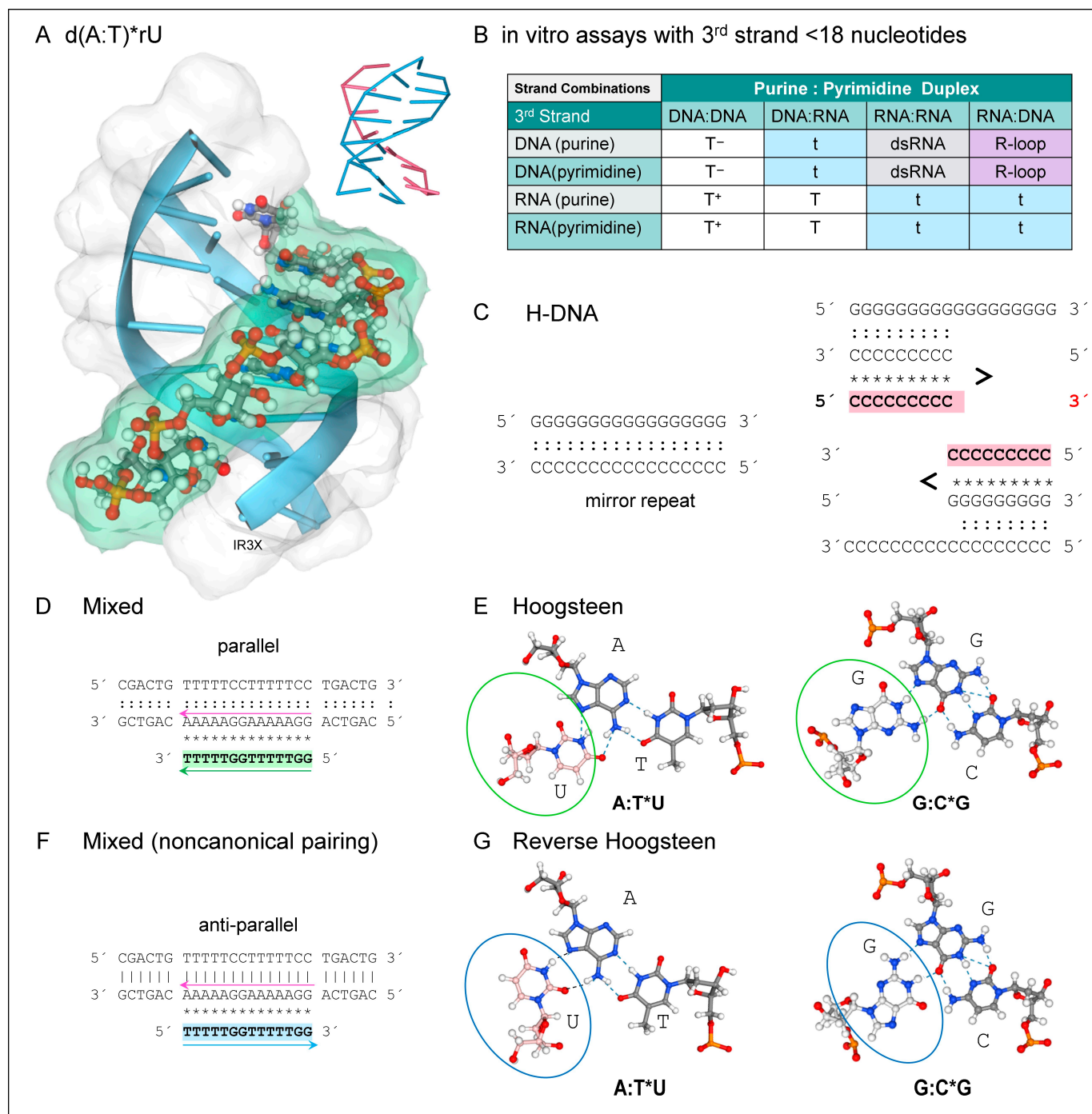

Figure 1. In a TPX, the orientation of the third stand is antiparallel to that of the matching pyrimidine or purine-rich strand within the double helix. To avoid confusion, TPXs are named with reference to the 5′ to 3′ polarity of the purine strand: polypyrimidine TPXs are called parallel, and polypurine

TPXs are named anti-parallel TPXs. In the text and figures, Watson-Crick hydrogen bonds in the duplex are shown as ":" and Hoogsteen bonds formed with the third strand are denoted by "*". DNA helical strands are designated by "d" and RNA strands by "r" (**A**) The canonical $d(A:T)_n$*rU TPX with the third strand shown as a ball and stick model in a green-tinted surface representation. The inset is of a wireframe model for the TPX viewed from behind. (**B**) The in vitro stability of TPXs depends on the ordering of the RNA and DNA strands. A helix with an RNA purine strand is not observed to form a TPX. Rather, the R:D helix is more energetically favored. (**C**) The $d(G)_n$ mirror repeat sequence can form a TPX with a purine strand. Under acid conditions, when cytosine is protonated, a TPX with a pyrimidine strand forms instead. (**D**) TPXs can dock a third strand composed of G and T bases in a parallel. (**E**) The parallel "mixed" TPXs form with Hoogsteen hydrogen bonds (**F**) The alternative anti-parallel orientation formed with the same sequences. (**G**) Here, reverse-Hoogsteen hydrogen bonds form between the duplex and the third strand. In panels (**C**,**D**,**F**), the third strand of the TPX has a colored background.

2. Genomes Encode Information by Both Shape and Sequence

The human genome encodes many possible outcomes. Initially, chromosomes were described as composed of centromeres, telomeres, and chromomeres [1]. The work on telomeres and centromeres revealed that they were composed of repetitive sequences, which play a role in chromosomal maintenance during cell division and differentiation. As technology advanced, each chromomere region was assigned a color as its functionality was defined. The labeling was likely a play on the name chromosome that was coined in 1888 by Wilhelm von Waldeyer-Hartzchroma. He used the Greek word "χρῶμα" (khrōma), meaning "color", as the root for the name. In these early maps of the chromaverse, heterochromatin was labeled black, and regions with actively transcribed genes were called red. Each domain within the nucleus was further shown to have its own location and unique sets of epigenetic marks [2]. High-resolution techniques, such as DNase I footprinting and ATAC-seq, then enabled the mapping of NFRs in each domain. These and other approaches identified the proteins bound, the sequence motifs recognized, and the functions performed. The work revealed the complexity of the chromaverse, demonstrating that the collage of colors within the composite maps was not fully defined by the linear flow of information from DNA to RNA to protein [3].

The ENCODE (Encyclopedia of DNA Elements of Human and Mouse Genomes) surprisingly revealed that many genomic regions within chromomeres were transcribed to produce non-protein-coding RNAs (historically referred to as ncRNAs). However, agreeing on the extent of this promiscuous expression took some time [4–6]. The current consensus is that over 75% of the human genome is read out as RNA, with 37,595 non-protein coding genes (ncRNAs) annotated compared to 20,225 transcripts that incorporate codons [7–9].

Recent high-throughput approaches have focused on ncRNAs and generated a wealth of new information. The ncRNAs vary in size, origin, copy numbers, protein binding partners, access to the nucleus, chromatin interactions, and even in definition [10–12]. At one end of this scale lie long ncRNA (lncRNAs) with low expression. Initially, lncRNAs were defined as longer than 200 nucleotides (nts) in length. Now it is suggested that 500 nts is a more appropriate size. The newer definition of small RNAs (sRNAs) includes a repetitive sequence like the Alu element (AE) repeat family of Small Interspersed Nuclear Elements (SINEs) and other transcripts such as microRNAs (miRNAs), piwi RNAs (piRNAs), transfer RNAs (tRNAs), ribosomal 5S rRNA, *7SK*, *7SL*, and vault and Y RNAs that with sizes in the 50–500 nt range [11].

Both tRNAs and miRNAs have been well characterized. Both are noncoding RNAs that impact mRNA translation into protein. The interactions of miRNA with transcripts are mediated by the argonaute (AGO) family of proteins, which also includes the PIWI clade that binds piRNAs [13,14]. In humans, there are four AGO proteins. AGO2 has nuclease

activity and slices mRNA that are fully matched to the miRNA. AGO3 has slicer activity that is most evident with miRNAs of ~14 nts [15]. AGO1 and AGO4 are not known to have slicer activity. Interestingly, AGO proteins can engage tRNA fragments and other small RNAs to perform similar functions to those involving miRNAs [12,16,17]. Understanding of these functions has grown as well. Initially, the pathways were thought of as only cytoplasmic. They targeted RNA for elimination or inhibited its translation. Then, the involvement of miRNAs and piRNAs in gene suppression within the nucleus was revealed in a number of model organisms [18]. More recently, other sRNA pathways have been identified that lead to gene activation [19,20].

With the exception of miRNAs and tRNAs, other roles for the noncoding part of the genome were not explored or were considered unlikely, as argued by Brenner [21]. Most notably, the low complexity repeat sequences that were not within codons were considered uninformative due to their high frequency in the genome. Furthermore, early models proposing a role for repeats in tissue specification were not experimentally verified (Britten, 1969 #303 [22]). Additionally, the repetitive sequences were considered detrimental because of the enhanced error rate associated with their replication and their highly recombinogenic nature [23]. While these events increased genetic diversity and sometimes produced advantageous outcomes retained by genomes through natural selection, they also increased disease risk. At best, the persistence of these repetitive sequences in the genome was considered as one of nature's quirky imperfections [21].

Quite separately, the ability of certain DNA sequences to form conformations other than the right-handed Watson–Crick B-DNA helix became apparent. Each structure was favored by a particular nucleotide motif, with polypurines favoring a three-stranded right-handed TPX [24–26] and guanine repeats favoring a four-stranded quadruplex composed of G-tetrads (GQs) held together by a combination of Watson–Crick and Hoogsteen hydrogen bonding schemes [27,28]. An alternating purine/pyrimidine motif favored left-handed Z-DNA formation [29–31], while a pair of cytosines could intercalate with another pair to form a four-stranded i-motif structure [32,33]. The sequences that form both B-DNA and these alternative conformations in vivo are called flipons [34]. They change biological outcomes by binding chromatin complexes that are specific for one conformation or the other [35]. Flipons are distributed nonrandomly throughout the genome. Recent evidence supports their important informational role in programming cellular responses.

G-flipons and Z-flipons have key roles in transcription and immunity. Each flipon type affects different outcomes. G-flipons regulate the onset of transcription by cycling between states. The complexes that they form remain locked and loaded until a particular event triggers the release of RNA polymerases from the pause position just downstream of the transcription start site. G-flipons also modulate many other processes, including replication and DNA repair [36–38]. Z-flipons play an important role in controlling innate immune responses to viruses and endogenous retroelements (EREs) that comprise autonomous Long Interspersed Nuclear Elements (LINEs), Long Terminal Repeats (LTRs), and non-autonomous SINE retrotransposons [39–42]. They also play a pivotal role in resetting promoters so that they are ready for the next round of transcription [43–50]. In addition to encoding Z- and G-flipons, EREs also form TPXs [51,52]. Most notably, SINEs represent the ultimate abstraction of genetic encoding by shape rather than sequence, packing many different flipon combinations into their 280 nt span, while not coding for any protein at all. Despite their small size, the genetic regulatory elements that they encode greatly impact a cell's molecular biology.

3. T-Flipons and TPX Forming Sequences

The biology of T-flipons is quite different from that of G- and Z-flipons. Only in the case of mirror repeats, in which a sequence is repeated with its order reversed (e.g., "ACG" becomes "GCA"), can TPXs form locally. Otherwise, the third strand is sourced from a distant site, with both sRNAs and lncRNAs involved in the fold. These interactions can approximate DNA segments that are separated by long distances or transcribed from different chromosomes. The bridges between these sites are strengthened by the many RNA copies transcribed from each locus, creating an RNA that concentrates protein into condensates, some of which undergo phase transitions to fulfill and facilitate particular functions, from modulating gene expression to coordinating cellular responses during differentiation. To understand these processes, we will first discuss the sequences prone to form TPXs.

4. DNA and Hybrid TPX

Right-handed DNA duplexes can form TPXs with either a DNA or RNA third strand (Figure 1A). The TPX stability arises from the additional stacking interactions and the increased number of hydrogen bonds formed between bases compared to B-DNA. TPXs can also form with RNA inserted into either the major or minor groove. Major groove TPXs are most common (Figure 1A). Experimental studies have revealed several factors that influence the stability of major groove TPXs. They exist in equilibrium with other folds, including DNA:DNA (D:D) and RNA:RNA (R:R) duplexes and RNA:DNA hybrids (D:R). They compete with R-loops, which form when a DNA strand is displaced from the double helix by the binding of an RNA complementary to the other strand [53–55]. The different outcomes are summarized in Figure 1B. Double helices (either R:D or R:R) with an RNA purine strand are more stable than helices with a DNA purine strand and will preferentially remain as dsRNA or form an R-loop instead of a TPX. In TPXs, RNA pyrimidine strands bind to D:D, R:R, and D:R duplexes, whereas a DNA pyrimidine strand will bind only to duplexes that have a DNA, but not an RNA, purine strand. Hence, TPX formed with RNA and a D:D are the most stable of all the possible combinations.

The formation of triplets in vitro requires high levels of divalent cations to shield the charge repulsion arising from the approximation of the phosphate backbones forming each strand. The repulsion can also be countered by lowering the pH to protonate the cytosines in a polypyrimidine third strand, promoting the formation of an additional hydrogen bond in each triplet. In contrast, the hydrogen bonding scheme in the major groove of long $d(A)_n$ repeat disfavors TPX formation. The bifurcated hydrogen bonds produce a high propeller twist of the bases, stabilizing a very rigid duplex structure. The zipper that forms can be disrupted by the insertion of other bases, enabling formation of a mixed sequence TPX with high thermal stability [56].

DNA-only TPXs arise from sequences with a mirror repeat (MR), in which the nucleotide order is reversed relative to the preceding repeat. The arrangement allows one of the repeats to fold back and hydrogen bond with a neighboring duplex region when the MR is 12 or more nts in length (Figure 1C). The formation of a TPX yields either an unpaired purine or pyrimidine DNA strand. A TPX incorporating the pyrimidine strand is favored under acidic conditions when the cytosine residues are protonated. TPXs formed in this way are referred to as H-DNA. They predispose to DNA damage and disease by stalling polymerases during cell division as the DNA strands are separated from each to undergo replication [26]. Over 85% of MRs in the genome are also perfect direct tandem homopurine repeats [57].

The "GT" TPX is another TPX variation with only G, T, or U in the third strand (Figure 1D–F). This motif is called "mixed" as the third strand can bind in either orientation

to the duplex (Figure 1D). This dual outcome is unique to this TPX type. The orientation of the third strand depends on the duplex sequence. As with a standard TPX, the polypyrimidine third strand can have the same 5′ to 3′ polarity as the purine strand of the duplex. This "GT" parallel third strand forms with Hoogsteen bonds (Figure 1E). In contrast, a polypyrimidine third strand may form a noncanonical TPX by binding in an anti-parallel orientation relative to the duplex purine strand. Like a standard anti-parallel purine helix, the third strand is stabilized by reverse Hoogsteen bonds (Figure 1G). Regardless of the orientation, the third strand can contain a mix of pyrimidine and purine residues that form G:C*T and A:T*T but not A:T*A triads (where "*" indicates Hoogsteen base pairing of the third strand and ":" a Watson-Crick hydrogen bond) [58]. By replacing "C" with "T", the design overcomes the pH-dependence of "C" protonation necessary to stabilize the TPX. Substitution 2—aminopyridine or 5-methylcytosine—also increases TPX stability [59].

Third strands composed only of "TCG" bases have also been explored as a way to target duplexes with a G repeat. In these TPXs, the strand orientation is parallel, with C:G*C⁺ triplets replaced by G:C*G. In contrast, a third strand composed of only "GAU" bases docks in the anti-parallel direction with A:T*U, A:T*A, and G:C*G triplets [60]. Other studies examined third strands that pair through noncanonical triplets. Examples of these triplets are shown in Figure 2A. Up to two noncanonical triplet substitutions, such as G:C*A, G:C*U, and A:U*C, are tolerated in d(T:A)*rU TPXs longer than 19 nucleotides. The results reveal that some sequence variations in the RNA third strand are tolerated and confirm results from earlier reports [61,62]. Examples of such noncanonical base pairing in TPXs formed by ribosomal DNAs with their functional outcome are shown in Figure 2B and 2C and are discussed below.
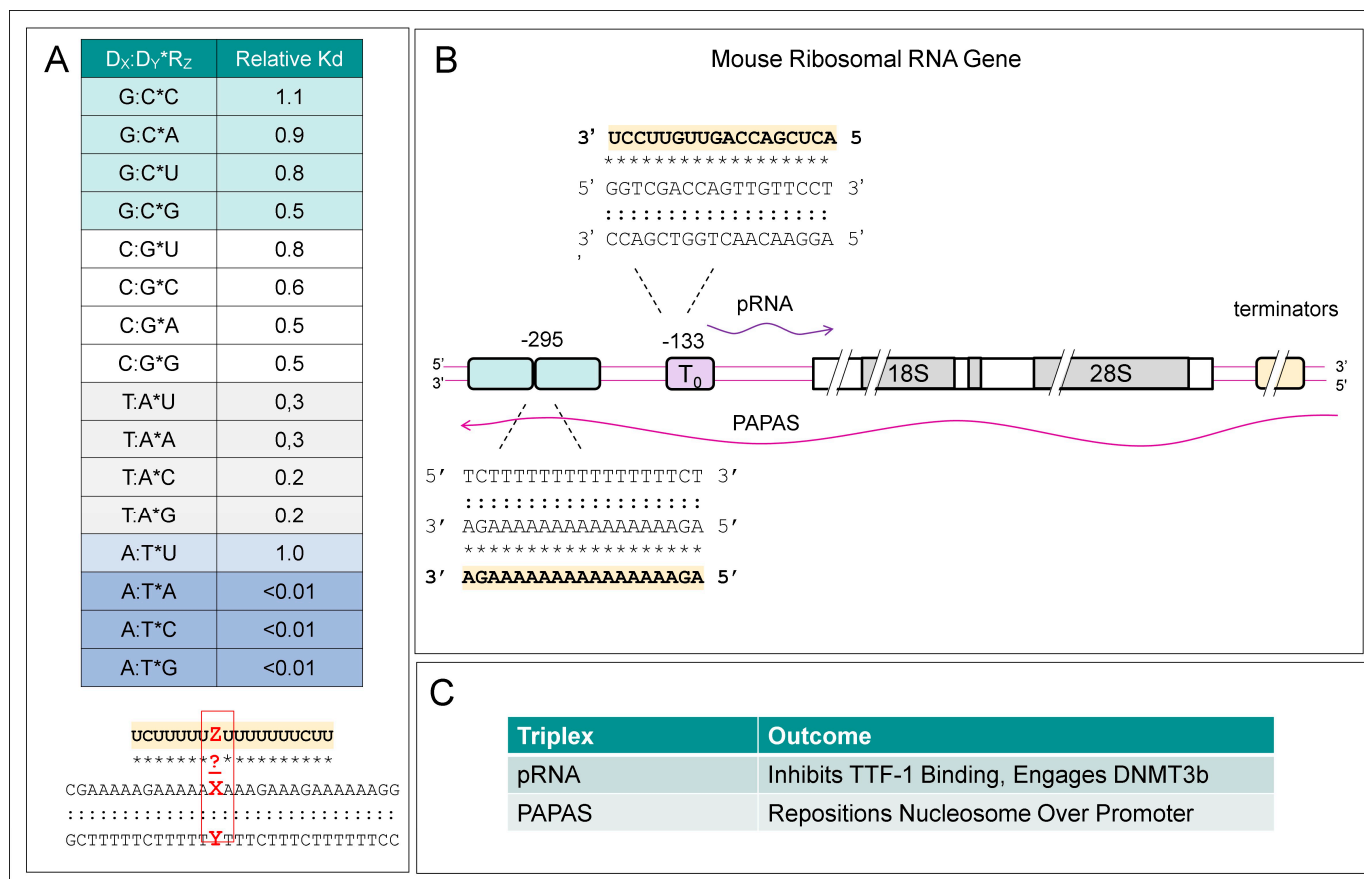

Figure 2. Noncanonical triplets and ribosomal gene regulation: (**A**) Substitution of a central triplet in a d(T:A)*U TPX (boxed in red) with a noncanonical triplet allows an estimate of their relative stability

(adapted from [62]). Watson-Crick basepairs are indicated by ":" and Hoogsteen basepairs by "*". (**B**) In mice, but not yet shown in humans, the sense pRNA lncRNA and the antisense PAPAS lncRNA form TPXs that suppress rRNA transcription, as explained in the text. (**C**) While pRNA induces DNA methylation by DNMT3b of the enhancer T_0, PAPAS repositions a nucleosome to occlude the promoter. The third strand of each TPX has a colored background.

There are many other base-pairing geometries compatible with triplet formation. A recent survey identified 108 predicted base triple families that belonged to 18 superfamilies. Of these, evidence for 68 was found in RNA structural databases [63]. The results suggest greater flexibility in triplet motifs than previously recognized. However, the number of these triples that can also assemble into stable triple-stranded helices is currently unknown.

5. Flipons That Fold in Many Different Ways

Like other flipons, TPX-forming sequences can fold into other alternative conformations. Some MR-like d(GCG GCG) (usually written as d(CGG CGG) where the mirror repeat is not apparent) can form other non-B-DNA structures, such as Z-DNA [64,65]. The c-MYC gene sequence provides two examples. One MR (the two repeats are underlined) can form GQ and H-DNA (5'-TGGGG<u>AGGGTGGGG</u>AGGGTGGGGAAGG-3') while another MR (the two repeats are underlined) forms GQ or H-DNA (5'-<u>GGGAGGGG</u>CGCTTAT<u>GGGGAGGG</u>-3') [66–68].

In other cases, each strand from a duplex can individually fold to produce different non-canonical structures. For example, the d(GGGGCC) repeat from the C9ORF72 gene that predisposes to amyotrophic lateral sclerosis can fold as a GQ (Figure 3A). The complementary strand instead can form a TPX (Figure 3B, the third strand is drawn in fuchsia). This sequence is a MR when written as CCG GCC. Alternatively, four C-rich strands can form i-motifs (highlighted in green). Each pair of strands can also pair to form duplexes as they enter and leave the i-motif region. These duplexes can also adopt an alternative structure by forming Z-DNA at both ends (Figure 3C). Besides adding a third strand to form a TPX, a pair of single-stranded d(AT) rich sequences can base pair to create helices with parallel strands instead of the anti-parallel alignment found in canonical Watson–Crick B-DNA (Figure 3D). It is currently unknown if and when such complex structures fold in vivo. However, these examples reveal the propensity of some sequences to fold in multiple ways, including those that can form non-canonical triplets. Both the folds and the outcomes then depend on the context.
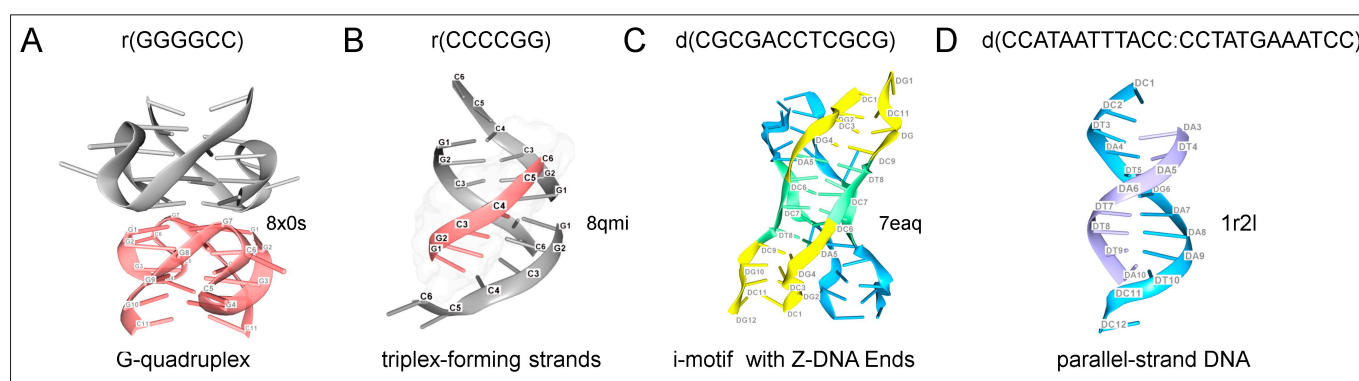

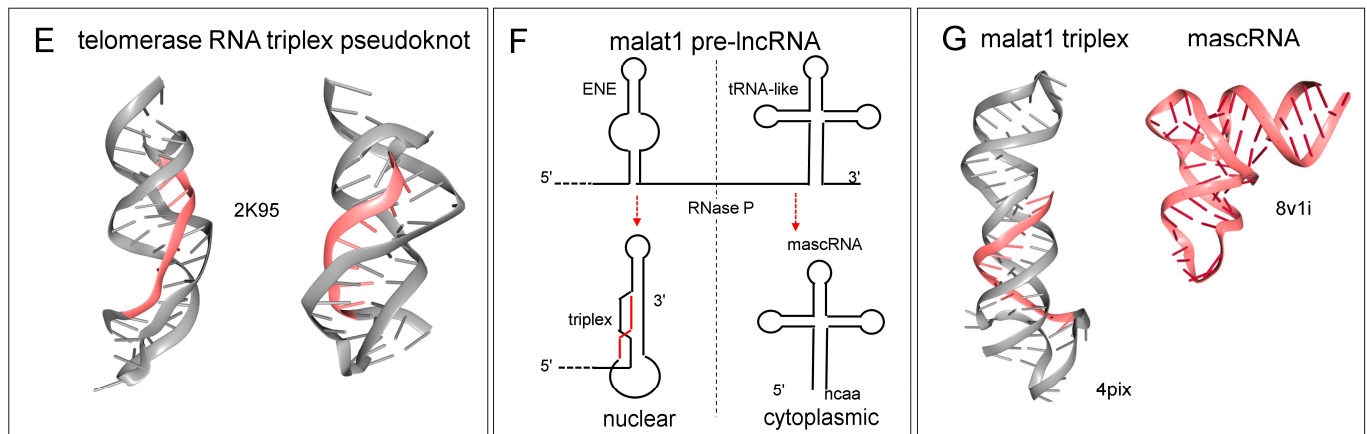| A r(GGGGCC)	B r(CCCCGG)	C d(CGCGACCTCGCG)	D d(CCATAATTTACC:CCTATGAAATCC)
8x0s	8qmi	7eaq	1r2l
G-quadruplex	triplex-forming strands	i-motif with Z-DNA Ends	parallel-strand DNA

Figure 3. *Cont.*

Figure 3. The diversity of folds formed by simple repeats: (**A**) One strand of the C9ORF72 repeat can fold with another segment to form a pair of GQs that stack on each other. (**B**) In contrast, the complementary strand can fold to form a TPX as the sequence CCGGCC is a MR. The third strand is colored fuchsia. (**C**) Four C-rich strands can form an i-motif (colored green) that can adopt complex folds, including one where each pair of strands (one pair colored blue and the other yellow) forms Z-DNA at both ends. (**D**) Two A-rich sequences can form a parallel-strand right-handed helix rather than folding with a third strand to form a TPX. (**E**) The telomerase pseudoknot also contains a TPX. Views from the front and back are shown. The third strand of the telomerase TPX is colored fuchsia. (**F**) The processing of the MALAT1 pre-lncRNA that produces a TPX at the end of MALAT1 and the tRNA-like mascRNA. The third strand of the MALAT TPX is colored fuchsia, (**G**) The structures of the MALAT1 TPX, tRNA-like mascRNA. The third strand of the MALAT TPX is colored fuchsia. Protein database (PDB) accession codes for the structures are given in each panel.

6. RNA Only TPXs and Other RNA Folds

RNA-only TPXs are also known and have recently been reviewed [69]. In cells, their role in RNA processing pathways varies. A TPX forms the catalytic core of the human U2–U6 spliceosome. A TPX fold also stabilizes the telomerase enzyme but is not part of the catalytic core [70,71] (Figure 3E). The TPX in MALAT1 RNA protects the 3′ end from nucleases [69]. Formation of the MALAT1 TPX depends on cleavage by RNase Z, a reaction that yields the mascRNA (Figure 3F with structures shown in 3G) [72]. A similar process is involved in the processing of the long NEAT1 pre-lncRNA. The tRNA-like molecules produced regulate inflammatory responses [73]. The MALAT1 TPX also binds METTL16 (methyltransferase 16, RNA N6-adenosine) through a protein-mediated VCR domain clamp. The bound METL16 then methylates adenosine 43 in the U6 snRNP, which is involved in splicing. How the different outcomes are related is not currently known [74–76].

7. Resolution of TPX

The use of RNA structural folds to scaffold the assembly of protein machines is quite widespread, as is well known from studies of ribosome assembly [77]. As with flipons, the RNA motifs rather than their sequence are critical to the assemblies created. Structural motifs are also important to their disassembly. The DExD/H-box RNA (DDX) helicases illustrate this principle [78]. The proteins initially undergo a low-affinity interaction with dsRNA. The deep grooves preclude the readout of their sequence. Instead, the helicases exploit structural features found in partially unwound dsRNA for targeting their interaction, with sequence-specific contacts are rarely made [79]. They instead make specific contacts with the phosphate backbone and the 2′OH group of the ribose sugar. The docking induces a conformational change that unwinds the RNA, resulting in a high-affinity interaction that reorients and separates the chains [79,80].

Paradoxically, many helicases are associated with particular sequence motifs [81]. This outcome may occur for several reasons. The helicase interactions may reflect the recognition of structural elements within a partially single-stranded RNA, such as K-turns [82,83], reverse K-turns [84], tetraloops and T-loops [85], and Z-turns [86] that nucleate RNA folding by bending and kinking the RNA backbone to create a distinct three-dimensional profile. The combination of distinct structural motifs bound by different helicase domains then codes for the interaction. Helicases may also bind to the partially single-stranded regions formed when flipons adopt an alternative conformation. These ssRNA sites are present at the junctions between a right-handed duplex and an alternative structure or in bulged regions or hairpins created to accommodate the flipon fold. The recognition by helicases of structural variations associated with G-flipons have been recently reviewed [36,87]. The sequence-specificity may arise for another reason. Helicases have different conformational states, with the transition from a closed to an open binding state triggered by another protein, often through a G-patch [88]. In this situation, the binding of an associated protein to a particular sequence motif can activate the helicase. These interactions with different sequence-specific proteins allow helicases to both enhance and suppress gene expression by dynamically remodeling the scaffolds bound.

8. Genome-Wide Prediction of TPXs by Computational Approaches

Following the identification and characterization of TPXs in vitro, interest turned to their in vivo relevance. Numerous studies have been facilitated by computational approaches designed to predict TPX-forming sequences in genomes. These tools have recently been reviewed [89]. Algorithms like those used in Triplexator, Triplex Domain Finder, and an updated PATO version use canonical Hoogsteen pairing rules while routinely limiting predictions to sequences that are 16 bases or longer [90–92]. The programs detect TPX sequences enriched in DNA simple repeats. They find fewer TPXs in either low-complexity or imperfect repeats. Later methods, such as 3 plex, predict triplet formation for third strands as short as 8 nts and incorporates noncanonical triplets. The approach also considers the alternative non-TPX RNA folds that a sequence can adopt [93]. The LongTarget method instead prioritizes TPXs of 50 bases in length while tolerating a 30% error rate, including those mispairings from non-canonical triplets. The approach is optimized for detecting TPXs formed by long noncoding RNAs (lncRNAs) and uses various experimentally derived rules to score triplets [94]. The authors note that many predicted long TPXs are composites of shorter TPXs separated by non-triplet forming residues. In this scenario, a long TPX may have shorter segments with an sRNA as the third strand. This allows triplets to arise in a combinatorial manner, analogous to how protein transcription factors act together to regulate gene expression [93]. TripletAligner adopted a different strategy to train an Expectation-Maximization model using experimental data from triplexRNA-seq and triplexDNA-seq data from HeLa cells, with validation based on RADICL-seq and RNA ends on DNA capture mappings. The experimental substantiation of a limited number of predictions was also performed. The approach confirmed the accuracy of short 7–8 nt motifs enriched in the predictions.

As noted in a recent review, the computational tools vary considerably in their output, making the comparison of the results challenging [89]. Even with Triplexator, the search results for a TPX-forming Oligo (TFO) and a Triplex Target Site (TTS) in a promoter vary depending on the parameter used. For example, predictions are significantly impacted by simultaneously decreasing the TPX length from 19 to 15 nts and increasing the allowed mispairing rate from 10% to 20%, while minimizing the effects of noncanonical triplets by raising the required G-content from 25% to 65%. When a set of 27,876 sequences 1000 bp upstream and 200 bp downstream of a transcription start site is analyzed with the adjusted

parameters, the number of predicted TPXs increases from 2596 to 6779 [90]. Interestingly, chromatin-associated RNAs were enriched for TTS, especially in GT-rich DNA sequences. As with any computational approach to a new problem, the initial TPX-focused methods captured classical motifs, while later methods detected other variants as the understanding of TPX-forming rules evolves.

9. Physical Mapping of TPXs In Vivo

Genome wide experimental studies to identify TPXs have also been developed. The chromatin-immunoprecipitation sequencing (ChIP-seq) approaches are currently limited due to a lack of TPX-specific antibodies. There is no antibody equivalent to the Z22 monoclonal antibody that allows for the detection of Z-DNA and Z-RNA, regardless of sequence [68], nor to the BG4 engineered antibody that binds to parallel-stranded G-quadruplex structures [95,96]. The three TPX reagents currently reported consist of an antiserum raised against d(A:T)*T [97,98] and the monoclonal antibodies JEL318 [99] and JEL466 [100] raised against d(T^{m5}C:GA)*d(^{m5}C$^+$T)]. All three reagents have different, almost non-overlapping sequence-specificities [100], which is rather surprising given that the bases in TPXs are buried within the TPX. If the differences are due to heterogeneity in the backbones of each TPX type, structural studies may add clarity. Another possibility is that the antibodies recognize junctions or different features in the polymers used for validation.

Other experimental approaches for TPX detections are based on the copurification of chromatin-associated RNAs and DNAs. Protocols involve the proximity ligation of DNA to RNA, followed by selective sequencing of the covalently linked DNAs and RNAs. The method maps interactions between sequences that arise from non-contiguous chromosomal loci. DNase I, RNase A, and RNase H incubation steps are used in the protocols to increase specificity as TPXs are more resistant to these enzymes than single-stranded or double-stranded nucleic acids. Proteases allow the isolation of complexes stabilized primarily by nucleic acid interactions. The methods have been adapted to study particular protein complexes that can be enriched with appropriate antibodies. Genome-wide maps have been prepared using other techniques [101,102]. An alternative approach is based on the sequencing of the RNA that is co-immunoprecipitated with DNA (DNA-ip) combined with the sequencing of DNA pulled down by chromatin-bound RNA ligated to biotin-tagged linkers (RNA-capture). Computational methods can then help identify likely TTS [101]. The strengths and weaknesses of each approach have recently been reviewed [25]. The proximity ligation methods often produce short 22–28 nt sequences that do not contain the TTS and may be difficult to map if they involve TPX formation by repetitive elements. The sequences identified by DNA-ip and RNA-capture approaches have a 60% overlap, with the RNA-based isolation approach showing less specificity, with 20% of peaks found in annotated R-loop regions [101].

A recently published chemoproteomic approach is based on the intercalation of benzo-quinoquinoxaline (BQQ) into TPXs [103]. The binding constants determined by Electrospray Ionization Mass Spectrometry for this compound in 1:1 and 1:2 complexes with TPX DNA were measured at $K_1 = 1.2 \times 10^5$ M^{-1}, $K_2 = 3.9 \times 10^4$ M^{-1}, respectively. In contrast, the values for the 1:1 and 2:1 complex for the DNA duplex were $K_1 = 3.2 \times 10^4$ M^{-1} and $K_2 = 3.7 \times 10^4$ M^{-1}, respectively, offering at best a 25-fold increase in specificity for TPX detection [104]. In the cell-based approach, a derivative of BQQ was prepared by adding a photoreactive cross-linking reagent for proximity-dependent labeling and a linker for the biotin-mediated pull-down of associated proteins. Of the 78 proteins found in two cell lines, 13 were helicases. There were 11 proteins previously published as TPX binding proteins and 16 others that showed a preference (based on published datasets) for polyT sequences,

with others showing a preference for mixed polypyrimidine motifs. Band-shift assays showed the specificity of the helicase DDX3X for triplets over duplexes and GQ [103].

These approaches can also be combined with other studies that utilize proximity ligation to identify DNA and RNA sequences or RNA sequences within 25 Å of each other. The motivation for these studies was to characterize chromatin-associated lncRNAs, their interactions, their relationship to RNAs transcribed at enhancers (eRNAs), and their sequence characteristics in an unbiased manner [102,105–107]. The analysis allowed the mapping of contacts within and between chromosomes, providing a high-resolution map of chromatin domains and nuclear structure. The overlay of ENCODE and other data for epigenetic marks, transcription binding sites, open chromatin regions, and flipon maps has provided new insights into non-canonical ways of encoding genetic information.

10. TPXs Forming Sequences and NFRs

Using these approaches, it is possible to ask how TPX formation influences the binding of nucleosomes to DNA. The analysis reveals that TTS are enriched in NFRs as they are not easily incorporated into nucleosomes [108,109]. Instead, TPXs enable the exact positioning of nucleosomes on DNA to create an NFR. This outcome is especially true for TPXs formed by adenosine-rich sequences (ARS). Even without forming TPX, ARS resist nucleosome formation. The duplexes are inherently rigid and bent due to the twisting of adenosine bases arising from the bifurcated hydrogen bonds formed in the major groove. These tensegrity struts formed resist deformation by bracing the successive layers of base pairs. Genome-wide, these features distorts the wrapping of nucleosomes around ARS, sometimes occluding binding sites for transcriptional factors [110]. The nucleosomes bound to ARS dissociate from DNA easily, unlike the stable nucleosomes present at −1 and +1 regions of an NFR [111]. They are enriched in NFRs and in promoters and play a role in phasing the downstream nucleosomes. However, DNA sequence alone cannot explain these outcomes in cells, as the recipient organism determines the phasing of a foreign DNA [112]. Further, recent time-course experiments suggest that the apparent enrichment of d(A:T) in fragile sites arises from the difference in sequence sensitivity to the nucleases used, rather than for other reasons [113].

ARS may also impact the positioning of nucleosomes involves by chromatin remodelers [114]. These sequences can render the movement of chromatin remodelers along DNA more energetically favorable in one direction, rather than the other [115]. In the yeast *Saccharomyces cerevisiae*, the ATP-dependent Remodeling the Structure of Chromatin (RSC) complex contributes to chromatin opening by preferentially displacing nucleosomes towards the 5′ end of a poly(dA:dT) tract and enhancing their disassembly [116,117]. Another complex, Imitation Switch 1a (ISW1a), which shrinks the NFR, preferentially moved nucleosomes in the opposite direction. A more elaborate system of periodic $d(A_n/T_n)$ clusters (PATCs) is believed to function similarly in *Caenorhabditis elegans*. PATC boundaries prevent the silencing of essential early development genes by blocking the transitive spread of transposon suppression via RNA interference (RNAi).

In mammalian genomes, the many repetitive sequences present also play a role in nucleosome phasing and epigenetic modification. SINEs localize the nucleosome over their A-rich regions and set the phase of adjacent nucleosome. Experimentally, these ARS can form TRX [101]. LINE-1s are also capable of TPX formation to create NFRs [52].

11. miRNAs and TPXs

Within cells, various factors can stabilize or promote TPX formation. Polyamines like spermidine and spermine counter the backbone charge and promote TPX formation at a neutral pH [118]. Other cationic proteins with suitably spaced charges can exhibit a similar

behavior. Notably, TPX formation can be stabilized by nucleosomes through interactions with the N-terminal basic tail of Histone H3 (H3) based on in vitro assays, but not Histone H4 [119]. The in vivo phasing of nucleosomes was mapped by micrococcal nuclease sequencing in human HeLa cervical carcinoma cells and overlapped with Triplexator predictions [120]. TTS peaked at 150 nt on either side of the midpoint of nucleosomal contact with DNA. The TTS also overlapped with predicted microRNA seed sequence matches [121] and with H3K4me3 (H3 with lysine 4 trimethylated) marks associated with active promoters and H2K27ac (H3 with lysine 27 acetylated) flags for active enhancers. Modeling shows that the increased negative charge density of a polypyrimidine D:D*R TPX markedly increased the ionic bonds formed with histone H3 arginines and lysines, stabilizing the TPX. These included Arg2 and Arg8 residues that undergo unique epigenetic modifications [122,123]. The data reveal that TPX formation can modulate epigenetic marks, which in turn modulate TPX formation.

Besides miRNA, other sRNAs can have high tissue-specific expression, potentially impacting gene expression. Some are present in the nucleus, while others, such as tRNAs, shuttle between the cytoplasm and nucleus [124]. Due to their size, sRNAs only have a single TPX-forming site. The sRNA may be produced by the cleavage of longer RNAs, facilitated by nucleases such as Drosha, Dicer, and Argonaute and Piwi proteins, RNase Z during tRNA processing, integrator complexes that cause premature RNA transcriptional termination, splicing complexes that remove introns, and the polyA machinery that cleaves and tails transcripts [17]. Consequently, many sRNAs arise from the post-transcriptional processing of various transcript classes, including exons, introns, rRNAs, tRNAs, snRNAs, and snoRNAs [12]. Of these sRNAs, the evidence that miRNAs form TPXs is strongest for those that lack a mixed purine and pyrimidine in their sequence [121,125]. Bioinformatic analysis suggests that many miRNAs originated from and coevolved with SINEs, LINEs and other ERE [126,127]. Both piRNA and tRNA fragments also play an important role in regulating these elements by targeting argonaute family member effectors to these sequences [128–130]. As noted above, SINEs [101,120], and LINE-1 [52]. are also capable of TPX formation, suggesting that the interactions between EREs are coaptive.

12. lncRNAs and TPXs

While the sRNAs involved in TPX formation are produced in trans from distant sites, lncRNAs produced in cis (i.e., within the same chromosomal domain) also induce TPX formation to regulate gene expression. The lncRNAs can also act by controlling promoter access through the placement of nucleosomes. One example is the regulation by ribosomal RNA (rRNA) expression in mice by the sense pRNA (promoter RNA) and the antisense PAPAS (promoter and pre-rRNA antisense), neither of which is well conserved phylogenetically [131–134] (Figure 2B,C). The lncRNAs suppress rRNA transcription in two different ways, both involving TPX formation. The pRNA TPX inhibits binding by the enhancer protein TTF-1 (transcription termination factor 1) but instead preferentially binds DNMT3b (DNA methyltransferase 3 beta) that methylates a cytosine in the enhancer T_0 at position 133 to decrease gene expression [133] (Figure 2B). The PAPAS TPX repositions nucleosomes by 24 nts to occlude the rRNA promoter [134]. It is unknown whether the lncRNAs remain bound to the chromosome after transcription, detach and form a bridge between the two chromosomal loci, or are processed to generate shorter TPX-forming fragments. Other examples of TPX-based promoter switches are known. For example, the *SPHK1* (sphingosine kinase 1) downstream promoter produces an antisense *KHPS1* lncRNA that initiates TPX formation at an upstream promoter, leading to enhanced expression of the sense transcript [135]. The human β-globin locus also employs a TPX

switch to downregulate ε- and γ-globin upstream promoters and to enhance expression of the downstream β-globin TPX-forming transcript [136].

The lncRNAs involved with gene suppression generally have low expression and are tightly associated with chromatin [10]. The transcripts almost always arise from elements identified by the ENCODE project as enhancers, with bidirectional transcription common [137,138]. They may be intergenic, overlap protein-coding genes, or only introns. Some are spliced, capped, and polyadenylated, while others are not. Identifying these lncRNAs and mapping their interactions is challenging experimentally. The binding of lncRNAs with neighboring regions is most often reported [102], leading to models in which lncRNAs bring regions within the same chromosome together to promote epigenetic modifications or topological domain formation [135,136]. The bridging by lncRNAs of nearby enhancers to promoters may switch gene expression either on or off or to vary the RNA isoform produced [101,136,139].

Other lncRNAs, such as MALAT1 and NEAT1, are highly expressed and interact in sequence-specific ways with loci on other chromosomes, and can scaffold protein assemblies in both the nucleus and the cytoplasm [11]. The loss of a lncRNA can cause disease. In the case of phenylketonuria, the enzyme PAH (phenylalanine hydroxylase) activity is decreased by loss of the lncRNA *HULC* (hepatocellular carcinoma up-regulated long noncoding RNA) without any mutation of *PAH*. *HULC* facilitates PAH-substrate and PAH-cofactor interactions. In contrast, another intronic lncRNA *Pair* (PAH activating lincRNA) inhibits PAH activity [140]. LncRNAs from both mice and humans have also been associated with other visible phenotypes [11,141].

13. TPXs and EREs

The phasing of nucleosomes is also of importance in regulating expression of AEs, initially providing a means to halt transcription and retrotransposition of these SINEs [142]. Consequently, of the more than ~1.3 million AEs in the human genome, only 17,249 are actively transcribed by RNA polymerase 3 (Pol3), despite the intact A and B promoter boxes within these sequences able to engage Pol3 [143,144]. The few AEs transcribed by Pol3 exhibit cell-specific expressions from upstream cell-specific enhancer regions. Their transcription increase the readout of genes by RNA polymerase 2 (Pol2), with the AEs acting as enhancer SINEs (eSINEs) and likely leads to Z-DNA and GQ formation in the NFRs created [38,145]. Treatment with retinoic acid also induces the expression of Pol3 eSINEs through the DR2 repeats present in many of these elements, resulting in the enhanced expression of neighboring Pol2-transcribed genes. The small repeat-induced RNAs (riRNAs) produced from the AEs play a role in negatively regulating the cytoplasmic levels of Pol2 transcripts through Ago3-mediated RNA interference pathways [146].

In general, AEs within mRNAs are not well targeted by miRNAs due to their secondary structure and the sequence changes that result from RNA editing [147]. Yet studies identify miRNAs that map predominantly to the AE sense strand [127,148]. The possibility that these miRNAs act in the nucleus to phase nucleosomes over Alu elements has not been addressed, although the enrichment of TPX formation in Alu has been demonstrated in experimental studies [101,120]. However, enhancer regions are enriched for AEs that are also engaged by AGO1 protein bound to RNA [107,149]. AGO1 knockdown cell disrupts chromatin organization and globally alters gene expression [150,151]. Interestingly, the KCNQ1OT1 lncRNA induces genome-wide silencing of AEs through TPX formation involving the ARS found in AEs by localizing HP1α (encoded by *CBX5*) to the site [152].

Another notable example relates to the murine EREs that play an important role in bootstrapping the early steps of embryogenesis by activating the expression of zygotic genes (Figure 4) [153–155]. Expression of LINEs pre-implantation is associated with TPX

formation, as detected with antiserum raised against d(A:T)*T (Figure 4C, adapted from the study of [153]). The later failure to repress LINE-1 transcription leads to developmental arrest [155]. The MTA (Mouse Transcript A) LTR drives the expression of embryogenic genes at the two-cell stage and is necessary to initiate the transition from maternal programming to embryonic gene activation [156].
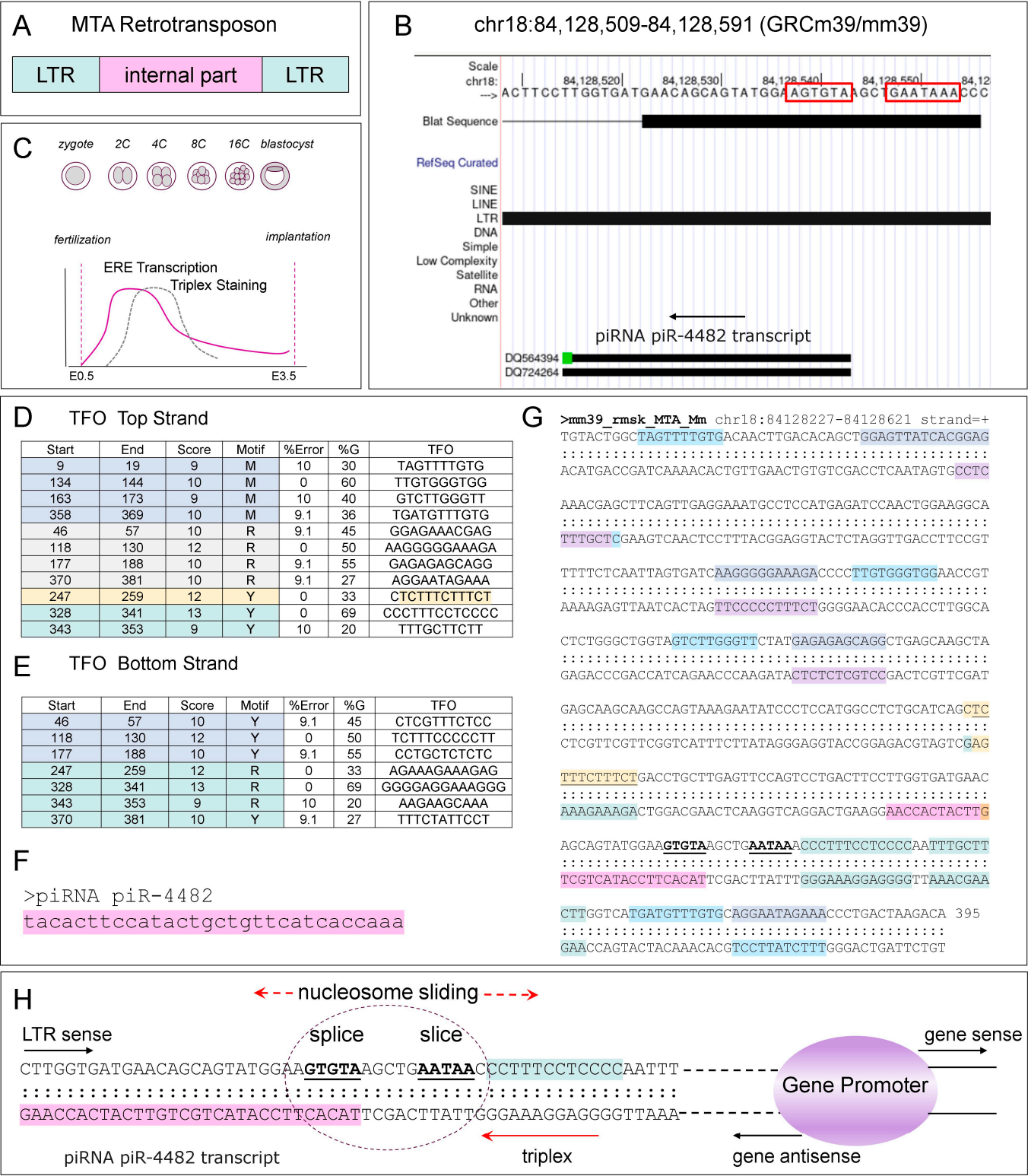

Figure 4. MTA Long Terminal Repeat (LTR) retrotransposon plays important roles in early embryonic development. (**A**) The structure of the MTA LTR. (**B**) The 369 nt MTA transcript shown contains a splice

donor site AGTGT and a AAUA cleavage site close to each other (both boxed in red). The antisense strand encodes the piR-4482 piRNA. (**C**) Like other EREs at the 2-cell stage, MTAs are highly expressed. Slightly later, TPX formation detected with a d(A:T)*T antiserum also peaks (adapted from the study of [153]). (**D**,**E**) show sequences and positions of TPX-forming sequences in the MTA sequence on the top and bottom strands. They are color coded by whether they have a pyrimidine (Y), purine (R), or mixed (M) purine and pyrimidine third strand. (**F**) The sequence of piR-4482 (fuchsia background color). (**G**) The sequence context of the potential TPX-forming sequences (using the same color coding as in (**D**,**E**), relative to the splice and cleavage sites (in bold font and underlined). (**H**) A dotted circle shows the location of the splice and cleavage sites in the MTA promoter region. The piR-4482 sequence (with a fuchsia background color) overlaps the splice site. A long TPX-forming sequence (position 328 to 353) is immediately 3′ to the cleavage site that are underlined and in bold font. These two sequence elements likely impact the switch from the MTA throughout the genome to the cellular promoters essential for bootstrapping zygotic gene expression.

In this pathway, the MTA promoter drives the production of a chimeric transcript by splicing the retroviral RNA to the transcript from the downstream gene. The MTA acts as an alternative promoter and encodes the first exon for a subset of genes [156]. Later, the MTA transcript undergoes cleavage rather than splicing, with the downstream gene promoter then driving gene transcription. The MTA element contains adjacent ERE splice and polyadenylation motifs (Figure 4A,B,H). The nucleosome placement may either favor splicing, or instead support transcript cleavage at the AAUAA site. Transcription from the MTA promoter likely favors the former, while antisense transcription from the downstream promoter would favor the latter. In both cases, chromatin remodelers that move nucleosomes in the direction of transcription finalize the placement of nucleosomes. The outcome depends on which promoter is currently active.

The MTA sequences encodes other regulatory elements that likely effect the switch in promoter usage. These include TPX-forming sequences predicted by the PATO program [92] (Figure 4C,D). A potential long TPX-forming sequence lies between positions 328 and 353 of the MTA sequence and just 3′ to the AAUAA cleavage site (Figure 4G,H). TPX formation could result from the fold back of the spliced transcript initiated from the MTA promoter. The TPX formed would facilitate use of the downstream cellular promoter by positioning nucleosomes to block splicing and promoting cleavage of the MTA transcript. The upstream transcript would be lost. The downstream promoter then is no longer subject to transcriptional interference mediated by processing of this RNA. The outcome is similar to the regulation of the rRNA gene by *PAPAS* (Figure 2). A positive feedback loop could then develop where the piR-4482 cleaved from the antisense transcript precludes the future use of the MTA promoter by directing its incorporation into constitutive heterochromatin (Figure 4F,G,H). The antisense transcript would also enable recruitment of tissue-specific enhancers to further amplify transcription from the cellular promoter.

14. lncRNAs, sRNAs, and Chromatin Condensates

The interplay between lncRNAs and sRNAs with promoter transcripts enables various regulatory interactions, including both cis and trans interactions, beyond the creation of NFRs [157]. The NFRs created with D:D*R specify the available transcription factor binding sites and flipons in the promoter and enhancer regions. The sense and antisense transcripts produced can then impact the expression of genes in the neighborhood by transcriptional interference, as described for the lncRNA upstream of the dihydrofolate reductase gene [158]. It was proposed in this case that TPX formation prevents TFIIB engagement with DNA. Other lncRNAs acting in cis also produce transcriptional interference, although TPX formation has not been experimentally assessed. The lncRNAs include *Chaserr* (*CHD2*-adjacent, suppressive regulatory RNA), which downregulates *Chd22* (Chromodomain Helicase DNA

Binding Protein 2) [159], and the *Airn*-induced H4K9me3 of the *Igfr2* gene. In both cases, the only requirement was the overlap of the lncRNA transcript with the gene promoter [160].

A subset of lncRNAs recruits the polycomb repressive complex 2 (PRC2) to promoters through TPX formation, with the PRC2 enhancer of zeste 2 polycomb repressive complex 2 subunit (EZH2) docking to a GQ formed by the lncRNA [161]. Examples of such lncRNAs include *MEG3*, which targets genes in the TGFβ (transforming growth factor beta) pathway [139], the heart and body wall-specific *Fendrr*, the radiation-induced *Particle*, and the WNT pathway regulator *HOTAIR* lncRNAs [162–164]. A computational analysis of super enhancers (SEs), defined as enhancers in close genomic proximity with unusually high levels of mediator binding, from 27 different human cell and tissue types identified 442 unique super-long noncoding RNA (super-lncRNA) transcripts, 70% of which were tissue-restricted and 80% of which had at least a single TPX-forming site. Triplexator predicted that, on average, each super-lncRNA can form D:D*R TPXs with 25 SEs, with a single SE targeted by multiple lncRNAs at many different, but evenly distributed locations (10–50 sites on average) [149]. Over 40% of SE sites were predicted to bind super-lncRNAs derived from EREs, with 34% from SINEs, 9% from LINEs, 4% from LTRs, and 2% from DNA transposons. A separate study revealed that the distribution of Alu elements within lncRNAs differed notably from coding genes, with Alus distributed throughout, with an increase in right arm monomers in the central region [165].

The outcomes also depend on the sequence features of the lncRNAs anchored through TPX formation (Figure 5A). Each lncRNAs delivers a set of protein-binding motifs that form a scaffold for assembling different cellular machines with various functionalities [166]. The lncRNAs can rapidly evolve by adding or subtracting sequence motifs to modify the condensates they nucleate [51]. They lack the sequence restraints that restrict the variation in protein-coding transcripts and represent a rapid way to evolve novel nuclear architectures [167]. The sequence overlap between different lncRNAs also provides functional redundancy. The shared TPX forming and protein binding motifs allows other lncRNAs to compensate for the loss of any one lncRNA.

Gene regulation may involve sRNAs processed from lncRNAs. For example, the H19 lncRNA, the first lncRNA identified, encodes miR-675. The gene is highly abundant in embryonic development and is important in skeletal muscle development [168–170]. The induction of gene silencing is possible with miRNAs that target sites in chromatin-associated RNA many kilobases away from an active promoter [171]. The miRNAs themselves may be produced by nearby ncRNAs. Indeed, it is possible to identify a subset of 303 miRNAs that are nuclear, produced from chromosomal regions with active chromosomal marks, and whose expression correlates with nearby genes, known as nuclear activating miRNAs (namiRNAs) [172,173]. These miRNAs can be overlapped with a set of high-confidence miRNAs (hc-miRNA, n = 504) [174] and with a set of miRNAs, called saRNA, that are proposed to activate gene expression. The saRNA were identified using RNA ChIP seq with Pol 2 antibodies. They have sequence matches within 20 bases of a promoter TATA box and also have a minimum free energy less than −27.6 kcal/mole. There are 72 miRNAs in the overlap of hc-miRNA, namiRNA, and saRNA (listed in Figure 5C).
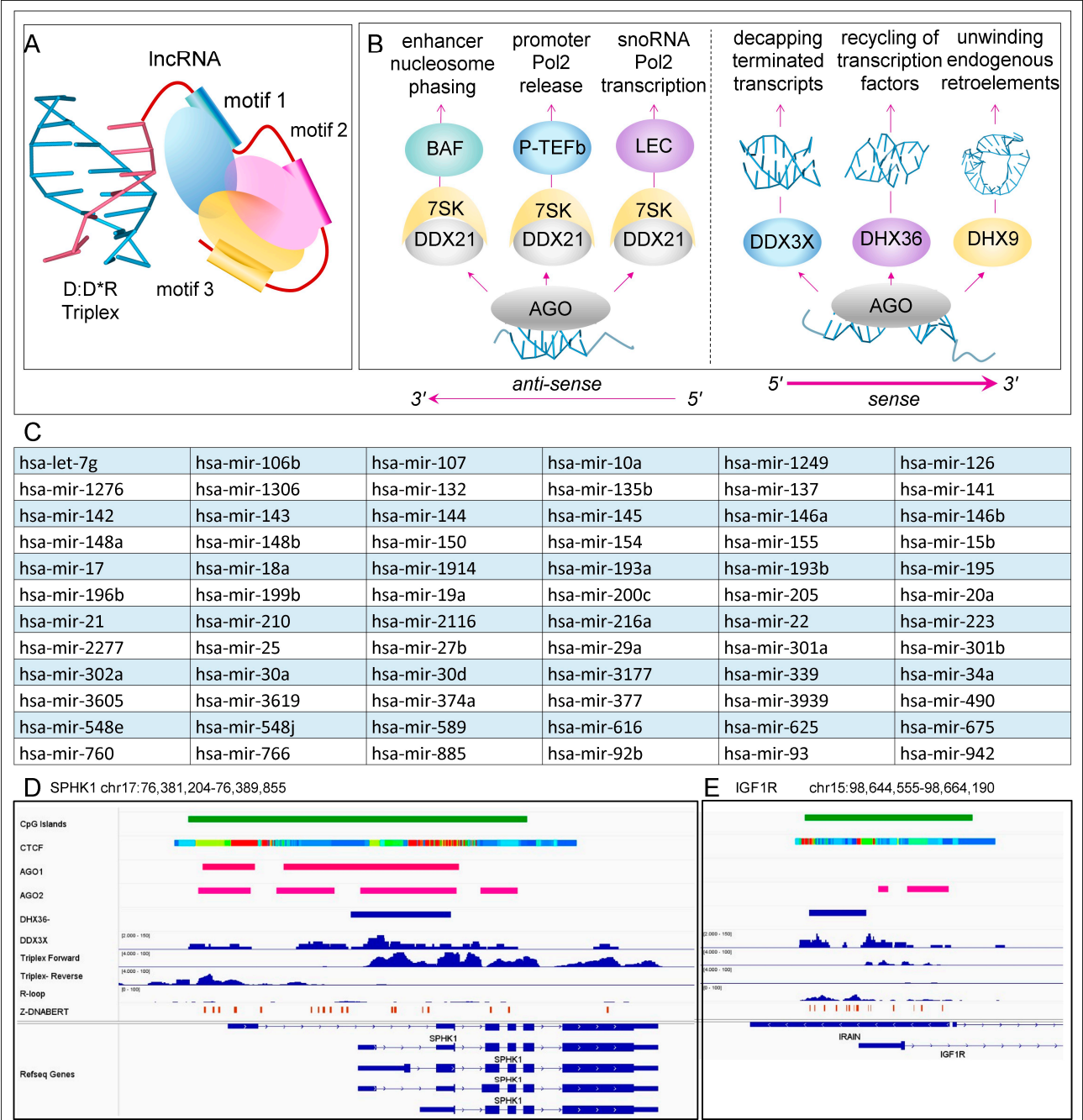
Figure 5. Noncoding RNA effectors: (**A**) lncRNAs anchored by TPX formation can localize different protein machines to a site through the motifs they contain. (**B**) Through various interactions with RNA helicases, AGO-bound miRNAs can activate Pol2 transcription. At enhancers, DDX2 acts through the Brahma-associated factor (BAF) complex to phase nucleosomes. At promoters, the DDX21 helicase disassembles the 7SK ribonucleoprotein to release Pol 2 from the pause site of cellular genes and also stimulates the transcription of snoRNAs through the little elongation complex (LEC). A different set of helicases (DDX3X, DHX9, and DHX36) enables the resolution of different flipon structures formed by RNA to release bound proteins and that decap prematurely terminated transcripts, enabling their removal by nucleases. (**C**) miRNAs that potentially activate transcription, derived by overlapping sets of well-validated miRNAs that are encoded by enhancers, bound by Pol2, and bind sequences within 20 nts of a promoter TATA box at minimum free energies of less than −27.6 kcal/mole. (**D**,**E**) Examples from the human *SPHK1* and *IGFR1* genes showing the overlap of helicase and argonaute protein binding with experimentally determined TPX using data from the ChIP-Atlas resource (https://chip-atlas.org/peak_browser (accessed on 8 April 2025)) and GSE120850 [101]. Genomic coordinates are from hg38.

15. Helicases and Enhancer–Promoter RNA Interactions

Interestingly, in the case of miR-34a, an interaction between the AGO2-bound miRNA activates the RNA helicase DDX21. The enzyme then induces the release of Pol2 from the pause site through an interaction with the 7SK complex that otherwise stalls transcript elongation [175]. In other studies, an interaction between 7SK and the Brahma-associated factor (BAF) chromatin remodeler delineates the NFR of enhancers and is particularly important for demarcating the SE [176]. Phasing nucleosomes in SEs is necessary to counter the convergent transcription arising from such closely spaced transcriptional units, both in cells and during development. The precise localization of nucleosomes is particularly important during embryogenesis, when the depletion of parental chromatin leads to pervasive transcription [177]. The nucleosome placement involves an embryonic version of the BAF [178]. Interestingly, DDX21 is also involved in the interaction between 7SK and the BAF. The finding is consistent with a model where AGO-bound miRNAs localize the complex (Figure 5B).

The different interactions involving helicases may also explain some of the divergent results on transcription obtained with nuclear miRNAs (Figure 5B). Previously the different outcomes were thought to depend on whether or not the target-bound miRNA activated the AGO2 slicer activity and transcript cleavage [179]. Instead, the outcome may depend on whether the miRNA complex localizes DDX12 close enough to a paused polymerase to disassemble the 7SK ribonucleoprotein and initiate transcript elongation. In contrast, other interactions that place DDX12 near BAF may help phase of nucleosomes in enhancers, or activate the little elongation complex that drives snoRNA transcription by Pol2 [180] (Figure 5B). Interestingly, the regulation of Pol1 transcription by a snoRNA complex also depends upon DDX21 [181].

At promoters, the miRNA guided engagement of AGO may enable the resolution of TRP by DDX3X, GQ by DHX36, and dsRNA by DHX9. The complexes formed then decap prematurely terminated transcripts, resolve R-loops, and recycle the factors necessary to reestablish a pre-initiation complex [28,182–184]. The role of helicases in modulating flipon conformation has also been recently reviewed [38,185], as has the setting of flipon conformation by miRNAs [186]. In all these cases, helicases enable setting and resetting of gene expression. The helicases involved mostly differ from those involved in the canonical cytoplasmic miRNA pathway, where the interaction of AGO2 with DDX6 usually initiates the engagement of the microRNA-induced silencing complex [187]. Other AGO proteins with helicases are known, but the regulatory interactions are less well characterized [188,189].

ENCODE data capture many of the interactions at promoters between flipons and helicases (Figure 5D,E). The overlap of AGO, DDX3X, and experimentally determined triplex formation is shown for the SPHK1 gene. There are no R-loops reported in this region [101], Both DDX3X and DHX36 localize over different regions of the insulin-like growth factor 1 receptor (IGFR1) gene. DDX3X and AGO2 overlap, while DHX36 is detected over R-loop regions where G-quadruplexes can form. Both of the gene promoters are in CpG-rich regions of the genome. They are both bound by the CCCTC-binding factor (CTCF) that facilitates the formation of chromatin loops.

16. Summary

The transactions between flipons and condensates influence gene expression by defining the NFR and forming RNA clouds that position protein complexes at enhancers and promoters (Figure 6). The RNA helicases color the chromatin in various hues through the interplay of lncRNAs with sRNAs and facilitate the transition from one chromatin state to another. Flipons play an important role in these processes. T-flipons set the stage by phasing nucleosomes, demarcating both the available transcription factor binding sites

and the G- and Z-flipons necessary to initiate and reset promoters for the next round of transcription. The RNAs produced from these open regions amplify these effects by increasing the number of available protein binding sites to capture the proteins released when promoters are reset. This ensures the proteins are available to initiate the next round of transcription. This process also facilitates the removal of proteins from DNA during the transition between active and suppressed states. At each stage, the RNA cloud also turns over. The churn is energetically costly, but facilitates promoter updates based on the current cellular state. The cycles created and counteracted constantly recolor chromatin to channel cell fate.
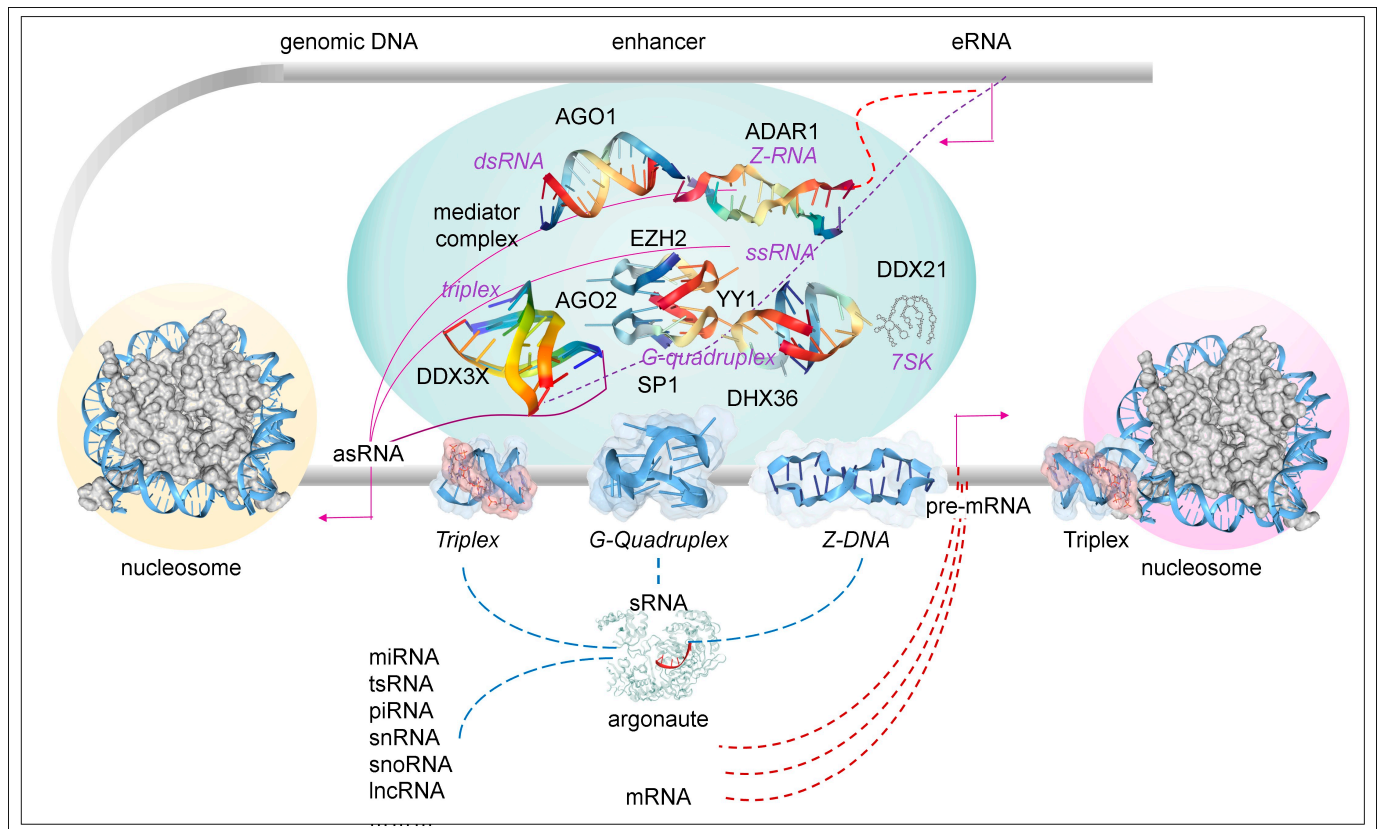

Figure 6. TPX Biology. Both lncRNAs and sRNAs play roles in the phasing of nucleosomes by chromatin remodelers. This process creates nucleosome-free regions that expose transcription binding sites and free flipons to adopt alternative conformations that are important for reinitiating transcription after each cycle [49]. The enhancer (red dashed lines) and antisense RNAs (purple solid lines) interact to form condensates that form various flipon structures and partially single-stranded regions to which helicases bind. The RNA cloud localizes chromatin-modifying complexes that color the histones with different epigenetic marks. Proteins that interact with flipons can partition between DNA and DNA binding sites to allow each to reset and to maintain each at a high local concentration. The turnover of condensates by helicases enables updates from inputs from elsewhere in the cell, including sRNAs, proteins, and signaling molecules. The processing of ncRNAs that generates argonaute bound sRNAs can occur in the nucleus (blue dashed lines) or in the cytoplasm (brown dashed lines). The different flipon RNA structures formed are tagged with purple, italicized labels. The arrows parallel to the genomic DNA show the direction of transcription. The placement of nucleosomes and flipons in enhancers are not shown but are similar to those drawn for promoters. Abbreviations: dsRNA, double-stranded RNA; ssRNA, single-stranded RNA; asRNA, antisense RNA; and sRNA, small RNA.

17. Future

The current view of the genome and the role of repeat sequences is undergoing rapid revision. New technologies and concepts are improving our understanding of how cells evolve and function. These insights will lead to novel strategies for treating disease by resetting disease states and preventing their progression. The targeting of TPXs is now aided by the ENCODE mapping of cell-specific regulatory sites [8] and the advances in nucleic acid chemistries for clinical applications [190]. Vector-based delivery systems for circRNAs also promise new therapeutic approaches. None of these technologies were available when TFO therapeutics were first investigated at the turn of the century. Everything was trial and error, with no clinical candidates emerging [191].

Now with our high-resolution maps of RNA and DNA interactions and our newfound understanding of nuclear dynamics, we have fresh insights on how to proceed. One current challenge is to develop strategies to reposition nucleosomes and recolor cells' epigenetic markings. This approach aims to change the enhancer-promoter pairs that drive gene expression. Therapeutics based on circRNAs are an interesting area to explore, given the increased stability and specificity of the TPX formed. Another strategy is to target the condensates as they recycle. Therapeutics that target helicases to a particular location in the RNA cloud could potentially enhance or suppress gene expression. Therapeutics that bind loops or bulges in DNA produced when flipons adopt different conformations may produce the same outcome by changing the active flipon in a particular NFR. The proposed strategies will benefit from recently developed computational tools to map RNA interaction motifs in chromatin [192]. All these approaches are sequence-specific and allow for prolonged dosing of the therapeutic, as already demonstrated for anti-sense and siRNA therapeutics. Given that the delivery of nucleic acid therapeutics to cells via gymnosis is possible, but only in limited amounts, therapeutics that activate existing positive feedback loops in a cell are more likely to see success in the clinic [193].

Conflicts of Interest: The author is the founder of InsideOutBio, Inc and declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Judd, B.H. Genes and chromomeres: A puzzle in three dimensions. *Genetics* **1998**, *150*, 1–9. [CrossRef] [PubMed]
2. Filion, G.J.; van Bemmel, J.G.; Braunschweig, U.; Talhout, W.; Kind, J.; Ward, L.D.; Brugman, W.; de Castro, I.J.; Kerkhoven, R.M.; Bussemaker, H.J.; et al. Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. *Cell* **2010**, *143*, 212–224. [CrossRef] [PubMed]
3. Crick, F. Central dogma of molecular biology. *Nature* **1970**, *227*, 561–563. [CrossRef] [PubMed]
4. van Bakel, H.; Nislow, C.; Blencowe, B.J.; Hughes, T.R. Most "dark matter" transcripts are associated with known genes. *PLoS Biol.* **2010**, *8*, e1000371. [CrossRef]
5. Eisen, M.B.; van Bakel, H.; Nislow, C.; Blencowe, B.J.; Hughes, T.R. Response to "The Reality of Pervasive Transcription". *PLoS Biol.* **2011**, *9*, e1001102. [CrossRef]
6. Eisen, M.B.; Clark, M.B.; Amaral, P.P.; Schlesinger, F.J.; Dinger, M.E.; Taft, R.J.; Rinn, J.L.; Ponting, C.P.; Stadler, P.F.; Morris, K.V.; et al. The Reality of Pervasive Transcription. *PLoS Biol.* **2011**, *9*, e1000625. [CrossRef]
7. Djebali, S.; Davis, C.A.; Merkel, A.; Dobin, A.; Lassmann, T.; Mortazavi, A.; Tanzer, A.; Lagarde, J.; Lin, W.; Schlesinger, F.; et al. Landscape of transcription in human cells. *Nature* **2012**, *489*, 101–108. [CrossRef]
8. Consortium, E.P.; Moore, J.E.; Purcaro, M.J.; Pratt, H.E.; Epstein, C.B.; Shoresh, N.; Adrian, J.; Kawli, T.; Davis, C.A.; Dobin, A.; et al. Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* **2020**, *583*, 699–710. [CrossRef]
9. Consortium, E.P.; Snyder, M.P.; Gingeras, T.R.; Moore, J.E.; Weng, Z.; Gerstein, M.B.; Ren, B.; Hardison, R.C.; Stamatoyannopoulos, J.A.; Graveley, B.R.; et al. Perspectives on ENCODE. *Nature* **2020**, *583*, 693–698. [CrossRef]

10. Li, Y.; Syed, J.; Sugiyama, H. RNA-DNA Triplex Formation by Long Noncoding RNAs. *Cell Chem. Biol.* **2016**, *23*, 1325–1333. [CrossRef]

11. Mattick, J.S.; Amaral, P.P.; Carninci, P.; Carpenter, S.; Chang, H.Y.; Chen, L.L.; Chen, R.; Dean, C.; Dinger, M.E.; Fitzgerald, K.A.; et al. Long non-coding RNAs: Definitions, functions, challenges and recommendations. *Nat. Rev. Mol. Cell Biol.* **2023**, *24*, 430–447. [CrossRef] [PubMed]

12. Velazquez-Flores, M.; Ruiz Esparza-Garrido, R. Fragments derived from non-coding RNAs: How complex is genome regulation? *Genome* **2024**, *67*, 292–306. [CrossRef] [PubMed]

13. Meister, G. Argonaute proteins: Functional insights and emerging roles. *Nat. Rev. Genet.* **2013**, *14*, 447–459. [CrossRef] [PubMed]

14. Ozata, D.M.; Gainetdinov, I.; Zoch, A.; O'Carroll, D.; Zamore, P.D. PIWI-interacting RNAs: Small RNAs with big functions. *Nat. Rev. Genet.* **2019**, *20*, 89–108. [CrossRef]

15. Zhang, H.; Sim, G.; Kehling, A.C.; Adhav, V.A.; Savidge, A.; Pastore, B.; Tang, W.; Nakanishi, K. Target cleavage and gene silencing by Argonautes with cityRNAs. *Cell Rep.* **2024**, *43*, 114806. [CrossRef]

16. Zhang, Y.; Shi, J.; Chen, Q. tsRNAs: New players in mammalian retrotransposon control. *Cell Res.* **2017**, *27*, 1307–1308. [CrossRef]

17. Di Fazio, A.; Gullerova, M. An old friend with a new face: tRNA-derived small RNAs with big regulatory potential in cancer biology. *Br. J. Cancer* **2023**, *128*, 1625–1635. [CrossRef]

18. Sarkar, N.; Kumar, A. Paradigm shift: microRNAs interact with target gene promoters to cause transcriptional gene activation or silencing. *Exp. Cell Res.* **2025**, *444*, 114372. [CrossRef]

19. Li, L.C.; Okino, S.T.; Zhao, H.; Pookot, D.; Place, R.F.; Urakami, S.; Enokida, H.; Dahiya, R. Small dsRNAs induce transcriptional activation in human cells. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 17337–17342. [CrossRef]

20. Janowski, B.A.; Younger, S.T.; Hardy, D.B.; Ram, R.; Huffman, K.E.; Corey, D.R. Activating gene expression in mammalian cells with promoter-targeted duplex RNAs. *Nat. Chem. Biol.* **2007**, *3*, 166–173. [CrossRef]

21. Brenner, S. Refuge of spandrels. *Curr. Biol.* **1998**, *8*, R669. [CrossRef] [PubMed]

22. Britten, R.J.; Davidson, E.H. Gene Regulation for Higher Cells: A Theory. *Science* **1969**, *165*, 349–357. [CrossRef] [PubMed]

23. Huang, R.; Zhou, P.K. DNA damage repair: Historical perspectives, mechanistic pathways and clinical translation for targeted cancer therapy. *Signal Transduct. Target. Ther.* **2021**, *6*, 254. [CrossRef] [PubMed]

24. Felsenfeld, G.; Davies, D.R.; Rich, A. Formation of a Three-Stranded Polynucleotide Molecule. *J. Am. Chem. Soc.* **1957**, *79*, 2023–2024. [CrossRef]

25. Leisegang, M.S.; Warwick, T.; Stotzel, J.; Brandes, R.P. RNA-DNA triplexes: Molecular mechanisms and functional relevance. *Trends Biochem. Sci.* **2024**, *49*, 532–544. [CrossRef]

26. Hisey, J.A.; Masnovo, C.; Mirkin, S.M. Triplex H-DNA structure: The long and winding road from the discovery to its role in human disease. *NAR Mol. Med.* **2024**, *1*, ugae024. [CrossRef]

27. Arnott, S.; Chandrasekaran, R.; Marttila, C.M. Structures for polyinosinic acid and polyguanylic acid. *Biochem. J.* **1974**, *141*, 537–543. [CrossRef]

28. Chen, M.C.; Tippana, R.; Demeshkina, N.A.; Murat, P.; Balasubramanian, S.; Myong, S.; Ferre-D'Amare, A.R. Structural basis of G-quadruplex unfolding by the DEAH/RHA helicase DHX36. *Nature* **2018**, *558*, 465–469. [CrossRef]

29. Pohl, F.M.; Jovin, T.M. Salt-induced co-operative conformational change of a synthetic DNA: Equilibrium and kinetic studies with poly (dG-dC). *J. Mol. Biol.* **1972**, *67*, 375–396. [CrossRef]

30. Wang, A.H.; Hakoshima, T.; van der Marel, G.; van Boom, J.H.; Rich, A. AT base pairs are less stable than GC base pairs in Z-DNA: The crystal structure of d(m5CGTAm5CG). *Cell* **1984**, *37*, 321–331. [CrossRef]

31. Herbert, A. *Flipons and the Logic of Soft-Wired Genomes*, 1st ed.; CRC Press: Boca Raton, FL, USA, 2024.

32. Gehring, K.; Leroy, J.L.; Gueron, M. A tetrameric DNA structure with protonated cytosine.cytosine base pairs. *Nature* **1993**, *363*, 561–565. [CrossRef] [PubMed]

33. Guneri, D.; Alexandrou, E.; El Omari, K.; Dvorakova, Z.; Chikhale, R.V.; Pike, D.T.S.; Waudby, C.A.; Morris, C.J.; Haider, S.; Parkinson, G.N.; et al. Structural insights into i-motif DNA structures in sequences from the insulin-linked polymorphic region. *Nat. Commun.* **2024**, *15*, 7119. [CrossRef] [PubMed]

34. Herbert, A. A Genetic Instruction Code Based on DNA Conformation. *Trends Genet.* **2019**, *35*, 887–890. [CrossRef] [PubMed]

35. Herbert, A. The Simple Biology of Flipons and Condensates Enhances the Evolution of Complexity. *Molecules* **2021**, *26*, 4881. [CrossRef]

36. Lightfoot, H.L.; Hagen, T.; Tatum, N.J.; Hall, J. The diverse structural landscape of quadruplexes. *FEBS Lett.* **2019**, *593*, 2083–2102. [CrossRef]

37. Spiegel, J.; Cuesta, S.M.; Adhikari, S.; Hansel-Hertsch, R.; Tannahill, D.; Balasubramanian, S. G-quadruplexes are transcription factor binding hubs in human chromatin. *Genome Biol.* **2021**, *22*, 117. [CrossRef]

38. Herbert, A. A Compendium of G-Flipon Biological Functions That Have Experimental Validation. *Int. J. Mol. Sci.* **2024**, *25*, 10299. [CrossRef]

39. Zhang, T.; Yin, C.; Fedorov, A.; Qiao, L.; Bao, H.; Beknazarov, N.; Wang, S.; Gautam, A.; Williams, R.M.; Crawford, J.C.; et al. ADAR1 masks the cancer immunotherapeutic promise of ZBP1-driven necroptosis. *Nature* **2022**, *606*, 594–602. [CrossRef]

40. Jiao, H.; Wachsmuth, L.; Wolf, S.; Lohmann, J.; Nagata, M.; Kaya, G.G.; Oikonomou, N.; Kondylis, V.; Rogg, M.; Diebold, M.; et al. ADAR1 averts fatal type I interferon induction by ZBP1. *Nature* **2022**, *607*, 776–783. [CrossRef]

41. Hubbard, N.W.; Ames, J.M.; Maurano, M.; Chu, L.H.; Somfleth, K.Y.; Gokhale, N.S.; Werner, M.; Snyder, J.M.; Lichauco, K.; Savan, R.; et al. ADAR1 mutation causes ZBP1-dependent immunopathology. *Nature* **2022**, *607*, 769–775. [CrossRef]

42. de Reuver, R.; Verdonck, S.; Dierick, E.; Nemegeer, J.; Hessmann, E.; Ahmad, S.; Jans, M.; Blancke, G.; Van Nieuwerburgh, F.; Botzki, A.; et al. ADAR1 prevents autoinflammation by suppressing spontaneous ZBP1 activation. *Nature* **2022**, *607*, 784–789. [CrossRef] [PubMed]

43. Wittig, B.; Dorbic, T.; Rich, A. The level of Z-DNA in metabolically active, permeabilized mammalian cell nuclei is regulated by torsional strain. *J. Cell Biol.* **1989**, *108*, 755–764. [CrossRef] [PubMed]

44. Wittig, B.; Wolfl, S.; Dorbic, T.; Vahrson, W.; Rich, A. Transcription of human c-myc in permeabilized nuclei is associated with formation of Z-DNA in three discrete regions of the gene. *EMBO J.* **1992**, *11*, 4653–4663. [CrossRef]

45. Wolfl, S.; Martinez, C.; Rich, A.; Majzoub, J.A. Transcription of the human corticotropin-releasing hormone gene in NPLC cells is correlated with Z-DNA formation. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 3664–3668. [CrossRef]

46. Kwon, J.A.; Rich, A. Biological function of the vaccinia virus Z-DNA-binding protein E3L: Gene transactivation and antiapoptotic activity in HeLa cells. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 12759–12764. [CrossRef]

47. Zhang, J.; Ohta, T.; Maruyama, A.; Hosoya, T.; Nishikawa, K.; Maher, J.M.; Shibahara, S.; Itoh, K.; Yamamoto, M. BRG1 interacts with Nrf2 to selectively mediate HO-1 induction in response to oxidative stress. *Mol. Cell. Biol.* **2006**, *26*, 7942–7952. [CrossRef]

48. Georgakopoulos-Soares, I.; Victorino, J.; Parada, G.E.; Agarwal, V.; Zhao, J.; Wong, H.Y.; Umar, M.I.; Elor, O.; Muhwezi, A.; An, J.Y.; et al. High-throughput characterization of the role of non-B DNA motifs on promoter function. *Cell Genom.* **2022**, *2*, 100111. [CrossRef]

49. Herbert, A. Flipons and small RNAs accentuate the asymmetries of pervasive transcription by the reset and sequence-specific microcoding of promoter conformation. *J. Biol. Chem.* **2023**, *299*, 105140. [CrossRef]

50. Herbert, A. The ancient Z-DNA and Z-RNA specific Zalpha fold has evolved modern roles in immunity and transcription through the natural selection of flipons. *R. Soc. Open Sci.* **2024**, *11*, 240080. [CrossRef]

51. Herbert, A. ALU non-B-DNA conformations, flipons, binary codes and evolution. *R. Soc. Open Sci.* **2020**, *7*, 200222. [CrossRef]

52. Maldonado, R.; Langst, G. The chromatin—Triple helix connection. *Biol. Chem.* **2023**, *404*, 1037–1049. [CrossRef] [PubMed]

53. Roberts, R.W.; Crothers, D.M. Stability and properties of double and triple helices: Dramatic effects of RNA or DNA backbone composition. *Science* **1992**, *258*, 1463–1466. [CrossRef] [PubMed]

54. Han, H.; Dervan, P.B. Sequence-specific recognition of double helical RNA and RNA.DNA by triple helix formation. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 3806–3810. [CrossRef]

55. Escudeé, C.; Francçois, J.-C.; Sun, J.-S.; Ott, G.; Sprinzl, M.; Garestier, T.s.; Heélene, J.-C. Stability of triple helices containing RNA and DNA strands: Experimental and molecular modeling studies. *Nucleic Acids Res.* **1993**, *21*, 5547–5553. [CrossRef]

56. Sandstrom, K.; Warmlander, S.; Graslund, A.; Leijon, M. A-tract DNA disfavours triplex formation. *J. Mol. Biol.* **2002**, *315*, 737–748. [CrossRef]

57. Cox, R.; Mirkin, S.M. Characteristic enrichment of DNA repeats in different genomes. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 5237–5242. [CrossRef]

58. Beal, P.A.; Dervan, P.B. Second Structural Motif for Recognition of DNA by Oligonucleotide-Directed Triple-Helix Formation. *Science* **1991**, *251*, 1360–1363. [CrossRef]

59. Escudé, C.; Sun, J.-S. DNA Major Groove Binders: Triple Helix-Forming Oligonucleotides, Triple Helix-Specific DNA Ligands and Cleaving Agents. In *DNA Binders and Related Subjects*; Topics in Current Chemistry; Springer: Berlin/Heidelberg, Germany, 2005; pp. 109–148.

60. Mills, M.; Arimondo, P.B.; Lacroix, L.; Garestier, T.; Klump, H.; Mergny, J.L. Chemical modification of the third strand: Differential effects on purine and pyrimidine triple helix formation. *Biochemistry* **2002**, *41*, 357–366. [CrossRef]

61. Sun, J.S.; Mergny, J.L.; Lavery, R.; Montenay-Garestier, T.; Helene, C. Triple helix structures: Sequence dependence, flexibility and mismatch effects. *J. Biomol. Struct. Dyn.* **1991**, *9*, 411–424. [CrossRef]

62. Kunkler, C.N.; Hulewicz, J.P.; Hickman, S.C.; Wang, M.C.; McCown, P.J.; Brown, J.A. Stability of an RNA*DNA-DNA triple helix depends on base triplet composition and length of the RNA third strand. *Nucleic Acids Res.* **2019**, *47*, 7213–7222. [CrossRef]

63. Abu Almakarem, A.S.; Petrov, A.I.; Stombaugh, J.; Zirbel, C.L.; Leontis, N.B. Comprehensive survey and geometric classification of base triples in RNA structures. *Nucleic Acids Res.* **2012**, *40*, 1407–1423. [CrossRef] [PubMed]

64. Fakharzadeh, A.; Zhang, J.; Roland, C.; Sagui, C. Novel eGZ-motif formed by regularly extruded guanine bases in a left-handed Z-DNA helix as a major motif behind CGG trinucleotide repeats. *Nucleic Acids Res.* **2022**, *50*, 4860–4876. [CrossRef] [PubMed]

65. Pan, F.; Xu, P.; Roland, C.; Sagui, C.; Weninger, K. Structural and Dynamical Properties of Nucleic Acid Hairpins Implicated in Trinucleotide Repeat Expansion Diseases. *Biomolecules* **2024**, *14*, 1278. [CrossRef] [PubMed]

66. Cooney, M.; Czernuszewicz, G.; Postel, E.H.; Flint, S.J.; Hogan, M.E. Site-Specific Oligonucleotide Binding Represses Transcription of the Human c-myc Gene In Vitro. *Science* **1988**, *241*, 456–459. [CrossRef]

67. Wang, G.; Vasquez, K.M. Naturally occurring H-DNA-forming sequences are mutagenic in mammalian cells. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 13448–13453. [CrossRef]

68. Esain-Garcia, I.; Kirchner, A.; Melidis, L.; Tavares, R.C.A.; Dhir, S.; Simeone, A.; Yu, Z.; Madden, S.K.; Hermann, R.; Tannahill, D.; et al. G-quadruplex DNA structure is a positive regulator of MYC transcription. *Proc. Natl. Acad. Sci. USA* **2024**, *121*, e2320240121. [CrossRef]

69. Brown, J.A. Unraveling the structure and biological functions of RNA triple helices. *Wiley Interdiscip. Rev. RNA* **2020**, *11*, e1598. [CrossRef]

70. Kim, N.K.; Zhang, Q.; Zhou, J.; Theimer, C.A.; Peterson, R.D.; Feigon, J. Solution structure and dynamics of the wild-type pseudoknot of human telomerase RNA. *J. Mol. Biol.* **2008**, *384*, 1249–1261. [CrossRef]

71. Ghanim, G.E.; Fountain, A.J.; van Roon, A.M.; Rangan, R.; Das, R.; Collins, K.; Nguyen, T.H.D. Structure of human telomerase holoenzyme with bound telomeric DNA. *Nature* **2021**, *593*, 449–453. [CrossRef]

72. Skeparnias, I.; Bou-Nader, C.; Anastasakis, D.G.; Fan, L.; Wang, Y.X.; Hafner, M.; Zhang, J. Structural basis of MALAT1 RNA maturation and mascRNA biogenesis. *Nat. Struct. Mol. Biol.* **2024**, *31*, 1655–1668. [CrossRef]

73. Gast, M.; Nageswaran, V.; Kuss, A.W.; Tzvetkova, A.; Wang, X.; Mochmann, L.H.; Rad, P.R.; Weiss, S.; Simm, S.; Zeller, T.; et al. tRNA-like Transcripts from the NEAT1-MALAT1 Genomic Region Critically Influence Human Innate Immunity and Macrophage Functions. *Cells* **2022**, *11*, 3970. [CrossRef] [PubMed]

74. Brown, J.A.; Kinzig, C.G.; DeGregorio, S.J.; Steitz, J.A. Methyltransferase-like protein 16 binds the 3′-terminal triple helix of MALAT1 long noncoding RNA. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 14013–14018. [CrossRef] [PubMed]

75. Satterwhite, E.R.; Mansfield, K.D. RNA methyltransferase METTL16: Targets and function. *Wiley Interdiscip. Rev. RNA* **2022**, *13*, e1681. [CrossRef]

76. Schievelbein, M.J.; Resende, C.; Glennon, M.M.; Kerosky, M.; Brown, J.A. Global RNA modifications to the MALAT1 triple helix differentially affect thermostability and weaken binding to METTL16. *J. Biol. Chem.* **2024**, *300*, 105548. [CrossRef]

77. Nomura, M. Assembly of bacterial ribosomes. *Science* **1973**, *179*, 864–873. [CrossRef]

78. Tapescu, I.; Cherry, S. DDX RNA helicases: Key players in cellular homeostasis and innate antiviral immunity. *J. Virol.* **2024**, *98*, e0004024. [CrossRef]

79. Bohnsack, K.E.; Yi, S.; Venus, S.; Jankowsky, E.; Bohnsack, M.T. Cellular functions of eukaryotic RNA helicases and their links to human diseases. *Nat. Rev. Mol. Cell Biol.* **2023**, *24*, 749–769. [CrossRef]

80. Toyama, Y.; Shimada, I. NMR characterization of RNA binding property of the DEAD-box RNA helicase DDX3X and its implications for helicase activity. *Nat. Commun.* **2024**, *15*, 3303. [CrossRef]

81. Jolma, A.; Zhang, J.; Mondragon, E.; Morgunova, E.; Kivioja, T.; Laverty, K.U.; Yin, Y.; Zhu, F.; Bourenkov, G.; Morris, Q.; et al. Binding specificities of human RNA-binding proteins toward structured and linear RNA sequences. *Genome Res.* **2020**, *30*, 962–973. [CrossRef]

82. Klein, D.J.; Schmeing, T.M.; Moore, P.B.; Steitz, T.A. The kink-turn: A new RNA secondary structure motif. *EMBO J.* **2001**, *20*, 4214–4221. [CrossRef]

83. Schroeder, K.T.; McPhee, S.A.; Ouellet, J.; Lilley, D.M. A structural database for k-turn motifs in RNA. *RNA* **2010**, *16*, 1463–1468. [CrossRef] [PubMed]

84. Li, B.; Liu, S.; Zheng, W.; Liu, A.; Yu, P.; Wu, D.; Zhou, J.; Zhang, P.; Liu, C.; Lin, Q.; et al. RIP-PEN-seq identifies a class of kink-turn RNAs as splicing regulators. *Nat. Biotechnol.* **2024**, *42*, 119–131. [CrossRef] [PubMed]

85. Ge, P.; Islam, S.; Zhong, C.; Zhang, S. De novo discovery of structural motifs in RNA 3D structures through clustering. *Nucleic Acids Res.* **2018**, *46*, 4783–4793. [CrossRef]

86. D'Ascenzo, L.; Vicens, Q.; Auffinger, P. Identification of receptors for UNCG and GNRA Z-turns and their occurrence in rRNA. *Nucleic Acids Res.* **2018**, *46*, 7989–7997. [CrossRef]

87. Troisi, R.; Sica, F. Structural overview of DNA and RNA G-quadruplexes in their interaction with proteins. *Curr. Opin. Struct. Biol.* **2024**, *87*, 102846. [CrossRef]

88. Studer, M.K.; Ivanovic, L.; Weber, M.E.; Marti, S.; Jonas, S. Structural basis for DEAH-helicase activation by G-patch proteins. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 7159–7170. [CrossRef]

89. Warwick, T.; Brandes, R.P.; Leisegang, M.S. Computational Methods to Study DNA:DNA:RNA Triplex Formation by lncRNAs. *Noncoding RNA* **2023**, *9*, 10. [CrossRef]

90. Buske, F.A.; Bauer, D.C.; Mattick, J.S.; Bailey, T.L. Triplexator: Detecting nucleic acid triple helices in genomic and transcriptomic data. *Genome Res.* **2012**, *22*, 1372–1381. [CrossRef]

91. Kuo, C.C.; Hanzelmann, S.; Senturk Cetin, N.; Frank, S.; Zajzon, B.; Derks, J.P.; Akhade, V.S.; Ahuja, G.; Kanduri, C.; Grummt, I.; et al. Detection of RNA-DNA binding sites in long noncoding RNAs. *Nucleic Acids Res.* **2019**, *47*, e32. [CrossRef]

92. Amatria-Barral, I.; Gonzalez-Dominguez, J.; Tourino, J. PATO: Genome-wide prediction of lncRNA-DNA triple helices. *Bioinformatics* **2023**, *39*, btad134. [CrossRef]

93. Cicconetti, C.; Lauria, A.; Proserpio, V.; Masera, M.; Tamburrini, A.; Maldotti, M.; Oliviero, S.; Molineris, I. 3plex enables deep computational investigation of triplex forming lncRNAs. *Comput. Struct. Biotechnol. J.* **2023**, *21*, 3091–3102. [CrossRef] [PubMed]

94. He, S.; Zhang, H.; Liu, H.; Zhu, H. LongTarget: A tool to predict lncRNA DNA-binding motifs and binding sites via Hoogsteen base-pairing analysis. *Bioinformatics* **2015**, *31*, 178–186. [CrossRef] [PubMed]

95. Biffi, G.; Tannahill, D.; McCafferty, J.; Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* **2013**, *5*, 182–186. [CrossRef]

96. Javadekar, S.M.; Nilavar, N.M.; Paranjape, A.; Das, K.; Raghavan, S.C. Characterization of G-quadruplex antibody reveals differential specificity for G4 DNA forms. *DNA Res.* **2020**, *27*, dsaa024. [CrossRef]

97. Stollar, B.D.; Raso, V. Antibodies recognise specific structures of triple-helical polynucleotides built on poly(A) or poly(dA). *Nature* **1974**, *250*, 231–234. [CrossRef]

98. Gorab, E.; Amabis, J.M.; Stocker, A.J.; Drummond, L.; Stollar, B.D. Potential sites of triple-helical nucleic acid formation in chromosomes of *Rhynchosciara* (Diptera: Sciaridae) and *Drosophila melanogaster*. *Chromosome Res.* **2009**, *17*, 821–832. [CrossRef]

99. Lee, J.S.; Burkholder, G.D.; Latimer, L.J.; Haug, B.L.; Braun, R.P. A monoclonal antibody to triplex DNA binds to eucaryotic chromosomes. *Nucleic Acids Res.* **1987**, *15*, 1047–1061. [CrossRef]

100. Agazie, Y.M.; Lee, J.S.; Burkholder, G.D. Characterization of a new monoclonal antibody to triplex DNA and immunofluorescent staining of mammalian chromosomes. *J. Biol. Chem.* **1994**, *269*, 7019–7023. [CrossRef]

101. Senturk Cetin, N.; Kuo, C.C.; Ribarska, T.; Li, R.; Costa, I.G.; Grummt, I. Isolation and genome-wide characterization of cellular DNA:RNA triplex structures. *Nucleic Acids Res.* **2019**, *47*, 2306–2321. [CrossRef]

102. Bonetti, A.; Agostini, F.; Suzuki, A.M.; Hashimoto, K.; Pascarella, G.; Gimenez, J.; Roos, L.; Nash, A.J.; Ghilotti, M.; Cameron, C.J.F.; et al. RADICL-seq identifies general and cell type-specific principles of genome-wide RNA-chromatin interactions. *Nat. Commun.* **2020**, *11*, 1018. [CrossRef]

103. Xu, H.; Ye, J.; Zhang, K.X.; Hu, Q.; Cui, T.; Tong, C.; Wang, M.; Geng, H.; Shui, K.M.; Sun, Y.; et al. Chemoproteomic profiling unveils binding and functional diversity of endogenous proteins that interact with endogenous triplex DNA. *Nat. Chem.* **2024**, *16*, 1811–1821. [CrossRef] [PubMed]

104. Rosu, F.; Nguyen, C.H.; De Pauw, E.; Gabelica, V. Ligand binding mode to duplex and triplex DNA assessed by combining electrospray tandem mass spectrometry and molecular modeling. *J. Am. Soc. Mass. Spectrom.* **2007**, *18*, 1052–1062. [CrossRef] [PubMed]

105. Nguyen, T.C.; Cao, X.; Yu, P.; Xiao, S.; Lu, J.; Biase, F.H.; Sridhar, B.; Huang, N.; Zhang, K.; Zhong, S. Mapping RNA-RNA interactome and RNA structure in vivo by MARIO. *Nat. Commun.* **2016**, *7*, 12023. [CrossRef] [PubMed]

106. Sridhar, B.; Rivas-Astroza, M.; Nguyen, T.C.; Chen, W.; Yan, Z.; Cao, X.; Hebert, L.; Zhong, S. Systematic Mapping of RNA-Chromatin Interactions In Vivo. *Curr. Biol.* **2017**, *27*, 602–609. [CrossRef]

107. Liang, L.; Cao, C.; Ji, L.; Cai, Z.; Wang, D.; Ye, R.; Chen, J.; Yu, X.; Zhou, J.; Bai, Z.; et al. Complementary Alu sequences mediate enhancer-promoter selectivity. *Nature* **2023**, *619*, 868–875. [CrossRef]

108. Espinas, M.L.; Jimenez-Garcia, E.; Martinez-Balbas, A.; Azorin, F. Formation of triple-stranded DNA at d(GA.TC)n sequences prevents nucleosome assembly and is hindered by nucleosomes. *J. Biol. Chem.* **1996**, *271*, 31807–31812. [CrossRef]

109. Brown, P.M.; Fox, K.R. DNA triple-helix formation on nucleosome core particles. Effect of length of the oligopurine tract. *Eur. J. Biochem.* **1999**, *261*, 301–310. [CrossRef]

110. Segal, E.; Widom, J. Poly(dA:dT) tracts: Major determinants of nucleosome organization. *Curr. Opin. Struct. Biol.* **2009**, *19*, 65–71. [CrossRef]

111. Kubik, S.; Bruzzone, M.J.; Jacquet, P.; Falcone, J.L.; Rougemont, J.; Shore, D. Nucleosome Stability Distinguishes Two Different Promoter Types at All Protein-Coding Genes in Yeast. *Mol. Cell* **2015**, *60*, 422–434. [CrossRef]

112. Struhl, K.; Segal, E. Determinants of nucleosome positioning. *Nat. Struct. Mol. Biol.* **2013**, *20*, 267–273. [CrossRef]

113. Chereji, R.V.; Bryson, T.D.; Henikoff, S. Quantitative MNase-seq accurately maps nucleosome occupancy levels. *Genome Biol.* **2019**, *20*, 198. [CrossRef] [PubMed]

114. Prajapati, H.K.; Ocampo, J.; Clark, D.J. Interplay among ATP-Dependent Chromatin Remodelers Determines Chromatin Organisation in Yeast. *Biology* **2020**, *9*, 190. [CrossRef] [PubMed]

115. de Boer, C.G.; Hughes, T.R. Poly-dA:dT tracts form an in vivo nucleosomal turnstile. *PLoS ONE* **2014**, *9*, e110479. [CrossRef] [PubMed]

116. Barnes, T.; Korber, P. The Active Mechanism of Nucleosome Depletion by Poly(dA:dT) Tracts In Vivo. *Int. J. Mol. Sci.* **2021**, *22*, 8233. [CrossRef]

117. Amigo, R.; Raiqueo, F.; Tarifeno, E.; Farkas, C.; Gutierrez, J.L. Poly(dA:dT) Tracts Differentially Modulate Nucleosome Remodeling Activity of RSC and ISW1a Complexes, Exerting Tract Orientation-Dependent and -Independent Effects. *Int. J. Mol. Sci.* **2023**, *24*, 15245. [CrossRef]

118. Hampel, K.J.; Crosson, P.; Lee, J.S. Polyamines favor DNA triplex formation at neutral pH. *Biochemistry* **1991**, *30*, 4455–4459. [CrossRef]

119. Maldonado, R.; Schwartz, U.; Silberhorn, E.; Langst, G. Nucleosomes Stabilize ssRNA-dsDNA Triple Helices in Human Cells. *Mol. Cell* **2019**, *73*, 1243–1254.e6. [CrossRef]

120. Jalali, S.; Singh, A.; Maiti, S.; Scaria, V. Genome-wide computational analysis of potential long noncoding RNA mediated DNA:DNA:RNA triplexes in the human genome. *J. Transl. Med.* **2017**, *15*, 186. [CrossRef]

121. Paugh, S.W.; Coss, D.R.; Bao, J.; Laudermilk, L.T.; Grace, C.R.; Ferreira, A.M.; Waddell, M.B.; Ridout, G.; Naeve, D.; Leuze, M.; et al. MicroRNAs Form Triplexes with Double Stranded DNA at Sequence-Specific Binding Sites; a Eukaryotic Mechanism via which microRNAs Could Directly Alter Gene Expression. *PLoS Comput. Biol.* **2016**, *12*, e1004744. [CrossRef]

122. Schwammle, V.; Sidoli, S.; Ruminowicz, C.; Wu, X.; Lee, C.F.; Helin, K.; Jensen, O.N. Systems Level Analysis of Histone H3 Post-translational Modifications (PTMs) Reveals Features of PTM Crosstalk in Chromatin Regulation. *Mol. Cell. Proteom.* **2016**, *15*, 2715–2729. [CrossRef]

123. Kohestani, H.; Wereszczynski, J. The effects of RNA.DNA-DNA triple helices on nucleosome structures and dynamics. *Biophys. J.* **2023**, *122*, 1229–1239. [CrossRef] [PubMed]

124. Ludwig, N.; Leidinger, P.; Becker, K.; Backes, C.; Fehlmann, T.; Pallasch, C.; Rheinheimer, S.; Meder, B.; Stahler, C.; Meese, E.; et al. Distribution of miRNA expression across human tissues. *Nucleic Acids Res.* **2016**, *44*, 3865–3877. [CrossRef] [PubMed]

125. Toscano-Garibay, J.D.; Aquino-Jarquin, G. Transcriptional regulation mechanism mediated by miRNA-DNA*DNA triplex structure stabilized by Argonaute. *Biochim. Biophys. Acta* **2014**, *1839*, 1079–1083. [CrossRef]

126. Piriyapongsa, J.; Marino-Ramirez, L.; Jordan, I.K. Origin and evolution of human microRNAs from transposable elements. *Genetics* **2007**, *176*, 1323–1337. [CrossRef]

127. Lehnert, S.; Van Loo, P.; Thilakarathne, P.J.; Marynen, P.; Verbeke, G.; Schuit, F.C. Evidence for co-evolution between human microRNAs and Alu-repeats. *PLoS ONE* **2009**, *4*, e4456. [CrossRef]

128. Girard, A.; Sachidanandam, R.; Hannon, G.J.; Carmell, M.A. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* **2006**, *442*, 199–202. [CrossRef]

129. Schorn, A.J.; Gutbrod, M.J.; LeBlanc, C.; Martienssen, R. LTR-Retrotransposon Control by tRNA-Derived Small RNAs. *Cell* **2017**, *170*, 61–71.e11. [CrossRef]

130. Stamidis, N.; Zylicz, J.J. RNA-mediated heterochromatin formation at repetitive elements in mammals. *EMBO J.* **2023**, *42*, e111717. [CrossRef]

131. Mayer, C.; Schmitz, K.M.; Li, J.; Grummt, I.; Santoro, R. Intergenic transcripts regulate the epigenetic state of rRNA genes. *Mol. Cell* **2006**, *22*, 351–361. [CrossRef]

132. Mayer, C.; Neubert, M.; Grummt, I. The structure of NoRC-associated RNA is crucial for targeting the chromatin remodelling complex NoRC to the nucleolus. *EMBO Rep.* **2008**, *9*, 774–780. [CrossRef]

133. Schmitz, K.M.; Mayer, C.; Postepska, A.; Grummt, I. Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. *Genes Dev.* **2010**, *24*, 2264–2269. [CrossRef] [PubMed]

134. Zhao, Z.; Senturk, N.; Song, C.; Grummt, I. lncRNA PAPAS tethered to the rDNA enhancer recruits hypophosphorylated CHD4/NuRD to repress rRNA synthesis at elevated temperatures. *Genes Dev.* **2018**, *32*, 836–848. [CrossRef] [PubMed]

135. Blank-Giwojna, A.; Postepska-Igielska, A.; Grummt, I. lncRNA KHPS1 Activates a Poised Enhancer by Triplex-Dependent Recruitment of Epigenomic Regulators. *Cell Rep.* **2019**, *26*, 2904–2915.e2904. [CrossRef] [PubMed]

136. Zhou, Z.; Giles, K.E.; Felsenfeld, G. DNA.RNA triple helix formation can function as a cis-acting regulatory mechanism at the human beta-globin locus. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 6130–6139. [CrossRef]

137. Li, W.; Notani, D.; Rosenfeld, M.G. Enhancers as non-coding RNA transcription units: Recent insights and future perspectives. *Nat. Rev. Genet.* **2016**, *17*, 207–223. [CrossRef]

138. Hon, C.C.; Ramilowski, J.A.; Harshbarger, J.; Bertin, N.; Rackham, O.J.; Gough, J.; Denisenko, E.; Schmeier, S.; Poulsen, T.M.; Severin, J.; et al. An atlas of human long non-coding RNAs with accurate 5′ ends. *Nature* **2017**, *543*, 199–204. [CrossRef]

139. Mondal, T.; Subhash, S.; Vaid, R.; Enroth, S.; Uday, S.; Reinius, B.; Mitra, S.; Mohammed, A.; James, A.R.; Hoberg, E.; et al. MEG3 long noncoding RNA regulates the TGF-beta pathway genes through formation of RNA-DNA triplex structures. *Nat. Commun.* **2015**, *6*, 7743. [CrossRef]

140. Li, Y.; Tan, Z.; Zhang, Y.; Zhang, Z.; Hu, Q.; Liang, K.; Jun, Y.; Ye, Y.; Li, Y.C.; Li, C.; et al. A noncoding RNA modulator potentiates phenylalanine metabolism in mice. *Science* **2021**, *373*, 662–673. [CrossRef]

141. Chen, L.L.; Kim, V.N. Small and long non-coding RNAs: Past, present, and future. *Cell* **2024**, *187*, 6451–6485. [CrossRef]

142. Tanaka, Y.; Yamashita, R.; Suzuki, Y.; Nakai, K. Effects of Alu elements on global nucleosome positioning in the human genome. *BMC Genom.* **2010**, *11*, 309. [CrossRef]

143. Conti, A.; Carnevali, D.; Bollati, V.; Fustinoni, S.; Pellegrini, M.; Dieci, G. Identification of RNA polymerase III-transcribed Alu loci by computational screening of RNA-Seq data. *Nucleic Acids Res.* **2015**, *43*, 817–835. [CrossRef] [PubMed]

144. Zhang, X.O.; Gingeras, T.R.; Weng, Z. Genome-wide analysis of polymerase III-transcribed Alu elements suggests cell-type-specific enhancer function. *Genome Res.* **2019**, *29*, 1402–1414. [CrossRef] [PubMed]

145. Policarpi, C.; Crepaldi, L.; Brookes, E.; Nitarska, J.; French, S.M.; Coatti, A.; Riccio, A. Enhancer SINEs Link Pol III to Pol II Transcription in Neurons. *Cell Rep.* **2017**, *21*, 2879–2894. [CrossRef] [PubMed]

146. Hu, Q.; Tanasa, B.; Trabucchi, M.; Li, W.; Zhang, J.; Ohgi, K.A.; Rose, D.W.; Glass, C.K.; Rosenfeld, M.G. DICER- and AGO3-dependent generation of retinoic acid-induced DR2 Alu RNAs regulates human stem cell proliferation. *Nat. Struct. Mol. Biol.* **2012**, *19*, 1168–1175. [CrossRef]

147. Hoffman, Y.; Dahary, D.; Bublik, D.R.; Oren, M.; Pilpel, Y. The majority of endogenous microRNA targets within Alu elements avoid the microRNA machinery. *Bioinformatics* **2013**, *29*, 894–902. [CrossRef]

148. Smalheiser, N.R.; Torvik, V.I. Alu elements within human mRNAs are probable microRNA targets. *Trends Genet.* **2006**, *22*, 532–536. [CrossRef]

149. Soibam, B. Super-lncRNAs: Identification of lncRNAs that target super-enhancers via RNA:DNA:DNA triplex formation. *RNA* **2017**, *23*, 1729–1742. [CrossRef]

150. Allo, M.; Agirre, E.; Bessonov, S.; Bertucci, P.; Gomez Acuna, L.; Buggiano, V.; Bellora, N.; Singh, B.; Petrillo, E.; Blaustein, M.; et al. Argonaute-1 binds transcriptional enhancers and controls constitutive and alternative splicing in human cells. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 15622–15629. [CrossRef]

151. Shuaib, M.; Parsi, K.M.; Thimma, M.; Adroub, S.A.; Kawaji, H.; Seridi, L.; Ghosheh, Y.; Fort, A.; Fallatah, B.; Ravasi, T.; et al. Nuclear AGO1 Regulates Gene Expression by Affecting Chromatin Architecture in Human Cells. *Cell Syst.* **2019**, *9*, 446–458.e446. [CrossRef]

152. Zhang, X.; Jiang, Q.; Li, J.; Zhang, S.; Cao, Y.; Xia, X.; Cai, D.; Tan, J.; Chen, J.; Han, J.J. KCNQ1OT1 promotes genome-wide transposon repression by guiding RNA-DNA triplexes and HP1 binding. *Nat. Cell Biol.* **2022**, *24*, 1617–1629. [CrossRef]

153. Fadloun, A.; Le Gras, S.; Jost, B.; Ziegler-Birling, C.; Takahashi, H.; Gorab, E.; Carninci, P.; Torres-Padilla, M.E. Chromatin signatures and retrotransposon profiling in mouse embryos reveal regulation of LINE-1 by RNA. *Nat. Struct. Mol. Biol.* **2013**, *20*, 332–338. [CrossRef] [PubMed]

154. Tohonen, V.; Katayama, S.; Vesterlund, L.; Jouhilahti, E.M.; Sheikhi, M.; Madissoon, E.; Filippini-Cattaneo, G.; Jaconi, M.; Johnsson, A.; Burglin, T.R.; et al. Novel PRD-like homeodomain transcription factors and retrotransposon elements in early human development. *Nat. Commun.* **2015**, *6*, 8207. [CrossRef] [PubMed]

155. Jachowicz, J.W.; Bing, X.; Pontabry, J.; Boskovic, A.; Rando, O.J.; Torres-Padilla, M.E. LINE-1 activation after fertilization regulates global chromatin accessibility in the early mouse embryo. *Nat. Genet.* **2017**, *49*, 1502–1510. [CrossRef] [PubMed]

156. Peaston, A.E.; Evsikov, A.V.; Graber, J.H.; de Vries, W.N.; Holbrook, A.E.; Solter, D.; Knowles, B.B. Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev. Cell* **2004**, *7*, 597–606. [CrossRef]

157. Napoli, S. Targeting Promoter-Associated RNAs by siRNAs. *Methods Mol. Biol.* **2017**, *1543*, 209–219. [CrossRef]

158. Martianov, I.; Ramadass, A.; Serra Barros, A.; Chow, N.; Akoulitchev, A. Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. *Nature* **2007**, *445*, 666–670. [CrossRef]

159. Rom, A.; Melamed, L.; Gil, N.; Goldrich, M.J.; Kadir, R.; Golan, M.; Biton, I.; Perry, R.B.; Ulitsky, I. Regulation of CHD2 expression by the Chaserr long noncoding RNA gene is essential for viability. *Nat. Commun.* **2019**, *10*, 5092. [CrossRef]

160. Latos, P.A.; Pauler, F.M.; Koerner, M.V.; Senergin, H.B.; Hudson, Q.J.; Stocsits, R.R.; Allhoff, W.; Stricker, S.H.; Klement, R.M.; Warczok, K.E.; et al. Airn transcriptional overlap, but not its lncRNA products, induces imprinted Igf2r silencing. *Science* **2012**, *338*, 1469–1472. [CrossRef]

161. Song, J.; Gooding, A.R.; Hemphill, W.O.; Love, B.D.; Robertson, A.; Yao, L.; Zon, L.I.; North, T.E.; Kasinath, V.; Cech, T.R. Structural basis for inactivation of PRC2 by G-quadruplex RNA. *Science* **2023**, *381*, 1331–1337. [CrossRef]

162. Chu, C.; Qu, K.; Zhong, F.L.; Artandi, S.E.; Chang, H.Y. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol. Cell* **2011**, *44*, 667–678. [CrossRef]

163. Grote, P.; Wittler, L.; Hendrix, D.; Koch, F.; Wahrisch, S.; Beisaw, A.; Macura, K.; Blass, G.; Kellis, M.; Werber, M.; et al. The tissue-specific lncRNA Fendrr is an essential regulator of heart and body wall development in the mouse. *Dev. Cell* **2013**, *24*, 206–214. [CrossRef] [PubMed]

164. O'Leary, V.B.; Ovsepian, S.V.; Carrascosa, L.G.; Buske, F.A.; Radulovic, V.; Niyazi, M.; Moertl, S.; Trau, M.; Atkinson, M.J.; Anastasov, N. PARTICLE, a Triplex-Forming Long ncRNA, Regulates Locus-Specific Methylation in Response to Low-Dose Irradiation. *Cell Rep.* **2015**, *11*, 474–485. [CrossRef]

165. Kim, E.Z.; Wespiser, A.R.; Caffrey, D.R. The domain structure and distribution of Alu elements in long noncoding RNAs and mRNAs. *RNA* **2016**, *22*, 254–264. [CrossRef]

166. Shaath, H.; Vishnubalaji, R.; Elango, R.; Kardousha, A.; Islam, Z.; Qureshi, R.; Alam, T.; Kolatkar, P.R.; Alajez, N.M. Long non-coding RNA and RNA-binding protein interactions in cancer: Experimental and machine learning approaches. *Semin. Cancer Biol.* **2022**, *86*, 325–345. [CrossRef]

167. Thakur, J.; Henikoff, S. Architectural RNA in chromatin organization. *Biochem. Soc. Trans.* **2020**, *48*, 1967–1978. [CrossRef]

168. Bartolomei, M.S.; Zemel, S.; Tilghman, S.M. Parental imprinting of the mouse H19 gene. *Nature* **1991**, *351*, 153–155. [CrossRef]

169. Cai, X.; Cullen, B.R. The imprinted H19 noncoding RNA is a primary microRNA precursor. *RNA* **2007**, *13*, 313–316. [CrossRef]

170. Dey, B.K.; Pfeifer, K.; Dutta, A. The H19 long noncoding RNA gives rise to microRNAs miR-675-3p and miR-675-5p to promote skeletal muscle differentiation and regeneration. *Genes Dev.* **2014**, *28*, 491–501. [CrossRef]

171. Ouvrard, J.; Muniz, L.; Nicolas, E.; Trouche, D. Small Interfering RNAs Targeting a Chromatin-Associated RNA Induce Its Transcriptional Silencing in Human Cells. *Mol. Cell. Biol.* **2022**, *42*, e0027122. [CrossRef]

172. Xiao, M.; Li, J.; Li, W.; Wang, Y.; Wu, F.; Xi, Y.; Zhang, L.; Ding, C.; Luo, H.; Li, Y.; et al. MicroRNAs activate gene transcription epigenetically as an enhancer trigger. *RNA Biol.* **2017**, *14*, 1326–1334. [CrossRef]

173. Liang, Y.; Zou, Q.; Yu, W. Steering Against Wind: A New Network of NamiRNAs and Enhancers. *Genom. Proteom. Bioinform.* **2017**, *15*, 331–337. [CrossRef] [PubMed]

174. Kozomara, A.; Griffiths-Jones, S. miRBase: Annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* **2014**, *42*, D68–D73. [CrossRef] [PubMed]

175. Ohno, S.I.; Oikawa, K.; Tsurui, T.; Harada, Y.; Ono, K.; Tateishi, M.; Mirza, A.; Takanashi, M.; Kanekura, K.; Nagase, K.; et al. Nuclear microRNAs release paused Pol II via the DDX21-CDK9 complex. *Cell Rep.* **2022**, *39*, 110673. [CrossRef]

176. Flynn, R.A.; Do, B.T.; Rubin, A.J.; Calo, E.; Lee, B.; Kuchelmeister, H.; Rale, M.; Chu, C.; Kool, E.T.; Wysocka, J.; et al. 7SK-BAF axis controls pervasive transcription at enhancers. *Nat. Struct. Mol. Biol.* **2016**, *23*, 231–238. [CrossRef]

177. Hainer, S.J.; Gu, W.; Carone, B.R.; Landry, B.D.; Rando, O.J.; Mello, C.C.; Fazzio, T.G. Suppression of pervasive noncoding transcription in embryonic stem cells by esBAF. *Genes Dev.* **2015**, *29*, 362–378. [CrossRef]

178. Ho, L.; Ronan, J.L.; Wu, J.; Staahl, B.T.; Chen, L.; Kuo, A.; Lessard, J.; Nesvizhskii, A.I.; Ranish, J.; Crabtree, G.R. An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 5181–5186. [CrossRef]

179. Gagnon, K.T.; Li, L.; Chu, Y.; Janowski, B.A.; Corey, D.R. RNAi factors are present and active in human cell nuclei. *Cell Rep.* **2014**, *6*, 211–221. [CrossRef]

180. Egloff, S.; Studniarek, C.; Kiss, T. 7SK small nuclear RNA, a multifunctional transcriptional regulatory RNA with gene-specific features. *Transcription* **2018**, *9*, 95–101. [CrossRef]

181. Xing, Y.H.; Yao, R.W.; Zhang, Y.; Guo, C.J.; Jiang, S.; Xu, G.; Dong, R.; Yang, L.; Chen, L.L. SLERT Regulates DDX21 Rings Associated with Pol I Transcription. *Cell* **2017**, *169*, 664–678.e616. [CrossRef]

182. Haberman, N.; Digby, H.; Faraway, R.; Cheung, R.; Chakrabarti, A.M.; Jobbins, A.M.; Parr, C.; Yasuzawa, K.; Kasukawa, T.; Yip, C.W.; et al. Widespread 3′UTR capped RNAs derive from G-rich regions in proximity to AGO2 binding sites. *BMC Biol.* **2024**, *22*, 254. [CrossRef]

183. Aktas, T.; Avsar Ilik, I.; Maticzka, D.; Bhardwaj, V.; Pessoa Rodrigues, C.; Mittler, G.; Manke, T.; Backofen, R.; Akhtar, A. DHX9 suppresses RNA processing defects originating from the Alu invasion of the human genome. *Nature* **2017**, *544*, 115–119. [CrossRef]

184. Chakraborty, P.; Huang, J.T.J.; Hiom, K. DHX9 helicase promotes R-loop formation in cells with impaired RNA splicing. *Nat. Commun.* **2018**, *9*, 4346. [CrossRef] [PubMed]

185. Herbert, A. To "Z" or not to "Z": Z-RNA, self-recognition, and the MDA5 helicase. *PLoS Genet.* **2021**, *17*, e1009513. [CrossRef] [PubMed]

186. Herbert, A.; Pavlov, F.; Konovalov, D.; Poptsova, M. Conserved microRNAs and Flipons Shape Gene Expression during Development by Altering Promoter Conformations. *Int. J. Mol. Sci.* **2023**, *24*, 4884. [CrossRef] [PubMed]

187. Bridge, K.S.; Shah, K.M.; Li, Y.; Foxler, D.E.; Wong, S.C.K.; Miller, D.C.; Davidson, K.M.; Foster, J.G.; Rose, R.; Hodgkinson, M.R.; et al. Argonaute Utilization for miRNA Silencing Is Determined by Phosphorylation-Dependent Recruitment of LIM-Domain-Containing Proteins. *Cell Rep.* **2017**, *20*, 173–187. [CrossRef]

188. Hock, J.; Weinmann, L.; Ender, C.; Rudel, S.; Kremmer, E.; Raabe, M.; Urlaub, H.; Meister, G. Proteomic and functional analysis of Argonaute-containing mRNA-protein complexes in human cells. *EMBO Rep.* **2007**, *8*, 1052–1060. [CrossRef]

189. Griffin, K.N.; Walters, B.W.; Li, H.; Wang, H.; Biancon, G.; Tebaldi, T.; Kaya, C.B.; Kanyo, J.; Lam, T.T.; Cox, A.L.; et al. Widespread association of the Argonaute protein AGO2 with meiotic chromatin suggests a distinct nuclear function in mammalian male reproduction. *Genome Res.* **2022**, *32*, 1655–1668. [CrossRef]

190. Egli, M.; Manoharan, M. Chemistry, structure and function of approved oligonucleotide therapeutics. *Nucleic Acids Res.* **2023**, *51*, 2529–2573. [CrossRef]

191. Duca, M.; Vekhoff, P.; Oussedik, K.; Halby, L.; Arimondo, P.B. The triple helix: 50 years later, the outcome. *Nucleic Acids Res.* **2008**, *36*, 5123–5138. [CrossRef]

192. Zhu, Q.; Schlick, T. A Fiedler Vector Scoring Approach for Novel RNA Motif Selection. *J. Phys. Chem. B* **2021**, *125*, 1144–1155. [CrossRef]

193. Stein, C.A.; Hansen, J.B.; Lai, J.; Wu, S.; Voskresenskiy, A.; Hog, A.; Worm, J.; Hedtjarn, M.; Souleimanian, N.; Miller, P.; et al. Efficient gene silencing by delivery of locked nucleic acid antisense oligonucleotides, unassisted by transfection reagents. *Nucleic Acids Res.* **2010**, *38*, e3. [CrossRef]