Review

RNA in pieces

Alex C. Tuck and David Tollervey

Wellcome Trust Centre for Cell Biology, King's Buildings, Mayfield Road, Edinburgh, EH9 3JR, UK

Eukaryotic genomes accommodate numerous types of information within diverse DNA and RNA sequence elements. At many loci, these elements overlap and the same sequence is read multiple times during the production, processing, localization, function and turnover of a single transcript. Moreover, two or more transcripts from the same locus might use a common sequence in different ways, to perform distinct biological roles. Recent results show that many transcripts also undergo post-transcriptional cleavage to release specific fragments, which can then function independently. This phenomenon appears remarkably widespread, with even well-documented transcript classes such as messenger RNAs yielding fragments. RNA fragmentation significantly expands the already extraordinary spectrum of transcripts present within eukaryotic cells, and also calls into question how the 'gene' should be defined.

The modular gene

Initial analyses of genes envisaged a simple reading whereby different types of genetic information were physically separate: transcription was driven by promoter elements located outside the transcribed region and the transcript would either specify an amino acid sequence or adopt a particular fold as a structural RNA. Transcripts themselves also appeared to be modular, comprising assemblies of distinct sequence elements. As analyses became more sophisticated, they increasingly revealed the use of alternative sites for transcription initiation, termination and splicing [1–3], which are now known to be widespread. The resultant transcripts were, however, seen as related sequence variants that modified the functions of the basic gene. Moreover, the modular notion of genes and transcripts largely assumed that each sequence element had a single function, with diversity arising from its inclusion or exclusion. For example, exons could be present or absent from a transcript, and promoter elements could be bound or unbound.

The genomic palimpsest

It is now apparent that multiple layers of information are superimposed within eukaryotic genetic sequences (Box 1). For example, a protein-coding transcript can concomitantly carry sequence-specific and structural information, governing its folding, protein binding, processing, localization and decay, as well as specifying an amino acid sequence. Similarly, at the DNA level, numerous layers of regulatory information pervade the transcribed region (Box 1), blurring the distinction between regulatory and transcribed sequences and refuting the notion of a modular gene.

Even the boundary between whether information is read from DNA or RNA is becoming blurred. Information initially assumed to be read from DNA is sometimes read at the RNA level, and vice versa: plant small interfering RNAs (siRNAs) recognize DNA targets via binding to specialized long non-coding RNAs (lncRNAs) synthesized by a dedicated RNA polymerase (Pol V) [4] and the nuclear mRNA cap-binding complex can bind to genomic coding regions to promote transcription initiation [5]. Furthermore, RNA signals can be recognized, in part, by proteins bound to chromatin, with recent examples provided by analyses of the effects of histone modifications on patterns of alternative splicing [6].

Thus, genetic information is arranged in an interleaved, overlapping fashion in both DNA and RNA. Rather than being modular, the genome resembles a 'palimpsest', an ancient parchment on which the original text has been overwritten numerous times (discussed further in Box 1). Multiple layers of genetic information can be embedded within a single sequence and, consequently: (i) a single sequence can perform multiple functions; and (ii) genetic information is not restricted in where it resides (transcribed or regulatory regions, DNA or RNA).

The many ways to use a sequence

This overlapping arrangement of genetic information contributes greatly to transcript diversity and complexity. A single locus can produce multiple transcripts that use shared sequences in distinct ways to fulfill a spectrum of fundamentally different biological functions (Figure 1).

The simplest incarnation of this concept is the handful of 'dual-functional' transcripts (Figure 1a). Here, transcripts identical in sequence and length perform alternative functions. For example, some have overlapping reading frames and encode two different proteins [7]. Other examples include the U1 small nuclear RNA (snRNA), a spliceosome component that can also protect pre-mRNAs from premature cleavage [8], and 7SK, a noncoding RNA that regulates multiple transcription factors [9]. In addition, it was known that introns within eukaryotic mRNAs (and mRNA-like species that lack protein-coding capacity) could encode small nucleolar RNAs (snoRNAs) and miRNAs, which are excised by processing [10].

However, recent high-throughput transcriptome sequencing studies reveal that multifunctional sequences occur very frequently. Eukaryotic genomes turn out to be pervasively transcribed, with many loci producing ensembles of interleaved transcripts [11,12]. Diverse functions for these are emerging; for example, lncRNAs, which

Corresponding author: Tollervey, D. (d.tollervey@ed.ac.uk).

 0168-9525/$ – see front matter doi:10.1016/j.tig.2011.06.001

frequently overlap or run antisense to protein-coding genes, can direct nucleosome assembly or provide a scaffold to recruit chromatin-modifying enzymes [13,14]. Even enhancer and promoter regions are transcribed [15], further eroding the dichotomy between regulatory and transcribed regions. Notably, the overlapping arrangement of genetic information at many genomic loci results in cohorts of transcripts that share common sequences but nonetheless show distinct functions (Figure 1b).

The most recent revelation is that direct RNA cleavage further expands the spectrum of functionally distinct transcripts [16–21] (Figure 1c). Stable fragments are derived from well-characterized classes of parent RNAs, notably tRNAs, mRNAs and snoRNAs. A growing body of evidence indicates that fragmentation does not simply reflect RNA degradation, but generates a *bona fide* class of transcripts that perform functions distinct from their parents. The current surge in deep-sequencing studies has been instrumental in the discovery of many RNA fragments and will undoubtedly lead to further progress in this field.

Here, we review the diverse origins and potential functions of these fragments. We conclude by considering how the interleaved arrangement of genetic information and the enormous complexity of the transcriptome impact upon the definition of a gene, and the implications for future research.

Small RNA fragments

Recent genome-wide studies have identified numerous large and small fragments arising from within annotated genes [17,19,22–24] (Figure 2). The small fragments [<30 nucleotides (nt)] have received particular attention, because they resemble miRNAs and, in some cases, bind the Argonaut (Ago) protein family (Box 2) [25]. Small fragments are generated by post-transcriptional cleavage of diverse parent transcripts by endonucleases (outlined in Box 2) and at least some have been shown to perform functions distinct from those of their parents.

snoRNA fragments

Well-characterized fragmentation-derived miRNA-like RNAs are generated from snoRNA-like precursors. The snoRNAs select sites of covalent RNA modification, with box H/ACA snoRNAs directing pseudouridylation and box C/D snoRNAs directing 2′-*O*-methylation. Many eukaryotes possess snoRNA-derived fragments (sdRNAs), with > 60% of snoRNAs represented (Figure 2a) [19,24,26–33]. The sdRNAs range from 15 to 35 nt and map to 5′, 3′ and central regions of snoRNAs. However, the lengths and distribution across the parent snoRNAs vary between species and even among mice subjected to a training regime [30,31]. Numerous factors apparently determine sdRNA abundance, including the poly(A) polymerase Cid14 (a cofactor for the exosome nuclease complex) and the RNA-binding proteins DiGeorge syndrome critical region gene 8 (DGCR8) and Loquacious (Box 2) [31,32].

The relationship between snoRNAs and miRNA-like RNAs has been investigated from both 'ends': identified human miRNAs were found to be derived from snoRNA-like precursors (termed primary miRNAs, 'pri-miRNAs'), whereas human snoRNAs were found to be fragmented into miRNA-like sdRNAs [26,28,29,33,34]. The snoRNA-like pri-miRNAs fall into both the H/ACA and C/D classes and can adopt folds resembling other pri-miRNAs; they also show typical snoRNA properties, including binding of characteristic proteins. Several miRNA-like sdRNAs, derived from both classes, have been shown to bind the silencing machinery and exhibit *trans*-silencing activity on endogenous targets [26,28,34].

However, whereas pri-miRNAs are cleaved by the Microprocessor complex (Drosha and DGCR8; Box 2), only a subset of H/ACA-derived sdRNAs require Microprocessor components for synthesis [26,31] and accumulation of box C/D sdRNA can be independent of DGCR8 and Dicer [31]. Thus, sdRNAs are diverse, but probably include many species that function as miRNAs.

Box 1. The layered arrangement of genetic information

Genetic information is arranged in an interleaved, overlapping fashion in eukaryotes (Figure I), resembling ancient parchments on which the original text was overwritten multiple times ('palimpsests'). DNA packaging into chromatin is partly influenced by local sequence. Sequence-specific binding recruits barrier elements against which nucleosomes are packed, chromatin remodeling factors and other non-histone proteins [85]. Moreover, the intrinsic sequence-dependent curvature of DNA directly influences nucleosome organization. Higher-order chromatin architecture is also sequence dependent. For example, certain DNA elements are tethered to nuclear pores and 'insulator' sequences interact to partition chromatin into domains [86,87]. Notably, some promoters can also act as insulators.

During transcription, information stored within DNA is transferred to RNA. Classical models envisaged that the transcription machinery primarily assembles at promoters. However, genome-wide profiling of transcription factors revealed that many bind transcribed regions [88,89]. Thus, regulatory information is not restricted to distinct sites, but pervades most genomic sequences.

Genetic information in the RNA transcripts is similarly interleaved and overlapping, but with additional capacity provided by complex features of secondary and tertiary structure. The superimposed layers of information within a transcript govern most aspects of its existence, exemplified by mRNAs as shown in Figure I. Primary mRNA transcripts undergo extensive processing, directed by sequence or structural elements. Cleavage, polyadenylation and splicing factors bind specific sequences and the conversion of adenosine to inosine is directed by three-dimensional folds [90,91]. mRNA localization is guided by structural elements in the protein-coding sequence as well as 3′ untranslated regions.

The primary role of the mRNA protein-coding sequence is as a template for translation, but additional layers of information regulate this process: (i) structures around the start codon can impede initiation; (ii) elongation might initially be slowed by codons with rare tRNAs, preventing downstream ribosomal traffic jams; and (iii) secondary structures and biased codon usage coordinate elongation rate with protein folding [92–95].

Transcripts are eventually degraded, with stabilities directed by numerous sequence elements. For example, AU- and GU-rich elements recruit decay factors and many sites are targets for RNA-guided silencing [96]. RNAi-related pathways use small miRNA guides, whereas cleavage in Staufen 1-mediated mRNA decay is programmed by lncRNAs [97,98]. Elements modulating stability reside in protein-coding as well as untranslated regions, for example Nrd1/Nab3-binding motifs, which promote decay [99]. Additionally, functionally distinct mRNA fragments are generated. Thus, many layers of information are superimposed within a transcript.

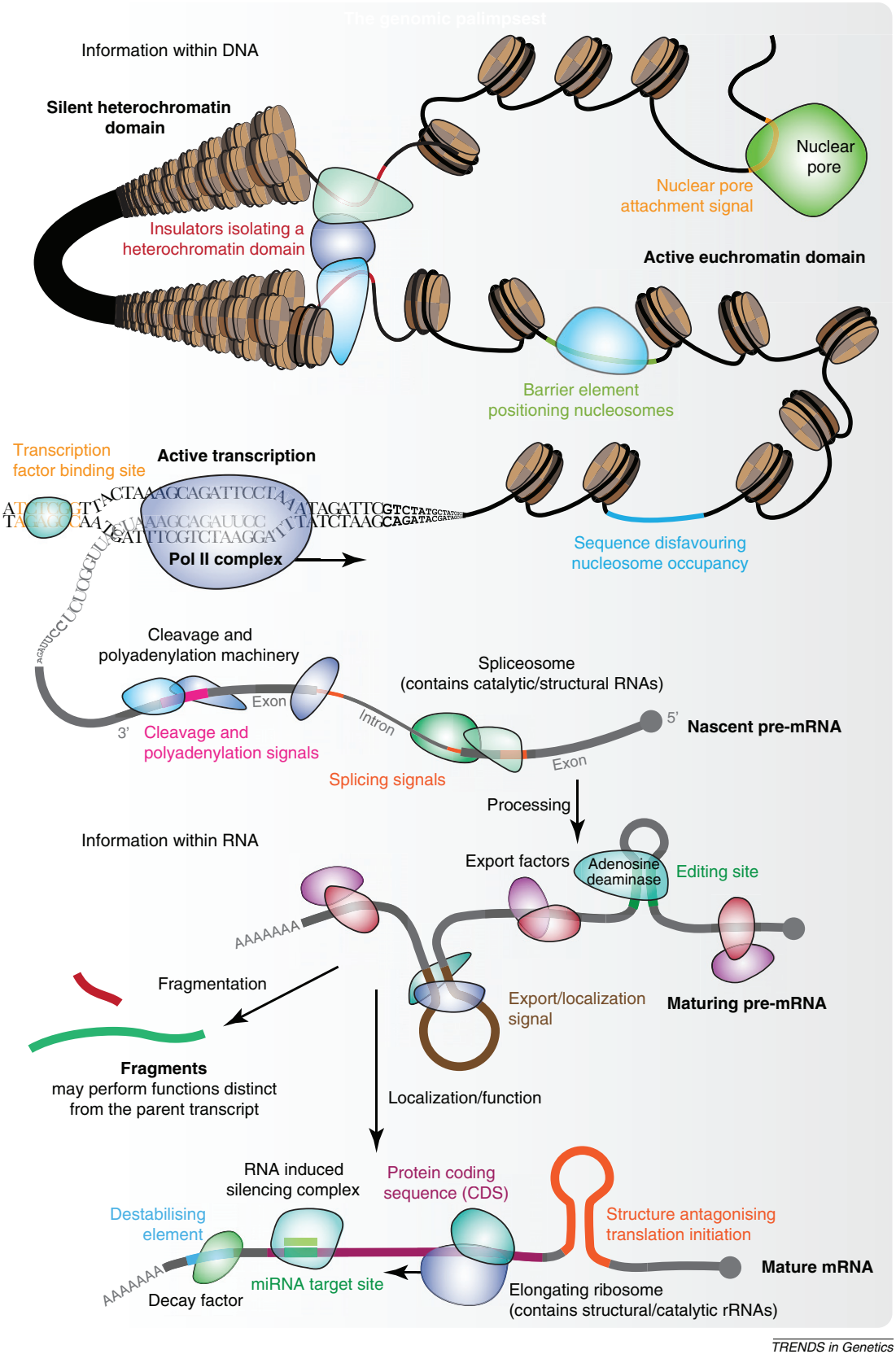

Figure I. DNA and RNA contain multiple overlapping and interleaved layers of information.

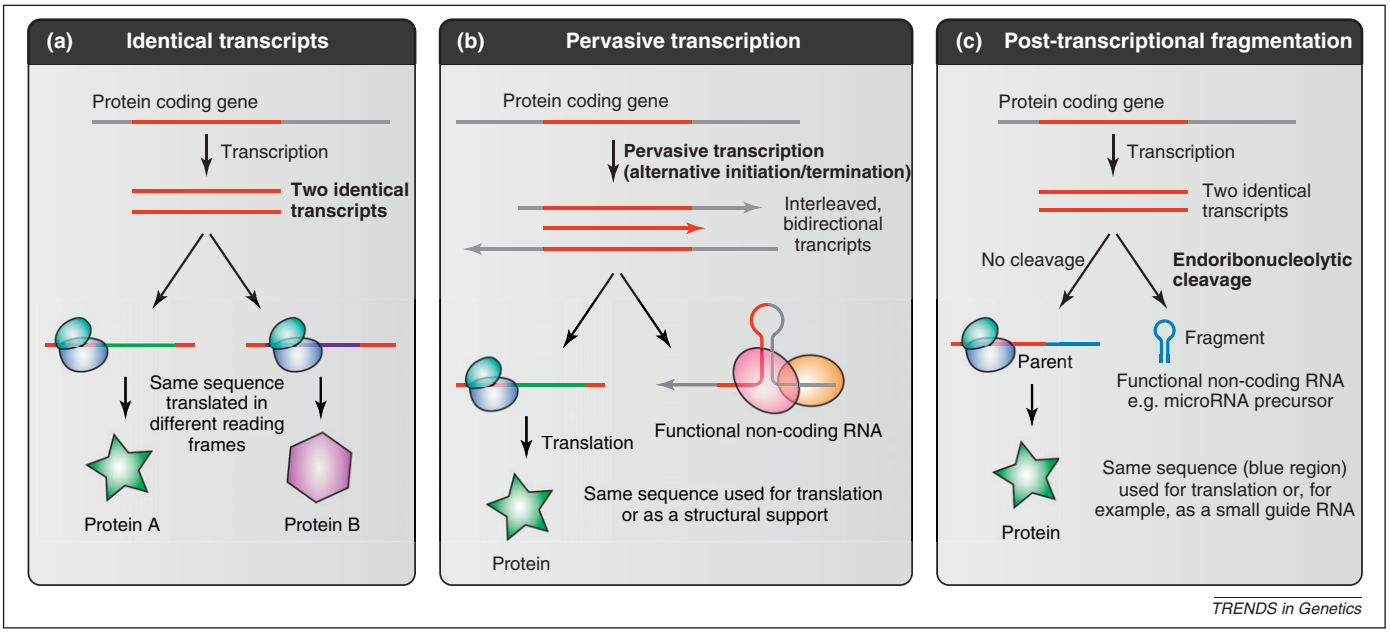

Figure 1. Two transcripts from the same locus can use the same sequence to different functional effects. The overlapping arrangement of genetic information enables a single sequence to encode multiple functions. This principle is embodied at many genomic loci, which generate ensembles of transcripts with shared sequences but disparate functions. This raises questions about how a specific function is assigned to a transcript, given the numerous possibilities. There are several explanations, illustrated by the various ways in which overlapping transcripts are generated: **(a)** Two transcripts identical in sequence and length might function differently, perhaps being translated in alternative reading frames (green or purple) to generate distinct proteins. Here, extrinsic factors are responsible for specifying which reading frame should be used. **(b)** Alternative transcription initiation and/or termination generate an ensemble of interleaved transcripts from a single genomic locus. Within this ensemble, a shared sequence (red) can perform distinct functions, perhaps contributing to an open reading frame (green) in one transcript and a structural feature in another. Here, the function of a sequence is governed by its context, with the different lengths and orientations of transcripts perhaps affecting their folding or recruitment of binding factors. **(c)** Many classes of transcript might act as precursors to shorter fragments, excised by post-transcriptional cleavage. These fragments might function in ways distinct from those of their parents. Thus, within the context of the shorter fragment, a shared sequence (blue) can perform an alternative role. This indicates that the length of a transcript might contribute to specifying which of several possible functions is performed by a particular sequence. Other post-transcriptional processes (such as splicing) can also generate alternative transcripts, but are beyond the scope of this review.

tRNA fragments

High-throughput sequencing detected small fragments of human and mouse tRNAs, some of which are highly abundant [16–20,24,35–38]. The tRNA regulatory fragments (tRFs) are classified as tRF-1, tRF-3 or tRF-5, depending on their origin [20] (Figure 2b). These probably arise from specific processing, because: (i) their abundance varies between cell lines and is distinct from parent tRNAs; (ii) tRF ends are precisely defined; (iii) there are sequence preferences for cleavage; and (iv) some specific functions have been uncovered [19,20]. Although we discuss the tRF classes separately, their biogenesis and probable functions overlap considerably.

tRF-3. tRF-3 species are 13–22 nt fragments derived from the 3′ end of mature tRNAs cleaved within the T-loop [16,19,20,24,35]. Examples studied in detail resemble miRNAs in exhibiting Dicer-dependent processing, association with Ago1–4 and *trans*-silencing activity (Box 2) [18,19,35]. Dicer binds dsRNA with a 2-nt 3′ overhang, cutting at a specific distance from the end of the duplex. One mouse tRF-3-like fragment arises when an isoleucine tRNA forms a long hairpin rather than the standard cloverleaf [16]. Other tRNAs can form duplexes with complementary RNAs. For example, a tRF-3 fragment arises by cleavage of $tRNA^{Lys}$ bound to the HIV-1 primer binding site. This tRF-3 species reduces HIV-1 replication in infected cells and might reflect a more general retroviral defence mechanism [19].

However, although tRF-3 RNAs resemble miRNAs, there are clear distinctions: Dicer-null mouse cells retain many tRFs [16] and tRF-3 RNAs do not associate with some miRNA-binding factors, such as Mov10 [18].

tRF-5. tRF-5 fragments arise from mature tRNA 5′ ends via D-loop cleavage. Similar to tRF-3 RNAs, some are Dicer dependent, cytoplasmic and able to associate with Ago proteins, albeit weakly [17,24]. The detection of long (31 nt) tRF-5 fragments suggests that compact tRNA folding allows Dicer to generate unusually long products [24]. However, *trans*-silencing activity has not yet been reported. One possibility is that tRF-5 and tRF-3 species resemble ‘mature’ and ‘star’ strands of precursor miRNAs (pre-miRNAs), respectively [17]. Selective stabilization of the ‘mature’ strand could explain why tRF-5 and tRF-3 fragments are not always either detected or functional [19]. Human miRNAs lack terminal modifications, whereas characterized fragments of $tRNA^{Gln}$ are 3′ modified, potentially explaining the weak association of these and other tRFs with Ago proteins [17].

tRF-1. tRF-1 fragments correspond to precursor (pre-)tRNA 3′ trailers, with 5′ ends generated by the endonuclease tRNase Z and 3′ ends matching Pol III termination sites (Box 2) [20]. tRF-1 species are cytoplasmic, whereas pre-tRNAs are 3′ matured in the nucleus, suggesting that tRF-1 RNAs are rapidly exported [39]. Fragments from a particular tRNA vary in length and precise 5′ end, so factors other than tRNase Z might contribute to tRF-1 production [18]. Indeed, some mouse small RNAs arise from 3′ trailers resembling pre-miRNAs, potentially processed by Dicer [40]. Additionally, Ago proteins might participate in processing, as their overexpression specifically enriches shorter

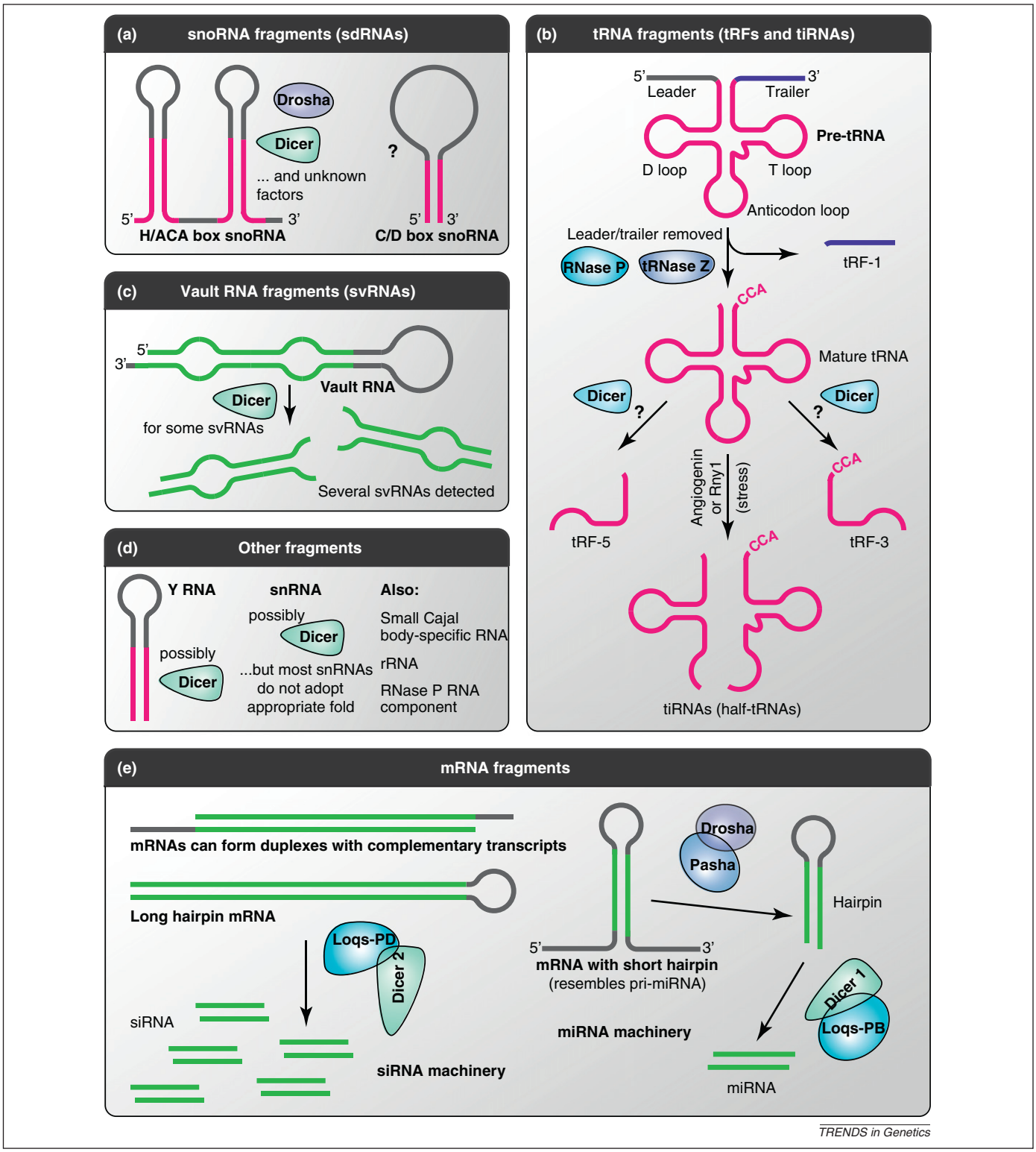

Figure 2. Products generated by post-transcriptional cleavage of diverse transcript classes. Recent studies have found that many well-documented classes of transcript can be post-transcriptionally cleaved to liberate shorter fragments, a phenomenon known as RNA fragmentation. To date, the participating endonucleases are thought to include the tRNA processing enzymes RNase P and tRNase Z, and the RNase III family members Drosha and Dicer (with their cofactors Pasha and Loqs), which generate small *trans*-silencing RNAs. Fragments from the following transcript classes have been documented: **(a)** small nucleolar RNAs (snoRNAs) target modifying enzymes to specific sites on transcripts, and fall into two classes (box H/ACA and box C/D). Fragmentation of some box H/ACA snoRNAs is catalyzed by Drosha and Dicer, but endonucleases excising fragments from other box C/D snoRNAs and other box H/ACA snoRNAs are currently unknown [31]. **(b)** Single-stranded, short regulatory tRNA fragments (tRFs) are generated by precursor (pre)-tRNA processing, which releases tRF-1 fragments (trailers), and mature tRNA cleavage, which releases tRF-3/5 fragments. Additionally, stress-induced cleavage by angiogenin or Rny1 releases longer tRNA halves (tiRNAs) [18,37,67]. **(c)** Vault RNAs (vRNAs) are a component of vault particles, which are ribonucleoprotein complexes linked to drug resistance. Several short fragments of vRNAs (svRNAs) have been detected, some of which are dependent on Dicer [44]. **(d)** Short fragments of other stable non-protein-coding RNAs have also been detected and, again, some are dependent on Dicer [19,24,27,30]. **(e)** mRNAs can adopt structures with long or short duplexes that are processed to small RNAs by the canonical small interfering RNA (siRNA) or miRNA biogenesis machinery. Abbreviations: pri-miRNA, primary miRNA; sdRNA, snoRNA-derived fragment; snRNA, small nuclear RNA.

Box 2. Endoribonucleases involved in RNA fragmentation

Analyses of eukaryotic RNA surveillance and turnover have largely focused on the roles of exonucleases. However, endonucleases also play important roles in RNA metabolism [100,101]. This is exemplified by RNA fragmentation, which involves the endonucleolytic cleavage of a parent transcript to release specific fragments, potentially with distinct functions. To date, most endonucleases implicated in RNA fragmentation also function in well-characterized RNA processing pathways.

RNAi and related pathways

miRNA biogenesis involves cropping of a pri-miRNA by Drosha and Pasha/DGCR8 to liberate a pre-miRNA, then excision of the mature miRNA by Dicer 1 and Loqs-PB (Figure Ia) [98]. Drosha and Dicer are RNase III endonucleases that make staggered cuts within a duplex. Any transcript could potentially provide a non-canonical substrate for cleavage by adopting a fold resembling pri-miRNAs or pre-miRNAs. This probably underlies the fragmentation of some tRNAs, snoRNAs, mRNAs and other structured RNAs.

siRNA biogenesis involves the processing of long duplexes by a complex between Dicer 2 and Loqs-PD (Figure Ib) [98]. Originally thought to derive exclusively from exogenous sources, siRNAs from endogenous intermolecular and intramolecular mRNA duplexes are now documented, revealing another source of RNA fragments.

The effectors of RNAi-related pathways are the Ago proteins, of which humans possess four. Each contains a PIWI domain (resembling an RNase H fold), but only Ago2 retains endonuclease activity. Ago2 cleavage can generate RNA fragments, which might function via Ago-binding to direct further cleavages, modify mRNA translation and stability or perturb the Ago-association of canonical miRNAs or siRNAs by competition.

pre-tRNA processing

Pre-tRNAs are processed by RNase P and tRNase Z to remove the 5′ leader and 3′ trailer, then a 3′ CCA triplet is added (Figure Ic) [68]. The liberated trailers constitute an additional class of tRNA-derived fragments. Both RNase P and a cytoplasmic isoform of tRNase Z can be directed by small guide RNAs to cleave specific targets, so might represent effectors programmed by RNA fragments.

Emerging endonucleases

Other endonucleases also participate in RNA fragmentation. For example, angiogenin and Rny1 cleave tRNAs during stress. Additional candidates include: (i) the PIN-domain proteins Rrp44, Swt1 and Smg6, for which roles in RNA surveillance are emerging; (ii) RNase MRP, which processes pre-rRNA and at least one cell cycle-regulated mRNA; (iii) Rnt1, the sole RNase III enzyme in *Saccharomyces cerevisiae*, which 3′ processes pre-rRNA, pre-snRNAs and pre-snoRNAs, but also cleaves pre-mRNAs; (iv) G3BP, which cleaves specific human mRNAs; (v) Ire1, involved in the unfolded protein response and decay of endoplasmic reticulum-targeted mRNAs; and (vi) RNase H, which cleaves nascent transcripts at sites of R-loop formation [100,101].

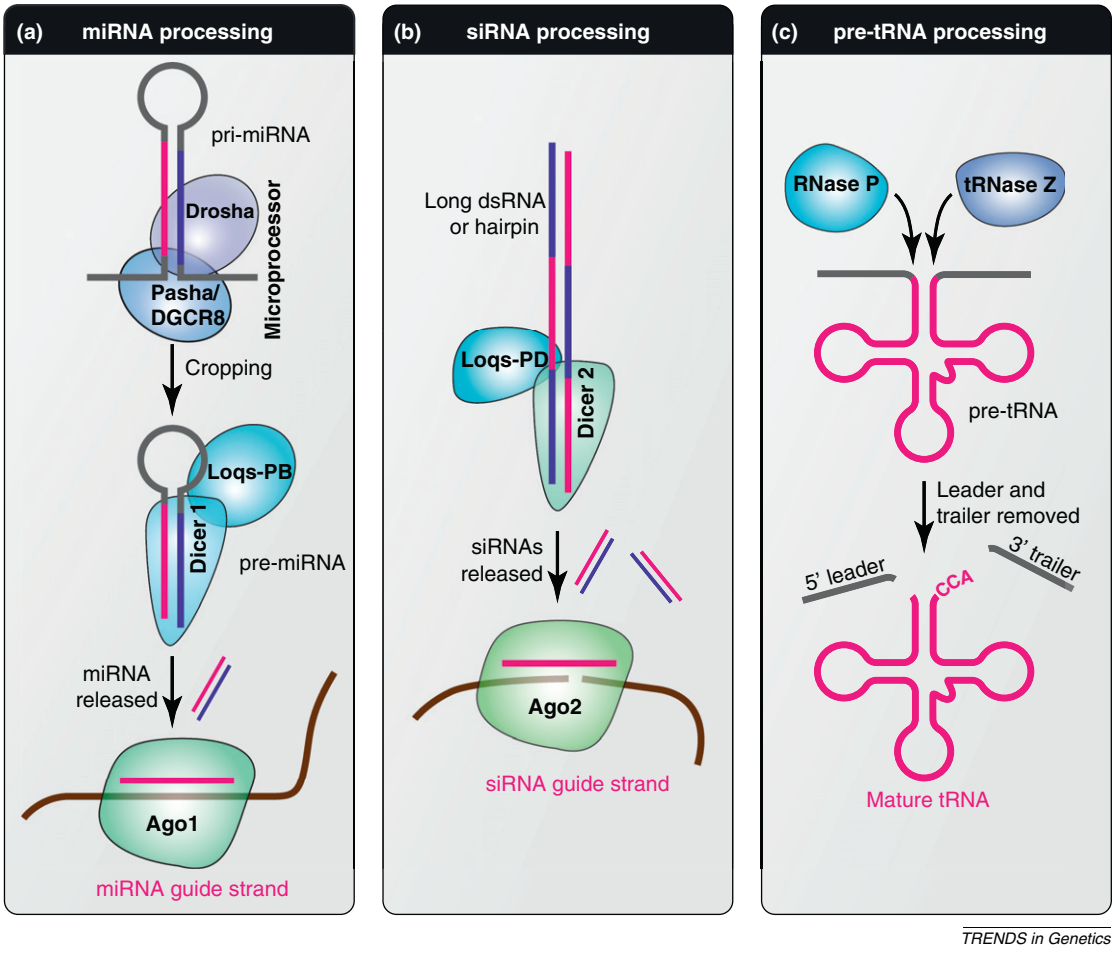

Figure I. Classical RNA processing pathways contribute to RNA fragmentation. Abbreviations: DGCR8, DiGeorge syndrome critical region gene 8; dsRNA, double-stranded RNA; pre-miRNA, precursor miRNAs; pri-miRNA, primary miRNA; pre-tRNA, precursor tRNA.

(20–21 nt) tRF-1 species [18]. This would resemble Ago2-mediated processing of the miR-451 precursor [41].

Tested tRF-1 RNAs show little *trans*-silencing activity [18,20], perhaps because they associate with Ago3/4, rather than with the slicing-competent Ago2, which preferentially binds duplex RNAs [18]. However, when duplexed with an antisense oligoribonucleotide, a tRF-1 RNA (cand45) bound Ago2 and elicited *trans*-silencing [18].

Rather than directing Ago proteins to specific targets, tRF-1 fragments might normally compete with miRNAs for Ago binding and/or alter miRNA distribution among Ago proteins, thus indirectly perturbing silencing activity [18]. Consistent with this, upregulating tRF-1 expression reduced the efficacy of miRNAs [18]. Conversely, loss of DGCR8 and Dicer increased tRF abundance, suggesting that miRNAs also antagonize tRFs [16]. In proliferative diseases such as cancer, tRNAs are overexpressed and miRNA profiles perturbed, whereas increased Pol III transcription can promote transformation [42,43]. Specifically, a tRF-1 RNA (tRF-1001) augments proliferation of a colon cancer cell line [20], and we speculate that this might be a more general property of tRFs.

Fragments of other structural RNAs

Fragments have also been detected for several other classes of RNA, including rRNA, snRNA, the RNA component of RNase P and the small, cytoplasmic vault and Y RNAs (Figures 2c and 2d) [19,24,27,30,38,44]. Some can be generated by Dicer or Drosha cleavage: Y RNAs adopt pre-miRNA-like folds; some regions of rRNAs form stable duplexes; and processing of human vault RNA 1 is Dicer dependent [19,27,30,44]. However, alternative pathways must exist as snRNAs that generate small RNAs generally lack a clear propensity to form Dicer-compatible hairpins [24].

Expression of these small RNAs is both regulated and distinct from the parent RNAs, supporting functional roles [30,44]. Small RNA fragments from snRNAs, vault RNAs (vRNAs) and rRNAs associate with Ago proteins and might therefore participate in RNA silencing [24]. Targets identified for one vRNA fragment include the mRNA encoding the drug metabolizing enzyme CYP3A4, perhaps explaining the chemotherapy resistance associated with vault particles [44]. However, no silencing activity was observed for other vRNA or Y RNA fragments [27,44].

mRNA fragments

Small mRNA fragments have been reported in diverse eukaryotes (Figure 2e) [16,19,24,45–53]. Some mRNAs with long duplexes or hairpins are processed by Dicer to generate endogenous siRNAs, often from repeat regions. For example, the *Drosophila CG4068* 3′ UTR produces siRNAs that silence the *mus308* gene [50]. Similarly, overlapping, oppositely oriented transcripts can generate '*cis*-natural antisense transcript siRNAs' (*cis*-NAT-siRNAs) from duplexed, complementary regions. This was discovered in *Arabidopsis*, where stress-induced antisense transcripts recruit Ago to their partner [54]. *Drosophila* produces *cis*-NAT-siRNAs from approximately 25% of convergent transcripts and, in mammals, they arise from convergent or divergent pairs [19,49,51,55]. In *Schizosaccharomyces pombe*, cell cycle-dependent read-though of transcription terminators on convergent genes leads to overlapping transcripts. These provoke siRNA-dependent heterochromatin formation, providing an autoregulatory system for siRNA components [56]. Bidirectional promoters, which are very common and produce divergent transcripts with a short overlap, are another source of double-stranded RNA [19,57], as are complementary transcripts from distant loci (such as gene–pseudogene pairs) [51]. Still more generally, pervasive transcription might provide antisense partners for the abundant mRNAs that generate siRNAs in the absence of hairpins or annotated antisense partners [19,49,58].

Endogenous miRNAs are processed from short hairpin structures, which are frequently located within mRNA introns. Some introns harbor full pri-miRNAs processed by Drosha and Dicer [59], whereas short 'mirtrons' produce a pre-miRNA directly by splicing and debranching (sometimes aided by exosome-mediated trimming), bypassing the need for Drosha [16,52,53,60,61]. Small Ago-associated RNAs are also excised from an exonic, pre-miRNA-like fold within human CYP46A1 [24] and Drosha apparently generates small RNAs from pre-miRNA-like folds in mouse mRNAs [45].

Together, these studies identify small RNA fragments that arise from specific post-transcriptional cleavage and interact with Ago proteins to function as mRNA silencers, perturb miRNA profiles, or stimulate proliferation. However, many small fragments are generated independently of Drosha and Dicer, do not associate with Ago and lack apparent silencing ability [16,62]. These might possess as yet untested functions.

Longer RNA fragments

In addition to the short fragments discussed above, longer (>30 nt) fragments are also generated from diverse transcripts [19,21,31,38,63–66]. These too are excised by post-transcriptional cleavage and can perform functions distinct from those of their parents.

Stress-induced fragments

The best-studied longer fragments are stress-induced, tRNA-derived RNAs (tiRNAs), found in diverse eukaryotes (Figure 2b). These are generated by endonucleolytic cleavage in the anticodon loop of mature tRNAs, catalyzed by Rny1 in yeast and angiogenin in mammals [19,64,65,67–69]. Under normal conditions, angiogenin is inhibited by binding to RNH1, whereas Rny1 is sequestered in the vacuole away from its substrates [65,67]. These enzymes have limited specificity, because cleavage occurs at various positions around the anticodon loop. Most tRNAs are susceptible to cleavage [21] although not all are equally cleaved [64]. Rny1 is an RNase T2 family member, and is predicted to leave 5′ hydroxyl and 2′-3′ cyclic phosphate groups. This is unusual for intracellular cleavage and significant because the 5′ hydroxyl group confers resistance to degradation by the 5′ exonucleases Xrn1 and Rat1/Xrn2. Moreover, the tRNA splicing ligase requires 5′ hydroxyl and 2′-3′ cyclic phosphate groups, suggesting that tRNA anticodon cleavage might be reversible under some circumstances [70].

Cleavage of tRNAs might help adaptation to stress, but this probably does not simply reflect reduced tRNA

availability as only a minority are cleaved [21]. Indeed, too much cleavage appears deleterious: *Dnmt2* mutants fail to methylate tRNAs with C38 and show increased tRNA cleavage but reduced stress tolerance [71]. Treating mammalian cells with angiogenin or synthetic tiRNAs inhibits translation [65] and promotes formation of stress granules (SGs); sites where translationally silenced mRNAs are sorted for re-initiation, decay or storage [72,73]. Denatured tiRNAs cannot inhibit translation, suggesting that structural features are important [66].

tiRNAs might also warn neighboring cells of imminent stress, giving them time to prepare. Angiogenin (a secreted enzyme) and tiRNAs (found in phloem sap) could both act as messengers.

Functional long fragments

Additional parent transcript classes produce long fragments. Deep-sequencing detected 30–40 nt fragments derived from human snoRNAs, snRNAs, rRNAs and mRNAs, whereas CAGE tags (cap analysis of gene expression; short sequences representing the 5′ ends of cDNAs) identified many transcripts of >200 nt [1,19,22,23,38,62]. Consistent with post-transcriptional cleavage, many CAGE tags do not coincide with hallmarks of transcription initiation and are derived from both protein-coding regions and 3′ untranslated regions (3′ UTRs). Some of these transcripts contain 5′ exons too short to have undergone splicing, confirming they were cleaved from already spliced transcripts and then capped [22,23,38,62]. Some mRNAs are cleaved by Ago2, targeted by miRNAs, or directly by Drosha [38,45,74,75]. However, the remainder must be generated by other endonucleases (Box 2).

Accumulating evidence suggests these mRNA fragments are functional. Some are capped, perhaps by the cytoplasmic capping complex, suggesting that they are stable, and many cleavages are conserved between humans and mice [23,62,76]. Inspection of CAGE tags and *in situ* hybridization reveals that expression of exonic fragments is tissue and developmental stage specific, and different from the parent mRNA, suggesting that they function in distinct ways [23,62]. Although some fragments potentially encode truncated proteins, many lack coding capacity and presumably function as non-coding RNAs. A few possibilities are illustrated by studies on the functions of 3′ UTRs, which by definition are not protein-coding.

Signals present in 3′ UTRs frequently control mRNA localization and stability, but separate functions have been found for several 3′ UTRs. The 3′ UTR of the *oskar* mRNA is expressed independently and is sufficient to restore Staufen accumulation to *oskar* mutant oocytes, perhaps providing a scaffold to localize regulatory proteins [62]. Other 3′ UTRs are tumor suppressors, whereas the 3′ UTRs of several muscle structural genes enhance myogenic gene expression [62]. In general, 3′ UTRs harbor protein-binding motifs and could therefore potentially sequester regulatory factors away from other targets.

Silencing by long fragments

Long fragments could target sequence-specific cleavage by RNase P or tRNase ZL, both of which can be programmed with guides resembling tRNA fragments [77,78]. However, such guides need not be derived from tRNAs, because tRNase ZL also binds rRNA and snRNA fragments. This mechanism potentially regulates many targets: tRNase ZL overexpression results in downregulation of 41 mRNAs, some of which were validated as targets for tRNase ZL primed with half-tRNAGlu or an rRNA fragment [77].

Despite being longer than canonical miRNAs/siRNAs, 30–40 nt fragments might still bind Ago to elicit *trans*-silencing. Ago2 binds pre-miRNAs or long RNAs, and the exosome and/or Ago might trim extended siRNAs [41,79,80]. Additionally, long initial fragments appear to be processed to smaller fragments [22,23,38]. Indeed, the discovery that even miRNAs generate smaller fragments hints at the existence of complex fragmentation hierarchies with many levels [81]. Many exon-derived small RNAs are generated independently of siRNA and/or miRNA processing factors, and sequential fragmentation of mRNAs to long then short species might provide a distinct pathway for small RNA generation [23]. An additional link between long fragments and RNA silencing is the observation that, in cells subjected to stress, Argonaute proteins accumulate in SGs, the structures induced by tiRNAs [82].

Overall, long fragments appear to be excised by specific, post-transcriptional cleavage of diverse transcripts. Accumulating evidence suggests that some function as scaffolds, translational inhibitors, tumor suppressors, transcriptional activators or RNA silencers, whereas others might be precursors to smaller fragments.

Functional fragments or pointless pieces?

We conclude that post-transcriptional fragmentation is widespread throughout eukaryotic transcriptomes and can generate fragments that function independently of the parent. However, a key question is how many of these fragments are functionally important? Are most just spurious degradation intermediates or evolutionary leftovers, maintained simply because counter selection is not strong enough? This resembles the debate over lncRNAs, with initial estimates of how many are functionally important ranging from almost all to almost none; and the truth appearing to be somewhere in between.

Many RNA fragments are generated by precise cleavage, conserved from mammals to protozoa, such as *Tetrahymena* [83]. They are expressed in a tissue- and condition-specific manner, their abundance is uncoupled from that of their parents and, within some transcript classes (e.g. snoRNAs and tRNAs), fragmentation appears to be a near ubiquitous phenomenon [21,31]. However, although these observations are strongly indicative of function, biological roles have only been directly demonstrated for a handful of examples. A global assessment of functionality is hampered by the fact that fragments might play widely disparate roles. Furthermore, many of these roles might only be apparent under certain conditions. This is analogous to the difficulties faced when investigating yeast genes, 80% of which are essential for viability only when tested in combination (synthetic lethal interactions).

Currently, the broadest functional studies of RNA fragments are those estimating the extent to which they participate in RNAi-related *trans*-silencing. However, within sequenced pools of Argonaute-bound small RNAs, reported

abundances of tRNA fragments range from approximately 0.01 to 10% of hits, with snoRNA fragments comprising approximately 0.2–1% [24,26,32,84]. The documented abundance of RNA fragments in whole-cell extracts similarly varies widely, from approximately 0.1 to 10% of total reads for snoRNA and tRNA fragments alike [19,24,31]. Furthermore, some fragments (such as snoRNA-derived RNAs) have known Ago-dependent functionality but are not always enriched in Ago immunoprecipitates [24]. Another complication is that, in *S. pombe* strains lacking the activities of the TRAMP and exosome RNA surveillance complexes, diverse RNA degradation products accumulate and bind Ago1 [32]. Thus, Ago association does not necessarily demonstrate biological function. A precise assessment of the participation of RNA fragments in *trans*-silencing therefore awaits the high-throughput identification of guide–target pairs, and a comprehensive list of RNA fragments. However, even if fragments do not have direct targets, the very fact that they are occupying RNA binding sites on proteins will perturb the binding of other transcripts, with potentially widespread consequences [18].

The multifunctional gene

The unexpected discovery of fragments derived from the best-characterized classes of transcripts illustrates the powerful and far-reaching consequences of the interleaved and overlapping arrangement of genetic information. Sequences with a single function might, in fact, be the exception rather than the rule, and the ability of a single sequence to encode multiple layers of information permits an almost unimaginable overall level of regulatory complexity and transcriptome diversity. These findings pose many pressing questions (Box 3). Notably, current understanding of RNA fragmentation is largely based on just a handful of high-throughput sequencing studies. Significant advances might therefore be made simply by reanalyzing the many additional data sets already in existence.

The gene as a system

Our initial 'modular' notion of a gene has been challenged by the realization that: (i) multiple layers of regulatory information permeate the transcribed region; (ii) eukaryotic genomes are pervasively transcribed, generating an ensemble of transcripts from any given locus; (iii) each of these transcripts might in turn undergo multiple rounds of cleavage to generate even greater complexity; and (iv) this panoply of transcripts can perform diverse biological roles.

Box 3. Outstanding questions

- What proportion of RNA fragments is functional, and what functions do they perform?
- Which endoribonucleases excise the many fragments for which the cleavage activity is currently unknown?
- How does the cell distinguish functional fragments from unwanted degradation products?
- For any given transcript, what determines the functions that an RNA region performs, out of the multiple possibilities?
- Has the limit of transcriptome complexity now been reached? Or do additional transcript classes await discovery and what might their nature be?

The overlapping nature of the genetic information and transcripts associated with a single locus limits the value of studies of any component in isolation. We therefore suggest that each gene must now be regarded as a system, comprising a genomic region with the corresponding network of control regions and ensemble of transcripts.

Acknowledgments

We thank Grzegorz Kudla and Agata Swiatkowska for critical reading of the manuscript. This work was funded by the Wellcome Trust.

References

1 Ni, T. *et al.* (2010) A paired-end sequencing strategy to map the complex landscape of transcription initiation. *Nat. Meth.* 7, 521–527
2 Ozsolak, F. *et al.* (2010) Comprehensive polyadenylation site maps in yeast and human reveal pervasive alternative polyadenylation. *Cell* 143, 1018–1029
3 Wang, E.T. *et al.* (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature* 456, 470–476
4 Wierzbicki, A.T. *et al.* (2008) Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. *Cell* 135, 635–648
5 Lahudkar, S. *et al.* (2011) The mRNA cap-binding complex stimulates the formation of pre-initiation complex at the promoter via its interaction with Mot1p *in vivo*. *Nucleic Acids Res.* 39, 2188–2209
6 Luco, R.F. *et al.* (2011) Epigenetics in alternative pre-mRNA splicing. *Cell* 144, 16–26
7 Xu, H. *et al.* (2010) Length of the ORF, position of the first AUG and the Kozak motif are important factors in potential dual-coding transcripts. *Cell Res.* 20, 445–457
8 Kaida, D. *et al.* (2010) U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature* 468, 664–668
9 Eilebrecht, S. *et al.* (2011) 7SK small nuclear RNA directly affects HMGA1 function in transcription regulation. *Nucleic Acids Res.* 39, 2057–2072
10 Brown, J.W.S. *et al.* (2008) Intronic noncoding RNAs and splicing. *Trends Plant Sci.* 13, 335–342
11 Costa, F.F. (2010) Non-coding RNAs: meet thy masters. *Bioessays* 32, 599–608
12 Chen, L-L. and Carmichael, G.G. (2010) Decoding the function of nuclear long non-coding RNAs. *Curr. Opin. Cell Biol.* 22, 357–364
13 Thebault, P. *et al.* (2011) Transcription regulation by the non-coding RNA SRG1 requires Spt2-dependent chromatin deposition in the wake of RNAP II. *Mol. Cell. Biol.* 31, 1288–1300
14 Hainer, S.J. *et al.* (2011) Intergenic transcription causes repression by directing nucleosome assembly. *Genes Dev.* 25, 29–40
15 Kim, T-K. *et al.* (2010) Widespread transcription at neuronal activity-regulated enhancers. *Nature* 465, 182–187
16 Babiarz, J.E. *et al.* (2008) Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev.* 22, 2773–2785
17 Cole, C. *et al.* (2009) Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. *RNA* 15, 2147–2160
18 Haussecker, D. *et al.* (2010) Human tRNA-derived small RNAs in the global regulation of RNA silencing. *RNA* 16, 673–695
19 Kawaji, H. *et al.* (2008) Hidden layers of human small RNAs. *BMC Genomics* 9, 157
20 Lee, Y.S. *et al.* (2009) A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). *Genes Dev.* 23, 2639–2649
21 Thompson, D.M. *et al.* (2008) tRNA cleavage is a conserved response to oxidative stress in eukaryotes. *RNA* 14, 2095–2103
22 Fejes-Toth, K. *et al.* (2009) Post-transcriptional processing generates a diversity of 5′-modified long and short RNAs. *Nature* 457, 1028–1032
23 Mercer, T.R. *et al.* (2010) Regulated post-transcriptional RNA cleavage diversifies the eukaryotic transcriptome. *Genome Res.* 20, 1639–1650
24 Burroughs, A.M. *et al.* (2011) Deep-sequencing of human Argonaute-associated small RNAs provides insight into miRNA sorting and reveals Argonaute association with RNA fragments of diverse origin. *RNA Biol.* 8, 158–177

25 Miyoshi, K. *et al.* (2010) Many ways to generate microRNA-like small RNAs: non-canonical pathways for microRNA production. *Mol. Genet. Genomics* 284, 95–103
26 Ender, C. *et al.* (2008) A human snoRNA with microRNA-like functions. *Mol. Cell* 32, 519–528
27 Meiri, E. *et al.* (2010) Discovery of microRNAs and other small RNAs in solid tumors. *Nucleic Acids Res.* 38, 6234–6246
28 Saraiya, A.A. and Wang, C.C. (2008) snoRNA, a novel precursor of microRNA in *Giardia lamblia*. *PLoS Pathog.* 4, e1000224
29 Scott, M.S. *et al.* (2009) Human miRNA precursors with box H/ACA snoRNA features. *PLoS Comput. Biol.* 5, e1000507
30 Smalheiser, N.R. *et al.* (2011) Endogenous siRNAs and noncoding RNA-derived small RNAs are expressed in adult mouse hippocampus and are up-regulated in olfactory discrimination training. *RNA* 17, 166–181
31 Taft, R.J. *et al.* (2009) Small RNAs derived from snoRNAs. *RNA* 15, 1233–1240
32 Buhler, M. *et al.* (2008) TRAMP-mediated RNA surveillance prevents spurious entry of RNAs into the *Schizosaccharomyces pombe* siRNA pathway. *Nat. Struct. Mol. Biol.* 15, 1015–1023
33 Ono, M. *et al.* (2011) Identification of human miRNA precursors that resemble box C/D snoRNAs. *Nucleic Acids Res.* 39, 3879–3891
34 Brameier, M. *et al.* (2011) Human box C/D snoRNAs with miRNA like functions: expanding the range of regulatory RNAs. *Nucleic Acids Res.* 39, 675–686
35 Yeung, M.L. *et al.* (2009) Pyrosequencing of small non-coding RNAs in HIV-1 infected cells: evidence for the processing of a viral-cellular double-stranded RNA hybrid. *Nucleic Acids Res.* 37, 6575–6586
36 Calabrese, J.M. *et al.* (2007) RNA sequence analysis defines Dicer's role in mouse embryonic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* 104, 18097–18102
37 Pederson, T. (2010) Regulatory RNAs derived from transfer RNA? *RNA* 16, 1865–1869
38 Bracken, C.P. *et al.* (2011) Global analysis of the mammalian RNA degradome reveals widespread miRNA-dependent and miRNA-independent endonucleolytic cleavage. *Nucleic Acids Res.* DOI: 10.1093/nar/gkr110
39 Liao, J-Y. *et al.* (2010) Deep sequencing of human nuclear and cytoplasmic small RNAs reveals an unexpectedly complex subcellular distribution of miRNAs and tRNA 3′ Trailers. *PLoS ONE* 5, e10563
40 Reese, T.A. *et al.* (2010) Identification of novel microRNA-like molecules generated from herpesvirus and host tRNA transcripts. *J. Virol.* 84, 10344–10353
41 Cifuentes, D. *et al.* (2010) A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science* 328, 1694–1698
42 Lu, J. *et al.* (2005) MicroRNA expression profiles classify human cancers. *Nature* 435, 834–838
43 Marshall, L. and White, R.J. (2008) Non-coding RNA production by RNA polymerase III is implicated in cancer. *Nat. Rev. Cancer* 8, 911–914
44 Persson, H. *et al.* (2009) The non-coding RNA of the multidrug resistance-linked vault particle encodes multiple regulatory small RNAs. *Nat. Cell Biol.* 11, 1268–1271
45 Chong, M.M.W. *et al.* (2010) Canonical and alternate functions of the microRNA biogenesis machinery. *Genes Dev.* 24, 1951–1960
46 Ghildiyal, M. *et al.* (2008) Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells. *Science* 320, 1077–1081
47 Jin, H. *et al.* (2008) Small RNAs and the regulation of *cis*-natural antisense transcripts in *Arabidopsis*. *BMC Mol. Biol.* 9, 6
48 Kawamura, Y. *et al.* (2008) *Drosophila* endogenous small RNAs bind to Argonaute-2 in somatic cells. *Nature* 453, 793–797
49 Okamura, K. *et al.* (2008) Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in *Drosophila melanogaster*. *Nat. Struct. Mol. Biol.* 15, 581–590
50 Okamura, K. *et al.* (2008) The *Drosophila* hairpin RNA pathway generates endogenous short interfering RNAs. *Nature* 453, 803–806
51 Watanabe, T. *et al.* (2008) Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* 453, 539–543
52 Berezikov, E. *et al.* (2007) Mammalian mirtron genes. *Mol. Cell* 28, 328–336
53 Ruby, J.G. *et al.* (2007) Intronic microRNA precursors that bypass Drosha processing. *Nature* 448, 83–86
54 Borsani, O. *et al.* (2005) Endogenous siRNAs derived from a pair of natural *cis*-antisense transcripts regulate salt tolerance in *Arabidopsis*. *Cell* 123, 1279–1291
55 Czech, B. *et al.* (2008) An endogenous small interfering RNA pathway in *Drosophila*. *Nature* 453, 798–802
56 Gullerova, M. *et al.* (2011) Autoregulation of convergent RNAi genes in fission yeast. *Genes Dev.* 25, 556–568
57 Wei, W. *et al.* (2011) Functional consequences of bidirectional promoters. *Trends Genet.* 27, 267–276
58 Hartig, J.V. *et al.* (2009) Endo-siRNAs depend on a new isoform of loquacious and target artificially introduced, high-copy sequences. *EMBO J.* 28, 2932–2944
59 Kim, V.N. *et al.* (2009) Biogenesis of small RNAs in animals. *Nat. Rev. Mol. Cell Biol.* 10, 126–139
60 Okamura, K. *et al.* (2007) The mirtron pathway generates microRNA-class regulatory RNAs in *Drosophila*. *Cell* 130, 89–100
61 Flynt, A.S. *et al.* (2010) MicroRNA biogenesis via splicing and exosome-mediated trimming in *Drosophila*. *Mol. Cell* 38, 900–907
62 Mercer, T.R. *et al.* (2010) Expression of distinct RNAs from 3′ untranslated regions. *Nucleic Acids Res.* 39, 2393–2403
63 Jöchl, C. *et al.* (2008) Small ncRNA transcriptome analysis from *Aspergillus fumigatus* suggests a novel mechanism for regulation of protein synthesis. *Nucleic Acids Res.* 36, 2677–2689
64 Fu, H. *et al.* (2009) Stress induces tRNA cleavage by angiogenin in mammalian cells. *FEBS Lett.* 583, 437–442
65 Yamasaki, S. *et al.* (2009) Angiogenin cleaves tRNA and promotes stress-induced translational repression. *J. Cell Biol.* 185, 35–42
66 Zhang, S. *et al.* (2009) The phloem-delivered RNA pool contains small noncoding RNAs and interferes with translation. *Plant Physiol.* 150, 378–387
67 Thompson, D.M. and Parker, R. (2009) The RNase Rny1p cleaves tRNAs and promotes cell death during oxidative stress in *Saccharomyces cerevisiae*. *J. Cell Biol.* 185, 43–50
68 Phizicky, E.M. and Hopper, A.K. (2010) tRNA biology charges to the front. *Genes Dev.* 24, 1832–1860
69 Thompson, D.M. and Parker, R. (2009) Stressing out over tRNA cleavage. *Cell* 138, 215–219
70 Schutz, K. *et al.* (2010) Capture and sequence analysis of RNAs with terminal 2′,3′-cyclic phosphates. *RNA* 16, 621–631
71 Schaefer, M. *et al.* (2010) RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage. *Genes Dev.* 24, 1590–1595
72 Anderson, P. and Kedersha, N. (2009) RNA granules: post-transcriptional and epigenetic modulators of gene expression. *Nat. Rev. Mol. Cell Biol.* 10, 430–436
73 Emara, M.M. *et al.* (2010) Angiogenin-induced tRNA-derived stress-induced RNAs promote stress-induced stress granule assembly. *J. Biol. Chem.* 285, 10959–10968
74 Karginov, F.V. *et al.* (2010) Diverse endonucleolytic cleavage sites in the mammalian transcriptome depend upon MicroRNAs, Drosha, and additional nucleases. *Mol. Cell* 38, 781–788
75 Shin, C. *et al.* (2010) Expanding the MicroRNA targeting code: functional sites with centered pairing. *Mol. Cell* 38, 789–802
76 Otsuka, Y. *et al.* (2009) Identification of a cytoplasmic complex that adds a cap onto 5′-monophosphate RNA. *Mol. Cell. Biol.* 29, 2155–2167
77 Elbarbary, R.A. *et al.* (2009) Modulation of gene expression by human cytosolic tRNase ZL through 5′-half–tRNA. *PLoS ONE* 4, e5908
78 Lundblad, E.W. and Altman, S. (2010) Inhibition of gene expression by RNase P. *New Biotechnol.* 27, 212–221
79 Tan, G.S. *et al.* (2009) Expanded RNA-binding activities of mammalian Argonaute 2. *Nucleic Acids Res.* 37, 7533–7545
80 Halic, M. and Moazed, D. (2010) Dicer-independent primal RNAs trigger RNAi and heterochromatin formation. *Cell* 140, 504–516
81 Li, Z. *et al.* (2009) Characterization of viral and human RNAs smaller than canonical microRNAs. *J. Virol.* 83, 12751–12758
82 Leung, A.K.L. *et al.* (2006) Quantitative analysis of Argonaute protein reveals microRNA-dependent localization to stress granules. *Proc. Natl. Acad. Sci. U.S.A.* 103, 18125–18130

83 Couvillion, M.T. *et al.* (2010) A growth-essential Tetrahymena Piwi protein carries tRNA fragment cargo. *Genes Dev.* 24, 2742–2747
84 Leung, A.K.L. *et al.* (2011) Genome-wide identification of Ago2 binding sites from mouse embryonic stem cells with and without mature microRNAs. *Nat. Struct. Mol. Biol.* 18, 237–244
85 Kaplan, N. *et al.* (2009) The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* 458, 362–366
86 Light, W.H. *et al.* (2010) Interaction of a DNA zip code with the nuclear pore complex promotes H2A.Z. incorporation and INO1 transcriptional memory. *Mol. Cell* 40, 112–125
87 Raab, J.R. and Kamakaka, R.T. (2010) Insulators and promoters: closer than we think. *Nat. Rev. Genet.* 11, 439–446
88 Farnham, P.J. (2009) Insights from genomic profiling of transcription factors. *Nat. Rev. Genet.* 10, 605–616
89 Ganapathi, M. *et al.* (2011) Extensive role of the general regulatory factors, Abf1 and Rap1, in determining genome-wide chromatin structure in budding yeast. *Nucleic Acids Res.* 39, 2032–2044
90 Chen, M. and Manley, J.L. (2009) Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat. Rev. Mol. Cell Biol.* 10, 741–754
91 Wulff, B-E. *et al.* (2011) Elucidating the inosinome: global approaches to adenosine-to-inosine RNA editing. *Nat. Rev. Genet.* 12, 81–85
92 Kertesz, M. *et al.* (2010) Genome-wide measurement of RNA secondary structure in yeast. *Nature* 467, 103–107
93 Ingolia, N.T. *et al.* (2009) Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. *Science* 324, 218–223
94 Lee, Y. *et al.* (2010) Translationally optimal codons associate with aggregation-prone sites in proteins. *Proteomics* 10, 4163–4171
95 Tuller, T. *et al.* (2010) An evolutionarily conserved mechanism for controlling the efficiency of protein translation. *Cell* 141, 344–354
96 Lee, J.E. *et al.* (2010) Systematic analysis of *cis*-elements in unstable mRNAs demonstrates that CUGBP1 is a key regulator of mRNA decay in muscle cells. *PLoS ONE* 5, e11201
97 Gong, C. and Maquat, L.E. (2011) lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3′ UTRs via Alu elements. *Nature* 470, 284–288
98 Czech, B. and Hannon, G.J. (2011) Small RNA sorting: matchmaking for Argonautes. *Nat. Rev. Genet.* 12, 19–31
99 Wlotzka, W. *et al.* (2011) The nuclear RNA polymerase II surveillance system targets polymerase III transcripts. *EMBO J.* 30, 1790–1803
100 Tomecki, R. and Dziembowski, A. (2010) Novel endoribonucleases as central players in various pathways of eukaryotic RNA metabolism. *RNA* 16, 1692–1724
101 Li, W.M. *et al.* (2009) Endoribonucleases – enzymes gaining spotlight in mRNA metabolism. *FEBS J.* 277, 627–641