Non-coding RNAs in homeostasis, disease and stress responses: an evolutionary perspective

Paulo P. Amaral, Marcel E. Dinger and John S. Mattick

Abstract

Cells and organisms are subject to challenges and perturbations in their environment and physiology in all stages of life. The molecular response to such changes, including insulting conditions such as pathogen infections, involves coordinated modulation of gene expression programmes and has not only homeostatic but also ecological and evolutionary importance. Although attention has been primarily focused on signalling pathways and protein networks, non-coding RNAs (ncRNAs), which comprise a significant output of the genomes of prokaryotes and especially eukaryotes, are increasingly implicated in the molecular mechanisms of these responses. Long and short ncRNAs not only regulate development and cell physiology, they are also involved in disease states, including cancers, in host–pathogen interactions, and in a variety of stress responses. Indeed, regulatory RNAs are part of genetically encoded response networks and also underpin epigenetic processes, which are emerging as key mechanisms of adaptation and transgenerational inheritance. Here we present the growing evidence that ncRNAs are intrinsically involved in cellular and organismal adaptation processes, in both robustness and protection to stresses, as well as in mechanisms generating evolutionary change.

Keywords: regulatory RNA; epigenetic inheritance; stress response; regulatory evolution; adaptation

INTRODUCTION

Regulation of gene expression lies at the centre of all aspects of cellular and organismal biology. During development, cells are instructed by endogenous information, modulated by stimuli from the surrounding cells and the environment, which together guide the programmes that determine cell fates and lead to progressive changes in organism morphology and physiology.

Physiological changes beyond normal variations, which can be caused by biotic challenges and abiotic alterations, can threaten cell function and organism survival. These trigger regulatory responses that are reflected in changes in gene expression in development and in the adult. The stresses include, for example, viral infections as well as changing environmental and microenvironmental conditions, such as suboptimal oxygen pressure, exposure to abnormal temperatures, high levels of radiation or chemical insults.

Regulatory processes are also involved in a variety of pathologically altered states (complex ‘regulatory diseases’, as opposed to simple structural ones, caused by monogenic protein-coding mutations), in particular in multifactorial diseases, which are influenced by environmental components [1–3]. The pathologies include different types of cancers, and at the molecular level may involve, for instance, defects in the regulation of chromatin structure, DNA methylation and RNA processing and metabolism.
Responses range from metabolic adaptation to alterations in cell growth to induced cell death, and individual variations are probably key determinants in differential susceptibility to diseases [6].

Given that protein-coding sequences comprise only a small fraction of the genomes of multicellular organisms [7], there is enormous potential for the expansion and evolution of regulatory sequences in non-coding regions. Nevertheless, studies of the mechanisms underlying organism development, stress response and disease have thus far focused on the roles and variations of protein-coding genes, from the onset of classic genetics to modern experimental approaches [8–11]. On the other hand, modern unbiased approaches, such as genome-wide association studies, to investigate genetic variation underlying complex traits, including cancers and other multifactorial diseases, show that most of the phenotypic variation cannot be satisfactorily accounted for by variations in coding regions, but rather are largely associated with non-coding DNA [12–15].

Although it is increasingly accepted that many changes underlying complex traits occur in cis-regulatory regions, such as promoters and enhancers [16], studies in previous decades have also established two interrelated elements in the pool of regulatory components that are potentially major targets. One is the emergence of epigenetic factors in the control of genome function and gene regulation. Complex traits can be influenced by epigenetic mechanisms, involving the modification of histones, DNA and RNA, which affect chromatin structure and gene expression [5,17]. These can be modulated by environmental signals and in principle be transmitted through mitotic or meiotic cell divisions [5,17]. The other major factor is the extensive population of regulatory non-coding RNAs (ncRNAs), which have emerged as an important component in the control of gene expression in all domains of life [12]. Thousands of ncRNAs have been identified and roles have been ascribed in a wide variety of cellular processes, influencing gene expression at many levels, including the site specificity of the chromatin-modifying enzymes that control epigenetic modifications, trajectories and memory [3,18].

In this article, we discuss the roles of regulatory RNAs in the mechanisms underlying homeostasis and disease and explore their involvement in stress responses, illustrating with recent examples and highlighting the intricate connection of RNA signalling with epigenetic mechanisms. We also discuss the evolutionary basis and implications of these RNA regulatory systems, arguing that the unique plasticity and functional versatility of ncRNAs enable them to serve as central targets and substrates for regulatory change. In turn, we posit that ncRNAs have been extensively recruited as tools for both stasis and phenotypic radiation during evolution.

**REGULATORY RNAs IN HOMEOSTASIS, DEVELOPMENT AND DISEASE**

**Properties of ncRNAs**

RNAs are now recognized as vastly heterogeneous and versatile molecules in the regulation of cellular processes. Through transcriptomic studies, thousands of long and short RNAs (sRNAs) have been discovered in the past decade in prokaryotes, eukaryotes and viruses, encompassing transcripts ranging in size from processed small RNAs of ~20–30 bases to long non-coding RNAs (lncRNAs) of hundreds or even thousands of bases (such as the RNA Air, ~108 kb). Small RNAs, such as microRNAs (miRNAs), are involved in the regulation of development, tissue identity and the robustness of biological systems [19,20]. Given the greater coverage of short and processed RNAs in the literature, here we explore mainly—but not exclusively—the processes involving lncRNAs.

RNAs with demonstrated regulatory functions originate from all parts of the genome, with large numbers and most known examples transcribed by RNA polymerase (RNAP) II, although other classes are transcribed by the different polymerases (see below). These transcripts often overlap protein-coding sequences in sense or antisense orientations and are also expressed from intergenic regions. Although some lncRNAs are processed into small RNA species, most are not [21]. Considering long intergenic (linc) RNAs alone, despite the fact that they are more numerous than protein-coding genes [22] and better represented than miRNAs in annotated repositories such as ENSEMBL, the functions of the vast majority have still not been investigated [21,23].

Nevertheless, the growing numbers of lncRNA studies have already revealed widespread roles for lncRNAs in normal cell processes, organismal development and physiology. LincRNAs commonly have
their own regulated promoters, which respond to physiological stimuli such as hormones [24,25] and accordingly can show exquisite spatiotemporal expression patterns [26,27]. Indeed, they are, in the main, produced in a more strictly regulated fashion than protein-coding genes [21,23,28,29]. The observation that cells are better defined by the complement of expressed ncRNAs than the complement of coding genes is significant; it suggests that ncRNAs are crucial for the regulatory processes that determine cell and tissue identity, architecture and function.

Interestingly, early studies indicated that IncRNAs often have highly conserved promoters, even though their primary sequences have on average lower overall conservation than coding sequences [30], indicating that although RNA sequences may be under constraint for their expression patterns, they are evolutionarily more plastic and/or may be more lineage-restricted and subject to positive selection for adaptive radiation [31–34]. Nevertheless, they can contain short sequence modules and secondary structures that may be conserved and that are presumably critical to their functions [18,35,36]. In addition, on average they are produced at lower levels than coding transcripts, with functionally characterized RNAs in some cases being present in only a few copies per cell [28,37,38], suggesting that their expression is finely adjusted to their specific regulatory roles.

The stability of IncRNAs too appears to be finely adjusted according to their regulatory functions, mechanisms of action and physiological state of the cell. Although some IncRNAs are highly stable, on average they are less stable than protein-coding mRNAs [39]. The use of the regulated biosynthesis and stability of small and long RNAs is clear in processes that require quick or dynamic changes in different species, such as in retinal light adaptation [40] and circadian rhythmic changes [18,41,42]. Moreover, some RNAs are only needed transiently—as illustrated by P15AS—which initiates repression of the associated gene through chromatin silencing, but is not required for maintenance of the repression [43]. This example also underscores the interaction of RNA regulation with epigenetic mechanisms as a means of establishing stable regulatory effects.

Interestingly, recent large-scale studies indicate that ncRNAs are predominantly nuclear and are often associated with chromatin [21,28], suggesting that most RNAs perform roles in nuclear processes, such as nuclear architecture, chromatin regulation, transcriptional regulation and RNA processing. Indeed, many RNAs are associated with specific chromatin-modifying and transcription factor complexes and regulate their activity [44,45], capable of acting in trans or cis, including mechanisms that involve the act of transcription [46–49], and potentially coordinating the expression of multiple genes in regulatory networks. However, RNAs can regulate gene expression at all levels and the molecular mechanisms of action involve interactions with other RNAs, DNA, proteins and even small molecules, in which the RNAs have diverse roles, such as molecular guides, scaffolds and decoys, based on the interaction specificity conferred by base pairing and their tridimensional structures [18,37,50]. All these properties are relevant for the processes discussed in this article.

**Biological functions of ncRNAs**

Although their modes of action are varied and still only beginning to be explored, regulatory RNAs are already known to play important roles in many different species. For example, short regulatory RNAs are involved in a variety of processes in several bacteria, including physiological and growth stage transitions, quorum sensing, toxin–anti-toxin systems, plasmid replication and regulation of photosynthesis [51–53]. In budding and fission yeasts, hundreds of sense, antisense and intergenic ncRNAs are involved in essential processes that range from meiotic control and locus pairing to pseudohyphal growth [54–57].

Although the list of processes extends to all classes of organisms, it is better represented in mammals, which have a substantially more extensive non-coding transcriptome. Indeed, the majority of IncRNAs with defined functions, so far, have been characterized in mammalian cells [58], including well-known RNAs, such as HOTAIR, Air and Xist [49,50,59]. Nevertheless, considerable progress has also been made in other model organisms, such as yeasts, worms, fruit fly and zebrafish, in which comparative analysis and transcriptomics have allowed the cataloguing of thousands of IncRNAs [29,34,60–62].

In addition to biological roles in adult tissues, IncRNAs have many functions in the regulation of developmental processes [18]. Their regulatory roles affect important ontogenetic processes, from imprinting to dosage compensation, as well as control of conserved developmental genes, and their broad biological significance is reflected in their precise
expression patterns at different stages of development [18,62,63]. For example, recent analysis in Caenorhabditis elegans shows the precise ncRNA expression associated with dauer formation, male identity, sperm formation and interaction with sperm-specific mRNAs [29].

On a cellular level, RNAs are involved in regulating central processes, such as proliferation, through the expression of independent IncRNA transcripts (such as Gadd7 RNA [64]) and also gene-associated RNAs, such as promoter-associated RNAs that regulate cell cycle genes [65]. These processes may involve the interplay between different ncRNAs, as exemplified by TUG1 and MALAT1, which regulate the expression of the key factor E2F1 during cell cycle through a mechanism of gene shuttling between the two nuclear RNA compartments, modulated by methylation/demethylation of Polycomb 2 protein (Pc2) [66].

Many other RNAs are involved in the control of differentiation and cell identity [67], including maintenance and establishment of pluripotency and lineage specification [63,67–69]. The importance of regulatory RNAs in tissue identity is illustrated by the control of human epidermal differentiation by the ncRNA TINCR (terminal differentiation-induced ncRNA), which regulates somatic tissue differentiation through a post-transcriptional mechanism, involving interaction with specific regulatory proteins and binding to differentiation mRNAs to ensure their expression [70]. Examples are also accumulating for ncRNAs with roles manifested in organism biology, including in neurological processes and specific behaviours [71–76].

ncRNAs underlying regulatory diseases

Given the functional versatility and highly regulated expression of ncRNAs that make them suitable for gene regulation, it is not surprising that they are implicated in a variety of pathological conditions. Indeed, roles for ncRNAs in diseases have been described since the discovery of the first mammalian IncRNA H19, which has described properties of both oncogene and tumour suppressor [77,78] and was more recently implicated in Polycomb regulation and cell signalling [79]. Disease-associated RNAs represent a significant fraction of RNAs characterized and functionally studied to date [58], with a constantly growing catalogue of implicated ncRNAs [80–82], including several implicated in neurological diseases [71,83]. Dedicated databases are already established for IncRNA-associated diseases, with one containing 475 IncRNA interaction entries, including 208 IncRNAs and 166 diseases, and many more with predicted associations [84].

Different types of RNAs have been genetically implicated in specific diseases. For example, deletion of the small nucleolar RNA (snoRNA) 116, which is required for proper postnatal growth and is located in an imprinted region linked to Prader–Willi syndrome, causes behavioural and metabolic changes (deficiency in motor learning and increased anxiety as well as hyperphagia) in a mouse model [76]. However, the involvement of most RNAs in disease has been through expression analysis and associations with regulatory processes. For example, antisense beta myosin heavy chain (β-MHC) transcripts are increased by both hypothyroidism and pressure overload and are proposed to regulate the myosin heavy chain developmental switch, but their functions and mechanism of action are unclear [85,86].

Alterations in regulatory processes are established causes of numerous diseases [1], including defects in the regulation of chromatin modification, transcription, splicing and mRNA translation, which may all be controlled by RNAs [18]. For example, RNAs often regulate transcription factors and are involved in the regulation or mis-regulation of disease-causing genes, including highly studied loci encoding proteins such as p53 [87–90], p15 [43] and PTEN [91–93].

In particular, RNAs have been implicated in regulating epigenetic processes in disease, such as the modification of heritable chromatin modifications, whose mis-regulation may be causative in disease aetiology [5]. For example, the epigenetic RNA regulator HOTAIR, which was first identified as a trans-acting RNA that recruits Polycomb complexes in human fibroblast cells [46], is involved in a number of different cancers, acting as an oncogene that increases metastatic progression and is a marker of poor prognosis [93–96]. Examples also include the well-studied IncRNAs involved in developmental processes such as imprinting and X-chromosome inactivation ([79,97] and see [98]).

Despite the strong links with chromatin and transcriptional regulation, many RNAs can regulate expression post-transcriptionally and alterations in this regulation may lead to disease. For example, the aforementioned RNA TINCR, expressed in epithelial cells, regulates the levels of key differentiation genes, many of which are mutated in human skin diseases [70]. Other RNAs act at the level of splicing
(such as MALAT and ZEB2NAT [99,100]) or as natural miRNA sponges [91,101–103], often affecting important disease-associated genes.

Although the functional bases of the associations of the RNAs with particular diseases are still mostly unknown, recent striking examples of lncRNAs mechanically involved in the pathology of diseases are emerging. One example is the involvement of the novel RNA DBE-T in facioscapulohumeral muscular dystrophy (FSHD), a disease associated with reduction in the copy number of a repeat sequence called D4Z4 mapping to 4q35. A reduction of the number of repeats to less than 11 is linked to disease development, but the mechanistic basis is not completely understood. Recently, it was shown that the reduction of the number of repeats causes expression of DBE-T RNAs, which act in cis by recruiting the Trithorax chromatin activator protein ASH1L, resulting in the de-repression of several FSHD disease-associated genes [104]. Interestingly, this mechanism parallels a process associated with α-thalassemia, in which an aberrant truncation in the LUC7L gene results in the expression of antisense transcripts causing the epigenetic silencing of the overlapping α-globin gene (HBA), involving hypermethylation of the CpG-rich promoter and triggering the onset of the disease [105]. Increasingly, mechanistic and genetic studies are revealing the extent of the involvement of novel regulatory RNAs with several different cancers and genetic diseases [106–108].

ncRNAs IN STRESS RESPONSES

Disease induces stresses that organisms respond to in order to regain homeostasis, such as oncogenic stress and physiological perturbations caused by metabolic disorders. Although physiological stresses and response mechanisms may have roles in normal development and homeostasis (e.g. the roles of physiological cell death or hypoxia in the development of different tissues [109,110]), many natural environmental and physiological changes beyond normal variation constitute disturbances that also trigger vital adaptive responses. Abiotic stresses include a range of environmental influences, such as thermal shocks and radiation exposure; nutritional imbalances and starvation as well as other insults, such as exposure to genotoxic substances. Biotic stresses directly caused by other organisms and viruses are also dealt with by a range of cellular defensive strategies and immune responses.

Although recent systematic analyses reveal a diversity of ncRNAs responding to specific stresses [65,111–117], equally notably, large numbers of ncRNAs have also been spontaneously involved in a large variety of stress responses (see examples of mammalian RNAs in Table 1). Study of the mechanisms of action of these RNAs reveals a range of different strategies, targeting different steps of gene regulation, at both chromatin and transcription [38,65,118–121] and post-transcriptional levels [64,87,122–125] (Figure 1).

In addition, in these responses, ncRNAs can also regulate and/or be targets of genes that are important in specific aspects of stress responses, such as signalling and apoptosis, including FAS, p53, p21, SOD1 and HIF1 (Table 1). Moreover, the establishment of certain cancers and other diseases also rely on a series of adaptations to stress conditions, such as to hypoxia and acidosis, and challenges by the immune system. Given the physiological importance of these responses, several of the RNAs may also have clinical importance, either by their involvement in disease aetiology or by their interference with therapeutic strategies, as exemplified by RNAs that enhance resistance to apoptosis (e.g. to doxorubicin treatment and other genotoxic stresses) found in different RNAs, including PCGEM1 [126,127], CUDR [128,129] and PANDA [65] (additional examples in Table 1).

In this section, we present evidence that both short- and long RNA species are implicated in abiotic stress responses in bacteria, unicellular eukaryotes, plants and animals, followed by an appraisal of their roles in the complex adaptive interactions between host and parasites. The implications of miRNAs in stress have been recently reviewed [114,130], highlighting the importance of small RNAs in adaptive responses.

Abiotic stresses

Although regulatory RNAs have been implicated in stress responses in many organisms, nowhere is this more pronounced than in prokaryotes. Bacteria and archaea are susceptible to substantial variations in their microenvironments and have developed a range of regulatory mechanisms to respond to environmental stressors. sRNAs (~50–300 nt) in particular coordinate responses to a variety of changes, in which a small number of key transcripts occur at the centre of global stress- and environmental-response regulons [166–168].
<table>
<thead>
<tr>
<th>RNA</th>
<th>Stress/stimuli</th>
<th>Features</th>
<th>Functional notes on stress and apoptosis</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADAPT33 (Chinese hamster)</td>
<td>Multiple stresses</td>
<td>Induced under conditions of cytoprotective adaptive response and hydrogen peroxide exposure. Altered electrophoretic migration in response to both hydrogen peroxide and cis-platinum treatment.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Involvement with apoptosis response indicated only by expression following apoptosis stimulation with hydrogen peroxide and staurosporine.</td>
<td>[13, 132]</td>
</tr>
<tr>
<td>αHIF (human)</td>
<td>Hypoxia</td>
<td>Natural antisense transcript to 'hypoxia-inducible factor I, alpha subunit' (HIF-1α). Promoter contains hypoxia response elements to which HIF-1α proteins can bind. Upregulated by prolonged hypoxia.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Hypoxic induction of the aHIF transcript occurs concomitantly with the decrease in HIF-1α mRNA. Proposed role in destabilization of HIF-1α mRNA.</td>
<td>[123, 124, 133]</td>
</tr>
<tr>
<td>ANRIL (CDKN2B-AS) (human)</td>
<td>DNA damage</td>
<td>Found upregulated following DNA damage, among many other lncRNAs including a small collection of known ones, such as CCND1 (below) and TUG1. Upregulation involves the transcription factor E2F1 in an Ataxia Telangiectasia Mutated (ATM)-dependent manner.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Elevated levels of ANRIL suppress-associated genes (INK4a, INK4b and ARF) at the late-stage of DNA damage response, allowing the cell to return to normal at the completion of repair.</td>
<td>[120]</td>
</tr>
<tr>
<td>APEG3 (rat)</td>
<td>Osmotic regulation</td>
<td>Antisense to Paternally expressed gene 3. Specifically expressed in vasopressinergic magnocellular neurons (VP-MCNs) in the supraoptic nucleus.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Involvement indicated only by expression, with both Peg3 and APEG3 expression in the VP-MCNs increasing during systemic hyperosmolality in vivo.</td>
<td>[134]</td>
</tr>
<tr>
<td>BACE1AS (human, mouse)</td>
<td>ER stress</td>
<td>Antisense to beta-secretase-1 (BACE1) gene. BACE1-AS and BACE1 are increased upon induction of endoplasmic reticulum (ER) stress and in biopsies of sporicul inclusion body myositis, as well as in brains of Alzheimer’s disease patients.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Regulates BACE1 expression in vitro and in vivo at post-transcriptional level.</td>
<td>[135, 136]</td>
</tr>
<tr>
<td>Bs (rat)</td>
<td>Drug treatment</td>
<td>Present only in rats. Expressed mainly in the central nervous system in differentiating cells but not in proliferating cells during brain development. Introns encode snoRNAs.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Localized in novel nuclear bodies, with cellular localization responding to drug treatment (actinomycin D or other drugs, such as leptomycin B and DRB) by shuttling to cytoplasmic stress granules.</td>
<td>[137–139]</td>
</tr>
<tr>
<td>CCND1-associated RNAs (human)</td>
<td>DNA damage</td>
<td>Transcribed from the promoter region of the cyclin D (CCND1) gene (a cell cycle regulator repressed by DNA damage signals) and induced by DNA damage signals.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AllostERIC control of theTLS (translocated in liposarcoma) protein, leading to inhibition of CREB (cAMP response element-binding protein) and histone acetyltransferase activities to repress cyclin DI.</td>
<td>[118]</td>
</tr>
<tr>
<td>CTN-RNA (mouse)</td>
<td>Transcription—inhibition/</td>
<td>‘CAT2 transcribed nuclear’ (CTN)-RNA. Produced through usage of alternative promoter and poly-A site of the CAT2 gene. Found diffusely distributed in nuclei as well as localized to paraspeckles, with RNA editing of its 3' UTR involved in its nuclear retention.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>stresses involving nitric oxide pathway</td>
<td></td>
<td>Under transcription—inhibition and physiologic stress CTN—RNA is cleaved to produce protein-coding mCAT2 mRNA released to the cytoplasm.</td>
<td>[140]</td>
</tr>
<tr>
<td></td>
<td>response</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CUDR (UCAI) (human)</td>
<td>Drug treatment, including doxorubicin and etoposide</td>
<td>‘Cancer upregulated drug resistant’. Upregulated in various human tumours. Overexpressed in doxorubicin-resistant squamous carcinoma cells, which are also more resistant to drug-induced apoptosis.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ectopic overexpression of different isoforms causes enhanced drug tolerance and resistance to apoptosis, as well as promotes transformation in vitro, enhancing cell motility and tumorigenic behaviour in vitro and in vivo.</td>
<td>[128, 129, 141]</td>
</tr>
<tr>
<td>GADD7 (adapt15) (Chinese hamster)</td>
<td>Multiple stresses</td>
<td>Induced by stresses that include UV radiation and DNA damage, lipotoxic stress, serum starvation, exposure to alkylating agents, hydrogen peroxide and heat shock.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Regulation of lipid-induced oxidative and ER stresses. Overexpression leads to a decrease in cell growth and is</td>
<td>[64, 142–144]</td>
</tr>
<tr>
<td>RNA</td>
<td>Stress/stimuli</td>
<td>Features</td>
<td>Functional notes on stress and apoptosis</td>
<td>Reference</td>
</tr>
<tr>
<td>---------------------------</td>
<td>-----------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>GAS5 (human, mouse, rat)</td>
<td>Starvation and growth arrest</td>
<td>Growth arrest-specific transcript. Highly unstable but stabilized by growth arrest, serum starvation and inhibition of protein synthesis. Gas5 is also a snoRNA host, but has functions as an lncRNA.</td>
<td>Necessary for both lipid- and general oxidative stress-mediated cell death. RNA represses the different steroid hormone receptors. Shown to function as a decoy repressor that interacts with the glucocorticoid receptor, preventing binding to regulatory targets, including apoptosis inhibitors. Controls apoptosis and cell cycle.</td>
<td>[21, 146–147]</td>
</tr>
<tr>
<td>H19 (Therians)</td>
<td>Serum stress, hypoxia and tumorigenesis</td>
<td>Imprinted and developmentally regulated. May act as both a miRNA precursor and as regulatory lncRNA. Regulates growth during development and has been implicated in cancer as both tumour suppressor and oncogene.</td>
<td>Growth advantage role of H19 RNA in serum stress (with cells cultured in serum-poor medium). Upregulated by hypoxia in cancer cells and knockdown impairs anchorage-independent growth after hypoxia recovery.</td>
<td>[78, 79, 148–150]</td>
</tr>
<tr>
<td>lincRNA-P21 (mouse)</td>
<td>DNA damage</td>
<td>Approximately 15 kb upstream to the p21 gene, transcribed in the antisense orientation. Also suggested to act as a translational repressor of specific mRNAs.</td>
<td>Interacts with hnRNP-K and represses many p53 target genes. Involved in p53-dependent apoptotic induction, but not cell cycle arrest.</td>
<td>[119, 151]</td>
</tr>
<tr>
<td>lincRNA-ROR (human)</td>
<td>DNA damage</td>
<td>Highly expressed in embryonic stem (ES) and induced pluripotent stem (iPS) cells, and able to promote reprogramming of differentiated cells to iPSCs. In cancer cell lines, p53 positively regulates RoR expression by direct interaction with a p53 response element in the putative RoR promoter.</td>
<td>Knockdown in ES and iPS cells leads to modest increase in apoptosis and activation of p53 pathways. In p53-expressing cancer cells, inhibits p53 response to DNA damage, including cell cycle arrest and apoptosis. Suppresses p53 translation through interaction with hnRNP L.</td>
<td>[87]</td>
</tr>
<tr>
<td>IncRNA-LET (human)</td>
<td>Hypoxia</td>
<td>Repressed by Hypoxia-induced HDAC3. Downregulated in hepatocellular carcinomas, colorectal cancers and squamous cell lung carcinomas.</td>
<td>Binds to NF90 and destabilizes it. Downregulation leads to hypoxia-induced cancer cell invasion, involving stabilization of NF90 protein.</td>
<td>[152]</td>
</tr>
<tr>
<td>LUST (human)</td>
<td>CD95-mediated apoptosis</td>
<td>Intronic RNA transcribed antisense to the putative tumour suppressor gene RBMS.</td>
<td>Ectopic expression affects splicing of RBMS and suppresses CD95-mediated apoptosis.</td>
<td>[153]</td>
</tr>
<tr>
<td>MSURI (mouse)</td>
<td>Oxidative stress</td>
<td>Sequence is comprised of 1.7 kb of the 18S RNA followed by a non-rRNA 3′-end of 31 bases, with a polyA signal. Upregulated in spinal cord of superoxide dismutase 1 (SOD1) mutant mice, which is linked with the development of amyotrophic lateral sclerosis.</td>
<td>Rescues mutant SOD1-mediated cell death. Reduces cellular oxidative stress, decreasing levels of hydroxyl radicals and oxidation of cellular proteins.</td>
<td>[154]</td>
</tr>
<tr>
<td>Neat1 (VINC) (placental mammals)</td>
<td>Viral infection</td>
<td>Abundant and widely expressed nuclear RNA, essential for formation and maintenance of paraspeckle structures. Upregulated upon mice infection with Japanese encephalitis or Rabies viruses, as well as in human T-cell lines infected with HIV. Upregulated in the nucleus accumens of human brains of heroin abusers.</td>
<td>Depletion in infected human T cell enhances HIV-1 virus production, through putative mechanism involving increased nucleus-to-cytoplasm export of unspliced HIV mRNAs.</td>
<td>[154–157]</td>
</tr>
<tr>
<td>NeST (Tmevpg1) (mouse)</td>
<td>Microbial infection</td>
<td>Present on a viral susceptibility locus, transcribed downstream to the Interferon-gamma (IFN-γ) in the antisense orientation. Expression required for inducible</td>
<td>Binds to WD5, inducing in trans the deposition of active chromatin marks at the IFN-γ locus. Increases Theiler's</td>
<td>[38, 158]</td>
</tr>
</tbody>
</table>
### Table 1: Continued

<table>
<thead>
<tr>
<th>RNA</th>
<th>Stress/stimuli</th>
<th>Features</th>
<th>Functional notes on stress and apoptosis</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>PS3 mRNA (human)</td>
<td>DNA damage and apoptosis</td>
<td>p53 RNA binds the Mdm2 protein in a regulated manner and promotes p53 translation and stabilization, also impairing Mdm2 E3 ligase activity. p53 RNA—Mdm2 interaction also controls Mdm2 SUMOylation and nuclear trafficking.</td>
<td>Virus persistence and reduces Salmonella lethality, regulating susceptibility to a viral and a bacterial pathogen. Induction of DNA damage with doxorubicin stimulates p53 mRNA interaction with Mdm2. p53 stabilization following genotoxic stress requires p53 mRNA—Mdm2 interaction, and its pro-apoptotic activity is dependent on binding to Mdm2.</td>
<td>[59, 160]</td>
</tr>
<tr>
<td>PANDA (human)</td>
<td>DNA damage</td>
<td>Transcribed bidirectionally to the p21 (CDKN1A), cyclin-dependent kinase inhibitor 1 gene. Induced along with p21 in human fetal lung fibroblast cells upon doxorubicin-induced DNA damage. Both are also positively regulated by p53.</td>
<td>Loss of PANDA affects the expression of different loci, including key apoptosis genes, and leads to the apoptosis of doxorubicin treated cells. Interacts with NF-YA and attenuates its occupancy on target gene promoters.</td>
<td>[65]</td>
</tr>
<tr>
<td>PCGEMI (human)</td>
<td>DNA damage and apoptosis</td>
<td>Prostate-specific and prostate cancer-associated, upregulated by cholesterol in prostate cancer cell lines.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRINS (human)</td>
<td>Multiple stresses</td>
<td>‘Psoriasis Susceptibility-Related RNA Gene Induced by Stress’ (PRINS), upregulated in psoriatic epidermis. It is also upregulated by multiple stresses, including UV-B irradiation, viral infection (herpes simplex virus), transnational inhibition, as well as exposure to microbial agents such as peptidoglycan, LPS and wall extract. In keratinocytes, PRINS regulates the expression of the G1P3 gene, an anti-apoptotic protein that is overexpressed in psoriasis.</td>
<td></td>
<td>[61–163]</td>
</tr>
<tr>
<td>PTENp51 (human)</td>
<td>DNA damage</td>
<td>PTEN pseudogen-encode antisense RNA. Two different isoforms independently regulate negatively the transcription and mRNA stability of tumour-suppressor PTEN.</td>
<td>Regulates cell cycle and apoptosis, as its disruption induces cell cycle arrest and sensitizes cells to doxorubicin.</td>
<td>[92]</td>
</tr>
<tr>
<td>SAF (human)</td>
<td>Fas-mediated apoptosis</td>
<td>Antisense RNA transcribed from intron I of the FAS gene, involved in membrane-mediated apoptosis.</td>
<td>Overexpression of Saf affects splicing of Fas and protects Jurkat cells from Fas-mediated, but not TNFα-mediated, apoptosis.</td>
<td>[122]</td>
</tr>
<tr>
<td>Uchll antisense (mouse)</td>
<td>Rapamycin-induced stress</td>
<td>Antisense/bidirectional to Uchll, a pro-survival gene. Regulates Uchll expression at the level of translation through a novel mechanism involving base complementarity and an embedded SINE B2 repeat. Stress-induced rapamycin treatment (inhibiting the mTOR pathway) causes shutting of the antisense RNA to the cytoplasm, promoting association of Uchll mRNA to polysomes for translation upregulation.</td>
<td></td>
<td>[164]</td>
</tr>
<tr>
<td>Urocortin antisense (rat)</td>
<td>Restraint stress in rats</td>
<td>Antisense to Urocortin (UCN) gene. UCN mRNA and antisense RNA is co-expressed with sense mRNA in many tissues and both are responsive to restraint stress. Anti-sense RNA expression was responsive to restraining tube-induced stress, with suggestion of antisense to play a role in regulating transcription or translation of UCN mRNA.</td>
<td></td>
<td>[165]</td>
</tr>
<tr>
<td>WRAP53 (human, mouse)</td>
<td>DNA damage (p53 regulation)</td>
<td>Divergently transcribed to the P53 gene, with a conserved 5'-antisense overlap; encodes a WD40 domain-containing protein. Expressed at low levels in multiple tissues, positively correlated with p53 expression. RNA positively regulates p53 mRNA and protein expression at a post-transcriptional level.</td>
<td>Knockdown causes significant decrease in p53 mRNA levels and suppresses induction of p53 upon DNA damage, whereas overexpression increases p53 mRNA and protein levels and potentiates sensitivity to p53-dependent apoptosis.</td>
<td>[90]</td>
</tr>
</tbody>
</table>

Species in which the RNAs have been identified are also indicated.
The accumulated examples of RNAs involved in prokaryote stress response, first reported several decades ago [169], are numerous and the mechanisms increasingly understood [51,53,166,170]. Classic instances of ncRNAs include DsrA, which has regulatory roles during exponential growth in low temperature [171]; OxyS, a trans-acting antisense RNA induced by oxidative stress that mediates adaptation to hydrogen peroxide and other stress responses [172] as well as heat shock-inducible antisense RNAs that regulate heat shock protein expression [173].

Interestingly, promoters of specific sRNA genes in bacteria are among the most stress sensitive [174], and novel mechanisms of stress response are emerging and connected to specific biological outcomes. For example, the S-adenosylmethionine riboswitches SreA and SreB function in trans to regulate the virulence factor PrfA, thus connecting response to nutrient availability in bacteria to virulence [175]. Indeed, RNAs in different prokaryotic groups can respond to an array of stresses, including RNAs in cyanobacteria, exemplified by the ~65 nt Yfr1 RNA, which forms a secondary structure containing a highly conserved sequence, present in most cyanobacterial lineages, and important to growth in multiple, yet specific, stress conditions such as upon iron limitation, oxidative stress and high salt conditions [176].

Unicellular eukaryotes, free-living or parasitic, are also highly dependent on environment adaptation and response to host defences. There are cases already identified of RNA-based responses to specific environmental factors, such as the XUT (Xrn1-sensitive unstable transcripts) antisense RNAs in yeasts, which accumulate in lithium-containing media and are thought to play a role in adaptive responses to changes in growth conditions [177], or classes of RNAs that respond (in this case downregulated) to starvation in Dictyostelium amoeba and are involved in the natural adaptation response through aggregation and cell specialization [178].

Aspects of normal life cycle in yeasts are also triggered by environmental variation and stress conditions, such as sporulation upon nutrient deprivation, comprising a series of important changes that lead to gametogenesis and the development of spores. The critical cell fate decision is triggered by extracellular and intracellular signals, but is mediated by the transcription of two overlapping lncRNAs. The lncRNAs control the expression of two central regulators of sporulation involving a transcription-dependent deposition of H3K4me2 and recruitment of Set3C deacetylase [179,180]. Other important adaptive processes, such as pseudohyphal growth and growth and metabolic adaptation in changed environments, such as upon glucose starvation, also involve regulation by different types of RNA in

Figure 1: Regulatory RNAs in stress response regulatory networks control gene expression in many levels. RNAP III transcripts include 7SK and SINE B2/Alu RNAs, and RNAP II targets include protein coding (PC) and non-coding (NC) RNA genes. Examples of stress responsive ncRNAs synthesized by different RNAPs and regulating gene expression are indicated (see Table I and main text for references).
yeasts [55,116,177]. Indeed, even heterogeneity among genetically identical yeast cells can be determined by the variable transcription of ncRNAs that control key regulatory genes, resulting in variated expression and phenotypic variation in clonal cell populations that may be important for adaptability of the population in fluctuating environmental conditions [55].

In plants, as sessile organisms exposed to environmental influences, RNAs have been identified as major components in response to variations and stresses, such as to drought, changed light regimes and nutrient and salt stresses, involving long RNAs, miRNAs and other endogenous small RNAs [112,113,181–188]. Environmental variations are part of plant natural life cycles and stresses can also be related to specific developmental processes. For example, the npc536 RNA promotes root growth under salt stress conditions [113], and at least two different IncRNAs are crucial in the phenomenon of vernalization in flowering species, which are exposed to prolonged periods of cold prior to flowering in spring. This involves the intersection of RNA regulation with chromatin-based epigenetic mechanisms, including the Polycomb silencing of the floral repressor FLOWERING LOCUS C (FLC). The intronic IncRNA Coldair, which is transcribed from the FLC locus, plays a role in the stable repression of FLC through specific recruitment of Polycomb components to target sequences [189], whereas the antisense transcript Coolair, which originates from the 3’-end of the FLC gene, has an early role in the silencing of FLC through a mechanism that involves antisense transcription [190]. More generally, given their extensive involvement in stress response and development, including through epigenetic mechanisms, ncRNAs have the potential to emerge as important molecules in the regulatory mechanisms that confer developmental and phenotypic plasticity in plants and other organisms [191,192].

Many examples of ncRNAs involved in systematic stress response have been described in animals ([193] and see below). In Drosophila, the heat shock locus hsr-ω (heat shock RNA omega) produces a set of nuclear-localized IncRNAs. hsr-ω RNAs are expressed in many tissues, including during development, and form localized nuclear complexes named ω-speckles, which are distributed in the interchromatin space in close proximity to the chromatin [194–197]. These ncRNAs are thought to regulate the availability of heterogeneous nuclear ribonucleoproteins (hnRNPs) in active (chromatin bound) and inactive (in ω-speckles) compartments and interfere with the processing of pre-mRNAs [198]. Interestingly, hsr-ω is also induced by other stressors, such as amides, benzamide and colchicine [199], and non-coding transcripts analogous to Drosophila hsr-ω transcripts are found in other species of flies interacting with Hsp90 [200], indicating broad roles of these RNAs in stress response.

In mammals, several non-coding transcripts have been similarly implicated in stress responses, such as heat shock, where ncRNAs orchestrate global changes at both post-transcriptional and transcriptional levels. RNAP III transcription of SINR repeat elements in both mouse (SINE B2) and humans (Alu) dramatically accumulate during heat shock and promote general transcriptional repression by direct interaction and inhibition of RNAP II (Figure 1), which is accomplished by interfering with the formation of the preinitiation complex in specific target gene promoters [201,202].

In addition, pericentromeric Satellite III (Sat III) repeats are activated upon thermal or chemical stress, generating large amounts of polyadenylated ncRNAs that accumulate in nuclear stress bodies [203–208]. Sat III RNAs have a key role in the recruitment and scaffolding of RNA processing and transcription factors, acting as nucleation centres for the assembly of the nuclear bodies [205,209], a theme found for other nuclear RNAs [210]. Interestingly, another RNAP II ncRNA, HSR1 (heat shock RNA-1), has been identified as an RNA ‘thermosensor’ in mammalian cells (despite unclear evolutionary origins [211]), which upon heat shock appears to undergo a conformational change, forms a complex with the translation elongation factor eEF1A and stimulates the trimerization of the transcription factor HSF1, promoting activation of heat shock response genes [212].

Altogether, these studies illustrate the extensive involvement of RNAs in response to environmental changes. Several other RNAP II IncRNAs are implicated in other forms of stress in mammalian cells by various mechanisms, many of which are involved in control of DNA damage response, cell cycle arrest and apoptosis (a common theme for many ncRNAs—see examples in Table 1).

Systematic involvement of classes of RNA in stress
In addition to the complementary roles of ncRNAs in stress response, several additional classes of RNAs
are systematically involved. This is illustrated by the extensive roles of RNAs in DNA damage, which encompass multiple levels of responses including DNA repair, cell cycle and apoptosis regulation. For example, in the filamentous fungus *Neurospora crassa*, specific long RNAs are produced upon DNA damage, predominantly from the ribosomal DNA locus, and are processed into small RNAs associated with the Argonaute (Ago) protein QDE-2, denominated qRNAs (QDE2-interacting RNAs). The qRNAs are proposed to have a role in DNA damage through inhibition of protein translation, which is supported by the increased sensitivity to DNA damage observed in RNAi mutants [213].

In mammals, miRNAs have a critical role in DNA damage response [214,215]. Components of the miRNA processing pathways are crucial for survival after UV irradiation, involving re-localization of Ago2 into stress granules in a cell-cycle-dependent manner, as well as expression change of several miRNAs that target key checkpoint genes, with impact in cell proliferation [216].

Recently, it was found that the components of the RNAi pathway DICER and DROSHA are more generally involved in response to DNA damage and oncogene-induced genotoxic stress in vertebrate cells, through the production of dedicated small RNA species named DDRNAs (DICER- and DROSHA-dependent small RNAs) [217]. The integrity of ‘DNA Damage Response’ foci is sensitive to RNAse treatment, and both RNAi components are required for their formation, with a direct role assigned for DDRNAs in the control of DNA damage response activation at sites of damage [217]. The involvement of RNAs may be even more extensive, as other proteins that have key roles in sensing and responding also interact with RNAs [218]. In addition to the mechanisms implicating small RNAs, there is also extensive evidence for systematic involvement of IncRNAs in stress response in eukaryotic cells, involving global changes in gene expression and regulation at both the transcriptional and post-transcriptional levels.

Broadly, stresses trigger profound effects on transcription, RNA processing and translation. In particular, there are entire response mechanisms for stress response based on RNAP I and III transcription. For example, 7SK is a highly abundant and ubiquitous RNAP III transcript, which is a global regulator of RNAP II transcription in vertebrates [219]. 7SK acts in complexes with the Hexim1 regulator and other proteins to repress specific groups of genes in different cell types, through a mechanism involving sequestration and inhibition of the elongation factor P-TEFb [219–221]. Multiple stress inducers, including UV radiation, actinomycin D and DRB treatment, cause dissociation of Hexim1 and 7SK from P-TEFb, leading to the transcription activation of multiple genes [219,220,222]. Interestingly, HIV uses an RNA regulatory element to activate its transcription by counteracting 7SK repression of P-TEFb, and stress with UV irradiation or actinomycin D increases transcription of the viral genome [220,223].

In addition to the above examples of RNAP II and RNAP III transcription (such as the SINE B2 and Alu RNAs, as well as the Sat III RNAs and stress granules [208]), there is evidence that RNAP I transcription and the nucleolus are major centres of stress response. Although increased RNAP I activity is associated with stimulation of cell growth and proliferation, a number of different stresses (including starvation, oxidative stress, toxic lesions and inhibition of protein synthesis) lead to a repression of rDNA transcription [224], whose epigenetic regulation has been linked with promoter-associated RNAs [225,226]. In addition, the state of RNAP I activity also connects with stress response signalling pathways [224,227–229], including control of p53 activity in response to DNA damage and other stresses.

Finally, ncRNAs may indeed commonly regulate general stress responses in eukaryotes at post-transcriptional levels, including translation. Although more general regulatory mechanisms are emerging, such as differential tRNA codon usage upon environmental perturbations, which may have significant impact in gene regulation during stress [230], there are also some specific mechanisms. The nuclear-enriched RNA antisense/bidirectional to the Uchl1 gene specifically regulates *Uchl1* expression in mouse through a novel mechanism that leads to upregulation in its protein synthesis [164]. Interestingly, the use of stress inducer rapamycin, which acts through mTORC1 kinase inhibition and leads to a general 5’-cap-dependent translation repression, causes shuttling of the antisense Uchl1 RNA from the nucleus to the cytoplasm, thus leading to association of the overlapping sense Uchl1 mRNA to active polysomes and translation upregulation. This targeting depends on the presence of a 5’-overlapping sequence and an inverted SINE B2
element on the RNA sequence. These elements are also present in several other antisense transcripts and were shown to confer regulatory activity using a reporter system, leading to the suggestion that stress-dependent nucleocytoplasmic shuttling of lncRNAs is a common strategy for post-transcriptional regulation [164]. Interestingly, different shuttling mechanisms in RNA regulation in stress are also suggested for other transcripts [137,140].

ncRNAs IN HOST–PATHOGEN INTERACTIONS, ATTACK AND DEFENCE

RNA in host defence

RNA-based defence systems are found in prokaryotes and eukaryotes, involving both small and long RNAs. In addition to potential RNA protective strategies, such as antisense RNA antitoxins in bacteria [231], an entire dedicated system was recently discovered in both bacteria and archaea in which small RNAs confer resistance to mobile genetic elements, such as plasmids and phages [232]. The RNAs are encoded in regions called CRISPRs (clustered regularly interspaced short palindromic repeats), which incorporate short sequences from invading genetic elements. These are transcribed and processed into small RNAs that guide protein effectors (composed of Cas proteins) to destroy invading genetic material through site-specific DNA cleavage [232]. In essence, this highly specific CRISPR/Cas system corresponds to an adaptive immune system in prokaryotes that provides protection against future infections [232–236]. In addition, CRISPR/Cas mechanisms are proposed to be involved in response to other stresses, such as DNA damage and protein mis-folding [232].

RNA-based systems have similarly evolved in eukaryotes for protection against both viruses and endogenous genetic mobile elements and are important mediators of viral immunity in yeast, worms and plants through various mechanisms. The RNA interference (RNAi) system is highly conserved in eukaryotes, involved in viral protection and endogenous regulatory processes, and the PIWI-associated RNA system is important for transposon control in both vertebrates and invertebrates [236,237]. Such systems are also based on specialized effector protein complexes and involve processing of long RNAs into small RNA precursors and targeting of specific nucleic acid sequences for silencing (usually through RNA degradation or DNA methylation and chromatin modification) [236–239].

A novel example of RNAi-mediated viral defence has been identified in worms. Expression of the Flock House virus in *C. elegans* is silenced by small RNAs processed from the virus RNA known as viRNAs (virus-derived, small-interfering RNAs) [240]. Remarkably, the processed viRNAs are transmitted transgenerationally in a non-Mendelian fashion, conferring resistance to infection to many subsequent generations, even in animals that are deficient in producing their own viRNAs. These observations suggest a powerful protection mechanism that may provide adaptive benefits to individuals and their derived lineages (see more on epigenetic transgenerational effects below).

In addition to the involvement in general defence systems against viruses in plants and animals, induction of specific endogenous small and long ncRNAs in response to pathogen infection has been observed in diverse organisms. These include *Dictyostelium* [241], mosquito disease vectors [242], plants [114,243–246] and mammals [38,155,247], indicating a role for ncRNAs in the immunity of these organisms. Indeed, several studies show that hundreds of small RNAs, such as miRNAs [248], and lncRNAs are differentially expressed in developing and challenged immune systems in different hosts. For example, in the malaria vector *Anopheles gambiae*, expression of different miRNAs responds to *Plasmodium* invasion and disruption of the RNAi pathway leads to increased sensitivity to infection [242], indicating a role for the small RNAs in host defence that may be widespread and diversified [249].

In mammalian systems, ncRNAs and their transcription are involved not only in basic processes of immunological diversification [250,251], but specific RNAs have also been implicated in cellular and systemic responses to pathogens. Initial large-scale studies already showed multiple RNAs responding to lipopolysaccharide (LPS) response in macrophages [252]. Different studies have expanded the repertoire of RNAs, including recent single-molecule RNA-seq analysis of transcriptome changes during LPS stimulation, which showed large numbers of regulated lncRNAs, encompassing thousands of regulated intronic RNAs [253]. Expression of hundreds of lncRNAs was also observed, for example, in cytotoxic CD8(+) T cells, many of which are lymphoid-specific and change dynamically with lymphocyte differentiation or activation [254], indicating a role...
in adaptive immunity. The examples of functional RNAs in immune cells include transcripts that act through epigenetic mechanisms that determine protection and susceptibility to viral and bacterial pathogens [38].

Finally, in eukaryotes, there are examples of specific lncRNAs controlling transposon activity, such as the yeast trans-acting cryptic antisense RNA in the silencing of the Ty1 repetitive element via Set1 histone methyltransferase [255], in a mechanism that may be regulated by environmental stress [256]. Together, these mechanisms in which RNAs act directly and specifically as defence tools (often involving the formation of small dsRNAs) add to the well-known general RNA-mediated PKR ('Protein kinase RNA-activated') response triggered by long dsRNAs upon viral infection in the cytoplasm of vertebrate cells [257,258].

**RNAs in pathogens**

Analogous to the employment of RNAs in defence, pathogens similarly deploy a range of RNA-based regulatory mechanisms for adaptation to host environments. As exemplified below, both individual ncRNAs and classes of RNAs are found in several pathogens, where they are used not only as regulatory virulence factors but are also important in defence to host response through different mechanisms. Indeed, although the genomes of pathogens are usually compact, they often encode a significant variety of small and long ncRNAs that have crucial functions. Viroids themselves may be regarded essentially as replicating regulatory ncRNAs [259], and both viroid- and viral-encoded ncRNAs regulate their own expression and replication, as well as manipulate the cellular environment and interfere with cellular and organismal physiology [259–261]. The variety of mechanisms deployed by such compact genomes highlights the versatile roles of regulatory RNAs.

A large number of miRNAs (over 200 identified to date) are encoded by viruses and play diverse roles [261,262]. For instance, during human cytomegalovirus (HCMV) infection, a virus-encoded miRNA (miR-UL112) represses the translation of a cell surface component of the MHC that is activated under high stress, preventing signalling to immune cells that would promote lysis of the infected cell [263]. In addition, viruses are able to explore host regulatory RNAs for their own advantage, such as by the Hepatitis C virus, which recruits a liver-specific miRNA to the 5′-terminus of the viral RNA genome that results in a stabilizing effect on the viral RNA [264,265]. Interestingly, viruses can also interfere with host miRNA signalling to affect cellular gene expression, as exemplified by the long U-rich ncRNAs encoded by primate herpesvirus saimiri, named HSURs, which contain three competitive binding sites of the host miRNAs [266].

In addition to virus-encoded miRNAs [261], longer ncRNAs occur in various species, including adenoviruses [267,268], flaviviruses [269], HCMV [270,271], Kaposi’s sarcoma-associated herpesvirus (KSHV) [272–276] and Epstein–Barr virus (EBV) [257]. For example, transcriptome analysis of the HCMV during different stages of infection of human fibroblast cells showed that 45% of the cloned cDNAs arise from non-coding regions of the genome and more than half are antisense to well-known general RNA-mediated PKR (‘Protein kinase RNA-activated’) response triggered by long dsRNAs upon viral infection in the cytoplasm of vertebrate cells [257,258].
rapidly accumulated in infected cells, accounting for \( \sim 20\% \) of viral RNA transcription [287], \( \beta 2.7 \) has a role in preventing stress response and apoptosis of the host cell through a mechanism involving specific binding of the RNA to components of the mitochondrial respiratory chain. This prevents re-localization of essential subunits associated with GRIM-19 (retinoid/interferon-induced mortality-19) in response to apoptotic stimuli. Such interaction is important for stabilizing the mitochondrial membrane potential and adenosine triphosphate production, preventing metabolic dysfunction, in a process that is essential for completion of the viral life cycle [271]. Remarkably, there is evidence that \( \beta 2.7 \) transcripts protect endothelial cells from rat aorta against apoptosis during ischaemia/reperfusion injury, with overexpression of the RNA reducing reactive oxygen species production [288].

Transcriptomic studies and functional analyses have identified many other examples of ncRNAs in non-viral pathogens, including species of Mycobacterium, Listeria, Salmonella, Staphylococcus and Pseudomonas genus. Indeed, bacterial genomes encode a complex and dynamic transcriptome with a variety of sRNAs and antisense transcripts [289–293], with increasing evidence that these have a broad importance for pathogenicity [294,295]. The diverse roles in virulence encompass biofilm formation, quorum sensing, neutralization of host defence, intracellular survival and pathogenesis-associated stress tolerance, toxin production and drug resistance [51,52]. One such example is the requirement of a small RNA in Neisseria gonorrhoeae, whose transcription has a \textit{cis}-regulatory role to induce pilin antigenic variation important in host immune response evasion [296].

This broad involvement is illustrated in the recent implications of ncRNAs in the switch from saprophyte to pathogen in \textit{Listeria} species, whose mechanisms have been largely elusive. An unbiased transcriptomic study showed a diversity of RNAs that are absent in non-pathogenic species and exhibit the same expression patterns as the known virulence genes [289]. The extensive use of RNAs in parasites is also highlighted by the emerging classes of RNAs in eukaryotic pathogens, such as Pinci1 ncRNA family in the fungal plant–pathogen \textit{Phytophthora infestans} expressed upon infection [297], ncRNAs in protozoans, such as \textit{Leishmania} and \textit{Plasmodium}, as well as other animal parasites such as schistosomes [298–302].
the emerging perception of the high evolutionary plasticity of RNAs and their molecular properties demand a different perspective and attitude when designing and interpreting functional RNA analyses. In fact, those properties may be intrinsic to the physiological and evolutionary relevance of these RNAs.

The participation of RNAs in complex regulatory circuits and specific environmental conditions indicates that some functions will be subtle and only apparent under specific environmental or experimental circumstances. For example, in *Drosophila*, the miRNA miR-7 acts to buffer regulatory networks against perturbation during development, and its critical function is only evident upon temperature instability stress [308]. Similar phenomena are also observed with mammalian ncRNAs, such as the rodent-specific BC1 RNAs, which evolved from a tRNA sequence [309] and whose biological role in knockout studies only becomes more apparent in field experiments, with an affected exploratory behaviour and anxiety phenotype, and with a lethal reaction upon auditory stress through epileptic activity and convulsive seizures [75,310].

The above cases show that RNAs are involved in adaptive responses to rapidly changing conditions in all classes of organisms, exploring their dynamic biosynthesis, stability and regulatory capacity. RNAs have been explored for a variety of molecular roles and processes throughout evolution, based on the same basic functional properties that have been selected in RNAs acting as molecular guides, scaffolds and even enzymes in ancestral functions in protein translation and gene expression (protein encoding, adapters in the transfer of aminoacids, peptide bond catalysis, splicing and RNA modification guiding). For example, the modularity and flexibility of RNA structures are employed in several mechanisms of adaptability.

In addition to the punctual studies of RNAs as switches and chemical ‘sensors’, in the shapes of RNA scaffolds, RNA ‘thermosensors’ and in riboswitches ([37,311–313] and examples above), recent high-throughput studies of RNA secondary structure show the abundance and dynamics of RNA structures throughout the transcriptome, with potential for the formation of thousands of RNA sensors [314]. These properties also allow a broad use of RNAs as molecular scaffolds, as illustrated above by RNAs that are central for the formation of nuclear bodies, commonly involved in stress response [210]. Interestingly, non-coding regions of miRNAs may also be involved in stress-responsive regulatory switches, such as the VEGFA mRNA 3′-UTR that integrates signals from interferon-gamma and hypoxia to regulate VEGFA translation though conformational change in mammalian cells—a mechanism analogous to riboswitches described in bacteria, fungi and plants [315].

This predisposition of RNAs to be used for regulatory processes is indicated by convergences in gene regulation and stress responses through analogous processes or mechanisms in different species, such as B2 and Alu SINE RNAs in mouse and humans [316] and hsr-0 and Sat III RNAs in flies and human cells [206], as well as in the small RNA-guided genome defence mechanisms in prokaryotes and eukaryotes.

**RNAs in the arms race and co-evolution**

As also shown above, the regulatory properties of RNAs have been widely explored in conditions of ecological pressure, including in the fast evolving arms race between host and pathogens. These mechanisms reflect the unique ability of RNA to manipulate gene expression and can have significant implications. Indeed, it is emerging that organisms can more broadly deploy regulatory RNAs to interfere with each other and in different ecological contexts. For instance, it has been observed that two well-known RNAs involved in stress response in *Escherichia coli* (DsrA and OxyS, mentioned above) can affect gene expression in *C. elegans* worms that feed on the bacteria [317]. The bacterial sRNAs have a direct impact on *C. elegans* physiology by affecting specific genes that control chemosensory behaviour and metabolism, remarkably interfering with the longevity of the worm. The authors suggested that the stress-induced ncRNAs in *E. coli* have a role in protecting them from overfeeding by the worm, with clear indications of the broad ecological roles of ncRNAs and their implications for symbiotic interactions and co-evolution.

The interference of parasites with host expression may also extend to humans, as suggested by the viral RNAs discussed above and other examples, such as the general downregulation of human macrophage expression of several RNAP III ncRNAs (including Alu and B1 RNAs, as well as the SRP signal recognition particle RNA) by *Leishmania* [318]. On the other hand, given the involvement of RNAs in host defence, it is also likely that cellular regulatory RNAs play an important role in shaping the
evolution of parasites. Indeed, the RNA systems that are at the base of cellular defence in all organisms, including (but not limiting to) the CRISPRs in prokaryotes and piRNAs in animals used as examples of adaptive immunity system against transposons, plasmids and viruses, have evolutionary importance as mechanisms driving genomic diversity and speciation [319].

**Intersection with epigenetic mechanisms**

While differential regulation of gene expression through epigenetic mechanisms is firmly established as central to differentiation and development, it is now increasingly clear that it is also involved in phylogenetic diversification and likely extensively employs ncRNAs [16,18,320].

Interestingly, while large numbers of orthologous ncRNAs can be identified in broad phylogenetic groups, they often have high rates of turnover, with potential important impacts on gene expression even among more closely related species [22,33]. Moreover, not only RNAs themselves can be carriers of epigenetic information (including through stress- and environmentally responsive RNA editing and modification), even through the germline, they are also intimately implicated in the regulation of chromatin-based epigenetic processes [3,321–323]. These include central ontogenetic ones, such as dosage compensation in mammals, despite the high plasticity of key RNAs involved [324]. On the other hand, developmental transcription factors and chromatin regulators are usually highly conserved, including complexes and domains that are associated with RNA regulation, such as Trithorax, Polycomb and chromodomain proteins [3]. Therefore, it is plausible that RNAs that regulate chromatin complexes such as Polycomb may have general and significant impacts in lineage-specific gene regulation.

Relevant examples of regulatory RNAs in epigenetic and evolutionary processes are already emerging. Small RNAs (usually transposon derived) in plants are implicated in epigenetic processes that underpin phenotypic differences between *Arabidopsis* ecotypes [325], as well as in hybrid vigour and compatibility in different species [326]. Remarkably, in fruit flies, transposon-derived piRNAs, which are important epigenetic regulators [239], are also centrally involved in the process of hybrid compatibility and dysgenesis [327,328]. Moreover, small regulatory RNAs, including piRNAs and siRNAs, are closely implicated in transgenerational epigenetic inheritance in both plants and animals [328–330], which may represent a widespread adaptive phenomenon, including in stresses and defence against aggressions and parasites [240,331–334].

Given the strong connections of epigenetic mechanisms and responses to the environment, including to stresses [335], the investigations of the roles of regulatory RNAs and the possible evolutionary implications are warranted. Moreover, transposons, whose transcription and activity are developmentally regulated and strongly responsive to stresses (such as heat shock, DNA damage, oxidative stress and viral infection) [292], can be the source or integrate the sequence of regulatory small and long RNAs ([336,337] and see previous examples) and have large impact not only in genome evolution but also in several epigenetic regulatory effects [338]. As these links between regulatory RNAs and molecular phenomena recently associated with evolutionary processes, including epigenetic regulation, stress response [339,340] and transposition [292], grow increasingly stronger, it will become fundamental to consider the centrality of RNA regulation in these processes.

We expect that similar processes and others that use the evolutionary plasticity and functional versatility of regulatory RNAs have been and continue to be exploited in the adaptive evolution of complex organisms.

**Key Points**

- ncRNAs play diverse roles in prokaryote and eukaryote physiological control.
- These roles are strongly implicated in response to diverse cellular and organismal stresses and are often disrupted in disease.
- RNAs are extensively used as regulatory tools in the interactions and arms race between host and pathogens.
- The regulatory properties and versatility of RNAs, as well as their participation in epigenetic regulation and stress response, have important evolutionary implications for adaptation and diversification.

**FUNDING**

P.P.A. is supported by a Royal Society Newton International Fellowship and is a Research Fellow at Corpus Christi College, University of Cambridge. M.E.D. and J.S.M. are supported by a Career Development Award and an Australia Fellowship, respectively, from the National Health and Medical Research Council of Australia.
References


56. Watanabe Y, Yamamoto M.

57. Chen HM, Neiman AM. A conserved regulatory role for

49. Nagano T, Mitchell JA, Sanz LA,

47. Latos PA, Pauler FM, Koerner MV,

45. Guil S, Esteller M. Cis-acting noncoding RNAs: friends and

Rinn JL, Kertesz M, Wang JK,

et al

S. pombe

et al


274

Amaral et al.


RNA in homeostasis, disease and stress responses


274. Pawlicki JM, Steitz JA. Primary microRNA transcript re-
279. Rossetto CC, Pari G. KSHV PAN RNA associates with
273. Zhong W, Ganem D. Characterization of ribonucleo-
278. Chandriani S, Ganem D. Array-based transcript profiling
275. Borah S, Darricarrere N, Darnell A,
272. Sun R, Lin SF, Gradoville L,
270. Zhang G, Raghavan B, Kotur M,
271. Reeves MB, Davies AA, McSharry BP,
276. Rossetto CC, Tarrant-Elorza M, Verma S,
269. Pijlman GP, Funk A, Kondratieva N,
278. McKenna SA, Kim I, Liu CW,
277. Zhang G, Raghavan B, Kotur M,
283. Thimmappaya B, Weinberger C, Schneider RJ,
281. Xu N, Segerman B, Zhou X,
282. Clarke PA, Schwemmle M, Schickinger J,
277. Avrova AO, Whisson SC, Pritchard L,
295. Gripenland J, Netterling S, Loh E,
298. Matrajt M. Non-coding RNA in apicomplexan parasites.
294. Wurtzel O, Sesto N, Mellin JR,
293. Zhao J, Sinclair J, Houghton J,
295. Cahoon LA, Seifert HS. Transcription of a cis-acting, non-
292. Patil V, Lescault PJ, Lirussi D,
291. Clarke PA, Schwemmle M, Schickinger J, et al. Binding of
287. Greenaway PJ, Wilkinson GW. Nucleotide sequence of the most abundantly transcribed early gene of human cyto-
296. Cahoon LA, Seifert HS. Transcription of a cis-acting, non-
297. Avrova AO, Whisson SC, Pritchard L, et al. A novel non-
299. Oliveira KC, Carvalho ML, Maracaju-Coutinho V, et al. Non-


