Original citation:

Permanent WRAP URL:
http://wrap.warwick.ac.uk/80135

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work of researchers of the University of Warwick available open access under the following conditions.

This article is made available under the Creative Commons Attribution 3.0 (CC BY 3.0) license and may be reused according to the conditions of the license. For more details see: http://creativecommons.org/licenses/by/3.0/

A note on versions:
The version presented in WRAP is the published version, or, version of record, and may be cited as it appears here.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk
CncRNAs: RNAs with both coding and non-coding roles in development

Karuna Sampath1,∗ and Anne Ephrussi2,∗

ABSTRACT
RNAs are known to regulate diverse biological processes, either as protein-encoding molecules or as non-coding RNAs. However, a third class that comprises RNAs endowed with both protein coding and non-coding functions has recently emerged. Such bi-functional ‘coding and non-coding RNAs’ (cncRNAs) have been shown to play important roles in distinct developmental processes in plants and animals. Here, we discuss key examples of cncRNAs and review their roles, regulation and mechanisms of action during development.

KEY WORDS: Bifunctional RNA, CncRNAs, Coding RNA, Development, Non-coding RNA, Regulatory RNA

Introduction
Ribonucleic acids (RNAs) serve important roles, most notably as intermediaries in the flow of genetic information from DNA to proteins. They can be found in various forms, for example as protein-coding messenger RNAs, as recruiters and machines for protein synthesis (e.g. transfer RNAs, ribosomal RNAs), as modifiers of ribosomal RNAs, and as regulators of RNA splicing, RNA stability and protein synthesis (e.g. small nuclear RNAs, small nucleolar RNAs and microRNAs). Other small RNAs function in epigenetic regulation and post-transcriptional gene silencing, and protect the genome from transposons (e.g. piwi-interacting RNAs). In addition to these well-studied groups of RNAs, transcriptomic analyses in a variety of organisms have identified hundreds of other RNAs (e.g. long non-coding RNAs) that are thought to function in numerous cellular and developmental processes (Bushati and Cohen, 2007; Houwing et al., 2007; Mercer et al., 2009; Pauli et al., 2011; Rinn and Chang, 2013; Weick and Miska, 2014; Hezroni et al., 2015).

Until recently, most RNAs were presumed to be exclusively protein coding or non-coding. However, studies of many bacteria, animals and plants have revealed an unusual group of RNAs that have both protein coding and non-coding roles. These RNAs have been referred to as ‘dual function’ or ‘bi-functional’ RNAs (Dinger et al., 2008; Uviling et al., 2011b); we refer to them hereafter as ‘coding and non-coding RNAs’ (cncRNAs) (Kumari and Sampath, 2015). The identification of such cncRNAs raises several interesting questions and poses a challenge to RNA classification. Here, we summarize the known features of a few exemplary cncRNAs and discuss their roles and regulation during plant and animal development.

1Division of Biomedical Sciences, Warwick Medical School, University of Warwick, Coventry CV4 7AJ, UK. 2Developmental Biology Unit, European Molecular Biology Laboratory, Meyerhofstraße 1, Heidelberg 69117, Germany.

∗Authors for correspondence (K.Sampath@warwick.ac.uk; ephrussi@embl.de)

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

1234
The non-coding activity of *osk* could be mediated either by sequestering an oogenesis inhibitory factor, or by a scaffolding function for a ribonucleoprotein (RNP) that assembles on *osk* RNA and promotes oogenesis. Evidence that the *osk* 3′UTR acts via sequestration was provided by mutations affecting Bru, which binds to Osk; mis-sense mutations in *bru* that reduce Bru levels suppress the *osk* null mutant phenotype, and this effect is even stronger with *bru* nonsense mutant alleles. However, mutations in regions of *osk* RNA that are not involved in Bru binding also affect egg-laying, indicating that additional unknown factors contribute to the non-coding functions of *osk* (Kanke et al., 2015).

**Squint**

The zebrafish protein Squint (Sqt; Nodal-related 1 – Zebrafish Information Network) is a secreted signaling morphogen that is essential for formation of the embryonic organizer and for specification of mesendoderm during gastrulation (Erter et al., 1998; Feldman et al., 1998; Rebagliati et al., 1998; Chen and Schier, 2001). However, maternal *sqt* RNA has an earlier non-coding function, independent of Sqt/Nodal signaling, during the formation of the zebrafish dorsal axis (Gore et al., 2005, 2007; Bennett et al., 2007; Lim et al., 2012). This non-coding activity was identified by overexpression and antisense knockdown studies: overexpression of mutant zebrafish *sqt* RNA that is incapable of encoding functional Sqt/Nodal protein expands the expression domains of dorsal progenitor genes such as *goosecoid* and *chordin*. Furthermore, the dorsal-inducing activity of *sqt* RNA does not require the Sqt-coding exons, and the 3′UTR of *sqt* RNA is sufficient for its dorsalizing activity.

It was also found that targeting maternal *sqt* RNA with antisense morpholinos that reproduce zygotic *sqt* mutant phenotypes causes mislocalization of *sqt* RNA and the complete loss of embryonic dorsal structures. This is in contrast to insertion mutations in the *sqt* locus that disrupt Sqt protein coding sequences but do not affect *sqt* RNA (Feldman et al., 1998; Feldman and Stemple, 2001; Aoki et al., 2002; Amsterdam et al., 2004; Gore et al., 2005; Lim et al., 2012). The precise molecular mechanism by which non-coding *sqt* RNA elicits its role is not known. Systematic analysis of mutants lacking *sqt* RNA can help address this question (Lim et al., 2013). Interestingly, the UTR-mediated non-coding activity of *sqt* RNA is independent of microRNA (miRNA) target sequences and the function of the miRNA processing gene *dicer* (*dicer1* – Zebrafish Information Network) (Lim et al., 2012). Therefore, binding to miRNA negative regulators of dorsal identity is unlikely to be the mode of action of non-coding *sqt* RNA. However, a requirement for the Wnt/β-Catenin pathway raises the possibility that *sqt* RNA might act as a scaffold for a factor that promotes nuclear accumulation of β-Catenin in dorsal progenitors. Alternatively, non-coding *sqt* RNA might sequester a negative regulator of dorsal identity.

**VegT**

VegT is a T-box family transcription factor that functions in mesoderm and endoderm formation during gastrulation in *Xenopus laevis* embryos (Zhang et al., 1998; Heasman et al., 2001; Kloc et al., 2005). However, in *Xenopus laevis* oocytes, vegT RNA plays an additional role in the organization of cytotkeratin filaments and germinal granules. Accordingly, depletion of vegT RNA in *Xenopus* oocytes by phosphorothioate antisense oligonucleotides results in disruption of the cytotkeratin network, mislocalization of maternal RNAs, and blocked formation of germinal granules (Heasman et al., 2001; Kloc et al., 2005). Exogenously provided vegT RNA was found to reconstitute and rescue the disrupted cytotkeratin network (Kloc et al., 2005). Based on these lines of evidence, vegT RNA is thought to be a structural matrix that stabilizes cytotkeratin filaments and anchors vegetal RNAs in frog eggs.

**Steroid receptor activator RNA**

The mammalian steroid receptor RNA activator (SRA; also known as SRA1) promotes myogenic differentiation and the conversion of non-muscle cells to myocytes by co-activation of the myogenic differentiation factor MYOD (also known as MYOD1). In addition to MYOD, SRA transcripts are believed to co-activate numerous nuclear receptors, and potentially regulate the proliferation and differentiation of a variety of cell types. SRA RNA has been found associated with the PRC2 Polycomb group and TrxG trithorax
transcriptional repressor complexes, and it directly interacts with the stem cell pluripotency factor NANOG (Wongtrakooongate et al., 2015). By contrast, SRA protein (SRAP) prevents SRA RNA-mediated co-activation and differentiation (Lanz et al., 1999; Hube et al., 2011), and is thought to act as a trans-activator of steroid hormone receptors in mammalian cell culture assays.

The activity of SRA RNA as a trans-activator of steroid hormone receptors and as a co-activator of MYOD was identified by studying SRA transcripts lacking a methionine initiation codon. Initial attempts to identify SRA transcripts with an extended 5’ sequence were unsuccessful, leading to the hypothesis that SRA functions as a non-coding RNA in this context. Subsequently, mutations in SRA revealed that its co-activator function resides in the transcript (Lanz et al., 2002). It was also shown that introducing multiple stop codons in SRA RNA did not abolish steroid receptor co-activation, and RNA activity was detected even in the presence of the translation inhibitor cycloheximide. Together, these findings suggested that SRA RNA has functions independent of SRAP (Lanz et al., 1999).

A recent study reporting the crystal structure of human SRAP suggests that the protein does not harbor a predicted RNA-recognition motif (RRM) but rather resembles the spliceosome complex protein PRP18 (also known as PRPF18) (McKay et al., 2014). Biochemical binding assays performed in vitro did not find a specific interaction between SRAP and SRA, nor was a specific response on estrogen receptor targets observed in cell culture assays. These observations have led to the proposal of an alternative model wherein SRAP is thought to stabilize internuclear interactions within a nuclear splicing complex, rather than directly binding and regulating SRA RNA (McKay et al., 2014). Genetic analysis of the sra locus with mutations that disrupt specific SRA RNA versus SRAP domains, which could resolve the roles and mechanisms by which the RNA and protein function during development, differentiation and homeostasis, is currently lacking.

p53
The tumor suppressor protein p53 (TP53) plays crucial roles in cell cycle regulation and in preventing mutations arising from DNA damage in the genome (Lane and Crawford, 1979). However, in addition to the well-known roles of p53 in protecting the genome, p53 mRNA was found to regulate the ubiquitin ligase MDM2, which is a negative regulator of p53 (Candeias et al., 2008). p53 mRNA directly interacts with the N-terminus of MDM2 to prevent its E3 ubiquitin ligase activity and thereby controls MDM2-mediated regulation of p53. Interfering with the p53 mRNA-MDM2 interaction following DNA damage prevents p53 stabilization and activation (Gajjar et al., 2012). Thus, p53 mRNA-MDM2 interactions are key for the genotoxic stress response. Interestingly, the sequences of p53 mRNA that interact with the N-terminus of MDM2 protein also encode the amino acids in TP53 that interact with and are poly-ubiquitylated by the MDM2 RING domain (Naski et al., 2009). Silent point mutations in this region of p53 weaken its interactions with MDM2, and reduce TP53 activity. This suggests that structural elements in p53 mRNA might harbor its non-coding activity.

IRS1
Insulin receptor substrate 1 (IRS1) is a major substrate and cytoplasmic docking protein for the insulin receptor and insulin-like growth factor receptor. IRS1 is thought to be an effector of insulin signaling, with roles in cell growth and proliferation. Deletion of Irs1 by conventional knockout strategies in mice led to growth retardation and compensated insulin resistance (Kido et al., 2000). IRS1 levels are generally low or absent in differentiating cells, and elevated IRS1 levels have been associated with cancers in mice and humans. Interestingly, a recent study found that Irs1 mRNA has a function in myoblasts that is independent of IRS1 protein (Nagano et al., 2015). The 5’UTR of Irs1 mRNA harbors sequences complementary to RNA encoding the cell cycle regulator retinoblastoma (RB; RB1). Overexpression of the 5’UTR region of Irs1 led to reduced Rb mRNA expression, whereas knockdown of Irs1 mRNA harboring the 5’UTR complementarity region led to increased Rb mRNA levels and enhanced myoblast differentiation. These findings suggest that Irs1 RNA has a novel role as a regulatory RNA that is independent of IRS1 protein (Nagano et al., 2015). It is not known whether Irs1 mRNA regulation is restricted to myoblasts, and what controls Irs1 RNA versus IRS1 protein functions is also unknown.

UBE3A
Finally, a recent study reports that transcripts encoding the E3 ubiquitin ligase UBE3A have a non-coding role in dendrite growth and spine maturation in hippocampal neurons; by contrast, UBE3A protein plays crucial roles in activity-dependent synapse development, plasticity, learning and memory (Sun et al., 2015; Valluy et al., 2015). This role was discovered using antisense short hairpin RNAs (shRNAs) that target Ube3a 3’ UTR variants in rat hippocampal neurons: the neurons showed significantly increased dendrite growth and complexity upon knockdown of the non-translated and intron-retaining Ube3a1 RNA, whereas shRNAs targeting the spliced coding Ube3a2/3 isoforms reduced dendrite complexity (Valluy et al., 2015). Antisense oligonucleotides can have off-target effects or, alternatively, they can uncover novel mechanisms not identified by protein-disrupting mutants. Therefore, it is crucial that the non-coding activities of ncRNAs identified by antisense approaches (e.g. shRNA, antisense morpholinos, phosphorothioate oligonucleotides) are independently validated by assays such as overexpression of non-translatable RNA, and by analysis of mutations that specifically disrupt the RNA (Lim et al., 2012, 2013; Kok et al., 2015; Rossi et al., 2015).

Plant ncRNAs
A number of ncRNAs have also been identified in plants (Table 1). Transcripts of the gene early nodulin 40 (enod40) in the legume Medicago trunculata contain two short open reading frames (ORFs), the products of which are required for cortical cell divisions in root cells (Yang et al., 1993; Crespi et al., 1994). A region of RNA secondary structure separates the two enod40 ORFs. This RNA segment is essential for enod40 activity and has a non-coding role in root nodule formation (Girard et al., 2003; Campalans et al., 2004). It was also shown that, in alfalfa, enod40 RNA is essential for a growth response in the root cortex (Sousa et al., 2001).

In soybean plants, ENOD40 peptides have been shown to regulate the turnover of the enzyme sucrose synthase (SUC1), which functions in sugar metabolism in roots and cotyledons. By using a combination of RNA structure prediction, comparison and structure probing, various regions of soybean enod40 RNA were identified to be key for root nodule formation. Of these, five domains are conserved amongst leguminous plants and are presumed to be required for the non-coding activity of enod40 RNA (Girard et al., 2003). Indeed, the deletion of an inter-ORF RNA region with predicted structure resulted in reduced activity of alfalfa enod40
without affecting the production of ENOD40 peptides (Sousa et al., 2001). Analyses in Arabidopsis thaliana and rice have identified RNAs similar to enod40 suggesting that cncRNAs exist in these plants as well (Kouchi et al., 1999). Furthermore, at least 50 miRNAs in the Arabidopsis transcriptome are predicted to encode short peptides (microRNA-encoded peptides; miPEPs) that appear to regulate the transcripts of primary transcript (pri-)miRNA. For example, precursor (pre-)miRNA for Medicago trunculata mir171b encodes a short peptide expression of which leads to transcriptional upregulation of the corresponding pri-miRNA, which in turn controls target genes involved in root development (Lauressergues et al., 2015). Interestingly, all the identified miPEPs were found to be conserved across flowering plants and are associated with ancient miRNA families. It is not known if miRNA-encoded peptides are present in animals.

cncRNA functions and mechanisms of action
The precise function and molecular mechanism of action is known for only a few cncRNAs. Nonetheless, it is emerging that all the mechanisms deployed by conventional non-coding RNAs are also represented amongst this hybrid class of RNAs.

Base pairing and roles as decoy and regulatory RNAs
Base pairing is a fundamental property of nucleic acids that is essential for processes such as codon-anticodon recognition during protein synthesis and microRNA recognition of target RNA sequences. Nucleotide complementarity also forms the basis for RNAs that function as ‘molecular sponges’, ‘decoys’ or ‘target mimics’, such as competing endogenous RNAs (ceRNAs) that share miRNA recognition sequences. ceRNAs bind to complementary sequences in miRNAs and prevent interactions between miRNAs and their bona fide targets (Franco-Zorrilla et al., 2007). For instance, the pseudogene PTENP1 harbors sequences in its 3′UTR that are complementary to miRNAs that target and repress PTEN mRNA (Poliseno et al., 2010). Indeed, cncRNAs can also function by base pairing with target RNAs and behave as decoys (Fig. 1). For example, non-coding Ube3a1 mRNA retains an intron and is thought to act as a decoy or molecular sponge for miRNA-134 (MIR134) that would otherwise target the spliced Ube3a2/3 transcripts and downregulate UBE3A ubiquitin ligase protein expression in dendrites (Vally et al., 2015).

Some cncRNAs appear to target and regulate other mRNA sequences directly (Fig. 2). For example, the 5′UTR of mammalian Irs1 RNA can base pair with Rb mRNA. The overexpression of Irs1 RNA, which contains a sequence element complementary to Rb mRNA, reduces RB levels by a mechanism independent of Dicer or Upf1 (i.e. independently of miRNAs or nonsense-mediated decay), and suppresses the differentiation of cultured skeletal muscle cells (Nagano et al., 2015). Interestingly, some bacterial small regulatory RNAs (sRNAs) also function in a similar manner. For example, Escherichia coli SgrS RNA regulates the glucose-phosphate stress response by base pairing with and blocking the translation of PtsG mRNA, which encodes a sugar phosphate transporter. SgrS is a cncRNA as it also encodes a short 43 amino acid peptide, SgrT, which inhibits the activity of the PtsG transporter protein (Wadler and Vanderpool, 2007; Rice and Vanderpool, 2011).

Structural roles, sequestration and scaffolding
Structural features in RNAs play crucial roles in their activity. RNA structural elements can either sequester or bind and deliver protein complexes (Figs 3 and 4). SRa RNA co-activation of MYOD, for example, is mediated via interactions between SRa RNA and the p68/p72 RNA helicases (Caretti et al., 2006). A sequestration and scaffolding function has also been proposed for osk 3′UTR sequences during early oogenesis, via their binding to Bru and some unknown factors (Kanke et al., 2015). In Xenopus eggs, vegt RNA forms aggregates that colocalize with cytoplasmic filaments, and the depletion of vegt destabilizes the cytoplasmic network and disrupts the anchoring of vegetal RNAs (Heasman et al., 2001; Kloc et al., 2005), suggesting that vegt RNA also plays a scaffolding or sequestering role. Finally, it was shown that the RNA sequence separating the two ORFs within enod40 RNA harbors a region of highly stable secondary structure that binds to the RNA-binding protein Mr1BP1, which is presumed to be a translational regulator (Crespi et al., 1994; Sousa et al., 2001; Campalans et al., 2004). Structural elements in RNAs can also act as sensors for environmental stimuli, as observed in some bacterial RNAs: these RNAs function as inactive nascent transcripts with secondary structures that are released upon binding of the RNA to ligands or metabolites, or upon increase in temperature, leading to activation of downstream gene expression. One such example is the S-adenosylmethionone-sensing S-box riboswitch in Bacillus subtilis (Henkin, 2008; Gottesman and Storz, 2011). Structural RNA sensors could facilitate rapid changes in gene expression in response to external signals during developmental transitions (Fig. 5).

Feedback regulation
Some cncRNAs mediate feedback regulation of the same pathway or process in which their protein product functions. This type of feedback is exemplified by the p53-MDM2 pathway (Candeias et al., 2008). A region of p53 mRNA binds to and regulates the E3 ubiquitin ligase and p53 tumor protein regulator MDM2, which inhibits p53 activity by controlling p53 translation, polyubiquitylation and degradation. The binding of p53 mRNA to MDM2 leads to accumulation of MDM2 at polysomes, stimulation of p53 synthesis, and inhibition of the E3 ligase activity of MDM2. Accordingly, mutations in p53 that reduce the affinity of p53 mRNA for MDM2 enhance the suppression of p53 activity by MDM2. Thus, p53 mRNA acts as a feedback switch that controls MDM2-mediated regulation of p53 by directly interacting with MDM2 (Candeias et al., 2008). Such regulation is also seen in the context of...
Fig. 2. cncRNAs as regulatory RNAs. Complementary sequence elements in cncRNAs can bind directly to other mRNAs and control their expression. In the example shown, the coding cncRNA is transcribed from one transcription start site (orange circle) and translated (left; yellow chain), whereas the non-coding cncRNA (right) is generated from an alternative transcription start site (magenta circle) such that the 5' UTR now includes a region of complementarity (red comb) to the target mRNA. Ribosomes are shown in green.

The regulation of coding versus non-coding roles

A unique feature of cncRNAs that sets them apart from other RNA classes is the exquisite regulation and partitioning of their coding and non-coding functions. This can be in the form of: (1) temporal separation of the coding and non-coding activities to different developmental stages; (2) spatial segregation of the coding and non-coding roles to distinct subcellular or cellular domains; or (3) activation of one function under particular physiological/environmental conditions via specific RNA elements.

How might segregation of the coding and non-coding functions be achieved? The analysis of some cncRNAs suggests that such partitioning requires extensive transcriptional and post-transcriptional regulation of the RNAs. For instance, two major SRA isoforms of 224 or 236 amino acids, respectively, are generated by an additional upstream exon in SRA RNA that contains two initiating methionine residues, from alternative transcriptional start sites. Alternative splicing can also generate different RNA isoforms, some of which may retain the coding activity whereas others might function as non-coding RNAs. This has been observed for SRA RNA and UBE3 RNA, whereby alternative splicing leads to retention of an intron and disruption of the ORF, generating the non-coding version of these transcripts (Lanz et al., 1999; Hube et al., 2011; Valuy et al., 2015). Thus, in addition to increasing the protein-coding capacity of genomes by generating peptide isoforms, alternative splicing can also generate a variety of non-coding RNA isoforms. In addition, isoform abundance can change as development and differentiation progress, and this can influence cell fate specification. For example, the ratio of coding and non-coding SRA RNA isoforms changes during muscle cell differentiation, with myotubes expressing two to five times higher levels of non-coding SRA RNA in comparison with myoblasts. This balance between the non-coding and coding SRA RNA isoforms influences MYOD activity and myogenic differentiation (Hube et al., 2011).

Temporal segregation of the coding and non-coding activities of cncRNAs is also evident, and has been found for osk, vegt and sqt RNAs, with the non-coding activity detected at earlier developmental stages: during early oogenesis for osk, in oocytes for vegt, and in early embryos for sqt. Post-transcriptional mechanisms, such as regulated polyadenylation, splicing and translational regulation, enable this precise temporal control of RNA activity. An example of a cncRNA that undergoes extensive post-transcriptional regulation is sqt. Unprocessed sqt pre-mRNA is found in zebrafish eggs and early embryos, whereas processed sqt RNA that is poly-adenylated and spliced is detected later, from the 16-cell stage (Gore et al., 2007; Lim et al., 2012; Kumari et al., 2013). In addition, maternal sqt RNA is translationally repressed in early zebrafish embryos by the RNA-binding protein Ybx1, which interacts with the sqt 3' UTR and the eIF4E translation initiation complex. The signaling activity of Sqt/Nodal protein is detected only from the 256-cell stage (Kumari et al., 2013), suggesting that the non-coding and coding activities of sqt are temporally segregated by post-transcriptional regulation of the RNA. In Drosophila oocytes, poly-adenylation of osk RNA also stimulates Osk protein translation and is regulated by the Orb and cytoplasmic polyadenylation element binding (CPEB) proteins (Castagnetti and Ephrussi, 2003). By contrast, translation of Osk protein is repressed during early oogenesis by Bru, which binds to BREs in the osk 3' UTR sequences and recruits Cup, an eIF4E-binding protein (Nakamura et al., 2004; Chekulaeva et al., 2006). Remarkably, one cluster of BREs also mediates translational activation, and both BRE-dependent repression and activation can occur in trans, presumably by co-assembly of osk mRNAs in cytoplasmic complexes (Reveal et al., 2010).

Spatial restriction of cncRNA activity can be achieved by the localization of RNA to distinct cellular or subcellular domains, as observed for sqt, osk and vegt. Maternal sqt RNA localizes to dorsal progenitor cells by the four-cell stage in zebrafish embryos (Gore et al., 2005). Sequences in the sqt 3' UTR and an intact microtubule cytoskeleton are required for localization of maternal sqt RNA in early embryos where it carries out its non-coding role in dorsal axis formation (Gore and Sampath, 2002; Gore et al., 2005; Gilligan et al., 2011; Lim et al., 2012). Interestingly, the sqt dorsal localization element (DLE) overlaps with the region of the sqt 3' UTR that is required for Ybx1-binding, ensuring that the coding...
Activities.

Translational control of the cncRNA regulates the coding versus non-coding activity (NCE) are distinct from elements required for RNA localization (LE).

The elements that carry out the non-coding activity (cncRNP granules) to specific subcellular locations via cytoskeletal transport tracks (green line). The elements that carry out the non-coding activity (NCE) are distinct from elements required for RNA localization (LE). Translational control of the cncRNA regulates the coding versus non-coding activities.

Potential of localized maternal sqt is shut off (Gilligan et al., 2011; Kumari et al., 2013). Drosophila osk RNA is localized at the posterior pole of the oocyte; this localization depends on a secondary structure formed from exonic sequences in the coding region upon splicing and on Osk protein itself (Ghosh et al., 2012; Ephrussi et al., 1991; Markussen et al., 1995; Rongo et al., 1995).

VegT RNA also shows a distinct localization pattern, being localized to the vegetal pole of Xenopus oocytes in a manner that depends upon sequences in the 3′ UTR and the protein Igf2bp3 (also known as Vg1-RBP and Vera) (Zhang and King, 1996; Bubunenko et al., 2002; Kwon et al., 2002).

Specific regions or elements of RNA also appear to play crucial roles in determining coding versus non-coding functions. For example, the non-coding activity of sqt and osk resides in elements within the 3′ UTR of these RNAs (Jenny et al., 2006; Lim et al., 2012), whereas an RNA element located between the two ORFs of enod40 harbors its non-coding activity (Girard et al., 2003). In each of these RNAs, the coding and non-coding roles can be clearly ascribed to distinct RNA segments. However, the non-coding activity of other cncRNAs appears not to be restricted to a discrete RNA region, but, instead, is intermingled with coding sequences and dispersed throughout the transcript, as in the case of SRA RNA (Lanz et al., 2002). In such RNAs, the protein-coding role may rely upon the primary sequence whereas RNA secondary/tertiary structure could engender non-coding activity. Distinguishing the coding and non-coding activities of such RNAs can be challenging.

Perspectives

The examples discussed above demonstrate that cncRNAs are emerging as key regulators of distinct developmental processes in animals and plants. It is noteworthy that all the known cncRNAs in multicellular organisms have been found in cells that are plastic, respond rapidly to their environment, and function in developmental processes (e.g. oocytes/early embryos, neurons in animals, root cells in plants). A recent study found that a significant proportion of non-coding RNAs are evolutionarily conserved and expressed in early embryos, suggesting that they might be involved in developmental processes (Necsulea et al., 2014). These observations lead to some key questions. For example, do cncRNAs represent an evolutionary link between non-coding and coding RNAs, or are they a more recent, derived group that arose independently in various clades?

Which function of cncRNAs emerged first – non-coding or coding? Did these RNAs acquire new functions whilst retaining their original role? Interestingly, the non-coding RNA Xist, which mediates X-chromosome inactivation in placental mammals, is proposed to have evolved by pseudogenization of an ancestral protein-coding mRNA and loss of protein function (Duret et al., 2006), which is distinct from cncRNA genes that have either retained or acquired both coding and non-coding roles. Furthermore, it is known that riboswitches and sRNAs in bacteria act as sensors that regulate gene expression under specific environmental conditions and control key processes such as vesicle versus dormant spore formation, or motile versus sessile biofilm states (Horler and Vanderpool, 2009). This suggests that cncRNAs constitute an ancient RNA class. It is not known how many cncRNAs are present in bacterial or other genomes. Predictions based upon splice variants suggest that ~300 cncRNAs exist in the human transcriptome (Dinger et al., 2008; Ulveling et al., 2011a). However, this is likely to be an underestimate, because, in addition to intron retention and alternative splicing, other post-transcriptional mechanisms (e.g. differential poly-adenylation, RNA modifications or decorations) and differential RNA structures can also bring about cncRNA activity.
To identify novel cncRNAs and determine their functions, systematic analyses of mutants that specifically disrupt the transcript versus those that affect the protein-coding capacity of genes needs to be performed. Determining the features shared by dual function RNAs (similar to those found in miRNAs and long intergenic non-coding RNAs) will also facilitate the identification of novel cncRNAs. Some of these characteristics might be spatially or temporally regulated, exemplified by the presence of structural features at specific developmental stages, in certain cell types, or under particular physiological and environmental conditions. For most cncRNAs, how the features or activities of the RNAs (coding and non-coding) are regulated, and how the RNAs switch from one role to the other is largely unclear. Teasing apart the coding and non-coding activities of cncRNAs and determining how they are regulated will be challenging, but should be facilitated by mutagenesis with conventional as well as new genome editing methods.

Recent work has shown that some metabolic enzymes also have ‘moonlighting’ roles, functioning in their normal, well-characterized capacity but also in an unrelated and unexpected role. For example, in addition to its well-established function in glycolysis, the enzyme glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was recently shown to have RNA-binding activity (Castello et al., 2012). In summary, these findings highlight that transcriptomes and genomes are far more complex than previously appreciated. Understanding the functions and mechanisms of action of cncRNAs will provide new insights into gene regulation. Finally, although the analysis of human diseases focuses primarily on the protein-coding capacity of the genome, it is plausible that mutations that affect non-coding functions of such cncRNAs could lead to disease states. Genome-wide analyses across phyla and throughout developmental stages, together with functional validation by classical mutagenesis, novel genome editing and RNA interrogation methods will hopefully identify novel cncRNAs and the mechanisms by which they function during plant and animal development.

Acknowledgements
We thank the members of our labs and many colleagues for discussions; Aniket Gore and Pooja Kumari for suggestions to improve the manuscript; and Jonathan Millar for coining the term ‘cncRNA’.

Competing interests
The authors declare no competing or financial interests.

Funding
K.S. is supported by Warwick Medical School and the Biotechnology and Biological Sciences Research Council; and A.E. by the European Molecular Biology Laboratory.

References
Horler, R. S. P. and Vandergool, C. K. (2009). Homologs of the small RNA SgrS are broadly distributed in enteric bacteria but have diverged in size and sequence. Nucleic Acids Res. 37, 5465-5476.


