

Coding and Non-coding RNAs, the Frontier Has Never Been So Blurred

Florent Hubé, Claire Francastel

▶ **To cite this version:**

Florent Hubé, Claire Francastel. Coding and Non-coding RNAs, the Frontier Has Never Been So Blurred. Frontiers in Genetics, 2018, 9, pp.140. 10.3389/fgene.2018.00140 . hal-02127256

HAL Id: hal-02127256

https://u-paris.hal.science/hal-02127256

Submitted on 13 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

frontiers in Genetics

OPINION
published: 18 April 2018
doi: 10.3389/fgene.2018.00140

Coding and Non-coding RNAs, the Frontier Has Never Been So Blurred

Florent Hubé [1,2]* *and Claire Francastel* [1,2]*

[1] Université Paris Diderot, Sorbonne Paris Cité, Paris, France, [2] Epigénétique et Destin Cellulaire, Centre National de la Recherche Scientifique UMR7216, Paris, France

Keywords: messenger RNA, non-coding RNA, bifunctional RNA, transcript isoform, UCSC genes, Affymetrix U133 Plus2 array, NCI-60 panel

In molecular biology, one of the founding dogmas states that genetic information, stored in the form of DNA molecules in the majority of living organisms, is translated into proteins via transient intermediaries, the RNAs. However, over the past 50 years, an increasing number of studies have highlighted that RNAs have a more significant and broader role. Whereas all the genomes of living organisms, whether single-cell or multi-cellular, have a high proportion of loci that are transcribed into RNA, this transcript is not necessarily translated into protein but can perform functions within the cell as the form of RNA. As a consequence, RNAs translated into proteins were named messenger RNAs (mRNA) as opposed to those that do not, which are referred to as non-coding or regulatory RNAs (ncRNA), clearly discriminating RNAs according to their protein coding capacity. Until recently, the coding/non-coding distinction appeared to be obvious, although it became more blurred in recent years. The first hints emerged when it became evident that the repertoire of genome-encoded RNAs is far more extensive and complex than previously thought. Indeed, many RNAs, here referred to as bifunctional RNAs (bifRNA), have both regulatory/non-coding and coding functions. The balance between non-coding and coding RNAs levels is modulated depending on the stage of development or differentiation, environmental cues, disturbances caused by a pathogen, etc. After having re-defined this striking but expanding class of bifRNAs, we will propose an estimate of the fraction of the human transcriptome that they may represent.

OPEN ACCESS

Edited by:
Philipp Kapranov, Huaqiao University, China

Reviewed by:
Zofia Szweykowska-Kulinska, Adam Mickiewicz University in Poznań, Poland
Enrique M. Muro, Johannes Gutenberg-Universität Mainz, Germany

***Correspondence:**
Florent Hubé
florent.hube@univ-paris-diderot.fr
Claire Francastel
claire.francastel@univ-paris-diderot.fr

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *28 February 2018*
Accepted: *04 April 2018*
Published: *18 April 2018*

Citation:
Hubé F and Francastel C (2018) Coding and Non-coding RNAs, the Frontier Has Never Been So Blurred. Front. Genet. 9:140. doi: 10.3389/fgene.2018.00140

THE CONCEPT OF BIFUNCTIONAL RNAs

When we refer to bifRNAs, we assume that the same molecule is able to perform both functions, namely encode for a protein and possess its own regulatory function, but this term is often misused (**Figure 1A**). In fact, this is the case for a subset of them, such as SRA (Steroid Receptor RNA Activator), the pioneer member of this family [for review see (Ulveling et al., 2011b)]. Other examples have been described; SgrS (SuGar transport-Related sRNA) RNA partially inhibits glucose transporters mRNAs through base-pairing, and encodes a small polypeptide that prevents glucose transport (Maki et al., 2010; Vanderpool et al., 2011; Wadler and Vanderpool, 2007). The interaction between E3 ubiquitin ligase Mdm2 and p53 proteins usually drives p53 for degradation by the proteasome. Upon stress, the p53 mRNA was shown to also interact with the Mdm2 protein to both promote p53 synthesis and prevent Mdm2 from targeting p53 degradation (Candeias et al., 2008; Naski et al., 2009). A competition between translation and the structural function of RNA has therefore to be envisioned, although the way this competition is controlled remains rather obscure to date. However, one can imagine that sub-cellular localization is one important aspect to take into account, the translation being accomplished only in the cytoplasm whereas many ncRNAs are sequestered in the nucleus. Other bifRNAs have the extraordinary ability to separate the two functions in space and time, as it is the case for Oskar. The site of Oskar RNA and protein localization within the oocyte determines where germ cells form in the primordial embryo and where the abdomen develops (Rongo et al., 1995). This localization is controlled by a feedback loop in which Oskar mRNA is locally translated into a protein, which in turn maintains the localization

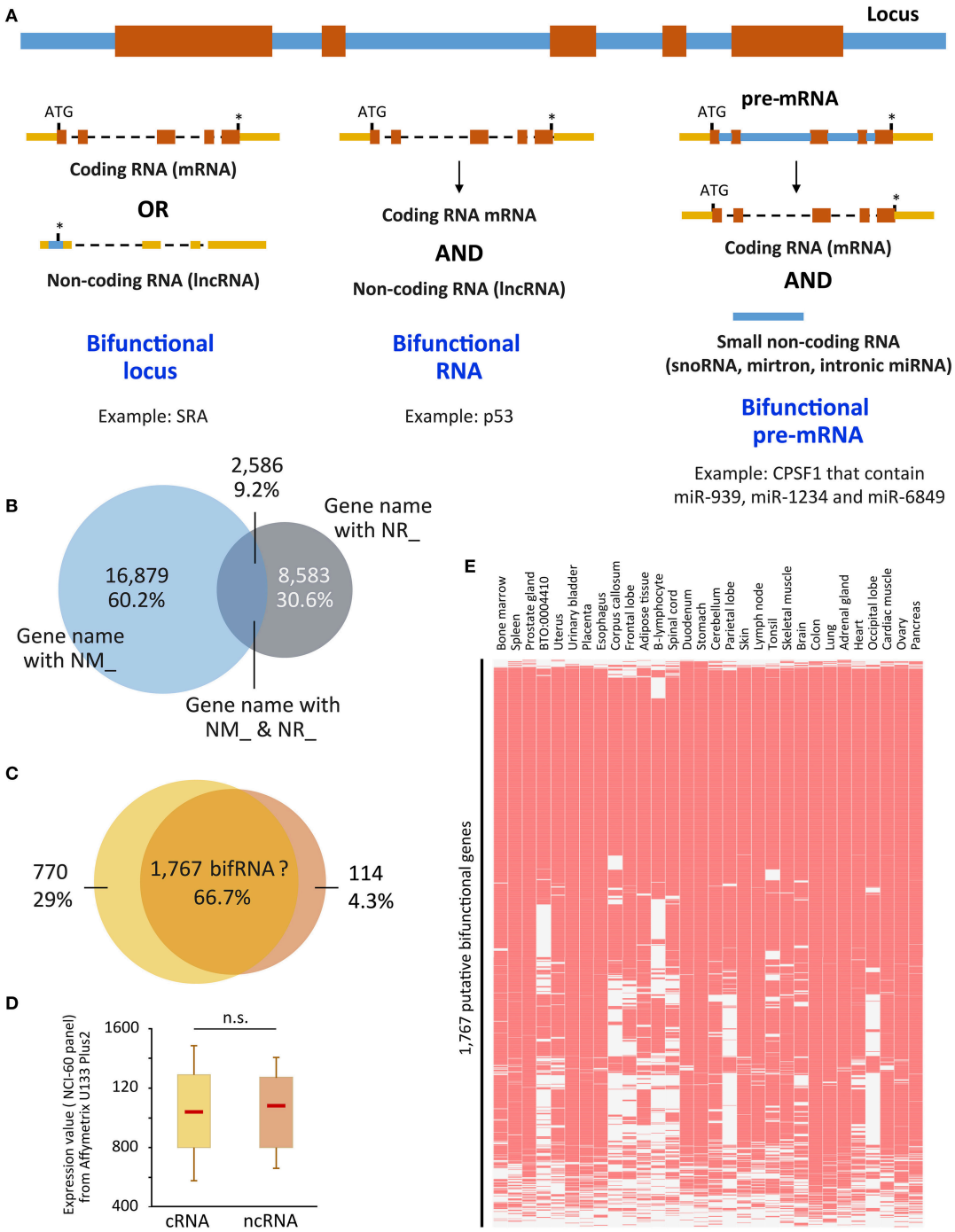


FIGURE 1 | (A) Representation of the three ways to achieve bifunctionality. Genomic/intronic regions are in blue, exonic regions are in orange (coding sequences) or yellow (non-coding sequences). **(B)** Overlap between genes that present NM_ and NR_ isoforms. "UCSC genes" track from UCSC Genome browser was retrieved and NM_ and NR_ transcripts were sorted using Galaxy tools. Venn diagram was performed using the gene names. 16,879 (60.2%) and 8,583 (30.6%) were related to only NM_ (coding) or NR_ (non-coding) transcripts, respectively, whereas 2,586 (9.2%) hold both coding and non-coding isoforms. **(C)** Putative bifunctional RNAs are recognized by probes from Affymetrix U133 Plus2 array. The 2,586 genes were recognized by 5,635 probes, of which 66.7% were not able to distinguish between coding and non-coding isoforms (NM_ and NR_, respectively). The 66.7% correspond to 1,767 genes. **(D)** Expression values attributed to each probe corresponding to the 770 and 114 exclusively coding and non-coding genes, was retrieved from the NCI-60 panel that referred to a panel of 60 Human Tumor Cell Lines Screen related to nine different cancer types [Geo DataSet Record GDS4296, (Barrett et al., 2005)] and plotted as boxplots. Significance was assessed using Wilcoxon-Mann-Whitney test. **(E)** The expression of the 1,767 genes which coding and non-coding properties are not distinguishable was assessed in the Jensen Tissues (https://tissues.jensenlab.org/) using Enrichr website (http://amp.pharm.mssm.edu/Enrichr/). ATG, start codon; *, stop codon; n.s., not significant; snoRNA, small nucleolar RNA, miRNA, microRNA; p53, protein 53 (Candeias et al., 2008); CPSF1, Cleavage and polyadenylation specificity factor 1 (Hube et al., 2017); cRNA, coding RNA; ncRNA, non-coding RNA. NM_ and NR_ are the prefixes used by RefSeq to describe protein-coding and non-protein-coding transcripts, respectively.

of its mRNA. Strikingly, it was also shown that Oskar could act as a non-coding RNA at earlier stages in oogenesis, independently of its coding capacity, where it may play a structural role for the assembly of cytoplasmic complexes essential for development of the oocyte (Jenny et al., 2006).

In fact, for many bifRNAs that were described here or elsewhere (Hube et al., 2006, 2011; Xu et al., 2008; Hashimoto et al., 2009; Ulveling et al., 2011a,b; Nagano et al., 2015; Nam et al., 2016; Williamson et al., 2017), it is not the same RNA molecule that carries both coding and non-coding functions. The use of alternative splicing or alternative promoter account for the numerous mechanisms that allow the production of coding or non-coding isoforms from the same loci. In this particular case, the terminology of bifRNA is incorrect and misleading; the genomic locus is bifunctional, not the RNA molecules produced.

A third possibility must be addressed to conclude this discussion about the concept of bifunctionality. For example, in the case of all the small nucleolar RNAs (snoRNAs) that have been described so far as being exclusively of intron origin, at least in humans, the host precursor RNA belongs to the bifunctional class as it can produce, after splicing, a mature mRNA and a mature snoRNA, with distinct functions; snoRNAs guide chemical modifications to other RNAs whereas mature mRNAs can then be translated into a protein product.

A FUZZY FRONTIER BETWEEN mRNAs AND ncRNAs

The first non-coding RNA ever identified was the alanine (Ala) transfer RNA (tRNA), purified from yeast, whose structure was published in 1965 (Holley et al., 1965). This Ala-tRNA, like the other tRNAs, is involved in protein synthesis by carrying the amino acid Ala to the protein chain being synthesized. Since this discovery, thanks to advances in sequencing techniques, thousands of non-coding RNAs have been identified in the genomes of prokaryotes and thousands in that of eukaryotes.

In prokaryotic organisms, the majority of the genome is coding. In eukaryotes, the opposite is true: most of the genome does not encode proteins, but contains colossal information in the form of panoply of regions transcribed into functional RNAs of different sizes and functions. Only about 2–5% of mammalian genomes contain information to produce proteins, whereas about 90% is transcribed over the lifespan into a large and complex transcriptome of ncRNAs. The repertoire of these transcripts is the subject of an active international search. It is still incomplete, but already makes it possible to propose a novel picture of the RNA world.

By nature, an mRNA is defined by the coding sequence it contains. Typically, upstream and downstream of their coding sequences, mRNAs also contain transcribed but untranslated regions (5′- and 3′-UnTranslated Region, 5′- and 3′-UTR, respectively) that are highly structured. Both UTRs are known to play key roles in post-transcriptional regulation, including the control of mRNAs transport, the translation efficiency, the subcellular localization and the overall stability of transcripts (van der Velden and Thomas, 1999; Bashirullah et al., 2001; Jansen, 2001) What needs to be emphasized here is that, even mRNAs defined to code for proteins have secondary structures or sub-structures that are functional in the RNA form. Conversely, regulatory RNAs with functional properties linked to their sequence and folding can also carry sequences that are translated into peptides under specific physiological conditions (Kondo et al., 2010; Kageyama et al., 2011; Magny et al., 2013; Zanet et al., 2015).

Thus, many mRNAs may act as regulatory RNAs whereas more and more regulatory RNAs, first classified as non-coding, are shown to hide small coding sequences *i.e.* less than the 300nt/100aa limit defined so far (Dinger et al., 2008; Ulveling et al., 2011b). A strict discrimination between these two classes of molecules appears even less realistic since, in eukaryotes, some ncRNAs have features comparable to that of mRNAs, such as polymerase II-dependant transcription and addition of a cap and a polyadenylated tail for instance (Dinger et al., 2008; Kondo et al., 2010; Ulveling et al., 2011b), and even their presence in polysome fractions (Ingolia et al., 2011). In various mammals, from mice to humans, many of these long ncRNAs contain coding sequences that may be expressed in specific contexts. How, in that case, could they be distinguished from mRNAs? It is to avoid this difficulty that we referred to them as bifRNAs (Dinger et al., 2008, 2011; Francastel and Hube, 2011; Ulveling et al., 2011a,b).

CAN WE ESTIMATE THE NUMBER OF BIFUNCTIONAL RNAS IN HUMAN?

While annotation systems are becoming more and more accurate, the previously automated annotated transcripts are being curated and verified manually. For example, the main features of the RefSeq collection (https://www.ncbi.nlm.nih.gov/books/NBK21091/) indicate, among others, non-redundancy, data validation, distinct accession series, and ongoing curation by NCBI staff and collaborators. In this type of nomenclature, "NM_" corresponds to validated protein-coding transcripts, while the prefix "NR_" indicates non-protein-coding transcripts also validated and curated. Each transcript, and more specifically each isoform produced, is associated with a unique identifier, which is itself linked to a gene locus. For example, the homo sapiens SRA1 gene has four isoforms that are deposited and accessible in NCBI website, variants 1 and 2 that are coding mRNAs and referred to as NM_001253764.1 and NM_001035235.3, respectively, whereas variants 3 and 4 are non-coding RNAs and identified as NR_045586.1 and NR_045587.1, respectively (Emberley et al., 2003; Chooniedass-Kothari et al., 2004, 2006). Actually, the diversity of SRA RNA transcripts in human cells is even more complex than described in PubMed since we have identified eight experimentally validated isoforms and at least 20 other isoforms from databases, all mainly non-coding transcripts (Hube et al., 2011), although they were not curated and thus, do not have an NR_ status.

We decided to use this characteristic (NM_ and NR_ criteria) to predict genes that could produce both coding and non-coding RNA isoforms. The methodology is explained in the legend of **Figure 1**. As seen in **Figure 1B**, 2,586 out of 28,048 gene loci are able to produce both NM_ and NR_ isoforms, i.e., almost 10% of the gene loci are able to produce both coding and non-coding isoforms that were validated and curated by NCBI. Going further in the analysis, we associated these transcripts to the "Affymetrix U133 plus 2 probes" and, as shown in **Figure 1C**, the majority of probes (66.7%) are unable to distinguish between coding and non-coding isoforms. It means that using the most up-to-date microarray technology, users are not able to distinguish between coding and non-coding genes for about 1,700 of them. Even if it represents only a small fraction of genes represented on the array (6%), data obtained from the array have to be used with caution, especially since the majority of these genes are genes expressed in almost all tissues [Jensen Tissues (Santos et al., 2015); **Figure 1E**] and are largely involved in metabolic pathways (KEGG hsa01100, $z = -3.385$, $p = 0.0003563$; not shown), which may lead to substantial biased interpretations. In addition, expression levels of these isoforms are not very different whether the transcripts are coding or not (**Figure 1D**) using dataset GDS4296 (Barrett et al., 2005).

This result leads to at least three very important conclusions. The first is that the non-coding versions of bifRNAs are not subjected to Nonsense-Mediated RNA Decay (NMD). Indeed, since these transcripts are detected in some cells/tissues, they escape the RNA surveillance mechanisms that otherwise rapidly degrades RNA with premature stop-codons (Zhang et al., 2009; Smith and Baker, 2015); Second, in contrast to what has been described for "conventional" ncRNAs supposed to date to be expressed at much lower levels than mRNAs (Ching et al., 2014; Wang et al., 2016), expression levels of coding vs. ncRNAs from the same locus are comparable regardless of their coding capacity (**Figure 1D**). Finally, almost two-third of the probes on expression arrays (**Figure 1C**) are not able to distinguish between coding and non-coding isoforms, which may lead to over- or mis-interpretation of data in some instances.

How to overcome this bias? An obvious but costly solution would be to develop new microarrays, taking into account all the annotated coding and non-coding isoforms. However, the constant discovery of new isoforms would rapidly overwhelm the process of probe design. The second solution may possibly come from the new technologies that are available today, such as high-throughput sequencing. However, at least in terms of isoform identification, this remains a challenge; the use of classical RNA-seq is still far from being accurate and the assembly of transcripts, although it is constantly being improved over time, still does not allow for true identification of the transcripts isoforms either. The latest generation of RNA-seq, the ultra-long reads method, seems to overcome most of these complications (Jain et al., 2017; Rhoads and Au, 2015), even if it remains to be improved since it is limited to several consecutive kbs, reveals error rates of about 10% and remains limited in the number of reads (Rhoads and Au, 2015).

CONCLUSION

Since the burst of studies on non-coding RNAs, i.e., for almost 20 years now, there is no longer any doubt about their existence and importance in cellular processes. Now that the scientific community has accepted the concept of coding and non-coding RNAs, we may have to take a step backward to reassess the possibility that these two categories of transcripts are more interdependent than thought. These new conclusions impose for a deeper examination into the functional significance of these dynamic bifRNAs and for increased efforts toward a more integrated view of transcriptome/proteome in a given cellular context.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This work was supported by AFM (Association Française contre les Myopathies)—Research Grant (#20534).

REFERENCES

Barrett, T., Suzek, T. O., Troup, D. B., Wilhite, S. E., Ngau, W. C., Ledoux, P., et al. (2005). NCBI GEO: mining millions of expression profiles–database and tools. *Nucleic Acids Res.* 33, D562–D566. doi: 10.1093/nar/gki022

Bashirullah, A., Cooperstock, R. L., and Lipshitz, H. D. (2001). Spatial and temporal control of RNA stability. *Proc. Natl. Acad. Sci. U.S.A.* 98, 7025–7028. doi: 10.1073/pnas.111145698

Candeias, M. M., Malbert-Colas, L., Powell, D. J., Daskalogianni, C., Maslon, M. M., Naski, N., et al. (2008). P53 mRNA controls p53 activity by managing Mdm2 functions. *Nat. Cell Biol.* 10, 1098–1105. doi: 10.1038/ncb1770

Ching, T., Huang, S., and Garmire, L. X. (2014). Power analysis and sample size estimation for RNA-Seq differential expression. *RNA* 20, 1684–1696. doi: 10.1261/rna.046011.114

Chooniedass-Kothari, S., Emberley, E., Hamedani, M. K., Troup, S., Wang, X., Czosnek, A., et al. (2004). The steroid receptor RNA activator is the first functional RNA encoding a protein. *FEBS Lett.* 566, 43–47. doi: 10.1016/j.febslet.2004.03.104

Chooniedass-Kothari, S., Hamedani, M. K., Troup, S., Hubé, F., and Leygue, E. (2006). The steroid receptor RNA activator protein is expressed in breast tumor tissues. *Int. J. Cancer* 118, 1054–1059. doi: 10.1002/ijc.21425

Dinger, M. E., Gascoigne, D. K., and Mattick, J. S. (2011). The evolution of RNAs with multiple functions. *Biochimie* 93, 2013–2018. doi: 10.1016/j.biochi.2011.07.018

Dinger, M. E., Pang, K. C., Mercer, T. R., and Mattick, J. S. (2008). Differentiating protein-coding and noncoding RNA: challenges and ambiguities. *PLoS Comput. Biol.* 4:e1000176. doi: 10.1371/journal.pcbi.1000176

Emberley, E., Huang, G. J., Hamedani, M. K., Czosnek, A., Ali, D., Grolla, A., et al. (2003). Identification of new human coding steroid receptor RNA activator isoforms. *Biochem. Biophys. Res. Commun.* 301, 509–515. doi: 10.1016/S0006-291X(02)03070-X

Francastel, C., and Hube, F. (2011). Coding or non-coding: need they be exclusive? *Biochimie* 93, vi–vii. doi: 10.1016/S0300-9084(11)00322-1

Hashimoto, K., Ishida, E., Matsumoto, S., Shibusawa, N., Okada, S., Monden, T., et al. (2009). A liver X receptor (LXR)-β alternative splicing variant (LXRBSV) acts as an RNA co-activator of LXR-β. *Biochem. Biophys. Res. Commun.* 390, 1260–1265. doi: 10.1016/j.bbrc.2009.10.132

Holley, R. W., Apgar, J., Everett, G. A., Madison, J. T., Marquisee, M., Merrill, S. H., et al. (1965). Structure of a ribonucleic acid. *Science* 147, 1462–1465. doi: 10.1126/science.147.3664.1462

Hube, F., Guo, J., Chooniedass-Kothari, S., Cooper, C., Hamedani, M. K., Dibrov, A. A., et al. (2006). Alternative splicing of the first intron of the steroid receptor RNA activator (SRA) participates in the generation of coding and noncoding RNA isoforms in breast cancer cell lines. *DNA Cell Biol.* 25, 418–428. doi: 10.1089/dna.2006.25.418

Hubé, F., Ulveling, D., Sureau, A., Forveille, S., and Francastel, C. (2017). Short intron-derived ncRNAs. *Nucleic Acids Res.* 45, 4768–4781. doi: 10.1093/nar/gkw1341

Hubé, F., Velasco, G., Rollin, J., Furling, D., and Francastel, C. (2011). Steroid receptor RNA activator protein binds to and counteracts SRA RNA-mediated activation of MyoD and muscle differentiation. *Nucleic Acids Res.* 39, 513–525. doi: 10.1093/nar/gkq833

Ingolia, N. T., Lareau, L. F., and Weissman, J. S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* 147, 789–802. doi: 10.1016/j.cell.2011.10.002

Jain, M., Tyson, J. R., Loose, M., Ip, C. L. C., Eccles, D. A., O'Grady, J., et al. (2017). MinION analysis and reference consortium: phase 2 data release and analysis of R9.0 chemistry. *F1000Res.* 6:760. doi: 10.12688/f1000research.11354.1

Jansen, R. P. (2001). mRNA localization: message on the move. *Nat. Rev. Mol. Cell Biol.* 2, 247–256. doi: 10.1038/35067016

Jenny, A., Hachet, O., Závorszky, P., Cyrklaff, A., Weston, M. D., Johnston, D. S., et al. (2006). A translation-independent role of oskar RNA in early Drosophila oogenesis. *Development* 133, 2827–2833. doi: 10.1242/dev.02456

Kageyama, Y., Kondo, T., and Hashimoto, Y. (2011). Coding vs non-coding: translatability of short ORFs found in putative non-coding transcripts. *Biochimie* 93, 1981–1986. doi: 10.1016/j.biochi.2011.06.024

Kondo, T., Plaza, S., Zanet, J., Benrabah, E., Valenti, P., Hashimoto, Y., et al. (2010). Small peptides switch the transcriptional activity of shavenbaby during Drosophila embryogenesis. *Science* 329, 336–339. doi: 10.1126/science.1188158

Magny, E. G., Pueyo, J. I., Pearl, F. M., Cespedes, M. A., Niven, J. E., Bishop, S. A., et al. (2013). Conserved regulation of cardiac calcium uptake by peptides encoded in small open reading frames. *Science* 341, 1116–1120. doi: 10.1126/science.1238802

Maki, K., Morita, T., Otaka, H., and Aiba, H. (2010). A minimal base-pairing region of a bacterial small RNA SgrS required for translational repression of ptsG mRNA. *Mol. Microbiol.* 76, 782–792. doi: 10.1111/j.1365-2958.2010.07141.x

Nagano, H., Yamagishi, N., Tomida, C., Yano, C., Aibara, K., Kohno, S., et al. (2015). A novel myogenic function residing in the 5' non-coding region of Insulin receptor substrate-1 (Irs-1) transcript. *BMC Cell Biol.* 16:8. doi: 10.1186/s12860-015-0054-8

Nam, J. W., Choi, S. W., and You, B. H. (2016). Incredible RNA: dual functions of coding and noncoding. *Mol. Cells* 39, 367–374. doi: 10.14348/molcells.2016.0039

Naski, N., Gajjar, M., Bourougaa, K., Malbert-Colas, L., Fahraeus, R., and Candeias, M. M. (2009). The p53 mRNA-Mdm2 interaction. *Cell Cycle* 8, 31–34. doi: 10.4161/cc.8.1.7326

Rhoads, A., and Au, K. F. (2015). PacBio sequencing and its applications. *Gen. Proteomics Bioinformatics* 13, 278–289. doi: 10.1016/j.gpb.2015.08.002

Rongo, C., Gavis, E. R., and Lehmann, R. (1995). Localization of oskar RNA regulates oskar translation and requires Oskar protein. *Development* 121, 2737–2746.

Santos, A., Tsafou, K., Stolte, C., Pletscher-Frankild, S., O'Donoghue, S. I., and Jensen, L. J. (2015). Comprehensive comparison of large-scale tissue expression datasets. *PeerJ.* 3:e1054. doi: 10.7717/peerj.1054

Smith, J. E., and Baker, K. E. (2015). Nonsense-mediated RNA decay—a switch and dial for regulating gene expression. *Bioessays* 37, 612–623. doi: 10.1002/bies.201500007

Ulveling, D., Francastel, C., and Hubé, F. (2011a). Identification of potentially new bifunctional RNA based on genome-wide data-mining of alternative splicing events. *Biochimie* 93, 2024–2027. doi: 10.1016/j.biochi.2011.06.019

Ulveling, D., Francastel, C., and Hubé, F. (2011b). When one is better than two: RNA with dual functions. *Biochimie* 93, 633–644. doi: 10.1016/j.biochi.2010.11.004

Vanderpool, C. K., Balasubramanian, D., and Lloyd, C. R. (2011). Dual-function RNA regulators in bacteria. *Biochimie* 93, 1943–1949. doi: 10.1016/j.biochi.2011.07.016

van der Velden, A. W., and Thomas, A. A. (1999). The role of the 5′ untranslated region of an mRNA in translation regulation during development. *Int. J. Biochem. Cell Biol.* 31, 87–106. doi: 10.1016/S1357-2725(98)00134-4

Wadler, C. S., and Vanderpool, C. K. (2007). A dual function for a bacterial small RNA: SgrS performs base pairing-dependent regulation and encodes a functional polypeptide. *Proc. Natl. Acad. Sci. U.S.A.* 104, 20454–20459. doi: 10.1073/pnas.0708102104

Wang, Y., Xue, S., Liu, X., Liu, H., Hu, T., Qiu, X., et al. (2016). Analyses of long non-coding RNA and mRNA profiling using RNA sequencing during the pre-implantation phases in pig endometrium. *Sci. Rep.* 6:20238. doi: 10.1038/srep20238

Williamson, L., Saponaro, M., Boeing, S., East, P., Mitter, R., Kantidakis, T., et al. (2017). UV irradiation induces a non-coding RNA that functionally opposes the protein encoded by the same gene. *Cell* 168, 843–855. doi: 10.1016/j.cell.2017.01.019

Xu, Q., Walker, D., Bernardo, A., Brodbeck, J., Balestra, M. E., and Huang, Y. (2008). Intron-3 retention/splicing controls neuronal expression of apolipoprotein E in the CNS *J. Neurosci.* 28, 1452–1459. doi: 10.1523/JNEUROSCI.3253-07.2008

Zanet, J., Benrabah, E., Li, T., Pélissier-Monier, A., Chanut-Delalande, H., Ronsin, B., et al. (2015). Pri sORF peptides induce selective proteasome-mediated protein processing. *Science* 349, 1356–1358. doi: 10.1126/science.aac5677

Zhang, Z., Xin, D., Wang, P., Zhou, L., Hu, L., Kong, X., et al. (2009). Noisy splicing, more than expression regulation, explains why some exons are subject to nonsense-mediated mRNA decay. *BMC Biol.* 7:23. doi: 10.1186/1741-7007-7-23

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Hubé and Francastel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.