



Editorial

Coding or non-coding: need they be exclusive?

“The molecular structure of proteins is determined by specific elements, the structural genes. They act by forming a cytoplasmic transcript of themselves, the structural messenger, which in turn synthesizes the protein”, from *Genetic Regulatory Mechanisms in the Synthesis of Proteins*, written by Francois JACOB & Jacques MONOD in 1961 [1].

Exactly 50 years later, the amazing discovery that RNA molecules serve as intermediates in the synthesis of proteins still holds true... in the case of protein-coding genes. With the boom of genome sequencing projects and advances in high throughput technologies, the first assumption to be seriously revised downwards is the actual number of coding-genes, that constantly decreased to reach the small number of about 19,000 in humans as opposed to 100,000 that was first announced about two decades ago. Since our acceptance that the remainder of the genome is not merely junk DNA, the number of non-coding regions associated with, thereafter named, non-coding transcripts reached the unpredictable number of more than 30,000 which includes non-coding RNA categorized as short, long, antisense and pseudogenes... and many more categories. With the number of distinct non-protein coding RNAs being comparable to that of messenger RNA, non-coding RNAs now emerge as key actors in fundamental processes of biology, in that they can fulfil most jobs primarily ascribed to proteins such as regulatory and architectural functions.

Even if looking with Candide's eyes¹ at the huge amount of data generated by genome consortiums one might realize that the rather Manichean classification into coding vs. non-coding transcripts is somewhat unrealistic with respect to the subtlety and complexity of genome regulation set up throughout evolution.

It is in that spirit, and as a logical extension of our review published in a previous issue of *Biochimie* [2], that this special issue has been designed. Only a few years ago, the concept that non-coding RNA may in fact hide a coding capacity was beyond rational. Needless to mention that the same applied to messenger RNA, that could hold a functional role in addition to protein-coding capacity.

This Special Issue is intended to provide a collection of evidence that, from Bacteria to Mammals, and as further emphasized by John Mattick in his keynote review, there are multiple messages in RNA. It groups more than twenty comprehensive reviews and reports, written by the foremost specialists in the field.

¹ In *Candide*, written by Voltaire in 1759, Voltaire describes Candide as a character very credulous and very naive who believes blindly in the philosophy of his preceptor Pangloss. After a voyage of discovery and the encounter of many people, Candide has become less candid and declares that one must think for oneself and form one's own opinion by focussing on evidence and not too much on theory.

This issue starts with reviews commenting on how functional RNA information can arise from pervasive transcription of the genome (Morillon's group), splicing of introns and mRNA maturation (Lai and Meister's groups), pseudogenes (Muro's group) or antisense transcription (Morris's group). Certain loci evolved from coding to non-coding or *vice versa* during evolution (Schmitz and Rougeulle's groups). Then, several reviews give examples of RNA with reported activities at both the protein and the RNA level in Bacteria (Vanderpool's group), plants (Crespi's group), amphibians and insects (Kloc's groups) to mammals (Candeias, Leedman and Leygue's groups), and the example of non-coding RNA that do contain a tiny ORF (Kageyama's group). Because the duality of information contained in RNA is not limited to messenger RNA, there are reviews on RNAs that hide both snoRNA and miRNA (Scott's group), RNAs that have both function as transfer and messenger RNA (Keiler's group) and the special case of ribozymes (Talini's group). Furling's group discuss the gain of dual function of RNA observed in diseases. Since genome projects generated huge amounts of data that await extensive analysis, we end with reports of a transcriptomic survey suggesting that many loci produce RNA with more than one function (Dinger's group), a computational analysis of the overlap between coding regions and regions that produce structured, potentially non-coding, RNA (Stadler's group) and finally a data mining approach to predict potential new bifunctional RNAs in humans (our own group).

More than a century after the discovery of RNA, and half a century after that the term 'messenger' RNA was used by François Jacob and Jacques Monod to designate the intermediate molecules between a gene and a protein, one has now to challenge and adapt established paradigms in light of recent evidence that multiple functions can be shared by the same RNA entity. Hempel's Raven Paradox reminds us that any theory, even if established, must be adaptable and prone to change, to assimilate new data. We therefore hope that, in light of the recent developments in the field presented here, the original definition for RNA, that is to store information, will remain the best definition in understanding that it does not have to be limited to a single one.

Acknowledgements

We wish to thank all the authors for having accepted our invitation and for their excellent contributions to this special issue, and the reviewers for their insightful and constructive comments. We are grateful to the Redaction de *Biochimie* (Richard Buckingham, Claude Forest and Malika Hassini) for having accepted our

proposition to set up a special issue on a subject still off the main stream of research in the field of RNA, for their help and efficient handling of the manuscripts until their final stage of publication.

References

- [1] F. Jacob, J. Monod, Genetic regulatory mechanisms in the synthesis of proteins, *J. Mol. Biol.* 3 (1961 Jun) 318–356.
- [2] D. Ulveling, C. Francastel, F. Hubé, When one is better than two: RNA with dual functions, *Biochimie* 93 (2011 Apr) 633–644.

Claire Francastel*, Florent Hubé
*Nuclear Architecture and Epigenetic Regulation of Differentiation,
Epigenetic and Cell Fate Units, Paris Diderot Paris
7 University, Bâtiment Lamarck – Case 7042,
35 rue Hélène Brion, 75013 Paris, France*

* Corresponding author.
E-mail addresses: claire.francastel@univ-paris-diderot.fr
(C. Francastel); Florent.hube@univ-paris-diderot.fr (F. Hubé)