



# Editorial: Bioinformatics of non-coding RNAs with applications to biomedicine: recent advances and open challenges

Alessandro Laganà<sup>1\*</sup>, Alfredo Ferro<sup>2</sup> and Carlo Maria Croce<sup>3</sup>

<sup>1</sup> Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA <sup>2</sup> Department of Clinical and Molecular Biomedicine, University of Catania, Catania, Italy <sup>3</sup> Department of Molecular Virology, Immunology and Medical Genetics, Comprehensive Cancer Center, The Ohio State University, Columbus, OH, USA

**Keywords:** ncRNA, miRNA, lncRNA, RNA editing, NGS, biomedicine

The recent advances in the functional characterization of non-protein-coding RNAs (ncRNAs) have represented a major breakthrough in the life sciences. Large-scale projects, such as ENCODE, have shown that the human genome is pervasively transcribed and that a large proportion of the mammalian transcriptome consists of ncRNA transcripts. ncRNA genes can be roughly classified into short ncRNAs (<200 nt) and long ncRNAs (>200 nt).

The first class includes well characterized, infrastructural molecules, such as rRNA, tRNA, and snoRNA, which have a housekeeping role in essential processes like splicing and translation, and regulatory ncRNA such as miRNA and piRNA, which are involved in post-transcriptional regulation of gene expression and in the silencing of transposable elements during germ line development, respectively.

The second class of ncRNA consists of longer transcripts that are still poorly characterized mostly due to their heterogeneity in size, structure, and biogenesis. LncRNA genes outnumber short ncRNAs and are probably more abundant than protein coding genes. Such RNAs exhibit various degrees of conservation and are often polyadenylated and tissue-specific. Accumulating evidence indicates that they likely have a broad range of functions, including chromatin remodeling, gene regulation, and protein transport and trafficking.

Genetic and epigenetic aberrations affecting ncRNA gene sequences and their expression have been linked to a variety of pathological conditions, including cancer, cardiovascular, and neurological diseases.

Recently, high-throughput sequencing techniques have enabled the study of entire transcriptomes at single nucleotide resolution, providing unprecedented details of their organization, expression, modifications, and structure. Bioinformatics tools constitute an essential resource for ncRNA research, providing a powerful means to organize, integrate, and analyze the huge amount of data generated by such technologies.

The aim of this Research Topic is to review current knowledge, introduce novel methods, and discuss open challenges of this exciting and innovative field in connection with the most important biomedical applications. We have collected five original research and methods articles and four reviews, spanning the full scope of the Research Topic.

Two excellent reviews focus on the discovery of ncRNA from NGS data. Kang and Friedländer (2015) surveyed computational tools to predict animal miRNAs from short RNA sequencing data (RNAseq). The authors covered the basics of miRNA prediction, reviewed several methods, described the algorithms, and discussed their strengths. They also described algorithms for specific cases, such as prediction from massively pooled data or in species without reference genomes, and discussed challenges and future directions of the field. Veneziano et al. (2015), instead, provided

## OPEN ACCESS

### Edited and reviewed by:

Richard D. Emes,  
University of Nottingham, UK

### \*Correspondence:

Alessandro Laganà  
alessandro.lagana@gmail.com

### Specialty section:

This article was submitted to  
Bioinformatics and Computational  
Biology, a section of the  
journal *Frontiers in Bioengineering and  
Biotechnology*

**Received:** 17 September 2015

**Accepted:** 25 September 2015

**Published:** 08 October 2015

### Citation:

Laganà A, Ferro A and Croce CM  
(2015) Editorial: Bioinformatics of  
non-coding RNAs with applications to  
biomedicine: recent advances and  
open challenges.  
*Front. Bioeng. Biotechnol.* 3:156.  
doi: 10.3389/fbioe.2015.00156

a more general state-of-the-art coverage of the computational approaches for the discovery and analysis of small and long ncRNA through NGS techniques.

The detection of miRNAs from NGS data becomes an even more challenging task when sequence variants, termed isomiRs, are taken into account. IsomiRs were initially considered sequencing artifacts, but evidence showed that they are functional variants with a specific biological role. Muller et al. (2014) introduced IsomiRage, a streamlined pipeline to identify and analyze isomiRs from next generation sequencing data. The tool is able to distinguish canonical miRNAs from templated and non-templated isomiRs, including 5'- and 3'-extended and trimmed variants.

Two articles of this collection concern RNA editing, a dynamic, widespread process that alters the sequence of RNA transcripts. In particular, A-to-I editing is the most common RNA post-transcriptional modification in human and involves the deamination of adenosine (A) to inosine (I), which is recognized as guanosine (G) by all cellular machineries. Editing may alter both coding and non-coding sequences, with important functional consequences. Nigita et al. (2015) presented a comprehensive state-of-the-art review of databases and computational approaches for the discovery and the analysis of RNA editing, with particular emphasis on ncRNA. They summarized current knowledge and discussed potential consequences of RNA editing on ncRNA, pointing out the lack of tools specifically designed for the detection of editing alterations in lncRNA sequences. This gap was actually addressed by a methods article in our collection by Picardi et al. (2014). They described a novel computational approach to reliably detect A-to-I editing events in human lncRNAs through NGS, based on their previously published package called REDIttools. In the presented article, the authors showed the potential of their tools in recovering A-to-I

candidates from RNAseq data and provided guidelines to improve RNA editing detection in non-coding RNAs, with specific focus on lncRNAs.

This collection also includes three articles concerning data integration and functional analysis. In the first one, Bonnici et al. (2014) introduced ncRNA-DB, a novel database of ncRNA interaction in human. The database integrates associations among ncRNA, protein coding genes and diseases. It can be searched by a web-based or a command line interface and is also accessible through a Cytoscape app called ncNetView. The second paper, by Alaimo et al. (2014), described ncPred, a novel tool that predicts ncRNA-disease association through tripartite network-based inference. The results of the experimental analysis show that the tool is able to predict more biologically significant associations than its competitors. In the third paper, Li et al. (2015) performed a large-scale integration of publicly available RNA binding protein (RBP) binding sites generated by high-throughput CLIP-Seq technology and identified thousands of RBP-lncRNA interactions. The authors reported combinatorial effects among RBPs and discovered hundreds of disease-related SNPs in RBP binding sites in lncRNA.

The collection also includes a review of design principles and computational tools for the design of synthetic RNAs for gene regulation (Laganà et al., 2014). The article provided guidelines for the design of siRNA, artificial miRNA, antagomiRs, miRNA sponges, and small guide RNA for CRISPRi, and presented strengths and limitations of the different technologies.

Bioinformatics of ncRNA is a vast and rich field and the papers that we selected for this Research Topic address some of its most exciting and pressing challenges. We believe that this volume represents a valuable and useful resource and hope it will be of interest to the many researchers involved in ncRNA research.

## REFERENCES

- Alaimo, S., Giugno, R., and Pulvirenti, A. (2014). ncPred: ncRNA-disease association prediction through tripartite network-based inference. *Front. Bioeng. Biotechnol.* 2:71. doi:10.3389/fbioe.2014.00071
- Bonnici, V., Russo, F., Bombieri, N., Pulvirenti, A., and Giugno, R. (2014). Comprehensive reconstruction and visualization of non-coding regulatory networks in human. *Front. Bioeng. Biotechnol.* 2:69. doi:10.3389/fbioe.2014.00069
- Kang, W., and Friedländer, M. R. (2015). Computational prediction of miRNA genes from small RNA sequencing data. *Front. Bioeng. Biotechnol.* 3:7. doi:10.3389/fbioe.2015.00007
- Laganà, A., Shasha, D., and Croce, C. M. (2014). Synthetic RNAs for gene regulation: design principles and computational tools. *Front. Bioeng. Biotechnol.* 2:65. doi:10.3389/fbioe.2014.00065
- Li, J.-H., Liu, S., Zheng, L.-L., Wu, J., Sun, W.-J., Wang, Z.-L., et al. (2015). Discovery of protein-lncRNA interactions by integrating large-scale CLIP-Seq and RNA-Seq datasets. *Front. Bioeng. Biotechnol.* 2:88. doi:10.3389/fbioe.2014.00088
- Muller, H., Marzi, M. J., and Nicassio, F. (2014). IsomiRage: from functional classification to differential expression of miRNA isoforms. *Front. Bioeng. Biotechnol.* 2:38. doi:10.3389/fbioe.2014.00038
- Nigita, G., Veneziano, D., and Ferro, A. (2015). A-to-I RNA editing: current knowledge sources and computational approaches with special emphasis on non-coding RNA molecules. *Front. Bioeng. Biotechnol.* 3:37. doi:10.3389/fbioe.2015.00037
- Picardi, E., D'Erchia, A. M., Gallo, A., Montalvo, A., and Pesole, G. (2014). Uncovering RNA editing sites in long non-coding RNAs. *Front. Bioeng. Biotechnol.* 2:64. doi:10.3389/fbioe.2014.00064
- Veneziano, D., Nigita, G., and Ferro, A. (2015). Computational approaches for the analysis of ncRNA through deep sequencing techniques. *Front. Bioeng. Biotechnol.* 3:77. doi:10.3389/fbioe.2015.00077

**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 © 2015 Laganà, Ferro and Croce. 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) or licensor 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.