isomiR-SEA: miRNA and isomiR expression level detection in seven RNA-Seq datasets

Gianvito Urgese, Giulia Paciello, Enrico Macii and Andrea Acquaviva, Elisa Ficarra

Department of Control and Computer Engineering, Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129 Turin, Italy

Why identifying and quantifying miRNAs and isomiRs?

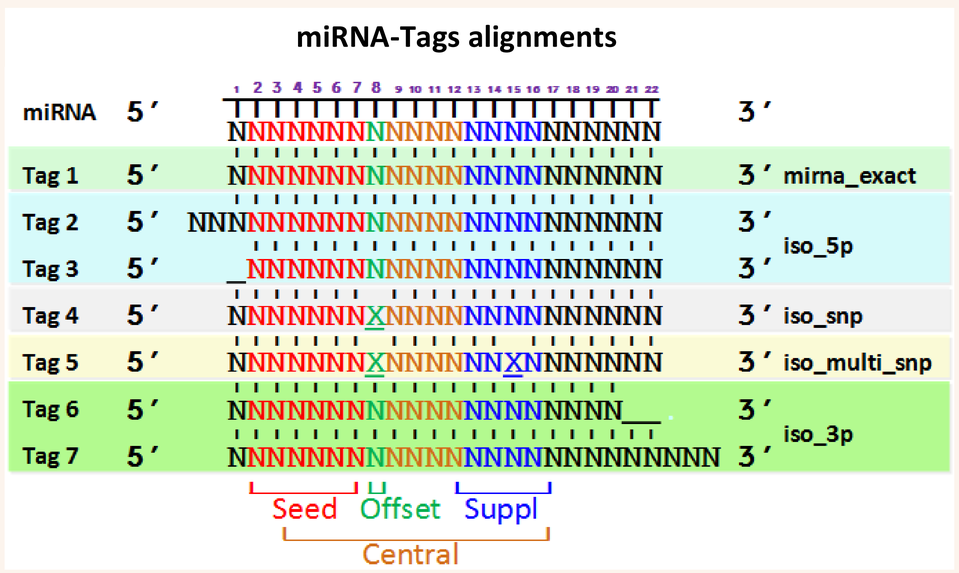
The role of miRNAs in multi-factorial diseases such as cancer is well assessed. Indeed, miRNAs can act by inhibiting or promoting both oncogenic or metastatic processes[1]. Conversely, the role of isomiRs has to be clarified. However, deregulated isomiR patterns were pointed out in some cancers as the gastric one. Furthermore, miRNAs are considered good bio-markers, making their adoption for disease characterization highly desirable.

The impact of NGS on miRNA/isomiR analysis

The analysis of data coming from transcriptome sequencing potentially accounts for an accurate identification and quantification of both known and novel miRNA and isomiR molecules. To this aim, several methodologies working on miRNA-Sequencing (miRNA-Seq) data were devised in the last decade. However, many of these tools do not provide any information concerning **isomiRs** and miRNA-mRNA **Interaction Sites**[2] spectrum.

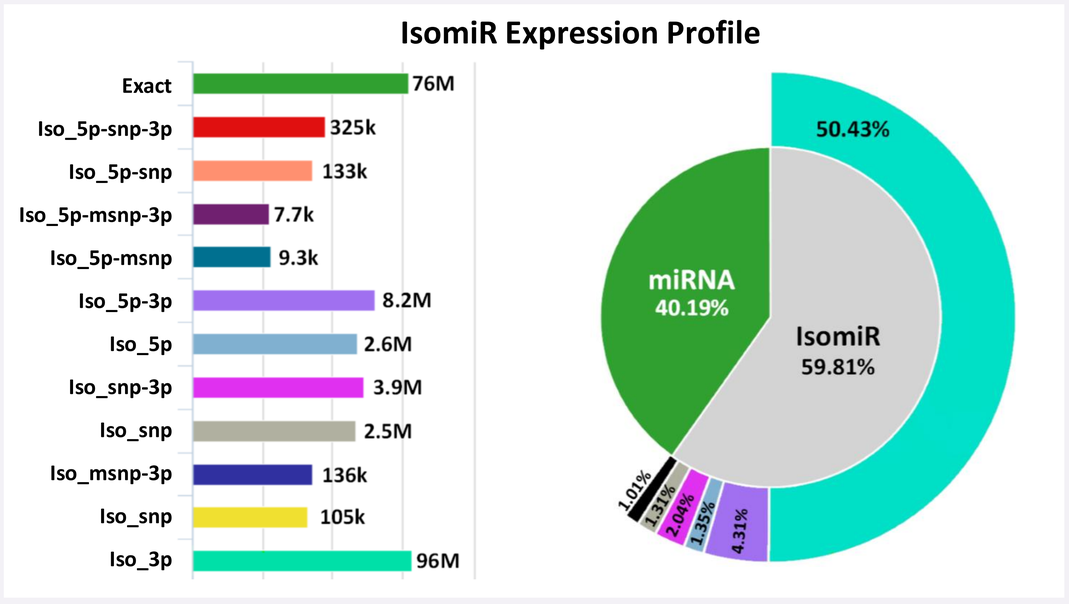
What's new in isomiR-SEA approach?

isomiR-SEA[3] implements a miRNA-specific alignment procedure that results in very accurate miRNA/isomiR expression levels and precise evaluations of miRNA-mRNA conserved **Interaction SIte**s. As first activity, **isomiR-SEA** identifies miRNA seeds within the tags. If the seed is found, the alignment is extended and the positions of the encountered mismatches recorded and analysed to distinguish among miRNAs and isomiRs, and to assess the conservation of the interaction sites.

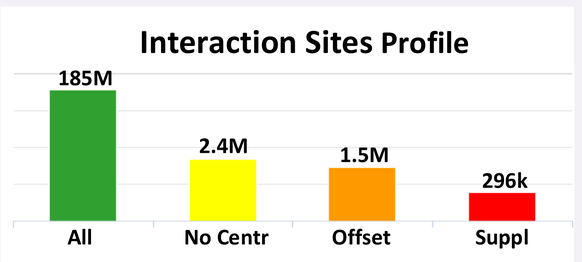


Experiment 1: Large scale analysis

We analysed 23 human miRNA-Seq Samples (GSE13370, GSE19812, GSE21279, GSE22918, GSE23090, and GSE26516) for reporting on the overall ***IsomiRs*** expression level and the conserved miRNA-mRNA ***Interaction Sites***.



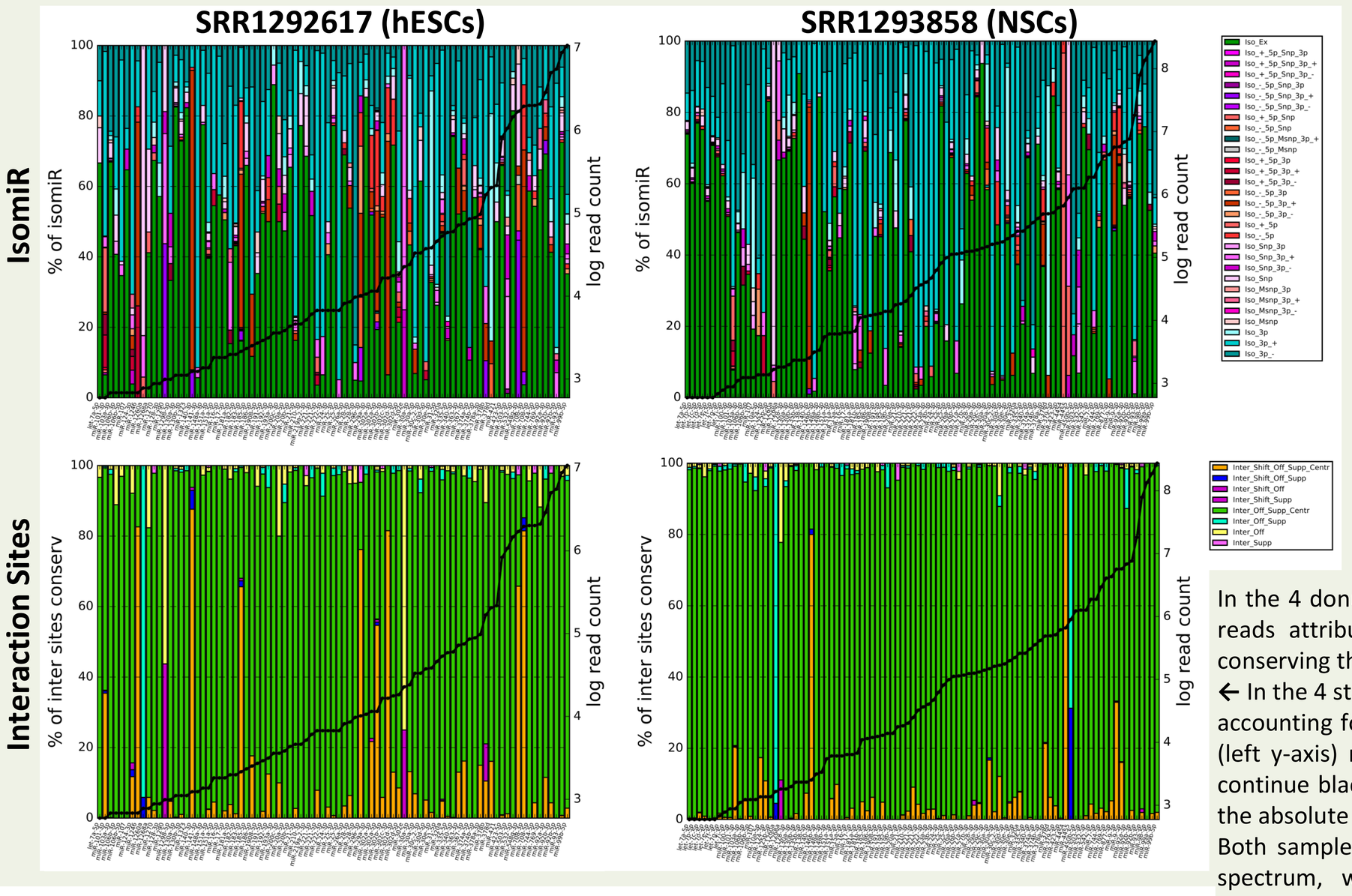
IsomiR: Reads attributed to miRNA exact sequences are about 40% of the 189M. The remaining reads come from isomiRs. A huge amount derives from 3' isomiR molecules (50%). IsomiRs 5' and SNP or combinations of them are in detected in lower percentages.

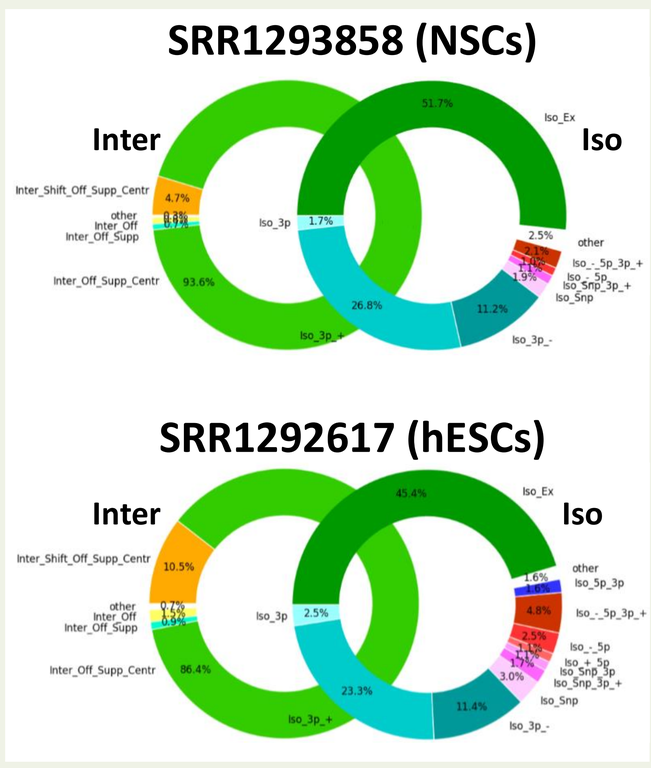


Interaction Site: The presence of isomiRs did not strongly impact on the conserved interaction sites. Indeed, only 4.2M of reads (2%) shown missed interaction sites.

Experiment 2: Two samples analysis

We analysed a human embryonic steam cells (hESCs) and a *neural steam cells (NSCs)* samples from Tan[4] for pointing out trends in both ***IsomiRs*** expression levels and conserved miRNA-mRNA ***Interaction Sites*** profiles. These trends cannot be appreciated from the overall miRNAs expression profile.





In the 4 donut charts ↑ we show the overall percentages of reads attributed to miRNAs/isomiRs (**Iso** label) and reads conserving the interaction sites (**Inter** label).

← In the 4 stacked barcharts we provide percentages of reads accounting for both **IsomiRs** and conserved **Interaction Sites** (left y-axis) referring to miRNAs labelled on the x-axis. The continue black line, with scale on the right y-axis, represents the absolute number of reads mapped on each miRNA.

Both samples are characterized by a heterogeneous **IsomiR** spectrum, while all the **Interaction Sites** are generally conserved with some shift caused by indels on the read 5p.

Conclusions

Analyses performed with isomiR-SEA proven the wide spectrum of information hidden into miRNAs-Seq data. Being designed around a miRNAs-specific alignment algorithm IsomiR-SEA is capable to:

i) Provide accurate miRNAs expression profiles;
ii) Distinguish and classify each read mapped on a miRNA as exact mature form or as ***IsomiR***;
iii) Report the spectrum of mapped reads conserving miRNA-mRNA ***Interaction Sites***.

The evaluation of isomiR-SEA results could help to gain novel insights into miRNAs mediated cell behaviours and to better characterize the miRNAs associated diseases and conditions.

[1]Etheridge, Alton, et al. "Extracellular microRNA: a new source of biomarkers." Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis 717.1 (2011): 85-90.

[2]Bartel, David P. "MicroRNAs: target recognition and regulatory functions." Cell 136.2 (2009): 215-233.

[3]Urgese, Gianvito, et al. "isomiR-SEA: an RNA-Seq analysis tool for miRNAs/isomiRs expression level profiling and miRNA-mRNA interaction sites evaluation." BMC bioinformatics 17.1 (2016): 148. http://eda.polito.it/isomir-sea/

[4]Tan, Geok Chin, et al. "5' isomiR variation is of functional and evolutionary importance." NAR (2014): gku656.