

The NonCode aReNA DB: a non-redundant and integrated collection of non-coding RNAs

Giorgio De Caro¹, Arianna Consiglio¹, Domenica D'Elia^{1✉}, Andreas Gisel^{1,2}, Giorgio Grillo¹, Sabino Liuni¹, Angelica Tulipano¹, Flavio Licciulli¹

¹CNR, Institute for Biomedical Technologies, Bari, Italy

²International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria

De Caro G *et al.* (2015) *EMBnet:journal* **21**(Suppl A), e834. <http://dx.doi.org/10.14806/ej.21.A.834>

The recent availability of high throughput technologies, like next generation sequencing (NGS) platforms, has provided the scientific community with an unprecedented opportunity for large-scale analysis of genome in a large number of organisms. However, among others, one of the most challenging task for bioinformaticians is to develop tools that provide biologists with an easy access to curated and non-redundant collections of sequence data.

Non-coding RNAs, for a long time believed to be not-functional, are emerging as the most large and important family of gene regulators. NonCode aReNA Database is a comprehensive and non-redundant source of manually curated and automatically annotated ncRNA transcripts. Originally developed as a component of a bigger project, composed by a datawarehouse for the functional annotation of ncRNAs from NGS data, NonCode aReNA DB is currently available

as a web-resource at <http://ncrnadb.ba.itb.cnr.it/>. Sequences have been classified in diverse biotypes and associated to Sequence Ontology terms. The database can be queried by using multi-criteria and ontological search through an easy-to-use web interface, and data exported as non-redundant collections of transcripts annotated in VEGA, ENSEMBL, RefSeq, miRBase, GtRNAdb and piRNABank. The database is updated through an automatic pipeline and last update was on January 2015. Presently NonCode aReNA DB contains 134,908 human ncRNAs classified in 24 biotypes, and next update will include transcripts of *Mus musculus* and *Arabidopsis thaliana*.

Acknowledgements

This work was supported by the Italian MIUR Flagship Project "Epigen".