

## **REPRODUCIBLE COMPUTATIONAL RESEARCH IN HIGH-THROUGHPUT PLANT BIOLOGY FROM AN OPEN SCIENCE POINT OF VIEW: CASE STUDIES IN PLANT GENOMICS**

PEDRONI L.\*, CATTIVELLI L.\*, FACCIOLI P.\*

\*) CREA Centro di Ricerca Genomica e Bioinformatica

*open science, computational reproducibility, bioinformatics, bigdata*

Open Science is a transparent and accessible knowledge developed and shared through collaborative networks. It is thus a new approach to the scientific process which has been made possible by new technologies and collaborative tools. Reproducibility is a fundamental component of a reliable and open research activity: experiments should always be verifiable, at all steps, through open data and open methods. This need is particularly challenging in genomics, where next generation sequencing technologies (NGS) have produced an explosion of data, thus increasing the need of applying complex computational analysis. This work is thus focused on a wide-spread issue: the difficulty of reproducing an accomplished by others computational work and of keeping our work as much reproducible as possible for others and for the future us. As a case study, the identification and characterization of intronless genes in barley are here reported. Several critical points have been highlighted both when trying to reproduce other authors' work on the same subject and our own developed pipeline. A reproducibility checklist to test our own work before publishing has thus been developed and a set of tools selected to be efficient, easily managed also to non bioinformaticians and aimed at ensuring reproducibility of complex genomic pipelines has been applied to the intronless genes analysis. The work presented can thus be taken as an example of high reproducible computational working mode.