

TOA: a software package for automated functional annotation in non-model plant species

Fernando Mora-Márquez¹, Víctor Chano¹, José Vázquez-Poletti², and Unai López de Heredia¹

¹Universidad Politécnica de Madrid

²Universidad Complutense de Madrid

July 30, 2020

Abstract

Functional annotation aims to assess the biochemical and biological functions of sets of genomic or transcriptomic sequences yielded after next-generation sequencing experiments. One common way to perform functional annotation of a set of sequences obtained from a next-generation sequencing experiment, is by searching for homologous sequences and accessing to the related functional information deposited in genomic databases. Functional annotation is especially challenging in de novo assemblies of transcriptomes of non-model organisms, like many plant species. In such cases, existing commercial and open access general purpose applications may not offer complete and accurate results. We present TOA (Taxonomy-oriented annotation), a user-friendly open-access application designed to establish functional annotation pipelines geared towards non-model plant species. TOA performs homology searches against proteins stored in the PLAZA platform databases, NCBI RefSeq Plant, Nucleotide Database and Non-Redundant Protein Sequence Database, and retrieves functional information for several gene ontology systems. The software performance was validated by comparing the runtimes, total number of annotated sequences and accuracy of the functional information obtained for several plant benchmark datasets with TOA and other open-access functional annotation solutions. TOA outperformed the other software in terms of number of annotated sequences and accuracy of the annotation, and constitutes a good alternative to improve functional annotation in plants. TOA is especially recommended for gymnosperms or for low quality sequence datasets of non-model plants.

Hosted file

Mora-Marquez-TOA-R1.doc available at <https://authorea.com/users/347523/articles/473180-toa-a-software-package-for-automated-functional-annotation-in-non-model-plant-species>