



**HAL**  
open science

## **ViSEAGO: Easier data mining of biological functions organized into clusters using Gene Ontology and semantic similarity**

Christelle Hennequet-Antier, Aurélien Brionne, Amélie Juanchich

### ► To cite this version:

Christelle Hennequet-Antier, Aurélien Brionne, Amélie Juanchich. ViSEAGO: Easier data mining of biological functions organized into clusters using Gene Ontology and semantic similarity. Journées Ouvertes Biologie, Informatique et Mathématiques JOBIM 2019, Jul 2019, Nantes, France. 1 p, 2019. hal-02786990

**HAL Id: hal-02786990**

**<https://hal.inrae.fr/hal-02786990>**

Submitted on 5 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# ViSEAGO: Easier data mining of biological functions organized into clusters using Gene Ontology and semantic similarity

Christelle HENNEQUET-ANTIER<sup>1§</sup>, Aurélien BRIONNE<sup>1§</sup> and Amelie JUANCHICH<sup>1§</sup>

<sup>1</sup> BOA, INRA, Université de Tours, 37380, Nouzilly, FRANCE

<sup>§</sup> Contributed equally to this work

Corresponding Author: [Christelle.hennequet-antier@inra.fr](mailto:Christelle.hennequet-antier@inra.fr)

## Abstract

The main objective of *ViSEAGO* workflow is to carry out a data mining of biological functions and establish links between genes involved in the study. We developed *ViSEAGO* in R to facilitate functional Gene Ontology (GO) analysis of complex experimental design with multiple comparisons of interest. It allows to study large-scale datasets together and visualize GO profiles to capture biological knowledge. The acronym stands for three major concepts of the analysis: *Visualization*, *Semantic similarity* and *Enrichment Analysis of Gene Ontology*. It provides access to the last current GO annotations, which are retrieved from one of NCBI EntrezGene, Ensembl or Uniprot databases for available species. *ViSEAGO* extends classical functional GO analysis to focus on functional coherence by aggregating closely related biological themes while studying multiple datasets at once. It provides both a synthetic and detailed view using interactive functionalities respecting the GO graph structure and ensuring functional coherence supplied by semantic similarity. *ViSEAGO* has been successfully applied on several datasets from different species with a variety of biological questions. Results can be easily shared between bioinformaticians and biologists, enhancing reporting capabilities while maintaining reproducibility.

*ViSEAGO* is publicly available on <https://forgemia.inra.fr/umr-boa/viseago>.