

R-SYST network : a French consortium for barcoding and species identification

Agnes Bouchez, Marc Buée, Philippe Chaumeil, Valerie Laval, Marion Le Saux, Marie-Anne Auger-Rozenberg, Alain Franc

▶ **To cite this version:**

Agnes Bouchez, Marc Buée, Philippe Chaumeil, Valerie Laval, Marion Le Saux, et al.. R-SYST network : a French consortium for barcoding and species identification. Journées annuelles de la Société Française de Systématique : "Systématique et sciences participatives", Oct 2013, Paris, France. 1 p. hal-01209971

HAL Id: hal-01209971
https://hal.science/hal-01209971

Submitted on 6 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.

R-SYST network : a French consortium for barcoding and species identification

Key words: Systematics, network, barcoding, identification, graph clustering

A. Bouchez [1], M. Buée [2], P. Chaumeil[3], V. Laval[4], M. Le Saux[5], M.A. Auger-Rozenberg[6], the R-SYST consortium and Alain Franc[3]

[1] *INRA* de Thonon les Bains, UMR CARRTEL , 74203 Thonon , FRANCE*

[2] *INRA* de Nancy, UMR IAM , 54280 Champenoux, FRANCE*

[3] *INRA *de Bordeaux, UMR BIOGECO, Pierroton, 33612 Cestas , FRANCE*

[4]*INRA*de Grignon, UMR1290 BIOGER-CPP, 78850 Thiverval-Grignon, FRANCE*

[5]*INRA * d'Angers-Nantes, UMR 1345 IRHS, 49070 Beaucouze, France*

[6]*INRA* d'Orléans, URZF 45, 45075 Orléans, FRANCE*

agnes.bouchez@thonon.inra.fr;
buee@nancy.inra.fr;
chaumeil@pierroton.inra.fr
vlaval@versailles.inra.fr;
marion.lesaux@angers.inra.fr;
Marie-Anne.Auger-Rozenberg@orleans.inra.fr;
Alain.Franc@pierroton.inra.fr;

R-SYST is a French national network of around a dozen research teams including technicians, researchers and engineers in the fields of molecular biology, genetics and bioinformatics who are involved in the molecular and morphological characterization of organisms. The network's ambition is to develop a dictionary between genetic and phenotypic variability, thus enabling a better characterization of these organisms, and to compare and study informations brought by morphology and molecular biology for systematics and taxonomy. The objective is to develop a series of computer tools (for scientists and non-specialist professionals) for the identification of species based on a network of expertise in the fields of morphological based taxonomy, molecular biology and bioinformatics. The strength of this network is that involved specialists have experience in the description, classification and recognition of organisms from a large panel of taxonomic groups. Currently five groups are included in this network : plant associated bacteria, freshwater micro-algae (diatoms), trees, invertebrates (insects, nematodes) and fungi. Common questions regarding identification are shared between specialists of the different taxa.

R-SYST consortium is also aiming at developing bioinformatics tools for analysis and sharing them via its web site. Our consortium wishes to propose a complete tool suite devoted to identification analysis. One part simply is to deliver user friendly access to classical tools such as alignment and phylogeny inference, and another part is to design tools specifically for species identification based on consistency (or not) between characters and clusters from sequences. One of them presented here is “declic”, which enables clusterization and visualization of a reference sequence data set up to a few thousands specimen. The skeleton of this tool is graph clustering (building communities on graphs), where a graph is produced from a pairwise genetic distance array using a threshold, sometimes referred to as “barcoding gap” in barcoding communities. It is expected that individuals belonging to a same taxa build a clique, and the tool permits an analysis between actual situation and this ideal one. Results will be shown on French Guiana tree flora, and on a freshwater diatom library. Accuracy and reliability of this tool will be discussed.

* Institut National de la Recherche Agronomique