

Martial BRIAND¹, Corinne RANCUREL², Ludovic COTTRET³ et Sébastien CARRERE³

¹ Institut de Recherche en Horticulture et Semences, UMR 1345, INRA, SFR 4207 QuaSaV, 42 rue Georges Morel, 49071 BEAUCOUZE CEDEX, France
² Institut Sophia Agrobiotech, UMR INRA-CNRS 1355/7254, 400 route des Chappes, F-06903, Sophia-Antipolis, France
³ Laboratoire des Interactions Plantes-Microorganismes, UMR INRA-CNRS 441/2594, Chemin de Borderouge, F-31320, Castanet-Tolosan, France

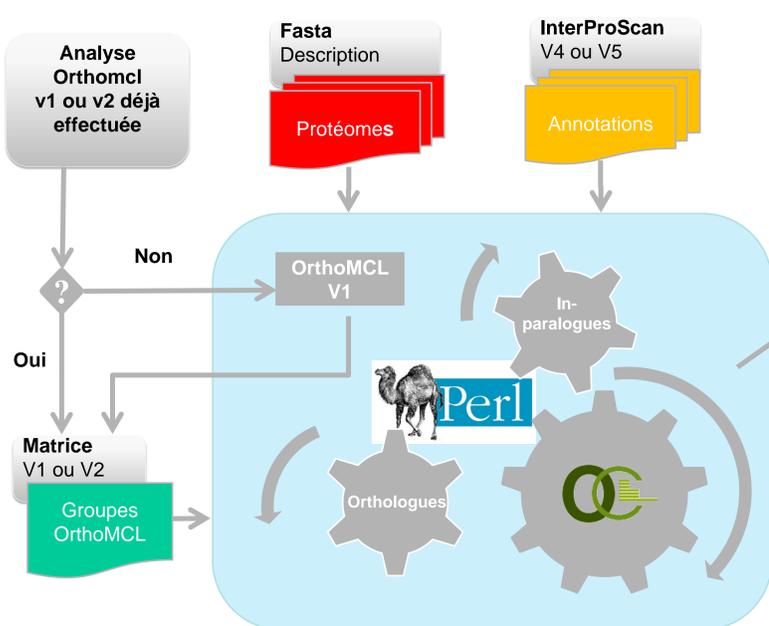
Principe

La multiplication des programmes de séquençage et l'automatisation des protocoles d'annotation structurale permettent aujourd'hui d'effectuer de nombreuses analyses de génomique comparée. OrthoMCL est un outil largement utilisé afin de comparer les protéomes de dizaines d'organismes et de construire des familles de gènes homologues. Cependant les résultats fournis par OrthoMCL (une matrice dont les lignes représentent les groupes de gènes homologues et les colonnes les identifiants des gènes constituant ce groupe et l'organisme dont ils sont issus) sont difficilement exploitables et interprétables pour un utilisateur ne maîtrisant pas un langage de script.

Nous avons développé une application web, Orthomcl-Companion, pour faciliter l'exploitation de ces résultats par les utilisateurs biologistes. Le site est accessible à tout membre de la Fédération Education-Recherche grâce aux informations de compte propres à chaque institution.



Analyse et extraction de données



Orthomcl-Companion permet d'accéder facilement aux protéines spécifiques de chaque organisme, mais aussi au core-protéome et au pan-protéome, ou encore de générer des matrices d'abondance et de présence/absence (profils phylogénétiques). En y associant des annotations InterPro, les données de sortie sont enrichies de tables d'occurrence de termes des différentes banques (ex: PFAM, GO).

Interprétation facilitée des résultats

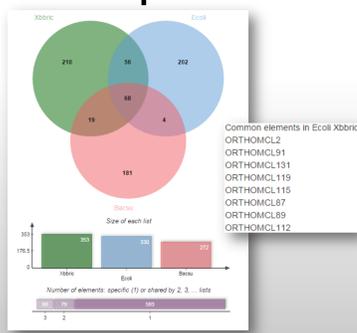
Tableaux et graphiques de synthèse

GROUP	NUMBER OF SPECIES	NUMBER OF PROTEINS	PROTEIN REVIEW	PROTEIN NAME	PROTEIN ID	PROTEIN DESCRIPTION	Remark
ORTHOMCL0	2	21	0	GO:0005498	GO:0005498	Plasma membrane	
ORTHOMCL1	3	14	0	GO:0005498	GO:0005498	Plasma membrane	
ORTHOMCL2	1	12	0	GO:0005498	GO:0005498	Plasma membrane	
ORTHOMCL3	3	4	0	GO:0005498	GO:0005498	Plasma membrane	

GROUP	Bacsu	Ecoli	Xbbric
ORTHOMCL0	0	19	2
ORTHOMCL1	11	2	1
ORTHOMCL2	0	11	1
ORTHOMCL3	0	10	1
ORTHOMCL4	0	10	1
ORTHOMCL5	5	2	0
ORTHOMCL6	5	1	0
ORTHOMCL7	5	1	0
ORTHOMCL8	4	1	0
ORTHOMCL9	1	3	0
ORTHOMCL10	1	2	0
ORTHOMCL11	3	1	0
ORTHOMCL12	3	1	0



Comparaisons

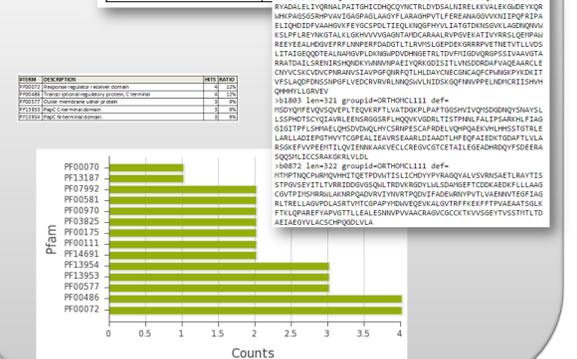


JVENN

Les fichiers générés utilisent des formats standards (textes tabulés, fasta) pour une exploitation ultérieure facilitée. Des représentations graphiques sous forme d'histogrammes et de diagrammes de Venn interactifs sont proposées. L'ensemble des résultats est accessible via une interface web et les utilisateurs ont accès à leur historique d'analyses.

Extraction des sous-ensembles

#seqid	groupid	length
d2146	ORTHOMCL103	412
d4398	PHOSPHO-CREB	229
d2878	ORTHOMCL103	1032
d1803		
d0872		
d4398		
d1757		
d2521		
d0718		
d2406		
d0940		
d4113	BASR-MONOMER	
d2964		
d4113		
d0532		



Protein family analyses

Ext JS

Perspectives

Nous envisageons de développer une interface offrant la possibilité à l'utilisateur :

- d'extraire les groupes auxquels appartiennent des gènes d'intérêts et les séquences de ces groupes
- d'extraire les groupes pour lesquels au moins une protéine possède une annotation particulière
- d'exécuter des analyses d'enrichissement de termes issus d'InterPro

Références

- Philippe Bardou, Jérôme Mariette, Frédéric Escudie, Christophe Djemiel and Christophe Klopp. jvenn: an interactive Venn diagram viewer. *BMC Bioinformatics* 2014, 15:293 doi:10.1186/1471-2105-15-293
- Li Li, Christian J. Stoeckert, Jr., and David S. Roos. OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. *Genome Res.* 2003 13: 2178-2189