

Marion Verdenaud¹; Sébastien Carrere¹; Sébastien Letort¹; Emeline Deleury²; Erika Sallet¹; Emmanuel Courcelle¹; Olivier Stahl³; Thomas Faraut³; Vincent Savoia⁴; Karine Gallardo⁴; Frédéric Debelle¹; Pascal Gamas¹; Jérôme Gozy¹

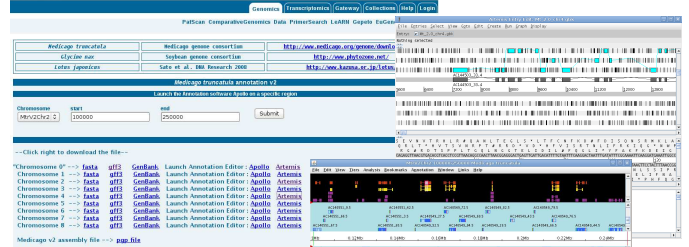
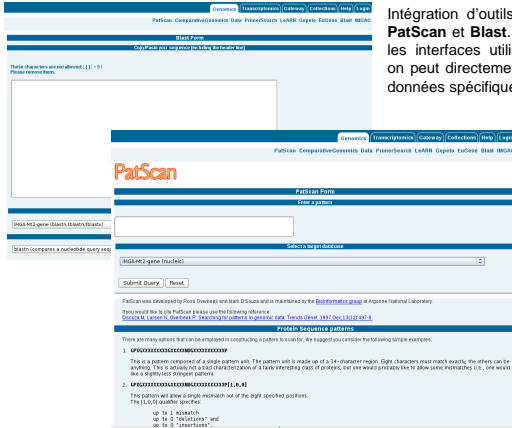
- ¹ Laboratoire des Interactions Plantes Micro-organismes (LIPM), UMR CNRS- F-31320 Auzeville
- ² Interactions biotiques en santé végétale, INRA/CNRS/Univ. Nice, F-06900 Sophia Antipolis
- ³ Laboratoire Génétique Cellulaire UMR444, INRA/ENVT, F-31320 Castanet Tolosan
- ⁴ Unité de Rech. en Génétique et Ecophysiologie des Légumineuses à Graines, INRA, F-21110 Bretenières

Objectifs

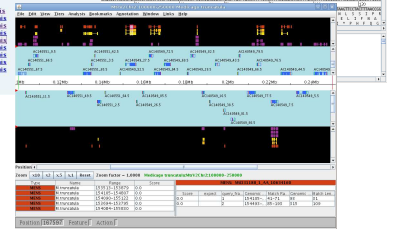
La biologie intégrative végétale peut être définie comme une intégration multidimensionnelle. Ainsi il convient, dans un premier temps, d'intégrer différents types de données « haut-débit ». Un second axe d'intégration doit permettre l'interprétation des données moléculaires à des niveaux d'organisation plus importants (cellule, tissu, etc.). La troisième dimension de l'intégration correspond au transfert de connaissances entre les espèces modèles et d'intérêts agronomiques, rendu possible par la comparaison des cartes génétiques et des génomes. « Legoo » fournit de nombreuses ressources le long de ces trois axes d'intégration. La particularité de « Legoo » est de se focaliser sur la gestion et la représentation des informations à partir de réseaux de gènes et de comparaisons inter-espèces, proposant ainsi de multiples points de vue sur les données et les connaissances acquises chez les légumineuses. L'intégration technique entre les différentes vues proposées est assurée par l'intermédiaire de web services BioMoby[6]. Ces derniers rendent également possible un accès programmatique, par la communauté bioinformatique, aux différentes ressources que nous mettons à disposition.

Présentation du portail

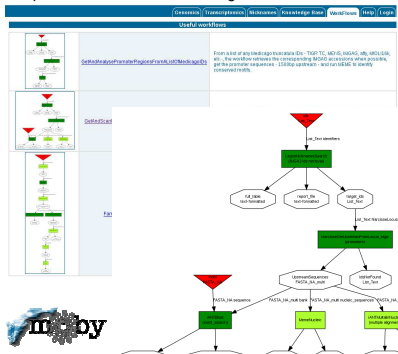
Intégration d'outils de référence tels que **PatScan** et **Blast**. Normalisation de toutes les interfaces utilisateurs avec lesquelles on peut directement interroger les jeux de données spécifiques légumineuses



Intégration de 2 outils de visualisation d'annotation : **Apollo**[2] et **Artemis**[1]. Apollo interroge une base de données CHADO, contenant l'annotation génomique ainsi que des comparaisons protéiques avec des banques de références. L'utilisation des versions Web Start d'Apollo[2] et d'Artemis[1] permet de simplifier au maximum la mise en œuvre par les biologistes puisque qu'aucune installation préalable des logiciels n'est nécessaire.



De nombreux **web services BioMoby**[6] découlant des outils du portail (outils de références, Nicknames, Knowledge, Narcisse) sont aussi disponibles. Ils peuvent être intégrés dans des workflows via **Remora**[5] ou un autre gestionnaire de workflows BioMoby. L'utilisateur peut ainsi créer des chaînes d'analyse, reproductibles et archivables. Sur le site nous proposons une section avec des workflows correspondant aux requêtes les plus fréquentes des utilisateurs de Legoo.

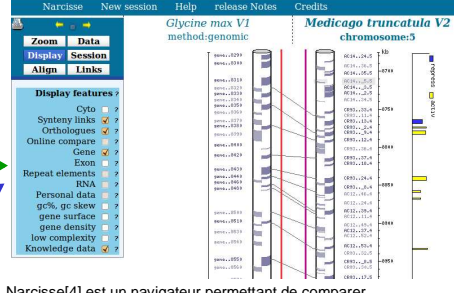


Outils de référence

Annotation génomique

Accès programmatique

Génomique comparée



Narcisse[4] est un navigateur permettant de comparer simultanément plusieurs génomes et ce à plusieurs niveaux de conservation (Poster 30; Letort et al.).

Interrogation directe de Knowledge via Narcisse possible à partir d'un gène ou d'une synténie. A l'issue de la requête, Narcisse proposera une nouvelle piste pour représenter l'ensemble des relations qui aura été renvoyées par le web service.

Structuration et représentation des données

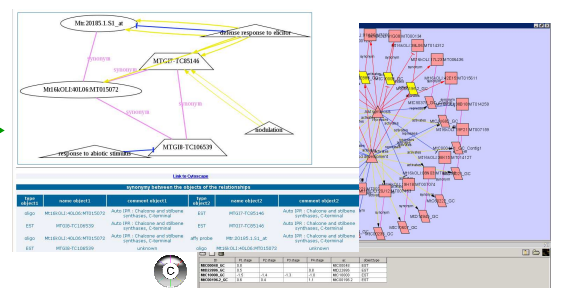
Intégration des « omics »



Pour notre légumineuse de référence *Medicago truncatula* on dispose de nombreux jeux de données utilisant des nomenclatures d'identifiants différentes, ce qui rend les comparaisons difficiles. Pour remédier à cela, **Nicknames**, pour un identifiant donné, va proposer la liste des identifiants « synonymes » dans les autres jeux de données. Ce lien de synonymie est déterminé à partir de similarités de séquences pré-calculées. De plus la majorité des séquences disponibles a été réannotée automatiquement avec **InterproScan** afin de proposer une annotation fonctionnelle homogène.

Accession	Accession	Accession	Accession	Accession	Accession	Accession	Accession	Accession	Accession
ME100001	ME100002	ME100003	ME100004	ME100005	ME100006	ME100007	ME100008	ME100009	ME100010
ME100011	ME100012	ME100013	ME100014	ME100015	ME100016	ME100017	ME100018	ME100019	ME100020

Nicknames est utilisé automatiquement pour gérer les relations de synonymie dans Knowledge. On peut donc interroger la base de données à partir de n'importe quelle nomenclature.



Références

1. T. Carver, M. Berriman, A. Tivey, C. Patel, U. Bohme, B.G. Barrell, J. Parkhill and M.A. Rajandream. Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database. (2008) *Bioinformatics*, **24**, 2672-2676.
2. S. Misra and N. Harris. Using Apollo to browse and edit genome annotations. (2006) *Curr Protoc Bioinformatics*, **Chapter 9**, Unit 9.5.
3. P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski and T. Ideker. Cytoscape: a software environment for integrated models of biomolecular interaction networks. (2003) *Genome Res*, **13**, 2498-2504.
4. E. Courcelle, Y. Beausse, S. Letort, O. Stahl, R. Fremez, C. Ngom-Bru, J. Gozy and T. Faraut. Narcisse: a mirror view of conserved synteny. (2008) *Nucleic Acids Res*, **36**, D485-490.
5. Carrere S. and Gozy J. REMORA: a pilot in the ocean of BioMoby web-services. (2006) *Bioinformatics* **22**(7):900-1
6. Wilkinson, MD, Links, M. BioMOBY: an open-source biological web services proposal. (2002). *Briefings In Bioinformatics* **3**(4). 331-341.