



GnpIS, a portal and information system to bridge genetics and genomics plant and fungi data



Delphine Steinbach¹, Aminah-Olivia Keliet¹, Erik Kimmel¹, Michael Alaux¹, Joelle Amselem¹, Sophie Durand¹, Cyril Pommier¹, Nacer Mohellibi¹, Nathalie Choisei¹, Daphné Verdelet¹, Isabelle Luyten¹, Nicolas Lapalu¹, Sébastien Reboux¹ and Hadi Quesneville¹
URGI (Unité de Recherche Génomique-Info), INRA, Centre INRA de Versailles, bâtiment 18, RD 10, 78000, Versailles, France ; urgi-contact@versailles.inra.fr

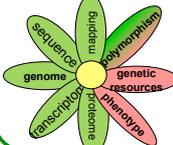
Summary : URGI (Unité de Recherche Génomique-Info) is an INRA bioinformatics unit dedicated to plants and pest genomics. Its mission is to develop and host a genomic and genetic information system called GnpIS, for plants of agronomical interest and their bioaggressors. It hosts a bioinformatics platform labeled and state-approved at institute and national level. URGI maintains an efficient computing environment and offers services covering database conception, software engineering, and bioinformatics. A focus is done on doing an **interoperability** between both the **tools** (a set of Oracle, PostgreSQL, MySQL databases and their interfaces in Java, Perl) and the **data** located in all these databases (for example on Wheat and Grapevine data). 2 new tools were developed to cross query these data, i) one tool (**a quick search tool**) based on Lucene (a high-performance, full-featured text search engine library, <http://lucene.apache.org>), ii) the other one (**an advanced search tool**) based on Biomart (a query-oriented data management system <http://www.biomart.org>) and Galaxy (<http://galaxy.psu.edu>). They are freely available at this url: <http://urgi.versailles.inra.fr/gnpis>, (D. Steinbach & al. paper in prep). These tools are developed according to **user needs**, geneticists or breeders in the frame of partnership through scientific projects. The goal of this portal and this information system GnpIS, is to help scientists in their research on **finding genes responsible for traits of agronomical interest** (QTLs), to explore diversity (SNPs) and to select the best allele for breeding, for example to resist to bioaggressors or to adapt plant to climate variations.

GnpIS description

GnpIS information system is composed of 8 modules: 5 genomic modules and 3 genetic modules and above all these modules, **A portal allowing to query through the different modules separately (via the flower) or simultaneously via datamarts or indexes (publication D. Steinbach & al. in prep. 2014).**

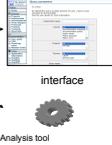
- 1) GnpSeq, the EST, mRNA sequence database which contains home-made clusters, contigs and annotations
- 2) GnpMap, the genetic mapping database
- 3) GnpGenome, a multispecies database based on Chado/Gbrowse GMOD tools, containing genomic sequences and their annotations
- 4) GnpArray, a MIAME compliant expression database
- 5) GnpProt, the proteomic database
- 6) GnpSNP, the polymorphism (SNP/DIP/STR) database
- 7) SiReGal, Genetic Resources database for accessions collections and phenotypic descriptors
- 8) Ephesis (development in progress), for genotypes, trials and phenotypes linked to environment database

Biologist Data

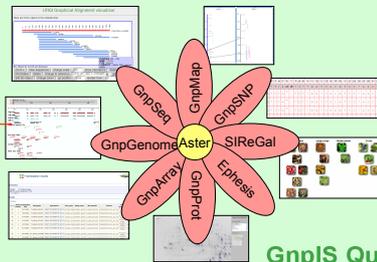


GnpIS data flow

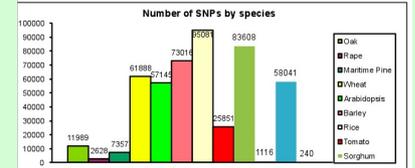
Data files



<http://urgi.versailles.inra.fr>



GnpSNP database summary 2010



GnpIS Quick search tool

Poplar

GnpIS Advanced search tool

Poplar

- Choose Poplar database
- Query by QTL
- Markers list
- Query by marker
- Query by gene name
- Query by clone (Bac) name

Grapevine

- Markers mapped on genome annotated and linked to GnpMap (CM)
- OTLs list
- OTLs mapped on genome annotated sequence
- External link with Gramene
- Query by gene prediction analysis software

Wheat

- Query by clone (Bac) name

Line card

Link an Aurone between: Accession (in Siregal) and Line (in GnpSNP)

Accession: ALRORE
Genotypepage: Cette accession a été génotypée voir ALRORE dans GnpSNP

Deposits number: APP (Géoplatte Valor, INRA)
 • GnpMap : IDDN.FR.001.260001.001.D.P.2003.000.10300
 • GnpSeq : IDDN.FR.001.21009.000.D.P.2002.000.10300
 • URGI-Extractor : © CeILL Licence

Publication: Géoplatte-Info (GPI) : a collection of databases and bioinformatics resources for plant genomics.
 D.Samson (Steinbach) & al. Nucleic Acids Research, 2003, Vol.31, No.1, 179-182.
 Full references: <http://urgi.versailles.inra.fr/index.php/urgi/Platform/Publications>

URGI Contact: urgi-contact@versailles.inra.fr
INRA URGI Web site: <http://urgi.versailles.inra.fr>

Acknowledgements:
 - for fundings: Genoplante grants from 2000-2006, ANR Genoplante (french national research agency) in 2007, 2008 and INRA
 - for data: all researchers for providing data and especially for wheat: Philippe Leroy, Etienne Paux, Frédéric Choulet, Catherine Ravel, Catherine Feuillet, Nicolas Guilhot, François Balthazier, Jacques Legouis, Pierre Sourdisse
 - Genetic resources: Jean-Marie Prosper
 - Grapevine: Anne-Françoise Adam-Blondin, Nathalie Choisei, for tree genomics: Christophe Plomion, Camille Pottvin, Véronique Jorge, Isabelle Bourgaill, Patricia Faivre Rampant,
 - for Maize: Alain Charcosset, Frédéric Sapet, Alain Murgueux, Pascal Perez, Mathieu Falque, - for Rice: Pierre Larmande, Emmanuel Guiderdoni, - for Pea: Judith Bustin, for Tomato: Nicolas Ranc, - for Arabidopsis: Fabienne Granier