GnpIS update: two new query tools to bridge plant genetics and genomics data

URGI

INRA — Institut National de la Recherche Agronomique

Delphine Steinbach[1], Erik Kimmel[1], Aminah-Olivia Keliet[1], Michael Alaux[1], Joelle Amselem[1], Sophie Durand[1], Cyril Pommier[1], Isabelle Luyten[1], Nacer Mohellibi[1], Daphné Verdelet[1], Sébastien Reboux[1] and Hadi Quesneville[1]
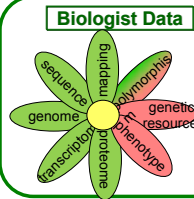
[1] URGI (Unité de Recherche Génomique-Info), INRA, Centre INRA de Versailles, bâtiment 18, RD 10, 78000, Versailles, France ; urgi-contact@versailles.inra.fr

Summary : URGI (Unité de Recherche Génomique-Info) is an INRA bioinformatics unit dedicated to plants and pest genomics. Created in 2002, one of its mission is to develop and to maintain a genomic and genetic information system called GnpIS, for INRA plants of agronomical interest and their bioagressors. It hosts a bioinformatics platform which belongs to the french ReNaBi bioinformatics platform network. It has a national bioinformatics platform label (IBISA 2009). The URGI unit maintains for its users, an efficient computing environment and offers services covering database conception, software engineering, and bioinformatics. Since 2008, a focus is done on doing an interoperability between both the tools (a set of Oracle, PostGreSql, MySql databases, their interfaces in Java, Perl) and the data located in all these databases (for example on Wheat and Grapevine data). We will present here the last databases developments, already available since January 2009 on our public web site at this url: http://urgi.versailles.inra.fr/gnpis. It is the first version of a new portal to bridge plant genetics and genomics data, and 2 new tools to query through all these databases transparently, one tool ("a quick search tool") based on Lucene (a high-performance, full-featured text search enginelibrary, http://lucene.apache.org,), the other one ("an advanced search tool") based on Biomart (a query-oriented data management system http://www.biomart.org). Examples will be shown on Wheat, GrapeVine and Poplar data.
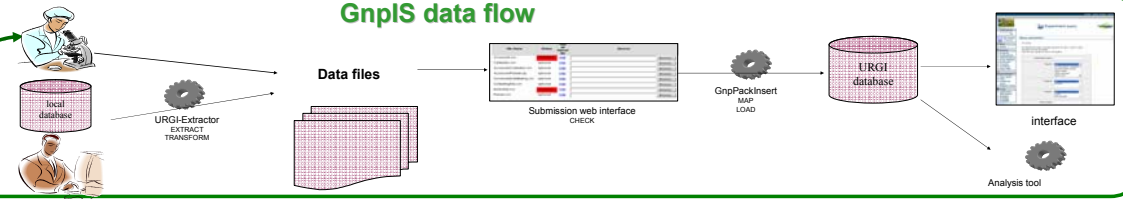
GnpIS description

The GnpIS information system is composed of 8 modules: 5 genomic modules and 3 genetic modules and above all these modules, *a new portal allowing to query through the different modules separately (via the flower) or simultaneously via datamarts or indexes (publication in prep. 2009).*

1) GnpSeq, the EST, mRNA sequence database which contains home-made clusters, contigs and annotations
2) GnpMap, the mapping database
3) GnpGenome, a multispecies database based on Chado/Gbrowse GMOD tools, containing genomic sequences and their annotations
4) GnpArray, the 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
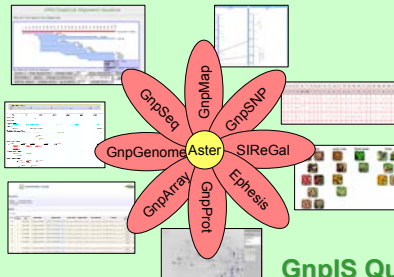8) Ephesis (development in progress), the genotypes, environment and experimentation database

GnpIS data flow

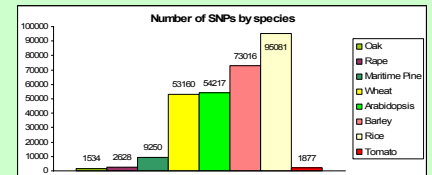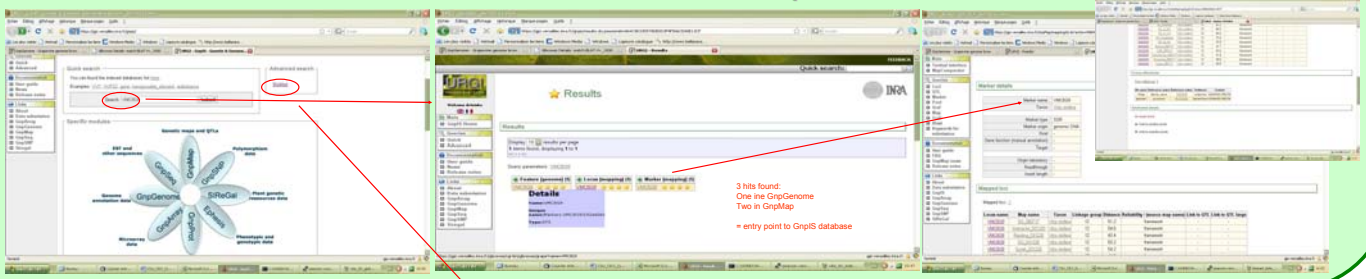
Biologist Data → local database → URGI-Extractor (EXTRACT TRANSFORM) → Data files → Submission web interface (CHECK) → GnpPackInsert (MAP LOAD) → URGI database → interface → Analysis tool

http://urgi.versailles.inra.fr

GnpSNP database summary

Number of SNPs by species

Species	Number
Oak	1534
Rape	2628
Maritime Pine	9250
Wheat	53160
Arabidopsis	54217
Barley	73016
Rice	95081
Tomato	1877

Flower modules: GnpSeq, GnpMap, GnpSNP, SiReGal, Ephesis, GnpProt, GnpArray, GnpGenome — Aster

GnpIS Quick search tool

3 hits found:
One ine GnpGenome
Two in GnpMap

= entry point to GnpIS database

GnpIS Advanced search tool

Poplar

Choose Poplar database

Query by QTL — QTLs list — QTLs mapped on genome annotated sequence

Query by marker — Markers list — Markers mapped on genome annotated (bp) sequence and linked to GnpMap (CM)

Query by gene name

External link with Gramene database

Query by gene prediction analysis software

Grapevine

Wheat

Query by clone (Bac) name

**Poplar data from V. Jorge et al. and public Genbank, Paper in prep.

** Grapevine data from Grapevine consortium, paper in prep.
See: Poster GnpGenome Grapevine for details

** Wheat Data from F. Choulet et al., paper in prep
See: Poster GnpIS Wheat for details.

GnpSNP Line card — Line details, Line complements

Siregal Accession card

Link between: Accession (in Siregal) and Line (in SNP)