

Gramene News

FAQ ... p3

Upcoming Outreach ... p3

Community News

Sequencing and Accessing the Maize Genome ... p 1

Gramene's Diversity Advanced Search ... p 2

Opportunities ... p 3

Recommended Reading ... p4

Maize is a staple food source for many people, especially in Africa. The graph below shows the top countries consuming corn for human nutrition. The countries are listed in order based on kg/capita/year, but you will notice that the calories from that supply differ according to the nutritional value of maize grown in or imported into that region.



Sequencing and Accessing the Maize Genome

Maize has an estimated 50,000-60,000 genes along 2.5 billion base pairs in 10 chromosomes. In comparison rice has about 40,000 non-transposable elements (genes/ loci), located on 400 Mbp in 12 chromosomes. Large regions in the maize genome do not contain genes - therefore sequencing efforts will be focused only on the gene-containing regions.

As a part of the Maize Sequencing project, the Maize Genome Browser at www.maizesequence.org was created to display annotated BAC clones and other data assembled in the process of sequencing and mapping. This browser provides entry points to the maize genome by searching or browsing by sequence accession, physical position, genetic position, and by conserved synteny with rice. Displayed features include predicted genes, markers, repeats, and expressed and conserved regions. A BLAST engine is also available.

The genome browser operates at three levels of genomic resolution. In **MapView**, the highest, most global perspective, the genome is displayed at the fingerprint contig level (based on the July 2005 release of the AGI (Arizona Genomics Institute) agarose FPC map). The maize core bin map is displayed alongside the physical map and the density of markers, clones, and accessioned BACs provide an overview

of genome organization as well as progress of the Project. **CytoView** provides detailed visualization of the AGI physical map. BAC clones, hybridized markers (including many classes of overgos), and simple sequence repeats are displayed here. Sequenced clones in GenBank are color-coded depending on their source and level of annotation. This view also serves as a launching point for **ContigView**, which displays the genome at the nucleotide level of resolution. ContigView is provided for clones that have reached HTGS_IMPROVED status and for clones publicly available outside of the current project. It displays primary sequence annotations such as *ab initio* gene prediction (FGeneSH), repeats, and alignment to EST, cDNA, and GSS sequences. In the future, ContigView will also serve as a launching point to view



places the maize views previously made available. It provides visual enhancements, such as a universal navigation bar, highlighted FPC clone-marker associations, and vertical shading of sequenced contigs and improved regions within BACs. Automated annotation of improved BACs is performed frequently in close collaboration with Gramene gene builds. Comprehensive deep links between Gramene and the maize browser provide for a seamless genome browsing experience.

The Maize Sequencing Consortium is composed of investigators from the Washington University School of Medicine, Genome Sequencing Center at St. Louis (lead institution), the Arizona Genome Institute, Iowa State University and Cold Spring Harbor Laboratory. This project is supported by a joint grant from the National Science Foundation, the U.S. Department of Agriculture and the Department of Energy, with the mission to sequence the maize genome, to finish the gene space of the genome, and to anchor the assembled sequence to the genetic map.

secondary annotations such as comparative maps (CMap), orthologue/paralogue/synteny analysis (Compara), and protein sequence characterization (Interpro/GO).

www.maizesequence.org is an extension to Gramene and re-

Article contributions by Sandra W. Clifton, Assistant Director, Genome Sequencing Center, Washington University School of Medicine, and Shiran Pasternak, Project Manager, www.maizesequence.org, Cold Spring Harbor Laboratory.

Gramene's Diversity Advanced Search: Highly-Flexible User Front-End Built on GDPC Middleware

Gramene's Genetic Diversity databases contain genetic data (i.e. SNPs, SSRs, sequences), phenotypic data (i.e. trait measurements), environment data (i.e. planting location, experiment), and germplasm data (i.e. stock, pedigree, passport). Currently three crops (maize, wheat, and rice) are stored with varying amounts of those data types. Several web-based tools provide access to these cereal databases to meet the needs of many users. In instances where more complex queries are required, Gramene's **Diversity Advanced Search** (www.gramene.org/diversity/gramene_gdpc.html) gives users more flexibility when retrieving data.

The Advanced Search (a.k.a. GDPC Browser) uses the GDPC middleware (www.maizegenetics.net/gdpc) to access databases. The middleware is a communications layer that makes it possible to retrieve data remotely via web services. The middleware also masks the complexities of various databases making it possible to query and integrate data from multiple data sources.

Although the middleware is important infrastructure, users of the Advanced Search only need to understand the various controls to construct a query and retrieve data. After choosing which crop database to access, users can build lists of data. This data is retrieved based on search criteria specified by the user. Once retrieved, the data can be manipulated in various ways including exporting to standard file formats.

Short example to get started...

- Go to the Diversity Advanced Search (www.gramene.org/diversity/gramene_gdpc.html)

- Click on "Get Maize Diversity Data."

NOTE: This may take a few minutes the first time.

- When it asks "Do you want to run the application?", click "Run." The application should start-up and show that it is connected to "Gramene Diversity Maize Database."

- Each labeled tab corresponds to a type of data that you can retrieve. First go to the tab labeled "Taxa."

- Check the box next to "Germplasm Type," and then select the germplasm type "Inbred." Also, check the box next to "Source," and then select the source "CIMMYT."

- Then click "Get Data" to retrieve all inbred taxa with source CIMMYT.

- Next go to the tab labeled "Loci."

- Check the box next to "Chromosome Name," and then select

chromosome name "1."

- Click "Get Data" to retrieve all loci on chromosome 1.
- Next click on the tab labeled "Genotype Experiments."
- The list labeled "Locus (working list)" will contain all the loci that you retrieve in the previous steps. Highlight all the loci in this list by first clicking (to highlight) any of the loci. Then press <Ctrl> <a> to select all loci in the list.
- Check the box next to "Polymorphism Type," and select polymorphism type "SNP"
- Click "Get Data" to retrieve all the genotype experiments associated with loci on chromosome 1 designed to score SNPs.
- Next click on the tab labeled "Genotypes." Items in both lists "Taxa (working list)" and "Genotype Experiment (working list)" result from the actions above. All items in both lists should already be highlighted.
- Click "Get Data" to retrieve all genotypes for the given taxa and genotype experiments.

Please feel free to experiment with other functionality of the Advanced Search. Notice that some search criteria may result in very large data sets, causing long wait times. Send any problems to tmc46@cornell.edu.

The screenshot shows the GDPC Browser window with the 'Genotypes' tab selected. The main table displays SNP data for various CML accessions across different chromosomes. Below the table are two working lists: 'Taxa (working list)' and 'Genotype Experiments (working list)'. The taxa list contains CML 10, 103, 108, 11, 14, 154Q, and 157Q. The genotype experiments list contains an1.1, d8.1, d8.2, d8.3, id1.2, id1.3, id1.6, id1.8, id1.9, and tb1.1.

	d8.1	d8.2	d8.3	id1.2	id1.3	id1.6	id1.8	id1.9	tb1.1	tl
CML10	N:N	A:A	A:G	A:A	C:C	C:C	T:T	C:C	C:C	A:A
CML103	A:A	A:A	A:G	A:A	C:C	C:C	T:T	C:C	C:C	A:A
CML108	A:A	A:A	G:G	A:A	C:C	C:C	T:T	C:C	C:C	A:A
CML11	A:A	A:A	A:G	A:A	C:C	C:C	T:T	C:C	C:C	A:A
CML14	N:N	A:A	G:G	A:A	C:C	C:C	T:T	C:C	C:C	G:G
CML154Q	N:N	A:A	A:G	G:G	T:T	T:T	T:T	C:C	C:C	A:G
CML157Q	A:A	A:A	A:G	A:A	C:C	C:C	T:T	C:C	C:C	A:A

Advanced Search after Retrieving SNP Data

Article contributed by Terry M. Casstevens, Institute for Genomic Diversity, and Edward S. Buckler, USDA-ARS, Institute for Genomic Diversity

Find proteins, maps, and QTL for your favorite species by using the Taxonomy Ontology.

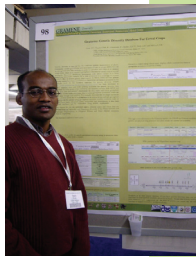
Upcoming Outreach:

Gramene will be at these upcoming meetings. You may attend these workshops and posters to learn about Gramene or get answers to questions. See the Calendar below for dates and locations.

- **CSHL Plant Genome Meeting**
- Posters
- **Maize Genetics Meeting** -
Demonstration with MaizeGDB
- **ITMI** - Poster
- **Biology of Genomes** - Posters
- **ASPB** - Demonstration with
TAIR and SGN
- **Crop Sciences** - Posters



Workshop and Poster
from PAG 2007



Gramene FAQ

Contact Gramene through the "Feedback" link at the top of any page to ask questions. Here are some recent questions that have been answered.

Q: Can I use the ontology term files provided on your website? How do I obtain a license agreement?

A: All of the data available at Gramene is in the public domain, and thus is freely available. Gramene does not require nor administer license agreements. However, as indicated on the web-page at www.gramene.org/plant_ontology/ the different ontologies come from various sources, and these sources should be acknowledged. The Gene Ontology (GO) terms and definitions are from the Gene Ontology Consortium, and the Plant Ontology (PO) terms are from the Plant Ontology Consortium. The Trait Ontology (TO), Environment Ontology (EO) and Taxonomy Ontology (Gr_Tax) are from Gramene. PO now offers Plant Growth and Development stages terms, in addition to Plant Structure terms, and Gramene is now using the PO terms instead of the GRO terms, which are no longer used.

Q: Is it possible to get markers for a particular rice QTL if I have a QTL accession id?

A: From the QTL detail page click "View Comparative Map" from QTL map to go to Maps/CMap for the map display. Your QTL will be highlighted, and you can find associated markers for that QTL.. Presently you have to search each QTL separately, but quite a number of users have requested QTL associated marker info and we are working on that. Look for changes, hopefully in May 2007.

Community Calendar

2007

- February 13.** Arkansas Rice Conference, 2007. Wynne, AR, USA
- March 11-16.** Agricultural Biotechnology: Today and Tomorrow. Ventura, CA, USA.
- March 12 to 14.** North American Wheat Workers' Workshop. Saskatoon, Saskatchewan, Canada
- March 15-18.** CSHL Plant Genome meeting. Cold Spring Harbor, NY, USA.
- March 22-25.** Maize Genetics Meeting, Illinois, USA.
- March 23-27.** 2nd International Conference on Plant Molecular Breeding. Sanya City, Hainan, P. R. China.
- April 11-13.** 5th European Conference on Evolutionary Computation Machine Learning and Data Mining in Bioinformatics. Valencia, Spain.
- April 16-20.** ITMI. Israel
- May 8-12.** Biology of Genomes, Cold Spring Harbor, NY, USA.
- May 31- June 3.** 9th Annual Plant Sciences Institute Symposium on Epistasis: Predicting Phenotypes and Evolutionary Trajectories. Ames, Iowa, USA.
- July 7-11.** ASPB Chicago, Illinois, USA.
- July 21-25.** ISCB: 15th Annual International Conference on Intelligent Systems for Molecular Biology (ISMB) & 6th European Conference on Computational Biology (ECCB). Vienna, Austria.
- August 13-17.** Computational Systems Bioinformatics, UC San Diego.
- November 4-8.** Crop Science, New Orleans, LA, USA
- ### 2008
- January 12-16, 2008.** PAG-XVI. San Diego, CA, USA

Opportunities

Here are opportunities for researchers and students. Please check with each organization to confirm due dates, as they are prone to change.

- **Biology of Genomes.** Abstracts due **2/14/07**. <http://meetings.cshl.edu/meetings/genome07.shtml>
- **ASPB** - Abstracts via the web at www.aspb.org/abstract. Mini-symposium abstracts due **2/14/07**. Others due **4/4/07**.
- **Monocots IV.** Symposium proposals due **6/1/07**. Abstracts for oral contributions and posters due **5/31/08**. www.monocots4.org/
- **ISMB Highlights and Industry Tract abstract** due **3/4/07**, Poster and PLoS abstracts due **4/6/07**. , Travel grant applications due 4/27/07) www.iscb.org/ismbecb2007/
- RiceCAP jobs page lists several opportunities at: www.rice-cap.uark.edu/jobs.htm
- Computational biology position at MaizeGDB. Application period ends 26 February 2007. <http://jobsearch.usajobs.opm.gov/getjob.asp?JobID=52116941&aid=3>

GRAMENE



A GENOMIC RESOURCE FOR CEREALS

BLAST
Genome Browser
CMap
Genetic Diversity
Genes & Alleles
Maps and CMap
QTL

Ontologies
Pathways
Literature
Proteins
Markers
GrameneMart

Gramene is a curated, open-source, web-accessible free data resource for comparative genome analysis in the grasses. Our goal is to facilitate the study of cross-species homology relationships using information derived from public projects involved in genomic and EST sequencing, protein structure and function analysis, genetic and physical mapping, interpretation of biochemical pathways, gene and QTL localization and descriptions of phenotypic characters and mutations.

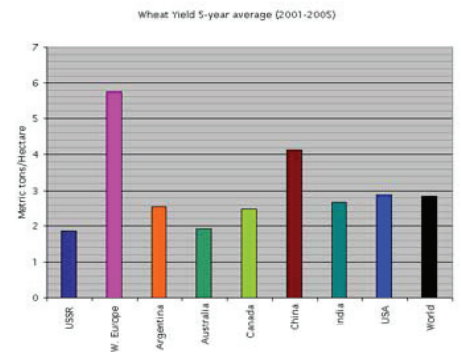
email gramene@gramene.org

For updates on releases and other information, join Gramene's mailing list or view the mailing list archive.

www.gramene.org

Wheat is a significant global crop, in part because it is adapted to many soil types, has a short growing season, offers good yield, and grows well in fairly dry and mild climates (although the highest yielding crops require optimal growing conditions) (see figure for 5-year average yield differences in top-growing countries). Global production has been steadily increasing, but the amount of land used globally to produce that wheat has remained level since 1961, indicating an increase in yield per hectare.

For more information on wheat and other species, see the Gramene Species Pages at www.gramene.org/species.



For world-wide rice news, check out IRRI's news page at <http://ricenews.irri.org/>

Recommended Reading

Genetic Diversity and Origin of Weedy Rice (*Oryza sativa* f. *spontanea*) Populations Found in North-eastern China Revealed by Simple Sequence Repeat (SSR) Markers. Cao *et al.* *Annals of Botany* (Lond). 2006 Dec; 98(6):1241-52. Epub 2006 Oct 20.

AgBase: a unified resource for functional analysis in agriculture. McCarthy *et al.* *Nucleic Acids Res.* 2006; 0:gkl936v1-D5.

Functional Classification, Genomic Organization, Putatively cis- Acting Regulatory Elements, and Relationship to Quantitative Trait Loci, of Sorghum Genes with Rhizome-Enriched Expression. Jang *et al.* *Plant Physiol.* 2006; 142:1148-1159.

The molecular genetics of crop domestication. Doebly *et al.* *Cell*, 2006, 127, pp.1309-1321

Genomic imprinting, methylation and molecular evolution of maize Enhancer of zeste (Mez) homologs. Haun *et al.* *The Plant journal : for cell and molecular biology*, 2007, 49, pp.325-337.

Enrichment of gene-coding sequences in maize by genome filtration. Whitelaw *et al.* *Science*, 2003, 302, pp.2118-2120.

Maize genome sequencing by methylation filtration. Palmer *et al.* *Science*, 2003, 302, pp.2115-2117.

Ensembl 2007. Hubbard *et al.* *Nucleic Acids Research*, 2006, Vol. 00, Database issue D1-D8.

The TIGR Rice Genome Annotation Resource: improvements and new features. Ouyang *et al.* *Nucleic Acids Research*, 2006, Vol. 00, Database issue D1-D5.

Gramene Staff

Lincoln Stein, PI
Susan McCouch, Co-PI
Doreen Ware, Co-PI
Pankaj Jaiswal, Co-PI, Curator
Ed Buckler, Co-PI
Chengzhi Liang, Coordinator
Junjian Ni, Curator
Immanuel Yap, Curator
Isaak Yosief Teclé, Curator
Dean Ravenscroft, Curator
Chih-Wei Tung, Curator
Ken Youens-Clark, Developer
Shulamit Avraham, Developer
Liya Ren, Developer
William Spooner, Developer
Payan Canaran, Developer
Sharon Wei, Developer
Terry Casstevens, Developer
Jim Thomason, Developer
Claire Hebbard, Outreach

To submit items for this newsletter, contact the editor at: cer17@cornell.edu