

## ALIMENTOS FUNCIONALES



Agricultural Research Service (ARS) developed the heart-healthy NuSun sunflower as a variety high in oleic acid, a monounsaturated fatty acid. That variety now accounts for about 77 percent of the sunflowers produced for oil seeds in the United States.



transgenic tomatoes that contain four to eight times as much lycopene, a carotenoid known for its strong antioxidant properties, as nontransgenic

### Nutrigenomics and Beyond: Informing the Future - Workshop Summary

Ann L. Yaktine and Robert Pool, Rapporteurs  
ISBN: 0-309-66921-9, 90 pages, 6 x 9, (2007)



Different plant and seed phenotypes were found to segregate in *Brassica rapa*.

**SenterNovem**

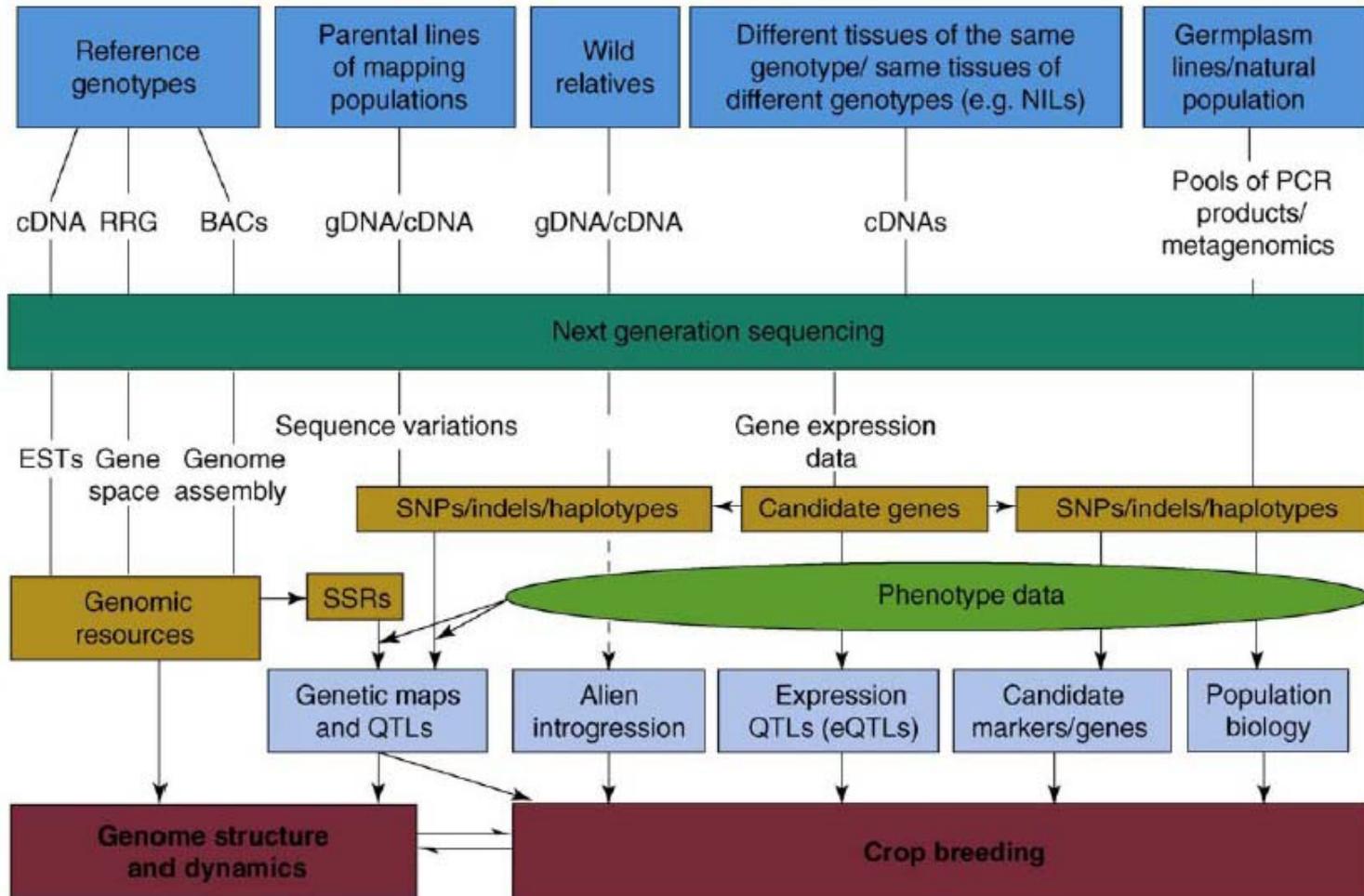


**Brassica vegetable nutrigenomics**

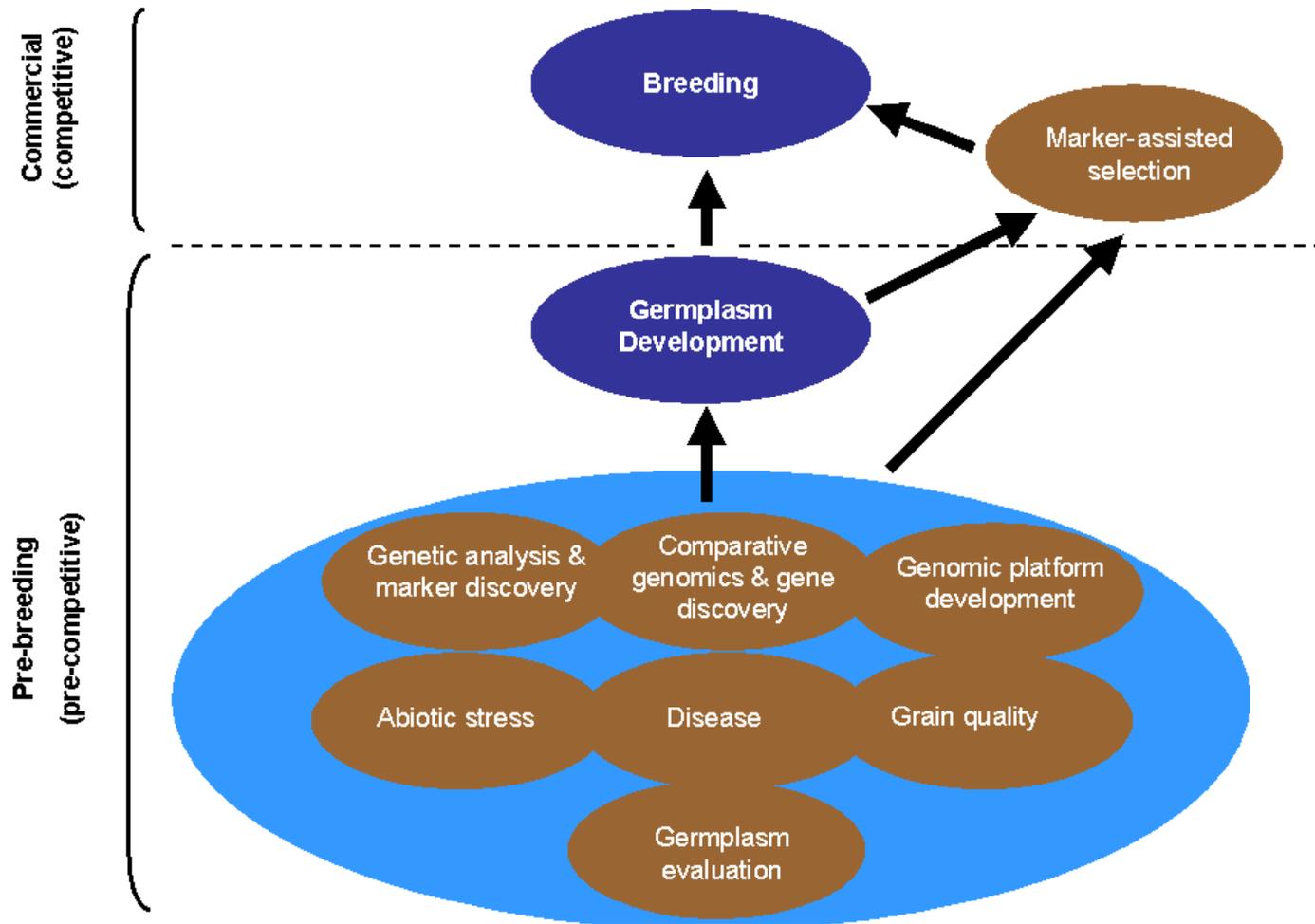
IOP Geomics

By combining metabolomics, genomics and bioinformatics, research groups from three different universities expect to identify genetic biomarkers for metabolites in *Brassica rapa* that make the crop more nutritious. Scientists are testing their approach first on a better-known relative of the turnip: the thale cress (*Arabidopsis*). Three years into the project, the research groups are getting closer and closer to their goal. "The industry is making a genuine commitment to the research."

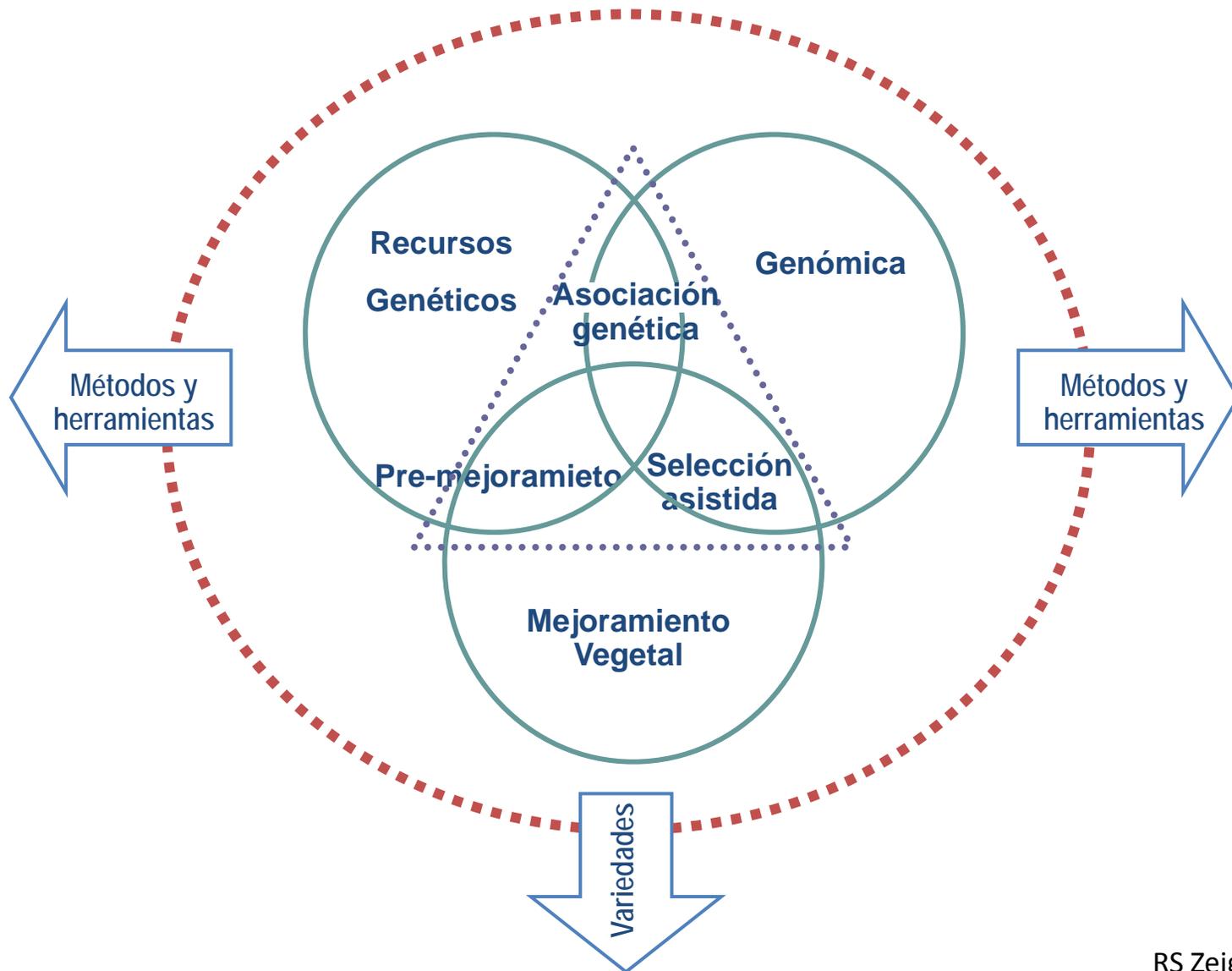
## Aplicaciones de las nueva generación de tecnologías de secuenciación al mejoramiento vegetal



# Pre-mejoramiento y mejoramiento



# Integración de la genómica y el mejoramiento vegetal



# Propósitos de los proyectos genómicos

- Establecer plataformas web que permitan integrar datos, información, herramientas de análisis.
- Ensamblar mapas genéticos y físicos
- Identificar y anotar el conjunto de genes codificados en el genoma
- Caracterizar la diversidad nucleotídica del genoma
- Reunir la información acumulada de expresión de genes (microarreglos)
- Proveer recursos y herramientas para la comparación de los genomas

## Plataformas web

# Proyectos genómicos en progreso y/o parciales

**The Arabidopsis Information Resource**

The Arabidopsis Information Resource (TAIR) maintains a database of genetic and molecular biology data for the model higher plant *Arabidopsis thaliana*. Data available from TAIR includes the complete genome sequence along with gene structure, gene product information, metabolism, gene expression, DNA and seed stocks, genome maps, genetic and physical markers, publications, and information about the Arabidopsis research community. Gene product function data is updated every two weeks from the latest published research literature and community data submissions. Gene structures are updated 1-2 times per year using computational and manual methods as well as community submissions of new and updated genes. TAIR also provides extensive linkouts from our data pages to other Arabidopsis resources.

**GDR | Genome Database for Rosaceae**

**Prunus Data**

Maps

- Maps available in CMap
  - Prunus: | Prunus bin map 2005 | TxE almond x peach F2 2004 | Myrobalan Plum x Almond-Peach hybrid 2004 | GxN almond x peach F2 2001 | Prunus Transcriptome Map | Prunus Resistance Map |
  - Almond: | Almond FxT 2000 | Almond FxB F1 1998 |
  - Apricot: | Apricot GxV F1 2002 | Apricot LxL F2 2003 |
  - Cherry (Sour Cherry): | PcerasusEB | PcerasusRS |
  - Peach: | Peach Axl 2005 | Peach DMF2 2005 | Peach peach x P. feigensteinii D01 PxF 2005 (Updated) | Peach JxT 2004 | Peach Du x D 1990 | GuardianxNemaguard 2007 | Lovell x Nemared 2006 | NJ Pillar x KV 77119 1998 |

**phytozome**

**Phytozome: a tool for green plant comparative genomics**

Explore a genome:  
 [Select an organism: v]  
 - or -  
 Find a gene family:  
 [Select node: v]

Species listed in the tree: *Populus trichocarpa*, *Ricinus communis*, *Medicago truncatula*, *Manihot esculenta*, *Glycine max*, *Cucumis sativus*, *Vitis vinifera*, *Arabidopsis lyrata*, *Zea mays*, *Arabidopsis thaliana*, *Sorghum bicolor*, *Carica papaya*, *Brachypodium distachyon*, *Mimulus guttatus*, *Oryza sativa*, *Sesajinella moellendorffii*, *Physcomitrella patens*, *Chlamydomonas reinhardtii*

**GRAMENE Home**

**Quick Start**

- GENOMES:** Browse assembled genomes for *Oryza sativa* indica and japonica, *O. glaberrima*, *O. rufipogon*, *O. barthii*, *O. brachyantha*, *O. nivara*, *Zea mays*, *Sorghum bicolor*, *Arabidopsis thaliana*, *Arabidopsis lyrata*, *Vitis vinifera*, *Populus trichocarpa*, and *Brachypodium distachyon*. Look for rice/maize syntenies. Narrow your search with GrameneMart. Search for sequence alignment with BLAST. Search by Gene Ontology.
- PROTEINS:** Search by PFam or ProSite or Browse by Gene Ontology using GO Slim.
- COMPARATIVE MAPS:** Browse genetic or physical maps for Wild Rice (*Oryza* sp. from OMAP), Rice, Maize, Wheat, Barley, Oats, Sorghum, and other grasses, or use the Comparative Map Viewer (CMap) to compare maps of different types and species. View map detail information.
- MARKERS:** Search for Genetic markers (RFLPs, SSRs, etc.), DNA Probes (Primers, Overgos, etc.), Genomic Regions (Clones, FPContigs, etc.), and Sequences (GSSs, ESTs, etc.). Search by species such as Sorghum, by type such as RFLP, or by species and type such as Rice SSR. Use the Simple Sequence Repeat Identification Tool (SSRIT).
- TRAITS:** Search the Genes or QTL database for important phenotype-related loci such as Rice Genes, Rice QTL, Maize QTL. Don't forget to explore traits in Ontologies.
- GENETIC DIVERSITY:** Search for SNP and SSR allelic variation on loci of rice, maize, arabidopsis, wheat and sorghum germplasms.
- BIOCHEMICAL PATHWAYS:** Search for a gene, protein, or pathway. Search for rice pathways associated with known reactions (e.g., starch biosynthesis) or get an overview of the metabolic network.
- LITERATURE:** Search the literature for your friends and topics of interest.
- SPECIES PAGES** provide overview information, pictures, and links to Gramene data for *Oryza*, *Zea*, *Triticum*, *Hordeum*, *Avena*, *Setaria*, *Pennisetum*, *Secale*, *Sorghum*, *Zizania*, and *Brachypodium*.
- WEB SERVICES:** Gramene offers many useful web services like distributed annotation services (DAS) for our genomes and genes, the GDPC for our diversity data, and semantic web services for our QTLs.





# Identificar y anotar el conjunto de genes codificados en el genoma



## Arabidopsis Gene Family Information [996 gene families ] and [8,331 genes].

We encourage users of TAIR to share their gene family data with the research community. To submit data for a gene family, please format your data as described in the Gene Family Data Submission page.

Please click on the links below to view the gene family of your choice. Alternatively, the gene families and genes are displayed in one tab delimited file in the FTP Downloads section.

Gene Family Name	Family Count/ Gene Count	Submitter
14-3-3 family	1 family 13 members	R Ferl Laboratory
ABC Superfamily	Subfamilies 8 Members 136	Paul Verrier Freddie Theodoulou Angus Murphy
ABI3VP1 Transcription Factor Family	1 families 11 members	AGRIS
AGC Kinase Gene Family	1 family 39 members	Laszlo Bogre, Laszlo Okresz
Aldehyde Dehydrogenase Gene Family	9 families 14 members	Dorothea Bartels Hans-Hubert Kirch
Amino Acid/Auxin Permease AAAP Family	1 family 43 members	John Ward
Acyl Lipid Metabolism Gene Families	174 families 610 members	Fred Beisson
Alfin-like Transcription Factor Family	1 family 7 members	AGRIS

## Locus: AT2G04880

**Date last modified:** 2007-04-17  
**TAIR Accession:** Locus:2045049  
**Representative Gene Model:** AT2G04880.1  
**Gene Model Type:** protein\_coding  
**Other names:** ATWRKY1, F1O13.1, F1O13\_1, WRKY1, ZAP1, ZINC-DEPENDENT ACTIVATOR PROTEIN-1  
**Description:** Encodes WRKY1, a member of the WRKY transcription factors in plants involved in disease resistance, abiotic stress, senescence as well as in some developmental processes. WRKY1 is involved in the salicylic acid signaling pathway. The crystal structure of the WRKY1 C-terminal domain revealed a zinc-binding site and identified the DNA-binding residues of WRKY1.  
**Other Gene Models:** AT2G04880.2 (splice variant)

**Map Detail Image:**

**Annotations:**

Category	Relationship Type	Keyword
GO Biological Process	involved in	regulation of transcription, DNA-dependent, salicylic acid mediated signaling pathway, positive regulation of transcription
GO Cellular Component	located in	nucleus
GO Molecular Function	functions in	zinc ion binding

**GRAMEN! Arabidopsis Genomes - *Oryza sativa***  
 Location: 43,397,017-43,400,616 Gene: S13  
 BLAST | BLAST | Deck & FADs | Feedback

**Gene: S13 (LOC\_Os04g54800)**  
 shikimate xylase, putative, expressed  
 Location: Chromosome 9, 38,297,017-38,400,616 reverse strand

**Transcripts:** There are 3 transcripts in this gene. [Hide transcripts](#)

Name	Transcript ID	Protein ID	Description
LOC_Os04g54800.1	LOC_Os04g54800.1	LOC_Os04g54800.1	protein_coding
LOC_Os04g54800.2	LOC_Os04g54800.2	LOC_Os04g54800.2	protein_coding
LOC_Os04g54800.3	LOC_Os04g54800.3	LOC_Os04g54800.3	protein_coding

**Transcript and Gene level displays**  
 In Ensembl a gene is made up of one or more transcripts. We provide displays at two levels:  
 • Transcript views which provide information specific to an individual transcript such as the cDNA and CDS sequences and protein domain annotation.  
 • Gene views which provide displays for data associated at the gene level such as orthologues and paralogues, regulatory regions and splice variants.  
 This view is a gene level view. To access the transcript level displays select a Transcript ID in the table above and then navigate to the information you want using the menu at the left hand side of the page. To return to viewing gene level information click on the Gene tab in the menu bar at the top of the page.

**Gene summary** [help](#) [Splice variants](#)

**Name:** S13 (Gramene Gene)  
**Gene type:** Known protein coding  
**Prediction Method:** Gene annotation by S13U through a process of automatic and manual curation

**Transcripts:**

**Configuring the display**  
 Tip: use the "Configure this page" link on the left to show additional data in this region.

# Caracterizar la diversidad nucleotídica del genoma

## 1001 Genomes A Catalog of *Arabidopsis thaliana* Genetic Variation



[Home](#) [Collaborators](#) [Accessions](#) [Tools](#) [Downloads](#) [Data Center](#) [Gallery](#) [About](#) [Help desk](#)

### The 1001 Genomes Vision

The 1001 Genomes Project has a simple goal: to discover the whole-genome sequence variation in 1001 strains (accessions) of the reference plant *Arabidopsis thaliana*. The resulting information will pave the way for a new era of genetics that combines large-scale association studies in wild strains with forward genetic analyses in experimental crosses, in order to identify alleles underpinning phenotypic diversity across the entire genome and the entire species. The analyses enabled by this project will have broad implications for areas as diverse as evolutionary sciences, plant breeding and human genetics.

### Collaborators

Name	Abbreviation	Contact	Collaborators	Funding
<b>DOE Joint Genome Institute</b>	JGI	<b>Len Pennacchio</b>	Joshua Heazlewood	DOE
<b>Max Planck Institute for Developmental Biology</b>	MPI	<b>Detlef Weigel</b>	Richard Clark (Utah) Karl Schmid (Uppsala)	Max Planck Society, DFG
<b>Salk Institute</b>	Salk	<b>Joe Ecker</b>	Michael Egholm Tim Harkins (454) Dan Rokhsar (JGI) Olivier Loudet	Applied Biosystems, Chapman Foundation, Roche
<b>Sainsbury Laboratory</b>	SL	<b>Jonathan Jones</b>	Richard Mott (WTCHG)	
<b>University of Lausanne</b>	UNIL	<b>Christian Hardtke</b>		
<b>University of Southern California</b>	USC	<b>Magnus Nordborg</b>		USC
<b>Wellcome Trust Center for Human Genetics</b>	WTCHG	<b>Richard Mott</b>	Paula Kover	BBSRC
<b>Waksman Institute of Microbiology, Rutgers</b>	WIM	<b>Todd Michael</b>		Charles and Johanna Busch Memorial Fund
<b>Bielefeld University, CeBiTec</b>	CeBiTec	<b>Bernd Weisshaar</b>	Thomas Altmann, IPK Gatersleben	UniBi & IPK
<b>University of Warwick</b>	Warwick	<b>Eric Holub</b>	Robin Allaby (Warwick) Jim Beynon (Warwick) Murray Grant (Exeter)	

# Reunir la información acumulada de expresión de genes

NCBI GEO DATASETS Gene Expression Omnibus

All Databases PubMed Nucleotide Protein Genome Structure OMIM PMC Journals Books

Search GEO DataSets for orzya [Go] [Clear] [Save Search]

Limits Preview/Index History Clipboard Details

Display Summary Show 20 Sort By Send to

All: 303 DataSets: 11 Platforms: 135 Series: 157

Items 1 - 20 of 135 Page 1 of 7 Next

1: GPL2025 record: [Rice] Affymetrix Rice Genome Array [ *Oryza sativa* ] Links

Summary: (Submitter supplied) Affymetrix submissions are typically submitted to GEO using the GEOarchive method described at [http://www.ncbi.nlm.nih.gov/projects/geo/info/geo\\_affy.html](http://www.ncbi.nlm.nih.gov/projects/geo/info/geo_affy.html) This array contains probes to query 51,279 transcripts representing two rice cultivars, with approximately 48,564 japonica transcripts and 1,260 transcripts representing the indica cultivar. This array contains probes to query 51,279 transcripts representing two rice cultivars, with approximately 48,564 japonica transcripts and 1,260 transcripts representing the indica cultivar. This unique design was created within the Affymetrix GeneChip Consortia Program and provides scientists with a single array that can be used for the study of rice. High-quality sequence data were derived from GenBank mRNAs, TIGR gene predictions, and the International Rice Genome sequencing project. [more...](#)

[6 related DataSets](#)  
[41 related Series](#)

EMBL-EBI ArrayExpress Search

All Databases Enter Text Here [Go] [Reset] [Advanced Search] [Give us feedback]

Databases Tools EBI Groups Training Industry About Us Help Site Index

Experiment, citation, sample and factor annotations [clear] Filter on [reset] Display options [reset]

Oryza Any species 25 experiments per page

Match whole words  Loaded in Gene Expression Atlas Any array  Detailed view

Submitter/reviewer login ArrayExpress Browser Help [Query]

ID	Title	Assays	Species	Date	Processed	Raw	Atlas
E-MEXP-2506	Transcription profiling of rice plants grown in different light and temperature cycles	78	<i>Oryza sativa</i>	2010-01-07			-
E-GEOD-19239	Transcription profiling of rice transgeneline carrying the maize resistance gene Rxo1 to Xanthomonas <i>oryzae</i> pv. o	12	<i>Oryza sativa</i>	2009-12-18		-	-
E-GEOD-18248	RNA-seq of rice degradome	2	<i>Oryza sativa</i>	2009-11-27		-	-
E-GEOD-18250	RNA-seq of Arabidopsis small RNA populations in rice total extract and purified AGO1 complexes	4	<i>Oryza sativa</i>	2009-11-27		-	-
E-GEOD-18251	RNA-Seq of rice small RNAs in rice AGO1 complexes and their targets	6	<i>Oryza sativa</i>	2009-11-27		-	-
E-GEOD-6187	Transcription profiling of rice varieties Nipponbare and 6-4 grown in either phosphorus deficient or fertilized soil	6	<i>Oryza sativa</i>	2009-11-09		-	-
E-GEOD-18361	Transcription profiling of rice cv. Nipponbare roots infected with Magnaporthe <i>oryzae</i> strain Guy11	12	<i>Oryza sativa</i>	2009-10-09			-
E-GEOD-11974	RNA-seq of rice seed	1	<i>Oryza sativa</i>	2009-09-16		-	-
E-GEOD-12317	RNA-seq of <i>Oryza sativa</i> small rnas from 4-week old seedlings	3	<i>Oryza sativa</i>	2009-09-11		-	-
E-GEOD-13152	RNA-seq of rice unique small RNA populations from grain	4	<i>Oryza sativa</i>	2009-09-11		-	-
E-GEOD-14606	Transcription profiling of rice in relation to infection with rice tungro spherical virus (RTSV)	4	<i>Oryza sativa</i> Indica Group	2009-09-11		-	-
E-GEOD-16350	RNA-seq of rice endogenous small RNAs of meristematic and a terminally differentiated tissue	12	<i>Oryza sativa</i>	2009-09-08		-	-
E-GEOD-14462	RNA-seq of rice reveals an expression alteration of small rna profiling in autotriploids derived from rice twin-seedli	2	<i>Oryza sativa</i>	2009-09-06		-	-
E-GEOD-16248	RNA-seq of <i>Oryza Sativa</i> small RNAs - run 1 and 2	2	<i>Oryza sativa</i>	2009-09-06		-	-
E-GEOD-11014	MicroRNA profiling of rice - reveals a diverse set of microRNAs and microRNA-like small RNAs in developing rice gr	6	<i>Oryza sativa</i>	2009-09-01		-	-
E-GEOD-11175	Transcription profiling of rice wild-type and dst mutant plants	6	<i>Oryza sativa</i>	2009-08-08			-
E-GEOD-14403	Transcription profiling of rice genotypes in response to calyculin stress	23	<i>Oryza sativa</i> Indica Group	2009-08-06			-

66 experiments, 745 assays. Displaying experiments 1 to 25. Pages: 1 2 3

-Proveer recursos y herramientas para la comparación de los genomas

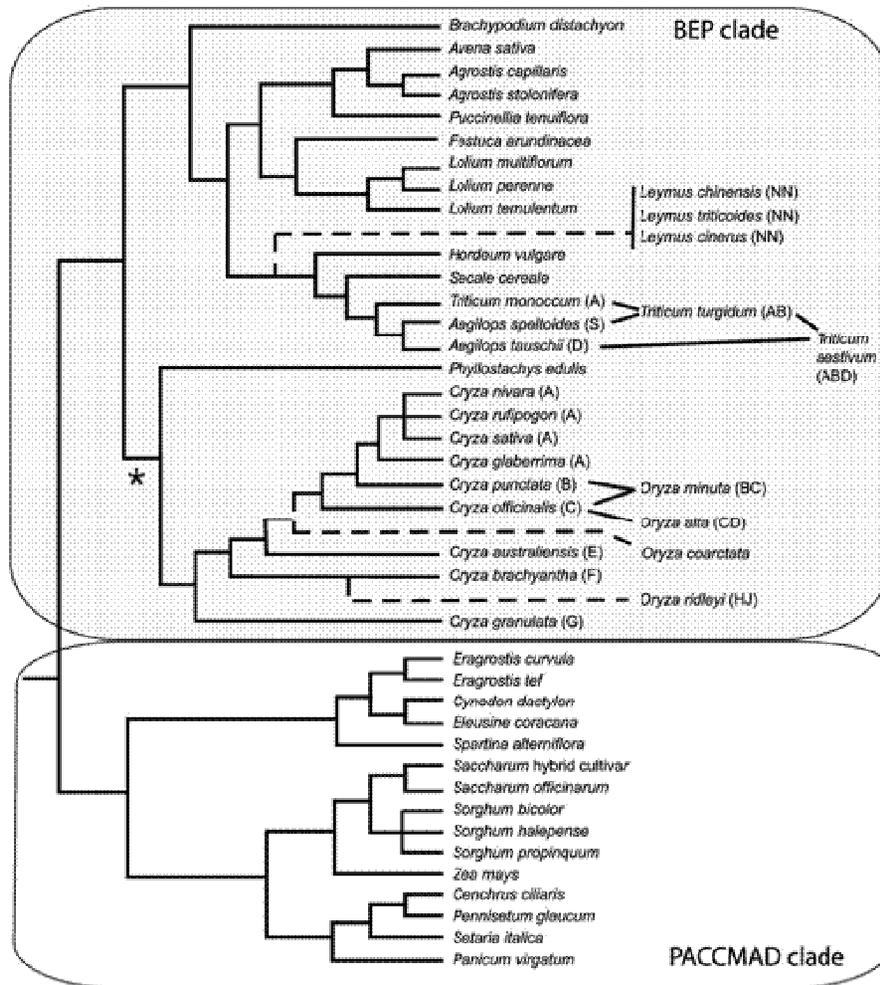
**PLANT  
PHYSIOLOGY**

**FOCUS ISSUE ON THE GRASSES**

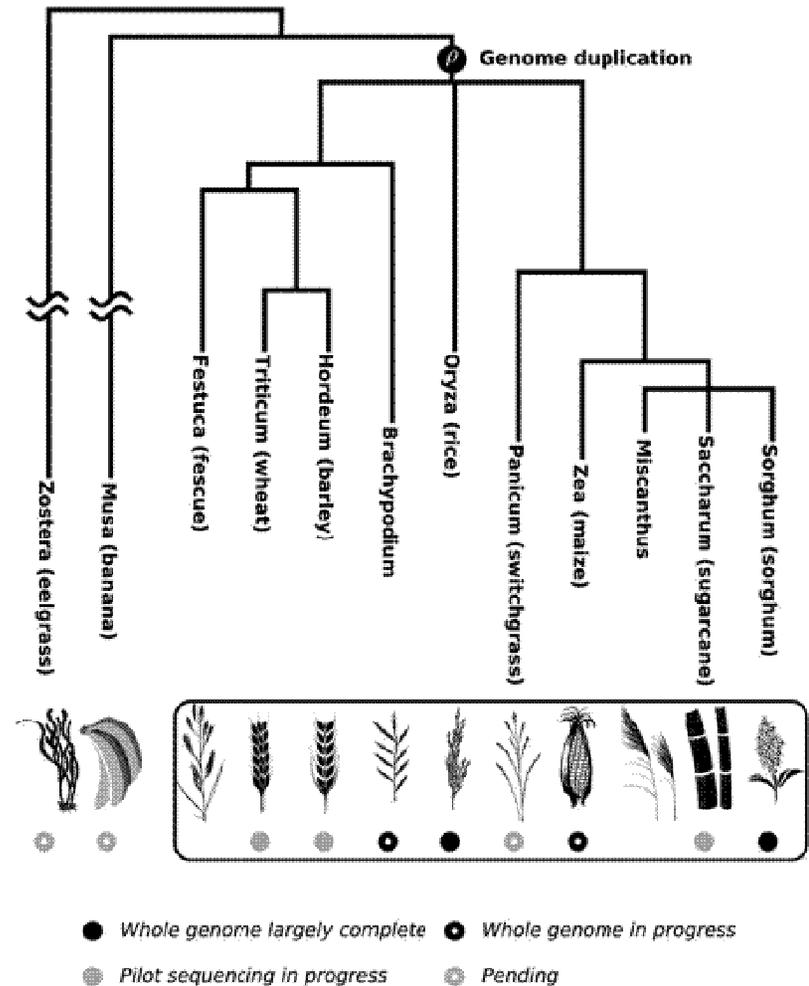
January 2009; Volume 149, Issue 1



**Phylogeny of grasses for which genome sequence data has been or will be generated in the near future**



**The phylogenetic relationship between selected cereals and two outgroup species is illustrated as described elsewhere**



**Buell, C. R. Plant Physiol. 2009;149:111-116**

Copyright ©2009 American Society of Plant Biologists

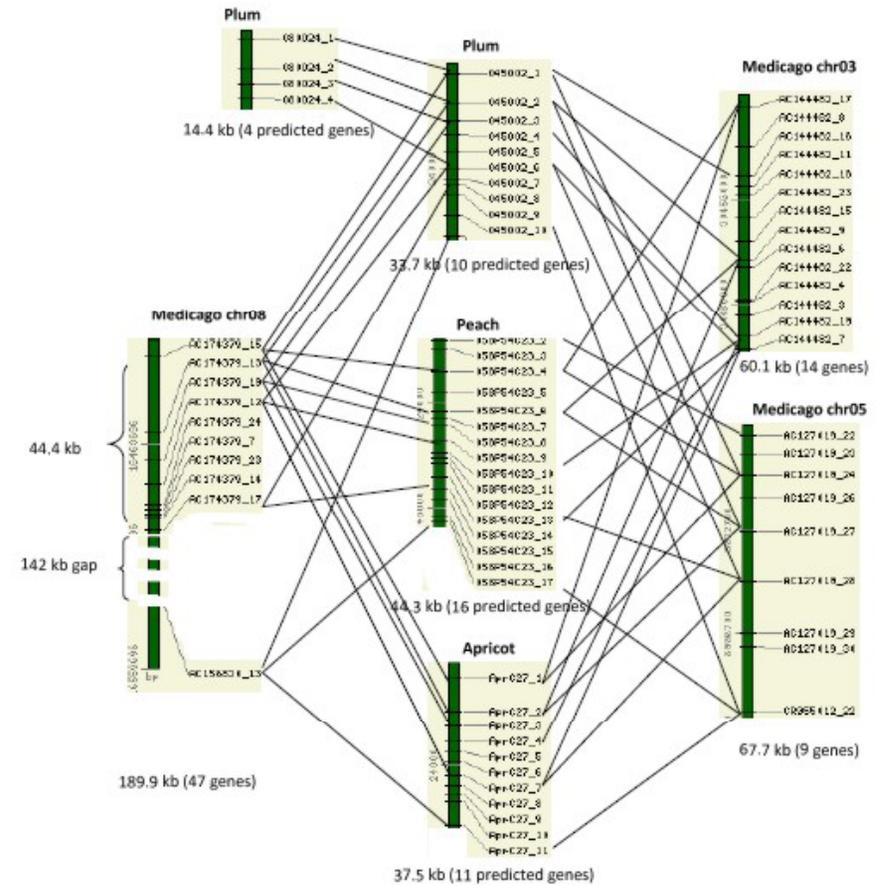
**Paterson, A. H., et al. Plant Physiol. 2009;149:125-131**

# -Proveer recursos y herramientas para la comparación de los genomas

**Table 1. Selected food and model legumes**

Crop legume species are grouped in the three classical legume subfamilies: Caesalpinioideae, Mimosoideae, Papilionoideae; and then by clade and tribe. Sequenced legume genomes are underlined. Model or major crop legumes are in bold text. Primary uses: s = seed; t = tuber or root; p = pod or pod wall; l = leaf; f = forage; m = model. D = Drought tolerant, C = cold tolerant, or F = flooding tolerant; P = perennial. \*, Varieties may contain toxins (alkaloids or cyanogenic glycosides) removable in preparation.

Clade	Tribe	Binomial	Common Name	Uses	Note
Cercidae	Cercideae	<i>Tylosema esculentum</i>	Maramba bean	s, t	D, P
Detarieae	Detarieae	<i>Detarium senegalense</i>	Sweet detar	s	P
Detarieae	Detarieae	<i>Tamarindus indica</i>	Tamarind	p	P
Urtizia	Caesalpinieae	<i>Ceratonia siliqua</i>	Carob	s, p	P
Caesalpinieae	Caesalpinieae	<b><i>Chamaecrista fasciculata</i></b>	Partridge pea	m	D, P
Caesalpinieae	Caesalpinieae	<i>Cordeauxia edulis</i>	Yeheb nut	s	D, P
Mimosoid	Mimoseae	<i>Parkia speciosa</i>	Petai	s, p, f	D, P
Mimosoid	Mimoseae	<i>Prosopis glandulosa</i>	Honey mesquite	s, p, f	D, P
Mimosoid	Mimoseae	<i>Desmanthus illinoensis</i>	Illinois bundleflower	s, f	D, P
Mimosoid	Mimoseae	<i>Inga edulis</i>	Ice-cream bean	p	P
Indigoferoid	Indigoferae	<i>Cyamopsis tetragonoloba</i>	Guar/cluster bean	s, p, f	
Genistoid	Genisteae	<i>Aspalathus linearis</i>	Roobios tea	l	D, P
Genistoid	Genisteae	<b>White lupin</b>	Lupinus	s	*
Genistoid	Genisteae	<i>Lupinus angustifolius</i>	Narrow-leaved lupin	s	*
Genistoid	Genisteae	<i>Lupinus luteus</i>	Yellow lupin	s	*
Genistoid	Genisteae	<i>Lupinus mutabilis</i>	Andean lupin; tarwi	s	C, *
Genistoid	Genisteae	<i>Lupinus polyphyllus</i>	Washington lupin	s, f	C, P, *
Dalbergioid	Aeschynomeneae	<i>Arachis hypogaea</i>	<b>Peanut/groundnut</b>	s	
Galegoid	Galegeae	<i>Glycyrrhiza glabra</i>	Licorice	t	P
Galegoid	Hedysareae	<i>Caragana arborescens</i>	Pea shrub	s, p	D, C, P
Galegoid	Ciceraceae	<i>Cicer arietinum</i>	<b>Chickpea</b>	s	
Galegoid	Trifolieae	<i>Trigonella foenum-graecum</i>	Fenugreek	s	
Galegoid	Trifolieae	<b><i>Medicago truncatula</i></b>	Barrel medic	f, m	
Galegoid	Vicieae/fabeae	<i>Lathyrus sativus</i>	Grass pea/chickling vetch	s, f	D, *
Galegoid	Vicieae/fabeae	<i>Lens culinaris</i>	<b>Lentil</b>	s	
Galegoid	Vicieae/fabeae	<b><i>Pisum sativum</i></b>	<b>Pea</b>	s, p, f, m	
Galegoid	Vicieae/fabeae	<i>Vicia faba</i>	<b>Fava bean/broad bean</b>	s	
Robinoid	Loteae	<i>Lotus tetragonolobus</i>	Asparagus pea	p	
Robinoid	Loteae	<b><i>Lotus japonicus</i></b>	Birdsfoot trefoil	f, m	P
Robinoid	Sesbanieae	<i>Sesbania</i> spp.	Agati	f, l, s, p	F, P
Millettioid	Phaseoleae	<i>Psoralea</i> spp.	Breadroot, prairie turnip	t	D, P
Millettioid	Phaseoleae	<i>Apios americana</i>	Potato bean; groundnut	t	P
Millettioid	Phaseoleae	<i>Cajanus cajan</i>	<b>Pigeonpea</b>	s, p	D, P
Millettioid	Phaseoleae	<i>Canavalia ensiformis</i>	Jack bean/velvet bean	s, p, f	*
Millettioid	Phaseoleae	<i>Lablab purpureus</i>	Hyacinth bean	s, p, f	
Millettioid	Phaseoleae	<b><i>Glycine max</i></b>	<b>Soybean</b>	s, m	
Millettioid	Phaseoleae	<i>Pachyrhizus erosus</i>	<b>Jicama/yam bean</b>	t	
Millettioid	Phaseoleae	<i>Phaseolus coccineus</i>	Scarlet runner bean	s, p	
Millettioid	Phaseoleae	<i>Phaseolus lunatus</i>	Lima bean	s	*
Millettioid	Phaseoleae	<b><i>Phaseolus vulgaris</i></b>	<b>Common bean</b>	s, p	
Millettioid	Phaseoleae	<i>Phaseolus acutifolius</i>	Tepary bean	s, p	D
Millettioid	Phaseoleae	<i>Macrotyloma geocarpum</i>	Hausa groundnut	s	D
Millettioid	Phaseoleae	<i>Psophocarpus</i> spp.	Winged bean	p, t	
Millettioid	Phaseoleae	<i>Vigna angularis</i>	<b>Adzuki bean</b>	s	
Millettioid	Phaseoleae	<i>Vigna aconitifolia</i>	Moth bean	s	
Millettioid	Phaseoleae	<i>Vigna mungo</i> and <i>radiata</i>	<b>Black gram; mung bean</b>	s	
Millettioid	Phaseoleae	<i>Vigna subterranea</i>	Bambara groundnut	s	D
Millettioid	Phaseoleae	<i>Vigna unguiculata</i>	<b>Cowpea/black-eyed pea</b>	s, p	

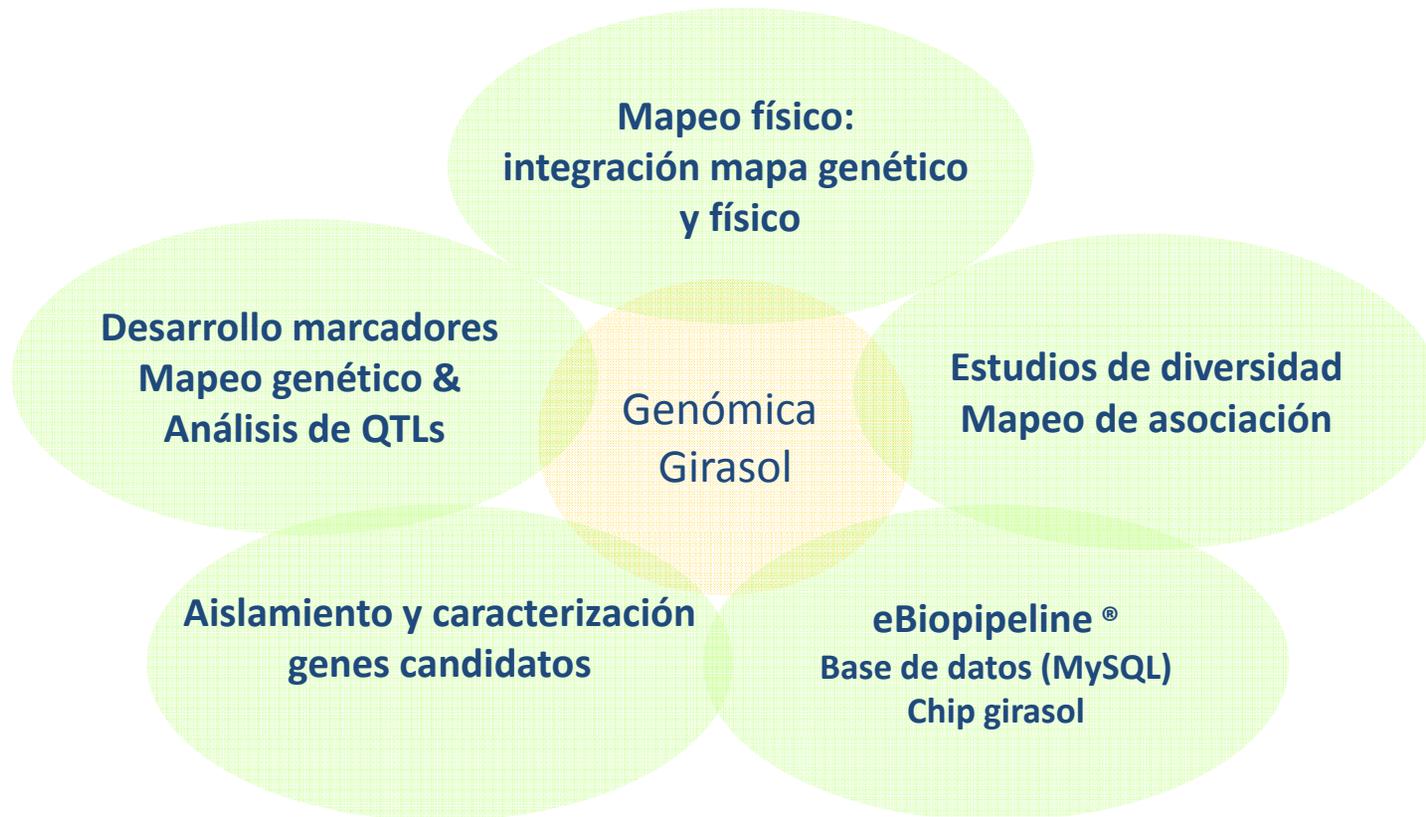


**Figure 2**  
**Conserved synteny between the *Prunus* BACs that contain disease-resistant genes and the genome *Medicago*.**  
 All the intervening genes in the syntenic regions, except those in the 142 kb gap, are also shown. The numbers on the left stand for base pair positions in the *Prunus* BACs or the *Medicago* linkage groups. The lengths of the syntenic regions and the total numbers of predicted genes in the regions are given below the bar.

# Genómica estructural y funcional del girasol cultivado

## *Areas de investigación – integración- finalidad*

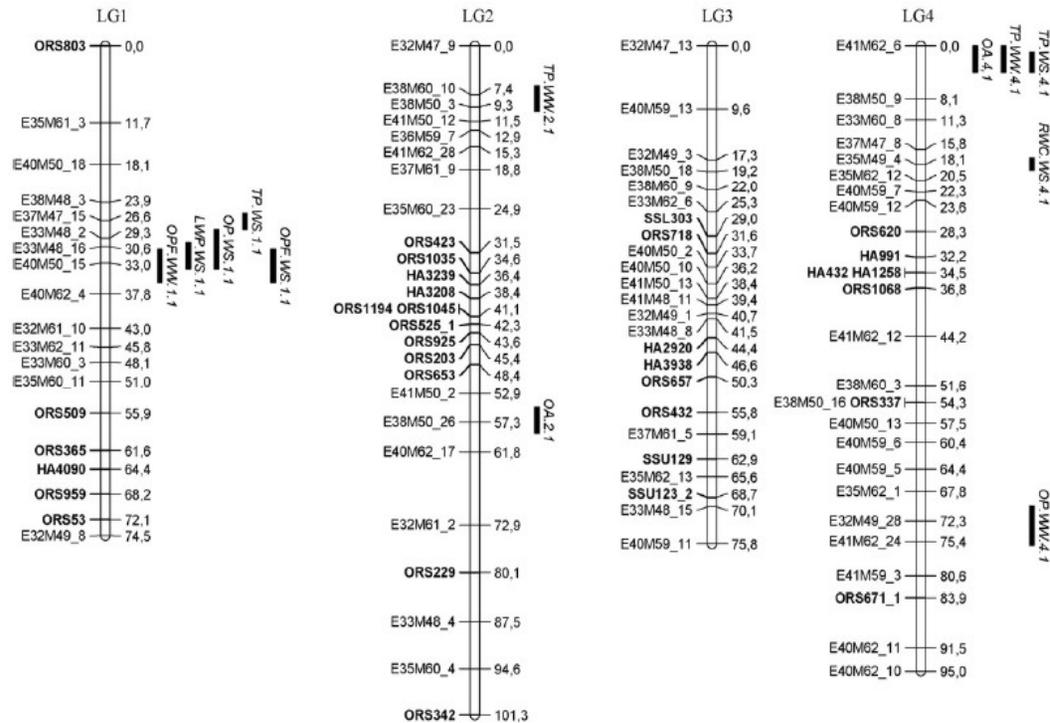
*Norma Paniego - Ruth Heinz - Esteban Hopp*



- SSR, SNP, EST
- Desarrollo poblaciones de mapeo
- Métodos de genotipificación

- Mapeo QTLs
- Perfiles transcripcionales y metabolitos
- Caracterización BACs e hibridación *in situ*

# Genómica girasol



## Microsatellite isolation and characterization in sunflower (*Helianthus annuus* L.)

Norma Paniego, Mercedes Echaide, Marianne Muñoz, Luis Fernández, Susana Torales, Paula Faccio, Irma Fuxan, Mónica Carrera, Rubén Zandomeni, Enrique Y. Suárez, and H. Esteban Hopp

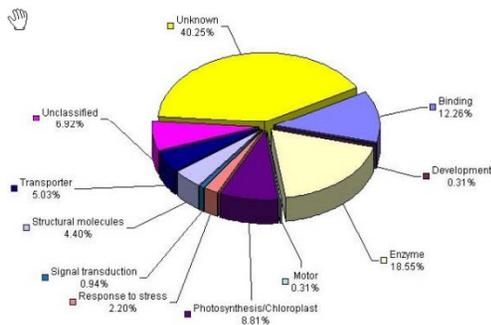
Genetic analysis of plant water status and osmotic adjustment in recombinant inbred lines of sunflower under two water treatments

S. Poormohammad Kiani<sup>a,c</sup>, P. Talia<sup>b</sup>, P. Maury<sup>c</sup>, P. Grieu<sup>c</sup>, R. Heinz<sup>b</sup>, A. Perrault<sup>a</sup>, V. Nishinakamasu<sup>b</sup>, E. Hopp<sup>b</sup>, L. Gentzbittel<sup>a</sup>, N. Paniego<sup>b</sup>, A. Sarrafi<sup>a,\*</sup>

**Plant Science**

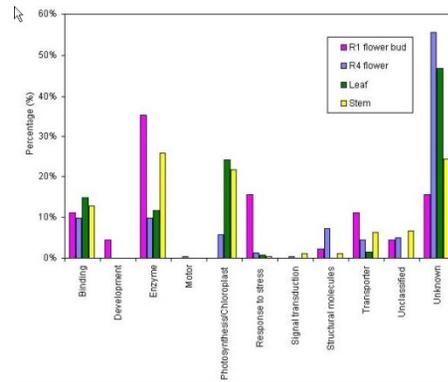
Volume 172, Issue 4, April 2007, Pages 773-787

# Genómica funcional girasol



**Expression analysis of ESTs from organ-specific cDNA libraries.** cDNA clones with significant similarity to protein sequences in SWALL were classified according to Gene Ontology annotation. Sequences with no hits to known protein sequences from BLASTX comparison were classified as unknown. ESTs with significant similarity according to BLASTX comparison but with no GO term definition associated to them were referred as unclassified. Functional analysis includes all non-redundant generated ESTs.

Fernández et al. *BMC Genomics* 2003 4:40 doi:10.1186/1471-2164-4-40



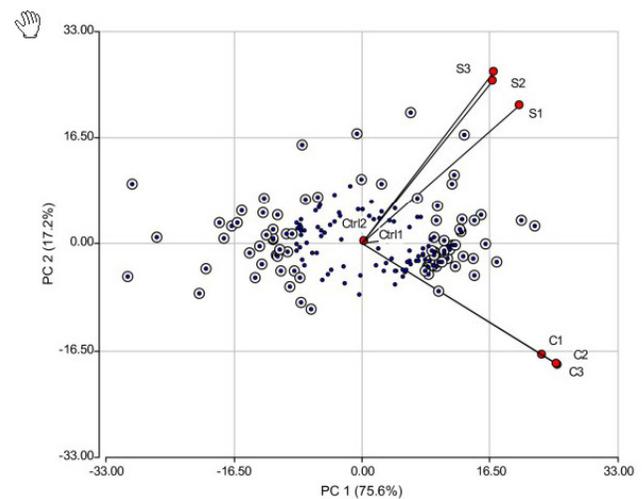
**Comparison of ESTs classified by predicted function among four organ-specific cDNA libraries.** Functional classification of all generated ESTs was done as described in Figure 2. Percentage of ESTs included in each functional class is compared among four differential cDNA libraries.

Fernández et al. *BMC Genomics* 2003 4:40 doi:10.1186/1471-2164-4-40

## Differential representation of sunflower ESTs in enriched organ-specific cDNA libraries in a small scale sequencing project

Paula Fernández<sup>1</sup>, Norma Paniego<sup>1</sup>, Sergio Lew<sup>2</sup>, H Esteban Hopp<sup>1</sup> and Ruth A Heinz<sup>\*1</sup>

*BMC Genomics* 2003, 4:40



**Bi-plot B.** Biplot of the expression matrix showing only those genes having p-values lower than 0.05 in the F-test. Genes with distance-to-the-origin greater than the 70th percentile of the distance-to-the-origin distribution are shown as dotted circles. The circled dots represent the 80 differentially expressed genes identified as differentially expressed among the evaluated treatments: control (Ctrl), cold (C) and salinity (S). Solid dots represent putative false positive genes.

Fernández et al. *BMC Plant Biology* 2008 8:11 doi:10.1186/1471-2229-8-11

## Transcriptomic identification of candidate genes involved in sunflower responses to chilling and salt stresses based on cDNA microarray analysis

Paula Fernández<sup>1</sup>, Julio Di Rienzo<sup>2</sup>, Luis Fernández<sup>1</sup>, H Esteban Hopp<sup>1,3</sup>, Norma Paniego<sup>1</sup> and Ruth A Heinz<sup>\*1,3</sup>

*BMC Plant Biology* 2008, 8:11 doi:10.1186/1471-2229-8-11



## CARACTERIZACIÓN DE GENES CANDIDATOS PARA RESISTENCIA A LA PODREDUMBRE HÚMEDA DEL CAPÍTULO DE GIRASOL

Lla VV, Peluffo L, Hopp E H, Vazquez Rovere C, Panlego N y Heinz, RA. Instituto de Biotecnología, CICVyA-INTA Castelar

La podredumbre húmeda del capítulo de girasol (PHC) es la forma más devastadora de la enfermedad causada por el hongo fitopatógeno necrotrofico *Sclerotinia sclerotiorum*, con una incidencia anual sobre la producción de la pampa húmeda del 10-20%. Debido a la complejidad de la base genética de la resistencia a este patógeno, han cobrado relevancia estrategias centradas en el estudio de los factores determinantes de la patogénesis y de las respuestas de defensa del hospedante. Las germinas (GP) y proteínas de tipo germinas (GLPs) han sido frecuentemente vinculadas con el funcionamiento de la pared celular y con los mecanismos de defensa frente a patógenos. En el caso de la podredumbre de tallo (PHT), la expresión constitutiva de una germina, el gen de oxalato oxidasa de trigo, en plantas de girasol transgénicas fue utilizada con éxito como aproximación para conferir resistencia a la invasión del hongo (1). Por otro lado, se ha descrito que la respuesta a PHT también estaría relacionada con la inducción de la síntesis de proteínas de transferencia de lípidos (LTPs), proteínas inhibidoras de poligalacturonasas (PGIPs) y defensinas (1,2,3,4). A partir del análisis de las colecciones de ESTs desarrolladas en el IB (5) se identificaron secuencias con similitud a genes de otras especies cuya función ha sido asociada con respuestas y/o defensas a estreses bióticos y/o abióticos. Entre estos candidatos se seleccionaron los transcritos correspondientes a un gen de una proteína de tipo germina (HaGLP1) y a un inhibidor de poligalacturonasa (HaPGIP1) para validar experimentalmente su función biológica y evaluar su contribución a la tolerancia a PHC.

### Objetivo: Caracterizar estructural y funcionalmente los genes candidatos HaGLP1 y HaPGIP1

#### Materiales y Metodos

**Extensión de transcritos mediante RACE:** Se extrajo RNA total de botones florales en el estado R3 (Trizol® - Invitrogen, USA) en dos líneas endocriadas representativas de los genotipos susceptibles (HA89) y moderadamente resistentes (RHA 801) a PHC. La extensión de los transcritos se realizó utilizando el 5' y 3' RACE Systems for Rapid Amplification of cDNA ends (Invitrogen, USA) siguiendo las instrucciones del fabricante.

**Análisis de Northern Blot:** Se extrajo RNA total de botones florales en los estados R1, R2, R3 (Trizol® - Invitrogen, USA) de las distintas líneas susceptibles (HA89) y resistentes (HA853, RHA275, RHA801) en estudio y se los hibridó con sondas específicas marcadas radiactivamente.

**Reconstrucción filogenética:** las secuencias de aminoácidos fueron alineadas con el programa Matt (6) y posteriormente convertidas en la secuencia nucleotídica correspondiente utilizando la subrutina transrig. El análisis filogenético se llevó a cabo con el algoritmo de reconstrucción por máxima parsimonia implementado en el software TNT (7).

#### Resultados

El análisis de secuencia de los transcritos obtenidos utilizando la técnica de RACE reveló la ausencia de sustituciones aminoácidas entre líneas resistentes y susceptibles para HaGLP1 (220 aa) y una divergencia del 0,6% para HaPGIP1 (331 aa). Ambos genes candidatos presentaron un intrón al ser amplificadas a partir de ADN genómico y exhibieron una población de transcritos de longitud variable (HaGLP1: 712-886 pb, HaPGIP1: 1116-1157pb), estando la totalidad de las diferencias de tamaño restringidas a la región 3' no codificante (Figura 1).

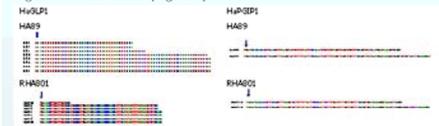


Figura 1. Análisis de secuencia de transcritos correspondientes a la extensión mediante 3' RACE de los genes HaGLP1 y HaPGIP1. Las flechas indican la posición del codón stop.

A partir de búsquedas en bases de datos de secuencia pública se identificaron cuatro entradas de girasol con similitud a HaGLP1 (GenBank AJ540203 (HaGLP2); TIGR Plant Gl: TC17527 (HaGLP3), 18217 (HaGLP4), 17648 (HaGLP5) y cuatro entradas correspondientes a la especie afín *Lactuca saliva* (TIGR Plant Gl: TC9227 (LsGLP1), 14591 (LsGLP3), 8238 (LsGLP4), 12192 (LsGLP5)). A fin de establecer su posición filogenética dentro de la familia de las Cupinas, las secuencias obtenidas se compararon con un conjunto de Germin y GLPs descritas en (8) y (9). El análisis de parsimonia incluyó 35 representantes del grupo provenientes de diversas especies vegetales y dio como resultado 1 único árbol de máxima parsimonia con una longitud de 5089 pasos (Figura 2). Las cinco GLPs de girasol se distribuyen en distintos clados del árbol, indicando que cada una de ellas pertenece a subfamilias diferentes. En concordancia con esto último, la similitud nucleotídica promedio entre las GLPs de girasol es de sólo 51,6% para la región codificante, mientras que al considerar la secuencia de aminoácidos esta cifra desciende a 44%.

**Conclusiones:** La proteína HaGLP1 pertenece a una diversa familia multigénica cuya caracterización funcional es aún incompleta en la mayor parte de las angiospermas. La gran variabilidad exhibida por las GLPs de girasol, así como las diferencias observadas en los perfiles de expresión de líneas resistentes y susceptibles a PHC, ofrecen interesantes perspectivas para explorar las bases genéticas de la resistencia a *Sclerotinia sclerotiorum*. Las investigaciones en curso prevén la obtención de plantas transgénicas para la sobre-expresión y/o silenciamiento de los genes candidatos aquí caracterizados a fin de evaluar su desempeño en ensayos de desafío frente al patógeno.

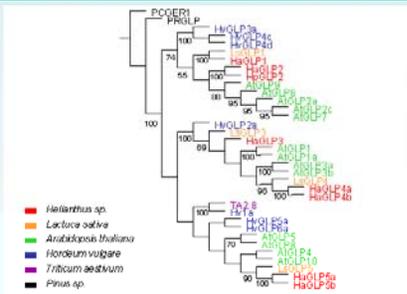


Figura 2. Árbol de máxima parsimonia obtenido en base a secuencias nucleotídicas. Los números sobre las ramas corresponden a las medidas de apoyo según el método de bootstrapping (1000 repeticiones).

El estudio de la abundancia de transcritos mediante la técnica de Northern Blot mostró la existencia de diferencias temporales y finales en los niveles de expresión de materiales resistentes (RHA801, RHA275, HA853) y susceptibles (HA89) (Figura 3).

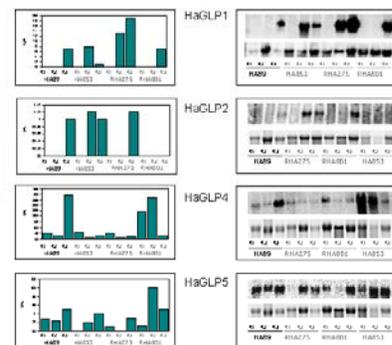
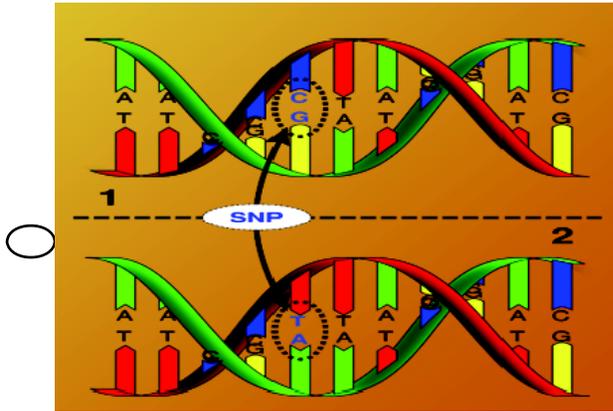


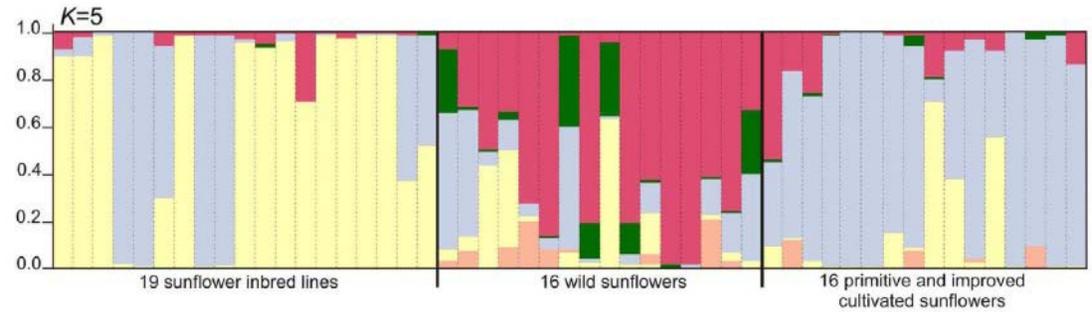
Figura 3. Análisis de expresión basado en abundancia de transcritos en líneas endocriadas de girasol susceptible (HA89) y moderadamente resistentes (RHA275, RHA801, HA853). Hibridación con sondas radioactivas específicas correspondiente a los genes HaGLP1, HaGLP2, HaGLP4 y HaGLP5. U.A.: Unidades arbóreas. R1, R2, R3: estados de desarrollo del botón floral.

# Diversidad nucleotídica

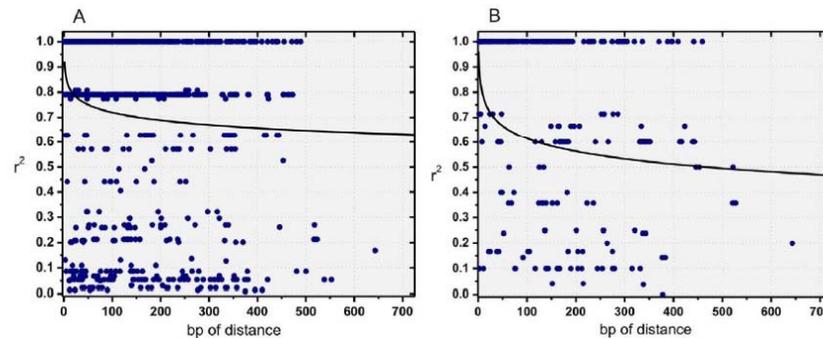


SNPs

InDel



**Figure 1**  
**Population structure in sunflower inbred lines.** Dash lines separate each individual, which is partitioned in K coloured segments that represent the individual's estimated membership fractions in K clusters. Black lines separate individuals from different groups. First group is composed by the 19 sunflower inbred lines (in order from left to right: HA52, HA61, HA89, HA292, HA303, HA369, HA370, HA821, HAR2, HAR3, HAR5, KLM280, PAC2, RHA266, RHA274, RHA293, RHA374, RHA801 and V94); the second and the third group are the individuals studied by Liu and Burke [46]. The inbred-lines group has mostly contributions of two clusters (yellow and light-blue).

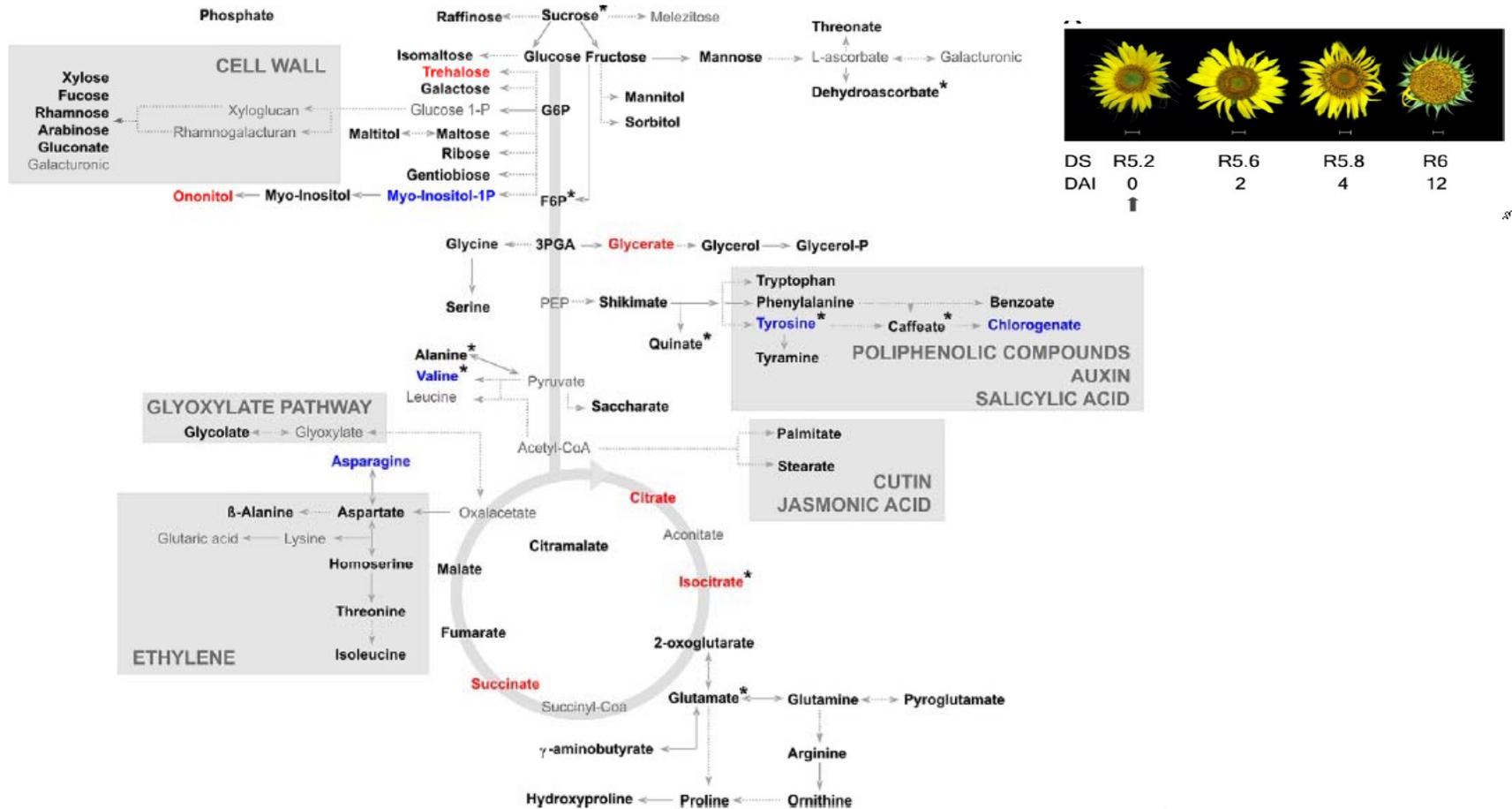


**Figure 2**  
**Linkage disequilibrium.** A: LD plot from 24 genes pooled together for the 19 inbred lines. The logarithmic trend line reaches a value of 0.64 at 643 bp. B: LD plot from the whole gene data calculated for the G1 subset of individuals identified in the STRUCTURE analysis (HA52, HA61, HA89, HA370, HAR3, HAR5, KLM280, PAC2, RHA266, RHA274, RHA293 and RHA374).

## Identification of Single Nucleotide Polymorphisms and analysis of Linkage Disequilibrium in sunflower elite inbred lines using the candidate gene approach

Corina M Fusari<sup>1</sup>, Verónica V Lia<sup>1,2</sup>, H Esteban Hopp<sup>1,2</sup>, Ruth A Heinz<sup>1,2</sup> and Norma B Paniego\*<sup>1</sup>

# Genómica funcional-metabolómica



Metabolic profiles of sunflower genotypes with contrasting response to *Sclerotinia sclerotiorum* infection

Lucila Peluffo <sup>a,b</sup>, Verónica Lia <sup>a,b,c</sup>, Carolina Troglia <sup>d</sup>, Carla Maringolo <sup>d</sup>, Paniego Norma <sup>a,b</sup>, Alberto Escande <sup>d</sup>, H. Esteban Hopp <sup>a,c</sup>, Anna Lytovchenko <sup>e</sup>, Alisdair R. Fernie <sup>e</sup>, Ruth Heinz <sup>a,b,c,\*</sup>, Fernando Carrari <sup>a,b,\*,1</sup>

# Métodos de genotipificación de alto procesamiento para SNPs

**STEPS:**

- PCR amplification of 2 individuals with different sequence for any candidate gene.
- Mix of the products in equimolar concentration of each one.
- Heat and re-anneal to form the Heteroduplex (if there is a SNP or InDel)
- Run in ABI 3130 after incubation with CEL1 or
- Run in dHPLC Agilent 1100 series

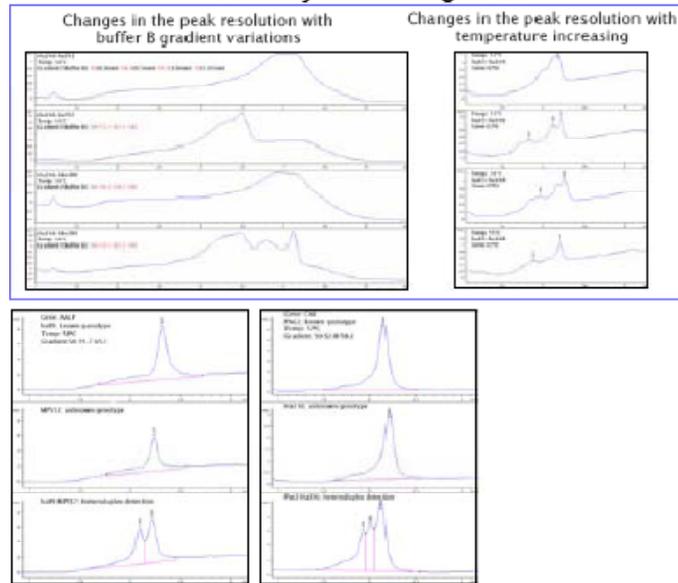
The heteroduplex detection with CEL1 is by capillary electrophoresis in ABI 3130XL: fluorescent primers. The run in the dHPLC is monitored by UV: common primers

- Quantification of the PCR products are needed.
- For dHPLC the fragment is recommended to be purified by QIAquick PCR Purification Kit (QIAGEN)
- Heat and re-anneal are taken on a thermocycler:  
95°C 2min  
95°C to 85°C (-2°C/seg)  
85°C to 25°C (-0.1°C/seg)  
4°C hold

•Incubation with CEL1 endonuclease (partially purified in our lab):  
45°C 20 min  
•Precipitation with EtOH 100% and dilution with Formamide and Mass Markers  
•Analysis with GeneMapper

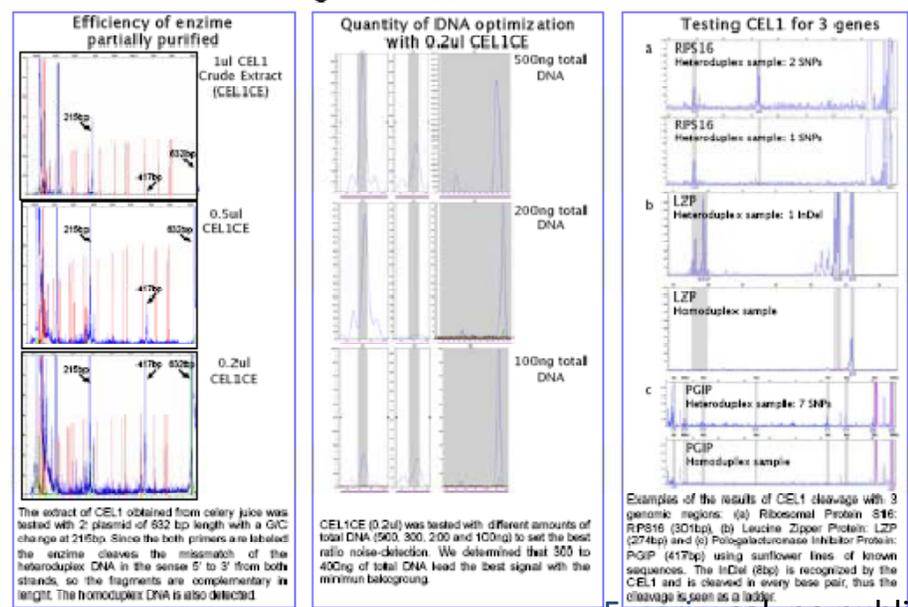
•Gradient of Buffer B (TEAA, EDTA, 25% Acetonitrile) and Temperature (50°C to 60°C) are determined with <http://insertion.stanford.edu/melt.html>  
•Analysis with HPCHEM

## Partially denaturing HPLC



dHPLC was used to genotype 22 sunflower for 2 genomic regions: (a) Arabidopsis Akirin-Like Protein: AALP and Chlorophyll a/b binding Protein: CAB. The results are shown for 2 individual tested against the control genotypes with each gene. In these cases both genotypes have differences in the sequences with the control, thus the mix of DNAs give a 2 or three peak detection in the HPLC.

## Cleavage with CEL1 endonuclease



# SNP Discovery by Massively Parallel Transcriptome Resequencing in Sunflower and Development of a Bioinformatic Pipeline and Database for Mining and Displaying SNPs in Next-Generation Sequence Assemblies

```

TACCATTAAACAGGTAAGCAATACGC TGAGAAGGTGTCGGAAAACACTGTGTAGCGGTATGGAAGCACTCGGGATAT ATACAGACGCAGATGCCAAAAC
383.SNPSTER4_77_300KC:3...
383.SNPSTER4_77_300KC:2...
383.SNPSTER4_77_300KC:1... :ta
383.SNPSTER4_77_300KC:2... :taccataacaggttaagcaata
383.SNPSTER4_77_300KC:3... :taccataacaggttaagcaata
383.SNPSTER4_77_300KC:2... :taccataacaggttaagcaataacgc
383.SNPSTER4_77_300KC:2... :taccataacaggttaagcaataacgc tgagaagg
383.SNPSTER4_77_300KC:1... caataacaggttaagcaataacgc tgagaagggtgtcgg
383.SNPSTER4_77_300KC:2... caggttaagcaataacgc tgagaagggtgtcggaaaact
383.SNPSTER4_77_300KC:3... caggttaagcaataacgc tgagaagggtgtcggaaaact
383.SNPSTER4_77_300KC:3... caataacgc tgagaagggtgtcggaaaactgttagcgc
383.SNPSTER4_77_300KC:2... atacgc tgagaagggtgtcggaaaactgttagcgggt
383.SNPSTER4_77_300KC:3... gagaagggtgtcggaaaactgttagcgggtatggaaa
383.SNPSTER4_77_300KC:3... aagggtgtcggaaaactgttagcgggtatggaaagca
383.SNPSTER4_77_300KC:1... gtgtcggaaaactgttagcgggtatggaaagcactc
383.SNPSTER4_77_300KC:3... cgaaaactgttagcgggtatggaaagcactcggg
383.SNPSTER4_77_300KC:3... ggaaaactgttagcgggtatggaaagcactcgggat
383.SNPSTER4_77_300KC:3... ggaaaactgttagcgggtatggaaagcactcgggat
383.SNPSTER4_77_300KC:1... gttagcgggtatggaaagcactcgggat atacagac
383.SNPSTER4_77_300KC:2... gttagcgggtatggaaagcactcgggat atacagac
383.SNPSTER4_77_300KC:3... gttagcgggtatggaaagcactcgggat atacagac
383.SNPSTER4_77_300KC:3... gcgggtatggaaagcactcgggat atacagacgc
383.SNPSTER4_77_300KC:3... gttagcgggtatggaaagcactcgggat atacagacgcgat
383.SNPSTER4_77_300KC:3... gttagcgggtatggaaagcactcgggat atacagacgcgat
383.SNPSTER4_77_300KC:1... atggaaagcactcgggat atacagacgcgatgc
383.SNPSTER4_77_300KC:3... ggaagcactcgggat atacagacgcgatgcc
383.SNPSTER4_77_300KC:1... gaaagcactcgggat atacagacgcgatgccaa
383.SNPSTER4_77_300KC:3... agcactcgggat atacagacgcgatgccaaaac
383.SNPSTER4_77_300KC:3... actcgggat atacagacgcgatgccaaaac
383.SNPSTER4_77_300KC:2... cttagcgggat atacagacgcgatgccaaaac
383.SNPSTER4_77_300KC:3... atat atacagacgcgatgccaaaac
383.SNPSTER4_77_300KC:1... cagacgcgatgccaaaac

```

```

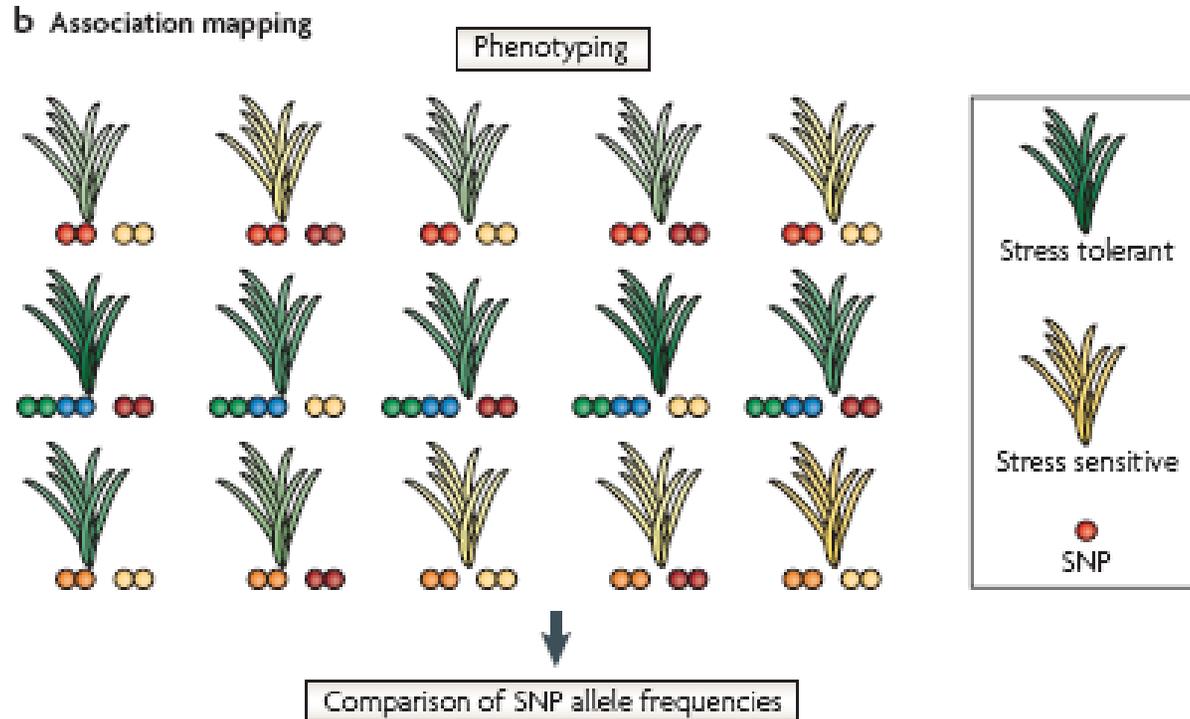
CL13332Contig1 TACCATTAAACAGGTAAGCAATACGC TGAGAAGGTGTCGGAAAACACTGTGTAGCGGTATGGAAGCACTCGGGATAT ATACAGACGCAGATGCCAAAAC
HA89 taccattaacaggttaagcaataacgc tgagaagggtgtcggaaaactgtgttagcgggtatggaaagcactcgggat atacagacgcagatgccaaaac
HA383 taccataacaggttaagcaataacgc tgagaagggtgtcggaaaactgtgttagcgggtatggaaagcactcgggat atacagacgcgatgccaaaac
RHA373 taccattaacaggttaagcaataacgc tgagaagggtgtcggaaaactgtgttagcgggtatggaaagcactcgggat atacagacgcagatgccaaaac
RHA415 taccattaacaggttaagcaataacgc tgagaagggtgtcggaaaactgtgttagcgggtatggaaagcactcgggat atacagacgcagatgccaaaac
HA434 taccataacaggttaagcaataacgc tgagaagggtgtcggaaaactgtgttagcgggtatggaaagcactcgggat atacagacgcgatgccaaaac

```

The top figure shows an alignment of reads from the genotype 'HA383' with three distinct positions of variation from the consensus unigene sequence. Our interface can compile the reads for each genotype into a genotype consensus sequence (bottom figure) and more easily demonstrate properties such as haplotypic agreement among collections of genotypes. 77,820 SNPs among 17,638 unigenes

## Realizar Mapeo por Asociación

Aproximación genómica para encontrar alelos (genes) involucrados en un carácter agronómico



Utiliza la variación tanto fenotípica como genotípica en un conjunto de individuos que pueden o no estar relacionados.

Ej. Poblaciones naturales o colecciones de germoplasma

## Realizar Mapeo por Asociación

Estrés Biótico

Tolerancia a la Podredumbre  
Húmeda del Capítulo (PHC)

134 líneas

16-02/17-03 banco de germoplasma

Caracterización fenotípica

Evaluación a campo **EEA Balcarce**  
Campañas 2008/2009 y 2009/2010

Incidencia de la enfermedad

Severidad

Período de incubación

Caracterización molecular

Polimorfismos SNPs e INDELS  
en Genes candidatos, **IB**.

Análisis de la estructura poblacional  
mediante SSRs, **IB - EEA Manfredi**

	SNPs/indels								Haplotipos	
indiv 1	A	A	T	A	C	G	A	--	T	1
indiv 2	A	A	T	A	C	G	A	--	T	
indiv 3	G	A	T	A	C	A	A	I/D	C	2
indiv 4	G	A	T	A	C	A	A	I/D	C	
indiv 5	G	A	T	A	C	A	A	I/D	T	3
indiv 6	G	A	T	A	C	A	A	I/D	T	
indiv 7	G	C	A	T	T	A	C	--	T	4
indiv 8	G	C	A	T	T	A	C	--	T	

**Figura 4.2. Determinación de Haplotipos a partir de sitios SNPs.**  
Sitios SNP encontrados en una región de 263 nt. Cuatro haplotipos pueden distinguirse a partir de estos sitios (1, 2, 3, 4). Para el análisis de asociación, en lugar de hacer nueve comparaciones, el número se reduce a menos de la mitad al considerar haplotipos (4). Tomado de Rafalski (2002).

## Realizar Mapeo por Asociación

### Tolerancia a la Podredumbre Húmeda del Capítulo (PHC)

#### Caracterización molecular

Polimorfismos SNPs e INDELs en Genes candidatos

#### Genes Candidatos (≈20)

**I.** Transcritos de expresión aumentada en estudios de interacción entre *Sclerotinia* y uno de sus huéspedes (*Brassica napus*) (Zhao y col. 2007)

**II.** Transcritos diferencialmente expresados en capítulos de la línea tolerante RHA801 a los dos días post-inoculación ≈ 17 (Peluffo y col. póster)

Búsqueda de ortólogos en girasol ≈ 38

Amplificación

Amplificación y secuenciación = 25

Evaluación del polimorfismo = 15 (8)  
(secuenciación 10 líneas)

Evaluación del polimorfismo ≈ 12  
(secuenciación 10 líneas)

Realizar Mapeo por Asociación

Tolerancia a la Podredumbre Húmeda del Capítulo (PHC)

Caracterización molecular

Polimorfismos SNPs e INDELS en Genes candidatos

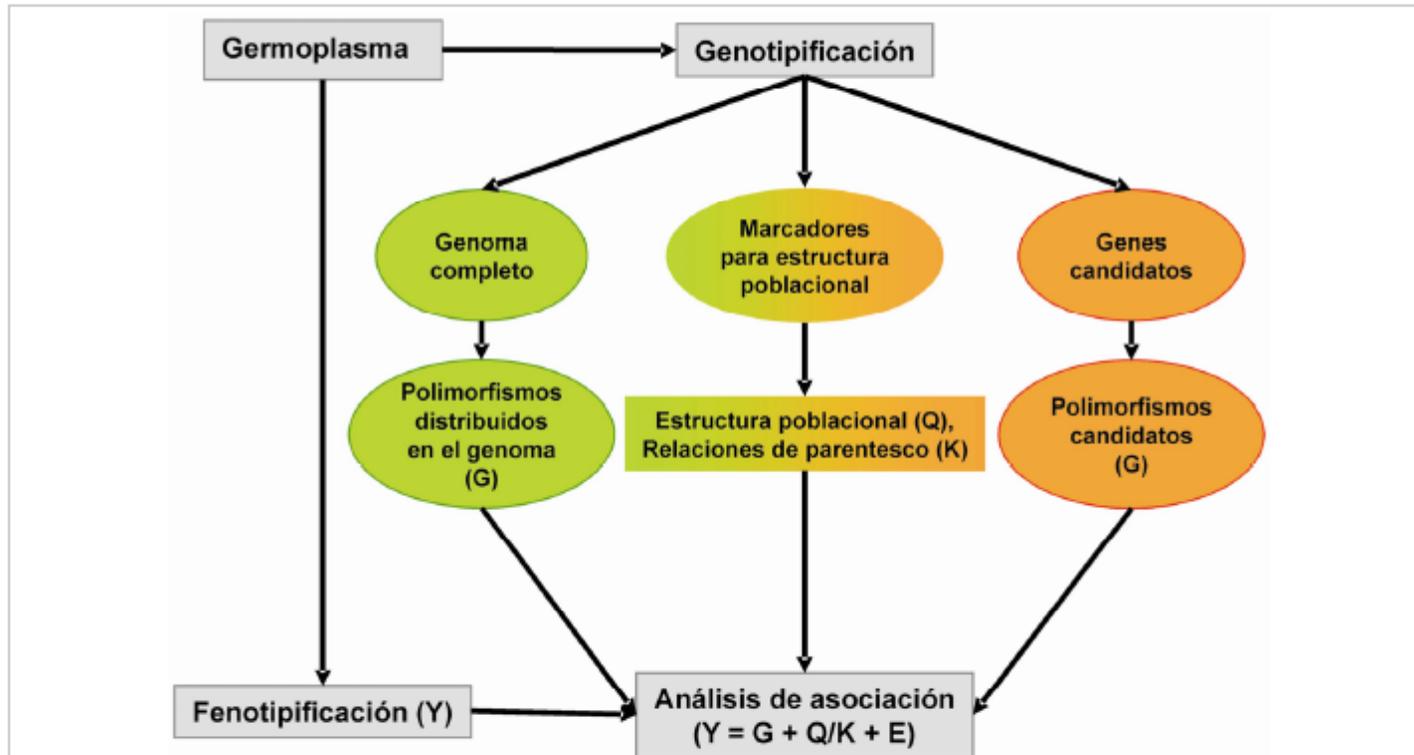
Genotipificación  
(3 individuos por línea)

Secuenciación

métodos de detección de heterodúplex mediante dHPLC  
(Oefner y Underhill. 1998)

Resolución de fragmentos de distinto tamaño mediante electroforesis capilar en sec. automático

## Posibles abordajes para el mapeo por asociación.



**Figura 4.1. Posibles abordajes para el mapeo por asociación.**

Componentes del mapeo por asociación: el germoplasma, la fenotipificación, la genotipificación y el análisis de asociación. Tanto para la estrategia de **genoma completo** (verde) como para la de **genes candidatos** (naranja) se obtiene un conjunto de polimorfismos (G). En ambos casos se realiza la inferencia de estructura poblacional (Q) y de las relaciones de parentesco (K) utilizando un conjunto independiente de marcadores moleculares.

El modelo estadístico del análisis de asociación intenta explicar la variación fenotípica (Y) en términos de la variación alélica (G), teniendo en cuenta la estructura y/o relaciones de parentesco de la población en estudio (Q y/o K). Un factor de error (E) completa el modelo.

## INTEGRANTES DEL GRUPO DE TRABAJO

### IB INTA Castelar

Fusari, CM  
Lia, VV  
Nishinakamasu, V  
Zubrzycki, J  
Puebla, AF  
Heinz, RA  
Paniego, NB  
Hopp, HE

### EEA INTA Balcarce

Trogia, C  
Maringolo, C  
Quiróz, F  
Escande, A

### EEA INTA Manfredi

Moreno, V  
Gieco, J  
Álvarez, D