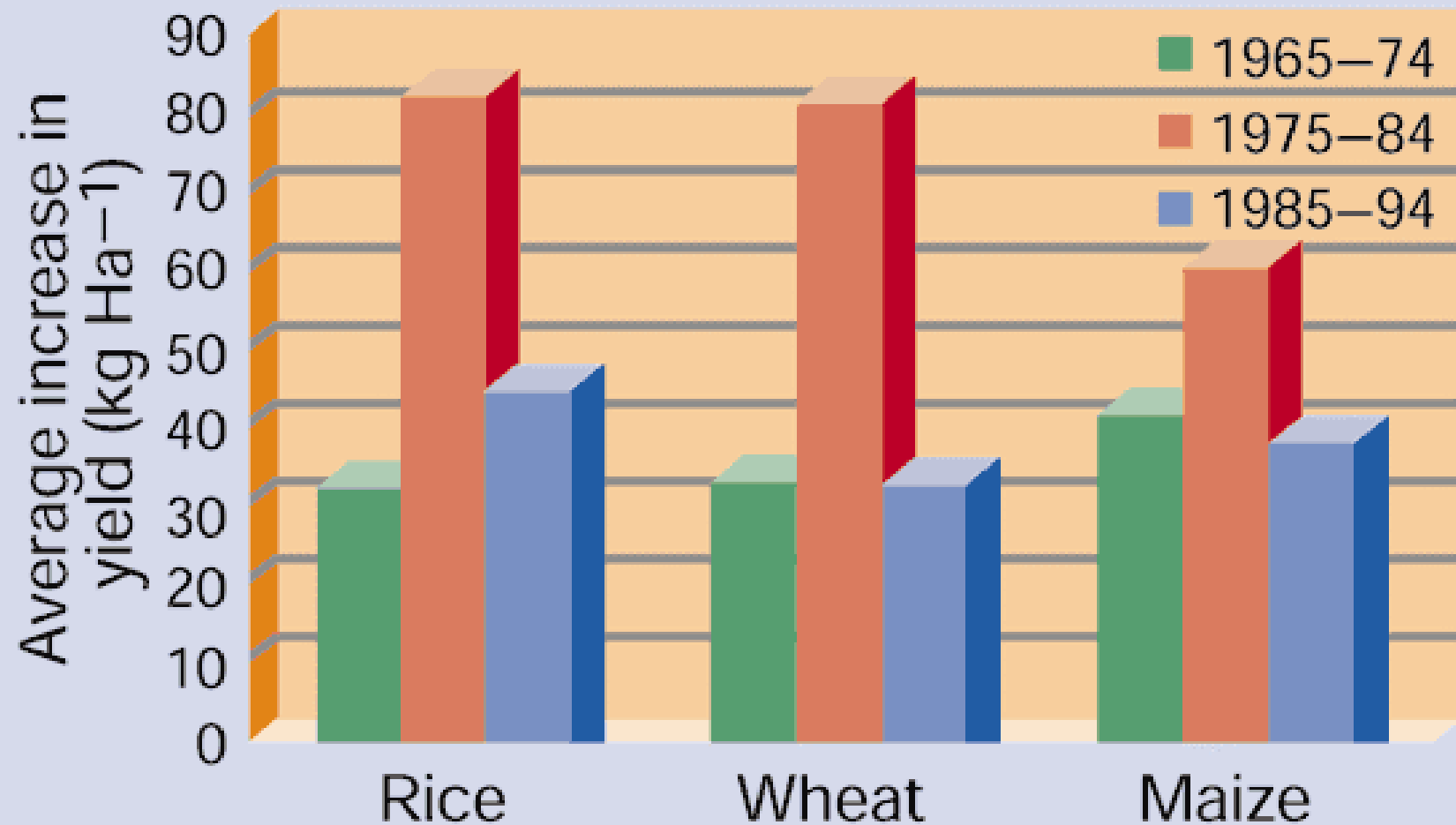


Genómica en Arabidopsis

la **Genómica**, el análisis a gran escala de la estructura y la función de los genes, es el resultado de una renovación metodológica y conceptual que es sólo comparable a la que supuso la **revolución industrial** de finales del siglo XVIII, que introdujo los sistemas de producción mecanizados.

En el área de la Biología Vegetal, la aproximación genómica encuentra su mejor aliado en ***Arabidopsis thaliana***, el **sistema modelo vegetal** por excelencia

El fin de la revolución verde del siglo XX



La duplicación de la producción agrícola mundial en los últimos 35 años, además de la aportación de la mejora genética, ha sido resultado de incrementos en la utilización de:

- la cantidad de fertilizantes (5 x)
- la superficie irrigada (1,7 x)
- la superficie cultivada (1,1 x)

Si se extrapolan las tendencias pasadas, para conseguir duplicar de nuevo la producción agrícola sería necesario incrementar

- 3 x la cantidad de fertilizantes
- 2 x la superficie irrigada
- 1,2 x la superficie cultivada

La próxima revolución verde tendrá que ser doblemente verde

Conway y Toenniesen 6 (1999) *Nature* **402** (supp), c55-c58

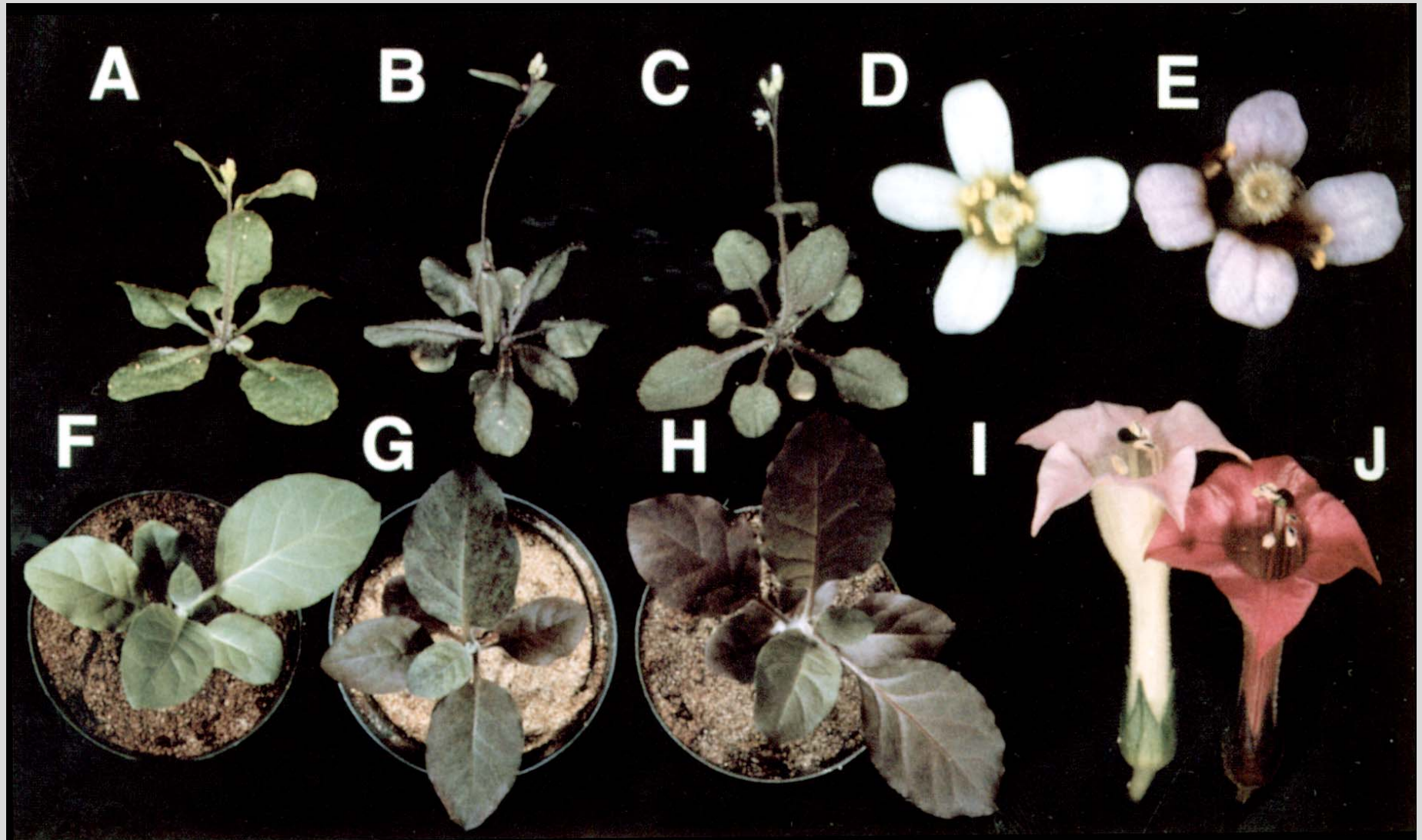
La biotecnología
es una metodología de mejora genética
basada en el conocimiento
de la función génica

Arabidopsis thaliana como sistema modelo



- Planta típica (resultados extrapolables)
- Tamaño y tiempo de generación reducidos
- Asequible al análisis genético
- Genoma pequeño
- Susceptible de manipulación genética

Overexpression of *PAP1* or *PAP2* Enhances Pigmentation in Arabidopsis and Tobacco



Genómica

Disciplina que tiene por objeto la **determinación de la estructura y función de todos los genes de un genoma** mediante técnicas de análisis global. Se basa en los importantes desarrollos que han tenido lugar en el campo de la analítica, robótica e informática aplicados a la biología molecular.

Genómica estructural

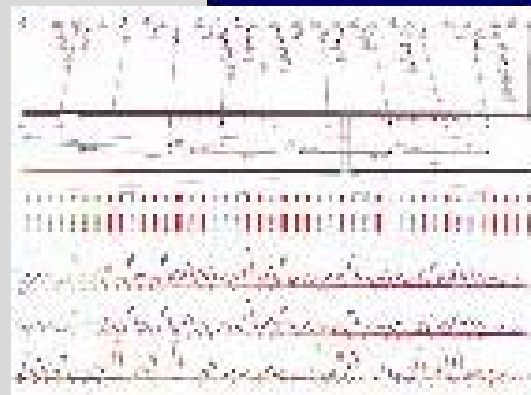
Su objetivo es la determinación de la secuencia de un genoma y la definición de sus genes

Genómica funcional

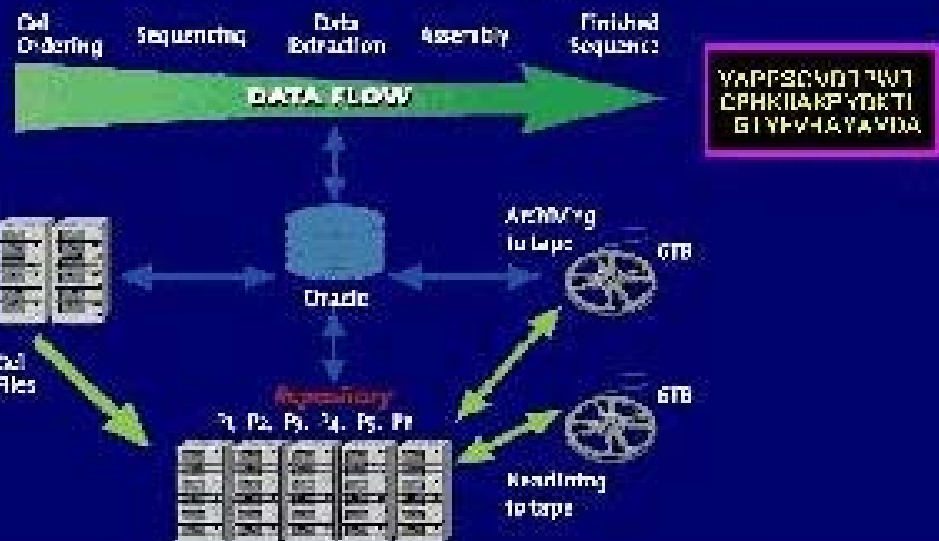
Su objetivo es el estudio de la función de los genes de un genoma

- *Transcriptómica*: análisis global del genoma a nivel de transcrito
- *Proteómica*: idem id a nivel de proteínas
- *Metabolómica*: idem id a nivel de metabolitos

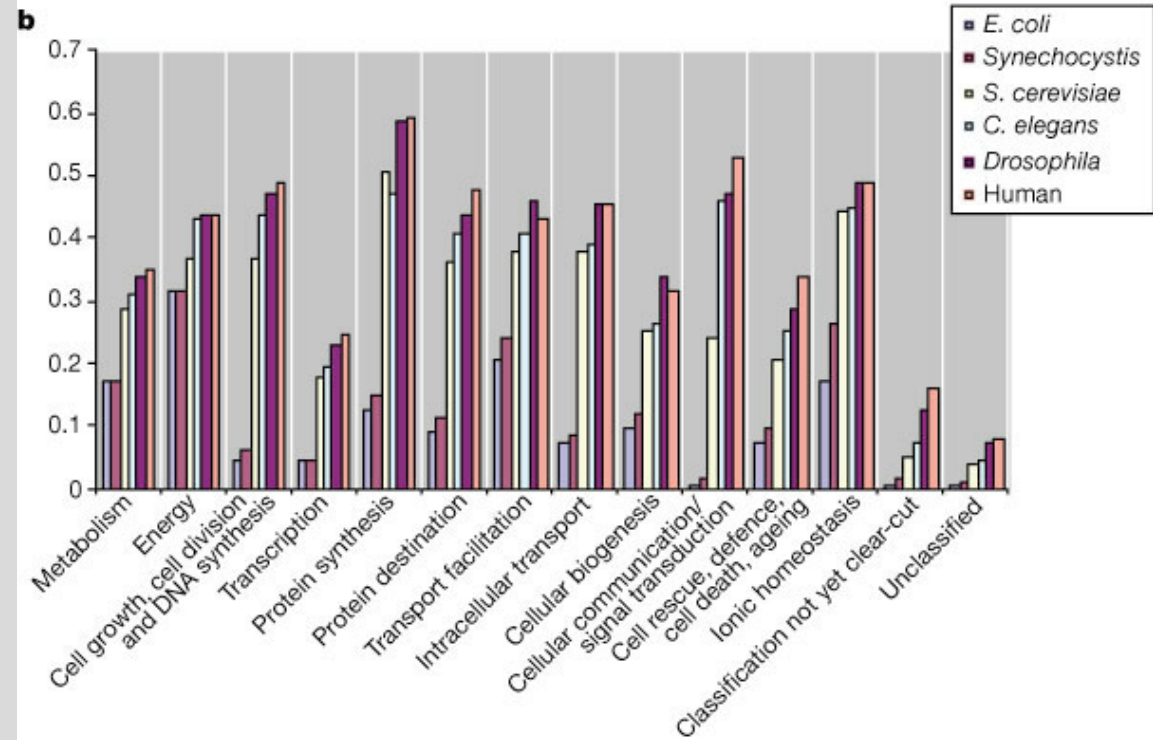
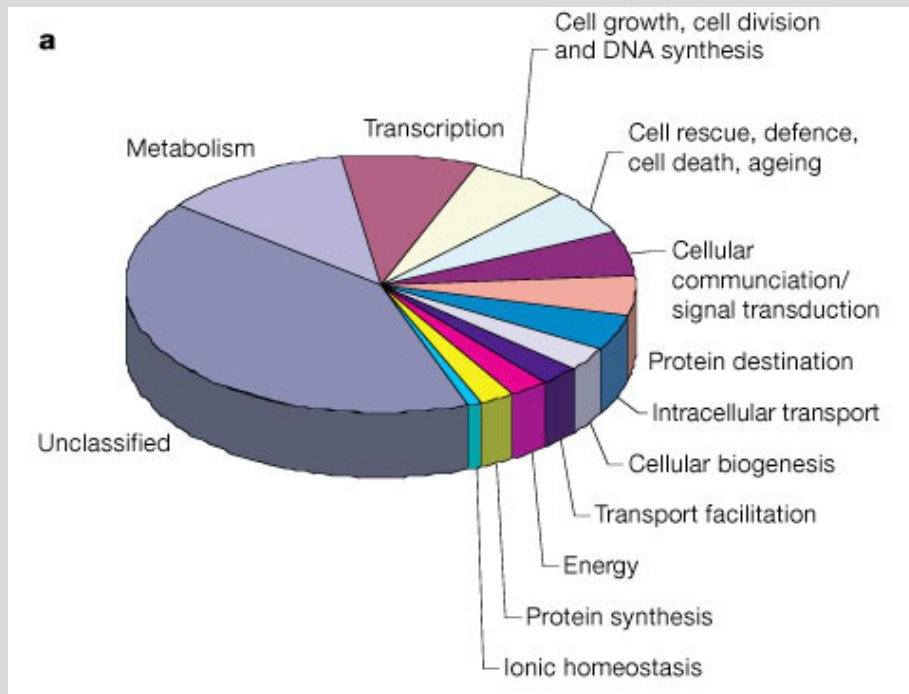
Genome sequencing factory at Sanger Centre (UK)



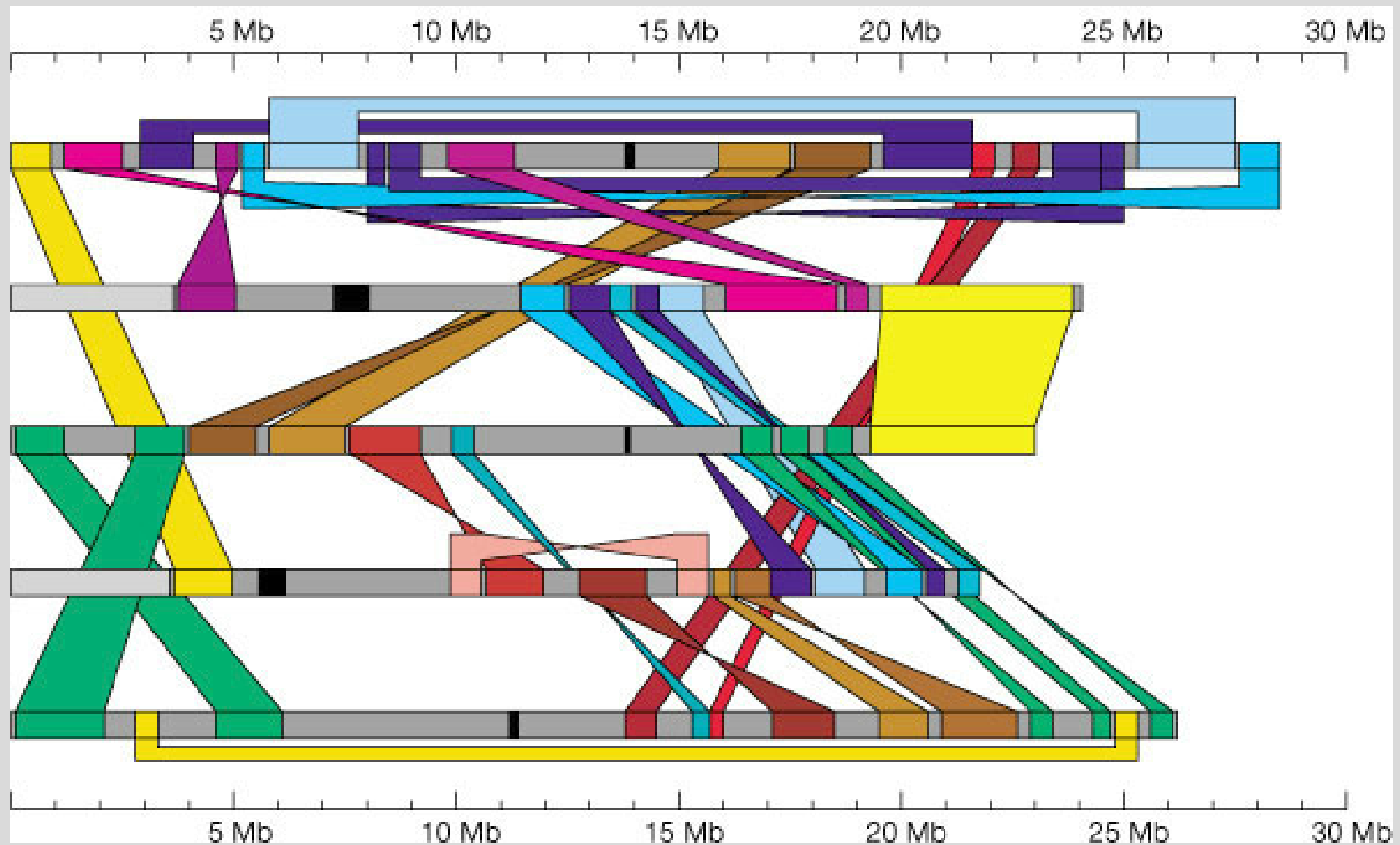
Data Storage Management



Clasificación de los genes de Arabidopsis en distintas categorías funcionales

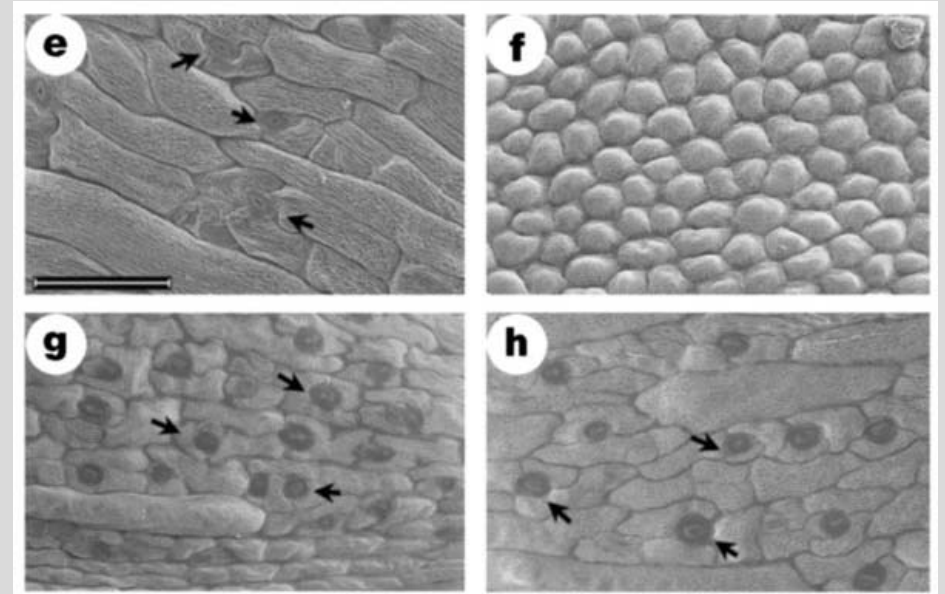
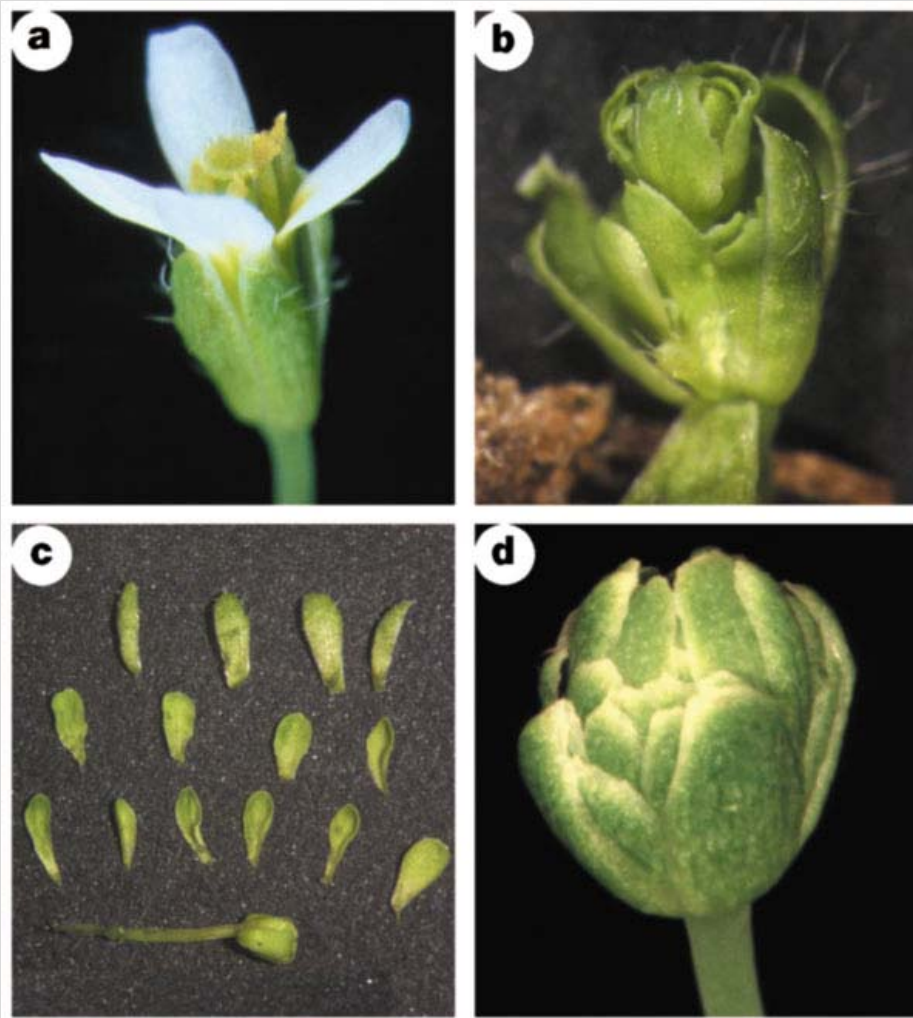


Duplicaciones en el genoma de Arabidopsis

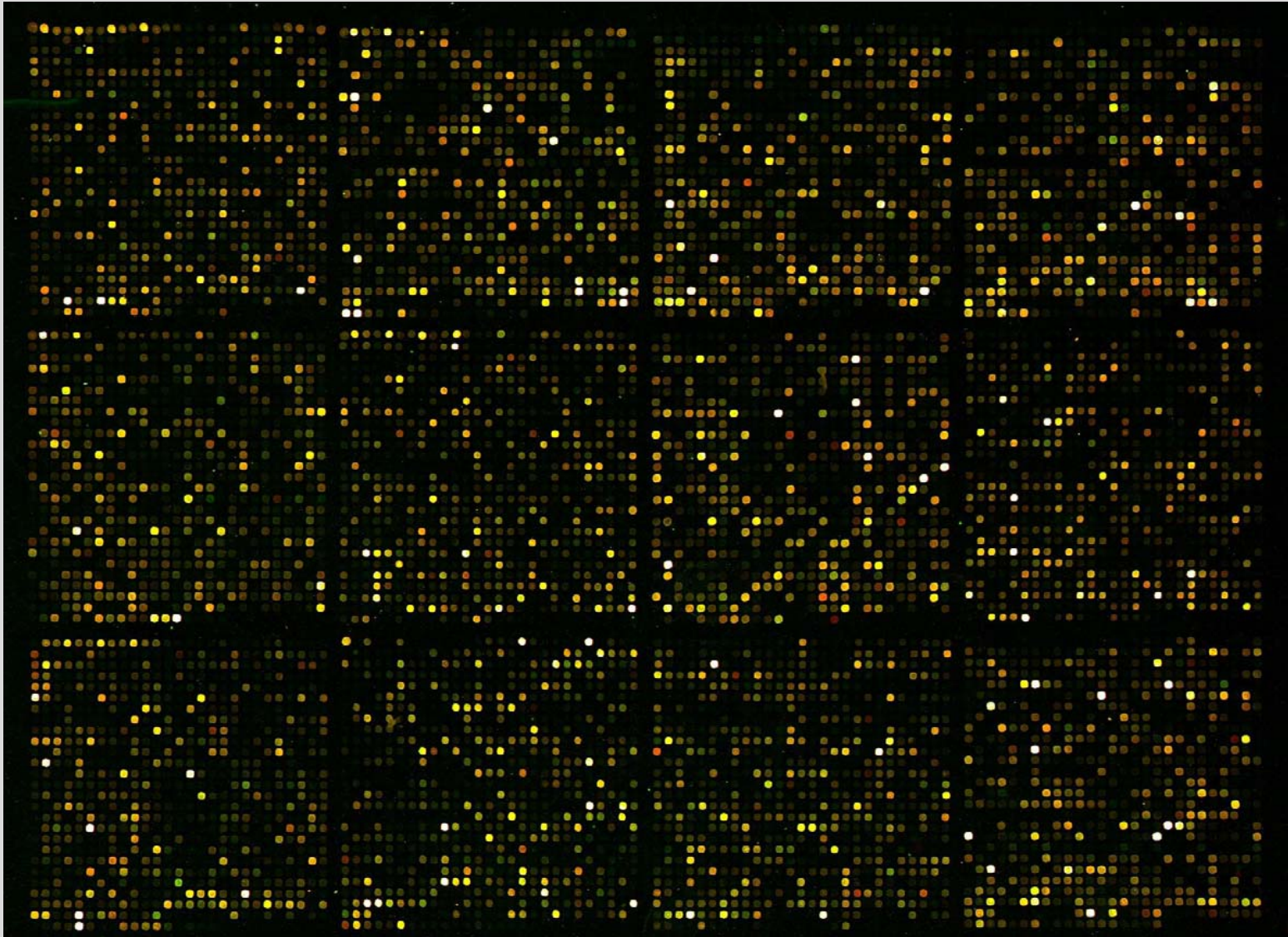


AGI (2000) *Nature* **408**:796

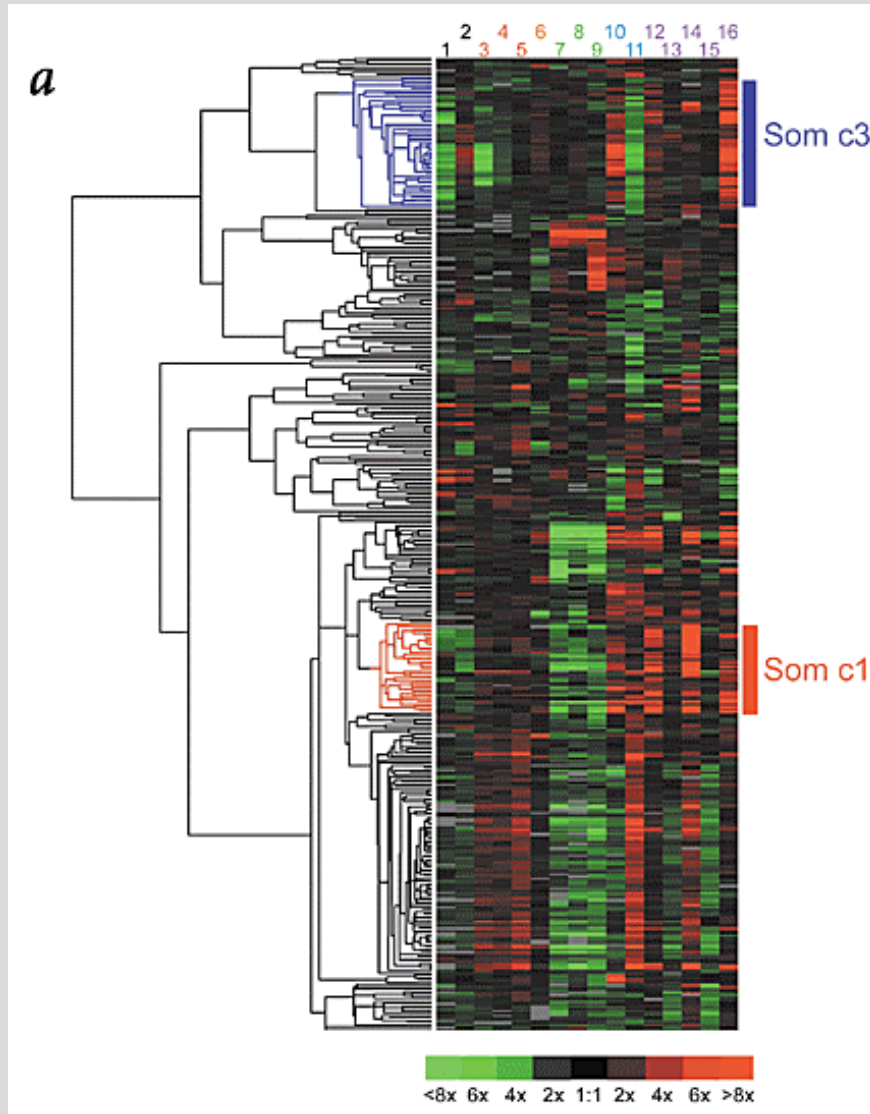
Redundancia funcional entre los genes reguladores SEP1, SEP2 y SEP 3



Transcriptómica: Chip de oligonucleotidos

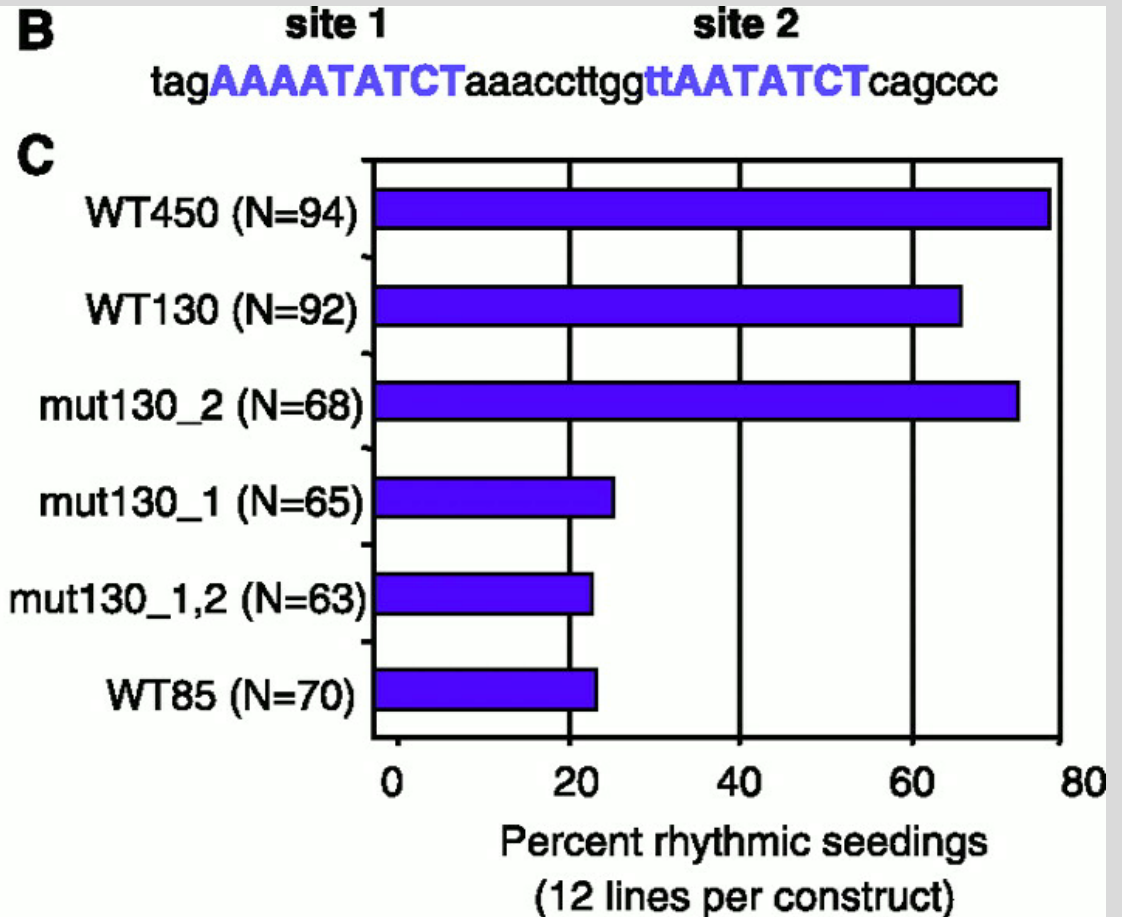
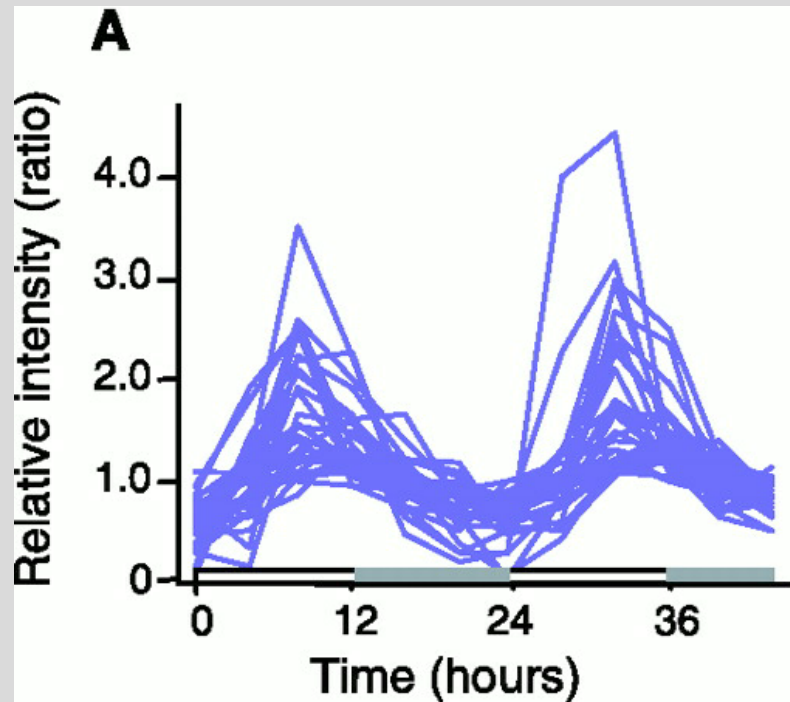


Agrupamiento de genes corregulados mediante métodos bioinformáticos



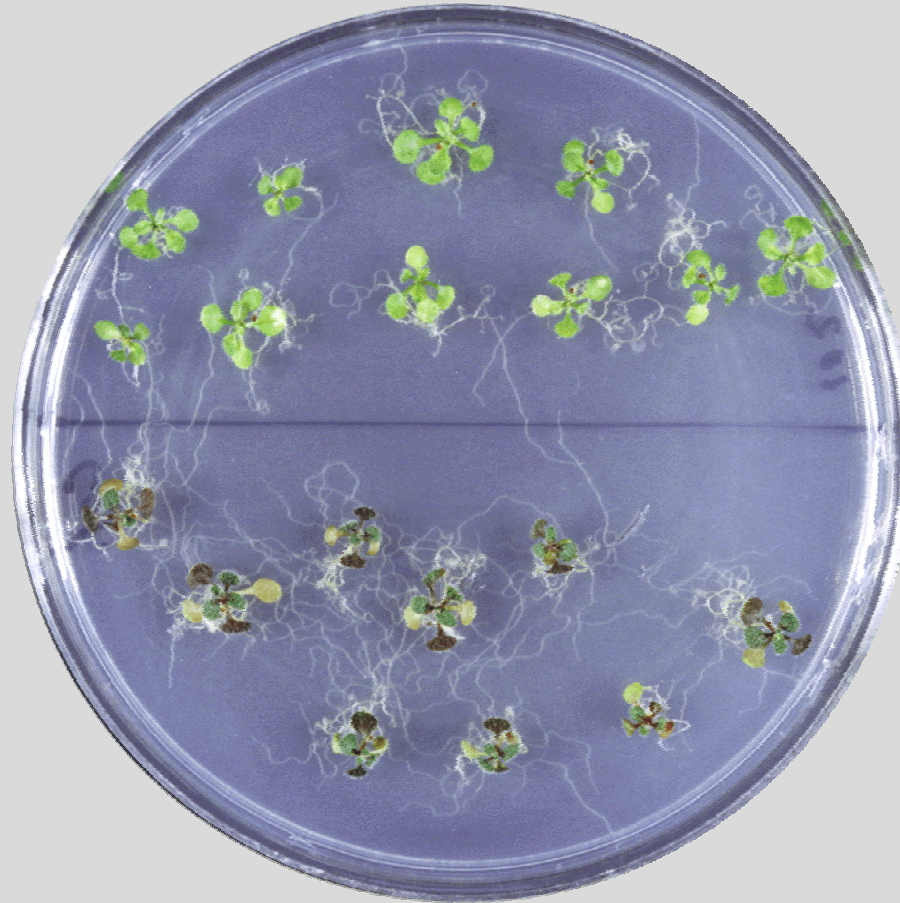
Maleck et al. (2000) *Nat. Genet.* **26**:403

Secuencias comunes en los promotores de genes corregulados



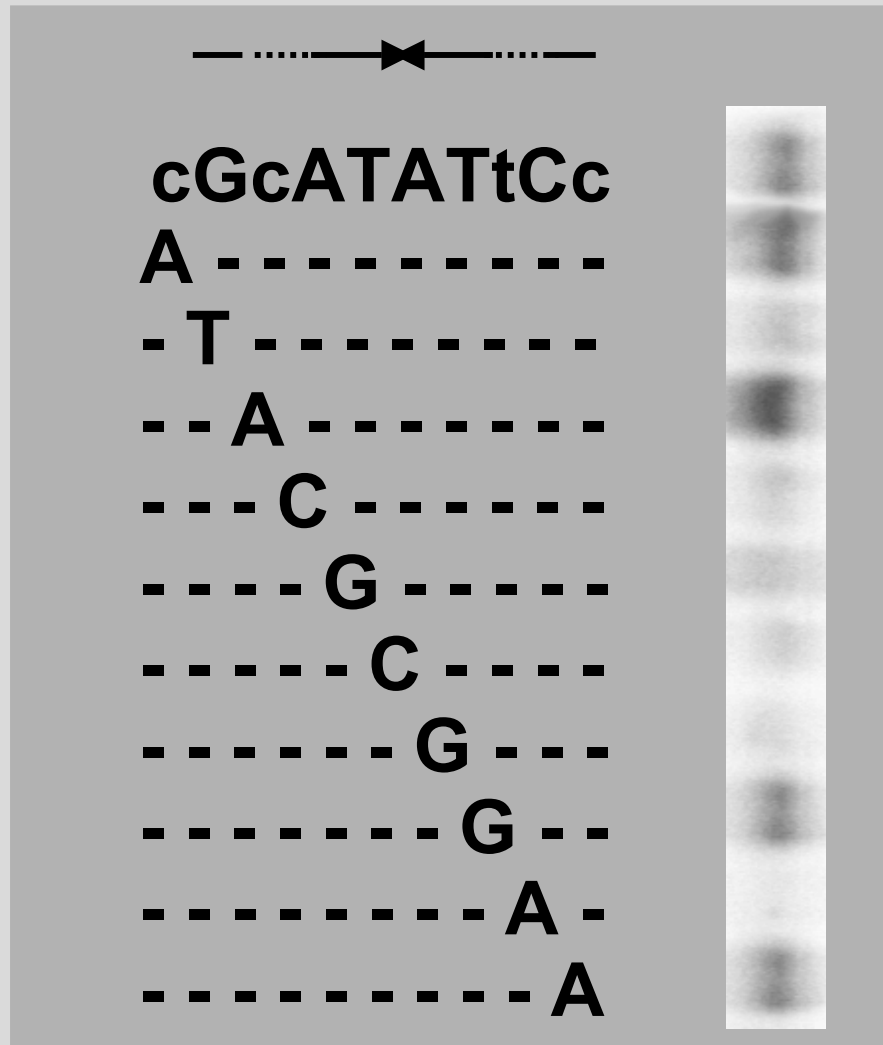
phr1 displays impaired Pi
starvation responses

phr1



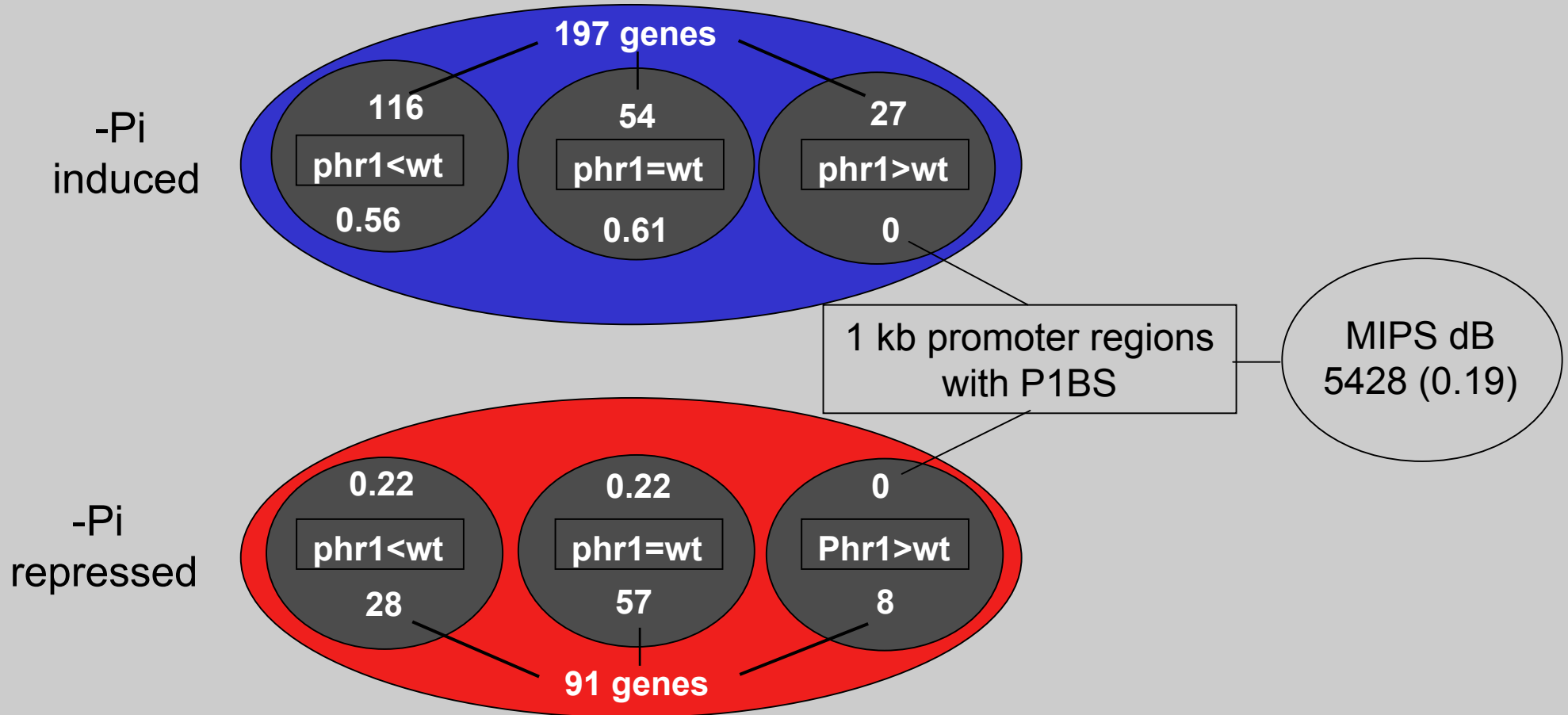
wt

PHR1 binds to a imperfect-palindromic sequence as a dimer



Transcriptome analysis of Pi starvation

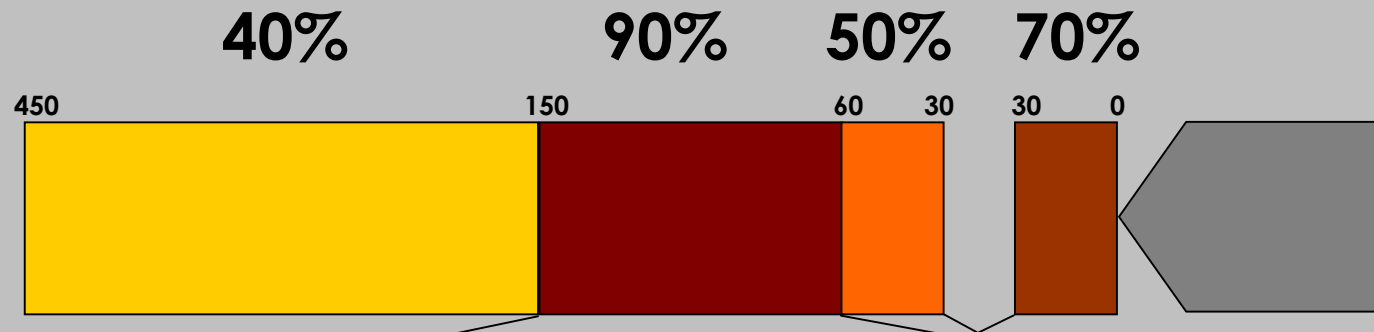
P1BS is overrepresented in promoters of Pi starvation inducible genes



Identificación de secuencias *cis*-reguladoras mediante análisis filogenético

- Las secuencias *cis*-reguladoras son de pequeño tamaño (4-8 bp), lo que complica su identificación en un conjunto de genes corregulados, dada la alta probabilidad de la presencia por azar de secuencias pequeñas en un promotor.
- Las secuencias no codificantes de genes ortólogos divergen de manera rápida durante la evolución, excepto los motivos que son funcionalmente relevantes, como lo son los motivos *cis*-reguladores importantes
- En especies próximas, las secuencias no esenciales aún no han divergido totalmente, siendo posible su alineamiento.

Análisis filogenético del promotor del gen *At4*, inducible por ayuno de fosfato. Alta conservación de P1BS y otras secuencias



TTTGGTAA --G**GCATATTCC**----GAAGATCCAAA ----35nt--TTT**CCATCATCT****c**CTTCCACTT

IPS1

TTTGGTAA -.G**GCATATTCC**----.G**AT**GATCCAAA ----36nt--TTT**CCATCATCT****g**CTTCCACTT

Mt4

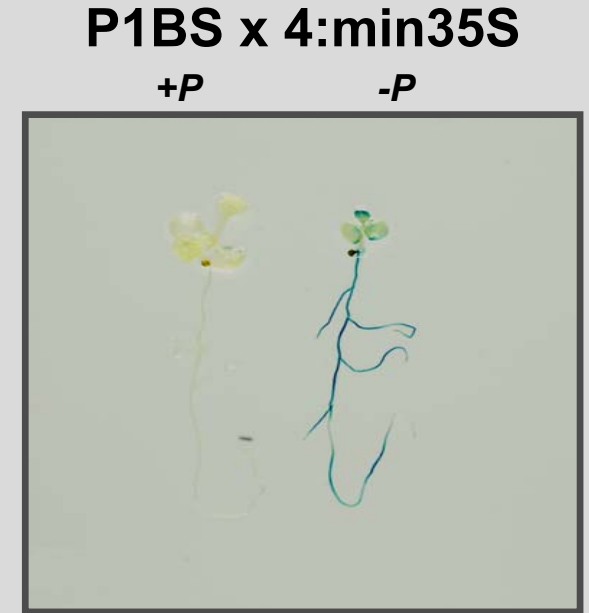
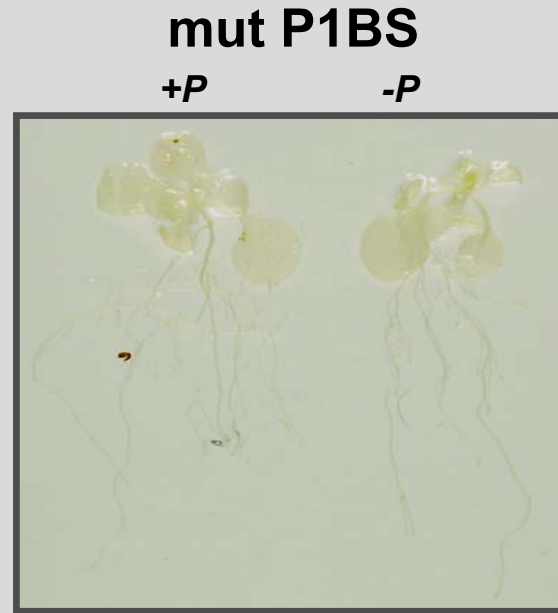
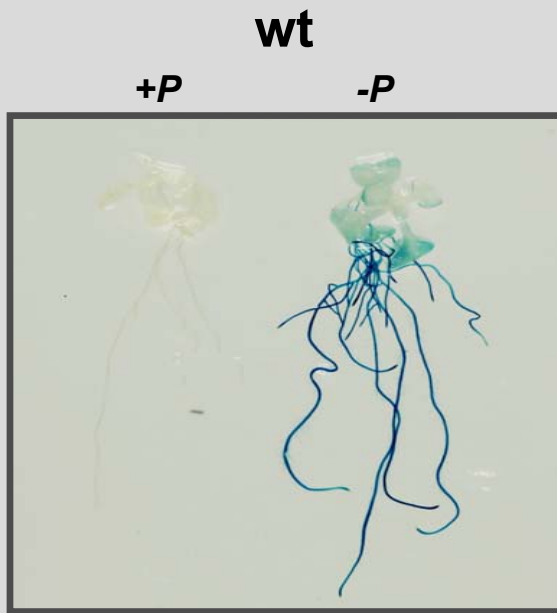
ATGTGGAACAAA-12nt-G**GCATATTCC**---..GAAGAT**gC**-----72nt----**GTCATCAT**-----

TPS1

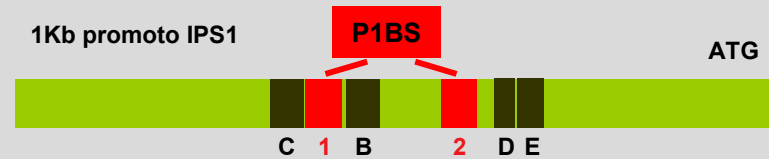
ATGTGGAACAAA-13nt-G**GCATATTCC**---..GAAGATCC -----397nt----**GTCATCAT**-----

P1BS

P1BS is a *bona fide* phosphate starvation response element



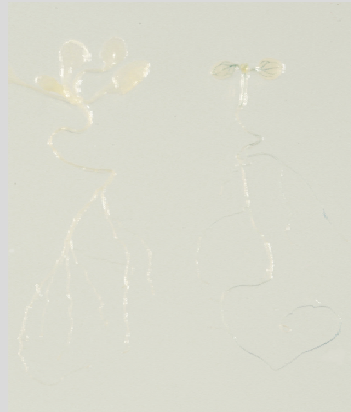
Conserved box B is also relevant for Pi responsiveness



wt



mut 1



mut 2



mut 2/B



mut 2/C



Phylogenetic footprinting and bioinformatics search for combinations of two motifs, detects novel cis-regulatory motifs important in Pi starvation control

- Conserved box B is not over represented in promoters of Pi starvation responsive genes
- *Combined PBS1 and box B is over-represented in promoters of Pi starvation responsive genes when restricting their distance to less than 10bps: 0.6 expected, 6 observed (in 215 PBS1 containing promoter regions; $p < 0.0001$)*

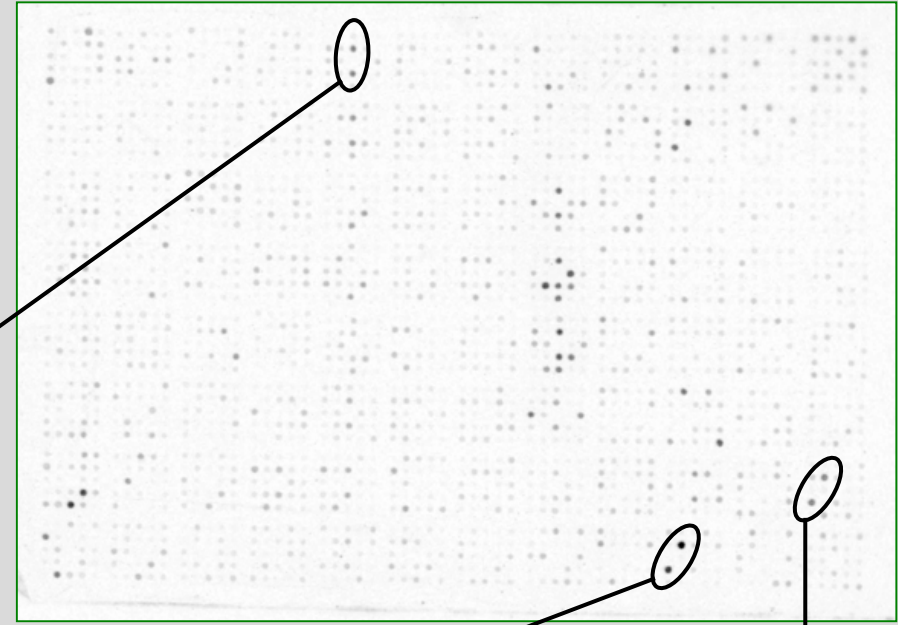
Transcriptome analysis of salt-treated *A. thaliana* (ColO)

Control : 4h. at 22°C

Salt treatment : 4h. 250 mM NaCl (22°C)



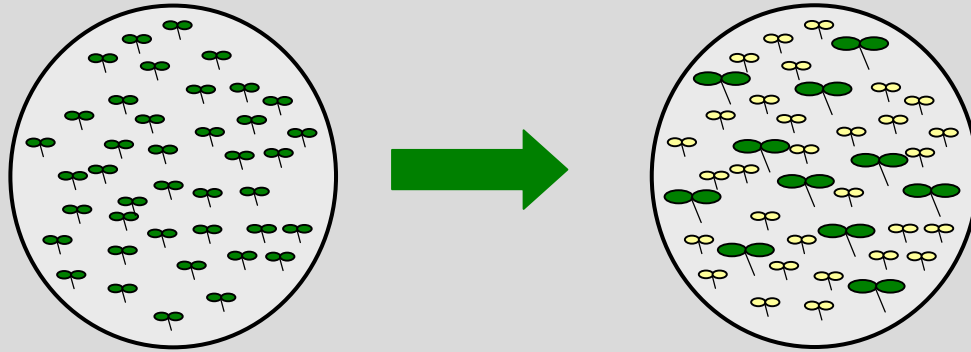
SIR1



ATHB7

ATHB12

Salt tolerance of transgenic lines with altered activity of *SIR1*

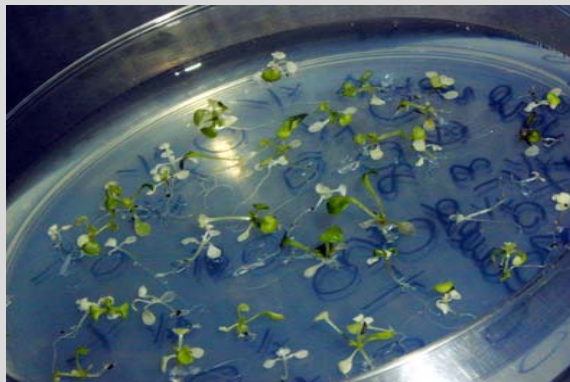


Germination on standard medium for 2 weeks

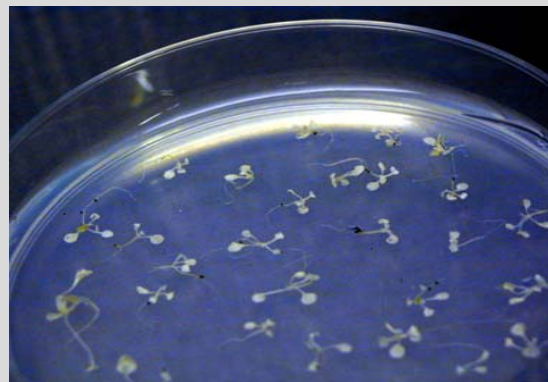
Transfer to NaCl added media:

- 50mM NaCl
- 100mM NaCl
- 200mM NaCl

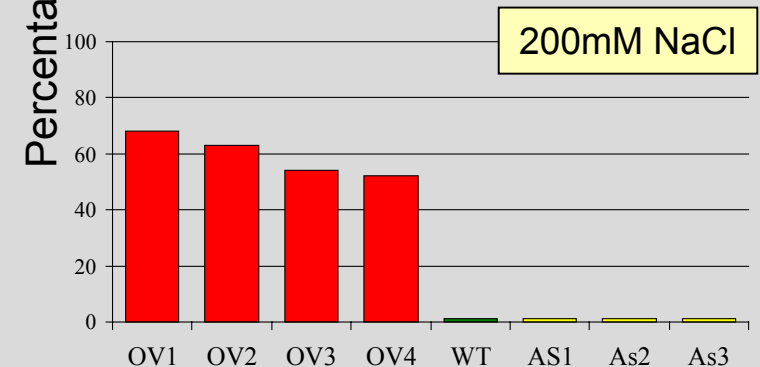
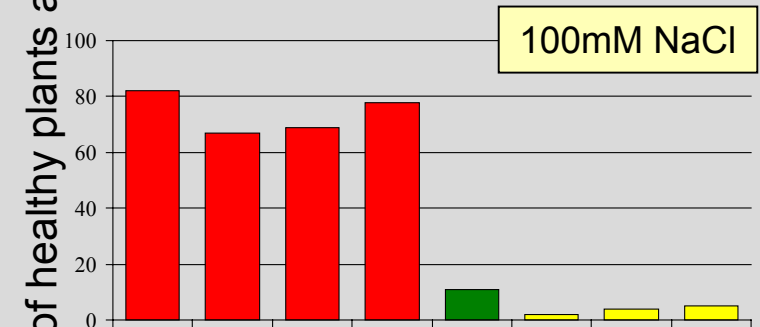
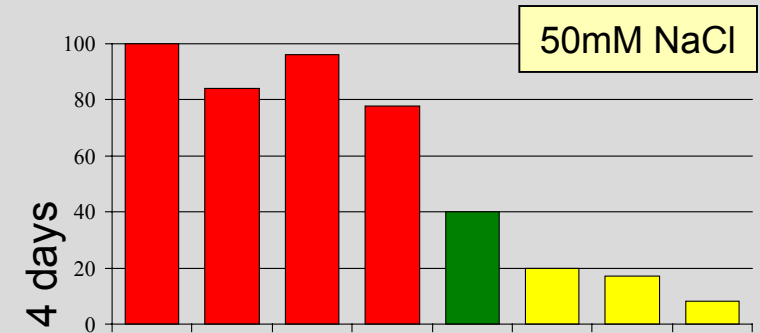
after 8 days on NaCl 200mM



SIR1 overexpression



WT



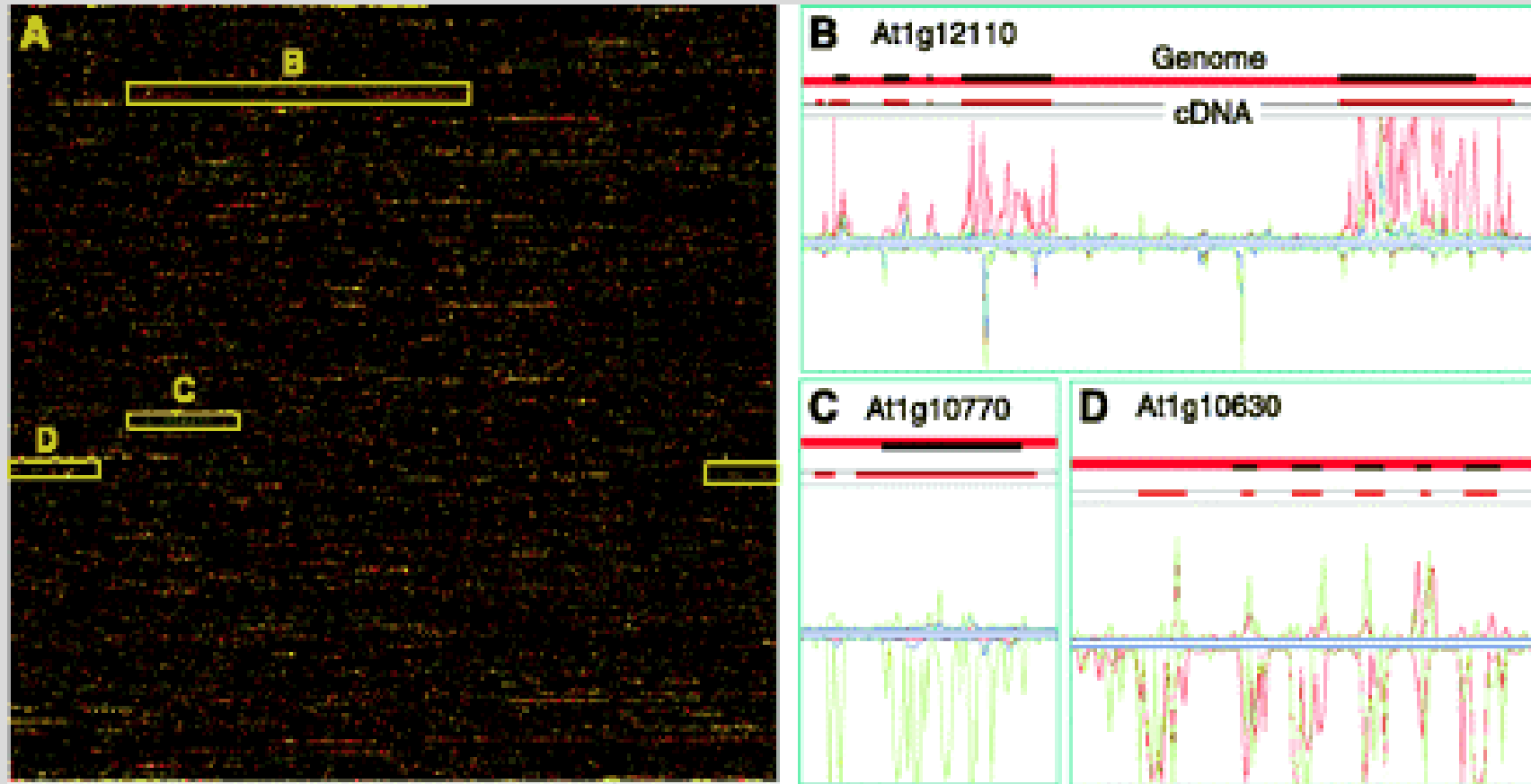
Transcriptómica/filogenómica

Se dispone de chips de DNA que representan el 90 % de los genes de Arabidopsis. Se ha comprobado la idoneidad de especies de la familia Brassicaceae para análisis filogenómicos de regiones promotoras

- Permite obtener información sobre el nivel de expresión de cada gen en distintos estadios de desarrollo, órganos y circunstancias medioambientales
- Permite definir grupos de genes corregulados e identificar secuencias reguladoras, especialmente con la ayuda de técnicas filogenómicas
- Permite definir fenotipos a nivel transcriptoma

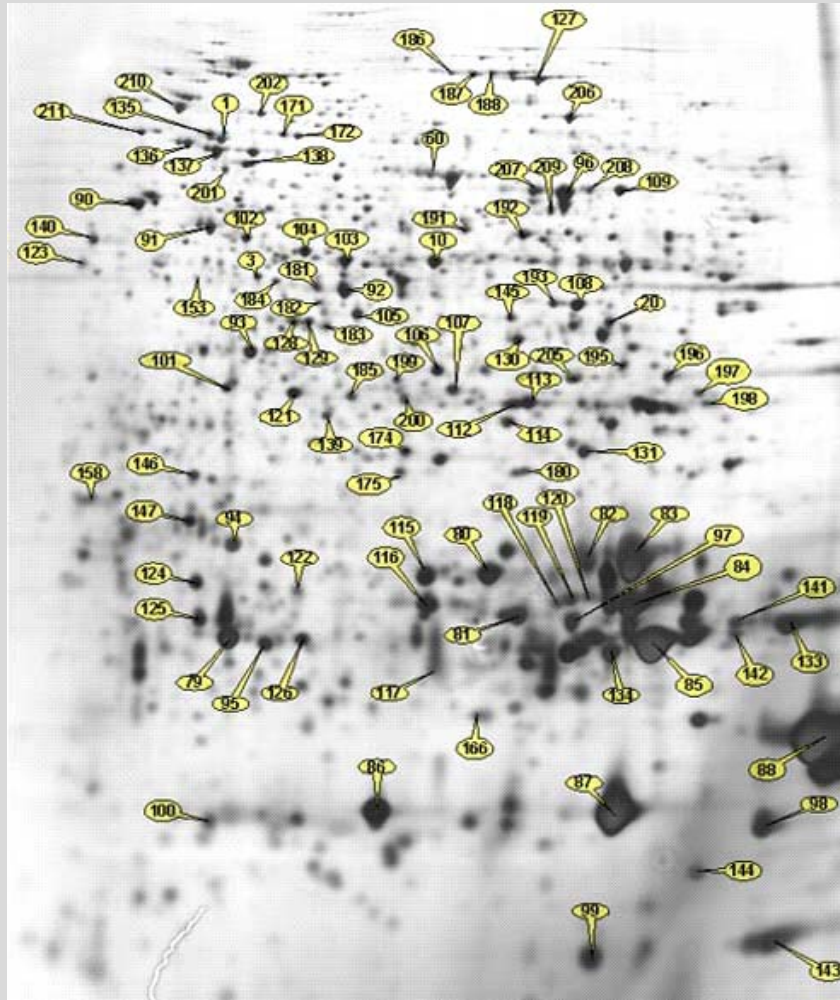
El genoma de Arabidopsis en un chip.

Definición de las unidades de transcripcion



El proteoma de una plántula de Arabidopsis

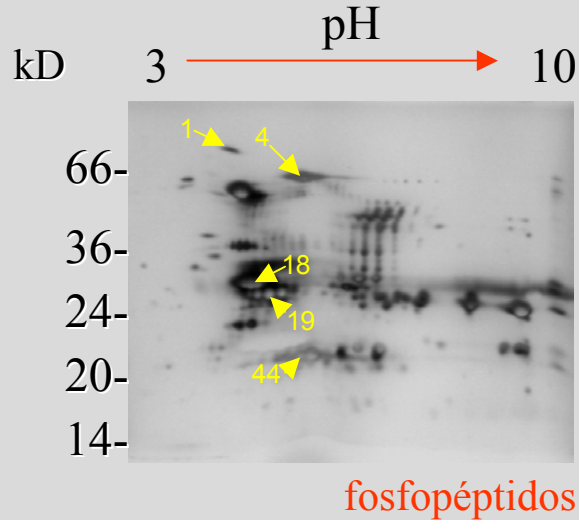
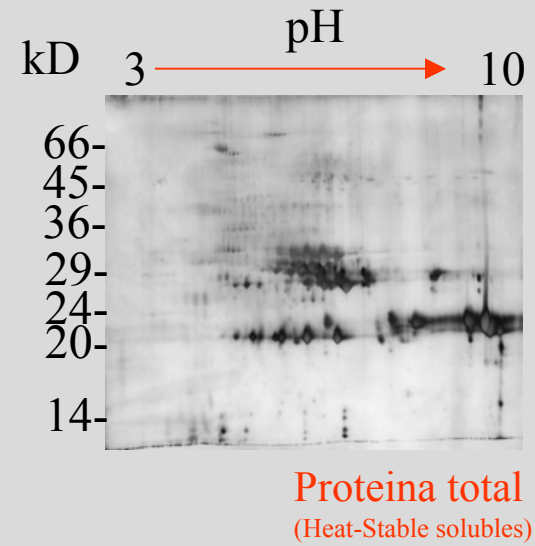
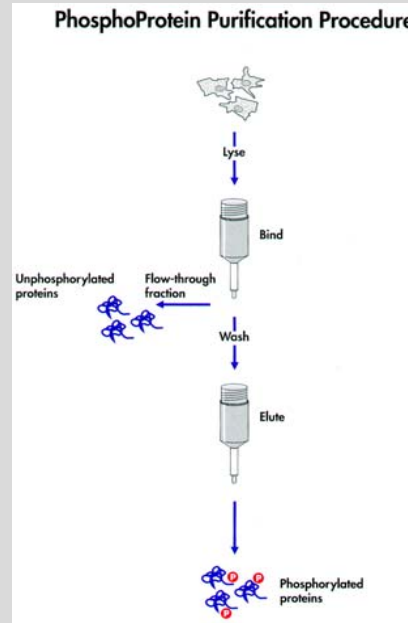
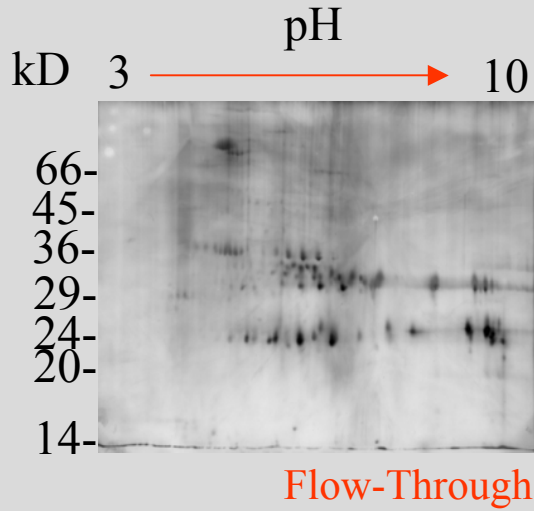
Proteínas identificadas in silico



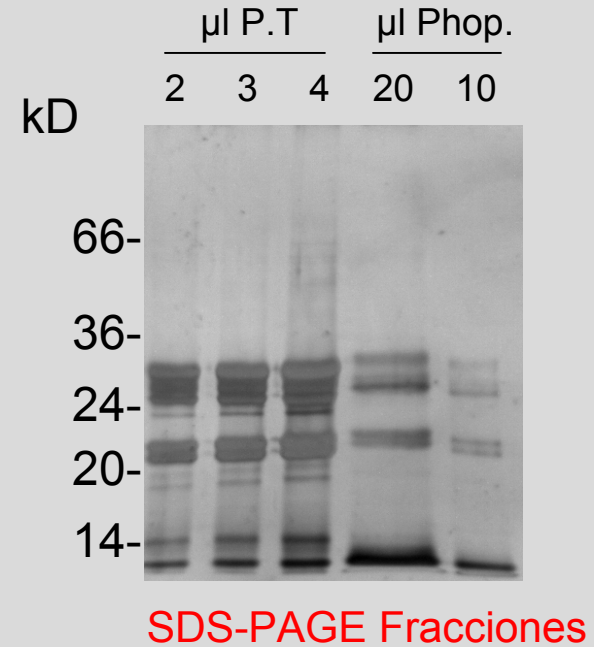
Experimento 1

Número	Proteína	Acceso (Swiss Prot-Ncbi-Mips)	Col0	3H	2H
7	Glutamine Synthase	Q9LV18	11.82	0	28.31
10	Ribulose biphosphate carboxylase large chain	O03042	13.14	26.78	0
32	Beta-Cruciferin 12S seed store protein precursor	P15456	17.30	45.12	0
33	Seed store protein	P15455			
34	Major latex protein	Q39132 or NP.172948	36.77	0	0
36	GAPDH Cytosolic	P25858	39.94	0	0
37	"	P25858	33.61	0	0
39	"	"			
39	"	"	35.10	0	0
40	"	"	34.33	0	0
56	LEA protein Group 5	Q96245 or BAA11016	22.46	59.72	0
60	Seed Maturation protein	Q9SIN3 or AAD22997	1.57	0	13.34
70	Seed store protein precursor	P15455			
76	Alfa-Cruciferin 12 S fragment	Q96318 or T04623	35.82	0	0
79	F411.20	O64873 or T02394	9.76	0	19.44
80	Alfa-Cruciferin 12 S fragment	Q96318 or T04623	72.57	0	17.57
83	"	"	3.91	0	12.89
86	Beta-Cruciferin 12S fragment	Q9ZWA9 or AAD10679	61.30	0	0
90	Disulfide Isomerase 1 precursor	Q9XI01 or AAD41430	26.16	0	0
91	LEA protein like	Q9LF88 or T47561	20.15	69.23	0
92	Mirosinase-Binding protein	O49326 or AAC04913	17.49	0	0
93	Actin 7	P53492	34.80	0	0
94	LEA like protein (LEA 32)	Q9ZRT9 or CAA10352	30.58	0	0
96	Beta-glucosidase	QLIF9 or BAB03050	28.46	0	10.11
97	Alfa- Globulin 12S fragment	P15455	33.26	0	0
99	Major Latex Protein	Q39132 or NP.172948	10.24	0	0
100	T1008.10 protein	Q9MO42 or NP.195750	44.76	0	0
101	F24B22.200. protein	Q9M382 or T47583	18.98	0	0
102	LEA protein	Q9SKP0 or AAD20140	27.61	0	0

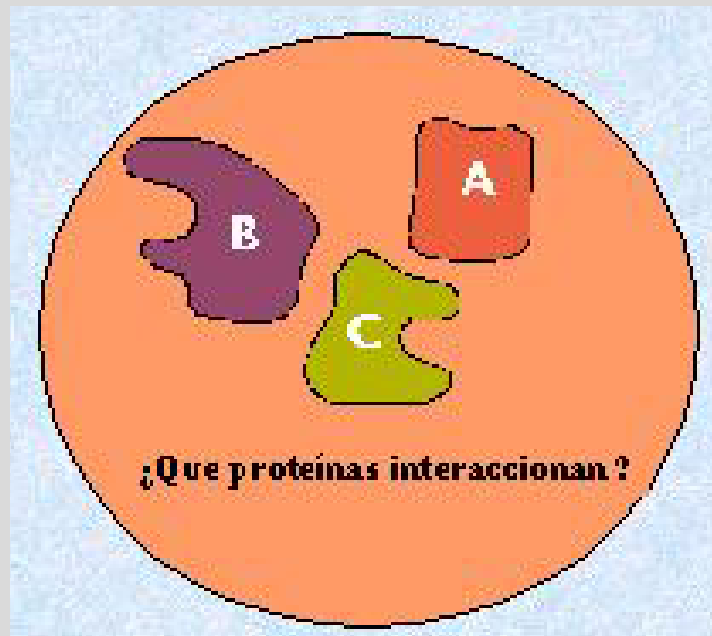
Fosfoproteoma



MALDI-TOF

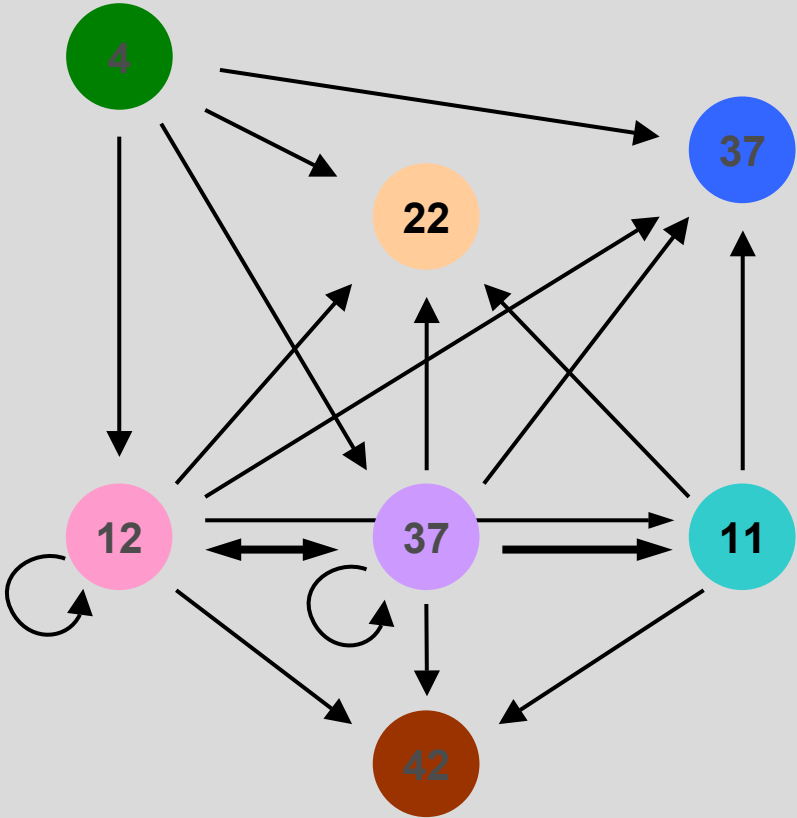


Estudios de interacciones entre proteínas. El método de los dos híbridos en levadura (Y2H)



Examples of TF:TF interactions

BD strain	AD strain	- H	LacZ	-U	Both orientations
RG_MYB1R 14	RG_MADS39	4	4	4	X
RG_MYB103	RG_MADS123	4	4	4	X
RG_ZFWRKY 4 	RG_ZFWRKY12	4	4	X	X
RG_ZFWRKY 4	RG_ZFWRKY22 	4	4	X	X
RG_ZFWRKY 4	RG_ZFWRKY37	4	4	X	X
RG_ZFWRKY 4	RG_ZFWRKY41	4	4	X	X
RG_ZFWRKY11	RG_ZFWRKY12	4	4	4	X
RG_ZFWRKY11	RG_ZFWRKY22	4	4	X	X
RG_ZFWRKY11	RG_ZFWRKY37	4	4	X	X
RG_ZFWRKY11	RG_ZFWRKY41	4	4	4	X
RG_ZFWRKY11	RG_ZFWRKY42	4	4	X	X
RG_ZFWRKY12 	RG_ZFWRKY12	4	4	4	X
RG_ZFWRKY12	RG_ZFWRKY22	4	4	4	X
RG_ZFWRKY12	RG_ZFWRKY37 	4	4	4	X
RG_ZFWRKY12	RG_ZFWRKY41	4	4	4	4
RG_ZFWRKY12	RG_ZFWRKY42	4	4	4	X
RG_ZFWRKY4 	RG_ZFWRKY12	4	4	4	4
RG_ZFWRKY41	RG_ZFWRKY22	4	4	4	X
RG_ZFWRKY41	RG_ZFWRKY37	4	4	4	X
RG_ZFWRKY41	RG_ZFWRKY41	4	4	4	X
RG_ZFWRKY41	RG_ZFWRKY42 	4	4	4	X

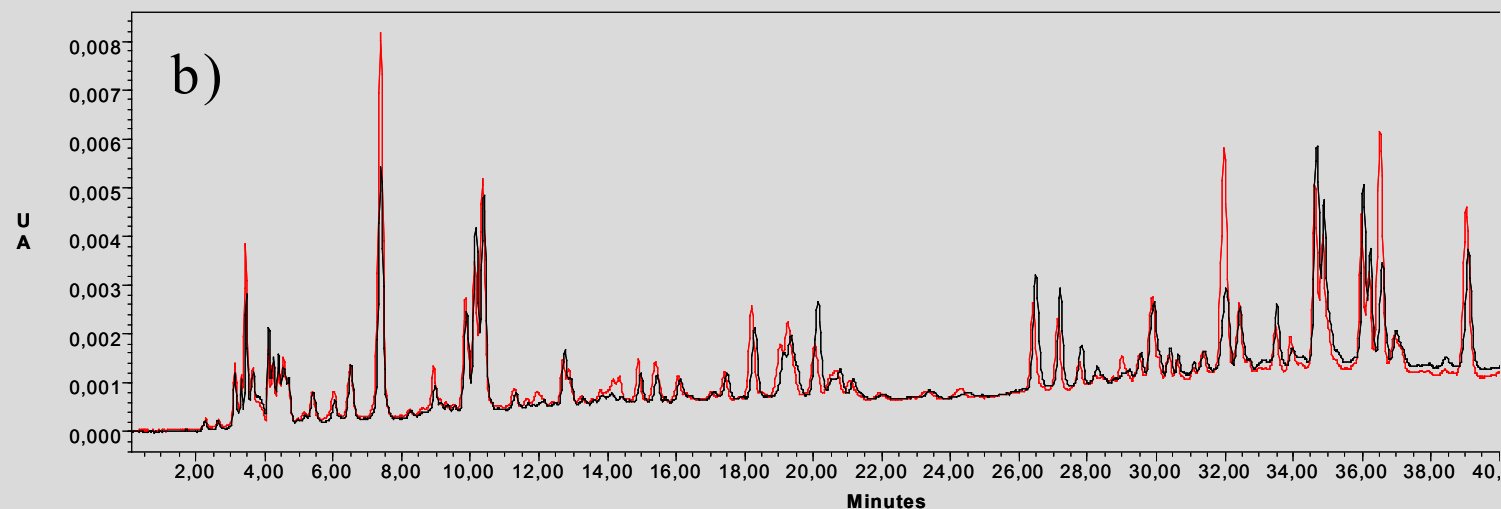
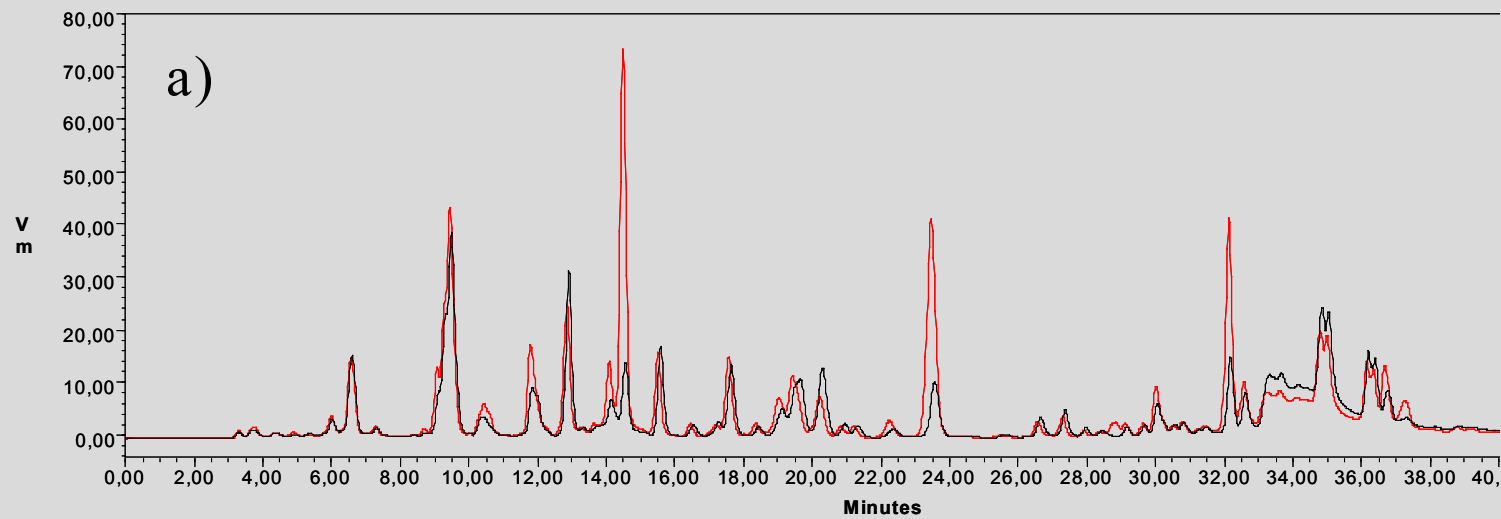


Proteómica

Se han puesto a punto métodos de fraccionamiento subcelular, y de separación y caracterización de proteínas de Arabidopsis. Se dispone de una colección de cDNAs de tamaño completo correspondiente al 50% de los genes de Arabidopsis (70% factores transcrip.)

- Provee información sobre el control traduccional y postraduccional
- Provee información sobre la localización subcelular del producto génico
- Provee información sobre interacciones entre las proteínas de un genoma
- Permite definir fenotipos a nivel de proteoma

HPLC-analysis of methanol soluble phenols in roots of Myb14



a) Fluorescence-chromatograms (excitation: 300nm, emission:400nm) and
b) UV-chromatograms (280nm)

Metabolómica

Se han desarrollado métodos cromatográficos (GC y HPLC) de alta resolución y sensibilidad que permiten analizar con gran eficacia más de 1000 metabolitos de Arabidopsis

- Permite establecer relaciones entre metabolitos y genes corregulados
- Provee información sobre el fenotipo a nivel de metaboloma

Genómica y genética inversa

La genómica funcional se apoya fundamentalmente en la metodología de la genética inversa, una aproximación experimental consistente en la determinación de la función de un gen a partir del estudio de los efectos fenotípicos de la manipulación de su actividad.

Estrategias para alterar la actividad génica:

- Disminución de la actividad:
Mutaciones por inserción, silenciamiento génico.
- Aumento de la actividad:
Plantas transgénicas que hiperexpresan el gen en cuestión

Búsqueda mediante ordenador de mutantes de inserción en Arabidopsis



[Home](#) | [About TAIR](#) | [Sitemap](#) | [Contact](#) | [Help](#) | [Order](#) | [Login](#) | [Logout](#)

[Search](#) | [Tools](#) | [Arabidopsis Info](#) | [News](#) | [Links](#) | [FTP](#) | [Stocks](#)

[TAIR Database](#)

[SeqViewer](#) | [MapView](#) | [AraCyc](#) | [BLAST](#) | [Wublast](#) | [FASTA](#) | [Pattern Matching](#) | [Restriction Analysis](#) | [Gene Hunter](#) | [Motif Analysis](#) | [Bulk Downloads](#)

NOTE: We are testing the hardware for WU-BLAST2 for faster searching. Please note that the server could be taken offline for maintenance at any time during this testing period. Please use our other [BLAST](#) service if this page is not accessible. The WU-BLAST2 fixes some alignment bugs in NCBI-BLAST and has a very nice summary graphic! Please try it out and report any bugs/problems to: informatics@arabidopsis.org.

Arabidopsis thaliana [WU-BLAST2](#) Search

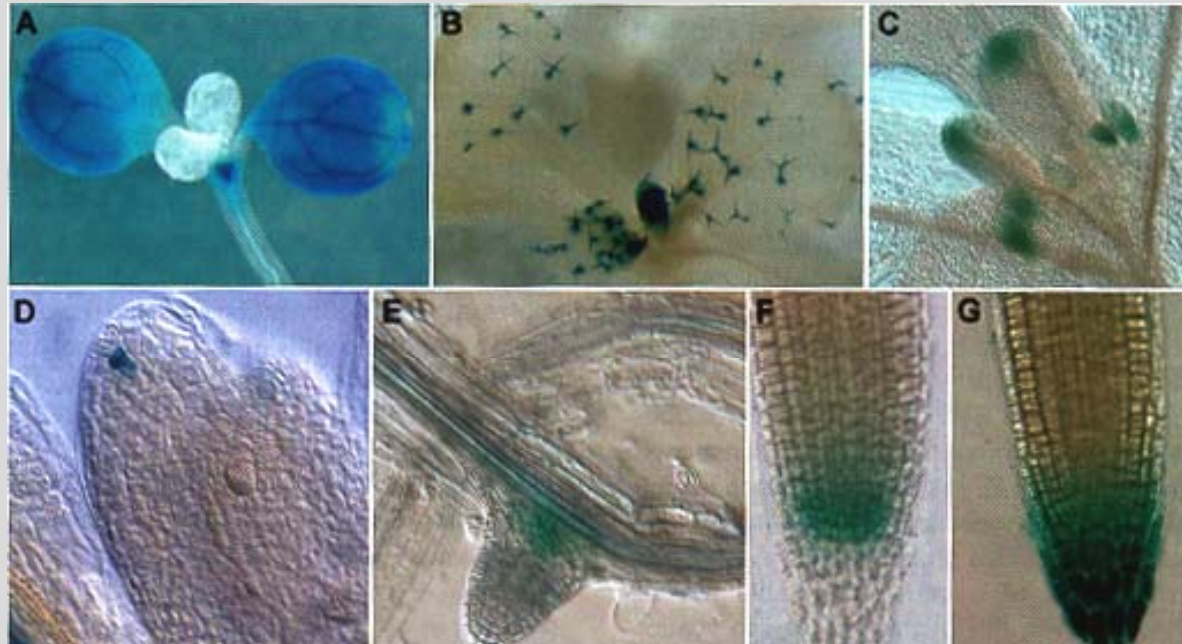
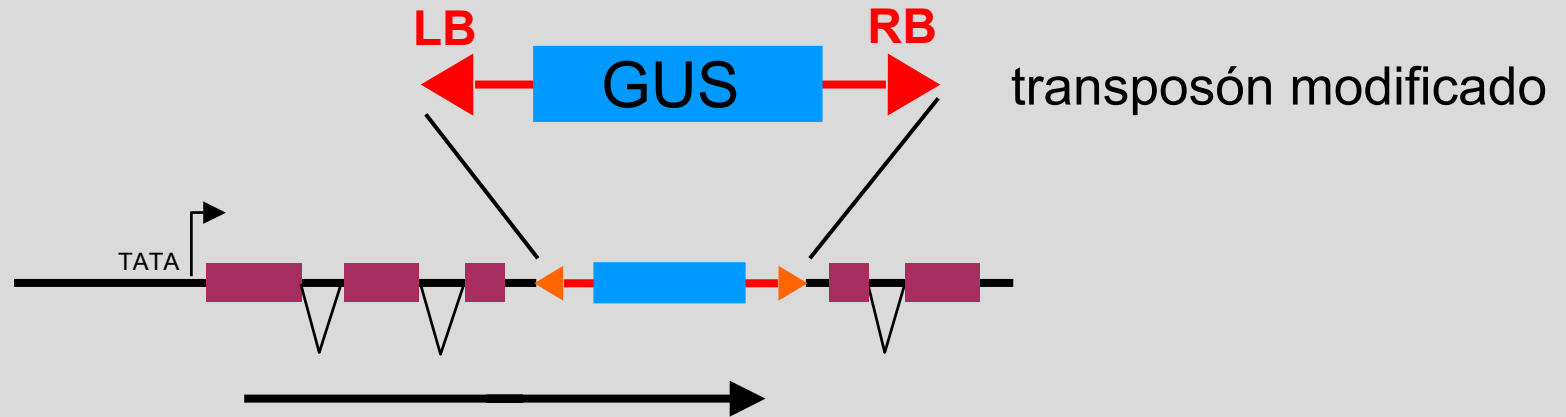
[BLAST](#) program:

[Datasets](#) **NEW** :

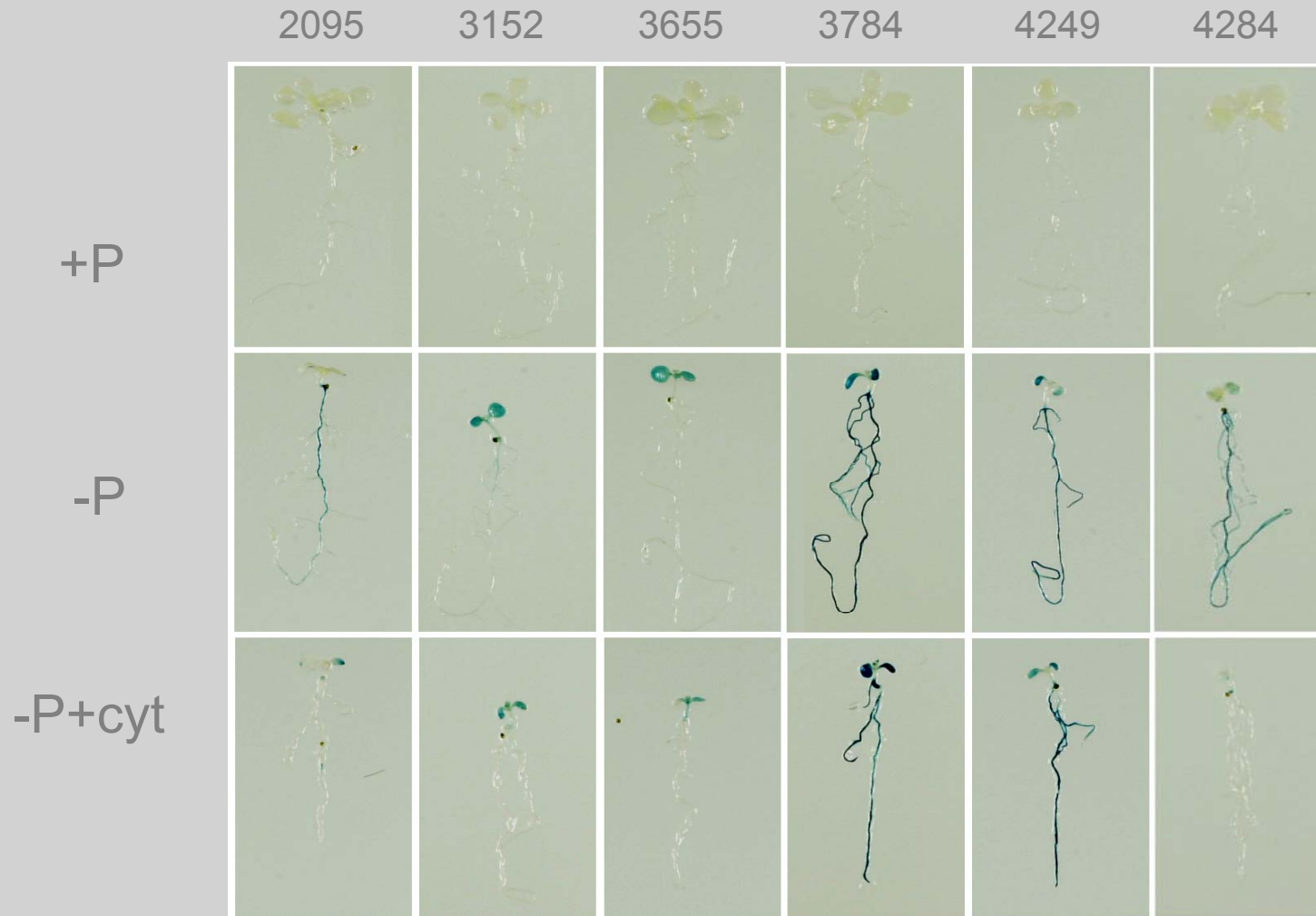
Input:

- ☒ [query sequence](#)
- ☐ locus name
(cds)
(At1g01030)

Transposones diseñados para visualizar la expresión génica

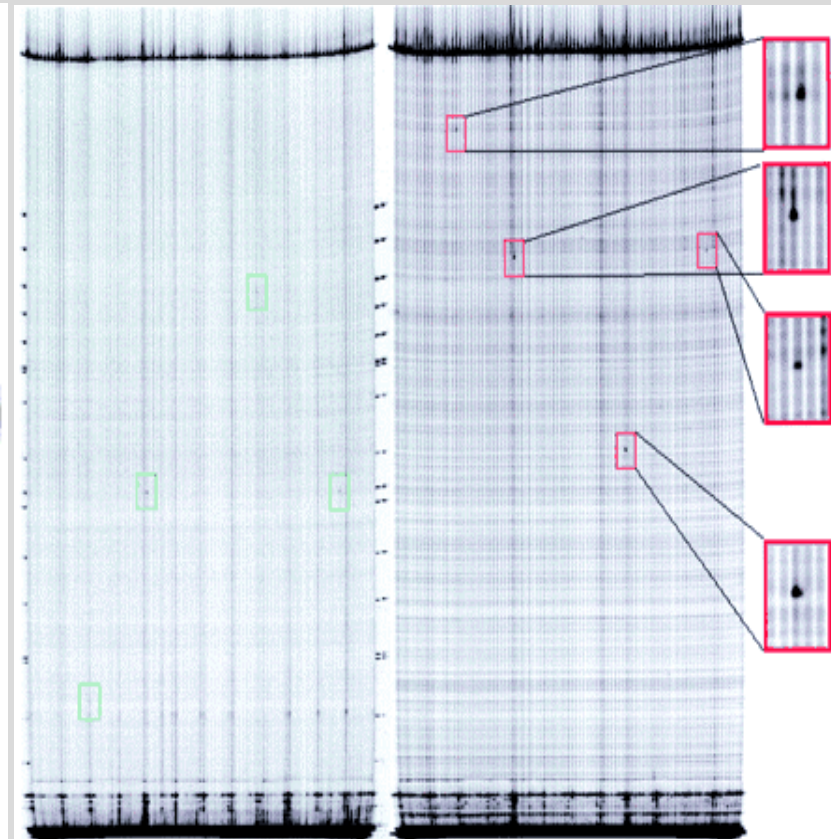
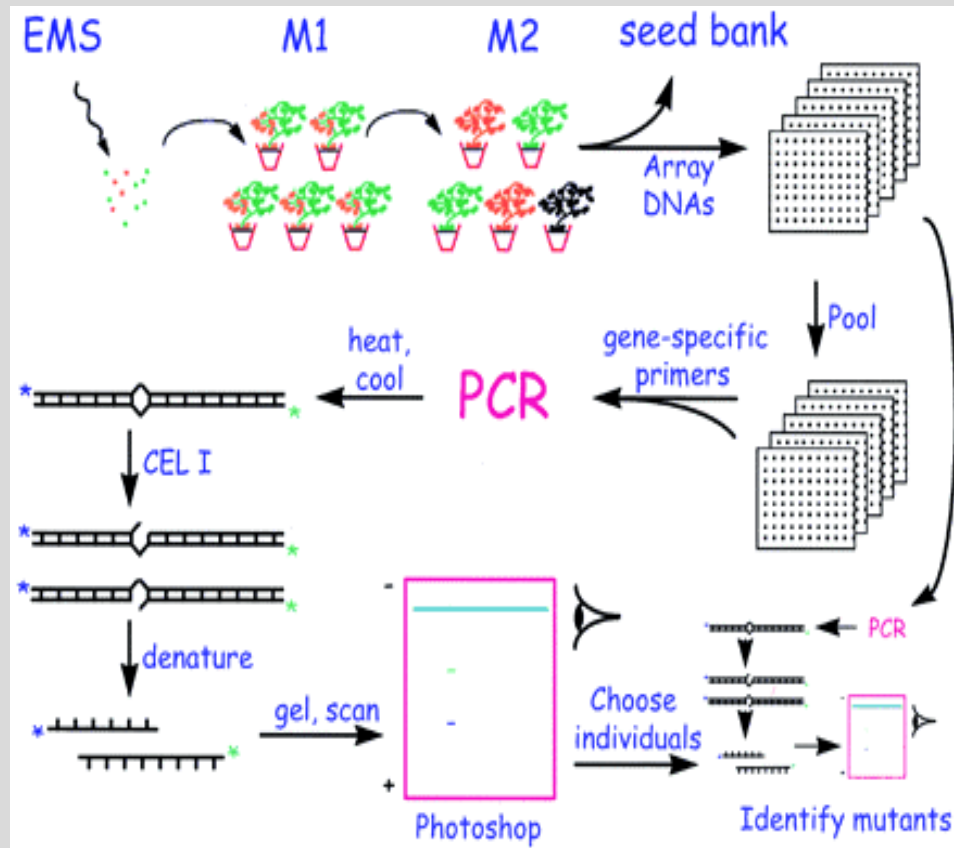


Identificación de mutantes de genes inducibles por ayuno de fosfato



Identificación de mutantes puntuales.

Reconocimiento de desapareamientos de una base por CEL I



Mutantes y su clonación

Existe una colección de más de 200.000 líneas independientes que contienen mutaciones de inserción en más de 20.000 genes de Arabidopsis. Asimismo, se han desarrollado métodos de identificación de mutantes puntuales, permitiendo identificar series alélicas de cualquier gen.

Por otra parte, se dispone de métodos de cartografiado de alta eficacia, y de la secuencia casi completa de un segundo ecotipo de Arabidopsis (Landsberg erecta), que facilita la identificación de marcadores polimórficos. De esta manera, la clonación posicional de cualquier gen se puede realizar en un espacio de tiempo corto (1-2 meses)

Representative goals of the 2010 project

- Identify cis regulatory sequences of all genes
- Identify regulatory circuits controlled by each transcription factor
- Determine biochemical function for every protein.
- Describe three-dimensional structures of members of every plant-specific protein family.
- Systems analysis of the uptake, transport, and storage of ions and metabolites.
- Describe globally protein-protein, protein-nucleic acid, and protein-other interactions at organ, cellular, and subcellular levels under various environmental conditions.
- Define a predictive basis for conservation versus diversification of gene function.
- Develop bioinformatics, visualization, and modeling tools that will facilitate access to all biological information about a representative virtual plant.

[Diapositiva 24](#)

[24](#)

Proyecto GEFA

proyecto integrado sobre GENómica Funcional en Arabidopsis

La genómica representa una renovación metodológica y conceptual que requiere una profunda transformación en la forma de organizar la actividad científica en Biología.

Arabidopsis es el sistema modelo por excelencia en biología vegetal, en el que la comunidad de científicos españoles es muy dinámica y activa, y representa un grupo ideal en el que ensayar la nueva forma de organizar la actividad científica en nuestro país.

El objetivo de esta propuesta es implantar en España plataformas de generación de recursos y provisión de servicios que posibiliten la aplicación de enfoques genómicos a las investigaciones en Arabidopsis.

Proyecto GEFA

Actividades

↗ Análisis del transcriptoma, mediante utilización de chips de DNA.

Investigador responsable: *Roberto Solano-CNB*

✂ Análisis del proteoma, el fosfoproteoma y el nucleoproteoma

Montse Pages-IBMB

✂ Definición y clonación del ORFeoma, ORFs del cromosoma 4

Pilar Carbonero-UPM; Crisanto Gutierrez-CBM

✂ Aislamiento de mutantes de inserción y secuenciación del sitio de inserción

Miguel Blazquez-IBMCP; Antonio Fdez Tiburcio-UB

↗ Generación e identificación de series alélicas de genes específicos.

José Miguel Mtnz. Zapater-CNB

↗ Cartografía génica automatizada, para facilitar la clonación posicional

José Luis Micoll-UMH

⌚ Explotación de la variabilidad natural en crucíferas, mediante la recolección en España y caracterización de ecotipos de *Arabidopsis* y de *Carrichtera*, y la realización de estudios filogenómicos para la identificación de regiones *cis*-reguladoras de la expresión génica.

Javier Paz-Ares-CNB

Rational plant improvement

The implications of genomics with respect to food, feed, and fibre production can be envisioned on many fronts. At the most fundamental level, the advances in genomics will greatly accelerate the acquisition of knowledge and that, in turn, will directly affect many aspects of the processes associated with plant improvement. Knowledge of the function of all plant genes, in conjunction with further development of tools for modifying and interrogating genomes, will lead to development of a robust genetic engineering discipline in which rational changes can be designed and modeled from first principles.