T. Page, C. Dean

Arabido

Science

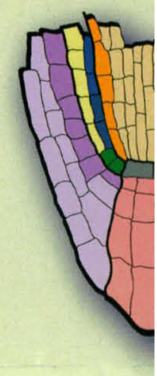
Genome Map

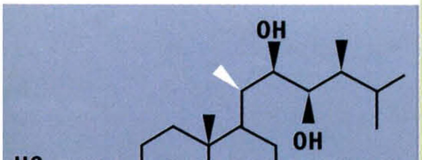
A Model System for Pla

Wild-type plant, 15-20cm in height at maturity.
(K. Sutliff)

Arabidopsis thaliana is a small plant in the mustard family that is being used in hundreds of laboratories worldwide as a model for studying plant biology. This simple angiosperm is becoming an important research tool for addressing fundamental questions of biological function and organization that extend across the major kingdoms of living organisms. Information obtained from *Arabidopsis* is also being used to develop improved crop plants and extend our understanding of cellular processes relevant to human health. *Arabidopsis* offers many advantages for genetic and molecular studies, including a short life cycle, small

genome
availab
tion. Th
was est
sis and
grams.
sortium
United





Characterizing the Ara

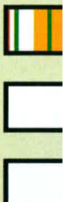
Key to Histogram

🟩 Completed sequence released to GenBank

🟧 Preliminary sequence released to GenBank

🟥 Clone library in preparation

Stat

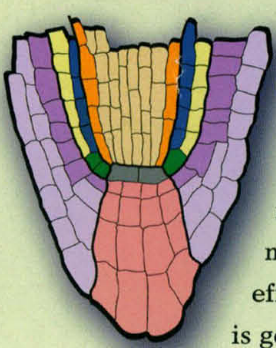


dopsis thalian

aps 9

r Plant Science

genome, ability to be transformed, widespread availability of mutants, and prolific seed production. The Multinational *Arabidopsis* Genome Project was established in 1990 to facilitate genome analysis and coordination of research and training programs. The *Arabidopsis* Genome Initiative is a consortium of laboratories in Europe, Japan, and the United States that are participating in large-scale sequencing efforts.



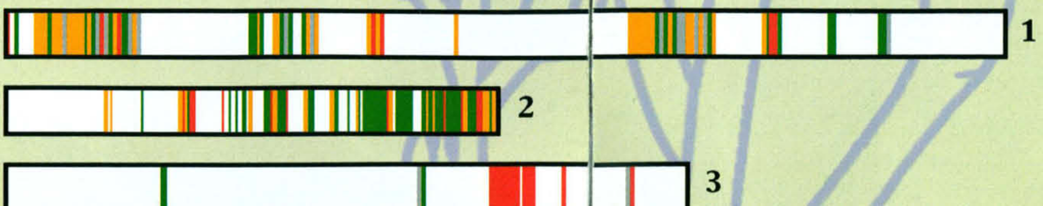
This chart illustrates recent advances in the molecular genetics and biology of *Arabidopsis*, with images along the margins showing the diversity of mutant phenotypes and research efforts. Summarized directly below is general information about the

Arabidopsis genome. Progress toward the goal
sequencing the entire genome by the end of t
year 2000 is shown in the histogram.

Three related maps of each chromosom
presented in this chart. The classical map show
estimated locations of mutant genes on the ba
recombination frequencies. The recombinant i
(RI) map includes cloned genes and molecula
markers linked to the physical map of overlapp
DNA segments that serve as the backbone of the sequencing initiative. These maps were constructed from data available through August 1998. Most of the mapped genes and a representative collection of molecular markers and cloned DNA segments are shown. To illustrate the

Arabidopsis Genome

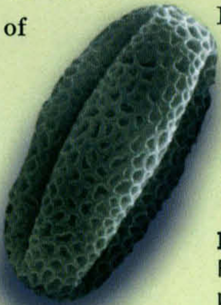
Status of Genome Sequencing Initiative



ana

ard the goal of
the end of the
m.
chromosome are
al map shows
es on the basis of
combinant inbred
nd molecular
of overlapping
ackbone of
ps were
rough
enes

A
he



richness of information available, a portion of the map has been expanded to display all identified genes within a 350-kb section of the genome.

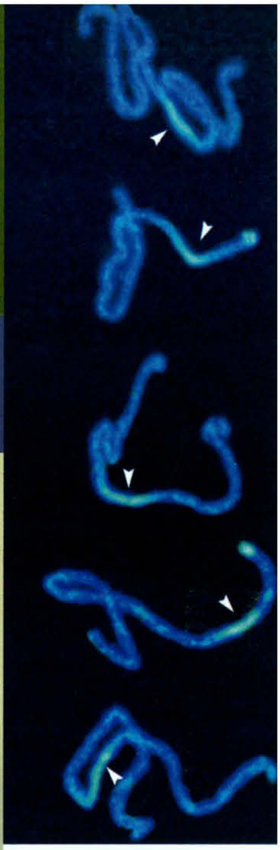
For current information on *Arabidopsis* biology, maps, sequencing efforts, and community resources, refer to the *Arabidopsis thaliana* Database (AtDB) at:

http://genome-www.stanford.edu/Arabidopsis/

Additional information can be obtained from the accompanying article published in the 23 October 1998 issue of *Science*.

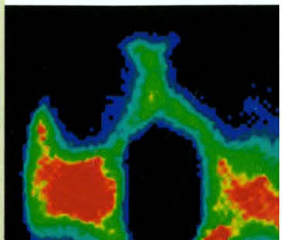
Inset Credits:
Flower (*K. Sutliff*); section of a root (*B. Scheres*); pollen grain (*D. Preuss*); embryo in a seed (*C. M. Liu*)

Estimated genome size	120 Mb
Number of chromosomes	5
Completed sequence in GenBank	>30 Mb
Estimated number of genes	20,000
Number of mutant loci identified	>1,000
Number of mapped mutant genes	>460
Number of mutant genes cloned	>110

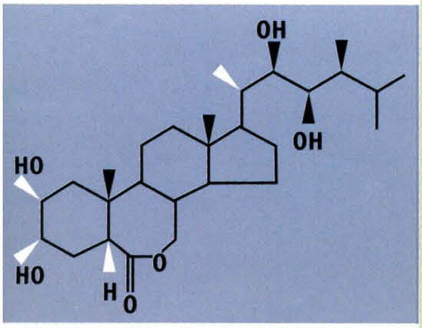


Karyotype of the five pachytene chromosomes.
(P. Fransz)

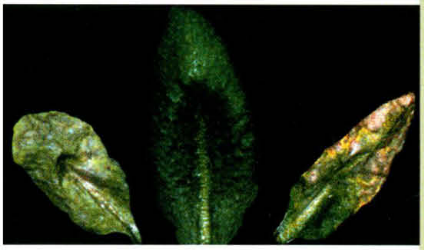
Mutant and normal rosettes in short-day photoperiod.
(T. Page, C. Dean)



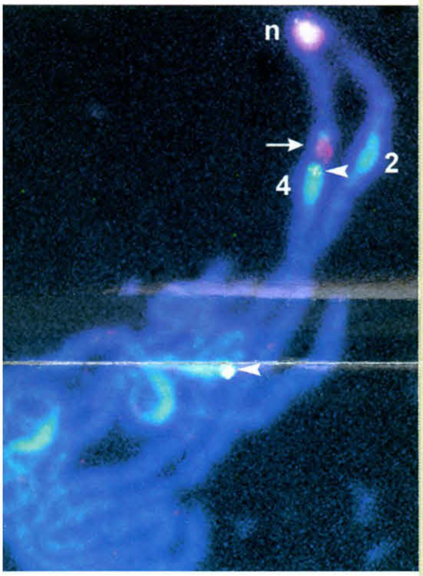
Wild-type plant, 15-20cm in height at maturity.
(K. Sutliff)



The plant hormone, brassinolide. *(J. Chory)*

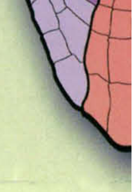


Leaves infected with bacterial pathogen.
(J. Dangl)



In situ hybridization of pachytene chromosomes.
(P. Fransz)

improved crop plants and extend our understanding of cellular processes relevant to human health. *Arabidopsis* offers many advantages for genetic and molecular studies, including a short life cycle, small



Characterizing the *Ara*

Key to Histogram

🟩 Completed sequence released to GenBank

🟧 Preliminary sequence released to GenBank

🟥 Clone library in preparation
OR sequencing in progress

▨ Clone verification in progress

Stat

8/14

Genetic and Physical M

Classical Map of Mutant Genes

■ Uncloned
🟧 Cloned
Mapped with mutant genes as reference points

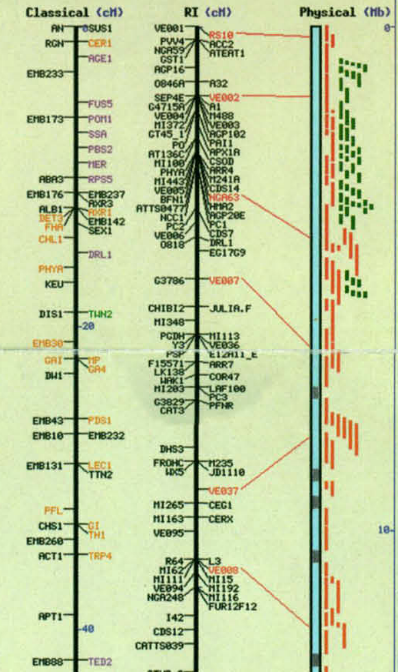
🟪 Uncloned
🟩 Cloned
Mapped relative to markers on the RI map

RI Map of Mo

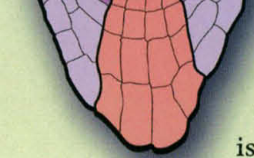
🟥 Location of re
indicated on

* *Length adjuste*

1



2



margins showing the diversity of mutant phenotypes and research efforts. Summarized directly below is general information about the

August 1998. Most of the mapped genes and a representative collection of molecular markers and cloned DNA segments are shown. To illustrate the

Arabidopsis Genome

Status of Genome Sequencing Initiative

al Maps

Map of Molecular Markers*

ocation of representative markers
ndicated on the physical map

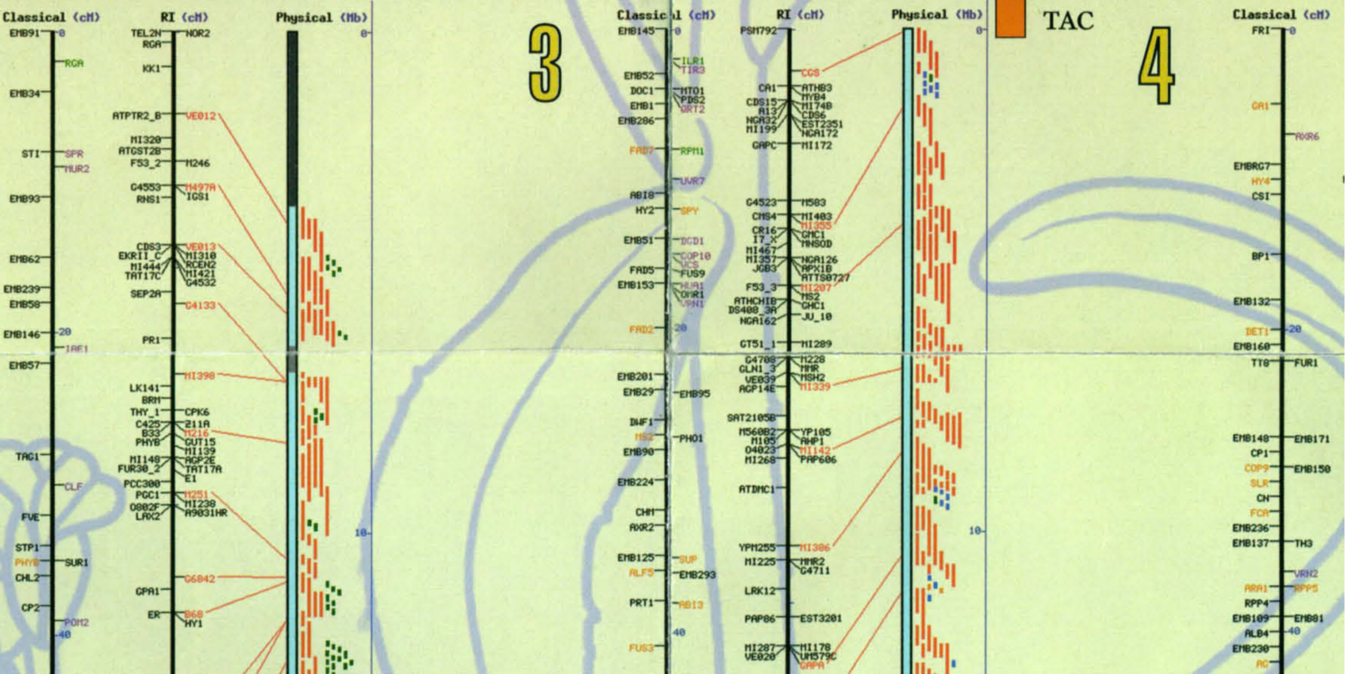
Length adjusted to match classical map

Physical Map

- 🟦 Regions covered by contigs
- ▨ Regions not covered by contigs
- ■ Nucleolar organizer regions

Clones Used to Construct the Physical Map

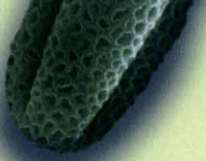
- 🟥 YAC
- 🟩 BAC
- 🟦 PAC
- 🟧 TAC



enes

NA

he

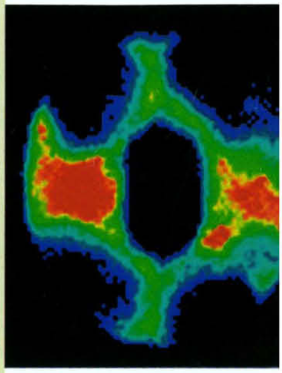


from the accompanying article published in the 23 October 1998 issue of *Science*.

Inset Credits:
Flower (*K. Sutliff*); section of a root (*B. Scheres*); pollen grain (*D. Preuss*); embryo in a seed (*C. M. Liu*)

Mutant and normal rosettes in short-day photoperiod.
(T. Page, C. Dean)

Estimated genome size	120 Mb
Number of chromosomes	5
Completed sequence in GenBank	>30 Mb
Estimated number of genes	20,000
Number of mutant loci identified	>1,000
Number of mapped mutant genes	>460
Number of mutant genes cloned	>110
T-DNA tagged lines available	>25,000
Molecular markers on RI map	>790
Nonredundant ESTs sequenced	>24,000



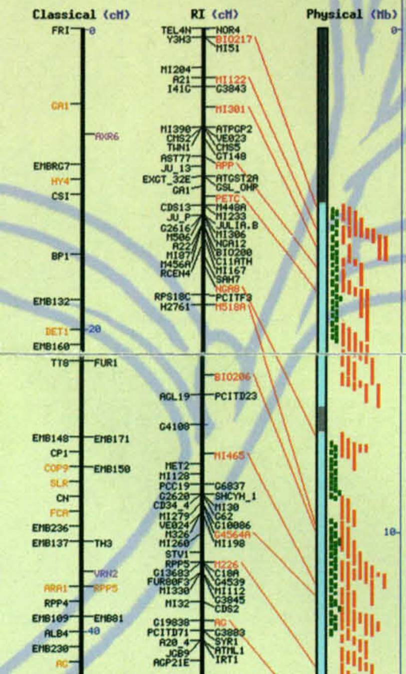
Calcium imaging in stomatal guard cells.
(K. Kuchitsu, J.I. Schroeder)

Dwarf plant with shortened internodes (*Science*, 10 Mar 1989).
(D. Patton, K. Feldmann)

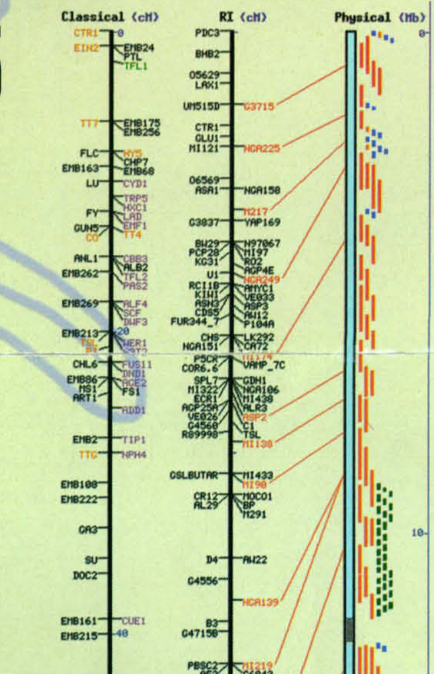
---

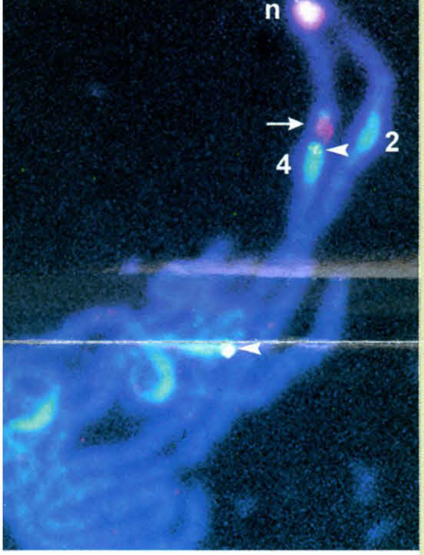
o Construct Map

BAC, bacterial artificial chromosome;
EST, expressed sequence tag;
PAC, P1 artificial chromosome;
RI, recombinant inbred;
T-DNA, DNA from *Agrobacterium* used to disrupt genes;
TAC, transformation-competent artificial chromosome;
YAC, yeast artificial chromosome.

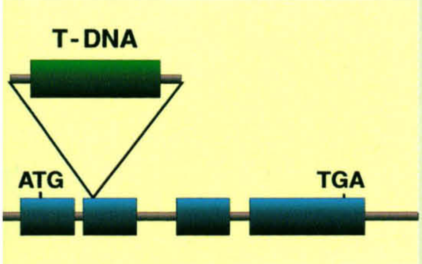


5



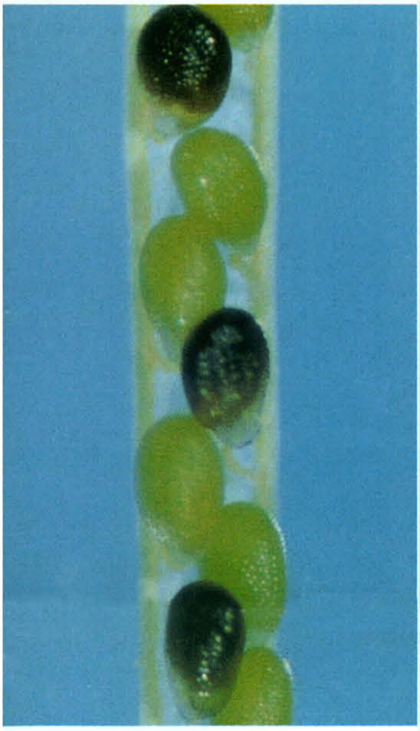


In situ hybridization of pachytene chromosomes.
(P. Fransz)

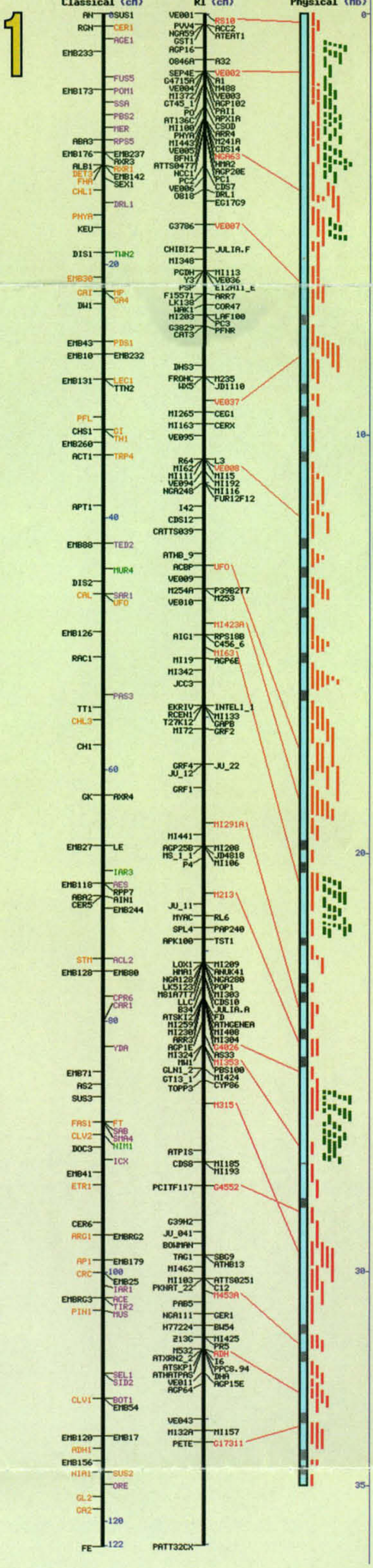


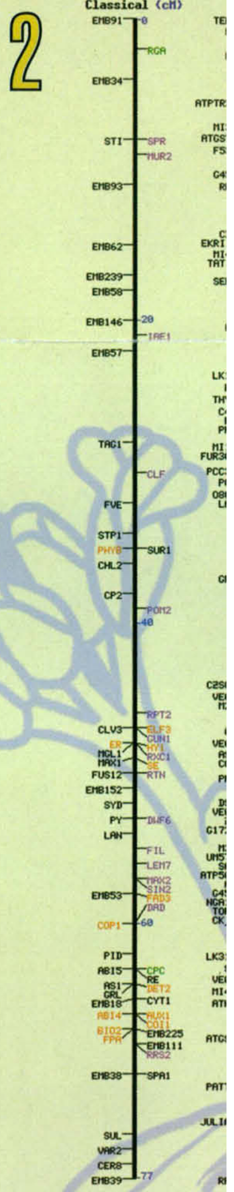
▲ T-DNA insertional mutagenesis.
(T. Nickle, D. Meinke)

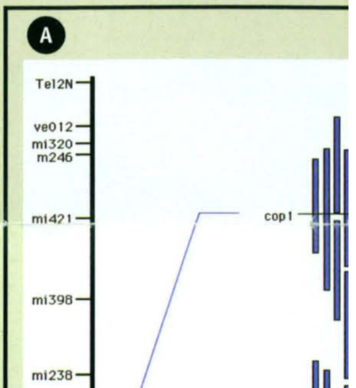
▼ Mutant seeds (dark) segregating in an immature silique.
(A. Peeters, M. Koornneef)

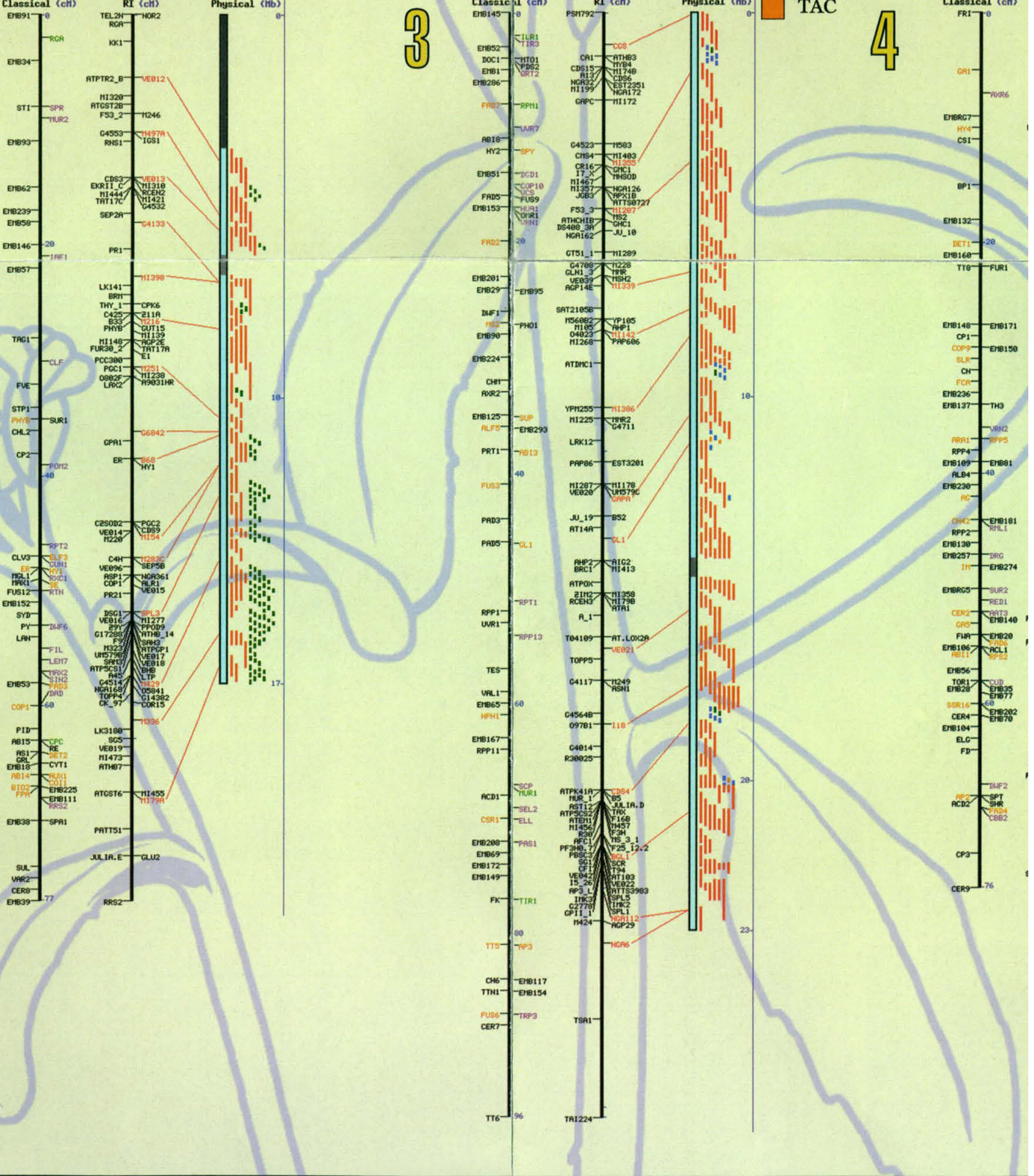


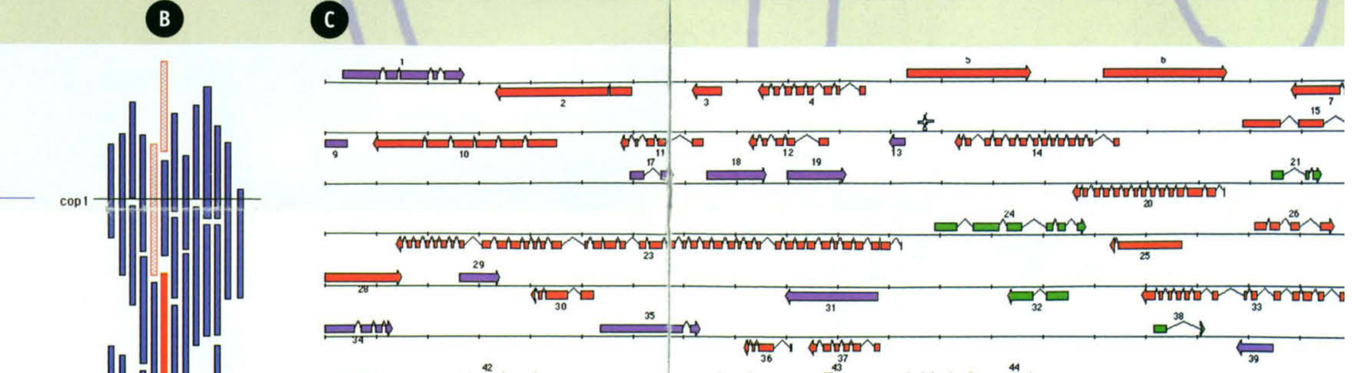
CREDITS

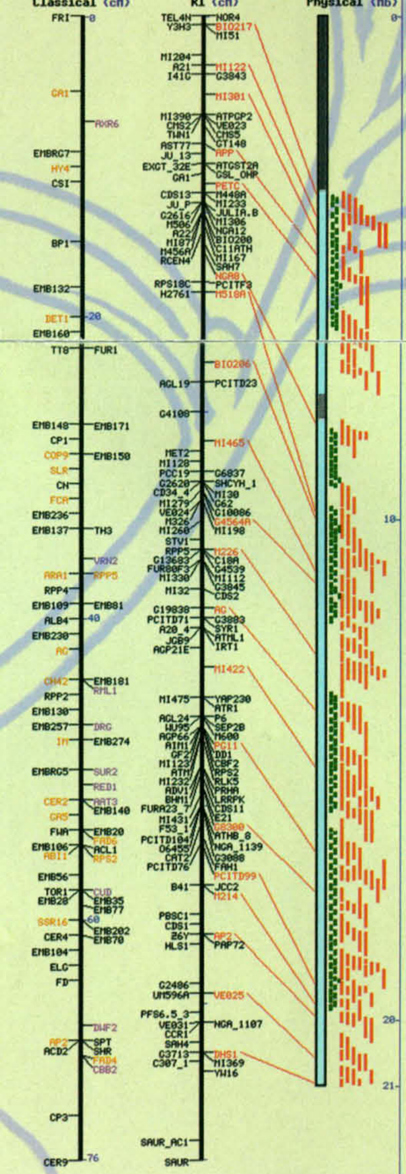


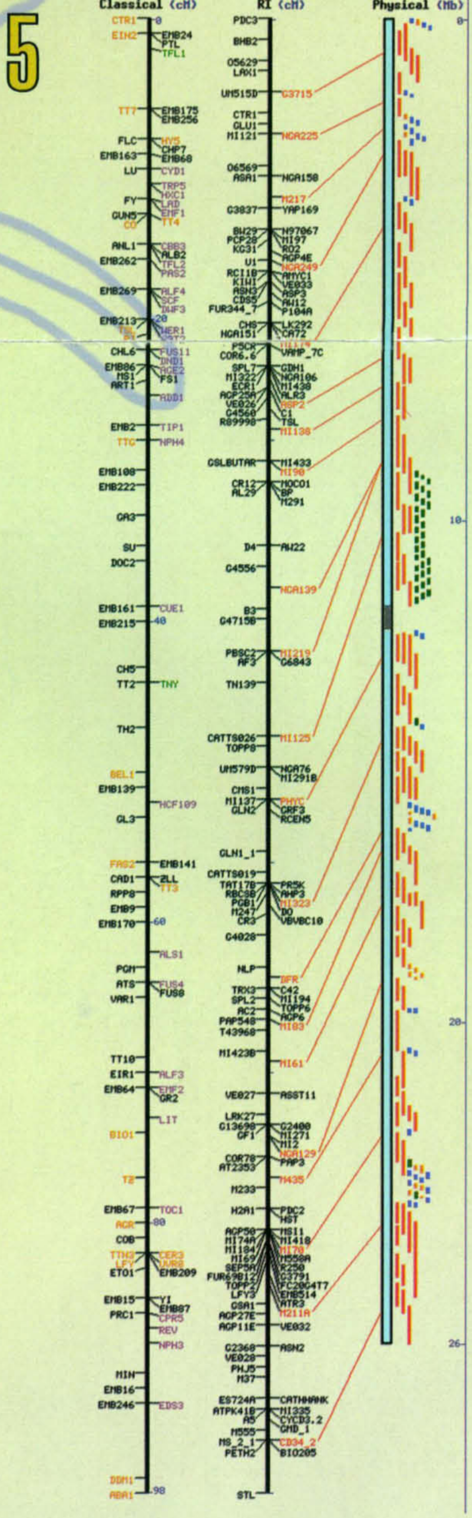




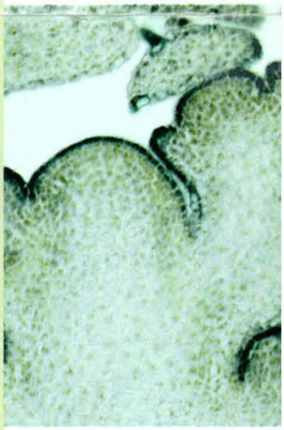

3
4
TAC
Classical (cM)
RI (cM)
Physical (Mb)
EMB91 0 RGA EMB34 STI SPR MUR2 EMB93 EMB62 EMB239 EMB58 EMB146 20 IAE1 EMB57 TAG1 CLF FVE STP1 PHYB SUR1 CHL2 CP2 POM2 40 RPT2 CLV3 ELF3 GUN1 ER HY1 MGL1 AXC1 MAX1 SE FUS12 RTH EMB152 SYD PY DWF6 LAN FIL LEM7 MAX2 SIN2 EMB53 FAD3 DAD COP1 60 PID ABI5 CPC RE AS1 DET2 GAL EMB18 CYT1 ABI4 AUX1 COI1 BIO2 EMB225 FPA EMB111 RRS2 EMB38 SPA1 SUL VAR2 CER8 EMB39 77
TEL2N NOR2 RGA KK1 ATPTR2_B VE012 MI320 ATGST2B F53_2 M246 G4553 M497A RHS1 IGS1 CDS3 VE013 EKRII_C MI310 MI444 RCEN2 TAT17C MI421 G4532 SEP2A G4133 PR1 MI398 LK141 BRM THY_1 CPK6 C425 211A B33 M216 PHYB GUT15 MI139 MI148 AGP2E FUR30_2 TAT17A E1 PCC300 PGC1 M251 MI238 0802F A9031HR LAX2 G6842 GPA1 ER B68 HY1 C2SOD2 PGC2 VE014 CDS9 M220 MI54 C4H M283C VE096 SEP5B ASP1 NGA361 COP1 ALR1 PR21 VE015 DSC1 SPL3 VE016 MI277 297 PPOD9 G17286 ATHB_14 F9 SAH3 M323 ATPCP1 UM579B VE017 SAH9 VE018 ATPSCS1 BH8 A45 LTP G4514 M429 NGA168 05841 TOPP4 C14382 CK_97 COR15 M336 LK3180 SG5 VE019 MI473 ATHB7 ATGST6 MI455 M179A PATT51 JULIA.E CLU2 RRS2
0 10 17
EMB145 0 ILR1 TIR3 EMB52 DOC1 MTO1 PDS2 EMB1 QRT2 EMB286 FAD7 RPM1 UVR7 ABI8 HY2 SPY EMB51 DCD1 COP10 FAD5 VCS FUS9 EMB153 HUA1 DHR1 VRN1 FAD2 20 EMB201 EMB29 EMB95 DWF1 MS2 PHO1 EMB90 EMB224 CHM AXR2 EMB125 SUP ALF5 EMB293 PRT1 ABI3 40 FUS3 PAD3 PAD5 GL1 RPT1 RPP1 UVR1 RPP13 TES VAL1 EMB65 60 NPH1 EMB167 RPP11 SCP ACD1 MUR1 SEL2 CSR1 ELL EMB208 PAS1 EMB69 EMB172 EMB149 FK TIR1 80 TT5 AP3 CH6 EMB117 TTN1 EMB154 FUS6 TRP3 CER7 TT6 96
PSM792 CGS CA1 ATHB3 MYB4 CDS15 MI74B A13 CDS6 NGA32 EST2351 MI199 NGA172 GAPC MI172 G4523 M583 CHS4 MI403 CR16 MI355 I7_X GMC1 MI467 MNSOD MI357 NGA126 JGB3 MPX1B ATTS0727 F53_3 MI207 ATHCHIB MS2 DS408_3A GHC1 NGA162 JU_10 GT51_1 MI289 G4708 M228 GLN1_3 MMR VE039 MSH2 ACP14E MI339 SAT2105B M560B2 YP105 M105 AHP1 04023 MI142 MI268 PAP606 ATDMC1 YPM255 MI386 MI225 MMR2 G4711 LRK12 PAP86 EST3201 MI287 MI178 VE020 UM579C GAPA JU_19 B52 AT14A GL1 AHP2 AIG2 BRC1 MI413 ATPOX 2IM2 MI358 RCEN3 MI79B ATA1 A_1 T04109 AT.LOX2A VE021 TOPP5 G4117 M249 ASN1 G4564B 097B1 I18 G4014 R30025 ATPK41A CDS4 MUR_1 B5 AST12 JULIA.D ATPSCS2 TAX ATEM1 F16B MI456 M457 R30 F3H AFC1 MS_3_1 PF3H0 F25_12.2 PBSC3 BGL1 SG1 SCR CF1 T94 VE042 AT103 I5_26 VE022 AP3_L ATTS3983 IMK3 SPL5 G2778 IMK2 GPII_1 SPL1 M424 NGA112 ACP29 NGA6 TSA1 TAI224
0 10 20 23
FRI 0 GA1 AXR6 EMBRG7 HY4 CSI BP1 EMB132 DET1 20 EMB160 TT8 FUR1 EMB148 EMB171 CP1 COP9 EMB150 SLR CN FCA EMB236 EMB137 TH3 VRN2 ARA1 RPP5 RPP4 EMB109 EMB81 ALB4 40 EMB230 AG CH42 EMB181 RPP2 RML1 EMB130 EMB257 DRG IM EMB274 EMBRG5 SUR2 RED1 CER2 AAT3 GA5 EMB140 FWA EMB20 EMB106 ACL1 ABI1 RPS2 EMB56 TOR1 CUD EMB28 EMB35 EMB77 SSR16 60 CER4 EMB202 EMB70 EMB104 ELG FD DWF2 AP2 SPT ACD2 SHR FAD4 CBB2 CP3 CER9 76


B
C
cop1
1 2 3 4 5 6 7 9 10 11 12 13 14 15 17 18 19 20 21 23 24 25 26 28 29 30 31 32 33 34 35 36 37 38 39 42 43 44





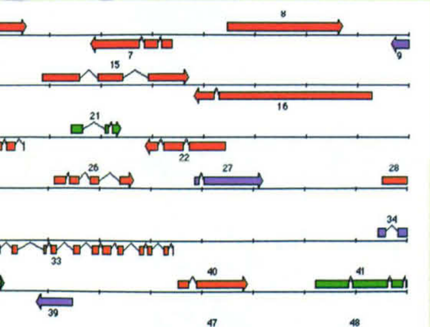
Dwarf plant with shortened internodes (*Science*, 10 Mar 1989).
(D. Patton, K. Feldmann)



In situ hybridization highlighting epidermis of shoot apex.
(A. Sessions, D. Weigel)

Mutant (red) and wild-type (green) embryos.
(L. Castle, D. Meinke)

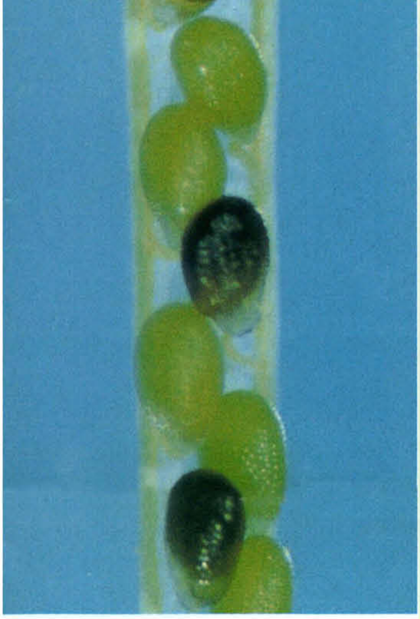
Mutant inflorescence (*Science*, 3 Jan 1997).

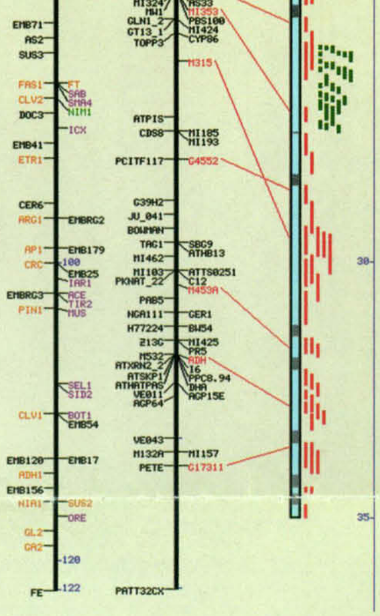


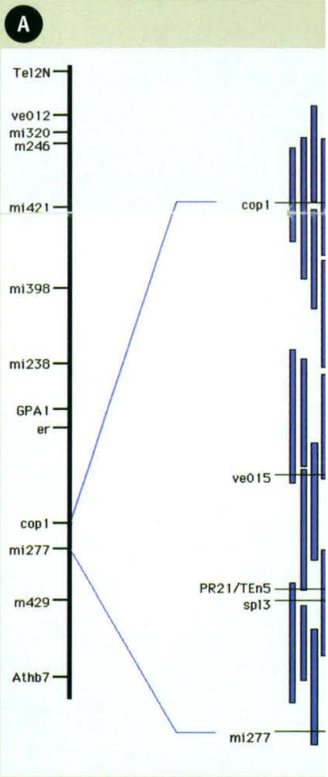
From chromosome to gene:

(A) RI map of chromosome 2, showing a subset of markers.

(B) A contig constructed from BAC clones for the cop1-mi277 interval. In red (solid and hatched) is a minimal set of BACs that can be ordered in an overlapping array to cover the region. Solid red indicates the region expanded in part (C).







D Gene Products With Si

3-Ketoacyl-CoA thiolase
30*S* ribosomal protein S5
Auxin-inducible protein, IAA 13
Auxin-repressed protein
Brassica napus hypothetical prote
Cadmium-induced protein
Calcium-binding EF-hand protein
Cellulose synthase
Chloroplast 50*S* ribosomal protei
Cinnamoyl-CoA reductase (2)
CONSTANS-like protein

CREDITS

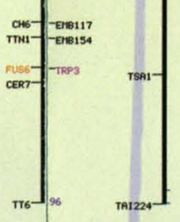
Science *Coordinators:*
Barbara R. Jasny and Pamela J. Hines

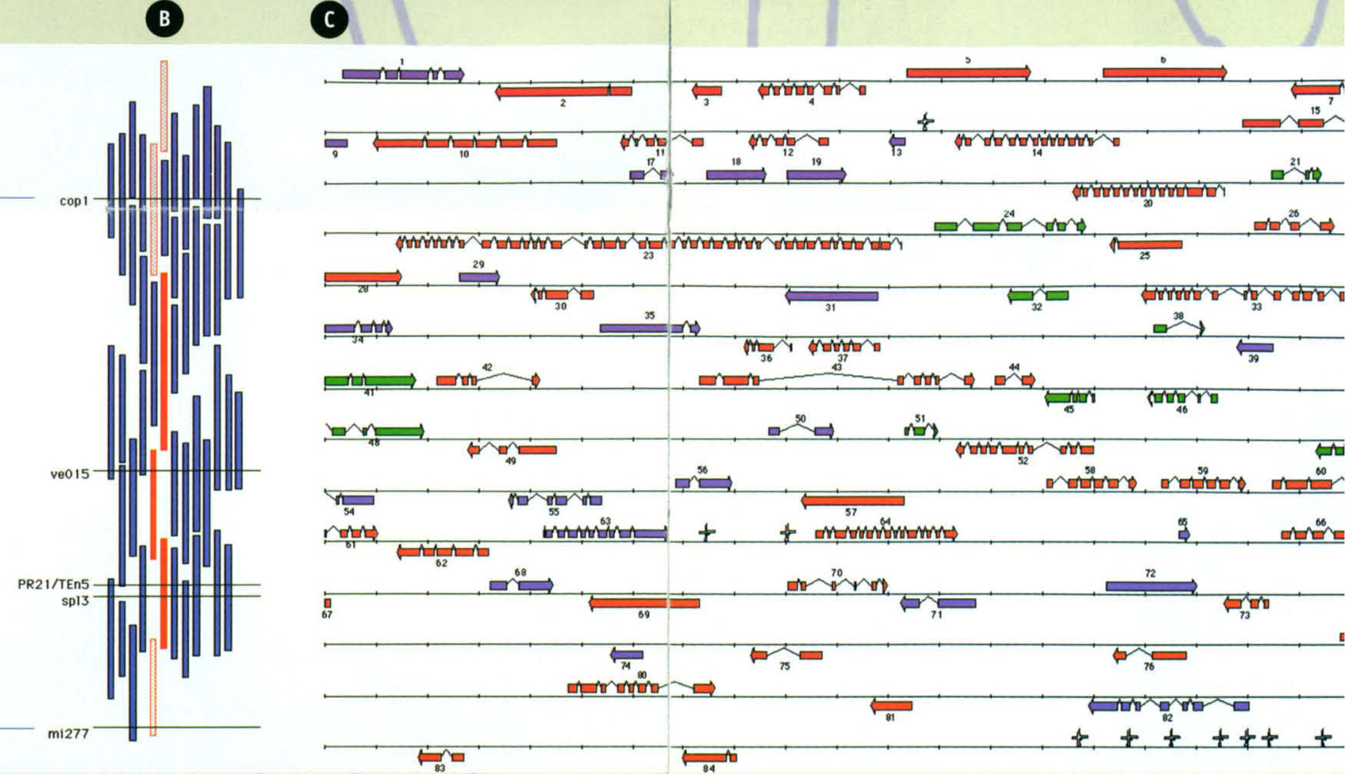
Authors:
Seung Yon Rhee, Shuai Weng, David Flanders, J. Michael Cherry, Stanford University, Stanford, CA, USA; Caroline Dean, Clare Lister, John Innes Centre, Norwich, UK; Mary Anderson, *Arabidopsis* Stock Centre, University of Nottingham, Nottingham, UK; Maarten Koornneef, Laboratory of Genetics, Wageningen, The Netherlands; David W. Meinke, Todd Nickle, Kelsey Smith, Oklahoma State University, Stillwater, OK, USA; Steven D. Rounsley, The Institute for Genomic Research, Rockville, MD, USA

Reviewers:
Elliot M. Meyerowitz, California Institute of Technology, Pasadena, CA, USA; Renate Schmidt, Max-Delbrück Laboratory, Cologne, Germany; Christopher R. Somerville, Carnegie Institution of Washington, Stanford, CA, USA

Art Direction:
Amy Decker Henry, Tracy Keaton Drew

Design and Production:
Tracy Keaton Drew





ucts With Similarity to Those Identified in Sequenced Region*

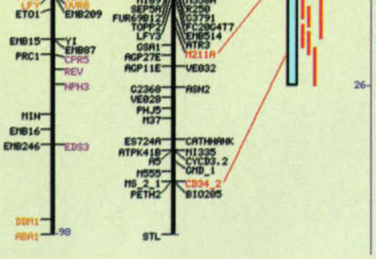
iolase
tein S5
rotein, IAA 13
rotein
othetical protein 2
protein
F-hand protein
osomal protein L28
ductase (2)
otein

CUC2 protein
CutA protein
Dimethylaniline monooxygenase
Disease resistance protein (5)
DNA-binding protein PD1
G9a protein
GTP-binding protein
Jasmonate-inducible protein
Mitochondrial carrier protein
Mitochondrial F1-ATPase, gamma subunit
Mlo protein

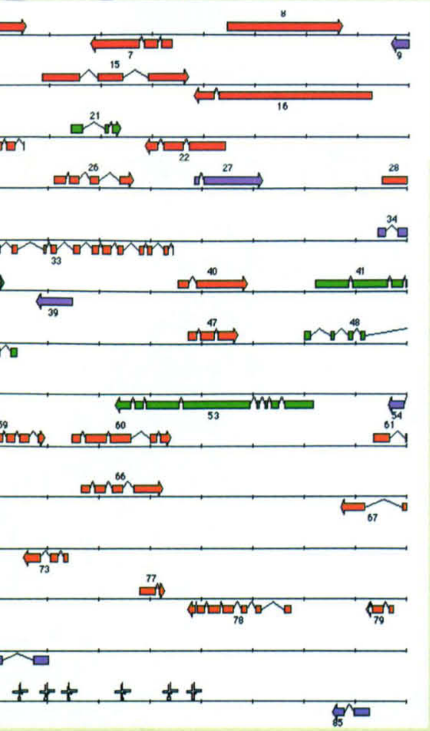
Myosin heavy chain
NAD(P)-dependent cholesterol dehydrogenase
Plastid protein
Polygalacturonase
Pre-mRNA splicing factor PRP 19
Proline-rich protein
Protein phosphatase 2C
Receptor kinase
Ribonucleoprotein
Ribosomal protein L17
RNA helicase

Serine carb
Serine prot
Splicing fac
Squamosa p
SWI/SNF co
Synaptobre
Thioredoxin
Tyrosine-spe
Tyrosyl-tRN
Ubiquitin-c

ociation for the Advancement of Science



Mutant (red) and wild-type (green) embryos.
(L. Castle, D. Meinke)



From chromosome to gene:

(A) RI map of chromosome 2, showing a subset of markers.

(B) A contig constructed from BAC clones for the cop1-mi277 interval. In red (solid and hatched) is a minimal set of BACs that can be ordered in an overlapping array to cover the region. Solid red indicates the region expanded in part (C).

(C) A 350-kb region that has been completely sequenced—protein coding genes are shown by solid boxes (exons) joined by thin lines (introns); the direction of transcription is shown by the arrows. tRNA genes are shown by ✛

Color coding of genes:
- 🟥 *Similar to known protein;*
- 🟩 *Unknown protein—contains a match to an EST sequence but shows no similarity to known proteins;*
- 🟪 *Hypothetical protein—identified by gene prediction software only.*

(D) List of gene products with similarity to those identified within the sequenced region.

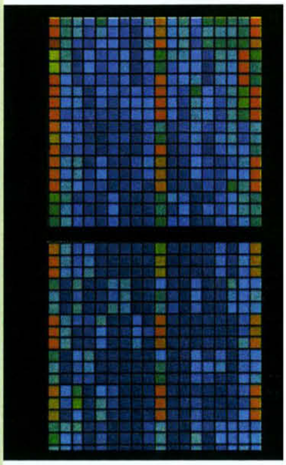
Serine carboxypeptidase
Serine proteinase (*stubble*)
Splicing factor U2AF large chain
Squamosa promoter-binding protein 1
SWI/SNF complex subunit BAF 170
Synaptobrevin (2)
Thioredoxin
Tyrosine-specific transport protein (TyrP)
Tyrosyl-tRNA synthetase
Ubiquitin-conjugating enzyme, E2-16kD

tRNA-Asp (2)
tRNA-Met
tRNA-Pro (9)

Hypothetical proteins (24)
Unknown proteins (10)

* *Numbers in parentheses indicate multiple forms of the gene.*

Mutant inflorescence (*Science*, 3 Jan 1997).
(D. Bradley et al.)



DNA arrays illustrating changing gene expression patterns.
(I. Brandstatter, D. Hearn, G. Karlin-Neumann)

omato BAC libraries • Human and Yeast GENEFILTERS™
Radiation Hybrid Mapping Panels • Colony Picking and Gridding Services
• Custom DNA • Custom Peptides and Antibodies
16