

Lo studio della funzione dei geni

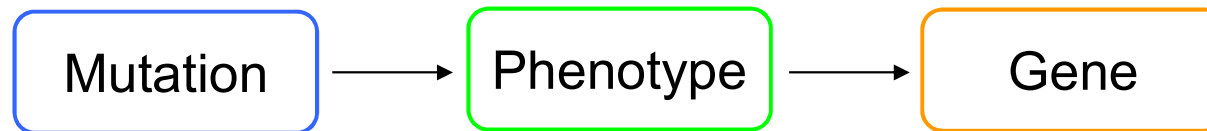
Forward and Reverse Genetics

- Forward genetics starts with identification of interesting mutants
 - Then aims to discover the function of genes defective in mutants
- Reverse genetics starts with a known gene and alters its function
 - Then aims to determine the role of the gene from the effects on the organism

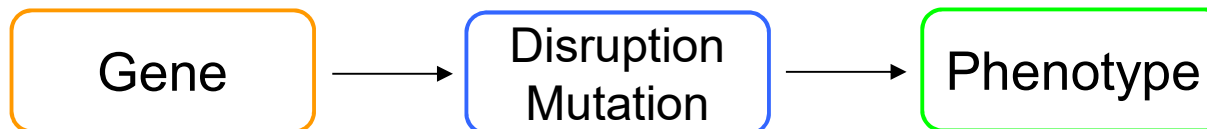


Forward and reverse genetics

- Forward genetics
: from phenotype to gene structure



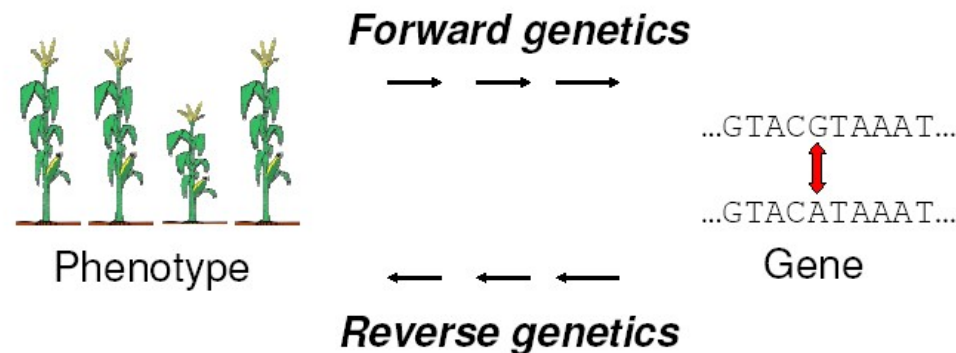
- Reverse genetics
: from gene to phenotype



Forward and Reverse Genetics

- Forward genetics starts with identification of interesting mutants
 - Then aims to discover the function of genes defective in mutants
- Reverse genetics starts with a known gene and alters its function
 - Then aims to determine the role of the gene from the effects on the organism

Forward vs Reverse



Basics of Forward Genetics

- Forward genetics usually starts with *mutagenesis* of an organism
 - Can use chemicals
 - e.g., ethyl methyl sulfonate (EMS)
 - often used in Arabidopsis
 - Can use radiation
 - e.g., X rays
 - Can use gene tagging
 - e.g., transposon tagging or T-DNA tagging
 - often used in Arabidopsis and maize
- Then *screen progeny* of mutagenized individuals for phenotypes of interest



Forward genetic

Select a biological process



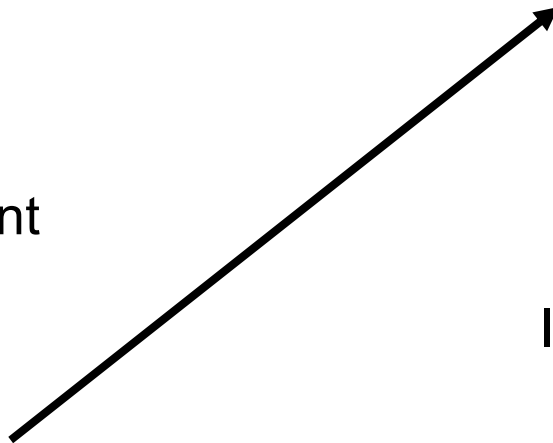
Generate a mutant population



Screen a large number of mutagenised M2 plants



Identify the mutated gene





Forward genetic

Generate or obtain a mutant population



Mutant: An organism that differs from the “normal” or wild type by one or more changes in its DNA sequence

Two major considerations in planning a mutant screen:

- How many plants can be reasonably screened?
- How many genes can mutate to cause the phenotype of interest?



Forward genetic

- **Chemicals agents:**

e.g., ethyl methyl sulfonate (EMS) or nitrosomethylurea (NMU)

- **Physical agents:**

e.g., X-rays, fast neutrons and accelerated ions

- **Biological agents:**

e.g., transposons and T-DNA

Point mutations

Large range of mutations including loss and gain of function

Very high efficiency with hundreds of mutations per genome

Difficult to find the mutation in the genome

Big insertions, deletions and rearrangements

Mostly loss-of-function mutants

Medium efficiency

Laborious identification of the mutations

Insertion of specific DNA sequences

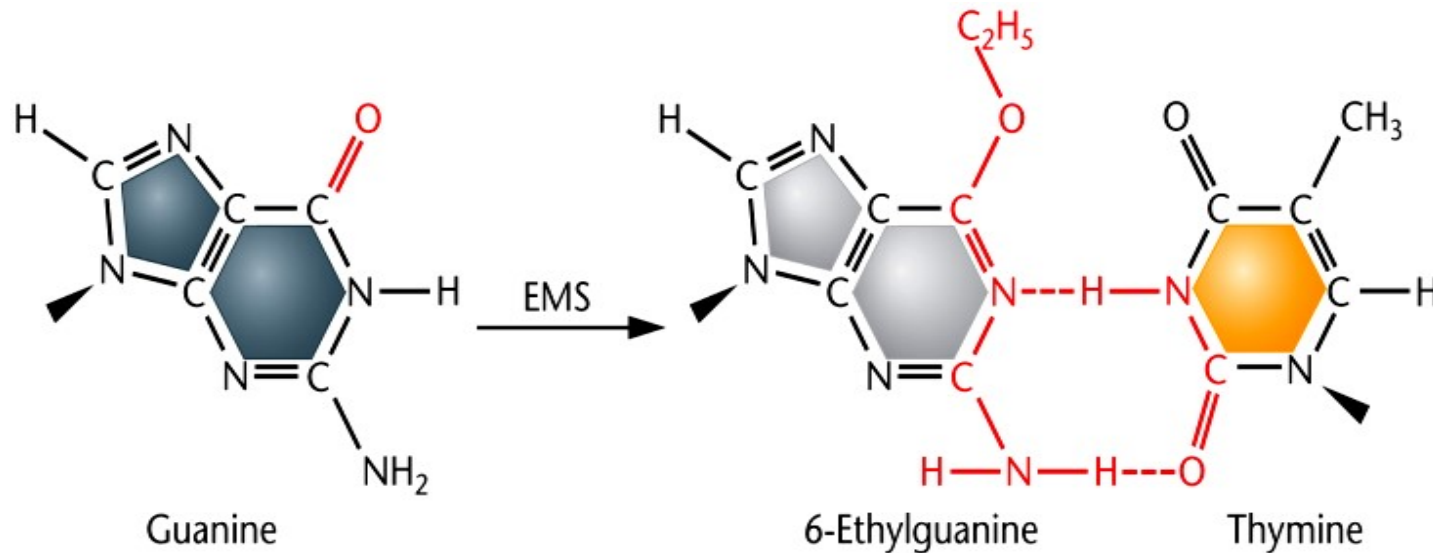
Mostly loss-of-function mutants

Low efficiency

Easy identification of the mutations



Chemical (alkylating) agents



- EMS causes addition of chemical groups primarily to Gs.
- The extra groups can cause altered base pairing resulting in transitions.
- EMS can also cause depurination (complete loss of the base) which leaves gaps that cause misincorporation during replication.

Protocollo di mutagenesi EMS

Weight 0.2 gram seeds (10,000 seeds total) (20 microgram/seed).

Wash in 0.1% Tween 15'

Put into 15 ml ddH₂O

Add 15-45 microliter (0.1% to 0.3%) EMS

Mix and incubate for 8-12 hrs (Rotating) in hood.

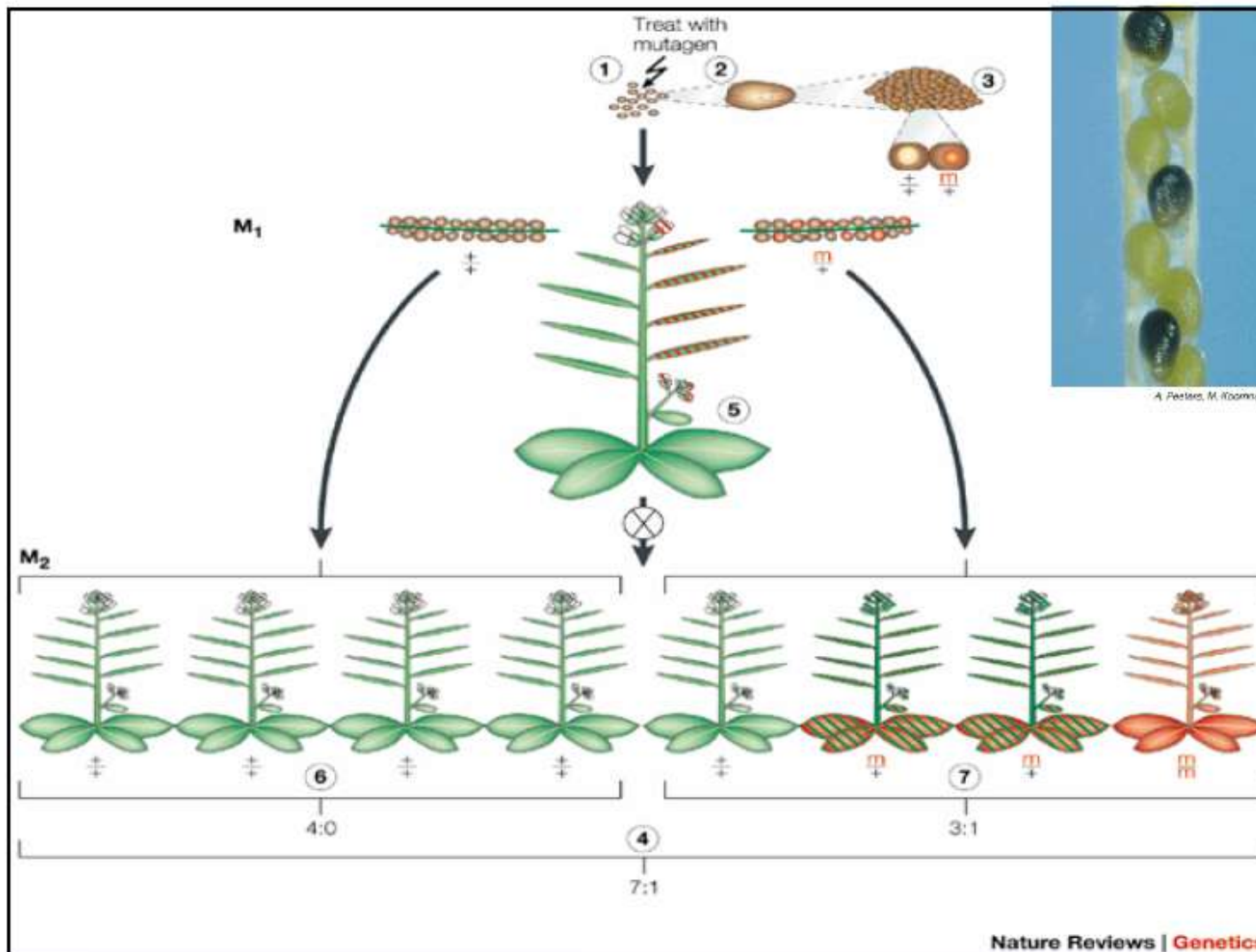
Remove the EMS (put EMS sol'n into 0.5M NaOH O/N, then dispose as regular waste)

Rinse the seeds once and then rinse them in 10 ml ddH₂O 2-4 hrs.

Put all seeds into 100 ml 0.1% Agar (Autoclave agar, cool on ice, and let solidify for several hours first)

Sow the seeds using a P1000 pipette (1 ml per pot).

Vengono introdotte mutazione puntiformi con transizioni
GC-AT e AT-GC



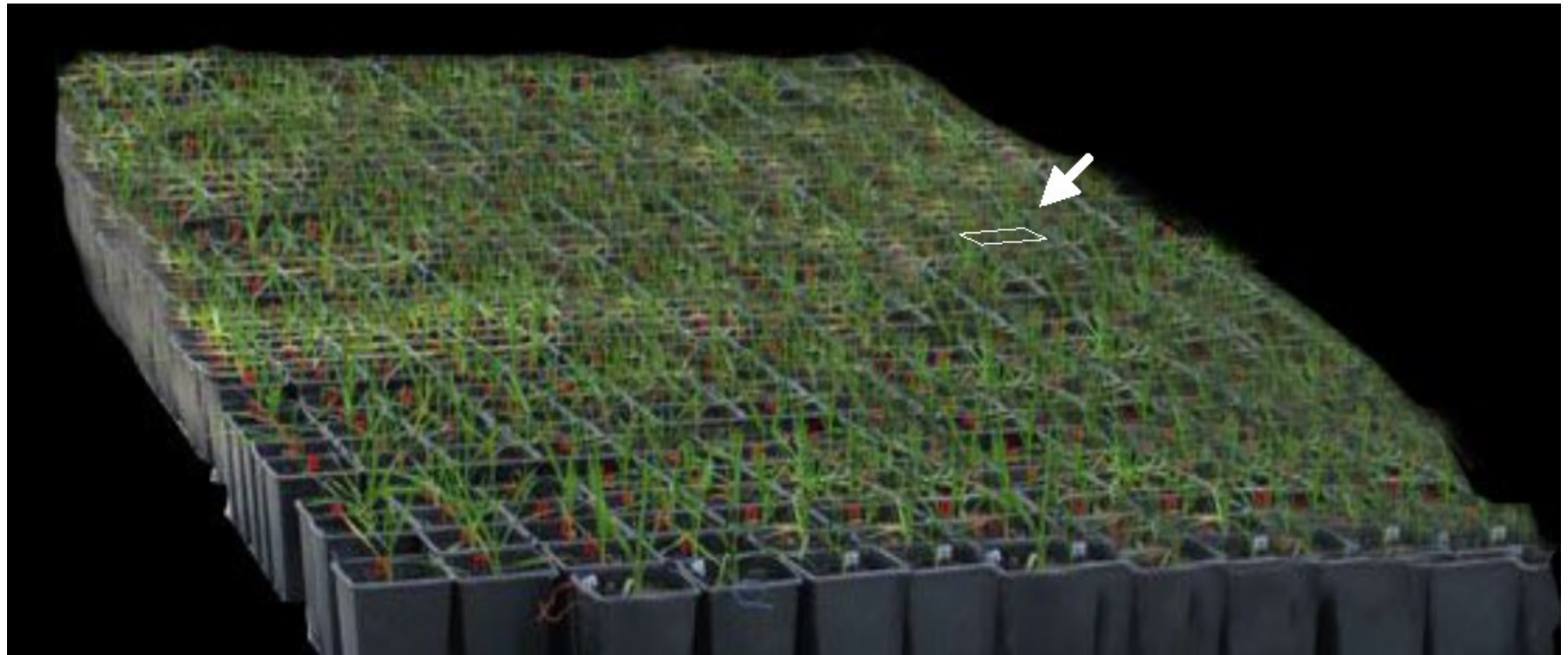
The targets of seed mutagenesis are the diploid cells of the fully developed embryo covered by the seed coat. To assess the effectiveness of mutagenesis, it is crucial to know how many of the targeted cells will eventually contribute to the next generation. The functional germ line can be defined as the number of cells in the SAM of the embryo that will contribute to the seed output (the genetically effective cell number (GECN)). In *Arabidopsis*, recessive mutants segregate in a ratio of 7:1 in an M_2 population, therefore, $GECN = 2$. Because the functional germ line consists of two cells at the time of mutagenesis, the developing M_1 is chimeric and consists of two sectors, which might vary in size. Mutations segregate 4:0 or 3:1, depending on the sector from which the M_2 seeds derive.



Ionizing radiation

- Radiation ionize water and result in the formation of oxygen radicals such as superoxide ($O_2^{\cdot-}$) and hydroxyl radicals ($\cdot OH$),
- Radicals are highly reactive and can cause single and double strand breaks and base modifications,
- Single strand breaks are easily repaired but double strand breaks often result in deletions, insertions and translocations.

Screening...



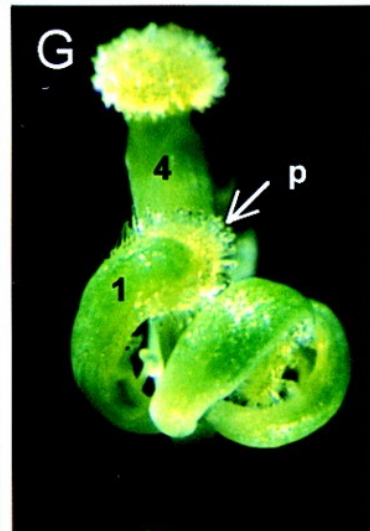
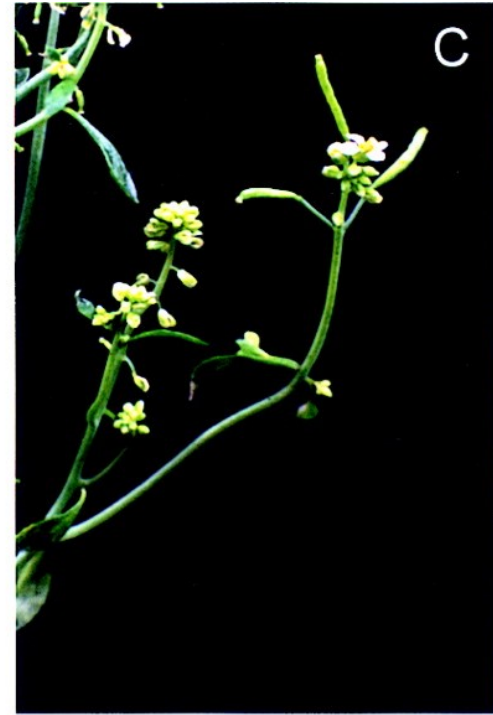
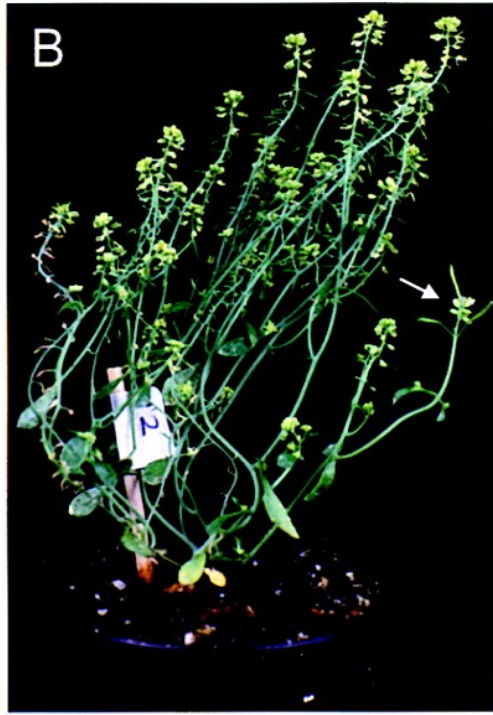


Forward genetic

You need a screenable phenotype



- Flower development
- Disease resistance
- Photomorphogenesis
- Responses to hormones etc.

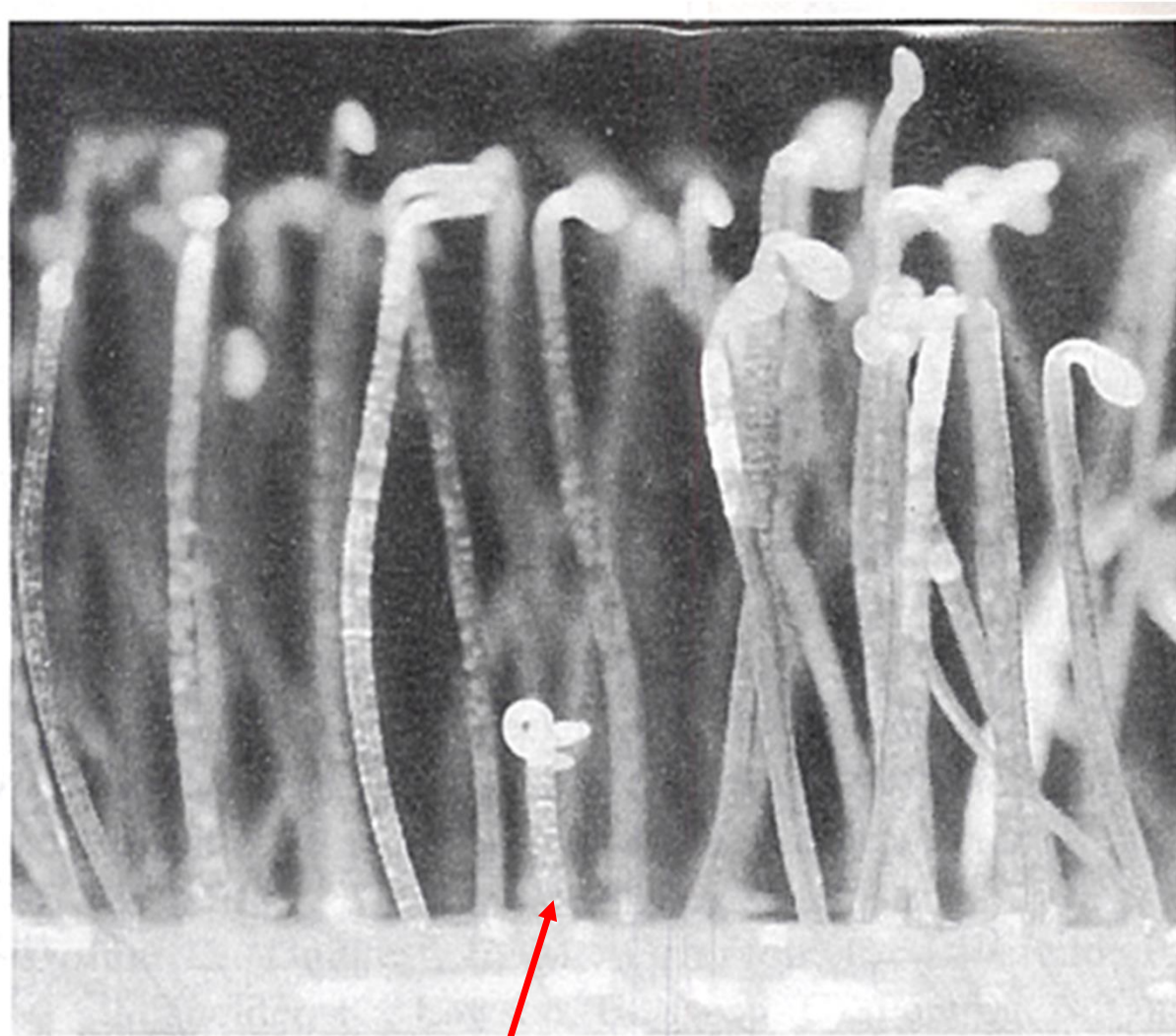


Recettori dell'etilene sono stati clonati per
isolamento di mutanti etilene -insensibili

(A)



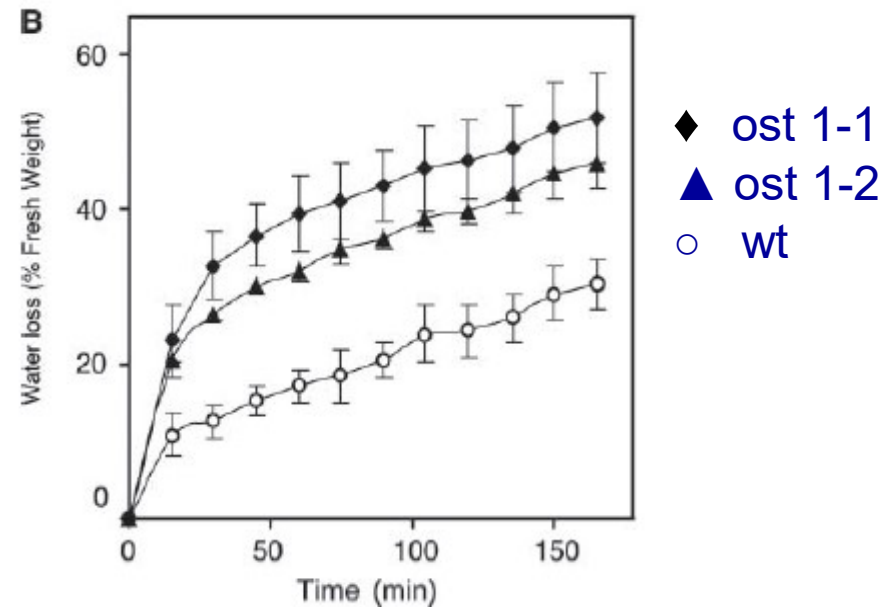
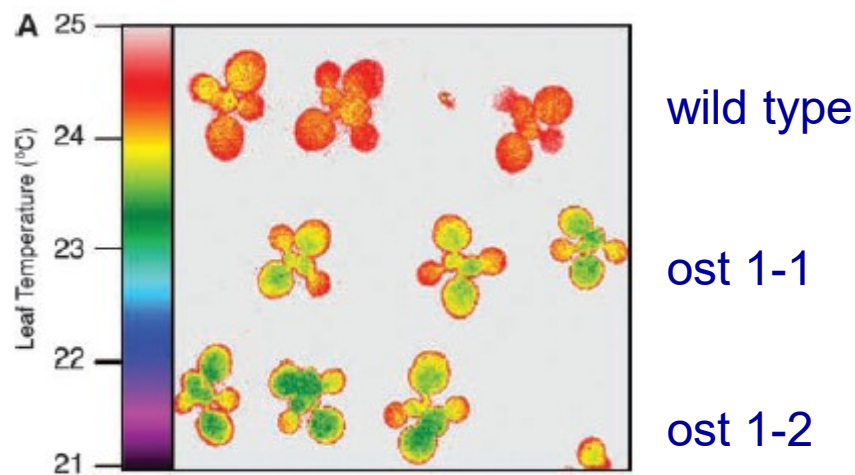
Isolamento di mutanti che presentano tripla risposta costitutiva (anche in assenza di etilene) ha permesso di identificare un regolatore negativo nella via di trasduzione



ctr1: constitutive triple response

Un sistema per isolare mutanti (*ost1*)

- Popolazione mutagenizzata con etil metansulfonato
- Stress idrico
- Analisi mediante termografia infrarossa
- Selezione dei mutanti



I mutanti *ost1* sono difettivi nella regolazione della traspirazione sotto stress idrico

Drawbacks of Forward Genetics

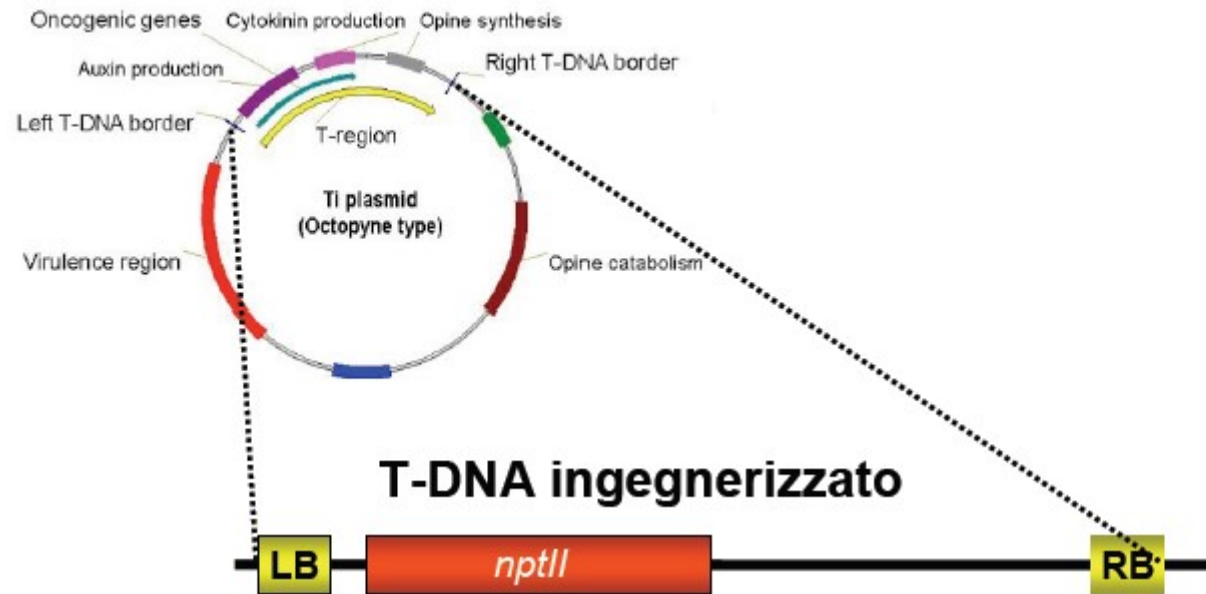
- Breeding must be possible
- You need a screenable phenotype
 - Easy with visible phenotypes
 - growth, color, development
 - not easy with metabolism
 - impossible for „essential genes“ (defects of which would be lethal)
- You need a dense genetic map and markers closely associated with the phenotype
 - Good for Arabidopsis, Maize, Rice
 - Not established for many tropical plants, hops (*Humulus lupulus*) and other speciality crops
- Access to genomic library (YAC, BAC) required for cloning the gene

Il T-DNA di *Agrobacterium* è stato utilizzato per generare collezioni di mutanti inserzionali

- Il T-DNA **si integra in maniera casuale** all'interno del genoma della cellula vegetale...
.. permettendo la generazione di knock-outs



- **L'integrazione è stabile** per un buon numero di generazioni
- Inserzioni singole o più inserzioni (testa testa, testa coda o coda-coda)



Il “T-DNA” viene ingegnerizzato con:

Un marcatore per la selezione delle linee T/F (Kan)

Mutagenesis by insertion tagging

Principle: A DNA fragment with a known sequence (transposons, *A.tumefaciens* T-DNA) is allowed to randomly insert into the genome (it usually causes a recessive, loss of function mutation)

Advantages:

- tags or marks the gene
- provides a powerful way to identify or fish the gene out

Disadvantages (like chemical and physical mutagenesis):

- cannot knock out essential genes,
- other redundant genes mask loss of disrupted gene
- may disrupt non-functional sub-region of gene.

Insertion tagging

[signature tagged mutagenesis (STM)]

-The advantage of insertion tagging over traditional forms of mutagenesis is that the interrupted gene becomes "tagged" with the insertion elements, hence the strategy is sometimes termed Signature-tagged mutagenesis (STM) .

Steps:

1. T-DNA or transposon tagging.
2. Techniques can be used to obtain the flanking DNA.
3. The sequence of the flanking DNA can then be used to interrogate sequence databases, allowing the tagged gene to be associated with its "parent" genomic clone phenotype.
4. If insertion also generates a mutant phenotype, the gene in the database can then be ascribed a tentative function.

Mutagenesis by T-DNA tagging

- Gene insertion is essentially random.
- It affects only the gene, where it incorporated (within the gene space).
- Disadvantage is the tendency to generate complex, multicopy integration patterns and sometimes deletions and gene rearrangements of surrounding genes.
- As *Arabidopsis* is a gene dense plant- having small introns, little intergenic space and therefore, about 80 % genome is thought to represent genome.
- - Insertions in around 90% of genes are present.
 - Over 300,000 transformants were made by T-DNA tagging
- As T-DNA is not a transposon, it has no ability to 'jump' following integration.....therefore, having advantage of generating stable insertions

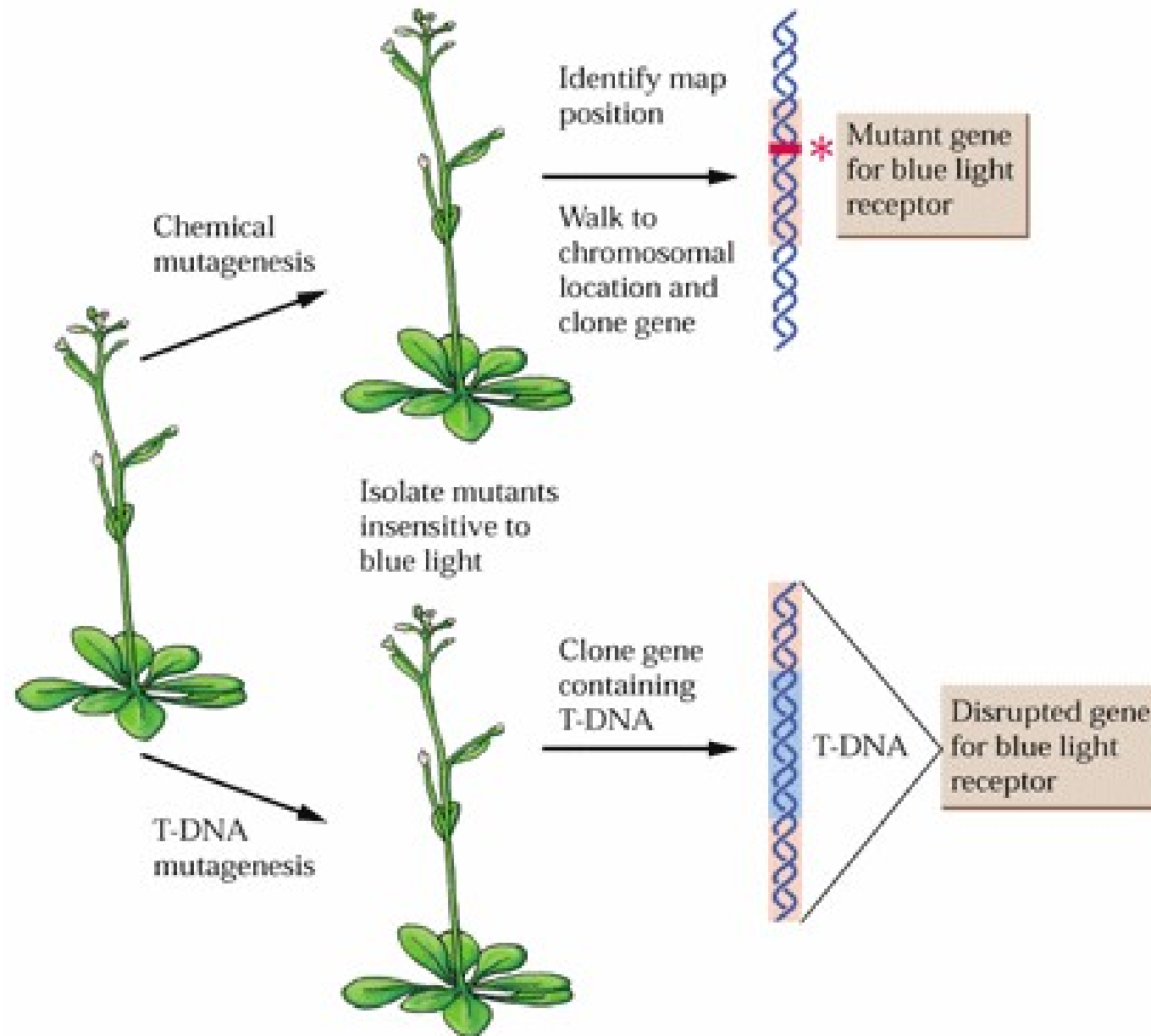
Flanking sequences **have been** characterized. Mutants can be used for reverse genetics (see ahead)







- DNA isolated and the junction fragment between the T-DNA and the plant DNA recovered. This allows the precise location of the T-DNA in the genome to be assessed. Called Flanking Sequence Tag (FST).
- These FSTs are present in databases, so insertions in a gene of interest can be recovered by searching the database.
- Mutant lines can be ordered from catalogue.

T-DNA insertion lines

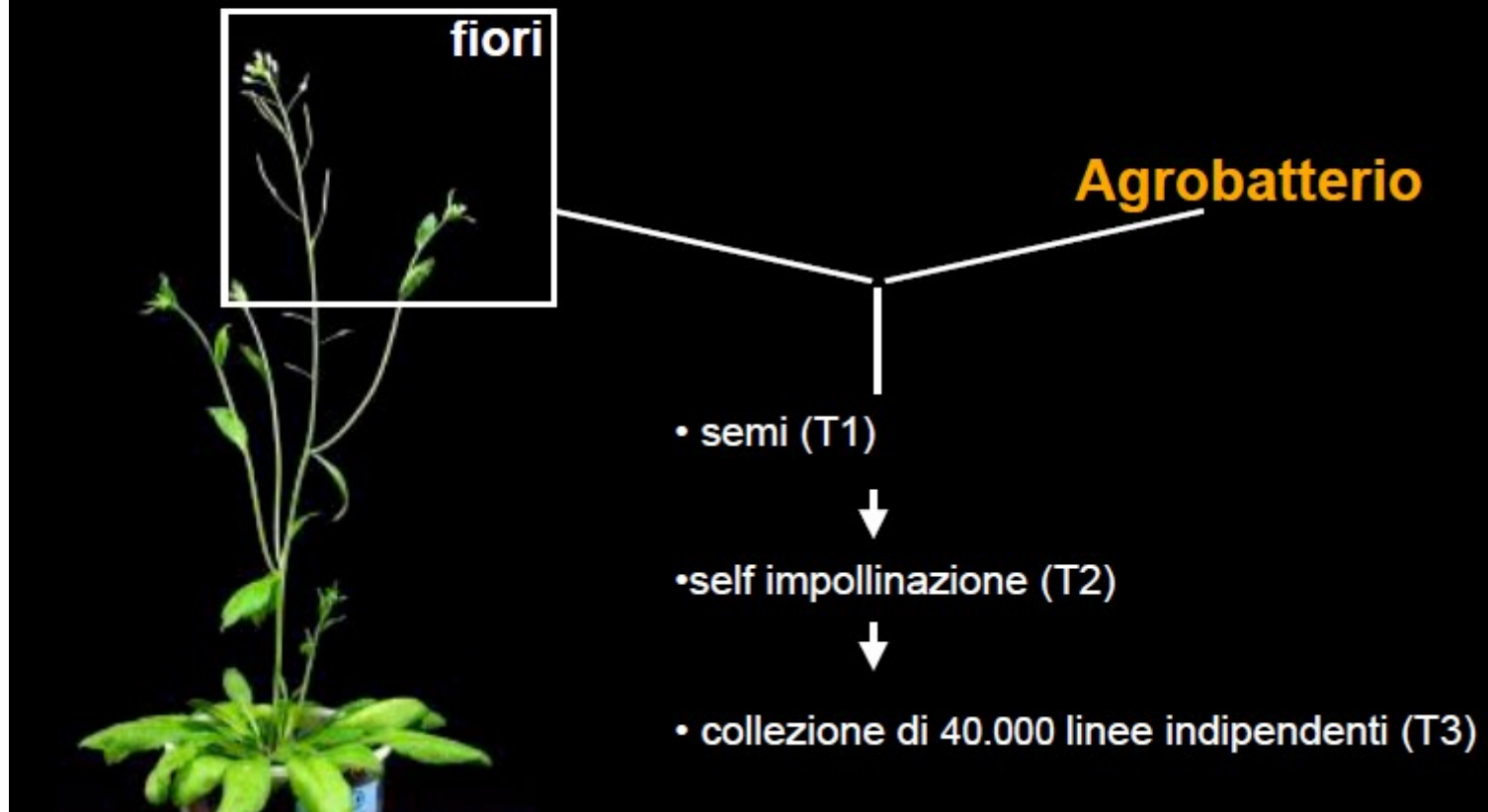
- Average number of independent insert 1.5 per diploid genome
- 57% of the transformants containing single insert
- 25% of the transformants containing double insert
- Bias observed against integration events in exons and introns in favour of promoters, 5' and 3' UTRs
- Number of integrations decrease from chromosome arms towards centromeres

(A) Chemical mutagenesis vs. T-DNA mutagenesis



<i>Name</i>	<i>Location</i>	<i>Result</i>
 <i>Knock-out</i>	coding region or promoter	null
 <i>Knock-down</i>	promoter or 3' UTR	reduced expression
 <i>Knock-on</i>	promoter	increased expression
 <i>Knock-about</i>	coding region	not null
 <i>Knock-knock</i>	coding region or promoter	multiple KOs in one plant
 <i>Knock-worst</i>	coding region or promoter	chromosomal rearrangement

Cellule di **Agrobacterio** contenenti il **T-DNA** ingegnerizzato vengono utilizzate per generare una popolazione di piante transgeniche (floral dip)



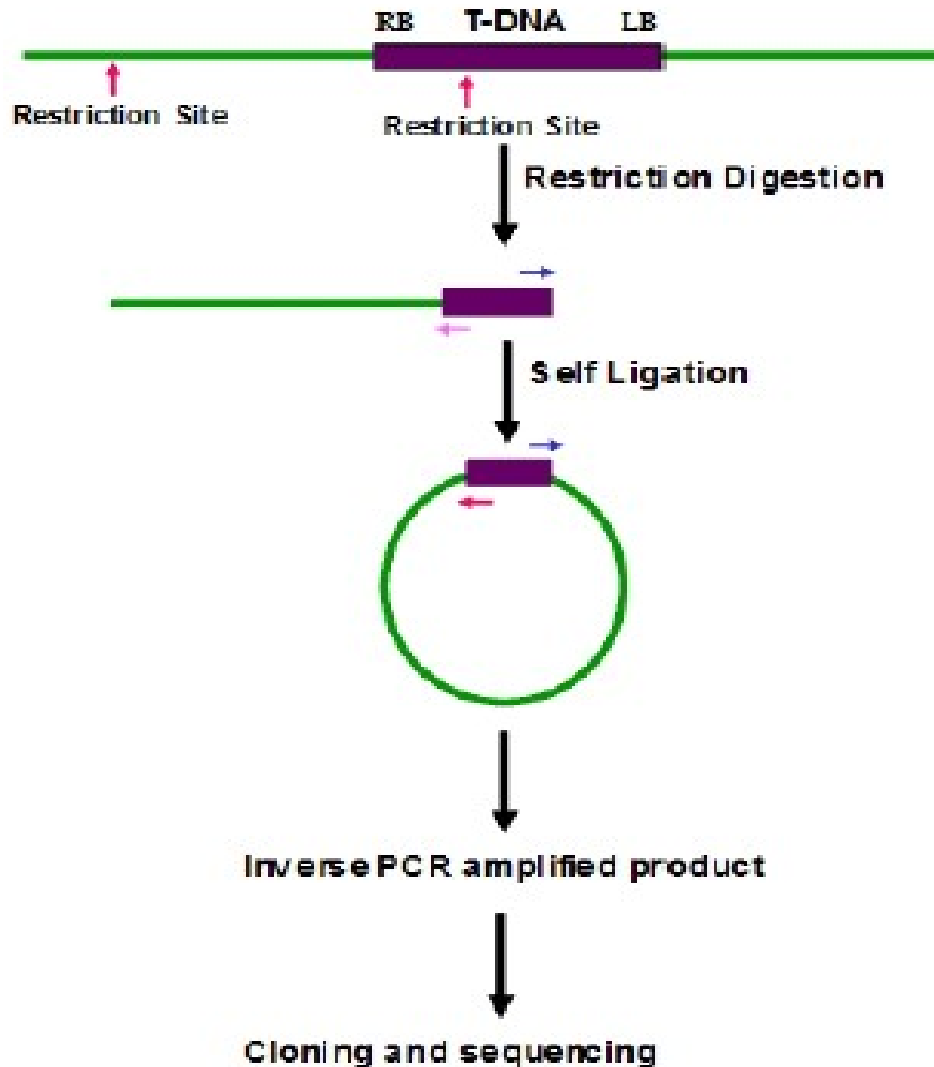


Identification of sequences flanking T-DNA (= mutated gene)

- Inverse PCR
- Thermal asymmetric interlaced (TAIL) PCR
- Plasmid rescue



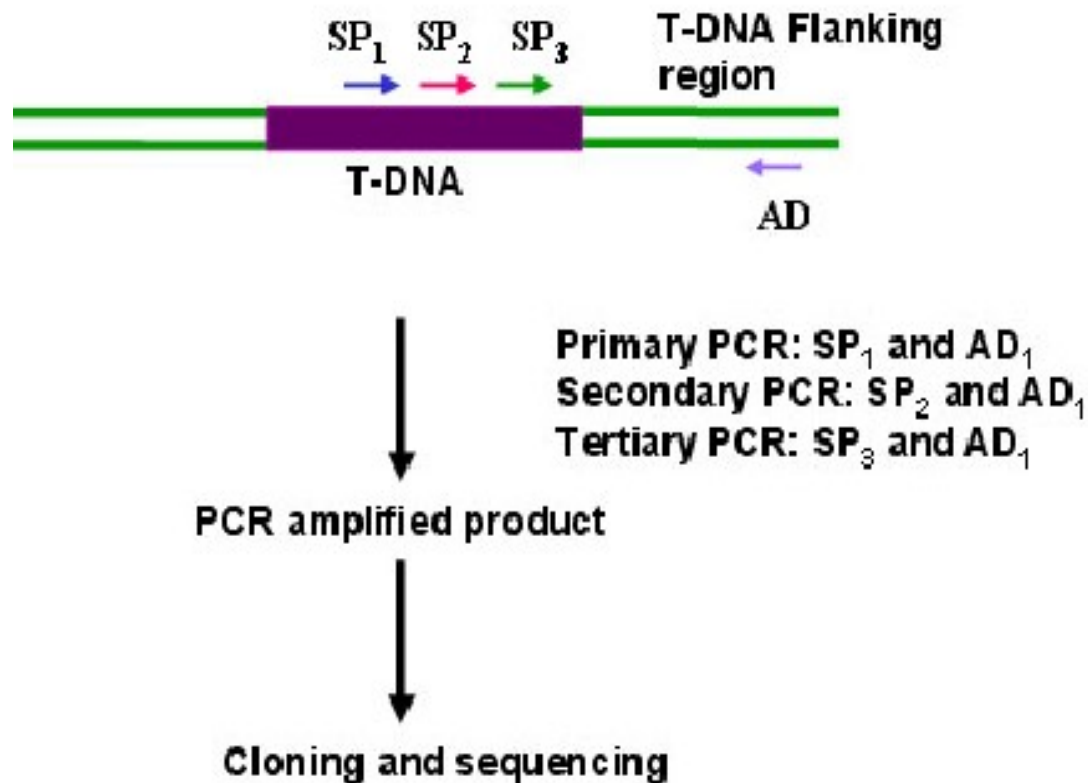
Inverse PCR



- 1) Restriction digestion of genomic DNA from mutant plant (preferably single cut within the T-DNA);
- 2) self-ligation;
- 3) circularized ligation products are used for PCR amplification using appropriate primers from the T-DNA region.



Thermal asymmetric interlaced (TAIL) PCR



Three consecutive rounds of PCR, performed with a set of three nested T-DNA specific primers (SP₁, SP₂ and SP₃) and a small, arbitrary primer.

AD is arbitrary degenerate primer indicated by purple arrow.

Arbitrary or degenerate primer

1.3.3. Primer degenerati

I primer degenerati sono primer la cui sequenza non è determinata univocamente, ma contiene una o più posizioni in cui possono essere presenti più nucleotidi in miscela. Ad esempio, la sequenza sottostante di un primer di 20 basi:

5'-ACGTATNTCCNCAAYATYGCT-3'

significa che il primer disegnato contiene delle posizioni ("degenerazioni" indicate con Y e N) che contengono in miscela le basi T + C (Y) o A + G + C + T (N). Le degenerazioni ovviamente riducono la specificità del primer aumentando le possibilità di appaiamento.

A cosa servono i primer degenerati?

I primer degenerati possono essere utilizzati per amplificare sequenze di DNA (ignote) da un organismo utilizzando per il disegno del primer la sequenza nota (omologa) proveniente da un altro organismo o dallo stesso organismo. In alcuni casi si possono usare i primer degenerati

Primer Design for TAIL-PCR

- **Specific primer (SP)**
 - Nested sequence specific primer complementary to vector sequence
 - High melting temperature, $T_m = 58-63^\circ\text{C}$
- **Arbitrary degenerate (AD) primer**
 - Relatively shorter
 - Lower melting temperature, $T_m = 47-48^\circ\text{C}$

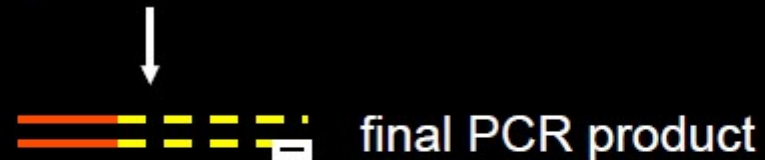
TAIL-PCR scheme



(A) Primary PCR with SP1 and AD

(B) Secondary PCR with SP2 and AD

(C) Tertiary PCR with SP3 and AD

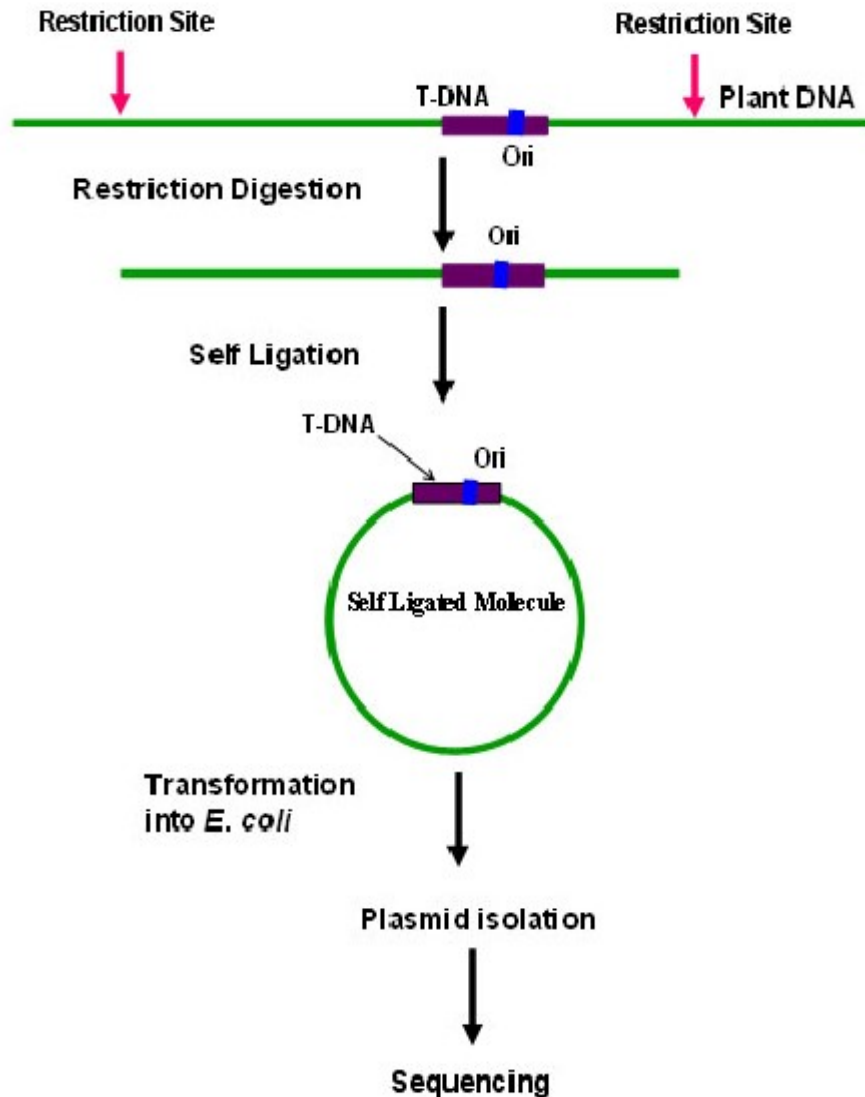


Agarose gel analysis

Direct sequencing

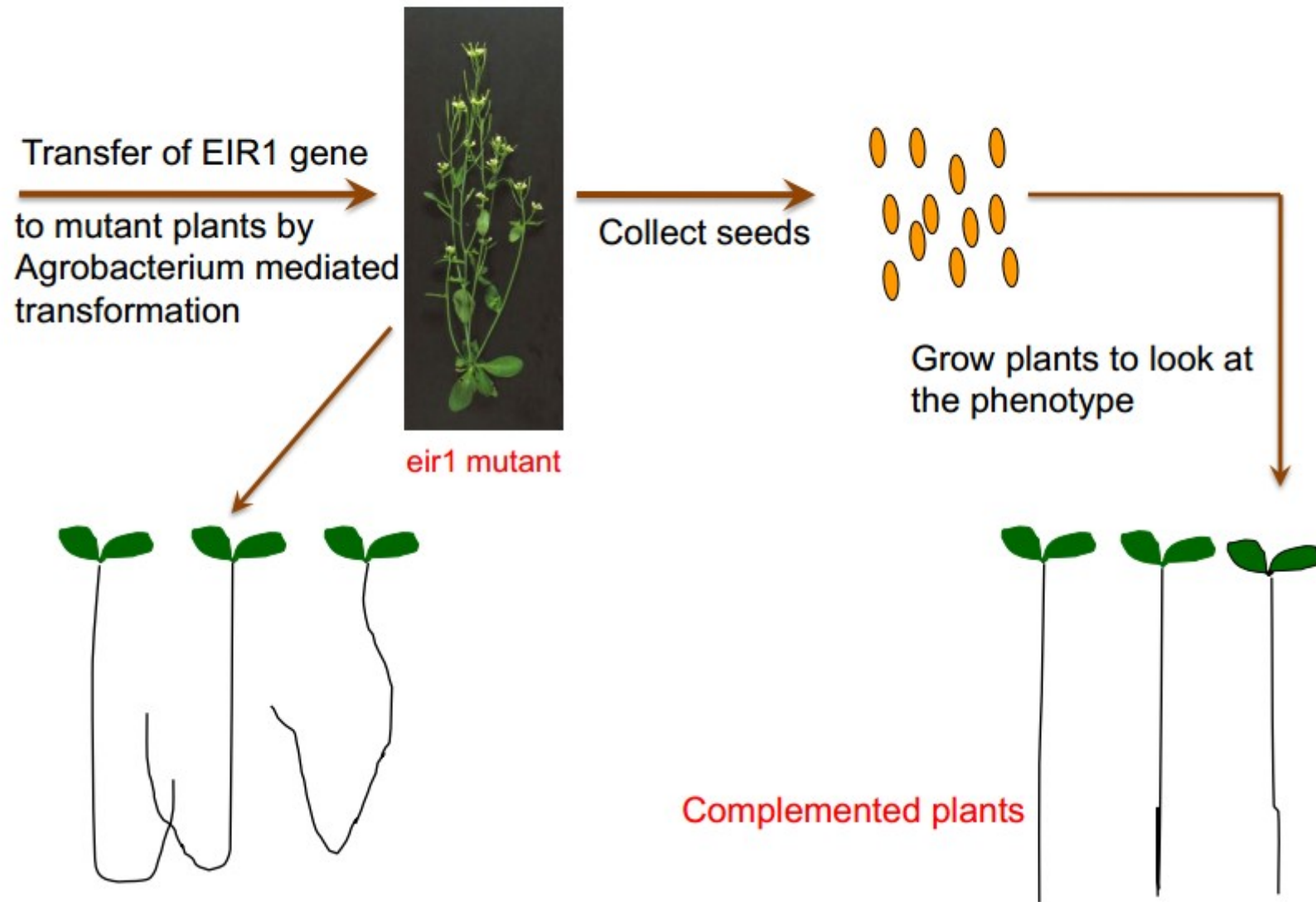


Plasmid rescue



- 1) Restriction digestion of genomic DNA from mutant plant (preferably with an enzyme which does not cut within the T-DNA);
- 2) self ligation;
- 3) transformation of *E. coli* with the self-ligated fragments

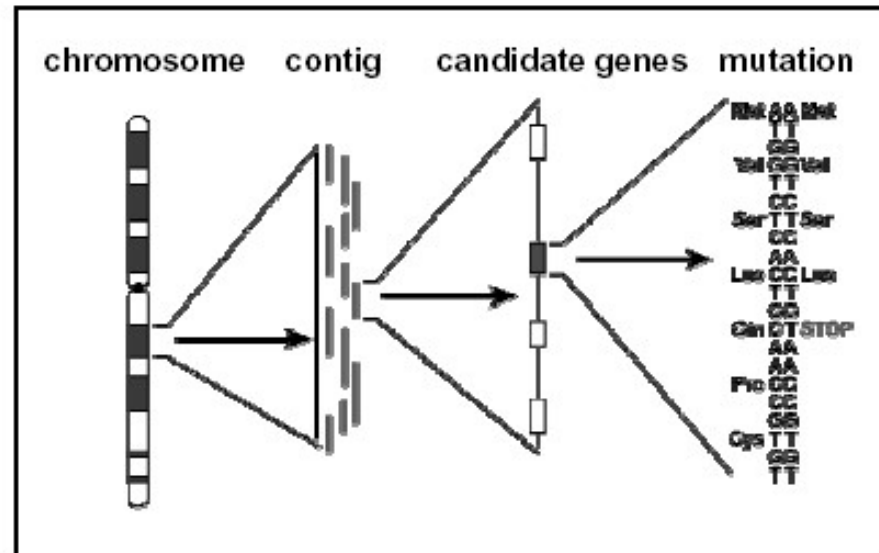
Final step: to validate the gene function we need to restore the mutant phenotype



From Phenotype to Gene

In case of chemical or physical mutagenesis

- Once an interesting mutant is found and characterized, we want to find the gene in which the mutation occurred
- Positional cloning
 - First use genetic mapping
 - Then use chromosome walking

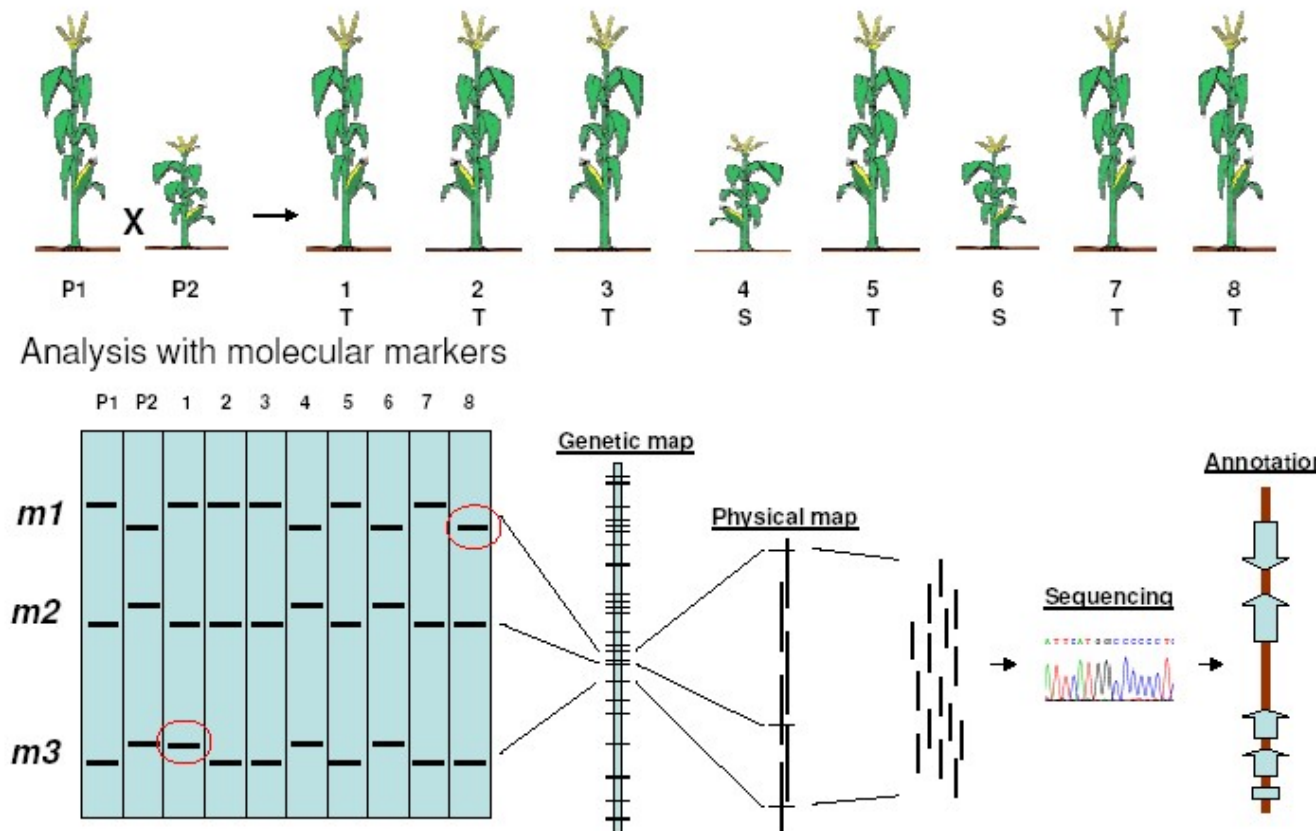


But now deep sequencing (it is becoming less and less expensive)

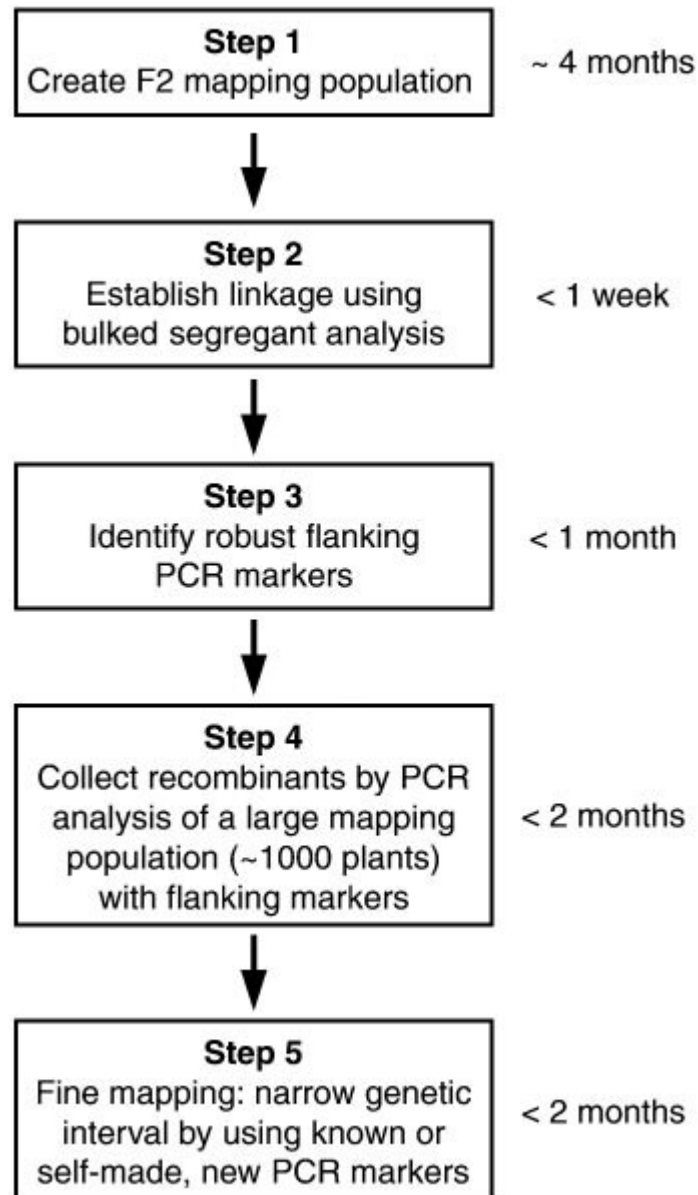
To isolate point mutations, deletions and translocation

Mapping will narrow down the genetic interval containing a mutation by successively excluding all other parts of the genome. Map positions are determined by measuring the recombination frequencies between the mutation and markers that lie in known position in the genome

Positional cloning (forward genetics)



MAP-BASED CLONING



Mapping with a high resolution requires a high density of genetic markers

- Several *Arabidopsis* accessions, or ecotypes, are sufficiently divergent to support the design of molecular markers at this high density.
- The most commonly used combination for mapping purposes is Landsberg *erecta* X Columbia (Ler X Col).
- These two accessions have been estimated to differ in 4 to 11 positions every 1,000 bp (Chang et al., 1988; Konieczny and Ausubel, 1993; Hardtke et al., 1996).
- Most existing mutations, including mutations causing visible phenotypes that can be employed as genetic markers have been induced either in a Col or a Ler background.

Molecular markers commonly used for mapping (SSLP, CAPS and dCAPS)

- They are codominant (both chromosomes of a plant may be genotyped)
- They are PCR-based and can be analyzed on agarose gels (easy to use and inexpensive)
- Several methods for the extraction of plant DNA from small tissue samples have been described
- These methods are cheap and suitable for high throughput applications in microtiter plates.

Molecular markers commonly used for mapping

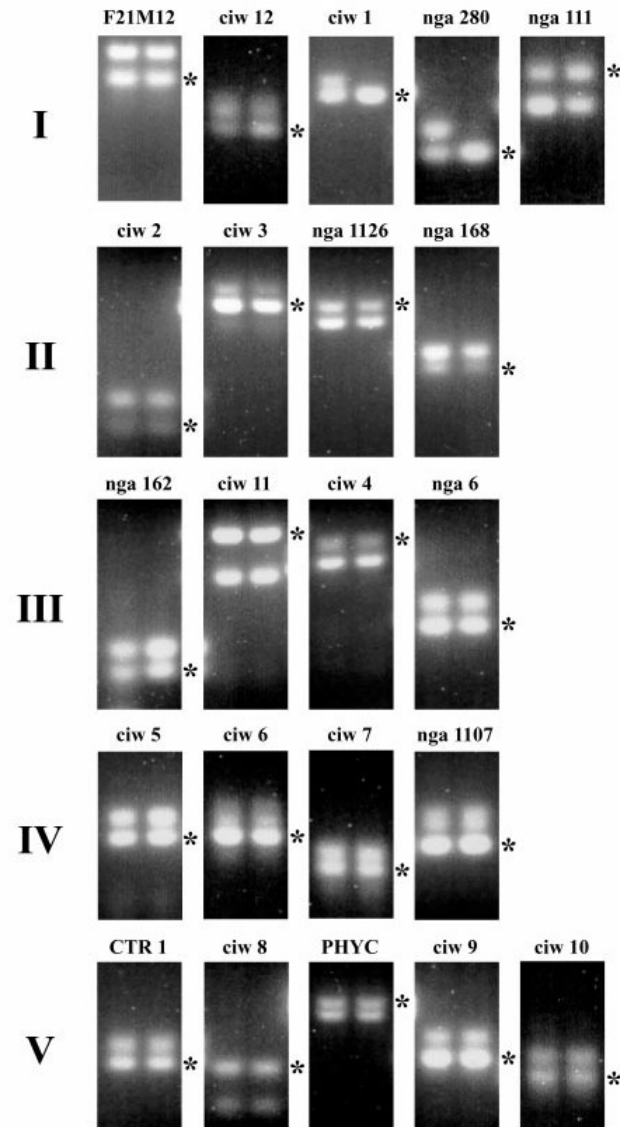
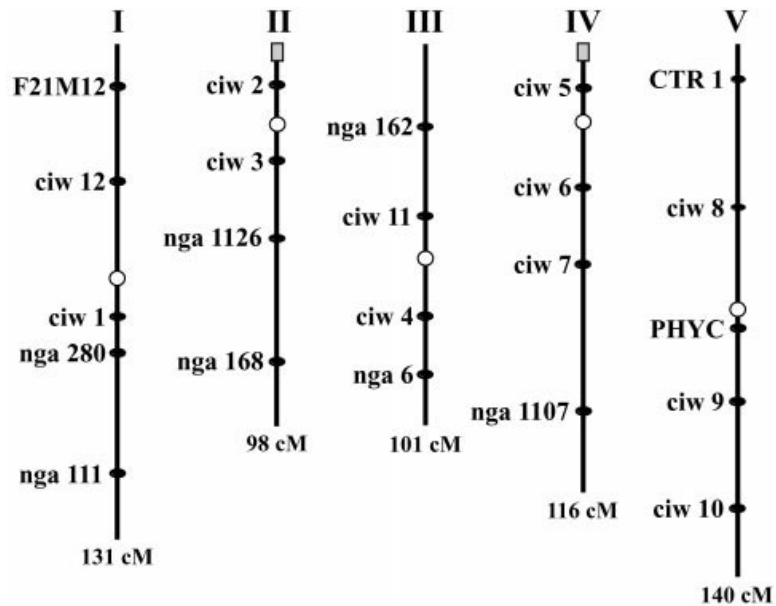
SSLP (simple sequence length polymorphisms) markers exploit the variability of short repetitive sequences for mapping purposes

CAPS (cleaved amplified polymorphic sequences; Konieczny and Ausubel, 1993), CAPS marker exploits polymorphic restriction sites for mapping purposes

dCAPS (Michaels and Amasino, 1998; Neff et al., 1998), markers can exploit almost all single nucleotide changes for mapping purposes. To achieve this a mismatched PCR primer is designed next to the polymorphic position such that an artificial restriction site is created with the sequence variant of just one accession.

BLUKED SEGREGANT ANALYSIS

- Su pool di DNA da diversi individui mutanti della popolazione segregante F2
- Controllo: DNA di individui F1 (eterozigoti in tutti i loci)



Three-point mapping

Once you have assigned your mutation to a chromosome, three-point mapping is almost always an obligatory step in the process of cloning our mutants.

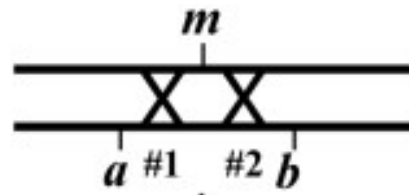
Even SNP mapping is usually preceded by three-point mapping with genetic markers.

We cross our mutation (m) into a strain with two linked morphological markers (a and b) that are on the same chromosome as m , to generate the $m/a b$ heterozygote.

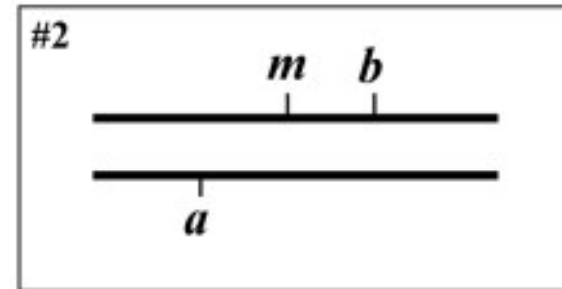
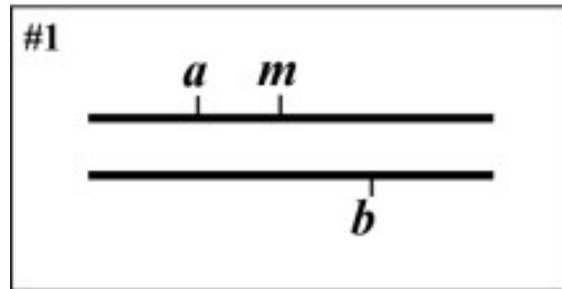
We then isolate and follow two classes of recombinant progeny; those that display the **A** phenotype only (**A non-B** recombinants) and those that display only the **B** phenotype only (**B non-A** recombinants). By seeing which of these two classes produce the mutant phenotype (**M**) and by scoring the percentages for each, we can determine whether our mutation lies to the left, to the right, or between our set of markers.

In the case where the mutation lies in between, we may then determine the approximate distance from each marker.

Three-point mapping

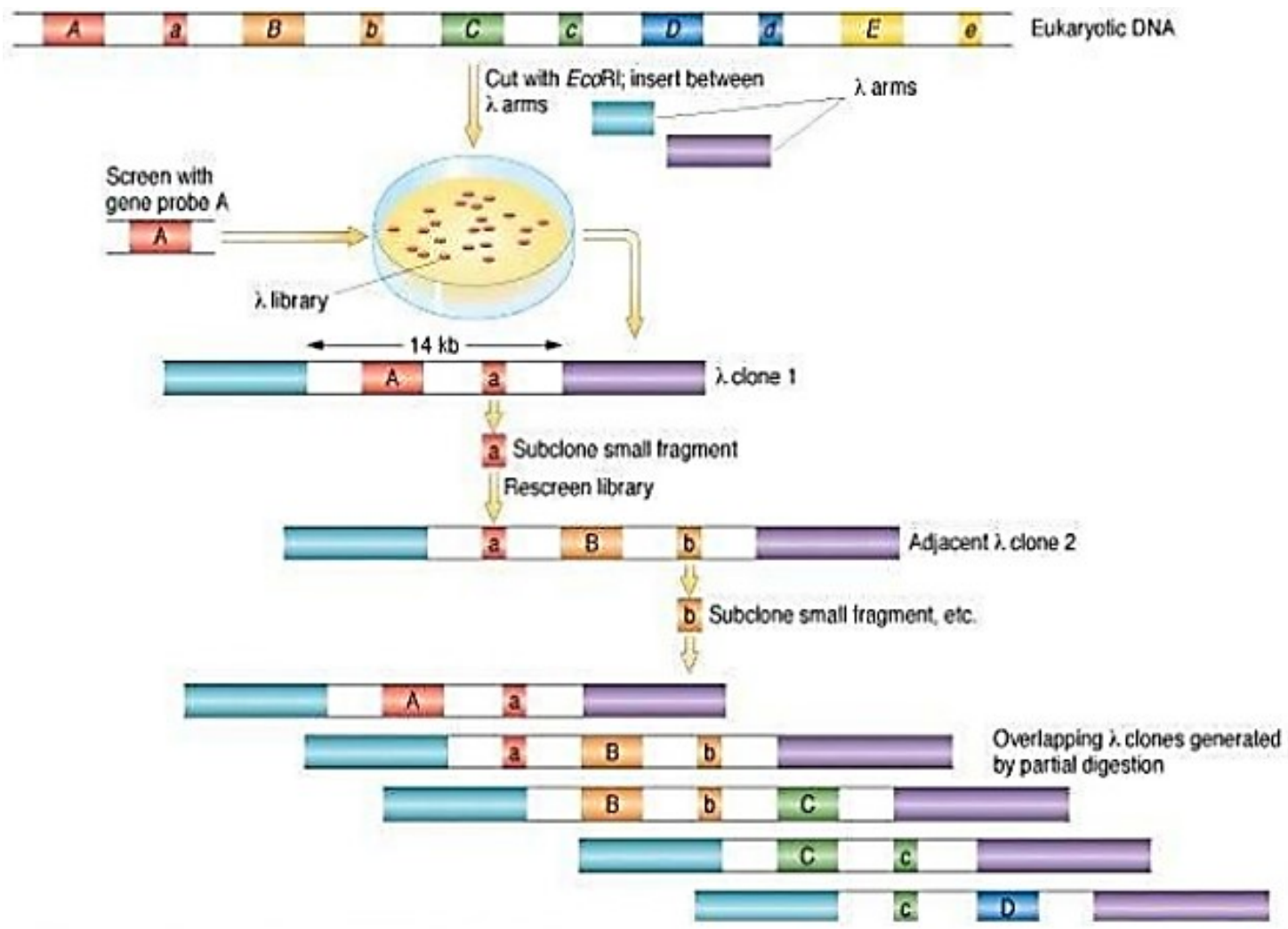


recombinant chromosomes



(recombinant + parental)





Identificazione di un gene R tramite clonaggio posizionale

- 1) Identificare i marcatori genetici fiancheggiati associati al gene R
- 2) Identificare i marcatori molecolari strettamente associati mediante analisi di un ampio gruppo di segreganti
- 3) Identificare i cloni YAC e BAC che contengono i marcatori fiancheggiati più vicini
- 4) Identificare i cloni cosmidici che si trovano tra i siti di ricombinazione
- 5) Trasformare un genotipo suscettibile con un singolo clone cosmidico; esaminare i trasformati per l'identificazione di piante resistenti alla malattia
- 6) Sequenziare la regione di sovrapposizione tra i cloni 1 e 2 per identificare il modulo di lettura aperta (ORF) del gene R

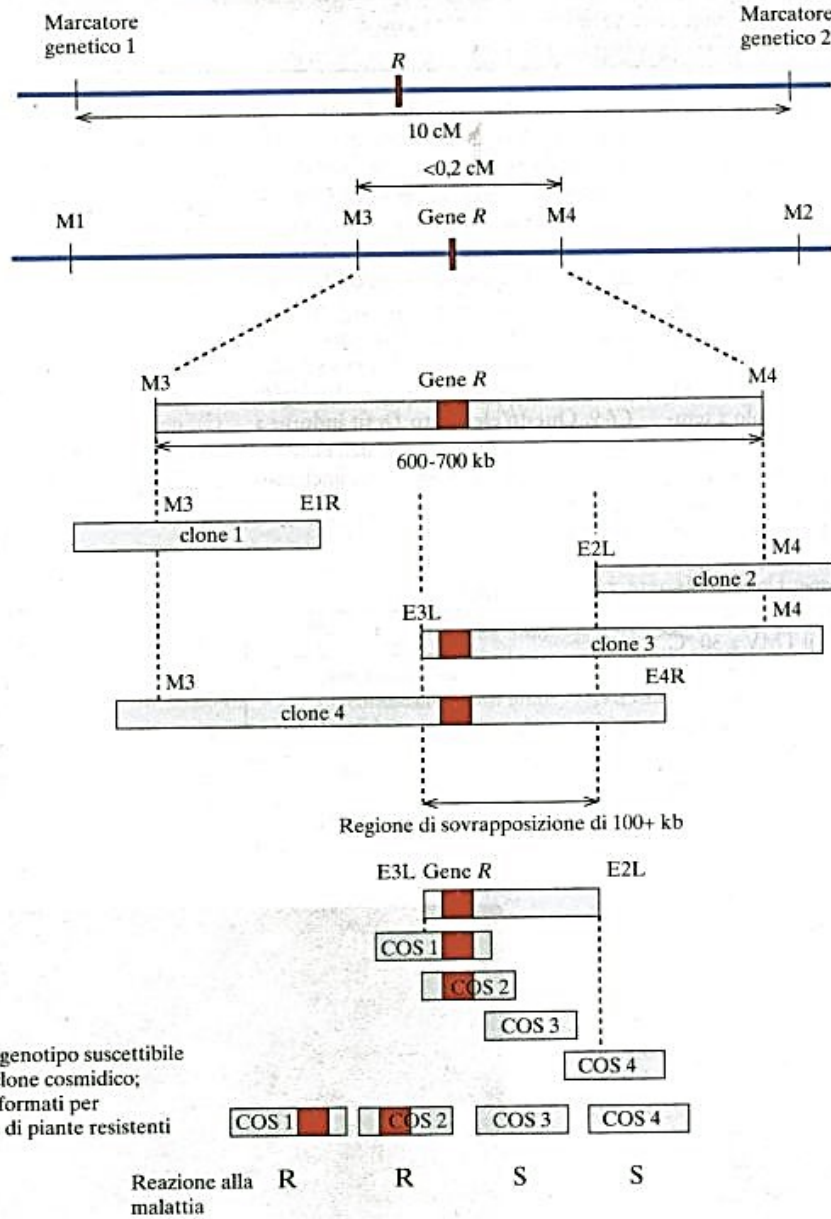


Figura 21.25 Identificazione di un gene R tramite clonaggio posizionale. R, resistente; S, suscettibile; YAC, cromosoma artificiale di lievito; BAC, cromosoma artificiale batterico.

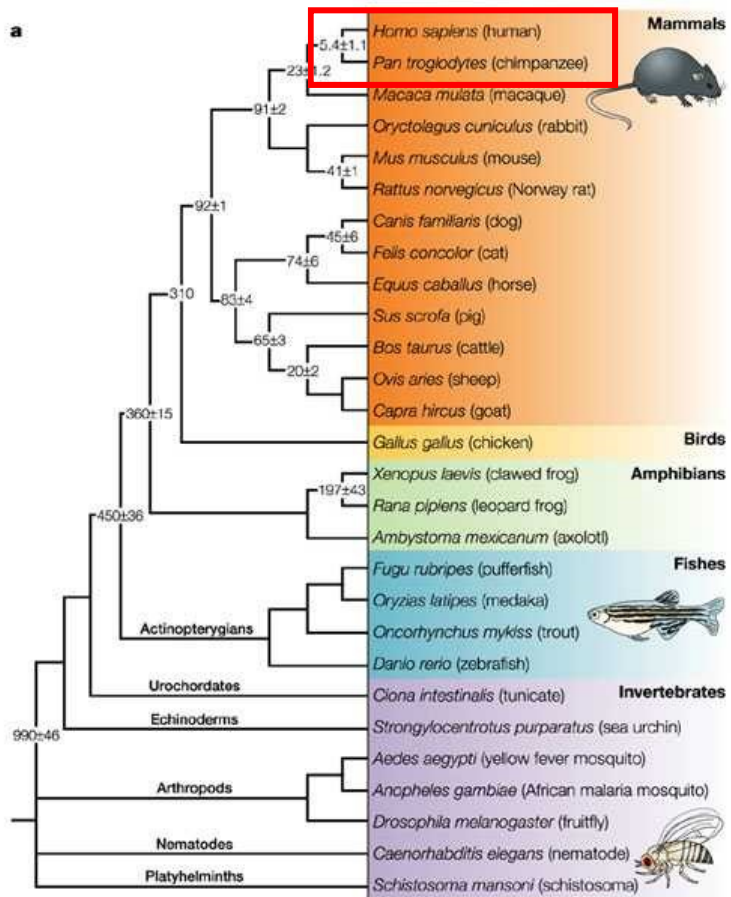
SINTENIA

- Dal greco = legati insieme
- Indica in genetica la presenza di due o più loci sullo stesso cromosoma
- Oggi il concetto è stato espanso per investigare l'**omeologia** (omologia residua tra cromosomi che in origine erano completamente omologhi)

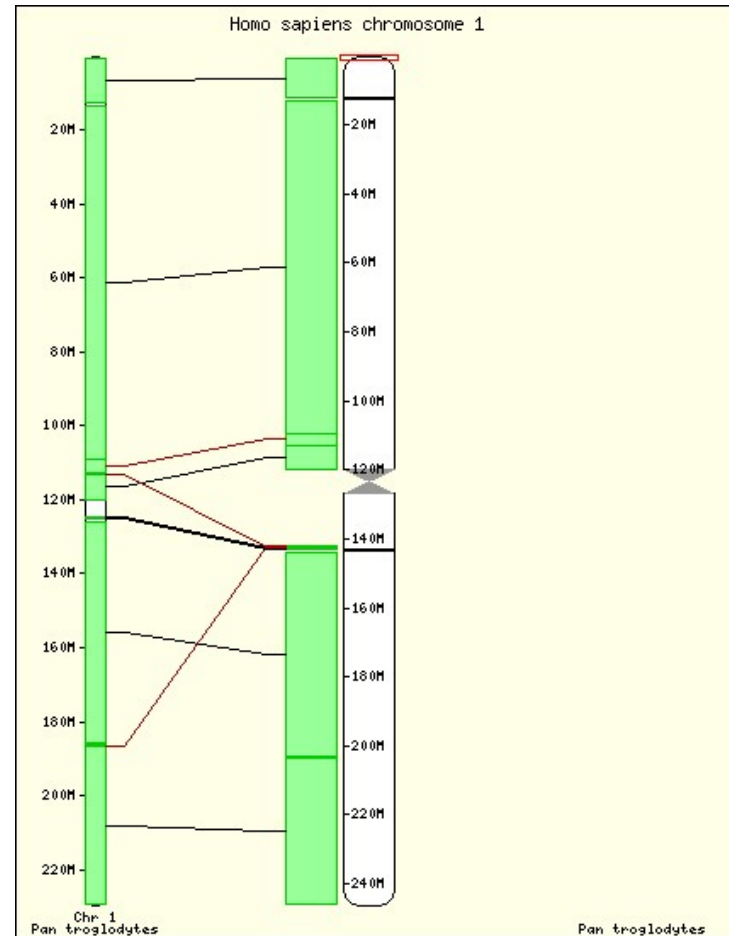
Nella maggior parte delle piante, l'evoluzione delle porzioni di genoma, piccole ma essenziali, che codificano per i geni ha proceduto con tempi relativamente lenti

-> sequenze di DNA intrageniche e organizzazione dei geni lungo i cromosomi sono riconoscibili

Molti fattori, come duplicazioni cromosomiche o segmentali, mobilità di sequenze di DNA (es. trasposoni), delezioni e riarrangiamenti localizzati, si sono sovrapposti a tale lenta evoluzione, causando molte deviazioni dalla co-linearità



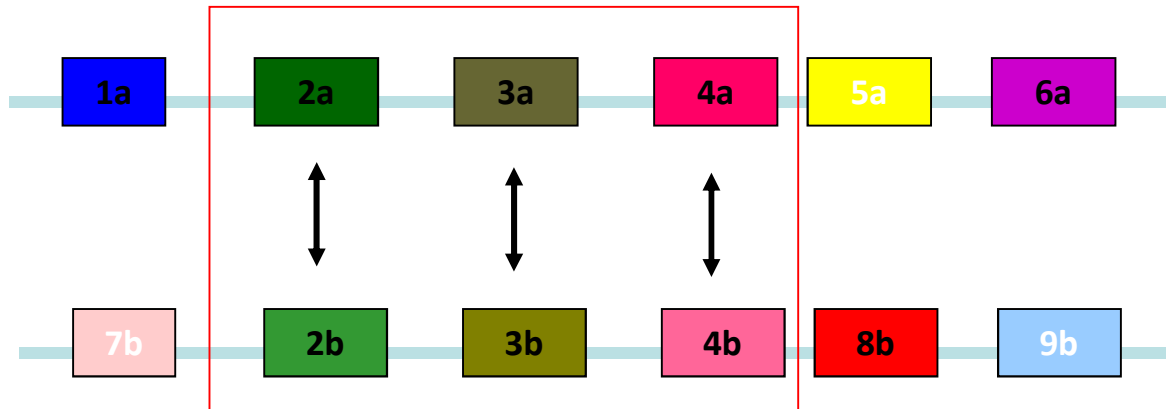
THE ORIGIN AND EVOLUTION OF MODEL ORGANISMS
 Hedges, *Nature Reviews Genetics* 3, 838 -849 (2002)



<http://www.ensembl.org/>

Blocchi di sintenia

Organismo A



Organismo B

SORGO (diploide) e CANNA DA ZUCCHERO (autopoliploide)

si sono separati circa 5 milioni di anni fa

-> alto grado di colinearità

-> potenziale uso del genoma di sorgo per identificare geni ortologhi nella canna da zucchero



C. Asnagli · F. Paulet · C. Kaye · L. Grivet · M. Deu
J.C. Glaszmann · A. D'Hont

Application of synteny across Poaceae to determine the map location of a sugarcane rust resistance gene



Ruggine (*Puccinia melanocephala*) su canna da zucchero cv CP72-1210

Può causare perdite fino al 40% del raccolto

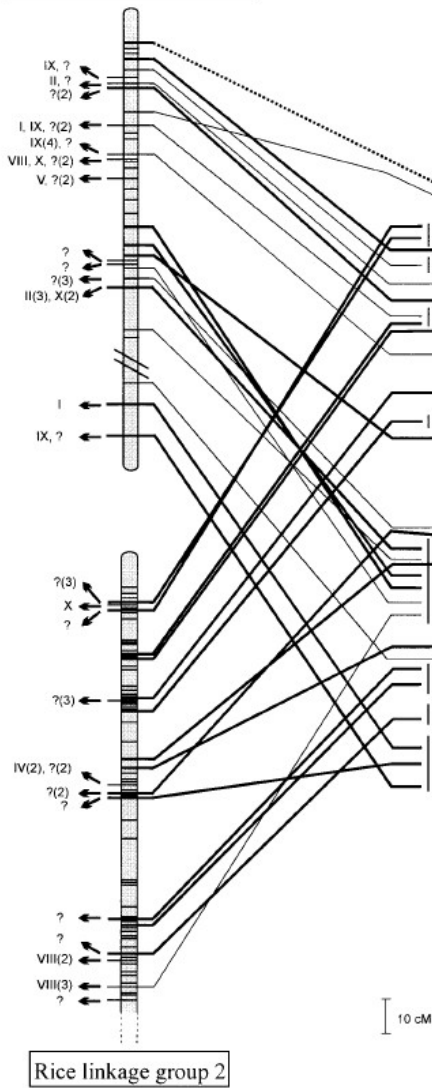
La cultivar R570 possiede un gene di resistenza

Table 1 Distribution of the 279 markers mapped on R570 linkage groups and locus coincidence with sorghum, maize and rice homoeologous segments

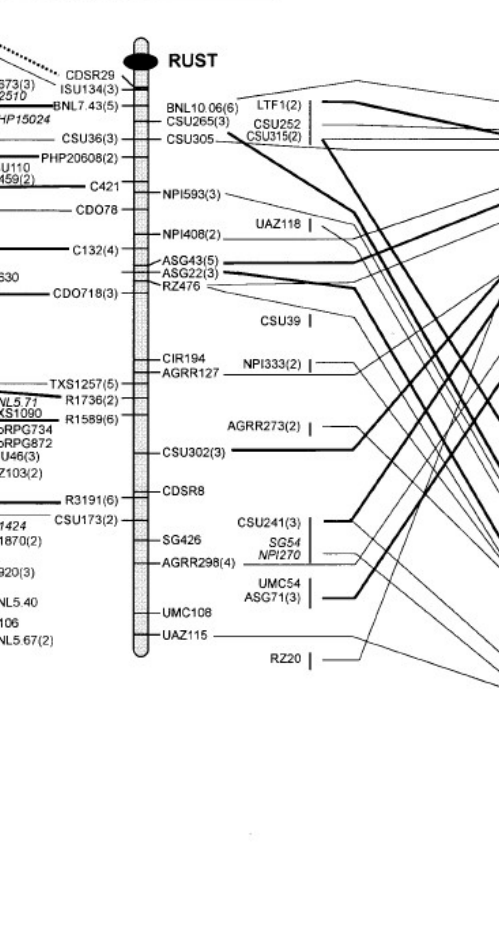
Linkage group ^a	Number of markers/linkage group	Number of probes involved	Locus coincidence with homoeologous regions			
			Sorghum D	Maize 4	Maize 5	Rice 2
I	11	5	2	2	2	–
II	12	6	2	–	4	–
IV	4	2	–	1	–	1
V	4	4	1	–	3	–
VII	122	53	17	11	14	14
VIII	10	7	1	3	1	2
IX	36	18	4	3	13	–
X	13	8	2	2	4	1
XI	5	2	–	1	1	–
Total	217		29	23	42	18
L	32	9				
U	30	34				
Total	279					

^a Roman numbers indicate sugarcane linkage groups (in accordance with Grivet et al. 1996); L indicates co-segregation groups yet unassigned to any defined linkage group; U indicates markers yet unlinked. The nomenclature of sorghum linkage groups is that of Pereira et al. (1994)

Sorghum linkage group D



Sugarcane linkage group VII



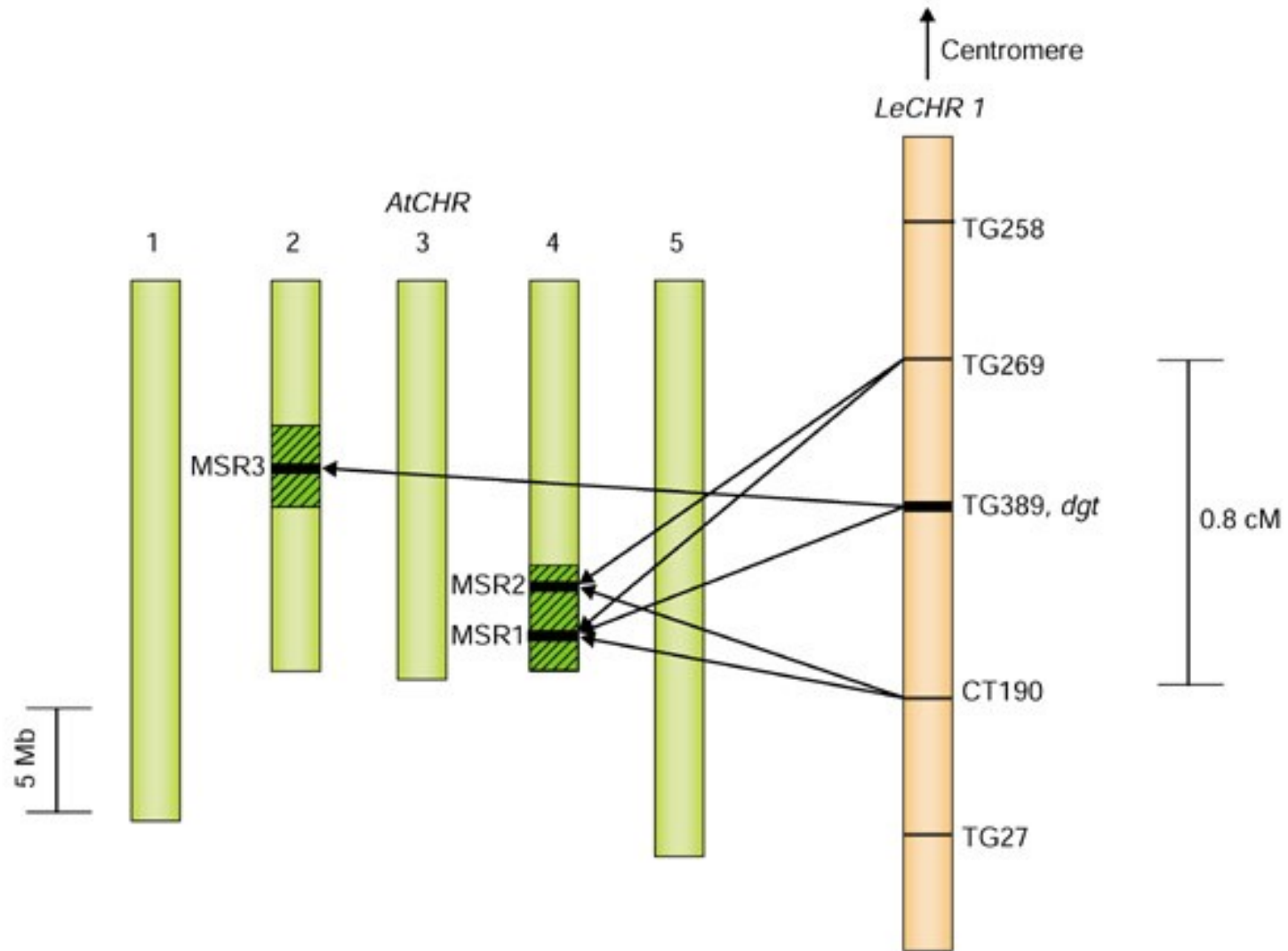
Maize linkage group 5



Maize linkage group 4

10 cM

Regioni microsintetiche tra arabidopsis e pomodoro

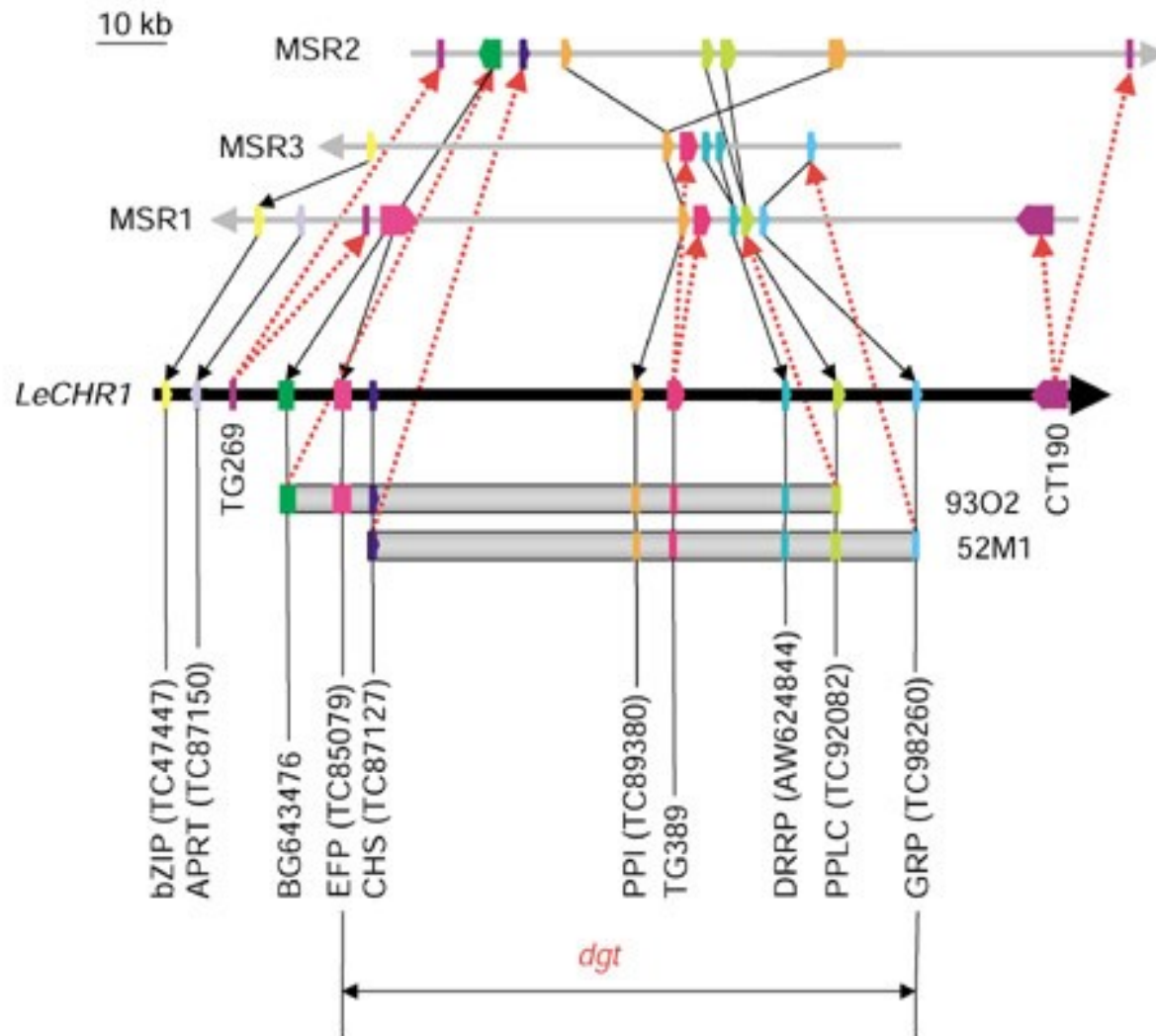


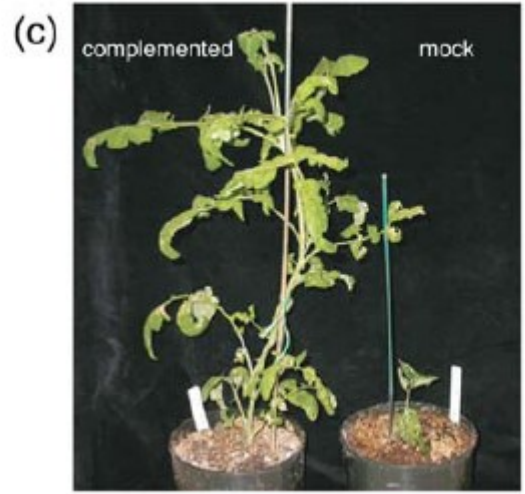
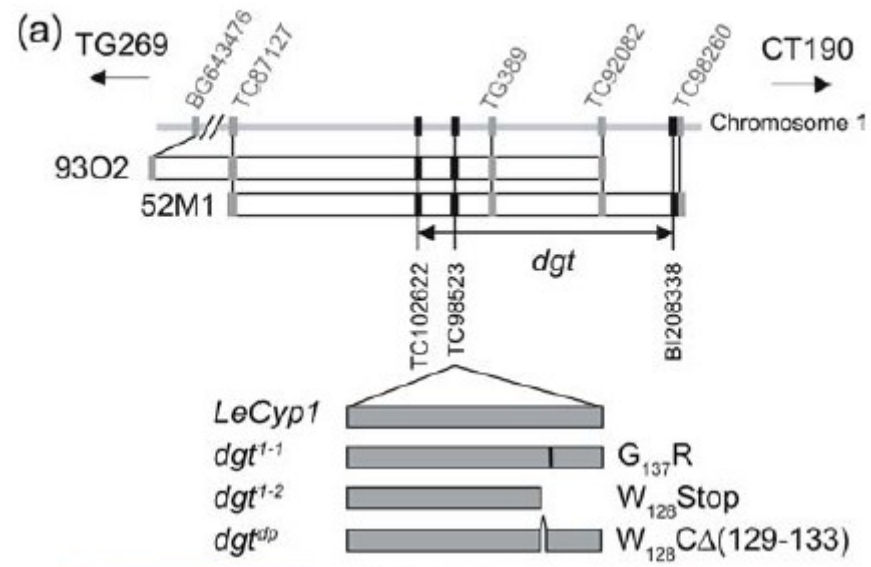
Il mutante *diageotropica* (*dgt*) di pomodoro

- Ridotta sensibilità all'auxina



Mappatura del locus *dgt* di pomodoro sulla base della microsintenia con arabidopsis





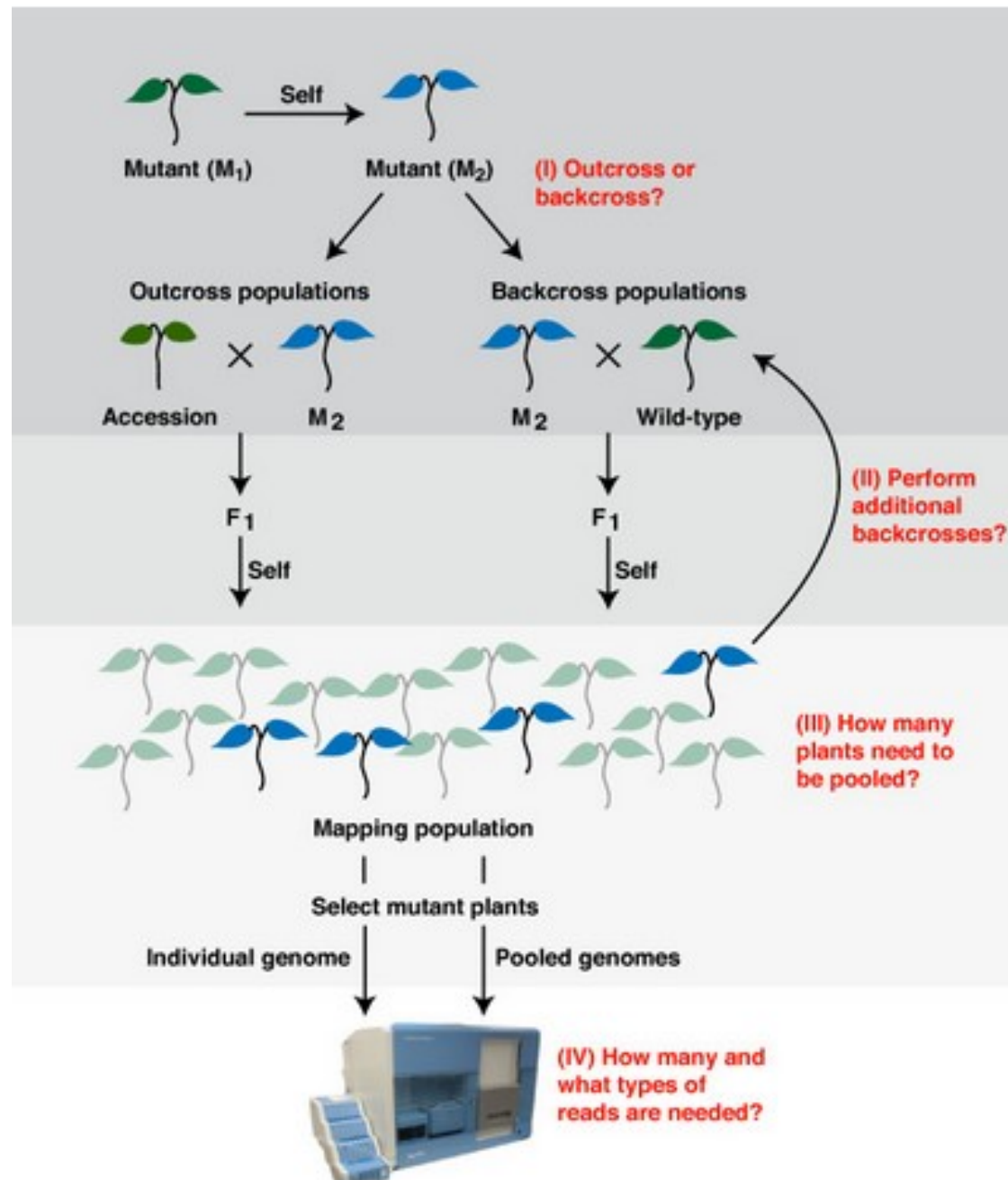
Mutation mapping by deep sequencing

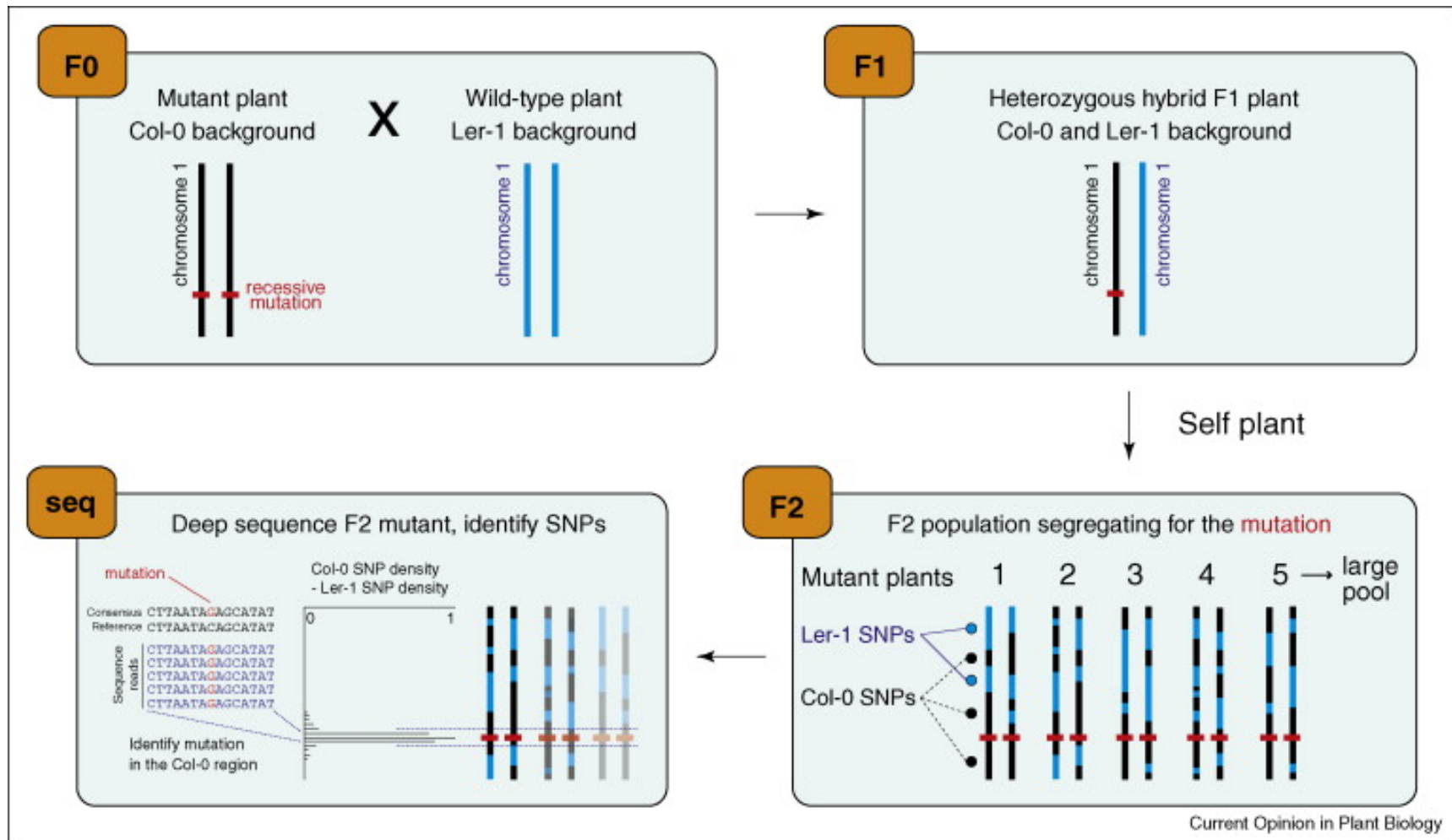
Forward genetic screens have proven extremely powerful in *Arabidopsis thaliana* for assigning genes to specific biological pathways .

The success of this approach is, in part, due to the highly accurate sequence of its compact genome, facile genetics, and extensive collection of mapping markers.

However, identifying the causative mutation commonly takes several months to years after generating a mapping population, so approaches to expedite this step will be highly valuable.

Deep sequencing of a pool of F2 individuals containing only mutant plants from a mapping population enables rapid mapping of enables rapid mapping of the mutation.





Identification of mutations by deep sequencing. A plant with Col-0 background that harbors a recessive mutation leading to a mutant phenotype is crossed to a wild-type Ler-1 plant. The heterozygous F1 hybrid plant is allowed to self-fertilize to produce a large pool of F2 plants that are segregating for the mutation. A large number of F2 plants that display the mutant phenotype are pooled and their gDNA subjected to deep sequencing. The density of single nucleotide polymorphisms (SNPs) inherent in the Ler-1 strain is subtracted from the density of SNPs indicative of the Col-0 background, identifying a discrete region on the chromosome in which only Col-0 marker SNPs are present. The deep sequencing data in this interval are then scoured for the potential causative mutation.

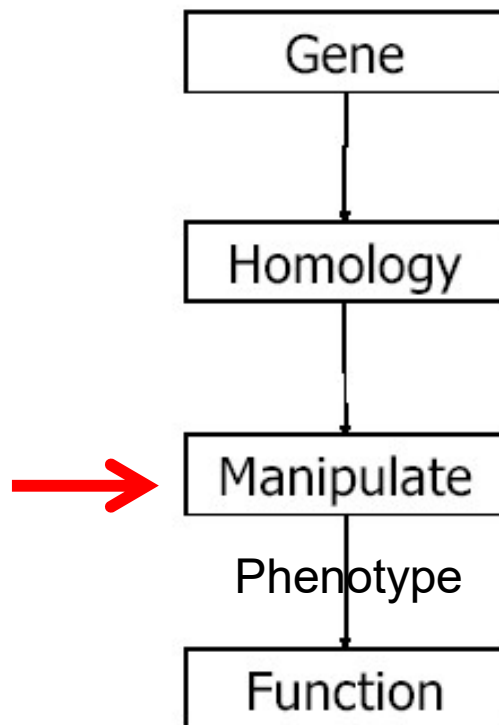
Table 1**Suggestions for the design of mapping-by-sequencing experiments**

	Outcross populations	Backcross populations	Direct sequencing	Deep candidate resequencing (dCARE)
Generation	F ₂	BC ₁ F ₂	BC ₁₋₃ F ₂ ^a	n/a
Mutants (n)	Approximately 150	Approximately 50	1	As many as possible
Optimal coverage	>25	Approximately 50	>25	n/a
Sequencing type	Paired-end	Paired-end	Paired-end	Single-end

^aDepending on mutation rate.

GENETICA INVERSA

'Reverse Genetics'



- Know a gene
 - genome sequence, EST etc
- and want to discover function
- Manipulate the gene
 - and do as for forward genetics

BUT

- Need to know roughly what to look for in phenotype (guess by homology)
 - But no idea of function of 30% of Arabidopsis genes
- Need gene not to have function covered by another gene
 - Multigene families (37% of Arabidopsis genes in families of >5 genes)

MANIPULATION OF THE EXPRESSION OF YOUR PLANT GENE FOR REVERSE GENETICS

- Gene disruption (Knock-out)
- Gene Knock-down
(you mostly buy mutants)

- Gene silencing (TARGETED) (Hairpin-loop, artificial miRNAs)
- Gene OVEREXPRESSION (35S promoter or inducible promoters)
- Genome editing
(you generate mutants)

- Gene overexpression or silencing in specific tissues (ex. GAL4-UAS)
(buy +generate)



Reverse genetics

Enabling tools for reverse genetics:

- Collections of mutations that reside in almost every gene
- Methods to rapidly knock out or knock or silence or overexpress genes of interest

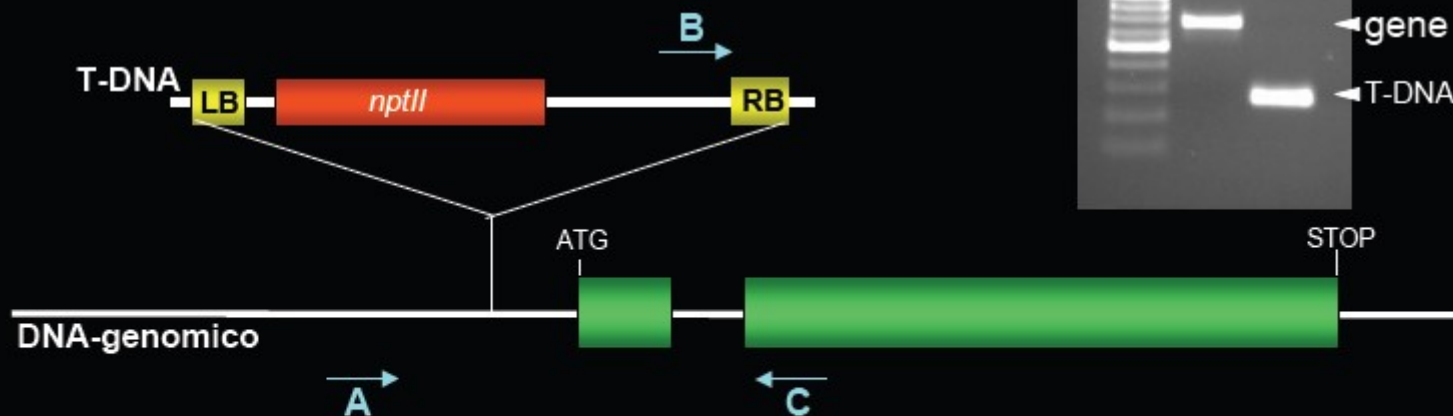
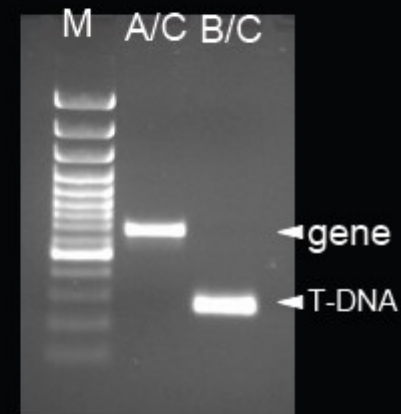
How to choose a gene:

- Your favorite gene family or gene function
- Candidate genes from other genomics approaches associated with your favorite biological process
- Candidate genes from other organisms associated with your favorite biological process

Approccio di Genetica Inversa:

- Conosco la sequenza del mio gene di interesse
- Conosco la sequenza del T-DNA utilizzato nella T/F

Disegno primers
specifici per il gene
ed il T-DNA

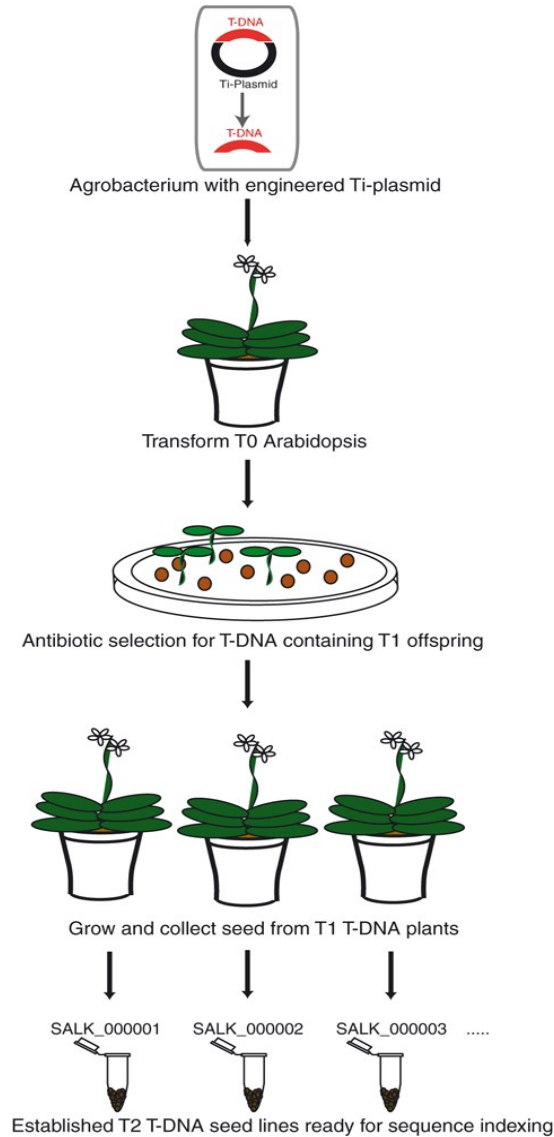


$A+C$ = sequenza genomica $B+C$ = sequenza T-DNA/genomica



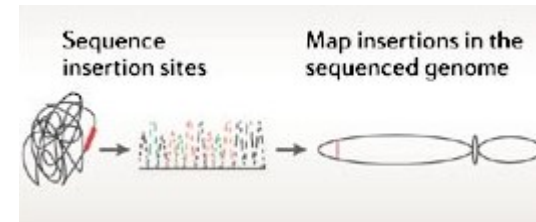
T-DNA insertional mutant collection for reverse genetics (commercially available)

Collection of mutants



Identification of insertion sites (TAIL PCR, Inverse PCR, plasmid rescue)

Catalogued mutants



Commercially available

Mutant collections

Submitted by	Background	Selectable marker	Promoter/enhancer/reporter gene	Population size	References
SALK	Columbia-0	Kanamycin	---	145589	Alonso et al. 2003
GABI-Kat ²	Columbia-0	Sulfadiazine	---	59455	Rosso et al. 2003
Czaba Koncz		Hygromycin	---	300	Szabados et al. 2002
Syngenta	C24	Hygromycin	GUS	1250	Sessions et al. 2002
INRA-Versailles	Ws, (Wassil-evskija)	Basta	GUS	1480	Balzerque et al. 2001
Le Clere and Bartel	Columbia	Basta	CaMV35S-cDNA	33100	LeClere and Bartel, 2001
Haseloff	C24	Kanamycin	GAL4-GFP	8000	Kiegle et al. 2000
Weigel	Columbia	Basta	Multimerised CaMV35S enhancers	>20000	Weigel et al. 2000
Sussman and Amasino ³	Ws-2	Kanamycin	Ap2::GUS	37800	Sussman et al. 2000
Jack	Columbia	Kanamycin	CaMV-GUS	11370	Campisi et al. 1999
Ehrhardt	Col-2 (CS907)	Basta	CaMV 35S -GFP	108	Cutler et al. 2000

La disponibilità della sequenza del genoma di *Arabidopsis* ha permesso di mappare sul genoma tutte le inserzioni

Tail-PCR e sequenziamento del DNA vicino al sito di inserzione

Ricerca omologia sul DNA genomico

Organizzazione in banche dati

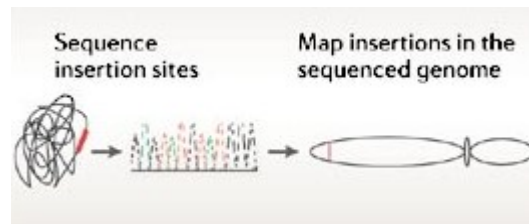
Per *Arabidopsis* il SALK raccoglie i risultati di più collezioni di mutanti (SAIL, Wisc) e la distribuzione è gratuita

<http://signal.salk.edu/about.html>

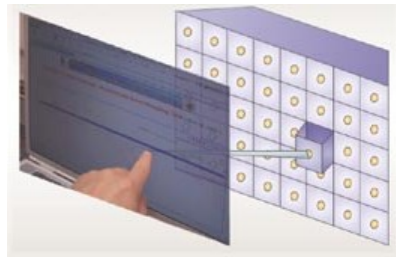


Then, what do you have to do?

Catalogued mutants



Select on the Internet and buy mutants in the gene or genes of interest



Molecular genotyping



SIGnAL

Salk Institute Genomic Analysis Laboratory

Funded by the National Science Foundation



T-DNA Express: Arabidopsis Gene Mapping Tool (May. 7, 2008)

1. Search : [\[?\]](#)

Type:

Query:

Chr:

Posn:

Display:

2. iSect Tool : [\[?\]](#)

3. Data Source :

- a. [Data Source, Detail and Summary.](#)
- b. [Gene Expression Atlas Data Source.](#)
- c. [Data Release Policy.](#)
- d. [FAQs.](#)



NASC

European Arabidopsis Stock Centre

Seed and information resources.

[About NASC](#) | [Address & Staff](#) | [Ask a Question](#) | [Background lines](#) | [Bioinformatics at NASC](#) | [Growing Arabidopsis](#) | [What is Arabidopsis?](#) | [Historical Documents](#)

Stock Catalogue

- Search Catalogue
- Browse Catalogue
- The Wiki
- About plant ontology
- How to Order
- Ordering FAQ
- Stock Overview
- Price Information
- Check Order Progress
- Seed Donation Form
- MTA FAQ
- Transcriptomics**
- Microarray Homepage
- Apply for Service

Stock Catalogue

NASC now maintains over 300,000 accessions of *Arabidopsis thaliana* representing over half a million genotypes



- Search Catalogue
- Browse Catalogue
- Browse by Ontology

[How to link to NASC stocks](#)
[How to reference NASC services](#)
The great annotation experiment:
[WeedsWorldWiki](#)
[Donate Seed](#)

News

Latest Additions to the Stock Catalogue:

[GABI-Kat lines](#) - donated by Bernd Weisshaar
[JIC Bancroft RI lines](#) - donated by Ian Bancroft.

Current News:

[NSF fund \\$50M iPlant Co-operative at Arizona](#)
Also see: [webcast](#) (choose archive button: top-right)
[International Arabidopsis Meeting \(ICAR\) 2008 - Montreal](#)
[MASC report 2007 available](#)
[WeedsWorld Wiki](#) - review that stock.
[Weed for Thought](#)
- a Dutch Ph.D. thesis considering NASC and TAIR.

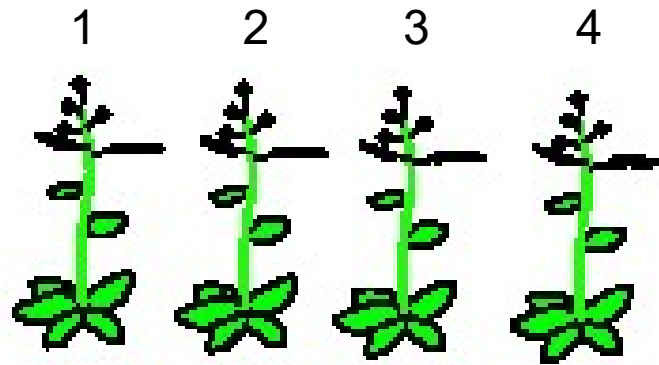
Transcriptomics

NASCarrays processes Affymetrix

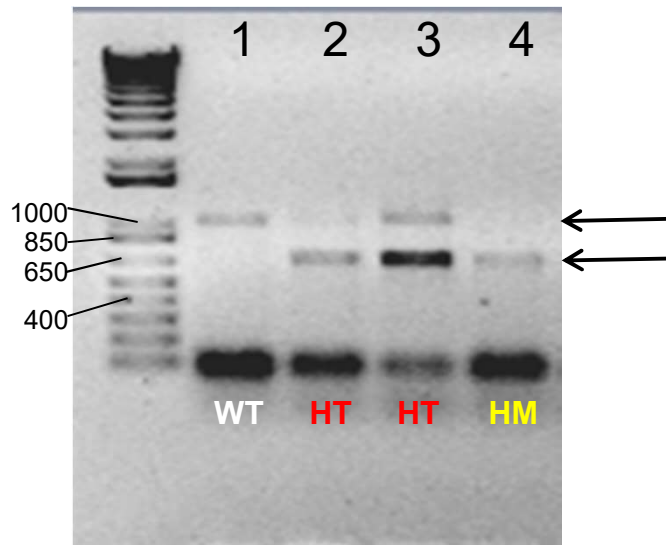
Genomics

NASC's well established genomic

Mutant genotyping

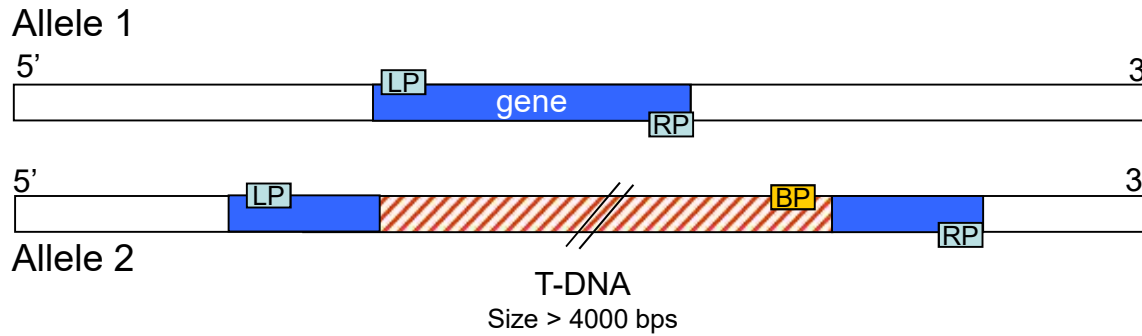


Genomic extraction

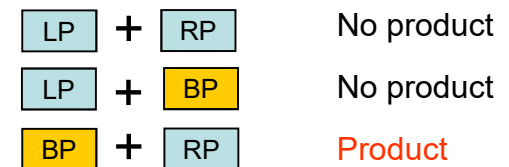
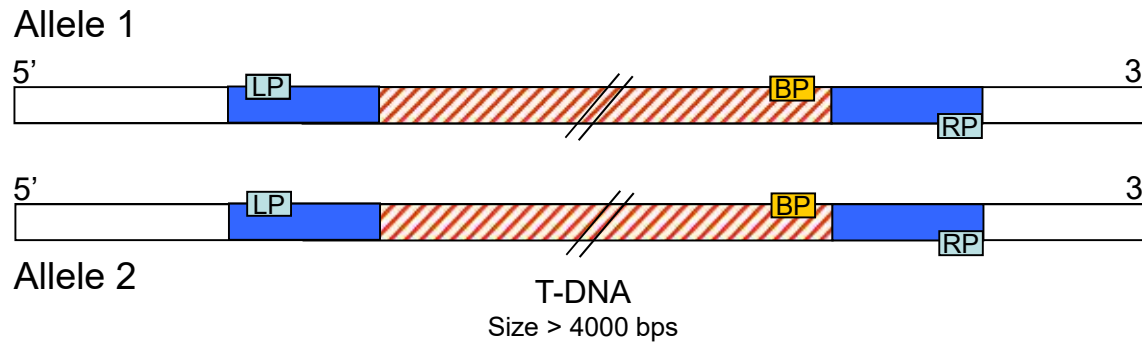
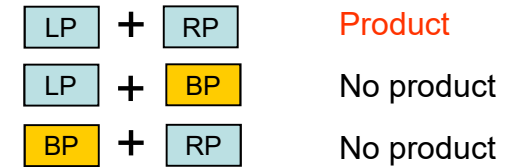
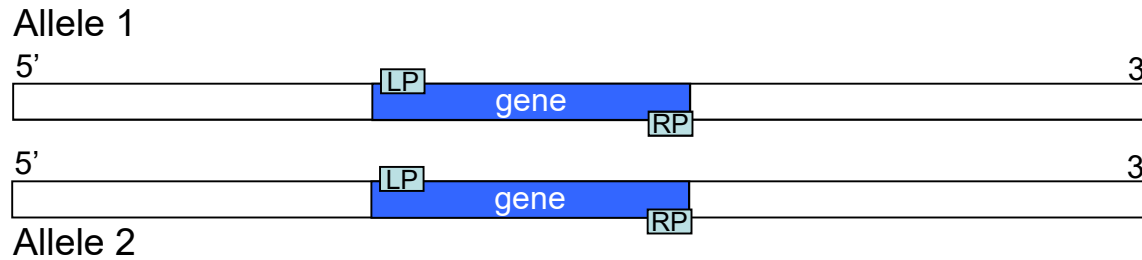
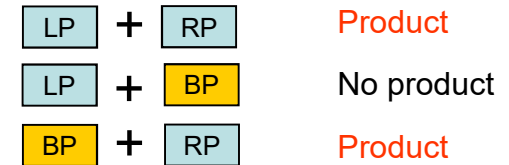


PCR

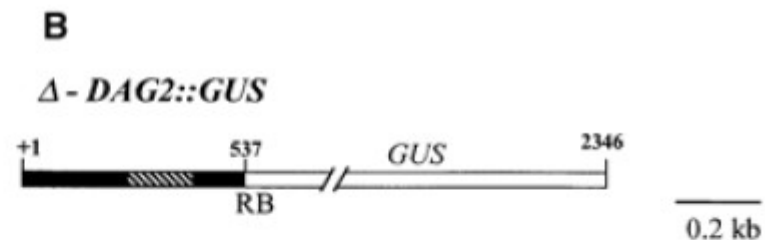
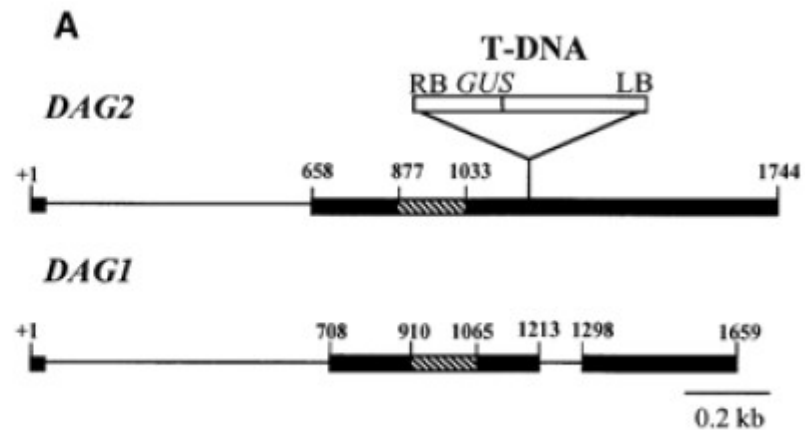
Mutant genotyping



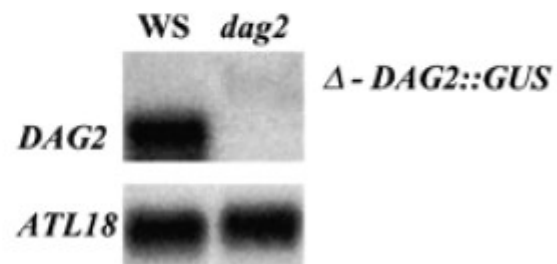
PCR primer pairs



Schematicamente i mutanti inserzionali possono essere rappresentati come nei disegni a fianco



Se il gene è distrutto non si dovrebbe avere la presenza del trascritto



La T/F con *Agrobacterium* ha una efficienza media di **1,5 inserzioni di T-DNA** per linea

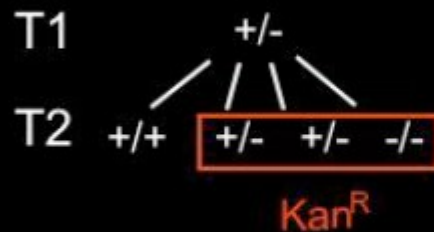
Più inserzioni nella mia linea?



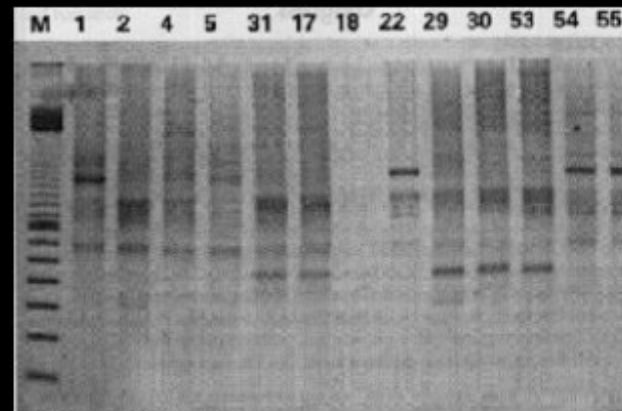
- in loci differenti

Rapporti di segregazione di linee eterozigoti sulla resistenza e Southern

- nello stesso locus
Southern



- = Allele selvatico
- = T-DNA



Southern Blot

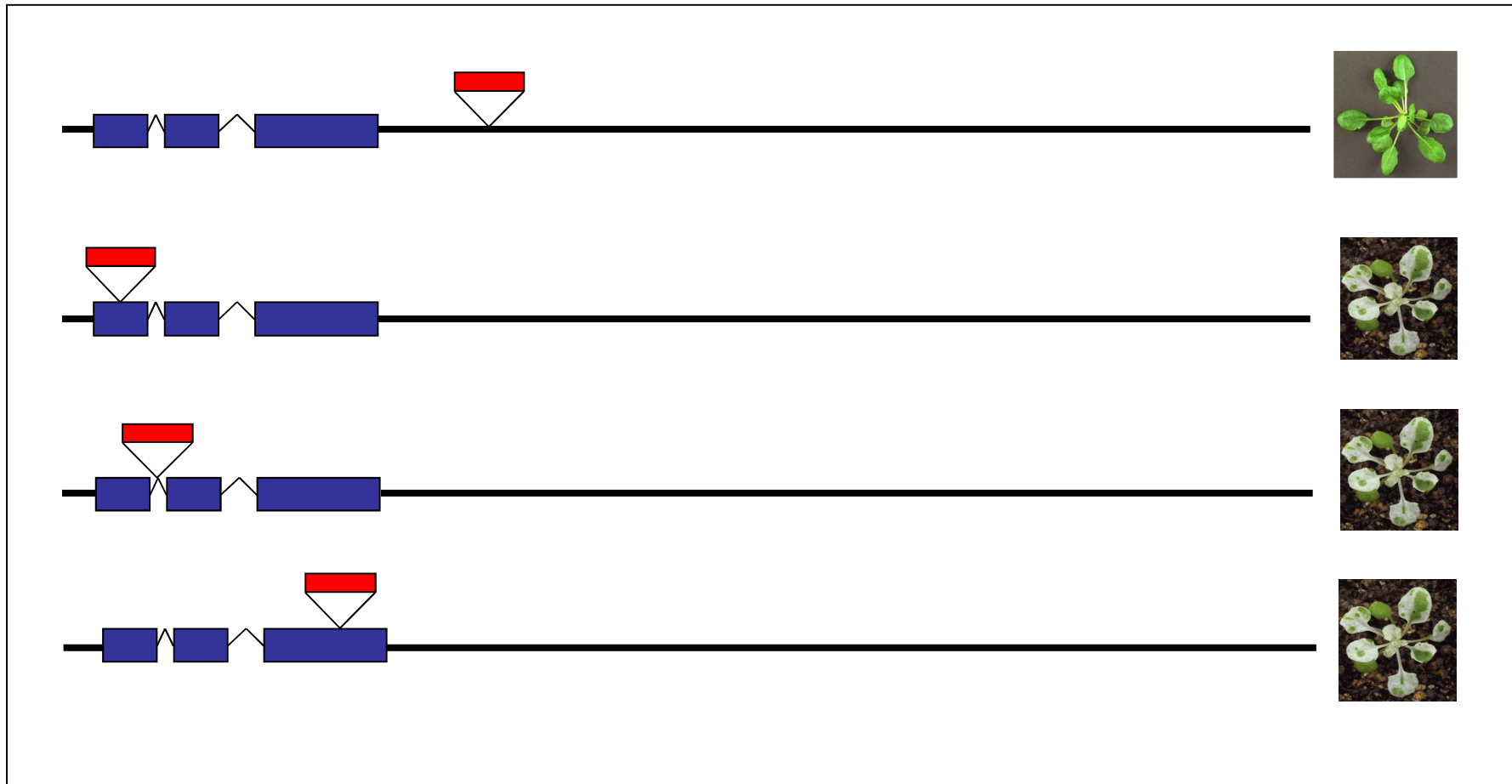
➔ Analisi del pattern di bande

Se ci sono più inserzioni in loci differenti si incrocia
la linea mutante con la rispettiva linea selvatica
(reincrocio)



Finchè non trovo un rapporto di segregazione di 3:1

How to choose the mutant of interest (you should have at least two alleles)



Select homozygous and propagate

Back crossing

IN general, but expecially for pleiotropic mutants, it is necessary to determine whether the phenotypes result from same or from two different mutations. This can be determined by multiple round of backcrossing or by co-segregation analysis

Backcross of the mutant to wt is important to clean up the mutant.

Theoretically, each back cross removes half of the unlinked secondary mutations.

The probability that a particular unlinked mutation remains after n rounds of back crossing is $(1/2)^n$

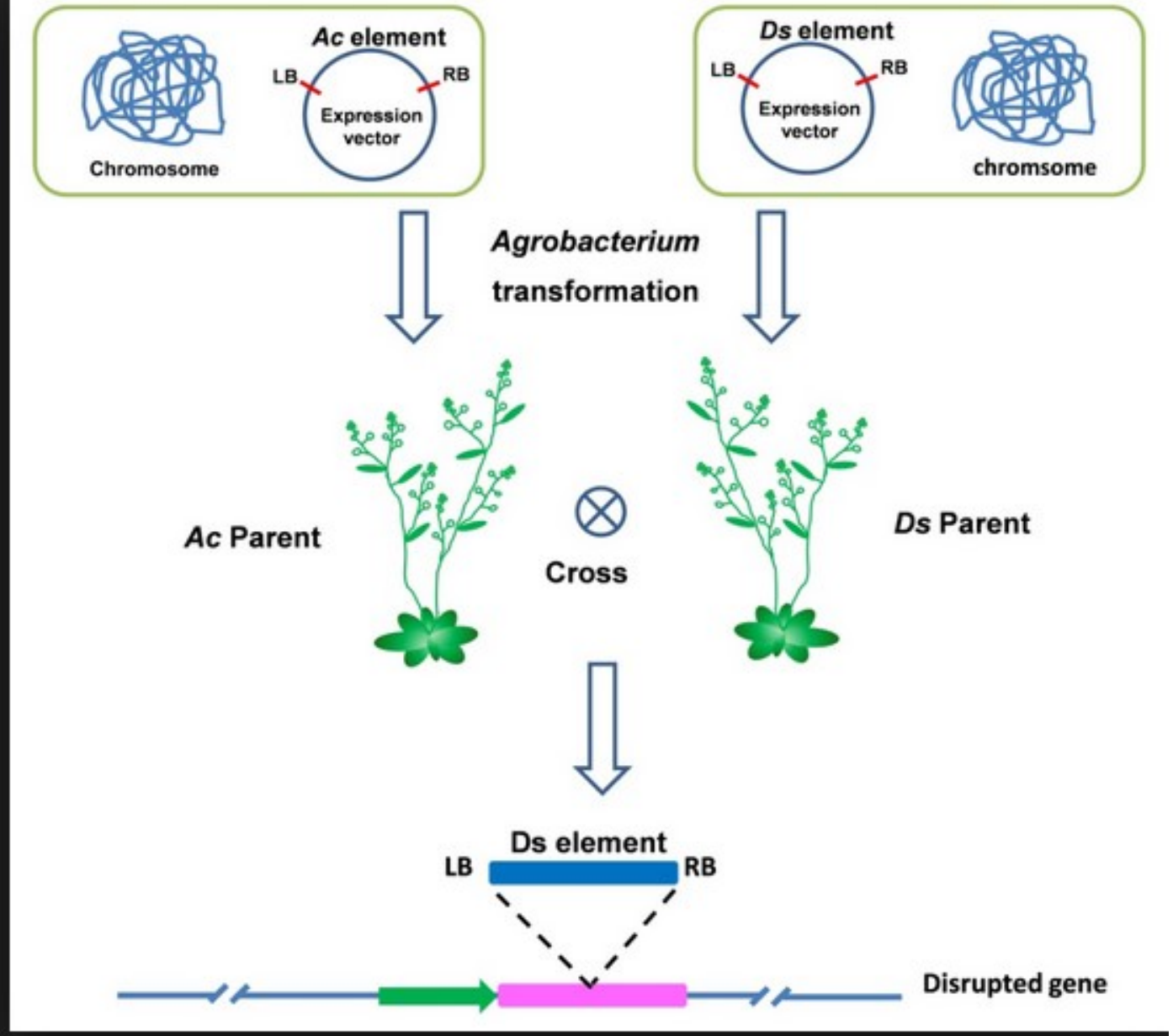
Analyze phenotype

Constraints on use of knockouts

Genes may have no obvious loss-of-function phenotype due to:

- presence of duplicate genes
- existence of alternative metabolic pathways and regulatory networks
- lethality

TRASPOSONI ED ACTIVATION TAGGING



Mutagenesis by transposon tagging

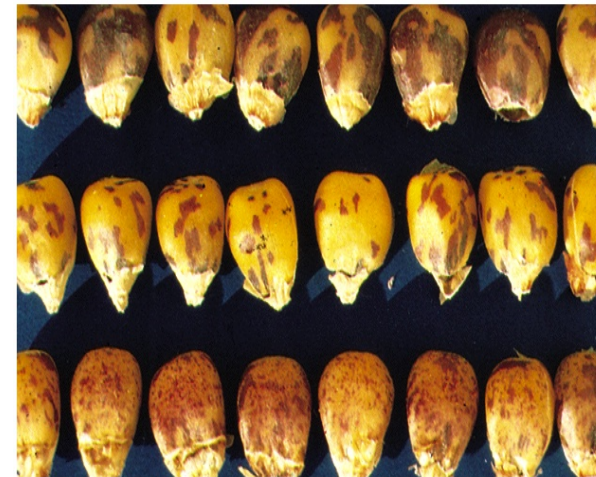
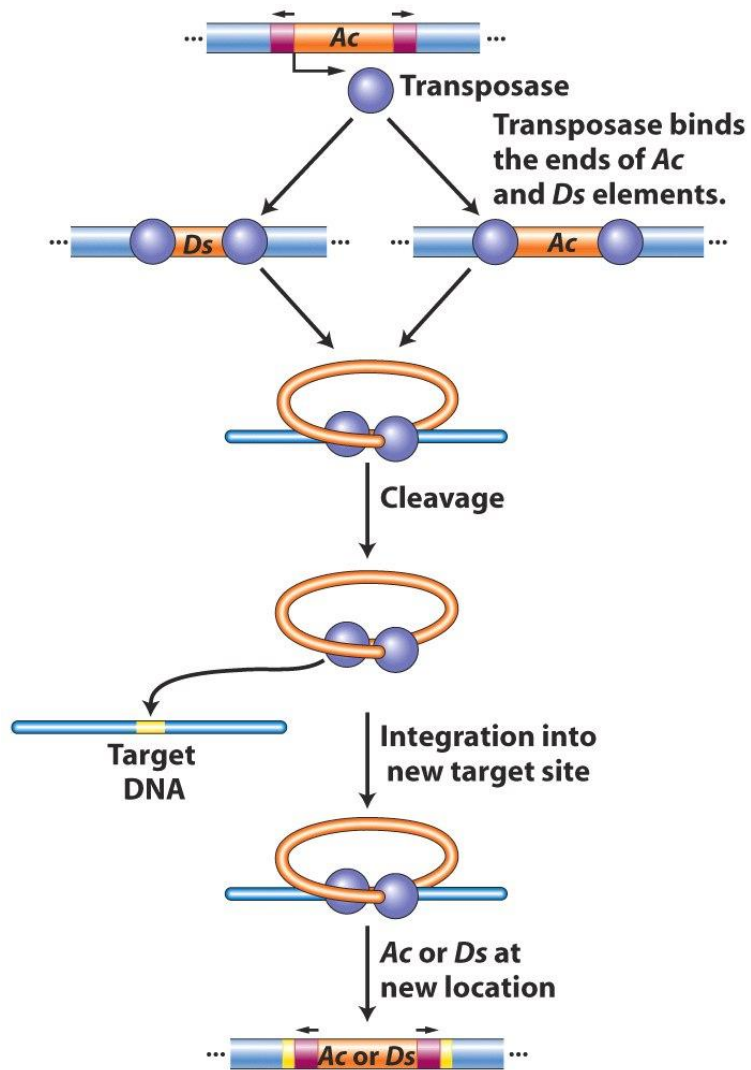
- Two maize-derived transposons are used
 - » Ac/Ds Transposon
 - » (En1 Transposon)
- Ds is the transposon (does not encode a transposase)
- Ac is an immobile element encoding the transposase
- Ds transposon is maintained immobile because Ac (encoding transposase) is maintained in a separate plant line
- To mobilize the Ds transposon you cross the plants and obtain an F1
- Eventually you segregate out Ac to have stable lines

MUTAGENESI TRAMITE TRASPOSONI

Esempio: il sistema Ac/Ds
(Activator/Dissociator)

Due linee transgeniche, che recano una l'elemento Ac e l'altra l'elemento Ds vengono incrociate.

Nella progenie la TRASPOSASI Ac mobilizza l'elemento Ds, generando piante F1 che sono MOSAICI GENETICI (Barbara McClintock)



Transposon tagging

Ac-Ds system: contain 11 bp terminal inverted repeat (TIR), create 8 bp target site duplication. The 4.5 kb *Ac* element codes for 3.5 kb mRNA for transposase.

En-Spm system: contain a 13 bp TIR and create a 3 bp target site duplication. *En* element is 8.3 kb long and contains 2 alternately spliced gene products, *TnpA* and *TnpD*. Both are required for transposition.

Robertson's mutator (*Mu*)/ *MuDR*: Very mobile, short (1.5 or 1.7 kb) in maize. Cause high rate of mutation, somatic instability. *MuDR* encodes the MURA transposase required for *Mu* transposition and MURB, a helper protein implicated in insertion. All *Mu* elements share 215-bp terminal inverted repeat (TIR) sequences and the mobile *Mu* elements contain a highly conserved 32-bp MURA transposase binding site. Characteristic 9-bp host sequence duplications are generated during *MuDR/Mu* germinal insertion.

MUTAGENESI TRAMITE TRASPOSONI

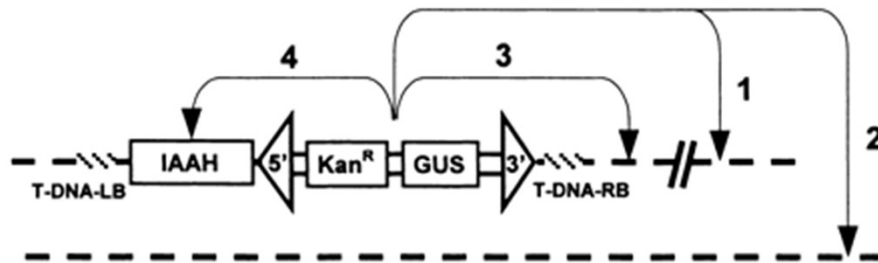
Gli elementi Ds traspongono preferibilmente in siti vicini al sito donatore, per cui le linee sono disegnate in modo da permettere:

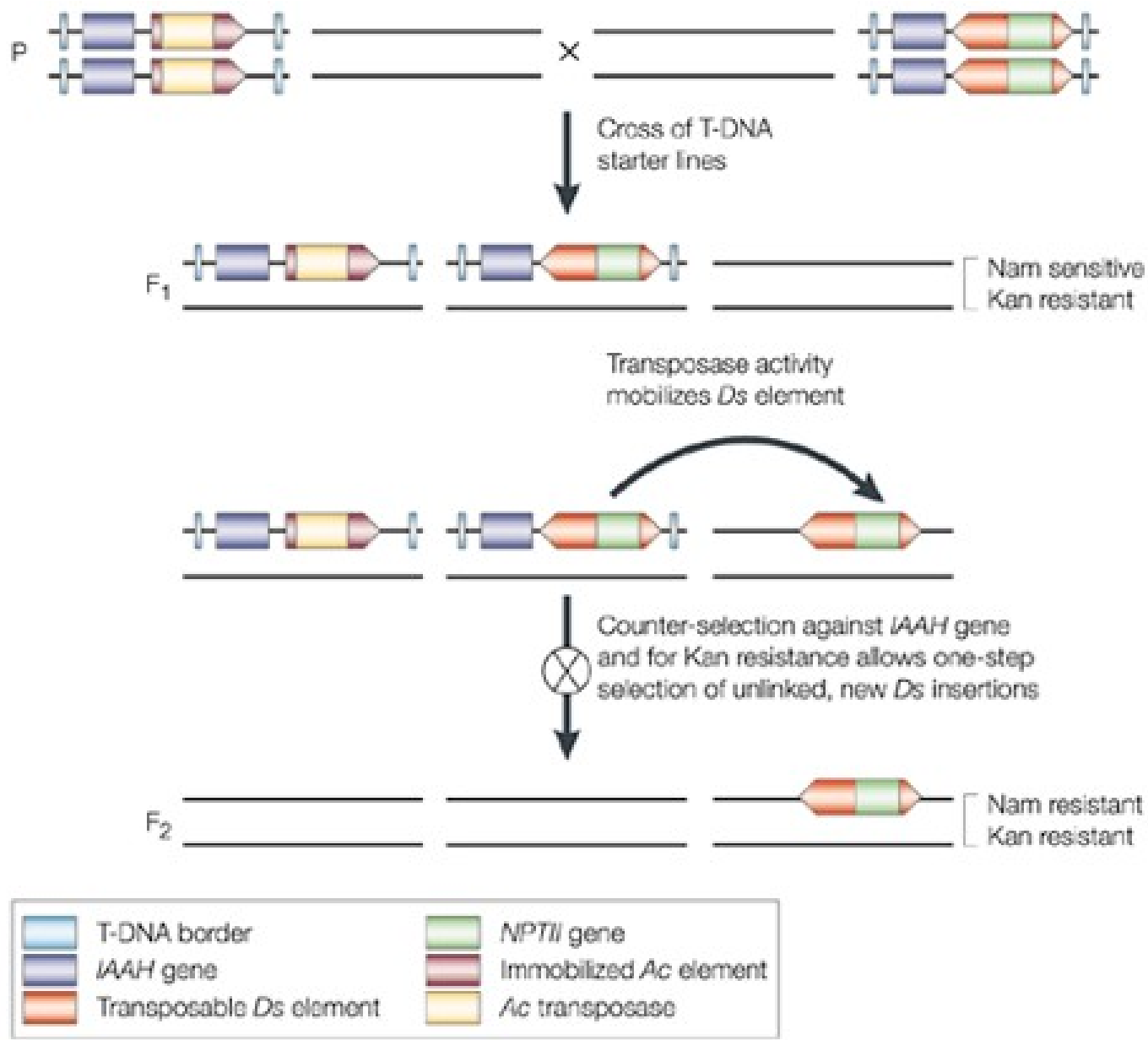
- a) la selezione positiva per la presenza di Ds originati da nuove trasposizioni
- b) la selezione negativa contro Ac e contro il locus Ds donatore.

Il T-DNA che contiene Ac reca anche il gene per l'indolacetammide idrolasi (IAAH), che conferisce sensibilità al naftalene acetammide (NAM).

L'elemento Ds porta il gene NPTII (resistenza alla kanamicina)

Il T-DNA per l'elemento Ds donatore porta anche il gene IAAH, in modo da selezionare contro la presenza del Ds trasposto vicino al Ds donatore.





VANTAGGI DELLA MUTAGENESI TRAMITE TRASPOSONI

- 1) La mutazione è reversibile: in presenza della trasposasi, il trasposone può escindersi nuovamente dal locus mutagenizzato, ripristinando il locus WT e confermando che il fenotipo è legato all'inserzione
- 2) Poiché gli eventi di trasposizione avvengono preferibilmente vicino al sito donatore, si può utilizzare questa strategia per effettuare mutagenesi mirate su determinate regioni cromosomiche

Constraints on use of knockouts

Genes may have no obvious loss-of-function phenotype due to:

- presence of duplicate genes
- existence of alternative metabolic pathways and regulatory networks
- lethality

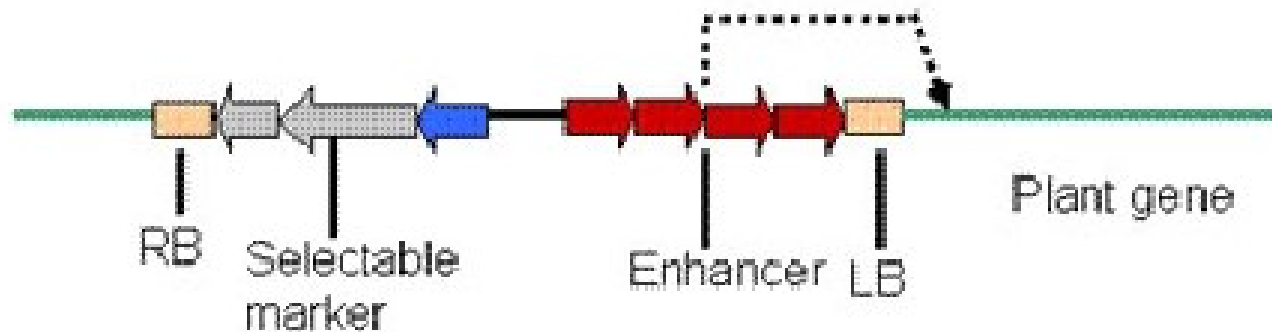
- Gene function studies can be performed also on the basis of expression patterns or **gain-of-function** phenotypes. These screens are particularly useful for genes that have redundant and multiple functions during development.
- Before 1992 there was no technique to isolate gain-of-function mutants.
- **Activation tagging technology** was originally developed by Rick Walden at the Max Planck Institute.
- This technology has made the production of gain-of-function mutants possible.

1) Generation of gain of function mutants by activation tagging

- In this technique instead loss of function in response to insertion element, gain of function takes place.
- The insertion element (T-DNA or transposon) carries an enhancer (or a a strong outward-facing promoter) and if, it integrates adjacent to an endogenous gene, the gene will be activated by the enhancer (or the promoter).
- This type of tagging causes over expression or ectopic expression of endogenous gene.



Activation TAG lines



The vector contains multimerized cauliflower Mosaic Virus (CaMV) 35S enhancer that can function in either orientation and can cause transcriptional activation of nearby genes resulting in dominant gain of function mutations.

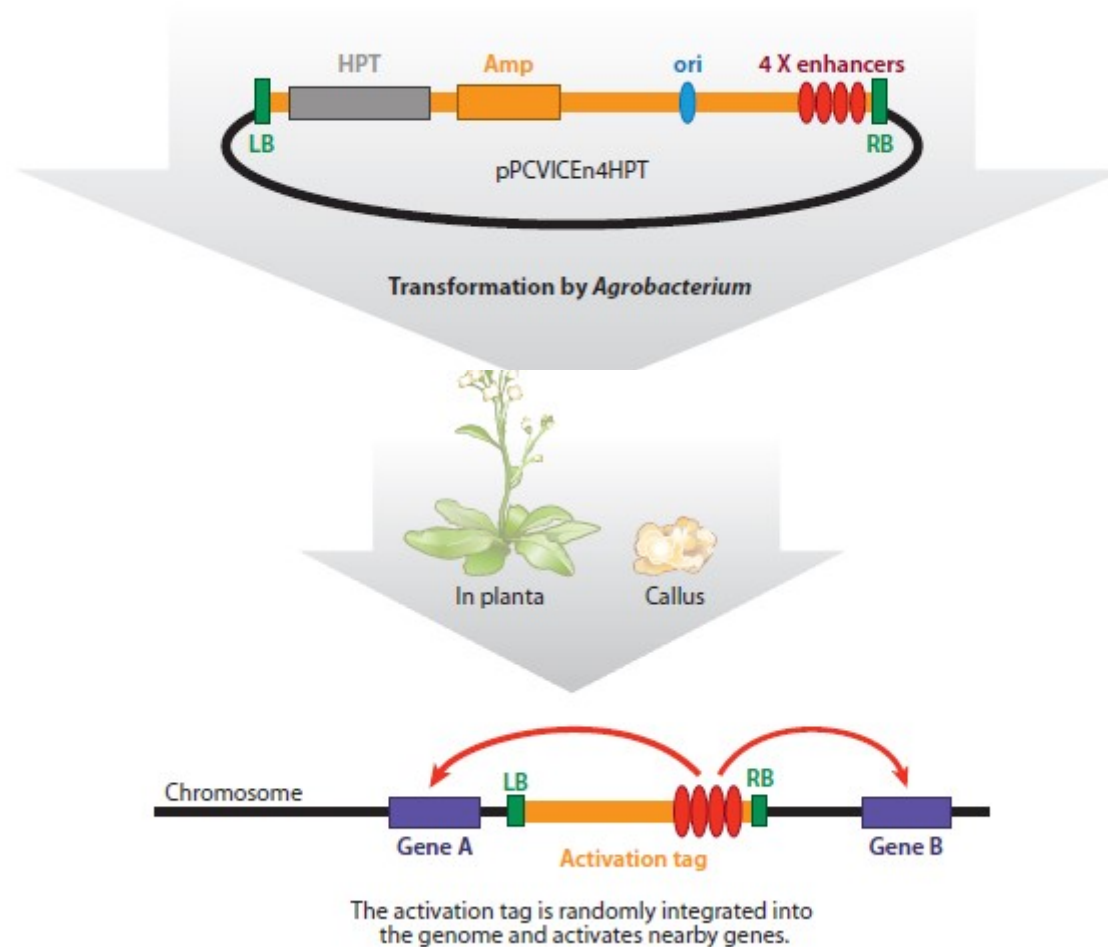


Figure 1

Overview of the activation-tagging system. The T-DNA region of the activation-tagging vector, pPCVEn4HPT, contains tetrameric *CaMV 35S* enhancers (4 X enhancers), left (LB) and right border (RB) sequences, the hygromycin resistance gene (HPT) for the plant selection marker, the ampicillin resistance gene (Amp) for bacterial selection, and the replication origin (ori) for stable maintenance in *E. coli*. Amp and ori are necessary for plasmid rescue. Plants and callus can be transformed by the activation-tagging vector through *Agrobacterium*. T-DNA (the activation tag) is randomly inserted into the plant chromosome. *CaMV 35S* enhancers in the activation tag activate the nearby genes (Genes A and B) in the activation-tagged mutant.

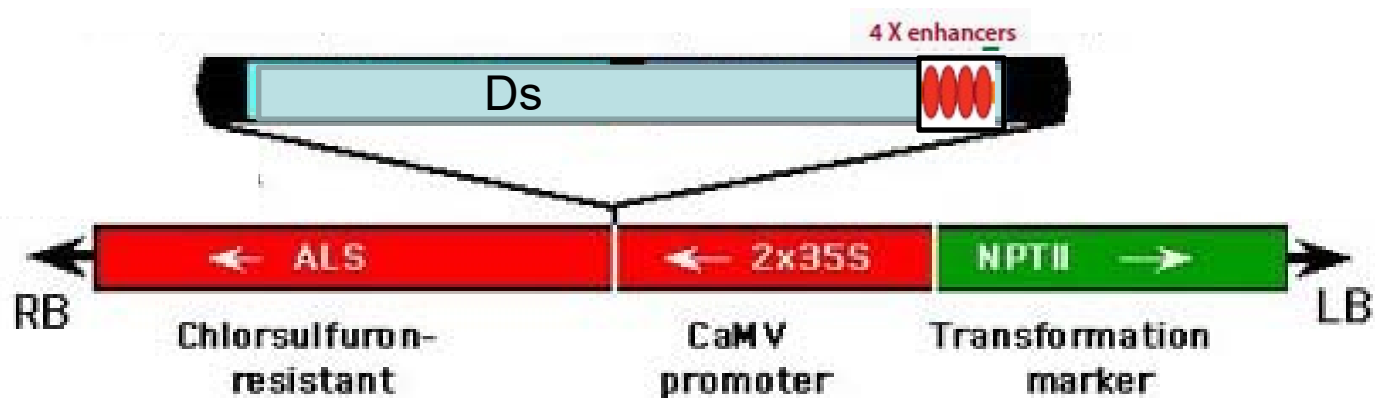
Transposon-based Activation Tagging

- Two transposons used:
 - » Ac/Ds Transposon
 - » En1 Transposon

Construct for transposon-based Activation Tagging

Figure 1: Structure of *Ac*- and *Ds*- T-DNAs

a. Transposon T-DNAs



b. Transposase T-DNAs

BamAc



Conclusion

- Using Activation tagging numerous genes that play a vital role in plant growth and development are being isolated and maintained.
- This method is extensively used for cloning several genes in all on-going plant genomics research.

2) Gain-of-function mutants generated by ectopic expression of cDNAs Driven by a Constitutive Promoter

Arabidopsis lines that overexpresses random cDNAs driven by the *CaMV* 35S promoter.

More than 30,000 *Arabidopsis* transgenic plants have been generated

Problem cDNA may not be full length

(risk of dominant-negative effects due to truncated proteins)

3) Gain-of-function mutants generated by the FOX Hunting System

The full-length cDNA over-expressing gene (FOX) hunting system is an alternative gain-of-function approach that uses fl-cDNAs

Overview of the cDNA overexpression and open reading frame (ORF) overexpression systems. (a) To generate *Agrobacterium* libraries, fl-cDNAs are cloned into an expression vector that contains the *CaMV 35S* promoter and *NOS* terminator. (b) As an alternative, ORFs can be amplified by polymerase chain reaction (PCR) and the resulting DNA fragments cloned into the same expression vector. (c) The expression vector is transformed into *Agrobacterium* and *Arabidopsis* plants transformed by floral dip infiltration with *Agrobacterium* that harbors the expression vector. T₀ plants are self-pollinated to generate independent T₁ FOX seeds. The phenotypic mutant lines (in this case, the E line) are identified from T₁ lines. The transgene of the E line is easily identified by PCR amplification with vector-specific primers and sequencing. (d) Wild-type *Arabidopsis* (1) and isolated mutants from *Arabidopsis* FOX lines (2-3) and rice FOX *Arabidopsis* lines (4-6). Abbreviations: LB, left border; RB, right border; Ter, terminator; UTR, untranslated region.

See figure in the next slide

d



Overexpression of your gene of interest
to look at the phenotype is also a
reverse genetics approach

**Usually the gene is placed under the CaMV
35S promoter**

Inducible promoter

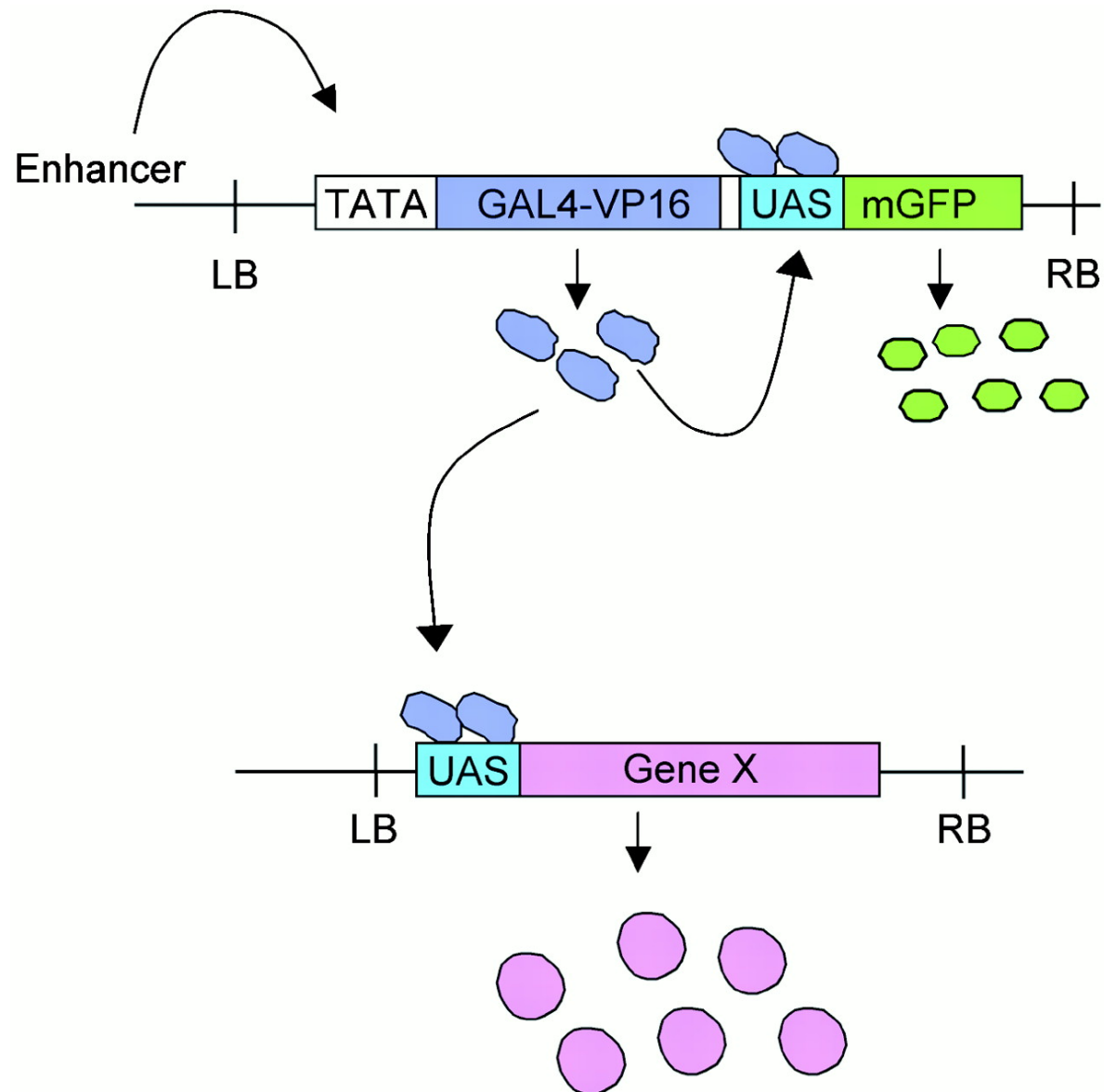
Dominant phenotype

Espressione ectopica in specifici tessuti per generare fenotipi dominanti

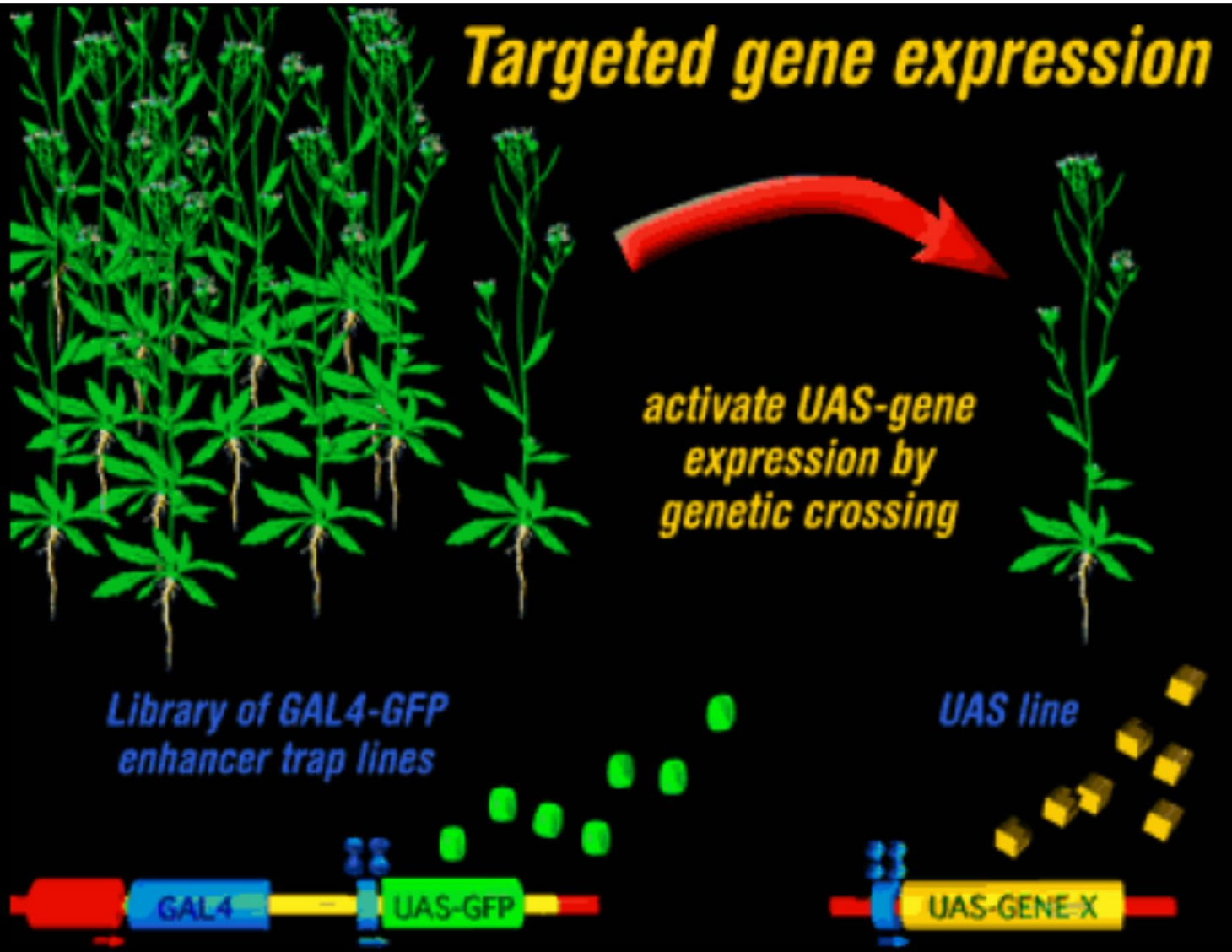
Sistema GAL4-UAS

- protocollo per attivare selettivamente l'espressione di un gene X in modo tessuto o cellula-specifico
- studio dell'effetto dovuto all'espressione ectopica di un gene X
- attivatore GAL4 e gene X sono clonati in linee diverse → si possono studiare anche i fenotipi letali
- le 2 linee vengono incrociate e il gene X verrà espresso solo in quelle cellule dove è espresso GAL4 → studio dell'effetto dell'espressione del gene X durante lo sviluppo

SISTEMA GAL4-UAS



Targeted gene expression





Starting material:
Root-specific over-expresser of GAL4:VP16 (GV)

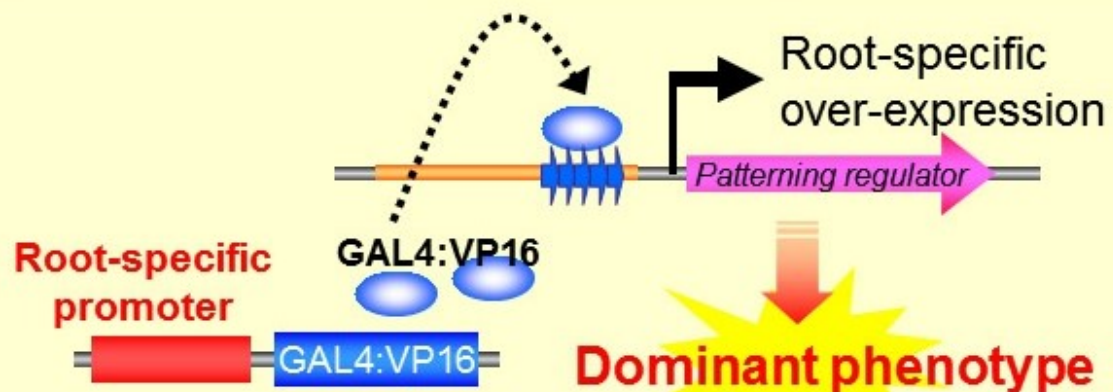


5x UAS

Agrobacterium-mediated
random integration of 5xUAS tag

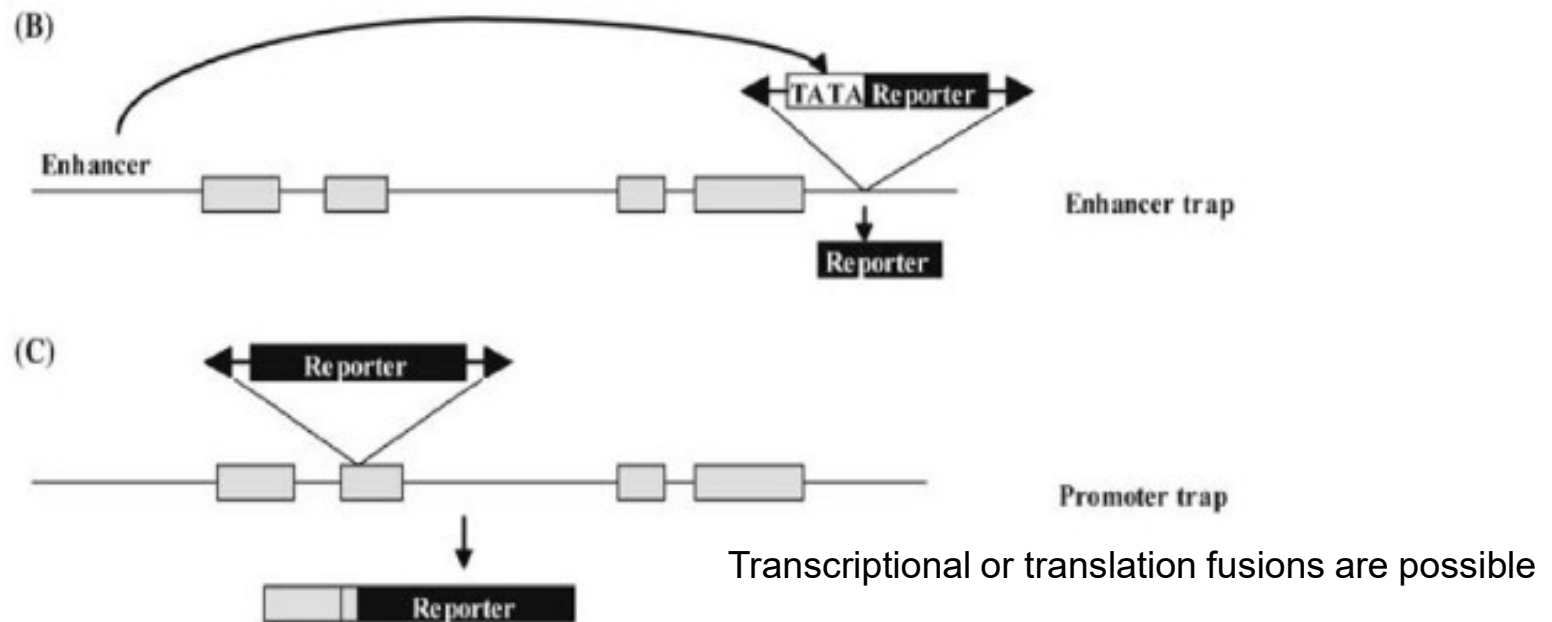
>30,000 lines
screened

Root-
specific GV
expression



Promoter/ enhancer trapping

To find genes and corresponding regulatory sequences with specific expression patterns



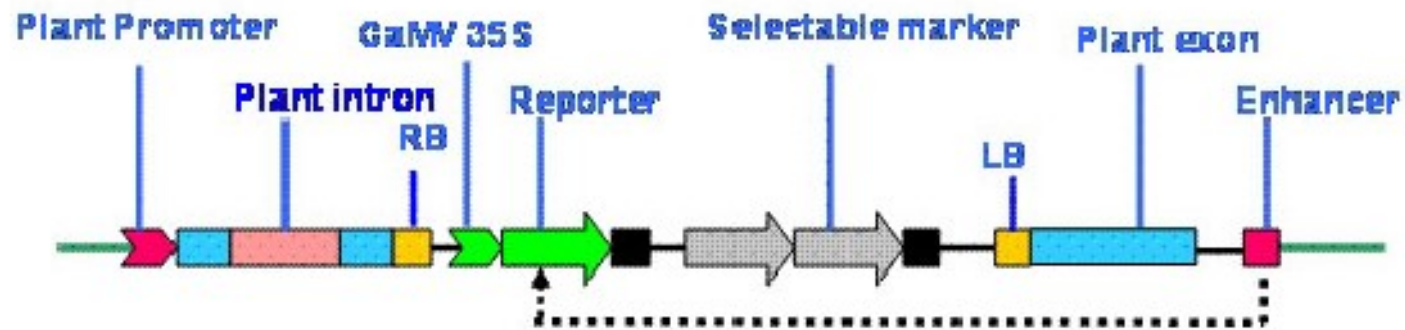
Enhancer/promoter trap vectors

-In these vectors, the insertion element contains a visible marker gene such as *Gus A* (Encoding β - *glucuronidase*) or GFP promoterless or downstream of a minimal promoter (that requires an enhancer for activity).

Note- The marker gene will be expressed only when the element insert in proximity of a transcriptional unit (minimal promoter construct) or within the transcription unit of a gene generating either a transcriptional or an in-frame translational fusion (promoterless construct).



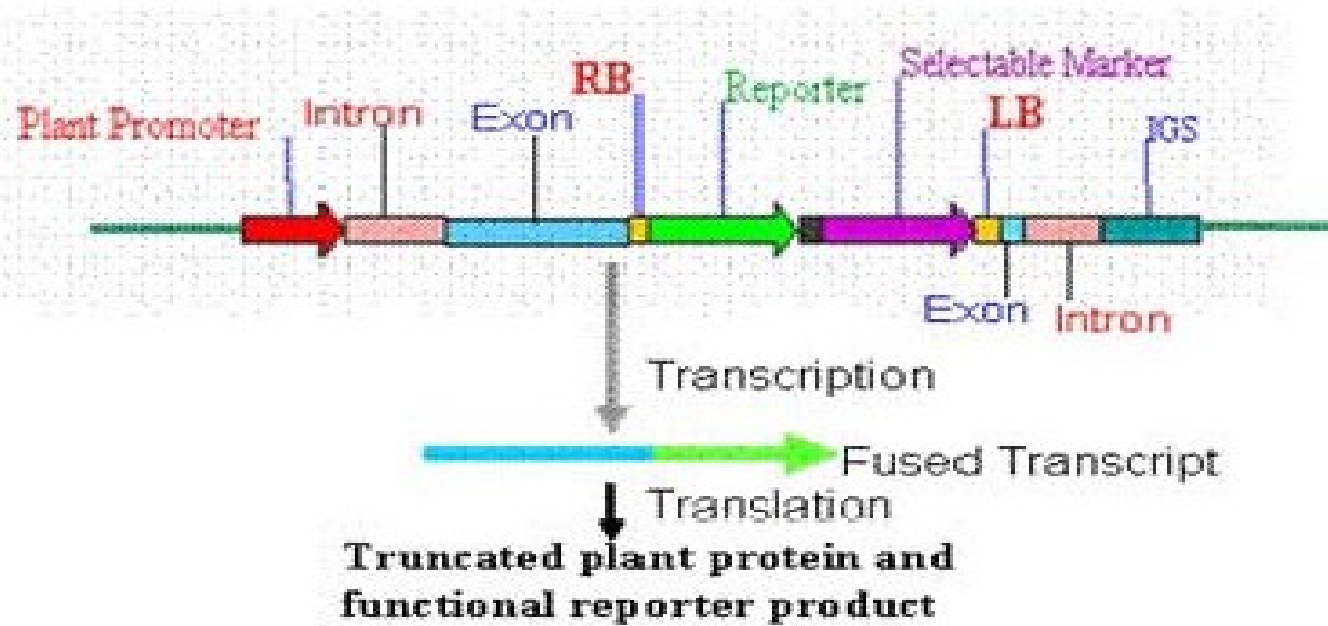
Enhancer Trap lines



The minimal promoter of the reporter gene is activated by a chromosomal enhancer element, resulting in over-expression of the reporter gene.



Promoter Trap lines



The promoterless reporter gene can be expressed when insertion of a promoter trap vector occurs in an exon so as to result in a transcriptional fusion and disruption of native gene product.

Enhancer trap in *Arabidopsis*

La tecnica dell' enhancer trap permette inoltre di identificare nuovi geni basandosi sul loro pattern di espressione.

L'analisi di linee enhancer trap di *Arabidopsis* si basa sulla selezione di linee che mostrano uno specifico pattern d'espressione e non su fenotipi mutati.

Un vantaggio di questa tecnica rispetto all'analisi di mutanti consiste nel fatto che anche se l'inserzione del T-DNA (enhancer trap) avviene in un gene essenziale la pianta sarà in quasi tutti i casi vitale essendoci nella generazione T_1 una condizione di emizigosi.

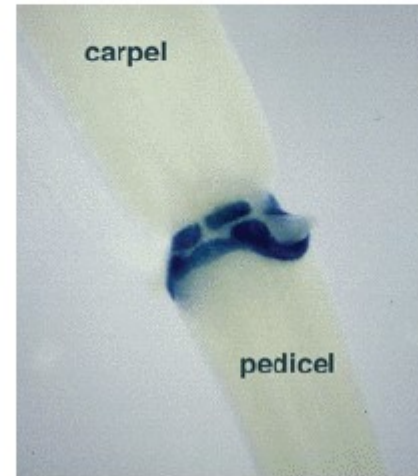
La possibilità di disporre di linee di *Arabidopsis* enhancer trap permette anche di ottenere, mediante incroci, piante in cui si può far esprimere un determinato gene in un limitato numero di cellule o tessuto.

Enhancer trap (GUS) in *Arabidopsis*

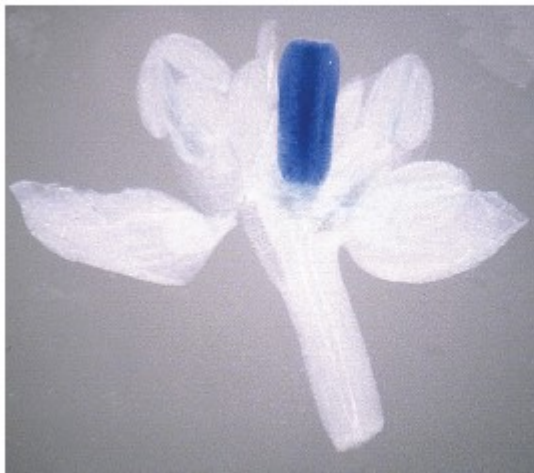
Selection of the desired mutants is based on the expression profile of the reporter gene



Espressione diffusa



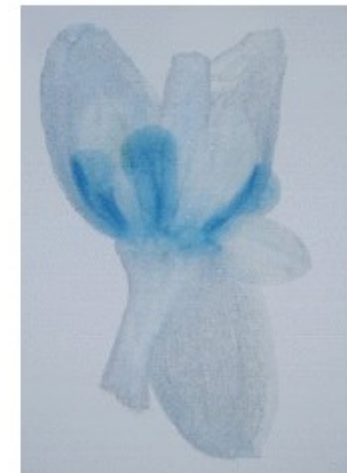
Zona di abscissione



Carpello specifica



Antera specifica



Petalo specifica

Enhancer trap (*GUS*) in *Arabidopsis*

- Primo progetto di enhancer trap su larga scala
- sono state generate più di 11,000 linee enhancer trap.
- semi di queste linee sono disponibili per la distribuzione sia al "Arabidopsis Biological Resource Center" (ABRC) presso l'università dell'Ohio sia al "Nottingham Arabidopsis Stock Center" (NASC).
- Il T-DNA con cui le piante sono state trasformate contiene il promotore minimale -60CaMV fuso al gene per la β -glucuronidasi (*GUS*).
- dopo l'integrazione, elementi enhancer che si trovano sul DNA genomico adiacente al T-DNA provocano un aumento della trascrizione a partire dal promotore minimale e si osserva un aumento del livello dell'espressione di *GUS*
- l'espressione di *GUS* risulta spesso essere cellula-, tessuto- o organo-specifica come specificato dall'enhancer. Questa espressione può essere visualizzata mediante la colorazione per il *GUS* usando il substrato cromogenico X-gluc
- 31% delle linee mostra colorazione a livello delle infiorescenze

Limitations

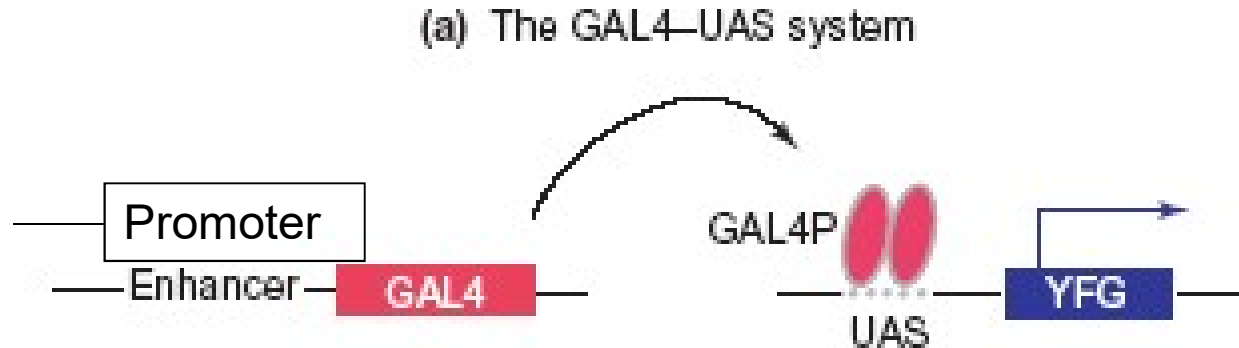
- gene trap strategy selects for insertions into genes -> is very useful in plants with large amount of non-genic DNA
- Dependency on 'in frame' insertions -> two-third of all "hits" on genes were not recognized
- expression of marker relies on the transcriptional activity of the surrounding gene, and therefore, non-expressed genes are not detected

Solution

- The use of ribosome entry sites (IRES) bypasses the need for in-frame design, it greatly increases the hit rate of gene traps.

A bipartite system for promoter/enhancer trapping

The bipartite GAL4-UAS system



- GAL4 is a transcriptional activator from yeast that recognizes a DNA sequence called the UAS (upstream activating sequence)
- The UAS to which Gal4 binds is $CGG-N_{11}-CCG$, where N can be any [base](#)
- We can use this to control expression of YFG in a tissue specific manner by using enhancer elements specific for the tissue we are interested in

GAL4-VP16 is an unusually potent transcriptional activator.

Formed by fusing a DNA-binding fragment of the yeast activator GAL4 to a highly acidic portion of the herpes simplex virus protein VP16 (ref. 11; also called Vmw65).

VP16 activates transcription of immediate early viral genes by using its amino-terminal sequences to attach to one or more host-encoded proteins that recognise DNA sequences in their promoters.

The hybrid protein (GAL4-VP16) activates transcription unusually efficiently in mammalian cells when bound close to, or at large distances from the gene.

A collection of transgenic plants have been obtained carrying T-DNA insertions containing both sequences for GAL4-based promoter/enhancer trapping and UAS-GFP



Mutants have been characterized for expression profile and made available [useful for reverse genetic approaches (see ahead)]

[NASC home](#)
[Help](#)

[Navigation](#)
[Search](#)
[Browse](#)
[Picture book](#)

[Ordering](#)
[How to order](#)
[Order by code](#)
[View selection](#)
[Order selection](#)
[Fees](#)
[FAQ / Help](#)

[Your Details](#)
[Create/Update](#)

[New charging system](#)

All single stocks are charged at £2.25 each with a fee of £8.50 per order

Haseloff GAL4-GFP Enhancer-trap lines

These lines have been produced, and donated, by Dr. Jim Haseloff, MRC Laboratory of Molecular Biology, Hills Road Cambridge, UK.

Ordering lines

Four sets of GAL4-GFP lines are now available. The sets correspond to those mentioned on [Jim Haseloff's web site](#) and fall into the following categories.

- Set 1 and 1a Lines *pre-screened for root expression* by Jim Haseloff and Sarah Hodge.
- Set KS Lines *pre-screened for shoot and floral expression* by Kirby Siemering and Marion Bauch.
- Set M Lines *prescreened for shoot and floral expression* by Marion Bauch.

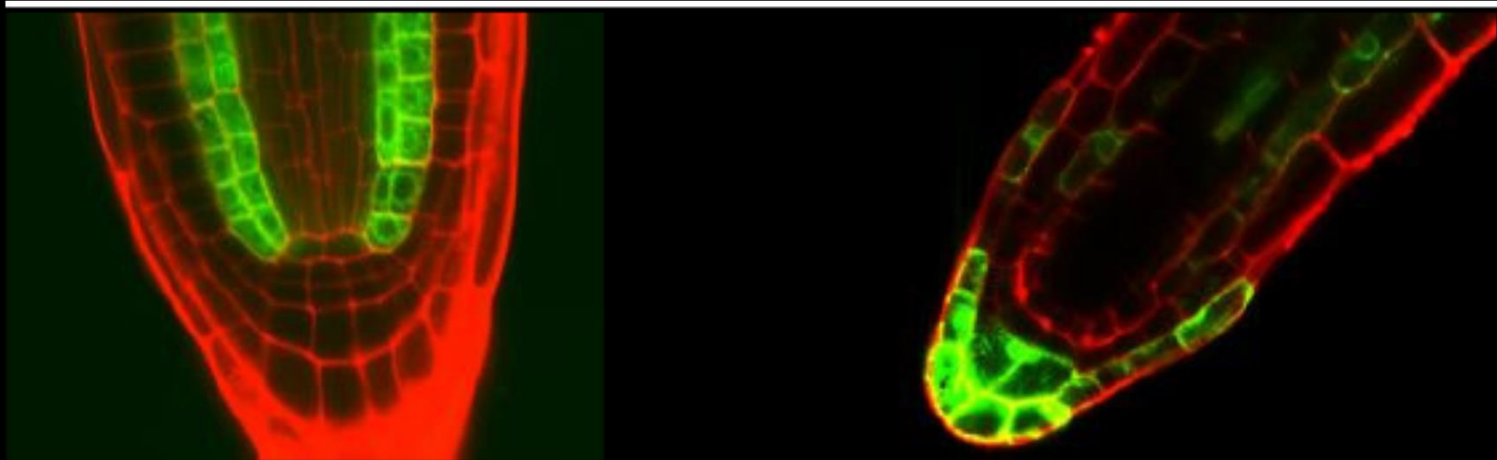
Individual lines may also be ordered, to do so follow the details link in the table below.

For detailed descriptions of these lines, and information relating to these lines, please refer to [Jim Haseloffs' web page](#).

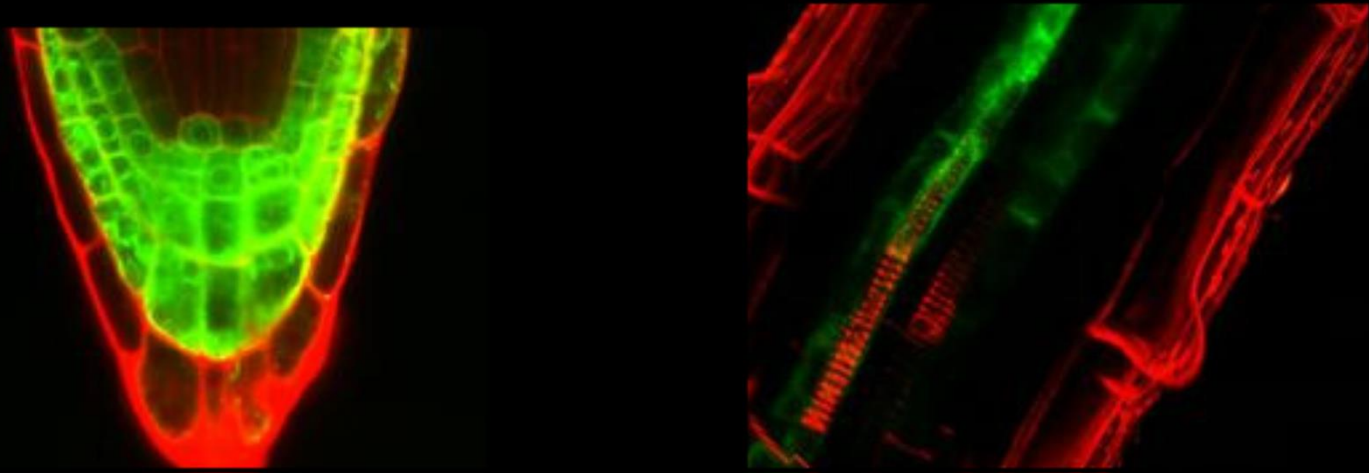
- [About the GFP lines](#)

Category	Stock No.	Price £	
GAL4-GFP Enhancer trap lines - set 1 (116 lines) details	N9089	122	<input type="checkbox"/>
GAL4-GFP Enhancer trap lines - set 1a (29 lines) details	N9340	33	<input type="checkbox"/>
GAL4-GFP Enhancer trap lines - set KS (31 lines) details	N9341	36	<input type="checkbox"/>
GAL4-GFP Enhancer trap lines - set M (70 lines) details	N9342	82	<input type="checkbox"/>

Add to Selection



Immagini prese al microscopio confocale di apici radicali di differenti linee di *Arabidopsis*



REVERSE GENETICS

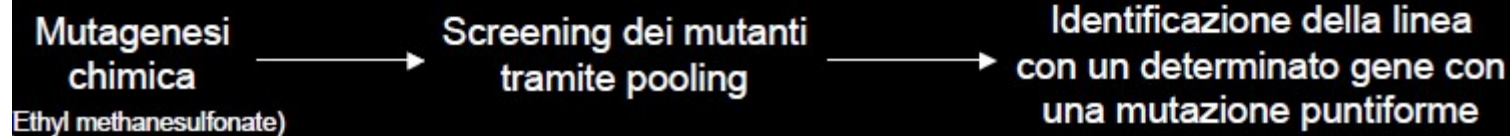
ALTRI APPROCCI

1. Come faccio a studiare i fenotipi letali?
2. Come posso studiare i fenotipi associati al mio gene di interesse se per questo non è ancora disponibile un mutante inserzionale?

- TILLING
- ANTISENSO
- RNAi

Il **T-DNA** ed i trasposoni **NON** sono facilmente applicabili a tutte le specie vegetali

TILLING (Targeting Induced Local Lesions IN Genomes)



Vantaggi:

- E' una procedura semplice e più ampiamente utilizzabile
- Non necessita di colture cellulari e non genera linee transgeniche
- Produce **serie alleliche** con linee K.O. o solamente "attenuate"

L'endonucleasi Cel-I identifica i mismatch dell'appaiamento tra wt e mutante

Primer gene specifici
A →
← B

A e B due marcatori differenti

PCR sul wt e
nel mutante

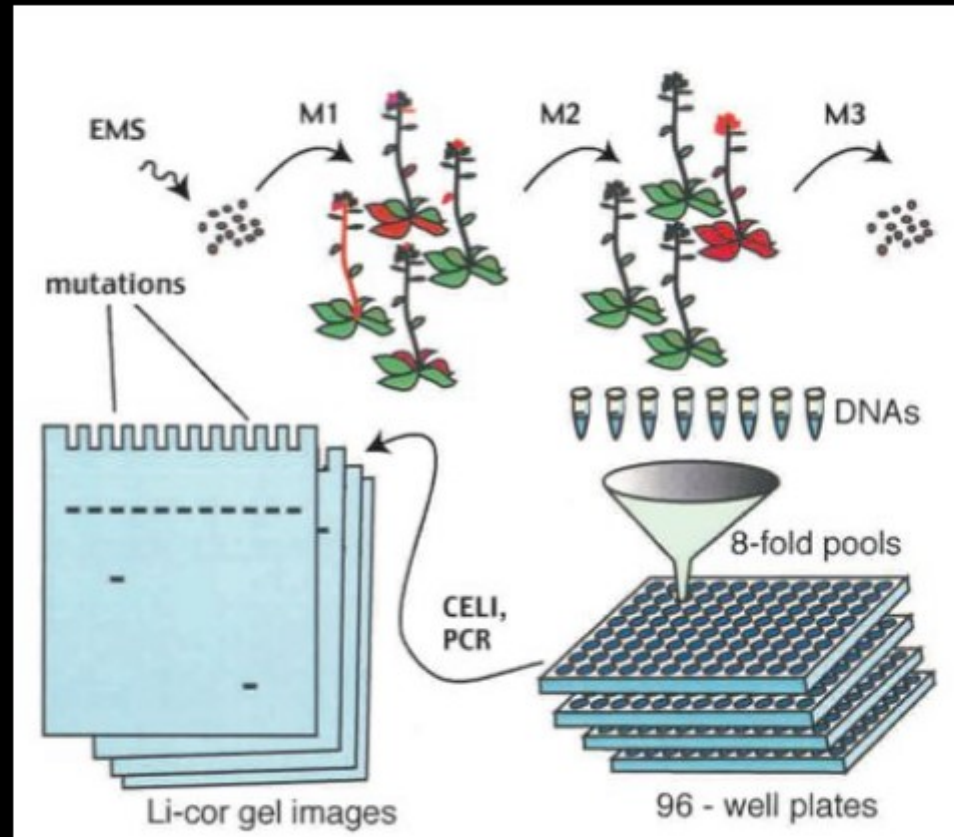


cicli di appaiamento dei
frammenti amplificati

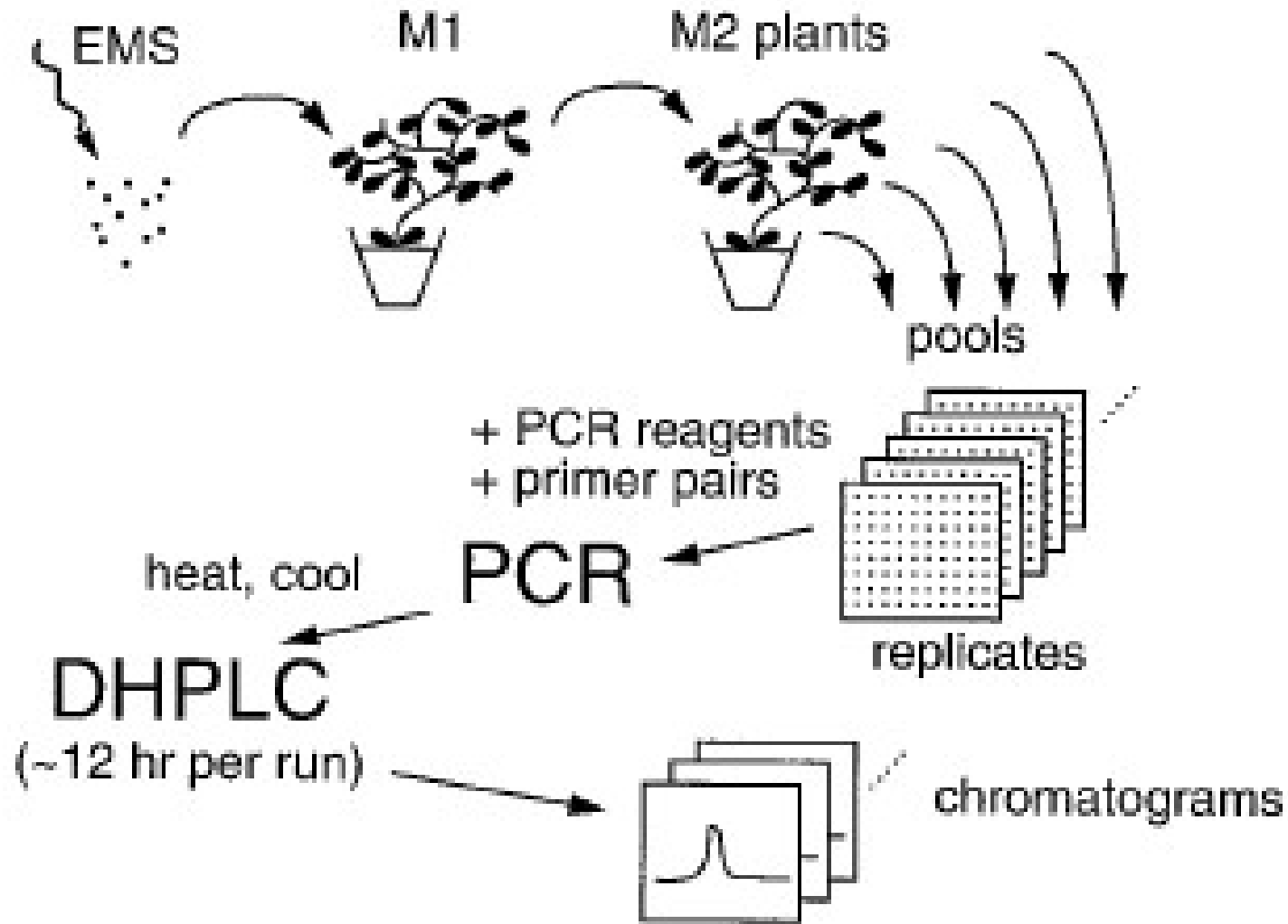
Heteroduplex



high-throughput TILLING - how it works

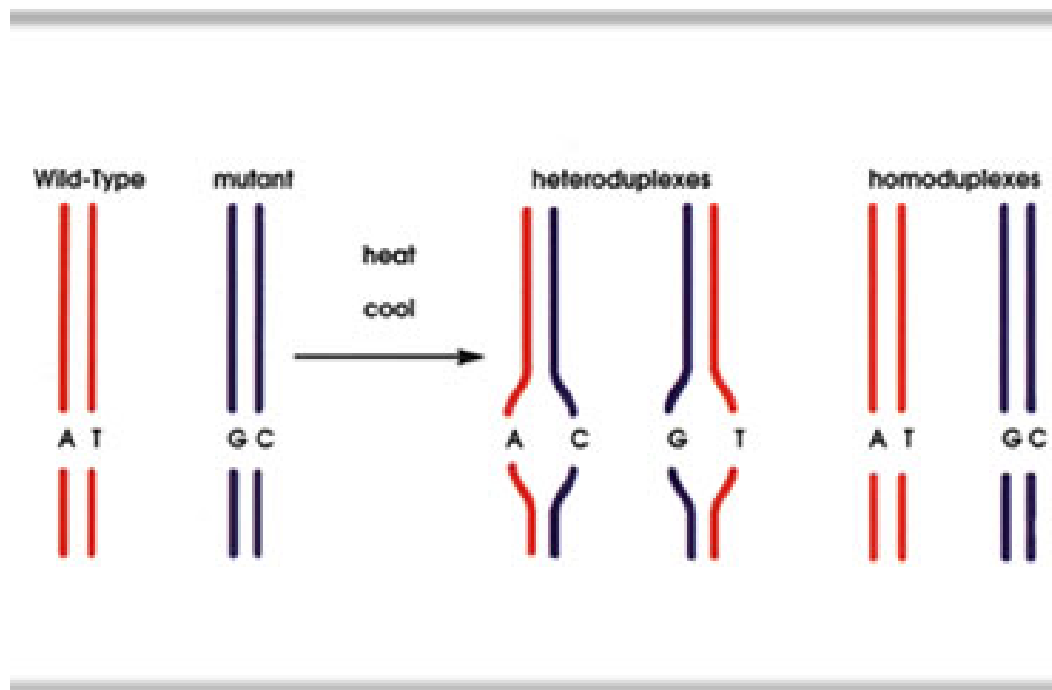


Targeting Induced Local Lesions IN Genomes (TILLING)

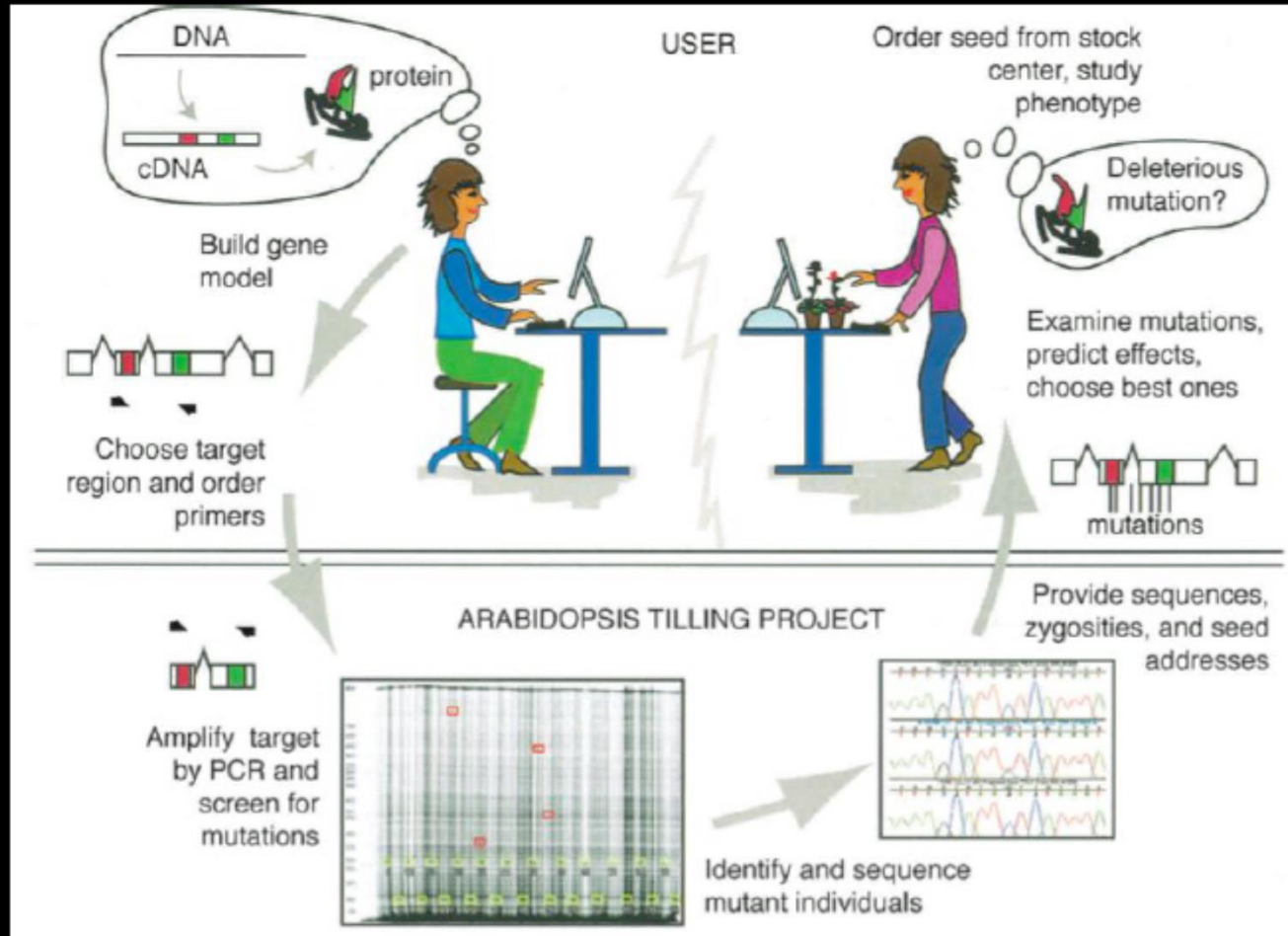


DHPLC= HPLC denaturante

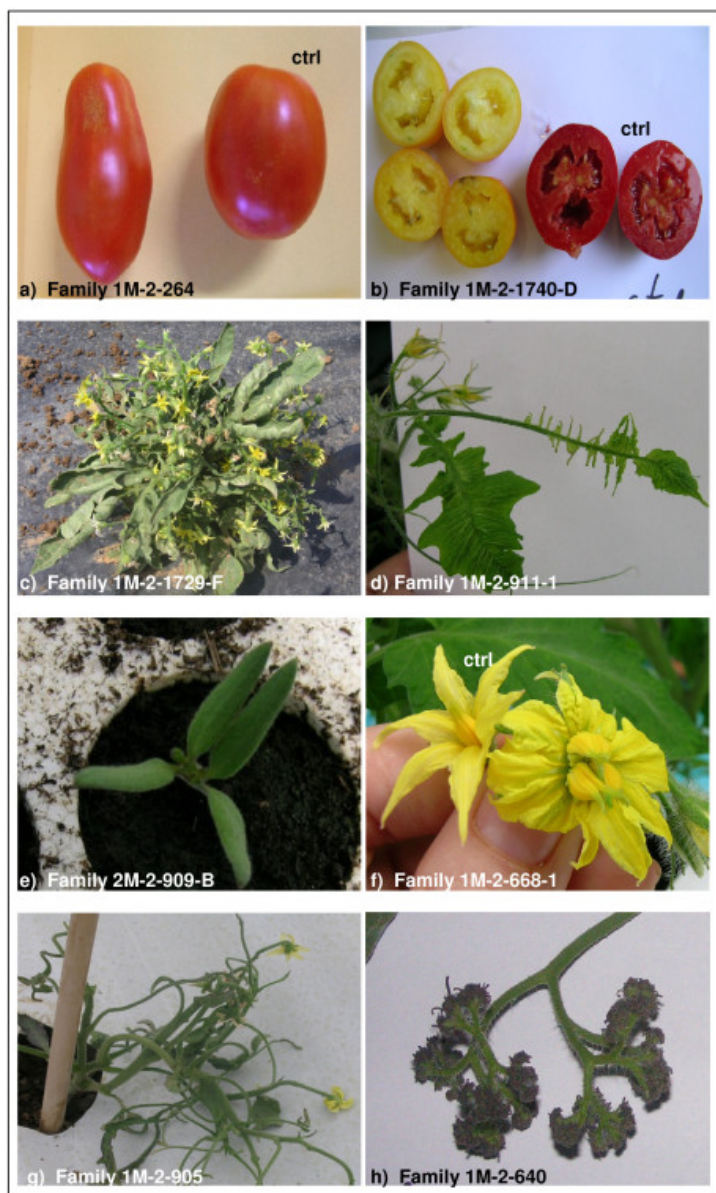
- Duplex si formano quando un frammento amplificato di DNA mutato ed uno non mutato vengono denaturati termicamente e lasciati ricombinare.
- Su una colonna cromatografica, l'eteroduplex è solitamente più veloce (meno trattenuto) dell'omoduplex
- Può essere impiegata per rilevare ogni tipo di mutazione (SNPs, inserzioni, delezioni e tandem repeat)



high-throughput TILLING - how it works



TILLING in pomodoro

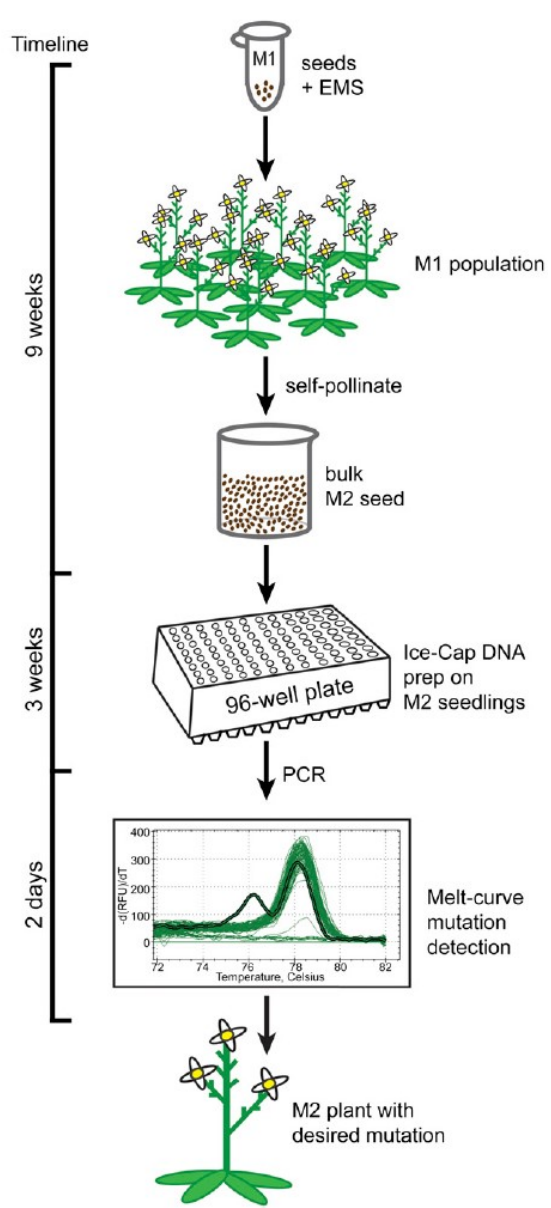


Mutation density in 0.7% EMS and 1% EMS Red Setter populations

Target gene	No. of screened M3 families	No. of identified mutations		Overall mutation density			
		0.7% EMS	1% EMS	0.7% EMS	1% EMS		
Name	Amplicon size (kb)	0.7% EMS	1% EMS	0.7% EMS	1% EMS	0.7% EMS	1% EMS
<i>Rab11a</i>	0.407	1,373	713	1	3	1/559 kb	1/97 kb
<i>PG</i>	2.587	2,791	963	7	2	1/1031 kb	1/1246 kb
<i>Exp1</i>	1.025	3,885	1,284	14	6	1/284 kb	1/219 kb
<i>RIN</i>	1.331	3,885	1,284	4	8	1/1293 kb	1/214 kb
<i>Gr</i>	1.409	3,885	1,284	5	3	1/1095 kb	1/603 kb
<i>Lcy-b</i>	1.274	3,801	1,252	4	3	1/1211 kb	1/532 kb
<i>Lcy-e</i>	1.414	3,630	1,185	6	0	1/855 kb	-
Total/mean	9.447			41	25	1/574 kb	1/322 kb

The accession numbers of the analyzed seven target genes are the following: *Rab11a* [GenBank:[AJ245570](#)], *PG* [GenBank:[M37304](#)], *Exp1* [GenBank:[AF548376](#)], *RIN* [GenBank:[AF448522](#)], *Gr* [GenBank:[DQ372897](#)], *Lcy-b* [GenBank:[CQ788383](#)], *Lcy-e* [GenBank:[Y14387](#)]. The number of screened M3 families, the number of identified mutations and the overall mutation density, estimated as described in Methods, are reported both for 0.7% and 1% EMS Red Setter populations.

Minoia *et al.* *BMC Research Notes* 2010 **3**:69 doi:10.1186/1756-0500-3-69



iTILLING: A Personalized Approach to the Identification of Induced Mutations in Arabidopsis¹[C][OA]

Susan M. Bush and Patrick J. Krysan*

Department of Horticulture (S.M.B., P.J.K.) and Genome Center of Wisconsin (P.J.K.), University of Wisconsin, Madison, Wisconsin 53706

Plant Physiology[®], September 2010, Vol. 154, pp. 25–35,

