Methods for Reverse genetics

References:

1. Alonso JM, Ecker JR. Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in Arabidopsis. Nat Rev Genet. 2006 Jul;7(7):524-36.
2. Waterhouse and Helliwell. Exploring plant genomes by RNA-induced gene silencing. Nat Rev Genet. 2003 4(1): 29-38
3. Krysan, Young, and Sussman. T-DNA as an insertional mutagen in Arabidopsis. Plant Cell 1999 11(12): 2283-90
4. Shan et al. 2013. Targeted genome modification of crop plants using a CRISPR-Cas system. Nat Biotechnol. 31(8):686-8.
5. Doudna and Charpentier E. 2014. The new frontier of genome engineering with CRISPR-Cas9. Science. 346(6213):1258096-1.

Genetic analysis
It's all about mutants and their phenotypes!

Forward genetics

- From mutant phenotype to gene, from gene to protein function

Reverse genetics

- From gene to mutant phenotype, to function

Reverse genetics

What kinds of manipulation do we usually do to our favorite gene (FG) ?

Knockout analysis:

1. **Find/Generate a knockout mutant in FG**
2. **Analyze the mutant to see if there is any defects**
3. **Connect the defects with biological processes**

Reverse genetics

Overexpression analysis/ectopic expression

1. Overexpress FG (endogenous promoter)
Ectopic expression (CaMV 35S promoter)

2. Analyze the overexpresser to see if there are any defects/phenotypes

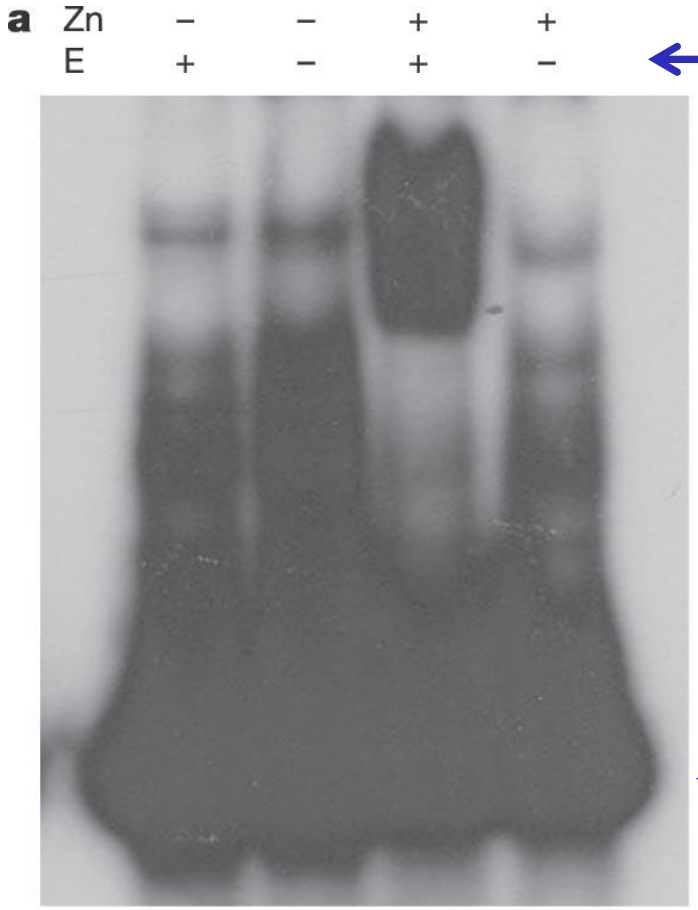
3. Connect the defects with biological processes

Over-expression of a gene of interest does not necessarily lead to a gain-of-function effect.

Why?

Story on *DAYSLEEPER*

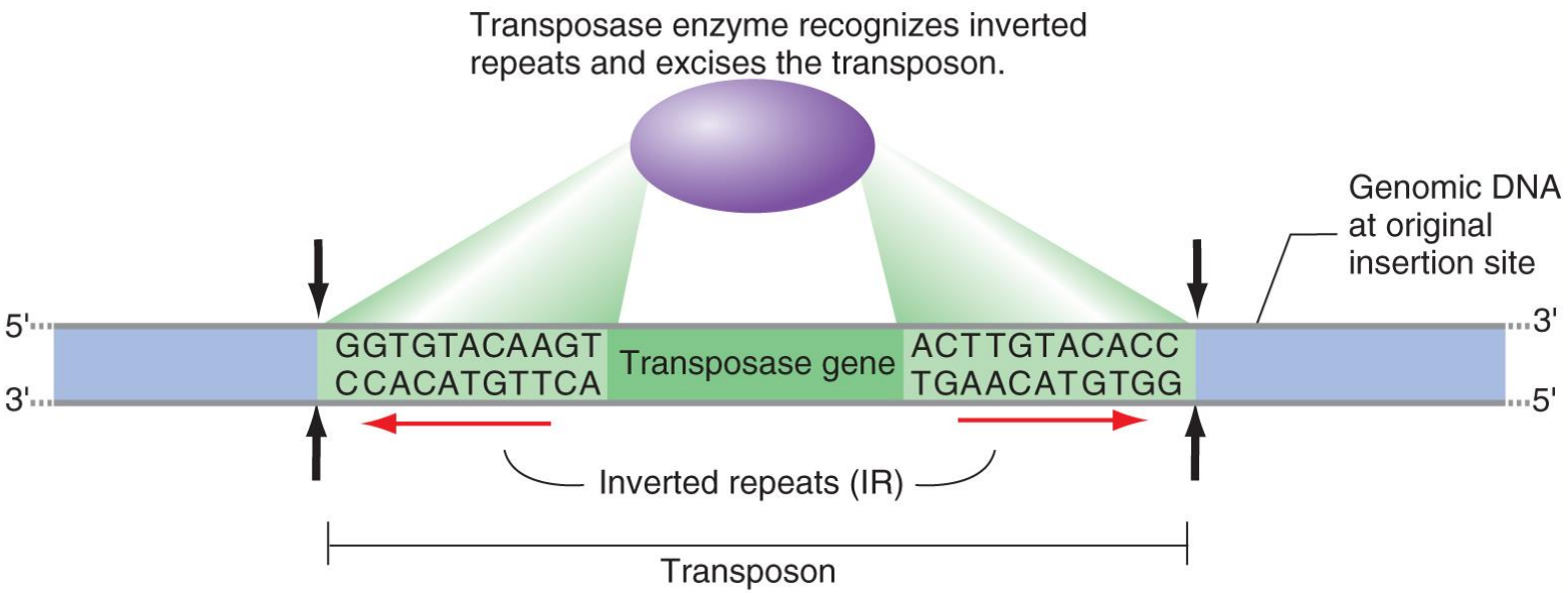
Background: DAYSLEEPER was found to bind to the promoter region of *Ku70*, which encodes a protein involved in **DNA repair**.



← DAYSLEEPER

← 25 bp Kubox motif

DAYSLEEPER encodes a hAT-like transposase



hAT-like transposon elements

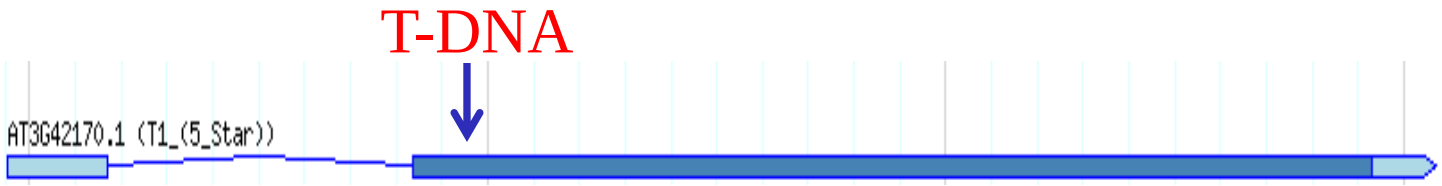
- In Arabidopsis, there are 246 hAT-like elements.
- Active hAT transposons: 8bp duplication of the insertion site and short terminal inverted repeats (TIR).
- Fossil elements lack the duplication of the insertion site and TIR, and often are transcriptionally silent.
- *Daysleeper* is a fossil element, but is expressed.

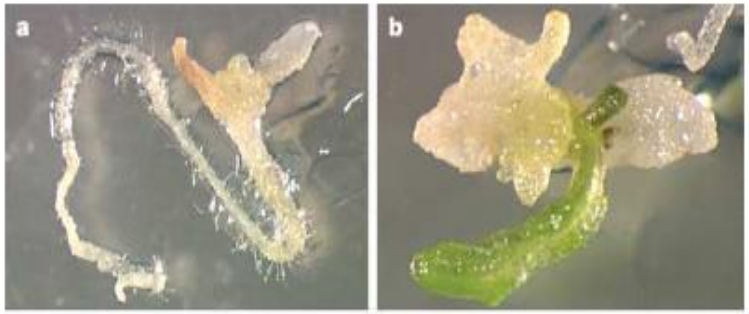
Q: DNA repair -----????----- Transposon

How do you find out the function of *DAYSLEEPER*?

Q: What would a mutant of a fossil transposable element look like?

DAYSLEEPER knockout mutant





Q: How would you prove that the mutant phenotype is caused by the mutation in *DAYSLEEPER*?

- **co-segregation**
- **transgene complementation**
- **obtaining second allele**

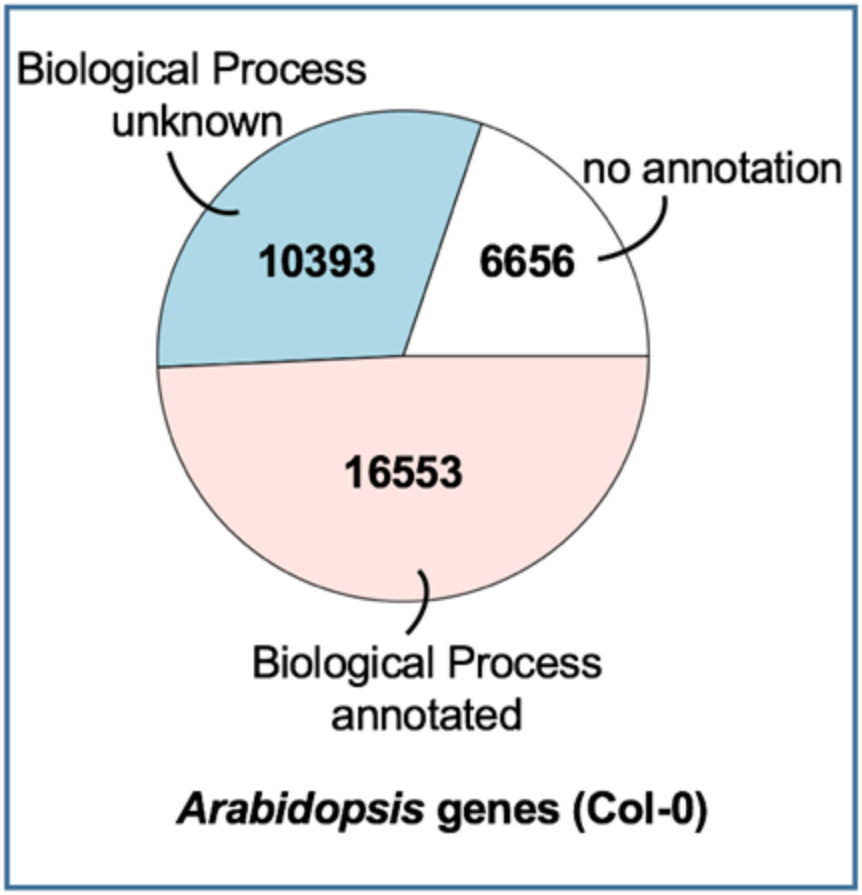
DAYSLEEPER over-expressers
Slower growth, delayed flowering, altered leaves, etc.

DAYSLEEPER is essential for plant development.

Fossil elements are not always fossils, they have the potential to evolve functions essential for plant growth and development.

Bundock P, Hooykaas P. An Arabidopsis hAT-like transposase is essential for plant development. Nature. 2005 Jul 14;436(7048):282-4.

Why do we need reverse genetics?



2018 Arabidopsis Annotated Gene Function

Why cannot forward genetics find mutants carrying mutations in all genes?

1. **Redundancy**
2. **Lethal mutations**
3. **Subtle or not obvious phenotypes**
4. **Mutant missed from forward genetic screens**

Why do we need reverse genetics?
(When to use reverse genetics?)

1. **Figure out function of YFG**
2. **Redundant genes**
3. **Essential genes**
4. **Assist forward genetics: second allele**

Why do we still do forward genetics?

- **Process more specific, it is less predictable in reverse genetics**

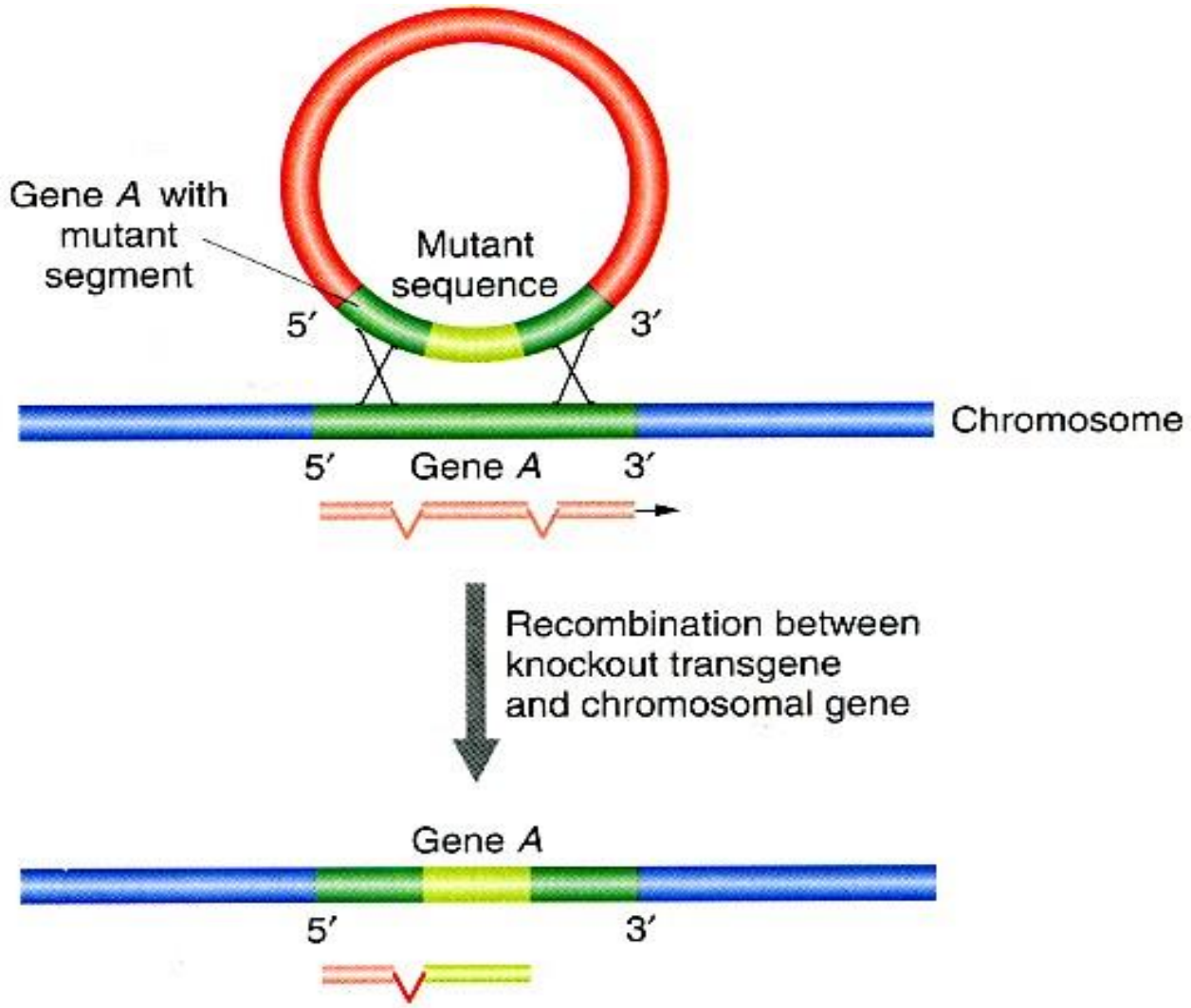
- **No previous knowledge needed for forward genetics**

- **Suppressor/enhancer screens that may lead to new insights**

With complete genome sequence information, we can pick and study our favorite genes by reverse genetics.

How do we knockout genes in plants?

Homologous recombination



Not for plants!

Reverse genetics:
Gene Knockout Strategies

1. RNAi-based silencing

2. T-DNA or transposon based insertional mutagenesis

3. CRISPR-CAS9-based gene editing

RNAi based methods

History:

Early 1990's, phenomena first found by plant scientists: co-suppression

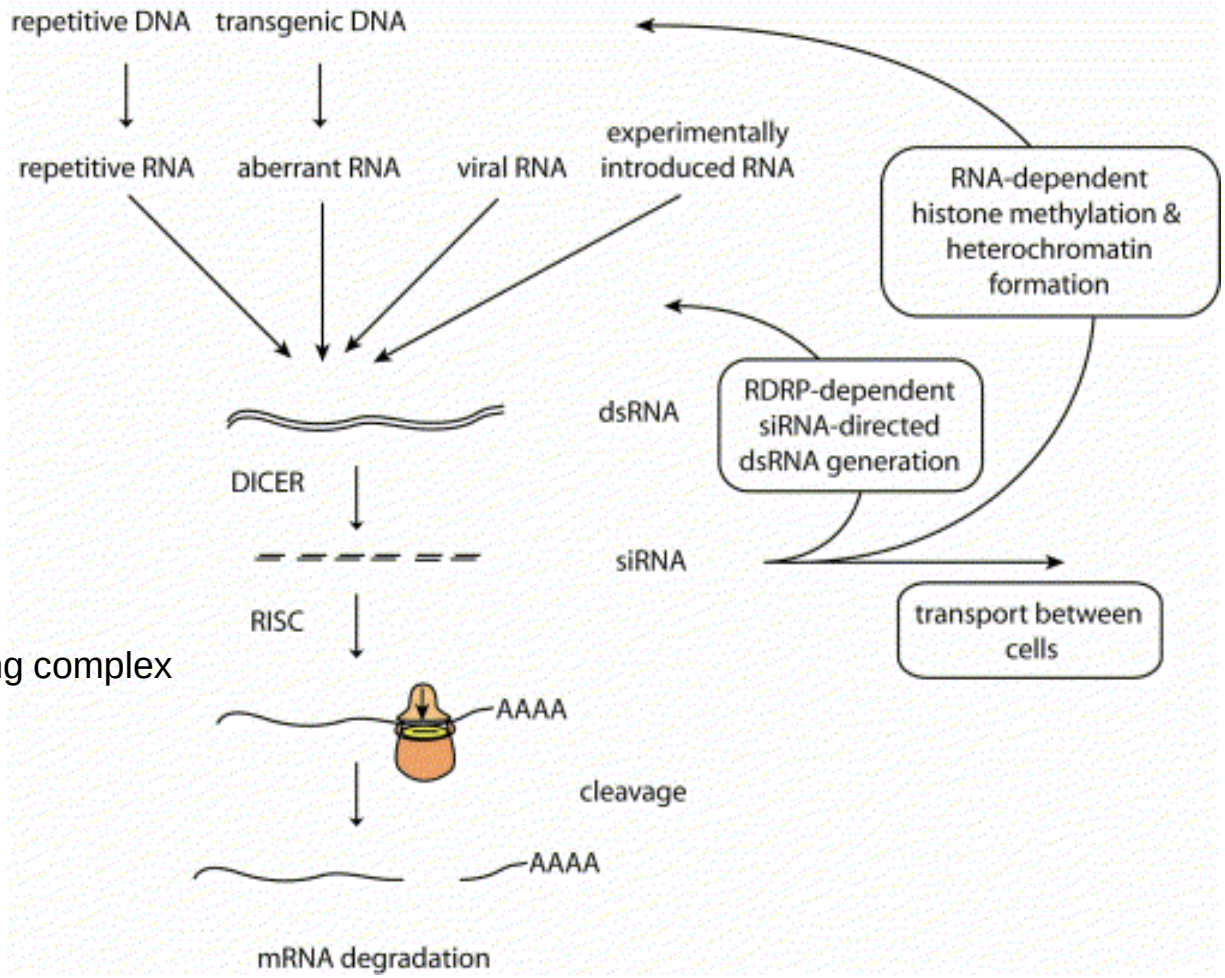
1998, in *C.elegans*, formally discover dsRNA as signal for RNA interference (Fire and Mello)

1999, small RNA species derived from mRNA detected (Baulcomb)

2001, discovery of dsRNA processing enzyme Dicer

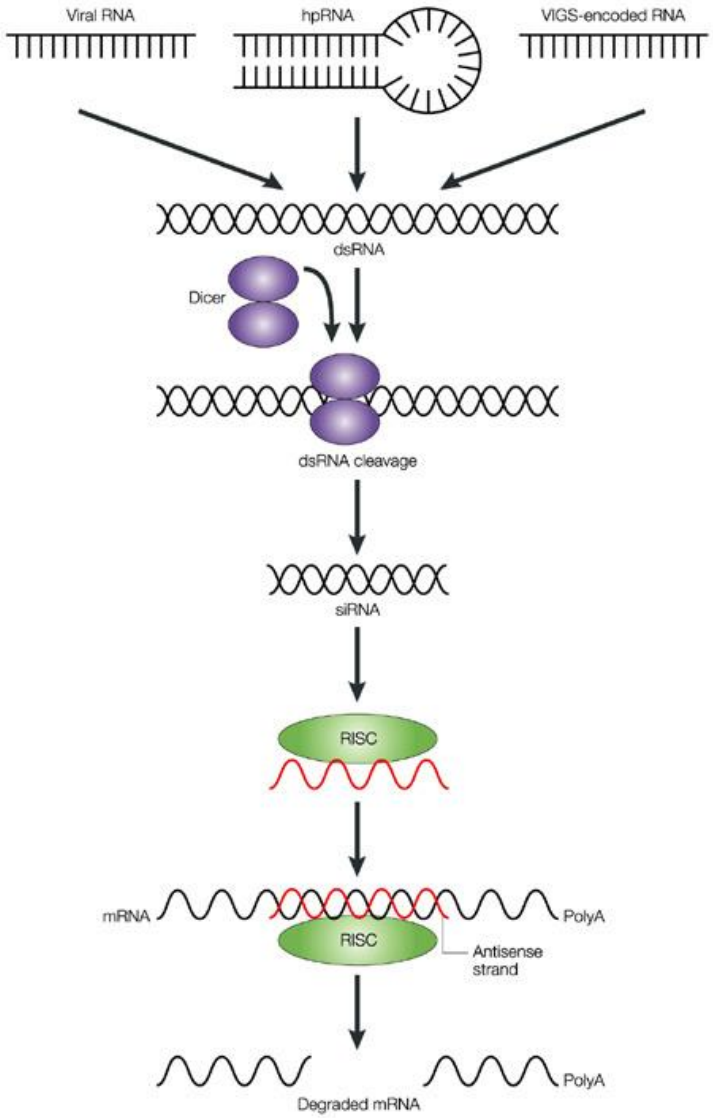
2006, A. Fire and C. Mello won Nobel prize in medicine because of their discovery of dsRNA as mediator of RNAi

RNAi: an ancient immune response against invasion of viruses and other genetic materials



RISC:
RNA-induced silencing complex

dsRNA-directed gene silencing mechanisms. Short dsRNA molecules can either be expressed by endogenous genes, invading viruses or by experimental means and are funnelled into one of two different silencing mechanisms. siRNAs that are perfectly complementary to their cognate mRNA species induce their endonucleolytic cleavage and degradation. Amplification of the RNAi signal by RDRP-dependent mechanisms, RNA-induced epigenetic control of gene expression as well as RNAi transfer between cells have been observed in some but not all species.



Nature Reviews | Genetics

Hairpin RNA-induced gene silencing

A typical T-DNA plasmid for the expression of hairpin RNAs (hpRNAs). A generic silencing precursor construct (pHANNIBAL) that enables hpRNA vectors to be easily constructed has different multiple cloning sites either side of the intron to enable the rapid insertion of target sequences in forward and reverse orientations. 35S, CaMV 35S promoter; Term, transcription termination sequence.

Silencing of the phytoene desaturase gene in Arabidopsis by hairpin RNA

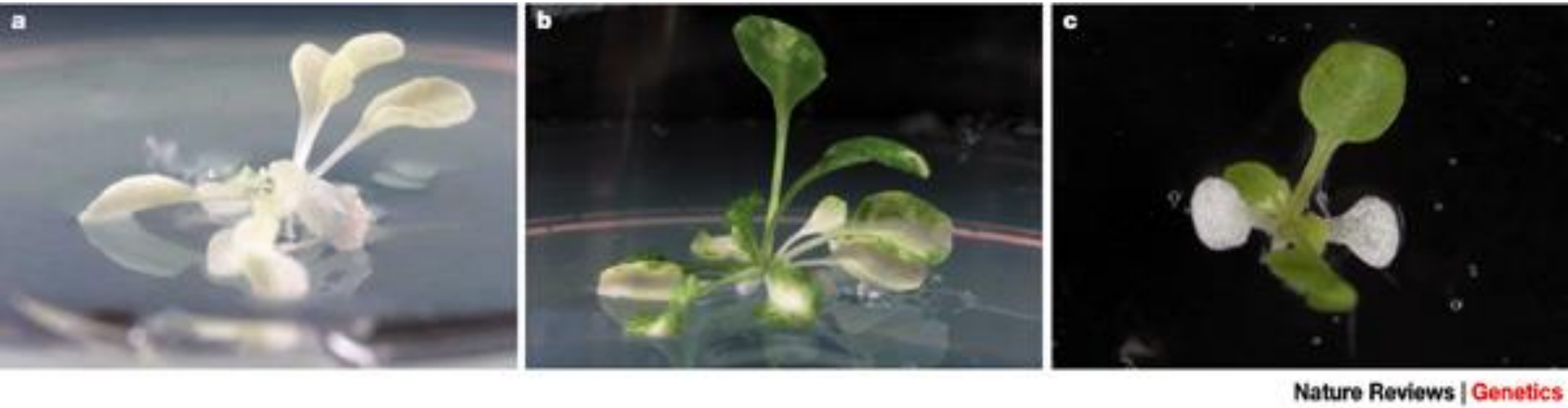
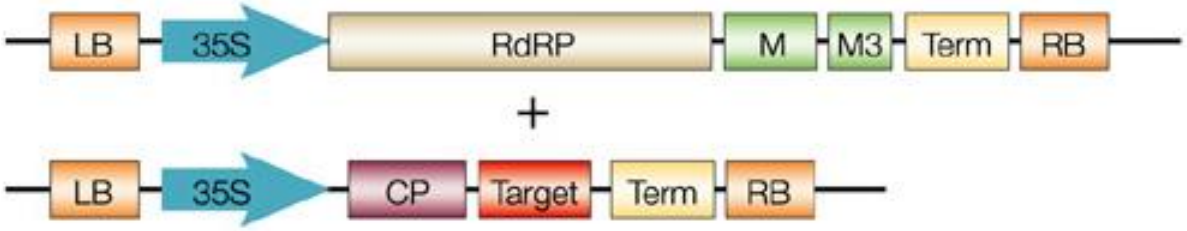


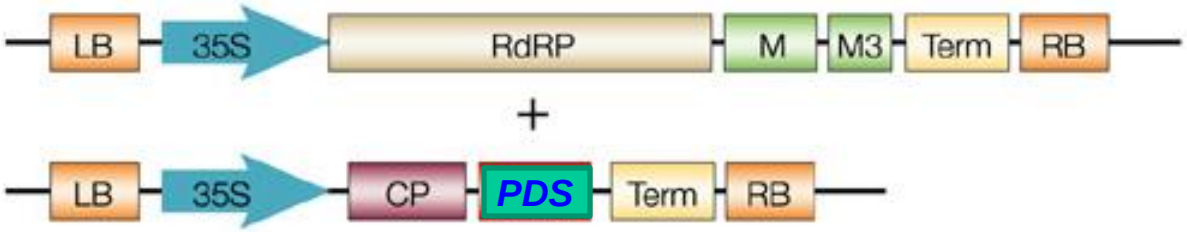
Figure 5 | **Degrees of silencing produced by hairpin-RNA-encoding transgenes.** The stable transformation of *Arabidopsis* plants with the same hairpin RNA (hpRNA) construct that is targeted against phytoene desaturase gives rise to lines that show a heritable photobleaching phenotype in: **a** | all tissues; **b** | sectors of tissue; or **c** | the cotyledons, but not the rest of the plant. Images courtesy of C.A.H. and P.M.W., CSIRO, Australia. Reproduced with permission from Ref. 63 © (2002) CSIRO Publishing.

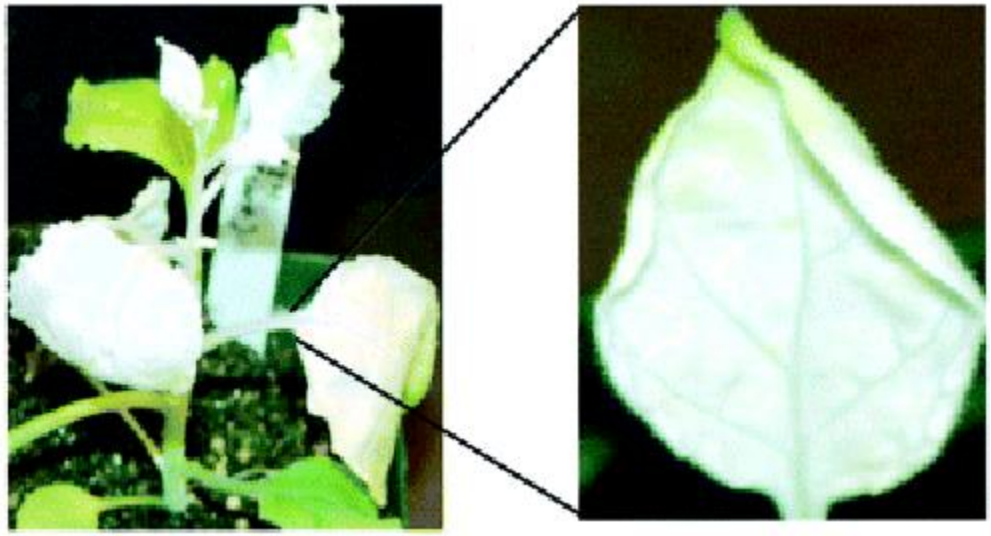
Virus-induced gene silencing (VIGS)



The tobacco rattle virus (TRV) virus-induced gene-silencing (VIGS) system. Two T-DNA plasmids that encode the TRV genome (one encoding TRV RNA1 and the other encoding TRV RNA2, which carries the inserted target sequence) are propagated separately in *Agrobacterium* and used to co-infect plant tissue. 35S, CaMV 35S promoter; CP, coat protein; M1,2,3, movement proteins 1, 2, 3; RdRP, RNA-dependent RNA polymerase; Term, transcription termination sequence.

Silencing of the phytoene desaturase gene in tobacco by TEV-based VIGS





Liu, *et al*. 2002. Plant Journal, 30: 415–429.

transient- vs. stable-integrated gene-silencing

	Advantages	Disadvantages
Viral-induced Gene silencing (VIGS)	● Rapid ● easy to use ● applicable to mature plants ● useful for species hard to generate transgenic plants	● Host range limitations ● restricted regions of silencing ● viral symptoms superimposed on silencing phenotype
Hairpin transgenes	● Not restricted by host range ● controllable tissue specificity ● range of degrees of silencing	● Require Transformation

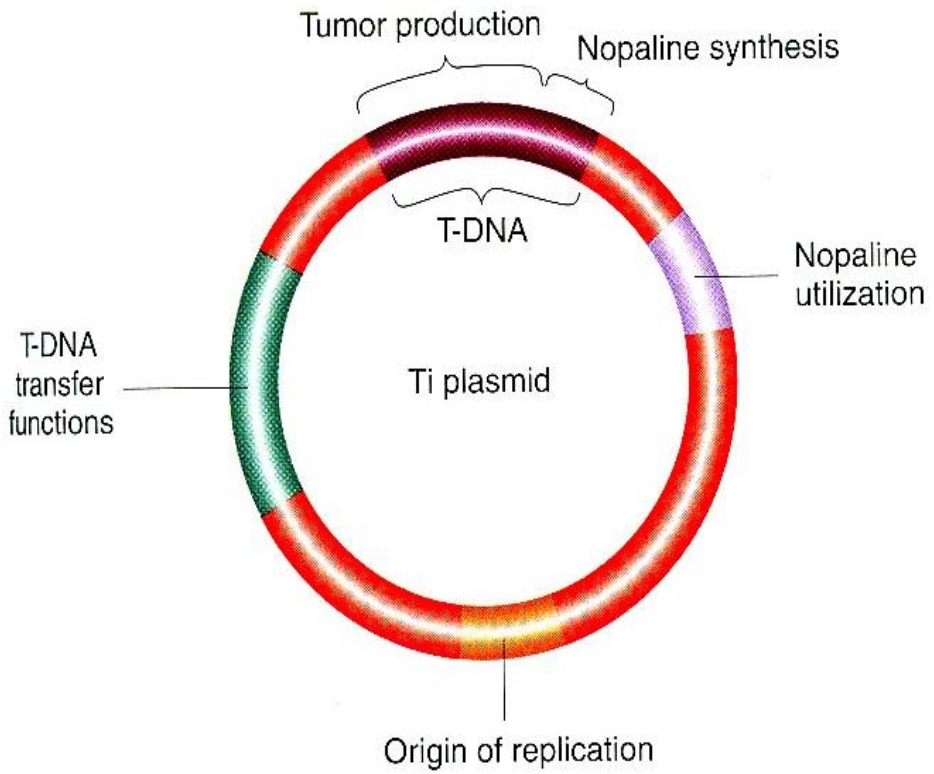
Reverse genetics:
Gene Knockout Strategies

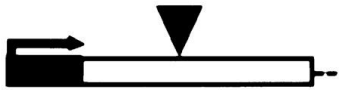
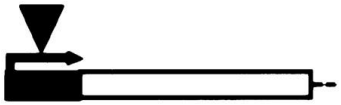
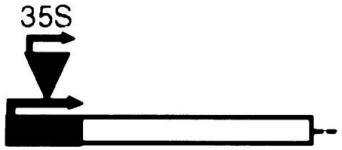
1. RNAi-based silencing

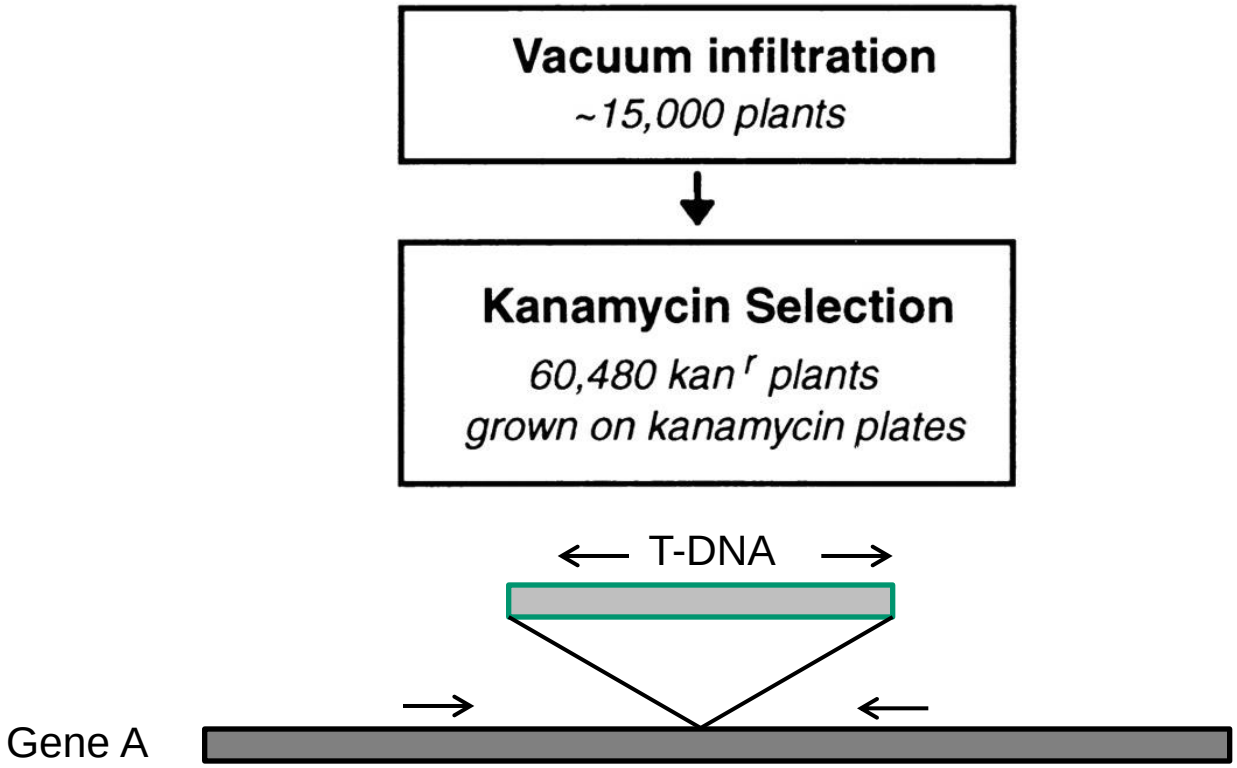
2. T-DNA or transposon based insertional mutagenesis

3. CRISPR-CAS9-based gene editing

Agrobacterium Ti plasmid-based transformation

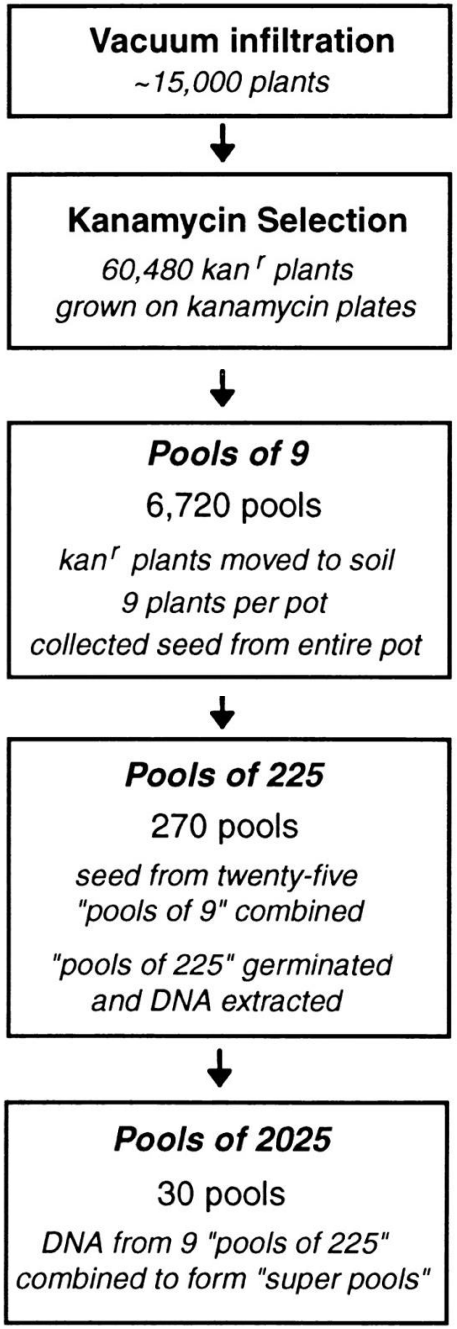


Name	Location	Result
 *Knock-out*	coding region or promoter	null
 *Knock-down*	promoter or 3' UTR	reduced expression
 *Knock-on*	promoter	increased expression



How do we test whether a plant has a T-DNA in Gene A?

How do you find a T-DNA insertion mutant in a population of 60,480 transgenic plants?



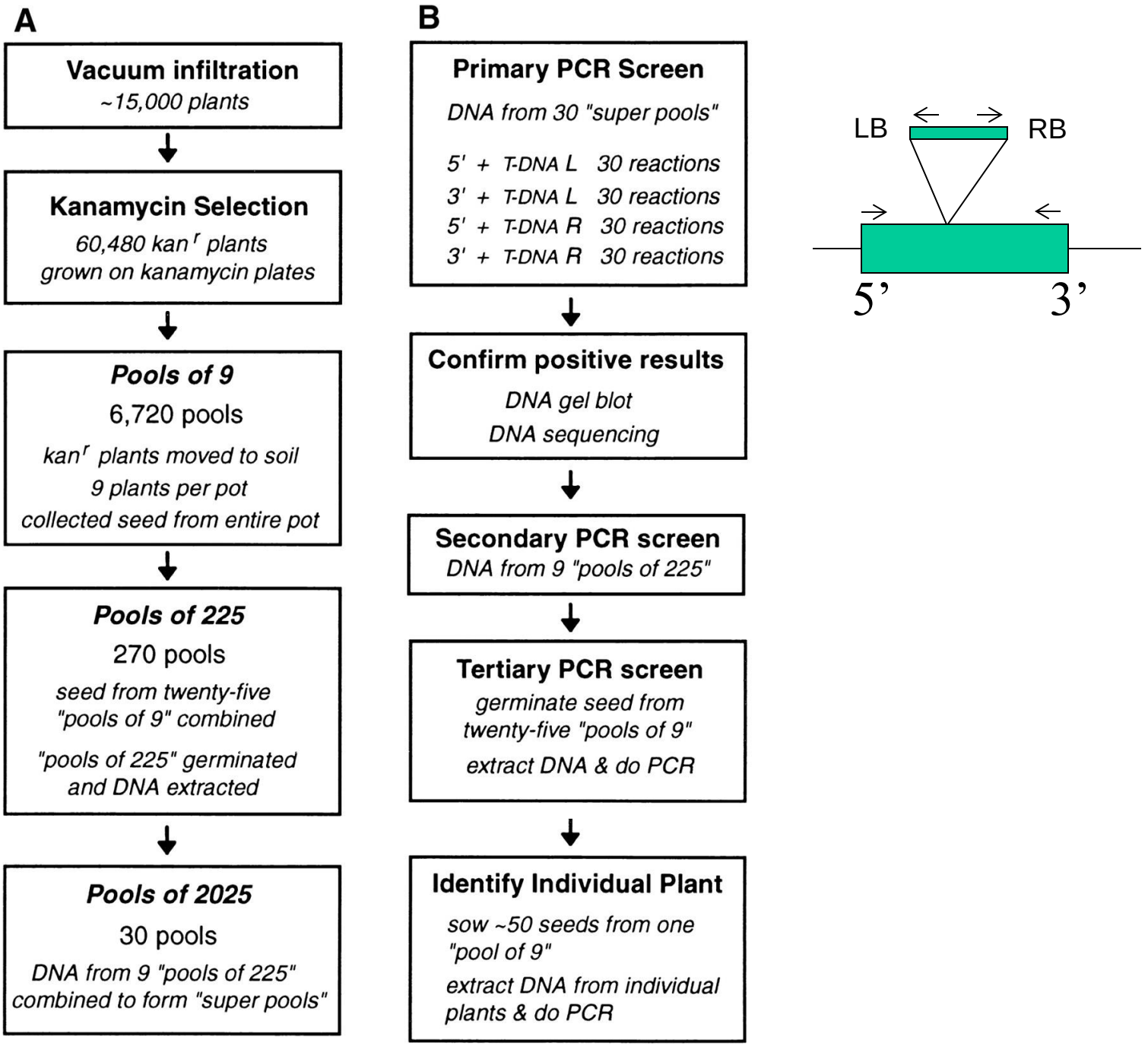
Pool 9 into one

↓

Pool 25 into one

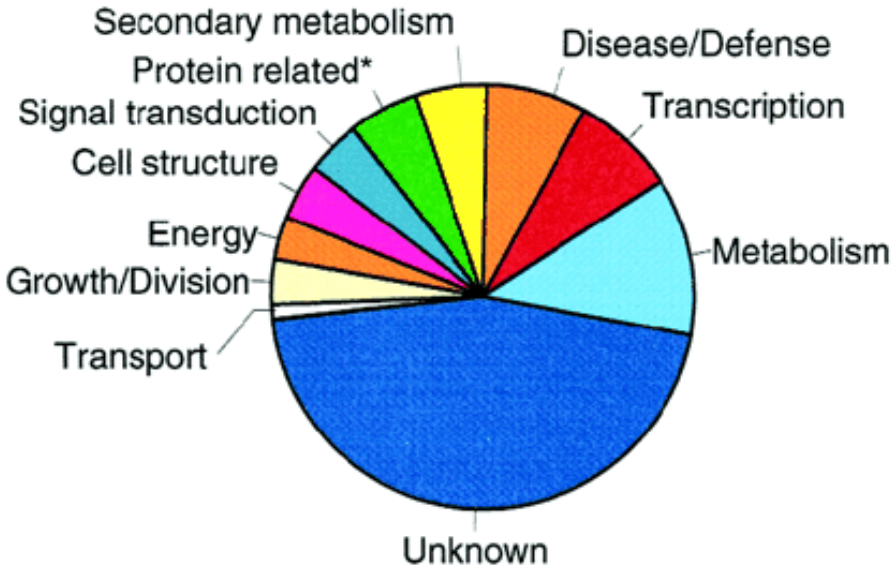
↓

Pool 9 into one


A
Vacuum infiltration
~15,000 plants
Kanamycin Selection
60,480 kan r plants
grown on kanamycin plates
Pools of 9
6,720 pools
kan r plants moved to soil
9 plants per pot
collected seed from entire pot
Pools of 225
270 pools
seed from twenty-five "pools of 9" combined
"pools of 225" germinated and DNA extracted
Pools of 2025
30 pools
DNA from 9 "pools of 225" combined to form "super pools"
B
Primary PCR Screen
DNA from 30 "super pools"
5' + T-DNA L 30 reactions
3' + T-DNA L 30 reactions
5' + T-DNA R 30 reactions
3' + T-DNA R 30 reactions
Confirm positive results
DNA gel blot
DNA sequencing
Secondary PCR screen
DNA from 9 "pools of 225"
Tertiary PCR screen
germinate seed from twenty-five "pools of 9"
extract DNA & do PCR
Identify Individual Plant
sow ~50 seeds from one "pool of 9"
extract DNA from individual plants & do PCR
LB
RB
5'
3'

Arabidopsis 2010 Project

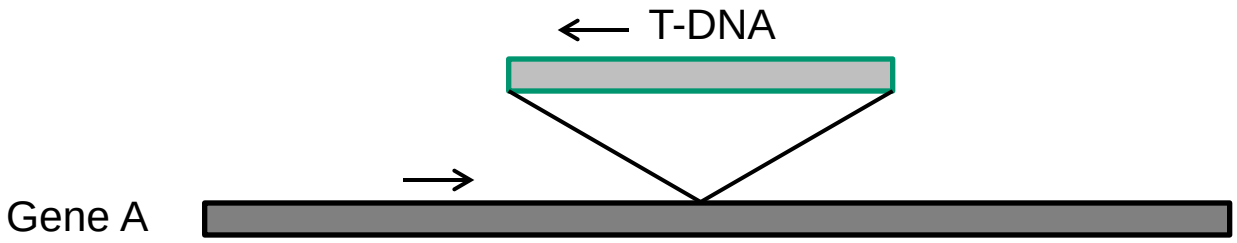
A research program proposed in 2000 to determine the function of every gene in Arabidopsis by 2010



Somerville C and Somerville, S. 1999. Plant Functional Genomics. Science 285(5426): 380-383.

Indexed T-DNA knockout lines

- Built using end-rescue and sequencing of individual T-DNA line in the population.



Ends of ~ 300,000 T-DNA lines have been sequenced.

- Major sources:
  - SALK Institute (SALK lines, USA)
  - Syngenta Inc. (SAIL lines, USA)
  - Wisc lines (UW, Madison lines, USA)
  - FLAG lines (French)
  - GABI lines (German)
  - SK lines (Canadian)

Indexed T-DNA knockout lines Search engine

T-DNA Express

http://signal.salk.edu/cgi-bin/tdnaexpress

Forward genetics in a reverse way

Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in *Arabidopsis*

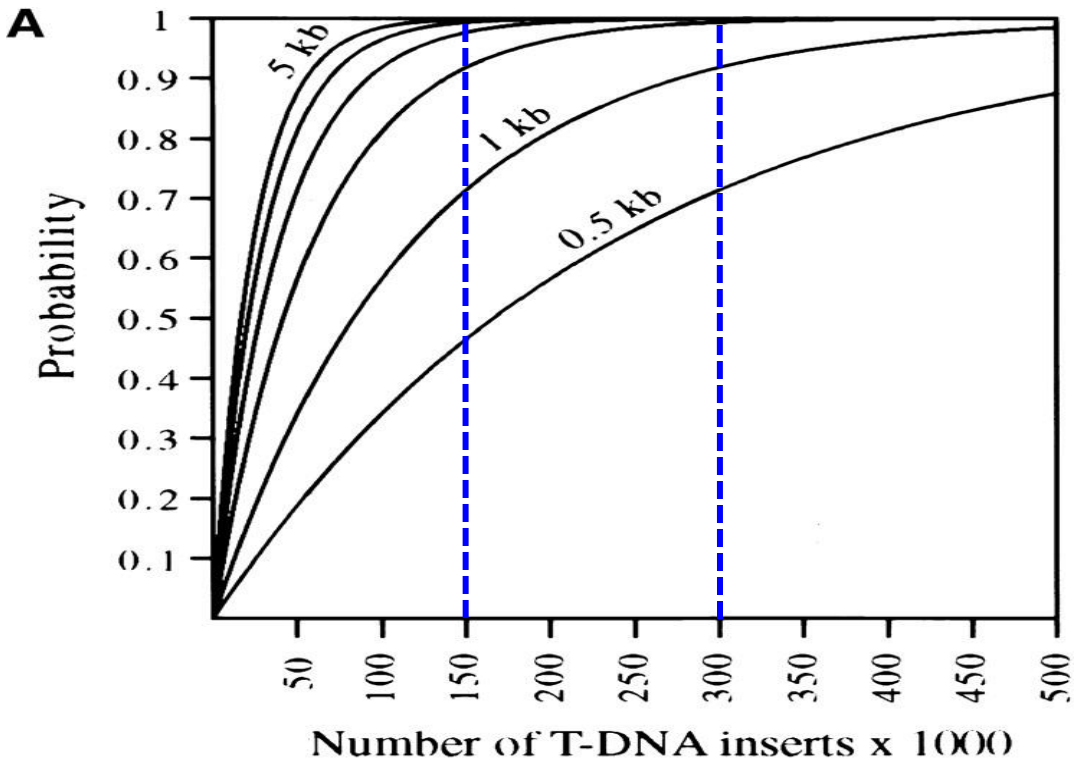
Jose M. Alonso and Joseph R. Ecker‡*

Abstract | Genome sequencing, in combination with various computational and empirical approaches to sequence annotation, has made possible the identification of more than 30,000 genes in *Arabidopsis thaliana*. Increasingly sophisticated genetic tools are being developed with the long-term goal of understanding how the coordinated activity of these genes gives rise to a complex organism. The combination of classical forward genetics with recently developed genome-wide, gene-indexed mutant collections is beginning to revolutionize the way in which gene functions are studied in plants. High-throughput screens using these mutant populations should provide a means to analyse plant gene functions — the phenome — on a genomic scale.

Alonso JM, Ecker JR. Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in Arabidopsis. Nat Rev Genet. 2006 Jul;7(7):524-36.

In Arabidopsis, about 300,000 T-DNA lines have been sequenced. T-DNA insertions are still not found in some genes.

Why?



Bigger genes have better chance being knocked out by T-DNA.

Reverse genetics:
Gene Knockout Strategies

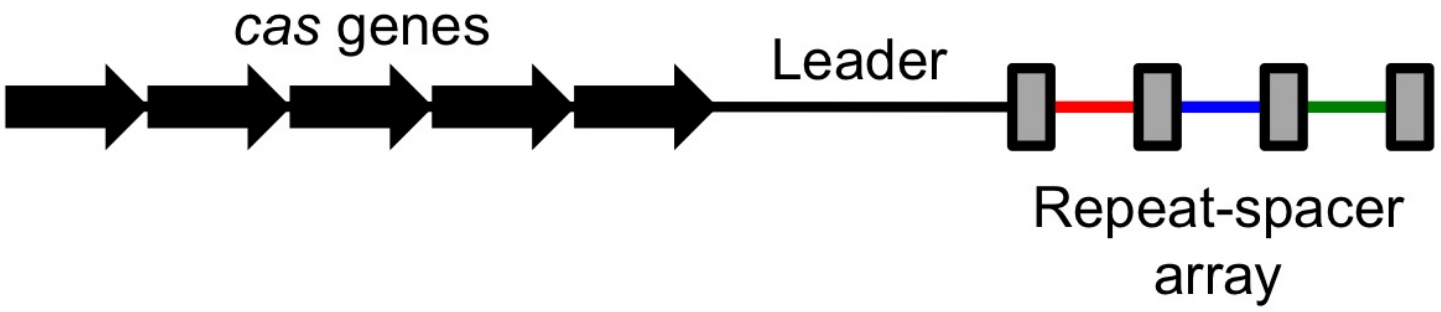
1. RNAi-based silencing

2. T-DNA or transposon based insertional mutagenesis

3. CRISPR-CAS9-based gene editing

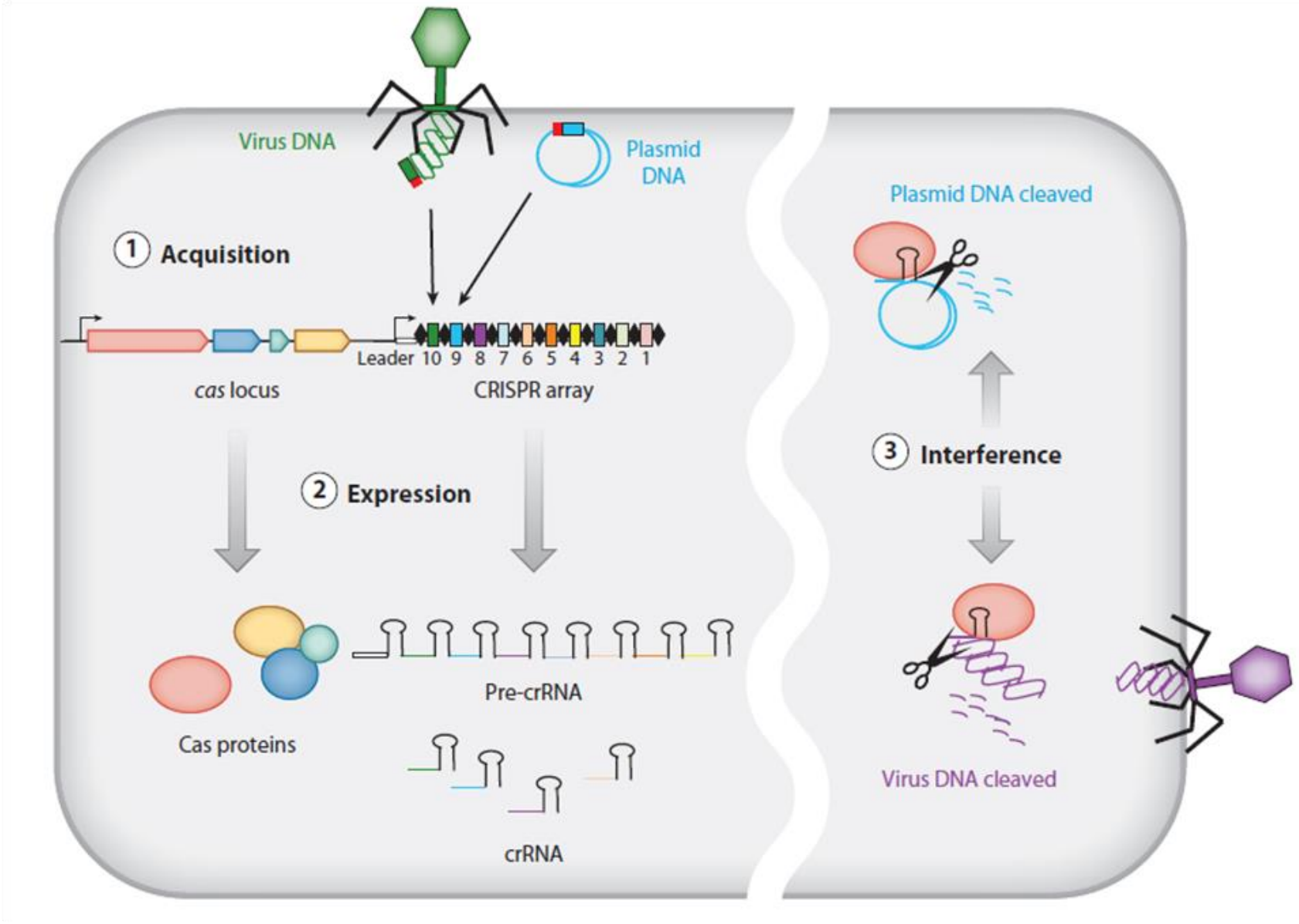
Genome Editing
Using the CRISPR-Cas9 System

CRISPR: clustered regularly interspaced short palindromic repeats



Cas: CRISPR-associated (*cas*) genes

Bacterial Immune System



CRISPR–Cas system

Cas9 (CRISPR associated protein 9)

Cas9 encodes a DNA endonuclease that associates with crRNA (also called guide RNA).

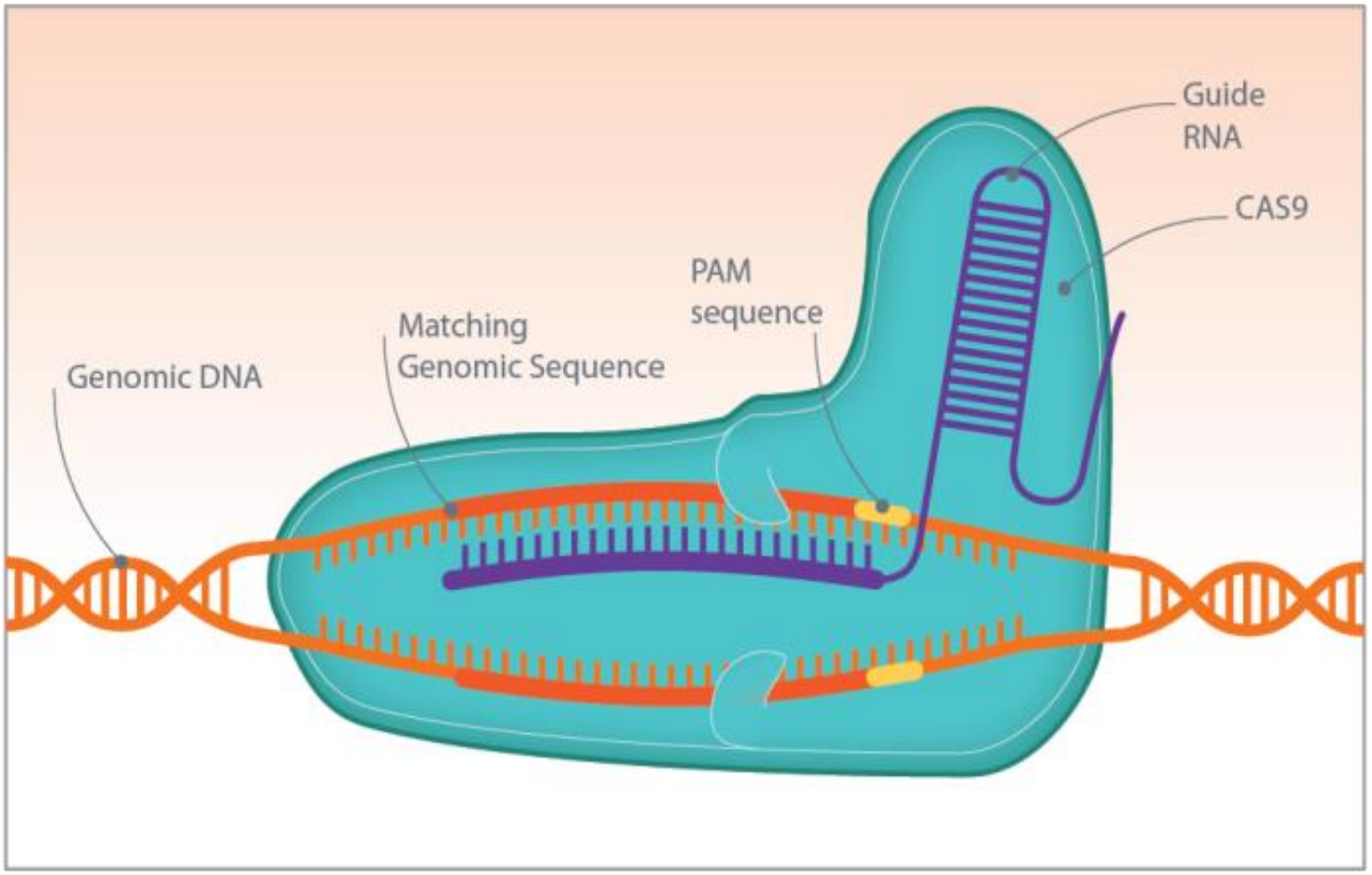
Cas9 unwinds foreign DNA to check if it is complementary to the 20 base pair spacer region of the guide RNA.

Cas9 cleaves the invading DNA if the DNA substrate is complementary to guide RNA.

Can we apply the CRISPR-Cas9 system to modify target genes in plants?
How?

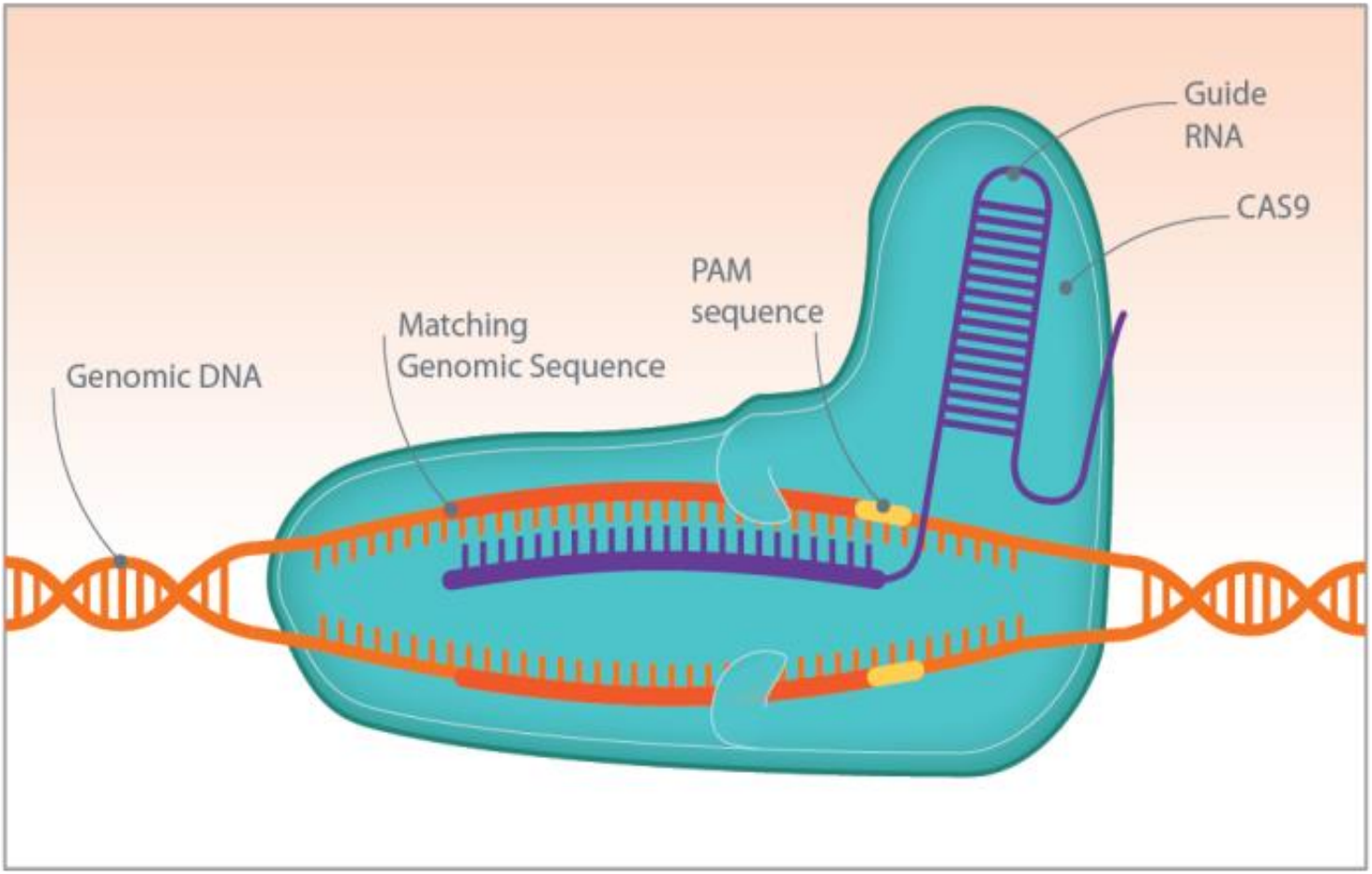
CRISPR-Cas9: a RNA-guided platform to cut at specified locations in the genome

1. Design Short guide RNAs with homology to target loci
2. Guide RNA + Cas9 are expressed in the cell
3. The Cas9 cleavage site is repaired by either nonhomologous end-joining or homologous DNA recombination in tandem with a donor
4. High efficiencies of knockout or knock-in



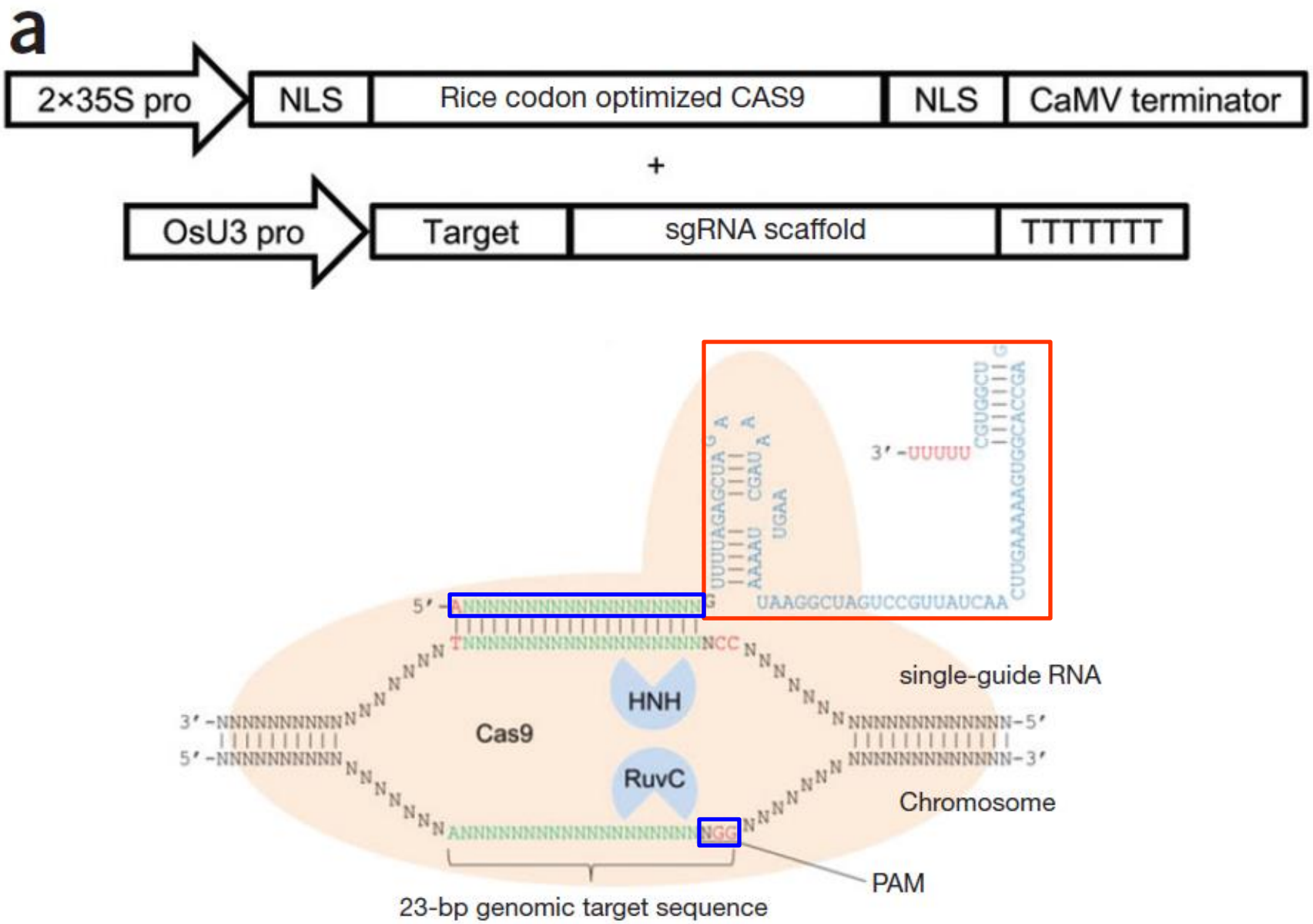
Guide RNA: crRNA + tracrRNA (trans-activating RNA)

PAM: protospacer adjacent motif, NGG

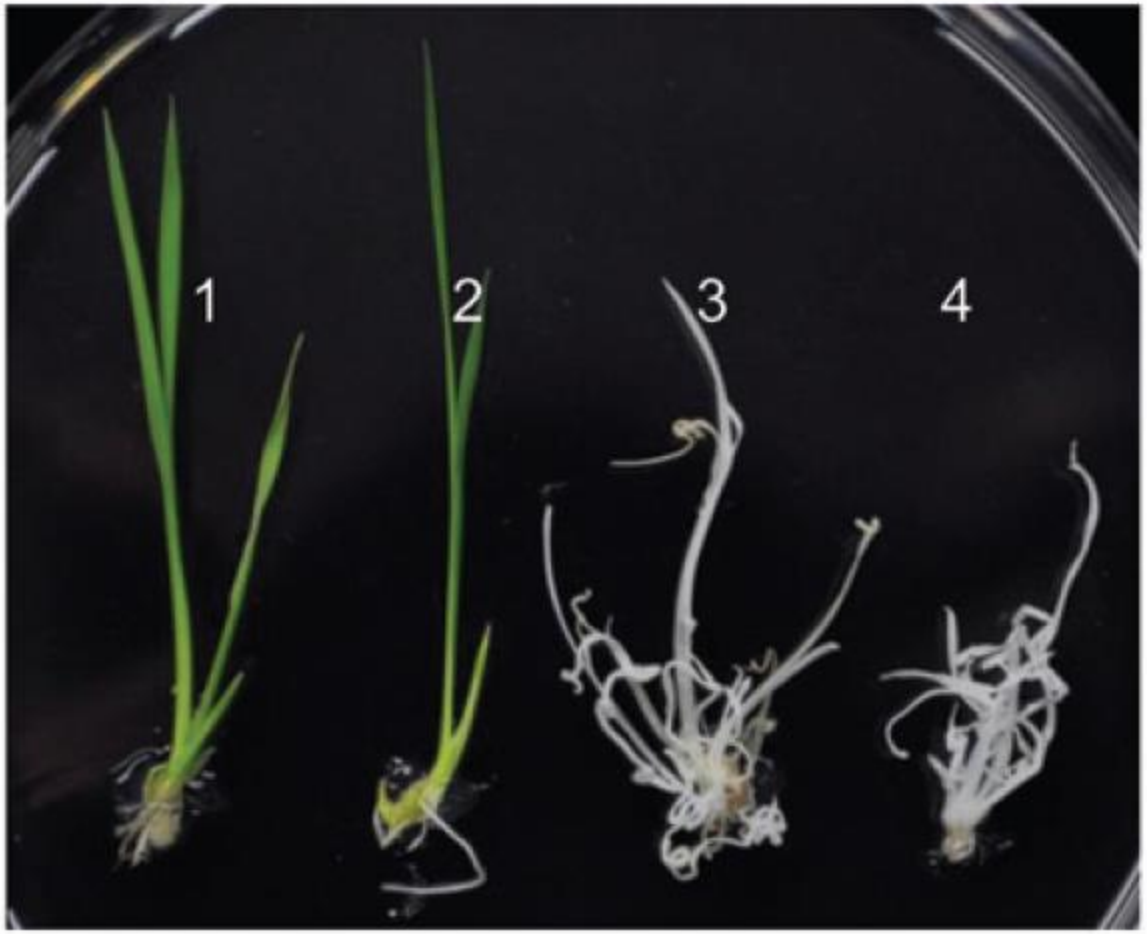


gRNA target sequence | PAM

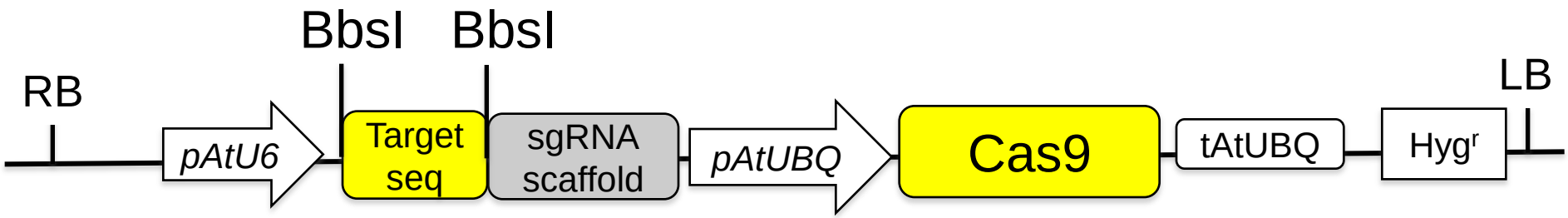
AGCTGGGATCAACTATAGCG NGG


a
2×35S pro
NLS
Rice codon optimized CAS9
NLS
CaMV terminator
+
OsU3 pro
Target
sgRNA scaffold
TTTTTTT
single-guide RNA
HNH
Cas9
RuvC
Chromosome
PAM
23-bp genomic target sequence

9 out of 96 T1 transgenic plants contain mutations in the rice phytoene desaturase gene



(1) Wild type rice plant
(2) Monoallelic mutant
(3) Biallelic mutant
(4) Biallelic mutant


BbsI
BbsI
RB
LB
pAtU6
Target seq
sgRNA scaffold
pAtUBQ
Cas9
tAtUBQ
Hygr

T-DNA vs. CRISPR-Cas9 system

(in knocking out genes)

Advantages of the CRISPR-Cas9 system

1. Targeted editing of a gene of interest
2. Applicable to crop plants
3. Easy to carry out
4. Target multiple homologous genes

https://www.youtube.com/watch?v=4YKFw2KZA5o