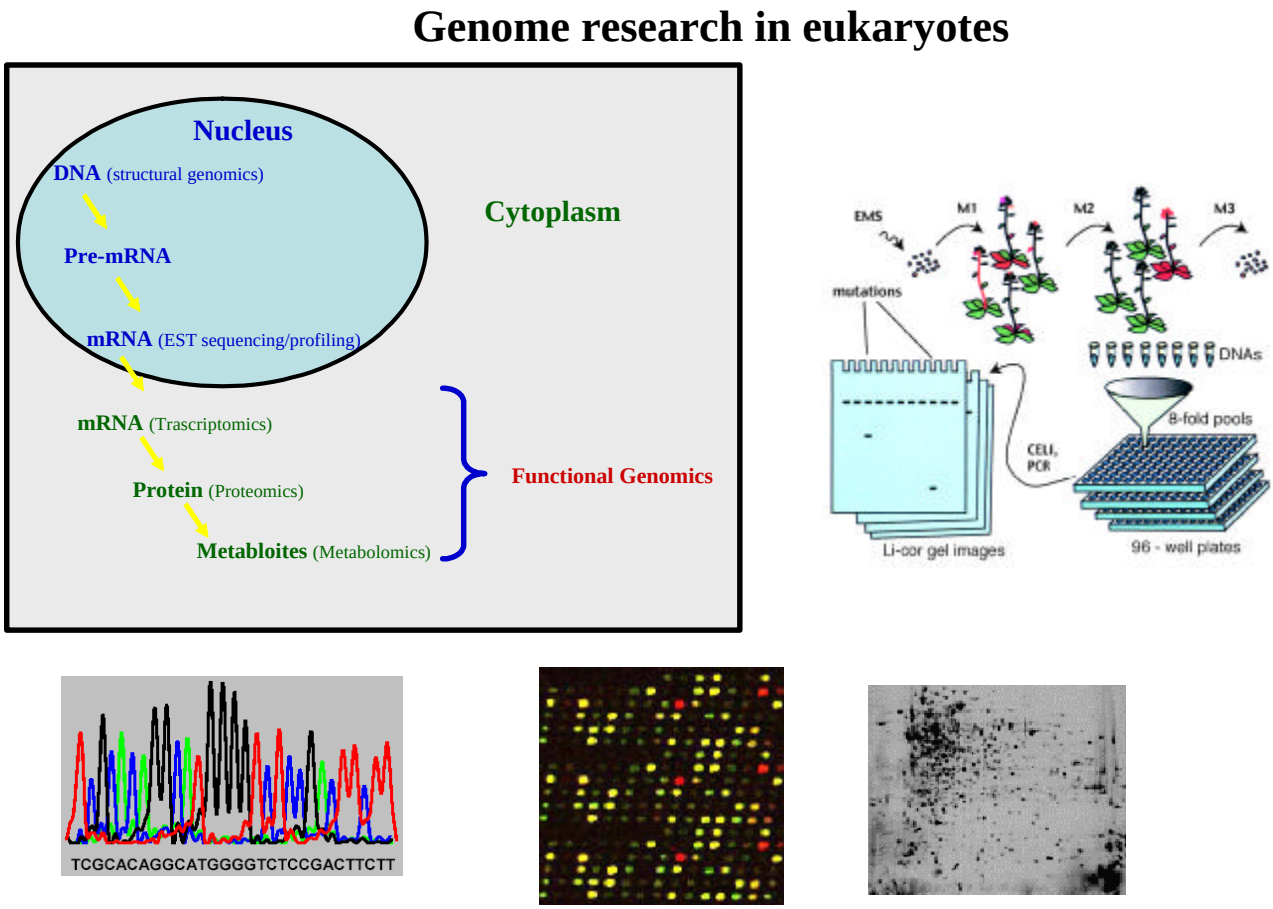
Functional Genomics

Genome and EST sequencing can tell us how many POTENTIAL genes are present in the genome

Proteomics can tell us about proteins and their interactions

The goal of functional genomics is to define the function of each and every gene in the genome.



Main categories of functional genomics

Forward genetics

Mutant phenotype leads to gene sequence and function

Reverse genetics

Mutant sequence (mutant genotype) leads to possible phenotype and function

Forward Genetics	Reverse Genetics
Known phenotype	Unknown phenotype
Unknown sequence	Known sequence

Do mutations in all genes lead to a phenotype?

Fine structure genetics

Tools for forward and reverse genetics

Insertional mutagenesis

Transposon tagging
T-DNA tagging

Sequence mutagenesis

Radiation mutagenesis
Chemical mutagenesis

Targeted gene mutagenesis

Sense or anti-sense expression
Homologous recombination
Virus induced gene silencing (VIGS)
RNA interference (RNAi)

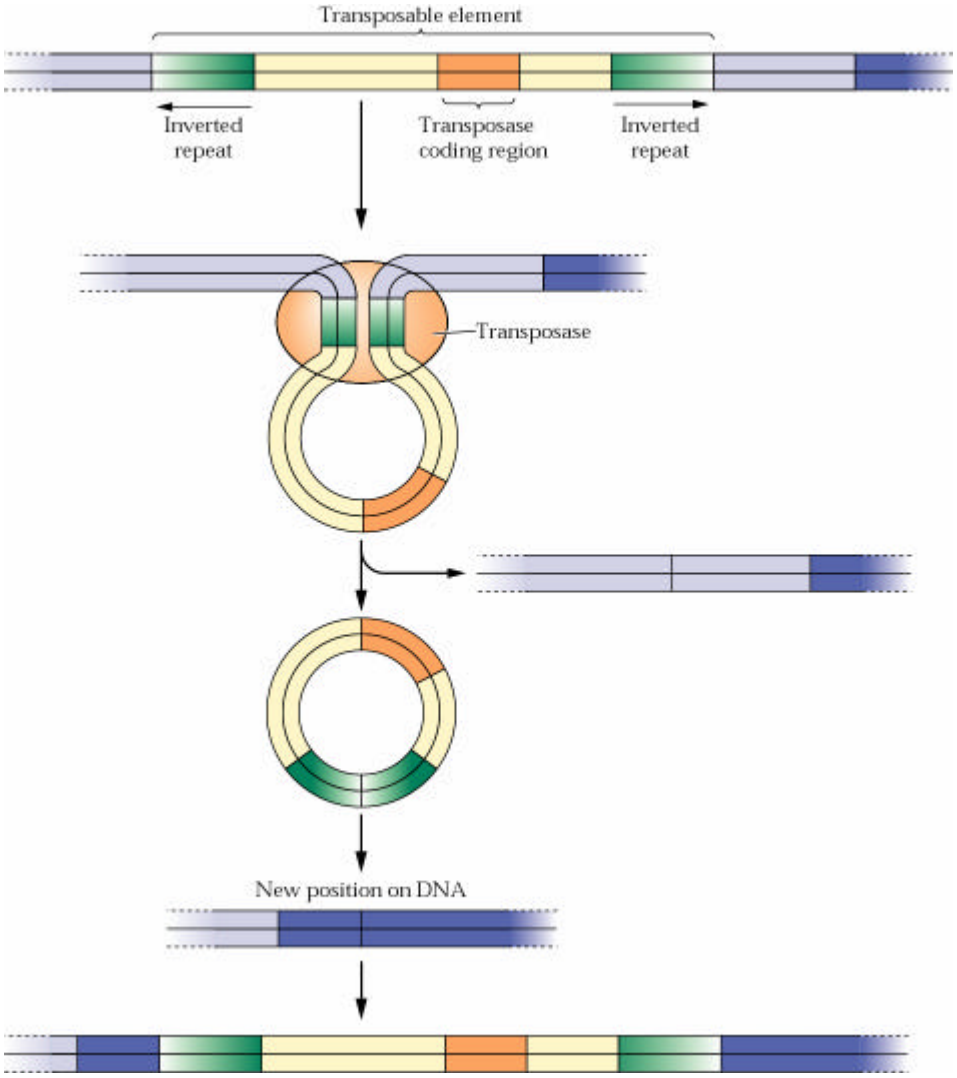
Insertional mutagenesis

Insertion of known DNA segment into a gene/sequence of interest

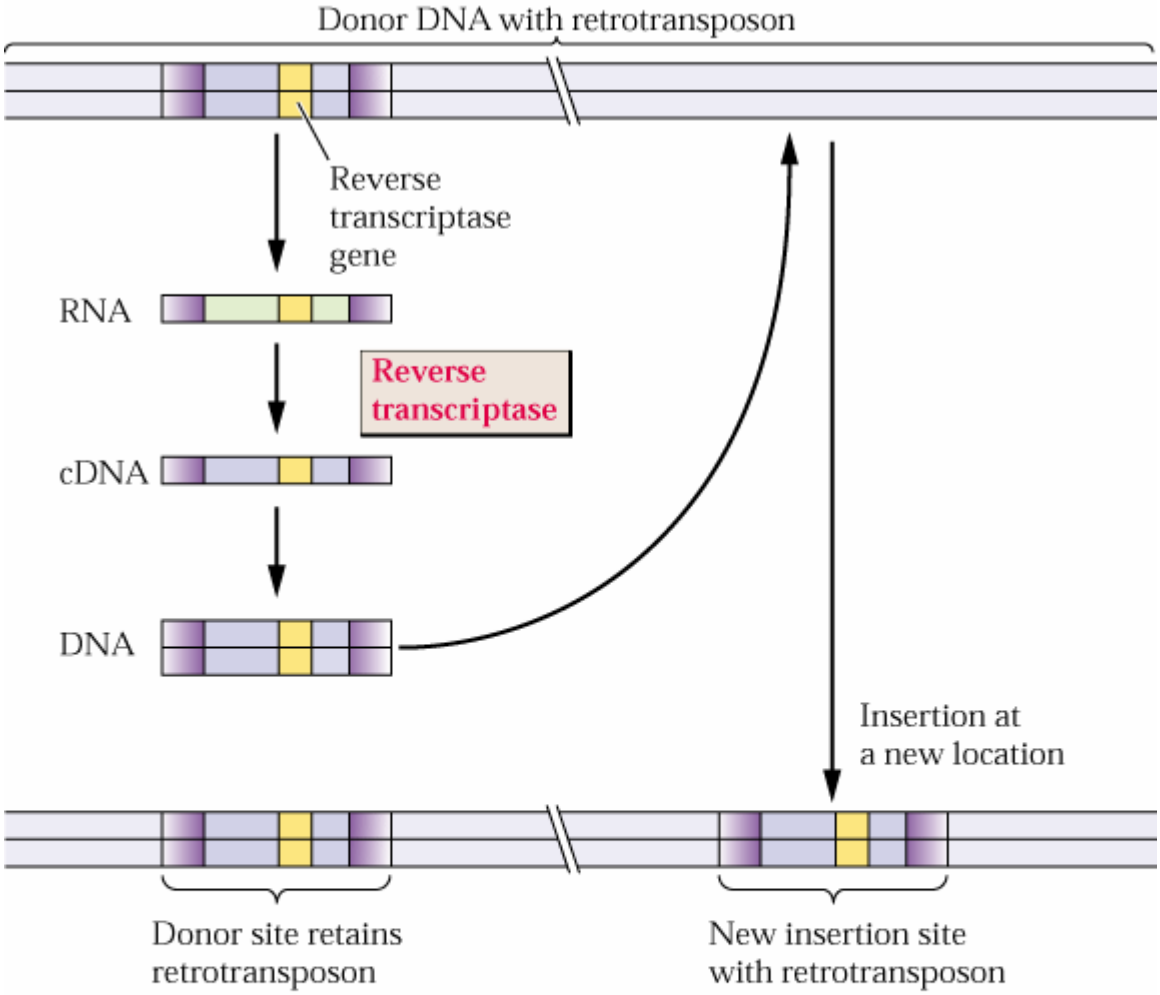
Transposon tagging

Transposable elements or transposons are DNA elements that have the ability to move from one chromosome site to another.

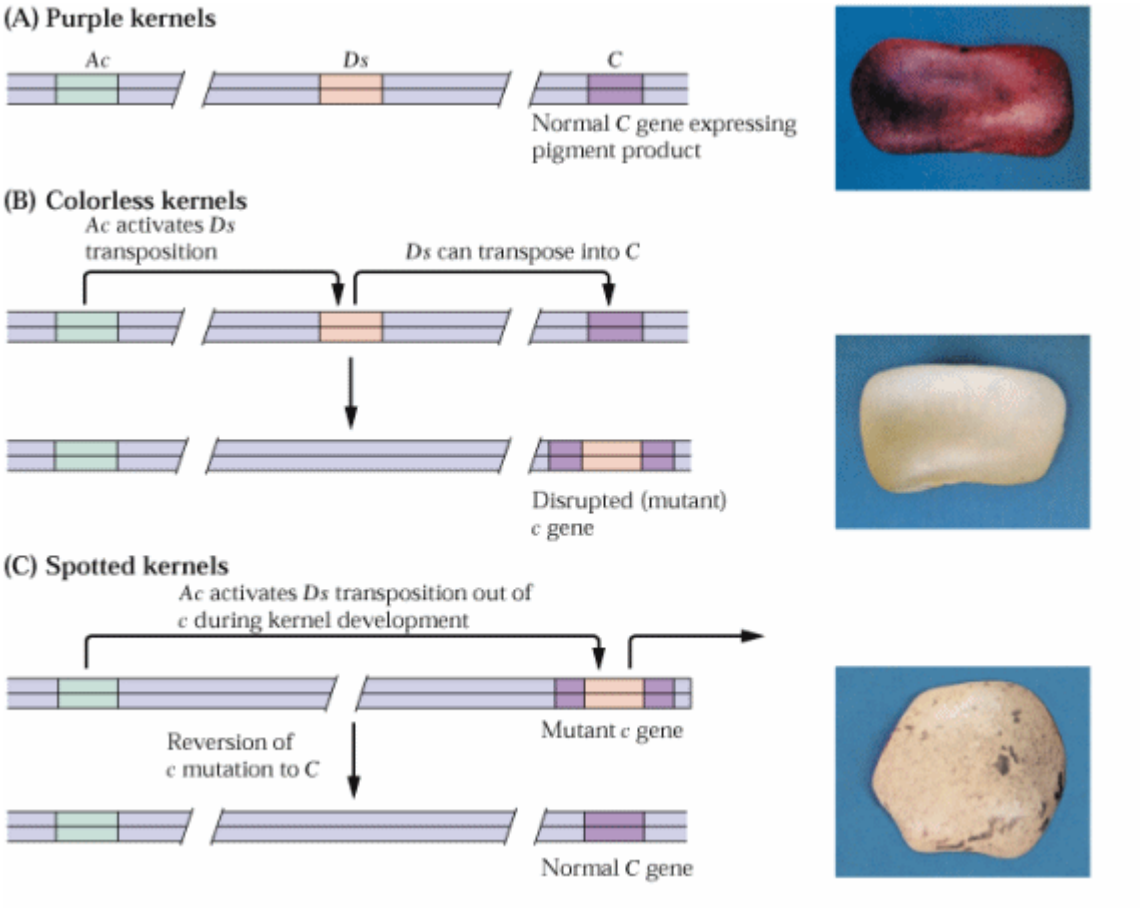
DNA transposons flanked by short inverted repeats move by excising from one chromosome site to another
Retrotransposons move via RNA intermediates



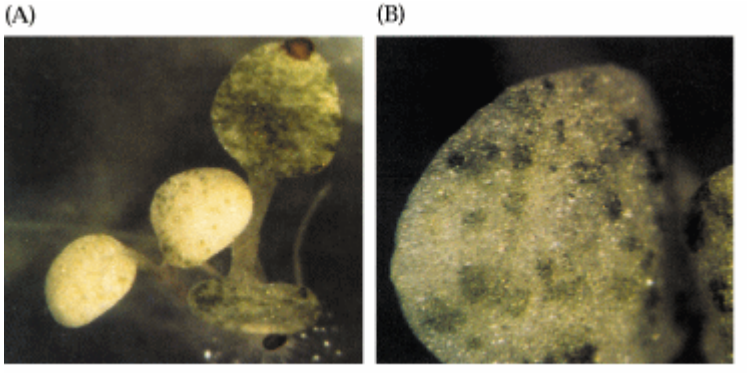
Biochemistry & Molecular Biology of Plants (ASPP)
Figure # 7.33 #736



Biochemistry & Molecular Biology of Plants (ASPP)
Figure # 7.34 #737



Biochemistry & Molecular Biology of Plants (ASPP)
Figure # 7.33 #739

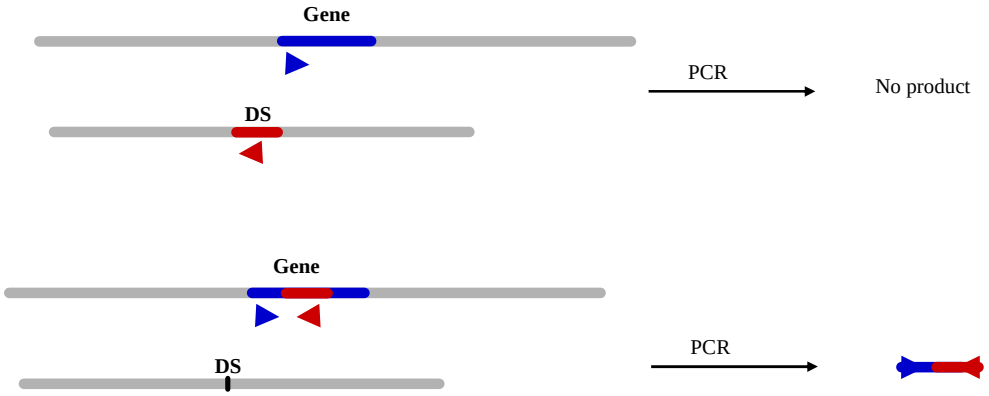


Biochemistry & Molecular Biology of Plants (ASPP)
Figure # 7.33 #740

Transposon tagging

In forward genetics application, transposons are used as random mutagens. If a mutation is generated by insertion of transposon, the transposon sequence can be used as a "mark" to identify and clone the tagged DNA.

In reverse genetics application, transposons are used as random mutagens in combination with PCR multiplexing to identify insertions in known sequences. Characterization and assignment of function to that sequence is the next step.



Advantages:
Efficient and cost-effective method to generate a large mutant population

Disadvantages:
Secondary transposition complicates gene identification
Not available in many species

T-DNA tagging

T-DNA is a segment of DNA from *Agrobacterium tumefaciens* tumor inducing (Ti) plasmid that is moved into the plant upon infection.

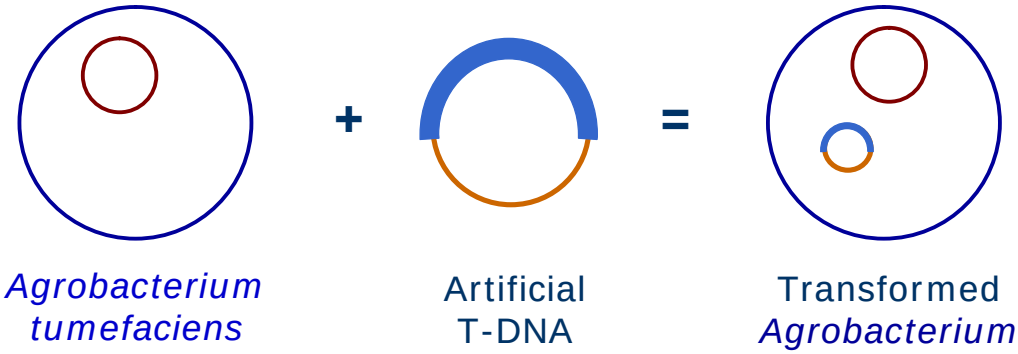
Agrobacterium tumefaciens has traditionally been used in some plant species for transformation of foreign DNA into the genome (i.e. RoundUp Ready gene) generating transgenic plants.

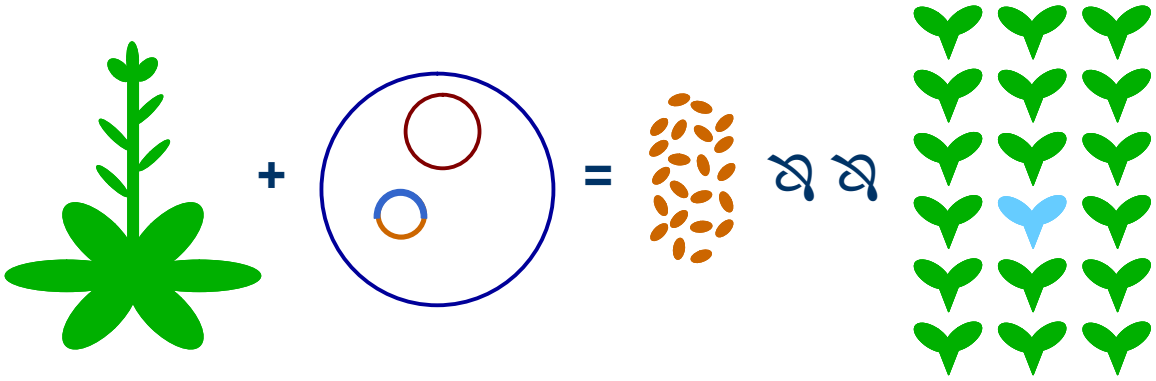
A marker selection gene (i.e. antibiotic resistance) is inserted between the borders of T-DNA so that transformed cells can be selected.

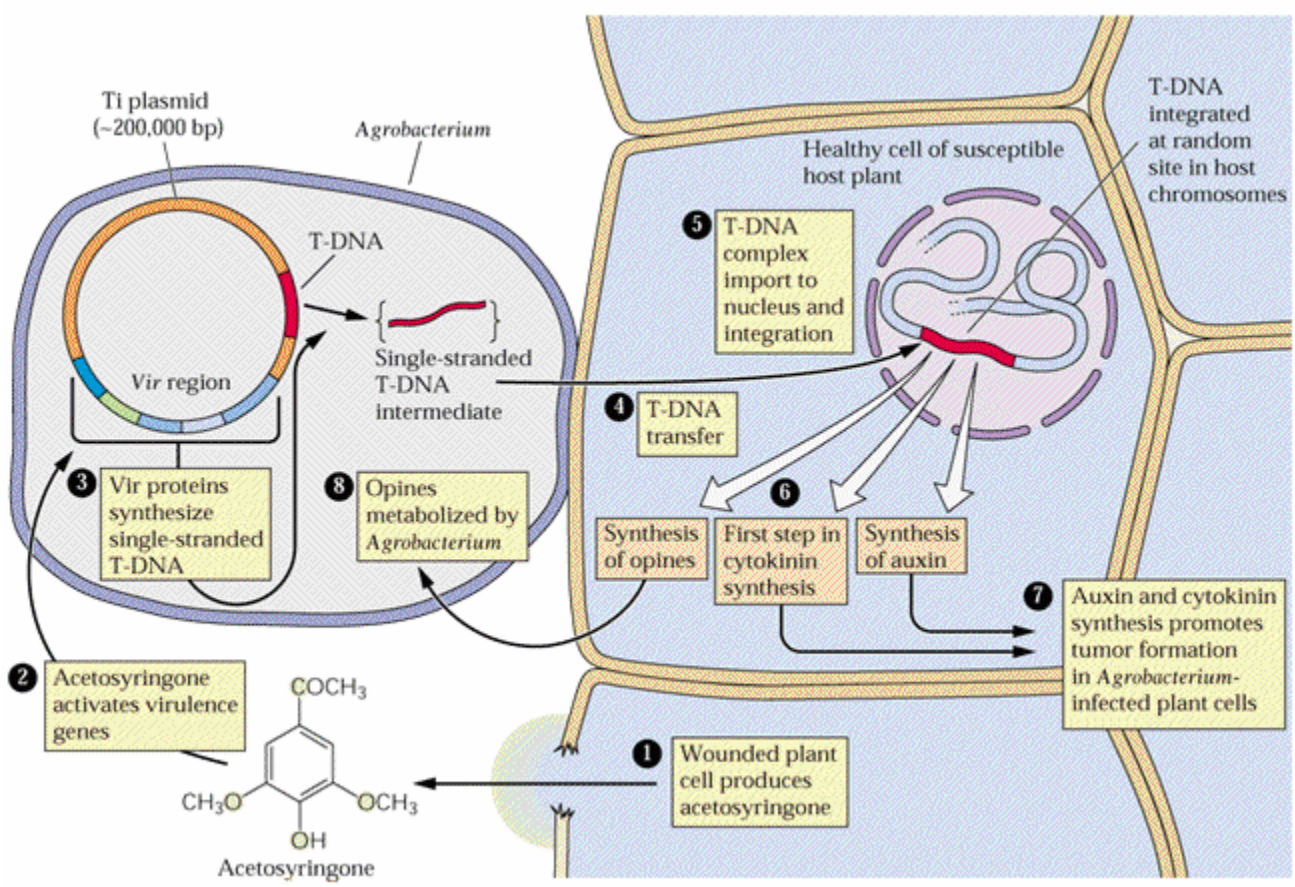
Advantages:
Effective interruption of genes
Low copy number (1.5)
Random insertion in the genome

Disadvantages:
Time consuming for transformation
Somatic variation caused by tissue culture process (i.e. high percentage of untagged mutations)
Not available in many species

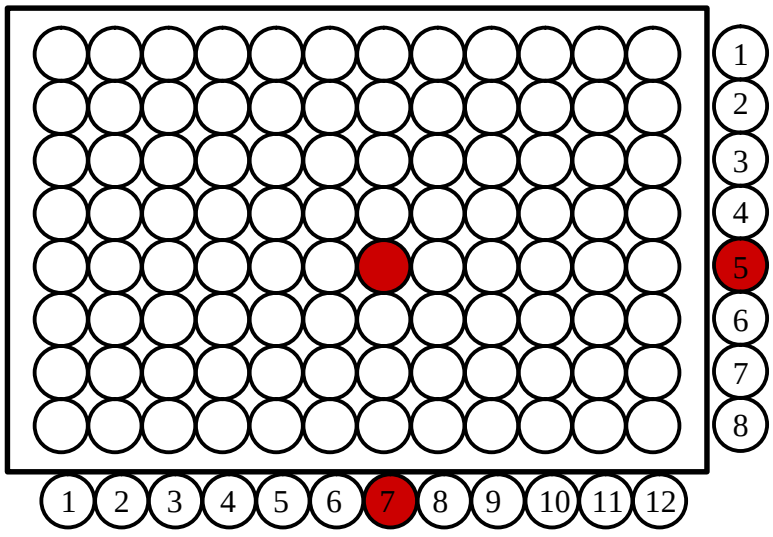

UGA4823036


+
=
Agrobacterium tumefaciens
Artificial T-DNA
Transformed *Agrobacterium*


+
=

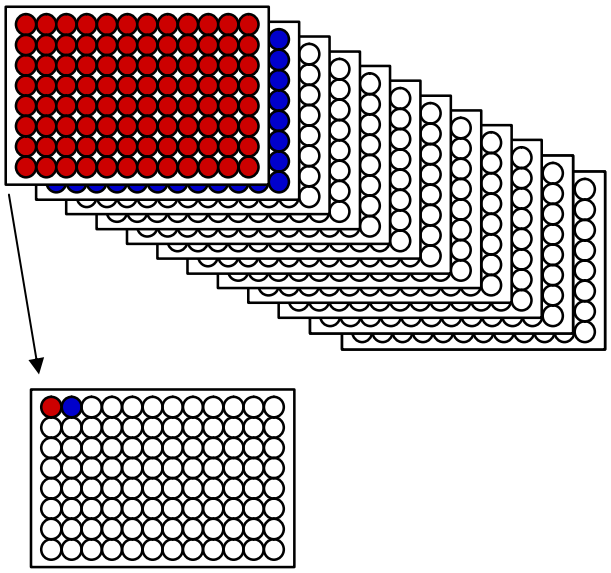


First pool



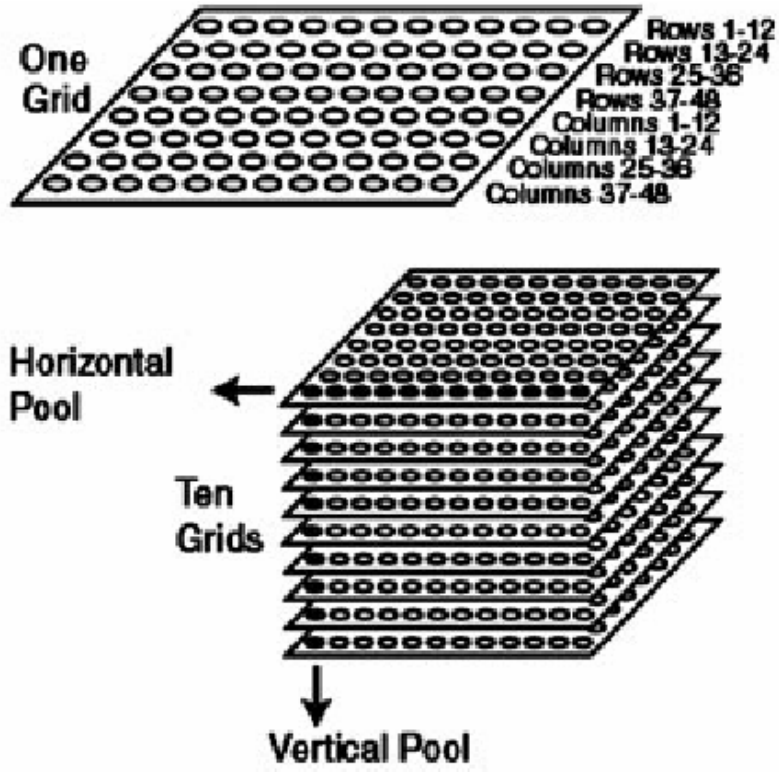
Individual reactions = 96 (12x8)
Pooled method = 20

Second pool

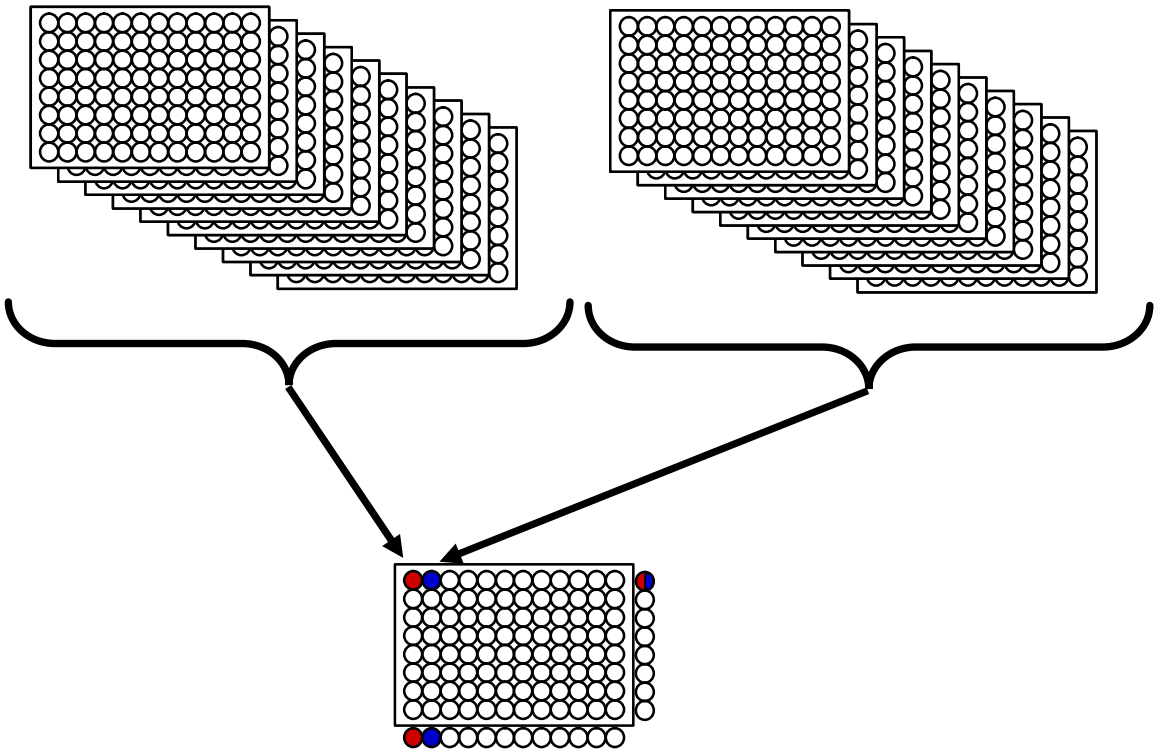


Individual reactions = 1152 (96x12)
Pooled method = 32 (12 + 20)

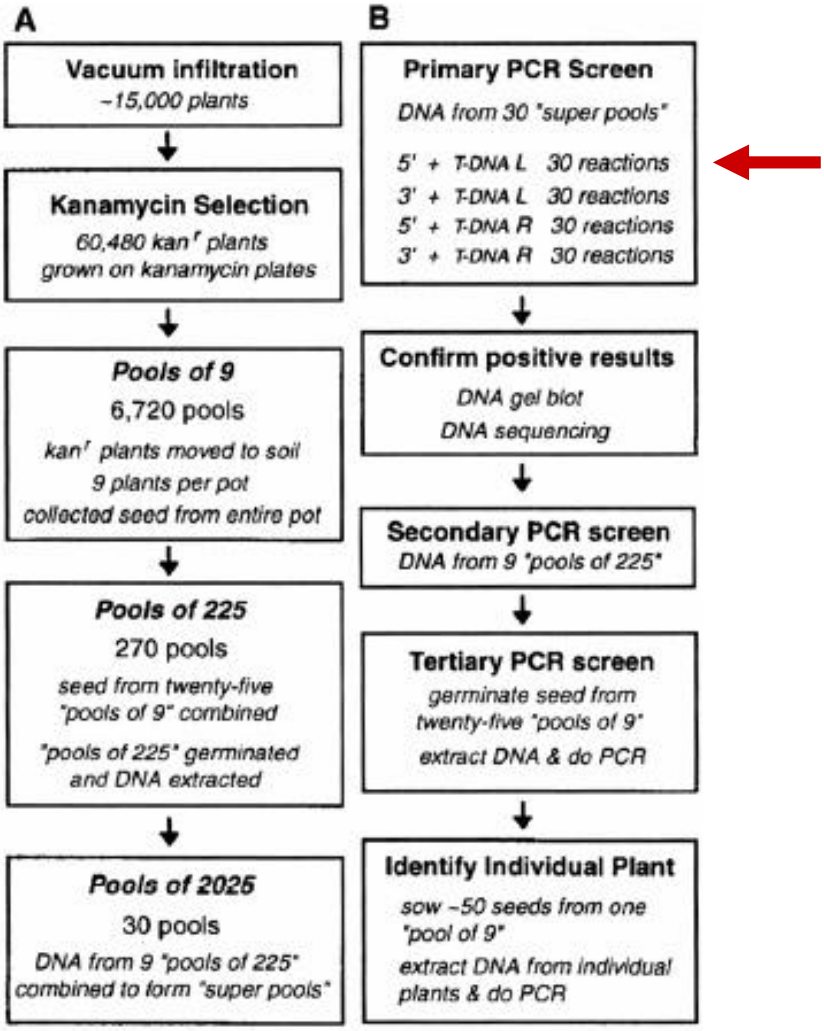
Higher order pooling



Higher order pooling



Organization and screeing of Arabidopsis T-DNA lines at Univ. of Wisconsin-Madison knockout facility



Sequence mutagenesis

Radiation mutagenesis
Chemical mutagenesis

Random mutagen are used to generate mutations in sequences through out the genome

Radiation mutagenesis

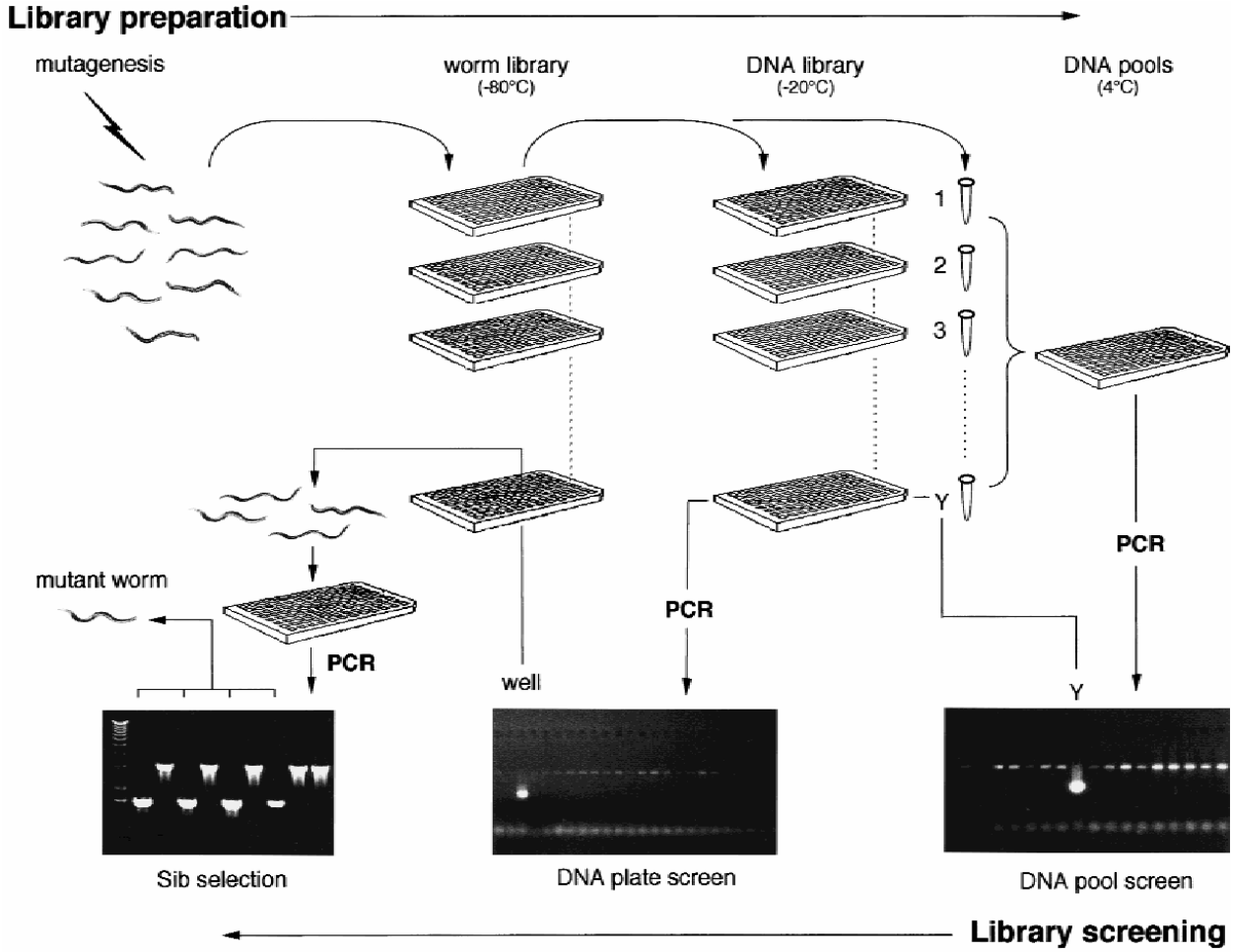
Ionizing radiation (i.e. fast neutron, gamma ray) are used to generate random mutations (breaks) in DNA segments. For example, fast neutron breaks the chromosomes leading to loss of DNA sequences (i.e. large deletions).

Chemical mutagenesis

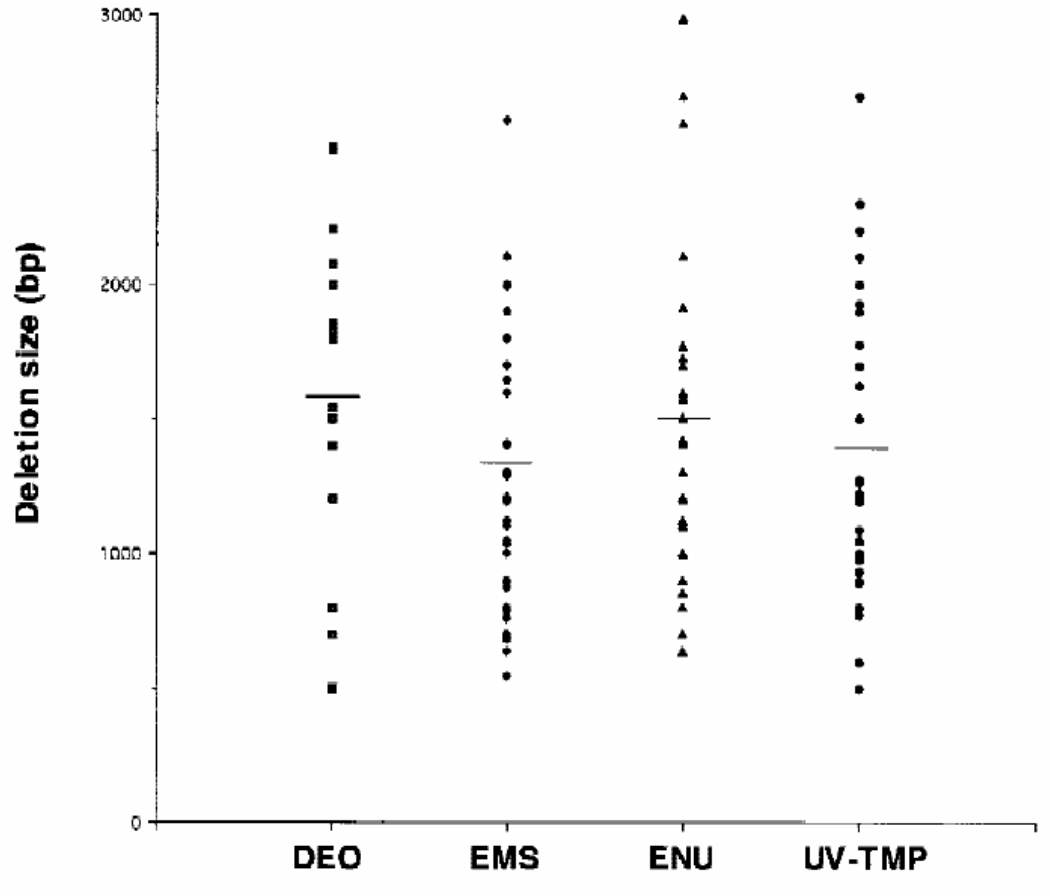
Chemicals (e.g. carcinogens) can also cause mutations in DNA sequences. Etheylmethane sulfonate (EMS) induces point mutations in DNA. EMS alkylates primarily guanine leading to mispairing: alkylated G pairs to T instead of C. The resulting mutations are mainly transitions (GC → AT). Diepoxybutane (DEB) and trimethylpsoralen + UV (TMP + UV) generally cause small deletions (100 bp to 1.5kb). These last chemicals generate DNA interstrand cross-links, which are repaired by the replication machinery by removal of the effected sequences.

High-throughput isolation of Caenorhabditis elegans deletion mutants

The nematode C. elegans is the first animal genome to be sequenced.
Four chemical mutants were used to induce detectable deletions
The deletions averaged in size about 1400 bp
Both reversed and forward genetics approaches can be applied to study of these mutants



Schematic of mutant construction and screening

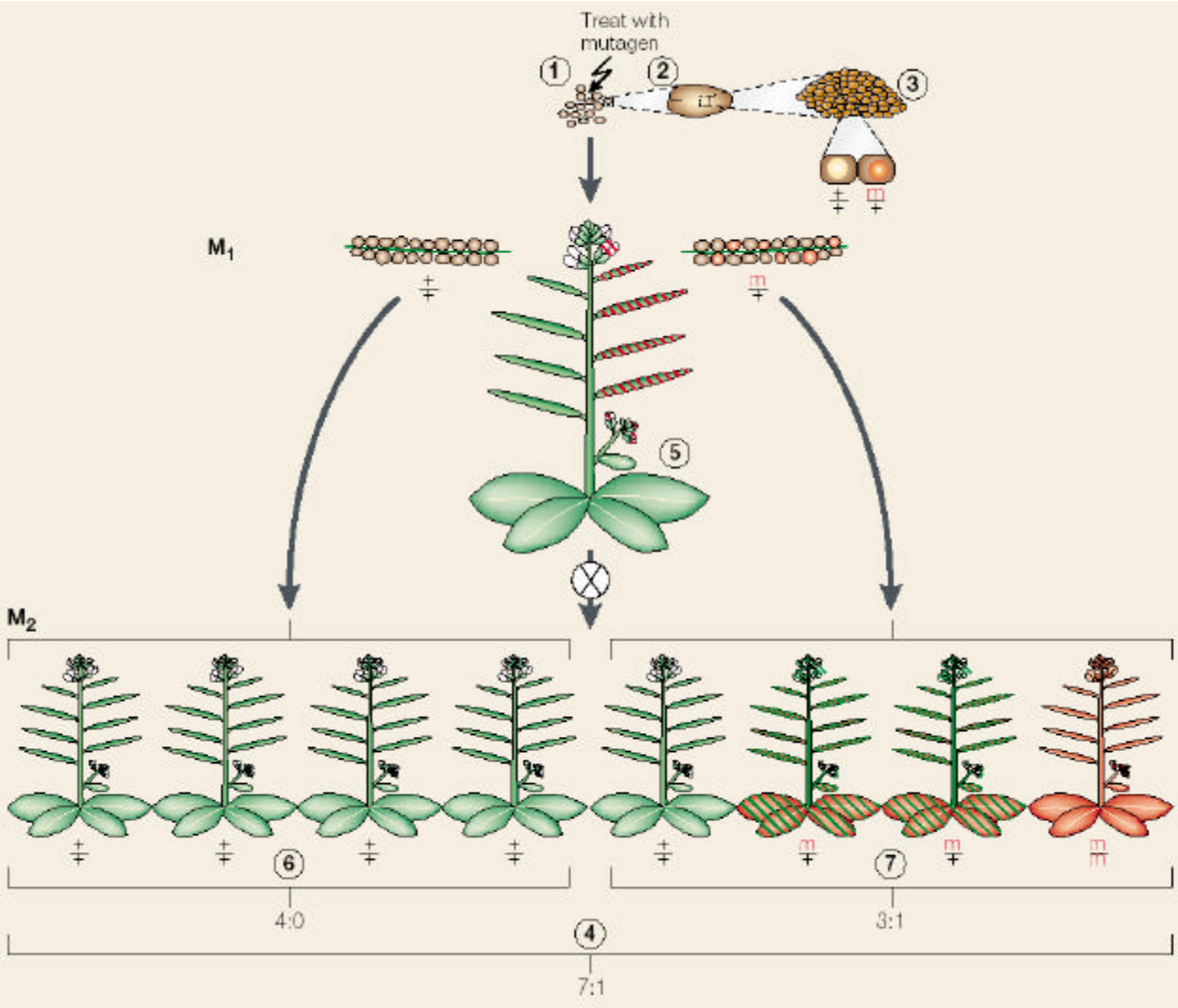


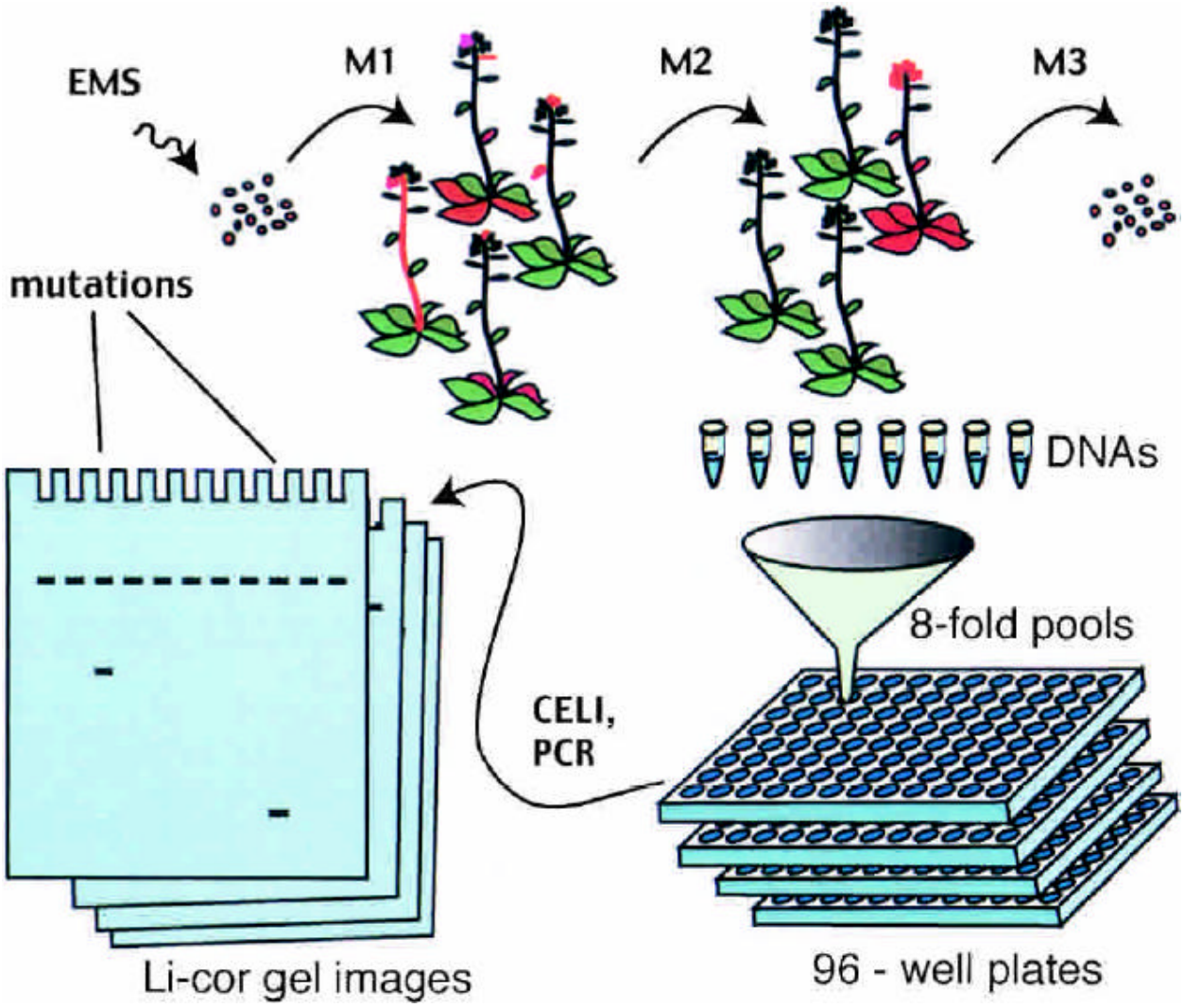
Deletion size range by mutagen

Table 2. Deletion Yield by Mutagen

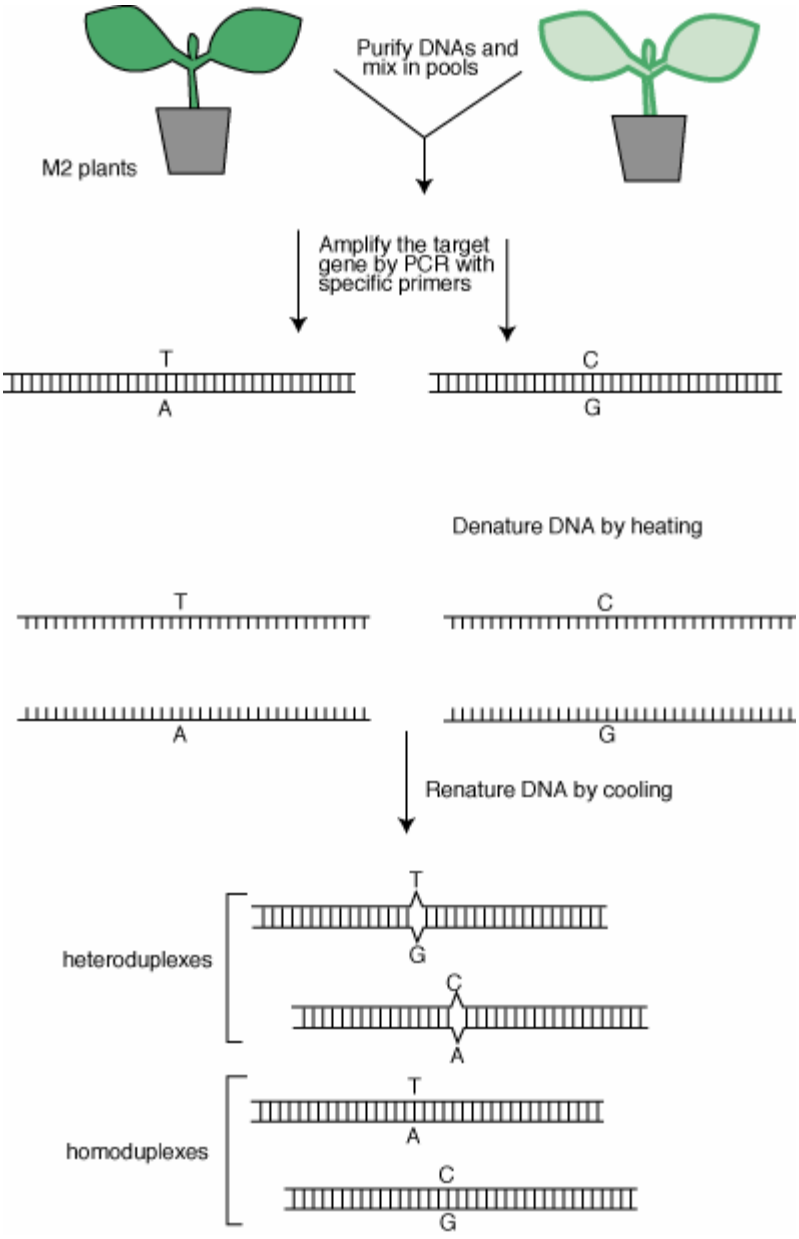
Mutagen	No. of libraries	No. of 1° screens	No. of deletion mutants	Ratio of deletions/ screens
DEO	3	171	12	0.070
EMS	9	479	47	0.098
ENU	4	208	23	0.111
UV-TMP	4	135	29	0.215

The total number of pool screens initiated (using a target window of 2800–3400 bp) and the total number of isolated deletions were tabulated for each mutant library.

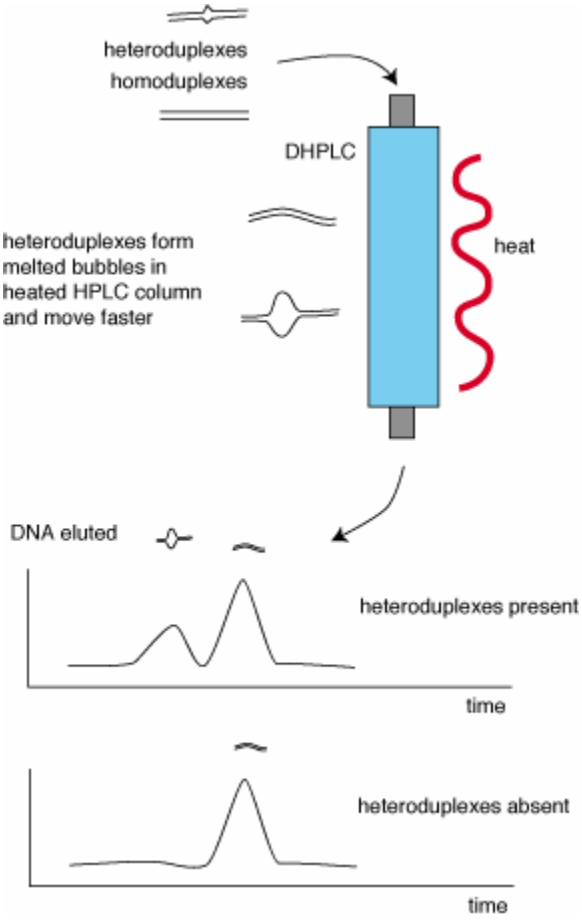

Treat with mutagen
1
2
3
M1
5
M2
6
4:0
4
7:1
7
3:1



Targeting Induced Local Lesions IN Genomes (TILLING)



TILLING analysis



TILLING detection

Table 2 Websites relevant to plant functional genomics

Plant genomic databases	
The Arabidopsis Information Resource (TAIR)	http://www.arabidopsis.org/agi.html
Rice Genome Project (RGP)	http://rgp.dna.affrc.go.jp/
Arabidopsis Transposon Insertion Database	http://formaggio.cshl.org/~h-liu/attdb/index.html
Arabidopsis Knockout Facility	http://www.biotech.wisc.edu/Arabidopsis
Torrey Mesa Research Institute T-DNA collection	http://tmri.org/pages/collaborations/garlic_files/
Arabidopsis SNP-site	http://www.arabidopsis.org/cereon/index.html
Project 2010	http://www.arabidopsis.org/workshop1.html
Microarray databases	
EBI: The ArrayExpress database	http://www.ebi.ac.uk/arrayexpress/
Affymetrix	http://www.affymetrix.com/products/Arabidopsis_content.html
TIGR Arabidopsis arrays	http://atarrays.tigr.org/
AFGC	http://afgc.stanford.edu/
GARNet	http://www.york.ac.uk/res/garnet/may.htm
Rice transcriptional database	http://microarray.rice.dna.affrc.go.jp
Proteomics databases	
Proteome Analysis at EBI	http://www.ebi.ac.uk/proteome/
Arabidopsis Membrane Protein Library	http://www.cbs.umn.edu/arabidopsis/
Database of *A. thaliana* Annotation	http://luggagefast.Stanford.EDU/group/arabprotein/
ExPASy *A. thaliana* 2D-proteome database	http://www.expasy.ch/cgi-bin/map2/def?ARABIDOPSIS
PlantsP: Functional Genomics of Plant Phosphorylation	http://PlantsP.sdsc.edu/
PPMdb: Plant Plasmamembrane DataBase	http://sphinx.rug.ac.be:8080/ppmdb/index.html