

Understanding Gene Function – By **Forward Genetics**

Create a **new phenotype or mutant**
(**due to loss function** or gain of function)
look for gene responsible for it !!!!

Which gene is
responsible ?

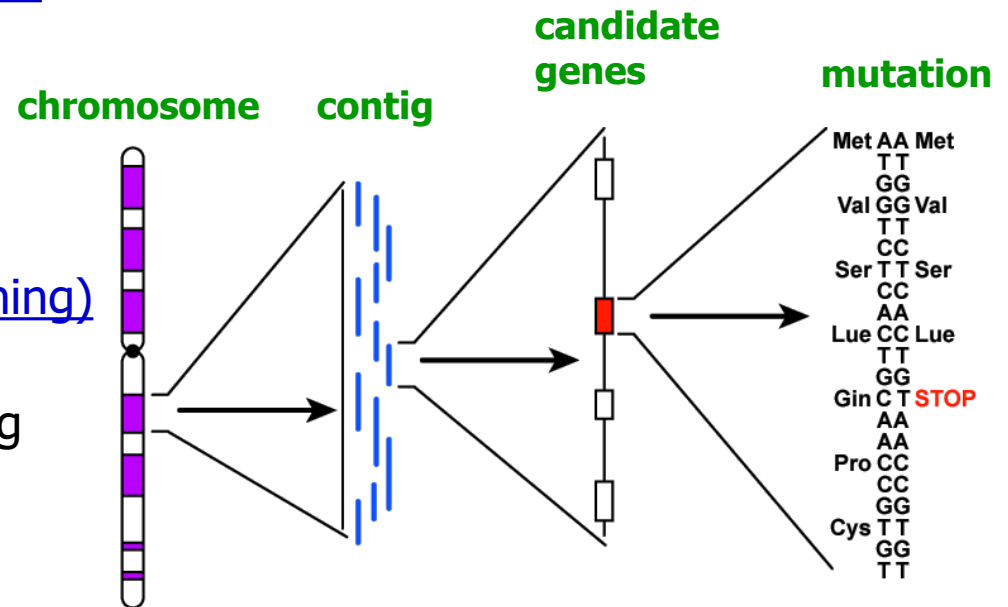


Altered phenotype or mutant
(**loss of function mutant**)

from Phenotype to Gene

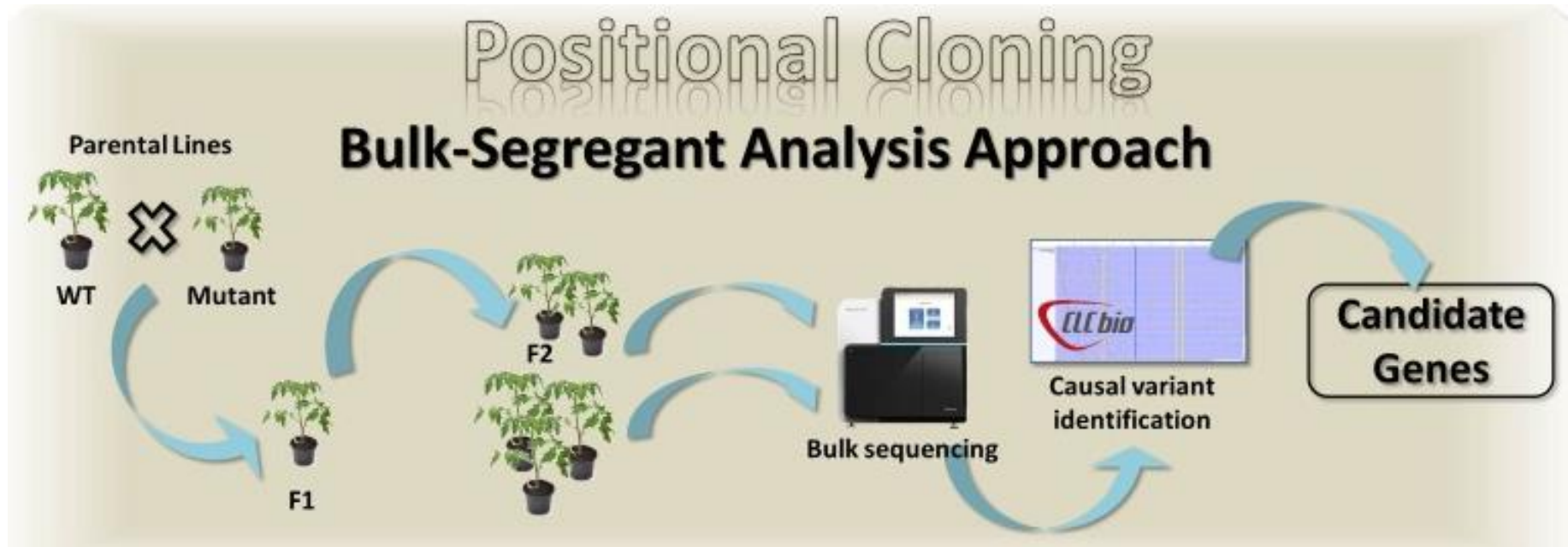
Mapping and identification of mutated gene

- Once an interesting Mutant is found and characterized, we look for the gene in which the mutation occurred
- Positional cloning (Map based cloning)
 - First use genetic mapping
 - Then use chromosome walking
 - Needle in hay stack

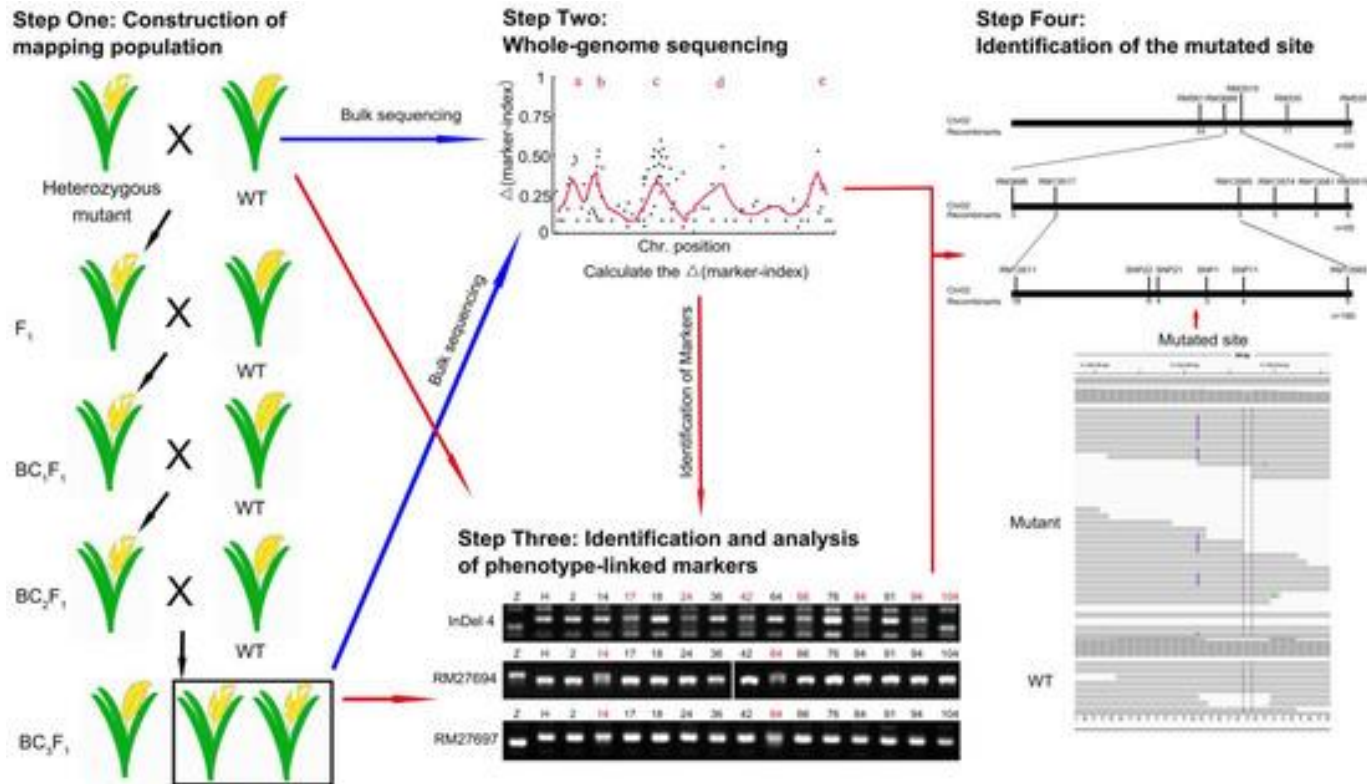


Map based cloning use genetic relationship between a gene and a marker as the basis for searching a gene. **Chromosome walking** is moving toward a gene using a probe for a marker near a gene to select genomic clone near the gene.

Sequencing-based Mapping Method for identification of mutated gene



Sequencing-based Mapping Method for effective identification of mutated gene



Step One: Cross the heterozygous mutant with WT of the other variety four times to generate the BC_3F_1 population.

Step Two: about 30–200 heterozygous plants of BC_3F_1 population were selected to extract DNA, and the DNA samples were mixed with equimolar of each, resequenced, and analyzed. The markers (mainly SNP and InDel) were identified and their marker indexes were calculated to identify the high-quality markers for identifying the candidate regions using the MutMap method.

Step Three: the BC_3F_1 population was used for the analysis of the recombinants at the phenotype-linked markers to identify and narrow the candidate region with the markers.

Step Four: using the resequencing genome data, the mutational sites (the $\Delta\text{marker-index}$ about 0.5) in the target region were identified.

How to create mutant ?

to generate large scale MUTANTS



Exa: Dwarf phenotype or mutant in Arabidopsis

Natural Genetic Variation is the basis for inventing genes

The variation may be produced by **Natural and Artificial mutations** as well as **Sexual crossing**.

NATURAL MUTANTS were generated during species evolution. Generally, the **ratio of natural mutation is very low** at only 10^{-5} - 10^{-8} in higher plants.

Large collection is still available during long evolutionary history. Some of such mutants were harmful or neutral and might be lost during evolution. **Others might exhibit higher resistance to various abiotic/biotic stresses or useful agricultural traits, which were valuable resources for breeding** and gene discovery.

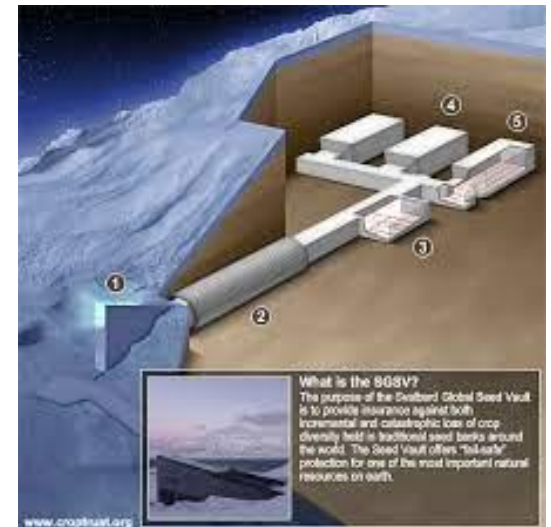
Exa: Dwarf germplasm Dee-geo-woo-gen (Mutant) from China used as donor for release of rice variety IR8.

Application of **cytoplasmic male sterile (CMS)** and **photoperiod-sensitive genic male sterile rice lines**, which are widely utilized to develop hybrid rice seeds for commercial release.

Nestled in the mountains of the world's northernmost community, one building braces for doomsday, both in design and function. The Norwegian archipelago of Svalbard — between continental Norway and the North Pole — houses the Svalbard Global Seed Vault, which stores and protects more than 1.1 million seed varieties from nearly every country. In an area that's home to more polar bears than people, the seed vault acts as a storage hub for humanity's future, thanks to its Arctic-specific construction intended to weather worst-case scenarios.

Norway's Global Seed Vault Steels Itself Against Climate Change

The secure vault at the top of the world safeguards 1.1 million seed varieties for the future



Genetic variation in Nature

is the basis for inventing genes

NATURAL MUTANTS is the outcome of **continuous conflicts** (perturbations and interactions) between **genetic determinants** and **randomness of the systems** (biological and environmental fluctuations) leading to evolution of diversity (**species evolution**).

Generally, the **ratio of natural mutation is very low** at only 10^{-5} - 10^{-8} in higher plants.

Large collection is still available during long evolutionary history. Some of such mutants were harmful or neutral and might be lost during evolution. **Others might exhibit higher resistance to various abiotic/biotic stresses or useful agricultural traits, which were valuable resources for breeding and gene discovery.**

Natural Genetic Variation

Resource to discover novel gene functions

Natural variation provides a relevant complementary resource to discover novel gene functions as well as those allelic variants **that specifically interact with the genetic background and/or the environment** or alleles showing small effects on phenotype, particularly for **traits related to plant adaptation**.

Natural variation is broadly defined as the **within-species phenotypic variation** caused by **spontaneously arising mutations** that have been **maintained in nature by any evolutionary process** including, among others, artificial and natural selection.

Thus, **natural variation embraces the enormous diversity** present within wild plant species as well as most of the genetic variants that are found in domesticated plants. Some of the phenotypic differences existing in wild or cultivated plants are due to **single-gene (monogenic) allelic variants**. However, most of the natural variation is **quantitative** and determined by **molecular polymorphisms at multiple loci and genes (multigenic)**, which are referred to as quantitative trait loci (QTL) and quantitative trait genes (QTGs).

Chemical or Physical (Irradiation) based mutation



Classical **chemical (EMS)** or **irradiation mutagenesis (high energy radiation)** is one of the **most powerful screen approaches** to uncover genes involved in certain genetic pathways.

The **goal** is to cause maximal genomic variation with a minimum decrease in viability.

Among the radiation-based methods, **γ -ray and fast neutron bombardment now supersedes X-ray in most applications** because **γ -ray bombardment is less destructive causing point mutations and small deletions** whereas **fast neutron bombardment causes translocations, chromosome losses, and large deletions**. Compared to chemical mutagens, **both types of radiation cause damage on a larger scale and severely reduces viability**.

Chemical mutagens are easy to use, do not require any specialised equipment, and provide a very high mutation frequency.

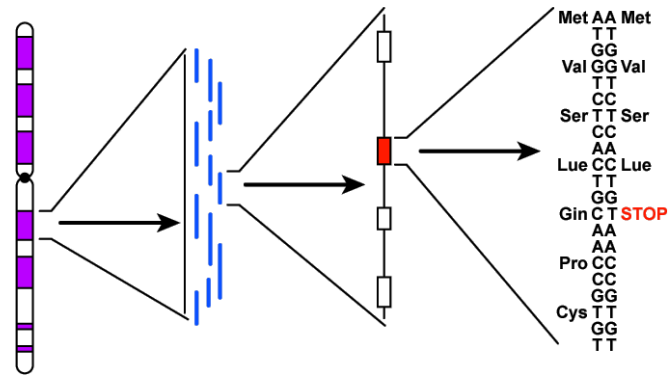
Compared to radiological methods, **chemical mutagens tend to cause single base-pair (bp) changes, or single-nucleotide polymorphisms (SNPs)**, rather than deletions and translocations.

Of the chemical mutagens, **EMS (ethyl methanesulfonate) is today the most widely used**. EMS **selectively alkylates guanine bases** causing the DNA-polymerase to favor **placing a thymine residue over a cytosine residue** opposite to the **O-6-ethyl guanine** during DNA replication, which **results in a random point mutation**. A majority of the changes (70–99%) in EMS-mutated populations are GC to AT base pair transitions.

Chemical or Physical (Irradiation) based technologies for creating mutation



However, the tedious works to hunt down the gene corresponding to a mutant allele by Map-base cloning make an intrinsic limitation of this approach.



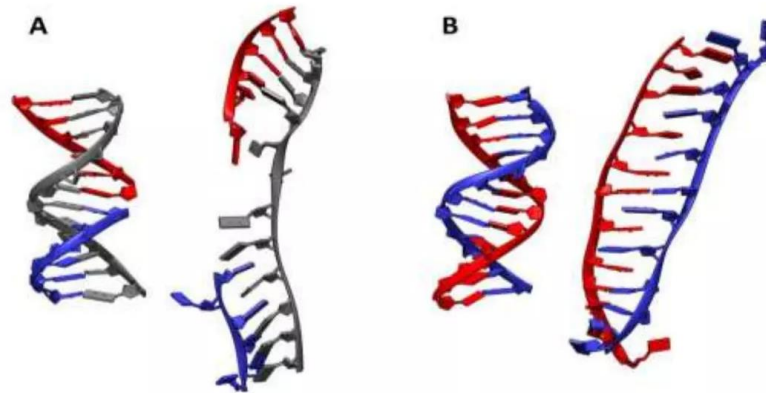
insertion sites in the genome is difficult to trace

Subtractive Hybridization to find out the mutated gene

Subtractive hybridization is a technique for identifying and characterizing differences between two populations of nucleic acids.

It detects differences between the RNA in different cells, tissues, organisms, or sexes under normal conditions, or during different growth phases, after various treatments (ie, hormone application, heat shock) or in diseased (or mutant) versus healthy (or wild-type) cells.

Subtractive hybridization also detects DNA differences between different genomes or between cell types where deletions or certain types of genomic rearrangements have occurred.

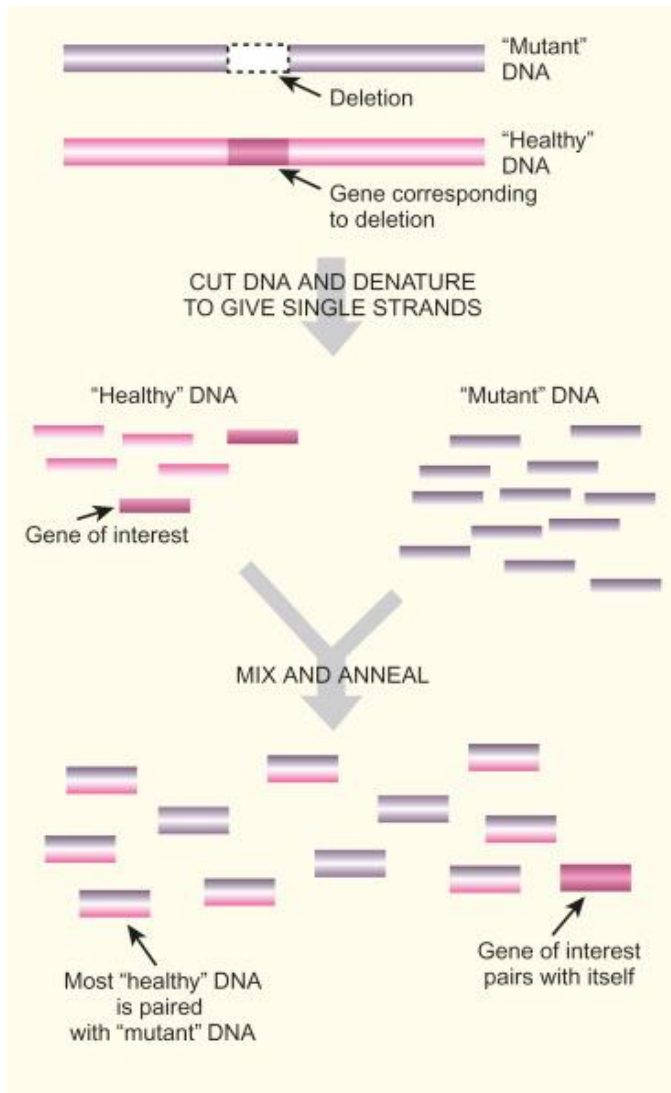


Subtractive Hybridization to find out the mutated gene

Subtractive hybridization can also be used to isolate the genes that are mutated. The total RNA is isolated from both samples (WT and mutated plants), then the mRNA is purified by hybridization to oligo(dT). The WT sample will contain mRNA from genes expressed under no mutation conditions. The experimental sample will contain mRNA from plants that are mutated.

Subtractive hybridization is performed with two samples of **mRNA (WT and mutated plants)**. The basic idea is that the standard mRNA is used to subtract out the corresponding mRNA molecules from the experimental sample. However, two mRNA molecules of the same sequence obviously cannot hybridize together directly. Therefore, the **standard mRNA (WT)** is first **converted to the corresponding double-stranded cDNA** by reverse transcriptase. The **cDNA is bound to a filter** and the **experimental sample of mRNA (mutated plants)** is incubated with the filter. mRNA corresponding to all non-mutated genes in the cDNA is retained by hybridization. Only the **mRNA from gene that is mutated (experimental) remains unbound**. This unbound mRNA (i.e, corresponds to mutated gene) can in turn be converted to cDNA so giving a sample of those genes that are mutated.

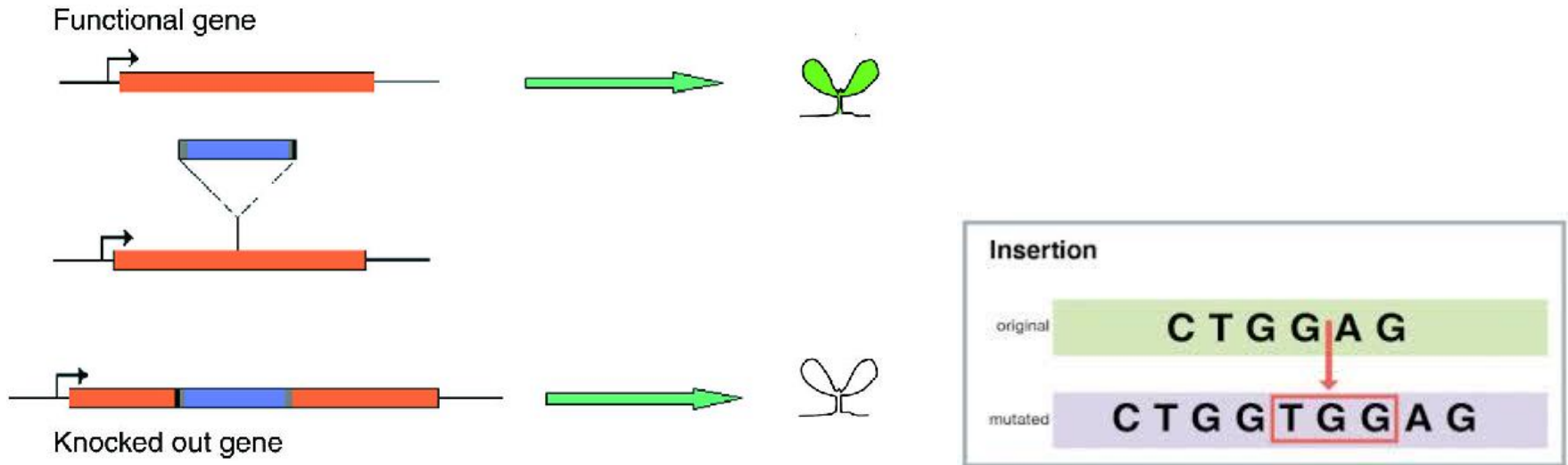
Subtractive Hybridization to find out the mutated gene



The mutant DNA (genome carrying mutant gene) is digested with restriction enzyme 2 and the wild-type DNA (wild type genome) is digested with restriction enzyme 1. Both samples are heated to separate the strands (denature), forming a pool of single-stranded fragments.

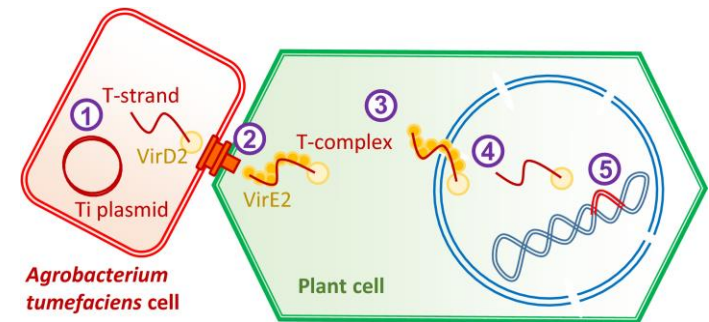
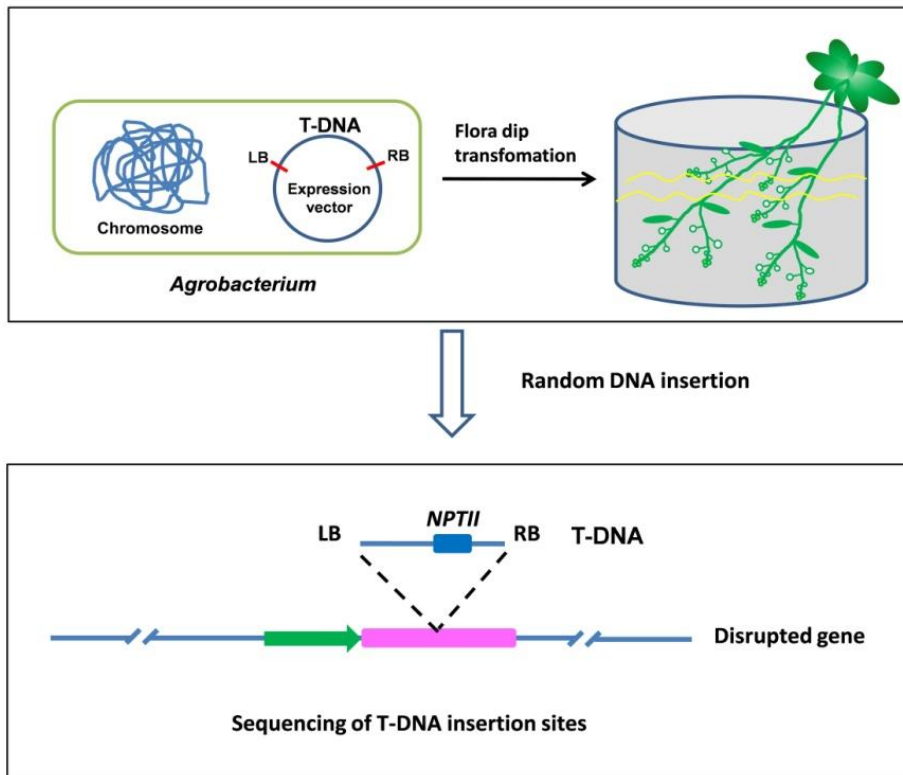
In order to ensure all the wild-type ssDNA (ss: single strand) is hybridized to mutant ssDNA and not to itself, a **large surplus of mutant ssDNA** is mixed with a small amount of wild-type DNA. The DNA is allowed to anneal yielding double-stranded DNA consisting of a **mixture of mutant:mutant, mutant:wild-type, and rare wild-type:wild-type molecules**. Since the ratio of wild-type DNA to mutant DNA is so low, theoretically the only molecule where two wild-type strands anneal should be the missing fragment in the mutant. Because two different restriction enzymes were originally used to digest the different samples of DNA, these desired DNA molecules are the only ones that can be digested with restriction enzyme 1. This allows them to be cloned into a vector cut with restriction enzyme 1 and captured.

Mutagenesis by Biological means



- Single-insertion elements
e.g., T-DNA in plants
- Transposable elements
e.g., AC/DS elements in plants

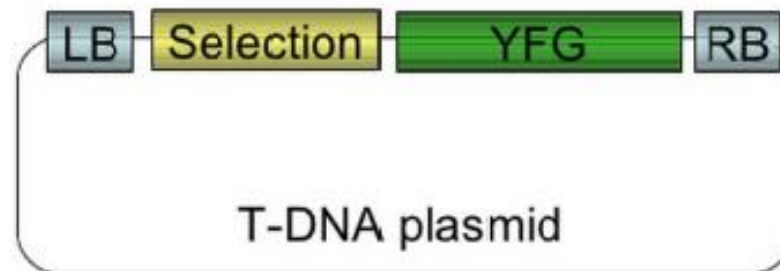
T-DNA insertion disrupts the native gene sequence causing Insertional Mutation



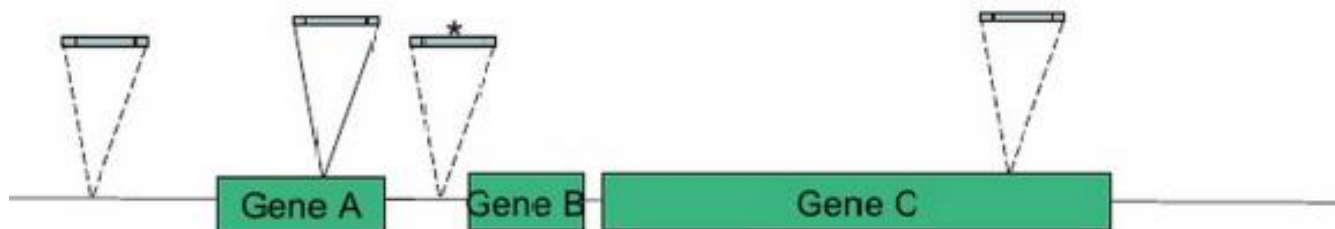
Agrobacterium evolved the capability to transfer a DNA fragment (T-DNA) of their own genome into the nuclear genome of plants. This natural process of inter-kingdom horizontal gene transfer (HGT) inspired the development of a wide spectrum of tools and technologies dedicated to genetic engineering of plants and some other eukaryotes.

Single-insertion elements (T-DNA in plants), Once inserted, can't move again
Tagged, therefore easier to isolate gene involved

T-DNA insertion in plant genome is always at Random locations

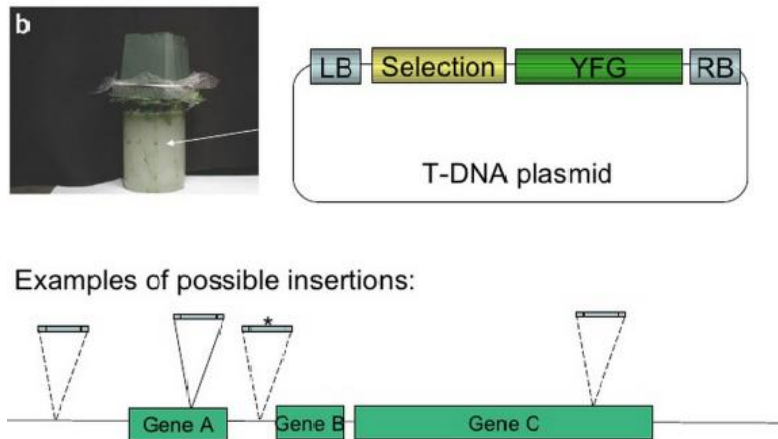


Examples of possible insertions:



Generate mutants

by T-DNA insertional mutagenesis (or T-DNA tagging)

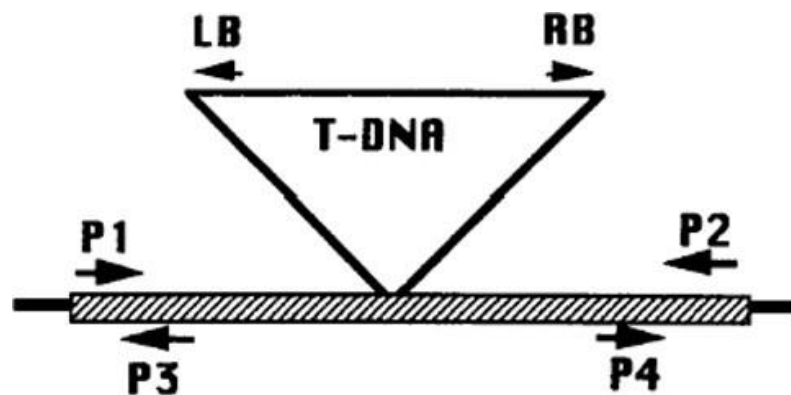


Mutagenesis in Arabidopsis is routine:

- introns are small
- there is very little intergenic material
- insertion of a piece of T-DNA on the order of 5 to 25 kb in length generally produces a dramatic disruption of gene function

- **T-DNA** can make effective interruption of genes
- Usually T-DNA insertion is low copy in nature (1.5)
- **T-DNA insertion is random** (no preference for any genomic locations)
- **Mutation will be stable** (as T-DNA once get inserted, can not move again)
- If a large enough population of T-DNA-transformed lines is available, one has a very good chance of finding a plant carrying a T-DNA insert within any gene of interest.

PCR-based identification of T-DNA insertion mutants



LB, primer homologous to LB of T-DNA,

RB, primer homologous to right border of T-DNA

P1, forward primer homologous to 5' end of gene,
P2, reverse primer homologous to 3' end of gene,

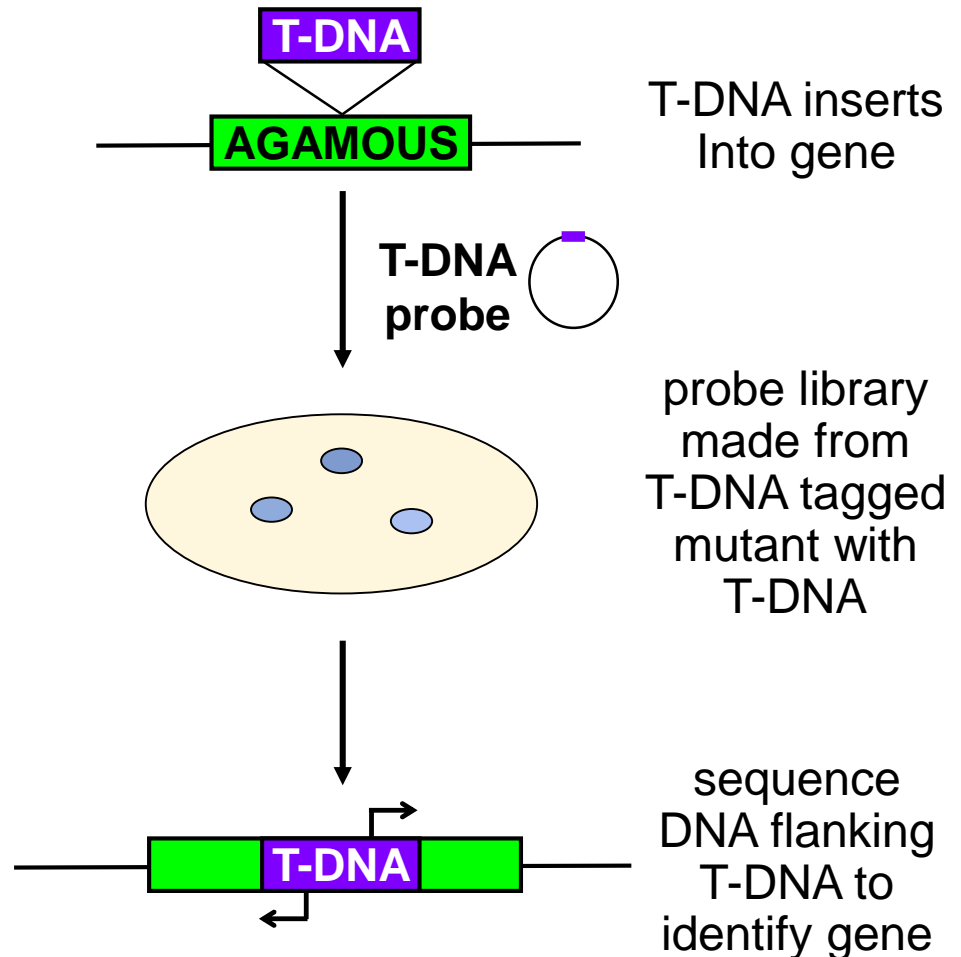
P3, reverse primer homologous to 5' end of gene,
and P4, forward primer homologous to 3' end of gene

Tens of thousands of T-DNA generated transformants have now been produced in *Arabidopsis*. These mutants contain insertions which are randomly distributed in the gene and genome.

T-DNA border-specific primers in combination with sequence specific primers to **identify insertion mutants** based on the generation of a **PCR** product among 6000 transformed lines.

Identification of Insertional mutant by Genomic library and PCR based detection

- T-DNA inserts into plant chromosome
- Screen for mutations that affect flower formation
- Make genomic library from mutant DNA
- Probe with T-DNA
- Identify mutant gene



Mutagenesis by mobile genetic elements

Barbara McClintock and JUMPING genes



A **transposable element (TE or transposon)** is a DNA sequence that can change its position within a genome, sometimes creating or reversing mutations and altering the cell's Genome size. Transposition often results in duplication of the TE. Barbara McClintock's discovery of these **jumping genes** earned her a Nobel Prize in 1983



BARBARA
McCLINTOCK'S



Discovered largely from cytogenetic studies in maize, but since found in most organisms.

1983 Nobel Prize in Physiology and Medicine)

profiles.nlm.nih.gov/LL/

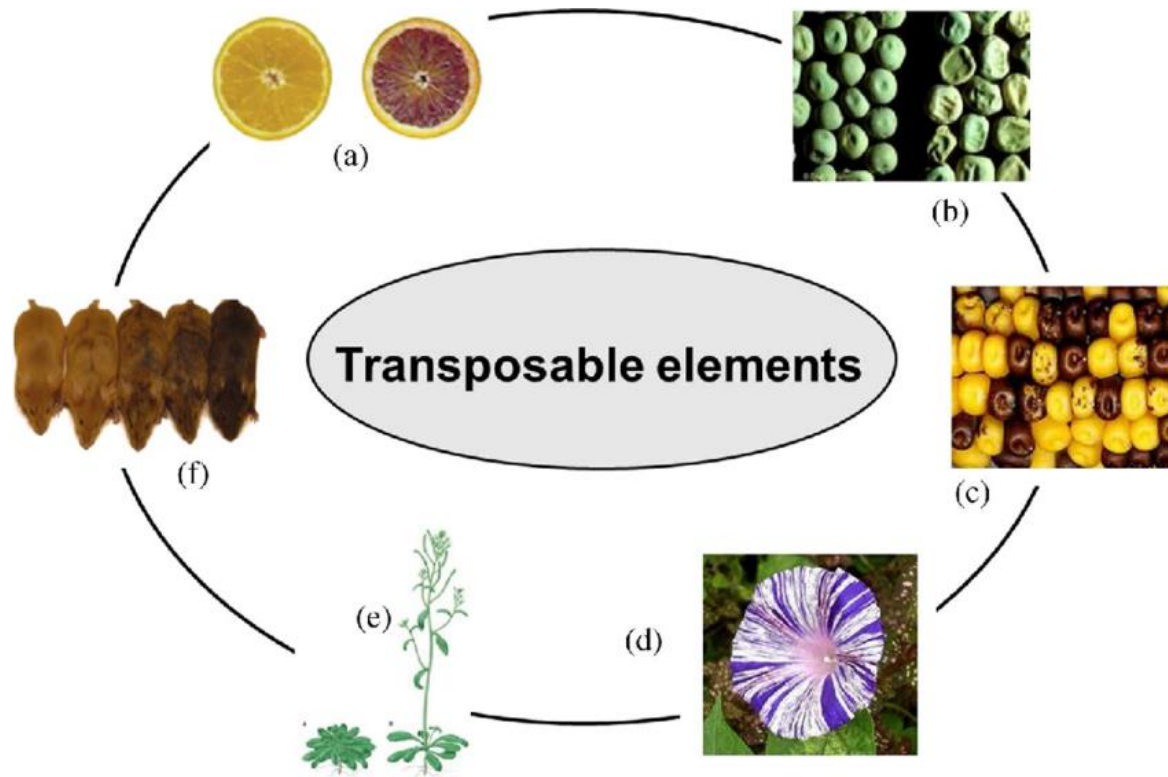


Corn (maize) varieties



“for her discovery of mobile genetic elements”

Effects of TE insertions on gene structure and function

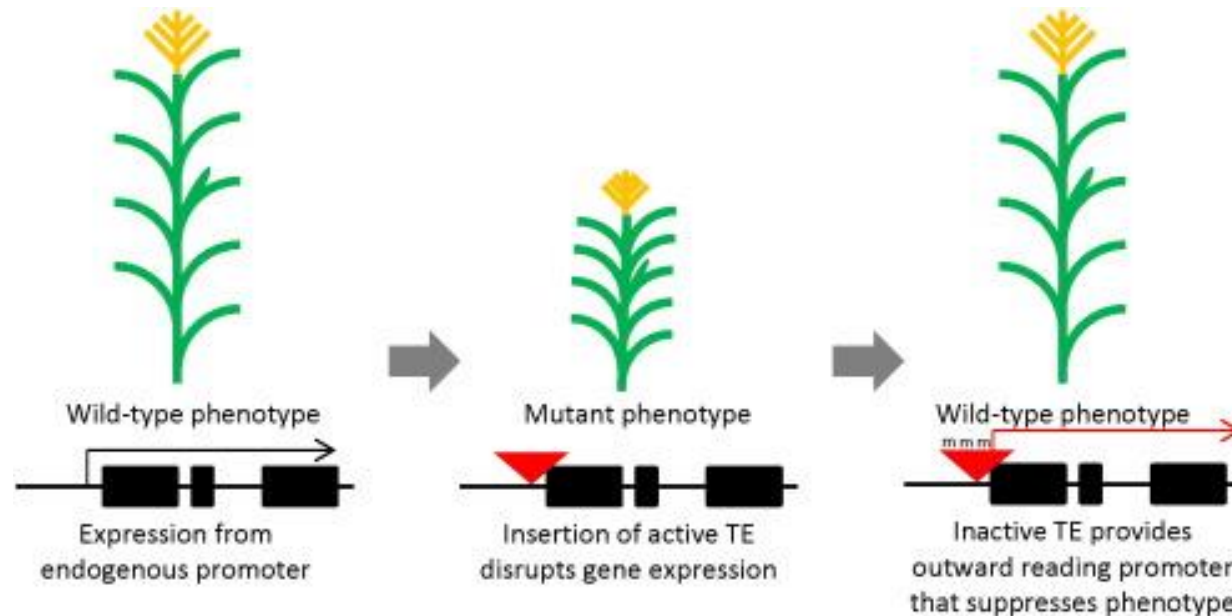


(a) Insertion of LTR retrotransposon upstream of the Ruby gene in the blood orange provides a new promoter controlling the expression of a gene for flesh colouration of the fruit

(b) Insertion of TE into the rugosa locus encoding a starch-branching enzyme in pea results in a wrinkled seed.

(c) Insertion of TE into a gene that produces anthocyanin pigments leads to its inactivation and changes in the colour: yellow seeds in maize, (d) White sectors in Petunia hybrida. (e) siRNA-controlled methylation of a TE insertion influence the expression of the FLC locus (a gene delaying flowering) and leads to earlier flowering in A. thaliana. (f) Alterations in the transcription levels of genes by antisense transcription from adjacent TE insertions as observed for the agouti colour gene in mice.

Transposable elements (TEs) contribute to **genomic innovations**, as well as **genome instability**



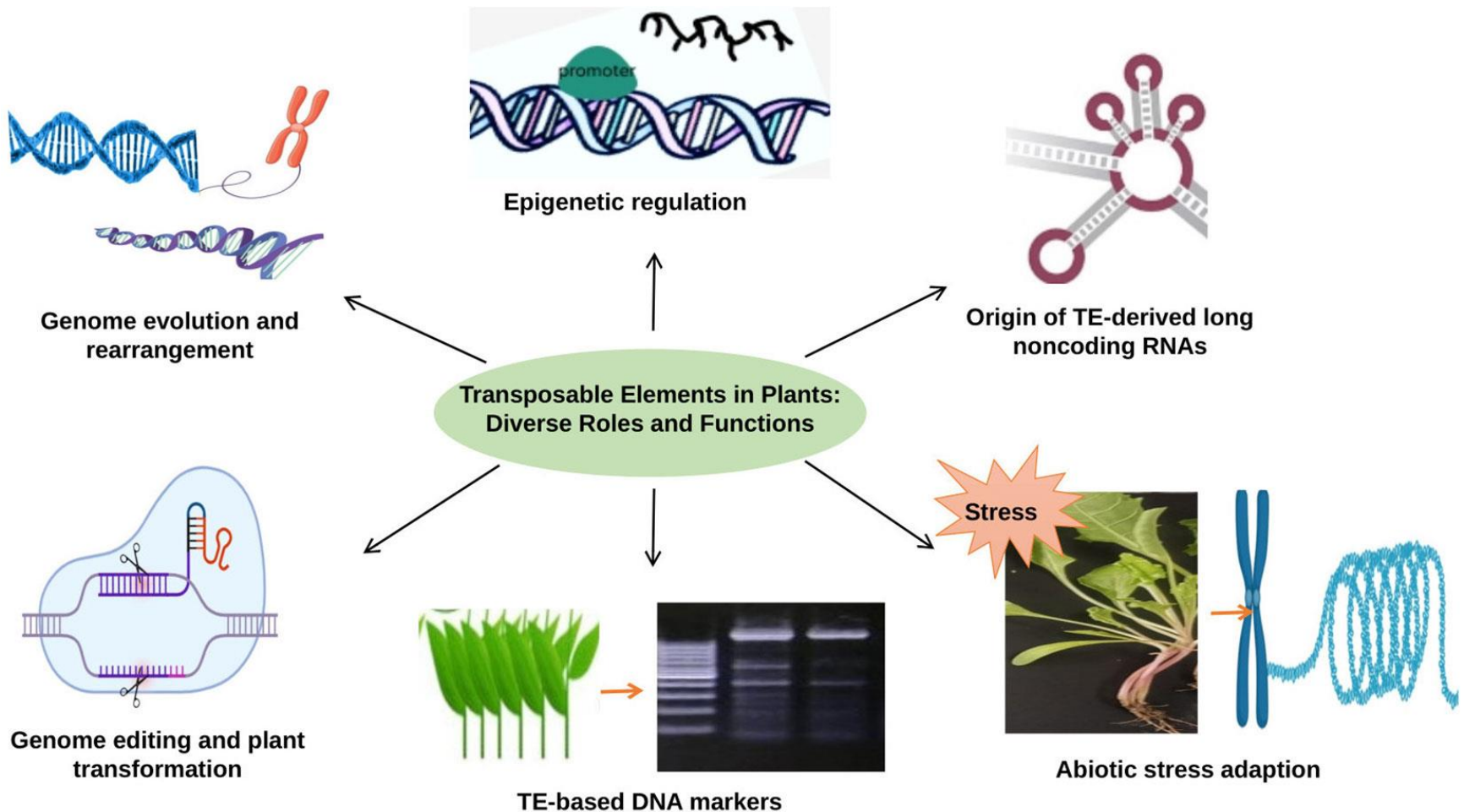
TE insertions within genes can **result in aberrant or novel transcripts**.

Second, **TEs can provide novel alternative promoters**, which can lead to new expression patterns or original coding potential of an alternate transcript.

Third, **TE insertions near genes** can influence regulation of gene expression through a variety of **mechanisms**. For example, TEs may provide novel *cis*-acting regulatory sites behaving as enhancers or insert within existing enhancers to influence transcript production.

Alternatively, **TEs may change chromatin modifications** in regions near genes, which in turn can influence gene expression levels.

Role of Transposons in Plants



Transposable elements (TEs) contributing to Plant Diversity



**ACTIVATION OF
CLASS 1 OR CLASS 2 TEs**



**EFFECTS OF TEs ON GENES
AND PHENOTYPES**

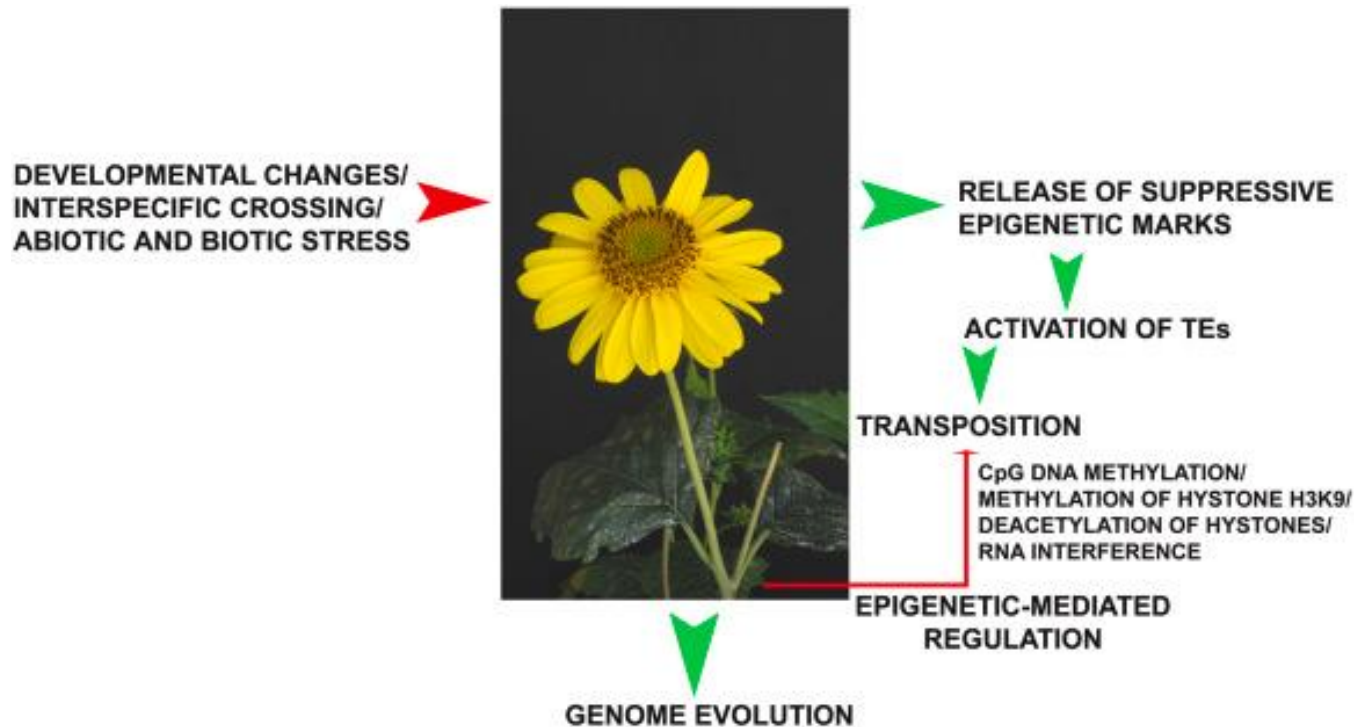


TEs have caused the genomes to evolve towards levels of greater complexity.

The hosts control the spread of TEs through epigenetic mechanisms.

Epigenetic mechanisms do not give a free ride to TEs though

Impact of Transposable Elements on the evolution of complex living systems and their Epigenetic control



Transposable elements (TEs) contribute to **genomic innovations**, as well as **genome instability**. Both plant and animal hosts have evolved **epigenetic mechanisms to reduce the impact of TEs**, both by directly silencing them and by reducing their ability to transpose in the genome.

However, TEs have also been co-opted by both plant and animal genomes to **perform a variety of physiological functions**, ranging from **TE-derived proteins acting directly in normal biological functions** to **innovations in transcription factor activity** and also **influencing gene expression**. Their presence, in fact, can affect a range of features at genome, phenotype, and population levels.

Transposable Elements (Transposons)

- DNA elements capable of moving ("transposing") about the genome
- Discovered by Barbara McClintock, largely from cytogenetic studies in maize, but since found in most organisms
- She was studying "variegation" or sectoring in leaves and seeds
- She liked to call them "controlling elements" because they affected gene expression in myriad ways

Insertional mutagenesis

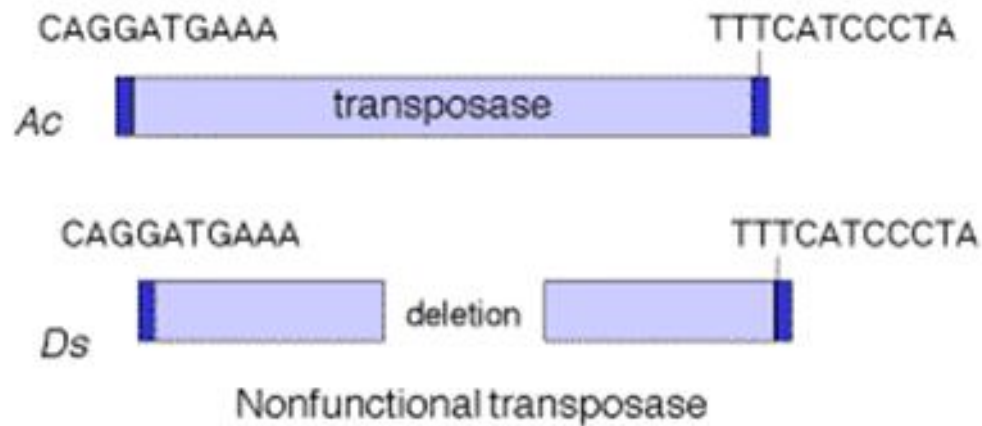
- Single-insertion elements

exa: T-DNA in plants

- Transposable elements

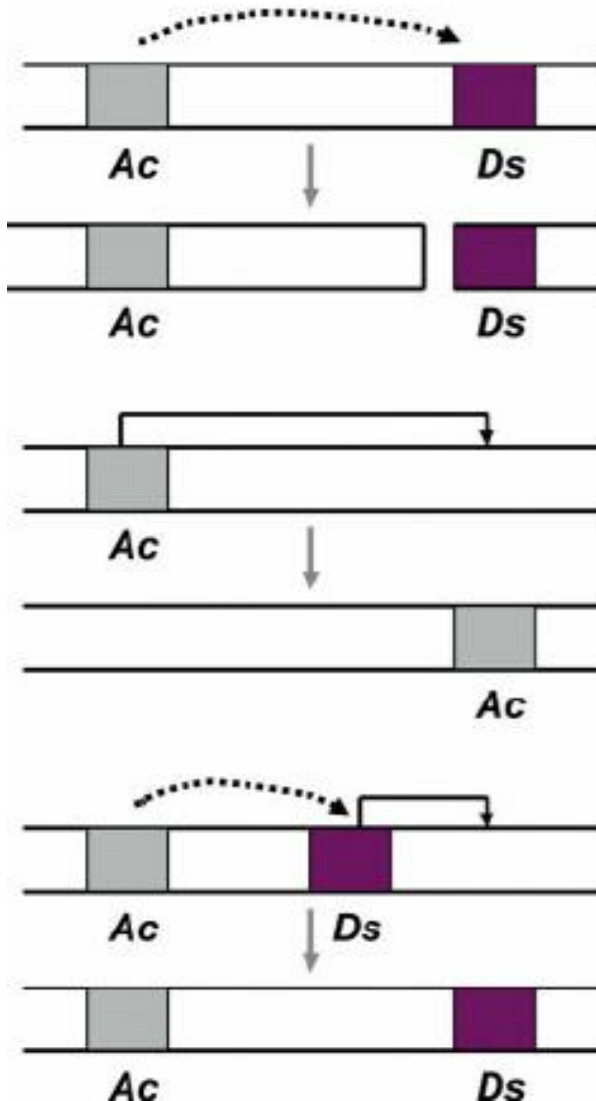
exa: Maize AC/DS elements in plants (**Transposon**)
Rice Tos17, a Copia like element (**Retrotransposon**)

Transposable Elements (Transposons)



Maize **autonomous** (*Ac*) and **nonautonomous** (*Ds*) transposable elements

Transposable Elements (Transposons)



Ac activates breakage at ***Ds***. Loci may be on different chromosomes.

Ac can promote its own transposition, or that of ***Ds***, to another site either on the same chromosome or on a different one.

Ds cannot move unless ***Ac*** is present in the same cell.

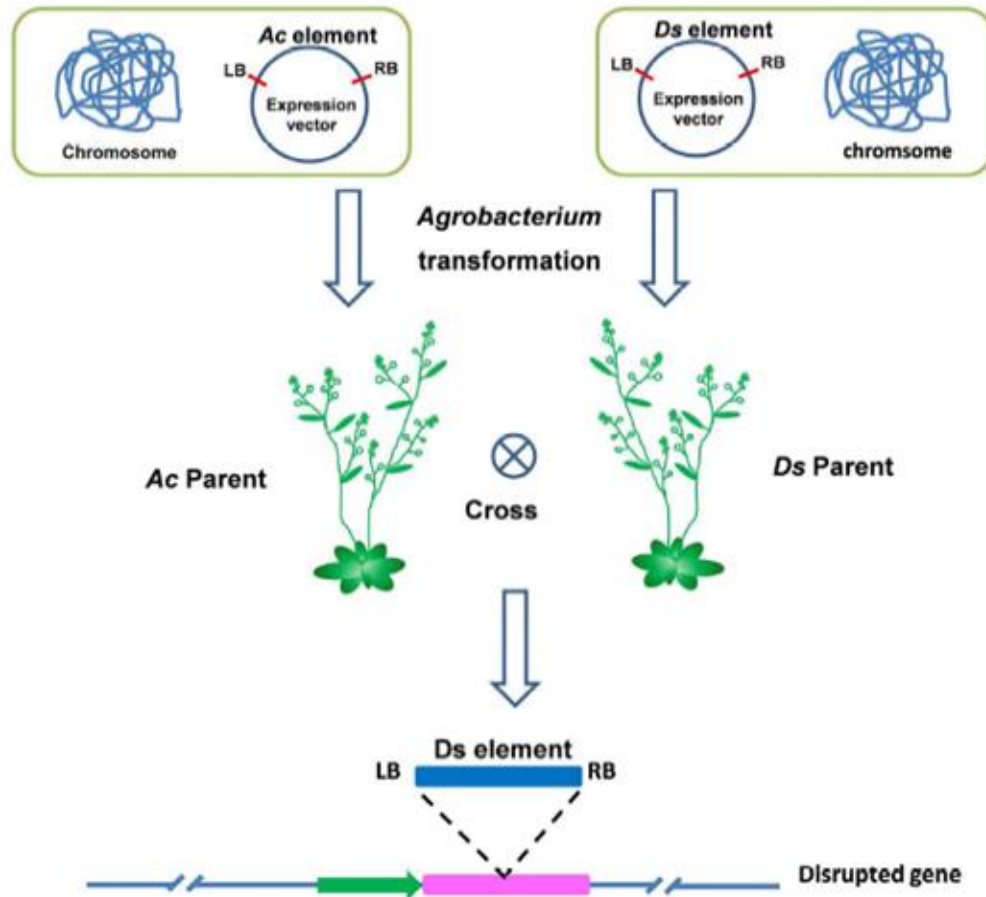
Ac is AUTONOMOUS
Ds is NONAUTONOMOUS



Ac and *Ds* elements

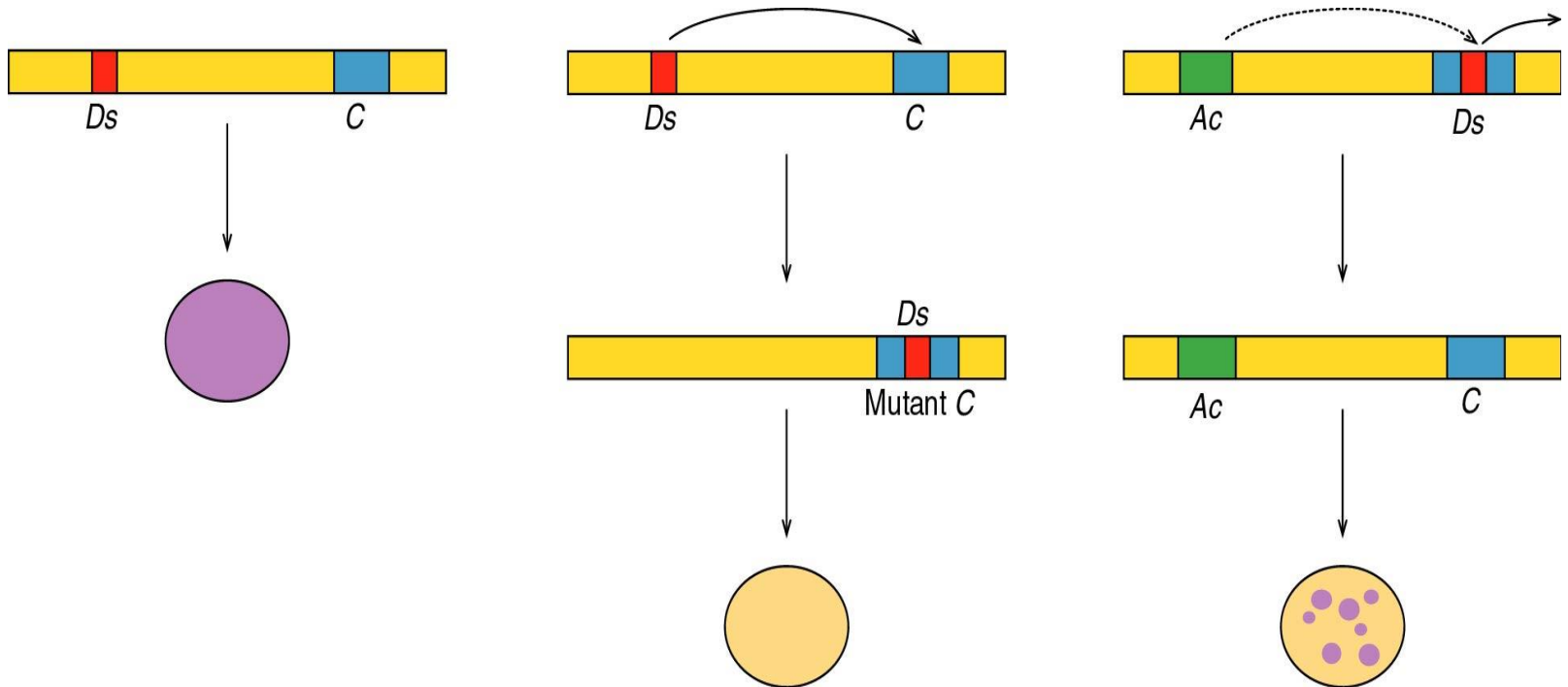
- *Ds* is derived from *Ac* by internal deletions
- *Ds* is not autonomous, requires *Ac* to move
- Element termini are an imperfect IR
- *Ac* encodes a protein that promotes movement - **Transposase**
- Transposase **excises element at IR**, and also cuts the target

How to develop a Stable insertional mutant lines, by combining two mobile elements (*Ac/Ds* system) ???



Scheme for the generation of insertion mutant lines by *Ds* transposon

Somatic Excision of *Ds* from *C*



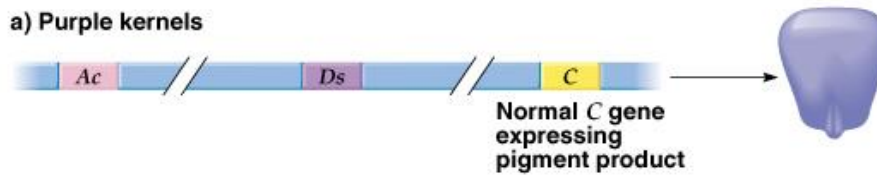
Wild type

Mutant

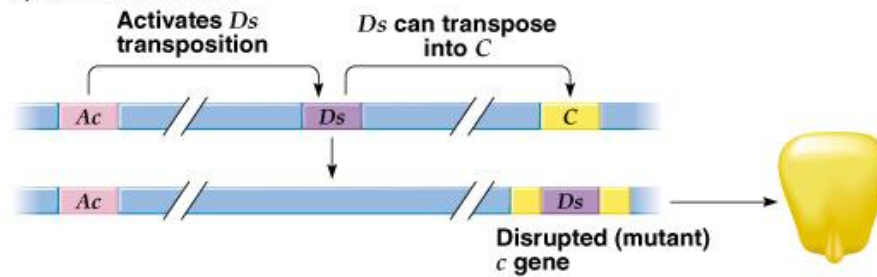
Sectoring

Transposon effects on corn kernel color

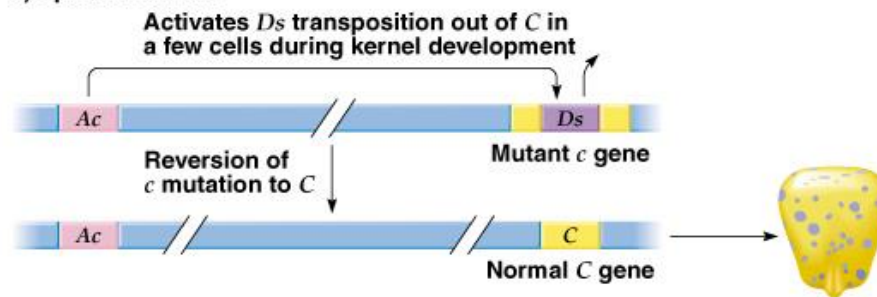
a) Purple kernels



b) Colorless kernels



c) Spotted kernels



Advantage of using T-DNA as insertional mutagen as opposed to Transposons ?

T-DNA insertions will not transpose subsequent to integration within the genome and are therefore **chemically and physically stable** through multiple generations.

The mobility of transposons is not necessarily a bad thing, however. In situations in which **multiple members of a gene family** are arranged in tandem along a chromosome, the ability of transposons to “hop” to nearby locations provides a convenient method for creating mutations within all of the members of the gene family within a single plant.

Limitations with Loss-of-function approaches

- Presence of duplicate genes, where loss of function in one copy will be compensated by other copies (many genes belong to gene families)
- Existence of alternative metabolic pathways and regulatory networks
- Not identify genes whose products required for survival of plants (Loss of these genes would lead to lethality)

Insertional mutagenesis

- Single-insertion elements

exa: T-DNA in plants

- Transposable elements

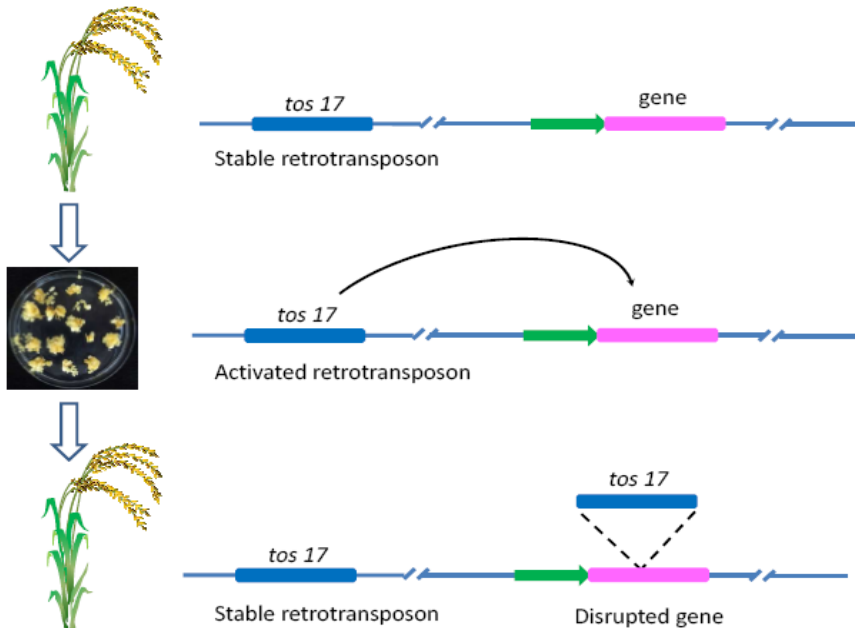
exa: Maize AC/DS elements in plants (**Transposon**)
Rice Tos17, a Copia like element (**Retrotransposon**)

Regulation of the activity of retrotransposons

- The expression of retrotransposons in animal and yeast is under the control of hormonal and developmental factors.
- Regulation of the expression of retrotransposons in plants is not yet fully understood because of the absence of comparative studies in different plant tissues.
- Transcription of most of the active plant elements characterized to date is largely quiescent (in a state or period of inactivity or dormancy) during normal development but can be induced by biotic and/or abiotic stresses, including cell culture, wounding, and pathogen attack.

Copia-like retrotransposons in rice for large scale insertional mutagenesis

(can duplicate and paste to elsewhere in the genome)



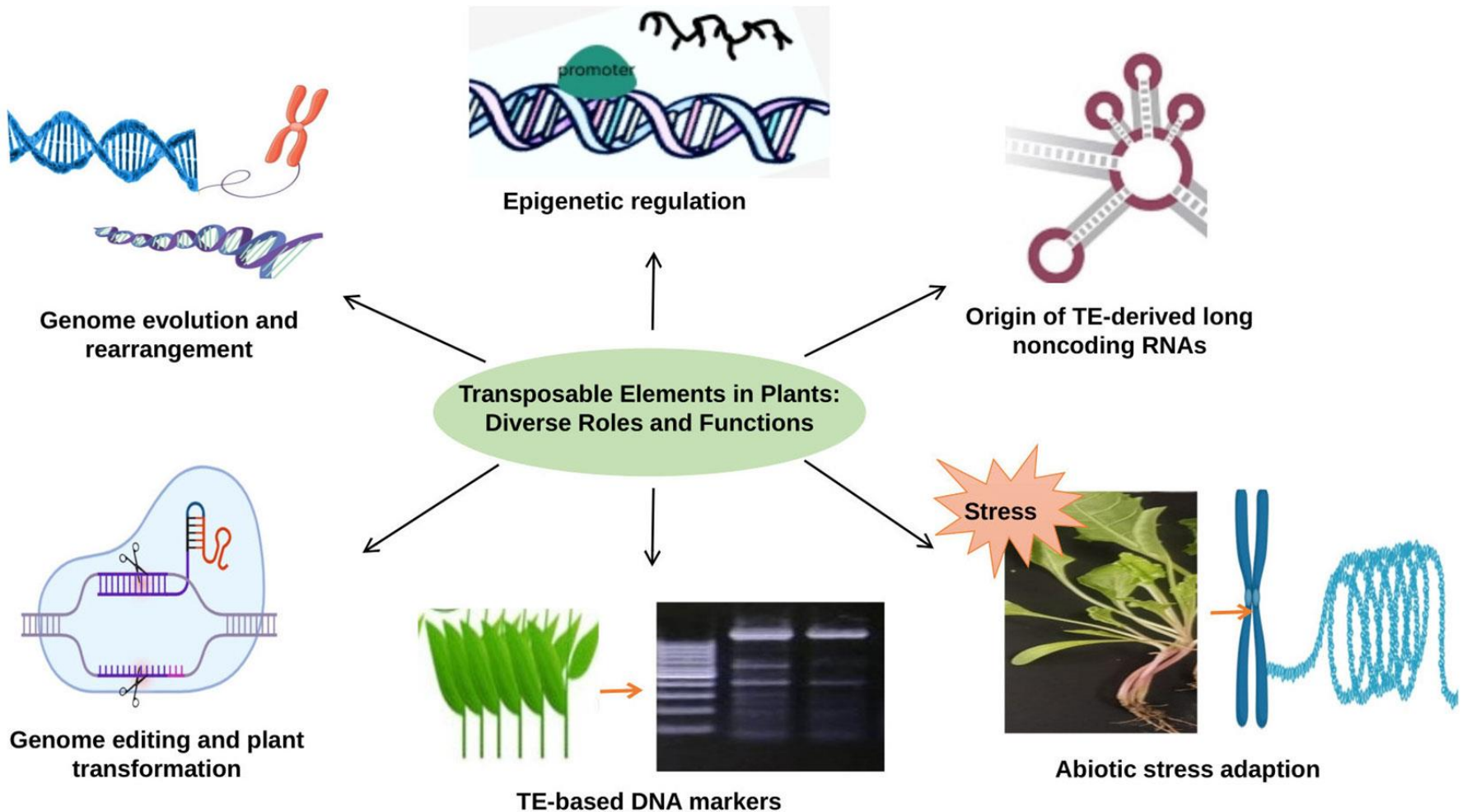
***Tos17* owns several special features** that make it suitable for engineering large scale insertional mutagenesis:

a) **copy number of *tos17* is quite low**, ranging from one to five among rice cultivars.

For example, the genome of cv. Nipponbare, the selected cultivar for the IRGSP (International Rice Genome Sequencing Project), contains only **two native copies of *tos17***.

- b) Transposition of *tos17* is **inactive under normal conditions** but only **activated in the callus** by tissue culture and then **becoming stable again in the regenerated plants**.
- c. Transposition site of *tos17* **prefers gene-dense regions** over centromeric heterochromatin regions with a three times **higher insertion frequency in genic regions than in intergenic regions**.
- d. Its size is just a little bit over 4kb and its insertion sequence is clearly known for **flanking sequencing**.

Role of Transposons in Plants



Role of Transposons in Plants

Transposable elements (**TEs**) form the **major part of 'junk DNA'** in all eukaryotic genomes. These DNA elements have the **potential to be mobile** and therefore to **induce genomic changes** and **reshape genomes over the course of life**.

TEs account for about **21% of the *Arabidopsis* genome**, **40% in rice**, **60% in tomato**, **80% in wheat** and **up to 85% in maize**.

Especially **retrotransposons constitute** the predominating part of plant species with big genomes, such as tomato and maize.

Although TEs are major drivers of genome evolution and **remnants of massive TE bursts are visible in plant genomes**, **transpositional activity is largely prevented** through epigenetic silencing by DNA methylation, histone modification and small RNAs in order to **maintain genome integrity**.

Nevertheless, there is **mounting evidence that TEs participate in the regulation of plant gene expression** upon changing environmental conditions. **Expression and transposition activity of quiescent TEs** upon plant abiotic and biotic stresses is well known.

Activation of TEs upon stress is often **mediated through** de-repression of the silenced epigenetic state or the activation by a transcription factor (TF).

Next, **stress-activated TEs** have the ability to alter the expression of genes flanking their insertion sites, which **leads to phenotypic plasticity** and **adaptation to stress**.

Role of Transposons in Plants

The **relationship between stress and TEs is complex**: some studies also report **TE repression** or **harmful effects of TE activation upon stress**. Moreover, the genomic context of the site of TE insertion defines their specific role in gene regulation.

TEs can exert a regulatory role in host gene expression in multiple ways. Insertion of a TE in or near a gene can lead to new transcription start sites and promoter behavior, disrupt existing and/or create novel regulatory motifs, or spread the chromatin state of the TE to the gene's genomic context.

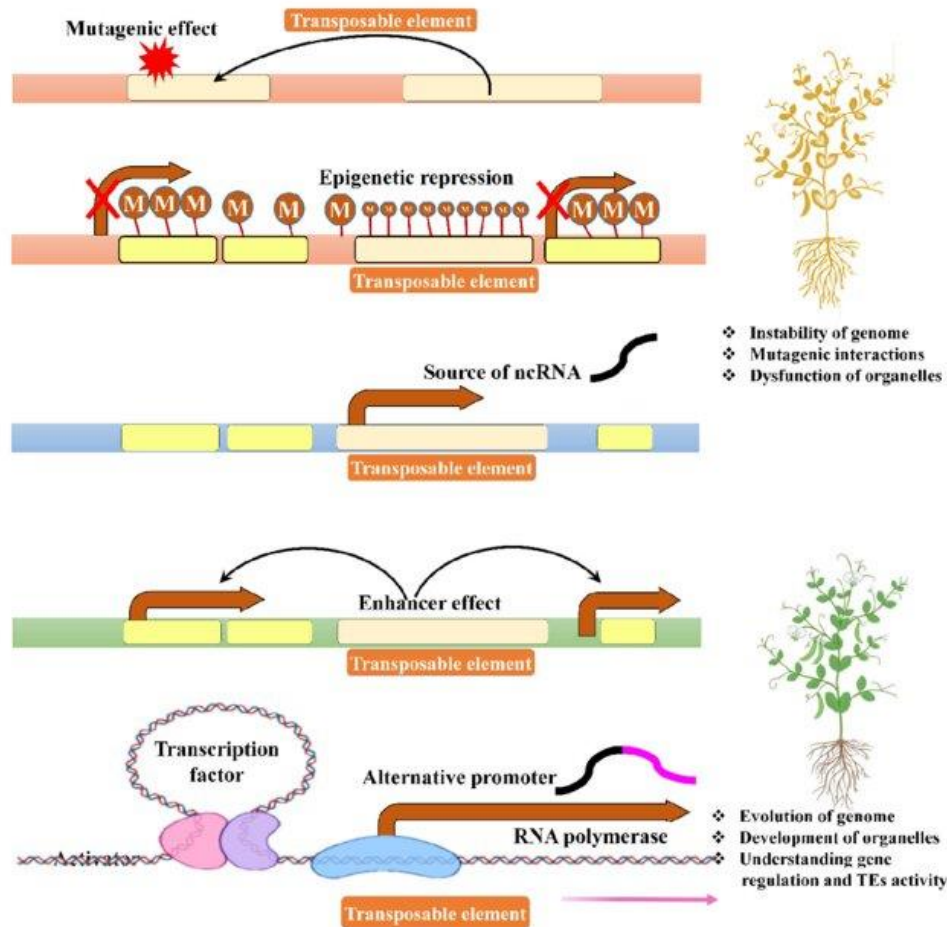
As a classroom example, the **emergence of the melanism phenotype in British peppered moths during the industrial revolution** is caused by a TE insertion into the first intron of the cortex gene that increases the abundance of the transcript.

In *A. thaliana*, the **ONSEN LTR retrotransposon is activated in response to heat stress** due to heat response factors recognizing a regulatory sequence in the promoter of ONSEN. As a consequence, the **insertion of ONSEN was shown to induce the transcriptional up-regulation of neighboring genes upon heat stress**.

In tomato, the **Rider Copia retrotransposon**, which is **triggered by drought stress and abscisic acid signaling**, contains several environment-responsive cis-regulatory motifs, such as Dehydration Responsive Elements (**DRE**), in its promoter. The above examples **illustrate how single TEs can drive the evolution and diversification of stress gene regulatory networks (GRNs)**.

TEs frequently contain transcription factor binding sites (TFBS) or regulatory motifs, which they spread through the genome by transposition.

Regulatory roles of transposable elements (TEs)



TEs are a rich source of host genome innovations. TE functions are either harmful or beneficial to the host genome, and their integration in the genome may induce deleterious mutations.

Silenced TEs, mostly covered with DNA methylation, can affect the expression of nearby genes. In contrast, active TEs can act as regulatory elements by producing noncoding RNA (ncRNA) and alternative promoters.

Are transposons, a blessing curse ?

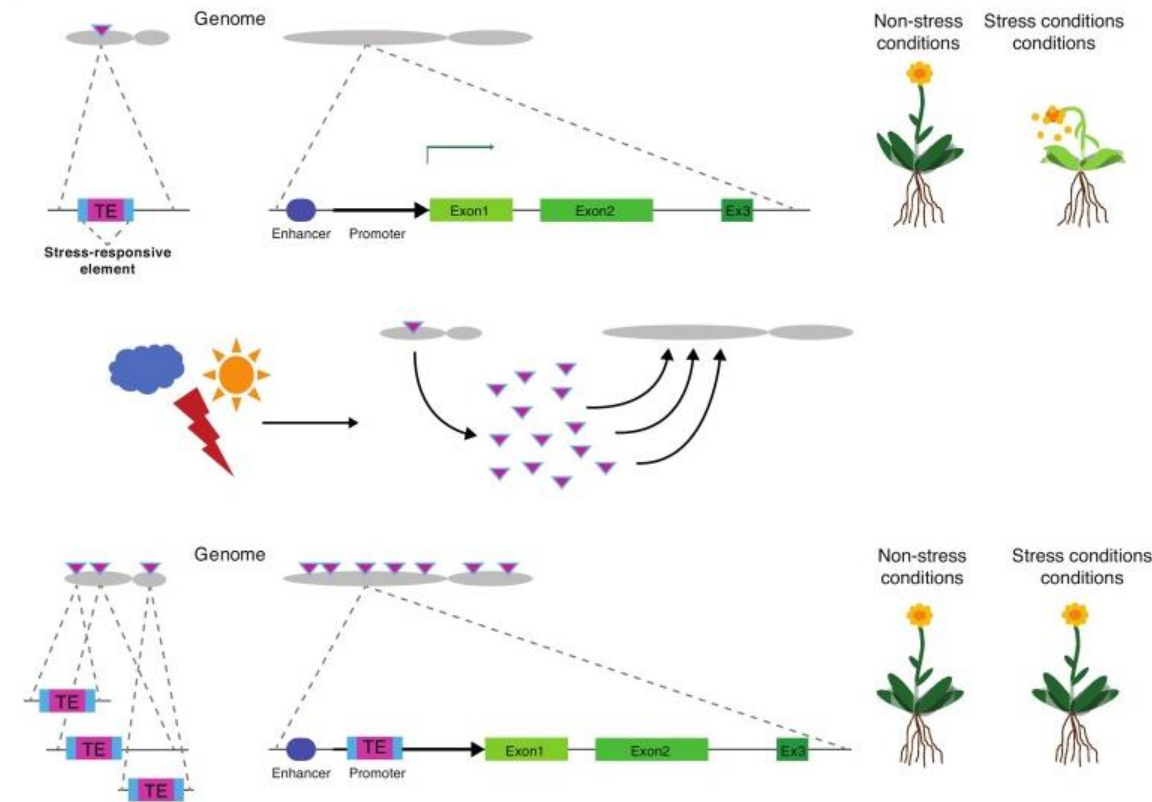
Transposon-based evolution of stress resistance

Stress-induced new insertions of a transposon create many deleterious genotypes,

while those with TE-linked enhanced stress resistance are enriched by natural selection or breeding over time.

Why Abiotic stress conditions would induce transposition of certain TEs?
and/or regulate genes close to residing transposons?

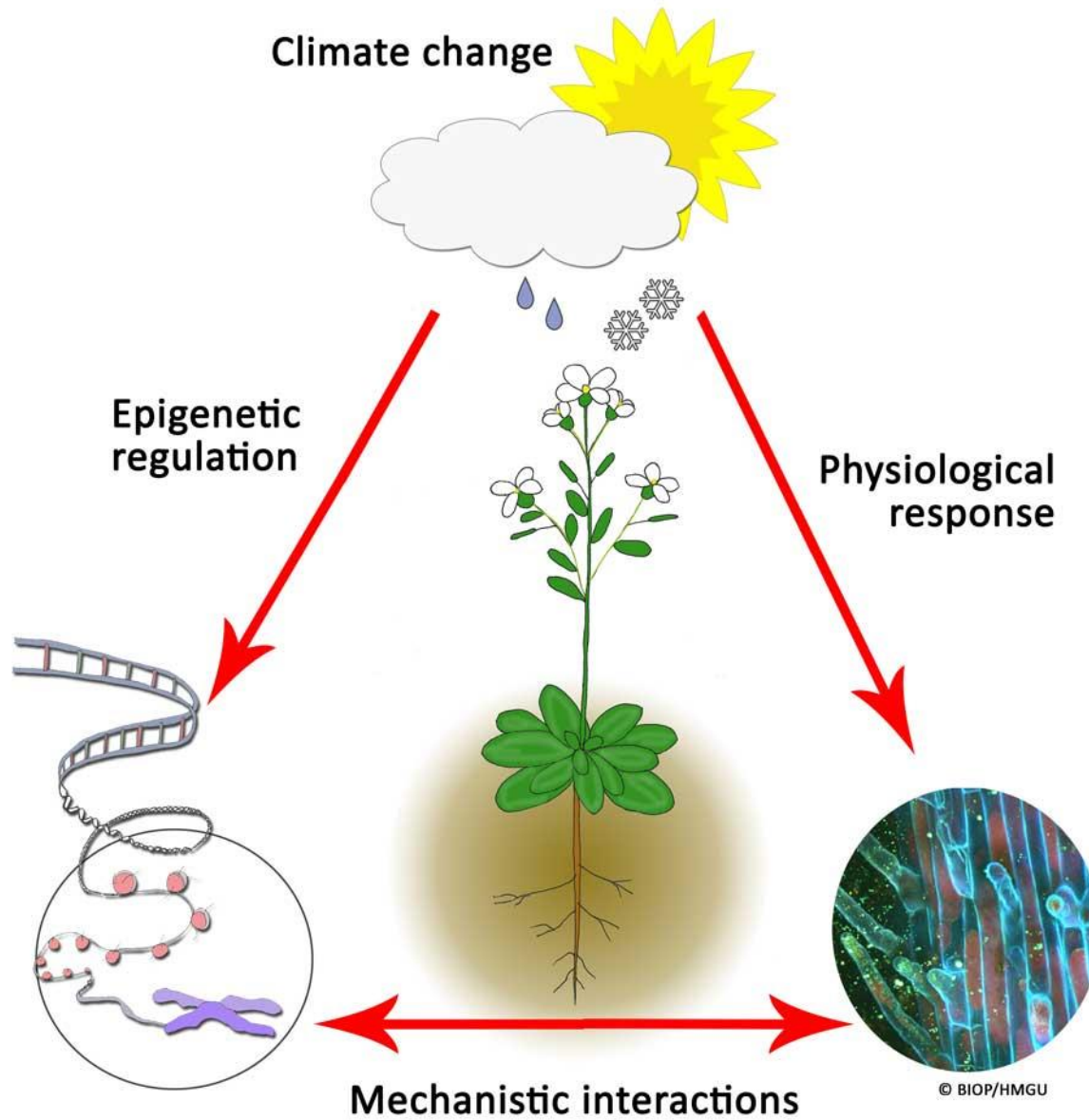
Consequences of TE insertions for adjacent genes



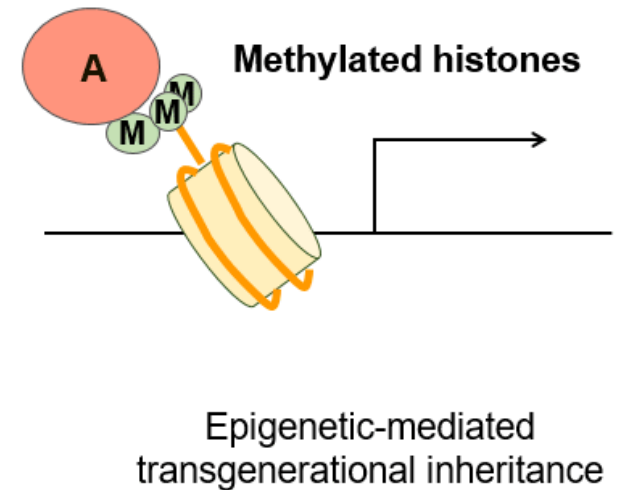
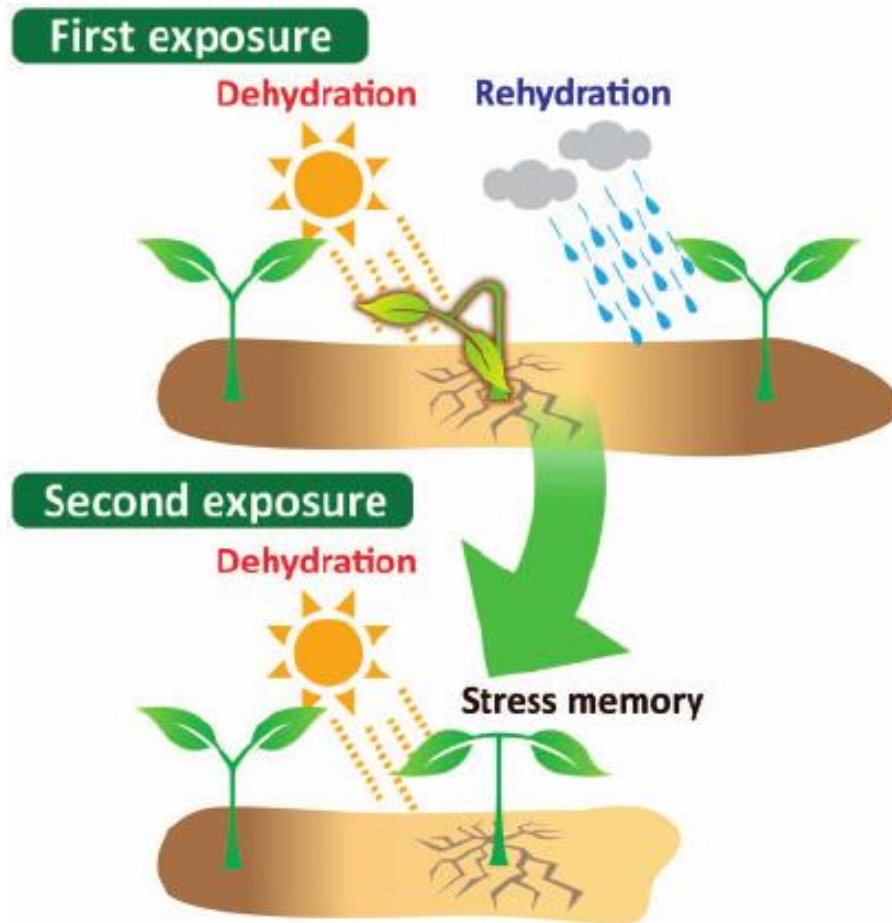
Environmental stress induces transcriptional activation of a TE containing a stress-responsive element (light-blue borders) already present in the genome.

Such **burst of TE expression** and **amplification of extrachromosomal DNA** result occasionally in new insertions at different genomic locations.

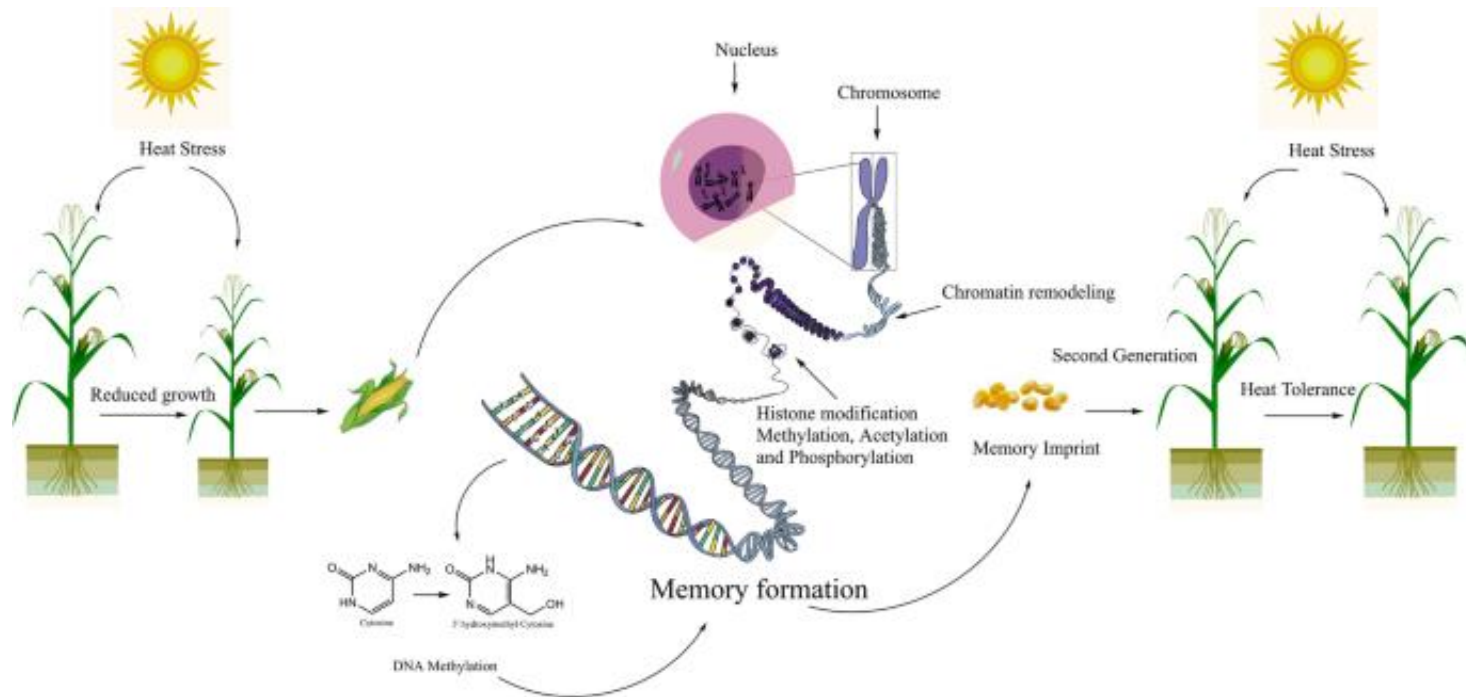
If inserted within the promoter of a protein-coding gene, the **stress-responsiveness of the TE** was maintained and the **gene product is beneficial for stress tolerance**, this combination increases resistance to the stress that initiated the TE burst in the past.



Plant Models of Transgenerational Epigenetic Inheritance

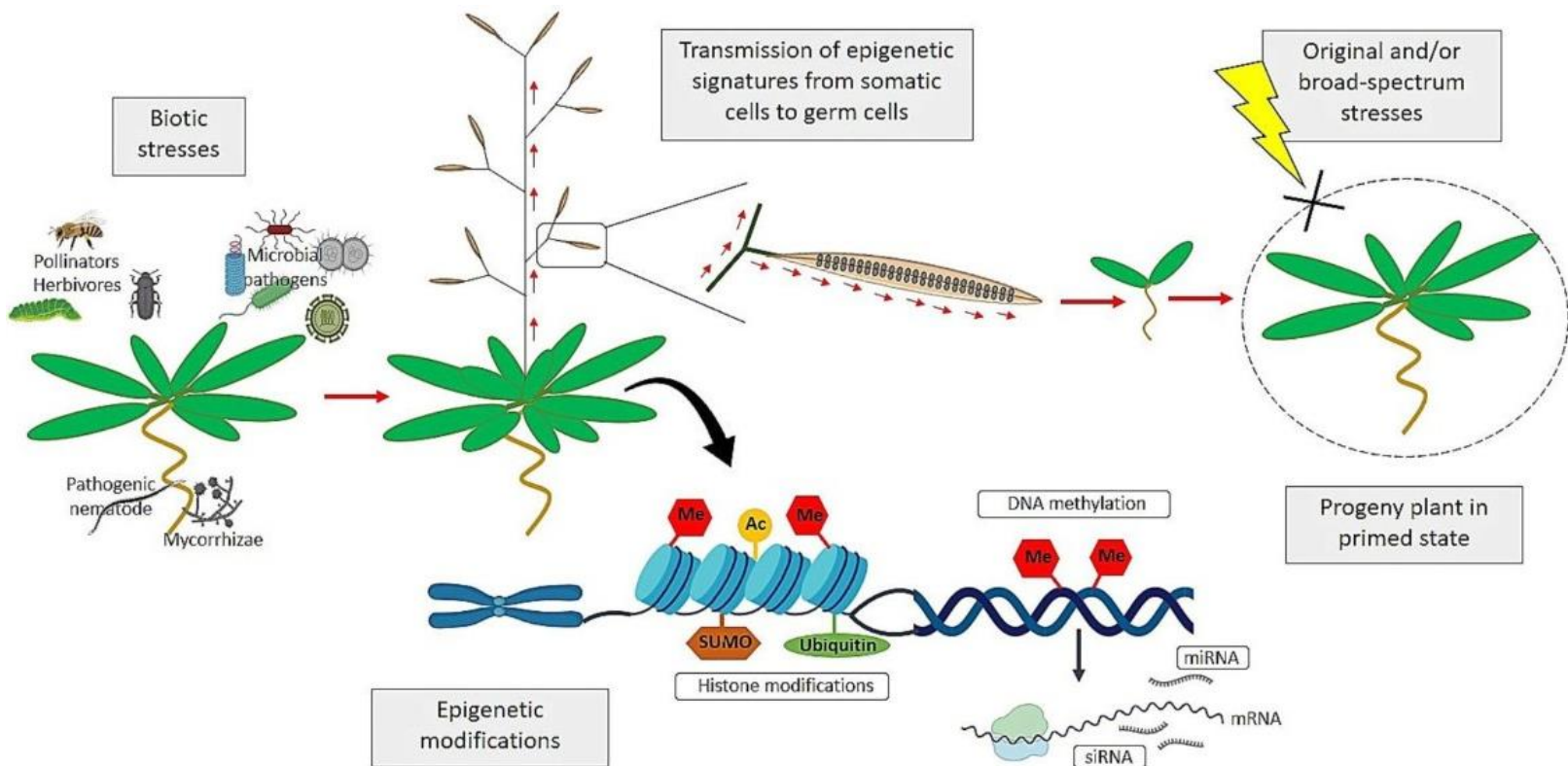


Epigenetics and transgenerational memory in plants under Heat stress



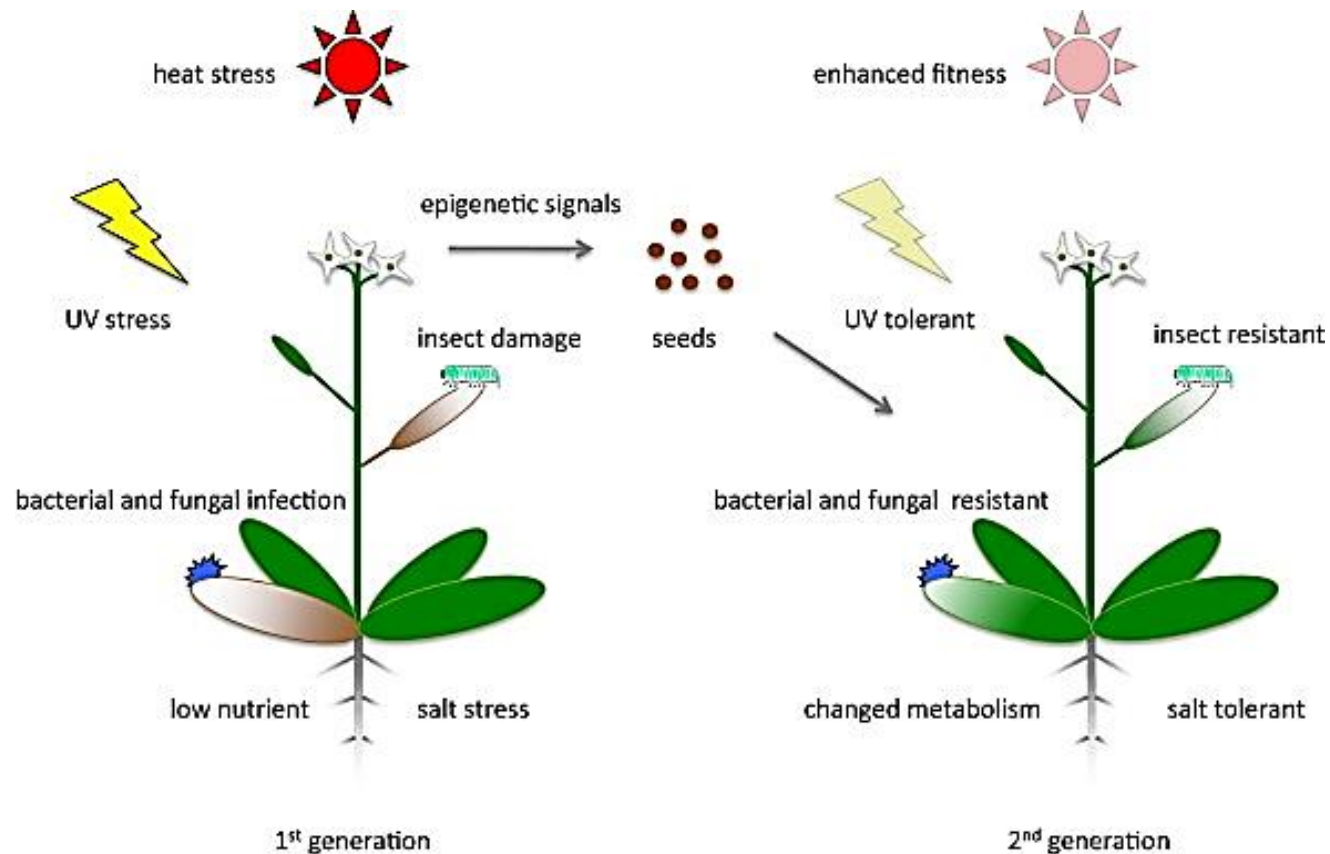
Responses to heat stress can be transient and can give the plants the ability for **short term acclimation**. The response to this **can also be in the form of a state** now popularly called **memory retention in epigenetic literature**. This epigenetic phenomenon predisposes the plant for an **effective response when it encounters the same stress again in its life cycle**, what is enigmatic is the possibility of **transgenerational transfer of this epigenetic memory** for a continued response to stress over generations

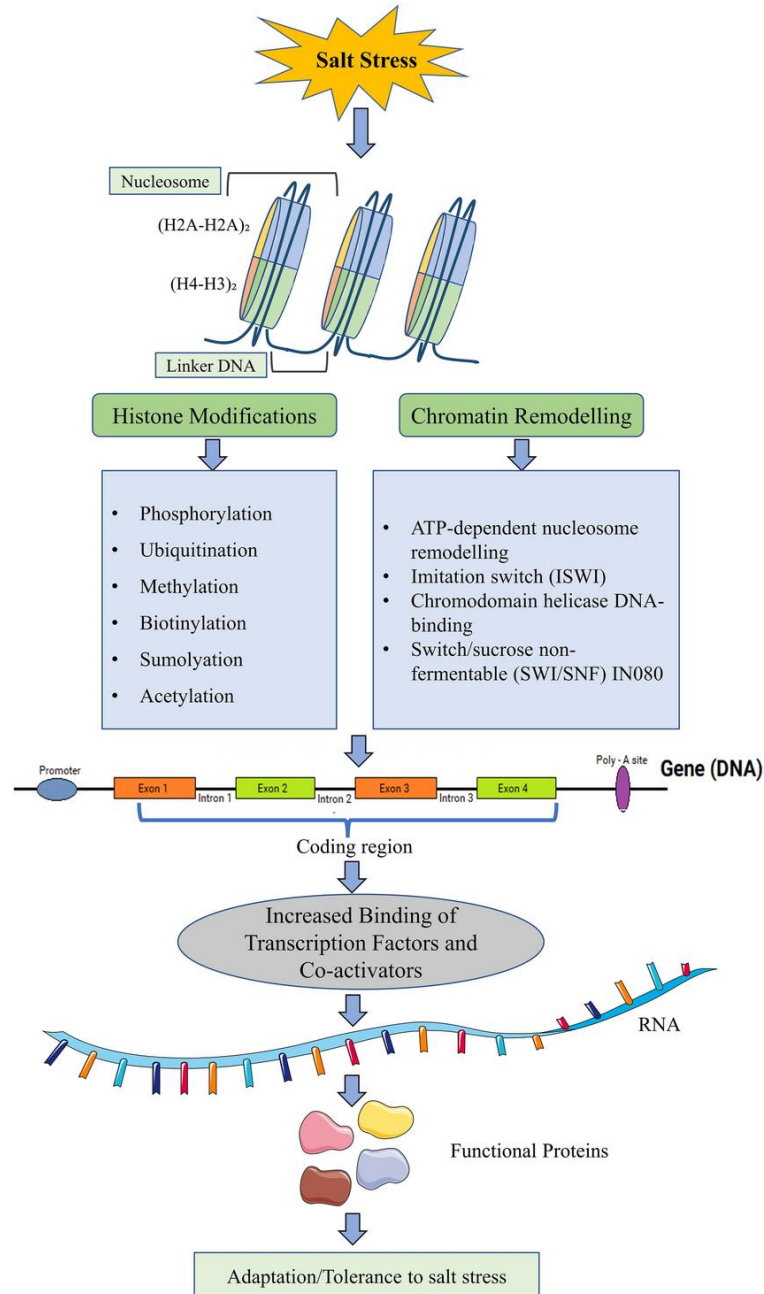
Biotic stress-induced epigenetic changes and transgenerational memory in plants



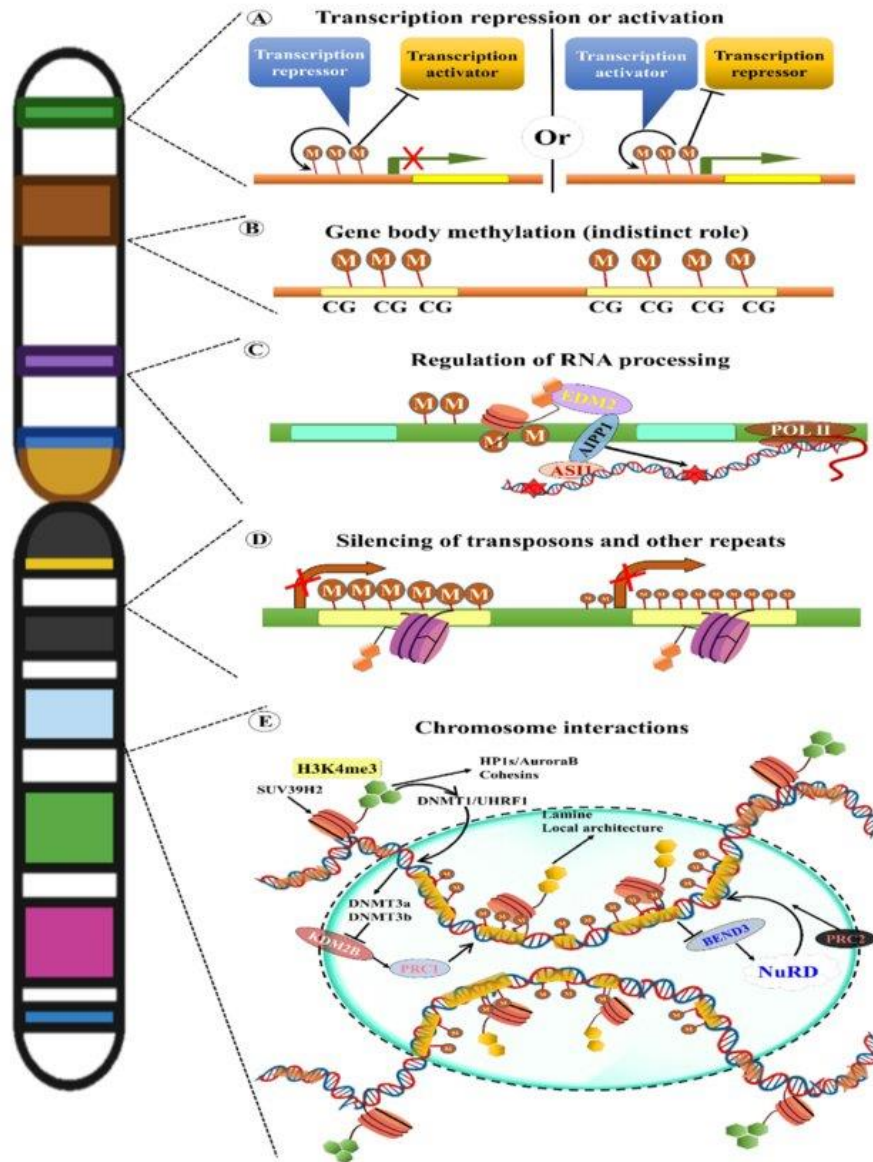
Plants are constantly exposed to hostile environments, which drive them to adopt several survival mechanisms including **epigenetic modifications**. Epigenetic modifications such as **DNA methylation, histone modifications, and small RNA mediated silencing** play significant roles in stress tolerance. Epigenetic modifications offer plants an advantage of **long-term stress adaptation** through either **prolonged gene regulation** or **transgenerational inheritance**. Among the epigenetic modifications, **DNA methylation is a major epigenetic mark that stably inherit to multiple generations**. Transgenerational epigenetic aspect of biotic stresses in plants is less explored compared to abiotic stresses.

Plant Models of Transgenerational Epigenetic Inheritance





Cellular functions of DNA methylation (m) in the plant



DNA methylation regulates transposon activation, gene regulation, and chromosome interactions. (A) Methylation in the gene promoter either represses or activates transcription. (B) Gene body methylations mainly occur in the CG context, although its function remains unclear