# CG920 Genomics

# Lesson 4

## Forward Genetics

Jan Hejátko

**Functional Genomics and Proteomics of Plants**,
CEITEC - Central European Institute of Technology
And
**National Centre for Bimolecular Research,**
Faculty of Science,

Masaryk University, Brno
hejatko@sci.muni.cz, www.ceitec.eu

# Outline

- Forward vs. Reverse Genetics

- Use of Libraries of Insertional Mutants in Forward Genetics
  - Searching in Libraries of Insertional Mutants According to:
    - anatomically or morphologically detectable phenotype
    - metabolic profile
    - expression of genes of interest

  - Identification of the Mutated Locus
    - plasmid rescue
    - iPCR

- Use of Libraries of Point Mutants in Forward Genetics
  - Positional Cloning

# Outline

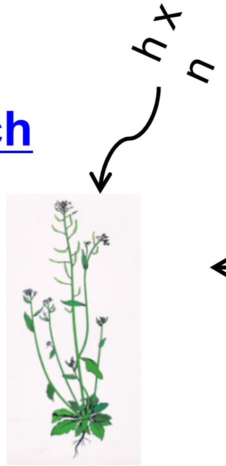
- Forward vs. Reverse Genetics

# „Classical“ genetics *versus* „reverse genetics“ approaches in functional genomics

## RANDOM MUTAGENESIS

### „Classical genetics“ approach

EMS



1. IDENTIFICATION OF PHENOTYPE

2. GENE MAPPING

3. GENE IDENTIFICATION
   - position cloning

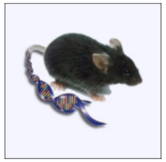
### „Reverse genetics“ approach

T-DNA

1. ISOLATION OF SEQUENCE-SPECIFIC MUTANT

2. IDENTIFICATION OF PHENOTYPE

3. PROOF OF CAUSAL RELATIONSHIP BETWEEN INSERTION AND PHENOTYPE



(retro)transposons

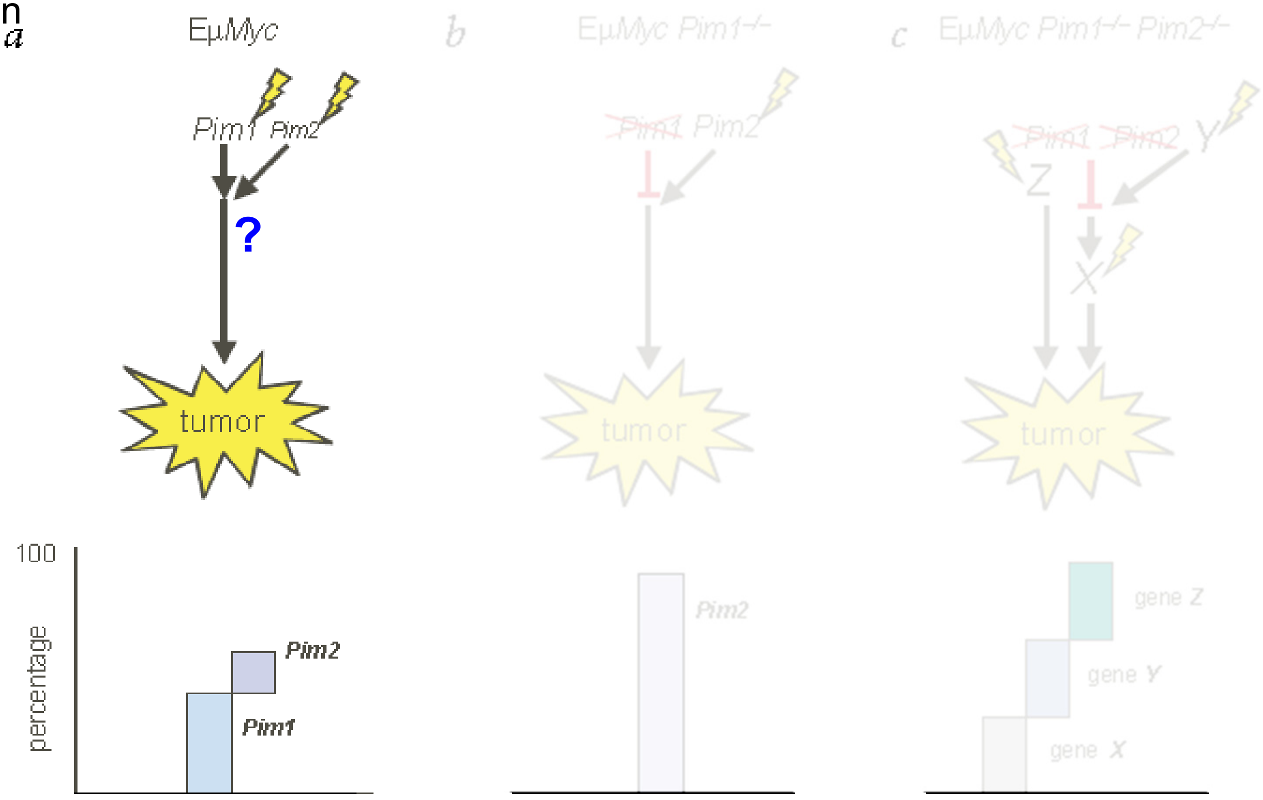
# Outline

- Forward vs. Reverse Genetics

- Use of Libraries of Insertional Mutants in Forward Genetics
  - Searching in Libraries of Insertional Mutants According to:
    - anatomically or morphologically detectable phenotype

# Insertional mutagenesis in forward genetics approaches

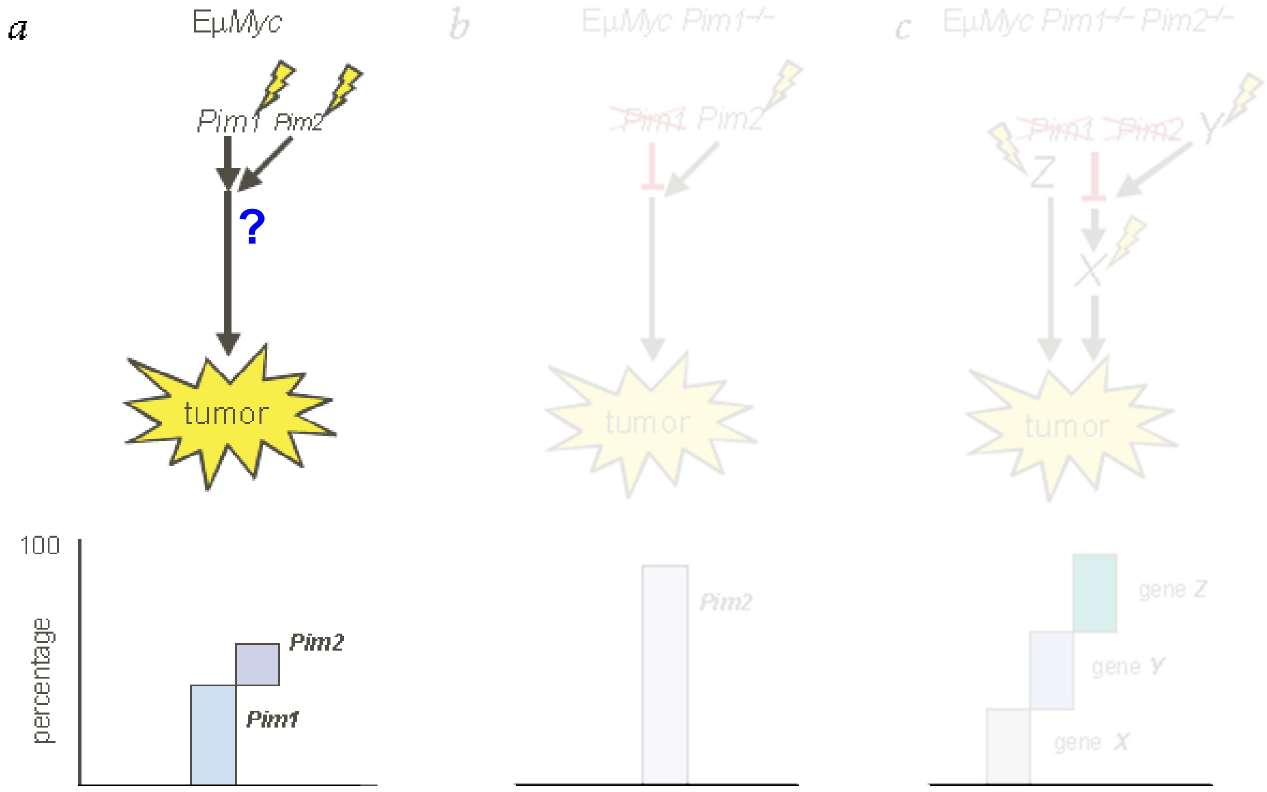
- Use of insertional mutagenesis for study of carcinogenesis
  - Infection of EμMyc mice by MoMuLV retrovirus leads to lymphomas formation, which arose due to activation of Pim kinases (40 % activation of Pim1, 15 % activation of Pim2), molecular targets of these kinases were unknown



Mikkers et al., Nature Gen (2002)

# Insertional mutagenesis in forward genetics approaches

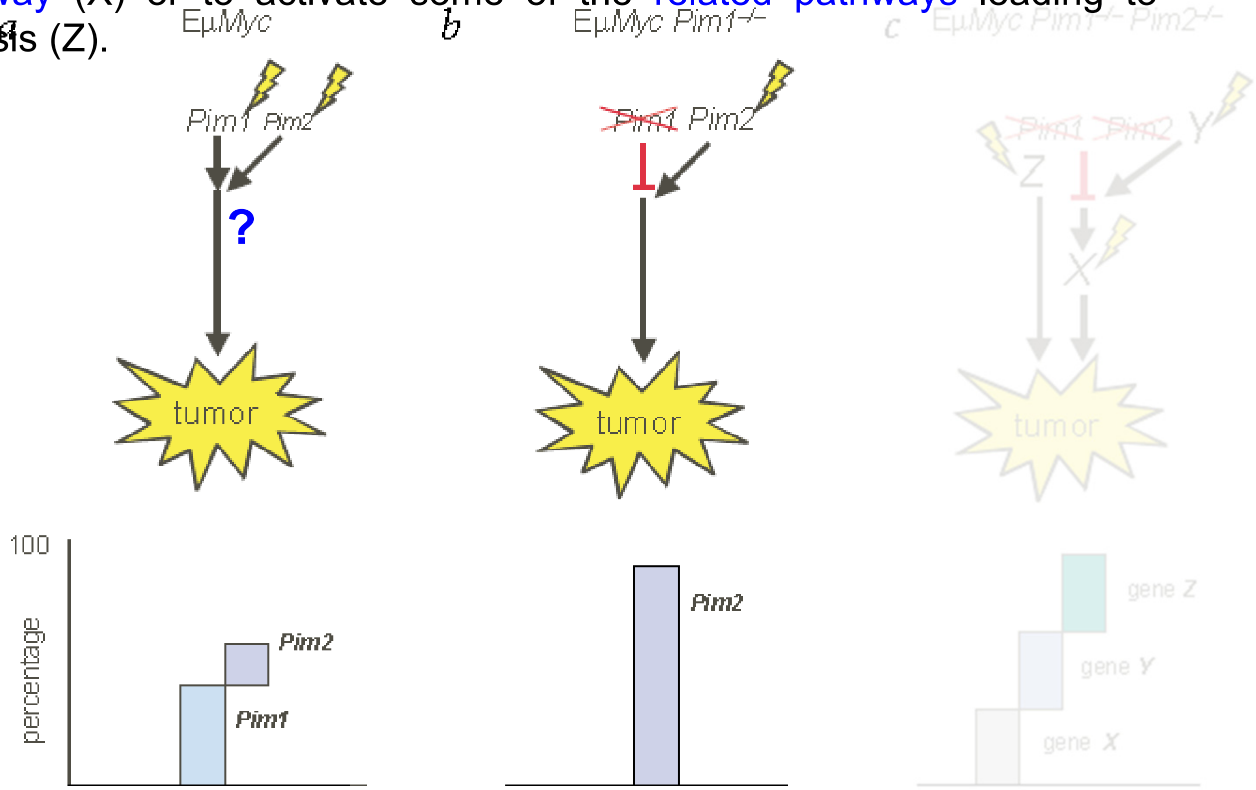
- Use of insertional mutagenesis for study of carcinogenesis
  - Infection of EμMyc *pim1* mutants by MoMuLV retrovirus leads to lymphomas formation, which in 90 % contain insertion nearby (activation) Pim2



Mikkers et al., Nature Gen (2002)

# Insertional mutagenesis in forward genetics approaches

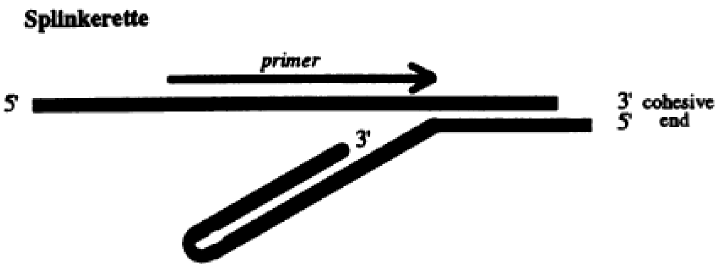
- Use of insertional mutagenesis for study of carcinogenesis
  - Infection of EμMyc double mutants *pim1, pim2* by MoMuLV retrovirus leads to lymphomas formation, which can be expected to activate either one of the signalling partner of Pim proteins (Y), one of the downsteram proteins of Pim signalling pathway (X) or to activate some of the related pathways leading to lymphomagenesis (Z).



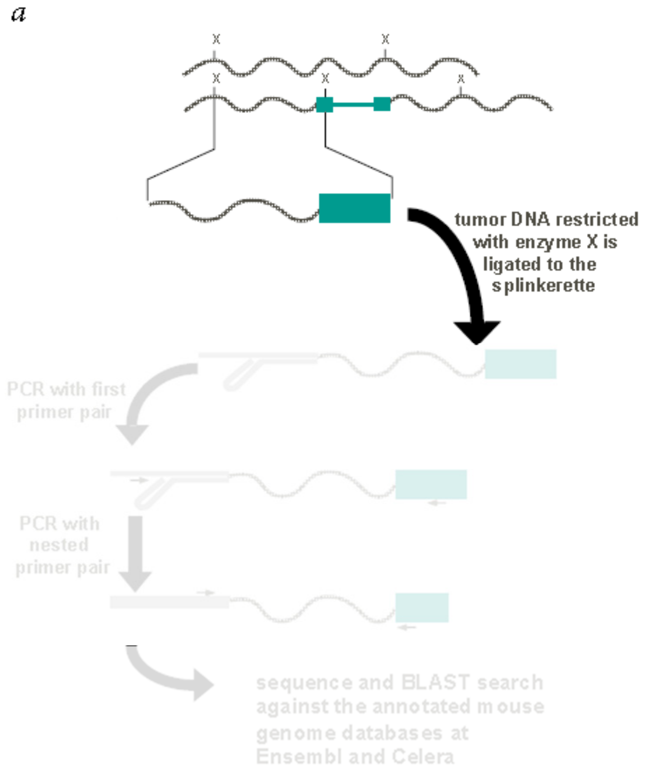
Mikkers et al., Nature Gen (2002)

# Insertional mutagenesis in forward genetics approaches

- Isolation of genomic regions adjacent to the insertion site of the provirus
  - Cleavage of genomic DNA and ligation of special linkers, so-called *splincerettes* (increasing the specifity of amplification)



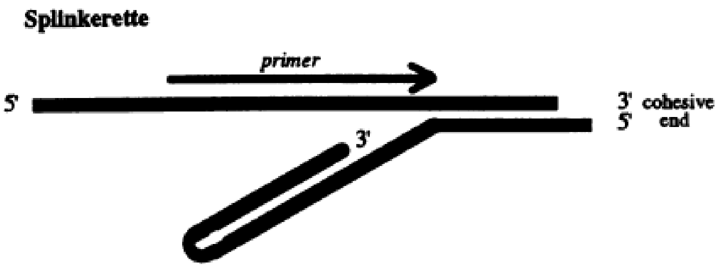
Devon et al., Nucl Acid Res (1994)



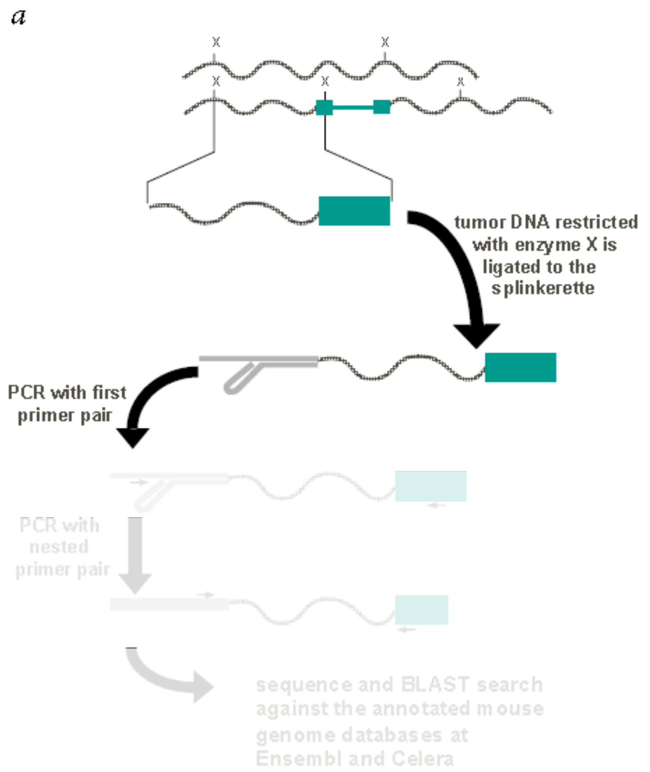
Mikkers et al., Nature Gen (2002)

# Insertional mutagenesis in forward genetics approaches

- Isolation of genomic regions adjacent to the insertion site of the provirus
  - First amplification using specific primers



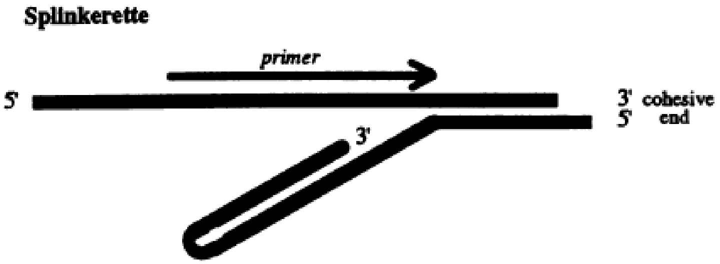
Devon et al., Nucl Acid Res (1994)



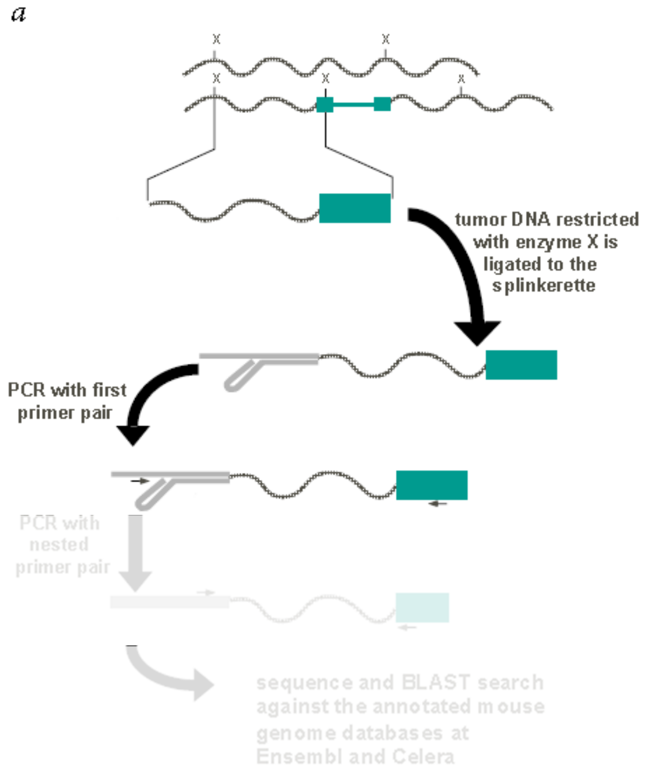
Mikkers et al., Nature Gen (2002)

# Insertional mutagenesis in forward genetics approaches

- Isolation of genomic regions adjacent to the insertion site of the provirus
  - Second amplification using nested primers (increasing the specifity)



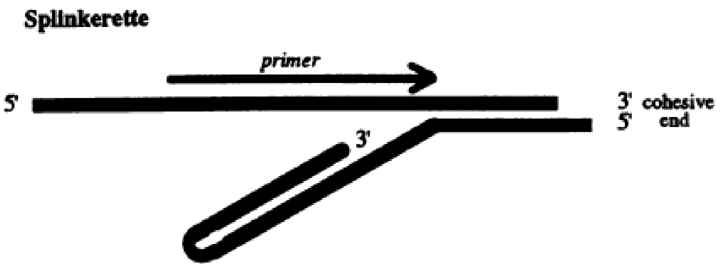
Devon et al., Nucl Acid Res (1994)



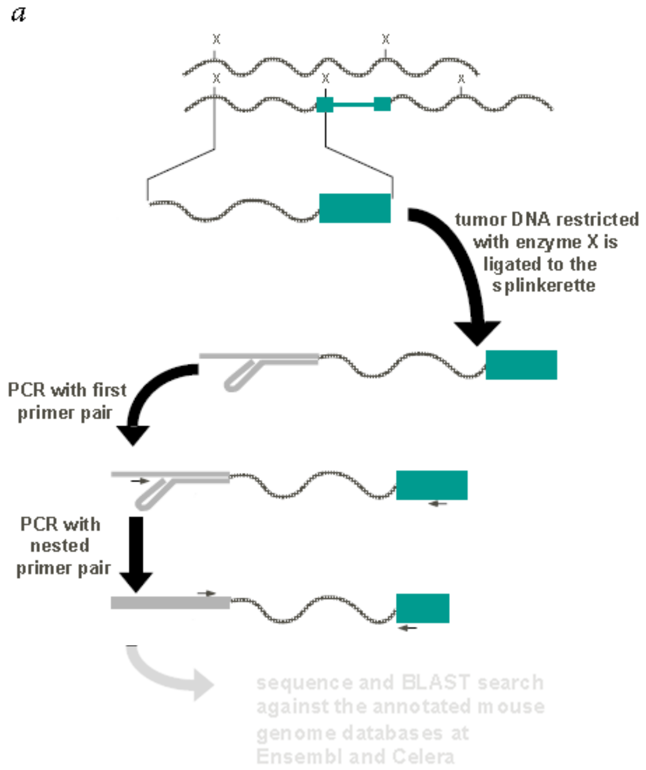
Mikkers et al., Nature Gen (2002)

# Insertional mutagenesis in forward genetics approaches

- Isolation of genomic regions adjacent to the insertion site of the provirus
  - Sequencing and localization of regions adjacent to provirus by searching in annotated databases of mouse genome



Devon et al., Nucl Acid Res (1994)



Mikkers et al., Nature Gen (2002)

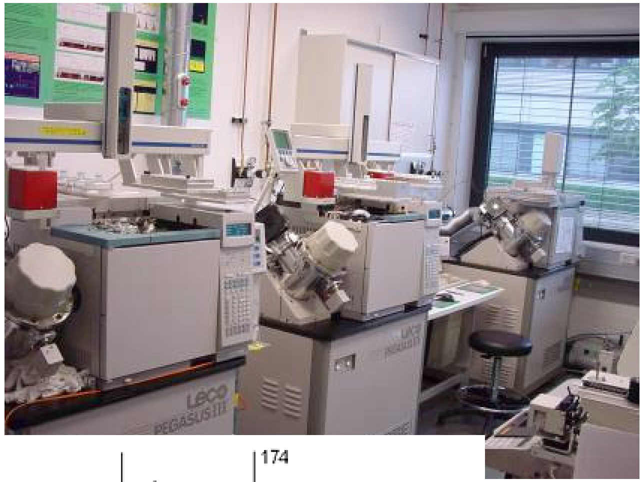
# Outline

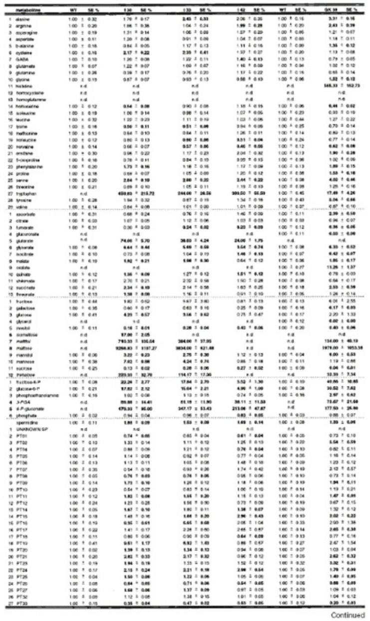
- Forward vs. Reverse Genetics

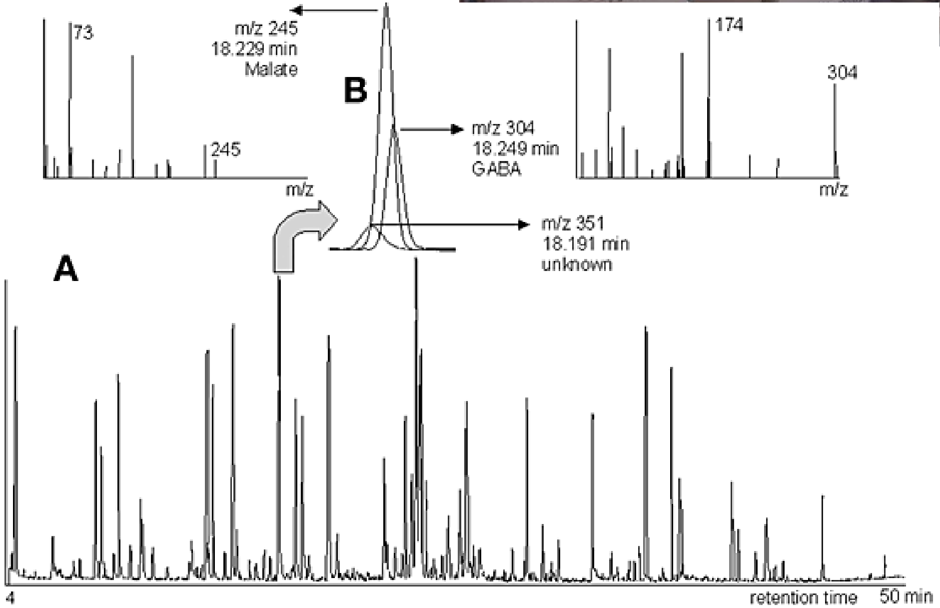
- Use of Libraries of Insertional Mutants in Forward Genetics
  - Searching in Libraries of Insertional Mutants According to:
    - anatomically or morphologically detectable phenotype
    - metabolic profile

# Metabolic profiling

- Metabolic profiling of plants
  - Automated analysis of metabolites (up to 25.000) by GC-MS techniques in libraries of T-DNA mutants

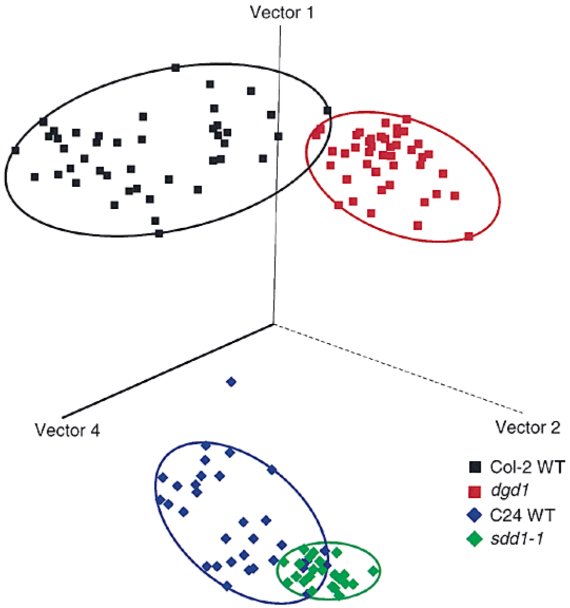






# Metabolic profiling

- Metabolic profiling of plants
  - Automated analysis of metabolites (up to 25.000) by GC-MS techniques in libraries of T-DNA mutants
  - Identification of interesting (even comercially interesting) mutants

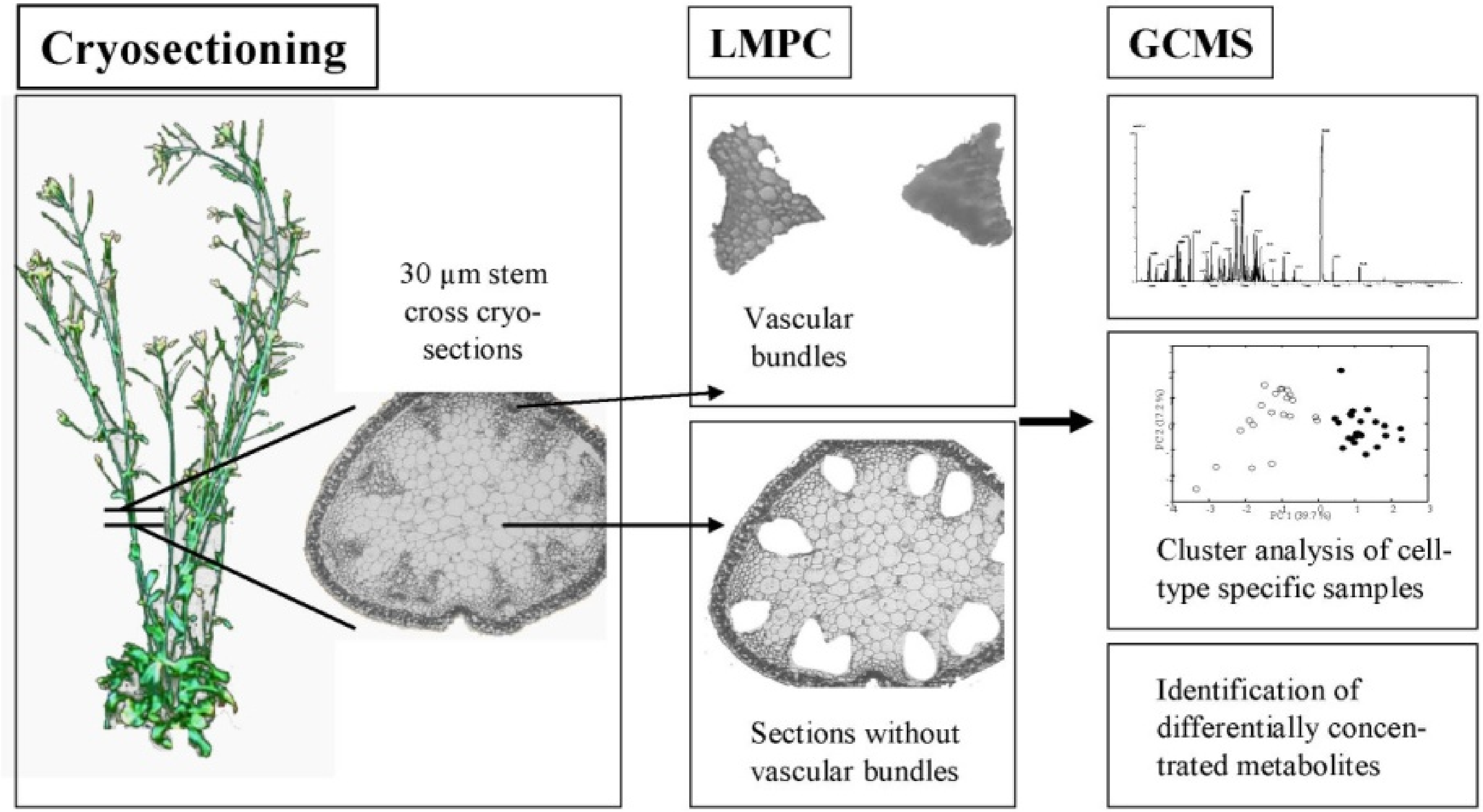


# Metabolic profiling

- Metabolic profiling of plants
  - Automated analysis of metabolites (up to 25.000) by GC-MS techniques in libraries of T-DNA mutants
  - Identification of interesting (even comercially interesting) mutants
  - Fast and easy isolation of genes through identification of sequences adjacent to T-DNA

# Metabolic profiling

- Metabolic profiling of plants
  - Possibility to use special techniques, e.g. microdissection



**Cryosectioning**

30 µm stem cross cryo-sections

**LMPC**

Vascular bundles

Sections without vascular bundles

**GCMS**

Cluster analysis of cell-type specific samples

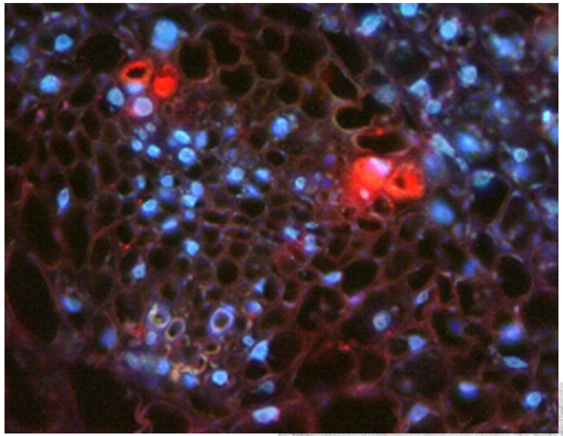
Identification of differentially concentrated metabolites

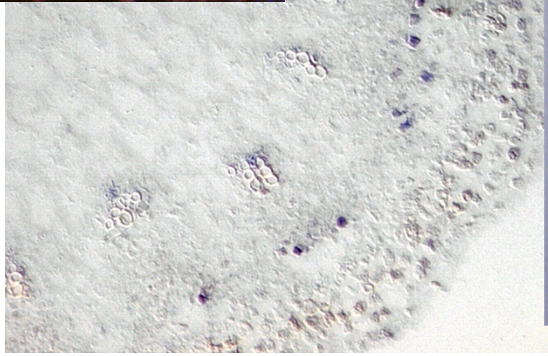
# Outline

- Forward vs. Reverse Genetics
- Use of Libraries of Insertional Mutants in Forward Genetics
  - Searching in Libraries of Insertional Mutants According to:
    - anatomically or morphologically detectable phenotype
    - metabolic profile
    - expression of genes of interest

# Expression profile

- Identification of mutants with a change in the expression profile
  - Analysis of *expression profile* (*pattern*) of the gene and identification of mutants with altered expression pattern

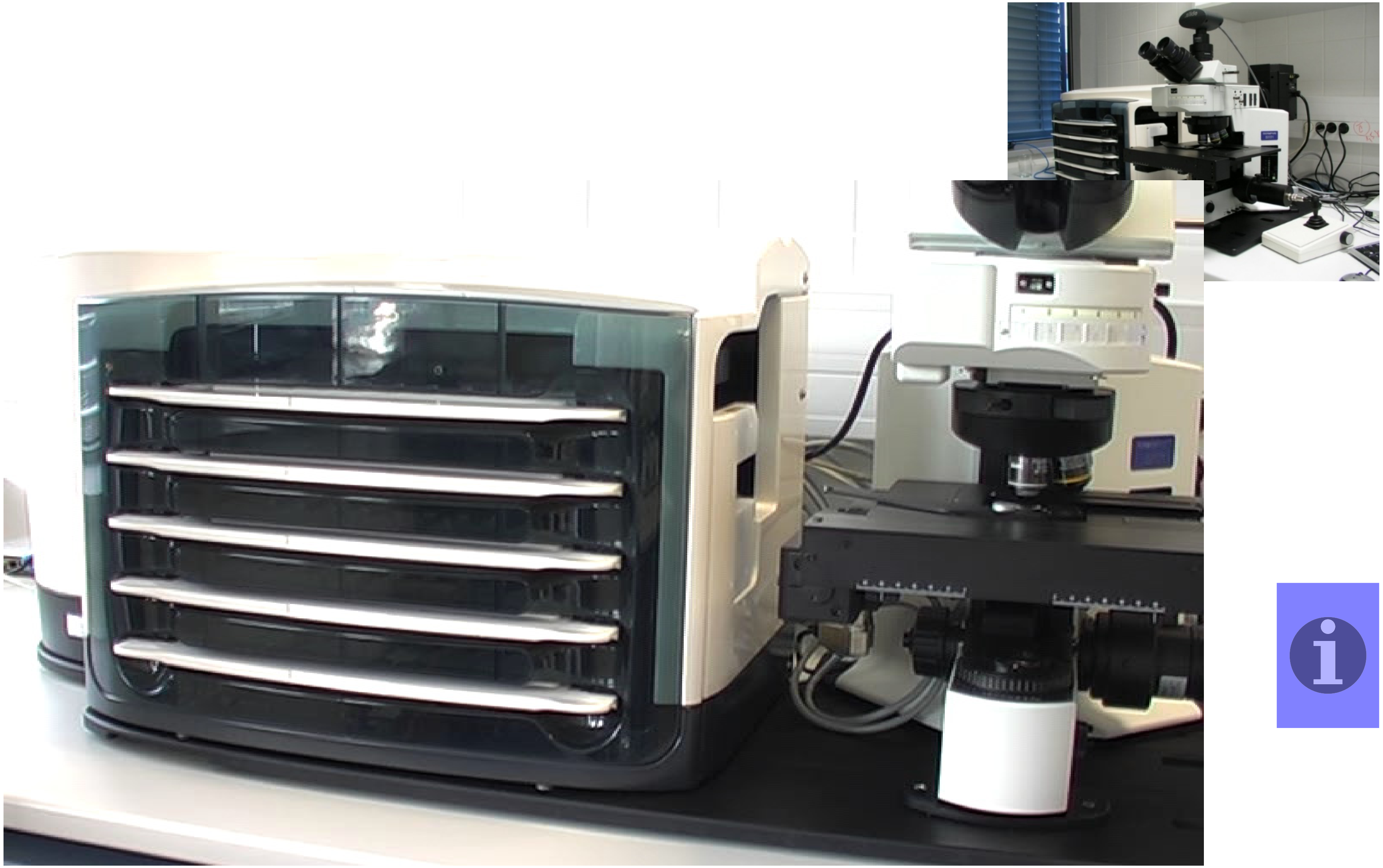




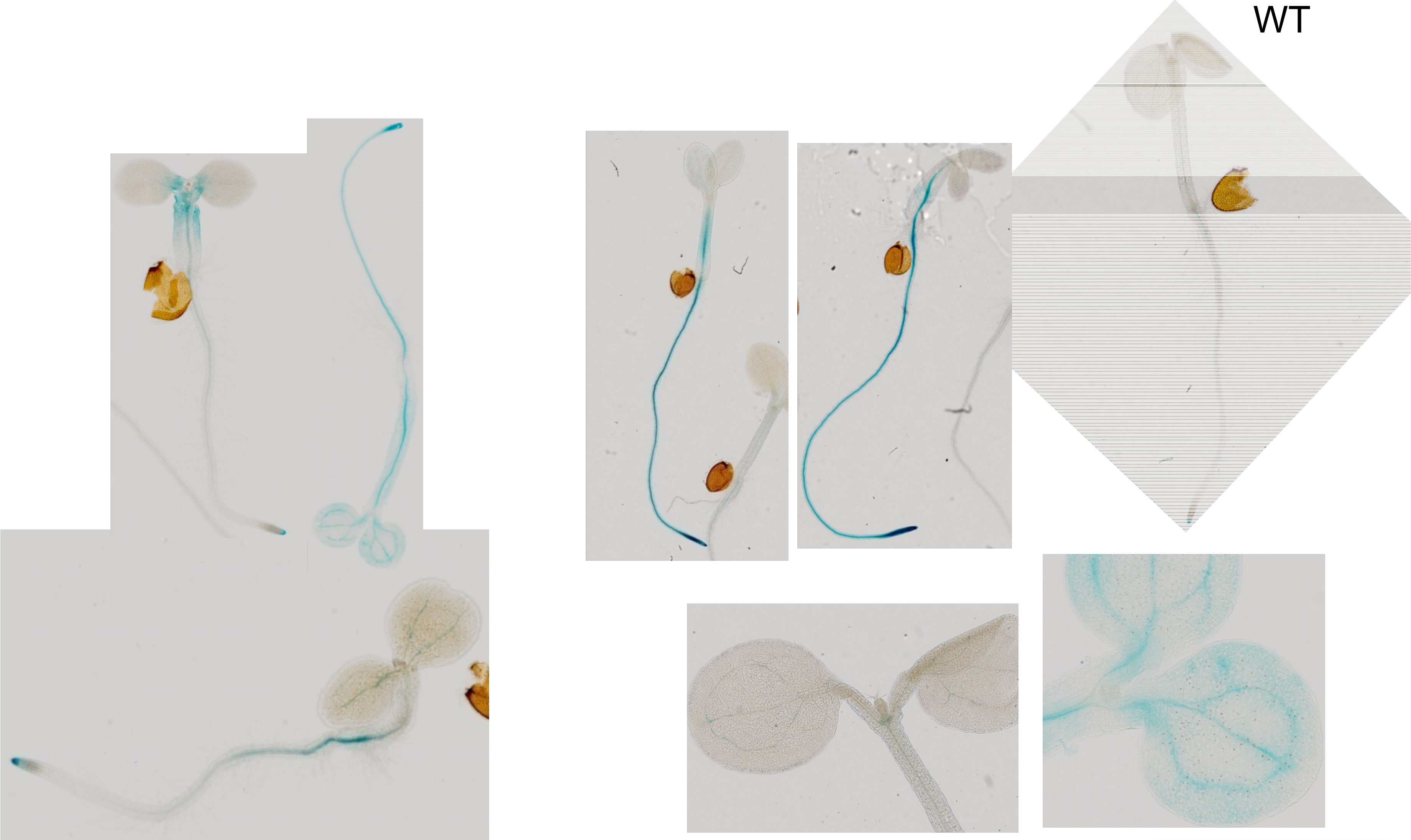
# Expression profile

- Identification of mutants with a change in the expression profile
  - Analysis of *expression profile* (*pattern*) of the gene and identification of mutants with altered expression pattern
  - Possibility of partial automation (virtual digital microscopy)

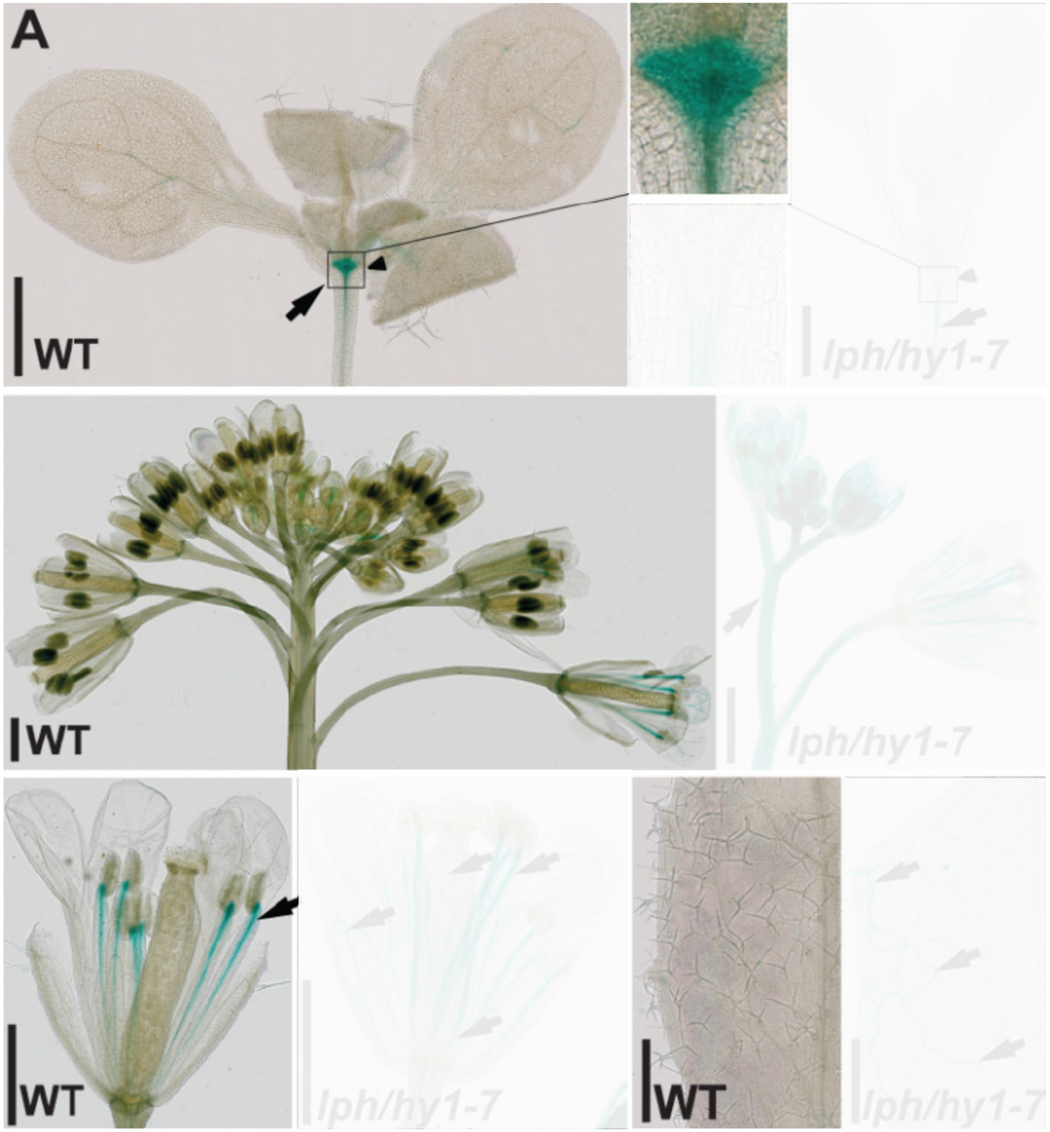
# Automated Microscopy Screening



# Expression profile



# *LPH* controls expression of *CKI1*

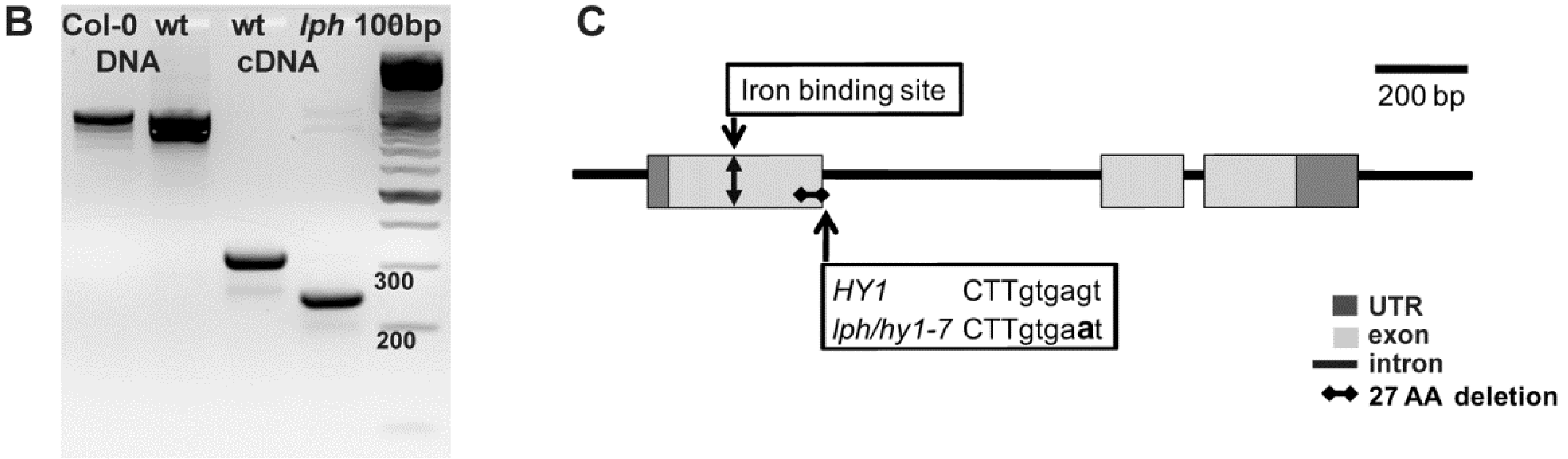


Dobisova et al., *Plant Physiol, 2017*

# *lph* is novel allele of HEME OXYGENASE 1

**A**

CCTACTGTTGAAGGTTACTTGAGGTTTCTTGTGGATAGTAAATTGGTTTATGATACTCTTGAACTGATTA
TTCAAGACTCCAATTTCCCAACTTgtgagtttttttttt



**D**

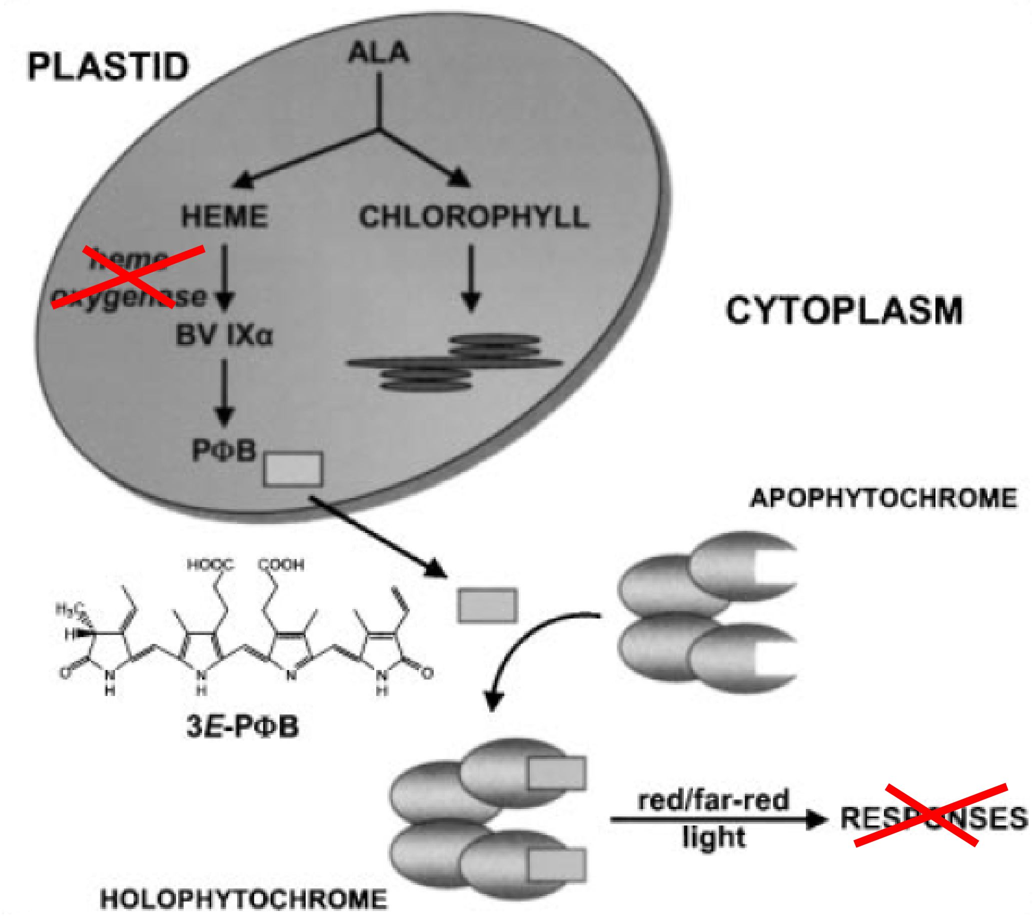
```
HY1         MAYLAPISSSLSIFKNPQLSRFQFSSSSPNPLFLRPRIQILSMTMNKSPSLVVVAATTAEKQKKRYPGESKGFVEEMRFVAMRLHTKDQAKEGEKETKS
lph/hy1-7   MAYLAPISSSLSIFKNPQLSRFQFSSSSPNPLFLRPRIQILSMTMNKSPSLVVVAATTAEKQKKRYPGESKGFVEEMRFVAMRLHTKDQAKEGEKETKS

HY1         IEERPVAKWEPTVEGYLRFLVDSKLVYDTLELIIQDSNFPTYAEFKNTGLERAEKLSTDLEWFKEQGYEIPEPTAPGKTYSQYLKELAEKDPQAFICHFY
lph/hy1-7   IEERPVAKWEPTVED--------------------------AEFKNTGLERAEKLSTDLEWFKEQGYEIPEPTAPGKTYSQYLKELAEKDPQAFICHFY

HY1         NIYFAHSAGGRMIGRKVAERILDNKELEFYKWDGELSQLLQNVREKLNKVAEEWTREEKNHCLEETEKSFKYSGEILRLILS
lph/hy1-7   NIYFAHSAGGRMIGRKVAERILDNKELEFYKWDGELSQLLQNVREKLNKVAEEWTREEKNHCLEETEKSFKYSGEILRLILS
```

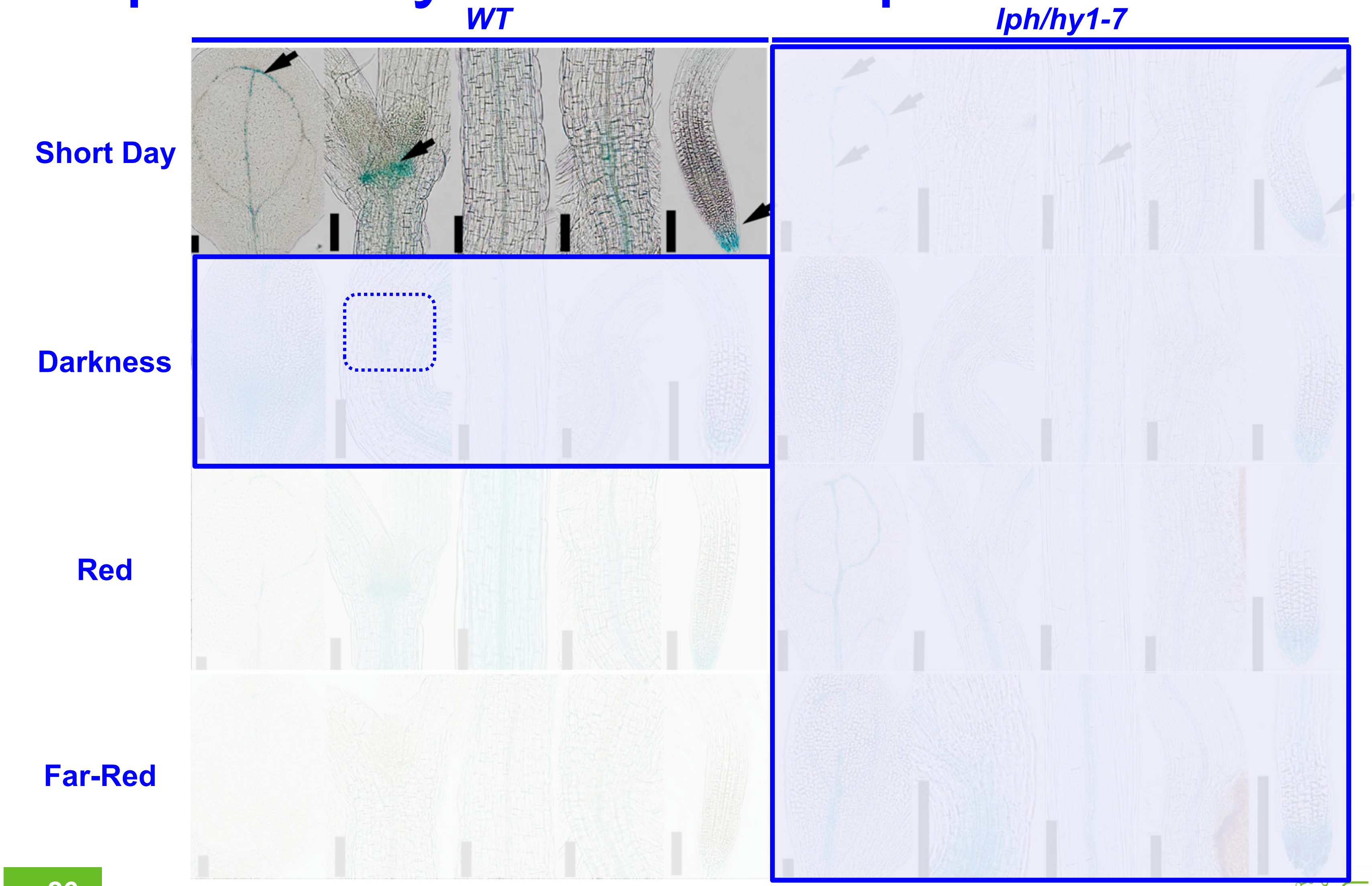
Dobisova et al., *Plant Physiol, 2017*

# *lph* is novel allele of HEME OXYGENASE 1



Terry et al., *Biochem Soc Trans*, 2002

# Light controls spatiotemporal specificity of *CKI1* expression



Dobisova et al., *Plant Phys, 2017*

# Outline

- Forward vs. Reverse Genetics

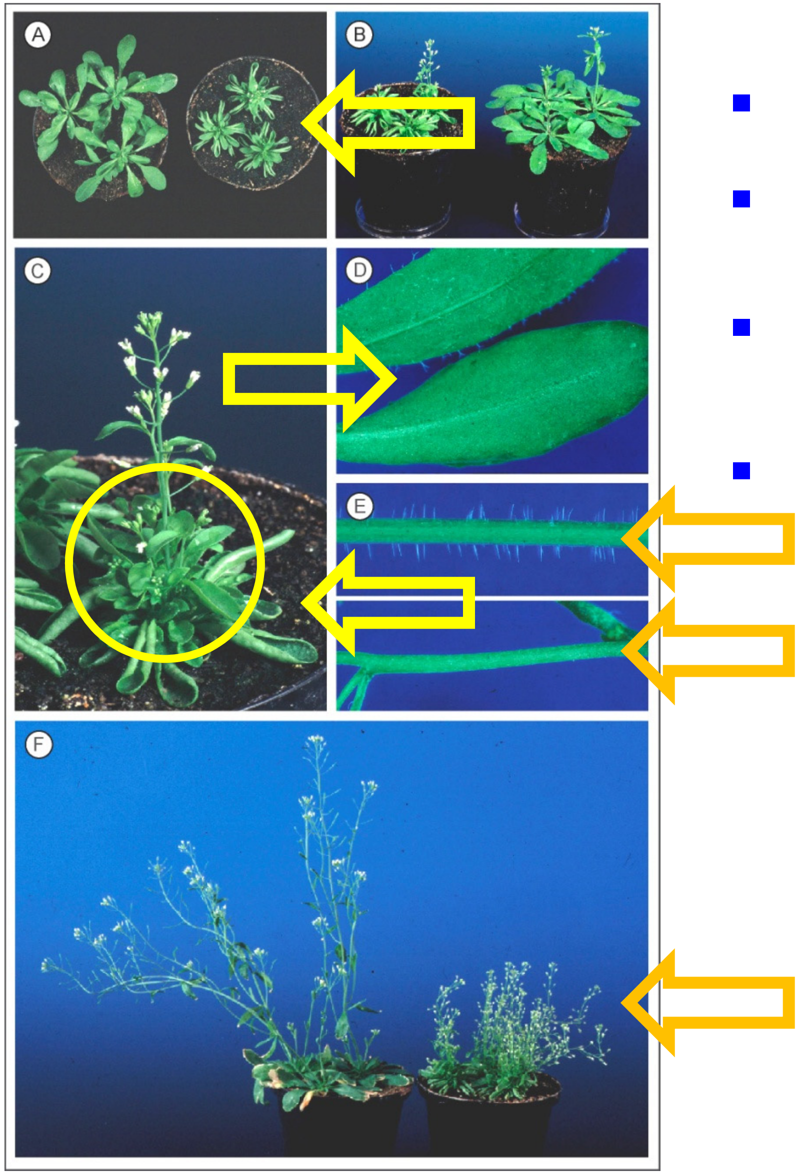
- Use of Libraries of Insertional Mutants in Forward Genetics
  - Searching in Libraries of Insertional Mutants According to:
    - anatomically or morphologically detectable phenotype
    - metabolic profile
    - expression of genes of interest

  - Identification of the Mutated Locus
    - plasmid rescue
    - iPCR

# Identification of mutated locus

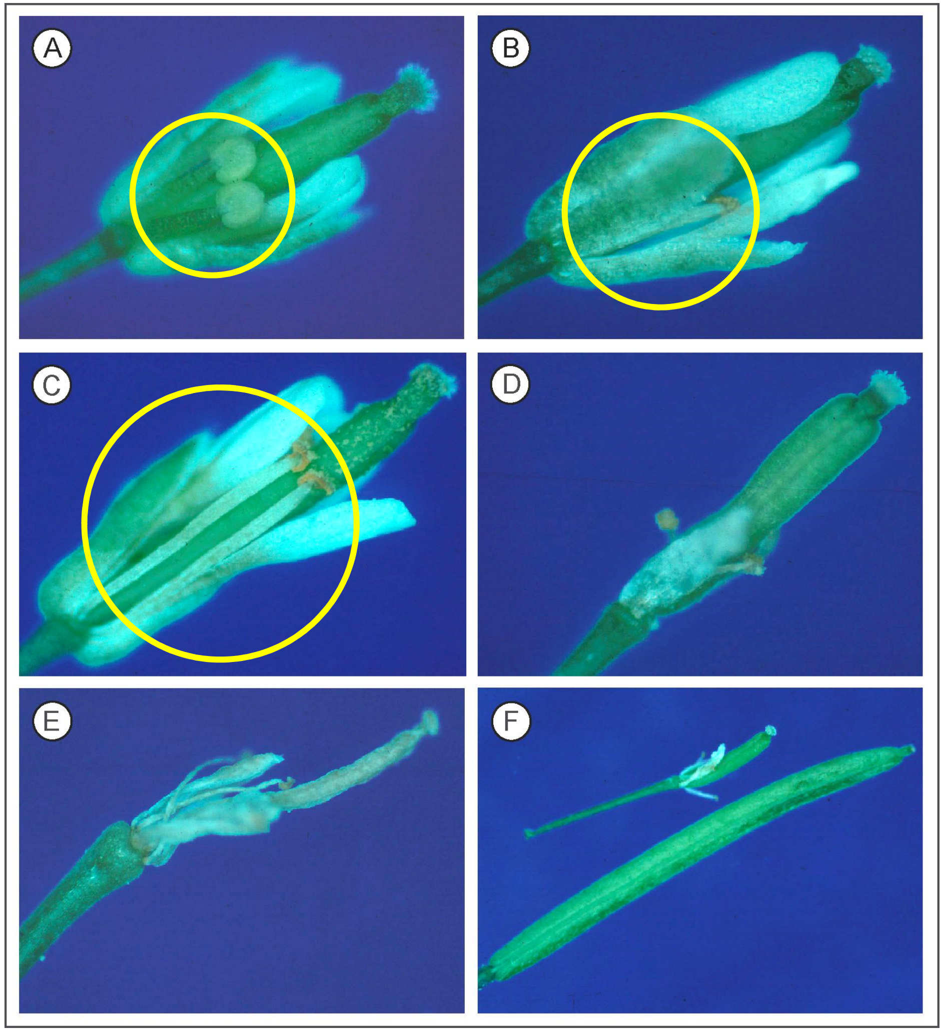
- Identification of chromosomal rearrangements responsible for bushy phenotype of *Arabidopsis*
    - Description of phenotype

# Identification of mutant



- Crinkled leaves
- Bushy phenotype (branching defective)
- No trichomes on leaves and stems
- Late senescence

# Identification of mutant



- Male sterility, defects in stamen filament elongation (A,B)

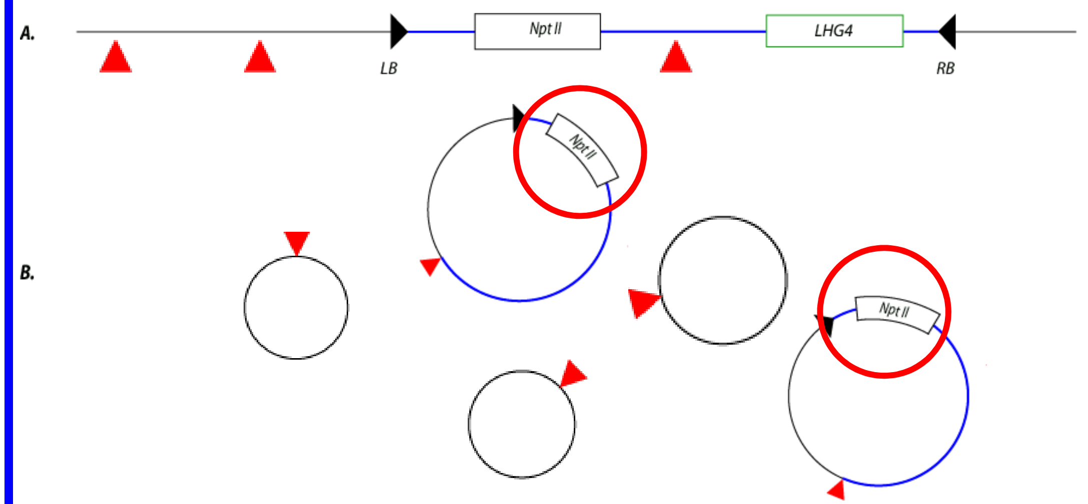
(compare with wild type C)

# Identification of mutated locus

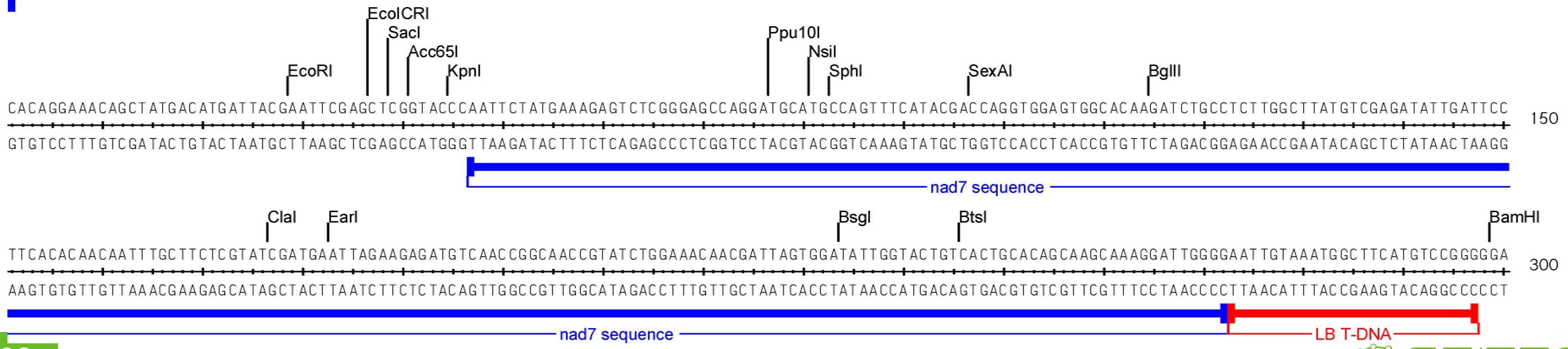
- Identification of chromosomal rearrangements responsible for bushy phenotype of *Arabidopsis*
    - Description of phenotype
    - Identification of T-DNA mutated region

# Identification of mutated locus

## 1. Identification of region of genomic DNA adjacent to the ***left border*** using ***plasmid rescue***

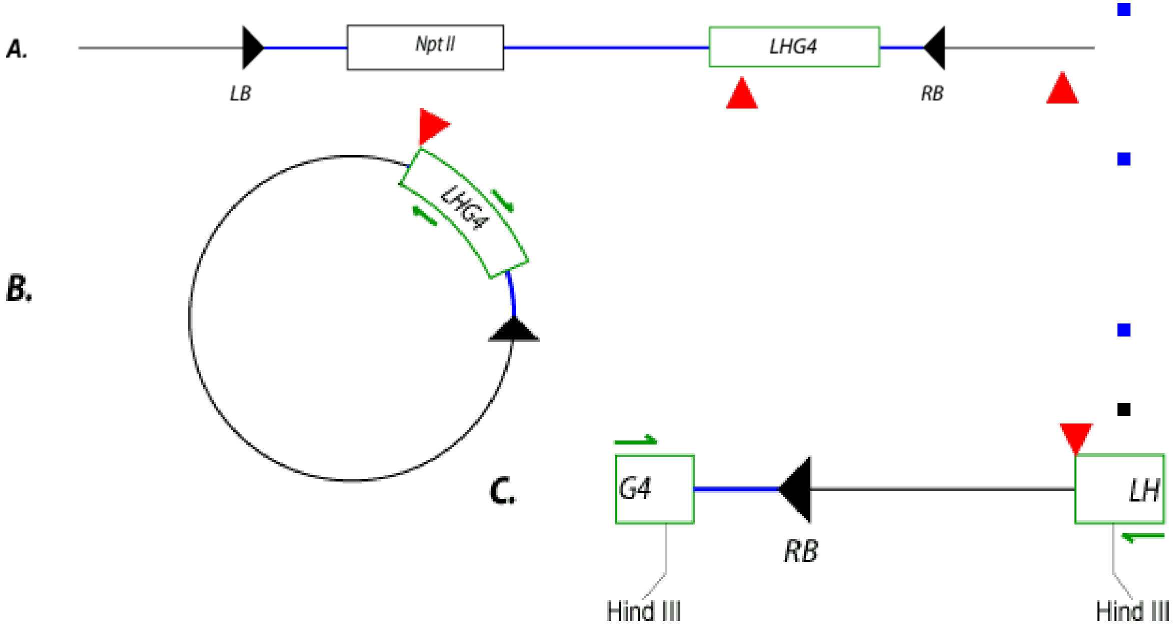


- Restriction digestion (*EcoRI*) of mutant genomic DNA
- Religation and transformation of *E. coli*
- Isolation of plasmid DNA from positively selected clones
- Identified sequence was identical to gene for *NAD7* coded by mtDNA

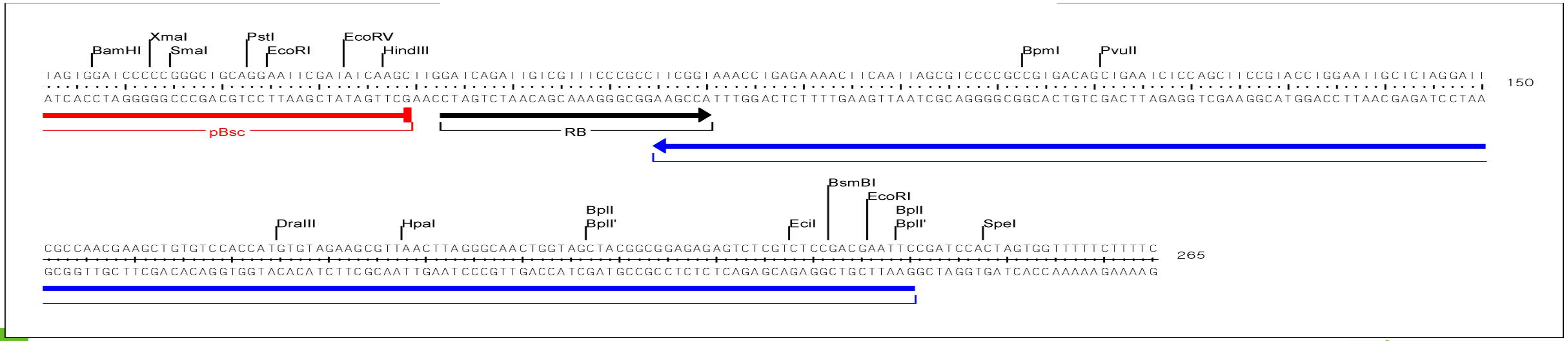


# Identification of mutated locus

2. Identification of region of genomic DNA adjacent to the ***right border*** using ***inversion PCR*** **(iPCR)**



- Restriction digestion (*EcoRI)* of mutant genomic DNA
- Purification, religation and PCR using T-DNA specific primers
- Cloning and sequencing
- Identified sequence was not homologous to any sequences with known function

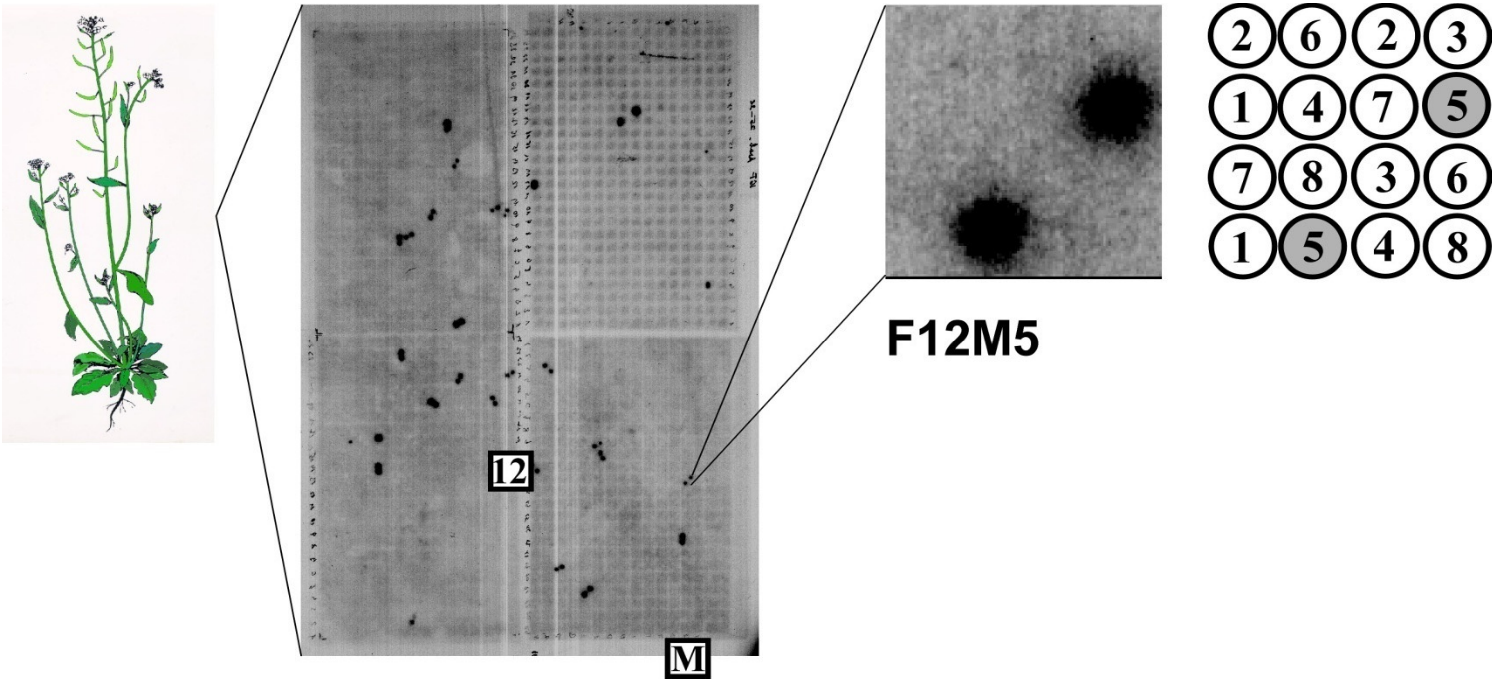


# Identification of mutated locus

- Identification of chromosomal rearrangements responsible for bushy phenotype of *Arabidopsis*
  - Description of phenotype
  - Identification of T-DNA mutated region
  - Localization of T-DNA insertion site in *Arabidopsis* genome

# Searching in library IGF-BAC

- Genome library containing 10.752 clones with an average size of an insert of 100 kb
- Bacterial clones arranged in the microtiter plates
- Library loaded onto nylon filters for hybridization with the radiolabeled probe



# Mapping with IGF-BAC database

## *I.* Sequences adjacent to the left border of T-DNA

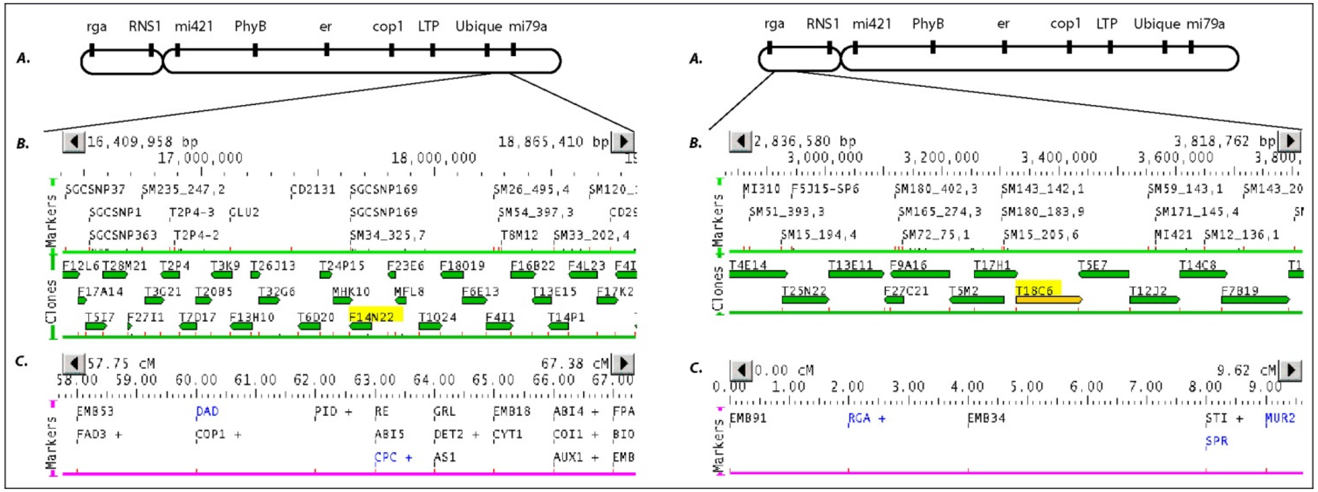
- 28 positively hybridizing clones in total
- 19 of them located on chromosome 2
- 18 of them similar with mtDNA

## *II.* Sequences adjacent to the right border of T-DNA

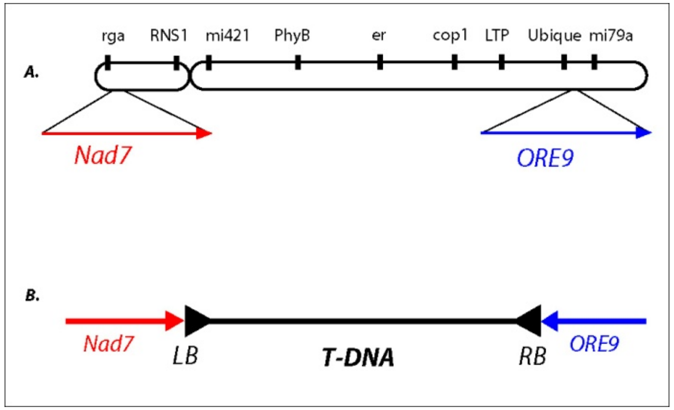
- 6 positively hybridizing clones in total
- all of them located on chromosome 2

# Localization of genomic T-DNA adjacent to both left and right T-DNA borders on chromosome 2

Sequences adjacent to ***right*** and ***left*** border of T-DNA



- There was probably an inversion of almost entire chromosome 2



# Outline

- Forward vs. Reverse Genetics

- Use of Libraries of Insertional Mutants in Forward Genetics
  - Searching in Libraries of Insertional Mutants According to:
    - anatomically or morphologically detectable phenotype
    - metabolic profile
    - expression of genes of interest

  - Identification of the Mutated Locus
    - plasmid rescue
    - iPCR

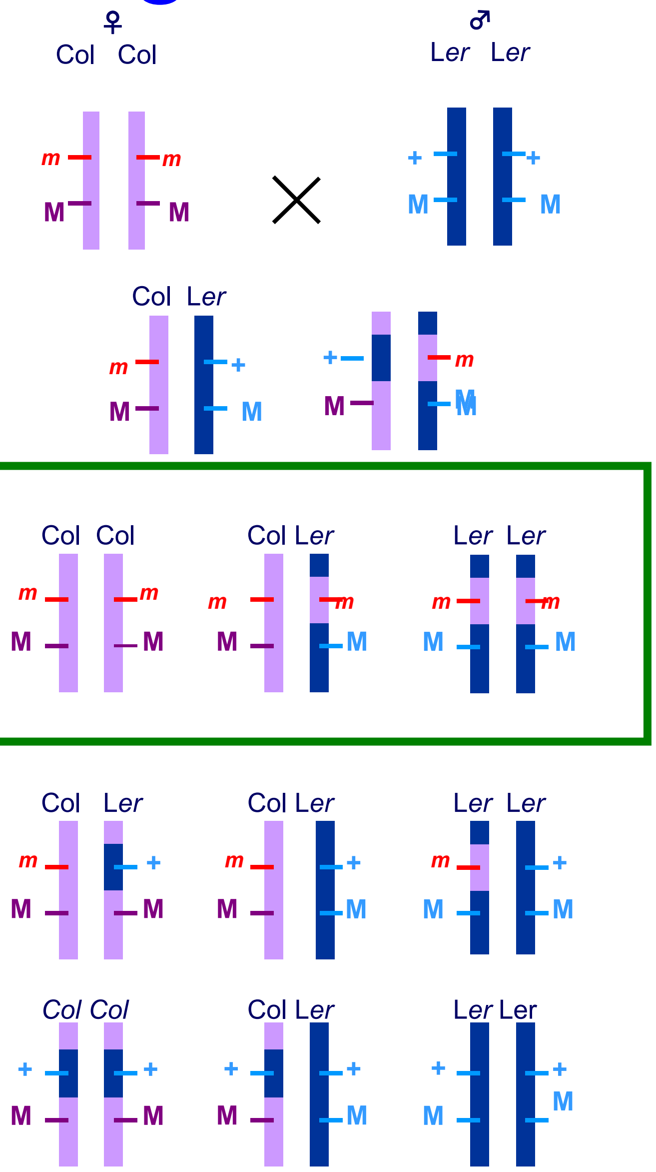
- Use of Libraries of Point Mutants in Forward Genetics
  - Positional Cloning

# Identification of mutated locus

- Positional cloning
  - Principle: co-segregation analysis of segregating population (mostly of offspring of backcrosses) with molecular markers
  - **SSLP** (Simple Sequence Length Polymorphism)
    - Polymorphism of genome (PCR products) length, amplified using specific primers
  - **RFLP** (Restriction Fragment Length Polymorphism)
    - Detection by Southern blot (PCR after digestion of the genomic DNA and ligation of adapters)
  - **CAPS** (Cleaved Amplified Polymorphic Sequence)
    - Restriction fragment length polymorphism, genome segments amplified by PCR
  - **RAPD** (Randomly Amplified Polymorphic DNA)
    - Polymorphism of length of randomly amplified genome segments, using short 8-10bp primers

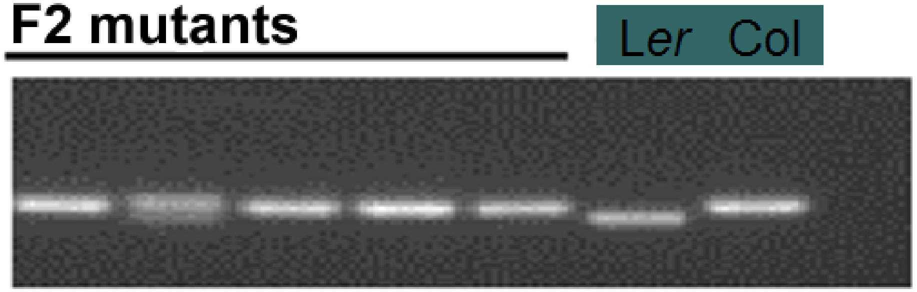
# Positional cloning

**Preparation of mapping population**

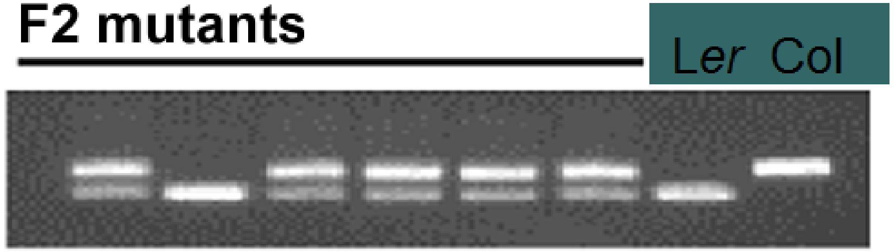


# **Recombinant analysis** – determining the percentage of recombination between mutation and molecular marker

**r [%] = number of chromosomes of *Col* / number of all the chromosomes × 100**



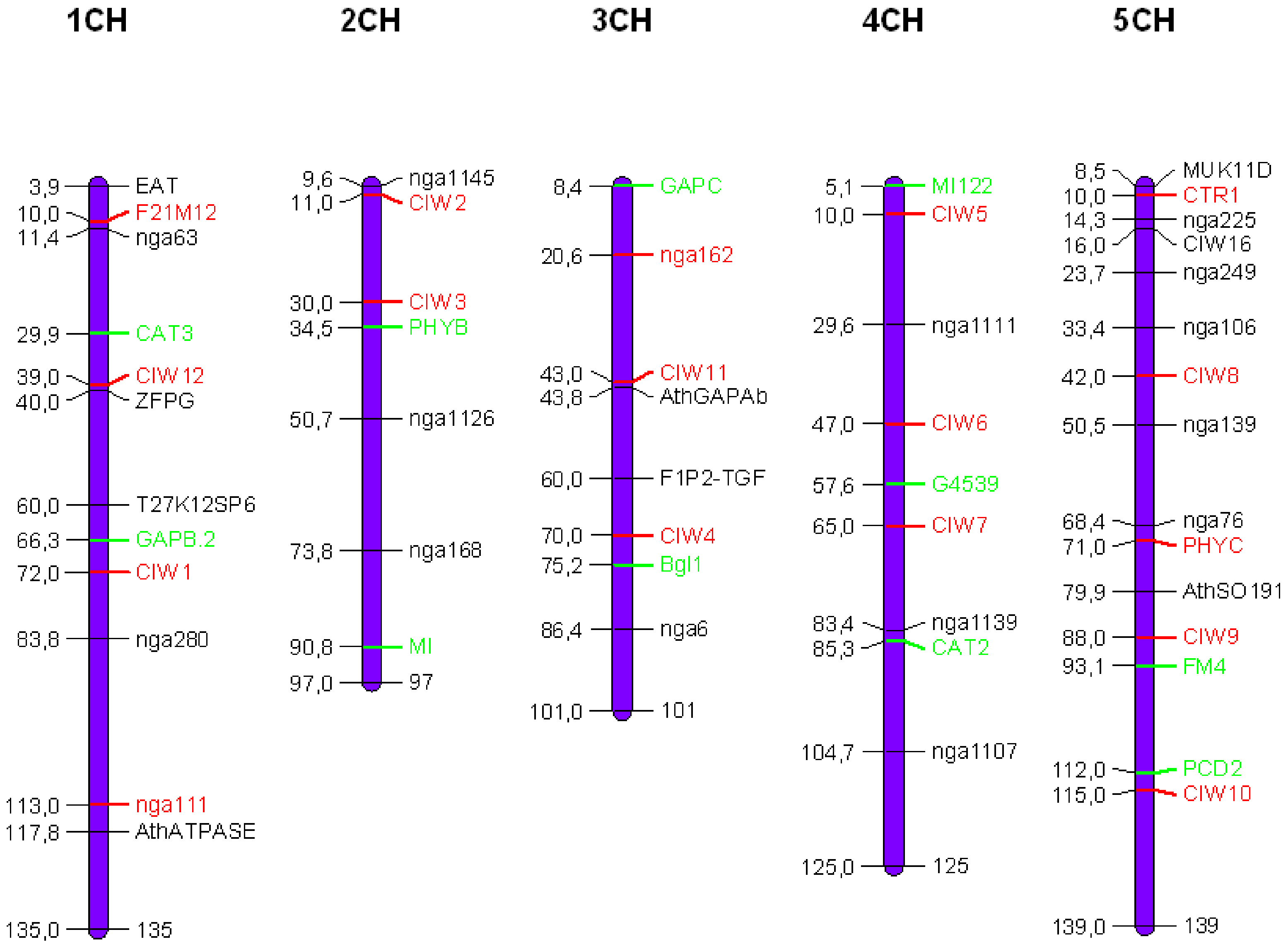
marker I – linked
5 mutants
1/10 × 100 = 10%



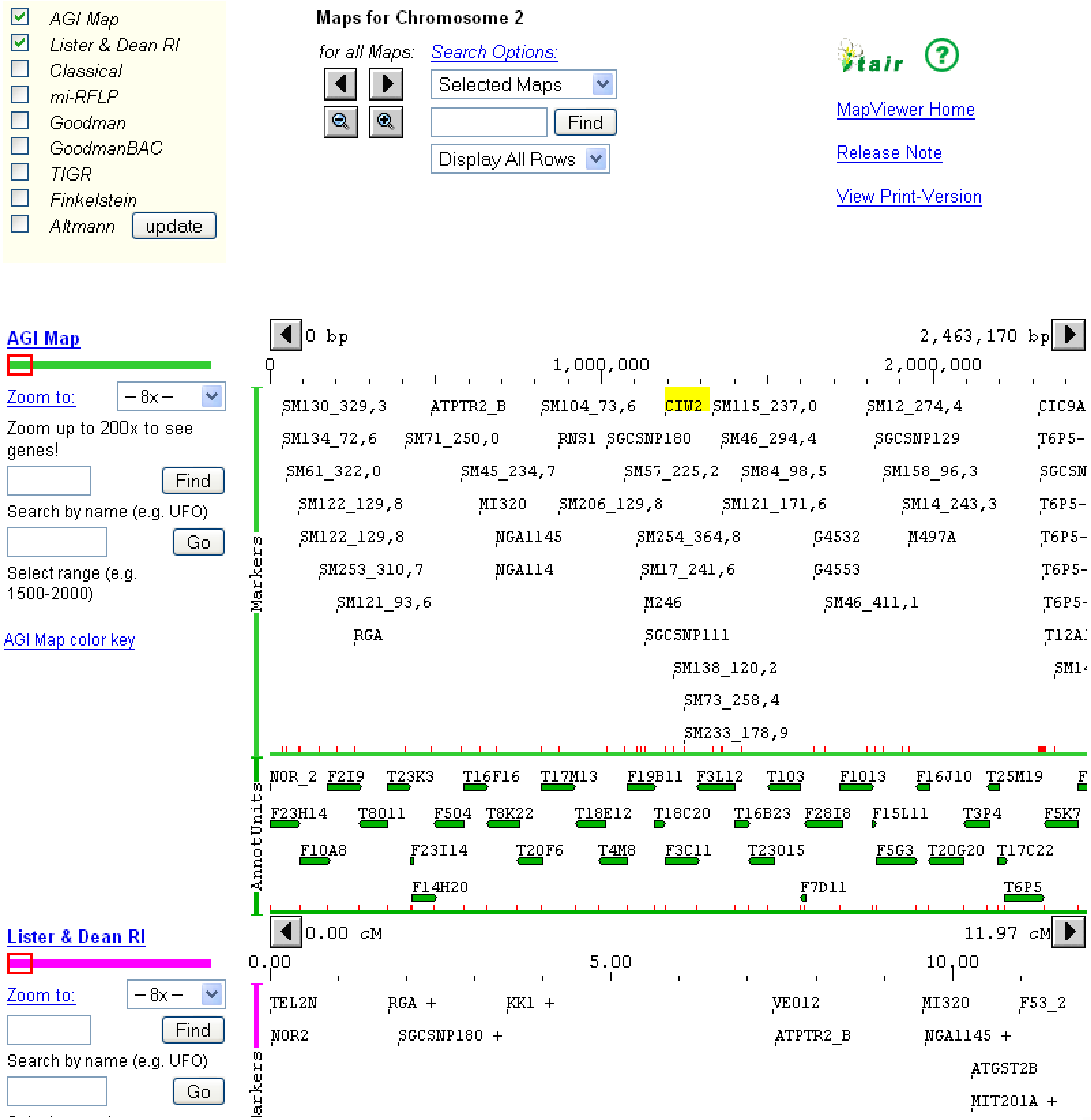
marker II - no linkage
6 mutants
7/12 × 100 = 58%

- Analysis of approximately 2000 mutant plants
- Determining the closest (still segregating) marker
- Identification of mutation by sequencing

# Map of DNA molecular markers



# Markers for fine mapping



☑ AGI Map
☑ Lister & Dean RI
☐ Classical
☐ mi-RFLP
☐ Goodman
☐ GoodmanBAC
☐ TIGR
☐ Finkelstein
☐ Altmann

Maps for Chromosome 2

for all Maps: Search Options: Selected Maps

Display All Rows

AGI Map

Zoom to: – 8x –

Zoom up to 200x to see genes!

Search by name (e.g. UFO)

Select range (e.g. 1500-2000)

Lister & Dean RI

Zoom to: – 8x –

Search by name (e.g. UFO)

# Key Concepts

- Forward genetics allows targeted screening for interesting phenotypes, whose association with a given gene/locus is unknown
  - Employs both insertional mutagenes as well as point mutations
    - **Inserional mutation**
      - (mostly) loss-of-function mutation
        - Identification via
          - iPCR
          - plasmid rescue
    - **Point mutation**
      - Both loss-of-function as well as
      - gain-of-function mutations
        - Identification via
          - map-based cloning
          - GWAS

# Discussion