

# CG020 Genomika

## Přednáška 4

### Genetika přímá

Jan Hejátko

**Funkční genomika a proteomika rostlin,**  
Středoevropský technologický institut (CEITEC)

a

**Národní centrum pro výzkum biomolekul,**  
Přírodovědecká fakulta,

MUNI  
SCI

Masarykova univerzita, Brno  
[hejatko@sci.muni.cz](mailto:hejatko@sci.muni.cz), [www.ceitec.eu](http://www.ceitec.eu)



# Osnova

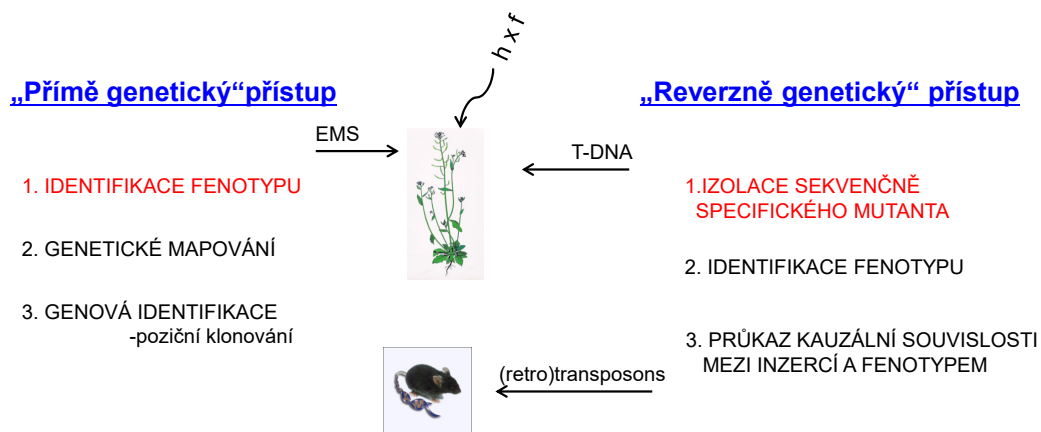
- Přímá vs. reverzní genetika
- Využití knihoven inzerčních mutantů v postupech přímé genetiky
  - vyhledávání v knihovnách inzerčních mutantů podle
    - anatomicky nebo morfologicky detekovatelného fenotypu
    - metabolického profilu
    - exprese zajímavých genů
  - identifikace mutovaného lokusu
    - plasmid rescue
    - iPCR
- Využití knihoven bodových mutantů v přímé genetice
  - poziční klonování
  - GWAS

# Osnova

- Přímá vs. reverzní genetika

## Přístupy „klasické“genetiky versus „reverzně genetický“ přístup ve funkční genomice

### NÁHODNÁ MUTAGENEZE

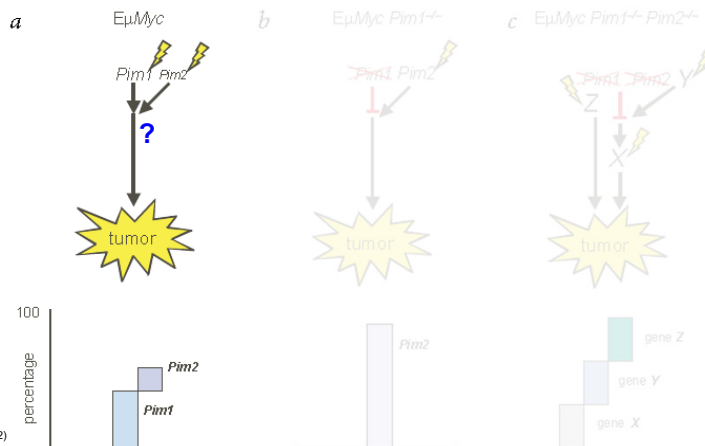


# Osnova

- Přímá vs. reverzní genetika
- Využití knihoven inzerčních mutantů v postupech přímé genetiky
  - vyhledávání v knihovnách inzerčních mutantů podle
    - anatomicky nebo morfologicky detekovatelného fenotypu

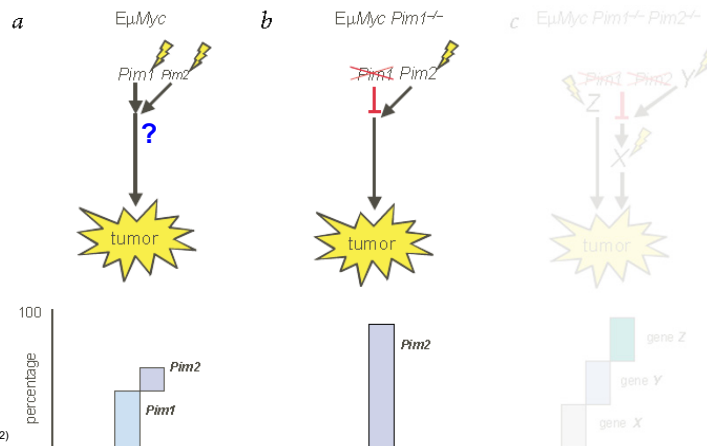
# Inzerční mutagenese v přímé genetice

- Využití inzerční mutagenese ve studiu kancerogeneze
  - Infekce EμMyc myší **retrovirem MoMuLV** vede k tvorbě lymfomů, které vznikly díky **aktivaci Pim kináz** (ve 40% aktivaci *Pim1* a v 15% aktivaci *Pim2*), molekulární **cíle těchto kináz** byly **neznámé**



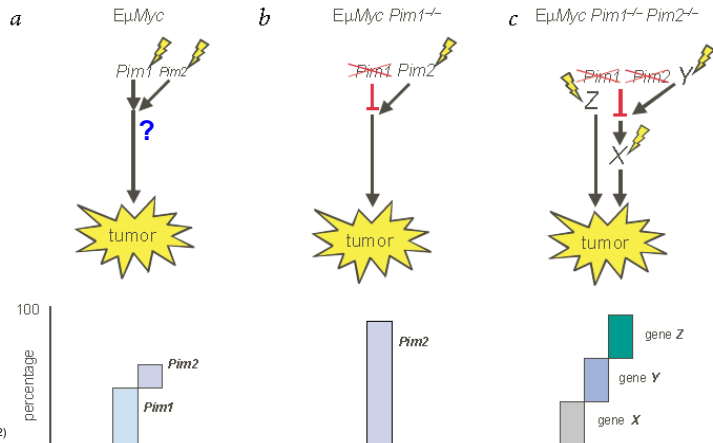
# Inzerční mutagenese v přímé genetice

- Využití inzerční mutagenese ve studiu kancerogeneze
  - Infekce EμMyc *pim1* mutantů retrovirem MoMuLV vede k tvorbě lymfomů, které obsahují v 90% inzerce v blízkosti (aktivaci) *Pim2*



# Inzerční mutagenese v přímé genetice

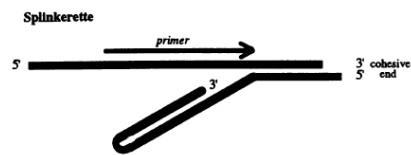
- Využití inzerční mutagenese ve studiu kancerogeneze
  - Infekce EμMyc dvojnásobných mutantů *pim1*, *pim2* retrovirem MoMuLV vede k tvorbě lymfomů, u kterých lze očekávat aktivaci buď některého ze **signálních partnerů Pim proteinů (Y)**, některého z **proteinů Pim signální dráhy (X)** nebo k **aktivaci některé z příbuzných drah vedoucích k lymfomagenezi (Z)**





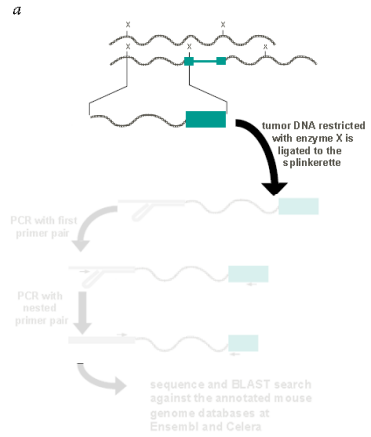
# Inzerční mutagenéze v přímé genetice

- Izolace genomových oblastí přilehajících k místu inzerce proviru
  - Štěpení genomové DNA a ligace speciálních linkerů, tzv. *splinkerett* (zvýšení specificity amplifikace)



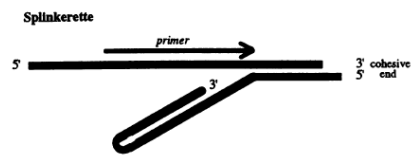
Devon et al., Nucl Acid Res (1994)

Mikkers et al., Nature Gen (2002)

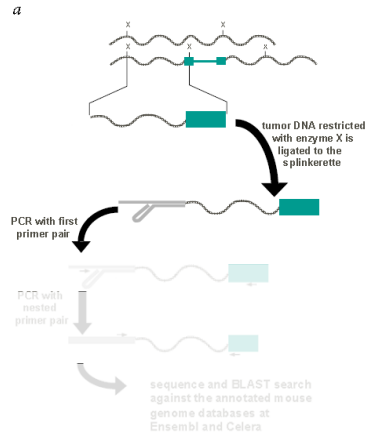


# Inzerční mutagenese v přímé genetice

- Izolace genomových oblastí přilhajících k místu inserce proviru
  - První amplifikace pomocí specifických primerů



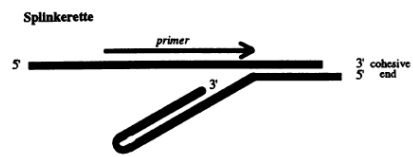
Devon et al., Nucl Acid Res (1994)



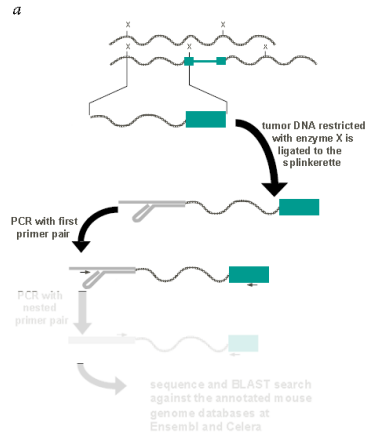
Mikkers et al., Nature Gen (2002)

# Inzerční mutagenese v přímé genetice

- Izolace genomových oblastí přilhajících k místu inserce proviru
  - Druhá amplifikace pomocí „nested“ primerů (zvýšení specifity)



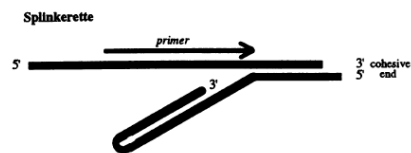
Devon et al., Nucl Acid Res (1994)



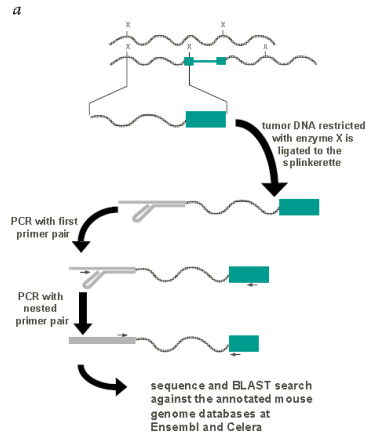
Mikkers et al., Nature Gen (2002)

# Inzerční mutagenese v přímé genetice

- Izolace genomových oblastí přiléhajících k místu inserce proviru
  - Sekvence a lokalizace oblastí přiléhajících k proviru vyhledáváním v anotovaných databázích myšního genomu



Devon et al., Nucl Acid Res (1994)



Mikkers et al., Nature Gen (2002)

In case of splinkerette, the primer is of the same sequence as the top strand and therefore it is unable to act as a primer until the complement of this strand has been synthesized (from the insert-specific primer at the right-hand side).

# Osnova

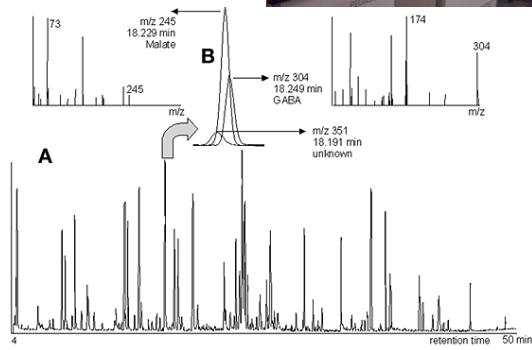
- Přímá vs. reverzní genetika
- Využití knihoven inzerčních mutantů v postupech přímé genetiky
  - vyhledávání v knihovnách inzerčních mutantů podle
    - anatomicky nebo morfologicky detekovatelného fenotypu
    - **metabolického profilu**

# Metabolické profilování

- Metabolické profilování rostlin
  - hromadná a automatizovaná analýza metabolitů (až 25.000) pomocí GC-MS technik v knihovnách T-DNA mutantů

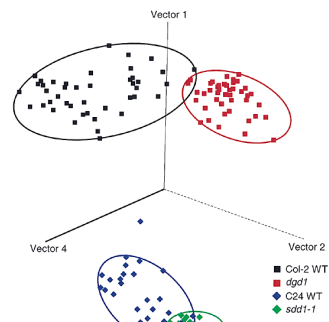


Retention Time (min)	m/z	Abundance	Identification
18.229	245	High	Malate
18.249	304	High	GABA
18.191	351	High	unknown



# Metabolické profilování

- Metabolické profilování rostlin
  - hromadná a automatizovaná analýza metabolitů (až 25.000) pomocí GC-MS technik v knihovnách T-DNA mutantů
  - identifikace (např. i komerčně) zajímavých mutantů



# Metabolické profilování

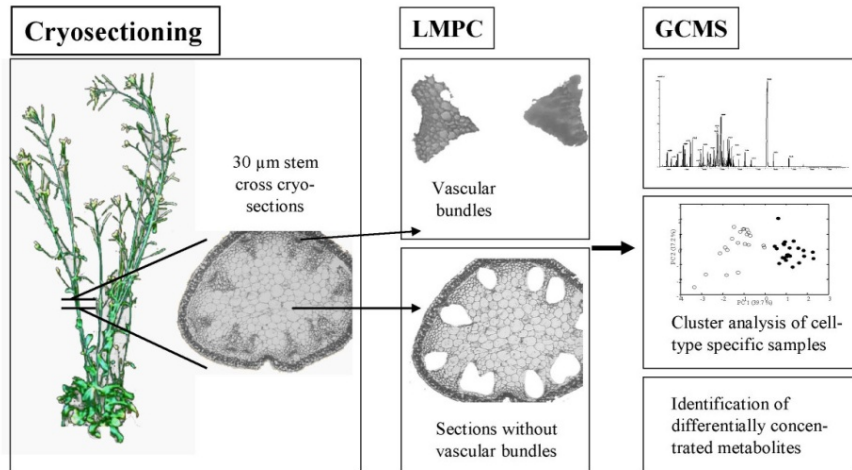
- Metabolické profilování rostlin
  - hromadná a automatizovaná analýza metabolitů (až 25.000) pomocí GC-MS technik v knihovnách T-DNA mutantů
  - identifikace (např. i komerčně) zajímavých mutantů
  - snadná a rychlá izolace genů pomocí identifikace T-DNA zasažených sekvencí





# Metabolické profilování

- Metabolické profilování rostlin
  - možnost využít i speciální techniky, např. **mikrodisekce**

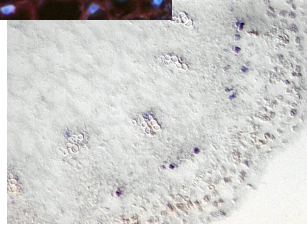
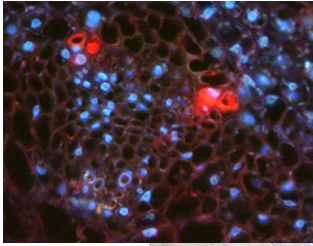


# Osnova

- Přímá vs. reverzní genetika
- Využití knihoven inzerčních mutantů v postupech přímé genetiky
  - vyhledávání v knihovnách inzerčních mutantů podle
    - anatomicky nebo morfologicky detekovatelného fenotypu
    - metabolického profilu
    - exprese zajímavých genů

# Expresní profil

- Identifikace mutantů se změnou expresního profilu
  - analýza expresního profilu (vzorce) daného genu a identifikace mutantů se změnou exprese



# Expresní profil

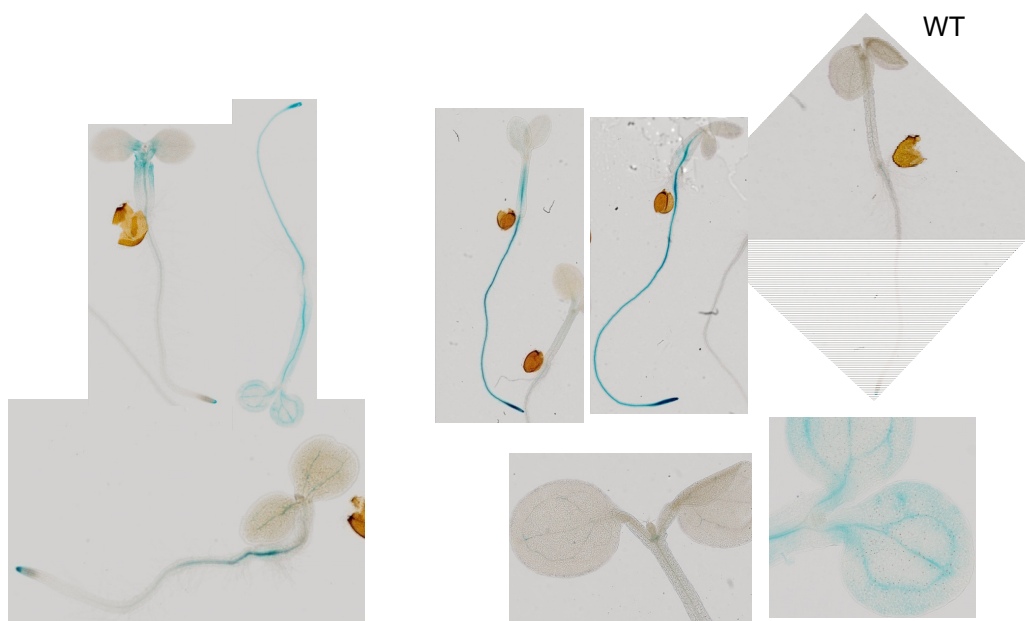
- Identifikace mutantů se **změnou expresního profilu**
  - analýza **expresního profilu** (vzorce) daného genu a identifikace **mutantů se změnou exprese**
  - možnost **částečné automatizace** (virtuální digitální mikroskopie)

# Vyhledávání pomocí automatické mikroskopie



Dobisova and Hejatko, *Methods in Mol Biol*, 2014

# Expresní profil



22

CEITEC

# Osnova

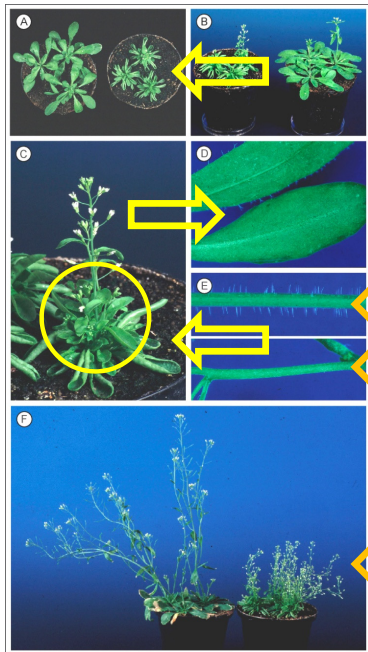
- Přímá vs. reverzní genetika
- Využití knihoven inzerčních mutantů v postupech přímé genetiky
  - vyhledávání v knihovnách inzerčních mutantů podle
    - anatomicky nebo morfologicky detekovatelného fenotypu
    - metabolického profilu
    - exprese zajímavých genů
  - identifikace mutovaného lokusu
    - plasmid rescue
    - iPCR

# Identifikace mutovaného lokusu

- Identifikace chromozomální přestavby zodpovědné za keříčkovitý fenotyp u *Arabidopsis*
  - popis fenotypu

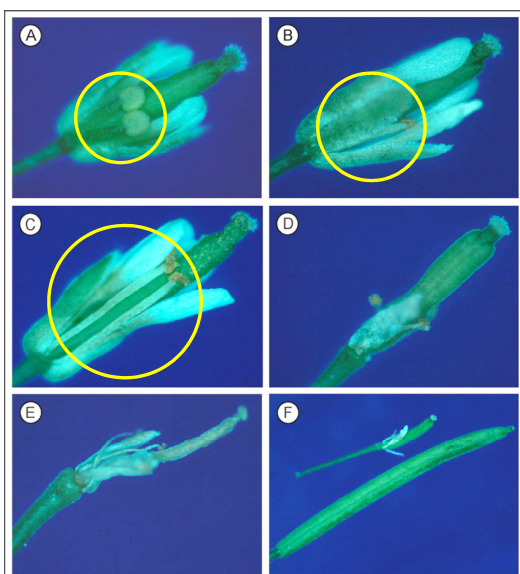


## Identifikace mutanta



- zvlňné listy
- keřčkovitý fenotyp (poruchy větvení)
- chybějící trychomy na listech a na stonku
- opožděné stárnutí

## Identifikace mutanta



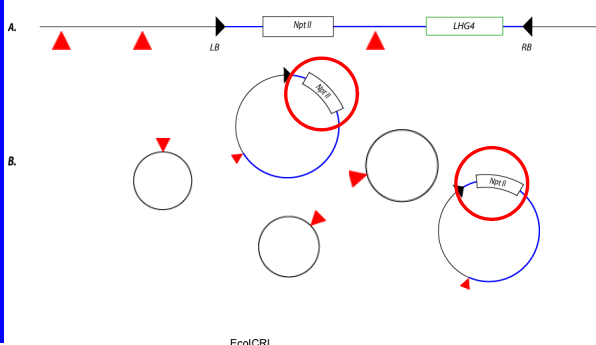
- samčí sterilita, poruchy v prodlužování tyčinek (A,B) (porovnej se standardním typem C)

# Identifikace mutovaného lokusu

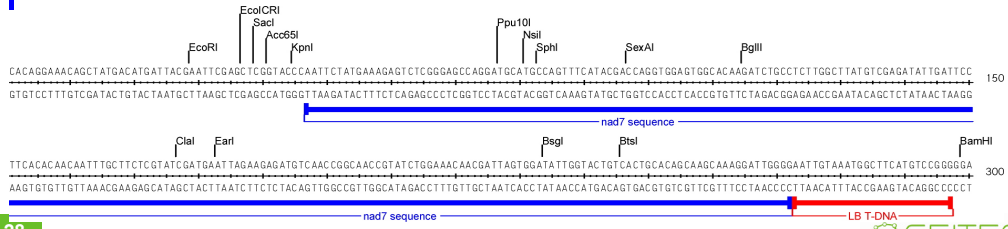
- Identifikace chromozomální přestavby zodpovědné za keříčkovitý fenotyp u *Arabidopsis*
  - popis fenotypu
  - identifikace T-DNA mutované oblasti

# Identifikace mutovaného lokusu

## 1. Identifikace oblasti genomové DNA přiléhající k *levé hranici* pomocí *plasmid rescue*

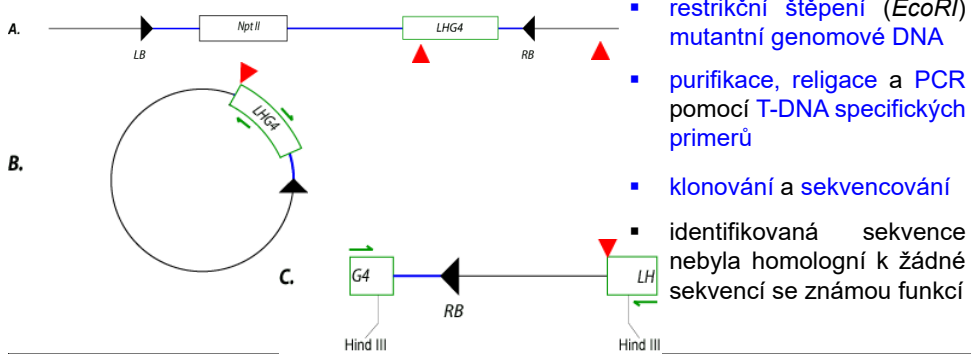


- restrikční štěpení (*EcoRI*) mutantní genomové DNA
- religace a transformace *E.coli*
- izolace plazmidové DNA z pozitivně selektovaných klonů
- identifikovaná sekvence byla identická s genem pro **NAD7** kódovaným na *mtDNA*

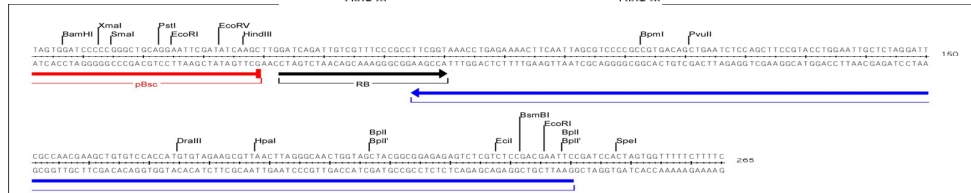


# Identifikace mutovaného lokusu

## 2. Identifikace oblasti genomové DNA přiléhající k **pravé hranici** pomocí **inverzní PCR** (iPCR)



- restrikční štěpení (*EcoRI*) mutantní genomové DNA
- purifikace, religace a PCR pomocí T-DNA specifických primerů
- klonování a sekvencování
- identifikovaná sekvence nebyla homologní k žádné sekvenci se známou funkcí

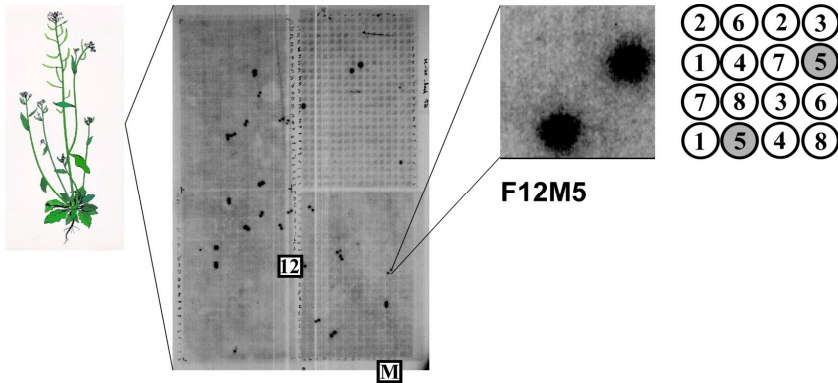


# Identifikace mutovaného lokusu

- Identifikace chromozomální přestavby zodpovědné za keříčkovitý fenotyp u *Arabidopsis*
  - popis fenotypu
  - identifikace T-DNA mutované oblasti
  - lokalizace T-DNA inzerce v genomu *Arabidopsis*

## Vyhledávání v knihovně IGF-BAC

- genomová knihovna obsahující 10,752 klonů s průměrnou velikostí inzertu 100 kb
- bakteriální klony uspořádané v mikrotitračních deskách
- knihovna nanášena na nylonové filtry pro hybridizaci s radioaktivně značenou sondou



## Mapování pomocí IGF-BAC databáze

### **I. Sekvence přiléhající k levé hranici T-DNA**

- celkem 28 pozitivně hybridizujících klonů
- 19 z nich lokalizováno na chromozomu 2
- 18 s podobností k mtDNA

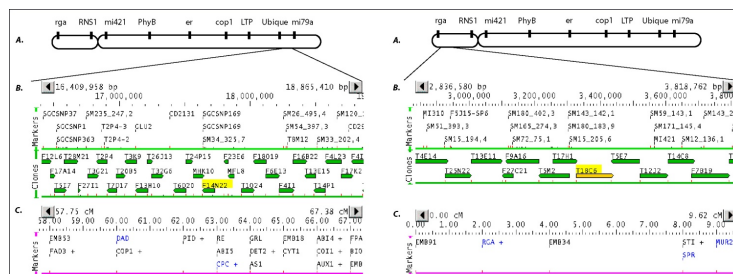
### **II. Sekvence přiléhající k pravé hranici T-DNA**

- celkem 6 pozitivně hybridizujících klonů
- všechny lokalizovány na chromozomu 2

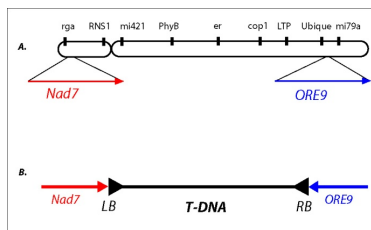


## Lokalizace genomové T-DNA přiléhající k levé i pravé hranici T-DNA na chromozomu 2

### Sekvence přiléhající k *pravé* a *levé* hranici T-DNA



- pravidelně došlo k inverzi téměř celého chromozómu 2



# Osnova

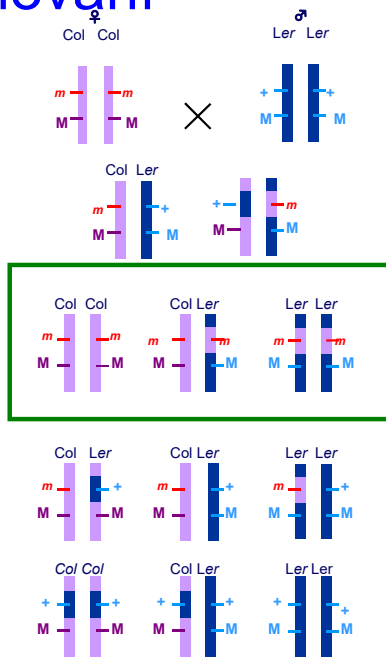
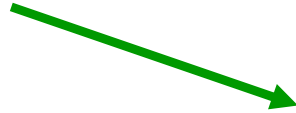
- Přímá vs. reverzní genetika
- Využití knihoven inzerčních mutantů v postupech přímé genetiky
  - vyhledávání v knihovnách inzerčních mutantů podle
    - anatomicky nebo morfologicky detekovatelného fenotypu
    - metabolického profilu
    - exprese zajímavých genů
  - identifikace mutovaného lokusu
    - plasmid rescue
    - IPCR
- Využití knihoven bodových mutantů v přímé genetice
  - poziční klonování

# Identifikace mutovaného lokusu

- Poziční klonování
  - podstatou je kosegregační analýza segregující populace (většinou potomstva informativního zpětného křížení) s molekulárními markery
  - **SSLP** (Simple Sequence Length Polymorphism)
    - **polymorfismus délky genomu** (PCR produktů) **amplifikovaného** pomocí specifických **primerů**
  - **RFLP** (Restriction Fragment Length Polymorphism)
    - **polymorfismus délky restrikčních fragmentů** úseků genomu, detekce pomocí Southern blotu (PCR po naštěpení genomové DNA a ligaci adaptorů)
  - **CAPS** (Cleaved Amplified Polymorphic Sequence)
    - **polymorfismus délky restrikčních fragmentů** úseků genomu amplifikovaných pomocí **PCR**
  - **RAPD** (Randomly Amplified Polymorphic DNA)
    - **polymorfismus délky náhodně** (pomocí krátkých primerů, 8-10 bp) **amplifikovaných úseků genomu**

# Poziční klonování

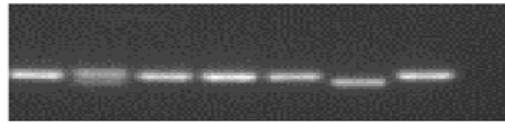
Příprava mapovací populace



**Rekombinantní analýza** – určení procenta rekombinace mezi mutací a molekulárním markerem

$$r [\%] = \frac{\text{počet chomozomů } Col /}{\text{počet všech chromozomů}} \times 100$$

**F2 mutanti**



marker I – ve vazbě  
5 mutantů  
 $1/10 \times 100 = 10\%$

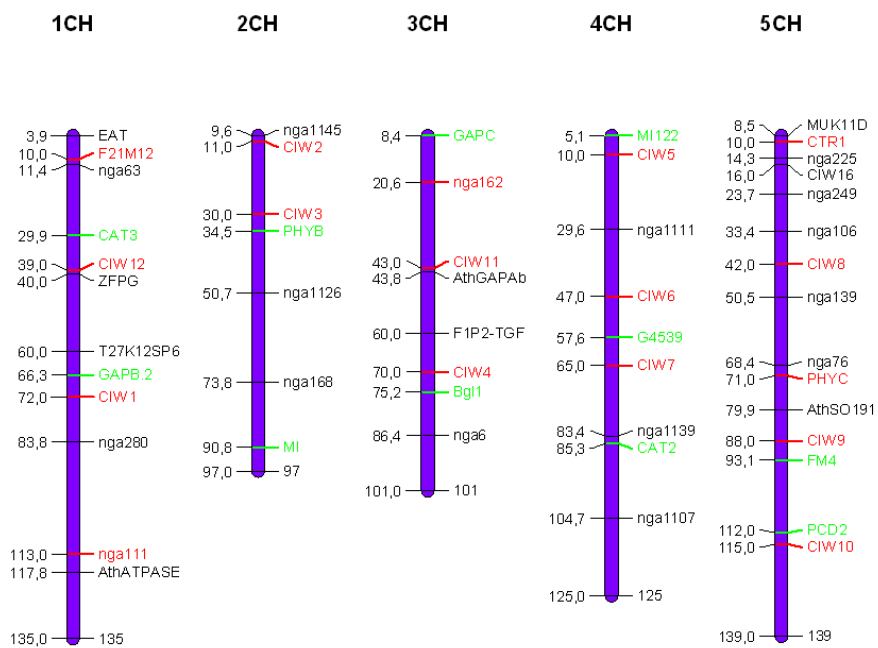
**F2 mutanti**



marker II - žádná vazba  
6 mutantů  
 $7/12 \times 100 = 58\%$

- Analýza cca 2000 mutantních linií
- Určení nejbližšího (ještě) segregujícího markeru
- Identifikace mutace pomocí sekvenování

## Mapa DNA molekulárných markerů



# Markery pro jemné mapování

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:

 [MapViewer Home](#)

[Release Note](#)

[View Print-Version](#)

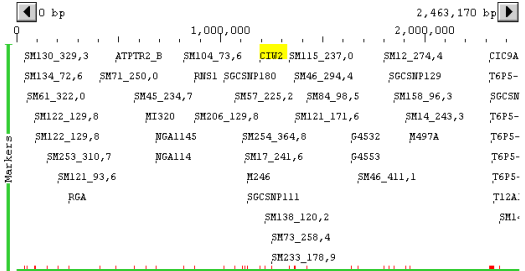
**AGI Map** 0 bp  2,463,170 bp

Zoom to:

Zoom up to 200x to see genes!

Select range (e.g. 1500-2000)

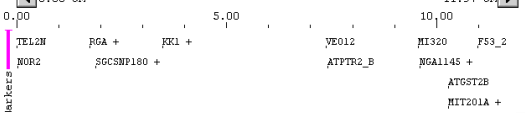
[AGI Map color key](#)




**Lister & Dean RI** 0.00 cM  11.97 cM

Zoom to:

Select range (e.g. UFO)



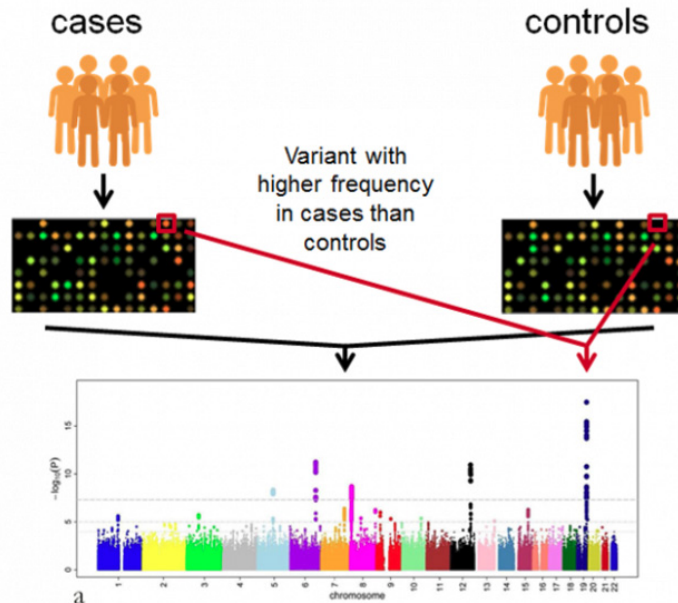


# Osnova

- Přímá vs. reverzní genetika
- Využití knihoven inzerčních mutantů v postupech přímé genetiky
  - vyhledávání v knihovnách inzerčních mutantů podle
    - anatomicky nebo morfologicky detekovatelného fenotypu
    - metabolického profilu
    - exprese zajímavých genů
  - identifikace mutovaného lokusu
    - plasmid rescue
    - IPCR
- Využití knihoven bodových mutantů v přímé genetice
  - poziční klonování
  - GWAS



# Genome Wide Association Study - GWAS



41

<https://www.ebi.ac.uk/training-beta/online/courses/gwas-catalogue-exploring-snp-trait-associations/what-is-gwas-catalog/what-are-genome-wide-association-studies-gwas/>

A typical GWAS study collects data to find out the common variants in a number of individuals, both with and without a common trait (e.g. a disease), across the genome, using genome wide SNP arrays. Variants associated with the disease, or within the same haplotype as a variant associated with a disease, will be found at a higher frequency in cases than in controls. Statistical analysis is carried out to indicate how likely a variant is to be associated with a trait.

As GWAS analyse common variants, usually typed on commercial SNP arrays (Figure 3), they do not generally identify causal variants. GWAS identify common variants which tag a region of linkage disequilibrium (LD) containing causal variant(s). Additional or follow-on studies are usually required to narrow the region of association and identify the causal variant

# Klíčové koncepty

- **Přímá genetika** umožňuje cíleně vyhledávat **zajímavé fenotypy**, jejichž **asociace s určitým genem/lokusem** není dosud známa
  - Lze využít jak **inzerční mutageny**, tak **mutace bodové**
    - **Inzerční mutace**
      - (většinou) mutace se **ztrátou funkce** (loss-of-function)
        - iPCR
        - plasmid rescue
      - **Bodové mutace**
        - Jak **mutace se ztrátou funkce** (loss-of-function), tak i
        - Mutace se **získanou funkcí** (gain-of-function)
          - **poziční klonování**
          - **GWAS**

# Diskuse