CG920 Genomics

Lesson 4

Forward Genetics

Jan Hejátko

Functional Genomics and Proteomics of Plants,

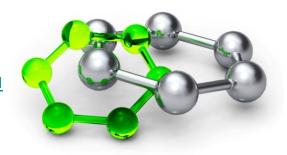
CEITEC - Central European Institute of Technology
And

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



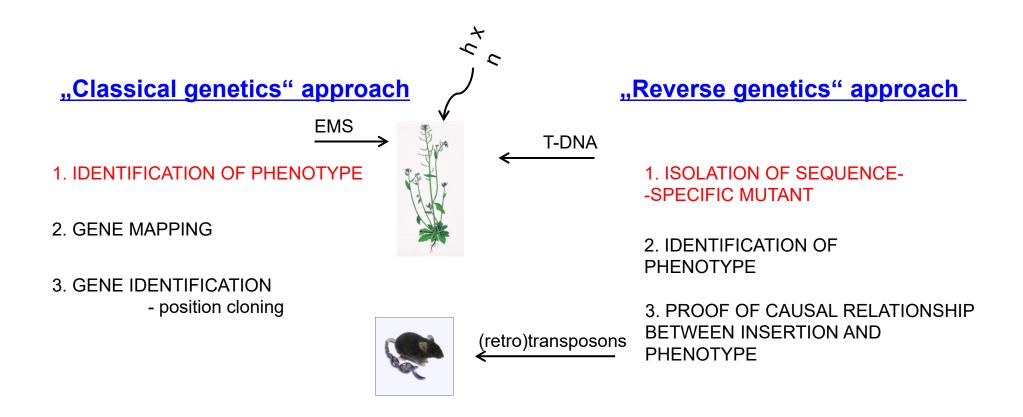
Outline

Forward vs. Reverse Genetics



"Classical" genetics *versus* "reverse genetics" approaches in functional genomics

RANDOM MUTAGENESIS



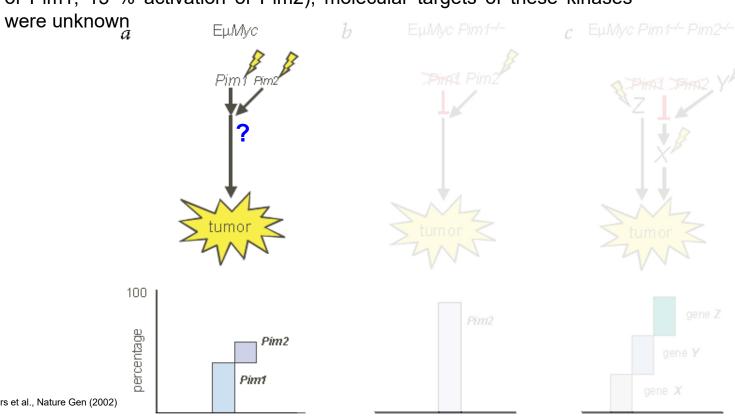
Outline

- 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

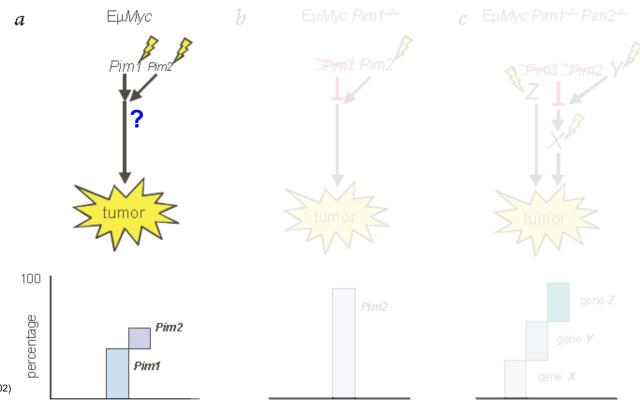


Use of insertional mutagenesis for study carcinogenesis

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

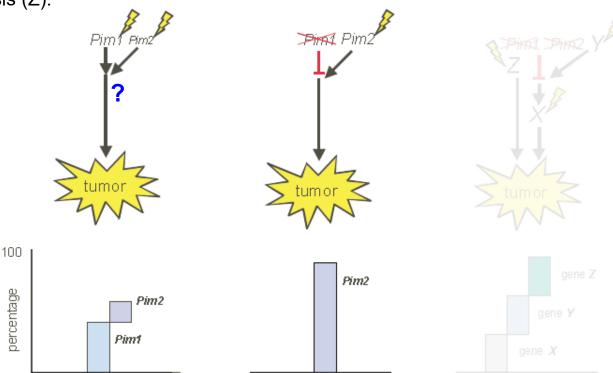


- Use of insertional mutagenesis for study of carcinogenesis
 - Infection of EμMyc pim1 mutants by MoMuLV retrovirus leads to lymphomas formation, which in 90 % contain insertion nearby (activation) Pim2



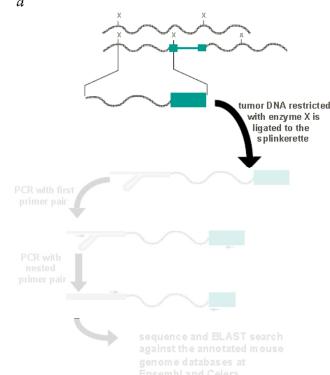
Use of insertional mutagenesis for study of carcinogenesis

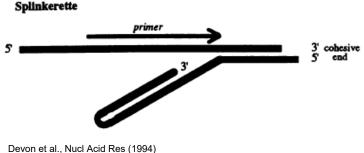
Infection of EμMyc double mutants *pim1*, *pim2* by MoMuLV retrovirus leads to lymphomas formation, which can be expected to activate either one of the signalling partner of Pim proteins (Y), one of the downsteram proteins of Pim signalling pathway (X) or to activate some of the related pathways leading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



 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)

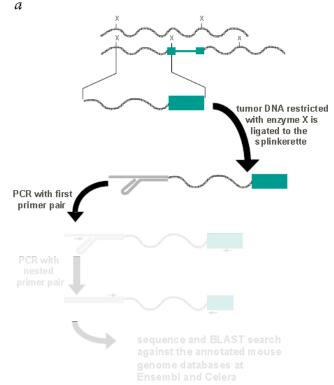






 Isolation of genomic regions adjacent to the insertion site of the provirus

First amplification using specific primers



Splinkerette

primer

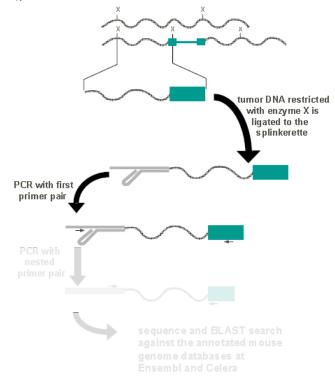
3' cohesive ond

Devon et al., Nucl Acid Res (1994)



 Isolation of genomic regions adjacent to the insertion site of the provirus

 Second amplification using nested primers (increasing the specifity)



Splinkerette

primer

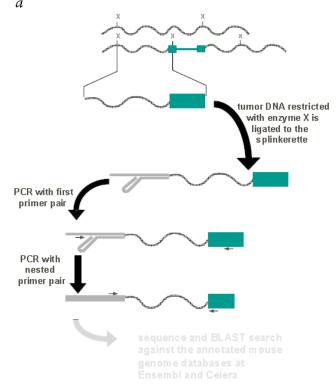
3' cohesive cnd

Devon et al., Nucl Acid Res (1994)



 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



Splinkerette

primer

3' cohesive cnd

Devon et al., Nucl Acid Res (1994)



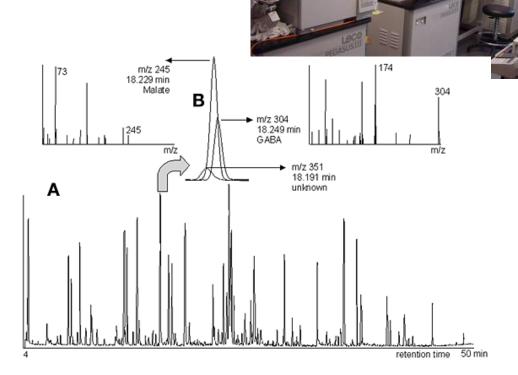
Outline

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

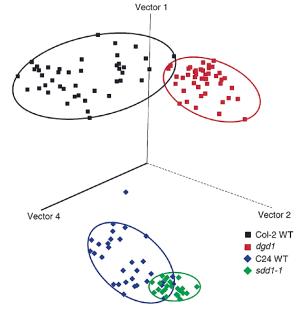
| retaboline | WY 58.7 | 110 565 | 130 56.5 | 10 103 | W1 16.5 | 95.H M.S. |
|---|-------------|--------------------------------------|-------------------------------|---------------------------|------------------------|-------------------------------|
| davive | 1.00 1 0.32 | 176 5 917 | 240 7 430 | 206 7 836 | 100 1 0.16 | 331 * 616 |
| argrine | 100 1 029 | 146 7 936 | 164 7 624 | 1,99 1 6,28 | 100 1 020 | 2.43 1 0.39 |
| акопун | 100 1 019 | 1.31 * 0.14 | 106 2 609 | 127 1 629 | 100 7 606 | 121 1 047 |
| b-statute | 100 1 011 | 120 1 006 | 117 5 617 | 134 1 607 | 100 1 000 | 136 1 632 |
| b-statute college | 100 1 616 | 217 1 922 | 230 1 641 | 137 1 627 | 100 7 020 | 119 1 046 |
| GARA | 100 1 619 | 120 1 000 | 122 1 611 | 1.40 4 6.13 | 100 4 0 13 | 0.79 ± 0.05 |
| glutamete | 100 1 607 | 122 4 957 | 169 1 647 | 116 1 000 | 100 1 004 | 100 1 010 |
| glidanas | 100 1 036 | 039 1 017 | 976 1 639 | 1 17 L 0 22 000 1 0 10 | 100 7 016 | C85 1 G16 |
| - gyone - hishine | 100 1 619 | 987 1 907 | 040 - 610 | *** | 100 7 000 | 166.33 2 163.73 |
| nomocystere . | 1.6 | 5.6 | 6.0 | 0.0 | 4.4 | 14 |
| 5 homogluturerar | 116 | nd | 4.6 | 0.4 | md. | 44 |
| honoserre | 100 I 612 | 100 7 9 14 | 990 7 508 | 1 58 A 0.95 | 100 1 0.06 | 6.44 1 G 62 G 90 1 G 10 |
| 6 holestw | 100 1 619 | 120 1 0 14 | 111 5 029 | 102 1 000 | 100 7 0.00 | 127 1 0 22 |
| 7 bone | 100 1 016 | 950 1 931 | 951 1 500 | 204 1 000 | 100 A 4.25 | 679 1 014 |
| 0 nethinine | 100 1 613 | 984 7 919 | 056 7 033 | 126 1 0 11 | 100 * 614 | 689 1 018 |
| noneucine . | 100 1 612 | 080 1 913 | 0.00 I 0.00 | 931 - 944 | 100 5 6.24 | 677 A 014 |
| novelne ordine | 100 8 614 | 966 1 9CF | 117 5 020 | 846 f 866 860 f 968 | 100 1 612 | 1.00 / 6.20 |
| - destana | 100 4 630 | 076 - 027 | 984 7 939 | 999 7 9 15 | 100 * 4.06 | 100 1 019 |
|) phenylsterine | 160 1 620 | 179 7 839 | 118 2 016 | 1.17 7 0.09 | 100 4 613 | 1.89 4 6.15 |
| s protee | 100 I 018 | 059 A 007 | 105 4 009 | 120 8 0 12 | 100 # €08 | 1.55 * 6.16 |
| 5 serre | 100 1 620 | 284 * 0.19 | 2.60 7 6.20 | 1.10 5 0.10 | 100 1 608 | 400 t 0.86 |
| f. treatine | 100 1 021 | 089 1 0 10 69083 1 215.72 | 1 05 E 0 11 244.00 7 24.56 | 20150 7 65 53 | 100 7 6 65 | 1749 1 426 |
| 7 Ingitation 6 Movine | 100 5 628 | 194 1 032 | 9 67 E 9 19 | 134 7 0 16 | 100 1 0 43 | 5.04 A G.86 |
| 7 value | 160 1 614 | 084 * 908 | 101 7 000 | 1.01 7 0.00 | 100 1 6.01 | 687 1 610 |
| soorbate | 1.60 * 631 | 066 * 924 | 976 2 919 | 140 7 909 | 100 2 6.11 | 239 ± 459 |
| CENN | 100 1 600 | 107 1 009 | 112 5 006 | 100 1 000 | 100 # 4.08 | 6M 1 017 |
| tumorate staturonate | 160 * 631 | 000 4 9 60 | 424 * 440 | 14.00 | 100 1 612 100 1 611 | 630 1 636 |
| Marron | 14 | 7440 1 539 | 2000 1 424 | 2500 / 120 | 44 | 24 |
| ghowale. | 160 1 600 | *** * *** | 5.60 T 6.00 | 164 7 434 | 100 7 6 08 | 630 : 652 |
| nochrie | 160 4 610 | 079 7 908 | 104 - 019 | 1.46 2 6.19 | 100 4 6 97 | 4.42 1 4.47 |
| relate a contribu | 1.00 5 610 | 1.02 * 0.21 | 100 1 430 | 064 1 0.12 | 100 7 6.06 | 186 1 617 |
| contracts contracts | 100 1 612 | 136 1 100 | 127 1 412 | 181 7 6 9 | 100 7 610 | 678 1 010 |
| 1 sheron | 100 1 617 | 270 1 925 | 230 1 050 | 150 1 028 | 100 / 608 | 664 1 0 17 |
| 2 succinete | 160 1 621 | 234 1 9.19 | 2 14 7 0 30 | 165 5 625 | 100 1 618 | 2.50 ± 6.56 |
| 2 fewrole | 1.00 1 619 | 130 1 400 | 196 2 011 | 091 1 019 | 100 1 606 | 124 + 014 |
| . bucken | 100 1 044 | 182 5 052 | 547 7 310 | 025 1 0.09 | 100 7 6 13 | 601 1 255 |
| galaction glaces | 100 1 041 | 429 1 447 | 356 7 852 | 025 7 0.06 | 100 7 8 16 | 220 1 133 |
| down | 110 1 041 | 14 | *** | 1.6 | 100 1 612 | 6.00 1 6.00 |
| inenti | 100 1 611 | 0.16 2 0.05 | 828 1 804 | 9.40 1 9.06 | 100 7 620 | E.40 ± 6.04 |
| S STATUTE . | 6.6 | 17.00 1 2.05 78000 1 100.64 | ** T 17.05 | 1.0 | 1.0 | 13600 1 45.13 |
| militar militar | n.d | 790.30 1 100.64 1000.80 1 1107.27 | 384.00 T 17.05 | ** | 14 | 1975.00 t 40.13 |
| rents | 100 7 600 | 322 - 422 | 275 7 430 | 1.12 1 0 13 | 100 1 604 | 1.00 + 1.50 |
| C manness | 160 7 636 | 742 5 658 | 424 1 674 | 086 7 018 | 100 7 611 | 119 5 0.66 |
| 1 sucrose | 100 1 625 | 0.12 - 4.02 | 9.20 = 9.00 | 927 1 938 | 100 : 000 | 104 1 641 |
| C PYNONY | - 64 | 205.35 1 30.29 | 194.17 T 17.36 | 540 / 130 | 100 0 0.10 | 10.30 / 7.34 40.00 / 10.00 |
| fuction FP | 100 1 006 | 29.24 1 2.77 17.62 1 2.12 | U.M. 1 270 | 5.52 × 1.50 | 100 7 6 10 | 91.50 1 16.85 31.52 1 7.42 |
| 2 phosphochanolymine | 100 1 616 | 102 1 008 | 110 - 015 | 974 7 975 | 100 2 0 35 | 297 1 542 |
| 3004 | nd. | D0.00 : 14.41 | 65.19 : 15.90 | 38.11 2 11.53 | 1.4 | 78.67 * 21.60 |
| 6 P-gluconate | 6.6 | C79.30 1 96.00 | 247.17 ± 53.43 | 21306 5 47.87 | 14 | 177.50 + 25.60 |
| phosphata | 100 2 602 | 0N 1 004 | | 082 4 685 | 100 4 900 | 100 1 007 |
| sperndre unenousse | 100 1 631 | 186 7 609 | 1.60 ± 4.00 | 140 1 4.14 | 1.00 1 0.00 | 139 1 444 |
| PTOI | 100 8 605 | 074 * 046 | 045 7 904 | 041 / 444 | 100 1 900 | 972 / 9.19 |
| 9 FT92 | 100 * 610 | 130 1 014 | 1.11 * 0.12 | 126 8 6 19 | 160 7 920 | 5.54 1 6.59 |
| F704 | 100 E 607 | 0.00 (0.00 | | 076 A 0.04 | 140 1 9 19 | 662 t 6.11 |
| P706 | 100 7 614 | 114 F 000 | 0.02 * 0.07 105 * 0.08 | 077 * 004 148 * 038 | 160 7 906 | 199 7 6.14 |
| P106 | 100 I 019 | 113 1 011 | 105 * 008 | 174 1 0 10 | 140 1 949 | 212 1 6.67 |
| PT00 | 100 7 605 | 674 1 849 | 0.76 1 8.05 | 096 1 006 | 140 1 0.10 | 079 7 016 |
| FT00 | 1.00 5 914 | 179 1 0.30 | 126 0 012 | 1.16 * 0.06 | 1.00 * 9.19 | 1.04 7 6.11 |
| | 100 4 025 | 054 7 007 | 0.60 7 0.14 | 100 1 010 | 100 0 0.56 | 1.19 7 6.01 |
| PTN1 | 100 7 612 | 120 1 000 | 1.66 7 0:30 | 079 1 0.09 | 100 1 004 | 147 1 646 |
| PTN2 | 100 7 624 | 147 1 0 20 | 180 7 030 | 136 1 0.00 | 100 1 000 | 132 7 612 |
| FT15 | 100 7 018 | 140 1 0 10 | 146 1 620 | 190 1 0.00 | 160 7 9 99 | 2.62 1 6.22 |
| 6 PTNS | 100 1 619 | 696 1 669 | 6.65 1 8.68 | 205 1 104 | 160 1 939 | 290 T 138 |
| 6 PT17 | 100 1 022 | 141 1 0.17 | 226 1 0 00 | 246 1 0.67 | 160 1 0.14 | 2:65 5 6,38 |
| T PTSS E PTSS | 1.00 1 0.11 | 060 1 000 | 690 F 609 | 064 7 659 | 100 7 0 13 | 977 * 6.16 247 1 1.54 |
| s PTNO | 100 7 041 | 130 1 033 | 134 5 632 | 094 1 000 | 100 1 927 | 100 7 0.04 |
| | 100 1 000 | 282 1 030 | 2.17 1 630 | 090 1 012 | 160 1 905 | |
| | | 184 1 0 79 | 1.55 = 0.15 | 150 1 612 | 100 1 932 | |
| | 100 5 019 | | | 200 7 0.54 | | 129 1 649 |
| 1 PT28 2 PT24 | 100 7 0.17 | 215 1 024 | 2.21 1 0.10 | | | |
| PT28 2 PT24 3 PT25 | 100 7 617 | 150 7 0.06 | 1.22 - 0.06 | 105 2 906 | 100 J 90F | 1.40 / 6.65 |
| 11 PT28 12 PT24 13 PT25 14 PT25 | 100 7 617 | 150 T 036 | 1.22 ± 606 9.71 ± 606 | 105 2 106 | 100 7 905 | 1.40 A 6.85 0.66 1 6.69 |
| 0 9724 11 9728 2 9724 0 9725 X 9725 6 9727 | 100 7 617 | 150 7 0.06 | 1.22 - 0.06 | 105 2 906 | 100 J 90F | 1.40 / 0.05 |





- 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





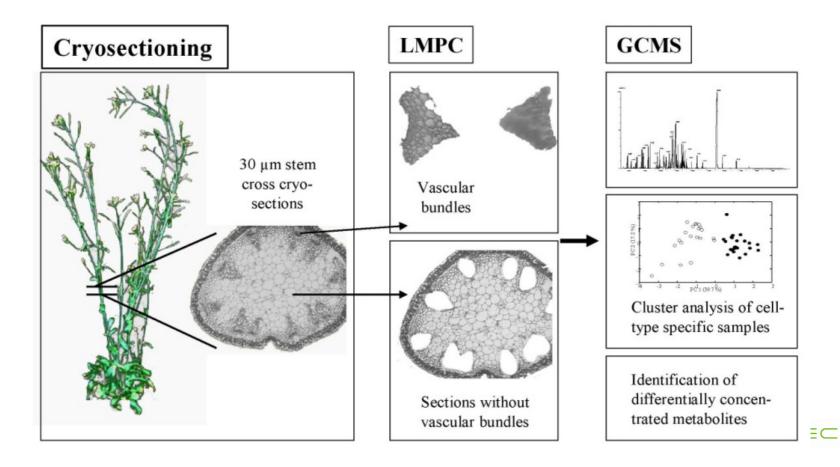


Metabolic profiling of plants

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



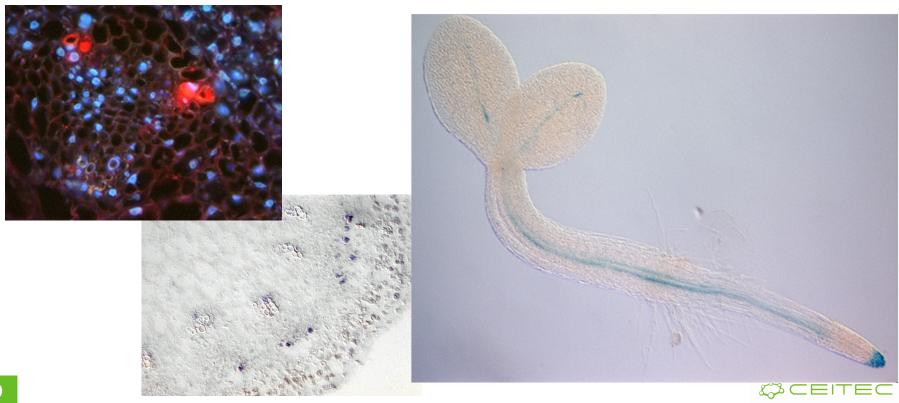
Outline

- 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

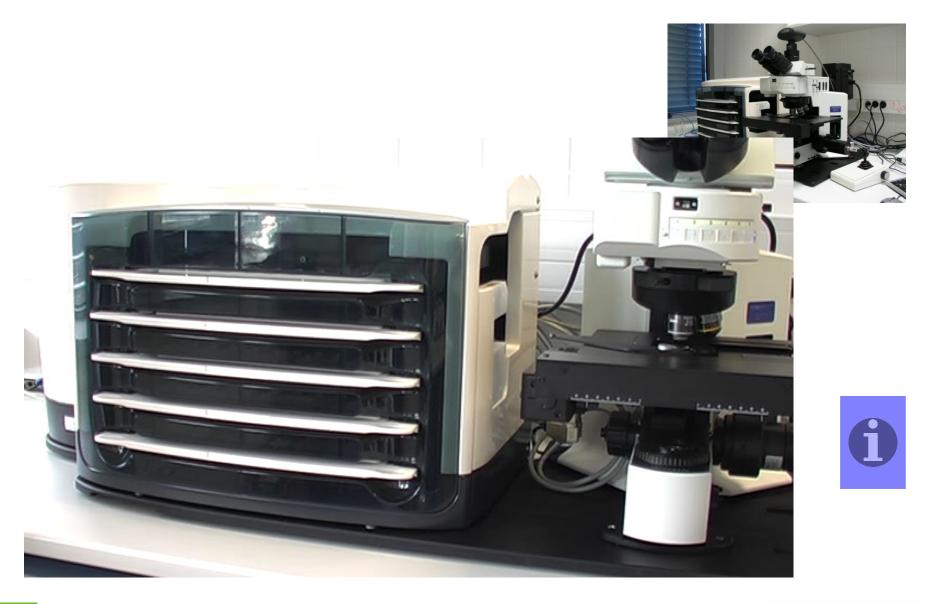


Expression profile

- 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)



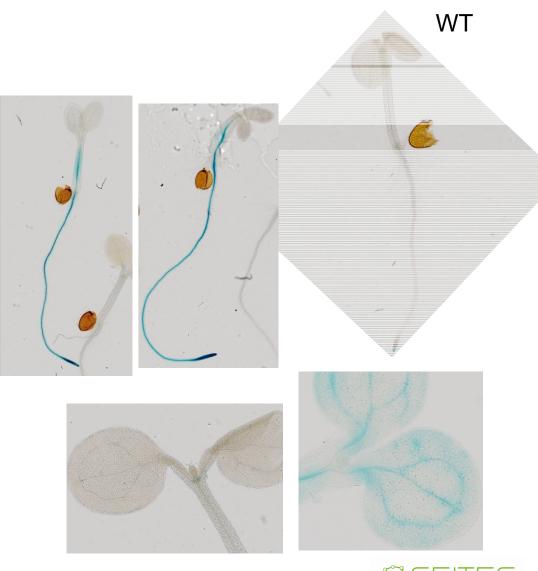
Automated Microscopy Screening





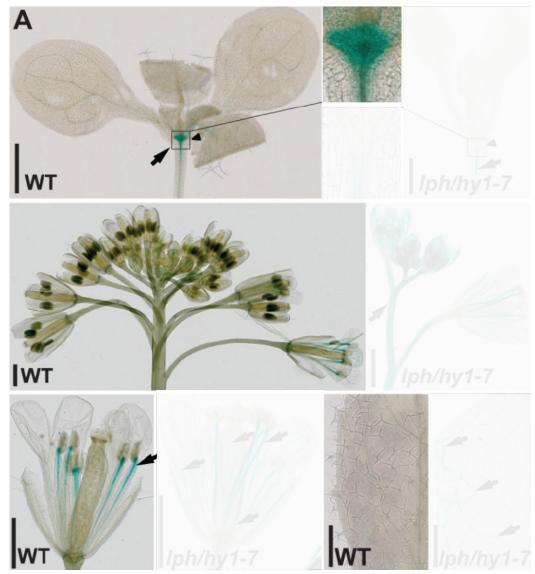
Expression profile







LPH controls expression of CKI1



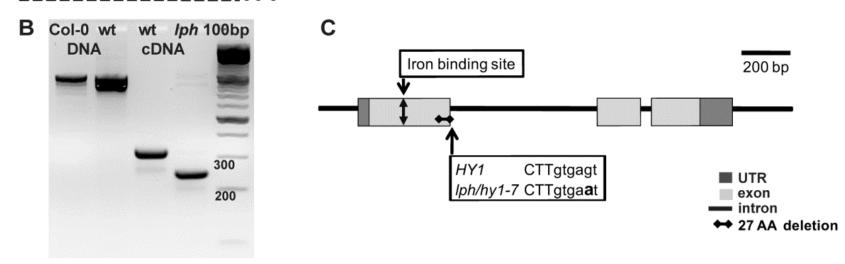




Iph is novel allele of HEME OXYGENASE 1

A <u>CCTACTGTTGAAGGTTACTTGAGGTTTCTTGTGGATAGTAAATTGGTTTATGATACTCTTGAACTGATTA</u>

TTCAAGACTCCAACTTGtgagttttttttt



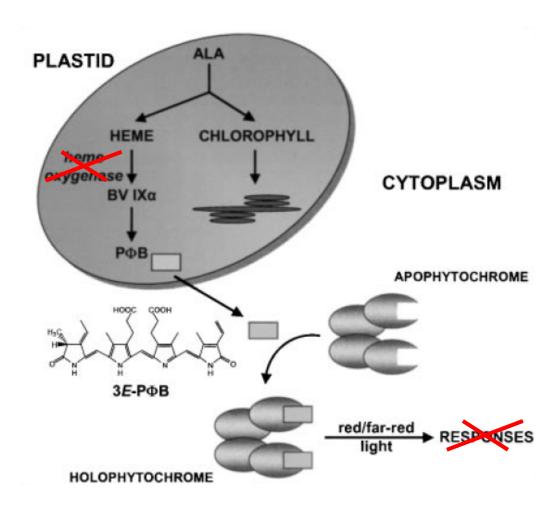
HY1 MAYLAPISSSLSIFKNPQLSRFQFSSSSPNPLFLRPRIQILSMTMNKSPSLVVVAATTAAEKQKKRYPGESKGFVEEMRFVAMRLHTKDQAKEGEKETKS 1ph/hy1-7 MAYLAPISSSLSIFKNPQLSRFQFSSSSPNPLFLRPRIQILSMTMNKSPSLVVVAATTAAEKQKKRYPGESKGFVEEMRFVAMRLHTKDQAKEGEKETKS

HY1 NIYFAHSAGGRMIGRKVAERILDNKELEFYKWDGELSQLLQNVREKLNKVAEEWTREEKNHCLEETEKSFKYSGEILRLILS 1ph/hy1-7 NIYFAHSAGGRMIGRKVAERILDNKELEFYKWDGELSQLLQNVREKLNKVAEEWTREEKNHCLEETEKSFKYSGEILRLILS



D

Iph is novel allele of HEME OXYGENASE 1



Terry et al., Biochem Soc Trans, 2002



Light controls spatiotemporal specificity of *CKI1* expression

lph/hy1-7 **Short Day Darkness** Red Far-Red 26 Dobisova et al., Plant Phys, 2017

Outline

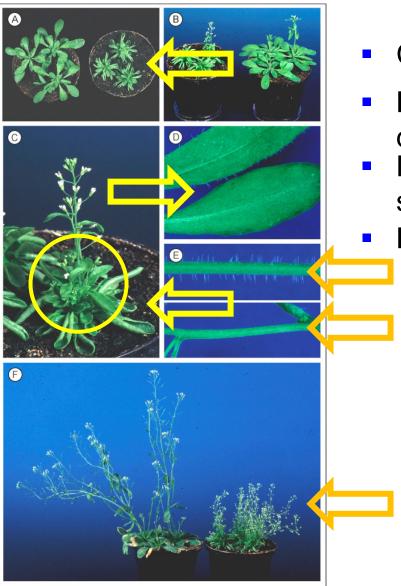
- Forward vs. Reverse Genetics
- Use of Libraries of Insertional Mutants in Forward Genetics
 - Searching in Libraries of Insertional Mutants According to
 - anatomically or morphologically detectable phenotype
 - metabolic profile
 - expression of genes of interest
 - Identification of the Mutated Locus
 - plasmid rescue
 - iPCR



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

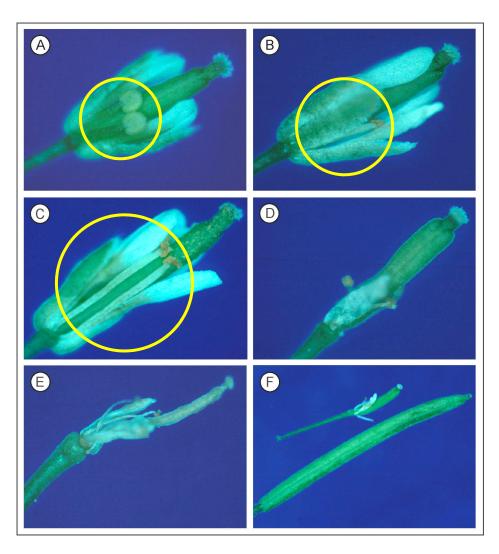


Identification of mutant



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

Identification of mutant



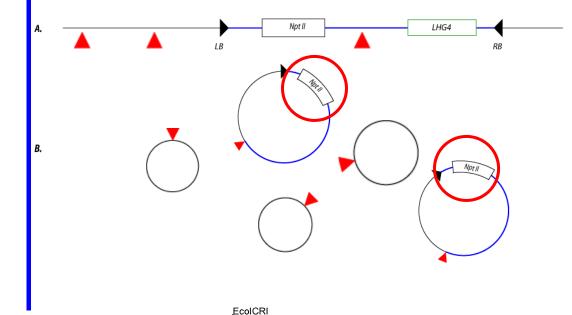
 Male sterility, defects in stamen filament elongation (A,B)

(compare with wild type C)

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



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

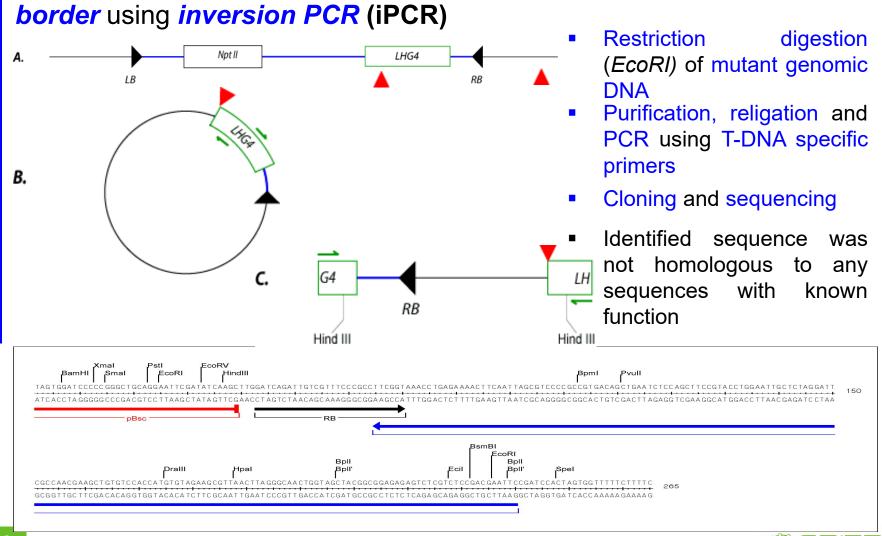


Clal

- 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



2. Identification of region of genomic DNA adjacent to the *right*

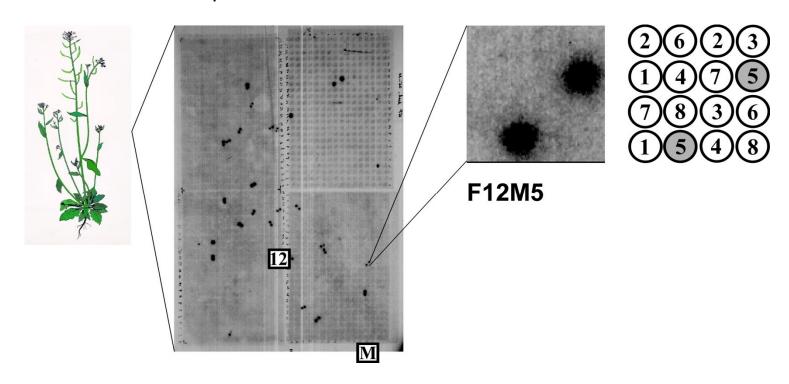


- Identification of chromosomal rearrangements responsible for bushy phenotype of *Arabidopsis*
 - Description of phenotype
 - Identification of T-DNA mutated region
 - Localization of T-DNA insertion site in Arabidopsis genome



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

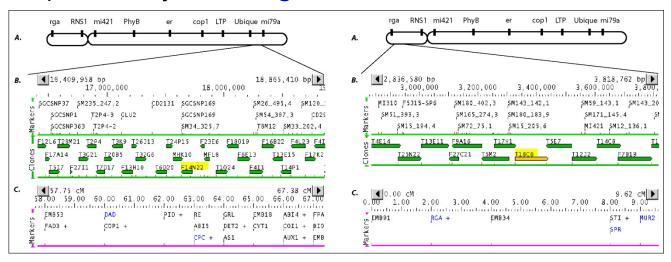
II. Sequences adjacent to the right border of T-DNA

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



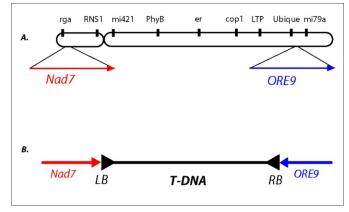
Localization of genomic T-DNA adjacent to both left and right T-DNA borders on chromosome 2

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



There was probably an inversion of almost entire

chromosome 2





Outline

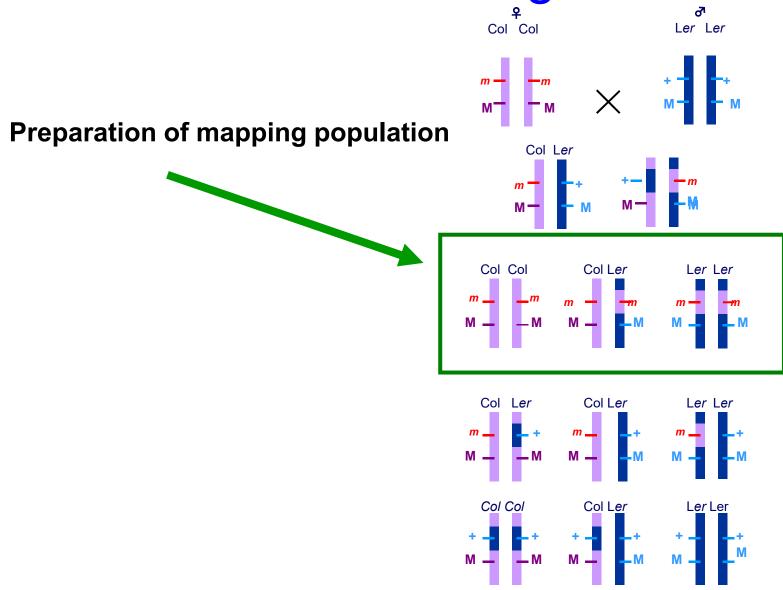
- Forward vs. Reverse Genetics
- Use of Libraries of Insertional Mutants in Forward Genetics
 - Searching in Libraries of Insertional Mutants According to
 - anatomically or morphologically detectable phenotype
 - metabolic profile
 - expression of genes of interest
 - Identification of the Mutated Locus
 - plasmid rescue
 - iPCR
- Use of Libraries of Point Mutants in Forward Genetics
 - Positional Cloning



- 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

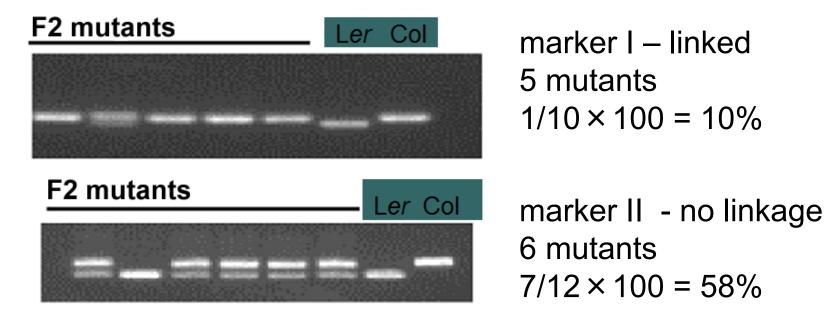


Positional cloning



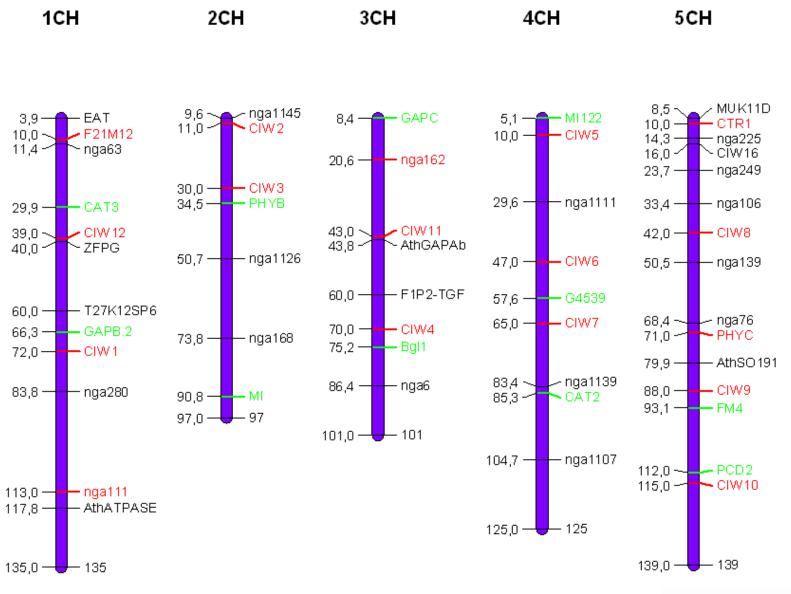


Recombinant analysis – determining the percentage of recombination between mutation and molecular marker r [%] = number of chromosomes of Col / number of all the chromosomes × 100



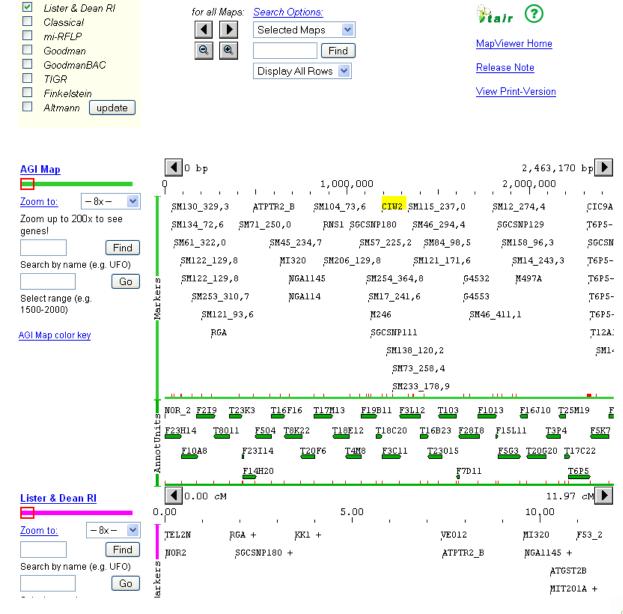
- Analysis of approximately 2000 mutant plants
- Determining the closest (still segregating) marker
- Identification of mutation by sequencing

Map of DNA molecular markers



Markers for fine mapping

AGI Map



Maps for Chromosome 2



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

