CG920 Genomics

Lesson 12

Systems Biology Tools

Model organisms, PCR and PCR Primer Design

Jan Hejátko

Functional Genomics and Proteomics of Plants,
Mendel Centre for Plant Genomics and Proteomics,
Central European Institute of Technology (CEITEC), Masaryk University, Brno
hejatko@sci.muni.cz, www.ceitec.muni.cz

EVROPSKÁ UNIE
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ
Tato prezentace je spolufinancována
Evropským sociálním fondem
a státním rozpočtem České republiky

Literature

- Literature sources for Chapter 12:

  - Wilt, F.H., and Hake, S. (2004). Principles of Developmental Biology. (New York ; London: W. W. Norton)
  - Roscoe B. Jackson Memorial Laboratory., and Green, E.L. (1966). Biology of the laboratory mouse. (New York: Blakiston Division) http://www.informatics.jax.org/greenbook/index.shtml
  - Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics 10, 48.
  - The Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408, 796-815.
  - Gregory, S.G., Sekhon, M., Schein, J., Zhao, S., Osoegawa, K., Scott, C.E., Evans, R.S., Burridge, P.W., Cox, T.V., Fox, C.A., Hutton, R.D., Mullenger, I.R., Phillips, K.J., Smith, J., Stalker, J., Threadgold, G.J., Birney, E., Wylie, K., Chinwalla, A., Wallis, J., Hillier, L., Carter, J., Gaige, T., Jaeger, S., Kremitzki, C., Layman, D., Maas, J., McGrane, R., Mead, K., Walker, R., Jones, S., Smith, M., Asano, J., Bosdet, I., Chan, S., Chittaranjan, S., Chiu, R., Fjell, C., Fuhrmann, D., Girn, N., Gray, C., Guin, R., Hsiao, L., Krzywinski, M., Kutsche, R., Lee, S.S., Mathewson, C., McLeavy, C., Messervier, S., Ness, S., Pandoh, P., Prabhu, A.L., Saeedi, P., Smailus, D., Spence, L., Stott, J., Taylor, S., Terpstra, W., Tsai, M., Vardy, J., Wye, N., Yang, G., Shatsman, S., Ayodeji, B., Geer, K., Tsegaye, G., Shvartsbeyn, A., Gebregeorgis, E., Krol, M., Russell, D., Overton, L., Malek, J.A., Holmes, M., Heaney, M., Shetty, J., Feldblyum, T., Nierman, W.C., Catanese, J.J., Hubbard, T., Waterston, R.H., Rogers, J., de Jong, P.J., Fraser, C.M., Marra, M., McPherson, J.D., and Bentley, D.R. (2002). A physical map of the mouse genome. Nature 418, 743-750.
  - Benitez, M. and Hejatko, J. Dynamics of cell-fate determination and patterning in the vascular bundles of Arabidopsis thaliana (submitted)

EVROPSKÁ UNIE
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ
Tato prezentace je spolufinancována
Evropským sociálním fondem
a státním rozpočtem České republiky

Outline

- Tools of **systems biology**
  - **Gene ontology** analysis
  - **Molecular Regulatory Networks Modeling**

- Model organisms
  - *Mus musculus*
  - *Arabidopsis thaliana*

- Selected **methods of molecular biology**
  - Preparation of transgenic organisms
  - PCR
  - Design and preparation of primers (Dr. Hana Konečná)

EVROPSKÁ UNIE
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ
Tato prezentace je spolufinancována
Evropským sociálním fondem
a státním rozpočtem České republiky

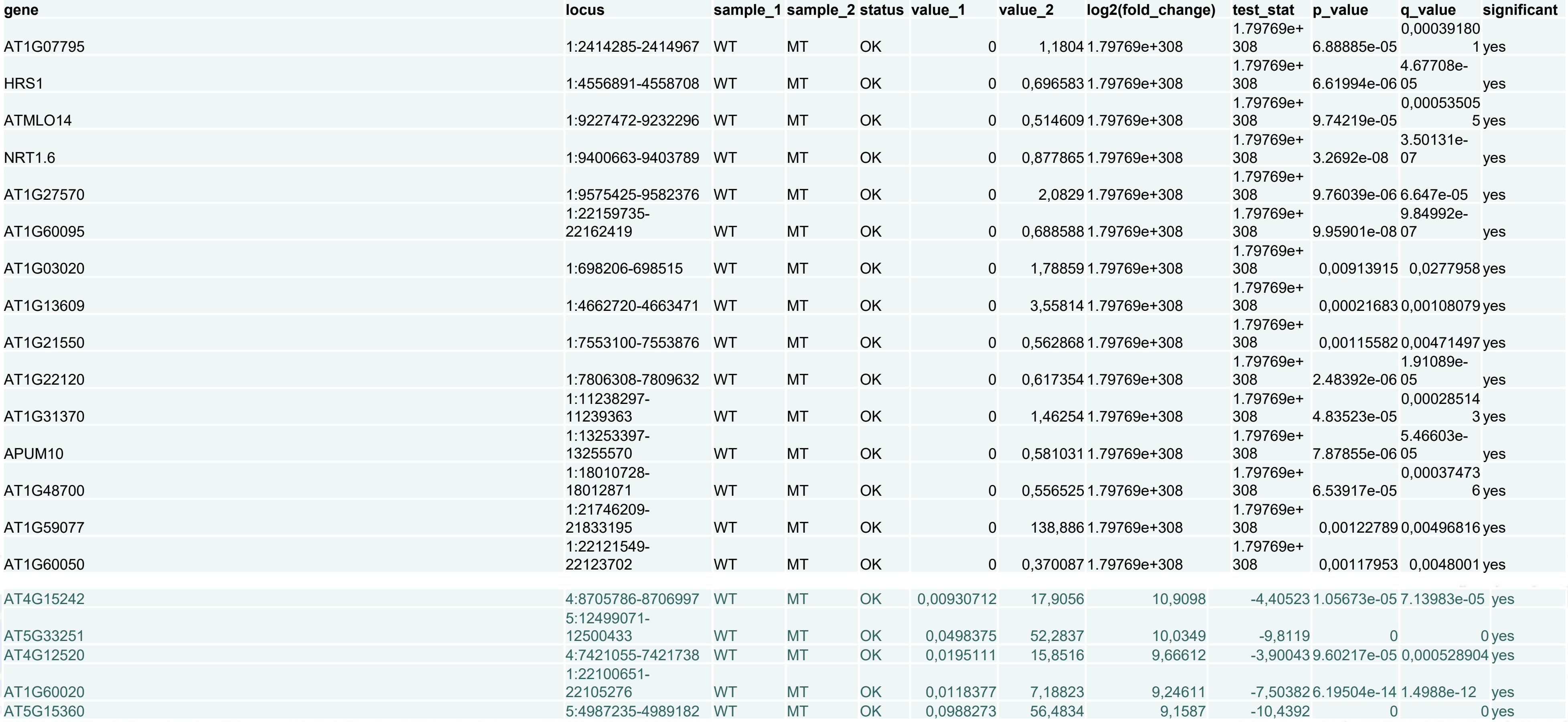
Outline

- Tools of **systems biology**
    - **Gene ontology** analysis

Results of –omics Studies vs Biologically Relevant Conclusions

- ☐ Results of **–omics studies** represent **huge amount of data**, e.g. genes with differential expression. But how to get any **biologically relevant conclusions** out of it?

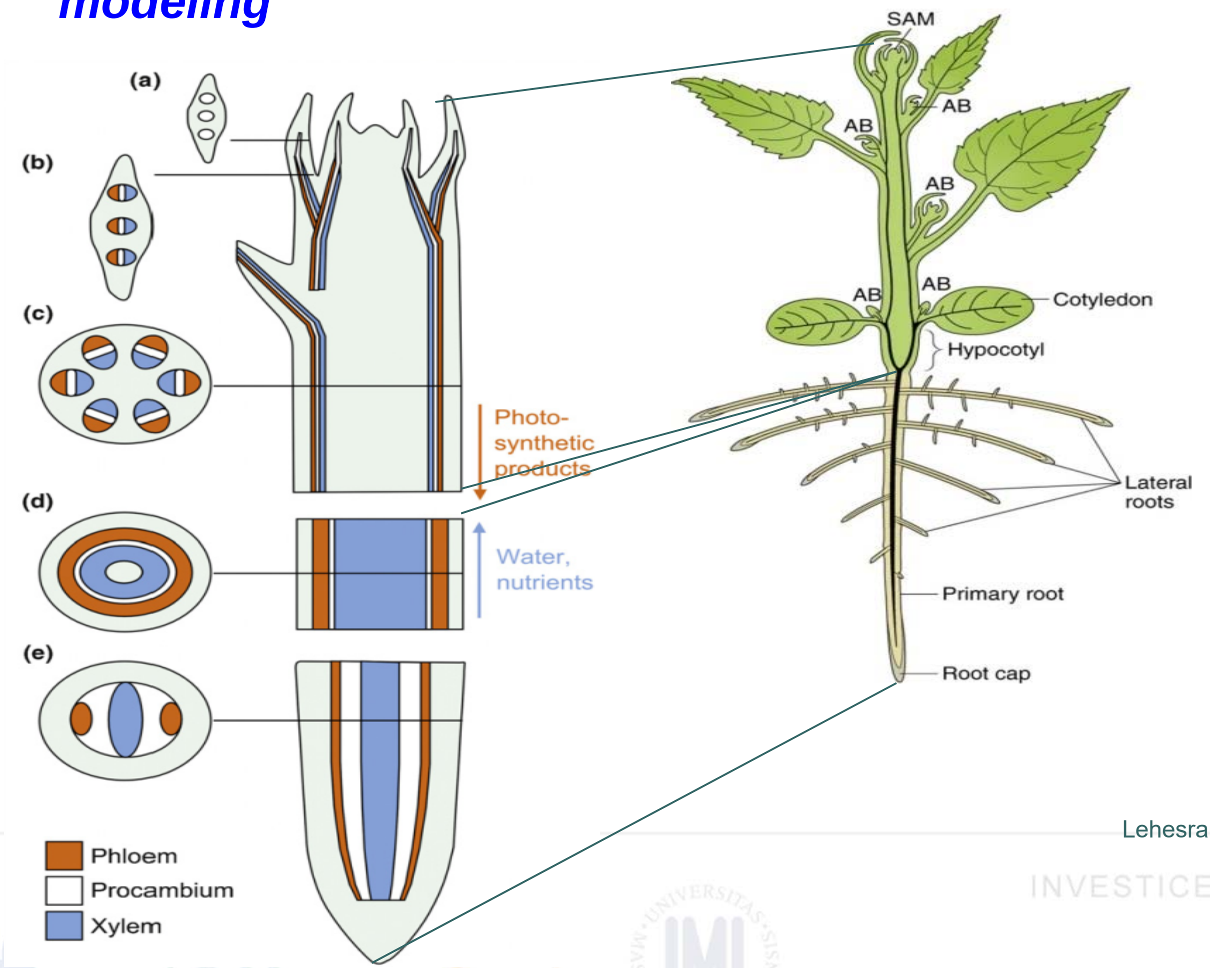
Ddii et al., *unpublished*



gene	locus	sample_1	sample_2	status	value_1	value_2	log2(fold_change)	test_stat	p_value	q_value	significant
AT1G07795	1:2414285-2414967	WT	MT	OK	0	1,1804	1.79769e+308	1.79769e+308	6.88885e-05	0,000391801	yes
HRS1	1:4556891-4558708	WT	MT	OK	0	0,696583	1.79769e+308	1.79769e+308	6.61994e-06	4.67708e-05	yes
ATMLO14	1:9227472-9232296	WT	MT	OK	0	0,514609	1.79769e+308	1.79769e+308	9.74219e-05	0,000535055	yes
NRT1.6	1:9400663-9403789	WT	MT	OK	0	0,877865	1.79769e+308	1.79769e+308	3.2692e-08	3.50131e-07	yes
AT1G27570	1:9575425-9582376	WT	MT	OK	0	2,0829	1.79769e+308	1.79769e+308	9.76039e-06	6.647e-05	yes
AT1G60095	1:22159735-22162419	WT	MT	OK	0	0,688588	1.79769e+308	1.79769e+308	9.95901e-08	9.84992e-07	yes
AT1G03020	1:698206-698515	WT	MT	OK	0	1,78859	1.79769e+308	1.79769e+308	0,00913915	0,0277958	yes
AT1G13609	1:4662720-4663471	WT	MT	OK	0	3,55814	1.79769e+308	1.79769e+308	0,00021683	0,00108079	yes
AT1G21550	1:7553100-7553876	WT	MT	OK	0	0,562868	1.79769e+308	1.79769e+308	0,00115582	0,00471497	yes
AT1G22120	1:7806308-7809632	WT	MT	OK	0	0,617354	1.79769e+308	1.79769e+308	2.48392e-06	1.91089e-05	yes
AT1G31370	1:11238297-11239363	WT	MT	OK	0	1,46254	1.79769e+308	1.79769e+308	4.83523e-05	0,000285143	yes
APUM10	1:13253397-13255570	WT	MT	OK	0	0,581031	1.79769e+308	1.79769e+308	7.87855e-06	5.46603e-05	yes
AT1G48700	1:18010728-18012871	WT	MT	OK	0	0,556525	1.79769e+308	1.79769e+308	6.53917e-05	0,000374736	yes
AT1G59077	1:21746209-21833195	WT	MT	OK	0	138,886	1.79769e+308	1.79769e+308	0,00122789	0,00496816	yes
AT1G60050	1:22121549-22123702	WT	MT	OK	0	0,370087	1.79769e+308	1.79769e+308	0,00117953	0,0048001	yes
AT4G15242	4:8705786-8706997	WT	MT	OK	0,00930712	17,9056	10,9098	-4,40523	1.05673e-05	7.13983e-05	yes
AT5G33251	5:12499071-12500433	WT	MT	OK	0,0498375	52,2837	10,0349	-9,8119	0	0	yes
AT4G12520	4:7421055-7421738	WT	MT	OK	0,0195111	15,8516	9,66612	-3,90043	9.60217e-05	0,000528904	yes
AT1G60020	1:22100651-22105276	WT	MT	OK	0,0118377	7,18823	9,24611	-7,50382	6.19504e-14	1.4988e-12	yes
AT5G15360	5:4987235-4989182	WT	MT	OK	0,0988273	56,4834	9,1587	-10,4392	0	0	yes

Molecular Regulatory Networks Modeling

- ***Vascular tissue*** as a developmental model for ***GO analysis*** and ***MRN modeling***

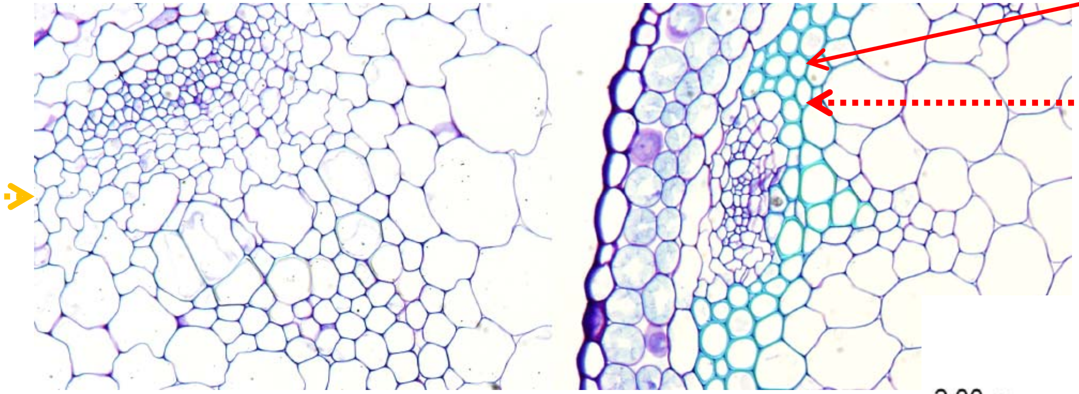


Lehesranta etal., *Trends in Plant Sci* (2010)

Hormonal Control Over Vascular Tissue Development

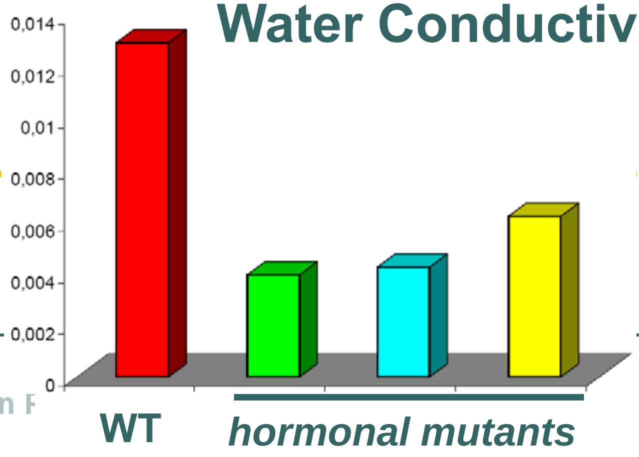
- Plant **Hormones Regulate Lignin Deposition** in Plant Cell Walls and **Xylem Water Conductivity**

WT *hormonal mutant*

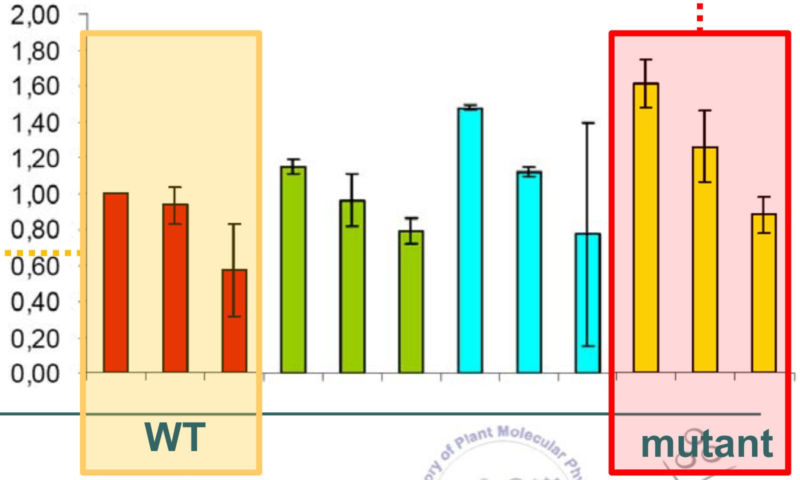


lignified cell walls

Water Conductivity

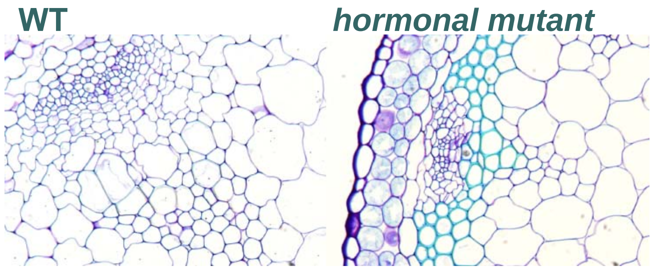


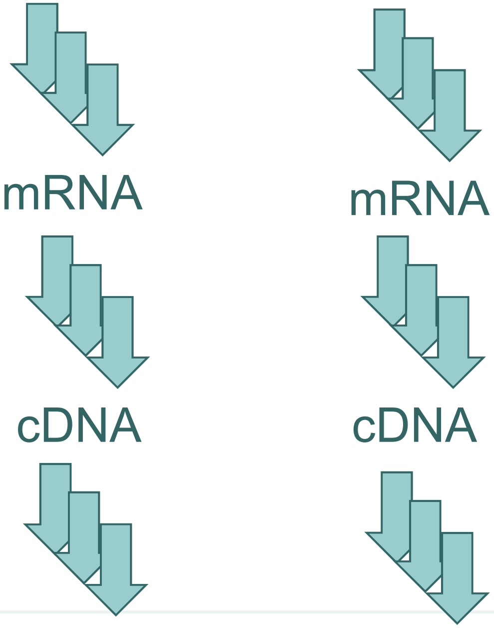
Acid-insoluble lignins



Hormonal Control Over Vascular Tissue Development

- ☐ ***Transcriptional profiling*** via ***RNA sequencing***





mRNA mRNA

cDNA cDNA

Library Preparation
~2 h [15 min hands-on (Nextera)]
< 6 h [< 3 h hands-on (TruSeq)]

Cluster Generation
~5 h (<10 min hands-on)

Sequencing by Synthesis
~1.5 to 11 days

CASAVA
2 days (30 min hands-on)

Sequencing by Illumina and
number of transcripts determination

Results of –omics Studies vs Biologically Relevant Conclusions

- Transcriptional profiling yielded more then **7K differentially regulated genes**…

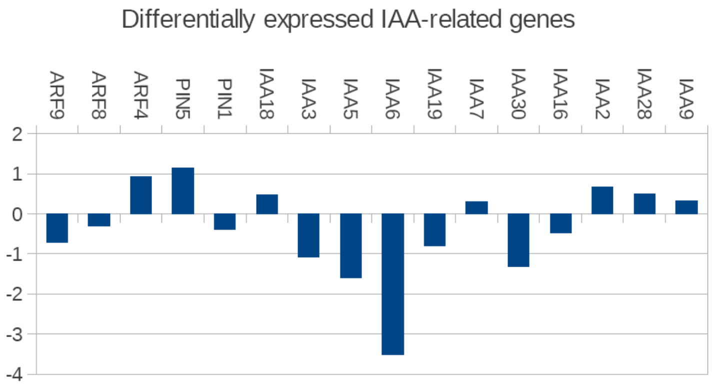
Ddii et al., *unpublished*

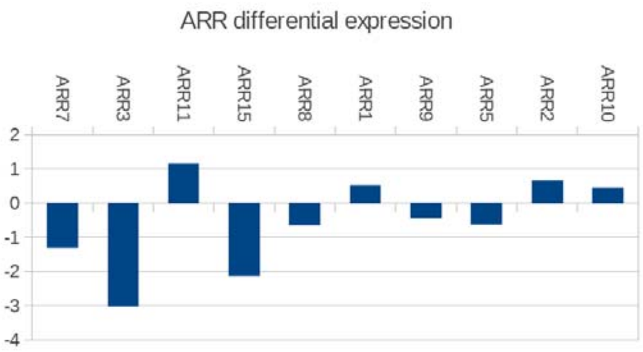
gene	locus	sample_1	sample_2	status	value_1	value_2	log2(fold_change)	test_stat	p_value	q_value	significant
AT1G07795	1:2414285-2414967	WT	MT	OK	0	1,1804	1.79769e+308	1.79769e+308	6.88885e-05	0,000391801	yes
HRS1	1:4556891-4558708	WT	MT	OK	0	0,696583	1.79769e+308	1.79769e+308	6.61994e-06	4.67708e-05	yes
ATMLO14	1:9227472-9232296	WT	MT	OK	0	0,514609	1.79769e+308	1.79769e+308	9.74219e-05	0,000535055	yes
NRT1.6	1:9400663-9403789	WT	MT	OK	0	0,877865	1.79769e+308	1.79769e+308	3.2692e-08	3.50131e-07	yes
AT1G27570	1:9575425-9582376	WT	MT	OK	0	2,0829	1.79769e+308	1.79769e+308	9.76039e-06	6.647e-05	yes
AT1G60095	1:22159735-22162419	WT	MT	OK	0	0,688588	1.79769e+308	1.79769e+308	9.95901e-08	9.84992e-07	yes
AT1G03020	1:698206-698515	WT	MT	OK	0	1,78859	1.79769e+308	1.79769e+308	0,00913915	0,0277958	yes
AT1G13609	1:4662720-4663471	WT	MT	OK	0	3,55814	1.79769e+308	1.79769e+308	0,00021683	0,00108079	yes
AT1G21550	1:7553100-7553876	WT	MT	OK	0	0,562868	1.79769e+308	1.79769e+308	0,00115582	0,00471497	yes
AT1G22120	1:7806308-7809632	WT	MT	OK	0	0,617354	1.79769e+308	1.79769e+308	2.48392e-06	1.91089e-05	yes
AT1G31370	1:11238297-11239363	WT	MT	OK	0	1,46254	1.79769e+308	1.79769e+308	4.83523e-05	0,000285143	yes
APUM10	1:13253397-13255570	WT	MT	OK	0	0,581031	1.79769e+308	1.79769e+308	7.87855e-06	5.46603e-05	yes
AT1G48700	1:18010728-18012871	WT	MT	OK	0	0,556525	1.79769e+308	1.79769e+308	6.53917e-05	0,000374736	yes
AT1G59077	1:21746209-21833195	WT	MT	OK	0	138,886	1.79769e+308	1.79769e+308	0,00122789	0,00496816	yes
AT1G60050	1:22121549-22123702	WT	MT	OK	0	0,370087	1.79769e+308	1.79769e+308	0,00117953	0,0048001	yes

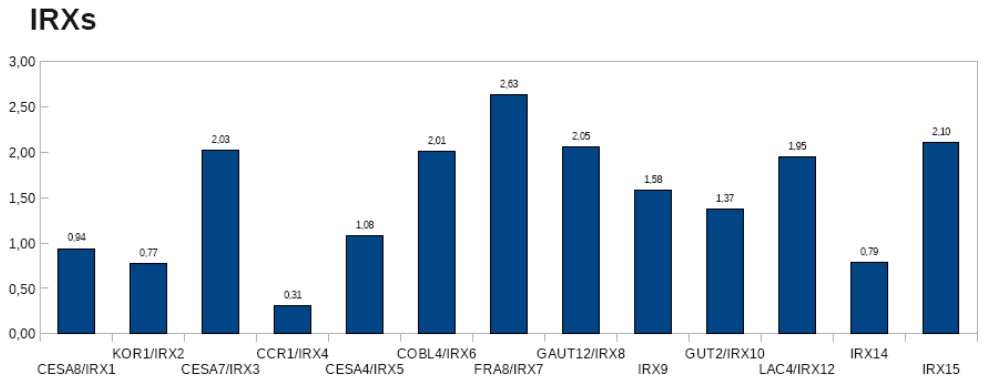
AT4G15242	4:8705786-8706997	WT	MT	OK	0,00930712	17,9056	10,9098	-4,40523	1.05673e-05	7.13983e-05	yes
AT5G33251	5:12499071-12500433	WT	MT	OK	0,0498375	52,2837	10,0349	-9,8119	0	0	yes
AT4G12520	4:7421055-7421738	WT	MT	OK	0,0195111	15,8516	9,66612	-3,90043	9.60217e-05	0,000528904	yes
AT1G60020	1:22100651-22105276	WT	MT	OK	0,0118377	7,18823	9,24611	-7,50382	6.19504e-14	1.4988e-12	yes
AT5G15360	5:4987235-4989182	WT	MT	OK	0,0988273	56,4834	9,1587	-10,4392	0	0	yes

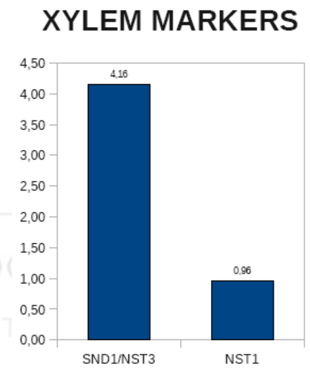
Gene Ontology Analysis

- One of the possible approaches is to study ***gene ontology***, i.e. previously demonstrated **association** of genes to **biological processes**


Differentially expressed IAA-related genes


ARR differential expression


IRXs

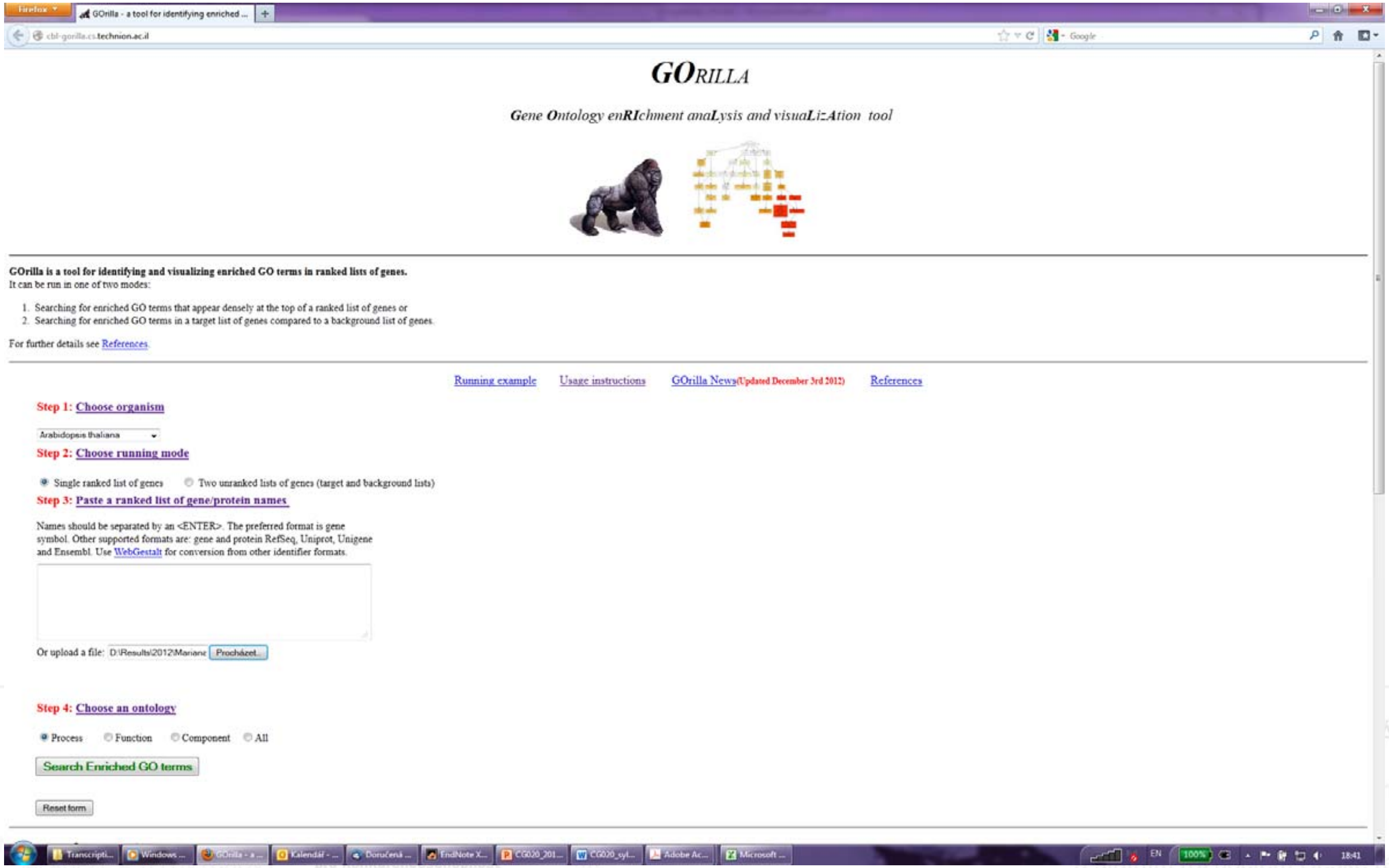

XYLEM MARKERS

Ddii et al., *unpublished*

Gene Ontology Analysis

- Several tools allow **statistical evaluation** of **enrichment** for genes **associated with specific processes**

Eden et al., *BMC Biinformatics* (2009)



GOrilla

Gene Ontology enRIchment anaLysis and visuaLizAtion tool

GOrilla is a tool for identifying and visualizing enriched GO terms in ranked lists of genes.
It can be run in one of two modes:

1. Searching for enriched GO terms that appear densely at the top of a ranked list of genes or
2. Searching for enriched GO terms in a target list of genes compared to a background list of genes.

For further details see References.

Running example Usage instructions GOrilla News(Updated December 3rd 2012) References

Step 1: Choose organism

Arabidopsis thaliana

Step 2: Choose running mode

● Single ranked list of genes ○ Two unranked lists of genes (target and background lists)

Step 3: Paste a ranked list of gene/protein names

Names should be separated by an <ENTER>. The preferred format is gene symbol. Other supported formats are: gene and protein RefSeq, Uniprot, Unigene and Ensembl. Use WebGestalt for conversion from other identifier formats.

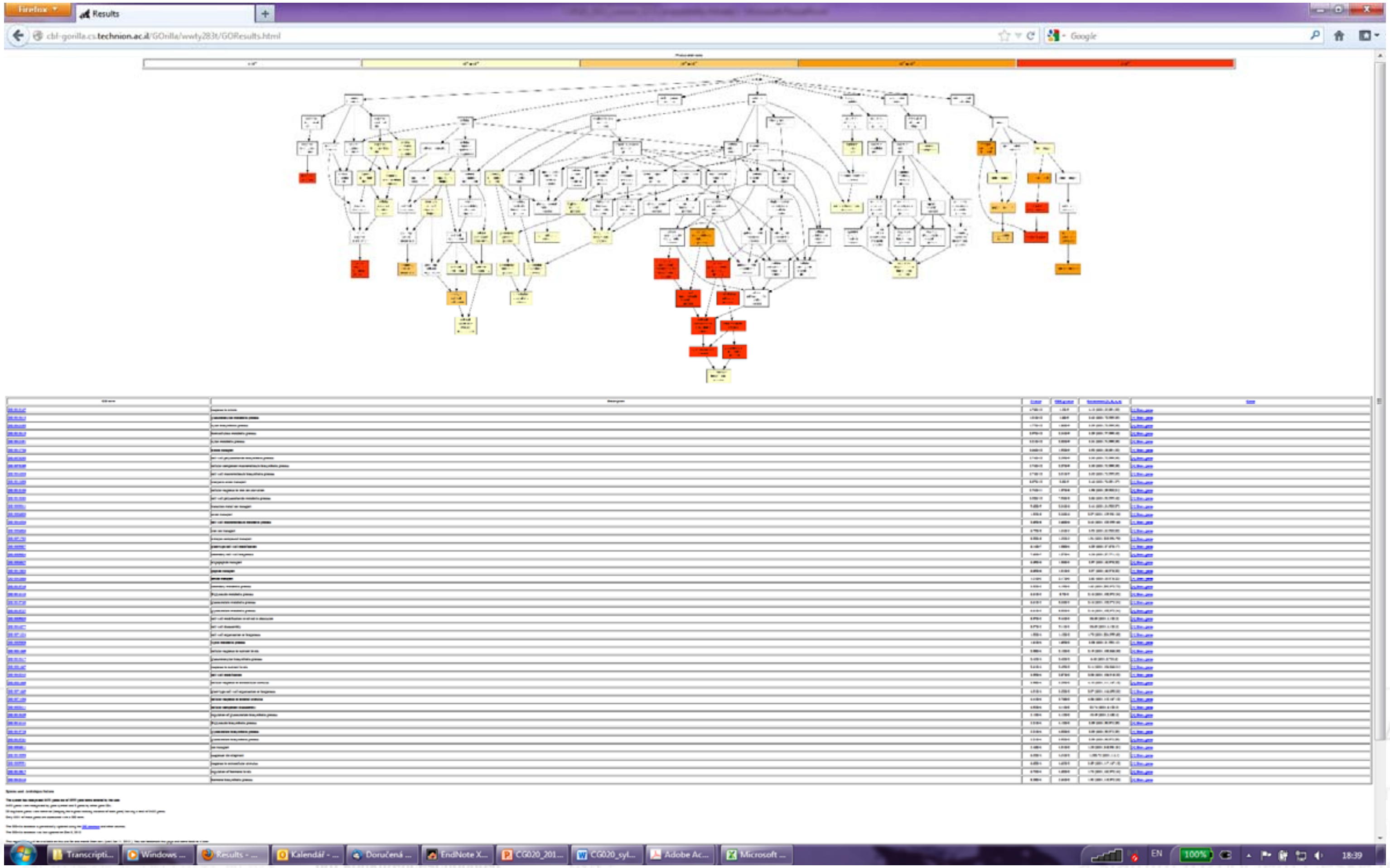
Or upload a file: D:\Results\2012\Mariane Procházet...

Step 4: Choose an ontology

● Process ○ Function ○ Component ○ All

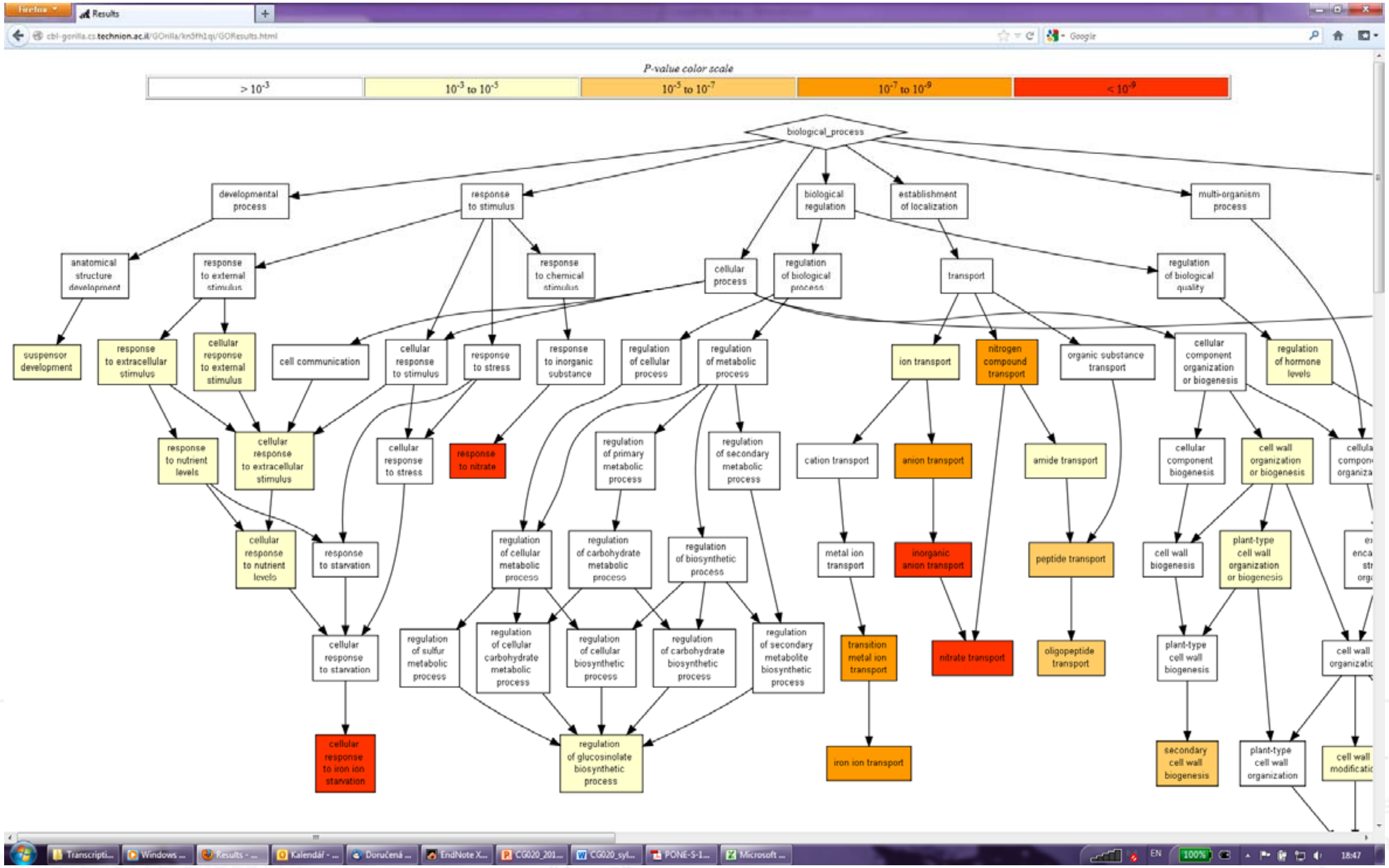
Gene Ontology Analysis

☐ Several tools allow **statistical evaluation** of **enrichment** for genes **associated with specific processes**



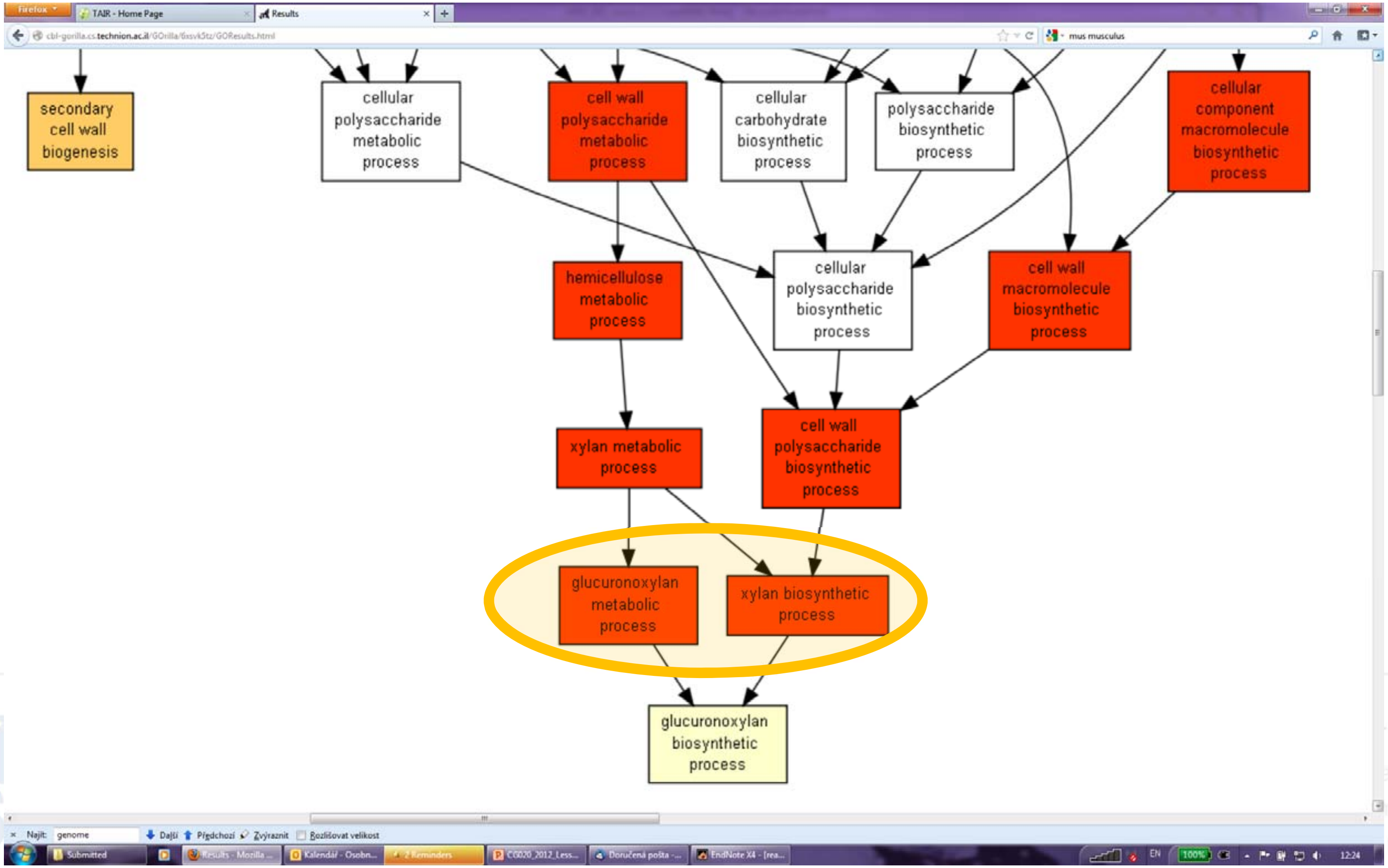
Gene Ontology Analysis

- Several tools allow **statistical evaluation** of **enrichment** for genes **associated with specific processes**



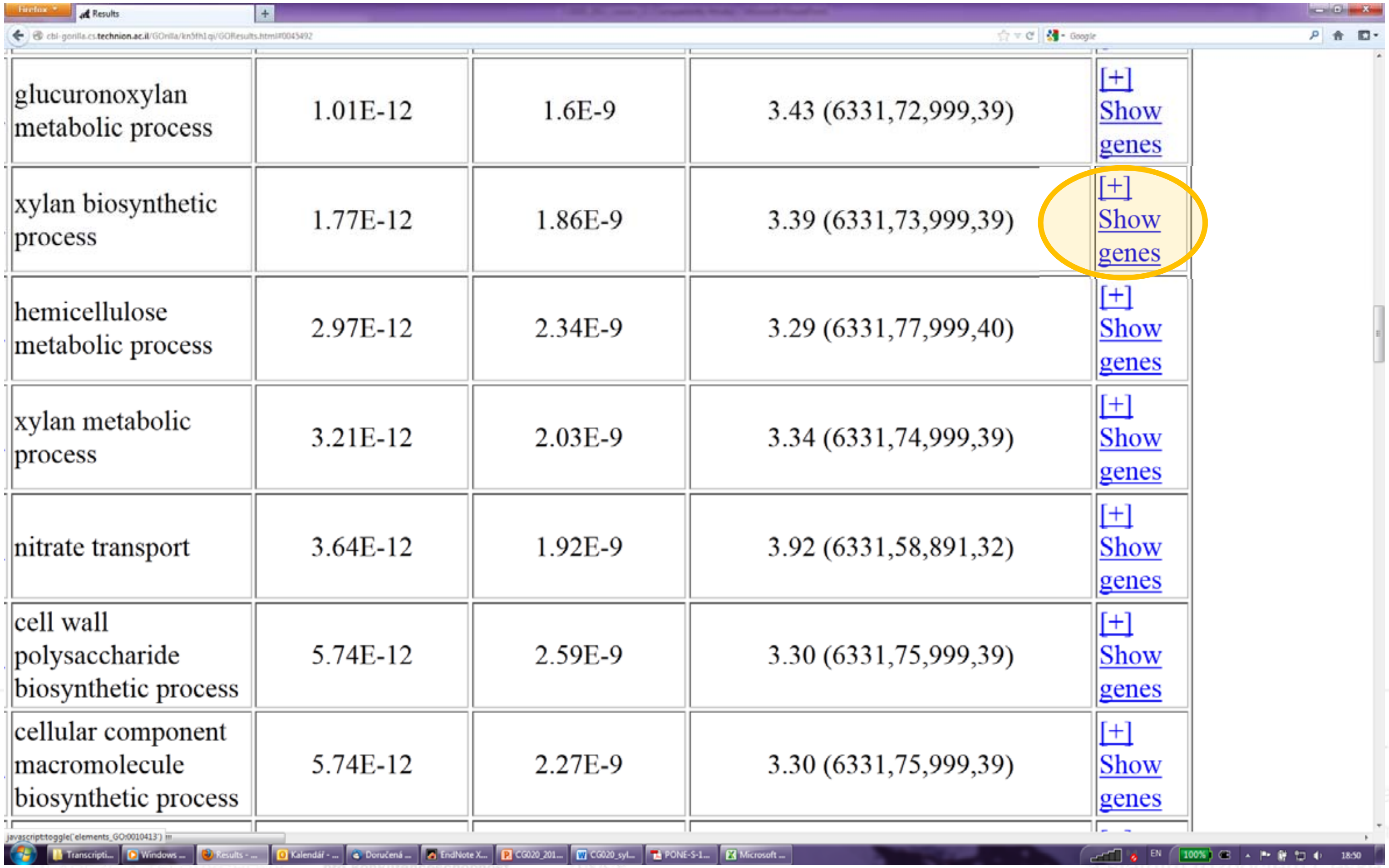
Gene Ontology Analysis

- □ Several tools allow **statistical evaluation** of **enrichment** for genes **associated with specific processes**



Gene Ontology Analysis

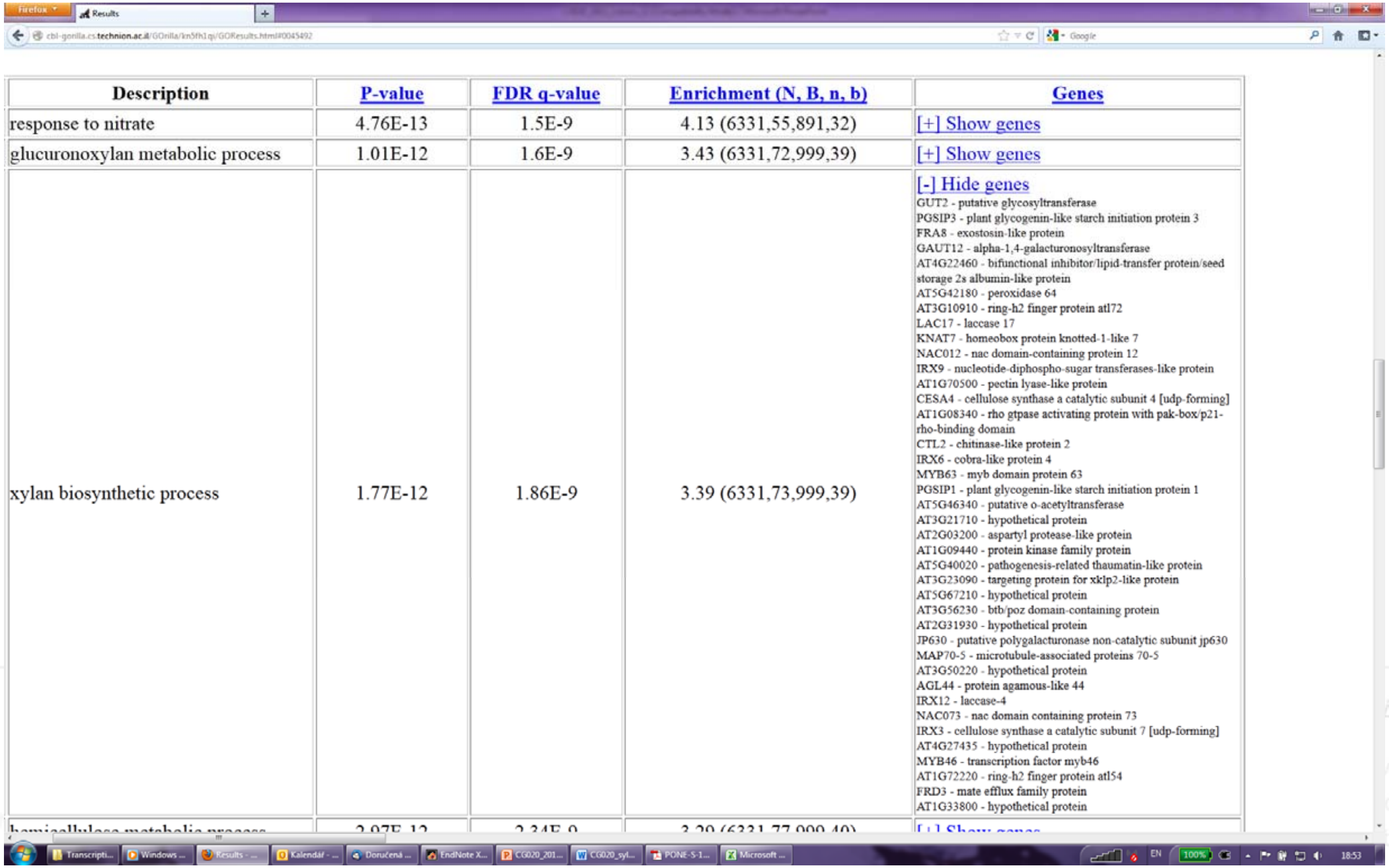
- Several tools allow **statistical evaluation** of **enrichment** for genes **associated with specific processes**



glucuronoxylan metabolic process	1.01E-12	1.6E-9	3.43 (6331,72,999,39)	[+] Show genes
xylan biosynthetic process	1.77E-12	1.86E-9	3.39 (6331,73,999,39)	[+] Show genes
hemicellulose metabolic process	2.97E-12	2.34E-9	3.29 (6331,77,999,40)	[+] Show genes
xylan metabolic process	3.21E-12	2.03E-9	3.34 (6331,74,999,39)	[+] Show genes
nitrate transport	3.64E-12	1.92E-9	3.92 (6331,58,891,32)	[+] Show genes
cell wall polysaccharide biosynthetic process	5.74E-12	2.59E-9	3.30 (6331,75,999,39)	[+] Show genes
cellular component macromolecule biosynthetic process	5.74E-12	2.27E-9	3.30 (6331,75,999,39)	[+] Show genes

Gene Ontology Analysis

- ☐ Several tools allow **statistical evaluation** of **enrichment** for genes **associated with specific processes**



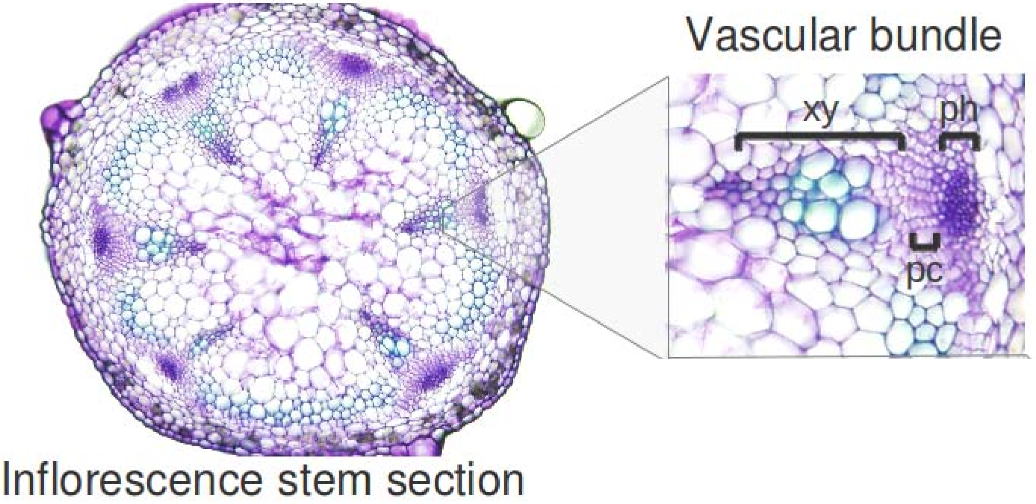
Description	P-value	FDR q-value	Enrichment (N, B, n, b)	Genes
response to nitrate	4.76E-13	1.5E-9	4.13 (6331,55,891,32)	[+] Show genes
glucuronoxylan metabolic process	1.01E-12	1.6E-9	3.43 (6331,72,999,39)	[+] Show genes
xylan biosynthetic process	1.77E-12	1.86E-9	3.39 (6331,73,999,39)	[-] Hide genes GUT2 - putative glycosyltransferase PGSIP3 - plant glycogenin-like starch initiation protein 3 FRA8 - exostosin-like protein GAUT12 - alpha-1,4-galacturonosyltransferase AT4G22460 - bifunctional inhibitor/lipid-transfer protein/seed storage 2s albumin-like protein AT5G42180 - peroxidase 64 AT3G10910 - ring-h2 finger protein atl72 LAC17 - laccase 17 KNAT7 - homeobox protein knotted-1-like 7 NAC012 - nac domain-containing protein 12 IRX9 - nucleotide-diphospho-sugar transferases-like protein AT1G70500 - pectin lyase-like protein CESA4 - cellulose synthase a catalytic subunit 4 [udp-forming] AT1G08340 - rho gtpase activating protein with pak-box/p21-rho-binding domain CTL2 - chitinase-like protein 2 IRX6 - cobra-like protein 4 MYB63 - myb domain protein 63 PGSIP1 - plant glycogenin-like starch initiation protein 1 AT5G46340 - putative o-acetyltransferase AT3G21710 - hypothetical protein AT2G03200 - aspartyl protease-like protein AT1G09440 - protein kinase family protein AT5G40020 - pathogenesis-related thaumatin-like protein AT3G23090 - targeting protein for xklp2-like protein AT5G67210 - hypothetical protein AT3G56230 - btb/poz domain-containing protein AT2G31930 - hypothetical protein JP630 - putative polygalacturonase non-catalytic subunit jp630 MAP70-5 - microtubule-associated proteins 70-5 AT3G50220 - hypothetical protein AGL44 - protein agamous-like 44 IRX12 - laccase-4 NAC073 - nac domain containing protein 73 IRX3 - cellulose synthase a catalytic subunit 7 [udp-forming] AT4G27435 - hypothetical protein MYB46 - transcription factor myb46 AT1G72220 - ring-h2 finger protein atl54 FRD3 - mate efflux family protein AT1G33800 - hypothetical protein

Outline

- Tools of **systems biology**
    - **Gene ontology** analysis
    - **Molecular Regulatory Networks Modeling**

Molecular Regulatory Networks Modeling

- ***Vascular tissue*** as a developmental model for ***MRN modeling***



Benitez and Hejatko, *submitted*

Molecular Regulatory Networks Modeling

- **Literature search** for published data and **creating small database**

Interaction	**Evidence**	**References**
A-ARRs –| CK signaling	Double and higher order type-A ARR mutants show increased sensitivity to CK. Spatial patterns of A-type ARR gene expression and CK response are consistent with partially redundant function of these genes in CK signaling. A-type ARRs decreases B-type ARR6-LUC. Note: In certain contexts, however, some A-ARRs appear to have effects antagonistic to other A-ARRs.	[27] [27] [13] [27]
AHP6 –| AHP	ahp6 partially recovers the mutant phenotype of the CK receptor WOL. Using an in vitro phosphotransfer system, it was shown that, unlike the AHPs, native AHP6 was unable to accept a phosphoryl group. Nevertheless, AHP6 is able to inhibit phosphotransfer from other AHPs to ARRs.	[9] [9]

Benitez and Hejatko, *submitted*

Molecular Regulatory Networks Modeling

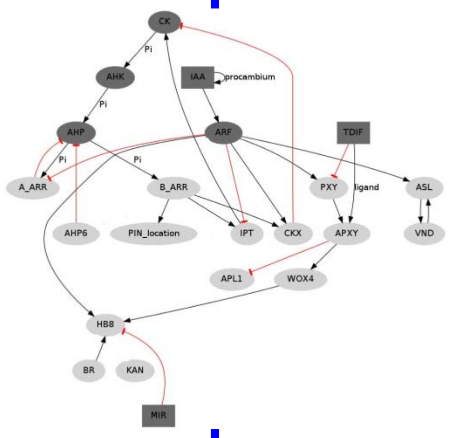
- Formulating ***logical rules*** defining the ***model dynamics***

Network node	Dynamical rule
CK	2 If ipt=1 and ckx=0 1 If ipt=1 and ckx=1 0 else
CKX	1 If barr>0 or arf=2 0 else
AHKs	ahk=ck
AHPs	2 If ahk=2 and ahp6=0 and aarr=0 1 If ahk=2 and (ahp6+aarr<2) 1 If ahk=1 and ahp6<1 0 else
B-Type ARRs	1 If ahp>0 0 else
A-Type ARRs	1 If arf<2 and ahp>0 0 else

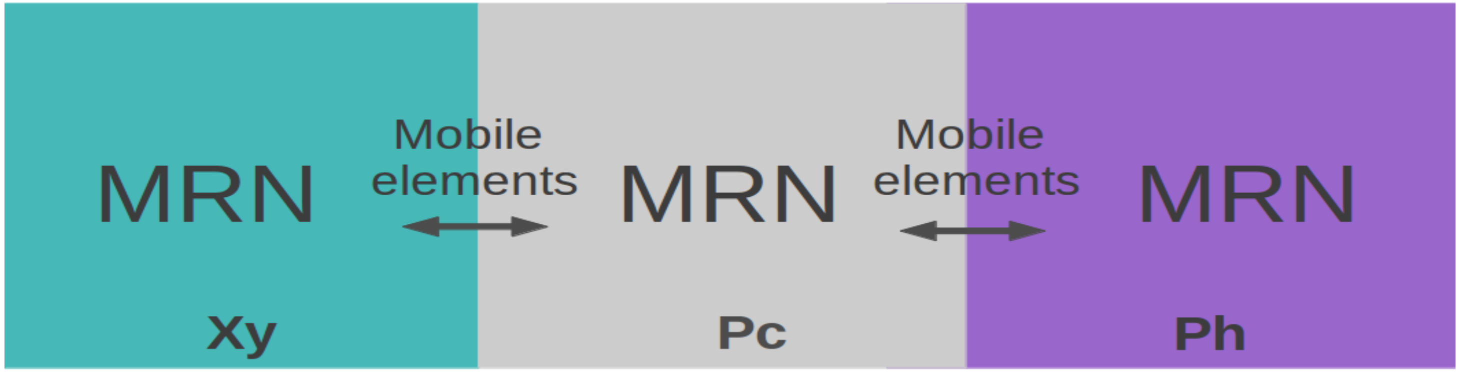
Benitez and Hejatko, *submitted*

Molecular Regulatory Networks Modeling

- ☐ Specifying ***mobile elements*** and their model behaviour



A



MRN

Mobile elements

MRN

Mobile elements

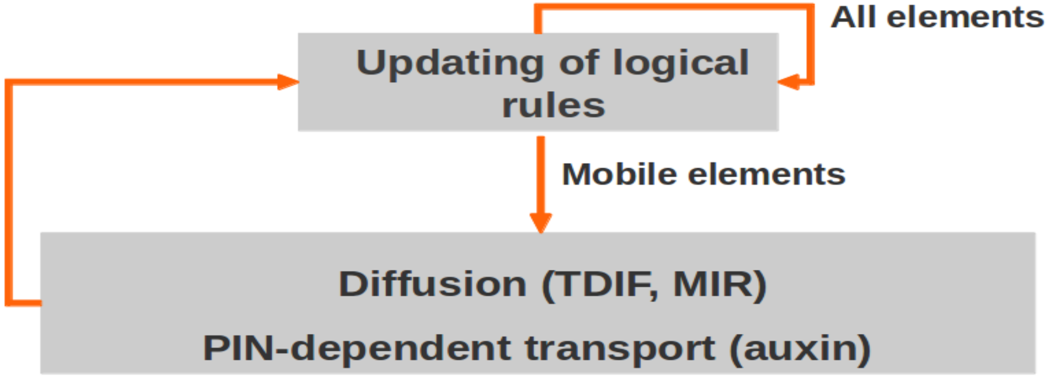
MRN

Xy

Pc

Ph

B



All elements

Updating of logical rules

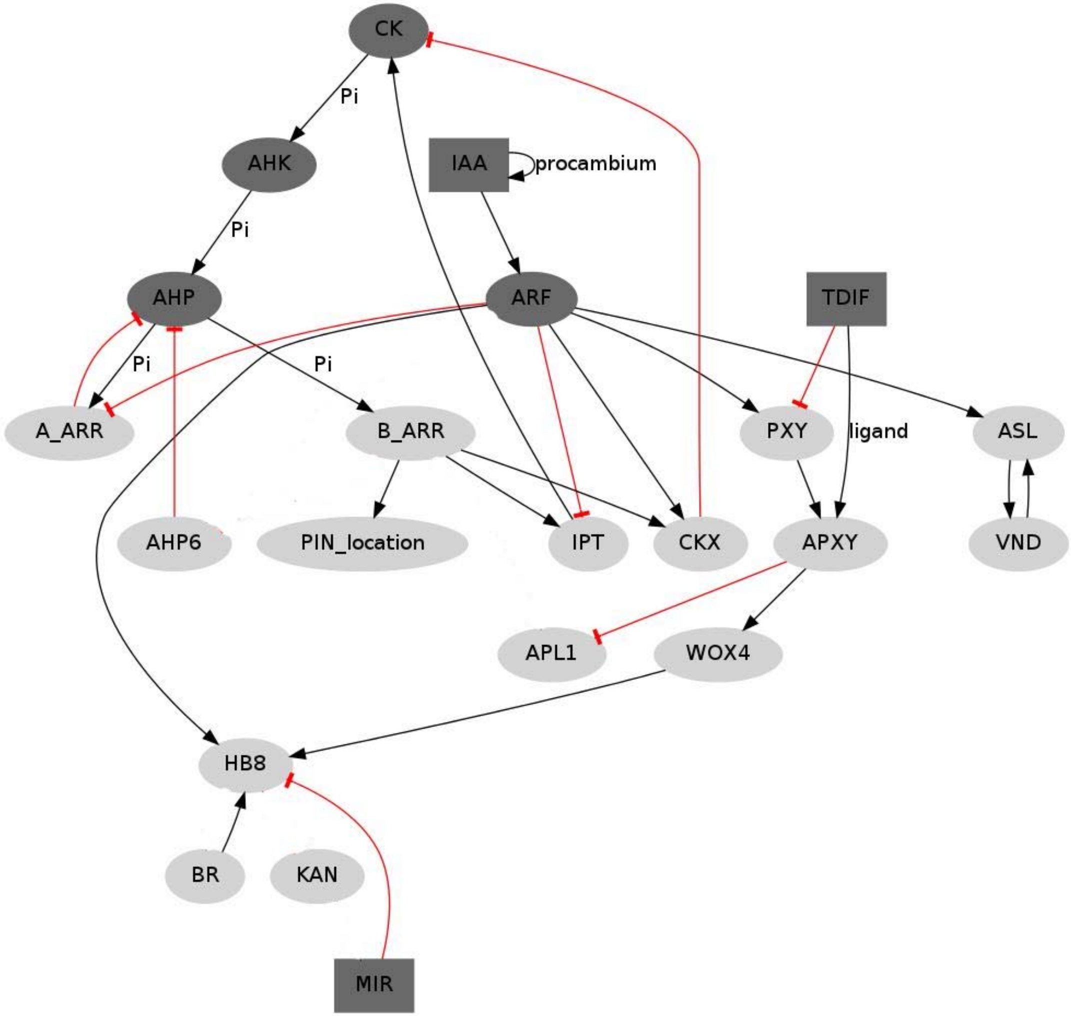
Mobile elements

Diffusion (TDIF, MIR)

PIN-dependent transport (auxin)

Molecular Regulatory Networks Modeling

- Preparing the ***first version*** of the model and its ***testing***



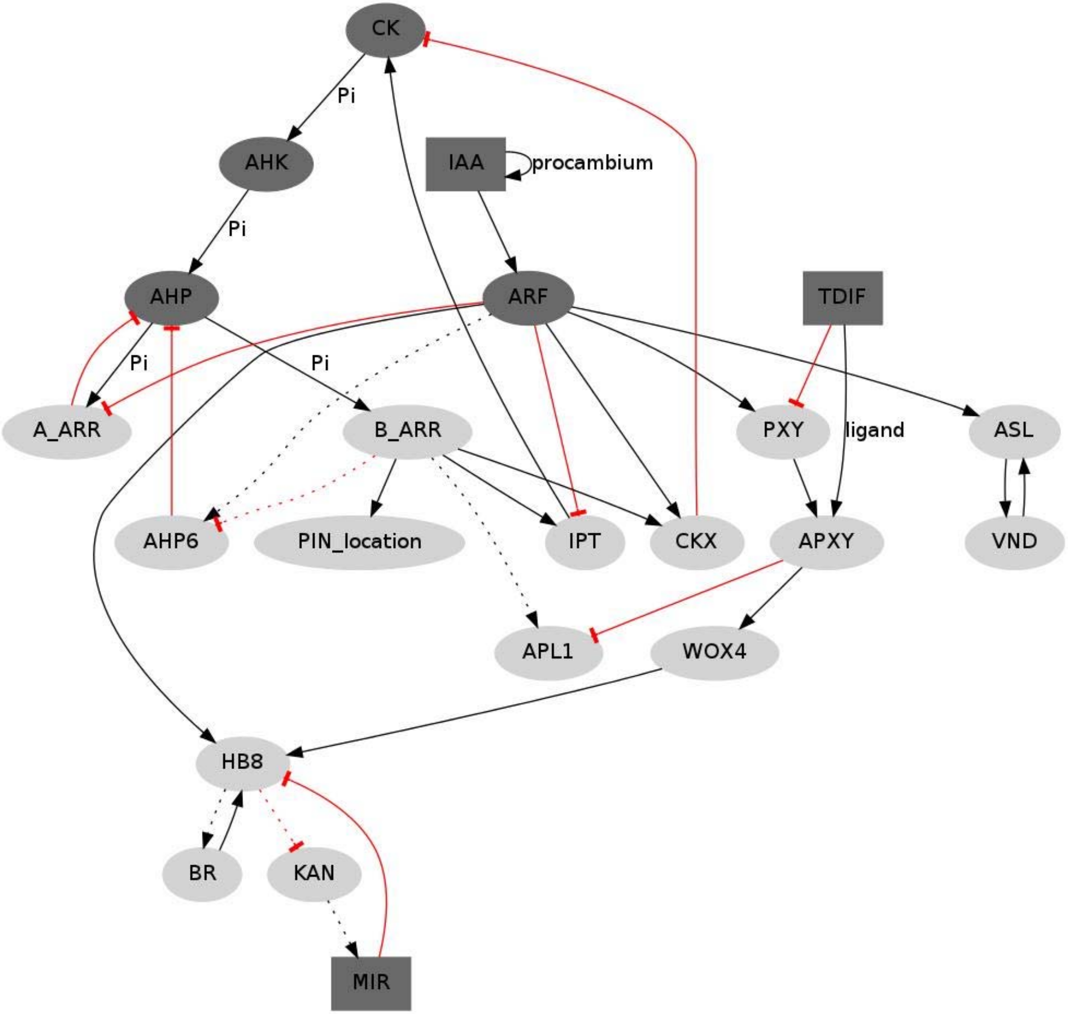
Molecular Regulatory Networks Modeling

- Specifying of missing interactions via ***informed predictions***

Interaction	**Evidence**	**References**
CK → PIN7 radial localization	Predicted interaction (could be direct or indirect) Informed by the following data:	
	During the specification of root vascular cells in Arabidopsis thaliana, CK regulates the radial localization of PIN7.	[18]
	Expression of PIN7:GFP and PIN7::GUS is upregulated by CK with no significant influence of ethylene.	[18,20]
	In the root, CK signaling is required for the CK regulation of PIN1, PIN3, and PIN7. Their expression is altered in wol, cre1, ahk3 and ahp6 mutants.	[19]
CK→ APL	Predicted interaction (could be direct or indirect)	
	Consistent with the fact that APL overexpression prevents or delays xylem cell differentiation, as does CKs.	[21]
	Partially supported by microarray data and phloem-specific expression patterns of CK response factors.	(TAIR, ExpressionSet:1005823559, [22])

Molecular Regulatory Networks Modeling

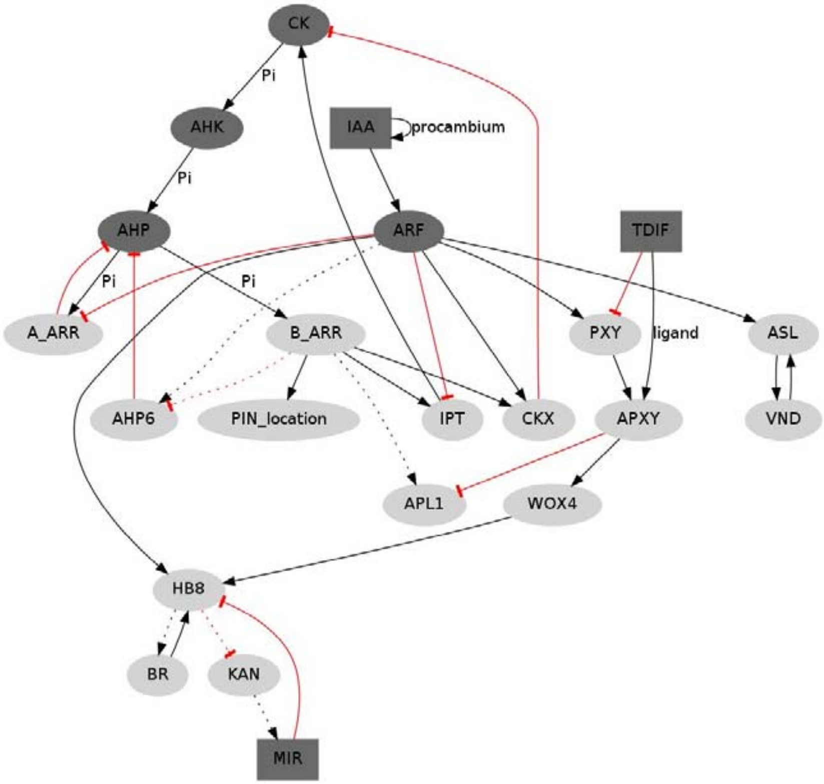
- ☐ Preparing the ***next version*** of the model and its ***testing***

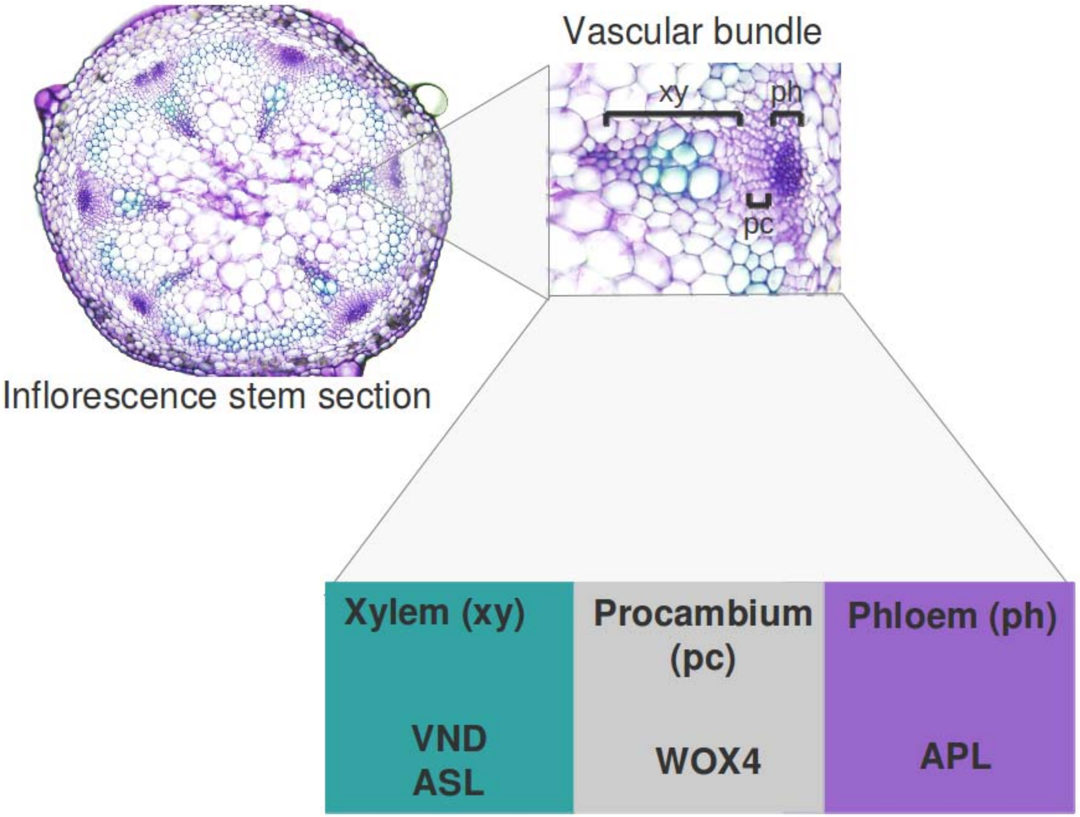


Benitez and Hejatko, *PlosONE, 2013*

Molecular Regulatory Networks Modeling

- ***Good model*** should be able to ***simulate reality***

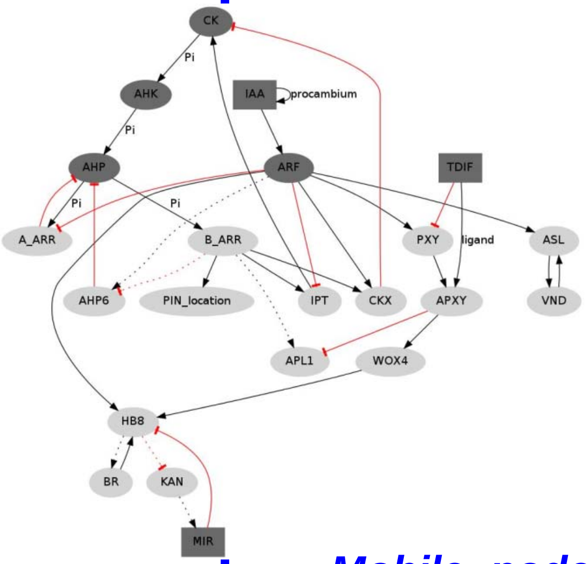




Benitez and Hejatko, *PlosONE, 2013*

Molecular Regulatory Networks Modeling

- Formulating ***equations*** describing the ***relationships*** in the model



logical rule function

state in the time *t*

Static **nodes:** $g_n(t+1)=F_n(g_{n1}(t),g_{n2}(t),\ldots,g_{nk}(t))$

state in the time *t+1*

Amount if TDIF or MIR165 in cell *i*

Mobile nodes**:** $g_{(t+1)T\,[i]}= H(g_{(t)\,[i]}+ D\,(g_{(t)\,[i+1]}+g_{(t)\,[i-1]} - N(g_{(t)\,[i]}))-b)$

state in the time *t+1*

constant corresponding to a degradation term

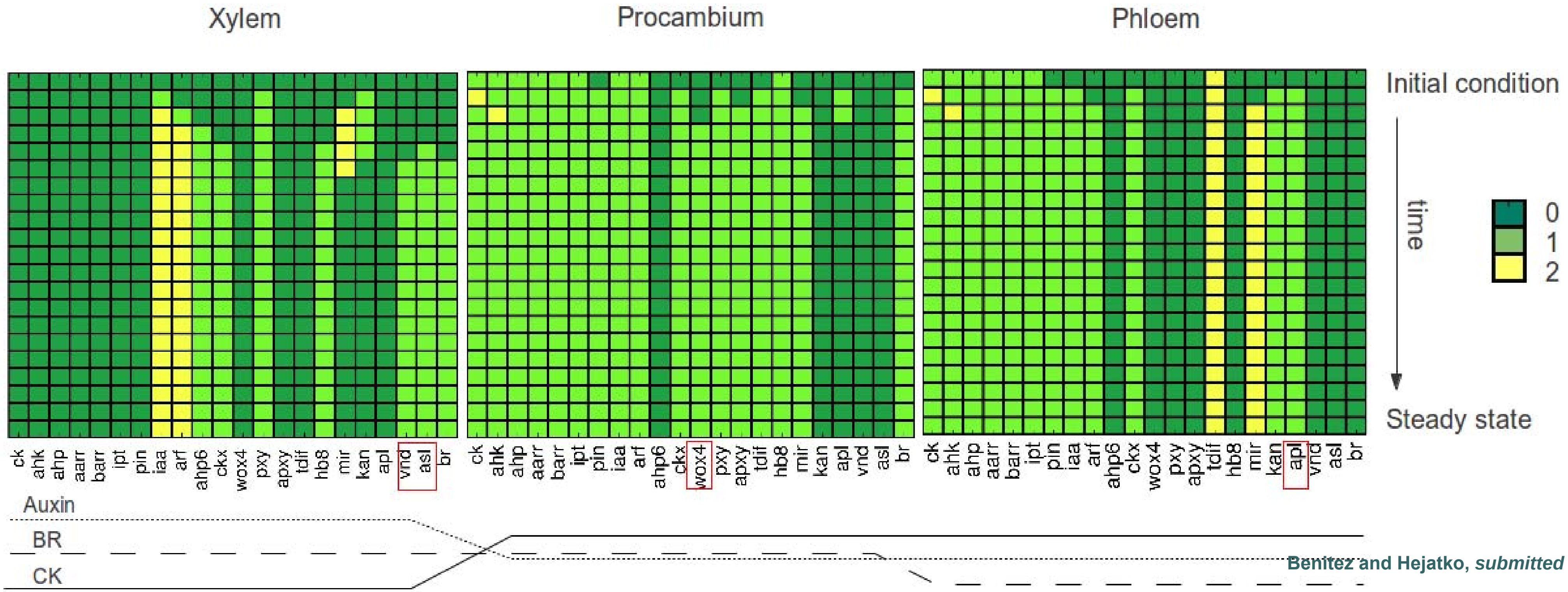
proportion of movable element

Molecular Regulatory Networks Modeling

- ***Good model*** should be able to ***simulate reality***

***Static* nodes:** $g_n(t+1)=F_n(g_{n1}(t),g_{n2}(t),\ldots,\ g_{nk}(t))$

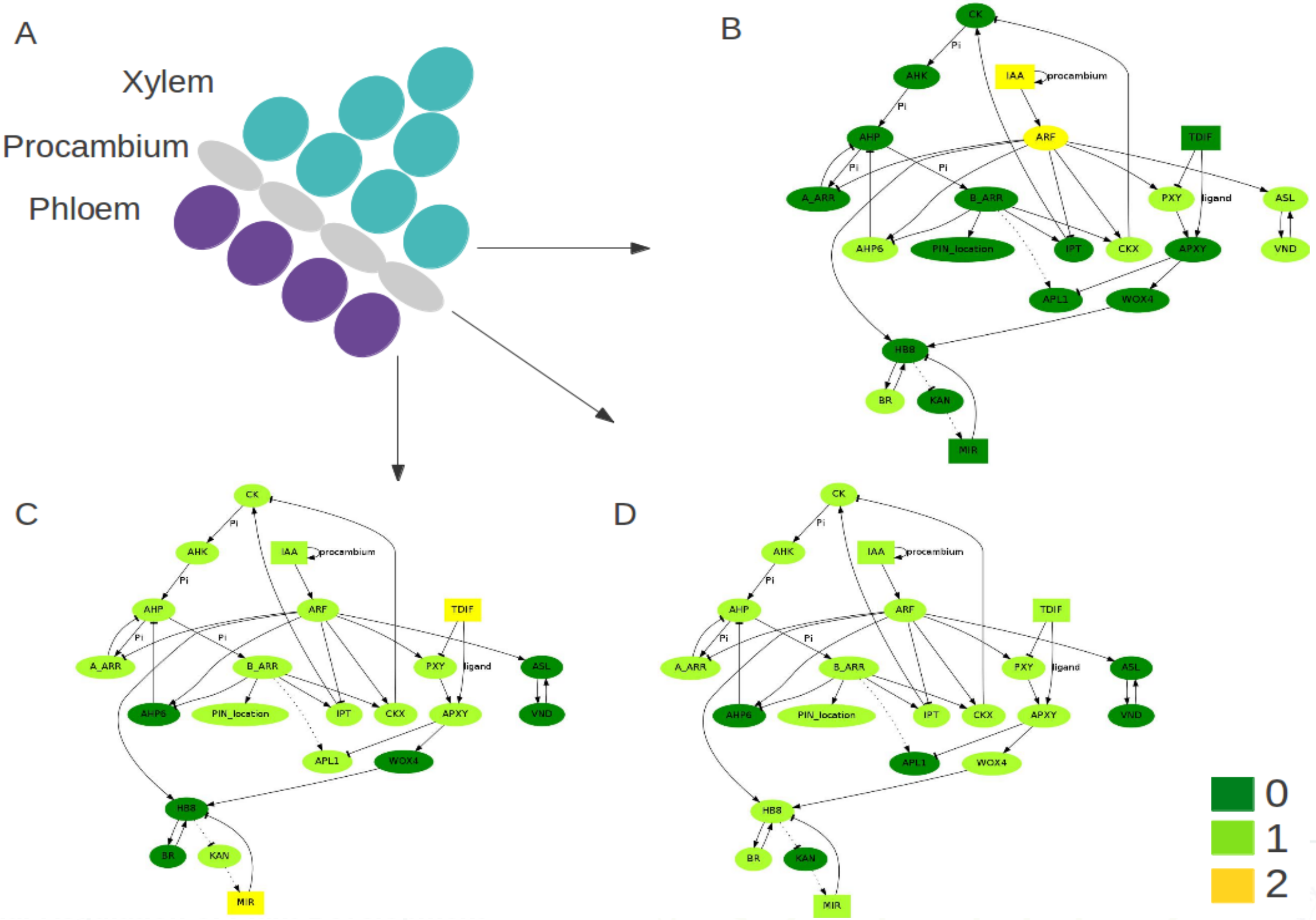
***Mobile nodes*:** $g_{(t+1)T\,[i]}= H(g_{(t)\,[i]}+ D\,(g_{(t)\,[i+1]}+g_{(t)\,[i-1]} - N(g_{(t)\,[i]}))-b)$



Benitez and Hejatko, ***submitted***

Molecular Regulatory Networks Modeling

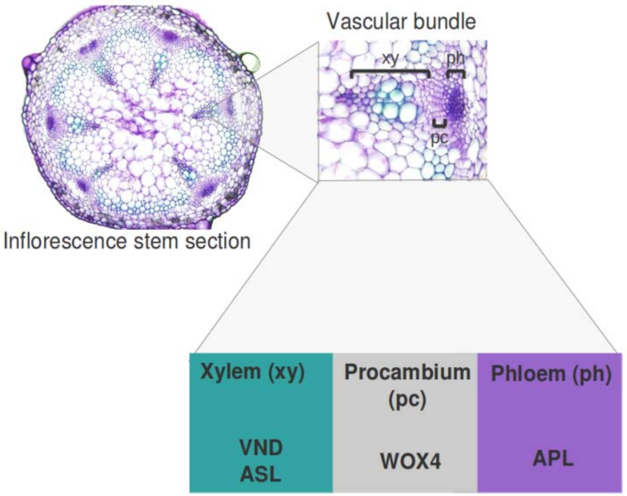
- The **good model** should be able to **simulate reality**

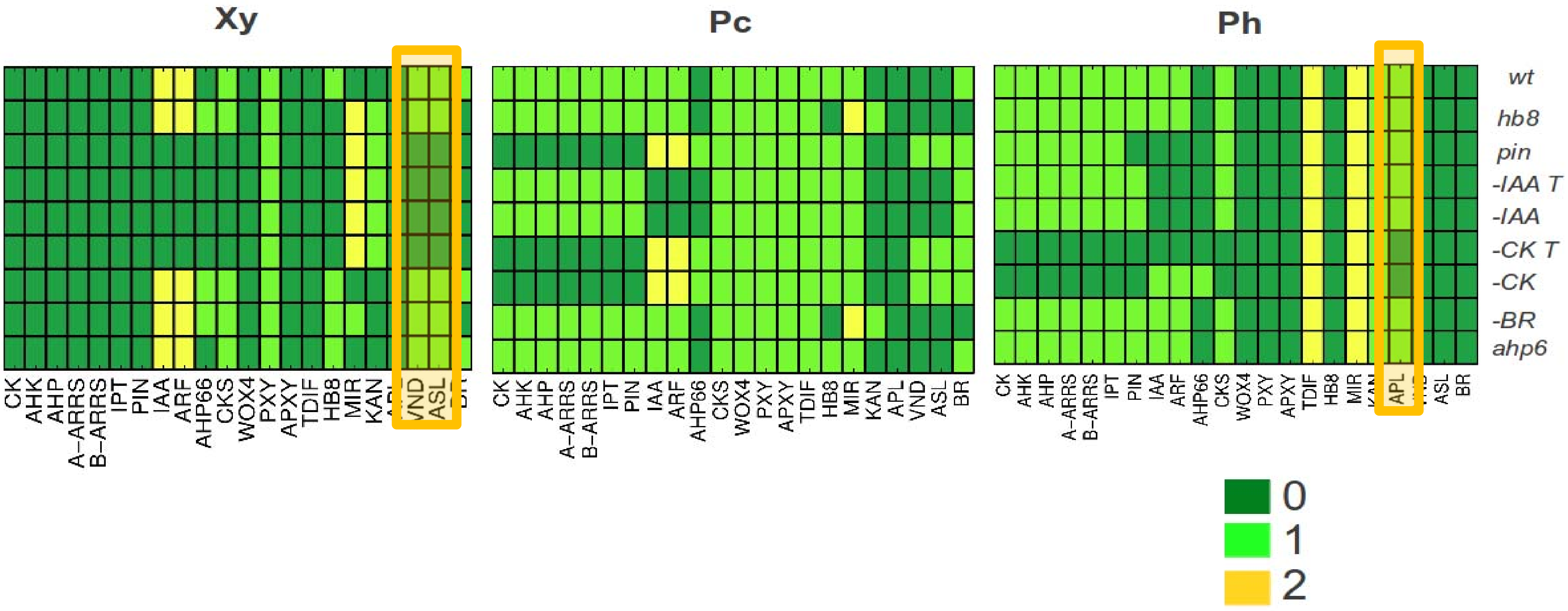


Benitez and Hejatko, *submitted*

Molecular Regulatory Networks Modeling

- Simulation of ***mutants***





Outline

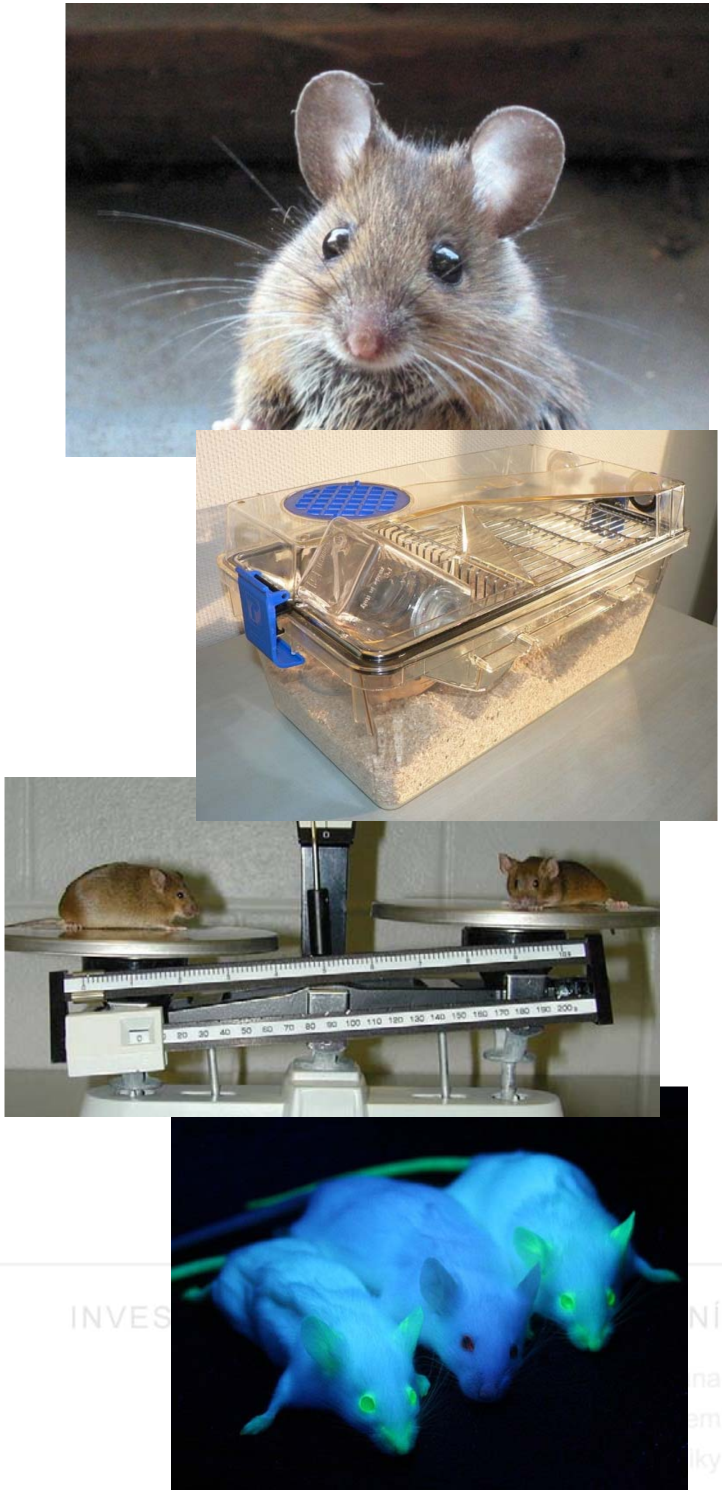
- Tools of **systems biology**
    - **Gene ontology** analysis
    - **Molecular Regulatory Networks Modeling**

- Model organisms
    - *Mus musculus*

Mus musculus

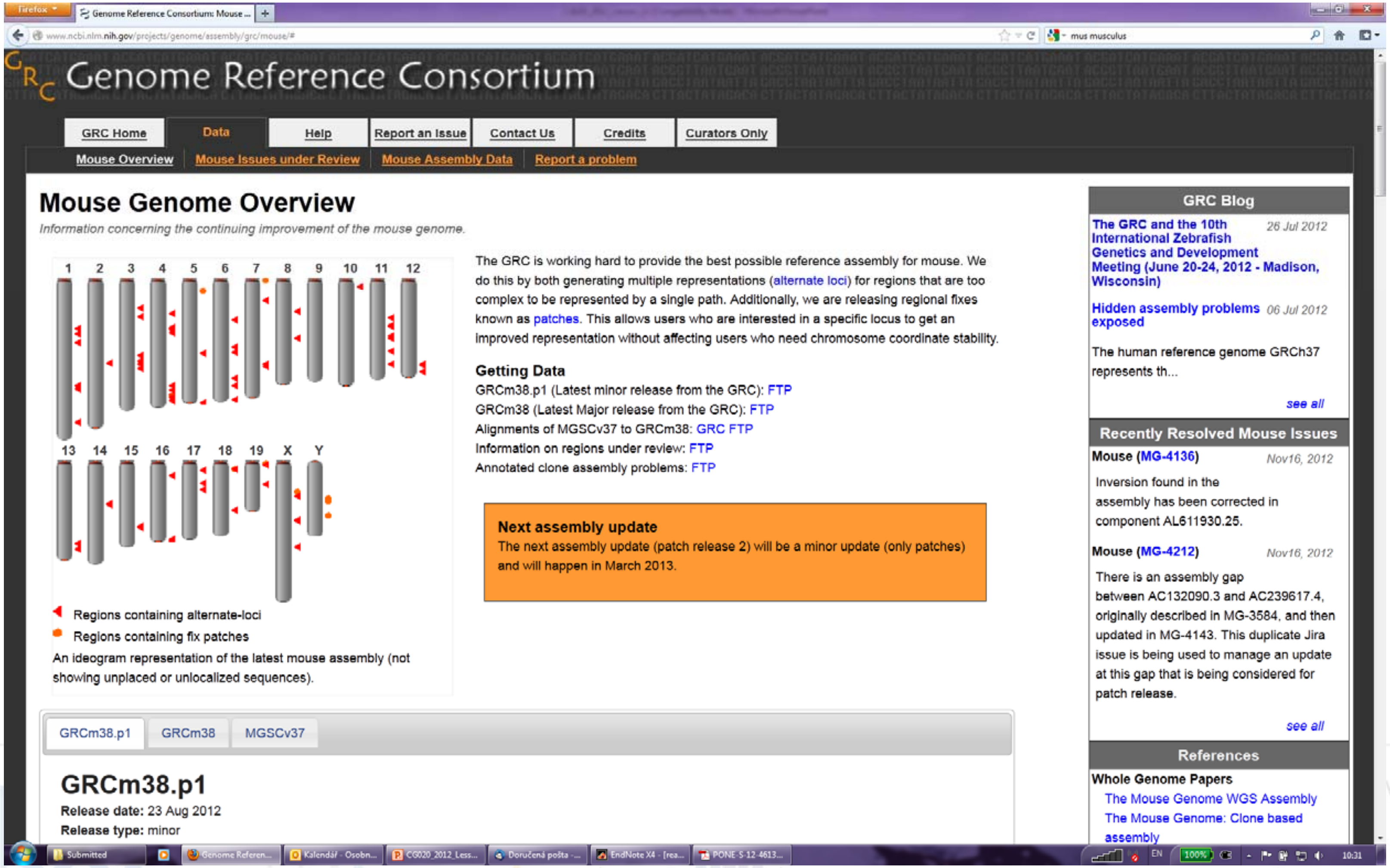
house mouse

- Low requirements for area
- Relatively large number of offspring (3-14, 6-8 on average)
- Genome size is close to the size of human genome (about 3000 Mbp), the number of genes as well (about 24K)
- 20 chromosomes (19+1)
- Suitable for a wide range of physiological experiments (anatomical and physiological similarity to human)
- Possibility to obtain (quite easily) KO mutants and transgenic lines



Mus musculus
house mouse

- Genome known since 2002 (http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/mouse/)



Mouse Genome Overview

Information concerning the continuing improvement of the mouse genome.

The GRC is working hard to provide the best possible reference assembly for mouse. We do this by both generating multiple representations (alternate loci) for regions that are too complex to be represented by a single path. Additionally, we are releasing regional fixes known as patches. This allows users who are interested in a specific locus to get an improved representation without affecting users who need chromosome coordinate stability.

◀ Regions containing alternate-loci

● Regions containing fix patches

An ideogram representation of the latest mouse assembly (not showing unplaced or unlocalized sequences).

Getting Data

GRCm38.p1 (Latest minor release from the GRC): FTP

GRCm38 (Latest Major release from the GRC): FTP

Alignments of MGSCv37 to GRCm38: GRC FTP

Information on regions under review: FTP

Annotated clone assembly problems: FTP

Next assembly update

The next assembly update (patch release 2) will be a minor update (only patches) and will happen in March 2013.

GRCm38.p1

Release date: 23 Aug 2012

Release type: minor

GRC Blog

The GRC and the 10th International Zebrafish Genetics and Development Meeting (June 20-24, 2012 - Madison, Wisconsin) *26 Jul 2012*

Hidden assembly problems exposed *06 Jul 2012*

The human reference genome GRCh37 represents th...

see all

Recently Resolved Mouse Issues

Mouse (MG-4136) *Nov16, 2012*

Inversion found in the assembly has been corrected in component AL611930.25.

Mouse (MG-4212) *Nov16, 2012*

There is an assembly gap between AC132090.3 and AC239617.4, originally described in MG-3584, and then updated in MG-4143. This duplicate Jira issue is being used to manage an update at this gap that is being considered for patch release.

see all

References

Whole Genome Papers

The Mouse Genome WGS Assembly

The Mouse Genome: Clone based assembly

Outline

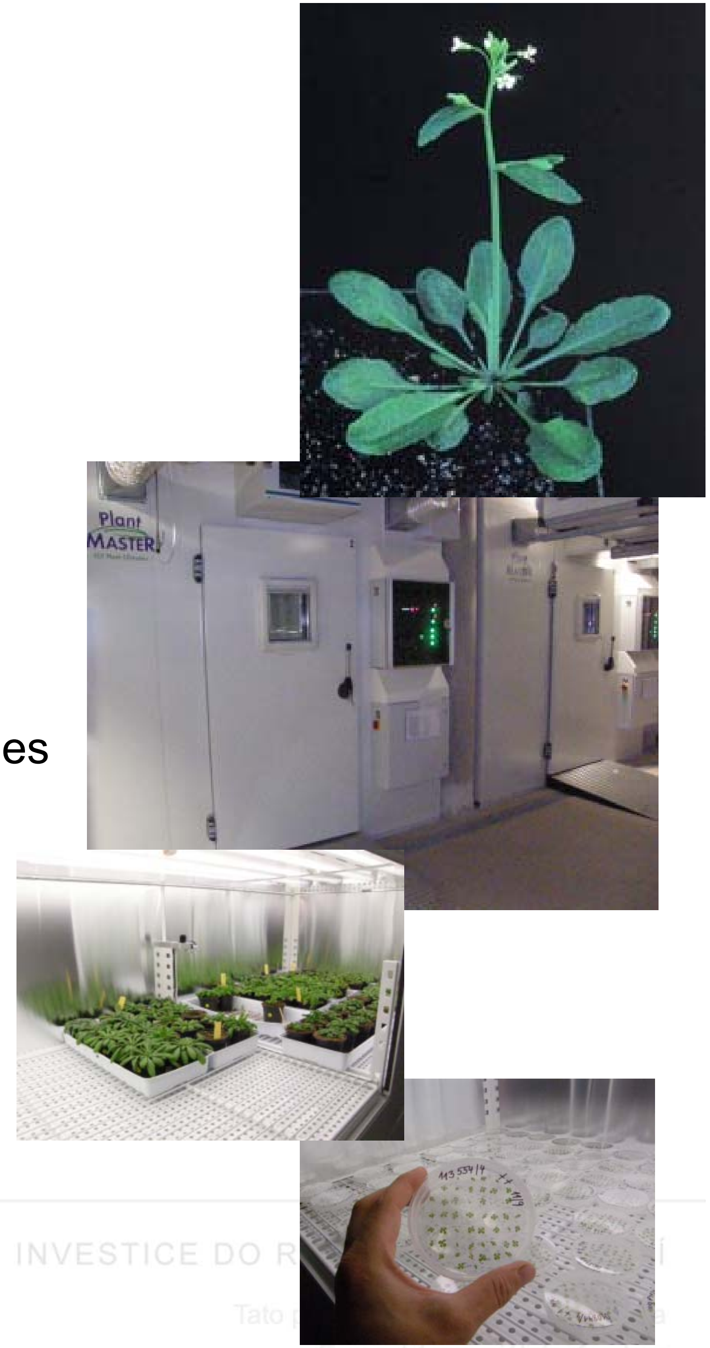
- Tools of **systems biology**
    - **Gene ontology** analysis
    - **Molecular Regulatory Networks Modeling**

- Model organisms
    - *Mus musculus*
    - *Arabidopsis thaliana*

EVROPSKÁ UNIE
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ
Tato prezentace je spolufinancována
Evropským sociálním fondem
a státním rozpočtem České republiky

Arabidopsis thaliana

mouse-ear cress



- Low requirements for cultivation area
- High number of seeds (20.000 per plant and more)
- Small and compact genome, (125 MBp, about 25.000 genes, average size 3 kb)
- 5 chromosomes
- Suitable for wide range od physiological experiments
- High natural variability (approximately 750 ecotypes (Nottingham Arabidopsis Seed Stock Centre))

Columbia 0

Landsberg 0

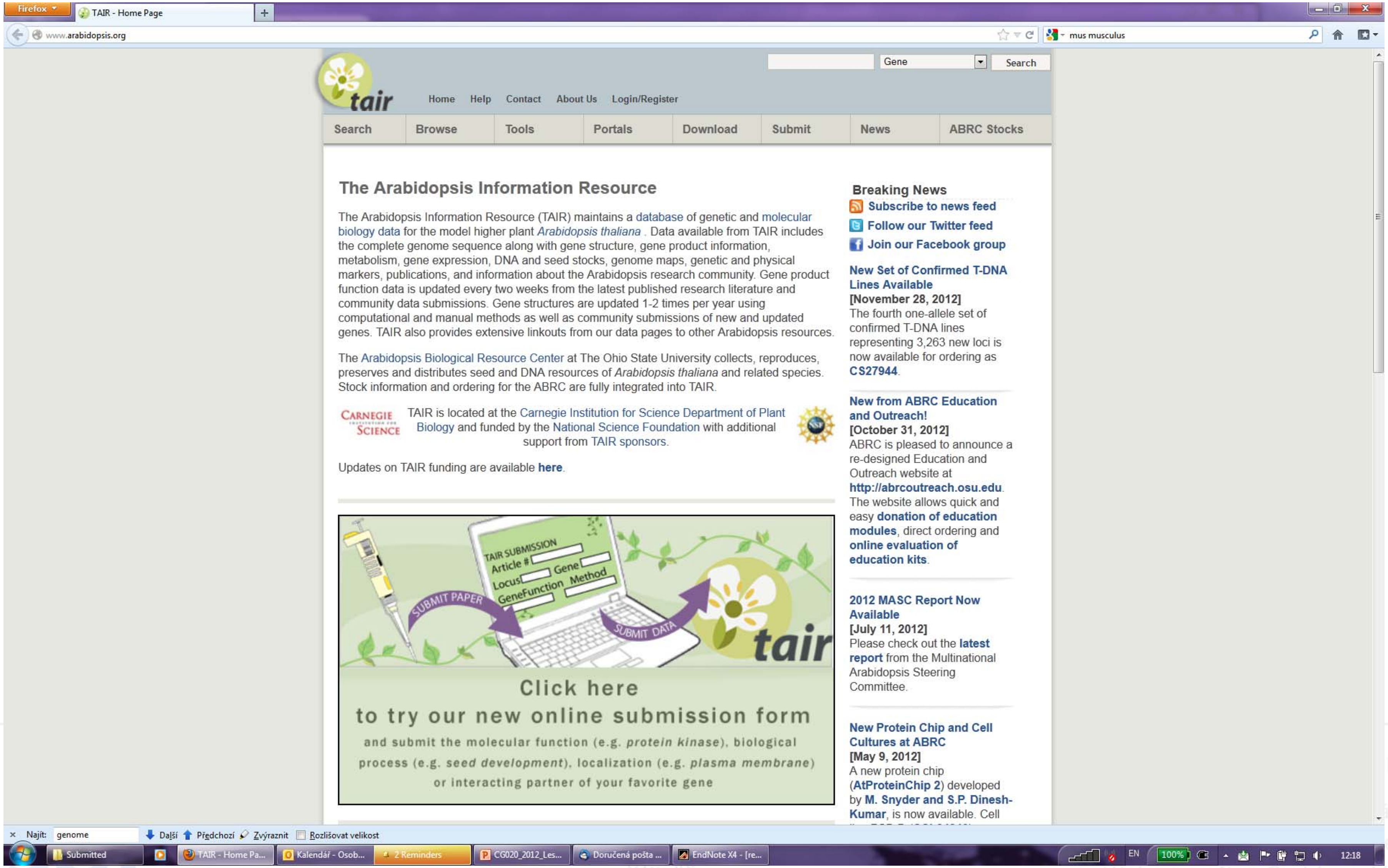
Wassilewskija 0

http://seeds.nottingham.ac.uk/

Arabidopsis thaliana

mouse-ear cress

- Genome known since 2000 (http://www.arabidopsis.org/)



The Arabidopsis Information Resource

The Arabidopsis Information Resource (TAIR) maintains a database of genetic and molecular biology data for the model higher plant *Arabidopsis thaliana* . Data available from TAIR includes the complete genome sequence along with gene structure, gene product information, metabolism, gene expression, DNA and seed stocks, genome maps, genetic and physical markers, publications, and information about the Arabidopsis research community. Gene product function data is updated every two weeks from the latest published research literature and community data submissions. Gene structures are updated 1-2 times per year using computational and manual methods as well as community submissions of new and updated genes. TAIR also provides extensive linkouts from our data pages to other Arabidopsis resources.

The Arabidopsis Biological Resource Center at The Ohio State University collects, reproduces, preserves and distributes seed and DNA resources of *Arabidopsis thaliana* and related species. Stock information and ordering for the ABRC are fully integrated into TAIR.

CARNEGIE SCIENCE TAIR is located at the Carnegie Institution for Science Department of Plant Biology and funded by the National Science Foundation with additional support from TAIR sponsors.

Updates on TAIR funding are available **here**.

Click here

to try our new online submission form

and submit the molecular function (e.g. *protein kinase*), biological process (e.g. *seed development*), localization (e.g. *plasma membrane*) or interacting partner of your favorite gene

Breaking News

Subscribe to news feed

Follow our Twitter feed

Join our Facebook group

New Set of Confirmed T-DNA Lines Available
[November 28, 2012]
The fourth one-allele set of confirmed T-DNA lines representing 3,263 new loci is now available for ordering as **CS27944**.

New from ABRC Education and Outreach!
[October 31, 2012]
ABRC is pleased to announce a re-designed Education and Outreach website at **http://abrcoutreach.osu.edu**. The website allows quick and easy **donation of education modules**, direct ordering and **online evaluation of education kits**.

2012 MASC Report Now Available
[July 11, 2012]
Please check out the **latest report** from the Multinational Arabidopsis Steering Committee.

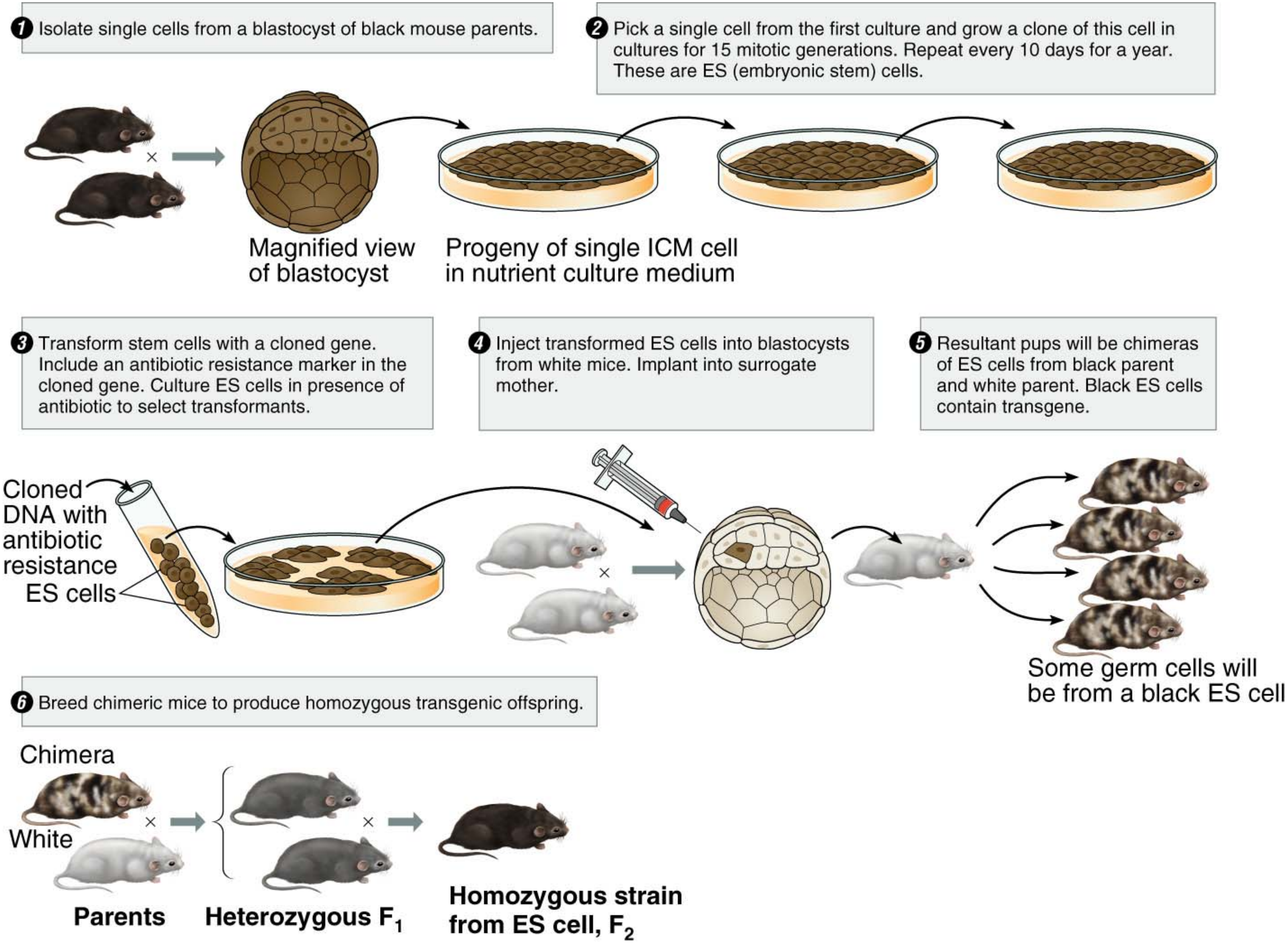
New Protein Chip and Cell Cultures at ABRC
[May 9, 2012]
A new protein chip (**AtProteinChip 2**) developed by **M. Snyder and S.P. Dinesh-Kumar**, is now available. Cell

Outline

- Tools of **systems biology**
  - **Gene ontology** analysis
  - **Molecular Regulatory Networks Modeling**

- Model organisms
  - *Mus musculus*
  - *Arabidopsis thaliana*

- Selected **methods of molecular biology**
  - Preparation of transgenic organisms



❶ Isolate single cells from a blastocyst of black mouse parents.

❷ Pick a single cell from the first culture and grow a clone of this cell in cultures for 15 mitotic generations. Repeat every 10 days for a year. These are ES (embryonic stem) cells.

Magnified view of blastocyst

Progeny of single ICM cell in nutrient culture medium

❸ Transform stem cells with a cloned gene. Include an antibiotic resistance marker in the cloned gene. Culture ES cells in presence of antibiotic to select transformants.

❹ Inject transformed ES cells into blastocysts from white mice. Implant into surrogate mother.

❺ Resultant pups will be chimeras of ES cells from black parent and white parent. Black ES cells contain transgene.

Cloned DNA with antibiotic resistance

ES cells

Some germ cells will be from a black ES cell

❻ Breed chimeric mice to produce homozygous transgenic offspring.

Chimera

White

Parents

Heterozygous F_1

Homozygous strain from ES cell, F_2

EVROPSKÁ UNIE
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ
Tato prezentace je spolufinancována
Evropským sociálním fondem
a státním rozpočtem České republiky

Transformation of *Arabidopsis* by *Agrobacterium tumefaciens*

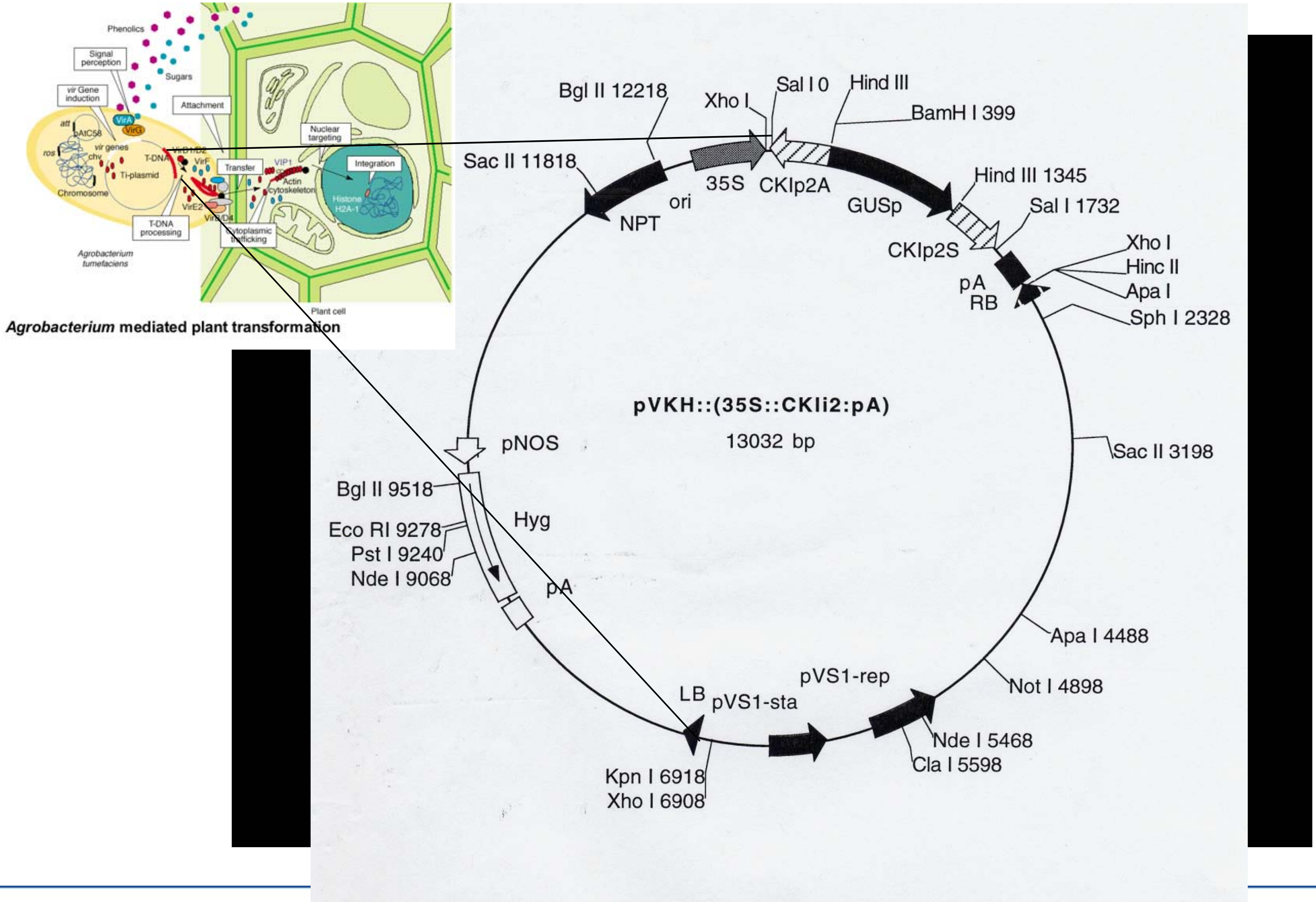
Crown gall of raspberry caused by *Agrobacterium tumefaciens*.

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Tato prezentace je spolufinancována
Evropským sociálním fondem
a státním rozpočtem České republiky

Transformation of *Arabidopsis* by *Agrobacterium tumefaciens*

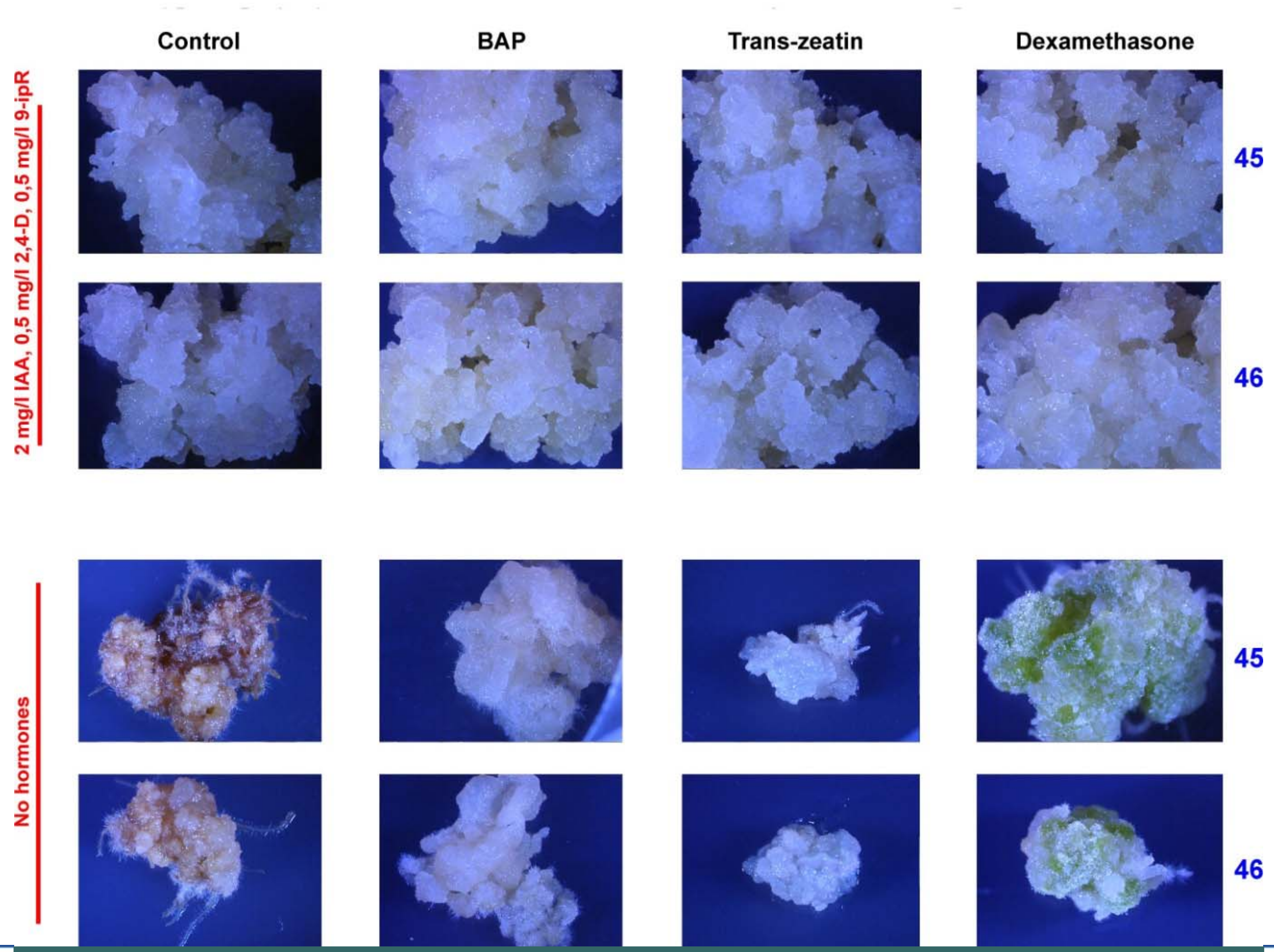
Transfer of bacterial DNA into plant cells



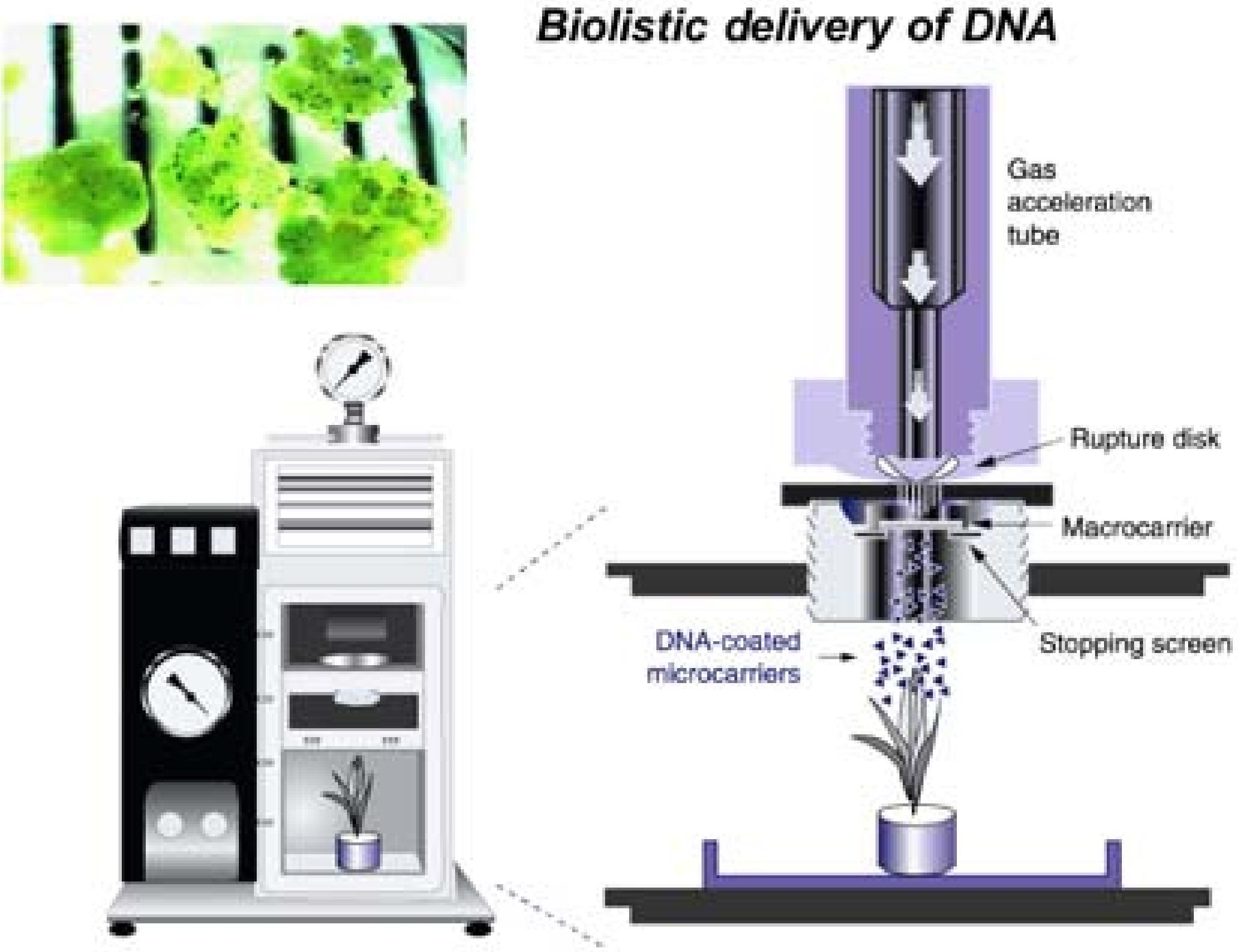
***Agrobacterium* mediated plant transformation**

Transformation by cocultivation of leaf discs

Transformation by cocultivation of calluses



Transformation by biolistic delivery of DNA



Transformation of inflorescence

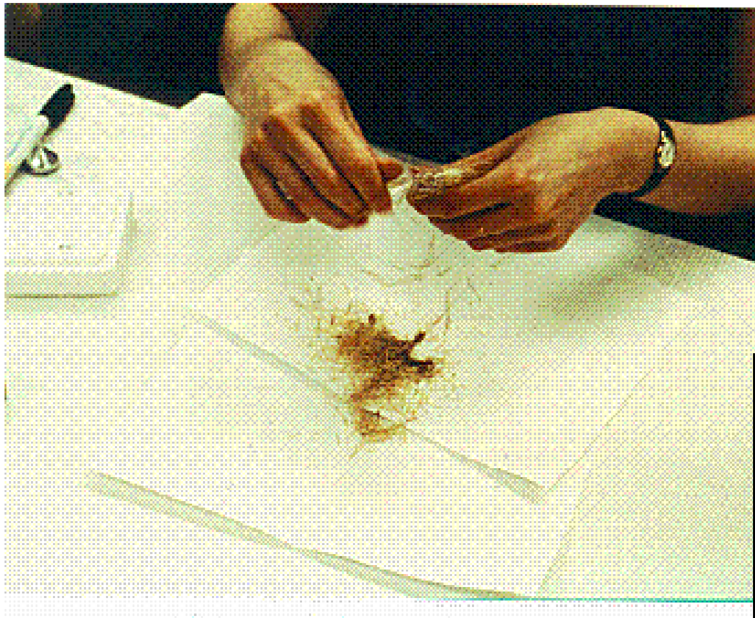
When plants have primary bolts 5-15 cm they are ready to infiltrate. Clipping of primary bolts is not necessary.

After infiltration, pots are placed on their sides to allow for drainage and are covered with plastic wrap. Plants are returned, in this state, to the growth chamber for 24 hours. After 24 hours, they are turned upright into a fresh flat.

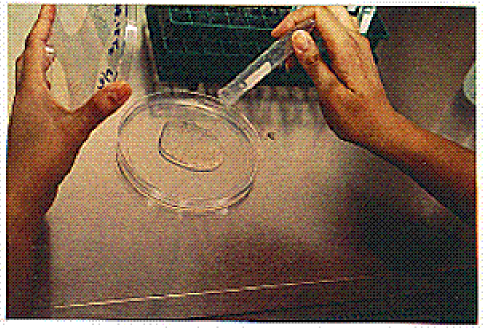
Plants are allowed to grow to maturity. They are staked to avoid seed loss and facilitate plant harvesting.
NOTE: Leaves degenerate within 2 weeks of infiltration. This is normal and does not affect seed set.

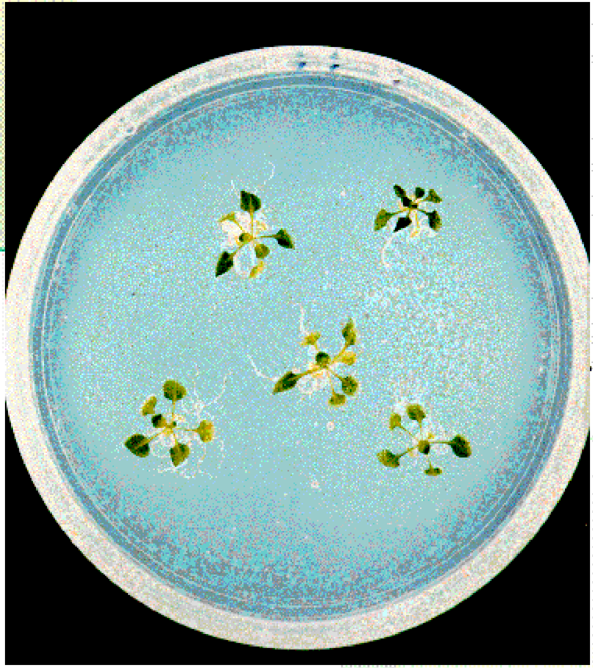
http://www.bch.msu.edu/pamgreen/green.htm

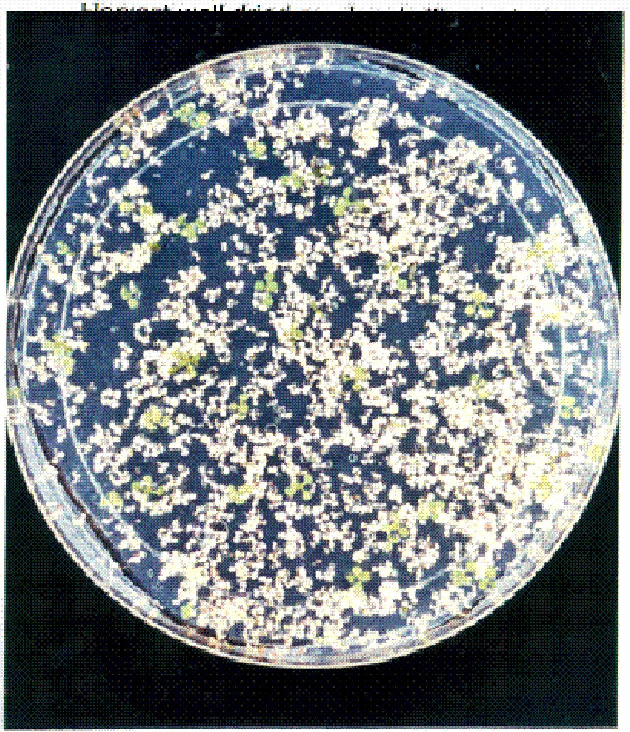
Transformation of inflorescence



Sterilize seed in bleach solution.







Transformed seedlings are green and have true leaves on selective medium (a 40mg/l kanamycin plate is shown).

Plant transformed seedlings in soil.

http://www.bch.msu.edu/pamgreen/green.htm

Outline

- Tools of **systems biology**
  - **Gene ontology** analysis
  - **Molecular Regulatory Networks Modeling**
- Model organisms
  - *Mus musculus*
  - *Arabidopsis thaliana*
- Selected **methods of molecular biology**
  - Preparation of transgenic organisms
  - PCR

EVROPSKÁ UNIE
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ
Tato prezentace je spolufinancována
Evropským sociálním fondem
a státním rozpočtem České republiky

PCR

Outline

- Tools of **systems biology**
  - **Gene ontology** analysis
  - **Molecular Regulatory Networks Modeling**
- Model organisms
  - *Mus musculus*
  - *Arabidopsis thaliana*
- Selected **methods of molecular biology**
  - Preparation of transgenic organisms
  - PCR
  - Design and preparation of primers (Dr. Hana Konečná)

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Tato prezentace je spolufinancována
Evropským sociálním fondem
a státním rozpočtem České republiky

Summary

- Tools of **systems biology**
  - **Gene ontology** analysis
  - **Molecular Regulatory Networks Modeling**

- Model organisms
  - *Mus musculus*
  - *Arabidopsis thaliana*

- Selected **methods of molecular biology**
  - Preparation of transgenic organisms
  - PCR
  - Design and preparation of primers (Dr. Hana Konečná)

Discussion