

Establishment of retrotransposon-mutagenized population of model legume *Lotus japonicus* and high throughput, deep sequencing-based insertion site identification

Dorian Urbański

PBD Annual meeting 2012,  
Copenhagen

Forward genetics:

Phenotypic screening → mutation mapping

VS.

Reverse genetics:

Altered gene sequence/expression → analysis of phenotype

TILLING facility (2003) is the only reverse genetics resource available for *Lotus japonicus* up to date

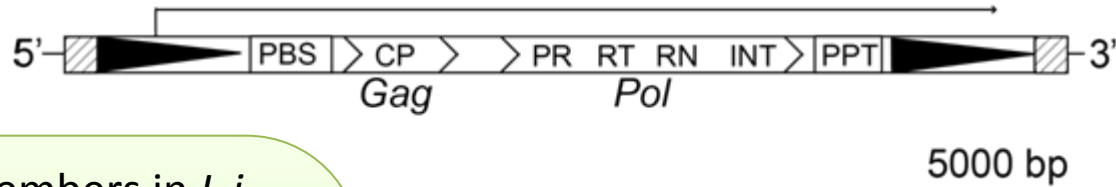
Development of insertional mutagenesis system that will allow robust mutagenesis and high-throughput/systematic annotation of mutants.

Establishment of mutant seed collection

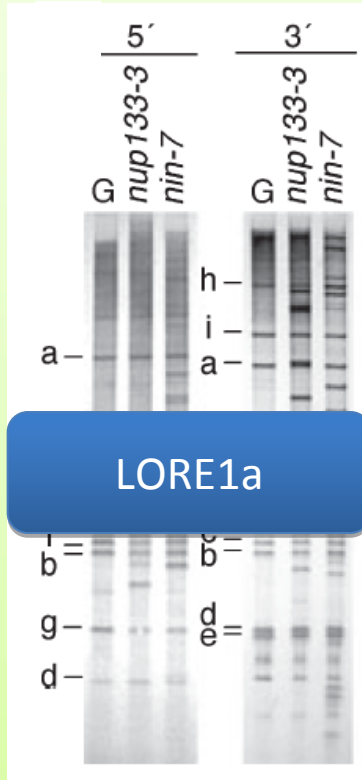
# Lotus retrotransposon 1 (LORE1)

A candidate for insertional mutagen

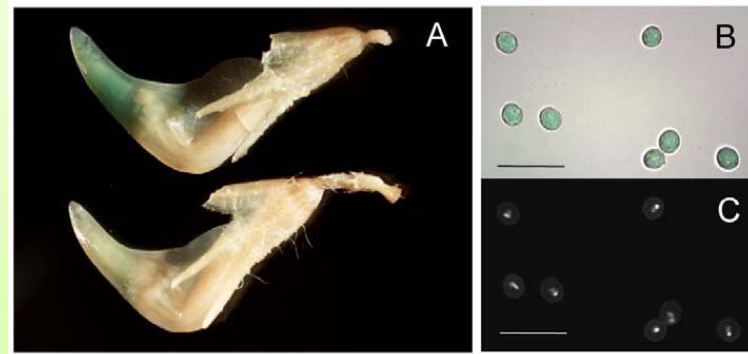
Ty3-gypsy



8 family members in *L.j.*  
var Gifu



1. De-repression during tissue culture
2. Activity in germlines



Madsen *et al.*, Plant Journal (2005)  
Fukai *et al.*, PLoS Genetics (2010)

Lotus endogenous element

Autonomous retrotransposon – „copy & paste” replication mechanism

No insertion site specificity

Active almost only in pollen  
Not active in sporophyte or tissue culture

Siblings from the same pod do not share  
the same insertions

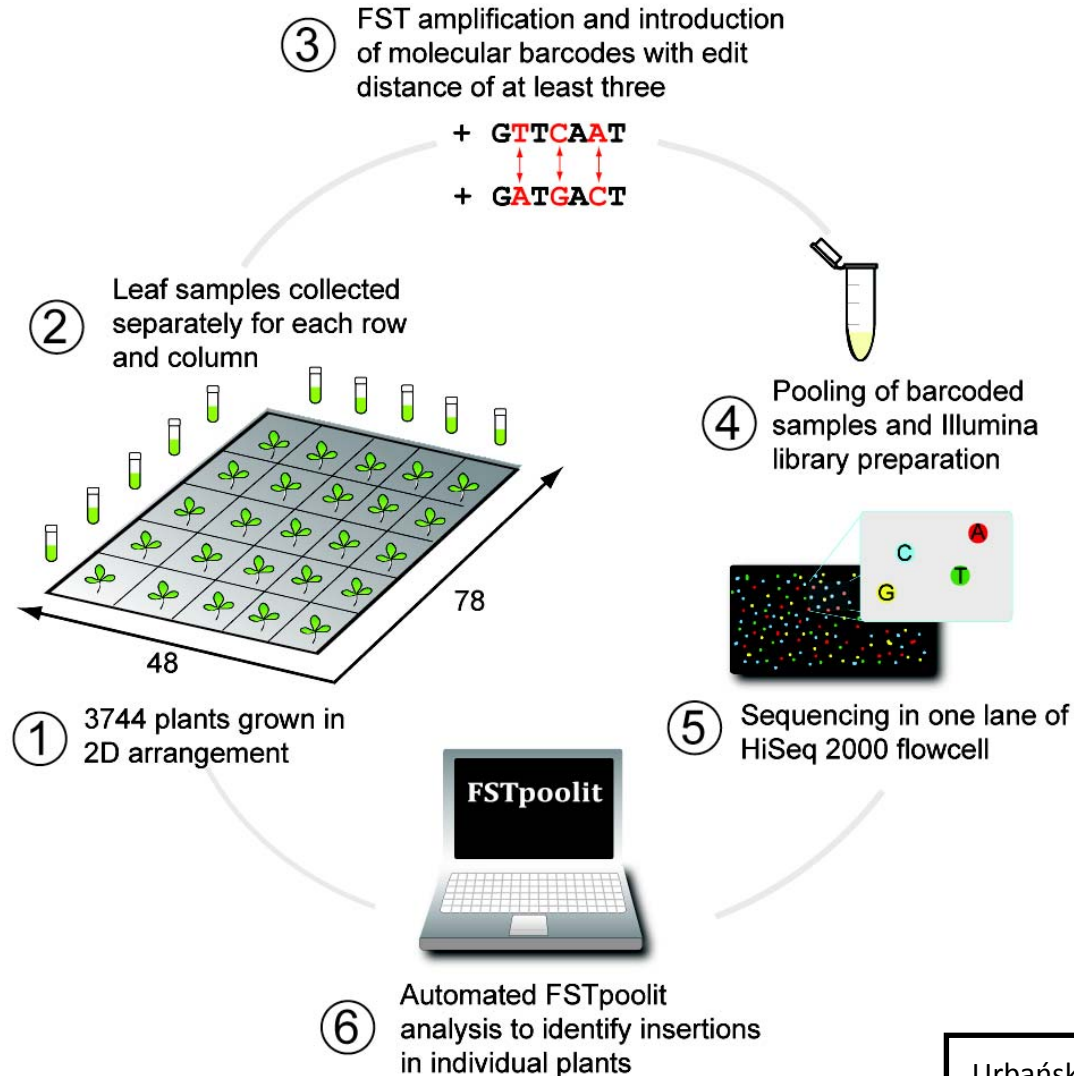
# Concept for reverse genetics pipeline - insertional mutagenesis

1) Establishment of mutant population

2) Enrichment and sequencing of insertion flanking regions

3) Computational analysis of insertion sites and mutant lines  
annotation  
Assignment of insertions to correct seed bag

## Flanking Sequence Tag pooling and identification



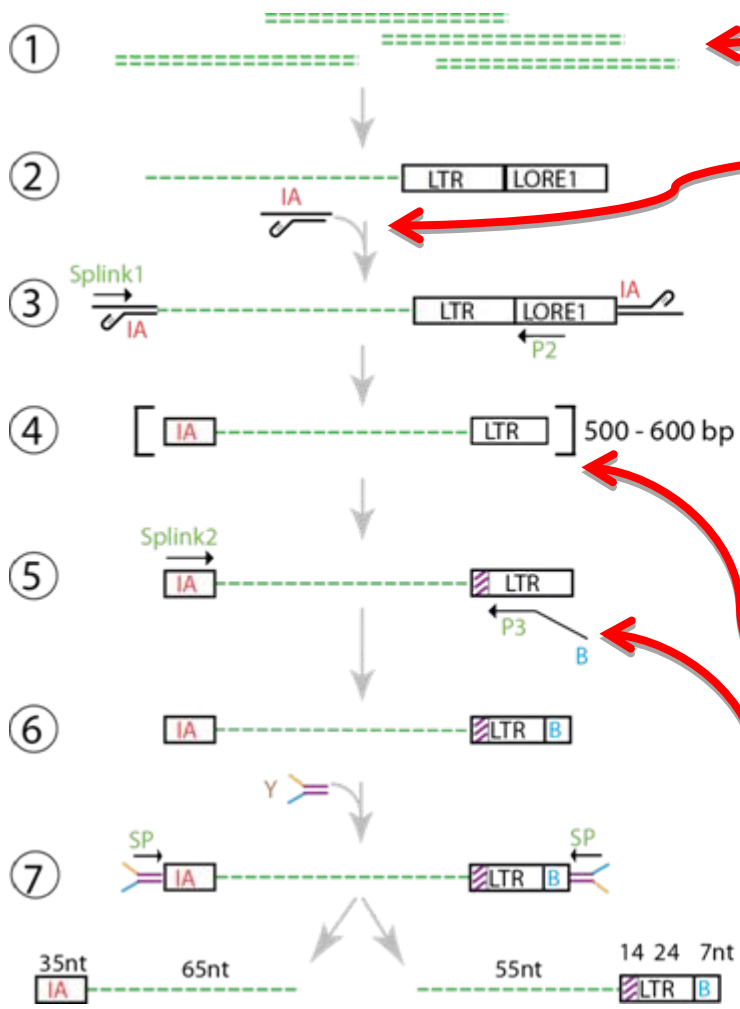
# FST amplification and tagging

No bias

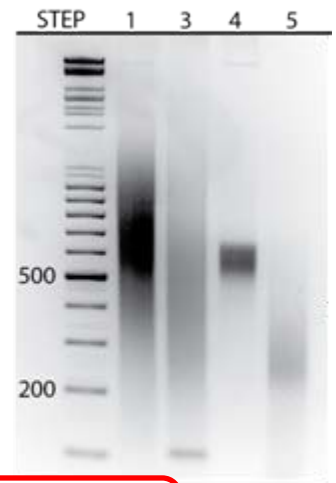
Reduced mis-priming and  
PCR artifacts

No unspecific products,  
control step

multiple tagging available



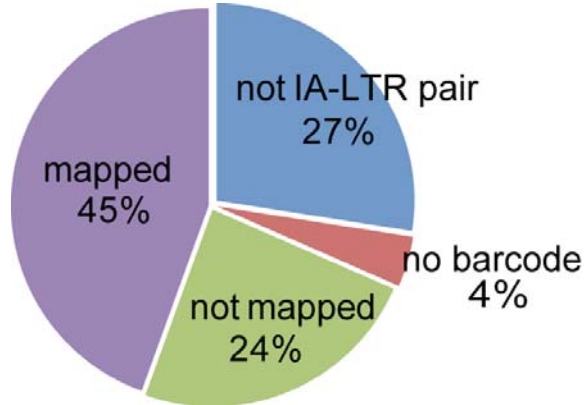
- LTR long terminal repeats
- IA Splinkerette intermediate adapter
- B unique barcode
- P1/Splink1 PCR primers
- Y Illumina Y adapter
- unique genomic sequence
- SP sequencing primer
- 14nt LTR specific sequence



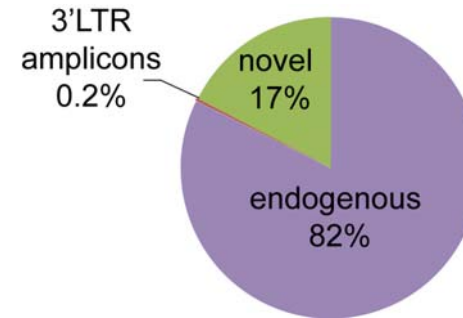
Sequencing analysis by custom made **FSTpoolit** bioinformatics pipeline

Urbański *et al.*, Plant Journal (2011)

## Sequencing reads representation



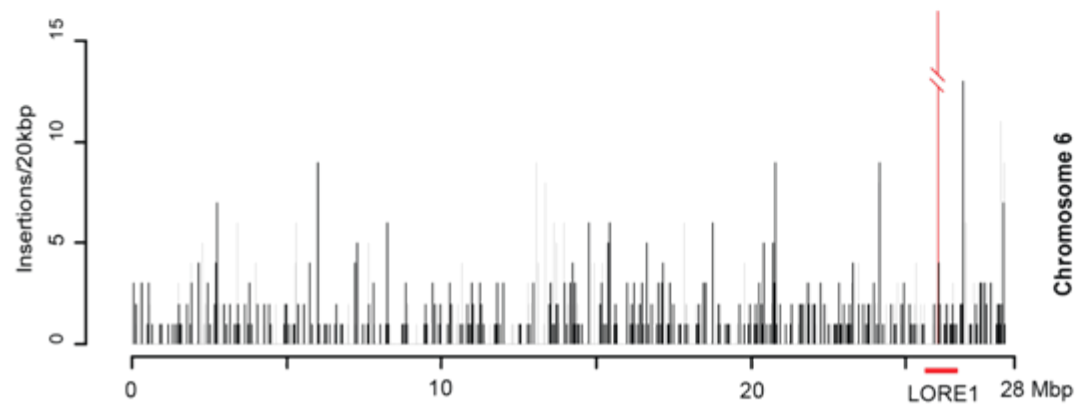
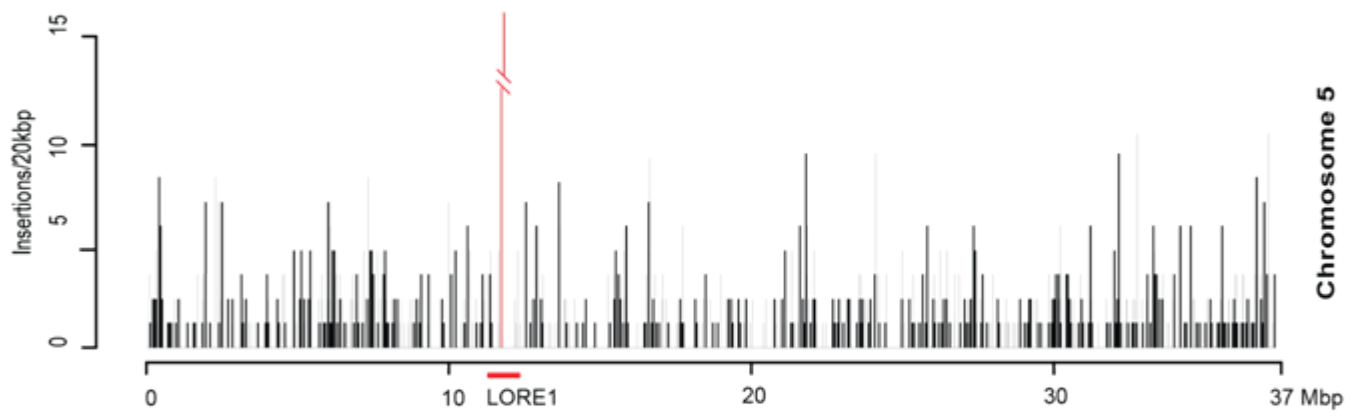
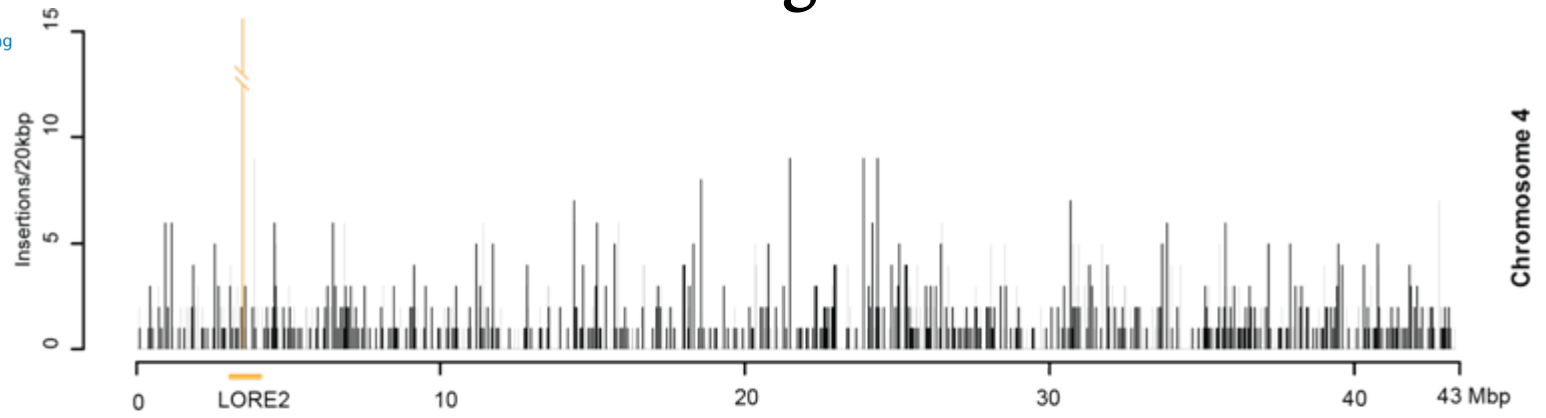
## novel vs. endogenous



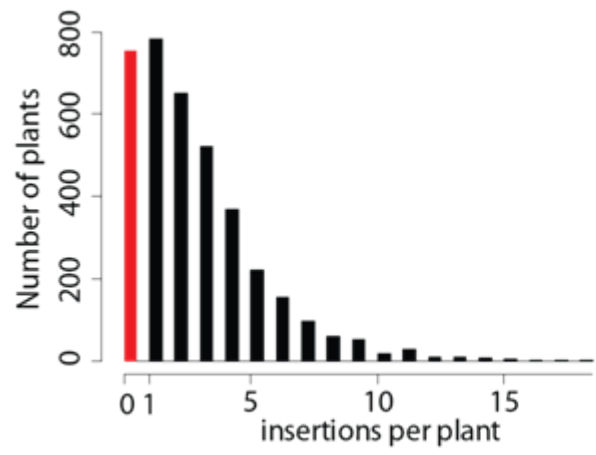
9 out of 121 mio reads - 8% - contribute to identification of novel insertions

Protocol	Generation	Number of investigated plants	New insertions detected compared to R <sub>1</sub>	Progeny plants with new insertions	Maximum number of insertions per plant	Mean insertion number acquired in generation
Southern blot	R <sub>1</sub>	1	–	1	–	3
Southern blot	R <sub>2</sub>	72	210	64 (88.9%)	10	2.9
<i>FSTpoolit</i>	R <sub>2</sub>	3744	8935	3054 (81.6%)	19	2.4
Southern blot	R <sub>3</sub>	90	427	89 (98.9%)	NA	1.9

# Even distribution along chromosomes

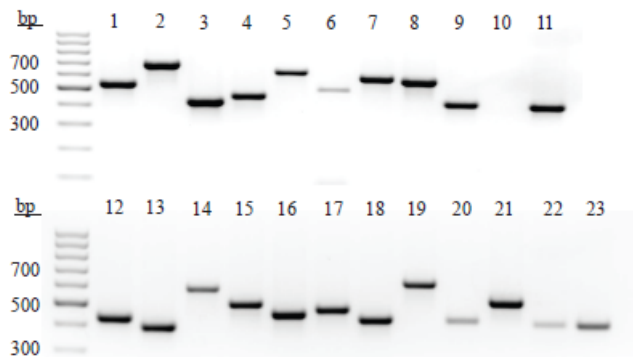


Urbański *et al.*, Plant Journal (2011)

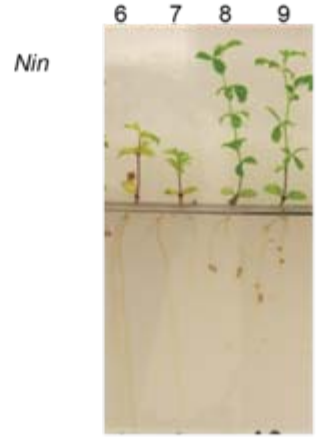
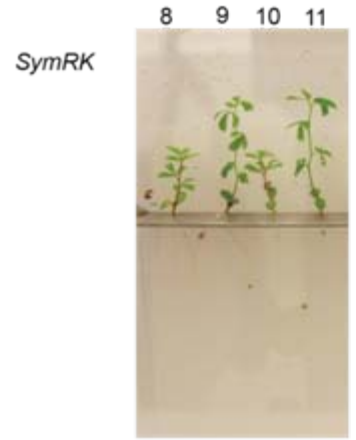


# High reliability

Sample no.	Chromosome	Position	Strand	Annotation	Description	Reads	predicted PCR product length (bp)	Confirmed?
1	chr0	2580678	R	exon	Ribonuclease III/ Dicer-like 4	658	535	YES
2	chr0	26154427	R	exon	E3 ubiquitin-protein ligase	1093	688	YES
3	chr6	8010860	F	exon	DNA topoisomerase, type IA, core	1639	421	YES
4	chr1	56996712	F	intron	Glutathione gamma-glutamylcysteinyltransferase	327	460	YES
5	chr2	31037122	F	exon	Armadillo repeat-containing kinesin-like protein	1784	627	YES
6	chr2	44369165	F	exon	Glutamate receptor	3775	499	YES
7	chr2	8951564	R	exon	Argonaute/Dicer protein, PAZ	2791	582	YES
8	chr3	10567763	R	exon	Cation/H <sup>+</sup> exchanger	352	569	YES
9	chr3	37524332	F	intron	Two pore calcium channel protein 1	502	423	YES
10	chr3	40800337	R	exon	Argonaute	721	674	NO
11	chr3	4223092	F	exon	Glycosyl transferase, family 14	541	416	YES
12	chr4	12854140	R	exon	Oxoglutarate/iron-dependent oxygenase	1369	437	YES
13	chr4	18167999	R	exon	No apical meristem (NAM) protein	4376	395	YES
14	chr4	26859208	F	exon	Mitochondrial substrate carrier	2043	593	YES
15	chr4	3316558	R	exon	Casp1p-like	2241	505	YES
16	chr4	36922800	F	exon	Plant specific eukaryotic initiation factor 4B	901	454	YES
17	chr4	37572099	F	exon	Uncharacterised conserved protein UCP009193	1739	478	YES
18	chr4	41247467	F	exon	Initiator tRNA phosphoribosyl transferase	2102	426	YES
19	chr4	41597366	F	exon	Ty3- retrotransposon gag protein	742	620	YES
20	chr4	6854996	R	exon	Retrovirus-related Pol polyprotein	71	420	YES
21	chr5	7885233	R	exon	Arginine biosynthesis protein ArgJ	948	516	YES
22	chr6	21547664	F	intron	Transcription initiation factor IIF, alpha subunit	573	410	YES
23	chr6	5309678	R	exon	Aminotransferase, class IV	358	407	YES



# New alleles of known nod<sup>-</sup> genes



*SymRK* - segregation 5:12:6

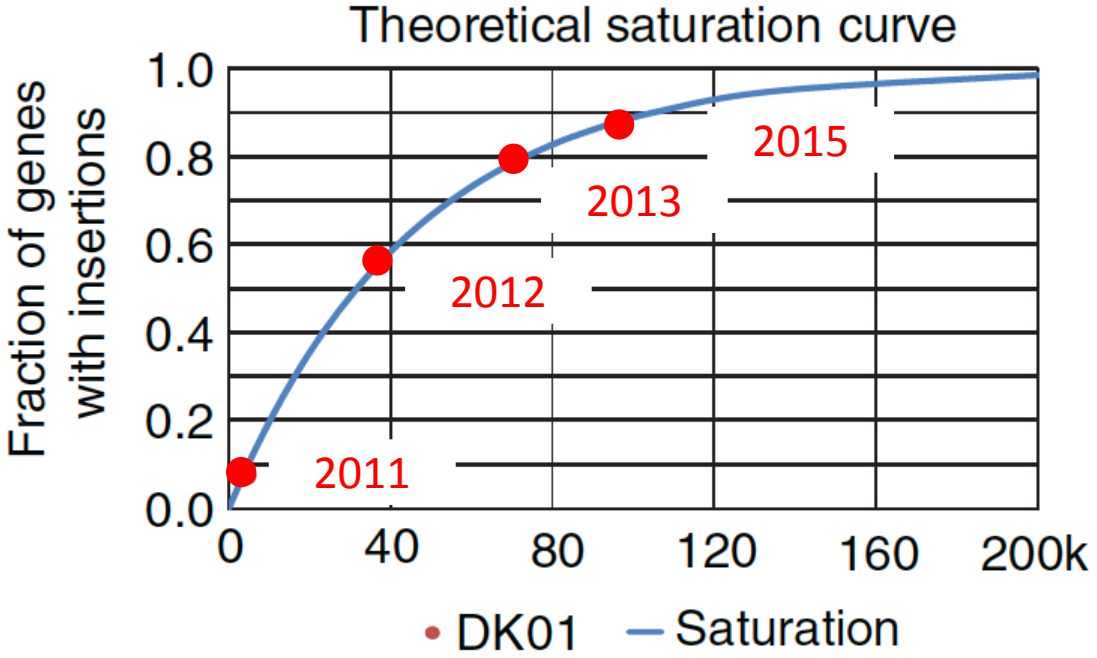
phenotype	+	+	+	-	+	-	-	-	-	-	-	-	-	+	-	-	+	-	-	-	-	-	-	-
plant no.	1	2	8	9	10	11	12	13	14	15	18	19	20	22	23	24	26	27	28	30	31	32	33	34
genotype	H	H	H	HT	H	WT	HT	HT	HT	HT	HT	HT	HT	HT	H	HT	WT	-	HT	HT	WT	WT	WT	WT

*Nin* - segregation 6:16:2

phenotype	-	-	-	-	-	+	+	-	-	-	-	+	-	-	-	+	+	-	-	-	+	-	-	-
plant no.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	25
genotype	HT	HT	HT	WT	HT	H	H	HT	HT	HT	HT	H	HT	HT	HT	H	H	HT	HT	HT	H	HT	WT	HT



2000 m<sup>2</sup> - 27 000 plants      March – November  
Barcoding and information storage in a database



In collaboration with Prof. Makoto Hayashi,  
National Institute of Agrobiological Science, JAPAN

- All mutant lines are available freely upon request
- The resource supports R3 generation (segregating population)

## CARB Centre

<http://www.carb.dk/resources.asp>

- BLAST
- Insertion list with genotyping primers
- LORE1 resource FAQ
- FSTpoolit software

## KAZUSA Institute

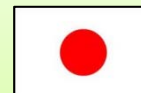
<http://www.kazusa.or.jp/lotus/>

- BLAST
- Genome browser with annotated LORE1 insertions

## Contact information



- [dou@mb.au.dk](mailto:dou@mb.au.dk)
- [ann@mb.au.dk](mailto:ann@mb.au.dk)
- [sua@mb.au.dk](mailto:sua@mb.au.dk)



- [makotoh@affrc.go.jp](mailto:makotoh@affrc.go.jp)
- [efu@affrc.go.jp](mailto:efu@affrc.go.jp)

batch	DK01
chromosome	chr1
position	286597
orientation	F
column_barcode	C_44, C92
row_barcode	R_36
plant_ID	30003390
seeds_stock	.
pool_coordinates	C_44#C_92#R_36
coordinate_abundances	50#123#185
abundances_sum	358
Forward_primer	TGGTGTAGGGATTGCGGGAAACAA
Reverse_primer	AATTGCCCCATTTCGAGCTTCCCAA
PCR_product_insertion_positive	435
PCR_product_wt	593
+/-1000_bp_insertion_flanking	CGTCG....



# LORE1 as a tool for forward genetics



nod<sup>-</sup>



fix<sup>-</sup>



photosynthesis

- New mutagen – LORE1 – retrotransposon with the unique feature of germinal transposition
  - a. Random distribution with few hot-spots
  - b. Preference for exonic insertions
- **FSTpoolit** is an efficient and accurate method for FST detection and annotation
- New reverse genetics resource established
  - a. Easy insertion identification by BLAST
  - b. Lines available off-hand and free of charge
  - c. 150 lines ordered (from Oct 2011)
  - d. Quick development predicted within next 3 years

## **Aarhus University**

- Jens Stougaard
- Stig Uggerhøj Andersen
- Anna Małolepszy
- Niels Sandal
- Finn Pedersen

## **Kazusa Institute JAPAN**

- Shusei Sato

## **National Institute of Agrobiological Sciences, JAPAN**

- Makoto Hayashi
- Eigo Fukai



Danmarks  
Grundforskningsfond  
Danish National  
Research Foundation



**MOBIL**