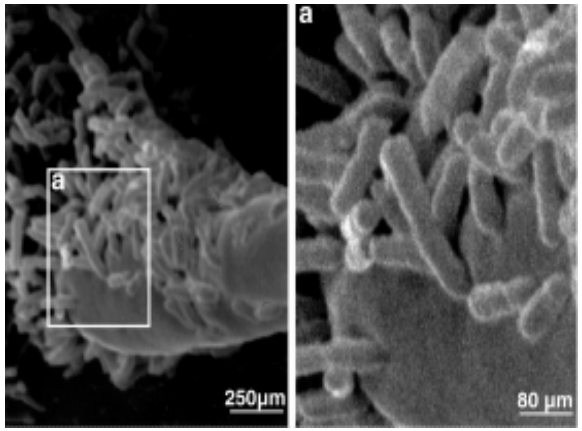
Creating tools for the control of transgene integration and for genome editing in plant species

Department of Molecular, Cellular and Developmental Biology, University of Michigan, Ann Arbor, MI

These days, virtually every plant species can be transformed…



… using various biological and physical means…

But …

several technical and biological barriers do not (in most cases) allow controlling the outcome of a transformation event …

What we got:

- **random integration**
- **integration of plasmid DNA, scrambled DNA, genetic 'contamination'**

What we want:

- **'clean' integration**
- **gene targeting (replacement, insertion, deletion)**
- **gene stacking**

Novel tools are needed for controlled DNA integration and genome editing

Double strand breaks can act as 'traps' for T-DNA integration (via NHEJ)

"Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells."
(Salomon 1998 EMBO J)

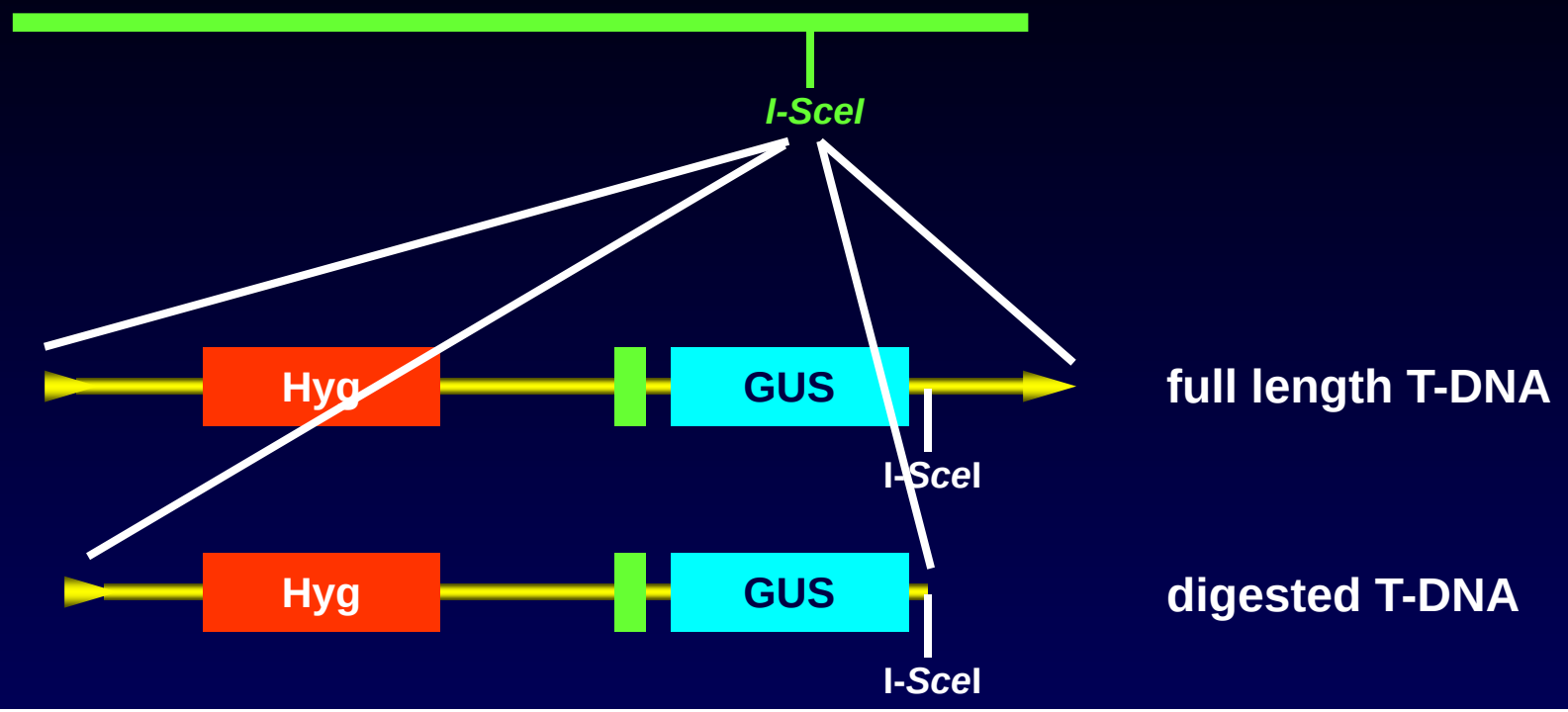
"Targeted integration of T-DNA into the tobacco genome at double-stranded breaks: new insights on the mechanism of T-DNA integration."
(Chilton 2003 Plant Physiol)

"Site-specific integration of *Agrobacterium tumefaciens* T-DNA via double-stranded intermediates."
(Tzfira 2003 Plant Physiol)

Altering the plant's DNA repair machinery leads to GT via HR

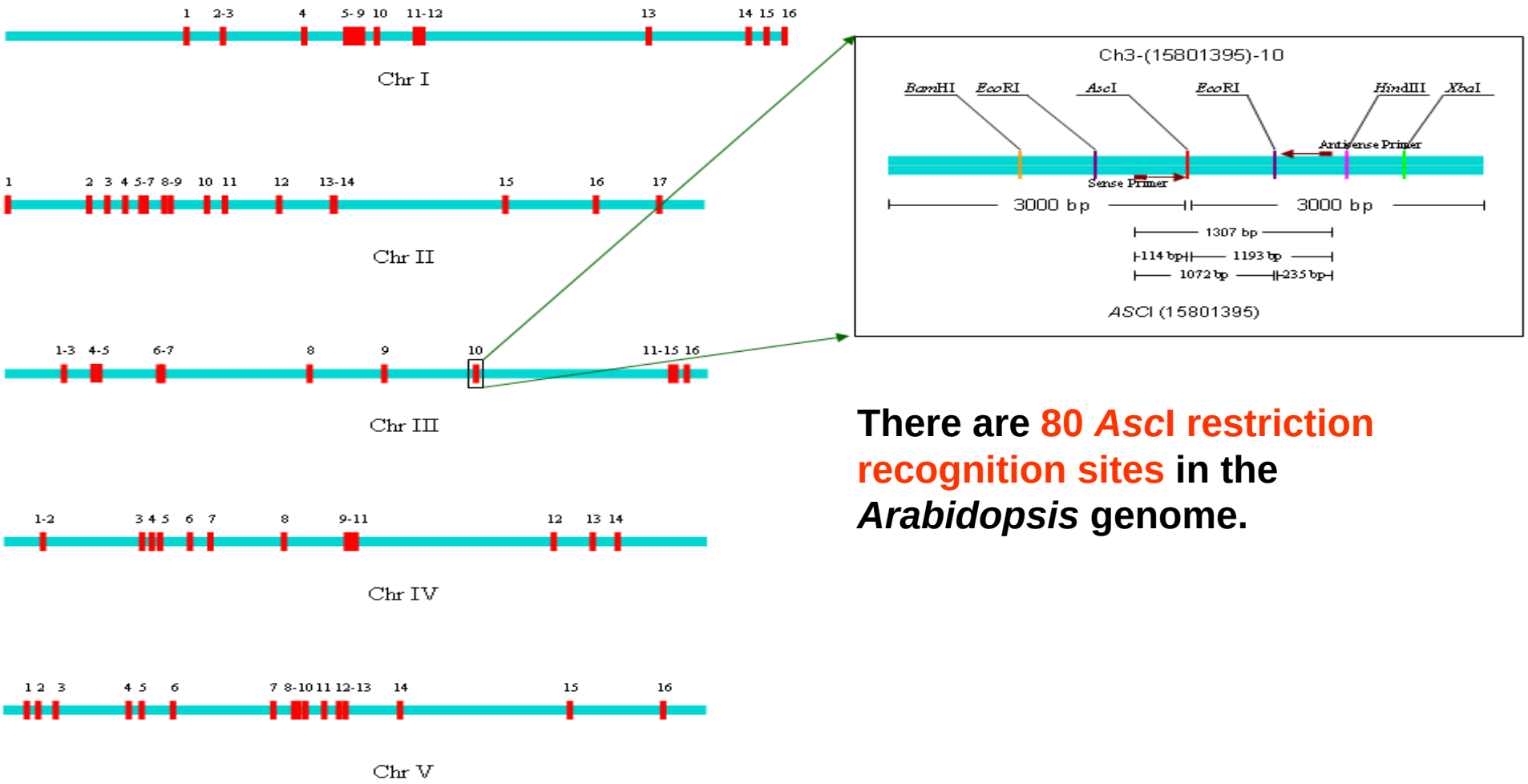
"High-frequency gene targeting in Arabidopsis plants expressing the yeast RAD54 gene."
(Shaked, 2005, PNAS)

T-DNA molecules integrate as ds molecules into DSBs

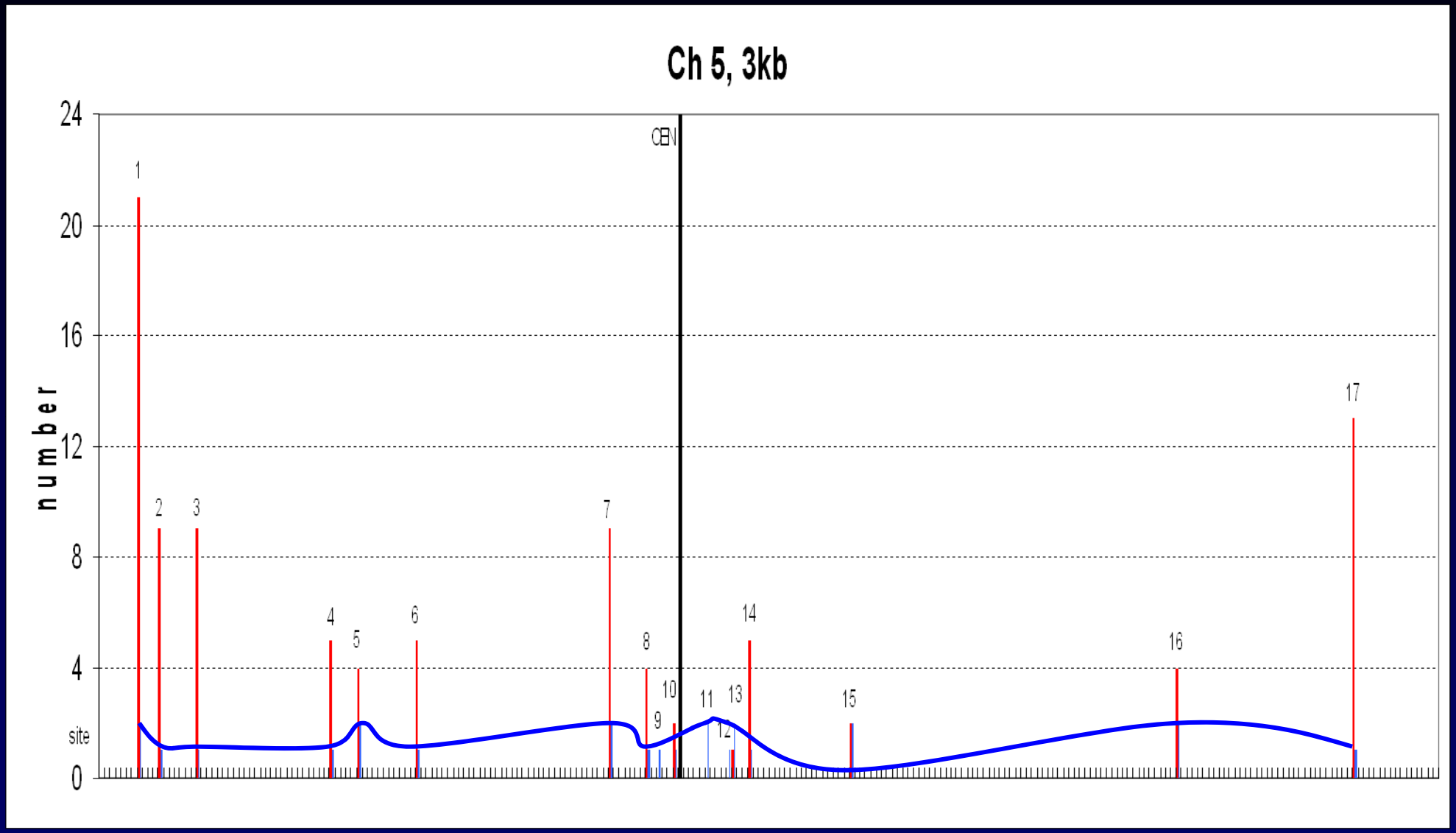



```
Plant DNA        5'tttggagaggacacgctcgacggtacctATTACCCTGTTATCCCTAggatccgtcgaagta
                   aaacctctcctgtgcgagctgccatggaTAATGGGACAATAGGGATcctaggcagcttcat

T-DNA            5'tggcaggatatattgtg..//..aattcATTACCCTGTTATCCCTAatgttacgtcctgta

Junction 332     5'tttggagaggacacgctcgacggtacctATTACCCTGTTATCCCTAatgttacgtcctgta
Junction 410     5'tttggagaggacacgctcgacggtacctATTACCCTGTTATCCCTAatgttacgtcctgta
Junction 522     5'tttggagaggacacgctcgacggtacctATTACCCTGTTATCCCTAatgttacgtcctgta
```

T-DNA insertion to DSBs is not limited to a single site



There are 80 *AscI* restriction recognition sites in the *Arabidopsis* genome.


Ch 5, 3kb
number
site
24
20
16
12
8
4
0
CEN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17

Infection of Arabidopsis plants with AscI expressing T-DNA resulted in integration into all AscI recognition sites.

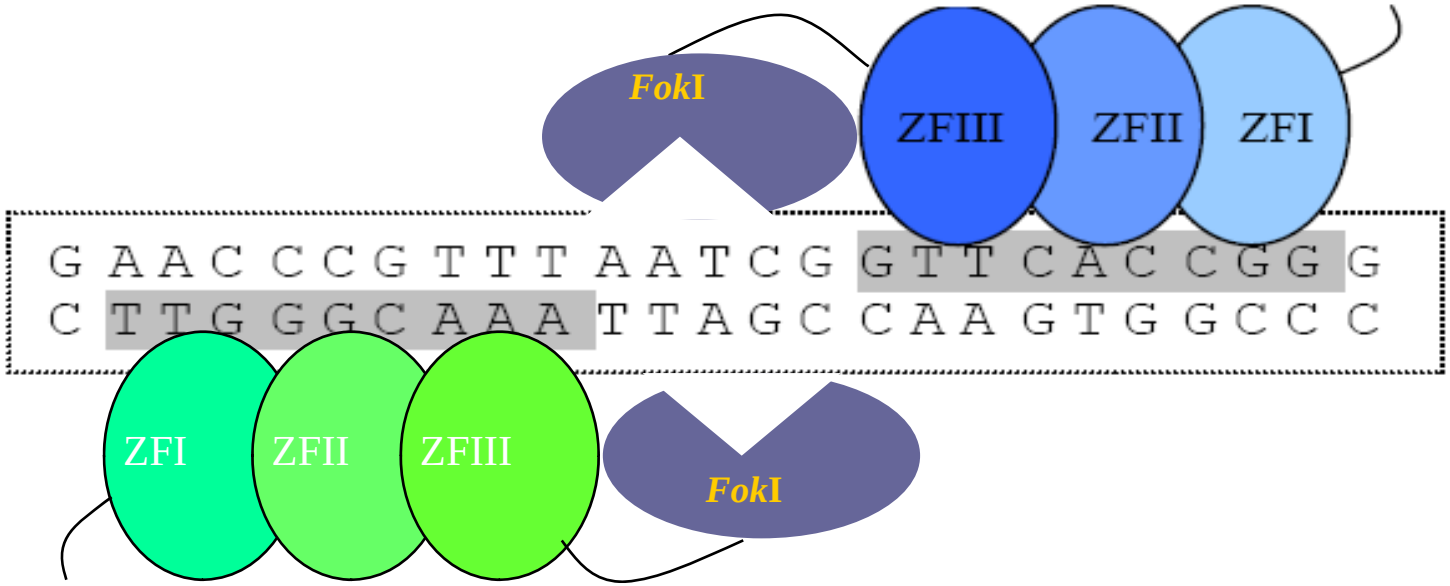
We analyzed approximately 940 insertions and found that ds T-DNA molecules can preferentially integrate into genomic DSBs.

T-DNA integration is not really a random process. The occurrence of DSBs is probably random.

The precise plant/T-DNA junctions indicates that T-DNA may integrate by a simple DNA repair.

But… engineering meganucleases for novel targets is tedious and nearly impossible.

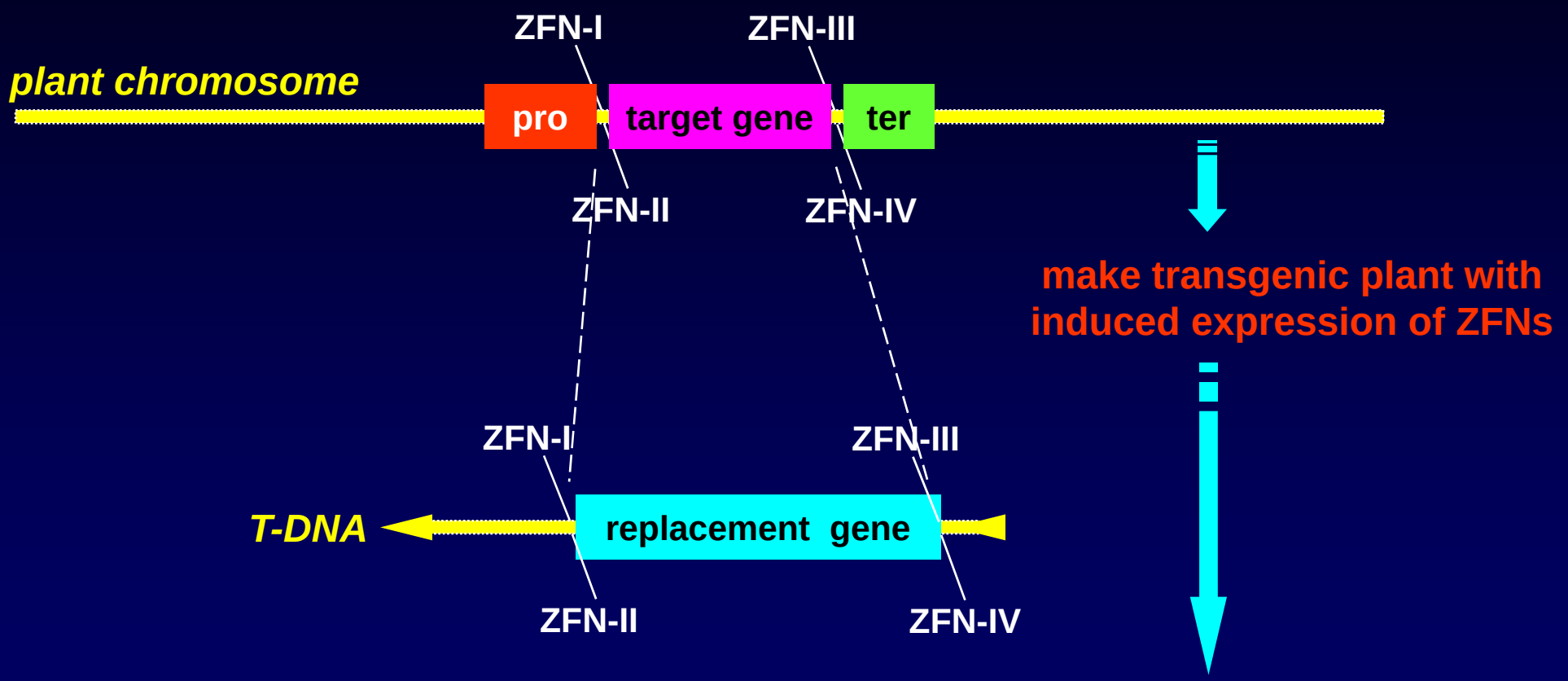
Zinc Finger Nucleases: a New Breed of Restriction Enzymes



ZFP + *Fok*I Endonuclease Domain = ZFN

“general” strategy for gene replacement/deletion in plant cells:

I. Express 4 ZNF for specific gene digestion *in planta*



II. replace native gene with partial T-DNA molecule

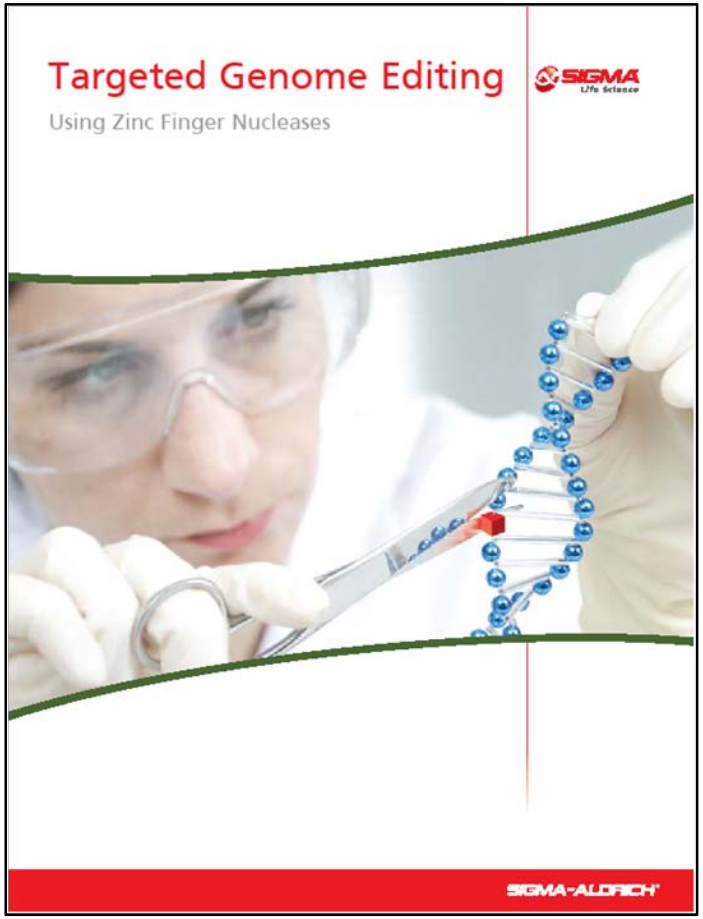
Challenges:

I. **Design 4 ZFN monomers for specific gene digestion *in planta***

From Sigma's web site…

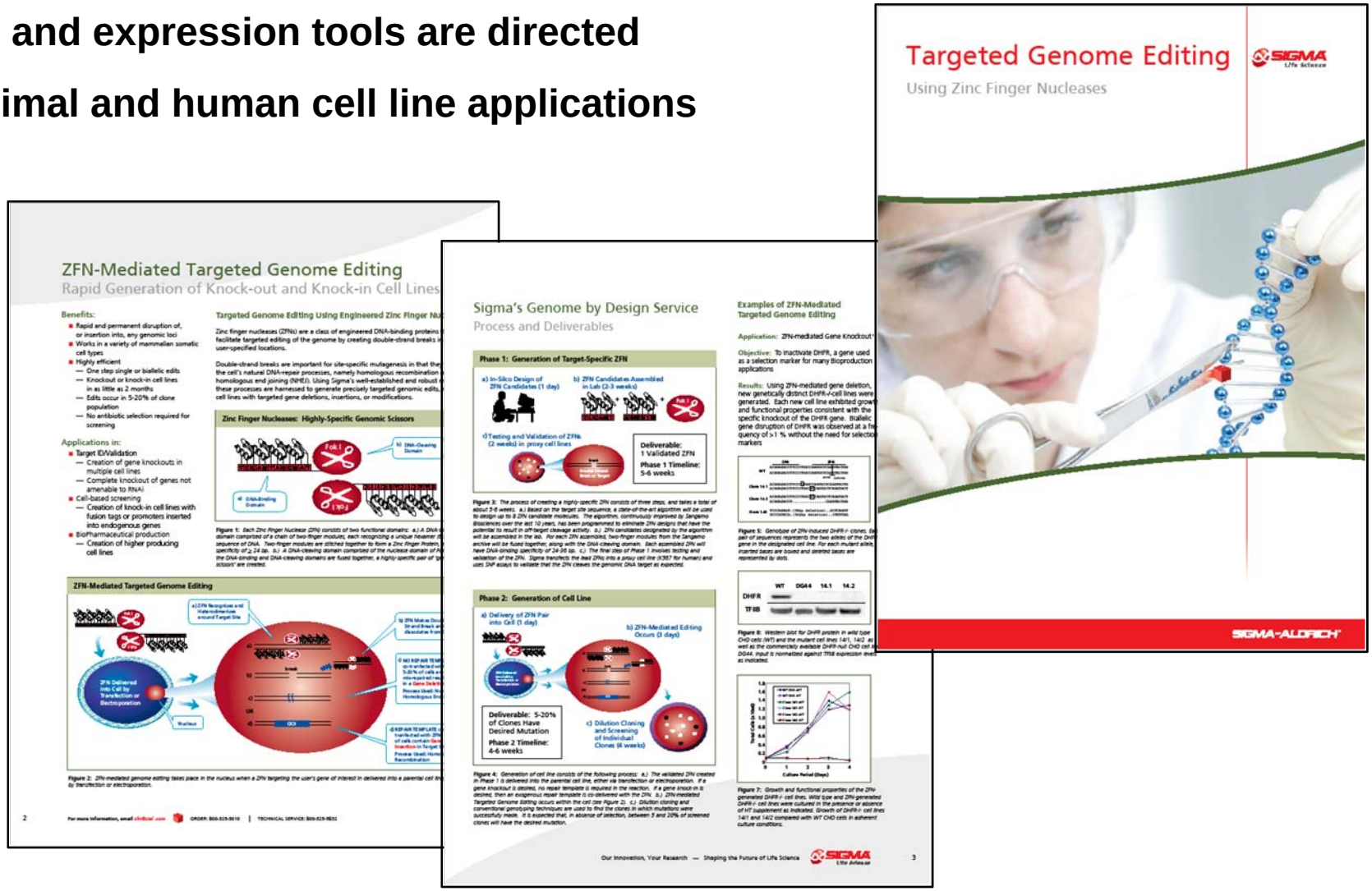
"Coming in the Fall of 2008 :

Sigma is pleased to announce our strategic relationship with Sangamo Biosciences, a pioneer in the development of Zinc Finger Nuclease (ZFN) technology, to become the exclusive supplier of Zinc Finger Nuclease-based reagents and services to the research community."



Great news for animal science…

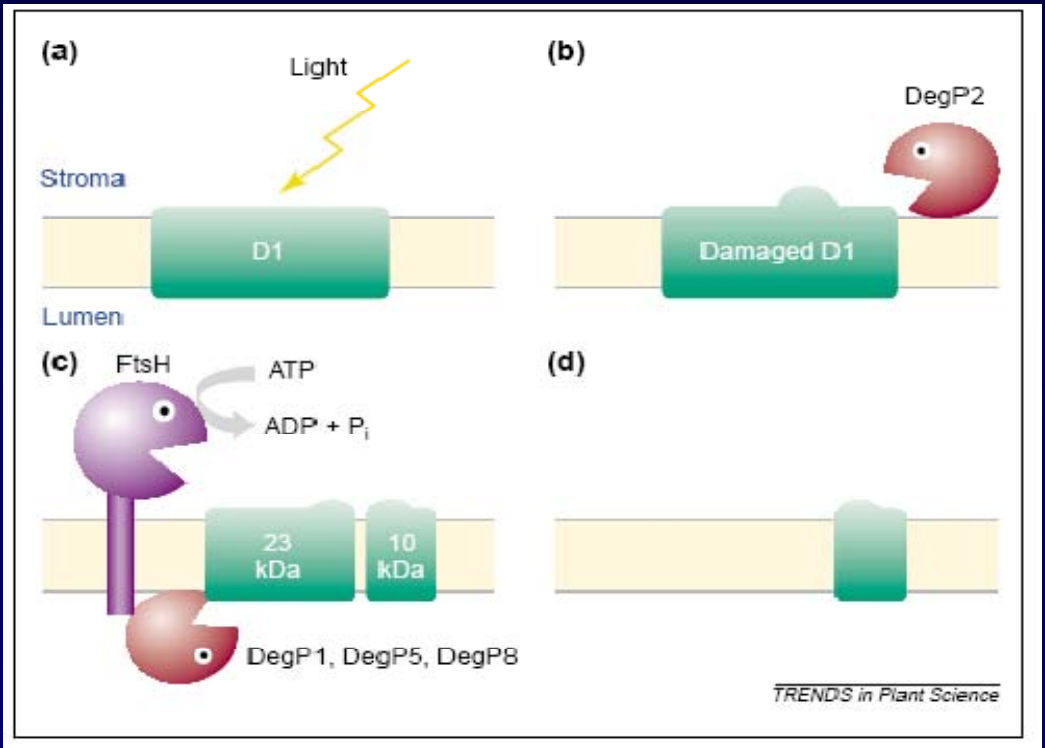
Validation and expression tools are directed toward animal and human cell line applications



Dedicated tools are still missing for plant research and biotechnology.

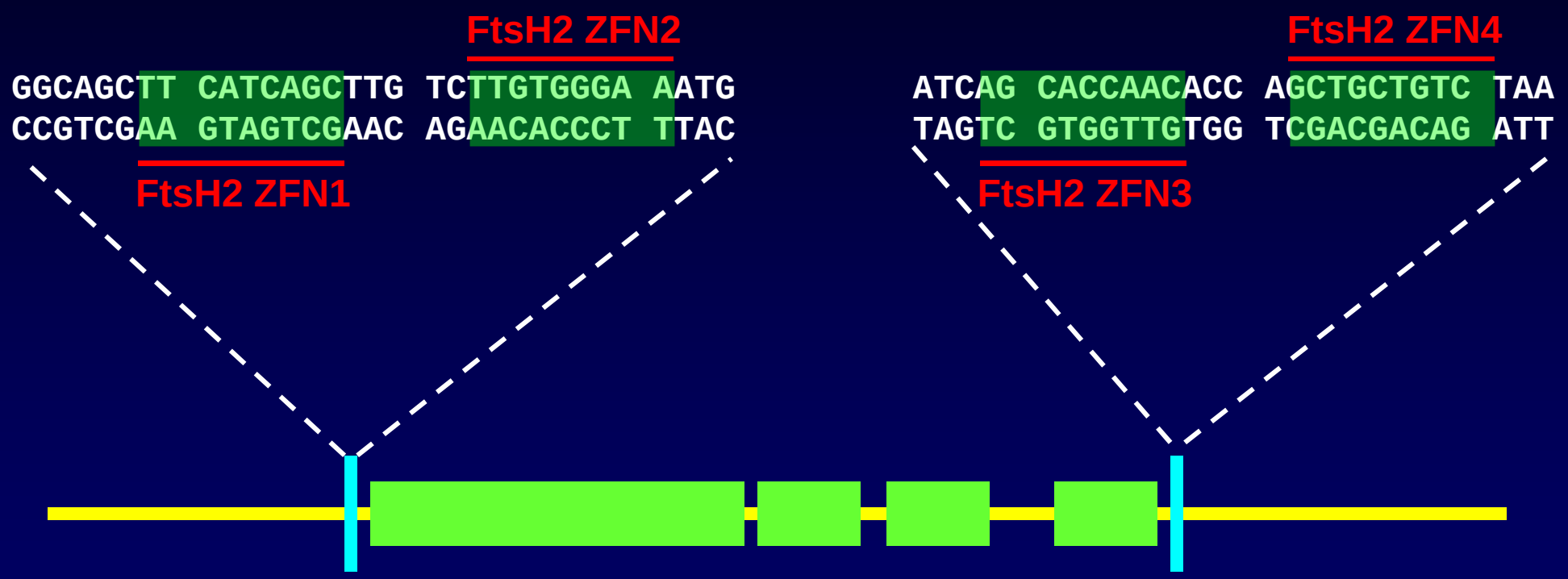
FtsH2 – At2g30950

- Encodes an ATP-dependent metalloprotease
- Involved in the turnover of the photosystem II D1 protein
- Mutants develop white leaf variegation phenotype



Adam and Clarke (2002) TiPS

Target Site Selection for AtFtsH2



At2g30950 FtsH2

4373 bp

I. Design 4 ZFN monomers for specific gene digestion *in planta*

Triplet Recognition Helices

DNA triplet	Recognition helix	DNA triplet	Recognition helix
	-1123456		-1123456
AAA	QRANLRA	GAA	QSSNLVR
AAC	DSGNLRV	GAC	DPGNLVR
AAGk	RKDNLKN	GAGk	RSDNLVR
AAT	TTGNLTV	GAT	TSGNLVR
ACA	SPADLTR	GCA	QSGDLRR
ACC	DKKDLTR	GCC	DCRDLAR
ACGk	RTDTLRD	GCGk	RSDDLVR
ACT	THLDLIR	GCT	TSGELVR
AGA	QLAHLRA	GGA	QRAHLER
AGC	N/A	GGC	DPGHLVR
AGGk	RSDHLTN	GGGk	RSDKLVR
AGT	HRTTLTN	GGT	TSGHLVR
ATA	QKSSLIA	GTA	QSSSLVR
ATC	N/A	GTC	DPGALVR
ATGk	RRDELNV	GTGk	RSDELVR
ATT	HKNALQN	GTT	TSGSLVR
TAGk	REDNLHT	TGGk	RSDHLTT

Andreas F. Kolb *TRENDS in Biotechnology*

TABLE ONE

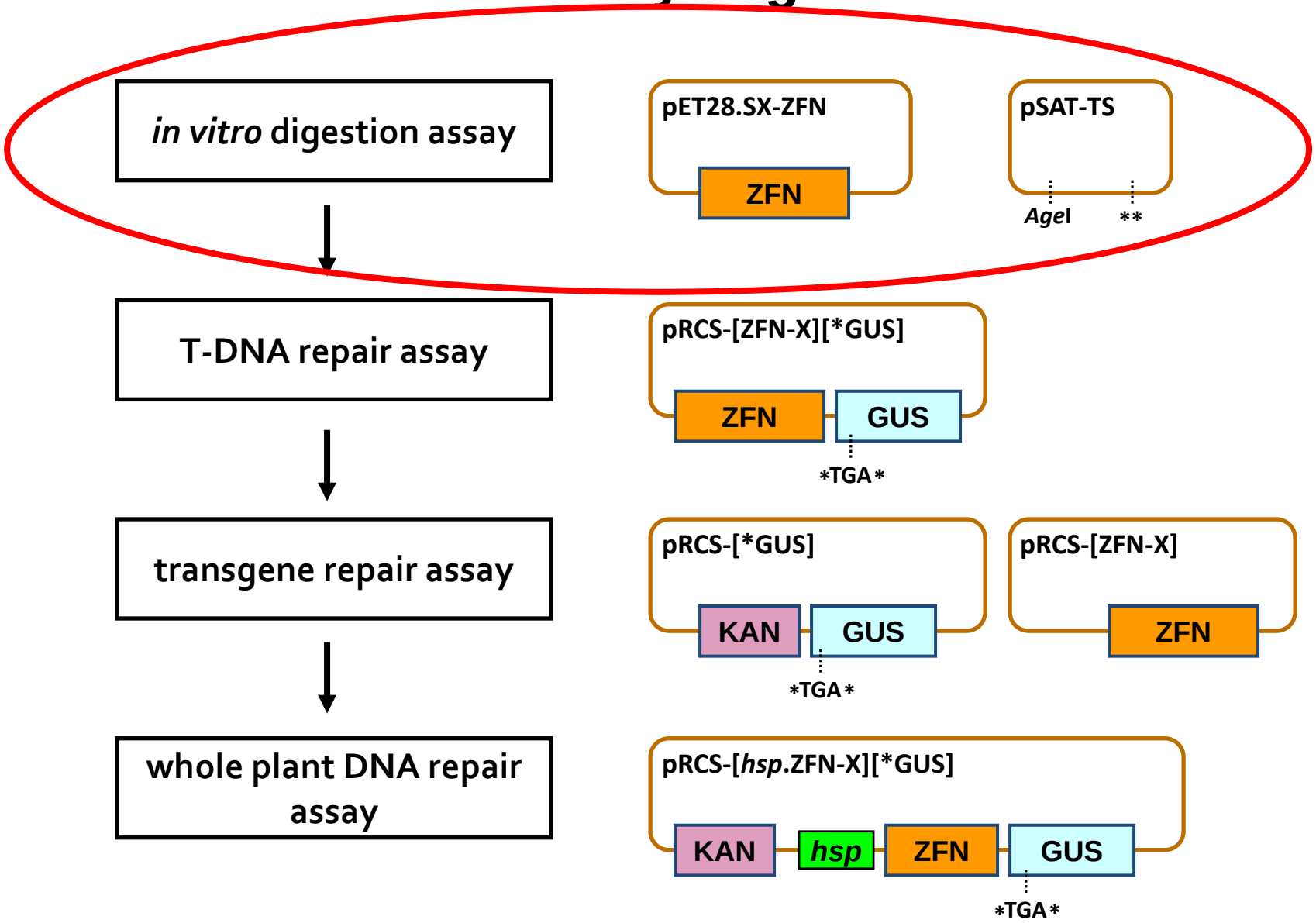
Summary of the most optimal zinc finger domain sequences identified in this study for binding to 5′-CNN-3′ subsites

The helices that bound to the 5′-CNN-3′ target sites are listed. P, helices derived from panning; M, helices derived by site-directed mutagenesis or *de novo* design.

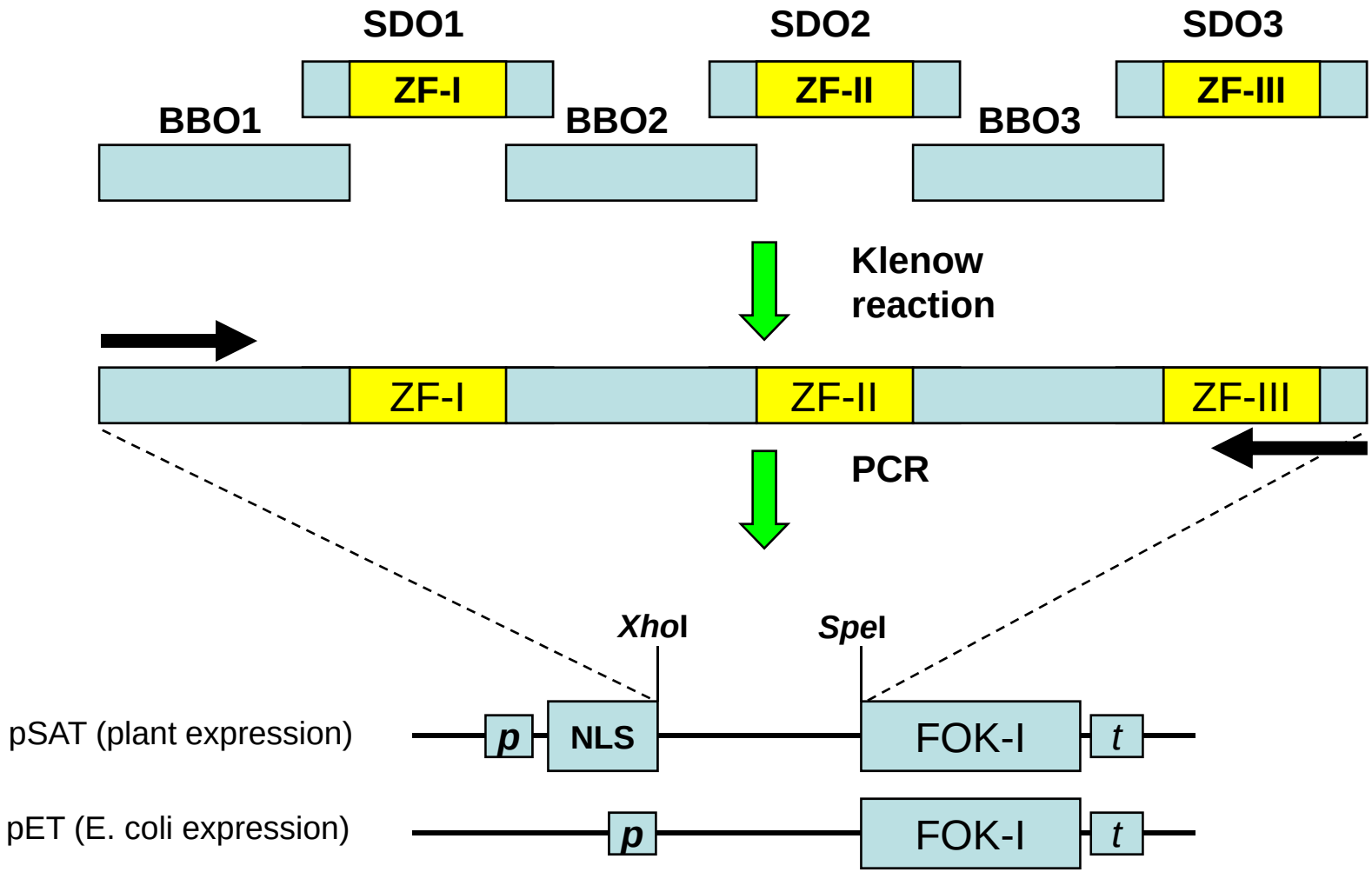
Target site 5′ → 3′	Finger-2 helix	Source	Fig. 2 panel
CAA	QSG-N-LTE	M	l
CAC	SKK-A-LTE	M	m
CAG	RAD-N-LTE	M	n
CAT	TSG-N-LTE	M	o
CCA	TSH-S-LTE	M	p
CCC	SKK-H-LAE	P	b
CCG	RND-T-LTE	M	q
CCT	TKN-S-LTE	M	r
CGA	QSG-H-LTE	M	s
CGC	HTG-H-LLE	P	f
CGG	RSD-K-LTE	M	t
CGT	SRR-T-CRA	P	h
CTA	QNS-T-LTE	M	u
CTC	None		
CTG	RND-A-LTE	P	k
CTT	TTG-A-LTE	M	v

Birgit Dreier *JOURNAL OF BIOLOGICAL CHEMISTRY*

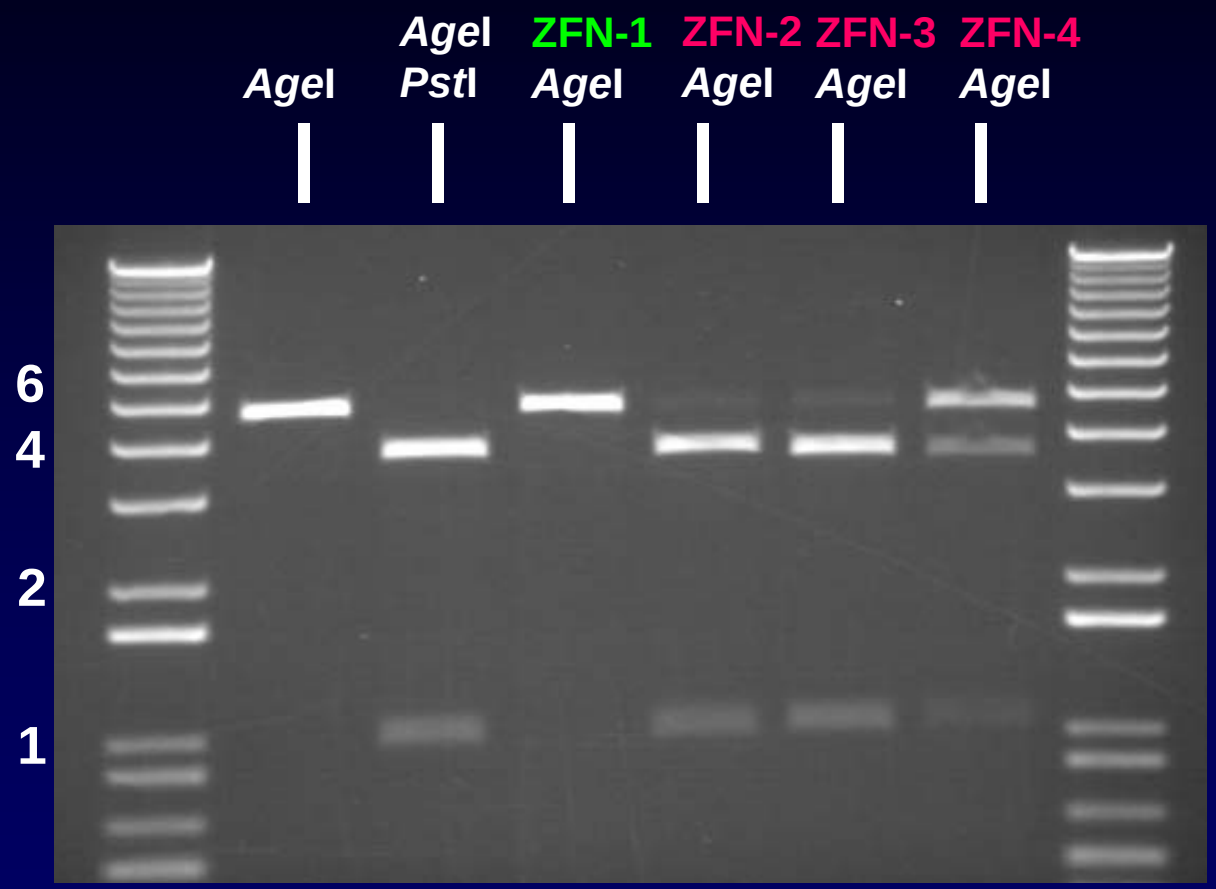
tools for analyzing novel ZFN

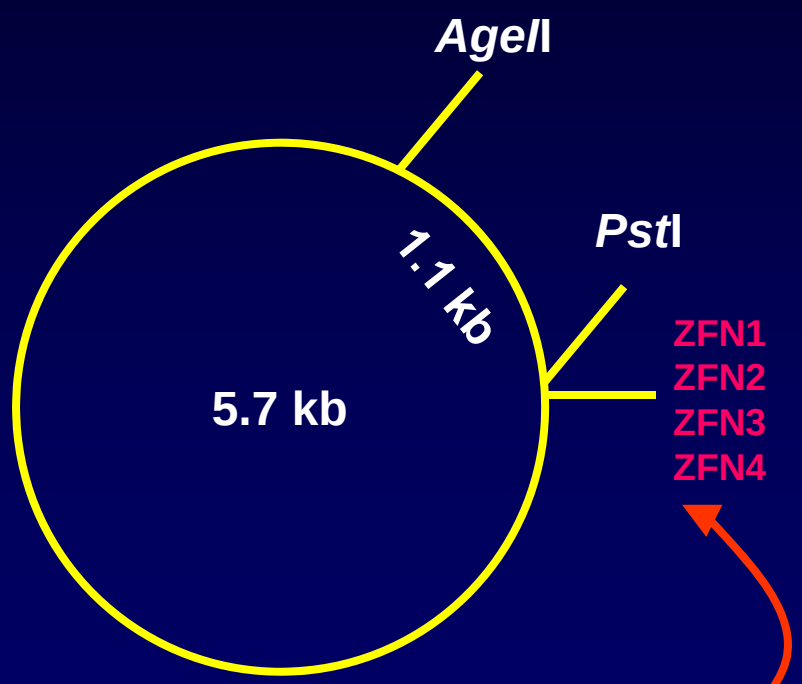


Fast and simple assembly of ZFs backbone



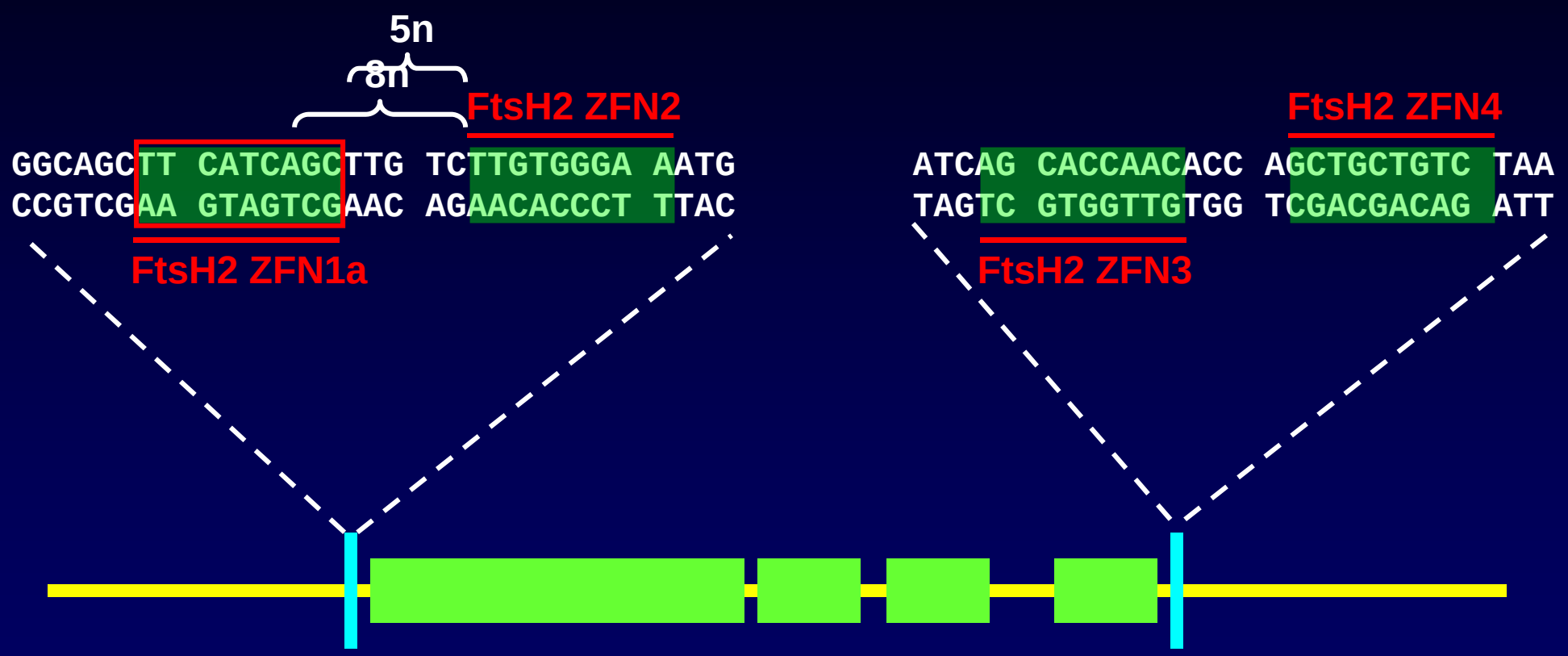
Activity of FtsH2 ZFNs





'palindrome'-like sequences

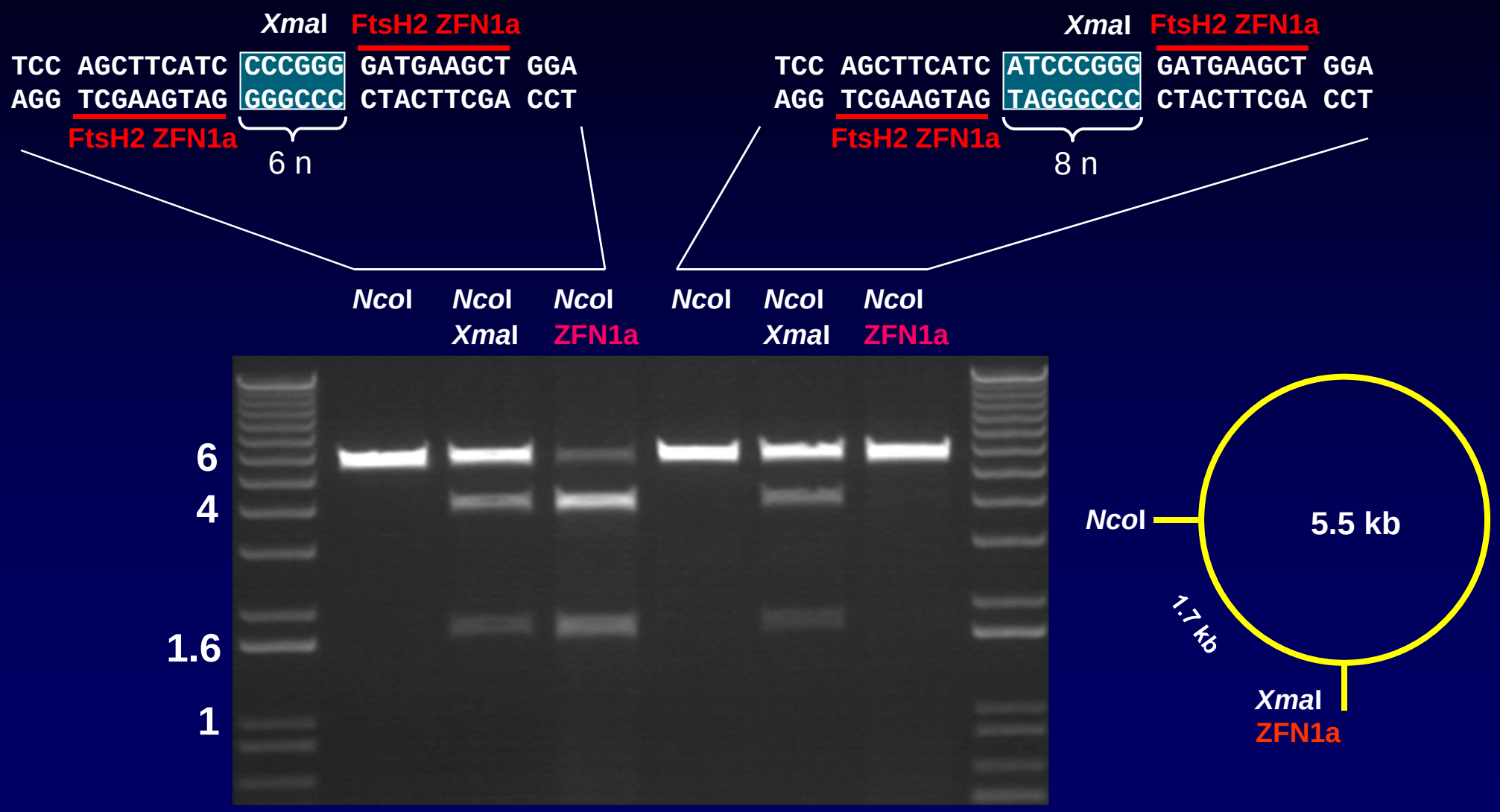
Target Site Selection for AtFtsH2



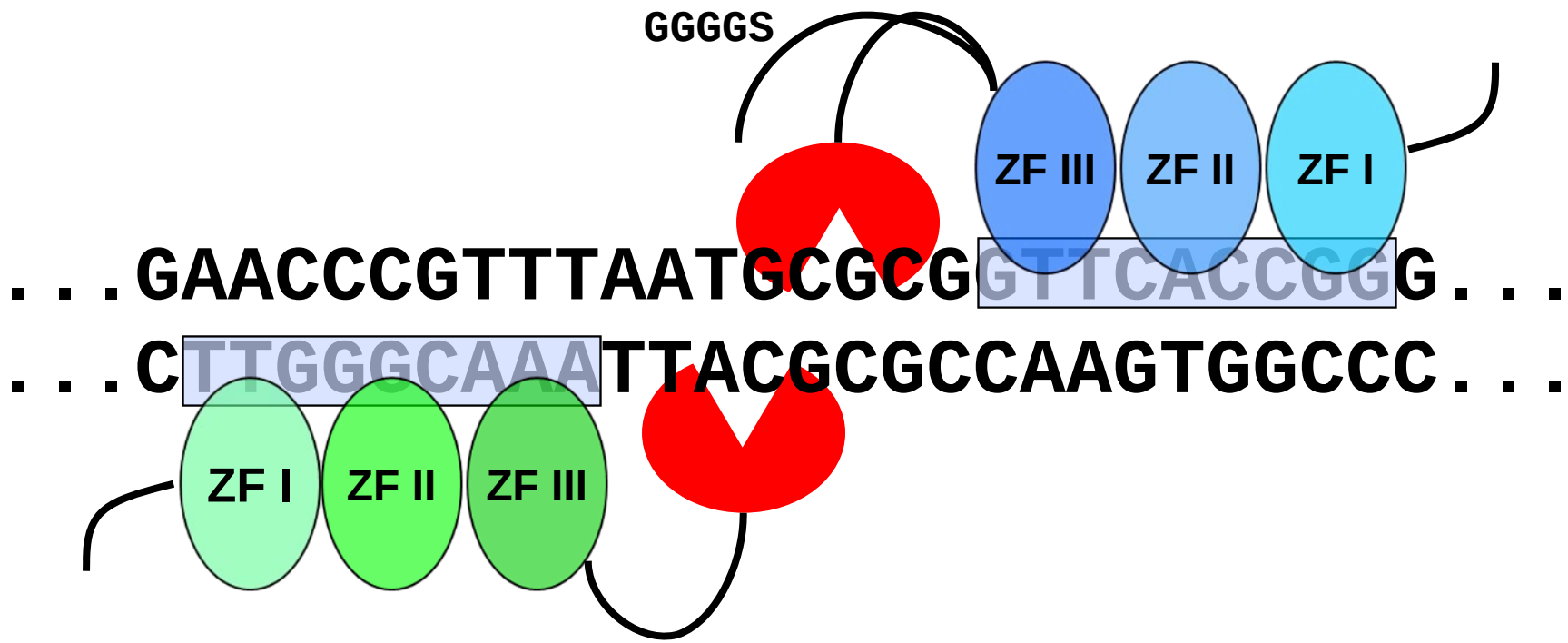
At2g30950 FtsH2

4373 bp

FtsH2 ZFN1a digests a target site with 6 but not 8 bp spacer



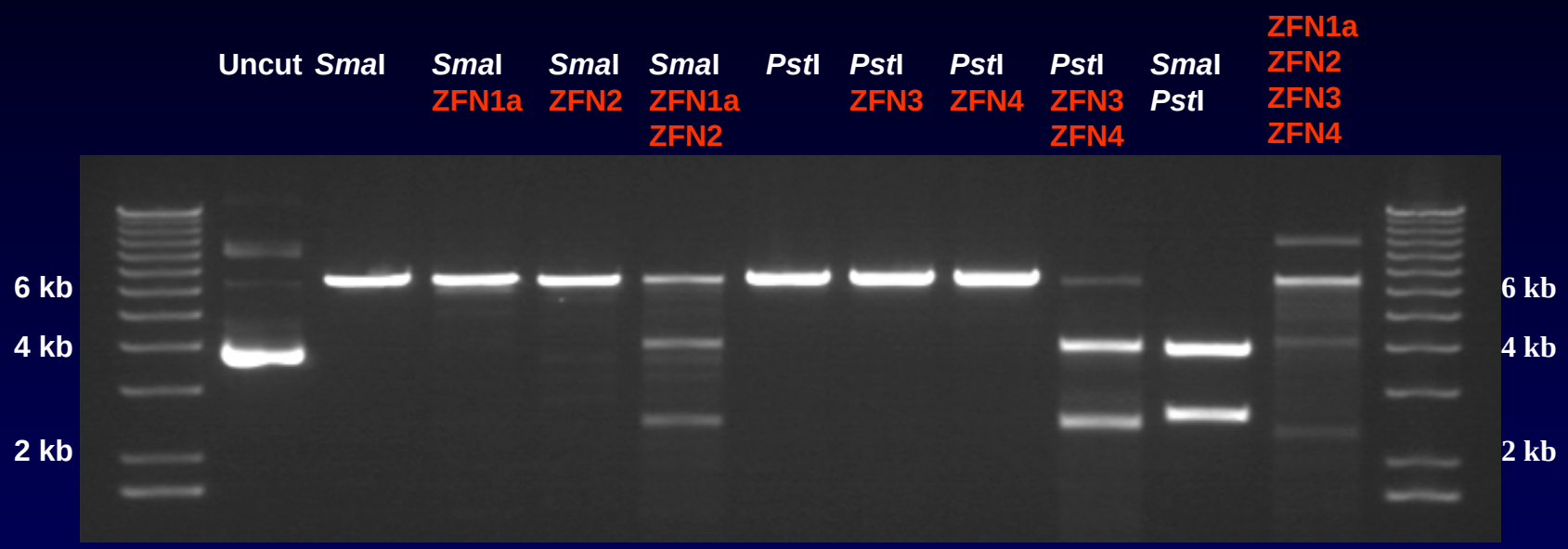
FtsH2 ZFN1a with $(Gly_4Ser)_2$ Linker

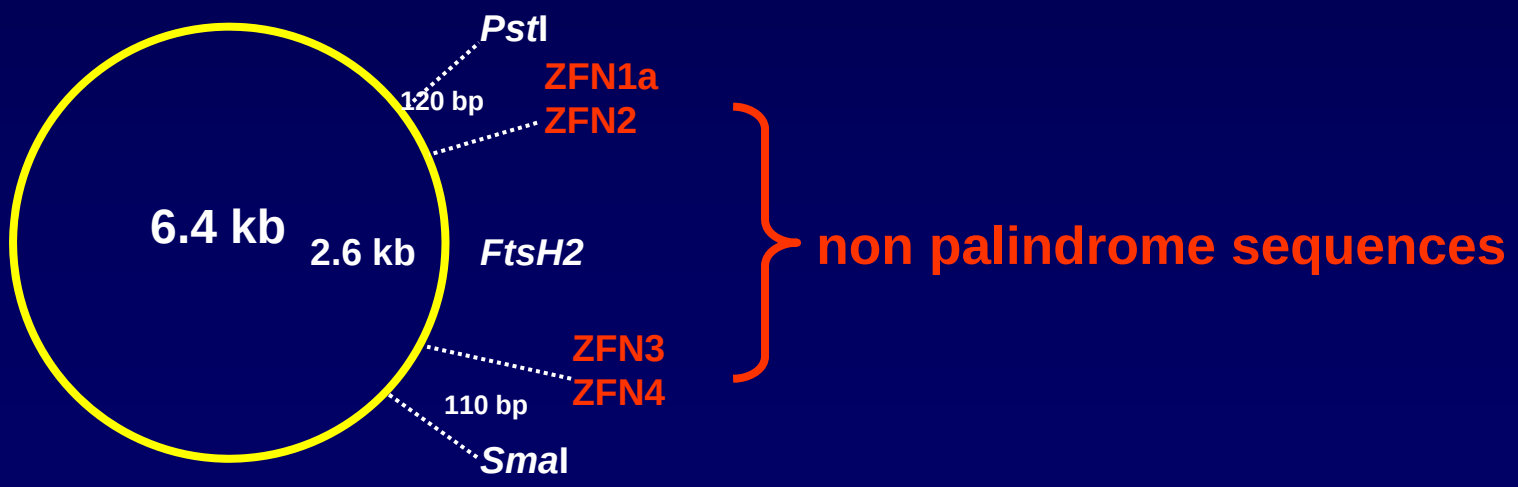


...GAACCCGTTTAA TGCGCGGTTCACCGGG...

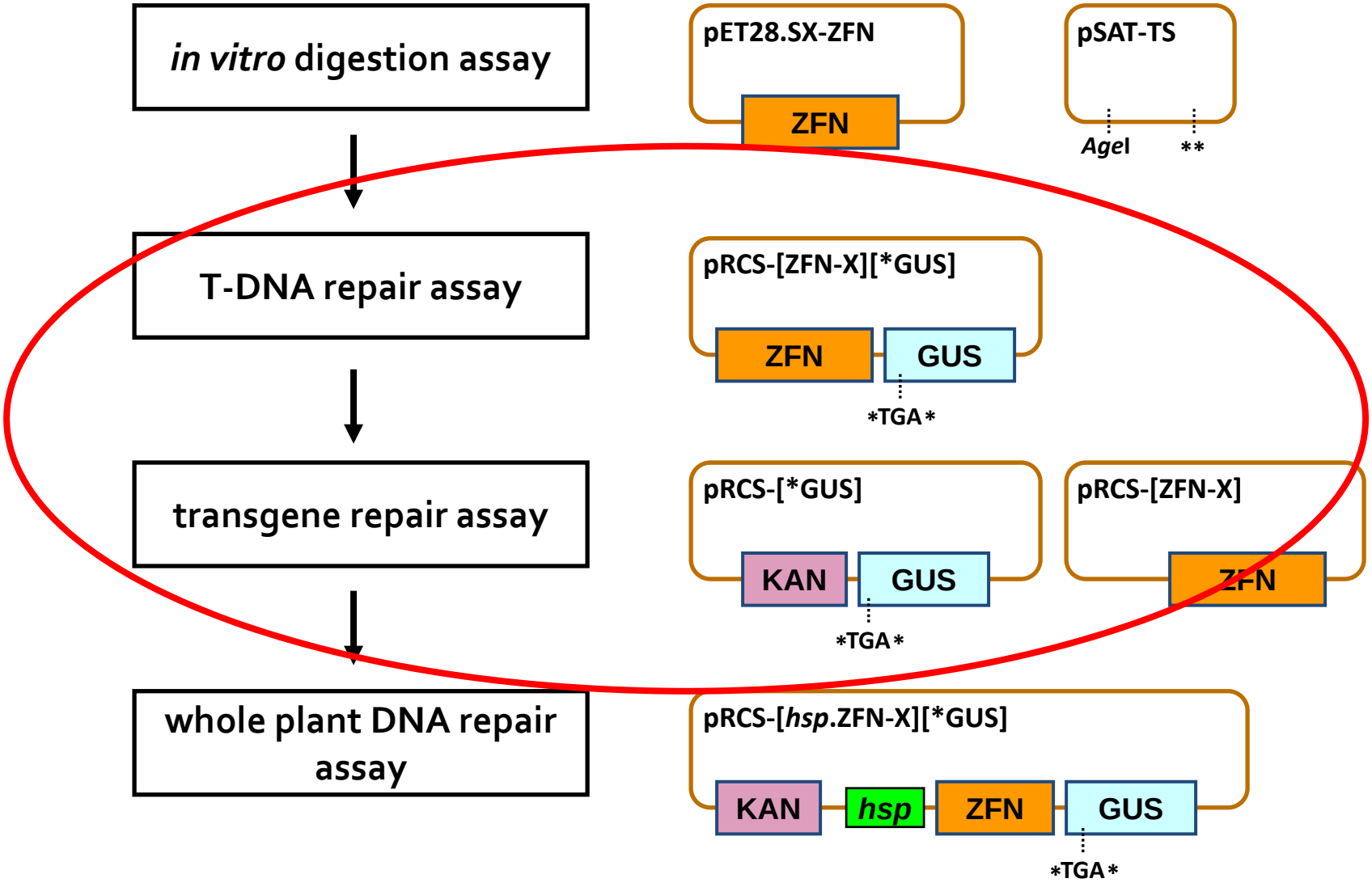
...CTTGGGCAAATTAC CGGCCAAGTGGCCC...

ZFN Pairs Digests Genomic FtsH2 Sequence



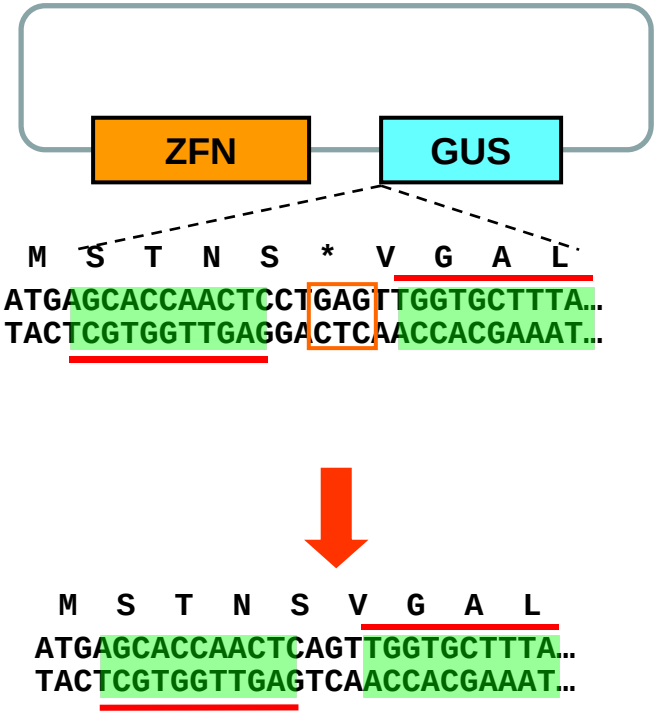


tools for analyzing of newly assembled ZFN

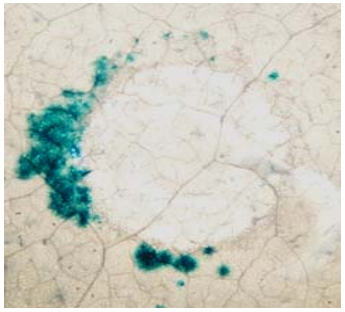


T-DNA repair assay

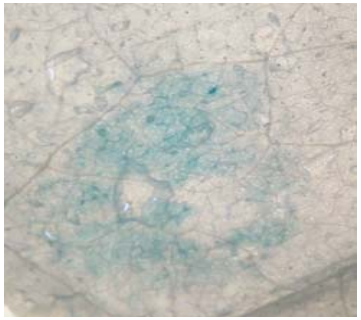
pRCS-[ZFN-X][*GUS]



- ZFN

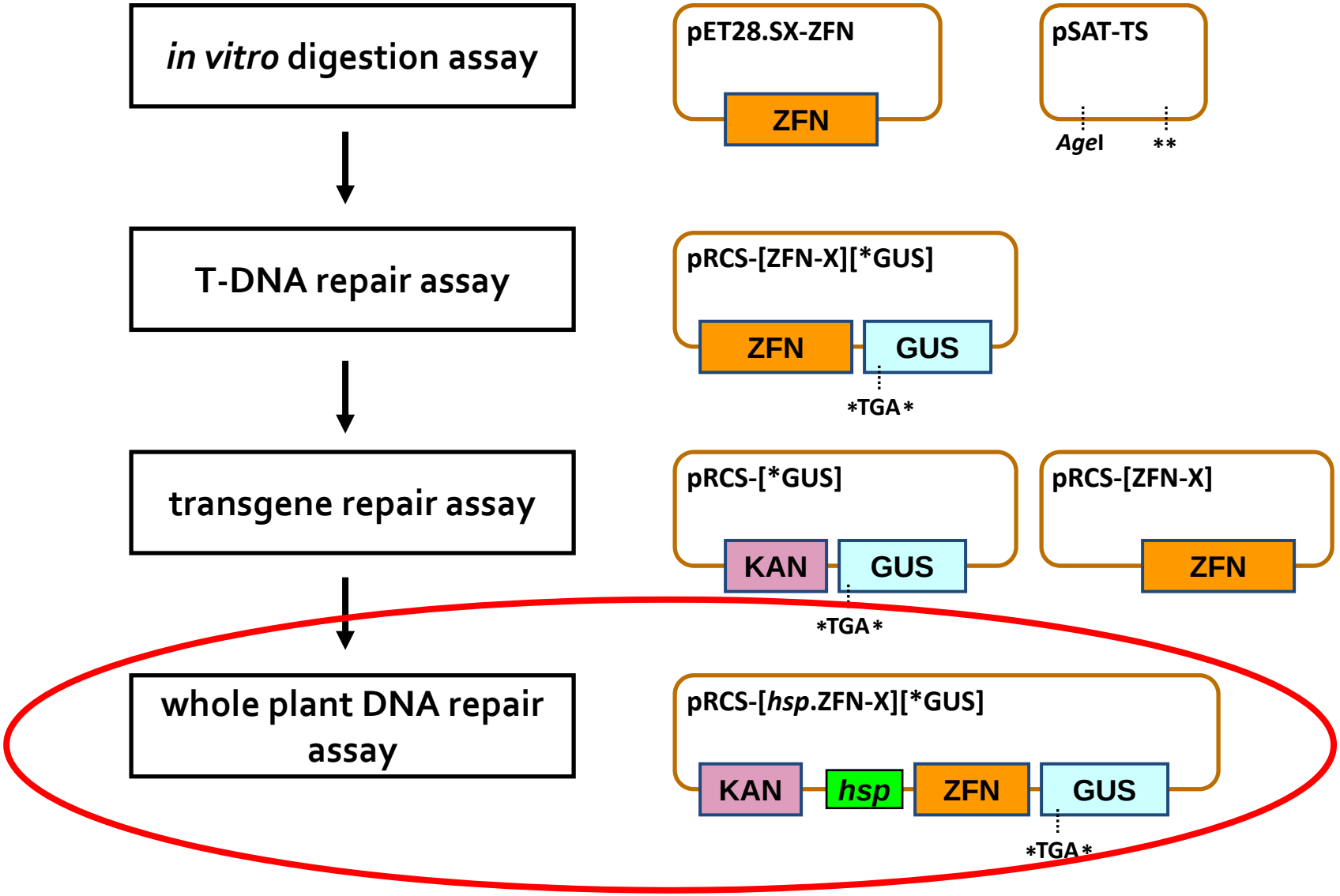


+ ZFN (QQR)

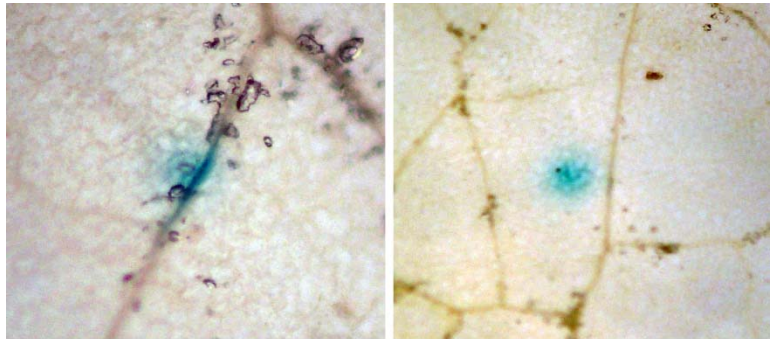


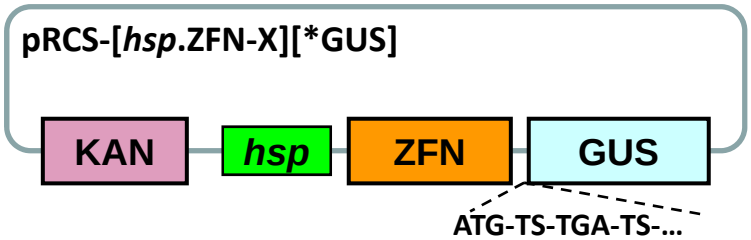
+ ZFN (ZFN3)

tools for analyzing of newly assembled ZFN



whole plant DNA repair assay






```
                    DdeI
ATGTTCTTCCCCTCCTGAGGGGAAGAATTA
ATGTTCTTCCCCTCCCGAGGGGAAGAATTA
ATGTTCTTCCCCTCCAGAGGGGAAGAATTA
ATGTTCTTCCCCTCCTGGGGGGAAGAATTA
ATGTTCTTCCCCTCTTGAGGGGAAGAATTA
ATGTTCTTCCCCTCCTGGAGGGGAAGAATTA
ATGTTCTTCCCCTCCTG--GGGAAGAATTA
ATGTTCTTCCC-----GAGGGGAAGAATTA
ACGAAC.....-82 bp.....TGTAGA
```

Challenges:

I. Design 4 ZFN monomers for specific gene digestion *in planta*

II. Validate ZFN monomers activity in model plants

III. Validate ZFN monomers activity in target plants

IV. Deliver 4 ZFN to target plants

V. Combine ZFN expression with second transformation cycle

VI. Develop methods for regenerating and selecting mutant lines

III. Validate ZFN monomers activity in target plants

Research conducted by Sangamo and several universities during the past 5-6 years resulted in significant progress in validating and using ZFN in various target species:

human cell lines

(gene correction, HIV-1 resistance, degradation of mutated human mitochondrial DNA, targeted gene edition

Zebra fish

(gene inactivation)

Drosophila

(targeting endogenous gene)

C. elegans

(targeting endogenous gene)

Chinese hamster

(targeting artificial gene)

III. Validate ZFN monomers activity in target plants

Similar effort should be made in plants if we wish to harness the power of ZFN for genome modification in crop plants, forest trees and other scientifically and biotechnologically important species.

Arabidopsis – mutagenesis of artificial sequence

Tobacco – inducing homologous recombination of transgene

canola – targeting of undisclosed native gene (DOW)

3-4 years before significant progress will be made with crop plants….

Challenges:

I. Design 4 ZFN monomers for specific gene digestion *in planta*

II. Validate ZFN monomers activity in model plants

III. Validate ZFN monomers activity in target plants

IV. Deliver 4 ZFN to target plants

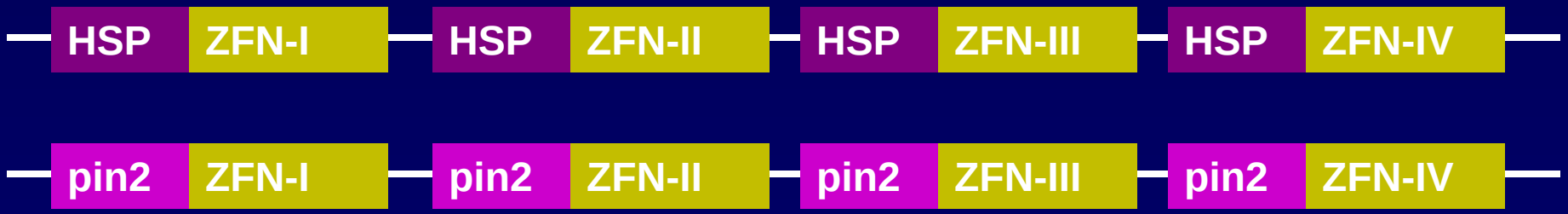
V. Combine ZFN expression with second transformation cycle

VI. Develop methods for regenerating and selecting mutant lines

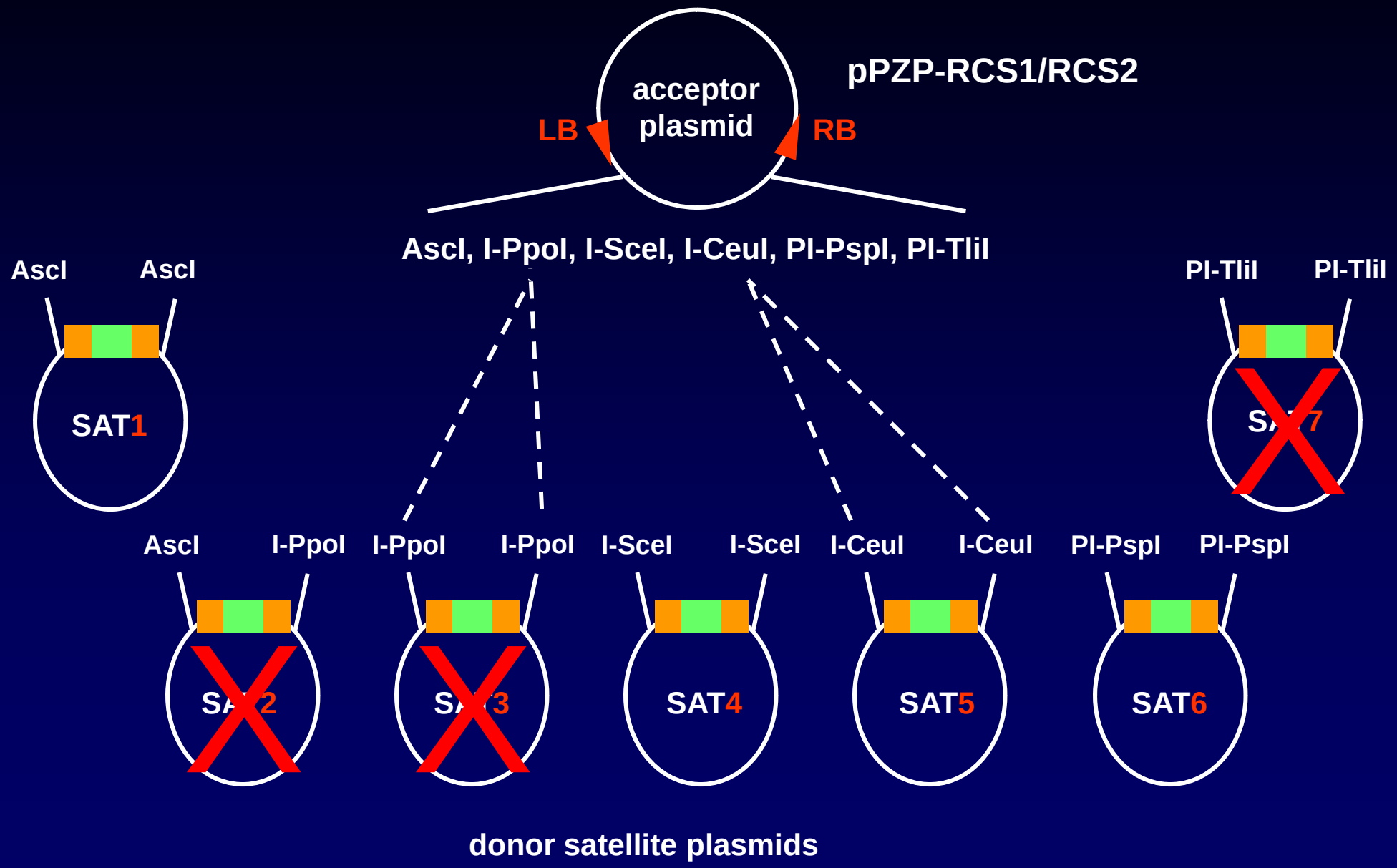
Multi gene delivery to plant cells is still challenging….

Expression of 4 ZFNs under Heat Shock (HS) or wound inducible (pin2) Promoters

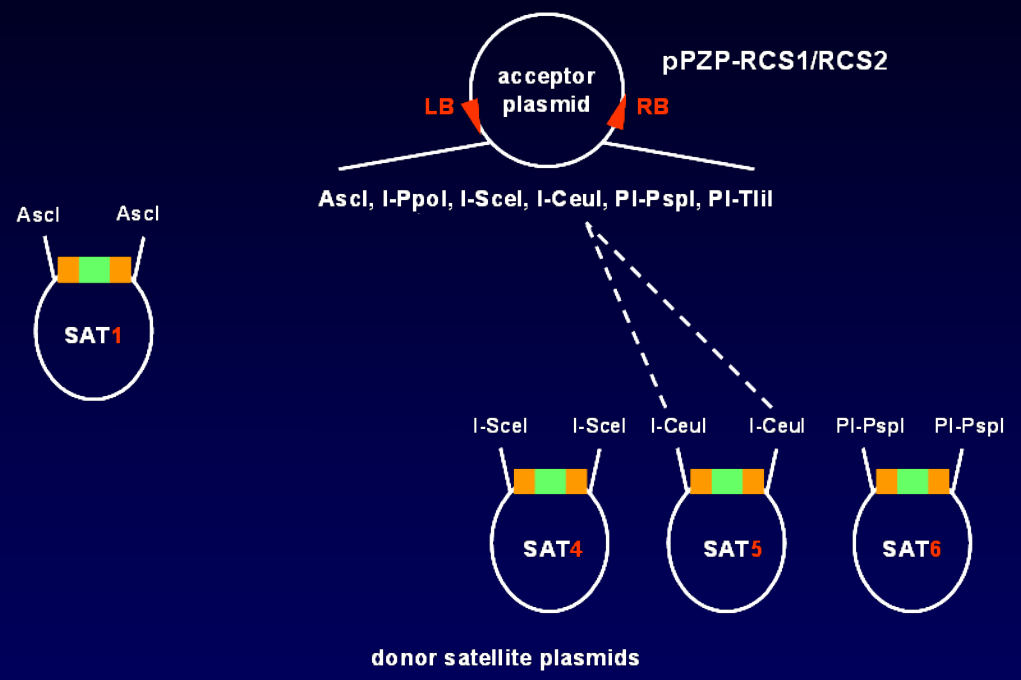
- **Introduce ZFNs under HSP and/or pin2**
- **Expose seedlings to 42ºC or wounding**
- **Combined with *Agrobacterium* infection to generate insertions/substitutions events**



Several expression cassettes can be mounted onto a single plasmid

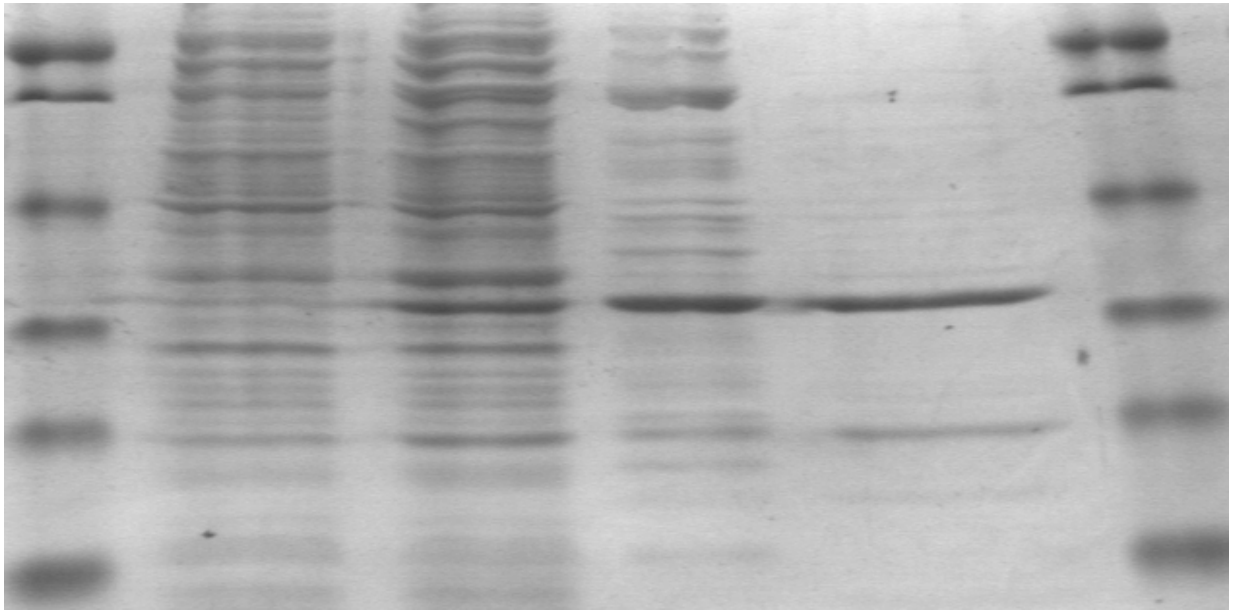


only four (useful !!) rare cutting enzymes are commercially available today



Can we use ZFNs for construction of plant transformation vectors ?

Purification of ZFN-IV Enzyme



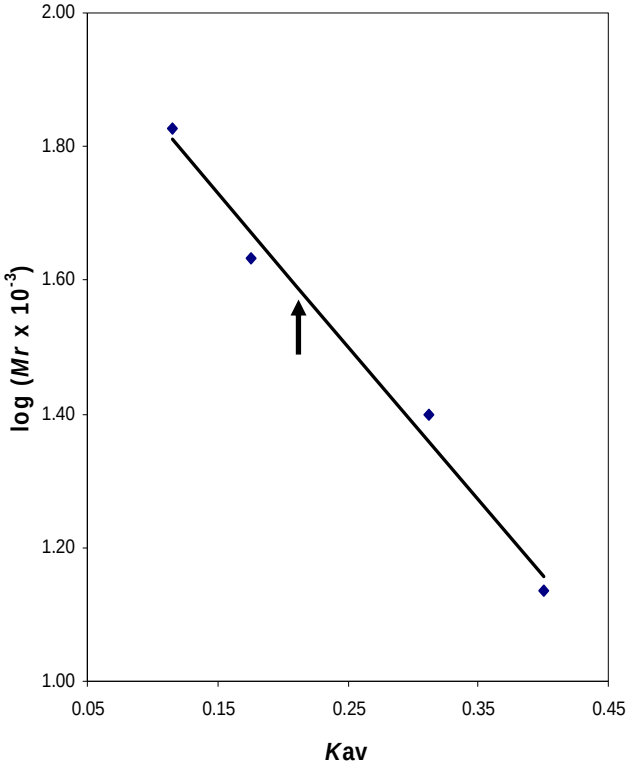
-IPTG

+IPTG

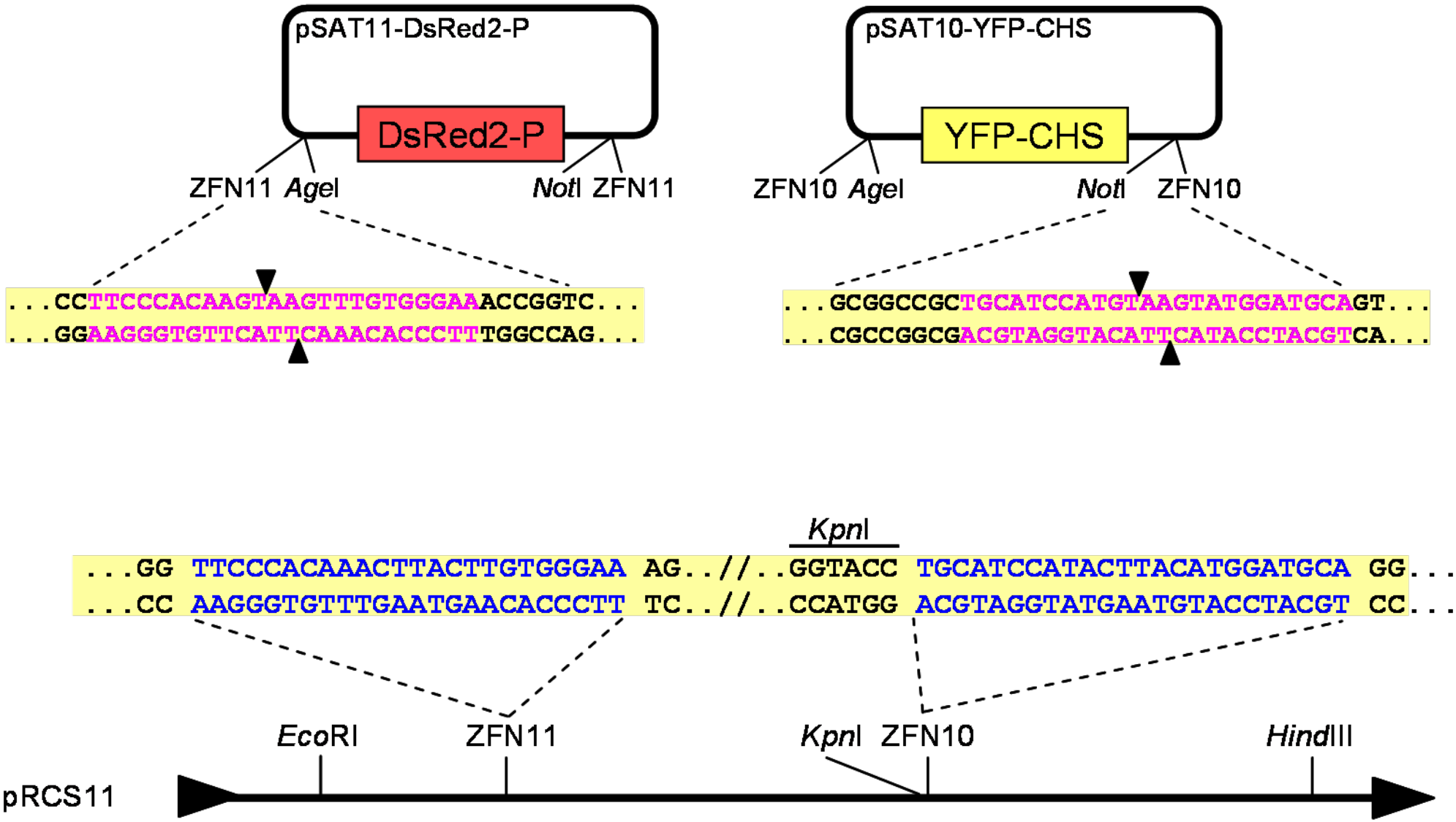
Lysate

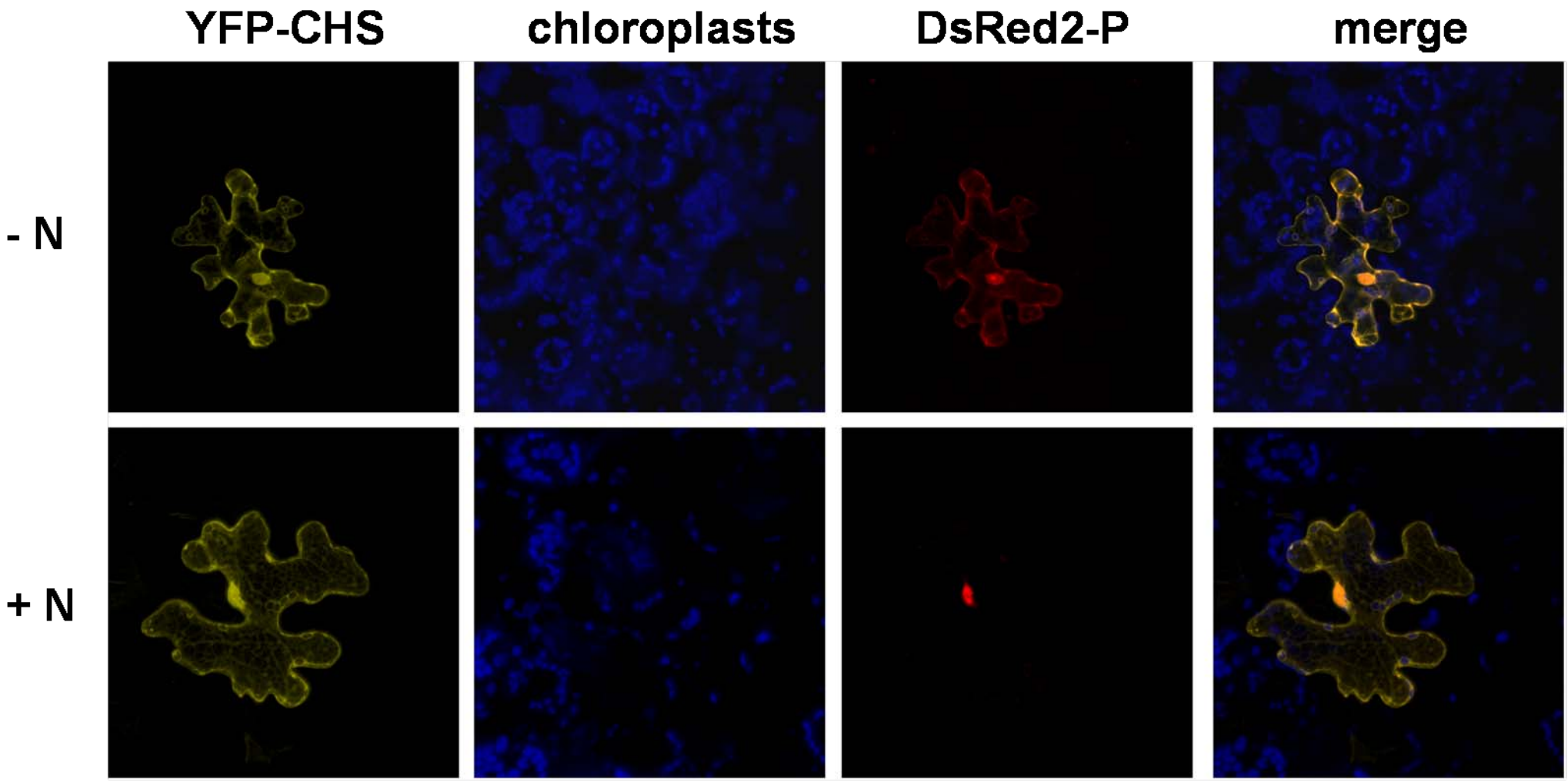
NiNTA

GF



Size-Exclusion Chromatography


pSAT11-DsRed2-P
DsRed2-P
ZFN11
AgeI
NotI
ZFN11
...CCTTCCCACAAGTAAGTTTGTGGGAAACCGGTC...
...GGAAGGGTGTTCATTCAAACACCCTTTGGCCAG...
pSAT10-YFP-CHS
YFP-CHS
ZFN10
AgeI
NotI
ZFN10
...GCGGCCGCTGCATCCATGTAAGTATGGATGCAGT...
...CGCCGGCGACGTAGGTACATTCATACCTACGTCA...
KpnI
...GG TTCCCACAAACTTACTTGTGGGAA AG..//..GGTACC TGCATCCATACTTACATGGATGCA GG...
...CC AAGGGTGTTTGAATGAACACCCTT TC..//..CCATGG ACGTAGGTATGAATGTACCTACGT CC...
EcoRI
ZFN11
KpnI
ZFN10
HindIII
pRCS11


YFP-CHS
chloroplasts
DsRed2-P
merge
- N
+ N

ZFN11 *Not*I *Age*I ZFN11

```
...GG TTCCCACAA ACTTAC TTGTGGGAA GCGGCCGC..//..ACCGGT TTCCCACAA ACTTAC TTGTGGGAA AG...
...CC AAGGGTGTT TGAATG AACACCCTT CGCCGGCG..//..TGGCCA AAGGGTGTT TGAATG AACACCCTT TC...
```

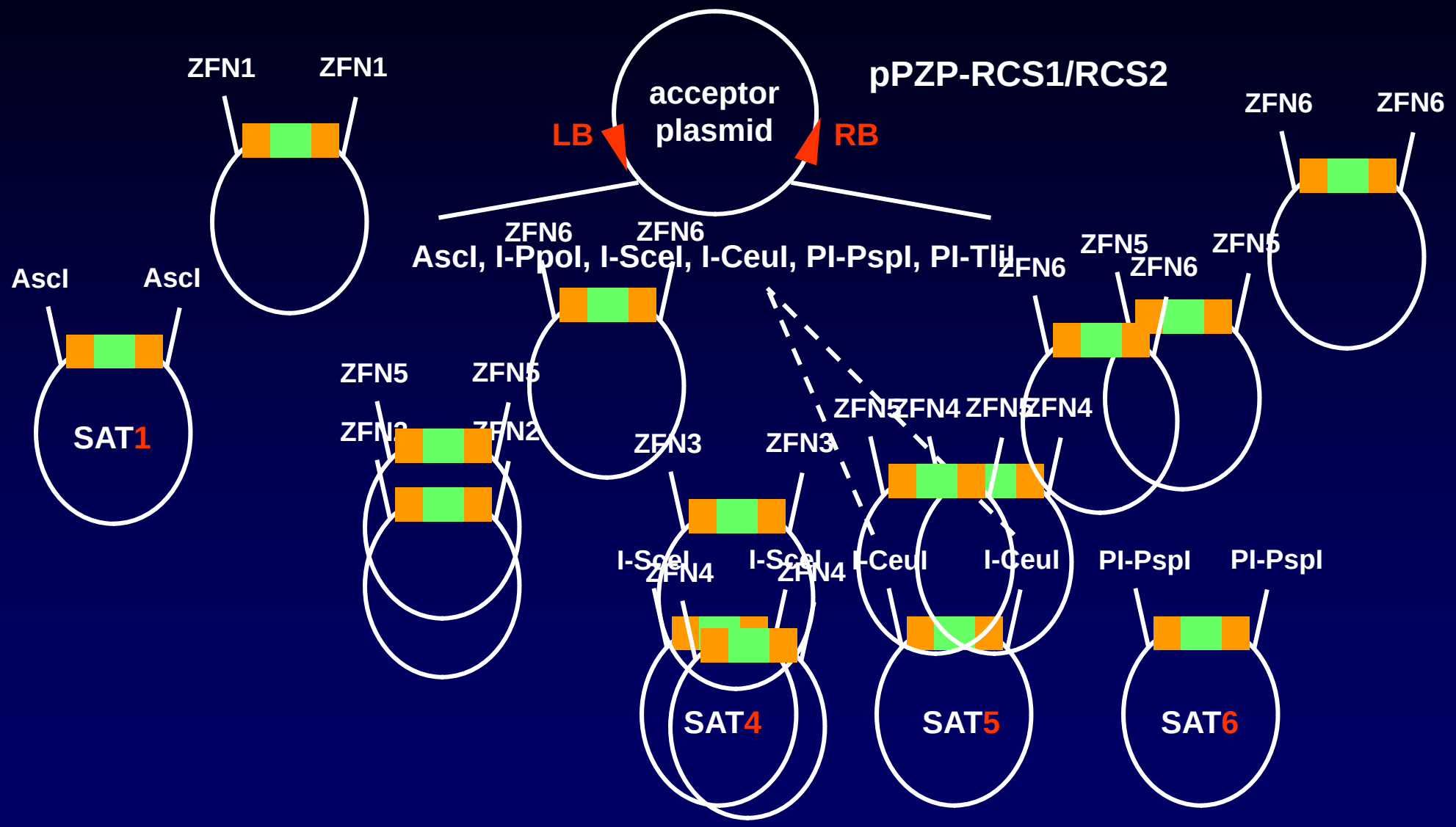
DsRed2-P expression cassette

*Kpn*I ZFN10 *Not*I *Age*I ZFN10

```
C GGTACC TGCATCCAT ACTTAC ATGGATGCA GCGGCCGC..//..ACCGGT TGCATCCAT ACTTAC ATGGATGCA GG...
C CCATGG ACGTAGGTA TGAATG TACCTACGT CGCCGGCG..//..TGGCCA ACGTAGGTA TGAATG TACCTACGT CC...
```

YFP-CHS expression cassette

ZFN can also be use for cloning of expression cassettes



Challenges:

I. Design 4 ZFN monomers for specific gene digestion *in planta*

II. Validate ZFN monomers activity in model plants

III. Validate ZFN monomers activity in target plants

IV. Deliver 4 ZFN to target plants

V. Combine ZFN expression with second transformation cycle

VI. Develop methods for regenerating and selecting mutant lines

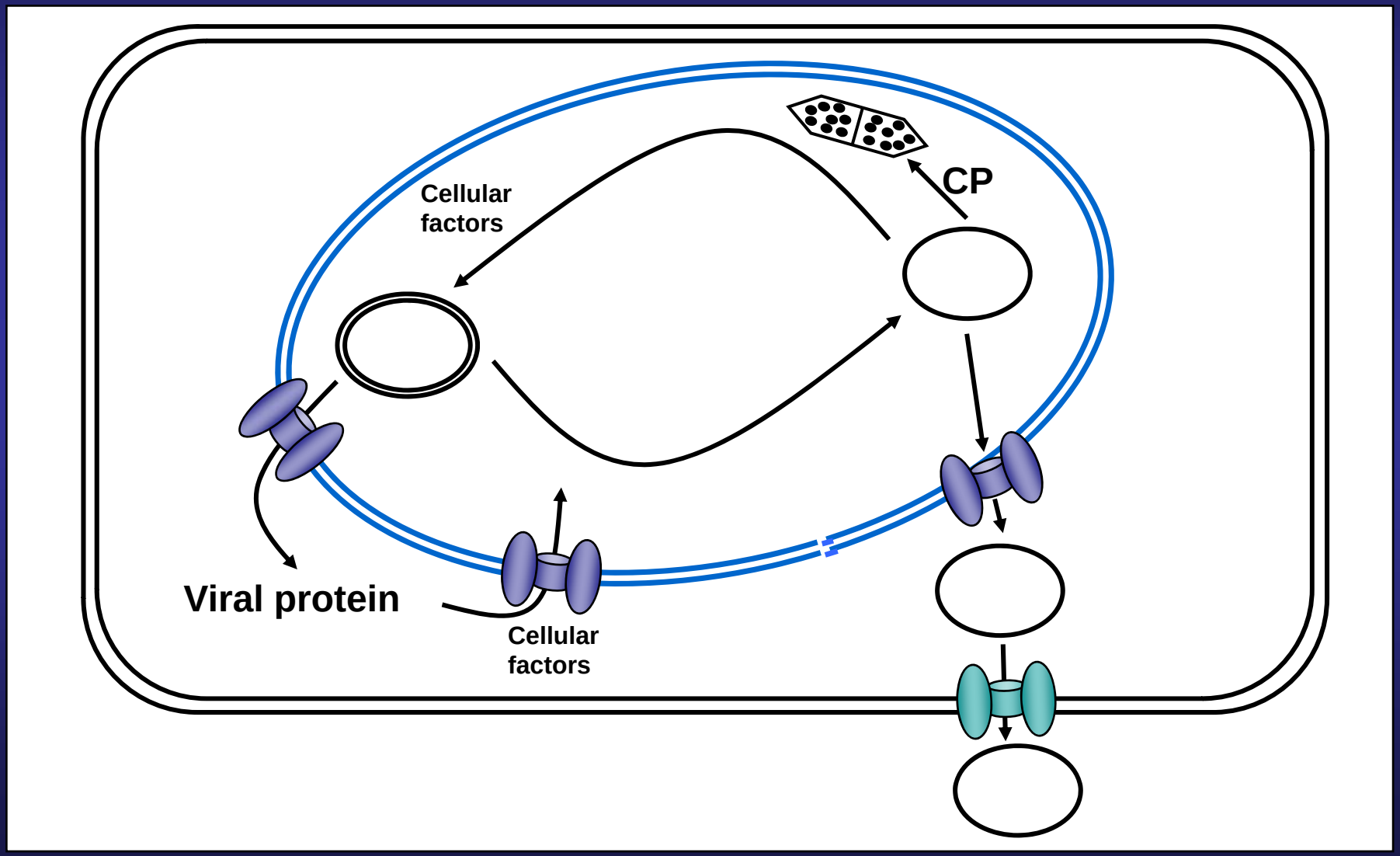
Protocols need to be developed for various species and different transformation methods….

Zinc-finger-nuclease-mediated resistance to plant viruses

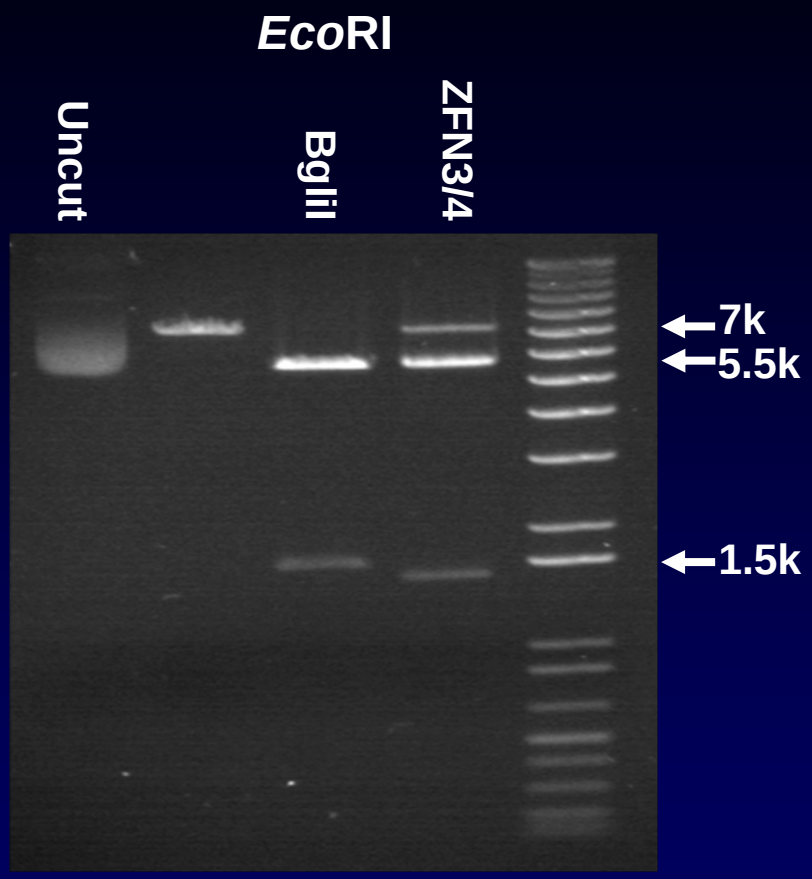
TYLCV infects tomato

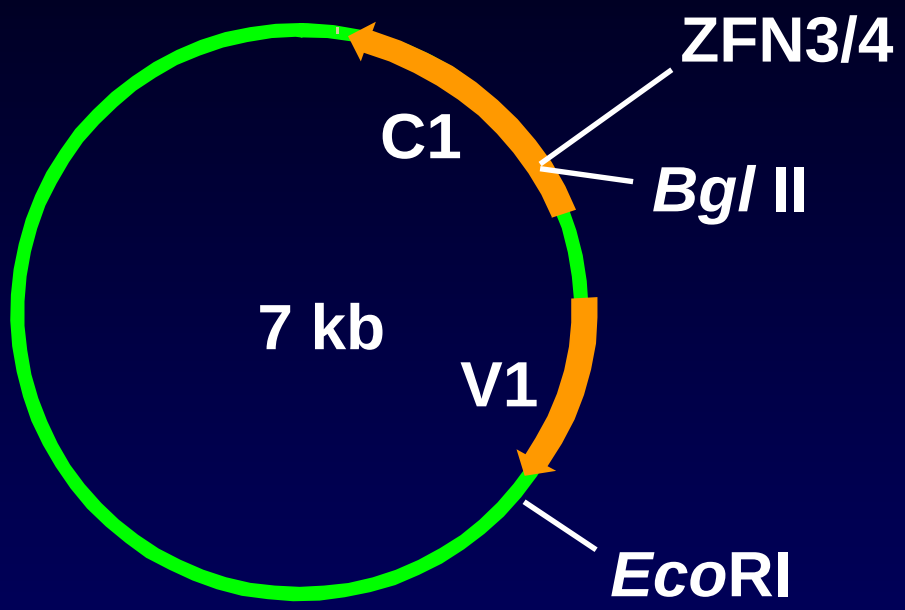
BDMV infects Bean plant

Geminivirus DNA replication cycle and movement from cell-to-cell



in-vitro digestion of BDMV- DNA B clone





ZF4

Bgl II

```
5'  ATC GAC GGC AGATCTG CTA GAG GAG
3'  TAG CTC CCG TCTAGAC GAT CTC CTC
```

ZF3

transgenic

Wild type

Possible broad range resistance to TYLCV

```
2-5rep              CGGTTCTTCGACCTGGTATCCCCAAACAGGTCAGCACATT    334
Rep[FL]             CGGTTCTTCGACCTGGTATCCCCAAACAGGTCAGCACATT    258
TYLCVX15656RC       CGGTTCTTCGACtTGGTATCCCCAAACAGGTCAGCACATT    253
FloridaAY530931R    CGGTTCTTCGACCTGGTATCCCCAAACAGGTCAGCACATT    419
Puerto_RicoAY134    CGaTTCTTCGACCTGGTATCCCCAAACAGGTCAGCACATT    419
GeziraAY044138RC    CGaTTCTTCGACCTGGTATCCCCAAgCAGGgCAGCACATT    422
[Cuba]AJ223505RC    CGGTTCTTCGACCTGGTATCCCCAAACAGGTCAGCACATT    412
Culiacan_MexicoD    CGGTTCTTCGACCTGGTATCCCCAAACAGGTCAGCACATT    419
SinaloaRC           CGGTTCTTCGACCTGGTATCCCCAAACAGGTCAGCACATT    419
Malaga_virus_AF2    CGaTTCTTCGACCTGGTATCCCCAAgCAGGgCAGCACATT    422
Mild[Aichi]AB014    CGaTTCTTCGACCTGGTATCCCCAAgCAGGgCAGCACATT    259
Mild[Portugal]AF    CGaTTCTTCGACCTGGTATCCCCAAgCAGGgCAGCACATT    422
Mild[ShizuokuaAB    CGaTTCTTCGACCTGGTATCCCCAAgCAGGgCAGCACATT    259
Mild[Spain7297]A    CGaTTCTTCGACCTGGTATCCCCAAgCAGGgCAGCACATT    422
MildX76319RC        CGaTTCTTCGACCTGGTATCCCCAggCAGGgCAGCACATT    259
Sardinia_virusX6    CGaTTCTTCGACCTGGTATCCCCAAcCAGGTCAGCACATT    415
Sardini-[Spain1]    aGaTTCTTCGAtCTGGTATCCCCAAcCAGaTCAGCACATT    417
Sardinia-[Sicily    aGaTTCTTCGAtCTGGTATCCCCAAcaAGGTCAGCACATT    415
SardiniaL27708RC    aGaTTCTTCGAtCTGGTATCCCCAAcCAGaTCAGCACATT    417
Consensus            g ttcttcga  tggtatcccca   ag  cagcacatt

2-5rep              TCCATCCAAACATTCAGGCAGCTAAGAGCTCAACAGATGT    374
Rep[FL]             TCCATCCAAACATTCAGGCAGCTAAGAGCTCAACAGATGT    298
TYLCVX15656RC       TCCATCCgAACATTCAGGCAGCTAAGAGCTCAACAGATGT    293
FloridaAY530931R    TCCATCCAAACATTCAGGCAGCTAAGAGCTCAACAGATGT    459
Puerto_RicoAY134    TCCATCCAAACATTCAGGCAGCTAAGAGCTCAACAGATGT    459
GeziraAY044138RC    TCCATCCAAACATTCAGGgAGCTAAatcCagttCAGAcGT    462
[Cuba]AJ223505RC    TCCATCCAAACATTCAGGCAGCTAAGAGCTCAACAGATGT    452
Culiacan_MexicoD    TCCATCCgAACATTCAGGCAGCTAAGAGCTCAACAGATGT    459
SinaloaRC           TCCATCCgAACATTCAGGCAGCTAAGAGCTCAACAGATGT    459
Malaga_virus_AF2    TCCATCCAAACATTCAGGgAGCTAAatcCagttCAGATGT    462
Mild[Aichi]AB014    TCCATCCAAACATTCAGGgAGCTAAatcCagttCAGAcGT    299
Mild[Portugal]AF    TCCATCCAAACATTCAGGgAGCTAAatcCagttCAGAcGT    462
Mild[ShizuokuaAB    TCCATCCAAACATTCAGGgAGCTAAatcCagttCAGAcGT    299
Mild[Spain7297]A    TCCATCCAAACATTCAGGgAGCTAAatcCagttCAGAcGT    462
MildX76319RC        TCCATCCAAACATTCAGGgAGCTAAatcCagttCAGAcGT    299
Sardinia_virusX6    TCCATCCgAACATTCAGGgAGCTAAatcgagctCcGAcGT    455
Sardini-[Spain1]    TCCATCCgAACATTCAGGgAGCTAAatcaagctCcGAcGT    457
Sardinia-[Sicily    TCCATCCgAACATTCAGGgAGCTAAatcgagctCcGAcGT    455
SardiniaL27708RC    TCCATCCgAACATTCAGGgAGCTAAatcaagctCcGAcGT    457
Consensus           tccatcc aacattcagg agctaa        c ga gt
```

Risks and limitations:

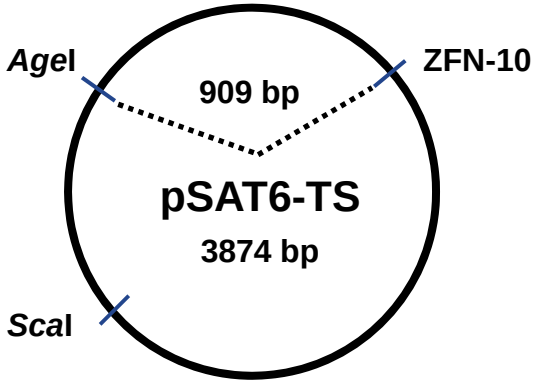
I. **Mutagenesis of off target sites**

A

RH1 - **GCA**	Q	S	G	D	L	R	R
RH2 - **GAT**	T	S	G	N	L	V	R
RH3 - **ATG**	R	R	D	E	L	N	V

```
TGCATCCAT-GTAAGT-ATGGATGCA
                 GG      T
                  C      C
                  A
```

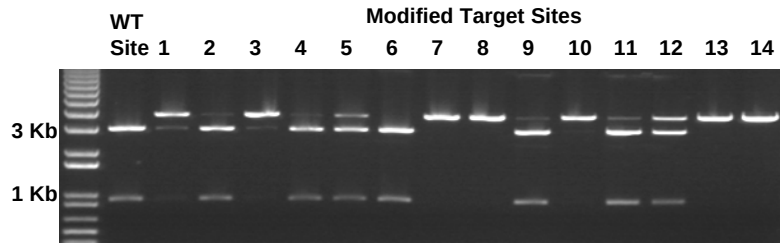
B



C

Construct	Sequence	Cutting Efficiency
WT	TGCATCCAT-GTAAGT-ATGGATGCA	100%
1	AGCATCCAT-GTAAGT-ATGGATGCT	8%
2	GGCATCCAT-GTAAGT-ATGGATGCC	97%
3	TGCATCCCT-GTAAGT-AGGGATGCA	2%
4	TGCATCCGT-GTAAGT-ACGGATGCA	98%
5	TGCATCCTT-GTAAGT-AAGGATGCA	79%
6	TGCATCCAC-GTAAGT-GTGGATGCA	100%
7	GGCATCCCT-GTAAGT-AGGGATGCC	0%
8	TGCATCCCC-GTAAGT-GGGGATGCA	0%
9	AGCATCCTC-GTAAGT-GAGGATGCT	97%
10	TGCATCCAT-GTAAGT-ATGCATGCA	0%
11	TGCATCCAT-GTAAGT-ATGGCTGCA	94%
12	TGCATCCAT-GTAAGT-ATGGACGCA	63%
13	TGCATCCAT-GTAAGT-ATGCCTGCA	0%
14	TGCATCCAT-GTAAGT-ATGCCCGCA	0%

D



Risks and limitations:

I. Mutagenesis of off target sites – PCR analysis of ‘known’ sites

II. Toxicity – new ZFN architecture

III. Transgene removal – breeding out ZFN cassettes (advantage of using a single transgene)

IV. Targeting complex genomes – multi copy genes, gene clusters

V. Limited access to the technology

VI. Public acceptence

Personal:

Andriy Tovkach, (Ph.D. student)

Avner Levi (Postdoc)

Vardit Zeevi (Tech)

Collaborator:

Avi Levi (Weizmann institute, Israel)

Supported by:

Basic Research and Development Cooperation

Consortium for Plant Biotechnology Research

Human Frontiers Science Program

DOW

The University of Michigan