TARGETED GENE INTEGRATION IN PLANTS: VECTORS, MECHANISMS AND APPLICATIONS FOR PROTEIN PRODUCTION

TAGIP

STREP

PUBLISHABLE FINAL ACTIVITY REPORT

Contract No. 018785

Period covered: from 1.12.2005 to 30.11.2008

Partic. Role*	Partic. no.	Participant name	Participant short name	Country
CO	1	Weizmann Institute of Science	Weizmann	IL
CR	2	Centre National de la Recherche Scientifique	CNRS-DR07	F
CR	3	Karlsruhe University	UKARL	D
CR	4	Czech Academy of Science – Institute of experimental Botany	ASCR-IEB	CZ
CR	5	University of Geneva	UNIGE	CH
CR	6	Biogemma	Biogemma	F
CR	7	Evogene	Evogene	IL
CR	8	Friedrich Miescher Institute	FMI	CH
CR	9	OSM-Dan Ltd.	OSM	IL

COORDINATOR	WEIZMANN		
Name	Prof. Avi Levy		
E-mail	Avi.levy@weizmann.ac.il		
Telephone	+972 8 934 2734	**Fax**	+972-8-9342734

Date of preparation : December 2008
Start date of project: **1.12.2005**
Duration: 36M

Revision [draft, 1]

CONTENT LIST

1. Project Execution

1.1 Executive Summary

State-of-the-art and beyond: One of the limitations of current transgenics technologies in plants is that the process of DNA transformation is not precise. There is no control on the location or copy number of transgene insertion, and as a result the transgene is often silenced in unpredictable ways, immediately after transformation or there is a slow "decay" in the expression of the transgene throughout generations. This is problematic when trying to develop a new generation of "GM" crops with stable transgene expression. This is even more problematic when one aims to maintain high levels of expression of the transgene, as in the case of the production of therapeutic proteins in plants (Molecular pharming). There are several reports in the literature, including by members of this group, on targeted or site-specific integration of transgenes in plants. Most of these reports are still in model species and targeting efficiency rates are still too low to make it a practical, routine application. Recent advances, achieved after this project was started, have shown that GT rates can be enhanced through Zinc Finger Nucleases (ZFNs) that can be engineered to cleave DNA at any genomic site and thus turn such DNA into hotspots for gene targeting. The ZFN technology is still in its infancy, feasibility was shown only for very few endogenous targets, it is not clear how specific is DNA break induction and its use is highly limited by intellectual property obstacles. The achievements of TAGIP, described below, can be combined with the ZFN technology to improve it, or can offer an alternative.

The goal of TAGIP was to develop efficient targeted integration of DNA into plant genomes in order to achieve high and stable protein expression in plants. We have addressed this goal in model species (Arabidopsis and tobacco) as well as in two crops (maize and tomato). In order to achieve this goal we have developed a number of strategies. Specific goals included the development of new assays and new strategies for gene targeting; upregulation of recombination in plants to achieve gene targeting more efficiently; and protein production in plants from transgene introduced at specific sites.

Achievements: The TAGIP consortium has achieved considerable progress. TAGIP achievements are described below:

We have built of a battery of constructs and transgenic plants for new gene targeting assays (based on mRFP expression in seeds); new GT systems and strategies, in Arabidopsis, tobacco, and maize, based on the use of the meganuclease I-SceI for excision of the vector from a genomic site and insertion into an engineered site also cleaved by I-SceI expression. Part of the system was to develop inducible I-SceI gene for use in maize (an achievement that lead to a patent application); preparing a series of mutants and double mutants, in genes involved in homologous or non-homologous recombination and assessing how these affect homologous recombination, DNA repair and gene targeting; preparing a system for over-expression of bottleneck genes in a specific and controlled manner; testing the importance of Cytosine methylation on the accessibility of the target to GT. The long term goal of the project is to apply the new knowledge to applications in molecular pharming, namely the production of therapeutic proteins in plants. So far, the Pharming effort has been focussed on preparing a system for site-specific DNA integration (via the Recombinase-Mediated Cassette Exchange) and for DNA integration via homologous recombination through RAD54-mediated stimulation in tobacco and tomato. We have not achieved protein production from targeted sites but we have made considerable progress towards this goal

Dissemination exposure and impact of TAGIP work was done through tens of refereed publications including in leading journals, through ~ 100 lectures in international conferences in seminars in Universities, in schools and community gathering places, in public debates on GMOs,

through posters at conferences and open days, through our web site, through a patent application and through the organization of an international workshop that attracted 120 people, including scientists, students, and industry. Overall, thousands of people, worldwide have been exposed to the work done by TAGIP.

In conclusion, the work done in TAGIP has contributed not only to the advancement of the technology of gene targeting and protein expression in plants, but also to the basic understanding of mechanisms that control DNA recombination and repair in plants, such as the role of resolvases, of non-homologous end-joining, of chromatin remodelling and of DNA methylation. We did not reach the stage of protein production from targeted integration events as originally planned, however, great progress was made towards these goals. Overall, the achievements of TAGIP have broad scientific, economical and social impact in plant biotechnology.

1.2 Project objectives

The goal of TAGIP was to develop efficient targeted integration of DNA into plant genomes in order to achieve high and stable protein expression in plants. We have addressed this goal in model species (Arabidopsis and tobacco) as well as in two crops (maize and tomato). In order to achieve this goal we have developed a number of strategies.

-Our first goal is to further improve gene targeting assays and targeting vectors for the identification of gene targeting events and for the optimization of the process. The new assays are based on the identification of the targeted event through visualization of fluorescent seeds in a way that enables high throughput screening of large seed populations. The new vectors that are being developed are either T-DNA vectors, or vectors "launched" via excision from within genomic sequences and targeted to a recombination hotspot. This work is done in Arabidopsis, tobacco and maize.

-Second, we want to up-regulate the rate of gene targeting on the one hand, and on the other hand to down regulate the competing pathway of DNA integration into the plant genome. Promoting GT will be achieved by over-expression of bottleneck proteins or by alterations in the chromatin structure, either through chromatin remodelling or through alterations in methylation in the target locus or through alteration in resolvase proteins. Preventing random integration will be achieved via silencing or via mutation of the more efficient non-homologous (or illegitimate) recombination machinery.

The third goal is to use TAGIP expertise as well as the novel knowledge obtained through TAGIP to achieve targeted integration of DNA for high protein expression in plants for pharming purposes. This will be performed in tomato and tobacco.

1.2 Contractors involved

Partic. Role*	Partic. no.	Participant name	Participant short name	Country
CO	1	Weizmann Institute of Science	Weizmann	IL
CR	2	Centre National de la Recherche Scientifique	CNRS-DR07	F
CR	3	Karlsruhe University	UKARL	D
CR	4	Czech Academy of Science – Institute of experimental Botany	ASCR-IEB	CZ
CR	5	University of Geneva	UNIGE	CH
CR	6	Biogemma	Biogemma	F

CR	7	Evogene	Evogene	IL
CR	8	Friedrich Miescher Institute	FMI	CH
CR	9	OSM-Dan Ltd.	OSM	IL

1.3 Work performed and end results

WP1: New vectors and delivery methods for GT

WP Leader: UKARL

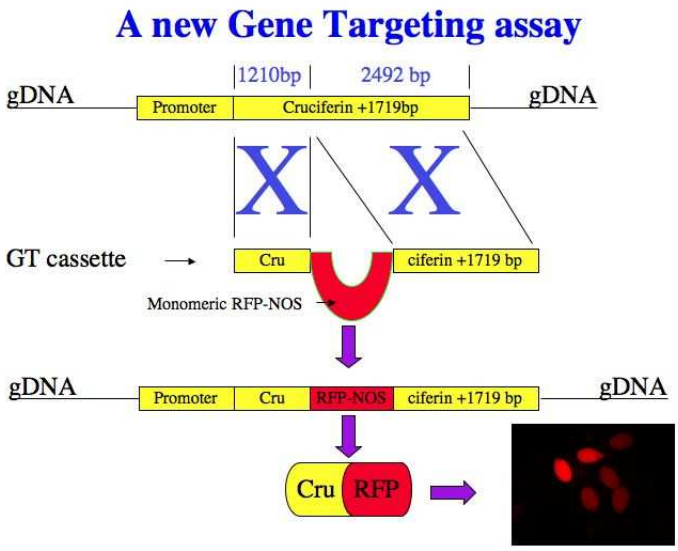
The goal of WP1 was to develop new vectors and assays for GT in plants. New RFP-based targeting vectors for Arabidopsis were set up providing a convenient assay for the development and improvement of Gene Targeting in plants (D1-1). In addition, we attempted to develop a positive-negative GT selection system based on visible seed markers (D1-2). The new assays we have developed before and during this project were based on selection of fluorescent seeds (green or red) upon targeting of the endogenous Cruciferin locus. Under this scheme, our initial vector contained an RFP gene under the control of the Napine promoter near each border of the T-DNA. Internally to the RFP gene are the regions homologous to the target. In between these regions is a GFP marker expressed under the Cruciferin promoter. The selection scheme was originally interpreted as follows: A seed that is red and green corresponds to a non-homologous insertion, a seed that is red only corresponds to the insertion of a truncated T-DNA and a seed that is green only is expected to be a GT event. To provide a proof of concept of this idea, a series of vectors have been built to target the endogenous AtRAD54 gene (the ortholog of RAD54 from yeast). For some reasons, which remain unclear to us, the positive selection marker, while present in the transformed plants, was silenced in all the constructs that we have made, including a series of controls that were aimed at determining the origin of the silencing. At this point, it is still not possible to use the positive-negative scheme for gene targeting of any locus in Arabidopsis. The work done here will serve as a basis to continue and develop this concept.

Epi-Alleles differing in the methylation status for the *PPO* locus were established in specific Arabidopsis lines (D1-3). Thus an unique system could established to address the question whether DNA methylation GT assay can influence GT in plants. The results on the effect of methylation and chromatin on GT rates are described in WP2 (D2-5).

A new targeting technique that relies on the excision of genomic sequences *in planta* was established. The strategy was applied to Arabidopsis, tobacco and maize. Transgenic plant lines with a 'Target Locus' (TL) or a 'TargetVector' (TV) integrated into the plant genome were combined by crossings. Induction of restriction endonuclease I-SceI activity should lead to the excision of the TV from the genome. Efficient excision could be achieved in Arabidopsis (D1-4) and by the use of an dexamethasone (Dex) inducible I-SceI gene in maize (D1-6) but not in tobacco. Due to time limitations no GT experiments could be performed in maize but multiple experiments were performed in Arabidopsis. Functional single copy lines of the respective constructs were selected and used to test the experimental system. We could show that in Arabidopsis excision of the linear targeting sequence is very efficient as measured by the HR-mediated repair of the resulting DSB. Surprisingly, we were not able to detect any GT events. Therefore we conclude that the linear targeting sequence is degraded before the HR machinery can act on it. Another possibility that cannot be excluded is that NHEJ factors access the linear sequence before HR proteins can bind resulting in random integration of the targeting sequence. As we could not demonstrate any GT events we did not further analyse the fate of the targeting

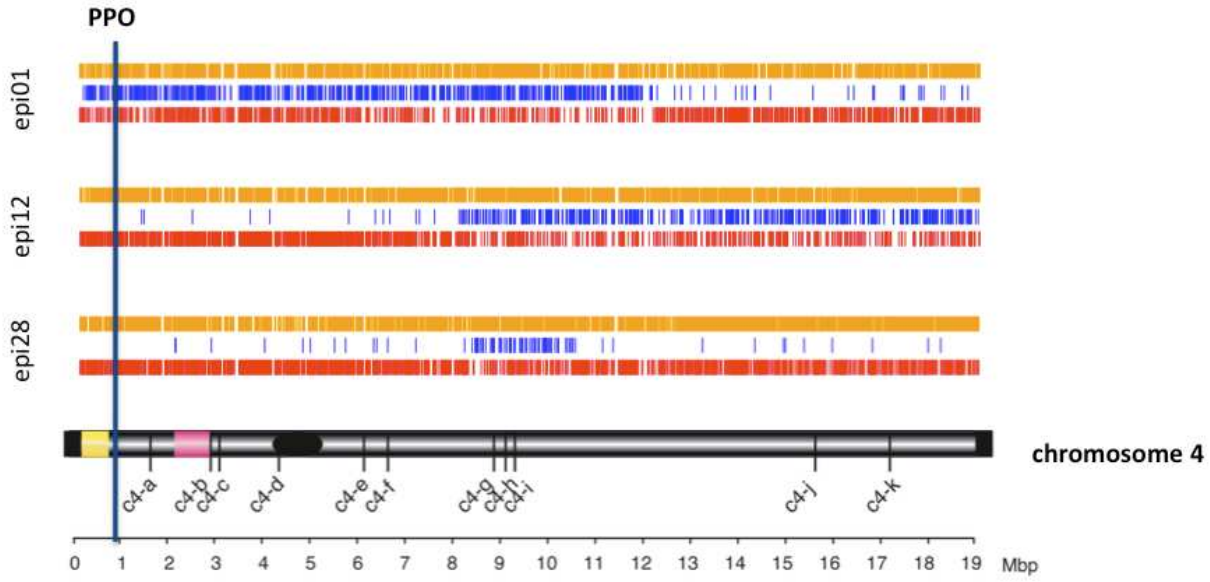
sequence. In summary we have to conclude that the excision-mediated delivery system is not yet optimized to become an appropriate tool for enhancing the GT frequency in plants.
Maize has been transformed with gene constructs that are designed to test gene targeting (GT) in maize to a defined genomic site (hot spot). Selected maize Target Lines (TL) have been crossed to lines that contain the repair region (Target Vector (TV) lines). Infrequent I-SceI-mediated excision of the repair DNA region from the TV loci was observed however, excision was not enhanced by the application of a chemical inducer of I-SceI activity. The weak excision of the repair region probably explains why GT events have not yet been observed (D1-7). Further work is required to validate the new delivery vector--hotspot recipient system. The frequency of Repair region excision might be enhanced either by optimisation of the application of the inducer or via constitutive strong expression of I-SceI. A patent application has been filed by P6 on the inducible I-SceI system.

Example of some highlights (WP1):



A gene targeting assay based on the activation of the promoter-less and ATG-less monomeric RFP upon homologous recombination with the Cruciferin target gene and the expression of a Cruciferin::mRFP protein fusion. Gene targeting events are identified as red fluorescent seeds.

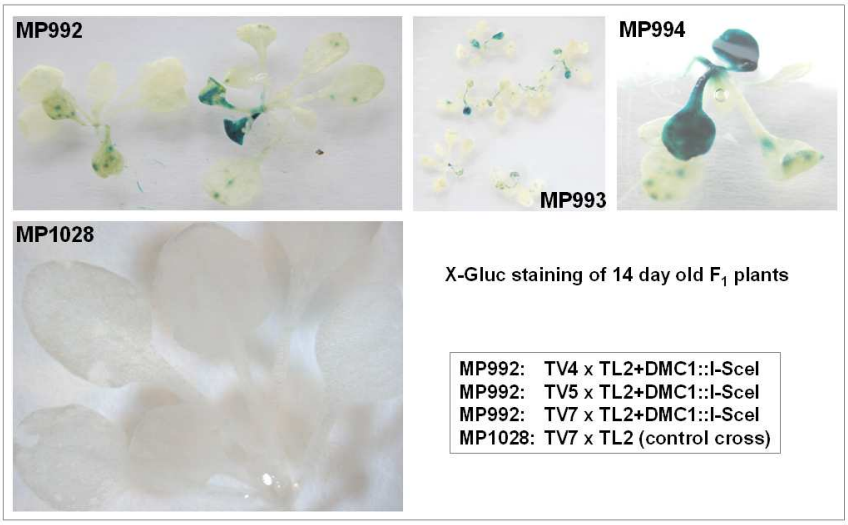
Characterization of epi alleles for Gene Targeting experiments



Distribution of DNA methylation polymorphisms on chromosome 4 for 3 analysed epiRILs. DNA methylation polymorphisms corresponding to hybridization signals are depicted as: non parental (orange bars), met1-3 like (blue bars), and WT-like (red bars). Positions of markers used for the validation assays are given (c4-a to k). Chromosome 4 drawing indicates the centromere (black), the heterochromatic knob (pink), the nucleolar organizer region (yellow). PPO locus (At4g01690) is indicated as a vertical bar.

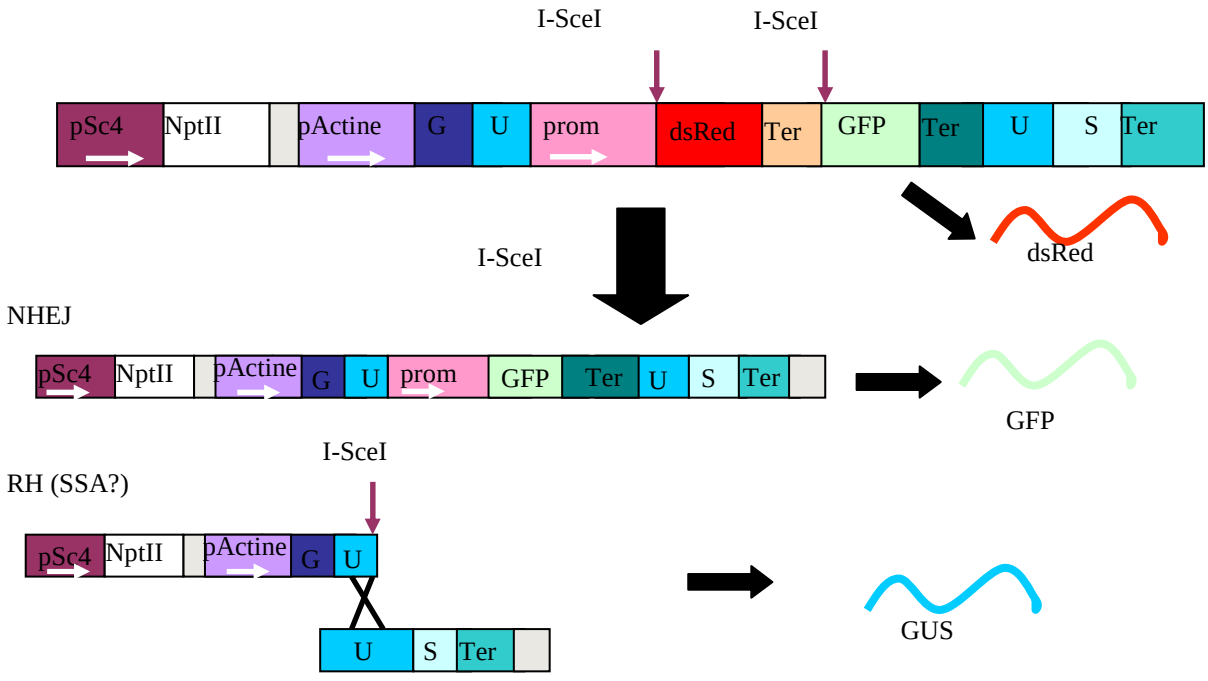
Part of the work was published in Cell: Mathieu O, Reinders J, Caikovski M, Smathajitt C, Paszkowski J. (2007) Transgenerational stability of the Arabidopsis epigenome is coordinated by CG methylation. Cell 130(5):851-62;
And in Genome Research: Reinders J, Delucinge Vivier C, Theiler G, Chollet D, Descombes P, Paszkowski J (2008), Genome-Wide, High-Resolution DNA Methylation Profiling Using Bisulfite-Mediated Cytosine Conversion. Genome Research 18 (3): 469-76

Development of an Excision system in planta for launching a Gene targeting vector:



Excision of targeting cassette via I-SceI mediated DSB induction. *Histochemical staining of F_1 seedlings for three different experimental crosses and one control clearly indicates that in many cells the targeting cassette has been excised after I-SceI expression and the resulting DSB in the Target Vector sequence has been restored via SSA-mediated HR.*

A system for inducible vector excision and targeting in maize:



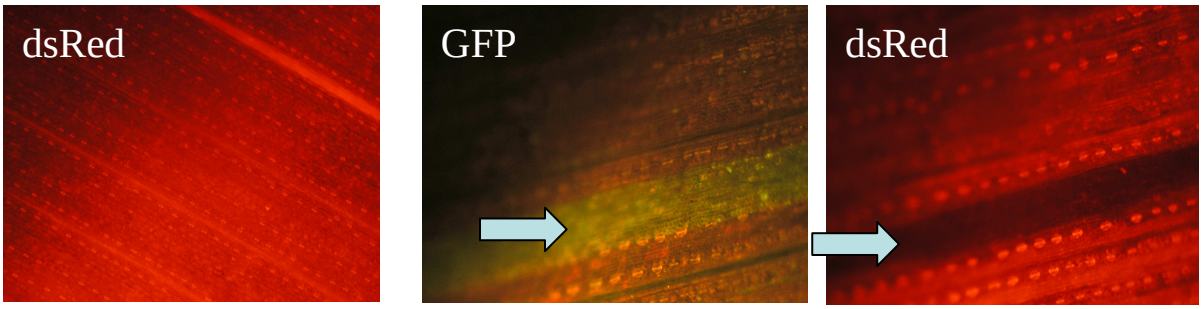
Scheme for monitoring HR or NHEJ

Plants containing just the reporter gene had leaves that uniformly expressed DsRed, whereas plants that contained both the reporter and the NLS-SynI-SceI-ZmMod gene had sectors that lacked dsRed activity.

I-*Sce*I activity in Maize

Reporter

Reporter + Optimised I-*Sce*I



Uniform DsRed expression

Non-uniform DsRed expression + GFP+ve

Sectors of maize leaves showing the phenotypes of recombination markers, GFP or RFP.

WP2: Altering DNA recombination pathways to enhance GT

WP Leader: CNRS

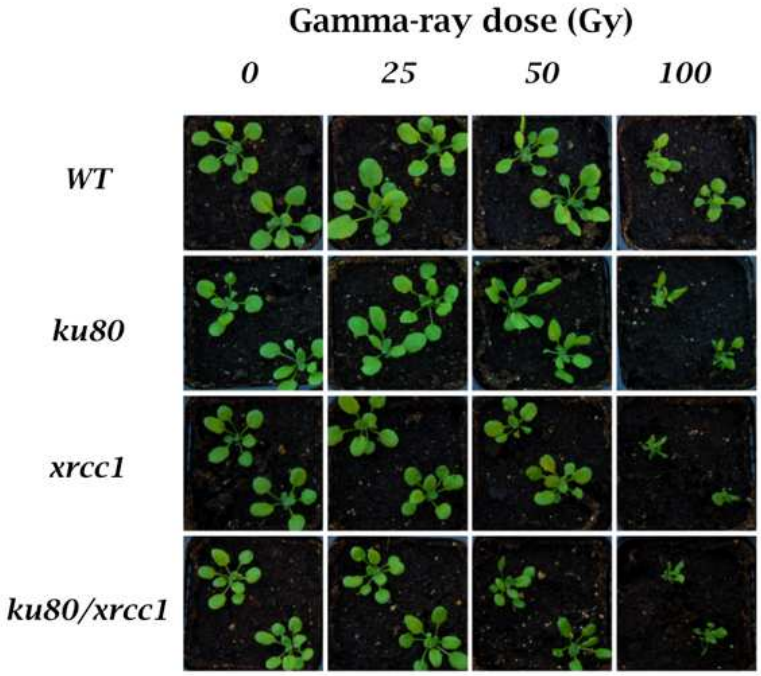
This workpackage aims at enhancing gene targeting through the favouring of homologous recombination (HR) with respect to non-homologous recombination (NHR) pathways in the plant. The work involves three complementary approaches and it is planned to combine those that are shown to have positive effects to look for synergistic effects on HR later in the project. The three approaches involve: 1) reducing the efficiency of NHR to favour HR; 2) Favour HR by expression of specific HR-implicated proteins; 3) Understand the roles of chromatin structure on HR and hopefully stimulate HR through the use of chromatin mutants. All three of these approaches include the characterisation and different mutant lines to define their roles in recombination and genome maintenance. Considerable progress has been made in all parts of this workpackage and all deliverables have been fulfilled (D2-1 to D25). Notwithstanding, as discussed at the mid-term review, unexpectedly low frequencies of gene targeting events and higher than expected numbers of false positives meant that the analytical use of the available gene targeting assays has advanced less than initially hoped. Thus in the second project period more than the planned amount of work concerned the testing and development of different gene targeting assays and vectors in large-scale experiments (WP1 and 2).

Highlights are:

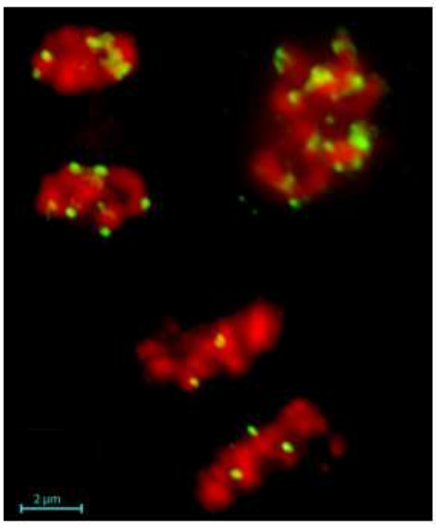
- characterisation of HR proteins and roles using the corresponding mutants (Mus81, Eme1a,b, Rad1, Ercc1, Rad51 paralogs) (D2-1 and D2-2).
- Characterization of recombination/repair mutants for DNA repair kinetics (D2-1 and D2-2).
- Identification of a new non-homologous recombination pathway in Arabidopsis (D2-1 and D2-2).
- purification of tagged proteins for in vitro recombination assays (D2-4)
- testing and validation of PPO locus for the study of the effects of epigenetic modifications on GT (D2-5).
- construction of inducible egg cell-specific expression system for GT and expression of yeast recombination genes in Arabidopsis - testing of GT in these lines (D2-3 and D2-5)
- tests of GT in NHR mutants (D2-5).
- screening of EpiRIL population for lines affecting DNA methylation levels and testing of promising candidates in GT assays (D2-5).

Examples of highlights:

Characterization of Arabidopsis non-homologous recombination mutants

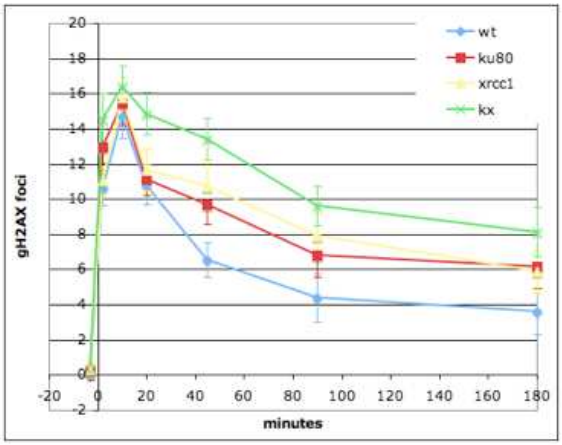


γ-H2AX foci : Kinetics of DNA double-strand break repair

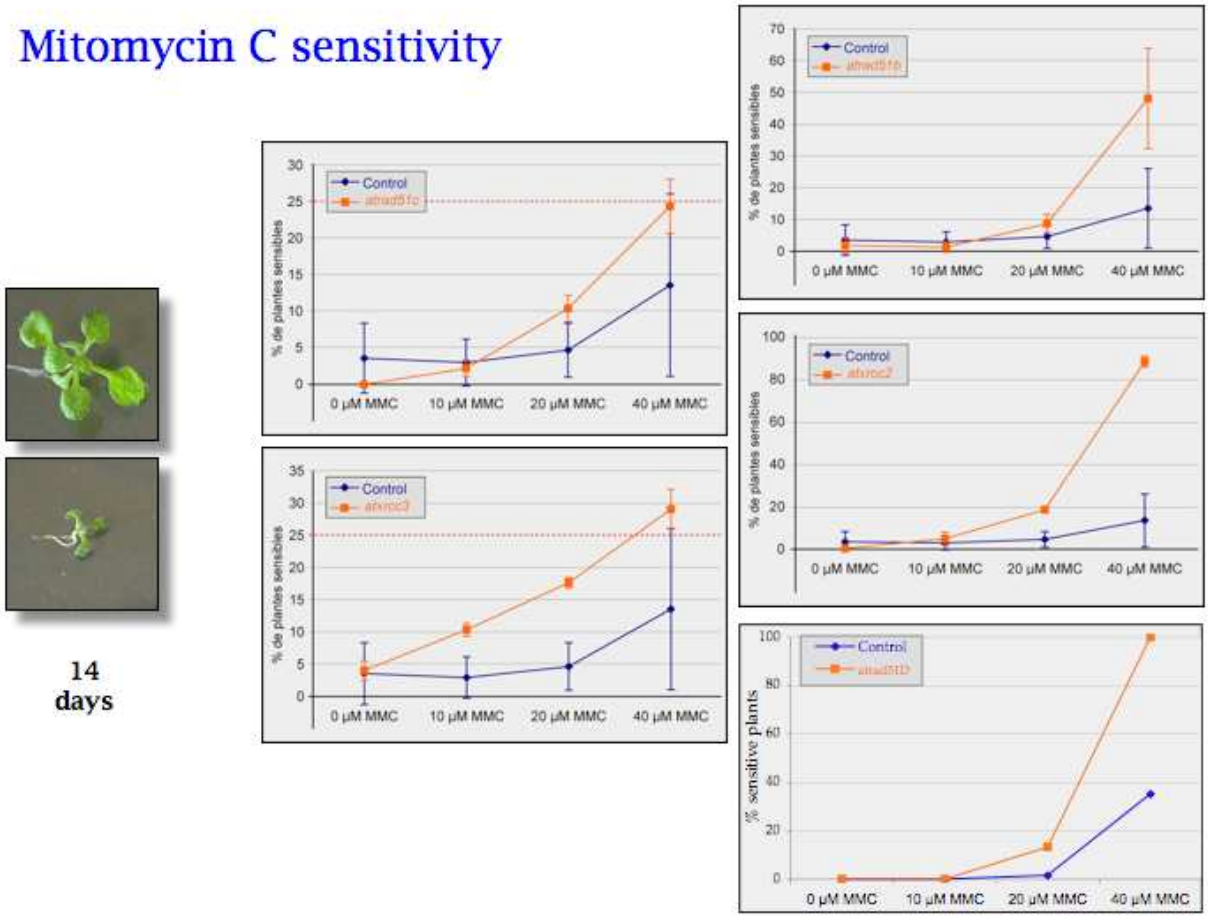


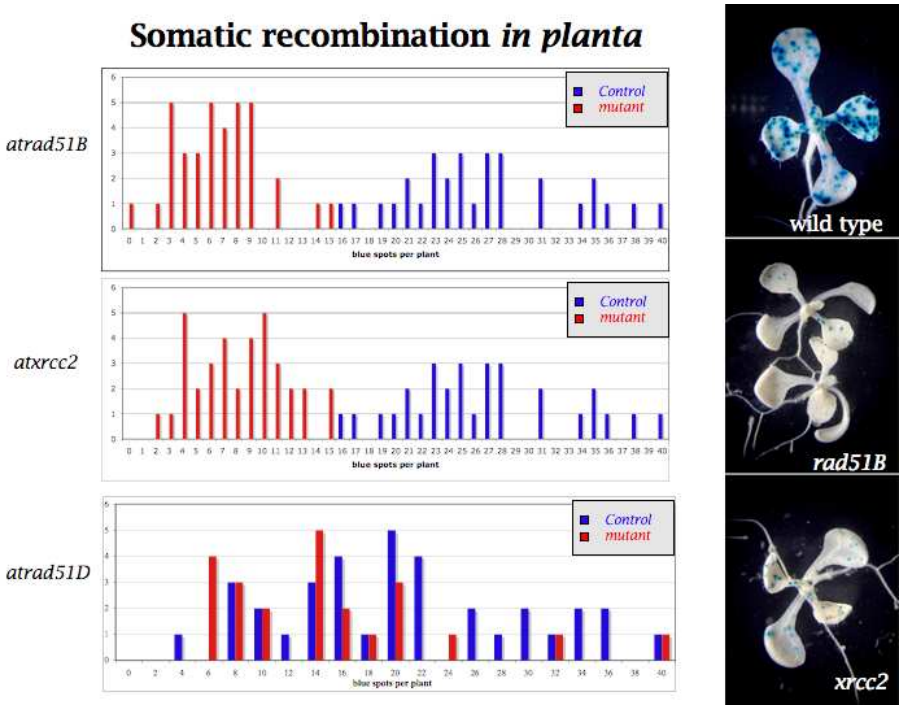
25 Gy.
Repair for 2 (upper) and 90 mins (lower)

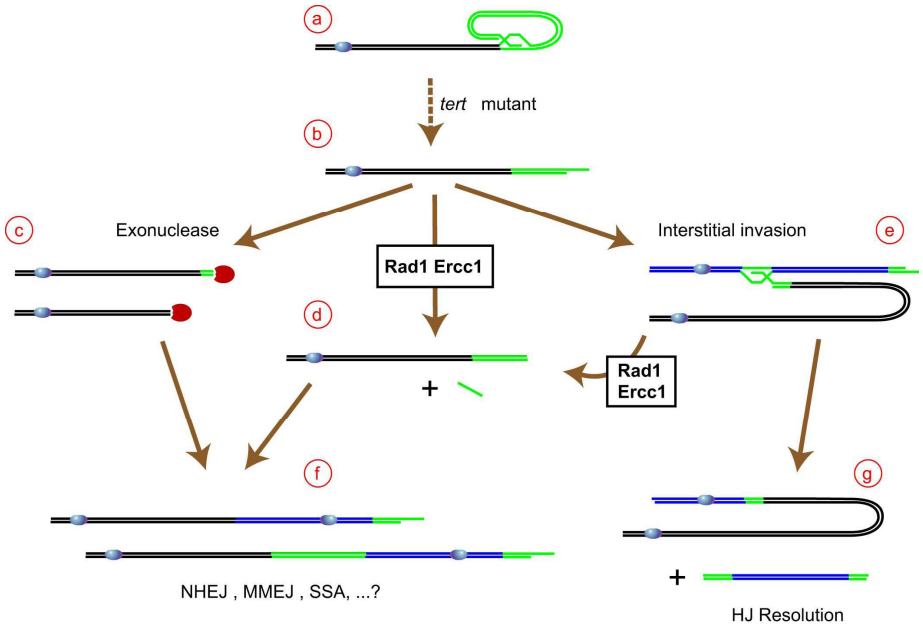
KU and Xrcc1 dependent pathways



Mutants and double mutants in homologous recombination-related genes

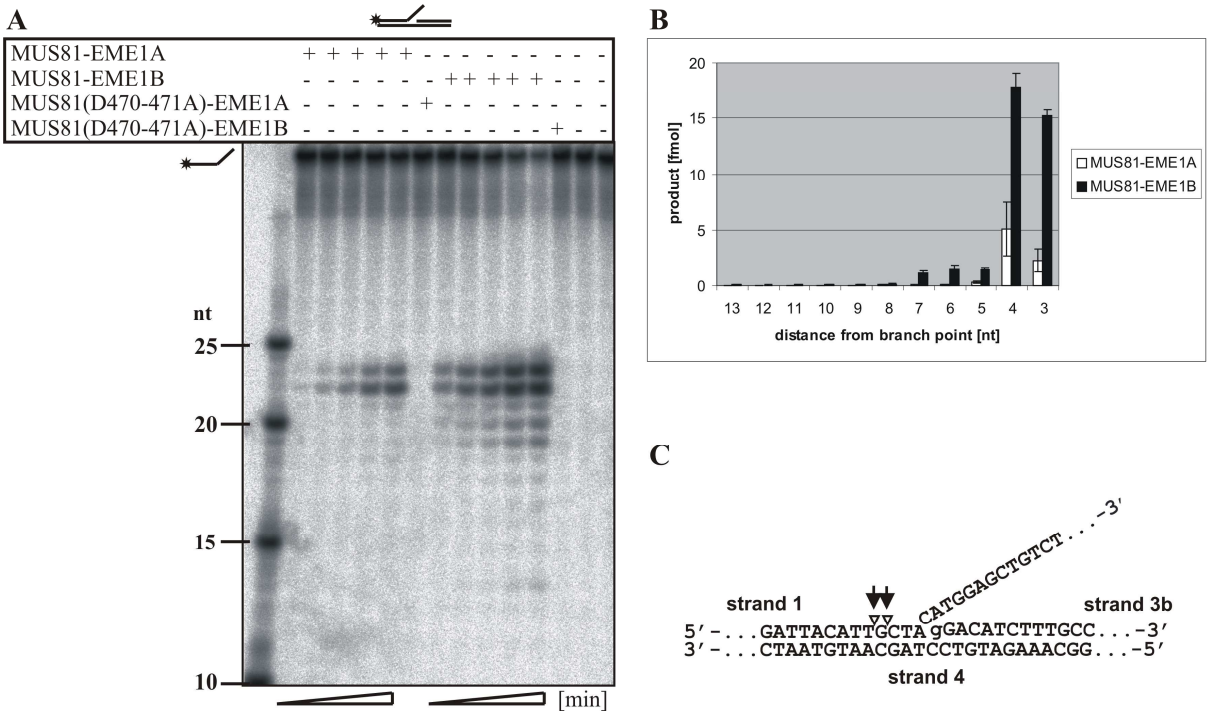






Model of the roles of AtERCC1/AtRAD1 in different fates of uncapped telomeres in Attert mutant Arabidopsis. *Erosion of telomeric repeat DNA in the Attert mutant leads to progressively more frequent loss of T-loop structure and uncapping of telomeres (a->b). Uncapped telomeres may be further eroded by exonucleases (b->c) or AtERCC1/AtRAD1 can cleave the G-strand overhang to leave a blunt end (b->d). Non-homologous, micro-homology mediated, or single-strand annealing recombination (NHEJ, MMJ, SSA) can fuse chromosomes of these structures (c->f and d->f). The G-strand overhang of structure (b) can also recombine with interstitial telomeric repeat sequences (e) and in certain invasion configurations (Chr 1R and 4R in Arabidopsis), resolution of this structure by Holliday-junction resolvase generates a dicentric plus an acentric chromosome (g). Cleavage of structure (e) by AtERCC1/AtRAD1 prior to the action of resolvase will produce structure (d), although the existence of this AtERCC1/AtRAD1-dependent process (e->d) cannot be verified in Arabidopsis, given that evidence for structure (e) is only found in the absence of AtERCC1/AtRAD1.*

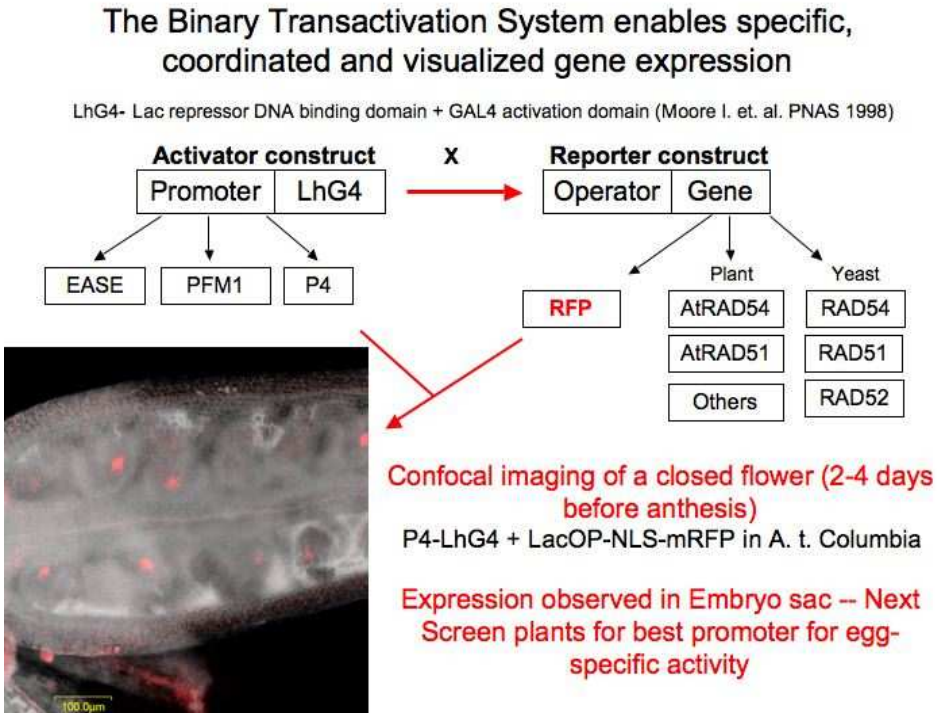
Identification and in vitro analysis of DNA cleavage specificities of the Arabidopsis Mus81/Eme1 complexes.



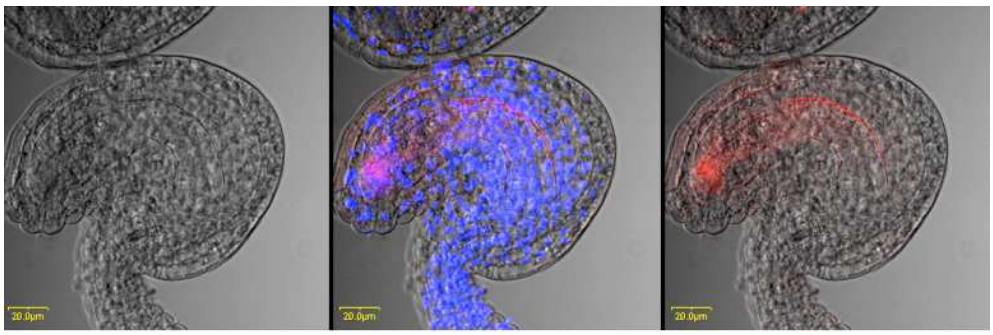
Processing of the 3'-flap structure by AtMUS81-AtEME1A and AtMUS81-AtEME1B. Representative autoradiography. The 3'-flap substrate was incubated with AtMUS81-AtEME1A or AtMUS81-AtEME1B respectively. A. Reactions were analyzed on a 20 % denaturing TBE-urea (7M) sequencing gel; the size of cleavage fragments is given in nucleotides. B. After quantification of the respective cleavage fragments the positions of endonucleolytic cleavage were mapped. The endonuclease data represent the mean of three independent experiments.C. Presentation of structure and sequence of the 3'-flap substrate with the corresponding cleavage sites of AtMUS81-AtEME1A (white arrowheads) and AtMUS81-AtEME1B (black arrowheads).

Results of GT enhancement in mutants and overexpressor plants

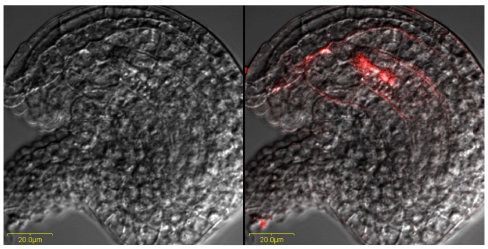
A system was designed for specific expression of HR-related genes in egg cell for GT via floral dipping in Arabidopsis. Testing this system with RAD54, RAD51, RAD52 genes confirmed the enhancing effect of RAD54 but did not show enhancement by RAD51 or RAD52.



A trans-activation system for specific expression in egg cells.



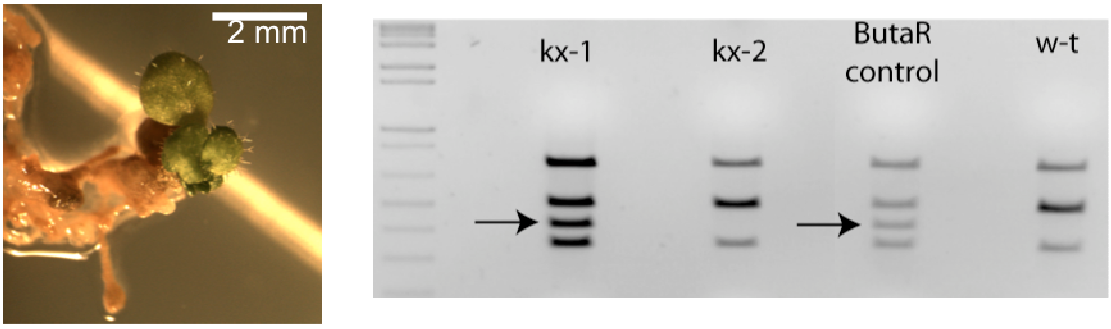
Transmission Transmission+RFP Transmission+DAPI+RFP



Transmission Transmission+RFP

EASE-LhG4 + LacOP-mRFP in ovules

Hypomethylation of the target did not significantly affect GT in epiRILs



Identification of a GT event with the root callus protocol. Photograph of green, Butafenacil-resistant callus (left) and NlaIII digested PCR products from DNA from 2 selected ButaR calli (lanes kx-1 and kx-2). The presence of an extra 85bp NlaIII fragment (arrowed) is diagnostic for gene targeting of the PPO locus. Controls with Butafenacil resistant and wild-type control DNA amplifications/digestions are also shown.

WP3: Application of GT for expression of foreign therapeutic proteins in plants

WP Leader: ASCR-IEB

The goal of this WP is to apply the new knowledge in DNA recombination to applications in molecular pharming, namely the production of therapeutic proteins in plants. Two strategies were used: site-specific DNA integration (via the Recombinase-Mediated Cassette Exchange) and DNA integration via homologous recombination through RAD54-mediated stimulation in tobacco and tomato. P4 (IEB) and P7 (Evogene) have constructed and characterized components for GT systems to produce therapeutic proteins. IEB is setting a system for Recombinase Mediated Cassette Exchange (RCME) in which lines with locus for integration between lox sites has to be first generated and selected, whereas Evogene is directly targeting the *Polyphenol oxidase* (*PPO*) locus to produce highly enriched protein of interest in tomato trichomes. In both systems GT is either only possible because of expression of *Cre* recombinase or is strongly increased by expression of ScRAD54 to facilitate selection of correct insertions. Both partners were challenged with problems mainly associated with expression of the genes that lead to the design of new constructs, improved transformation protocols and accommodation of new methodological approaches. Some modifications were thus implemented and progress has been slower than expected. The specific achievements were:

- Production of tomato plants expressing GT-enhancing ScRAD54. (P7, D3-1 and D3-5)

We describe here the work that has been done to optimize gene targeting in tomato, based on ScRAD54-mediated enhancement of recombination. Our modified approach was to develop an efficient protoplast gene targeting assay in tomato that enables to screen a large population of cells by flow cytometry (FACS analysis) and thus to assess in a quantitative manner the rate of GT. We describe our advances with this assay in ScRAD54-expressing tomato protoplasts. To show the effect of scRAD54, this gene was introduced into MicroTom plants and T2 plants were analyzed for scRAD54 expression by Western blot. Unfortunately, no commercially available anti-scRAD54-specific antibody could be found. We therefore analyzed the expression level by quantitative RT-PCR. We could show that almost all of the 32 independent events express the

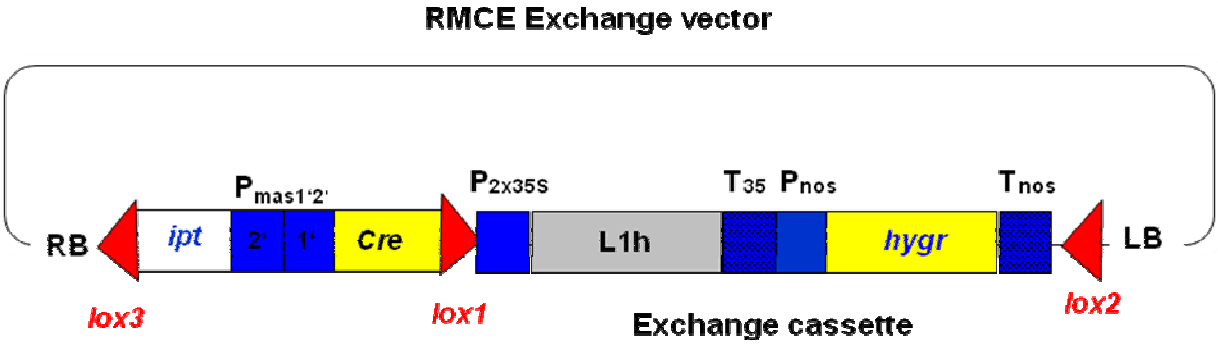
gene. T3 plants and seeds were produced for all events expressing the gene. Protocols for the extraction of protoplasts from MicroTom leaves as well as for their efficient transformation was developed, leading to up to 70% transformation efficiency. In addition, to be able to analyze the single recombination events, a special growth medium for single cells sorted by flow cytometry was developed, which would permit growth of the sorted cells and PCR analyses of the different recombination events. The first targeting site of choice was the PPO gene, which is highly expressed in tomato trichomes and which is therefore the ideal candidate to be targeted for the production of a protein. In fact, the transformation efficiencies were always tested using the "positive control" – the 100aa PPO-GFP fusion protein. Unfortunately, although the 100aa PPO-GFP fusion protein is very fluorescent, the full length PPO-GFP fusion protein, which will be produced upon the recombination between the targeting vector and the targeted site, is not fluorescent. Therefore we needed to look for an alternative targeting site. We chose TAG1, a gene similar to Cinnamic acid 4-hydroxylase, for which we showed that it is highly expressed in MicroTom protoplasts. We also chose this gene, as it contains a large intron quite close to the 5' side, permitting the preparation of a construct, which contains a large enough 5' end for recombination (~1300 bp), but a quite short (~240 aa) protein part upstream to the GFP in the fusion protein produced upon recombination. Unfortunately, this fusion protein is not fluorescent either. Our third approach was to use an IRES (Internal Ribosome Entry) stretch between the TAG1 and the GFP ORF, based on the data obtained by P4, that seemed to show that IRES permits the expression of a selection gene in tobacco. Unfortunately, the IRES did not work in our hands, either. Although we prepared all the necessary elements, we could not show the proof of concept of the scRAD54 effect on the recombination efficiency in tomato, due to the missing reporter system. However, the experience gained has enabled to further advance towards our goal to target HGH1 or human interferon beta to the PPO locus, as originally planned.

-Tomato plants expressing HPV16L1h targeted at the PG locus (P4, D3-4)

Due to initial problems with transformation transgenic tomato lines expressing reporter and production protein HPV16 L1h we decided to test L1 expression transiently to verify its functionality when used in RMCE. Transient expression takes advantage of mRNA transcription of desired protein from the same vector that is used for transformation before it becomes integrated into genome. Transient expression levels are usually higher than that of transgenes and so the system is well suited for characterization of potential products. Approximately 1 g of fresh leaves collected from 4 weeks old plants were vacuum infiltrated with Agrobacteria culture for 10 minutes and left for 3 days in humid chamber for protein expression. Expressed protein was detected in crude extracts prepared from infiltrated leaves by competitive (sandwich) ELISA with mab 1.3.5.15 HPV16 L1 as primary - binding and mab E2 secondary – detection antibody. Mab E2 is specific for HPV16 VLP's and was generated by P4. Testing showed that plant expressed HPV16-L1h is similar in all respects to reference HPV16VLP's prepared in baculovirus system. This shows on side capacity of tobacco and tomato to express fully functional protein and on the other full functionality of HPV16 L1h expression cassette on RMCE integration vector. While the project has progressed and tomato transformation became better calibrated, this system is going to be used in stably transformed plants at targeted sites via RMCE (see below and D3-7)

- Establishing the Recombinase-Mediated Cassette Exchange (RMCE) system in tomato for HPV16-L1 expression. (D3-7)

Construction of an exchange vector proceeded in 3 stages: a) Construction of selection and expression cassettes and of dual Cre/ipt cassette, b) Testing of expression cassette for VLP's production and c) assembly of RMCE exchange vector in pGreenII.



a) Construction of selection and expression cassettes.

a.1. Coding sequence of hptII from pCAMBIA 1300 (provided by P1) was PCR cloned into pNOS cassette (JII, John Innes Institute, UK) under Pnos and NOSpolyA control to create HYGR cassette.

a.2. L1h coding sequence from plasmid 881 (Dr. Müller, DKFZ, Heidelberg, Germany) was PCR cloned into pKA121 (P4 lab) under P2x35S and CaMVpolyA control to create p2x35S-L1h cassette. pKA121 is a p2x35S cassette pJIT60 (JII) subcloned in a binary pGreenI 0029 (JII) for testing transient expression of cloned sequences.

a.3. Coding sequence of Cre recombinase was proprietary optimized for expression in tomato/tobacco and custom synthesized by GeneArt (Regensburg, Germany) to enable SalI/BamHI direct cloning.

Pmas1'2' dual promoter was PCR cloned with Tg7polyA from pAP2034 (Velten, Schell). In primer flanking 2' promoter of Pmas1'2' was added NdeI site and in primer flanking Tg7polyA PmeI site. Coding sequence of IPT with native polyA signal was directly PCR amplified from C58 T-DNA in plasmid 1346 (Dr. Vlasak, IPMB, Ceske Budejovice, Czech Republic) with primers enabling NdeI and PmeI cloning. Dual pMAS1'2'ipt cassette was assembled in pGEM3 vector by joining ipt fragment and Pmas1'2' fragment through NdeI.

Optimized Cre fragment was then directly cloned into this construct as SalI/BamHI fragment under Pmas1'. Final dual cassette pMAS1'cre2'ipt is available as a single PmeI fragment for further cloning.

b. Binary RMCE vector pKA180d was assembled in pGreenII in following steps:

b.1. pKA157 is pGreenII vector with modified T-DNA between RB and LB and cloned lox1 and lox2 sites pointing inward (deliverable D3.3). In this vector lox1 site next to RB was replaced between AccIII and StuI sites by dual lox site oligonucleotide. This oligonucleotide codes lox1 and lox3 sites pointing outward with PmeI restriction site in between. Final vector pKA180a was used for subcloning selection and expression cassettes.

b.2. p2x35S-L1 cassette from pKA121-L1h was inserted between SacI and SpeI sites in polylinker of pKA180a giving rise to pKA180b.

b.3. HYGR cassette was inserted as a EcoRV fragment between L1h cassette and lox1. The clone with L1h and HYGR cassettes arranged head to tail was selected for further cloning because it provides possibility to replace CaMVpolyA and Pnos for IRES sequence for bicistronic translation of L1h and hptII. The intermediate vector is pKA180c.

b.4. pMAS1'cre2'ipt cassette was cloned into PmeI site between lox1 and lox3 in pKA180c to provide RMCE exchange vector pKA180d to be used in transient exchange of exchange cassette for target cassette in plants.

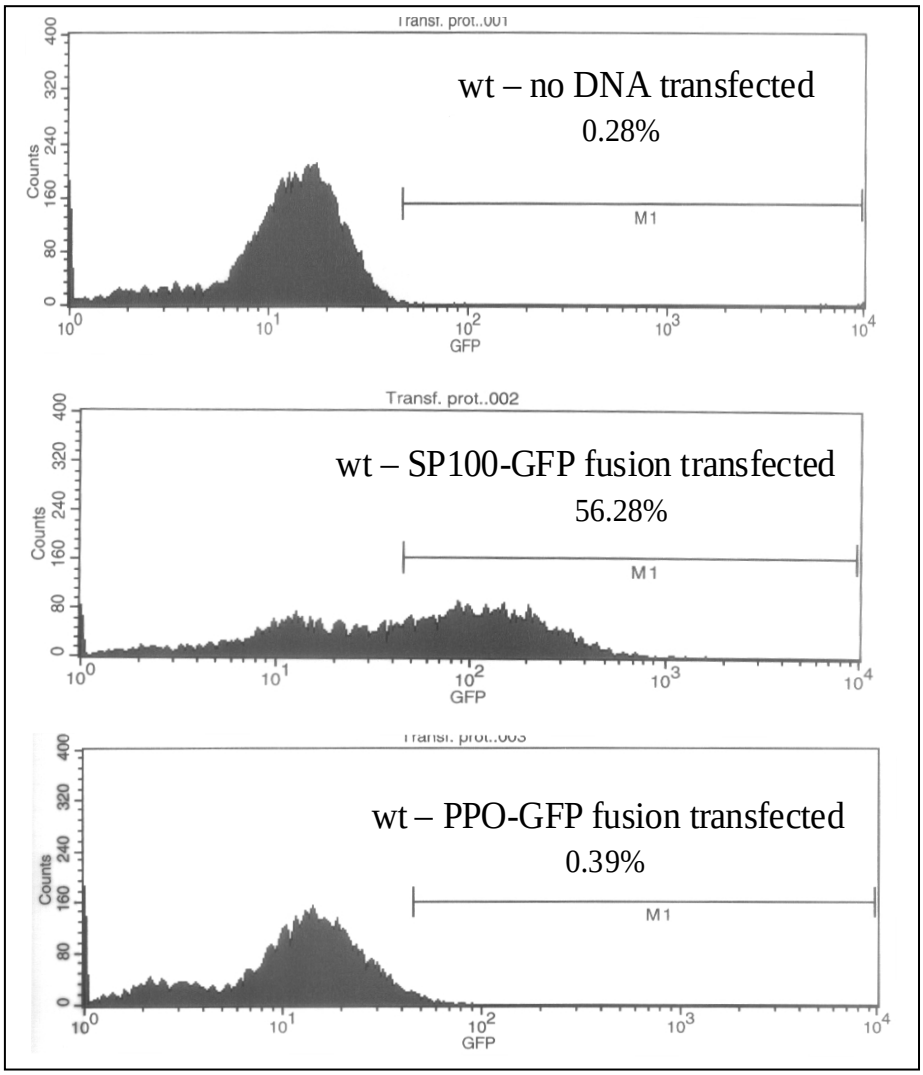
-Tomato lines with lox sites at loci whose expression is characterized (P4,D3-3 and D3-6)
Tomato lines with inserted target sequence with lox sites in reverse orientation are prerequisite for Recombinase Mediated Cassette Exchange (RMCE). For introduction of target sequence into tomato genome is constructed binary vector pKA140 (deliverable D3-3) with glucuronidase uidA (GUSi) gene and fusion gene codA::nptII (CODNPT) for positive/negative selection. When integrated, target cassette enables selection of transformants and checking influence of integration context on GUS expression level. Integrated negative selection is used to eliminate lines that escape exchange of target sequence for expression cassette with gene of interest during second transformation. By Cre mediated cassette exchange can be suitable target master lines easily converted to production lines without extensive testing of integration.
The work is divided into three stages: (a) Developing of efficient tomato transformation protocol and transformation with pKA140 target binary vector; (b) Growing transgenic plants (T_0 and T_1 generations) and selection of lines with highest GUS expression; (c) Selection of single copy lines and analysis of genome integration site.

(a) MicroTom was selected for production of bioactive proteins (in TAGIP HPV16 L1 capsid protein) because of quicker life span and modest space requirement in comparison to other varieties. When testing transformation protocols provided by P1 and P7 we found that use of "armed" *Agrobacterium* strain Agl1 is mandatory for efficient transformation with pGreenII derived binary vectors used as binary backbone for RMEC constructs. This strain harbouring pKA140 was used to transform MicroTom cotyledons and currently first transformants are regenering (M18).

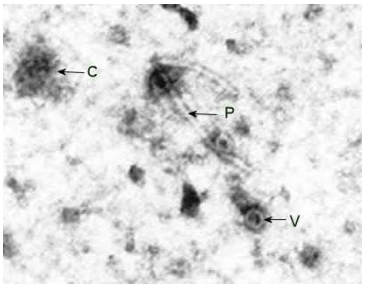
Because of persisting problem with regeneration of MicroTom transformants by M18 Nicotiana tobaccum was adopted as alternative host of RMCE system. By M36 several tens of tobacco (T1) master lines are available and tested for GUS expression. Finally by the end of the project first regenerated MicroTom tomatoes appeared. Currently (M36) 8 tomato (T0) lines are available.

Examples of highlights:

-FACS system for analysis of homologous recombination in tomato protoplasts. (P7, new task)

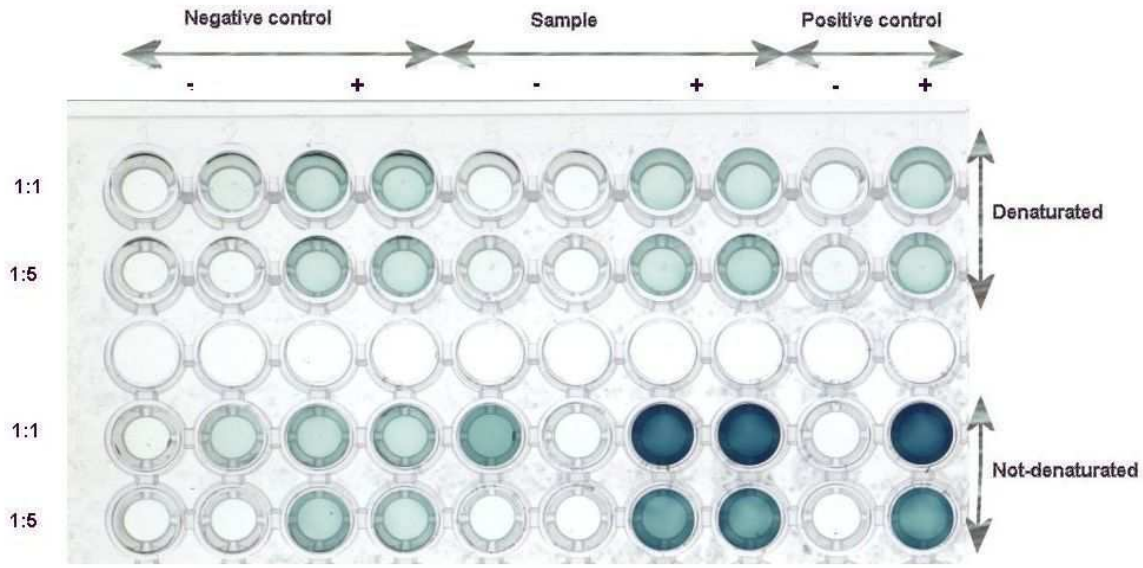


Flow cytometry analysis of GFP in tomato protoplasts. Protoplasts were prepared from wt MicroTom plants either transformed with no DNA, with 35S-100aaPPO-GFP (SP100-GFP) or the full length PPO-GFP fusion construct. Histogram plots were gated in a region showing highest fluorescence with SP100-GFP transfected and lowest with mock-transfected cells. Numbers indicate the % of GFP-positive cells obtained in each transfection.



EM-picture showing expression of L1h VLP's along with substructures from pKA121-L1h vector in crude extracts from infiltrated tobacco leaves.

Detection of L1 expression in crude extract from pKA180d infiltrated tobacco leaves.



Expression of HPV16 L1 VLP's from RMCE integration vector pKA180d as detected by competitive ELISA. (+) Slots coated with primary antibody (mab 1.3.5.15 HPV16 L1), (-) Slots without coating (only PBS - blank). Heat (65°C, 10 min) denatured and not-denatured samples were tested. Sample – extract from Nicotiana benthamiana infiltrated leaves. Negative control - extract from non-infiltrated leaves. Positive control – VLP's from baculovirus expression system. Peroxidase-conjugated anti-rabbit antibodies and ABTS were used for visualization. Positive signal was detected only in slots with not-denaturated sample and positive control. Due to selectivity of secondary E2 antibody majority of expressed protein assembled into VLP's.

2. Dissemination and Use

Dissemination exposure and impact of TAGIP work was done through tens of refereed publications including in leading journals, through ~ 100 lectures in international conferences in seminars in Universities, in schools and community gathering places, in public debates on GMOs, through posters at conferences and open days, through our web site, through a patent application and through the organization of an international workshop that attracted 120 people, including scientists, students, and industry. Overall, thousands of people, worldwide have been exposed to the work done by TAGIP.

Utilization and disseminating knowledge We have published our work in the scientific literature, as well as in the non-scientific press via press releases done at the onset of the project. This is the best way to attract the attention of the scientific and biotech community to the new developments that will emerge from TAGIP. Approx 30 publications have already been published and we expect that this number will keep growing as results will mature. The reagents (mutants, lines, clones, vectors..) developed through the work will also be disseminated, either through MTAs or freely or through licensing after patenting. In addition, the work done in TAGIP has been very widely presented at scientific meetings or in seminars in various institutions by all partners. Lectures were also given in schools, in community centres, and in public debates on GMOs. Altogether ~ 100 lectures describing TAGIP work were presented.
Our work is also advertised through web sites, the central TAGIP site
http://www.tau.ac.il/~amieldro/tagip/abstract.htm and in other related sites at partners institutions and companies. Finally, partners were encouraged to protect their intellectual property and apply for patents. One such patent application was already filed by one of the SMEs (P6). It is possible that additional patents will follow.

One highlight of the TAGIP dissemination process is the International Plant DNA recombination and repair International meeting, in Giens, May 31-June 3rd 2007, chaired by P2 (Charles White) and co-organized by TAGIP partners, P1, P3 and P9. Funding from EMBO was also successfully obtained by P1 and enabled to invite prominent speakers and to support registration fees of all students. The meeting that was essentially initiated by TAGIP embers, attracted 120 scientists from all over the world, including ~ 10 biotech companies.

Overall the project has received a very broad international exposure. It has contributed strongly to the scientific understanding of DNA recombination in plants through publication in journals such as Cell, Plos Genetics, Genome research, Genetics, etc.. It has involved SMEs who have also been strengthened as a result of their TAGIP activity. Note also that during the period of the TAGIP project the two SMEs have also broadened their collaboration to other fields.
There were some changes in the program, due to technical difficulties that were encountered, however, most deliverables were fully met and for a few deliverables, progress was made towards future delivery.
In summary, we consider that the materials and results derived from the TAGIP project have been and will be highly useful to the academy and industry and have made a significant **impact in the field of plant biotechnology for agriculture and health**.

Overview table : Exploitable knowledge and its use

Exploitable Knowledge (description)	Exploitable product(s) or measure(s)	Sector(s) of application	Timetable for commercial use	Patents or other IPR protection	Owner & Other Partner(s) involved
EK1 mRFP-Cru vector	Vector used as assay	Optimization of GT for industry and academy	Immediate	None	P1
EK2 Positive-Negative seed selection	GT Vector to target any gene	Mostly research in Arabidopsis	2 years	None	P1
EK3 dexamethasone (Dex) inducible I-SceI gene	Clones and transgenic plants	Agriculture Biotech	Immediate	Patent Pending	P6
EK4- ISceI meganuclease-based GT	Clones and transgenic plants	Agriculture Biotech	2 years	Not yet	P3 and P6
EK5- Transactivation system for expression in egg cells	Clones and transgenic plants	Mostly research in Arabidopsis	Immediate	Not yet	P1
EK6- Mutants in homologous and non-homologous recomb.	Plants and recombination and genome stability data	Agriculture Biotech and research in Arabidopsis	Immediate	None	P1, P2, P3, P5
EK7- Epi-RILs	Plants and DNA methylation data	Research in Epigenetics and recombination	2 years	Not yet	P5
EK8- Plants overexpressing Rec genes	transgenic plants	Agriculture Biotech and research	2 years	Not yet	P1, P2, P3, P5
EK9- Isolation of the PPO locus	Sequence data	Agriculture Biotech for molecular Pharming	Immediate	Not yet	P7
EK10- Vectors for studying GT in tomato protoplasts	Vectors	Agriculture Biotech for molecular Pharming	2 years	Not yet	P7
EK11- Vectors for RMCE	Vector	Agriculture Biotech	2 years	Not yet	P4
EK12 Vectors and plants expressing HPV16 L1h	Vectors and plants	Agriculture Biotech for molecular Pharming	2 years	Not yet	P4
EK13 Vector, *Arabidopsis* and maize lines for measuring intrachromosomal RH and NHEJ *in planta*	Vectors and plants	Agriculture Biotech and research to optimize GT	Immediate	Not yet	P6