Universität für Bodenkultur Wien
University of Natural Resources
and Life Sciences, Vienna

Master Thesis

Attempt to Inactivate *Sfh8*, Encoding a Membrane Interacting Protein in Tomato, via CRISPR/Cas9 And de Novo Induction of Meristems

submitted by
Anna WULGARAKIS, BSc

in the framework of the Master programme

Phytomedizin

in partial fulfilment of the requirements for the academic degree

Diplom-Ingenieurin

Vienna, September 2022

Supervisor:

Ao.Univ.Prof. DI Dr. Gerhard Adam
Institute of Microbial Genetics

Department of Applied Genetics and Cell Biology

Attempt to Inactivate *Sfh8,* Encoding a Membrane Interacting Protein in Tomato via CRISPR/Cas9 And de Novo Induction of Meristems

Anna Wulgarakis

A Master thesis submitted to the

University of Natural Resources and Life Sciences, Vienna

Department of Applied Genetics and Cell Biology

in the study field of

Phytomedicine

(UH 066 422)

with a major in

Plant Biotechnology

supervised by

Ao.Univ.Prof. DI Dr. Gerhard Adam

The experimental part was performed during a research stay in Greece

at the

University of Crete, Department of Biology

in the group of

Assoc.Prof. Dr. Panagiotis Moschou

Crete, Greece

April to November 2021

Vienna, 21 September 2022

Declaration of Authorship

I hereby declare that I have authored this master thesis independently, and that I have not used any assistance other than that which is permitted. The work contained herein is my own except where explicitly stated otherwise. All ideas taken in wording or in basic content from unpublished sources or from published literature are duly identified and cited, and the precise references included. I further declare that this master thesis has not been submitted, in whole or in part, in the same or a similar form, to any other educational institution as part of the requirements for an academic degree. I hereby confirm that I am familiar with the standards of Scientific Integrity and with the guidelines of Good Scientific Practice, and that this work fully complies with these standards and guidelines.

Anna Wulgarakis

Acknowledgements

I would like to thank the following people in particular for helping me with my project: Dr. Gerhard Adam and Dr. Gerlinde Wiesenberger, who not only inspired me to dive into the field of molecular biology in the first place but also supported my idea of going abroad and always gave me the most valuable tips and cues. Dr. Panagiotis Moschou, who allowed me to work in his group for 8 months and let me free hand in the overall approach and planning of the project. Dr. Androniki Bibi, who very generously shared her lab with me and supported me with friendship, positivity and an incredibly peaceful work environment.

Special thanks to:

Glykeria (Rita) Mermigka, PhD

Irini Dervisi, PhD

Vasiliki Michalopoulou, PhD

Artemis Perraki, PhD

Anthony James, PhD

Vassiliki Kathidou, MSc

Marialena Vathi

And last but not least my family. Without their financial support I would not have been able to undertake this journey to Greece. A special thanks in that regard must go out to my dear aunt Monika.

Contents

Abstract [EN]

In the roots of legume plants (plants from the *Fabaceae* family) rhizobia bacteria form a unique, symbiotic, nitrogen-fixing organ called a nodule. Homologs of genes (e.g. *Sfh8*) that are associated with nodule formation can also be found in non-leguminous plants pointing to a possible ancestral role for these genes in plant physiology. My thesis work is part of a project that tries to investigate the role and function of the *Sfh8* gene in a non-leguminous plants like tomato (*Solanum lycopersicum*). As an initial step tomato mutants should be generated. The gene knock-out procedure was done using state of the art molecular biology techniques like CRISPR/Cas9 and Golden Gate cloning. Simultaneously a new protocol was established that has the potential of making classical tissue culture obsolete in the future. De Novo Induction of Meristems uses developmental growth regulators to induce the growth of a potentially altered meristem on a young plantling on soil. Guide RNA construction was achieved via primer annealing and subsequent integration into a multiple module transformation vector by type II restriction enzymes. Additional modules were the Cas9 enzyme, two growth regulators, WUSCHL2 and IPT and luciferase as a non-invasive selection marker. The final assembled plasmid was then electroporated into Agrobacterium LBA4404 which were then injected into the cut site of the apical meristem of a Moneymaker tomato plantling. After 2-3 weeks growth of new branches from the injection site could be observed. Due to a lack of time, I could not confirm successful transformation anymore. This evaluation will be done by the host lab. Additionally, as a back-up plan a *Sfh8* T-DNA mutant (background: MicroTom) was found in an online database, and the initial characterization of the obtained plants was started.

Keywords

Sfh8, CRISPR/Cas9, Nodulin-like Genes, Plant Microbiome, Tissue Culture, De Novo Induction of Meristems, Developmental Regulators, WUSCHEL, IPT, Agrobacterium mediated Plant Transformation, Golden Gate Cloning

Abstract [DE]

In den Wurzeln von Hülsenfrüchten (Familie *Fabaceae*) bilden Rhizobiumbakterien ein einzigartiges, symbiotisches, stickstofffixierendes Organ, das als Knöllchen bezeichnet wird. Homologe von Genen (z. B. *Sfh8*), die mit der Knöllchenbildung assoziiert sind, können auch in nicht-leguminösen Pflanzen gefunden werden, was auf eine mögliche Rolle dieser Gene in der evolutionären Pflanzenphysiologie hinweist. Meine Diplomarbeit ist Teil eines Projekts, das versucht, die Funktion des *Sfh8*-Gens in einer nicht-leguminösen Pflanze wie Tomate (*Solanum lycopersicum*) zu untersuchen. Als erster Schritt sollten Tomatenmutanten kreiert werden. Das Gen knock-out Verfahren wurde unter Verwendung modernster molekularbiologischer Methoden wie CRISPR/Cas9 und Golden Gate Cloning durchgeführt. Gleichzeitig wurde ein neues Protokoll etabliert, dass das Potenzial hat, klassische Gewebekultur in Zukunft zu ersetzen. De Novo Induction of Meristems verwendet Entwicklungsregulatoren, um Wachstum eines potenziell genetisch veränderten Meristems in einer Jungpflanze auf Erde zu induzieren. Die Guide RNAs wurden durch Primer-annealing konstruiert und anschließend in einen Transformationsvektor mit mehreren Modulen integriert. Die Konstruktion dieser Vektoren wurde von Typ II Restriktionsenzymen durchgeführt. Weitere Module waren das Cas9 Enzym, die zwei Wachstumsregulatoren, WUSCHL2 und IPT sowie Luciferase als nicht-invasiver Selektionsmarker. Die Vektoren wurden in Agrobacterium LBA4404 elektroporiert und in die Schnittstelle des Apikalmeristems einer Moneymaker Tomatenpflanze injiziert. Nach 2-3 Wochen konnte das Wachstum neuer Strukturen an der Injektionsstelle beobachtet werden. Eine erfolgreiche Transformation konnte aus Zeitmangel nicht mehr bestätigt werden. Diese Auswertung erfolgt durch das Gastlabor. Zusätzlich wurde als Backup-Plan eine *Sfh8* T-DNA Mutante (Hintergrund: MicroTom) in einer Onlinedatenbank gefunden, und mit der ersten Charakterisierung der gewonnenen Pflanzen begonnen.

Introduction

In the 21st century, scientific discovery and understanding are playing a more important role than ever in meeting environmental, human health and economic challenges. For decades humans were manipulating their crop plants selectively to meet changes in society and the environment [1][2][3][4]. The discovery of the DNA structure in 1953, followed by the invention of recombinant-DNA technology twenty years later, laid the foundation for direct genetic engineering so that the "nature" of an organism could be precisely altered within a single generation [5]. These advanced technologies have the potential to greatly improve the quality of human live: promoting health, preventing and curing diseases, finding new energy sources, ensuring food availability and safety - which at the same time could lead to greater measures to protect the environment.

In the past, novel genetic variation in crop plants could be achieved using methods such as insertion of transfer DNA (T-DNA), transposon insertion or T.I.L.L.I.N.G. (Targeting Induced Local Lesions in Genomes). The latter causes random point mutations across the genome via the chemical mutagen ethyl methanesulfonate (EMS). The disadvantage of these methods is the lack of control over the mutations introduced into the genome (background-mutations) plus the need for huge populations and tedious screening processes [6]. Since the discovery of the CRISPR system however, we now have a tool with high precision and high specificity. It enables scientists to target precise sequences in the genome and introduce mutations at any chosen site.

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)

To accomplish knock-out or knock-down of a target gene typically transformation vectors are used which encode domesticated Cas9. Molecular domestication accomplished to eliminate the enzyme's undesirable natural tendency to cut all sequences that are even remotely similar to the target (off-target-effect) [7]. Only a short RNA sequence must be synthesised for target recognition which makes the CRISPR/Cas9 system relatively cheap, extremely versatile and an easy to implement technology [8]. The importance of DNA targeting lies in creating damage at specific sites in genomes. The CRISPR system creates double strand breaks (DSBs) which most frequently are repaired by end-joining repair mechanisms, resulting in insertion/deletion

mutations (indels) [9]. In its natural form it consists of three essential components, the Cas9 protein and two short RNAs called CRISPR RNA (crRNA) and trans-activating CRISPR RNA (tracrRNA). The two RNAs together with Cas9 nuclease build an effector complex ready to fight viral invadors. The taget sites are recognised via a protospacer adjacent motif (PAM region). It is located at 3′ downstream of this target and is obligatory in order to recognize and cleave the site [10]

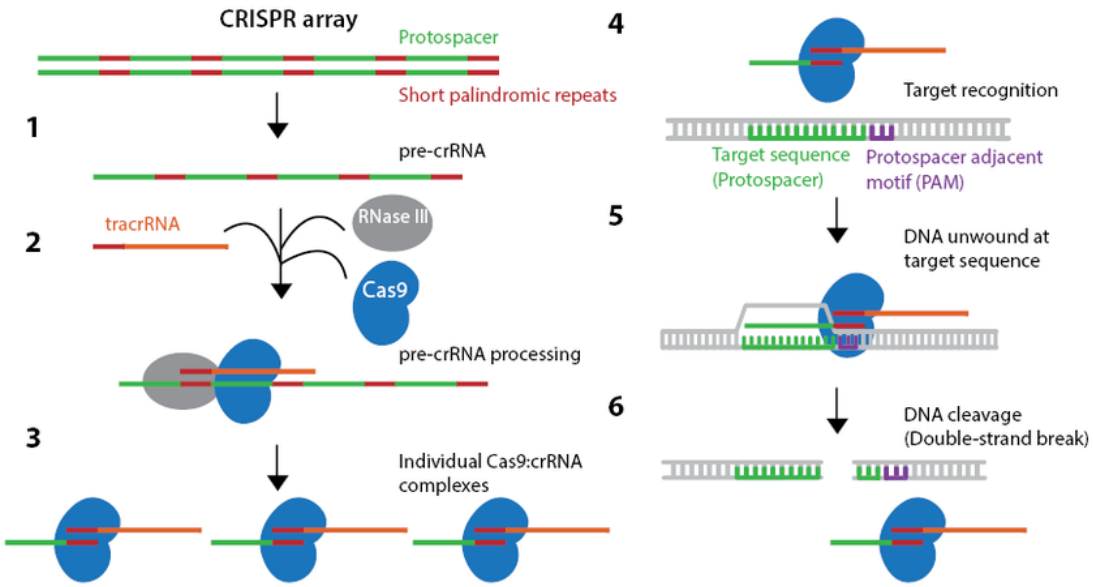


Fig. 1: Overview of a natural, bacterial CRISPR system.

Picture obtained from *addgene.org*. (1) sections of virus DNA (spacer) sandwiched between palindromic sequences together build a CRISPR array. After this array has undergone transcription it is called pre-crRNA. (2) A small tracrRNA, that binds to the palindromic repeats, and the nuclease Cas9 (CRISPR-associated nuclease protein) gets involved. Ribonuclease III (RNase III) cleaves the CRISPR array. (3) tracrRNA, crRNA and Cas9 form effector complexes. (4) Bacterial defence mechanism ready to target invader. (5) If the sequence is complementary to the crRNA, the nuclease enzyme recognises a short sequence that is unique to the viral genome, the so called PAM (protospacer-adjacent motif) region and introduces a DSB a few base pairs upstream this region (6). The DSB will neutralise the virus and prevent infection.

To make use of the CRISPR/Cas system in molecular biology for genome editing the crRNA and tracrRNA were joined together by a linker to form one single molecule. This molecule is called single guide RNA (sgRNA) and can be synthesised in a lab [11]. This now two component

system can detect any chosen site of about 20 base pairs as a target for editing. After a DSB has been introduced the host cell will try to repair the damage either by homology-directed repair (HDR) or by non-homologous end joining (NHEJ). The latter is more common in eukaryotes and leads to either nucleotide deletion or insertion [9]. Together with DNA sequencing gene editing, especially when the edit results in a frameshift or premature stop codon offers the possibility to gain insight into how an organisms genome dictates development and growth. The CRISPR/Cas9 system has been shown to work in a variety of important crop plant species and is simpler, more efficient and more flexible than former systems like zinc finger nucleases (ZFNs) or transcription activator-like effector nucleases (TALENs) which can also be engineered to cut DNA [10].

During the cause of this thesis vectors containing the CRISPR/Cas9 system were used in an attempt to knock-out the *Sfh8* gene in *Solanum Lycopersicum* (SoLyc) to study its effect on the plant root and subsequently its microbiome.

WHY SFH8 ?

The family of Phosphatidylinositol/phosphatidylcholine transfer proteins (PITPs) derive from highly conserved genes that can be found in all eukaryotic cells. Their function is associated with phospholipid metabolism, membrane trafficking, and polarized membrane growth [12]. They are capable of transporting either phosphatidylinositol (PtdIns) or phosphatidylcholine (PtdCho) between membrane bilayers and can be divided into fungal/plant PITPs and the metazoan PITPs depending upon their primary sequence similarities [12]. Interestingly, despite this difference the resulting proteins are virtually indistinguishable in biochemical assays. It is further believed that PITPs in higher plants facilitate developmental pathways for polarised membrane biogenesis in such a way, that they influence the unique symbiosis that permits nitrogen fixation, plus regulate stress responses [12]. The symbiotic nitrogen fixation is a partnership between polarized membrane structures (root hair) of leguminous plants and N_2-fixing rhizobia bacteria in the soil. Bacteria secrete so called nodulation factors to which the plant responds by deforming growing tips of root hair [13]. Through local invagination of the root hair plasma membrane an infection thread is formed and the resulting structures entrap

the bacteria [14]. Nodulins, organ specific plant proteins, define the plants contribution to symbiosis and are only expressed during nitrogen fixation [15].

The first protein identified from this family was Sec14p (*SECretory – SEC14*) in the yeast *Saccharomyces cerevisiae*. It plays an important part in the transport of proteins from the trans Golgi network [16]. Along with *SEC14*, the genome of *S. cerevisiae* encodes five additional Sec Fourteen Homologues (SFH): Sfh1p, Sfh2p, Sfh3p, Sfh4p, and Sfh5p, with Sfh1p sharing the highest homology to Sec14p [17] [18]. Genome sequence databases revealed that Sec14p can also be found in the model plant *Arabidopsis thaliana*. In fact, *Arabidopsis* encodes at least 31 potential Sec14p-like proteins in its genome (*Fig. 2*) [12]. Further, functional analysis of AtSfh1p could uncover that it is a novel Sec14p-nodulin, two-domain protein that regulates a complex developmental membrane morphogenetic pathway, coordinating membrane trafficking, phosphoinositide signaling, Ca^{2+} signaling and regulation of the actin and microtubule cytoskeletal systems [17].

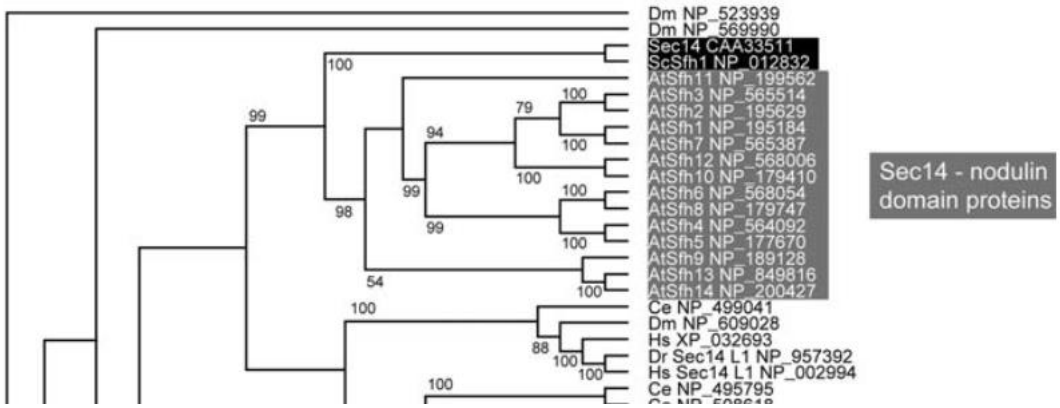


Fig. 2: Cladogram of the Sec14p family in *Arabidopsis*.

Picture from Phillips (2006) [12]. Partial listing of known members of the eukaryotic Sec14p family. When known, common designations are given along with the corresponding accession numbers.

Its expression is highly root specific and therefore essential for coordinated root hair development, regulating polarized membrane growth from a precise position on the root epidermal cell plasma membrane [12].

In 1997, Dr. Philipp Kapranov and his team discovered a highly expressed late nodulin gene *LjNODI6* in the model legume plant *Lotus japonicus* which corresponds to a particularly abundant mRNA species present in nitrogen-fixing root nodules [19]. Further studies [20] pointed to a relationship between the 15.5 kD protein Nlj16 and the predicted protein products of four members of a previously undescribed gene family in *L. japonicus*. They encode novel PITP-like proteins (LjPLPs) all sharing a two-domain structure consisting of an N-terminal PITP-like domain joined to a C-terminal domain composed of amino acid sequences identical to, or highly related to, nodulin Nlj16. It is believed that the latter domain functions as a specific plasma membrane targeting module.

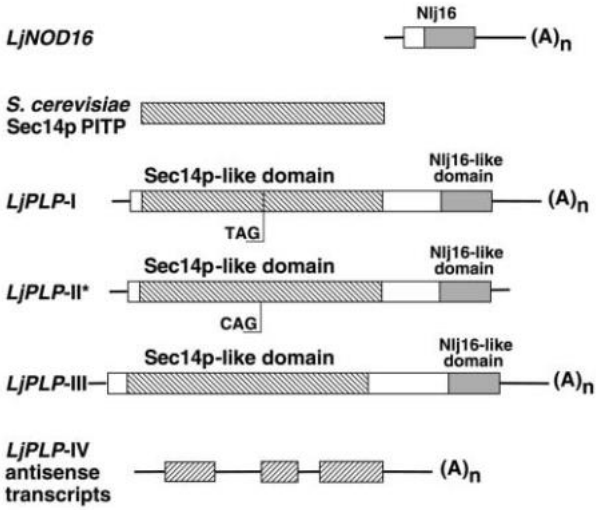


Fig. 3: Scheme of all four LjPLPs cDNAs.

Picture from Kapranov et al. (2001) [21]. boxes = coding regions, lines = 5' and 3' untranslated regions (UTRs), hatched area = Sec14p/Sec14p-like domains, shaded boxes = Nlj16-like domains, open boxes = no apparent sequence similarity with the other cDNAs.

Conceptual translation of the open reading frame (ORF) of LjPLPIV revealed amino acid sequences sharing a high degree of similarity with Sec14p-like proteins from other plant species, namely SFH.

Kapranov's results [20] led him to believe that the new gene family of developmentally regulated genes (LjPLPs), encoding Sec14p-like proteins, exert a dominant negative effect directed at inactivating the expression of specific LjPLPs in nodules.

Sec14-Nlj16–like nodulin proteins in plants tie together three developmentally important attributes, root hair development, lipid signaling, and nodulation. Arabidopsis mutants lacking the AtSfh1 Sec14-nodulin elaborate short, distorted root hairs characterized by loss of tip-directed phosphatidylinositol 4,5-bisphosphate gradients, disorganized cytoskeleton networks, and delocalized Ca^{2+} signaling [22].

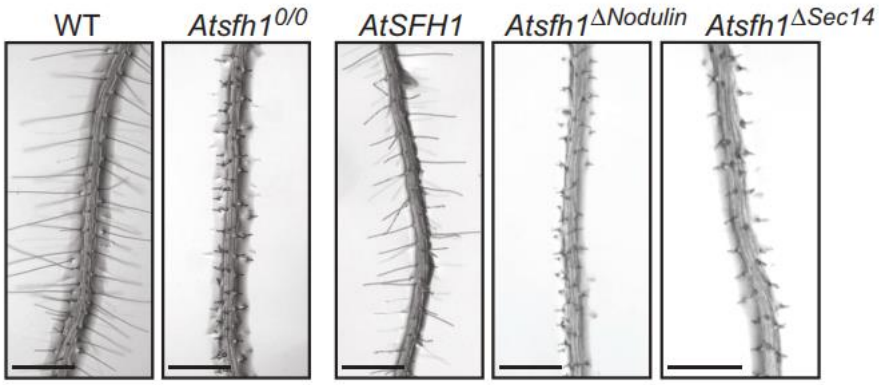


Fig. 4: Bright-field images of root hair of transgenic seedlings.

Picture from Ghosh et al. (2015) [21]. Scale bars: 1 mm.

Rhizobia bacteria infect the root hair of the host plant. A cytosolic Ca^{2+} gradient is necessary to sustain growth. The protoplast of the tip growing cell has a typical polarized organization. The cell enlargement is promoted by deposition of plasma membrane and cell wall components via a vesicle-rich tip [23].

Taking all of these important findings together one can easily see the importance of the understanding of Sec14p-like proteins and it's possible application for non-leguminous plants and the idea of manipulating its root microbiome in the future.

ERASMUS

My thesis lab work was performed in the larger framework of the Greek project BIOME. I was given the unique opportunity to stay and work in Panagiotis Moschou's lab at the University of

Crete for 8 months and was involved in the initial phase of the BIOME project. My stay was kindly funded by the European Union student exchange programme ERASMUS+ (*https://erasmus-plus.ec.europa.eu/*).

The following excerpt from the original project program description was translated and summarised from Greek to English by me.

THE PROJECT "BIOME"

is a cooperation of:

- Nikolaos P. Nikolaidis: Professor and Director of the Laboratory of Hydrogeochemical Engineering and Remediation of Soils (Technical University of Crete)
- Kavroulakis Nektarios (ELGO): Researcher at the Institute of Olive Tree Subtropical Crops and Viticulture (IOSV), Plant Pathology laboratory.
- Kidonakis Filippos (VIANAME): Messara Biological Development (https://www.bianame.gr/index.php)

The program is dedicated to crop improvement and the investigation of sustainable agriculture that is environmentally friendly and less harmful to human and animal health. The company VIANAME has recognised that modern biotechnological tools and inventions need to be implemented into domestic [Greek] agriculture in order to remain competitive and increase productivity. One of these tools for the future could be the manipulation of the root microbiome in order to minimize the impact on the environment and health due to lower inputs of biocides, fertilizers and water.

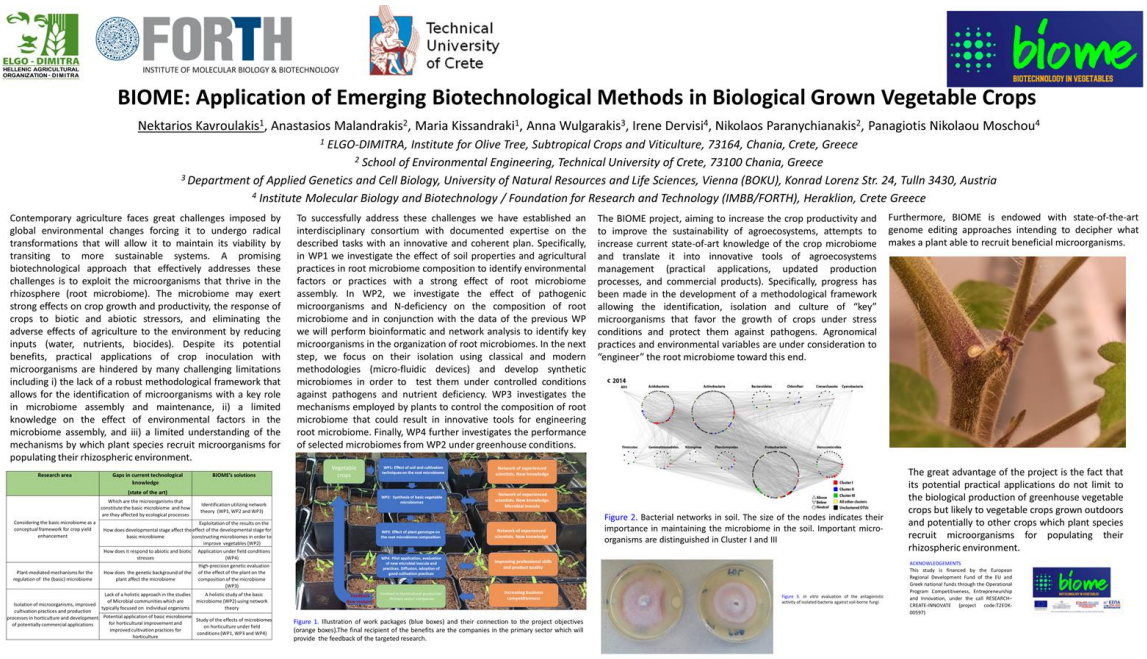
BIOMEs main objectives are:

- The understanding of the effect of soil and cultivation practices on microbiome composition of selected horticultural crops (tomato, cucumber, pepper).

- The isolation (or acquisition from international collections) of essential microorganisms of the healthy microbiome.
- The synthesis of microbiomes and the investigation of their effect on the plant under laboratory conditions.
- The investigation of the role of the plant in the composition of the root microbiome (tomato).
- The selection of microbiomes that favour the growth of the plant and give resistance to the cultivation and their investigation under field conditions.

The central idea of BIOME is the modification of the root microbiome to create genotypes with potentially modified microbiomes. After the establishment of such genotypes the effect on microbiome affiliations on plant productivity can be studied. These results will guide the selection of tomato plants from existing collections of genetic material which may show increased productivity under normal conditions or under stress.

[It should be noted that varieties produced via T-DNA insertions or CRISPR/CAS9 technology cannot be commercialized yet due to the European Union restrictions on plants derived from this type of mutagenesis. However, mutations from EMS can be used directly for commercialisation by VIANAME or even as genetic material in crop improvement through crossbreeding.]



BIOME: Application of Emerging Biotechnological Methods in Biological Grown Vegetable Crops

Nektarios Kavroulakis[1], Anastasios Malandrakis[2], Maria Kissandraki[1], Anna Wulgarakis[3], Irene Dervisi[4], Nikolaos Paranychianakis[2], Panagiotis Nikolaou Moschou[4]

[1] ELGO-DIMITRA, Institute for Olive Tree, Subtropical Crops and Viticulture, 73164, Chania, Crete, Greece

[2] School of Environmental Engineering, Technical University of Crete, 73100 Chania, Greece

[3] Department of Applied Genetics and Cell Biology, University of Natural Resources and Life Sciences, Vienna (BOKU), Konrad Lorenz Str. 24, Tulln 3430, Austria

[4] Institute Molecular Biology and Biotechnology / Foundation for Research and Technology (IMBB/FORTH), Heraklion, Crete Greece

Contemporary agriculture faces great challenges imposed by global environmental changes forcing it to undergo radical transformations that will allow it to maintain its viability by transiting to more sustainable systems. A promising biotechnological approach that effectively addresses these challenges is to exploit the microorganisms that thrive in the rhizosphere (root microbiome). The microbiome may exert strong effects on crop growth and productivity, the response of crops to biotic and abiotic stressors, and eliminating the adverse effects of agriculture to the environment by reducing inputs (water, nutrients, biocides). Despite its potential benefits, practical applications of crop inoculation with microorganisms are hindered by many challenging limitations including i) the lack of a robust methodological framework that allows for the identification of microorganisms with a key role in microbiome assembly and maintenance, ii) a limited knowledge on the effect of environmental factors in the microbiome assembly, and iii) a limited understanding of the mechanisms by which plant species recruit microorganisms for populating their rhizospheric environment.

Research area	Gaps in current technological knowledge (state of the art)	BIOME's solutions
Considering the basic microbiome as a conceptual framework for crop yield enhancement	Which are the microorganisms that constitute the basic microbiome and how are they affected by ecological processes	Identification utilizing network theory (WP1, WP2 and WP3)
	How does developmental stage affect the basic microbiome	Exploitation of the results on the effect of the developmental stage for constructing microbiomes in order to improve vegetables (WP2)
	How does it respond to abiotic and biotic stresses	Application under field conditions (WP4)
Plant-mediated mechanisms for the regulation of the (basic) microbiome	How does the genetic background of the plant affect the microbiome	High-precision genetic evaluation of the effect of the plant on the composition of the microbiome (WP3)
Isolation of microorganisms, improved cultivation practices and production processes in horticulture and development of potentially commercial applications	Lack of a holistic approach in the studies of Microbial communities which are typically focused on individual organisms	A holistic study of the basic microbiome (WP2) using network theory
	Potential application of basic microbiome for horticultural improvement and improved cultivation practices for horticulture	Study of the effects of microbiomes on horticulture under field conditions (WP1, WP3 and WP4)

To successfully address these challenges we have established an interdisciplinary consortium with documented expertise on the described tasks with an innovative and coherent plan. Specifically, in WP1 we investigate the effect of soil properties and agricultural practices in root microbiome composition to identify environmental factors or practices with a strong effect of root microbiome assembly. In WP2, we investigate the effect of pathogenic microorganisms and N-deficiency on the composition of root microbiome and in conjunction with the data of the previous WP we will perform bioinformatic and network analysis to identify key microorganisms in the organization of root microbiomes. In the next step, we focus on their isolation using classical and modern methodologies (micro-fluidic devices) and develop synthetic microbiomes in order to test them under controlled conditions against pathogens and nutrient deficiency. WP3 investigates the mechanisms employed by plants to control the composition of root microbiome that could result in innovative tools for engineering root microbiome. Finally, WP4 further investigates the performance of selected microbiomes from WP2 under greenhouse conditions.

Figure 1. Illustration of work packages (blue boxes) and their connection to the project objectives (orange boxes).The final recipient of the benefits are the companies in the primary sector which will provide the feedback of the targeted research.

The BIOME project, aiming to increase the crop productivity and to improve the sustainability of agroecosystems, attempts to increase current state-of-art knowledge of the crop microbiome and translate it into innovative tools of agroecosystems management (practical applications, updated production processes, and commercial products). Specifically, progress has been made in the development of a methodological framework allowing the identification, isolation and culture of "key" microorganisms that favor the growth of crops under stress conditions and protect them against pathogens. Agronomical practices and environmental variables are under consideration to "engineer" the root microbiome toward this end.

Figure 2. Bacterial networks in soil. The size of the nodes indicates their importance in maintaining the microbiome in the soil. Important micro-organisms are distinguished in Cluster I and III

Figure 3. In vitro evaluation of the antagonistic activity of isolated bacteria against soil-borne fungi

Furthermore, BIOME is endowed with state-of-the-art genome editing approaches intending to decipher what makes a plant able to recruit beneficial microorganisms.

The great advantage of the project is the fact that its potential practical applications do not limit to the biological production of greenhouse vegetable crops but likely to vegetable crops grown outdoors and potentially to other crops which plant species recruit microorganisms for populating their rhizospheric environment.

ACKNOWLEDGEMENTS

This study is financed by the European Regional Development Fund of the EU and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH-CREATE-INNOVATE (project code:T2EDK-00597)

Fig. 5: Project poster of "BIOME"

This poster has been created by Dr. Panagiotis Moschou after my stay at UOC and summarises the key points of the project tiers.

MONEYMAKER

Since the BIOME project demanded the use of crop plants (see first objective) the question arose which variety to use. For the purpose of simplicity the decision was made to use the English heirloom "Moneymaker". For one, it is a real variety and not an F1 hybrid which eliminates the problem of genetic segregation in the next generation and secondly, plenty of evidence can be found, that this variety is indeed transformable. Sadly, this cannot be said from most local Greek varieties which obviously would have been the preferred choice of the company VIANAME. Moneymaker shows high regeneration capability in classical tissue culture transformation and therefore, should also respond well in a tissue culture free method [24], which will be explained in the chapter "*De Novo Induction of Meristems*". Moneymaker usually also shows early ripening of the fruits and good productivity, meaning sufficient seeds can be harvested for seed selection later on.

SEC14-LIKE PROTEINS IN TOMATO

In Böhme et al. (2004) a protein in Arabidopsis was described (CAN OF WORMS1 [COW1]; other name: AtSFH1) that belongs to a subfamily of proteins showing a two-domain arrangement in which the N-terminal Sec14p coding region is linked to a C-terminal Nlj16 coding region (previously described as LjPLP-I–IV in *Lotus Japonicus*). LjPLP-IV protein shows 65% similarity with COW1 encouraging the suspicion that it has a similar function or interferes with tip growth or is a target for rhizobia bacteria [25].

```
          1                                                                                                            112
LjPLP-IVp -----------msellssgtdsfkkkainasnmlrnsltrkgrrs-skvmsveiedvhdaeelkaveefrqalisddllpakhddyhmmlrflkarkfeidkskqmwsdmlk
COW1 p    maetkpeiemseeerkivkisslkkkainasnrfknsfkkkgrrsssrvmsvpieddidaedlqaldafrqalildellpsklddlhmmlrflrarkfdiekakqmwsdmiq

.............................................................................................................................

          113                                                                                                          224
LjPLP-IVp wrkefgadtiveefefkeidevlkyypqghhgvdkegrpvyieqlgqvdatklmqvttmdryikyhvkefektfdlkfaacsiaakkhidqsttildvqgvglksfnkhare
COW1 p    wrkdfgadtiiedfdfeeidevmkhypqgyhgvdkegrpvyierlgqidankllqvttmdryvkyhvkefektfkvkfpscsvaankhidqsttildvqgvglknfsksare

.............................................................................................................................

          225                                                                                                          336
LjPLP-IVp lvtriqkvdgdnypetlnrmfiinagsgfrilwntvksfldpkttakinvlgnkydtklleiidaselpeflggtctctdqggcmrsdkgpwkdeeilrmvqngahkcsrkp
COW1 p    llqrlckidnenypetlnrmfiinagsgfrllwstvksfldpkttakihvlgnkyhsklleividaselpeffggactcedkggcmrsdkgpwndpevlkiainreakcspis

.............................................................................................................................

          337                                                                                                          448
LjPLP-IVp eshgeeekpisedktskfd------------enltpqvspvcdqvpaaklsknv--dvipv-----------------------------------aipvgvpvavpaadk-
COW1 p    edehkhvdq--grstsgfeslerikkktdednvyekqiatidksmdmawlaktqkaenfpiskglecyvrkgapkkgdgllvggvmafvmgivamvrlskdvprklteaaly

.............................................................................................................................

          449                                                                                                      556
LjPLP-IVp ----------naskkvg--qndttskefttvmkrmaeleekmttmnhqpatmppekeemlnatisradvlekqlmdtkkaledslakqevlsayvekkkqkkktffcc
COW1 p    gnsvcyeestkskqnqgqfaapvssseymlmvkrmaeledkcmfldlkpahvesekeeklqaalnrvqvleqeltetkkaleealvsqkeilayiekkkkkkklffgf

.............................................................................................................................
```

Fig. 6: Own repeat of protein sequence alignment of LjPLP-IV with COW1.

Sequences obtained from NCBI. Alignment Statistics: Length: 556, Identical Sites: 291, Pairwise Identity: 52%.

To find a homologous gene in tomato a Basic Local Alignment Search Tool (BLAST) was carried out in Sol Genomics Network (*https://solgenomics.net/*) using the genomic sequence of LjPLP-IV (AF367434_1).

Untitled_sequence vs Tomato Genome cDNA (ITAG release 2.40)

SubjectId	id%	Aln	evalue	Score	Description
Solyc01g109870.2	76.92	210/273	9e-35	150	SEC14-like protein (AHRD V1 **-- Q9T026_ARATH); contains Interpro domain(s) IPR001251 Cellular retinaldehyde-binding/triple function, C-terminal Length=2433
Solyc11g040280.1	75.37	205/272	4e-28	128	SEC14-like protein (AHRD V1 **-- Q9T027_ARATH); contains Interpro domain(s) IPR001251 Cellular retinaldehyde-binding/triple function, C-terminal Length=1797
Solyc11g040280.1	89.36	84/94	2e-25	119	SEC14-like protein (AHRD V1 **-- Q9T027_ARATH); contains Interpro domain(s) IPR001251 Cellular retinaldehyde-binding/triple function, C-terminal Lengt ...
Solyc01g109860.2	75.64	208/275	4e-28	128	Phosphatidylinositol/phosphatidylcholine transfer protein-like (AHRD V1 **-* Q9LRX7_ARATH); contains Interpro domain(s) IPR001251 Cellular retinaldehyde-binding/triple function, C-terminal
Solyc10g053970.1	75.09	202/269	5e-27	124	SEC14-like protein (AHRD V1 **-- Q9T026_ARATH); contains Interpro domain(s) IPR001251 Cellular retinaldehyde-binding/triple function, C-terminal Length=1872
Solyc10g053970.1	84.38	81/96	4e-18	95.3	SEC14-like protein (AHRD V1 **-- Q9T026_ARATH); contains Interpro domain(s) IPR001251 Cellular retinaldehyde-binding/triple function, C-terminal Lengt ...
Solyc10g053900.1	73.90	201/272	2e-21	106	SEC14-like protein (AHRD V1 **-- Q9T027_ARATH); contains Interpro domain(s) IPR001251 Cellular retinaldehyde-binding/triple function, C-terminal Length=1710
Solyc10g053900.1	84.47	87/103	2e-20	102	SEC14-like protein (AHRD V1 **-- Q9T027_ARATH); contains Interpro domain(s) IPR001251 Cellular retinaldehyde-binding/triple function, C-terminal Lengt ...
Solyc09g060090.2	83.33	65/78	2e-11	73.1	SEC14 cytosolic factor family protein (AHRD V1 **-- D7KN07_ARALY); contains Interpro domain(s) IPR001251 Cellular retinaldehyde-binding/triple function, C-terminal Length=2007

Fig. 7: BLAST hit with the highest E-value

Red arrow: Solyc01g109870.2. Description: SEC14-like protein with Interpro domain(s), Cellular retinaldehyde-binding/triple function, C-terminal Length=2433; Sequence identity: 76.92%.

BLAST-ing the genomic sequence of Solyc01g109870.2 in NCBI obtained a locus (Accession: XM_004230820) named:

"PREDICTED: *Solanum lycopersicum*

phosphatidylinositol/phosphatidylcholine transfer protein

SFH8 (LOC101265778), mRNA"

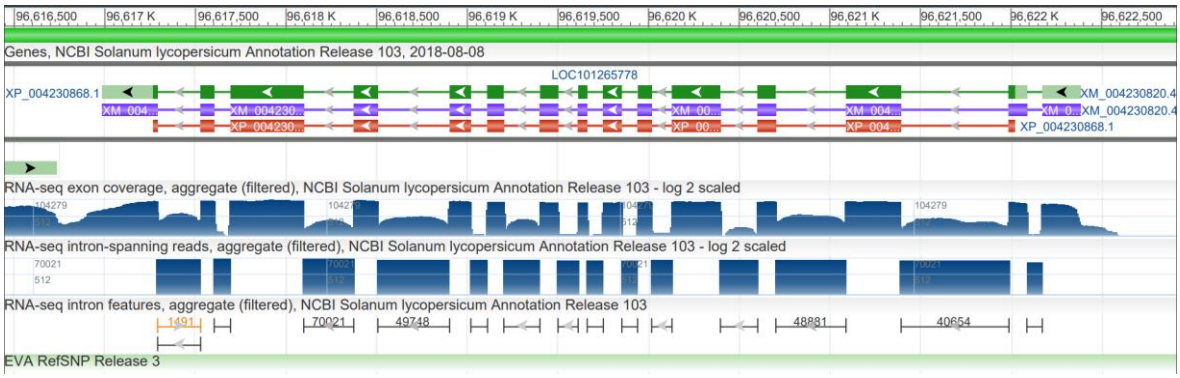


Fig. 8: Genome data viewer by NCBI.

Assembly: SL3.1 (GCF_000188115.5), Chr 1 (NC_015438.3), Exon count: 15.

By knocking out the *Sfh8* gene in tomato it is hypothesised that one could observe a similar phenotype as previously seen in *Arabidopsis* (e.g. distorted or shortened root hair) and potentially a different microbiome composition. It should also be noted that the goal of this project is crop research and not necessarily industrial use (yet).

Plasmid assembly was accomplished using state of the art Golden Gate cloning technique.

Golden Gate Cloning

An essential element in synthetic biology, or genetic engineering, is the potential of assembling many diverse DNA fragments. Over the past decades the majority of this has been done via homologous recombination which provides independence from any restriction sites present in the fragments to assemble [26]. However, optimising phenotypes requires the ability of generating various coding sequences and many variants of regulatory sequences [26]. Like in other divisions of engineering one first step would be the standardisation of a system's basic parts. It took 15 years from NOMAD [27] to BioBrick standards [28] to finally establish a system that is able to even overcome restrictions like the assembly of multiple DNA fragments in a single step. The essential characteristic of Golden Gate cloning is the ability of type II restriction enzymes to cleave outside their recognition site thereby removing restriction sites during the cleavage process. What's left are fragments flanked by compatible sequence overhangs that can then be ligated seamlessly [26]. Restriction-ligation makes it possible to assemble multiple fragments in one step (one tube reaction) with extremely high efficiency. Additionally, these compatible DNA fragments can then be assembled in a linear manner simply depending on the individual design of the overhangs, which can be any four+ nucleotide sequence of choice [29].

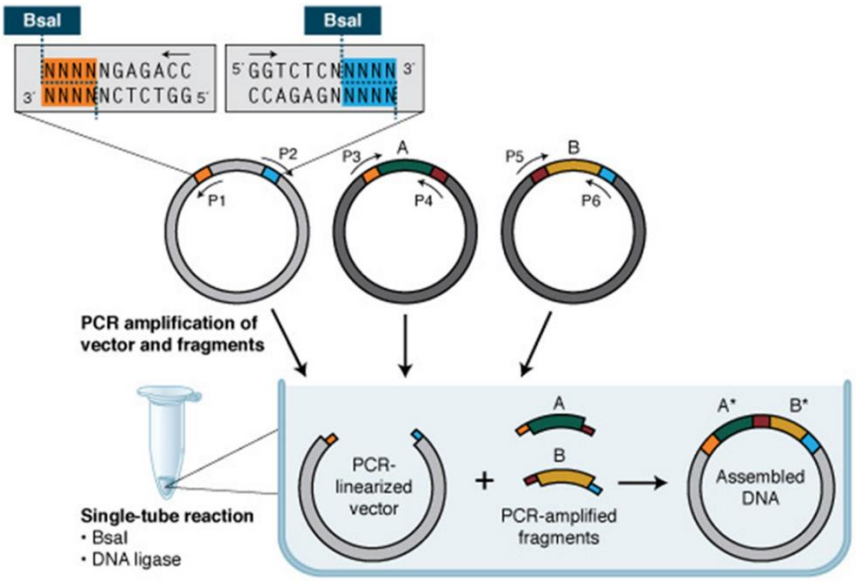


Fig. 9: Scheme of plasmid assembly in a one-tube-reaction

using a type II restriction enzyme (BsaI) plus DNA ligase. Picture from *https://www.neb-online.de/*. Fragments, including matching restriction sites, coding for any desired sequence can be produced via PCR (P1-P6 = primers). The enzyme will cut outside its recognition site (*Fig. 10*) leaving behind so called "sticky ends" which are 4 base pair overhangs that correspond to those on the fragments created via PCR. DNA ligase will seamlessly "glue" the fragments and the vector together.

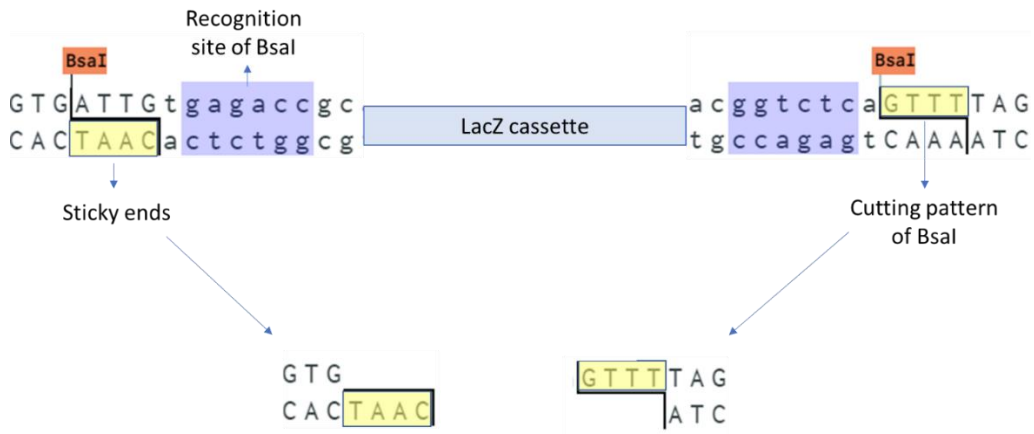


Fig. 10: Principle of Golden Gate Cloning.

BsaI type II restriction enzyme cuts outside its recognition sites (highlighted in blue) and therefore removes the sequence. It leaves behind a specific 4 base pair overhang – aka "sticky ends" (highlighted in yellow) which can be utilised to construct inserts which carry the

complementary sequence for seamless ligation. This way, self-ligation of the vector is also not possible.

The plasmids were delivered to the plant via *Agrobacterium tumefaciens*.

AGROBACTERIUM-MEDIATED TRANSFORMATION

The naturally occurring soil phytopathogen *Agrobacterium tumefaciens* (from the family of *Rhizobiaceae*) infects plant wounds and causes crown gall disease by delivering a piece of its large virulence plasmid (tumor inducing Ti-plasmid), the so called transfer DNA (T-DNA) into a host plant cell via a bacterial type IV secretion system (T4SS) [30]. By replacing the oncogenes with a gene of interest the system can be easily utilised in plant biotechnology to manipulate traits and gain insight in how plants function. The transformation process requires two genetic components, the border sequences at the ends of the T-region (25 bp imperfect repeats) and the virulence (vir) region, coding for components of the bacterial protein machinery [31]. Vir gene products nick the T-DNA region at its left (LB) and right border (RB) and transfer the T-DNA into plant cells. In modern plant biotechnology binary vector systems are used for Agrobacterium-mediated transformation. These vectors are small plasmids containing a (multiple) cloning site and selectable markers between the left and right border, that can replicate both in *E. coli* and *Agrobacterium*. To ensure transformation without tumorigenicity, modified (disarmed) Ti-plasmids are used. They lack the T-DNA but retain the entire vir-region. Unfortunately, only a small number of Ti-plasmids from different *Agrobacterium* species have been disarmed so far [32]. Generally, *A. tumefaciens* is well suited to transform various host cells including dicots, monocots, angiosperm and gymnosperm species [33]. Two types of transformation methods can be distinguished, a) transient transformation where expression usually only lasts for a couple of days [34] and b) stable transformation, where the T-DNA is integrated into the hosts genome and subsequently passed on to the next generation [35] [36]. Mutations introduced into the plant genome by transformation are usually transmitted to the progeny in a Mendelian segregation pattern 3:1. However, it can happen that the T-DNA is inserted into two or more sites of the plant genome. Ondřej et al. (1999) [37] could show that in *Arabidopsis*, amongst 49 lines tested, 72% segregated in a monohybrid or dihybrid Mendelian

segregation, 6% of the lines segregated for more than two T-DNA inserts and 22% showed non-Mendelian patterns of T-DNA inheritance.

De Novo Induction of Meristems through Developmental Regulators

Plant development is a complex network of transcription factors, hormones, enzymes, microRNAs, and many more, which all together regulate developmental processes such as embryo, apical root and shoot meristem, leaf, flower and seed formation. The response of these processes to a changing environment even has become an interesting new research area [38] [39].

The development from a single cell zygote into an embryo by asymmetric cell division marks the beginning of the life cycle of a plant. In dicotyledon plants like e.g. *Arabidopsis* the embryo consists of a basal root meristem, a central region hypocotyl and two seed leaves flanking a shoot apical meristem (SAM) which can be seen as the pantry of undifferentiated stem cells eventually yielding the adult root architecture [40]. However, many species in plants are not strictly dependent upon fertilisation. They can either naturally produce asexually derived embryos in the seed (apomixis) or can be induced to do so in tissue culture [41]. In tissue culture plant growth regulators initiate the formation of a callus or somatic embryogenesis. Therefore, finding the regulatory genes behind this phenomenon is a compelling approach to circumventing traditional methods.

It was discovered that the expression of maize (*Zea mays*) *Wuschel2* (*Wus2*) and *Baby Boom* (*Bbm*) in monocots promote somatic cells to form embryos that develop into whole plants. Co-delivering transgenes together with these developmental regulators (DR) expedites the production of transgenic plants and might make tissue culture more or less redundant [35]. Tissue culture is the biggest bottleneck in the production of a transgenic plant. It works only in a handful of species, is time consuming, requires more labour and equipment as well as financial resources and it often results in unexpected changes to genomes [42].

The following two growth regulators have been chosen for transformation in this experiment.

Wuschel2 (WUS2)

The bifunctional homeodomain transcription factor *Wus2*, which is expressed in the organising centre of the shoot apical meristem (SAM) and is an essential factor for de novo establishment of the shoot stem cell niche. It induces the direct somatic embryo formation in tissue culture and works synergistically with other developmental growth regulators, such as *Bbm*.

Isopentenyl Transferase (*ipt*)

Ipt is a gene on the T-DNA from *Agrobacterium tumefaciens*. It encodes an enzyme that catalyses formation of isopentenyl-adenosine-5-monophosphate, the first intermediate in cytokinin biosynthesis. The integration of a T-DNA into a plant host cell has two consequences, a) tumour formation and b) the regeneration of physiologically abnormal shoots that are derived from transformed protoplasts [43]. Using the *ipt* gene under the control of a more active promoter (e.g. 35S) increases cytokinin levels in transgenic plants, showing enhanced shoot organogenesis [44].

Breaking the Tissue Culture Bottleneck with Developmental Regulators

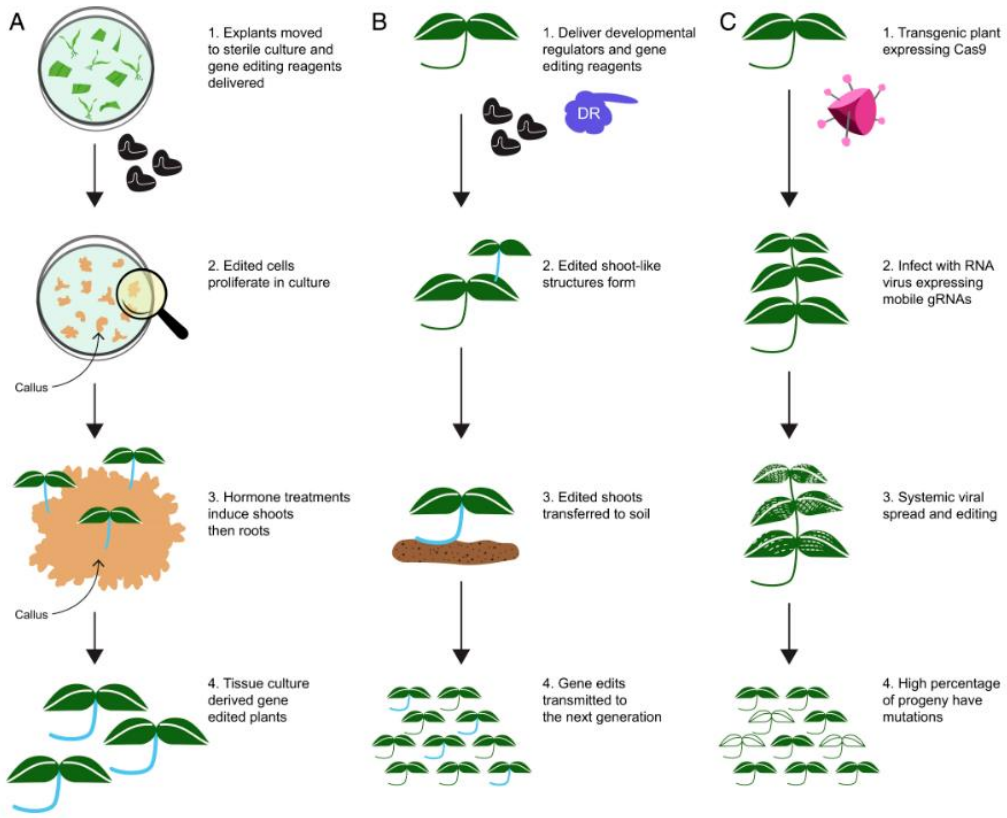


Fig. 11: Approaches for genetic variation in plants.

Picture from Nasti et al. (2021) [9]. (A) Gene editing through tissue culture. (B) Gene editing through de novo meristem induction. Agrobacterium delivers plasmids coding for developmental regulators (DR). DR trigger growth of a new meristem from edited somatic cells. The meristem ultimately forms a shoot which is excised and roots are induced. Progeny of the plant (seeds) can be selected for the mutation of interest. (C) Gene editing with RNA viruses.

DR in concert with cytokinin biosynthesis can be used to induce the growth of new meristems on soil-gown plants. Co-delivered reporter genes, such as Luciferase for example, make it possible to recover shoots that are either transgenic or gene edited. When the shoots produce flowers and seeds the gene edits are transmitted to the next generation. In Maher et al. (2020) [35] proof of concept was carried out in *Nicotiana benthamiana*. The possibility of creating gene-edited shoots sidesteps the process of tissue culture and will accelerate crop plant production in the future. Natural genetic diversity will be supplemented with synthetic diversity. Increased

phenotypic diversity could equip crop plants with enhanced stress responses as well as enhance or introduce secondary metabolic pathways to produce metabolites for food or energy production.

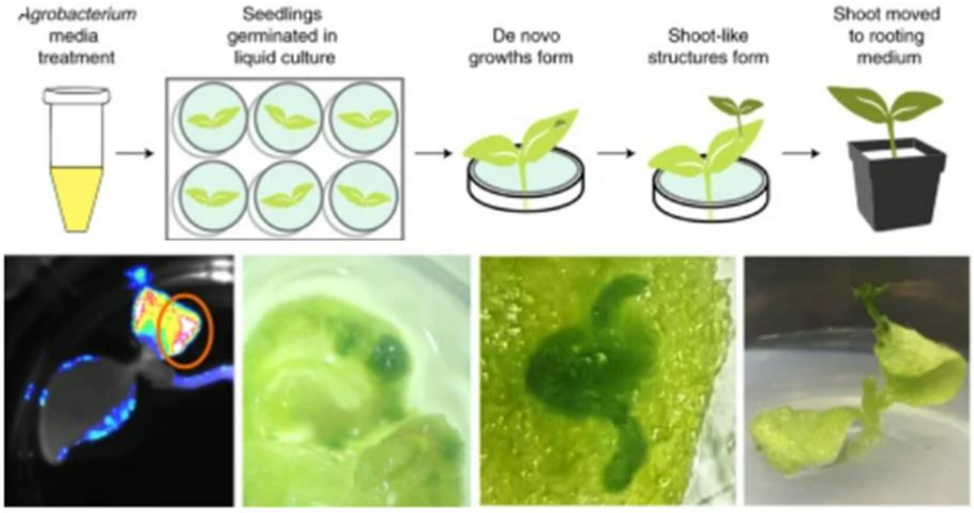


Fig. 12: Ectopic delivery of DR induces the growth of new shoot-like structure.

Picture from Maher et al. (2020) [35]. Fast-TrACC (Fast Treated Agrobacterium Co-Culture) delivers reagents to seedlings and uses a luciferase reporter to monitor and calibrate the efficiency of transformation.

Back-up Plan

Since all of the above mentioned techniques and desired crop plants were a new addition to the labs commonly used procedures, there was no guarantee that the projects delivery, that is, the production of a tomato mutant with potentially altered root hair and microbiome inhabitation, could be fulfilled in time. For this reason, a parallel aim was to comb online genome databases and potentially find a line that already has a mutation in the target gene (see chapter *MicroTom Mutant*).

Materials & Methods

Software

The Benchling Life Sciences R&D Cloud:

Benchling [Biology Software]. (2021). Retrieved from *https://benchling.com*.

MicroTom Mutant

One deliverable of the project included the creation of a genetically modified tomato line compliant to the EU Directive 2001/18, regulating GM crops for environmental release [45]. This excludes the use of the novel CRISPR/Cas technique as well as T-DNA insertion lines and leaves EMS treatment of seedlings as the only option. EMS treatment itself is generally a very quick procedure but creates many unwanted background mutations and it takes years to get rid of them by continuous selection and back crossing. It was therefore the aim to find a tomato mutant seed bank online that already lists a mutant with a mutation in the target gene.

This search was unsuccessful because most listed EMS lines are not sequenced and can only be selected by phenotype. T-DNA insertion lines however, do generally mention "flanking regions", which can be generated via TAIL-PCR and can therefore be aligned to a target region. One such line was found on the website *https://tomatoma.nbrp.jp/*, a tomato mutant archive from the University of Tsukuba in Japan. It was ordered as a backup plan. Line Information can be seen in *Tab. 1*.

Tab. 1: Information about the T-DNA tagged line from the TOMATOMA database

Strain ID	TOMJPT00175
Strain Type	T-DNA-tagged lines
Hit Region	SL2.50ch01:9468501-9470300
Flanking Sequence	ATTTAAGGTAT-GTTAAGACTTTT-TTCTT AGACTTGTTGAAGGACGTTTTCCTAATT AAAAATTAAAAATGAAAATAGACAAAG GGGTAAGGGGGAAAGATGGTGGTCTC GTATCAATGTGACGGGCATTGGTACGA GTACCGGTGTTATAAAAAGGAAAACAC TATTATAGAATGTGGATTGAAATTTGAG AATGCCAAAGCATATGGGCATTAGCTG ATATATTATTGACTTGAATTCCTTGTGT GATTGTGTTTTATTTATTTCACCCGTATA GTTGAGATAATTGAGGTGGTTATGTAT TATATTCATATTGATTGATTGAGATGCA TCATCATTCCCTCTATTGAAACAATATTG TGCACATGCATCGAGATGAGACTGAGT ATAAGTTGGGCACGTGGAGATCGTCCG TGCTGGGGATGGTGAGATGTTAAGATT GTAATTTGGGCACGTGGAGATCGTCCA TGCGAAAATTGTTTGATATTATGATAGT GCGTTGAGATCGTCCGCACAGACACGT GGAGATCGTCCGTGTCGGTATATGGAC CTCGCGAGTCCCCCATGGGTCATGAACT CTCGATGTATTTCCGAGGAGTATCATGT ATATACGGTTGAGTGAGTACTGGGTAT TCTGAGACATATCATTACATGCATCATA TTGCATTGCATTTCATCCATCATTCATTC CTGATGACTWATGTTTCGTTGGTGTTG AATATTAAATGGTTGKAATTCCTCTTTA ATTKAT

To check the T-DNA insert the following primer sequences were provided by TOMATOMA:

FWD: AGAACTCGTCAAGAAGGCGA

REV: CTGAATGAACTGCAGGACGA

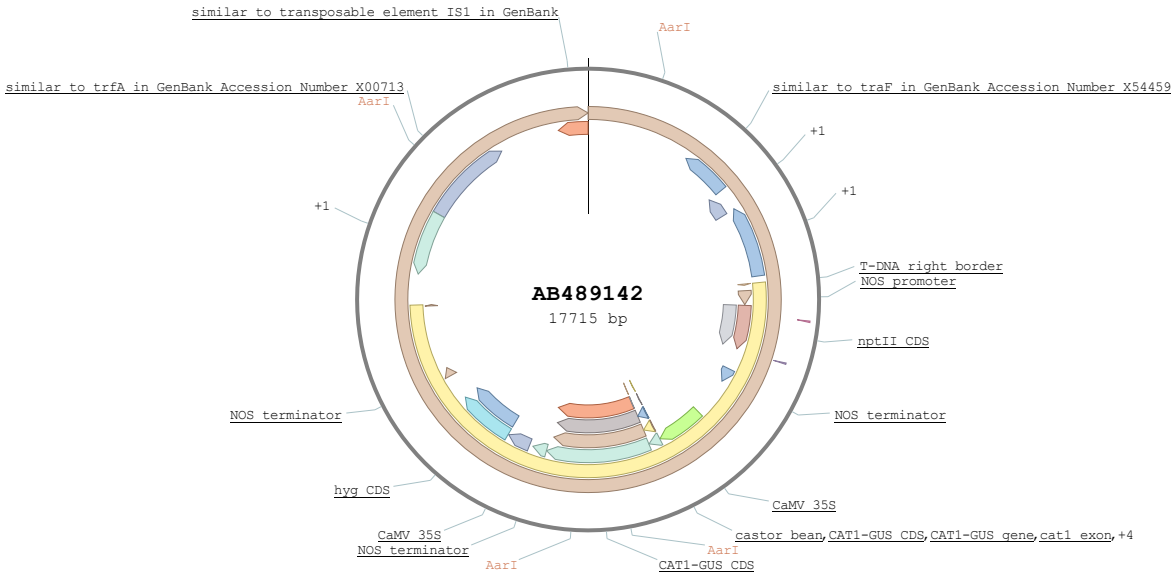


Fig. 13: T-DNA vector map from Japan

of the plasmid that was used to create the insertion mutant. In yellow: T-DNA including *nptII* gene. Length: 9046 bp.

For propagation purposes, seeds were washed with ethanol and placed directly onto soil to germinate. The seedlings were left to grow in a growth chamber (16h light/8h dark, constant 25°C, humidity 60%) and re-potted into bigger pots after about 4 to 5 weeks.

Moneymaker Plants

All plant seeds used were derived from university-intern propagation (green house manager, former group members). For pure propagation purposes the seeds were washed with ethanol and placed directly onto soil. They were left to germinate in an air conditioned green house and re-potted into bigger pots after 4 to 5 weeks. At that point the plants were also tied to a wooden stick to ensure upright growth. Side shoots were continuously pinched out during apical shoot growth. When fruit carrying branches grew too long and heavy (BBCH scale: Principal growth Stage 5: Inflorescence emergence, [46]), they were tied up with strings and mounted to the ceiling. This not only ensures straight growth and leaves room for aeration, which is particularly important in a very hot climate like Crete, it also makes harvesting a lot easier and saves space.

Seed Collection

Mature fruit could be harvested after roughly 12 weeks (BBCH scale: Principal growth stage 8: Ripening of fruit and seed). Fruit were left outside until the mesocarp became soft. At that point it is easy to squeeze them into a fine sieve and separate them from the rest of the fruit. Under running water the seeds were separated from the pericarp as much as possible and then placed on baking paper in order for them not to stick to any tissue during the drying process. After a few days dry seeds could be collected into microtubes and stored on 4 degrees in the dark.

Quick DNA Extraction from Plants

Plant leaf discs were taken with the lid of a sterile 1.5 ml microtube, placed into liquid nitrogen and grinded directly in the tube. 400 μL of freshly prepared extraction buffer (200 mM Tris HCl pH 7.5, 250 mM NaCl, 25 mM EDTA, 0.5 % SDS) was added, mixed vigorously and centrifuged for 15 min at 13 000 rpm. 300 μL of the supernatant were then transferred into a new microtube together with 300 μL of isopropanol. After inversion and a 2 min incubation at room temperature the tubes were centrifuged for 10 min at 13 000 rpm. The supernatant was then discarded and the pellet washed with 70% ethanol. The ethanol was discarded after a 5 min centrifugation step and the pellet was then left to dry under a fume hood by placing the tubes upside-down on a paper towel. The dry pellet was then dissolved in 30-50 μL of TE buffer.

PCR Protocols

- Kapa2G Fast HotStart ReadyMix PCR kit (Molecular BioProducts, Inc. & Sigma-Aldrich Co. LLC) *https://www.sigmaaldrich.com/GR/en/product/roche/2gfhsrmkb*

- Taq DNA polymerase by MINOTECH IMBB-FORTH© *https://minotech.gr/index.php/products/taq-dna-polymerase*

- Phusion® High-Fidelity DNA polymerase by New England BioLabs Inc. (NEB) *https://international.neb.com/products/m0530-phusion-high-fidelity-dna polymerase#Product%20Information*

PRIMERS & GUIDE RNAs (gRNAs)

All primers were provided by ©Macrogen Inc.

https://dna.macrogen-europe.com/eng/member/login.jsp?backURL=/eng/order/oligo/order_step0.jsp

All primer- and gRNA sequences were designed with Benchling Biology Software.

Tab. 2: List of all primer- and gRNA sequences designed for this work

name	sequence	description
Seq Primer (G1:Ex3) FWD	tcttggttctcaacagtcgataca	For sequencing of exon 3 of the *Sfh8* gene
Seq Primer (G1:Ex3) REV	aggggagggtaggaagctaagca	
Seq Primer (G2:Ex13) FWD	tctccatatgcctgcatgcatct	For sequencing of exon 13 of the *Sfh8* gene
Seq Primer (G2:Ex13) REV	tgccttgtttttggttatgtcctgt	
Sfh8 G1 (BsaI) FWD	ATTGtaacacggcatgtaactgag	Guide 1 for targeting exon 3 with BsaI overhangs
Sfh8 G1 (BsaI) REV	AAACctcagttacatgccgtgtta	
Sfh8 G2 (BpiI) FWD	GTGAagcatcccgaatgtcctcga	Guide 2 for targeting exon 13 with BpiI overhangs
Sfh8 G2 (BpiI) REV	CTAAtcgaggacattcgggatgct	
Guide Scaffold (AarI) FWD	agcgcccacctgccaggggaccgctactaGAATTCGAGCTCGGAGCATC	Primers for amplification of the guide scaffold with the addition of AarI overhangs
Guide Scaffold (AarI) REV	tttcatcacctgcgtcaccggTCTGAGCGTAATGCCAACTTTGTAC	
SV40 NLS FWD	agaggaaggtttgacgtcgacga	Sequencing/PCR of guide scaffold in pTRANS_Cas9 & pTRANS_Luc
NOS REV	attctccgctcatgatcggcgc	
HSP term FWD	cagagctctggtgacggac	
T-DNA FWD (from JP)	agaactcgtcaagaaggcga	Verify T-DNA insert in MicroTom mutant from Japan
T-DNA REV (from JP)	ctgaatgaactgcaggacga	
Flanking Seq FWD	caaacaagtcctcggaacaagt	

Primer stock concentration: 100 µM

Primer Annealing

2 primers (e.g. G1 FWD & G1 REV), that when annealed together create one double stranded gRNA sequence, were diluted in annealing buffer (10 mM Tris, pH 7.5 - 8.0, 50 mM NaCl, 1 mM EDTA) from stock concentration to a 1:50 ratio, then measured with a spectrophotometer and if necessary, adjusted to an equal concentration. Both primers were then mixed together (1:1) in a microtube and placed into a heat block on 95°C for 5 min. The tube was then left to slowly cool down at room temperature for at least 1h or overnight before further use. Heating followed by cooling facilitates hybridization.

Guide Scaffold Cloning Vectors

To minimise off-target effects and reach maximum efficiency of mutagenesis the vectors used allow generation of customized dual-gRNA expression constructs via one-step Golden Gate cloning due to insertion of annealed oligonucleotides with different overhangs (*Fig. 16*) [47].

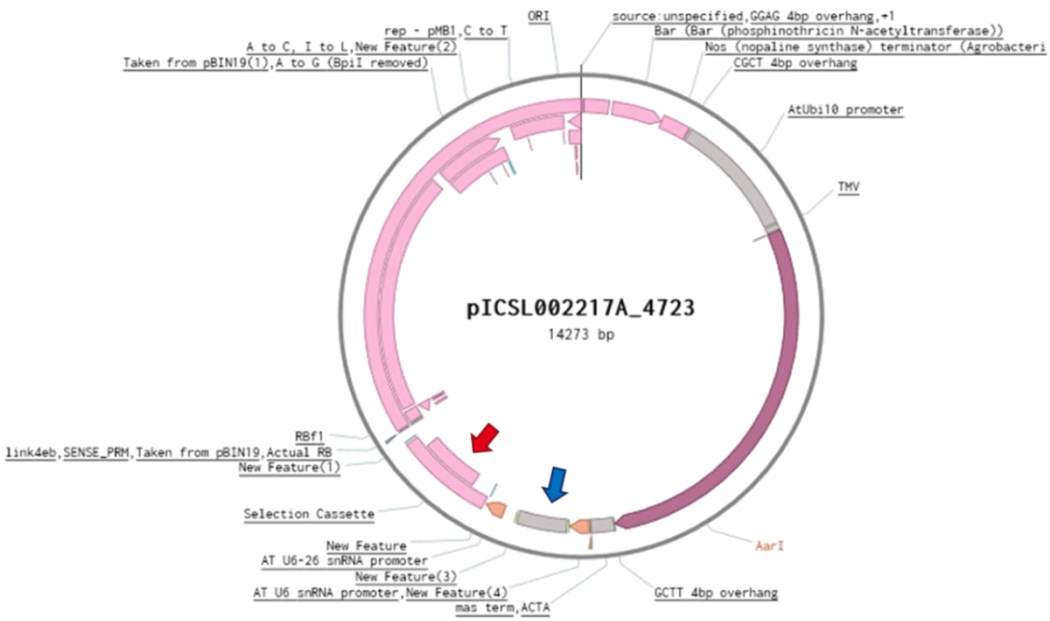


Fig. 14: Vector map of pICSL002217A (17A).

Vector includes a LacZ cassette (blue arrow) and an RFP cassette (red arrow). These two can each be replaced with one gRNA sequence. The vector already contains the necessary U6 promoter (yellow) and "guide tail" for both inserts.

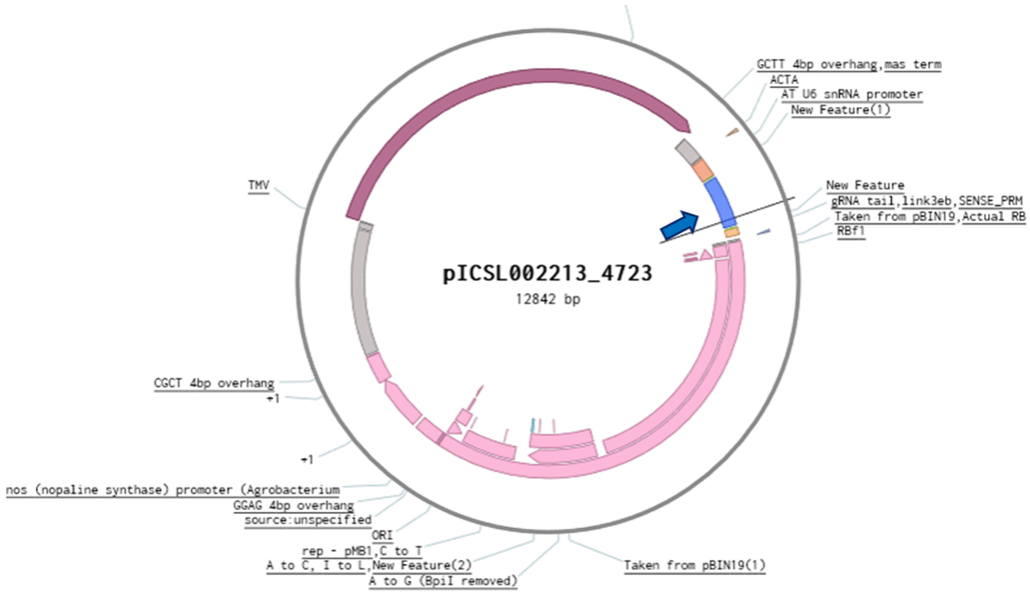


Fig. 15: Vector map of pICSL002213 (13).

Vector includes one LacZ cassette (blue arrow). It can be replaced with one gRNA sequence. The vector already contains the necessary U6 promoter and "guide tail".

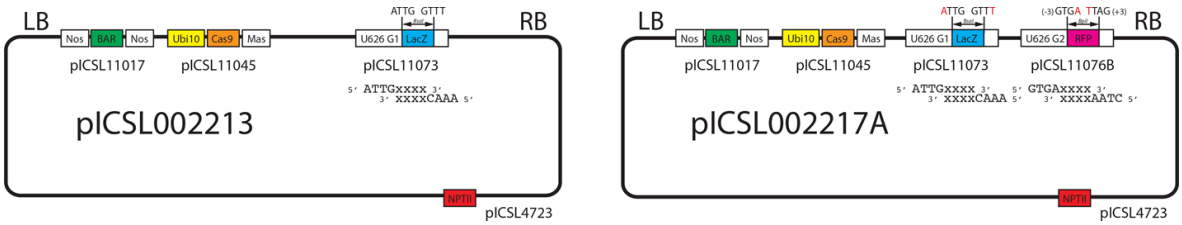


Fig. 16: Schematic representation of cloning vectors 17A and 13.

Replacement cassettes with different overhangs for one-step Golden Gate assembly. LacZ with BsaI and RFP with BpiI overhangs.

Blue-White Selection

Screening was performed by adding 40 µL of X-gal solution (20 mg/mL) and 40 µL of 100 mM IPTG solution onto an LB agar plate and spread evenly with a sterile spatula. Preparation of the plates was performed in a laminar flow hood.

Digestion-Ligation (dig-lig) Reaction

For seamless insertion of 2 guide sequences into vector 17A two slightly different reactions were performed, one followed by another.

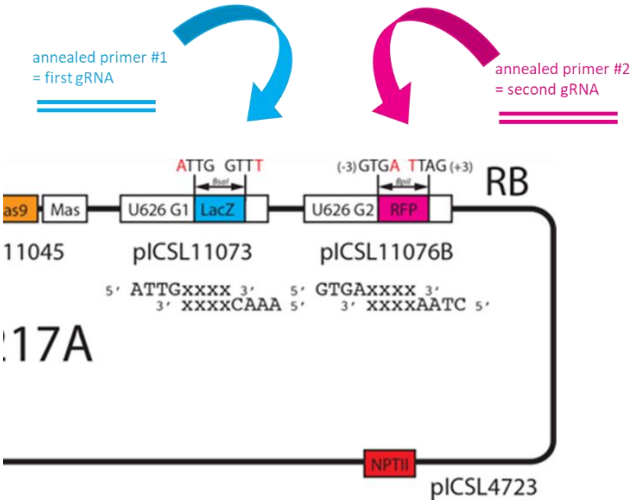


Fig. 17: Assembly of cloning vector.

First dig-lig reaction performed with vector and first gRNA. Duration: 3h. Second dig-lig reaction performed directly after, with vector + first gRNA and second gRNA. Duration: again 3h.

First reaction was performed using 100 ng of vector, 1 µL of oligo mix 1, 1 µL of NEB ligase buffer (10x), 1 µL of BSA (10x), 0.5 µL of BsaI enzyme and 0.5 µL of NEB ligase in a 10 µL reaction. Mixture was placed into a thermocycler and cycled after the following scheme: 37°C – 20", 37°C – 3", 16°C – 4" [15x], 16°C – 30", 50°C – 5", 80°C – 15" for a total of 3 hours.

After that, the second reaction was performed using 1 µL of oligo mix 2, 0.5 µL of NEB ligase buffer (10x), 1.5 µL of BSA (10x), 1 µL of ATP (10 mM), 0.5 µL of BpiI enzyme and 0.5 µL of NEB ligase in a 5 µL reaction. This mixture was cycled using the same program as described above.

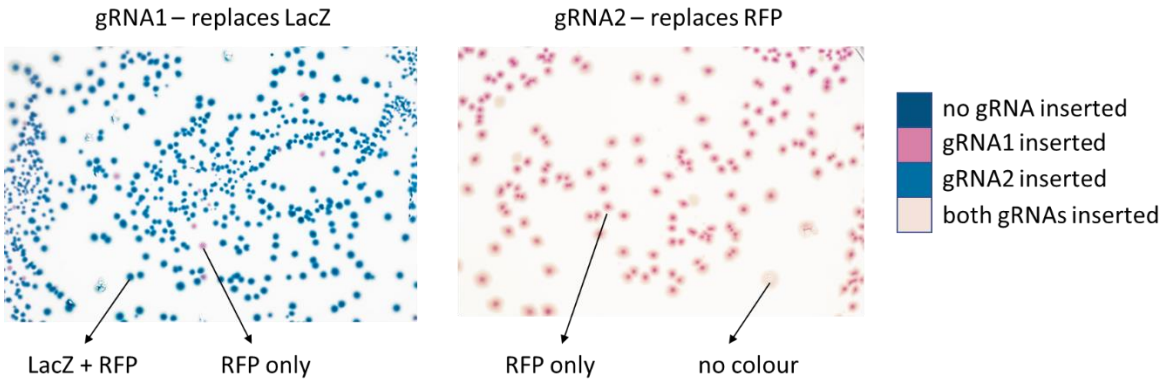


Fig. 18: Schematic colour key for selection.

gRNA1 replaces LacZ in first reaction. RFP is still functioning in the vector and therefore, shows successfully transformed colonies in red. With those the second reaction shall be performed. When gRNA2 replaces RFP no colour will can be displayed anymore. After successful insertion of both gRNAs the colonies to pick shall be white-ish.

Vector System for Plant Transformation with Growth Regulators

All vectors used during this procedure were designed by *https://www.voytaslab.com/* [35] and can be obtained from Addgene on the basis of a material transfer agreement for non-commercial research purposes *https://www.addgene.org/Daniel_Voytas/*.

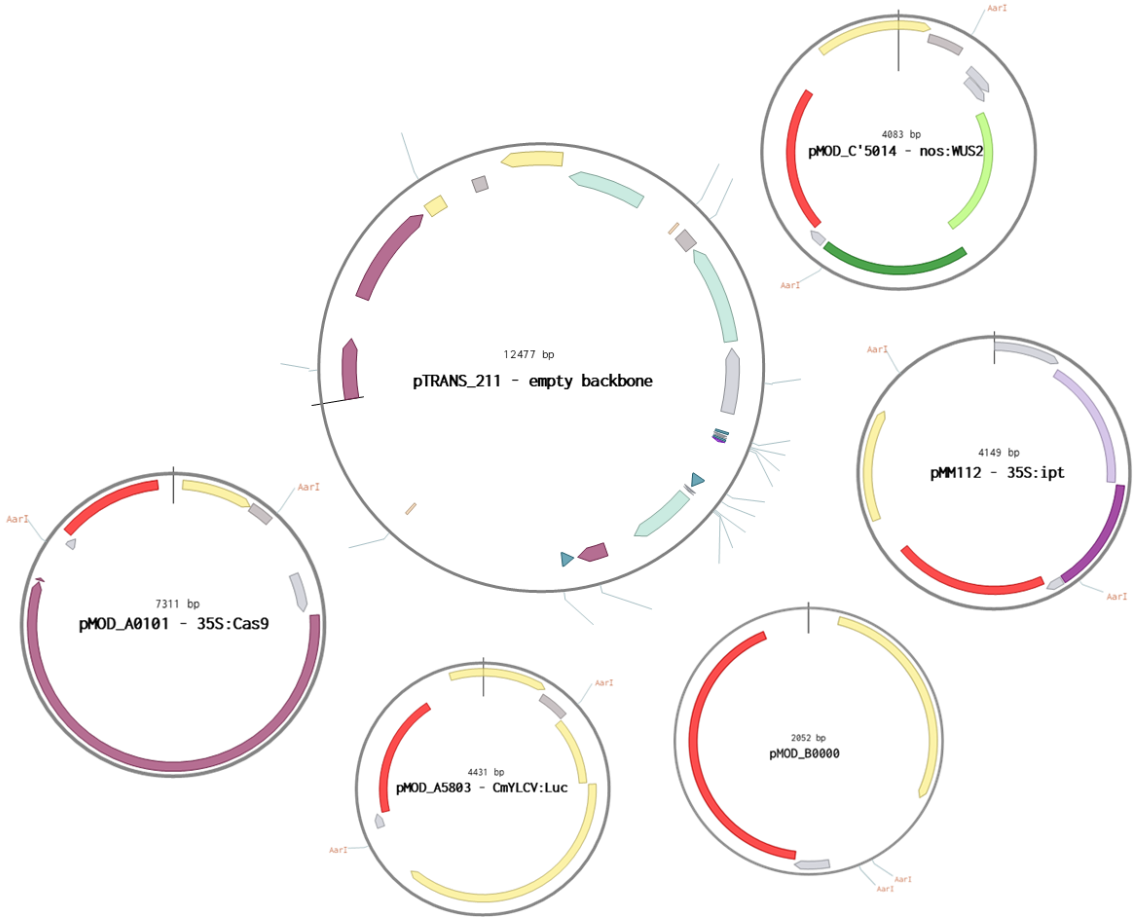


Fig. 19: Vectors and modules used for plant transformation.

Golden gate cloning system making use of PaqCI (AarI) restriction enzyme. pTRANS_211 – empty backbone for integration of 4 modules (A, B, C, D), includes kanamycin resistance for bacterial and hygromycin resistance for plant selection; pMOD_C'5014 – nos:WUS2 C module coding for the transcription factor *Wuschel2* under a nopaline synthase promoter, includes ampicillin resistance; pMM112 – 35S:ipt D module coding for the isopentenyl transferase under a 35S promoter, includes ampicillin resistance; pMOD_B0000 empty B module accepting the guide scaffold, includes ampicillin resistance; pMOD_A5803 – CmYLVC:Luc A module coding for Luciferase under the Cestrum yellow leaf curling virus promoter, includes ampicillin resistance; pMOD_A0101 – 35S:Cas9 A module coding for domesticated Cas9, includes ampicillin resistance.

Digestion-ligation reaction was performed following "Golden Gate Assembly Protocol using PaqCI® (NEB #R0745) and T4 DNA Ligase (NEB #M0202)" from New England BioLabs *https://international.neb.com/protocols/2021/01/11/golden-gate-assembly-protocol-using-paqci-neb-r0745-and-t4-dna-ligase-neb-m0202*.

Preparation of Electro-competent Cells

E. coli – DH10B

An over-night (o/n) bacteria culture was grown in 5 ml of salt-free Luria broth (LB^{-NaCl}) with the addition of Streptomycin c = 25 µg/ml on 37°C. 250 µL of this o/n culture were added to 50 ml of fresh LB^{-NaCl} the next day. This mixture was left to grow for 2-3h until OD_{600} reached 0.4 - 0.6. It was left to chill on ice for 15 min, then poured into 4 x 50 ml Falcon tubes and centrifuged for 15 min at 4000 g. The supernatant was discarded and the pellet resuspended in 10 ml of ice-cold water. The content of 2 tubes was then combined into one, topped up to 50 ml with ice-cold water and centrifuged again as above. Supernatant was discarded again, two tubes combined to one and once more centrifuged as above. For a final step, the pellet was resuspended in ice-cold 7% DMSO and 30 µL aliquoted into microtubes, which were kept on -80°C until further use.

A. tumefaciens – LBA4404

5 ml of LB^{-NaCl} were inoculated with one colony of LBA 4404 and grown on a shaker on 28°C until OD_{600} reached 0.2 – 0.3. 40 µL of this o/n culture were then transferred into a 2L flask containing 200 ml of LB^{-NaCl} plus Streptomycin (300 mg/ml) and grown on a shaker on 28°C for 28 – 30h (LBA4404 grows very slowly especially when antibiotics are added). Once the culture has reached an OD_{600} of 0.8 the culture was chilled on ice for 15 min and processed as described above. Aliquots of 50 µL were kept on -80°C until further use.

Electroportation

50-100 ng of DNA were used to mix gently with 20 µL of DH10b *E. coli* cells (or 50 µL of LBA4404). The cells were first taken from -80°C and left to thaw on ice for about 10 min. The

mixture was then transferred to a pre-chilled electroporation cuvette. The cuvette was properly wiped dry and then pulsed using the following standard settings for most *E. coli* strains: 1.8 kV, 25 µF and 200 Ω (LBA4404: 2.5 kV, 25 µF and 200 Ω). 1 ml of LB was added to the cells immediately after the pulse was delivered. The mixture was then transferred back into a microtube and left for recovery on 37°C (LBA4404: 28°C) with shaking for 1h (LBA4404: 3h). To ensure that individual, medium-sized colonies could be picked on the next day 50 µL of the transformation was plated on a pre-warmed selective plate. The rest was centrifuged, resuspended in 50 µL of fresh LB and plated onto a second plate. On some occasions even plating only 50 µL yielded too many colonies - then the mixture was diluted e.g. 1:10 using LB medium.

Preparation of Growth- and Infiltration Medium for LBA4404

Growth medium and infiltration medium were prepared following Maher et al. (2020) [35].

One transformed colony of LBA4404 was left to grow in growth medium (10 mM MES pH 5.6, 20 µM acetosyringone, 50 µg/ml kanamycin and 60 µg/ml streptomycin – in LB) for 12h at 28°C with shaking. Rifampicin was dismissed when preparing the growth medium because it slows down growth and reduces transformation efficiency. Cells were collected by centrifugation and resuspended in infiltration medium (10 mM MES pH 5.6, 150 µM acetosyringone and 10 mM MgCl2). They were then left at room temperature for 2-4h until OD_{600} reached 0.2 – 0.3.

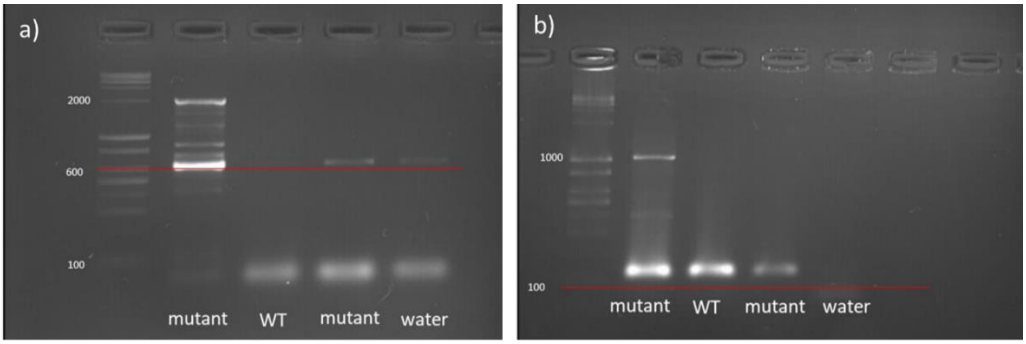
Infection

4- to 5-week-old tomato plants were infected with the before freshly prepared Agrobacterium infiltration medium. The apical meristem was cut off and very small quantities of the solution were injected directly in, as well as around the cut site using a 1 ml syringe with a 30 gauge needle. To reduce the risk of growing a huge callus, the injection of big quantities should be avoided because a bigger callus is slow in generating new shoots.

Results

MicroTom Mutant

To test the seeds received from Japan (University of Tsukuba) plant material was extracted from young seedlings and tested via PCR for the T-DNA insertion. The goal was to see whether the insert has a) not gone lost and b) is located in the *Sfh8* target gene.



Ladder: Long rage

Fig. 20: PCR of Japanese MicroTom mutant

with recommended primers provided by TOMATOMA. a) The two primers bind to the *nptII* gene (ubiquitin-NPTII fusion: ABC61923.1) located on the T-DNA insertion. PCR yielded the expected 622 bp fragment from both mutant DNA samples and no fragment from the wild type (WT) plant. (The slight band in the water control was most likely caused by contamination.) b) As a reference gene the elongation factor *EF 1-alpha* (LOC544055) was used. This is an essential protein playing a key role during the elongation cycle of protein biosynthesis and should be intact in all samples, mutant and wild type. PCR yielded a 123 bp product from all samples, as was expected.

To show that the insertion is located somewhere within the *Sfh8* gene, the aim was to amplify a sequence with one primer binding to the beginning of the gene (exon 3) and one close to the end (exon 13). If the T-DNA is located somewhere in between, this would show a shift on an agarose gel (=one band longer/higher than the other). Polymerase: Phusion® High-Fidelity by NEB. Expected band size of WT: 3760 bp. Expected band size of mutant: unknown, due to the absence of the plasmid map at that point in time.

Fig. 21: PCR of the *Sfh8* gene - mutant vs. WT.

In the WT is one clear band (expected size) visible. In the mutant (the expectation was a bigger band) no band can be seen. Therefore, it is assumed that the T-DNA is too long to amplify even via Phusion® polymerase.

As a next step the Japanese supplier (TOMATOMA) was asked for information on the T-DNA. Their vector map (*Fig. 13*) shows a T-DNA length of 9046 bp. Even though Phusion® polymerase is a high processivity enzyme (shorter extension times, more robust amplification resulting in the ability to amplify long templates) the maximum length that can be achieved is around 6-8 kb. Therefore, the initial plan had to be discarded.

Another possibility to check the insert for correct location placement would be to use different primer combinations and see if one of them yields a band. By using a primer that binds to the *Sfh8* gene combined with one that binds to insert (*nptII* gene), a fragment would only be produced if the insert is located in its expected place, namely the *Sfh8* gene.

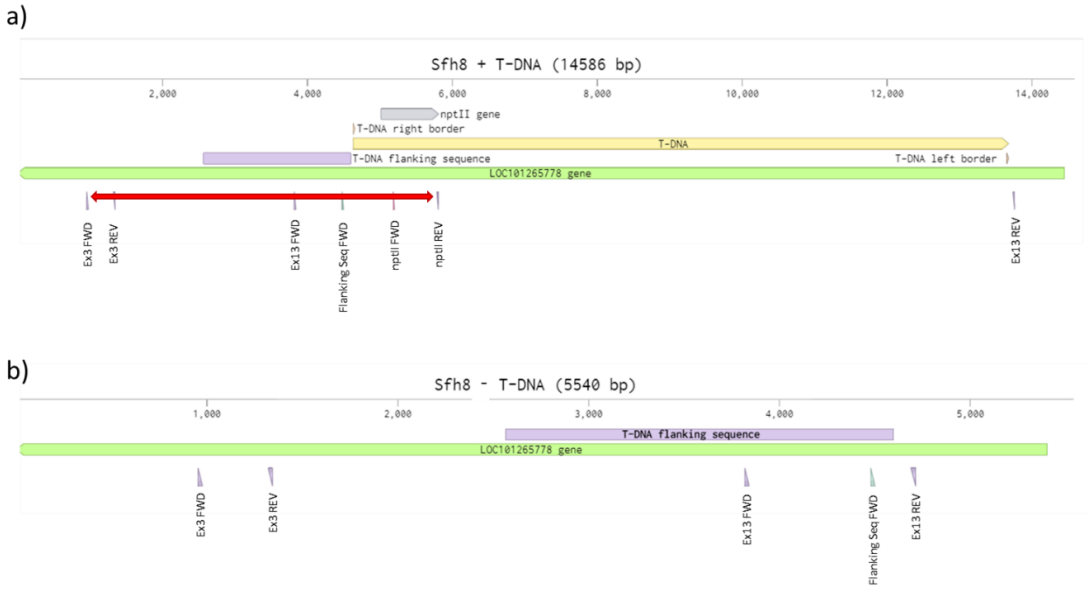


Fig. 22: Example of *Sfh8* gene a) with and b) without the T-DNA insert.

The picture is an example of how a T-DNA insertion inside a gene could look like. In reality it is not known in which direction the insert has been integrated and therefore it often results in trial and error when it comes to picking the right primers (*Tab. 2*). In addition, there is a good chance that the fragments might still be too long to be amplified. a) In a mutant plant, the red arrow indicates one possible amplicon that could be achieved with a primer that binds to the gene and one that binds to the insert. b) In a WT plant amplification is not possible because the nptII primer (primer that binds to the T-DNA insert) has nowhere to bind.

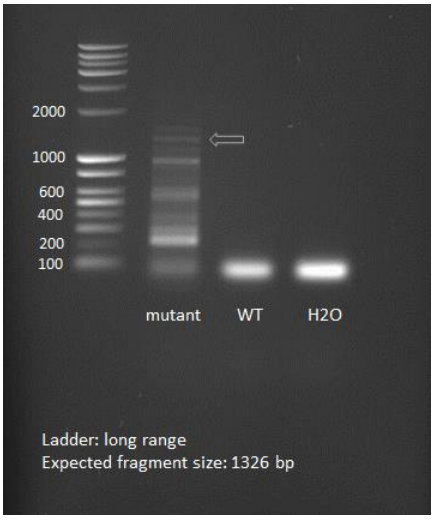


Fig. 23: Amplification with Flanking Seq. FWD & nptII FWD.

Only one primer combination gave a visible band with the correct size pointed to by the arrow.

The additional bands in the "mutant" (*Fig. 23*) are likely caused by unspecific binding and/or an incorrect annealing temperature. Therefore, gradient PCR was performed using different annealing temperatures because there is a chance that the temperature calculated by *https://tmcalculator.neb.com/* is too unspecific for this particular primer pair.

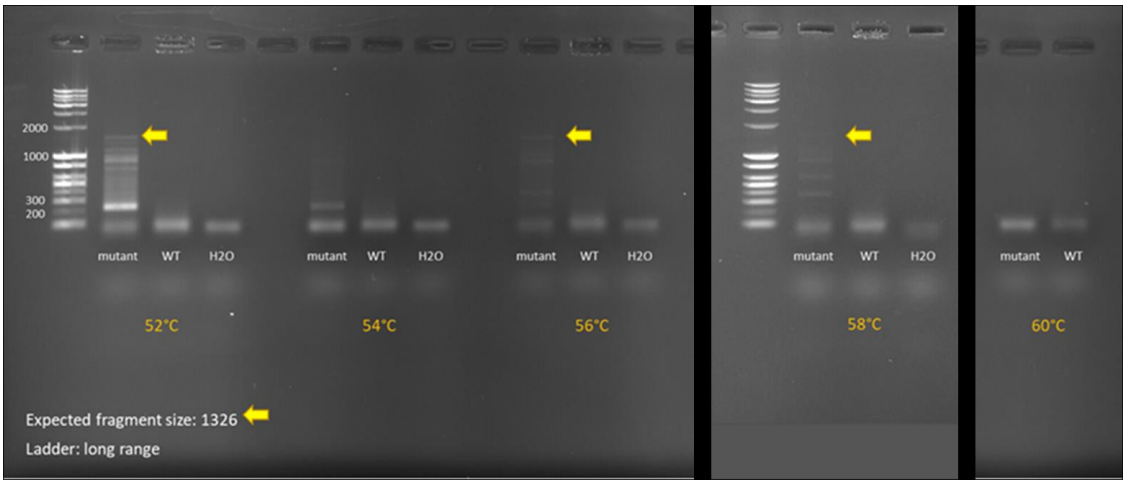


Fig. 24: Gradient PCR with different annealing temperatures.

In none of the applied temperatures did the result differ greatly from the initial result in *Fig. 23*.

Nevertheless, since there are only bands produced in the mutant samples within all repetitions, it is fair to assume that the T-DNA insertion is located within our target gene.

Moneymaker Plants

Because there was only a very limited amount of seeds available in the beginning (all derived from university-intern propagation) seed propagation was the first step to do. Additionally, the principal growth stages, as described by Lorenz et al. (1994) [46], were described to better calculate general growth, flowering time, fruit production and more.

Tab. 3: Documentation of Moneymaker plant growth

Picture	**Date**	**Age in days**	**Principal Growth Stage**	**In accordance with BBCH scale?**
	16.04.21	3d - Seed imbibition complete	0: Germination	yes
	22.04.21	9d - Emergence: cotyledons break through soil surface	0: Germination	yes
	27.04.21	14d – First true leaves appear	1: Leaf development	yes ± 2d
	10.05.21	27d - 1st secondary apical primary side shoots visible	2 - Formation of side shoots	yes
	13.05.21	30d - 1st secondary apical primary side shoots visible	2 - Formation of side shoots	yes
	17.05.21	34d - 1st tertiary apical side shoots visible	2 - Formation of side shoots	yes
In the Moneymaker variety apical side shoot formation occurs concurrently with the emergence of the inflorescence, hence there is no distinction of principal growth stages 3 and 4.				

	31.05.21	48d - First inflorescence visible (first bud erect)	5 - Inflorescence emergence	yes ± 2d
	08.06.21	56d – continuous stem elongation	5 - Inflorescence emergence	yes
	29.06.21	77 – 1st/2nd/3rd fruit has reached typical size and form	7: Development of fruit	yes
 l: plants without climbing aid vs. r: plants with climbing aid: straighter growth, room for aeration, easy harvest, saves space.				
	→ continuous harvest	81d – 89d: fruits show typical fully ripe colour	8: Ripening of fruit and seed	yes

Because of limited space in the greenhouse the plants had been removed before principal growth stage 9: Senescence.

Construction of the Guide Scaffold

The initial plan of primer annealing and dig-lig reaction (with vector 17A) at the same time did not bring the expected result. The reaction was repeated for 3 times until the task was finally split into two separate steps. First, vector linearization and then, ligation of the desired module into the vector. The goal was to cut out the LacZ cassette before the ligation step of the guide sequence. To exclude the possibility that neither the BsaI restriction enzyme nor one of the competent cells used for propagation of the vector in the first place, were the cause of the problems, 1) 17A propagated by ©NEB$_{stable}$ cells, 2) 17A propagated with DH10b cells, 3) BsaI enzyme from our lab stock, 4) BsaI enzyme from a neighbour lab stock and 5) a known, functional control vector were put to the test.

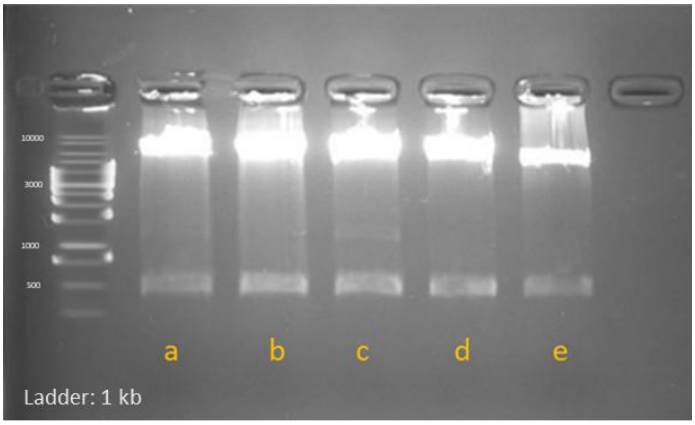


Fig. 25: Vector linearization and test of BsaI restriction enzyme

Digest for 2h at 37°C. a) 17A propagated in ©NEB$_{stable}$ cells, digested with BsaI from our lab stock; b) 17A propagated in ©NEBstable cells, digested with BsaI from a neighbour lab stock; c) 17A propagated in DH10b cells, digested with BsaI from our lab stock d) 17A propagated in DH10b cells, digested with BsaI from a neighbour lab stock; e) control vector. 17A total vector size: 14.273 bp; 17A - Size of LacZ: 596 bp; Control vector: pICH86988 – total size: 9078 bp, Size of LacZ-alpha: 596 bp.

The result clearly shows that the restriction enzyme as well as all vectors are functional. The enzyme in all cases was able to detect the correct recognition sites and cut LacZ out of the vector. Nevertheless, there might still be some uncut or nicked vectors in the mix. The upper band (*Fig. 25* c,d) were cut from the gel and used for the following ligation reaction (Protocol

by NEB: DNA ligation with T4 DNA-ligase-M0202 [*https://international.neb.com/protocols/0001/01/01/dna-ligation-with-t4-dna-ligase-m0202*]).

After an overnight incubation at 16°C in a thermo-cycler the reaction was used for electroporation with DH10b electro-competent cells. After another night of incubation at 37°C the plasmid was purified out of the cells and digested with ApaI. This enzyme cuts within the LacZ cassette and should therefore give a different pattern when compared to a vector without the insert.

Fig. 26: Virtual vs. real digest of ligation reaction with linearised 17A vector and G1 (guide 1).

a) Virtual digest with ApaI: 1) 17A without guide insertion, 2) 17A with guide insertion. b) Digestion of 10 colonies plus one empty vector as a negative control (c). Number 6 shows the expected pattern for a positive sample.

Only sample #6 shows the desired pattern. Since the lowest band is hardly visible the digest was repeated using a different restriction enzyme.

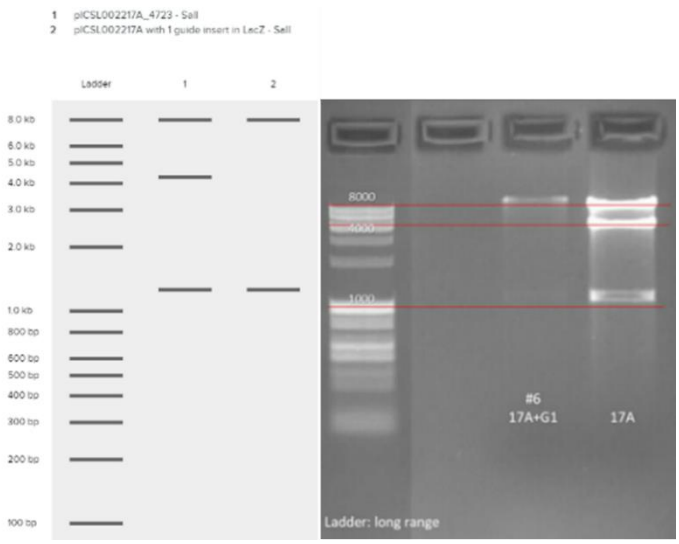


Fig. 27: Virtual vs. real digest of the positive sample #6

Left lane: vector with correct insert (= 1st guide insertion). Right lane: empty 17A vector as a negative control. The lower band in the left lane (17A + G1) is again rather faint but was nevertheless sent for sequencing.

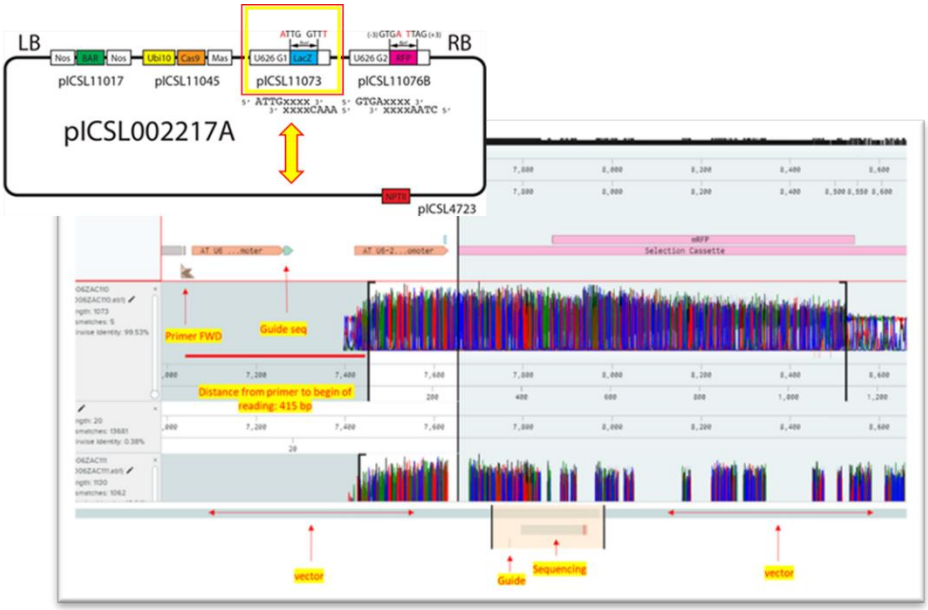


Fig. 28: Sequencing result of sample #6.

The sequencing result clearly shows a 415 bp gap between the primer binding region (Primer FWD) and the start of the reading, indicating a deletion of the U6 promoter region, the guide sequence and guide tail.

VECTOR TROUBLESHOOTING

Since all attempts of cloning vector 17A failed so far and trials did rarely show the expected colour scheme (*Fig. 18*) it became clear that the vector itself must be the cause of the encountered difficulties. In total 4 very similar vectors were available, all derived from the same source (*Fig. 16*). They only differed in the amount of replacement cassettes. To find out if one of them could be used instead of 17A, digests were performed to see if the vectors still match their in silico vector map.

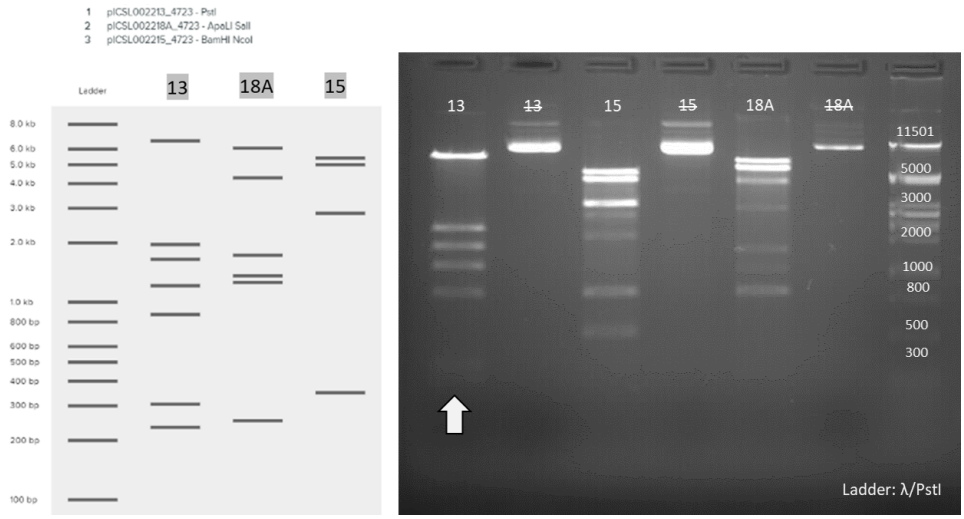


Fig. 29: Virtual vs. real digest of the remaining cloning vectors.

From left to right: plCSL002213 (13), ~~13~~ – undigested vector, plCSL002218A (18A), ~~18A~~ – undigested vector, plCSL002215 (15), ~~15~~ – undigested vector. Only vector 13 (white arrow) shows a matching digestion pattern when compared to the virtual digest.

Based on these results (*Fig. 29*), vector 13 was used for further experiments, yet the same difficulties were encountered. Again, the bacteria deleted the same sequence after ligation (*Fig. 28*). One possible solution to this problem is to skip the usage of bacteria altogether and perform "blind" PCR from vector 13. "Blind", because at this point there is no evidence the guide sequence has for sure been ligated into the vector. At the same time, these primers will add the AarI overhangs that are needed for the next cloning step.

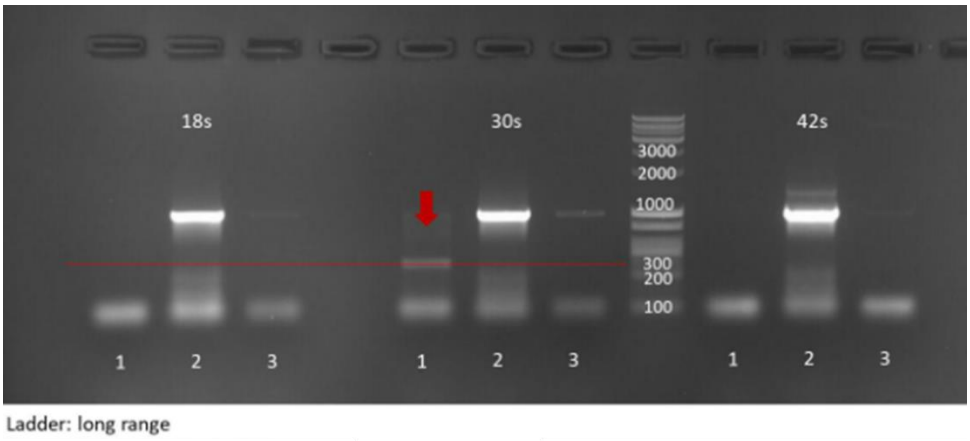


Fig. 30: Amplification of guide scaffold using 3 different extension times.

Expected size of guide scaffold: 373 bp (red arrow); Expected size of LacZ: 944 bp; 1= Vector 13 with guide (G1) insert; 2= linearised vector 13; 3= negative water control. Primers: Guide Scaffold (AarI) FWD + Guide Scaffold (AarI) REV (*Tab. 2*).

An extension time of 30 seconds yielded a 373 bp product pointed to by a red arrow in *Fig. 30*. The length matches the expectation of the desired guide scaffold. It was cut from the gel, re-PCR was performed with the high-fidelity Taq ®Phusion and then sent for sequencing.

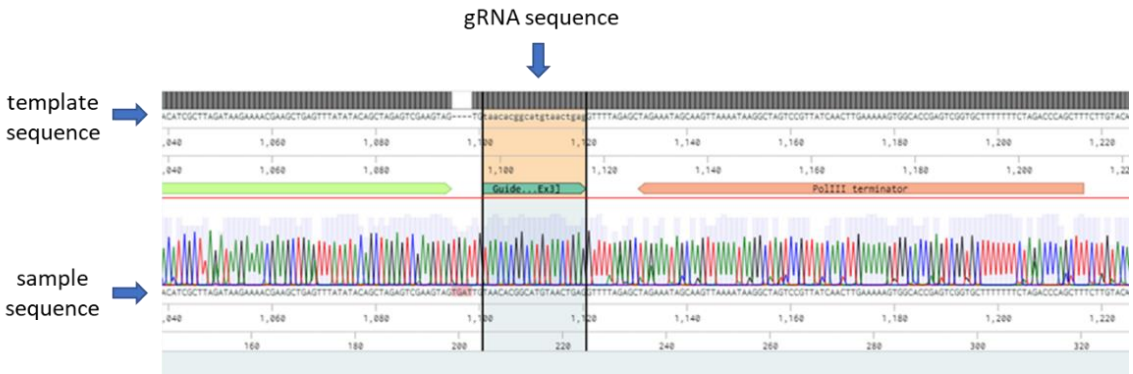


Fig. 31: Sequencing result of the guide sequence.

The highlighted section shows the read of the designed 20 bp gRNA sequence compared to the in silico template. All bases are correct.

This fragment could now be used as a single module for the plant transformation cloning vector system.

VECTOR SYSTEM FOR PLANT TRANSFORMATION WITH GROWTH REGULATORS

The Voytas vector system (*Fig. 19*) has the capacity to incorporate 4 modules in an A-B-C-D-manner depending on the overhangs. Using the one-tube reaction method and the "Golden Gate Assembly Protocol using PaqCI®" by NEB, 2 vectors (*Fig. 32*) were assembled, which could then be used for plant transformation.

Tab. 4: Vector design with 4 modules

Module	Position	pTRANS_Cas9	pTRANS_Luc
pTRANS_211	empty backbone	x	x
pMM112 – 35S:ipt	D	x	x
pMOD_0101 – 35S:Cas9	A	x	
pMOD_5014 – nos:WUS2	C	x	x
pMOD_5803 – CmYLCV:Luc	A		x
pMOD_guide scaffold	B	x	x

Because the modules came pre-designed every vector could incorporate 4 modules in an A-B-C-D-manner. Because Cas9 and Luciferase are both place A modules, 2 slightly different vectors were designed (see *Tab. 4*). The first vector contained modules in the following order: Cas9, gRNA, WUS2 and ipt (*Fig. 32*: Cas9 Plasmid). Obviously, Cas9 and the guide RNA must be on the same vector and introduced simultaneously. The second vector contained Luciferase, gRNA, WUS2 and ipt (*Fig. 32*: Luc Plasmid). This vector has the potential to induce the growth of a new meristem but cannot introduce a mutation because the CRISPR system is incomplete with Cas9 missing. Nevertheless, due to the injection of both vectors into the same site, selection via Luciferase should be possible. Bioluminescence will distinguish potentially altered meristems from random new growth initiated by the plant itself. However, this experimental setup is flawed and will be discussed in more detail in the last chapter (see *Future Outlook*).

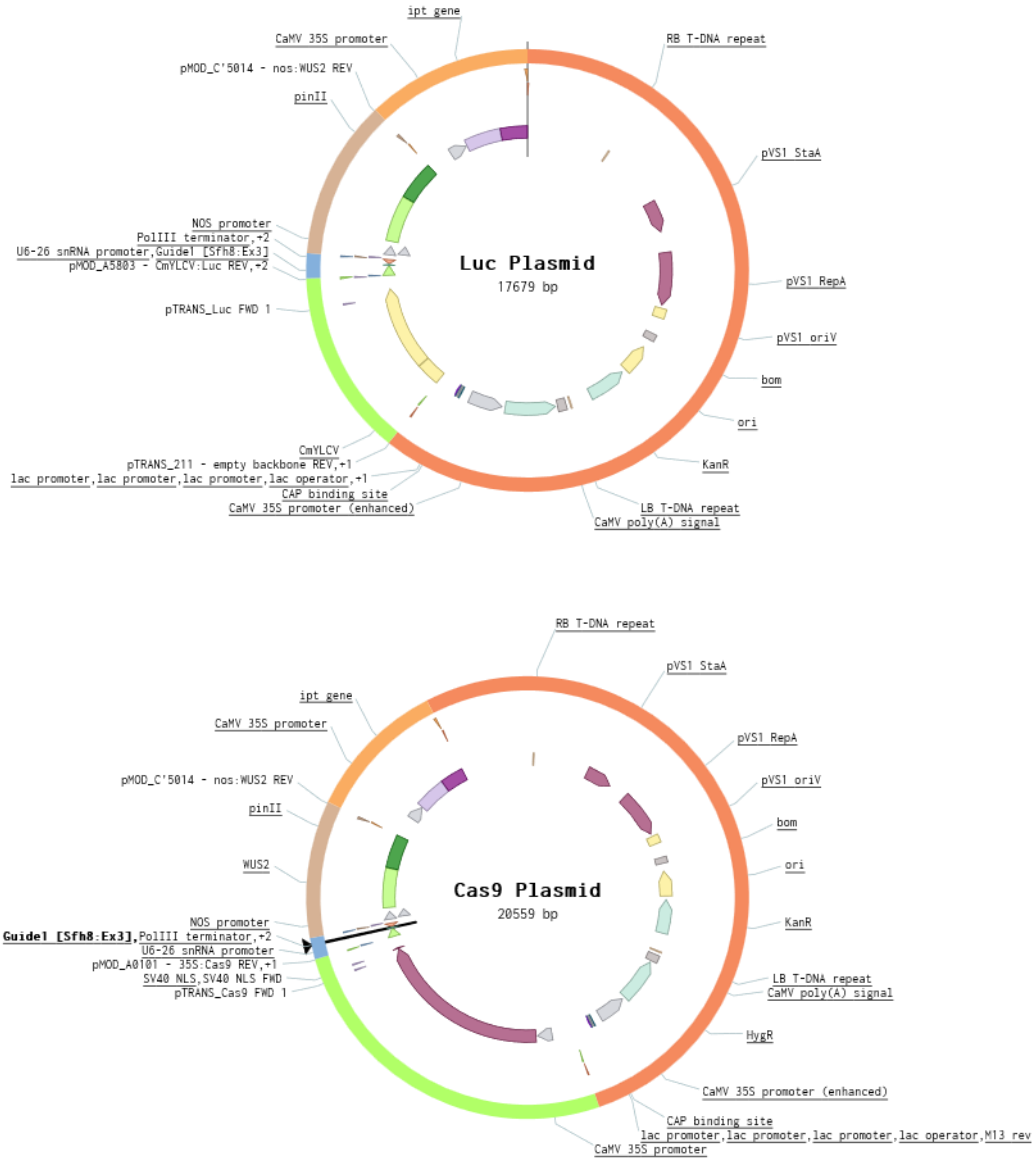


Fig. 32: Benchling vector maps of ready assembled plant trans-formation vectors.

Created with the "assembly wizard" tool.

To examine whether the vectors had been assembled correctly during the digestion-ligation reaction, the vectors were digested using a restriction enzyme that leaves a distinctly different pattern compared to an empty vector, without any inserts.

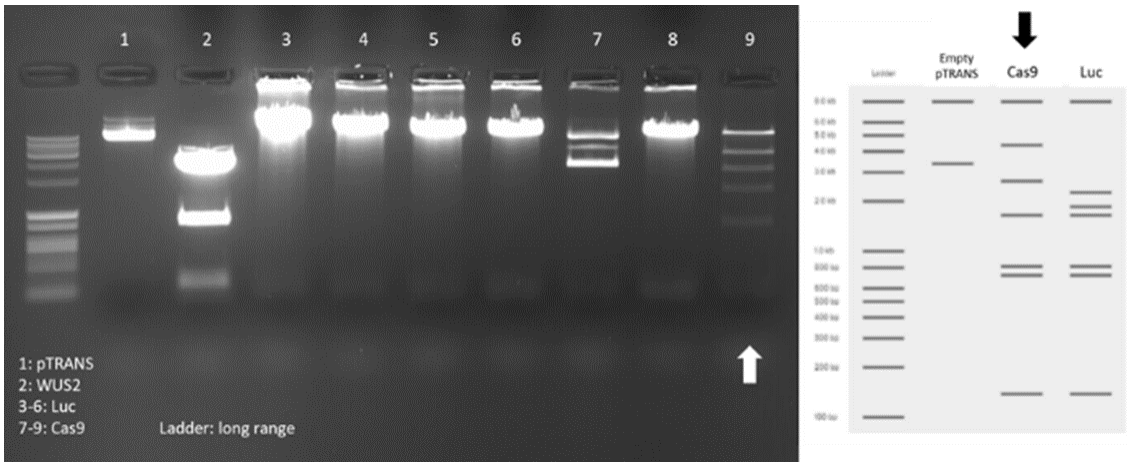


Fig. 33: Restriction digest of modules and vectors with PstI.

1) digest to confirm pTRANS empty backbone, 2) digest to confirm module WUS2, 3-6) pTRANS_Luc vector, 7-9) pTRANS_Cas9 vector. The white arrow points to the only sample that was identified as positive.

Only sample #9 shows an identical digestion pattern when compared to the in silico digest. This sample was evaluated as positive.

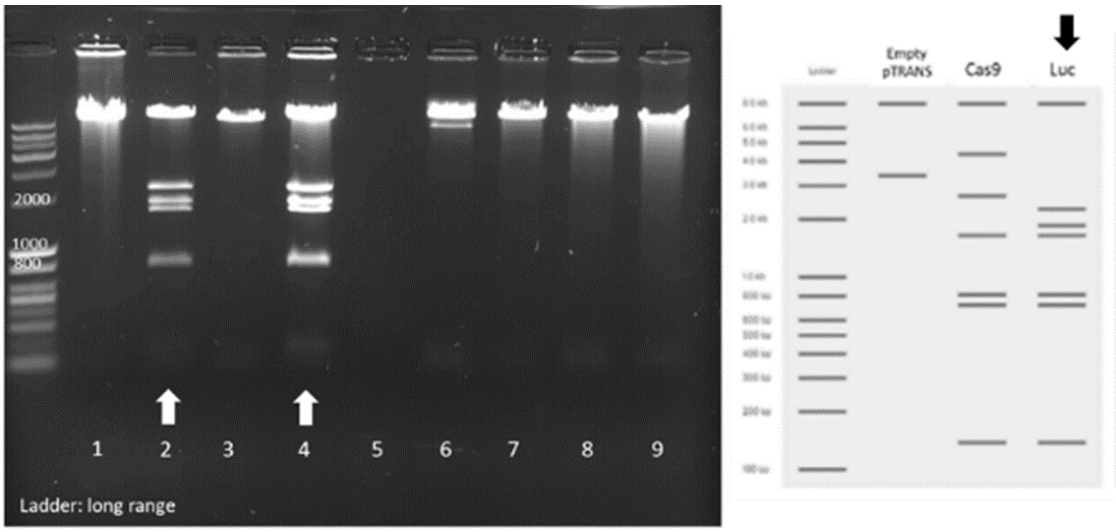


Fig. 34: Restriction digestion of pTRANS_Luc with PstI.

1-9) digest of pTRANS_Luc samples. Positively identified samples are marked with an arrow.

Only sample #2 and #4 show an identical digestion pattern when compared to the in silico digest. These samples were evaluated as positive.

Naturally, the next step would be to send the vectors for sequencing and verify the gRNA sequence for its accuracy, though since this has already been done before, during the guide scaffold design (*Fig. 31*), and no further bacteria had been used for the creation of the pTRANS vectors, it is unlikely that a base change has occurred. Nevertheless, PCR of the location, in which the guide sequence was inserted, was performed.

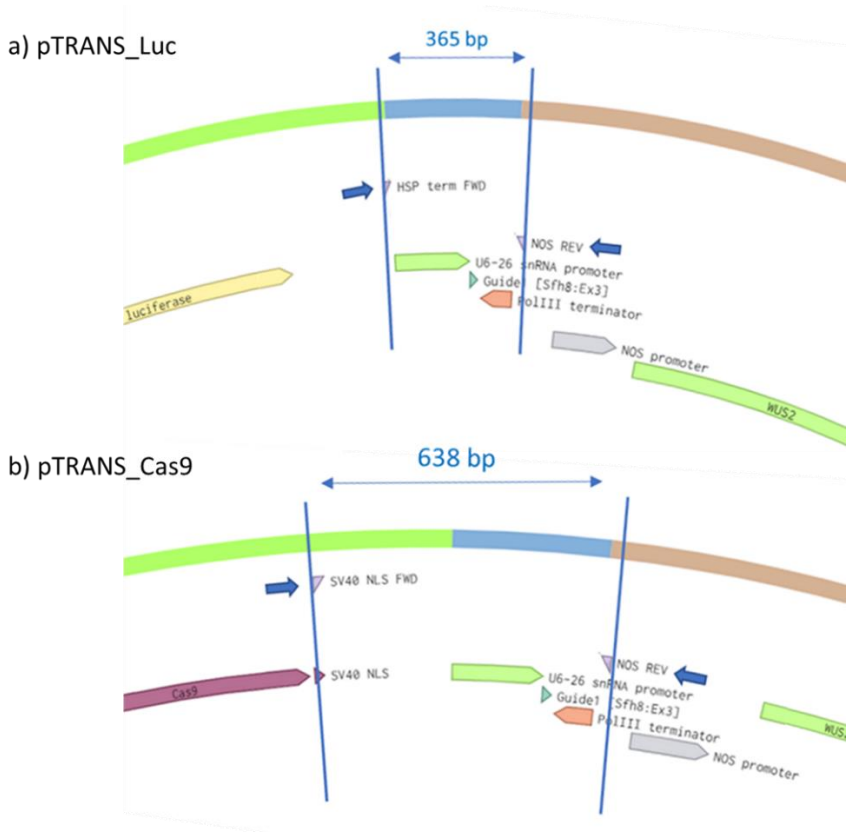


Fig. 35: Schematic representation of the two plant transformation vectors

and the location of the guide scaffold. The blue arrows point to the primers that have been used (*Tab. 2*) for PCR. The resulting products have a noticeable size difference (365 vs. 638 bp) compared to vectors with no insert: pTRANS_Luc $^{-insert}$: 47 bp; pTRANS_Cas9 $^{-insert}$: 320 bp.

Using different primers, each specific to the previous module terminator a) Luciferase – HSP term, b) Cas9 – SV40 nuclear localization signal (NLS), results in 2 different product sizes that are easily distinguishable and considering prior results, are sufficient proof of gRNA-insert accuracy.

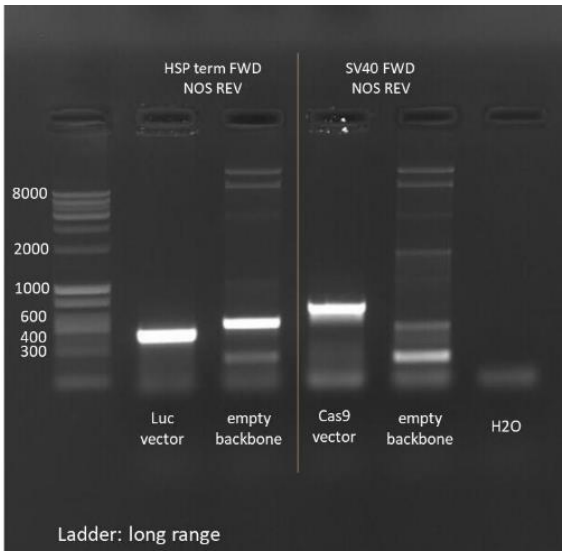


Fig. 36: PCR results of the two plant transformation vectors.

The visible bands on the gel resemble in silico calculated band lengths (*Fig. 35*). Empty vectors would theoretically produce the following band sizes: pTRANS_Luc $^{-insert}$: 47 bp; pTRANS_Cas9 $^{-insert}$: 320 bp.

pTRANS_Cas9 coding for the CRISPR/Cas9 system plus a gRNA will target exon 3 in the *Sfh8* gene in an attempt to either knock it out or down (has to be established later on). pTRANS_Luc can be used for selection of transformed meristems (see *Breaking the Tissue Culture Bottleneck with Developmental Regulators*). Both vectors were electroporated into Agrobacterium LBA4404.

Back-Transformation

Agrobacterium has a higher tendency to alter or delete sequences and has higher nuclease activity than *E. coli* (DH10b), so that direct minipreps out of Agrobacterium do not give clear bands. In order to check the efficiency of Agrobacterium for plant transformation, back-transformation into *E. coli* with subsequent minipreps and vector digestion was performed.

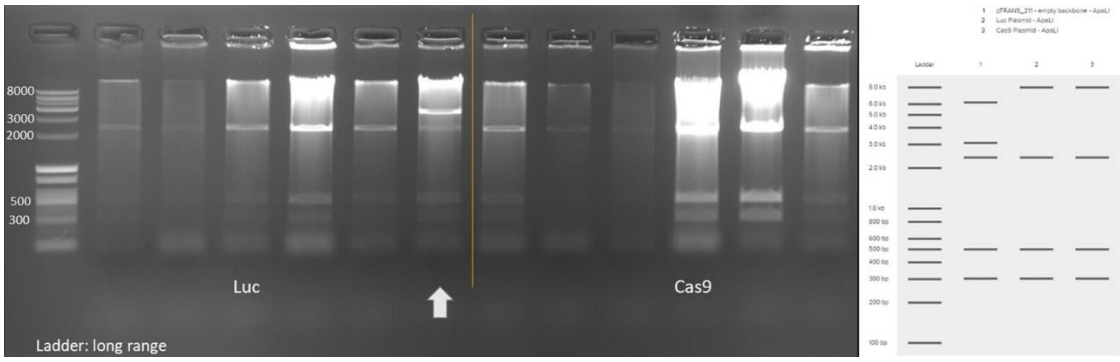


Fig. 37: Vector digest with ApaLI.

In the case of pTRANS_Luc all but one sample showed a correct pattern. In pTRANS_Cas9 all samples were correct.

The efficiency of pTRANS_Luc respectively pTRANS_Cas9 amounts to 83.33% and 100%.

Infection of Moneymaker Plants

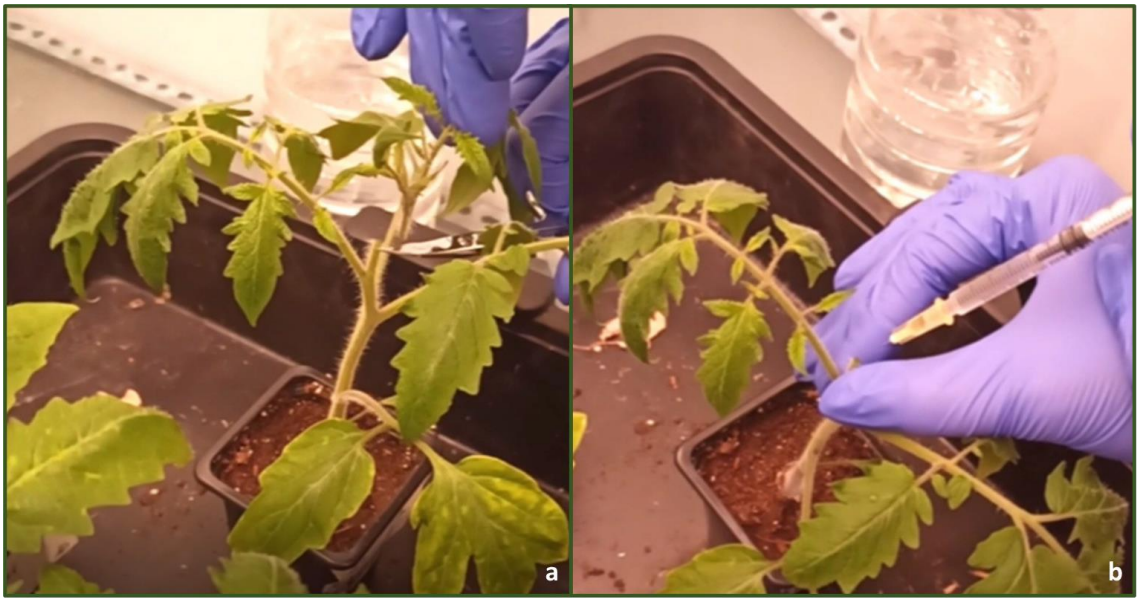


Fig. 38: Agrobacterium-mediated transformation

of 4 week old Moneymaker tomato plants. a) removal of the apical meristem with sterilised, sharpened scissors, b) infection medium containing either the pTRANS_Luc or pTRANS_Cas9 was injected into the cut site in very small amounts.

At this point only 3 plants had a sufficient age (~ 3 to 4 weeks) for transformation. These were used as first trials. 2 plants received both of the transformation vectors. 1 plant received only pTRANS_Luc. Since both of the vectors code for the developmental regulators (WUS2 and ipt) growth of a new meristem was expected in all plants. However, since only pTRANS_Cas9 carries the CRISPR system with a gRNA the expectation was to (possibly) find a mutation in only 2 plants. Screening for Luciferase should be possible in all plants.

This experimental setup was flawed however, since a positive and a negative control (Agrobacterium without a vector) were missing. The initial plan foresaw the knock-out of the JOINTLESS gene as a positive control but due to the before mentioned cloning difficulties and a subsequent lack of time, the construction of such vectors had to be dismissed at that point. Future experiments should include such a control though, or similar, to better determine and analyse different outcomes and results.

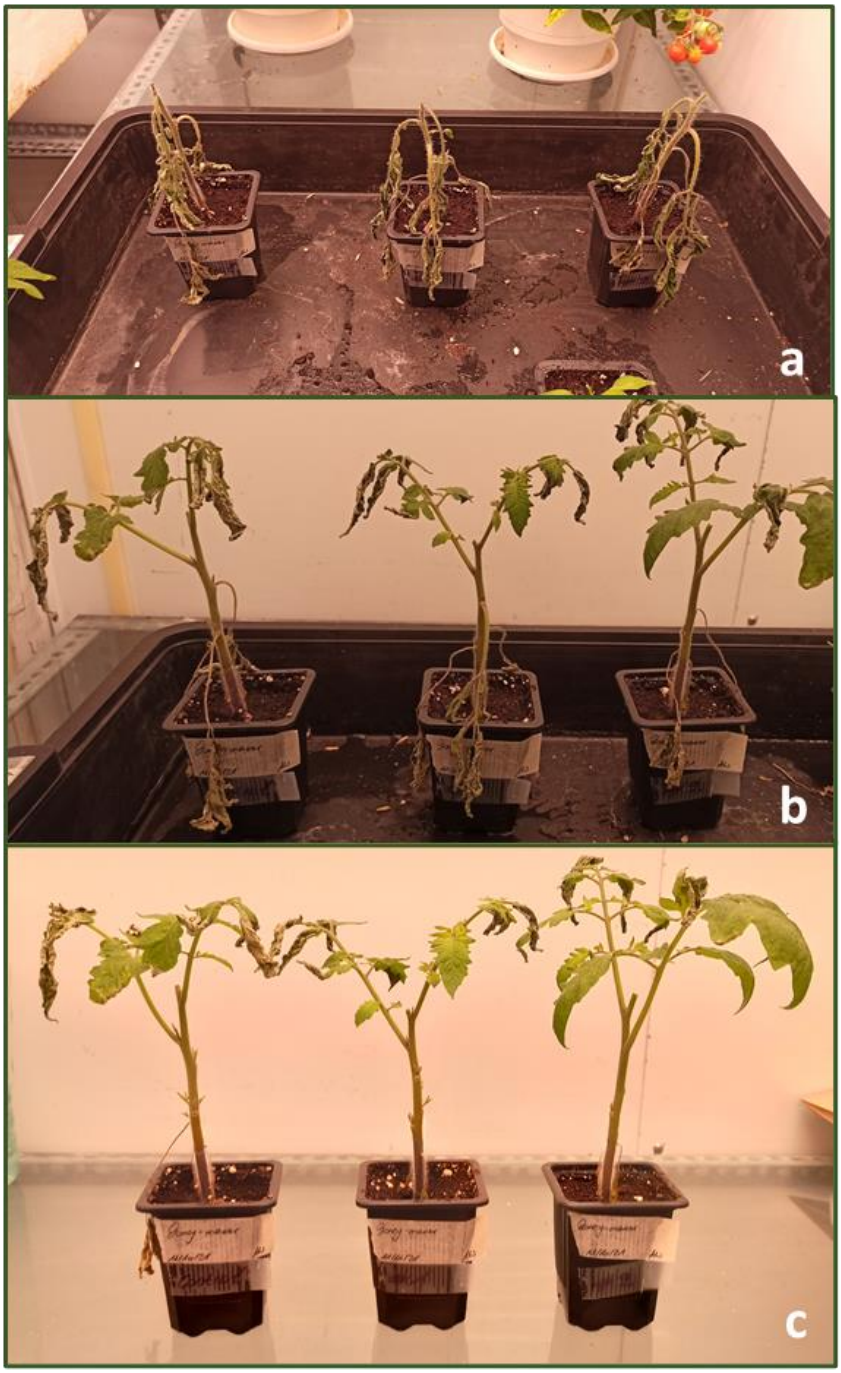


Fig. 39: Moneymaker plants

a) 5d post inoculation (pi); b) 6d pi and c) 7d pi. After the removal of the apical meristem the plant is left with a wound, losing eminently more water than usual. After the first few days of the infection the plants were watered every day and left in a tray with water to counterbalance for the loss.

For an unexpected reason (probably a human mistake), the plants run out of water over the first weekend and wilted considerably. It was possible to resurrect them over the following 2 days (*Fig. 39a,b,c*).

Fig. 40: Moneymaker plants 13d pi.

The plants were re-potted into bigger pots and fertiliser was added to the irrigation water.

Fig. 41: Moneymaker plants 2 ½ weeks pi.

With growth of a potentially altered new meristem. Visible on all plants.

Discussion

Contingency Plan – Identification of T-DNA Mutant

It is a future task of the program to extract and sequence DNA from soils in close proximity to plant roots during different growth stages. These samples will then be clustered into operational taxonomic units (OTU) representing molecular variance and can be seen as an indicator for community assembly and density dynamics [48]. Therefore, an important next step should be to measure mRNA levels of *Sfh8* and check whether the T-DNA insertion caused a complete

knock-out or just a knock-down. This would make a considerable difference, should this gene be involved in the establishment of a plants microbiome composition. The evidence so far does not provide a satisfactory answer to these questions.

One possibility to examine the close proximity and therefore the location of the insert is Thermal Asymmetric InterLaced PCR (TAIL-PCR). This method makes use of nested, insertion-specific primers and provides flanking sequences in the end. TAIL-PCR is a fast and cost-effective method [49] to examine numerous samples at the same time and I suspect that this is the method by which TOMATOMA is able to provide flanking sequences on their website. A more precise alternative to TAIL-PCR or the here performed regular PCR would be sequencing of the insertion site. This would show the exact location of the insert and confirm that the template DNA sequence is present as expected.

To confirm the presence of *Sfh8* mRNA, one possibility would be Northern blotting. By isolating RNA from different target tissues it is possible to compare gene activity amongst different plant organs (and/or growth stages), such as leaves, roots, flowers [50,51]. The comparison will be possible since even in knockouts the mRNA levels are usually not zero, just significantly lower. Theoretically, if an antibody exists, one could also do quantitative Western blotting (or In-Cell Western assays for quicker results). This way the SFH8 protein can be detected and compared over relative protein abundance before and after the knock-out (WT vs. mutant) or again between various tissues and organs. However, one thing to keep in mind with this method is that there is a chance of cross-reactions with other SFHs. The most sensitive method surely is Mass Spectrometric analysis (MS) for relative and absolute quantification of proteins [52]. Obviously, carrying out MS is dependent upon availability but it would offer a range of possibilities such as protein expression profiling, post-translational modifications, protein-protein interactions, structural and functional proteomics plus proteome mining for target identification/validation [53]. One of these options should be considered for future research to properly determine the function of the *Sfh8* gene and its involvement in the composition of the tomato microbiome.

VECTOR TROUBLESHOOTING

One of the main problems encountered during the cloning process was the malfunction of the first-step vectors 13 and 17A (see *Fig. 17*, *Fig. 18* and *Construction of the Guide Scaffold*). The unreliable colour scheme (randomly red or blue or no colour at all) took a long time to trouble-shoot and identify the main problems that made it impossible in the end to clone the desired guide sequences into these vectors. The following description is an attempt to explain why they did not work as expected and the initial idea of using them for the construction of the guide scaffolds had to be discarded as a whole.

With the insertion of the first guide sequence into the place of LacZ the selection should be based on classic "blue-white selection". When the LacZ cassette is replaced, no α-complementation can take place and no functional β- galactosidase will be produced anymore. Therefore, there will be no reaction with X-gal (a chromogenic substrate added to the medium) to give the distinct blue colour. To collect positive colonies, white colonies should be collected, or in the case of 17A red colonies because of the additional, functional RFP cassette on the vector (see *Fig. 18*). One reason why there rarely were red coloured colonies could be, that the RFP cassette including restriction sites, were mutated and therefore dysfunctional. This assumption was further supported by colleagues from the Swedish group at the University of Uppsala, who were working with the same vectors at the time.

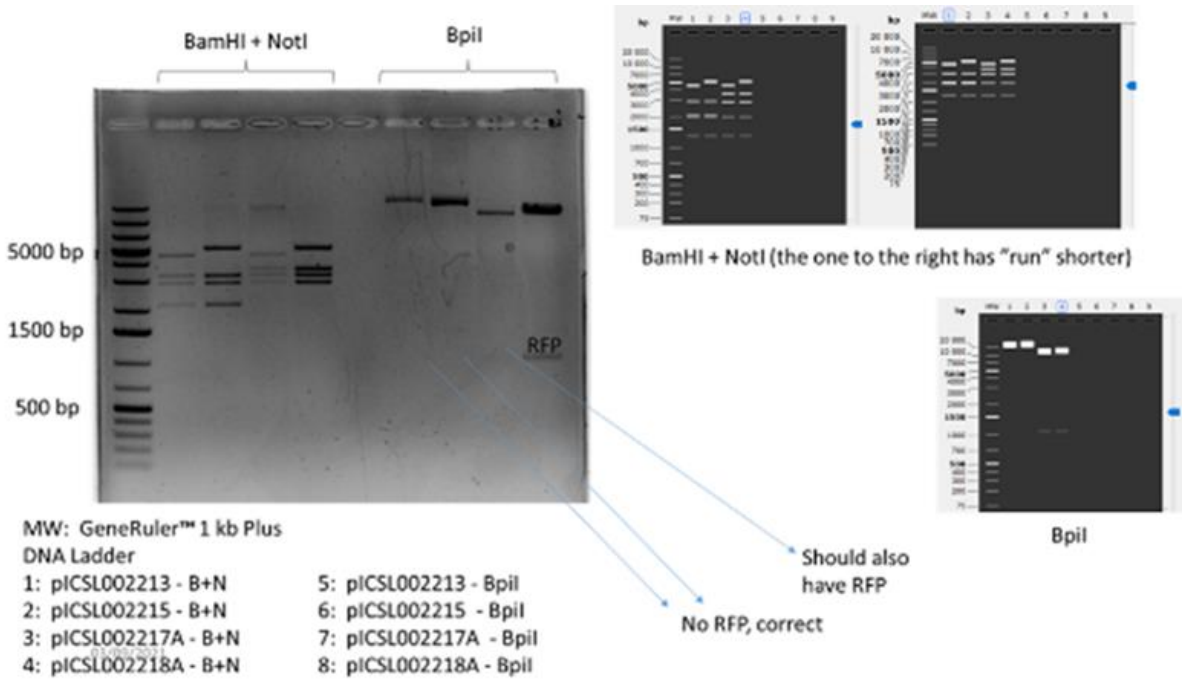


Fig. 42: Digest of cloning vectors 13, 15, 17A and 18A.

All available vectors were derived from the same source. Picture by Tornkvist, A. University of Uppsala. In the case of 17A, there is no band for the RFP cassette visible, meaning it was not possible to cut it out with BpiI. Additionally, no digestion pattern, except of vector 13, matches the virtual digest of the vector map, matching findings from *Fig. 29*.

These, and previous results (*Fig. 29*), led to the conclusion that the mutation sites of the vectors must have been mutated at some point previously to our experiments and as a result, cannot be recognised anymore by the BpiI enzyme or various other enzymes used throughout this work. This was further confirmed by digestion patterns that often showed one band less than they should have.

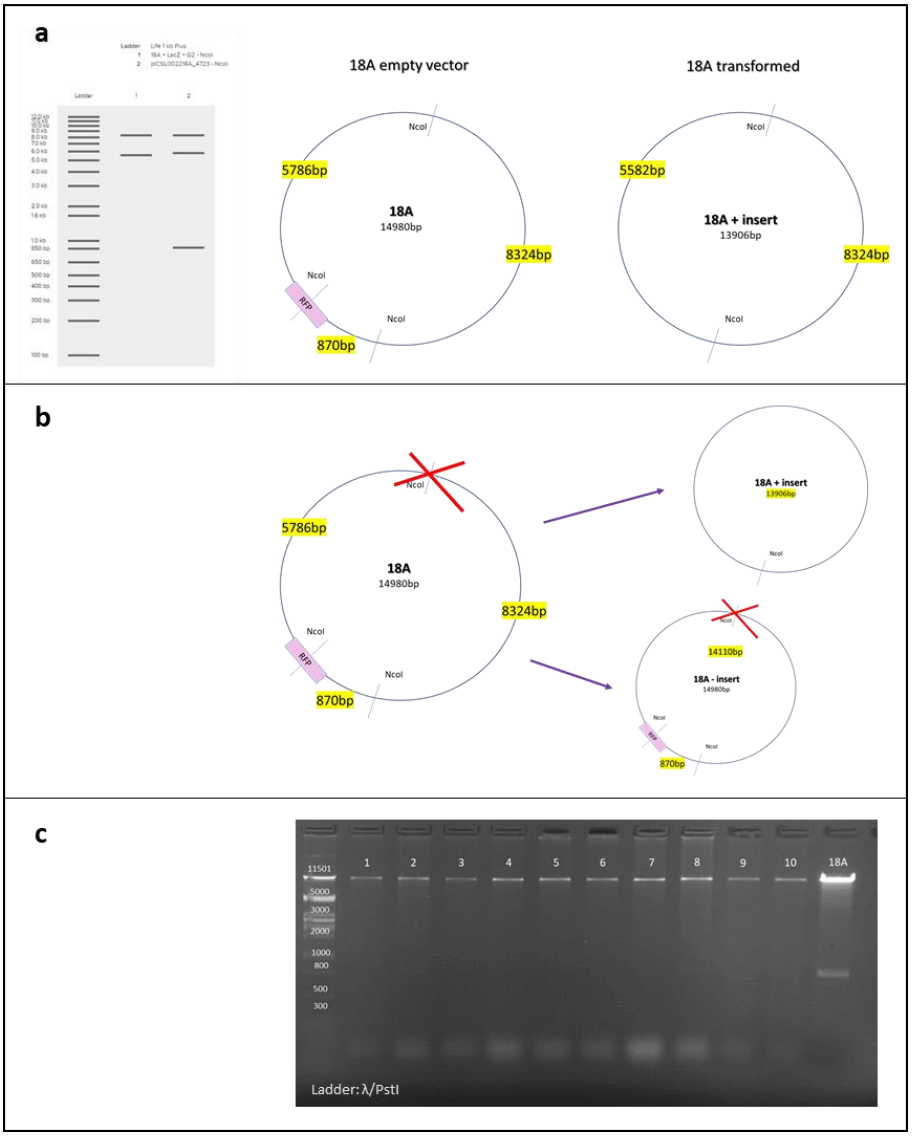


Fig. 43: Schematic representation of the 18A vector and the consequences of missing restriction sites.

a) The 18A empty vector matching it's in silico vector map. It contains 3 NcoI restriction sites, one located within the RFP cassette that will be replaced with the guide sequence during the restriction-ligation reaction. After that, the vector will be left with only 2 restriction sites (18A transformed) and will, after successful transformation, produce a distinctly different pattern vs. an empty vector. Highlighted numbers represent resulting band sizes. b) Representation of the same vector under the assumption that one restriction site is mutated and can therefore not be recognised anymore. It will subsequently show one band less on a gel after digestion. c) Agarose gel picture with one band less than expected compared to a virtual digest. (a) Only 2 bands can be seen in all samples indicating that NcoI only made 2 incisions instead of the expected 3.

The restriction enzymes used varied depending on the vector, the insert or replacement cassette and current availability of the enzyme. Despite this, the results yielded from restriction digests often showed one band less than expected (*Fig. 43c*). Assuming that one recognition site on the vector is missing, therefore naturally resulting in one visible band less on a gel, it is easy to see why it was not possible to confirm positive gRNA insertion. Mutations of this kind occur naturally with time. This is especially problematic in bacteria because of how fast they grow and divide, leading to frequent chances for mutation through random chance or natural selection [54]. The mutations might not only be limited to several recognition sites but can also effect the RFP or LacZ cassette, which in turn explains the random colour scheme obtained.

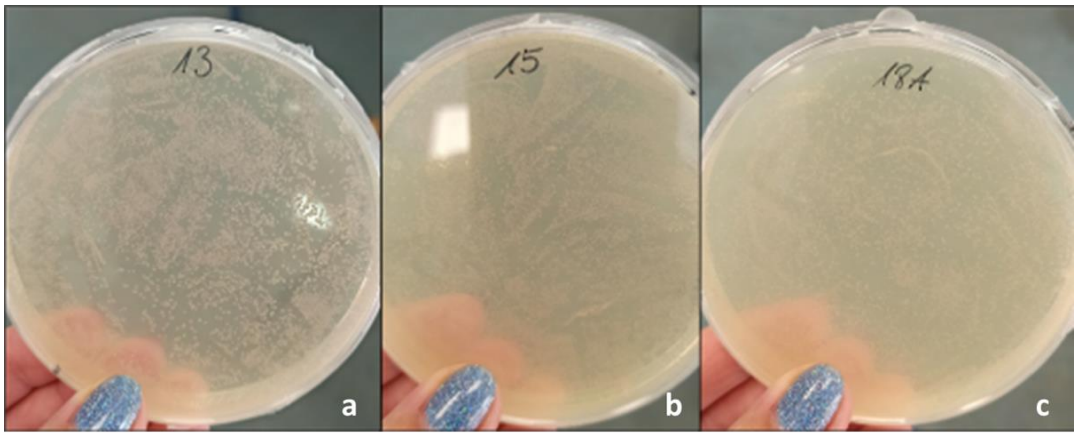


Fig. 44: Incorrect colour from transformation vectors

after only vector propagation in *E.coli* DH10b (no cloning took place at this point in time). All colonies show a distinct red-ish colour. The plates contain no additives (such as X-gal) and have been left in the freezer for a few hours after o/n incubation. This usually enhances the red colour from RFP. a & b) vector 13 and 15 have only LacZ and no RFP and should therefore not be red. c) vector 18A has LacZ plus RFP and should show a red colour since X-gal was not added to the plate.

Construction of the Guide Scaffold

Another reason for mutation could be the RecA (bacterial DNA recombination protein) homologous recombination pathway of *E. coli*. It is the process of deletion of repeated DNA sequences which, when occurring in the genome, usually contribute significantly to genetic

instability in procaryotes as well as eucaryotes and are known to even cause genetic diseases in the latter [55]. Even though all strains used for cloning are *recA* mutants, this phenomenon still occurs to a significant degree depending on length and how far the repeats are away from each other [56]. As previously shown in *Fig. 28* *E. coli* deleted a direct repeat sequence after potentially successful cloning of vector 17A. Deletions occur more frequently in larger repeats than shorter ones nevertheless, they are a common source of spontaneous mutation [55]. Vector 17A and 18A (*Fig. 16*) both contain an identical U6 promoter and guide tail in close proximity, which should generally be avoided when designing vectors.

Future Outlook

Due to a lack of time (my stay there ended after 8 months) and not enough tomato plants with the correct age for transformation I was only able to try agrobacterium-mediated transformation once with 3 plants. In the future this experiment should be repeated with more replicants plus a positive and a negative control (see *Infection of Moneymaker Plants*). Should the knock-out of *Sfh8* not work or result in unexpected plant death there is no way to verify the cause. That's why it is imperative to use a control that results in a known phenotype. I initially planned to knock out the JOINTLESS gene as a control but due to the previously explained cloning problems and subsequent troubleshooting process I lost a good portion of time and sadly, did not get to this point. However, in the future such "safety nets" should be considered when repeating the experiment.

A flawless Luciferase protocol should also be set up in the future since it is the first indication whether the transformation worked (formation of callus) or not (formation of regular side branch). In a positive case the developmental regulators (delivered by the agrobacterium) promote the induction of a meristem from edited somatic cells. The meristem will form a shoot, which can be excised and later on, the formation of roots can be induced. Progeny of the plant can then be screened for mutations via PCR primers that flank the expected cut site.

Conclusion

My main task throughout this cause was the establishment of a new method. However, even the insertion of only 2 guides turned out to be technically challenging. All steps of the method

have been performed at least once however, but the final result is still open and at present it is unclear whether knockout of *Sfh8* could be achieved.

As a backup strategy a T-DNA cultivar could presumably be identified but also this result is preliminary.

The advantage of the new method is that it should work in most economically relevant crop plants species and is way faster than traditional introgression of a mutation plus has an antibiotic resistance selection marker. The new method only leads to transient expression of the editing construct and should not be a stable transformation. It will just leave a small indel in the genome which is almost impossible to prove that this did not naturally exist in the germplasm.

Hopefully future work can show that the tomato plants have been successfully transformed (mutation introduced) and that either one of those mutant plants allows testing of the hypothesis of the *Sfh8* gene being involved in the composition of the root microbiome.

Literature

[1] Corinto GL. Nikolai Vavilov's centers of origin of cultivated plants with a view to conserving agricultural biodiversity. Human Evolution 2014;29:285–301.

[2] Schlegel RHJ. History of Plant Breeding. Boca Raton, FL 33487-2742: CRC Press Taylor & Francis Group; 2018. https://doi.org/https://doi.org/10.1201/b22334.

[3] Doebley JF, Gaut BS, Smith BD. The molecular genetics of crop domestication. Cell 2006;127:1309–21. https://doi.org/10.1016/J.CELL.2006.12.006.

[4] Meyer RS, Purugganan MD. Evolution of crop species: genetics of domestication and diversification 2013. https://doi.org/10.1038/nrg3605.

[5] Institute of Medicine and National Research Council. Globalization, biosecurity, and the future of the life sciences. The National Academies Press; 2006. https://doi.org/10.17226/11567.

[6] Nasti RA, Zinselmeier MH, Vollbrecht M, Maher MF, Voytas DF. Fast-TrACC: A Rapid Method for Delivering and Testing Gene Editing Reagents in Somatic Plant Cells. Frontiers in Genome Editing 2021;2:1–9. https://doi.org/10.3389/fgeed.2020.621710.

[7] Urnov F. The domestication of Cas9. Nature 2016 529:7587 2016;529:468–9. https://doi.org/10.1038/529468a.

[8] Belhaj K, Chaparro-Garcia A, Kamoun S, Patron NJ, Nekrasov V. Editing plant genomes with CRISPR/Cas9. Current Opinion in Biotechnology 2015;32:76–84. https://doi.org/10.1016/j.copbio.2014.11.007.

[9] Nasti RA, Voytas DF. Attaining the promise of plant gene editing at scale. Proceedings of the National Academy of Sciences 2021;118. https://doi.org/10.1073/PNAS.2004846117.

[10] Bhushan K. Evolution and molecular mechanism of CRISPR/Cas9 systems. Genome Engineering via CRISPR-Cas9 System, Elsevier; 2020, p. 15–25. https://doi.org/10.1016/b978-0-12-818140-9.00002-7.

[11] Singh V. An introduction to genome editing CRISPR-Cas systems. Genome Engineering via CRISPR-Cas9 System, Elsevier; 2020, p. 1–13. https://doi.org/10.1016/b978-0-12-818140-9.00001-5.

[12] Phillips SE, Vincent P, Rizzieri KE, Schaaf G, Bankaitis VA, Gaucher EA. The diverse biological functions of phosphatidylinositol transfer proteins in eukaryotes. Crit Rev Biochem Mol Biol 2006;41:21–49. https://doi.org/10.1080/10409230500519573.

[13] Irving HR, Boukli NM, Kelly MN, Broughton WJ. Nod-Factors in Symbiotic Development of Root Hairs. In: Root Hairs Cell and Molecular Biology, ed. RW Ridge and EM Emonds. Root Hairs, Springer Japan; 2000, p. 241–65. https://doi.org/10.1007/978-4-431-68370-4.

[14] van Spronsen PC, Grønlund M, Bras CP, Spaink HP, Kijne JW. Cell biological changes of outer cortical root cells in early determinate nodulation. Mol Plant Microbe Interact 2001;14:839–47. https://doi.org/10.1094/MPMI.2001.14.7.839.

[15] Mylona P, Pawlowski K, Bisseling T. Symbiotic Nitrogen Fixation. Plant Cell 1995;7:869–85. https://doi.org/10.1105/TPC.7.7.869.

[16] Mousley C, Tyeryar K, Vincent-Pope P, Bankaitis V. The SEC14-superfamily and the regulatory interface between phospholipid metabolism and membrane trafficking. Biochimica et Biophysica Acta 2007;1771:727–36.

[17] Li X, Routt SM, Xie Z, Cui X, Fang M, Kearns MA, et al. Identification of a Novel Family of Nonclassic Yeast Phosphatidylinositol Transfer Proteins Whose Function Modulates Phospholipase D Activity and Sec14p-independent Cell Growth. Molecular Biology of the Cell 2000;11. https://doi.org/10.1091/mbc.11.6.1989.

[18] Schnabl M, Oskolkova O v., Holič R, Brežná B, Pichler H, Zágoršek M, et al. Subcellular localization of yeast Sec14 homologues and their involvement in regulation of phospholipid turnover. Eur J Biochem 2003;270:3133–45. https://doi.org/10.1046/J.1432-1033.2003.03688.X.

[19] Kapranov P, de Bruijn FJ, Szczyglowski K. Novel, Highly Expressed Late Nodulin Gene (LjNOD16) from Lotus japonicus. Plant Physiology 1997;113:1081–90. https://doi.org/10.1104/PP.113.4.1081.

[20] Kapranov P, Routt SM, Bankaitis VA, de Bruijn FJ, Szczyglowski K. Nodule-specific regulation of phosphatidylinositol transfer protein expression in Lotus japonicus. Plant Cell 2001;13:1369–82. https://doi.org/10.1105/tpc.13.6.1369.

[21] Ghosh R, de Campos MKF, Huang J, Huh SK, Orlowski A, Yang Y, et al. Sec14-nodulin proteins and the patterning of phosphoinositide landmarks for developmental control of membrane morphogenesis. Molecular Biology of the Cell 2015;26:1764–81. https://doi.org/10.1091/mbc.E14-10-1475.

[22] Vincent P, Chua M, Nogue F, Fairbrother A, Mekeel H, Xu Y, et al. A Sec14p-nodulin domain phosphatidylinositol transfer protein polarizes membrane growth of Arabidopsis thaliana root hairs. Journal of Cell Biology 2005;168:801–12. https://doi.org/10.1083/jcb.200412074.

[23] Timmers ACJ. Infection of Root Hairs by Rhizobia: Infection Thread Development with Emphasis on the Microtubular Cytoskeleton. Root Hairs 2000:223–39. https://doi.org/10.1007/978-4-431-68370-4_14.

[24] Rai GK, Rai NP, Kumar S, Yadav A, Rathaur S, Singh M. Effects of explant age, germination medium, pre-culture parameters, inoculation medium, pH, washing medium, and selection regime on Agrobacterium-mediated transformation of tomato. In Vitro Cellular and Developmental Biology - Plant 2012;48:565–78. https://doi.org/10.1007/S11627-012-9442-3.

[25] Böhme K, Li Y, Charlot F, Grierson C, Marrocco K, Okada K, et al. The Arabidopsis COW1 gene encodes a phosphatidylinositol transfer protein essential for root hair tip growth. Plant J 2004;40:686–98. https://doi.org/10.1111/J.1365-313X.2004.02245.X.

[26] Weber E, Engler C, Gruetzner R, Werner S, Marillonnet S. A modular cloning system for standardized assembly of multigene constructs. PLoS ONE 2011;6. https://doi.org/10.1371/journal.pone.0016765.

[27] Rebatchouk D, Daraselia N, Narita JO. NOMAD: a versatile strategy for in vitro DNA manipulation applied to promoter analysis and vector design. Proc Natl Acad Sci U S A 1996;93:10891. https://doi.org/10.1073/PNAS.93.20.10891.

[28] Knight T. Idempotent Vector Design for Standard Assembly of Biobricks. MIT Libraries 2003:1–11.

[29] Engler C, Marillonnet S. Golden Gate cloning. Methods in Molecular Biology 2014;1116:119–31. https://doi.org/10.1007/978-1-62703-764-8_9.

[30] Cascales E, Atmakuri K, Sarkar MK, Christie PJ. DNA Substrate-Induced Activation of the Agrobacterium VirB/VirD4 Type IV Secretion System. Journal of Bacteriology 2013;195:2691–704. https://doi.org/10.1128/JB.00114-13/SUPPL_FILE/ZJB999092623SO1.PDF.

[31] Hwang H-H, Yu M, Lai E-M. Agrobacterium -Mediated Plant Transformation: Biology and Applications . The Arabidopsis Book 2017;15:e0186. https://doi.org/10.1199/tab.0186.

[32] Kiyokawa K, Yamamoto S, Sakuma K, Tanaka K, Moriguchi K, Suzuki K. Construction of Disarmed Ti Plasmids Transferable between E. coli and Agrobacterium Species.pdf. Applied and Environmental Microbiology 2009:1845–51.

[33] De Cleene M, De Ley J. The host range of crown gall. The Botanical Review 1976;42:389–466. https://doi.org/10.1007/BF02860827.

[34] Wu HY, Liu KH, Wang YC, Wu JF, Chiu WL, Chen CY, et al. AGROBEST: An efficient Agrobacterium-mediated transient expression method for versatile gene function analyses in Arabidopsis seedlings. Plant Methods 2014;10. https://doi.org/10.1186/1746-4811-10-19.

[35] Maher MF, Nasti RA, Vollbrecht M, Starker CG, Clark MD, Voytas DF. Plant gene editing through de novo induction of meristems. Nature Biotechnology 2020;38:84–9. https://doi.org/10.1038/s41587-019-0337-2.

[36] Gelvin SB. Agrobacterium-mediated plant transformation: the biology behind the "gene-jockeying" tool. Microbiol Mol Biol Rev 2003;67:16–37. https://doi.org/10.1128/MMBR.67.1.16-37.2003.

[37] Ondřej M, Kocábek T, Rakouský S, Wiesnerová D. Segregation of T-DNA inserts in the offspring of Arabidopsis thaliana after Agrobacterium transformation. Biologia Plantarum 1999;42:185–95. https://doi.org/10.1023/A:1002100400045.

[38] Gray SB, Brady SM. Plant developmental responses to climate change. Developmental Biology 2016;419:64–77. https://doi.org/10.1016/J.YDBIO.2016.07.023.

[39] Babé A, Lavigne T, Séverin JP, Nagel KA, Walter A, Chaumont F, et al. Repression of early lateral root initiation events by transient water deficit in barley and maize.

Philosophical Transactions of the Royal Society B: Biological Sciences 2012;367:1534–41. https://doi.org/10.1098/rstb.2011.0240.

[40] Yruela I. Plant development regulation: Overview and perspectives. Journal of Plant Physiology 2015;182:62–78. https://doi.org/10.1016/j.jplph.2015.05.006.

[41] Boutilier K, Offringa R, Sharma VK, Kieft H, Ouellet T, Zhang L, et al. Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. Plant Cell 2002;14:1737–49. https://doi.org/10.1105/TPC.001941.

[42] Neelakandan AK, Wang K. Recent progress in the understanding of tissue culture-induced genome level changes in plants and potential applications. Plant Cell Rep 2012;31:597–620. https://doi.org/10.1007/S00299-011-1202-Z.

[43] Kunkel T, Niu QW, Chan YS, Chua NH. Inducible isopentenyl transferase as a high-efficiency marker for plant transformation. Nature Biotechnology 1999 17:9 1999;17:916–9. https://doi.org/10.1038/12914.

[44] Smigocki AC, Owens LD. Cytokinin gene fused with a strong promoter enhances shoot organogenesis and zeatin levels in transformed plant cells. Proc Natl Acad Sci U S A 1988;85:5131. https://doi.org/10.1073/PNAS.85.14.5131.

[45] Hilbeck A, Meyer H, Wynne B, Millstone E. GMO regulations and their interpretation: how EFSA's guidance on risk assessments of GMOs is bound to fail. Environmental Sciences Europe 2020;32:1–15. https://doi.org/10.1186/S12302-020-00325-6/FIGURES/1.

[46] Lorenz D, Eichhorn K, Bleiholder H, Klose R, Meier U, Weber E. Growth stages of mono-and dicotyledonous plants. BBCH Monograph. Federal Biological Research Centre for Agriculture and Forestry 1994;49(2):66–70.

[47] Adikusuma F, Pfitzner C, Thomas PQ. Versatile single-step-assembly CRISPR/Cas9 vectors for dual gRNA expression. PLOS ONE 2017;12:e0187236. https://doi.org/10.1371/JOURNAL.PONE.0187236.

[48] Wei Z, Gu Y, Friman VP, Kowalchuk GA, Xu Y, Shen Q, et al. Initial soil microbiome composition and functioning predetermine future plant health. Science Advances 2019;5:1–12. https://doi.org/10.1126/sciadv.aaw0759.

[49] Singer T, Burke E. High-throughput TAIL-PCR as a tool to identify DNA flanking insertions. Methods Mol Biol 2003;236:241–72. https://doi.org/10.1385/1-59259-413-1:241.

[50] Brown T, Mackey K, Du T. Analysis of RNA by Northern and Slot Blot Hybridization. Current Protocols in Molecular Biology, Volume 67, Issue 1 2004:4.9.1-A.4.35. https://doi.org/https://doi.org/10.1002/0471142727.mb0409s67.

[51] Strategies for Detecting mRNA | Thermo Fisher Scientific - AT n.d. https://www.thermofisher.com/at/en/home/references/ambion-tech-support/northern-analysis/tech-notes/strategies-for-detecting-mrna.html (accessed August 20, 2022).

[52] van de Merbel NC. Protein quantification by LC-MS: A decade of progress through the pages of Bioanalysis. Bioanalysis 2019;11:629–44. https://doi.org/10.4155/BIO-2019-0032/ASSET/IMAGES/LARGE/FIGURE4.JPEG.

[53] Graves PR, Haystead TAJ. Molecular Biologist's Guide to Proteomics. Microbiology and Molecular Biology Reviews 2002;66:39–63. https://doi.org/10.1128/MMBR.66.1.39-63.2002.

[54] Watford S, Warrington SJ. Bacterial DNA Mutations. StatPearls 2021.

[55] Lovett ST, Gluckman TJ, Simon PJ, Sutera VA, Drapkin PT. Recombination between repeats in Escherichia coli by a recA-independent, proximity-sensitive mechanism. Mol Gen Genet 1994;245:294–300. https://doi.org/10.1007/BF00290109.

[56] Kong D, Masker W. Deletion between direct repeats in T7 DNA stimulated by double-strand breaks. Journal of Bacteriology 1994;176:5904–11. https://doi.org/10.1128/JB.176.19.5904-5911.1994.

Figures

Tables