



Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Faculty of Natural Resources and
Agricultural Sciences

CRISPR Cas9 system for plant genome editing within the European Union

Ida Eriksson

Department of Plant Biology

Master's thesis • 30 hec • A2E

Agricultural programme - Soil/Plant • ISSN 1651-5195, No 149

Uppsala, 2015

CRISPR Cas9 system for plant genome editing within the European Union

Ida Eriksson

Supervisor: Dr Jens Sundström, Swedish University of Agricultural Sciences,
Department of Plant Biology

Examiner: Prof. Christina Dixelius, Swedish University of Agricultural
Sciences,
Department of Plant Biology

Credits: 30 hec

Level: A2E

Course title: Independent Project/Degree Project in Biology – Master's thesis

Course code: EX0565

Programme/education: Agricultural programme - Soil/Plant

Place of publication: Uppsala

Year of publication: 2015

Number: 149

ISSN: 1651-5195

Online publication: <http://stud.epsilon.slu.se>

Keywords: CRISPR Cas9, site-directed mutagenesis, GMO, EU GMO legislation

Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Faculty of Natural Resources and Agricultural Sciences
Department of Plant Biology

Abstract

CRISPR (clustered regularly interspaced short palindromic repeat) Cas9 systems is a novel technique for genome editing. In this thesis a start-up attempt using CRISPR Cas9 to induce site-specific mutations in *Arabidopsis thaliana* has been made. By this means, a guide RNA sequence targeting the specific site of interest has been created, amplified and ligated into a vector containing a pcoCas9-expression cassette. The work performed shows that the CRISPR Cas9 method is easy to set up. This thesis also evaluates the current EU legislation on genetically modified organisms and the impact of extending it to CRISPR Cas9 as well as seven other new plant breeding methods. Subjecting CRISPR Cas9 as well as the majority of the other discussed plant breeding techniques to the EU GMO legislation requires considerable amendment and reworking, as the condition of traceability cannot be met. To avoid making amendments for each new technique that may be developed in the future and to increase legal certainty in this field, this thesis proposes a rewriting of the EU GMO legislation into a technology neutral biosafety legislation. Rather than focusing on the plant breeding techniques used in a particular case, the legislators should define a GMO by the phenotypic criteria an organism displays after induced genetic modification.

Sammanfattning

CRISPR (clustered regularly interspaced short palindromic repeat) Cas9 är en riktad mutationsteknik som under de senaste åren revolutionerat molekylärbiologin. Tekniken kan bland annat användas som en växtförädlingsteknik, men dess praktiska tillämpning inom EU är ännu oklar då Europeiska Kommissionen inte beslutat om CRISPR Cas9 samt sju andra växtförädlingstekniker kommer att omfattas av Eus nuvarande lagstiftning för genmodifierade organismer eller ej. Föreliggande arbete behandlar därav både CRISPR Cas9-tekniken som sådan och EUs lagstiftning för GMO.

I den tekniska delen av detta examensarbete har ett arbete påbörjats för att inducera riktade mutationer i *Arabidopsis thaliana*. Konkret innebär det att en guidesekvens (sgRNA) skapats, amplifierats och ligerats in i en vektor som kan uttrycka Cas9 för växter. Den färdiga vektorn kan användas för vidare transformation in i *Agrobacterium* och sedan *Arabidopsis*. I den regulatoriska delen som behandlar EUs lagstiftning för GMO, har en litteraturgenomgång av rådande lagstiftning gjorts. Frågor om krav för att genomföra fältförsök med *Arabidopsis* muterade med CRISPR Cas9 har också skickats till Jordbruksverket för att få en bild av deras syn på tekniken i förhållande till dagens GMO-lagstiftning. I den regulatoriska delen finns ett fokus på spårbarhetslagen, då majoriteten av de nya växtförädlingsmetoderna som kommissionen arbetar med genererar produkter som varken kan detekteras eller med säkerhet särskiljas som GMO. Analysen av litteraturstudien har lett fram till slutsatsen att hela lagstiftningen bör ändras till att vara produktbaserad snarare än teknikbaserad. En sådan ändring gör att lagens syfte blir tydligare samt att det i framtiden blir lättare att handlägga och bedöma effekterna av nya tekniker.

INTRODUCTION	5
PART I	6
BACKGROUND	6
CRISPR Cas – inheritable prokaryotic immune system	6
Adaptation	7
Expression.....	7
Interference	8
The CRISPR Cas9 system as a genetic engineering tool.....	9
Interrupting stamen development in Arabidopsis by use of CRISPR Cas9	11
Male sterility in hybrid seed production	12
Aim/hypothesis	13
MATERIALS AND METHOD	13
Vector plasmids	13
Mini- and Maxiprep samples of vectors plasmids	14
Finding CRISPR sites (sgRNA sequences).....	14
Designing PCR primers	15
Primers to assemble gRNA.....	15
Primers to use with NEBuilder Hifi DNA Assembly Cloning Kit	15
PCR	16
Assembly of gRNA (gRNA expression cassette)	16
Ligation.....	20
Ligation method A – T4 ligase.....	20
Ligation method B - NEBuilder Hifi DNA Assembly Cloning Kit.....	20
Transformation of ligated products into chemically competent <i>E. coli</i>.....	22
DNA Analysis	22
Colony PCR for transformed <i>E. coli</i>	22
Sequencing of colony PCR products	22
RESULTS	23
Designed PCR primers.....	23
Assembly of gRNA	23
Ligation and transformation	23
Sequencing	24
DISCUSSION	27
CONCLUSION.....	29

PART II	30
BACKGROUND	30
The EU legislation for genetically modified organisms (GMOs)	30
Cultivation of GMO within the EU territory	30
Detection and identification – the requirement for a policy	31
New plant breeding techniques	32
Field trials with CRISPR Cas9-generated plant mutants in Sweden	34
RESULTS AND DISCUSSION.....	35
CONCLUSION.....	37
ACKNOWLEDGEMENTS	37
REFERENCES	38
APPENDIX A.....
APPENDIX B.....
APPENDIX C.....
APPENDIX D.....

Introduction

Plant breeding comprise of many methods to improve crop yield and quality in order to meet the increasing demand for food, fibre and energy associated with a growing global population. The methods are also used to reduce environmental impact by allowing for lower consumption of pesticides, fertilisers and water.

Classical plant breeding was first used in late 19th century, see *e.g.* Lehrman *et al.*, (2014) and references within, and implemented Darwinian and Mendelian theories about selection and heritage of desired traits into elite plant varieties. By increasing mutation rates through use of X-rays or toxins, it was found possible to acquire new traits within plants. This technique of inducing random mutations came to be called mutagenesis. A century later, in early 1980, the first transgenic plant was created. This was the starting point for development of transgenic (GMO) methods to alter plant DNA. Recently, new genome editing techniques of site-directed mutagenesis have evolved. Two examples of such techniques are zinc finger nuclease (ZFN) as well as oligonucleotide directed mutagenesis (ODM) (Lusser *et al.*, 2011). These two are growing in importance and are increasingly seen as very real alternatives to the more traditional plant breeding methods.

In 2007, a working group was established by the European Commission to evaluate whether eight new techniques used to alter DNA, including ZFN and ODM, should fall within the EU GMO legislation (Directive 2001/18/EG and Regulation (EC) No 1829/2003). Seven out of these eight techniques can be used to make alterations of plant DNA. A report was published by the group in 2011 but no conclusion was reached regarding the EU GMO legislation for the techniques in question (Lusser *et al.*, 2011). The group did however state that “crops resulting from most of the techniques cannot be distinguished from conventionally bred crops and detection is therefore not possible”.

While gene technology allowing for changes to be made at specific sites in the genome, mutagenic technologies instead generate a number of changes that are more or less random (Miraglia *et al.*, 2004). According to the EU legislation, detection methods used to determine whether a specific crop is, in fact, a GMO, need to be available before that GMO crop is allowed on the market. Bearing this in mind, it is very understandable that mutagenesis as a plant breeding method is excluded from the EU GMO legislation (Directive 2001/18/EG) seeing as no such methods are available but any changes in the genome are indistinguishable from spontaneous mutations. As stated above, several of the new techniques, including ZFN and ODM allow for the induction of site-specific mutations, but the resulting crops are not distinguishable from crops bred using traditional mutagenic methods (Lusser *et al.*, 2011).

Since 2011 when the report by Lusser *et al.* was published, a new technique allowing for site-specific mutations, the CRISPR (clustered regularly interspaced short palindromic repeat) Cas9 method has been developed. This new technique has several advantages compared to other site-specific mutation methods such as ZFN, mostly thanks to the former being more easy to use and the ease of multiplexing *i.e.* genome editing at different loci simultaneously (Cong *et al.*, 2013; Mali *et al.*, 2013). Despite the continuous development of site-specific mutations techniques however, the

European Commission is still undecided as to whether all, some or any of the techniques should be subject to the traditional GMO legislation.

This thesis will seek to explain and evaluate both the CRISPR Cas9 method itself as well as the impact of potential changes in the EU legislation whereby the method is considered to fall within the scope of GMO legislation. Part I, focusing on the method, is structured by way of a laboratory analysis where the CRISPR Cas9 method is set up to induced site-specific mutations in *Arabidopsis thaliana* in a continued work. Part II draws on the scientific conclusions from part I and offers an outline of the current EU legislation on GMO. The report also offers a discussion of the current legislation as well as of the feasibility and the impact of extending its reach to the CRISPR Cas9 method.

Part I

Background

Plant breeding is, as previously stated, vital in order to improve crop plants and at the same time contribute to a sustainable food production. A new plant breeding technique allowing for site-directed mutation in the plant genome is the CRISPR Cas9 method. Before explaining how the CRISPR Cas9 method can be used for prokaryote genome editing, the inheritable prokaryotic immune system has to be outlined as the CRISPR technique is based upon this natural phenomena (Jinek *et al.*, 2012).

CRISPR Cas – inheritable prokaryotic immune system

Bacteria and archaea can resist viral and plasmid challenges by integrating short fragments of foreign nucleic acid into the genome at a specific locus consisting of short (20-50 bp) DNA repeat sequences (Sorek *et al.*, 2013). These sequences are separated by spacer sequences of equal lengths; creating a repeat-spacer-repeat pattern that makes a CRISPR locus, figure 1a. There is a common repeat sequence within a CRISPR locus but the repeats can vary in both length and sequence when comparing two or more CRISPR loci, leaving a high diversity among CRISPR repeats (Kunin *et al.*, 2007).

A prokaryotic chromosome can contain one or more CRISPR loci (Godde & Bickerton, 2006). Within species that contain two or more CRISPR loci, there is a leader sequence rich in AT flanking one side of each locus (Jansen *et al.*, 2002), figure 1a. The leader sequence consists of 300-500 base pairs (bp). In turn, CRISPR-associated sequence (Cas) genes have been found on either side of the CRISPR loci. There are several different types of Cas genes that codes for proteins within different families. However, the Cas1- and Cas2 genes are conserved in all genomes that contain CRISPR loci.

There are three classified types of CRISPR Cas systems in prokaryotes and all three systems mediate immunity to invading viruses in a common three-stage process of adaptation, expression and interference (reviewed by Makarova *et al.*, 2011). The difference between the systems is what other Cas genes, besides Cas1 and Cas2, that act mainly in the stages of expression and interference. No further comparison of the different CRISPR Cas systems will be made in this report. Instead, in order to

understand how the CRISPR method can be used as a genome engineering tool in eukaryotes, the stages of adaptation, expression and interference for the type II CRISPR Cas inheritable immune system in prokaryotes will be outlined below.

Adaptation

In the adaptation stage, short pieces of viral DNA are recognised, processed and integrated in the chromosomal CRISPR locus at the leader side (Barrangou *et al.*, 2007), figure 1b. The integrated sequences will bring out the function of a genetic memory and prevent the host from being infected by viruses with the same sequence. However, what sequence from the viral genome that is chosen to be integrated is not arbitrary. Instead, it has been found by analysing newly added spacer sequences that they contain a short motif downstream of the proto-spacer that is specific to the CRISPR Cas variant within the host (Mojica *et al.*, 2009). Hence, the selection of spacer precursors (proto-spacers) from the invading DNA appears to be determined by the recognition of a proto-spacer-adjacent motif (PAM) with a sequence of 5' - NGG- 3'. How the PAM sequence is recognised remains still unknown. What is known is that the GG motif is required for proto-spacer plasmid DNA elimination by CRISPR Cas in bacterial cells (Jinek *et al.*, 2012; Saprunauskas *et al.*, 2011). However, the integration of the proto-spacer is followed by a duplication of a repeat, leaving a new spacer-repeat sequence within the CRISPR locus (Makarova *et al.*, 2006).

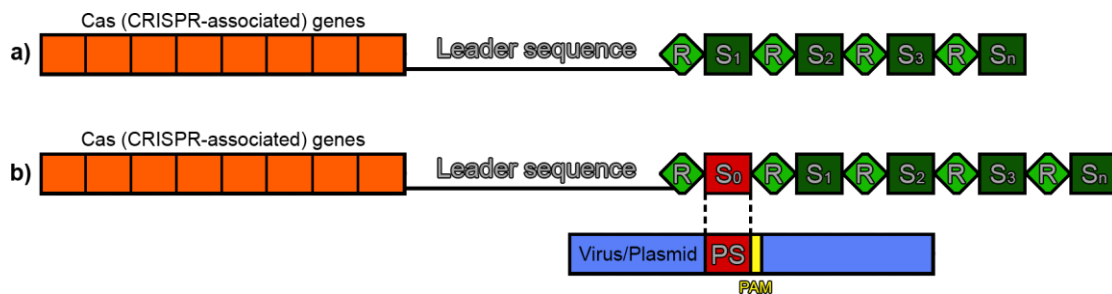


Figure 1. **a)** A CRISPR locus in a prokaryote chromosome with the characteristic repeat(R)-spacer(S_x)-repeat pattern. **b)** A prokaryote can recognise a PAM sequence within an invading virus/plasmid and integrate foreign nucleic acid from it, *i.e.* a proto-spacer (PS), in its CRISPR locus. The integration of the proto-spacer creates a new spacer (S₀) and a repeat sequence (R). The figure is modified from van der Oost *et al.* (2009).

Expression

Following adaptation, a second stage called expression comes next in the CRISPR Cas mediated immunity (Makarova *et al.*, 2011). In the leading sequence flanking the CRISPR locus, there is a promoter sequence that synthesises a primary transcript of the CRISPR locus (Brouns *et al.*, 2008; Pul *et al.*, 2010). The primary non-coding transcript that is synthesised is called a pre-CRISPR RNA (pre-crRNA) (Deltcheva *et al.*, 2011). In turn, trans-activating CRISPR RNA (tracrRNA) that is complementary to the pre-crRNA and which pairs with its repeat sequence, cleaves the pre-crRNA into mature CRISPR RNAs (crRNAs) by the activities of endogenous RNase III and in the presence of a DNA endonuclease referred to as the Cas9 protein.

Interference

During the third and final stage of interference, Cas9 protein cleaves the double-stranded DNA (dsDNA) of the invading virus. The cleaving process starts by activation of the Cas9 which occurs through a base-pairing between tracrRNA and crRNA in the repeat sequence (Jinek *et al.*, 2012). In turn, the Cas9 protein needs to be guided to the right location on the virus in order for the cleavage to be site-specific. This guidance occurs through a base pairing between the crRNA and the proto-spacer DNA at a location where the short PAM sequence, previously mentioned to also have a role in the adaptation stage, is present downstream of the target DNA, figure 2. Cas9 is thereby both required for the maturation of crRNA in the expression stage and dsDNA cleavage in the interference stage. tracrRNA can in a similar way be described to have multiple functions of processing pre-crRNA by the enzyme RNaseIII in the expression stage (Deltcheva *et al.*, 2011) and activating Cas9 resulting in tracrRNA:crRNA-guided dsDNA cleavage (Jinek *et al.*, 2012).

When the Cas9 protein has been activated and guided to a specific position in the invading nucleic acid, the cleavage of the dsDNA will take place. At this stage, the two nuclease domains that the Cas9 contains *i.e.* RuvC-like and HNH nuclease domain (Makarova *et al.*, 2006), play a major role. Sapranaukas *et al.*, (2011) found that mutations in the RuvC and/or the HNH-motif abolish the ability of Cas9 to interfere with the plasmid DNA. The HNH-domain has thereby been considered essential in the DNA degradation step, as the HNH-domains are common components in nucleases that act on double-stranded DNA. It is today known that the HNH-domain cleaves the complementary strand three bp upstream of the PAM of the invading nucleic acid (Jinek *et al.*, 2012). The RuvC-domain cleaves the noncomplementary strand at one or several sites at a location of three to eight bp upstream of the PAM sequence (figure 2).

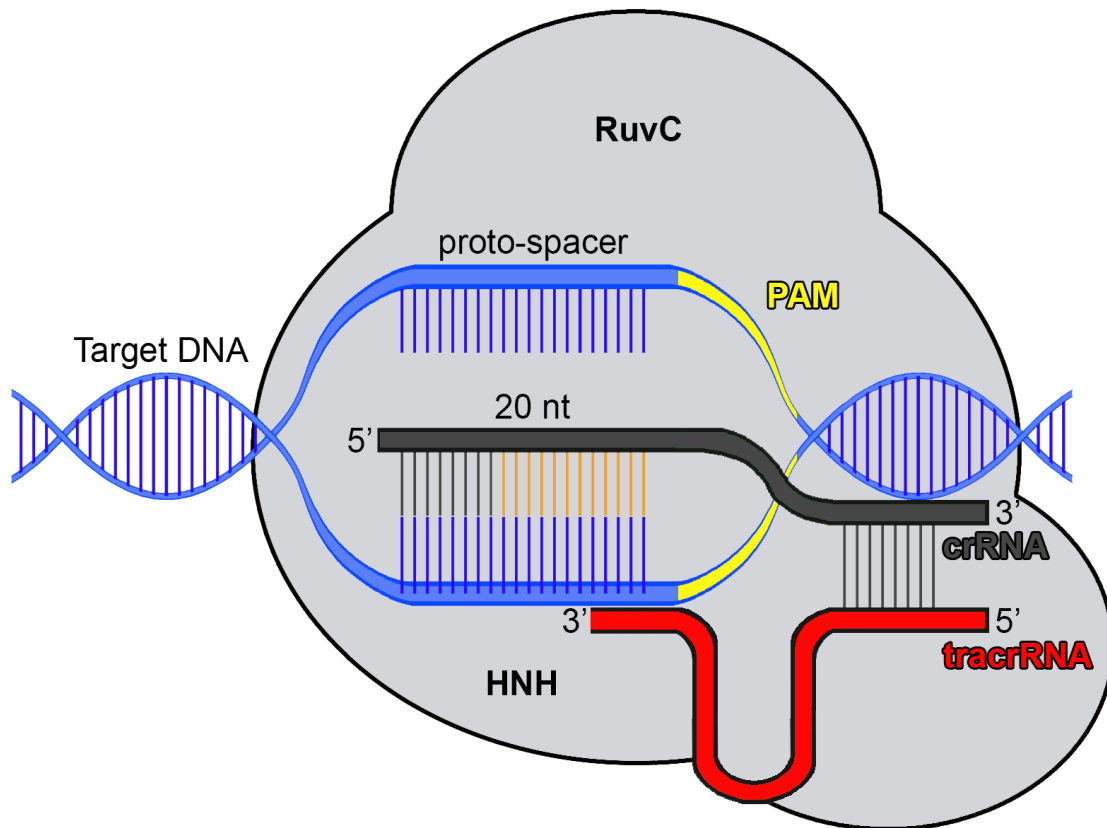


Figure 2. A model of the interference stage in the native prokaryotic CRISPR Cas defense system. Cas9 protein (grey) is activated by base pairing of crRNA (black) and tracrRNA (red). The Cas9 protein is in turn guided to a specific cleavage site in the invading nucleic acid through base pairing of crRNA and the proto-spacer DNA in the presence of an adjacent PAM sequence (yellow). The HNH- and RuvC-domain in the Cas9 protein will cleave the dsDNA and hence destroy the invading nucleic acid. The figure is modified from Bortesi & Fischer (2015) and Jinek *et al.* (2012).

The CRISPR Cas9 system as a genetic engineering tool

The two-RNA structure that is formed between crRNA and tracrRNA in prokaryotes to guide the Cas9 protein to a specific location in the invading DNA, was in 2012 found to guide Cas9 and induce an equal dsDNA cleavage in eukaryotes when engineered as a single RNA chimera (Jinek *et al.*, 2012). At time for first design and trial, this structure was called a crRNA-tracrRNA chimera, but is today known as a single-guide RNA (sgRNA).

A sgRNA contains a recognition sequence of 20 nucleotides to target the complementary DNA strand (Jinek *et al.*, 2012), figure 3. Out of these 20 nucleotides, the ones located 5 to 12 positions adjacent to a U-shaped structure fusing the 3' end of crRNA to the 5' end of tracrRNA make the seed sequence. The seed sequence is particularly important for target recognition, but does not leave the remaining nucleotides within the target sequence as useless. On the contrary, the first nucleotide in the target sequence *i.e.* the nucleotide located 20 positions adjacent to the U-shaped structure, is important for transcription. The first nucleotide in the gRNA sequence should be a guanine (G) as the U6 RNA polymerase III prefers having a G as the first base in its promoter (Ran *et al.*, 2013). The G-nucleotide and the U6 promoter are in

other words important for the expression of sgRNA. Finally, in order for the Cas9 cleavage to take place, the target sequence has to be located adjacent to a PAM sequence of 5' -NGG- 3'. Adding all together, the general sequence template to use when designing and building sgRNAs for Arabidopsis is 5' -G N₁₉ NGG- 3' (Mali *et al.*, 2013).

A Cas9 protein programmed with a sgRNA induces a double strand break (DSB), but it is not the cleavage itself that generates the desired genetic mutation. Instead, the mutation is an outcome of the repair pathway of the broken DNA strand (Maresca *et al.*, 2013; Rouet *et al.*, 1994). DSBs can be repaired by homologous recombination (HR), also commonly discussed in terms of homologous directed repair (HDR), or non-homologous end joining (NHEJ) (Rouet *et al.*, 1994). NHEJ is the major repair pathway (Ferreira & Cooper, 2004) and may result in random insertion or deletion of nucleotides when the two DNA ends are ligated together, as the joining of the DNA ends occurs irrespective of end sequences (Rouet *et al.*, 1994). Knowing this, it is understandable that a DSB caused by CRISPR Cas9 and repaired by NHEJ will result in a mutation. This mutation will in turn be indistinguishable to any other mutation resulting from the repair of a DSB by NHEJ (Lusser *et al.*, 2011).

Natural cell repair of broken DNA through HR/HDR will, in contrast to NHEJ, result in a precise repair of the DNA (Rouet *et al.*, 1994). This precision is a result of the cell's use of homologous sequences around the DSB that work as templates in the repair process. When it comes to HR in the context of genetic modification, exogenously introduced repair templates in form of double-stranded DNA or a single-stranded DNA oligonucleotide can be used to induce a targeted point mutation, targeted genomic deletion or targeted insertions of genetic elements that are associated with genomic deletions (Chen *et al.*, 2011). The only requirement is that the exogenously introduced repair templates have homology sequences flanking the site of modification. Worth mentioning is that DNA templates with a transgene also can be inserted through NHEJ and not exclusively within the repair process of HR. Moreover, the HNH and RuvC nuclease domains within the Cas9 protein can be mutated to exploit new functions (Barrangou *et al.*, 2007). This can, for instance, generate site-specific single-strand DNA breaks instead of double-strand breaks. However, using CRISPR Cas9 to induce DSBs that are repaired by NHEJ without the presence of a repair template is the technique that is relevant for this thesis.

Since the first discovery of the possibility to use a Cas9 protein programmed with a sgRNA to induce site-specific DSBs, it has been found that multiple guide RNAs (gRNAs) with different sequences can be introduced into a single CRISPR array (Cong *et al.*, 2013; Mali *et al.*, 2013). Using several different gRNAs allows for simultaneous genome editing at different sites. The use of the CRISPR Cas9 method can thereby be described as a technique which is easy to use, inexpensive compared to other methods and as a technique that has a wide range of purposes (Bortesi & Fischer, 2015). It is therefore understandable why research groups across the world are working with the CRISPR Cas9 method, especially when knowing that CRISPR Cas9 can induce site-specific, endogenous, genetic changes in a wide range of species including human (Jinek *et al.*, 2013), mouse (Cong *et al.*, 2013), zebrafish (Hwang *et al.*, 2013) and plants (Feng *et al.*, 2013; Li *et al.*, 2012; Nekrasov *et al.*, 2013; Shan *et al.*, 2013; Xie *et al.*, 2014).

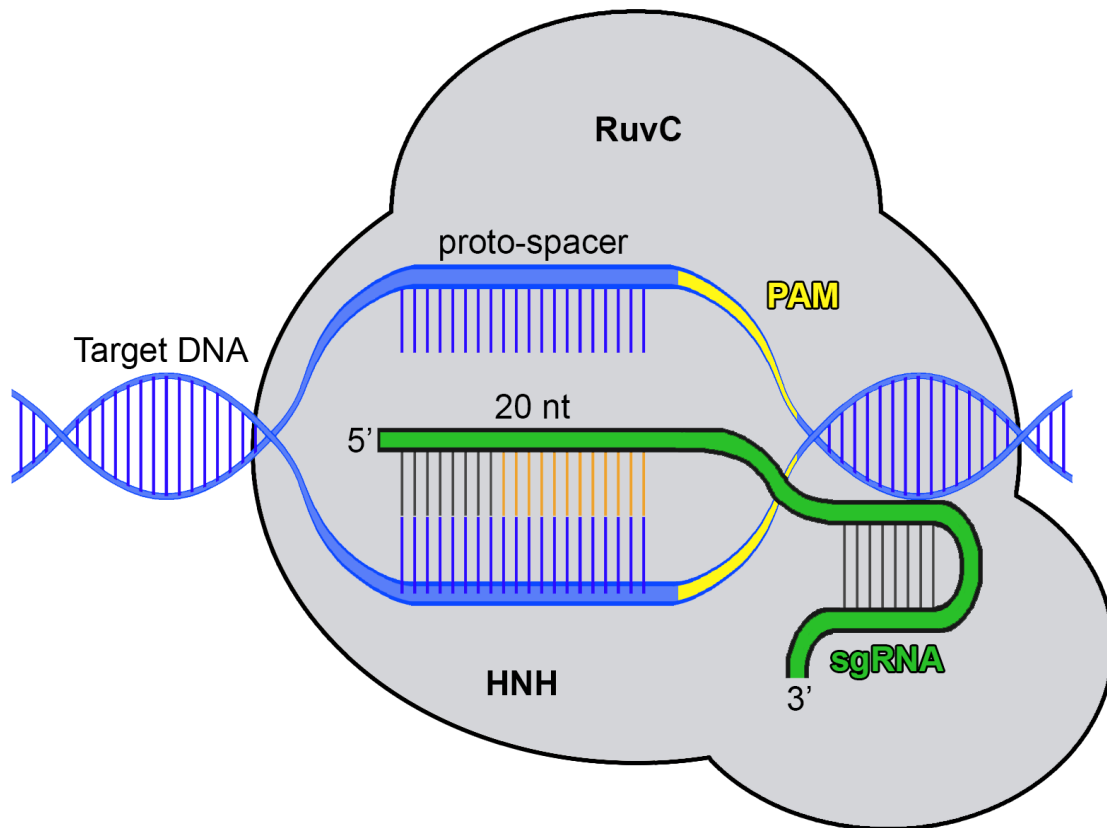


Figure 3. A model of the CRISPR Cas9 system for eukaryotic genome editing. A Cas9 protein (grey) is programmed with a sgRNA (green). The sgRNA target sequence, binding to the proto-spacer, consists of 20 nucleotides; 5-12 nucleotides out of these make the seed sequence (orange) that is particularly important for target recognition. The presence of a PAM sequence of 5' -NGG- 3' (yellow) is required for Cas9 cleavage of the dsDNA by the nucleases RuvC and HNH. The figure is modified from Bortesi & Fischer (2015) and Jinek *et al.* (2012).

Interrupting stamen development in *Arabidopsis* by use of CRISPR Cas9

The structure of the mature flower of *Arabidopsis thaliana* is typical for *Brassicaceae* *i.e.* an important family of flowering plants (Angiosperms). It has a sterile perianth consisting of four sepals and four petals that surrounds the reproductive organs (Smyth *et al.*, 1990). The reproductive organ does in turn consist of six stamens and two fused carpels that build up the pistil.

The development process of the mature flower of *Arabidopsis* can be divided into 14 distinct stages from floral initiation to anthesis (Smyth *et al.*, 1990). The process starts with the initiation of a floral buttress on the apical meristem and ends with the opening of a one millimetre long bud. Total time taken for all 14 stages to proceed is 13.25 days and approximately 1.9 buds are being initiated on average each day. Looking specifically at stamen development, it starts in floral stage 5 with the appearance of stamen primordia (Smyth *et al.*, 1990). In later stages, the primordia will bulge out and form a filament with anthers on top. Within an anther, four pollen sacks surrounded by a distinct cell layer called the tapetum layer will develop. The tapetum layer functions as a nurturing cell-layer for the developing pollen grains, which, in turn, are released during anthesis.

Development of stamen primordia specification depends on the interaction of a number of genes as well as on a correlated action of the mitochondrion (Carlsson *et al.*, 2008). Alves-Ferreira *et al.*, (2007) have shown that genes located in the tapetum are active during stamen development. The functions of individual genes are although unknown. A way to study gene function, including function of genes being closely related to each other, is to induce point mutation(s) within the protein-coding region of the genes of interest, *e.g.* through using the CRISPR Cas9 method. In case of genes with putative functions during tapetum development (figure 4) CRISPR Cas9 can be used to induce individual gene knockouts so as to interrupt stamen development and thereby get a deeper understanding of genes regulating male fertility.

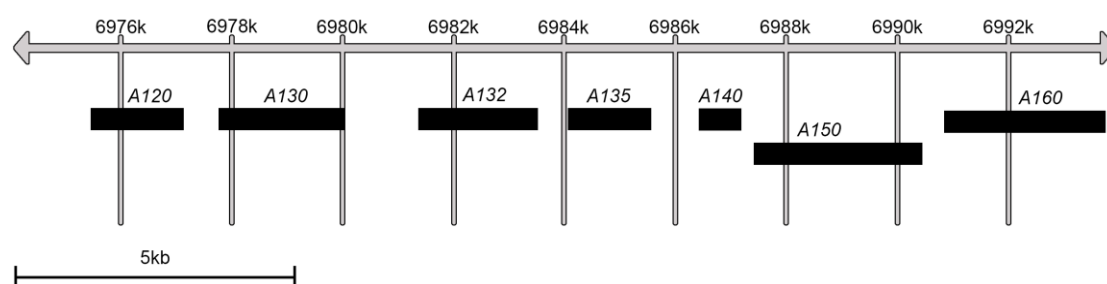


Figure 4. Genes located on chromosome 1 in Arabidopsis and implicated in stamen development. Here named A120, A130, A132, A135, A140, A150 and A160.

Male sterility in hybrid seed production

Hybrid seed production involves cross-pollination of two homozygote parental lines. This will generate a heterozygote progeny that displays a superior performance called heterosis or hybrid vigour. Heterosis in the progeny is exhibited as better plant growth and a higher biomass yield compared to the parental lines and/or conventional lines (Birchler *et al.*, 2003).

In order for cross-pollination to occur, self-pollination of the female plant has to be prevented. This can be accomplished through manual removal of anthers or use of chemicals. In fact, it was found already in 1954 that subjecting plants to α -naphthaleneacetic acid resulted in exclusive production of pistillate flowers (Wittwer & Hillyer, 1954). However, chemicals may cause damage to the rest of the plant and manual removal of anthers is time- and labour intensive (Perez-Prat & van Lookeren Campagne, 2002). It may be possible to manually remove anthers in monoecious plants such as corn, where the male and female flowers are located at separate parts of the plant. This is possible but difficult in plants with bisexual flowers, *e.g.* oilseed rape, as the anthers then have to be removed from each flower.

Male-sterility of a female plant, as an outcome of an induced genetic defect to nuclear or cytoplasmic genes, is to be preferred before manual emasculation or use of chemicals (Horner & Palmer, 1995; Perez-Prat & van Lookeren Campagne, 2002). Causing gene defects within hybrid seed production does although require knowledge about genes being involved in stamen development; highlighting the value of inducing site-directed mutations discussed in previous section. Making such trials will

give a deeper understanding of how male-sterility can be induced and may thereby enhance hybrid seed production.

Aim/hypothesis

The aim of part I is to make a start-up attempt for using CRISPR Cas9 to induce site-specific mutations in *Arabidopsis thaliana*. To be more precise, the aim is to transform a binary vector containing a pcoCas9-expression cassette for plant expression of Cas9 and a sgRNA sequence targeting gene *A120* (figure 4) into *E. coli* (DH5 α). Whether the aim is fulfilled or not will be checked by DNA sequencing of a colony from the transformed *E. coli*. The aim of part I is also to design sgRNAs to target the remaining genes that are illustrated in figure 4. These gRNAs should be possible to use in CRISPR arrays targeting a single site within Arabidopsis, but also possible to use in CRISPR arrays that allowing for simultaneous editing at several sites.

A satisfying result of this aim will make the foundation of a greater research project in which the created binary vector (containing pcoCas9-expression cassette and sgRNA targeting gene *A120*) can be transformed into *Agrobacterium*. A confirmation of the usability of this method will also allow creation of additional vectors containing the remaining gRNAs that have been designed. However, once having a vector containing desired gRNA sequence(s) cloned into *Agrobacterium*, it can be transformed into Arabidopsis. Plants of the T₀-generation will be heterozygote with regards to the CRISPR Cas9 transgene, whereby homozygote T₁-plants will be selected and further used, see appendix C for more information. By making *Agrobacterium* transformation into Arabidopsis, disruption/knockout of genes will be made as a result of NHEJ of induced CRISPR Cas9 DSB(s). This will in turn allow testing of the hypothesis that knockout of abovementioned genes reduce male fertility.

Materials and Method

Vector plasmids

Two vector plasmids have been used. Both vectors were ordered from Addgene.

1. pUC119-gRNA (Addgene plasmid # 52255)

The pUC119-gRNA vector contains an Arabidopsis U6 promoter that targets the AtPDS3 gene, target site 1, to drive sgRNA expression (Addgene, 2015b).

2. pFGC-pcoCas9 (Addgene plasmid # 52256)

pFGC-pcoCas9 is a binary vector that contains a pcoCas9-cassette for plant expression of Cas9 (figure 5). It also contains a multiple cloning site for insertion of guide RNA as well as two selection markers. The kanamycin resistant gene is used for bacteria selection at lab while the BAR-gene, a glufosinate resistance gene, is used for selection in plants (useful for confirming that the CRISPR Cas9 transgene is absent in chosen progenies of the T₁-generation, appendix C).

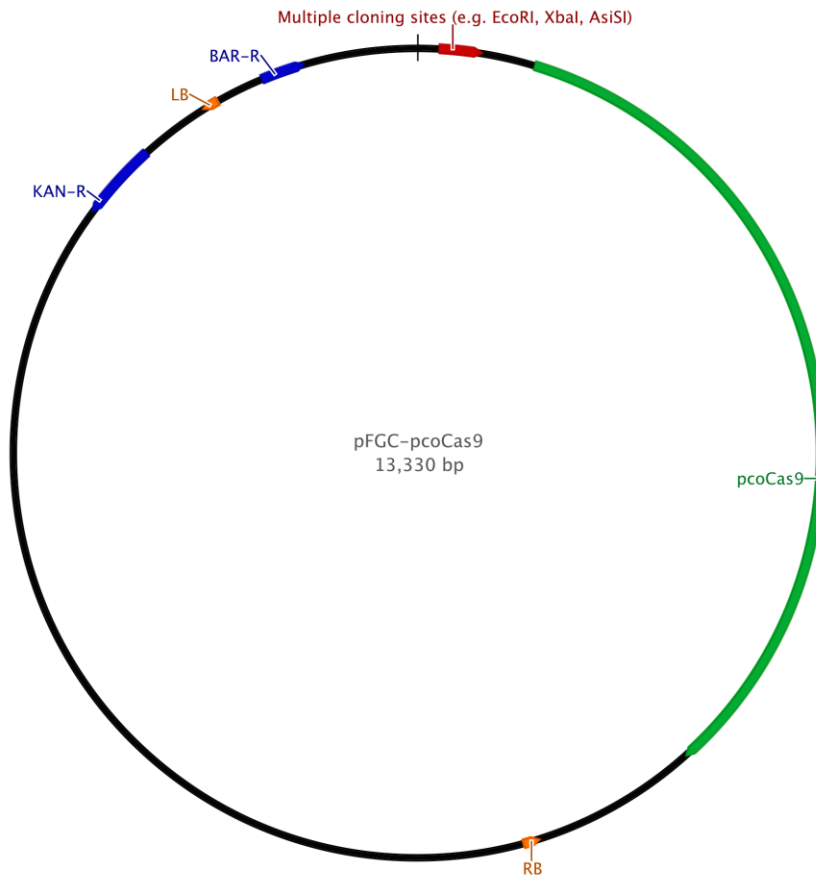


Figure 5. The binary vector pFGC-pcoCas9 (Addgene plasmid # 52256). The figure is modified from (Addgene, 2015a).

Mini- and Maxiprep samples of vectors plasmids

The vector plasmids arrived as stab cultures (pFGC-pcoCas9/DH5 α and pUC119-gRNA/DH5 α). Bacteria containing the vectors were re-streaked on lysogeny broth (LB)-agar plates. Vector pUC119-gRNA/DH5 α was grown on a LB_{ampicillin}-plate and vector pFGC-pcoCas9/DH5 α on a LB_{kanamycin}-plate. Single colonies were picked from these plates and incubated overnight in liquid LB with corresponding antibiotics added. In turn, two Miniprep samples of vector pUC119-gRNA/DH5 α as well as three Miniprep samples of vector pFGC-pcoCas9/DH5 α were made. User manual was followed (Life Technologies).

An additional Maxiprep sample of vector pFGC-pcoCas9/DH5 α was made approximately two months after making the Minipreps. User manual was followed (Life Technologies).

Finding CRISPR sites (sgRNA sequences)

To find CRISPR sites *i.e.* gRNA sequences targeting genes *A120*, *A130*, *A132*, *A135*, *A140*, *A150* and *A160*, the programme Geneous 8.0.5 was used. For each gene, full-length cDNA and full-length genomic information for Arabidopsis were downloaded from the website arabidopsis.org. A translation of the cDNA was then performed in the Geneous programme. Translation frames were chosen to avoid stop codons within the translated sequences.

When having a translated cDNA sequence for a gene, the cloning function in Geneous, which includes an option of finding CRISPR sites, was used. Target sequence was set at G(N19). PAM sequence used was NGG. When searching for endogenous CRISPR sites with these target- and PAM sequences, a number of results were shown. CRISPR sites located in exons, in the beginning of the protein-coding region of genes and with a low off-target frequency were chosen and further used in creation of polymerase chain reaction (PCR) primers.

Designing PCR primers

Primers to assemble gRNA

Sets of four PCR primers, two forward and two reverse, were designed in Geneous 8.0.5 for each gene. These were designed to work with a PCR method published by Li *et al.* (2013). Out of the four primers in each set, one forward and one reverse primer (U6_F2 and U6_R1) were gene specific *i.e.* based on the CRISPR site found. The other forward and reverse primers included in each set, U6_F1 and U6_R2, were designed to work for all genes. Sequences for these non-specific primers were built with regards to the multiple cloning site of vector pFGC-pcoCas9.

Two different types of the non-specific forward and reverse primers were designed in order to have the option of creating CRISPR arrays that target a single genomic site as well as arrays that target multiple sites. This is further explained in the section of ligation, method A. Sequence templates used in the primer designing process are listed in table 1 and adapted from Li *et al.* (2013).

Table 1. Naming of PCR primers and templates used in sequence building.

Primer name	Templates used in sequence building (5' to 3')
U6_F1 type 1	NNN + <i>EcoRI</i> + AGAAATCTCAAATTCG
U6_F1 type 2	NNN + <i>XbaI</i> + AGAAATCTCAAATTCG
U6_F2 *	The 5'-most N23 of a gRNA target site + GTTTAGAGCTAGAAATAGC
U6_R1 *	Reverse complement of the 5'-most N23 of a gRNA target site + AATCACTACTTCGTCTCT
U6_R2 type 1	NNN + <i>XbaI</i> + TAATGCCAACTTTGTACA
U6_R2 type 2	NNN + <i>AsiSI</i> + TAATGCCAACTTTGTACA

* The last three numbers in targeted gene, *e.g.* 140 in gene *A140*, complete primer name. For some genes, more than one primer was design as two CRISPR sites were chosen. In cases of multiple primers targeting the same gene, the tree last numbers in gene name are following by a letter, *e.g.* U6_F2_120a.

Primers to use with NEBuilder Hifi DNA Assembly Cloning Kit

When having a gRNA expression cassette, it can be run in an additional PCR reaction to work with the NEBuilder Hifi DNA Assembly Cloning Kit (ligation). As this method was tried, further explained in the section 'ligation method B', two additional primers had to be designed. These have been named NEB_U6_F1 type 1 and NEB_U6_R2 type 1 and modified from primer U6_F1 type 1 and U6_R2 type 1 respectively. NEB_U6_F1 type 1 and NEB_U6_R2 type 1 were ordered from Invitrogen (Life Technologies).

The NEBuilder Hifi DNA Assembly Cloning Kit can also, apart from being used for ligation of a single gRNA expression cassette, be used to assemble up to five DNA fragments in a single linear vector (New England BioLabs, 2015). The only requirement is that the DNA fragments, PCR-amplified, have overlapping ends of 15-30 bp. This method is further described in the section ‘ligation method B’, but a description of primer design follows below.

For the purpose of making CRISPR arrays simultaneously targeting five sites, several new primers were created. The first primers were created in the Geneious programme and named U6_F1 NEB_x and U6_R2 NEB_x. U6_F1 NEB_x were based on the common sequence in U6_F1 type 1 and U6_F1 type 2, *i.e.* AGAAATCTCAAATCCG and U6_R2 NEB_x on the common sequence in U6_R2 type 1 and U6_R2 type 2, *i.e.* TAATGCCAACTTTGTACA. In addition to these sequences, individual sequences of 20 nucleotides were added to make the complete primer sequences, see appendix A table 1. These primers will be used in five separate PCR reactions with the gRNA expression cassettes that should be ligated together (further explained in the section ‘ligation method B’).

In order to make remaining primers to use for multiple assemble of gRNA, the programme NEBuilder Assembly Tool 1.8.0 was used. A FASTA sequence of vector pFGC-pcoCas9 was uploaded in the NEBuilder Assembly Tool programme. The vector was virtually cut by *EcoRI* and *XbaI*. In turn, sequences for gRNA expression cassettes flanked by primer sequences U6_F1 NEB_x and U6_R2 NEB_x were uploaded. By using preferences set to a minimum primer length of 18 bp and a minimum overlap of 15 bp, primer sequences that will overlap the gRNA expression cassettes were created. This will enhance the ligation. All primers were ordered from Invitrogen (Life Technologies).

PCR

Assembly of gRNA (gRNA expression cassette)

The vector pUC119-gRNA was used as a PCR template to assemble individual expression cassettes including specific gRNA sequences for all genes. The complete assembly of a single gRNA expression cassette required three PCR reactions; round 1, round 2.1 and round 2.2. All rounds, *i.e.* the three PCR reactions, are further explained below in a general approach with respect to primers used. The PCR reactions and the assembly of a gRNA expression cassette are graphically illustrated in figure 8 on next page to help the reader understand the PCR reactions described. Details of which primers that were used to assemble gRNA expression cassettes for each gene are listed in table 2.

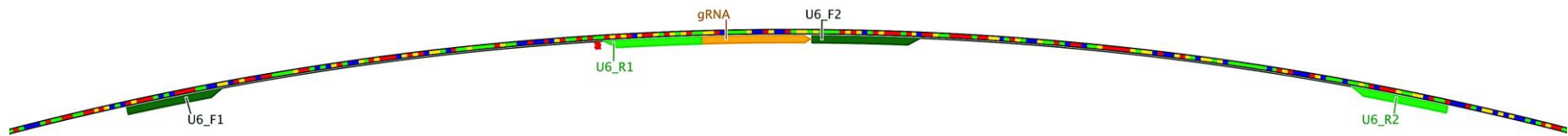


Figure 8. Illustration of how primers U6_F1, U6_R1, U6_F2 and U6_R2 assemble a gRNA (CRISPR site) in PCR reactions 1, 2.1 and 2.2. The whole sequence region *i.e.* primers as well as the assembled gRNA make the gRNA expression cassette. This gRNA expression cassette can, due to the flanking restriction sites in primers U6_F1 and U6_R2 (table 1), be ligated to vector pFGC-pcoCas9. This is further explained on next page. A part of the sequence in between U6_F1 and U6_R1 is truncated because of space constraints.

Phusion High-Fidelity DNA Polymerase was used in all PCR reactions and the user manual was followed when making the mastermix.

PCR Round 1

Two separate reactions are included in the first PCR round. In the first reaction, the left part of the gRNA expression cassette will be generated while the right part of the expression cassette is generated in the second reaction, figure 8. The defining gRNA sequence will be amplified in both reactions as its sequence is included in primer U6_R1 and U6_F2.

Following two reactions were set up:

18 µl mastermix + 0.5 µl 10µM U6_F1+ 0.5 µl 10µM U6_R1 + 1 µl pUC119-gRNA

18 µl mastermix + 0.5 µl 10µM U6_F2+ 0.5 µl 10µM U6_R2 + 1 µl pUC119-gRNA

PCR round 1 was brought out according to protocol in figure 6 below.

98°C	98°C	62°C	72°C	72°C	12°C
30''	10''	20''	30''	10'	∞
× 35					

Figure 6. Illustration of conditions in PCR round 1. The figure shows conditions for initial denaturation, the 35 cycles of denaturation, annealing and extension as well as the final extension. Single apostrophe (') represent minutes and double apostrophes (') represent seconds.

PCR products from round 1 were purified by use of a PCR purification kit (Life Technologies) before used in PCR round 2.1. Concentration measurement of the purified products was made as well.

PCR Round 2.1

PCR round 2.1 joins the right and the left part of the gRNA expression cassette together so as to get a complete cassette.

Purified PCR products from round 1 was used in following reaction:

5 µl (U6_F1+U6_R1) + 5 µl (U6_F2+U6_R2) + 10 µl mastermix

PCR round 2.1 was brought out according to protocol in figure 7 below.

98°C	98°C	62°C	72°C	12°C
3'	30''	30''	10''	∞
× 1				

Figure 7. Illustration of conditions in PCR round 2.1. The figure shows conditions for initial denaturation, the single cycles of denaturation, annealing and extension as well as the final extension. Single apostrophe (') represent minutes and double apostrophes (') represent seconds.

PCR Round 2.2

In the final PCR round, the complete gRNA expression cassette generated in PCR round 2.1, will be amplified so as to get a high concentration to use for ligation.

The following reaction was set up with products from PCR round 2.1:

17 μ l mastermix + 0.5 μ l U6_F1 + 0.5 μ l U6_R2 + 2 μ l product from PCR round 2.1.

Protocol run was the same as for PCR round 1 shown in figure 6.

The final PCR product from round 2.2 was purified using a PCR purification kit (Life Technologies). The purified product was in turn electrophoresed on 1% agarose gel (TBE-buffer) with a few drops of ethidium bromide (EtBr) added. By doing this, the product size could be checked.

A PCR product from round 2.2 will further be referred to as a gRNA expression cassette, highlighting the presence of assembled gRNA as well as the region in between forward primer U6_F1 as well as reverse primer U6_R2 within the product. For a graphical illustration, see figure 8.

Table 2. Primers used in PCR reactions to assemble gRNA sequences.

Gene (gRNA target)	Primers used in PCR reaction
<i>A120</i>	U6_F1 type 1 U6_F2 120a or U6_F2 120b U6_R1 120a or U6_R1 120b U6_R2 type 1
<i>A130</i>	U6_F1 type 2 U6_F2 130a or U6_F2 130b U6_R1 130a or U6_R1 130b U6_R2 type 2
<i>A132</i>	U6_F1 type 1 U6_F2 132a or U6_F2 132b U6_R1 132a or U6_R1 132b U6_R2 type 1
<i>A135</i>	U6_F1 type 2 U6_F2 135 U6_R1 135 U6_R2 type 2
<i>A140</i>	U6_F1 type 1 U6_F2 140 U6_R1 140 U6_R2 type 1
<i>A150</i>	U6_F1 type 2 U6_F2 150 U6_R1 150 U6_R2 type 2
<i>A160</i>	U6_F1 type 2 U6_F2 150 U6_R1 150 U6_R2 type 2

Ligation

Two different ligation methods, A and B, were used in order to ligate the gRNA expression cassette into vector pFGC-pcoCas9/DH5 α . gRNAs targeting all seven abovementioned genes were assembled through PCR reactions. As a proof of concept, the gRNA expression cassette containing assembled gRNA targeting gene *A120* (made from primers U6_F1_type 1, U6_F2_120a, U6_R1_120a and U6_R2_type 1) was from this point further used. This sample will be referred to as gRNA120a expression cassette.

Ligation method A – T4 ligase

Preceding ligation, the gRNA120a expression cassette and vector pFGC-pcoCas9 (MiniPrep sample with a concentration of 122.4 ng/ μ l) were set up in separate fast digest restriction enzyme reactions. Enzymes used in the reactions were *EcoRI* and *XbaI* (Life Technologies). User manual was followed. The optional step of heat inactivation was not performed. Instead, direct following enzyme digestion, both samples were electrophoresed on 1% EtBr agarose gel. Clear bands shown were cut from the gel and purified by use of a gel purification kit (Life Technologies).

Concentration measurement for the two gel-purified samples was made post purification. In turn, these were ligated together using T4 ligase (Life Technologies). User manual was followed for making reaction buffer. The ligation reaction were brought out at 16°C for approximately 16 hours.

Ligation method A to insert two gRNA expression cassettes in a single vector

As mentioned in the section of ‘primers designed to assemble gRNA’, two different types of the non-specific forward and reverse primers, U6_F1 and U6_R2 were created in order to allow insertion of two gRNA expression cassettes in a single vector. Trials with multiple insertions of gRNA sequences were not performed in this experiment. However, following method can be used in future to insert two gRNA expression cassettes into a single vector for Cas9 expression.

1. Make a complete ligation reaction (described above for gRNA120a) for any gRNA expression cassette made from primers U6_F1_type 1 and U6_R2_type 1 (table 2).
2. Set up a fast digest restriction enzyme reaction in which the ligated sample is digested by enzymes *XbaI* and *AsiSI*. An additional enzyme digest reaction would be set up as well in which any gRNA expression cassette made from primers U6_F1_type 2 and U6_R2_type 2 reacts with *XbaI* and *AsiSI*.
3. Make a T4 ligation (described above). In this reaction, the gRNA expression cassette digested by *XbaI* and *AsiSI* will be ligated to the end of the first introduced gRNA expression cassette as well as to the linear vector. This will generate a circular structure.

Ligation method B - NEBuilder Hifi DNA Assembly Cloning Kit

Setting up a PCR reaction with the gRNA120a expression cassette, started ligation method B. The reaction was as follows:

18 μ l mastermix + 0.5 μ l 10 μ M NEB_U6_F1 type 1 + 0.5 μ l 10 μ M NEB_U6_R2 type 1 + 1 μ l gRNA120a

The PCR product was electrophoresed on 1% EtBr agarose gel to check the product length before moving on to next step.

Next step was a fast digest restriction enzyme reaction for vector pFGC-pcoCas9 (MaxiPrep sample with a concentration of 2 µg/µl). Enzymes used in the reaction were *EcoRI* and *XbaI* (Life Technologies). The reaction proceeded during 20 minutes before it was heat inactivated. The sample was in turn electrophoresed on 1% EtBr agarose gel to check that the reaction had generated a clear band indicating a successfully digestion. The sample was in turn gel-purified (Life Technologies).

Following the gel-purification, the NEBuilder Hifi DNA cloning was performed. The assembly protocol was followed for setting up the reaction. Products used in the reaction, apart from the assembly mix, were 2µl of digested vector pFGC-pcoCas9 as well as 0,5µl gRNA120a (product from additional PCR with NEB primers). The assembly reaction proceeded for 60 minutes at 50°C.

Ligation method B to insert five gRNA expression cassettes in a single vector

The NEBuilder Hifi DNA Assembly Cloning Kit can be used for simultaneous insertion of several sequences into a linear vector as previously mentioned. Primers to make such reaction have been created but no trial has been made. However, following can be done in future to insert five gRNA expression cassettes into vector pFGC-pcoCas9:

1. Set up a fast digest reaction for vector pFGC-pcoCas9 with restriction enzymes *EcoRI* and *XbaI*.
2. NEB-PCR reaction 1: Set up a PCR reaction with the gRNA expression cassettes and primers listed in NEB-PCR reaction 1 in table 3.
3. NEB-PCR reaction 2: Use PCR products from NEB-PCR reaction 1 and make additional PCR reactions with primers listed in NEB-PCR reaction 2 in table 3.
4. Following the gel-purification of vector pFGC-pcoCas9, the NEBuilder Hifi DNA cloning can be performed. Use 2µl of digested vector pFGC-pcoCas9 as well as 0.5 µl of each product from NEB-PCR reaction 2 for an assembly reaction. The assembly reaction should proceeded for 60 minutes at 50°C.

Table 3. Additional PCR reactions for assembled gRNA expression cassettes in order to work with NEBuilder Hifi DNA cloning.

gRNA expression cassette for associated genes:	Primers used in NEB-PCR reaction 1	Primers used in NEB-PCR reaction 2
<i>A120</i>	NEB_U6_F1 type 1 U6_R2_NE_B_A	NEB_U6_F1 type 1 gRNA_A_rev
<i>A130</i>	U6_F1 NEB_B U6_R2_NE_B_B	gRNA_B_fwd gRNA_B_rev
<i>A132</i>	U6_F1 NEB_C U6_R2_NE_B_C	gRNA_C_fwd gRNA_C_rev
<i>A135</i>	U6_F1 NEB_D U6_R2_NE_B_D	gRNA_D_fwd gRNA_D_rev
<i>A150</i>	U6_F1 NEB_E NEB_U6_R2 type 1	gRNA_E_fwd NEB_U6_R2 type 1

Transformation of ligated products into chemically competent *E. coli*

Samples collected from ligation reactions in method A and B were transformed into chemically competent *E. coli* (DH5 α). Following protocol was used for the transformation:

1. Thaw DH5 α cells on ice.
2. Incubate for 10 minutes.
3. Add 2 μ l sample (ligation product) to 40 μ l DH5 α cells.
4. Incubate on ice for 30 minutes.
5. Heat shock for 30 seconds at 42°C.
6. Incubate on ice for 2 minutes.
7. Add 250 μ l SOC medium.
8. Incubate at shaker for 60 minutes at 37°C.
9. Spread sample on pre-warmed LB-plates with correct selection medium.
10. Incubate overnight at 37°C.

A positive control with pUC19 was made as well as a control with direct transformation of undigested vector pFGC-pcoCas9.

DNA Analysis

Colony PCR for transformed *E. coli*

Eleven colonies of transformed *E. coli* were re-streaked on a LB_{kanamycin}-plate and incubated overnight at 37°C. Following incubation, a part of each of the colonies were resuspended in 20 μ l dH₂O, microwave treated for 5 seconds and used in PCR reactions to amplify the genomic region of inserted gRNA120a expression cassettes. Eleven PCR reactions were set up as follows:

18 μ l mastermix + 0.5 μ l NEB_U6_F1 type 1 + 0.5 μ l NEB_U6_R2 type 1 + 1 μ l resuspended bacteria

The PCR was run according to protocol in figure 9 below.

98°C	98°C	62°C	72°C	72°C	12°C
30''	10''	15''	30''	7'	∞
× 35					

Figure 9. Illustration of conditions in colony PCR. The figure shows conditions for initial denaturation, the 35 cycles of denaturation, annealing and extension as well as the final extension. Single apostrophe (') represent minutes and double apostrophes (') represent seconds.

PCR samples were in turn electrophoresed on 1% EtBr agarose gel to see product length.

Sequencing of colony PCR products

Remaining bacteria from two colonies that gave rise to clear bands when its associated colony PCR product were electrophoresed on agarose gel, were prepared for sequencing by making overnight cultures (37°C) in LB-medium. MiniPreps (Life Technologies) of these overnight cultures were made and sent to GATC Biotech for sequencing.

Results

Designed PCR primers

All primers designed within this project were ordered from Invitrogen (Life Technologies). Names and sequences for all primers are listed in appendix A table 1.

Assembly of gRNA

PCR products showed straight and clear bands when electrophoresed on agarose gel, figure 10. Obtained product concentrations of assembled gRNA expression cassettes varied between 19.6 and 93.8 ng/ μ l.

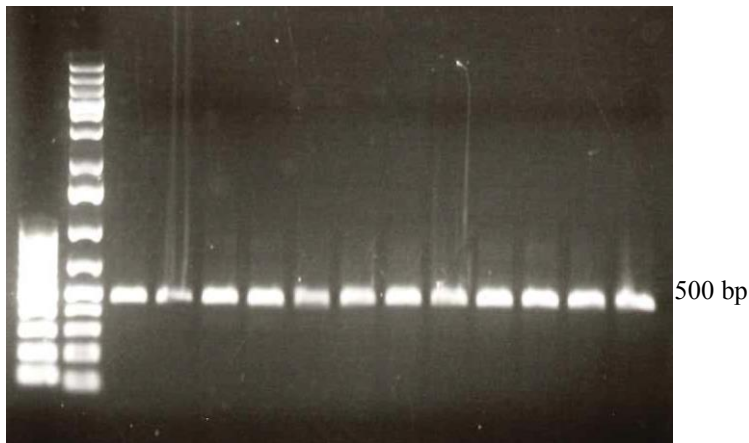


Figure 10. Assembled sgRNA expression cassettes when viewed under UV-light (after being electrophoresed on agarose gel). From left to right: 100 bp DNA ladder and 1 kb DNA. These two ladders are followed by two gRNA expression cassettes for each of the genes *A120* (b), *A130* (b), *A132* (b), *A140*, *A150* and *A160*.

Ligation and transformation

Products that were gel-purified and used in the ligation method A can be seen in figure 11. Method A and transformation of associated ligation product into *E. coli* did not result in colony growth. Colonies were seen for the positive control (pUC19) transformed at same occasion. Ligation method B on the other hand and transformation of this ligation product (vector for Cas9 expression ligated with the gRNA120a expression cassette), resulted in colony growth. Digested vector sample that was used in this method can be seen in figure 12. Colony PCR were made on an *E. coli* colony generated from ligation method B.

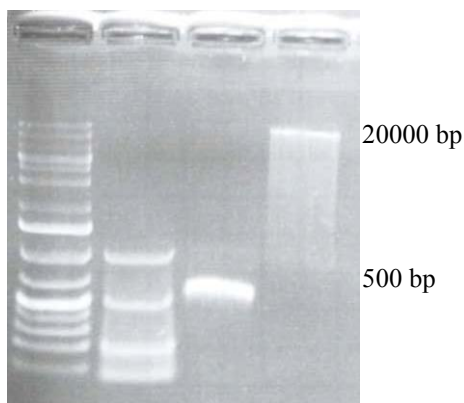


Figure 11. Digestion products from enzyme reactions with *EcoRI* and *XbaI* when viewed under UV-light (after being electrophoresed on agarose gel). From left to right: 1 kb DNA ladder, 50 bp DNA ladder, digested sample of gRNA120a expression cassette and digested vector pFGC-pcoCas9.



Figure 12. Undigested and digested plasmid Maxiprep samples of vector pFGC-pcoCas9 when viewed under UV-light (after being electrophoresed on agarose gel). From left to right: 1 kb DNA ladder, digested sample of Maxiprep *i.e.* vector pFGC-pcoCas9 cleaved with *EcoRI* and *XbaI* and an undigested vector sample.

Sequencing

The sequencing of the plasmid in the *E. coli* colony is shown in figures 13a-d on following pages, named GATC_120a.seq. Figure 13a shows an overview of the whole region that was sequenced. The forward and reverse primers used to ligate the gRNA120a expression cassette into vector pFGC-pcoCas9, NEB_U6_F1 type 1 and NEB_U6_R2 type 1, are illustrated in the same figure. The CRISPR site for gene *A120* can also be seen. Further, vector pUC119-gRNA containing the template gRNA is also illustrated in the figure as well as an extraction of vector pFGC-pcoCas9.

Figure 13b-d shows enlargements of three parts of figure 13a. Starting with figure 13b, this figure illustrates the CRISPR site found in Geneous for *A120* and following PAM-sequence. The complete sequence is in the desired form of 5' -G N₁₉ NGG- 3', seen in the figure by GCCGGGGAAGGCCCTGGATTTGG. Moreover in figure 13b, the vector sequence for pUC119-gRNA (unspecific gRNA) can be seen as well as part of the sequenced plasmid (GATC_120a.seq). The alignment of the three sequences shows that the sequenced *E. coli* colony has a sequence that is a perfect match of the desired CRISPR site rather than a sequence that matches the unspecific gRNA within the vector.

Figure 13c illustrates the binding region of the forward primer. In the figure, it can be seen that the forward primer sequence matches the plasmid sequence as well as the sequence for vector pUC119-gRNA at 14 bases upstream in the primer direction (right). In figure 13d on the other hand, the binding region of the reverse primer is illustrated. In this figure, a sequence extraction for vector pFGC-pcoCas9 is inserted. Looking at this sequence and also the sequence for vector pUC119-gRNA, there is a perfect match of three bases in position 848-850. Before this point, the *E. coli* sequence matches the vector pUC119-sequence. After this point, there is a match between the *E. coli* sequence and the extraction for the pFGC-pcoCas9 vector. Position 848-850 is located within the binding region of the reverse primer, indicating that the reverse primer ligates the ends of the gRNA120a expression cassette and vector pFGC-pcoCas9.

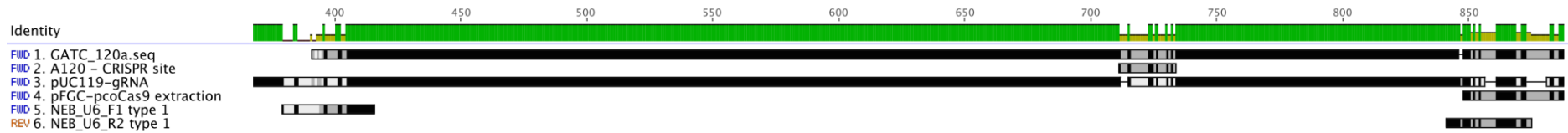


Figure 13a. An overview of a sequenced region within *E. coli*, in which vector pFGC-pcoCas9 including a gRNA120 expression cassette has been inserted. The forward and reverse primers used to ligate the gRNA120a expression cassette into vector pFGC-pcoCas9, NEB_U6_F1 type 1 and NEB_U6_R2 type 1 are shown as well as the CRISPR site for gene *A120*. Vector pUC119-gRNA containing the template gRNA and an extraction of vector pFGC-pcoCas9 can also be seen in the alignment.

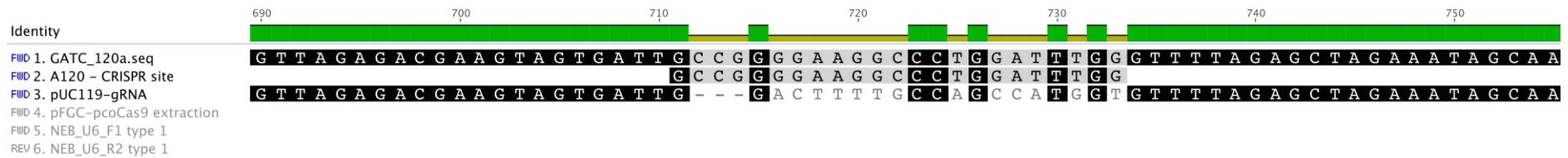


Figure 13b. Enlargement of figure 13a. The CRISPR site found in Geneous for gene *A120* and following PAM-sequence is illustrated as A120 - CRISPR site. A sequenced region within *E. coli*, in which vector pFGC-pcoCas9 including a gRNA120 expression cassette has been inserted is shown as well as the pUC119-vector sequence.

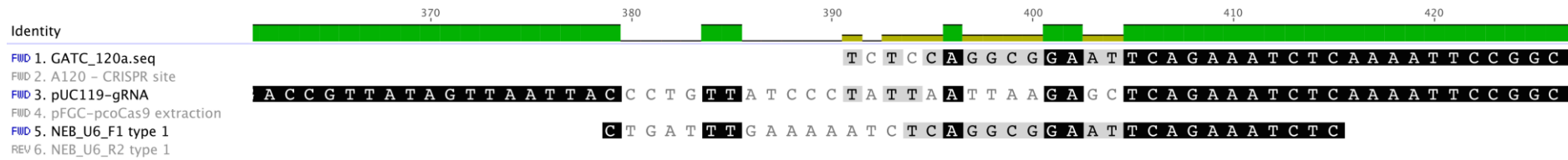


Figure 13c. Enlargement of figure 13a. The binding region of the forward primer is illustrated as well as the sequenced *E. coli*. pUC119-gRNA vector sequence is also illustrated.

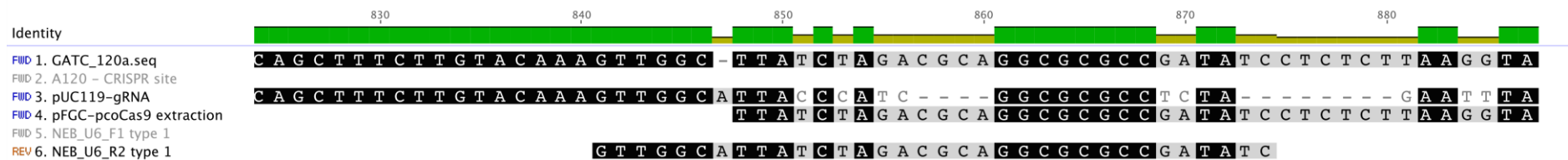


Figure 13d. Enlargement of figure 13a. The binding region of the reverse primer is illustrated as well as the sequenced *E. coli*. pUC119-gRNA vector sequence is also illustrated. Moreover, a sequence extraction for vector pFGC-pcoCas9 can be seen.

Discussion

The aim of this laboratory work was to transform a vector that express Cas9 for plants as well as a gRNA sequence targeting gene *AI20* into a chemically competent strain of *E. coli* cells. This aim has been achieved as the GATC sequencing result of a transformed *E. coli* colony matches the sequence of the CRISPR site for *AI20*, figure 13b. Another aim was to design primers for additional five previously mentioned Arabidopsis genes (figure 4). These primers should be possible to use for assembling and ligation of both single and multiple gRNA expression cassettes into vector pFGC-pcoCas9. This has been done as well, appendix A table 1. However, more can be said and discussed about the method and the obtained results.

Starting with the designed primers to work with the PCR method used to assemble gRNA expression cassettes, all primers generated gRNA expression cassettes despite variations in concentration. The generated products had predicted sizes when checked on agarose gel (figure 10) indicating that the generated products are complete gRNA expression cassettes. It can thereby be concluded that these primers were correctly designed and that the PCR method published by Li *et al.* (2013) to assemble gRNA expression cassettes has a high usability and generates desired products.

Comparing the ligation methods, it appears that method B works while method A does not. This statement assumes that the absence of *E. coli* colonies on LB-plates after transformation of ligation products from method A is a result of a failure within the ligation method itself rather than any other factor. This is further supported by the fact that colony growth was seen for the positive control in the same transformation trials. However, a major difference between ligation method A and B that cannot be neglected is the concentration of the plasmid (pFGC-pcoCas9) samples used in the ligation reactions.

In method A, ligation reactions were set up with Mimiprep plasmid samples with a concentration of 122.4 ng/ μ l. In method B on the contrary, a Maxiprep plasmid sample with a concentration of 2 μ g/ μ l was used. This concentration difference can also be seen when comparing digested vector samples in figure 11 and 12. The brighter band in figure 12 indicates that there is a much higher concentration of digested vector compared to the lighter band in figure 11. The concentration difference was not compensated by an equivalent difference in sample volumes used in the reactions, whereby the plasmid concentration was higher in method B. A low plasmid concentration may therefore be the reason why ligation method A did not work while method B did. Moreover, factors such as number of bases overlapping in the ligation reaction or details such as exclusion of heat inactivation of the enzyme digestion reaction in method A may explain why this method did not work.

Leaving the reasoning of why ligation method A did not work and instead focusing on the method B that did work, this method successfully generated colony growth of *E. coli* after transformation. Repeats of the same method but with other gRNA expression cassettes targeting different genes are thereby expected to generate colony growth as well. There is although a risk when working with ligation that fragments expected to be ligated together, in fact, are not. In this case, the gel purification of the digested vector allowed for removal of undigested vector. However, the risk remained that a fragment of undigested vector may have been kept, transformed into *E. coli* and

causing colony growth. Making colony PCR and sample sequencing was thereby required in order to show that the sample transformed into *E. coli* was a pFGC-pcoCas9 vector having inserted gRNA120a expression cassette rather than being an undigested vector sample. Figure 13b proves, as previously mentioned, that the plasmid carrying the gRNA120a expression cassette was successfully integrated in *E. coli*.

The sequencing of the plasmid in the *E. coli* colony does further, apart from showing that gRNA120a has been successfully integrated, also show that the ligation of pFGC-pcoCas9 and the gRNA120a sequence occurs within the binding region of the primers, figure 13d. This was expected as primer NEB_U6_R2 type 1 was designed to have a sequence matching both the flanking region of the gRNA expression cassette and the vector. However, it is always good to prove it. There is no extraction of the Cas9-sequence in figure 13c, making it harder to prove that the GATC sequence is a ligation result from the gRNA expression cassette and the vector for Cas9 plant expression. Although, seeing in figure 13c that the primer sequence matches the gRNA expression cassette sequence at a region where the gRNA-vector sequence does not match, generates an acceptable proof of a successful ligation within the binding region of the primer. Especially since it is known that the forward primer, just as the reverse primer, is designed to have a sequence matching the flanking region of the gRNA expression cassette as well as the vector for Cas9.

Returning to the topic of primer design, nothing has been discussed when it comes to primers created to insert multiple gRNA expression cassettes into a single vector. These primers can only be assumed to work as no trials have been made to prove it. As ligation method A did not work for insertion of a single gRNA expression cassette, this method will neither work for insertion of multiple gRNA expression cassettes. The primers may therefore work theoretically, but most likely not in reality. As ligation method B with the NEB cloning kit worked, this method should preferably be used in future experiments aiming to make insertions of multiple gRNA sequences into vector pFGC-pcoCas9. A conclusion of whether or not designed primers work can be reached after such trial.

Transformation of vector pFGC-pcoCas9 with the inserted gRNA120a expression cassette into *Agrobacterium* is the next step to continue the work that has been started in this project. Once having growing colonies of *Agrobacterium*, these will be transformed into Arabidopsis by using the floral dip method (Clough & Bent, 1998). The gRNA will bind to the target site in *A120* and induce a DSB. As the cell repairs the broken DNA strands by NHEJ (repair by HR is unlikely to occur as no DNA template is inserted), there will be deletion of nucleotide(s). This will generate a mutation within the plant's genome that is heritable. Moreover, making the same transformation into Arabidopsis for a vector that contains multiple gRNA expression cassettes is expected to induce multiple DSBs and thereby generate progenies that are complete infertile.

Continuing the started work in this project will not only allow for study of genes involved in stamen development of Arabidopsis. Working with the CRISPR Cas9 method will make it possible to investigate aspects such as its mutation efficiency and off-target frequency; aspects that preferably should be further investigated to better understand and discuss pros and cons with the CRISPR Cas9 method. It would also be

valuable to make experiments to *e.g.* investigate whether CRISPR Cas9-induced mutations can be tissue specific and make a gene expressed in one part of a plant while it is silent in another part. Published research of such experiments has not been found within the literature study of this thesis, but is highly interesting when it comes to usability of the CRISPR Cas9 technique.

Conclusion

The aim of Part I is achieved as the modified sequence and the vector pFGC-pcoCas9 was verified in *E. coli*. Primers created and the method used to assemble a gRNA expression cassettes works. The primers and the method can thereby be used in future experiments. Following assembly of gRNA, ligation method B should be used to ligate an assembled gRNA expression cassette into vector pFGC-pcoCas9. Making trials with this ligation method and associated primers are recommended to generate vector collections that contain both single and multiple gRNA expression cassettes.

The work that has been started in this project proves that it is easy to design sgRNAs. The CRISPR Cas9 method can hence be described as a method with great potential that is easy to use, supposing that the continued work of *Agrobacterium* transformation into Arabidopsis is not more complicated than any similar transformation with another vector. Using CRISPR Cas9 may in the small perspective of this project allow for a deeper understanding of genes involved in stamen development of plants. In a greater perspective, knowledge about how CRISPR Cas9 can be used to generate plants of Arabidopsis that have reduced male fertility can be applied to crop plants. Findings from this work can thereby be used to enhance hybrid seed production and improve crop yield and quality.

Part II

Background

The EU legislation for genetically modified organisms (GMOs)

There is a comprehensive legal framework in place for authorisation, traceability and labelling of genetically modified organisms (GMOs) within the European union; Regulation (EC) No 1829/2003 of the European Parliament and of the Council of 22 September 2003 on genetically modified food and feed (European Parliament, 2003a) and Directive 2001/18/EC of the European Parliament and of the Council of 12 March 2001 on the deliberate release into the environment of genetically modified organisms and repealing Council Directive 90/220/EEC (European Parliament, 2001).

For the purpose of Directive 2001/18/EC, following definition is used to describe a genetically modified organism (GMO):

“Genetically modified organism (GMO) means an organism, with the exception of human beings, in which the genetic material has been altered in a way that does not occur naturally by mating and/or natural recombination”.

In the Annex I of Directive 2001/18/EG, techniques that give rise to GMOs, techniques that are not considered to result in GMOs and techniques of genetic modification that are excluded from the Directive can be found. Techniques that are considered to generate GMOs are recombinant nucleic acid techniques as well as techniques that involves introduction of heritable material that are created outside the organisms. Examples of such methods and materials are micro-injections, macro-injections and micro-encapsulation. Cell fusion or hybridisation techniques are also considered to be techniques that generate GMOs, as there is formation of new heritable genetic material that does not occur naturally.

When it comes to techniques that are not considered to give GMOs, three methods are listed in Directive 2001/18/EG; in vitro fertilisation, natural processes *e.g.* conjugation, transduction and transformation and finally polyploidy induction. A requirement is although that these three listed techniques do not involve use of recombinant nucleic acid molecules or any GMOs made from techniques that are not excluded from the Directive. Two techniques that are listed to give GMOs but excluded from Directive 2001/18/EG are mutagenesis and cell fusion of plants cells.

Cultivation of GMO within the EU territory

A GMO intended for placing on the market within the EU or to be cultivation within the EU territory, has to undergo an individual risk assessment in accordance with Regulation (EC) No 1829/2003 and Annex II in Directive 2001/18/EG before it can be authorised. The risk assessment includes evaluation on effects on human health and the environment with the aim to ensure a high level of protection.

The scientific risk assessment of GMOs is carried out by European Food Safety Authority (EFSA) in cooperation with the Member States' scientific bodies (European Commission, 2015). Based on conclusions stated by EFSA regarding potential effects on human and animal health and the environment of a particular GMO, the European

Commission will make a draft decision for GMO authorisation. Member States of the European Union vote on this draft decision within the Standing Committee. A qualified majority is needed for a decision to be reached. If there is no qualified majority for or against the draft decision at this stage, the draft will be tabled for an additional voting in the Appeal Committee. If there is no qualified majority for or against the draft decision at this stage either, the Commission has to take a decision on authorisation without support of Member States in relevant committees.

When it comes to authorisation of GMOs, there is rarely any qualified majority reached within the Standing Committee or the Appeal Committee; most arguments opposing the authorisation reflect national concerns rather than scientific considerations (European Commission, 2015). The European Council did thereby come to the conclusion that improvements were needed in the legal framework of GMOs (Council of the European Union, 2008). Directive (EU) 2015/412 has for this purpose recently amended article 26 in Directive 2001/18/EG. Directive (EU) 2015/412 makes it possible for Member States to restrict or prohibit the cultivation of GMOs in their territory despite these being authorised for cultivation and placing on the market at Union level.

Reasons stated by Member States' for restricting or preventing cultivating of an authorised GM crop should, according to Directive (EU) 2015/412, not affect the scientific risk assessment in the system of Union authorisation of GMOs. Member States can make their decisions on grounds with respect to environmental policies that are clearly distinct from the assessment of health- and environmental impacts evaluated in accordance with the authorisation procedure stated in Directive 2001/18/EC and in Regulation (EC) No 1829/2003. Member States can also make decisions on grounds of socioeconomic impacts as well as public policy and cultural traditions. However, details of if and how Directive (EU) 2015/412 can be interpreted in Swedish legislation for GMOs are unknown.

Detection and identification – the requirement for a policy

As there is legislation for what techniques that gives GMOs, detection methods to trace and identify GMOs has to be available. Detection and detection methods is for this purpose defined as a methods that allow determination of the existence of a change in the genetic material (Lusser *et al.*, 2011). This can be done by reference to an appropriate comparator. Identification of GMOs on the other hand, means that detected existence of a genetic change has to be identified as a genetic modification caused by a certain technique.

As mentioned in the introduction and in the section above, mutagenesis as a plant breeding method is excluded from the EU GMO legislation as there is no detection method available that identifies and distinguishes a genetic modification caused by mutagenesis from a spontaneous mutation. Detection methods allowing for identification of GMOs as such or in products have to be available at all stages of GMO placing on the market according to Regulation (EC) No 1829/2003 and Regulation (EC) No 1830/2003 (European Parliament, 2003b). Examples of such methods are DNA-based analysis, protein-based analysis and metabolite-based analysis (Lusser *et al.*, 2011). DNA-bases analysis, particularly DNA amplification-

based methods, is the most applicable for detection and identification of genetic modifications.

DNA amplification-based methods (PCR)

The most commonly used DNA amplification-based method to amplify specific DNA, is PCR technique (Lusser *et al.*, 2011; Miraglia *et al.*, 2004). To detect and identify whether or not there is GMO present in a sample, a conventional qualitative PCR method can be used (Lusser *et al.*, 2011). A multiplex PCR technique can also be used to amplify two or more loci simultaneously (Henegariu *et al.*, 1997). However, PCR primers have to be designed, whereby information on the modified sequence is required.

If a food sample is to be investigated for presence of GMO, a quantitative PCR method is most commonly used (Buh Gašparič *et al.*, 2010; Miraglia *et al.*, 2004). In a quantitative real-time PCR, primers as well as fluorescent probes are used. Primers will bind to the inserted genomic material (GMO) and the probes will further emit fluorescent light that corresponds to the amount of synthesised DNA in the PCR (Miraglia *et al.*, 2004). The amount of GMO in the sample can thereby be measured, which is relevant as Regulation (EC) No 1830/2003 states that GMO products containing more than 0,9 % approved GMO product has to be labelled at all stages of placing on the market. Products containing less than 0,9 % approved GMO products do not need to be GMO-labelled.

Detecting and identifying GMOs may be possible when having information on the modified sequence. This kind of information can for instance be adapted from available GMO Detection Databases (Dong *et al.*, 2008), but it inevitably evokes a question of whether or not it is possible to detect GMO without this kind of prior information. Tengs *et al.*, (2007) has published a microarray-based method for detection of unknown genetic modifications. Fragments equal or greater than 140 bp can be detected without prior knowledge of the transgene sequence. A differential PCR may also be used to detect unauthorised GMO (Cankar *et al.*, 2008). This method has been found to work when searching for donor organisms of the Cauliflower 35S promoter.

New plant breeding techniques

In 2007, a working group was established by the European Commission to evaluate whether eight new techniques of genetic modification, seven of which can be used to alter plant DNA, should fall within the EU GMO legislation (Lusser *et al.*, 2011). All seven techniques are listed in table 4 below together with short summaries of intended changes/effects when using the techniques and possibilities for detection. The new technique of CRISPR Cas9 is also listed in the table.

Table 4. Short summary of intended changes/effects and possibilities for detection of seven new plant breeding techniques. Information in the table is adapted from (Lusser *et al.*, 2011). Intended changes/effects and possibilities for detection is also given for the CRISPR Cas9 method.

Name	Intended changes/effects	Possibilities for detection and identification
Zink finger nuclease (ZFN) technology	ZFN-1: Deletion of one or few nucleotides as a result of NHEJ repair of an induced DSB. No repair template is provided to the cell. ZFN-2: Induced mutation through HR repaired of a DSB. Template DNA homologous to the targeted area, except for the desired base pair alteration that should be induced, is delivered to the cell with the ZFN. ZFN-3: A constructed recombinant DNA molecule (transgene) is delivered to the cell with the ZFN. Transgene integration by HR repair of DSB.	ZFN-1 and ZFN-2: Mutations can be detected but not distinguished from natural mutations or mutations caused by other mutation techniques. ZFN-3: Mutations can be detected with PCR. Identification requires prior knowledge.
Oligonucleotide directed mutagenesis (ODM)	Site-specific nucleotide substitution. Oligonucleotides are homologous to target sequence with the exception for nucleotide(s) to be modified. These are inserted to a cell and target the homologous sequence. Substitution occurs as the cell repairs the mismatches.	Mutations cannot be detected unless there is information provided on nucleotides close to the mutation-site. Identification is not possible.
Cisgenesis and intragenesis	Cisgenesis: DNA-fragment from a plant species (from itself or from a cross-compatible plant species) is inserted. Inserted gene includes introns and regulatory sequences <i>i.e.</i> the gene is unchanged. Intragenesis: DNA-fragment from a plant species (from itself or from a cross-compatible plant species) is inserted. The inserted DNA is a new combination of DNA fragments.	Detection and identification is possible if information regarding mutation site is known (PCR). Searching for mutations without information provided (<i>e.g.</i> through genome sequencing) has not been proved to both detect and identify technique-induced mutations.
RNA-dependent DNA methylation (RdDM)	Genes encoding RNAs, homologous to plant sequences, are inserted in plant cells. These genes generate small double stranded RNAs that induce methylation and thereby inhibit transcription of the homologous sequence.	Detection and identification is not possible no matter if prior information given or not.
Grafting (on GM rootstock)	A chimeric organism is created through attachment of vegetative component (non-GM) above ground to a rootstock of another plant (GM).	Detection and identification is not possible as the DNA sequence of the non-GM scion is not modified.
Reverse breeding	A homozygous parental line is generated from double-haploid (DH) plants. This requires suppression of meiotic crossovers as well as fixation of non-recombinant chromosomes in DH.	Detection and identification is not possible as the product cannot be distinguished from a product produced from conventional breeding.

Agro-infiltration (agro-infiltration “sensu stricto”, agro inoculation, floral dip)	Agro-infiltration “sensu stricto” and agro inoculation: Temporary expression of specific coding sequence. This is done without integration of introduced DNA in the plant genome. Floral dip: Seedlings are stably transformed. No need for plant regeneration phase. Properties of a transgenic plant are obtained.	Agro-infiltration: Detection and identification is only possible in agro-infiltrated plants. Detection and identification is not possible in the products of these plants unless the inserted fragment is stable. Floral dip: Detection and identification is possible through PCR. Unless no information is available, identification is not possible.
CRISPR (clustered regularly interspaced short palindromic repeat) Cas9	Site-specific mutation(s) are induced in the NHEJ repair process of an endogenous DSB. Deletions of	CRISPR Cas9 will cause similar genetic mutations as ZFN-1 and ZFN-2. Mutations can be detected but not distinguished from natural mutations or mutations caused by other mutation techniques.

ZFN-3, cisgenesis and intragenesis as well as agro-infiltration (floral dip) are the only methods out of all seven plant breeding techniques where it is possible to detect and identify the products (Lusser *et al.*, 2011). This does however require prior information about the introduced modification.

Field trials with CRISPR Cas9-generated plant mutants in Sweden

Swedish legislation for GMOs is stated in Miljöbalken (MB) 1998:808 (Regeringskansliet/Lagrummet, 2015b). When it comes to field trials with plants bred with techniques that are classified as GM-techniques (Directive 2001/18/EG), permission has to be received before such trials can be performed according to MB 13:12 and the Swedish constitution (FU) 2002:1086 (Regeringskansliet/Lagrummet, 2015a). Application for field trials should be sent to the Swedish Board of Agriculture. Information to be included in the application is regulated by FU 2:2-6. However, FU 4 § p. 2 states that crops bred with mutagenesis or cell fusion *i.e.* techniques listed as excluded from legislation within Directive 2001/18/EG are not regulated by MB 13.

As mentioned in the introduction of this report, the EU is still undecided as to whether all, some or any of the new plant breeding techniques should be subject to the traditional GMO legislation. Considering this in the context of field trials with GMO in Sweden, no information is available on whether or not an application has to be sent in and approved before field trials with plants bred with any of these new techniques can be performed. Relating this to Part I of this report, it remains unknown what rules to follow for future field trials with the Arabidopsis-plants having site-specific mutations induced by CRISPR Cas9. In order to seek answer, a formal letter has been sent to the Swedish Board of Agriculture. The letter can be found in Appendix B. No English translation is available, but the questions asked are:

1. Are plants, in which mutations are induced by use of CRISPR Cas9, subjected to the EU GMO legislation?
2. Are there any authorisation requirements for making field trials with these plants?

3. Are there any regulations to be followed when making the field trials? If such are missing, is it still possible to perform field trials?

The letter was sent to the Board of Agriculture in June 2015. In the same period of time, a letter from the European Commission was sent to all competent authorities for Directive 2001/18/EC. In this letter, it is written that some Member States within the EU has given permission for field trials with oilseed rape produced with ODM *i.e.* one of the new plant breeding techniques that is being legally evaluated. The European Commission is further asking the Member States to await a legal interpretation before authorising any products produced with the new plant breeding techniques. Finally, it is written that a legal analysis will be presented before the end of 2015. The letter sent from the European Commission can be found in appendix D.

Results and Discussion

Regulation (EC) No 1829/2003 and Directive 2001/18/EG make a legal framework for GMOs within the EU and regulates cultivation of crops bred with techniques that alter the genetic material in a way that does not occur naturally. Comparing a crop regarded as a GMO to a crop being conventionally bred, there is an induced genetic modification within the genome of the GMO that can be detected, identified and distinguished from a spontaneous mutation. In crops conventionally bred, induced mutations are indistinguishable from such mutations.

The reason for having legislation for GMOs is to protect humans, animals and the environment from potential cumulative long-term effects that may arise after an intended release of GMO on the market. An example of such long-term effect could be an uncontrolled reproduction and survival of the genetically modified organism that may disturb the natural ecosystem. The legislation does also serve the purpose of taking ethical aspects into consideration before releasing GMOs.

Since Directive 2001/18/EG was agreed on, new plant breeding techniques have evolved. Despite the European Commissions' establishment of a working group to evaluate whether these new techniques should fall within the scope of the EU GMO legislation, no recommendation has been issued. The legal outcome of these new plant breeding techniques is therefore unknown. One reason behind this unsatisfying situation is the fact that the new plant breeding techniques generate products that cannot be distinguished from spontaneous mutations or mutations induced by any of the two techniques listed in Annex I B as excluded from Directive 2001/18/EG. Only ZFN-3, cisgenesis, intragenesis and agro-infiltration (floral dip) can be detected and identified. The new plant breeding techniques seem to fall within the EU legislation in that they involve modification of genetic material in a way that does not naturally occur, but they offer problems for the legislators in that the changes to the genome created are untraceable. If crops bred using these techniques should be classified as GMOs, it therefore follows that the current legislation requires considerable amendment and reworking.

As previously mentioned, the intention of the law is to protect humans, animals and the environment from unknown effects of GMOs. This aim is further supported by the Directive (EU) 2015/412 that allows Member States of the European Union to restrict

or prohibit the cultivation of GMO crops with regards to environmental factors. Subjecting the new plant breeding techniques to the EU GMO legislation puts the legislation into action. The new plant breeding techniques, as previously mentioned, induce mutations similar to mutations that naturally occur in all living organisms; subsequently plants produced with the new techniques cannot in a convincing way be argued to have any more unknown effects than other wild-type plants. Subjecting the new plant breeding techniques to the current EU GMO legislation means that an organism classified as GMO is in theory identical to a wild-type organism - where mutations have occurred through evolution. Consequently the GMO cannot be argued to have a higher competitiveness than the wild-type organism and as such the GMO in turn cannot be said to disturb the ecosystem more than any other organism.

It is my belief that there is a risk of plant breeding innovations being prevented if the new techniques including CRISPR Cas9 are placed under the scope of the EU GMO legislation. Applications for field trials for GMOs are expensive, especially as new cultivars cannot be used within the EU territory, which reduces the number of trials performed. I thereby propose that the European Commission, if the new plant breeding techniques are to be included in Directive 2001/18/EG, list these in Annex I B. Although, another important aspect that deserves attention in this context, is whether there is a need of the whole legislation *i.e.* Directive 2001/18/EG to be rewritten. Considering there is a continuous development of new techniques in plant breeding; the CRISPR Cas9 method is merely one example of a technique that has evolved since the Commission started working with the area in 2007, there would be a potential gain in making the legislation product-based rather than based on techniques. GMOs could, for this purpose, be defined by all traits displayed after induced genetic modification rather than by technique used.

Having a product-based GMO law would, in my opinion, to a greater extent make sense in context of the aim of the legislation. Release of organisms that are genetically modified to such extent that they have negative effects on human and animal health should be prevented by legislation. These organisms would also be prevented by legislation if they can reproduce out of human control and express a higher competitiveness in environments that they are not intended for, as this would risk disturbing the natural ecosystem. Organisms that are not a risk for human and animal health and/or do not have a beneficial competitiveness, should be seen as similar to wild type organisms and thereby not be prevented by the law. This decision should be made with regards to the organism itself and its features rather than by technique used to produce it. A product-based legislation would also make it easier to handle new techniques for genome editing that may evolve in the future. No amendments of the law are needed as the techniques themselves are not to be subjected to it.

Despite my analysis of the EU GMO legislation and how the new plant breeding techniques should be considered in its context, the question remains whether or not an application has to be sent in and approved before making field trials with Arabidopsis mutated by use of CRISPR Cas9. Considering the letter sent from the European Commission and its demands to Member States to await the legal outcome before approving any field trials with the new plant breeding techniques, it is understandable why I have not received any answer to the questions that I sent. The Swedish Board of Agriculture has however asked for complementary information. This is interpreted as a first step in the procedure of preparing a reply. The letter with complementary

information can be found in appendix C. Meanwhile, it is clear that the delay in decisions on this matter by European policy makers not only hinders applied research and technology development but also directly influence curiosity driven- or basic science in a negative manner.

Conclusion

Techniques that induce genetic modifications of organisms in a way that do not occur naturally by mating and/or natural recombination are listed in Directive 2001/18/EG. Products derived from these techniques, GMOs, are subjected to Regulation (EC) No 1829/2003. The EU GMO legislation can thereby be described to regulate GMO products on the basis of what techniques that have been used to generate these rather than considering the genetic modification induces as such. In turn, having a legislation being based on techniques requires considerable amendment and reworking for any new technique that may evolve. The EU GMO legislation is thereby proposed a change so as to regulate GMOs on basis of all traits displayed after induced genetic modification rather than by the technique used.

Making the EU GMO legislation product-based would make it easier to decide whether new plant techniques should fall under the scope of the legislation. This would, as an example, reduce the uncertainties regarding whether or not an application has to be approved before making field trials with plants being mutated by use of the CRISPR Cas9 method. Reducing uncertainties for the EU GMO legislation will likely enhance the development of new techniques and thereby contribute to great innovations with the area of plant breeding.

Acknowledgements

I would like to express my deepest gratitude to my supervisor Jens Sundström for his endless support and expert guidance throughout my project. Without his incredible patience and ingeniousness, this thesis would not have been possible. This thesis has not only been a meaningful journey that has supported my education of plant breeding techniques; it also marks the end of my studies within Agronomy. Through Jens contribution to this thesis as well as his engagement in a student teaching project that I have been involved in for the last few years, I have been privileged to work alongside the very best, in preparation for taking the first few steps in my career.

I would also like to extend a big thank you to all those that have supported me in this adventure including my lab colleagues, study tutors and peers.

Finally a very big thank you to my fiancé, Oskar Björling and to my family; with their unconditional love and support I have been encouraged each day to push myself that little bit further in this area of Agronomy that I am so very much fascinated by, for that I am so very thankful

References

- Addgene. (2015a). Addgene, pFGC-pcoCas9 (Plasmid #52256).
<https://www.addgene.org/52256/>.
- Addgene. (2015b). Addgene, pUC119-gRNA (Plasmid #52255).
<https://www.addgene.org/52255/>.
- Alves-Ferreira, M., Wellmer, F., Banhara, A., Kumar, V., Riechmann, J. L., & Meyerowitz, E. M. (2007). Global Expression Profiling Applied to the Analysis of Arabidopsis Stamen Development. *PLANT PHYSIOLOGY*, *145*(3), 747–762.
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., Horvath, P. (2007). CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science*, *315*, 1709–1712.
- Birchler, J. A., Auger, D. L., & Riddle, N. C. (2003). In search of the molecular basis of heterosis. *The Plant Cell*, *15*(10), 2236–2239.
- Bortesi, L., & Fischer, R. (2015). The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnology Advances*, *33*(1), 41–52.
- Brouns, S. J. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J. H., Snijders, A. P. L., Dickman, M. J., Makarova, K. S., Koonin, E. V., van der Oost, J. (2008). Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. *Science*, *321*, 960–964.
- Buh Gašparič, M., Tengs, T., La Paz, J. L., Holst-Jensen, A., Pla, M., Esteve, T., Žel, J., Gruden, K. (2010). Comparison of nine different real-time PCR chemistries for qualitative and quantitative applications in GMO detection. *Analytical and Bioanalytical Chemistry*, *396*(6), 2023–2029.
- Cankar, K., Chauvensy-Ancel, V., Fortabat, M.-N., Gruden, K., Kobilinsky, A., Žel, J., & Bertheau, Y. (2008). Detection of nonauthorized genetically modified organisms using differential quantitative polymerase chain reaction: application to 35S in maize. *Analytical Biochemistry*, *376*(2), 189–199.
- Carlsson, J., Leino, M., Sohlberg, J., Sundström, J. F., & Glimelius, K. (2008). Mitochondrial regulation of flower development. *Mitochondrion*, *8*(1), 74–86.
- Chen, F., Pruett-Miller, S. M., Huang, Y., Gjoka, M., Duda, K., Taunton, J., Collingwood, T. N., Frodin, M., Davis, G. D. (2011). High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nature Methods*, *8*(9), 753–755.
- Clough, S. J., & Bent, A. F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *The Plant Journal*, *16*(6), 735–743.
- Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., Zhang, F. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*, *339*(6121), 819–823.

- Council of the European Union. (2008). Council Conclusions on Genetically Modified Organisms (GMOs) 2912th Environment Council meeting, Brussels, 4 December 2008. *EN*.
- Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., Eckert, M. R., Vogel, J., Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, 471(7340), 602–607.
- Dong, W., Yang, L., Shen, K., Kim, B., Kleter, G. A., Marvin, H. J., Guo, R., Liang, W., Zhang, D. (2008). GMDD: a database of GMO detection methods. *BMC Bioinformatics*, 9(1), 260.
- European Commission. (2015). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. Reviewing the decision-making process on genetically modified organisms (GMOs). *COM(2015) 176 Final*.
- European Parliament. (2001). Directive 2001/18/EC of the European Parliament and of the Council of 12 March 2001 on the deliberate release into the environment of genetically modified organisms and repealing Council Directive 90/220/EEC. *Official Journal of the European Communities*, L106/1.
- European Parliament. (2003a). Regulation (EC) No 1829/2003 of the European Parliament and of the Council of 22 September 2003 on genetically modified food and feed. *Official Journal of the European Communities*, L268/1.
- European Parliament. (2003b). Regulation (EC) No 1830/2003 of the European Parliament and of the Council of 22 September 2003 concerning the traceability and labelling of genetically modified organisms and the traceability of food and feed products produced from genetically modified organisms and amending Directive 2001/18/EC. *Official Journal of the European Communities*, L268/24.
- Feng, Z., Zhang, B., Ding, W., Liu, X., Yang, D.-L., Wei, P., Cao, F., Zhu, S., Zhang, F., Mao, Y., Zhu, J.-K. (2013). Efficient genome editing in plants using a CRISPR/Cas system. *Cell Research*, 23(10), 1229–1232.
- Ferreira, M. G., & Cooper, J. P. (2004). Two modes of DNA double-strand break repair are reciprocally regulated through the fission yeast cell cycle. *Genes & Development*, 18(18), 2249–2254.
- Godde, J. S., & Bickerton, A. (2006). The Repetitive DNA Elements Called CRISPRs and Their Associated Genes: Evidence of Horizontal Transfer Among Prokaryotes. *Journal of Molecular Evolution*, 62(6), 718–729.
- Henegariu, O., Heerema, N. A., Dlouhy, S. R., Vance, G. H., & Vogt, P. H. (1997). Multiplex PCR: critical parameters and step-by-step protocol. *Biotechniques*, 23(3), 504–511.
- Horner, H. T., & Palmer, R. G. (1995). Mechanisms of Genic Male Sterility. *Crop Science*, 35, 1527–1535.

- Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., Peterson, R. T., Yeh, J.-R. J., Joung, J. K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nature Biotechnology*, *31*(3), 227–229.
- Jansen, R., Embden, J., Gaastra, W., Schouls, L., & others. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology*, *43*(6), 1565–1575.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, *337*, 816–821.
- Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., & Doudna, J. (2013). RNA-programmed genome editing in human cells. *Elife*, *2*, e00471.
- Kunin, V., Sorek, R., & Hugenholz, P. (2007). Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biol*, *8*(4), R61.
- Lehrman, A., Chatzopoulou, S., Feng, L., Forabosco, F., Jonas, E., Karantininis, K., Levander, F., Nicolia, A., Rydhmer, L., Röcklinsberg, H., Sandin, P., Sundström, J., Zhu, L.-H. (2014). *Shaping our food - an overview of crop and livestock breeding*. Uppsala: Mistra Biotech, SLU. Ebook.
- Li, J.-F., Norville, J. E., Aach, J., McCormack, M., Zhang, D., Bush, J., Church, G. M., Sheen, J. (2013). Multiplex and homologous recombination–mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nature Biotechnology*, *31*(8), 688–691.
- Li, T., Liu, B., Spalding, M. H., Weeks, D. P., & Yang, B. (2012). High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nature Biotechnology*, *30*(5), 390–392.
- Lusser, M., Parisi, C., Plan, D., Rodriguez-Cerezo, E., Institute for Health and Consumer Protection, & Institute for Prospective Technological Studies. (2011). *New plant breeding techniques state-of-the-art and prospects for commercial development*. Luxembourg: Publications Office. Retrieved from
- Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I., & Koonin, E. V. (2006). A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biology Direct*, *1*(1), 7.
- Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F. J. M., Wolf, Y. I., Yakunin, A. F., van der Oost, J., Koonin, E. V. (2011). Evolution and classification of the CRISPR–Cas systems. *Nature Reviews Microbiology*, *9*(6), 467–477.
- Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E., Church, G. M. (2013). RNA-Guided Human Genome Engineering via Cas9. *Science*, *339*(6121), 823–826.
- Maresca, M., Lin, V. G., Guo, N., & Yang, Y. (2013). Obligate Ligation-Gated Recombination (ObLiGaRe): Custom-designed nuclease-mediated targeted

- integration through nonhomologous end joining. *Genome Research*, 23(3), 539–546.
- Miraglia, M., Berdal, K. G., Brera, C., Corbisier, P., Holst-Jensen, A., Kok, E. J., Marvin, H. J. P., Schimmel, H., Rentsch, J., van Rie, J. P. P F., Zagon, J. (2004). Detection and traceability of genetically modified organisms in the food production chain. *Food and Chemical Toxicology*, 42(7), 1157–1180.
- Mojica, F. J. M., Diez-Villasenor, C., Garcia-Martinez, J., & Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*, 155(3), 733–740.
- Nekrasov, V., Staskawicz, B., Weigl, D., Jones, J. D. G., & Kamoun, S. (2013). Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. *Nature Biotechnology*, 31, 691–693.
- New England BioLabs. (2015). NEBuilder HiFi DNA Assembly Cloning Kit.
- Perez-Prat, E., & van Lookeren Campagne, M. M. (2002). Hybrid seed production and the challenge of propagating male-sterile plants. *Trends in Plant Science*, 7(5), 199–203.
- Pul, Ü., Wurm, R., Arslan, Z., Geißen, R., Hofmann, N., & Wagner, R. (2010). Identification and characterization of *E. coli* CRISPR-*cas* promoters and their silencing by H-NS. *Molecular Microbiology*, 75(6), 1495–1512.
- Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, 8(11), 2281–2308.
- Regeringskansliet/Lagrummet. (2015a). Förordning (2002:1086) om utsättning av genetiskt modifierade organismer i miljön. Ändrad t.o.m. SFS 2012:260.
- Regeringskansliet/Lagrummet. (2015b). Miljöbalk (1998:808). Ändrad t.o.m. SFS 2015:232.
- Rouet, P., Smih, F., & Jasin, M. (1994). Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Molecular and Cellular Biology*, 14(12), 8096–8106.
- Sapranaukas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., & Siksnys, V. (2011). The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Research*, 39(21), 9275–9282.
- Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., Zhang, K., Liu, J., Xi, J. J., Qiu, J.-l., Gao, C. (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. *Nature Biotechnology*, 31(8), 686–688.
- Smyth, D. R., Bowman, J. L., & Meyerowitz, E. M. (1990). Early flower development in *Arabidopsis*. *The Plant Cell*, 2(8), 755–767.
- Sorek, R., Lawrence, C. M., & Wiedenheft, B. (2013). CRISPR-Mediated Adaptive Immune Systems in Bacteria and Archaea. *Annual Review of Biochemistry*, 82(1), 237–266.
- Tengs, T., Kristoffersen, A. B., Berdal, K. G., Thorstensen, T., Butenko, M. A., Nesvold, H., & Holst-Jensen, A. (2007). Microarray-based method for detection of unknown genetic modifications. *BMC Biotechnology*, 7(1), 91.

- Van der Oost, J., Jore, M. M., Westra, E. R., Lundgren, M., & Brouns, S. J. J. (2009). CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends in Biochemical Sciences*, *34*(8), 401–407.
- Wittwer, S. H., & Hillyer, I. G. (1954). Chemical induction of male sterility in cucurbits. *Science*, *120*(3126), 893–894.
- Xie, K., Zhang, J., & Yang, Y. (2014). Genome-Wide Prediction of Highly Specific Guide RNA Spacers for CRISPR–Cas9-Mediated Genome Editing in Model Plants and Major Crops. *Molecular Plant*, *7*(5), 923–926.

Appendix A

Appendix Table 1. Primers used.

Oligo Name	Sequence (5' to 3')
gRNA_A_rev	GTGTCTCAGGGAGTACTAC
gRNA_B_fwd	AGTAGTACTCCCTGAGAC
gRNA_B_rev	GCCTCCTAAACTGTGAAG
gRNA_C_fwd	TTCTTCACAGTTTAGGAGG
gRNA_C_rev	GTCCAAATTTTTCCATGGC
gRNA_D_fwd	TGCCATGGAAAAATTTGGAC
gRNA_D_rev	GTCTCTTTTCCTTGCCCTAATAATG
gRNA_E_fwd	TTAGGGCAAGGAAAAGAG
NEB_U6_F1 type 1	CTGATTTGAAAAATCTCAGGCGGAATTCAGAAATCTC
NEB_U6_R2 type 1	GATATCGGCGCGCCTGCGTCTAGATAATGCCAAC
U6_F1 NEB_B	AGTAGTACTCCCTGAGACACAGAAATCTCAAAATTCCG
U6_F1 NEB_C	TTCTTCACAGTTTAGGAGGCAGAAATCTCAAAATTCCG
U6_F1 NEB_D	TGCCATGGAAAAATTTGGACAGAAATCTCAAAATTCCG
U6_F1 NEB_E	TTAGGGCAAGGAAAAGAGACAGAAATCTCAAAATTCCG
U6_F1 type1	GCGGAATTCAGAAATCTCAAAATTCCG
U6_F1 type2	GCGTCTAGAAGAAATCTCAAAATTCCG
U6_F2 120a	GCCGGGGAAGGCCCTGGATTTGGGTTTTAGAGCTAGAAATAGC
U6_F2 120b	GCTGAAGCATGCATCGTCAAGAGGGTTTTAGAGCTAGAAATAGC
U6_F2 130a	GGATACGGCCGTGGCCACAATGGGTTTTAGAGCTAGAAATAGC
U6_F2 130b	GACCAAGGATGATCGCTTCATGGGTTTTAGAGCTAGAAATAGC
U6_F2 132a	GATCACTTGGGAATAGGTAGTGGGTTTTAGAGCTAGAAATAGC
U6_F2 132b	GGGCCGGACCCAGATTCTGGCGGGTTTTAGAGCTAGAAATAGC
U6_F2 135	GTATTCCTACAGGAAGATTCGGGGTTTTAGAGCTAGAAATAGC
U6_F2 140	GACTTAGGAAAAGTTTTTGAAGGGTTTTAGAGCTAGAAATAGC
U6_F2 150	GGGAAGACAGATAGTACTCCCGGGTTTTAGAGCTAGAAATAGC
U6_F2 160	GGGGTCTACATTGTCTACATGGGGTTTTAGAGCTAGAAATAGC
U6_R1 120a	CCAAATCCAGGGCCTTCCCCGGCAATCACTACTTCGTCTCT
U6_R1 120b	CCTCTTGACGATGCATGCTTCAGCAATCACTACTTCGTCTCT
U6_R1 130a	CCATTGTGGCCACGGCCGTATCCAATCACTACTTCGTCTCT
U6_R1 130b	CCATGAAGCGATCATCCTTGGTCAATCACTACTTCGTCTCT
U6_R1 132a	CCACTACCTATTCCCAAGTGATCAATCACTACTTCGTCTCT
U6_R1 132b	CCGCCAGAATCTGGGTCCGGCCCAATCACTACTTCGTCTCT
U6_R1 135	CCCGAATCTTCTGTAGGAATACAATCACTACTTCGTCTCT
U6_R1 140	CCTTCAAAAATTTTTCTAAGTCAATCACTACTTCGTCTCT
U6_R1 150	CCGGGAGTACTATCTGTCTTCCAATCACTACTTCGTCTCT
U6_R1 160	CCCATGTAGACAATGTAGACCCCAATCACTACTTCGTCTCT
U6_R2 type1	GCGTCTAGATAATGCCAACTTTGTACA
U6_R2 type2	GCGGCGATCGCTAATGCCAACTTTGTACA
U6_R2_NEB_A	GTGTCTCAGGGAGTACTACTTAATGCCAACTTTGTACA
U6_R2_NEB_B	GCCTCCTAAACTGTGAAGAATAATGCCAACTTTGTACA
U6_R2_NEB_C	GTCCAAATTTTTCCATGGCATAATGCCAACTTTGTACA
U6_R2_NEB_D	GTCTCTTTTCCTTGCCCTAATAATGCCAACTTTGTACA



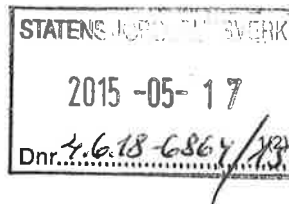
Appendix B

Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Department of Plant Biology
Associate Professor Jens Sundström

2015-06-09

Jordbruksverket



Förfrågan med anledning av planerade fältförsök med *Arabidopsis* som muterats med hjälp av CRISPR Cas9

Vi arbetar med CRISPR Cas9-metoden för att orsaka riktade mutationer i två lokus i *Arabidopsis* (At1g20120 och At1g20150). Genen i lokus At1g20120 kodar för ett lipas och genen i lokus At1g20150 kodar för ett peptidas. Bägge dessa proteiner antas vara viktiga under ståndarutvecklingen, varvid förväntad genotyp efter mutation är en nedsatt hanfertilitet. Vi vill i detta forskningssteg få bekräftat att de mutationer som induceras med hjälp av CRISPR Cas9 kommer att resultera i nedsatt hanfertilitet även vid fältförsök.

Försöket med CRISPR Cas9 och mutation som orsakar nedsatt hanfertilitet är intressant då det vid en framtida tillämpning har potential att underlätta förädlingsprocessen för hybridutsäde, i första hand för oljeväxter. Hybrid sorter har en högre avkastning, högre odlingssäkerhet och bättre enhetlighet, vilket gör dem viktiga för produktiviteten inom såväl nationell som internationell växtodling.

Genom den riktade mutation som induceras av CRISPR Cas9 kommer vi att ta fram tre olika typer av mutanter:

1. En enkel mutant med 1-2 nukleotiders deletion i första exonen i lokus At1g20120.
2. En enkel mutant med 3-4 nukleotiders deletion i första exonen av lokus At1g20120.
3. En mutant med 30 kb. deletion. CRISPR Cas9 kommer att orsaka ett dubbelsträngsbrott i At1g20120 samt i At1g20150, vilket resulterar i en deletion från första exonen av At1g20120 till och med At1g20150. Deletion av de gener som ligger i lokus At1g20130, At1g20132, At1g20135 och At1g20140 kommer således att ske.

De tre mutanterna erhålles genom att genspecifika guide-RNA (gRNA) designas och klonas in i den binära vektorn pFGC-pcoCas9. Cas9-vektorn kommer därefter att föras över till *Agrobacterium* för att sedan användas i en transformation av *Arabidopsis*. Transformationen kommer att ske genom en så kallad floral dip vilket innebär att vävnad från en sekundärt producerad blomstjälk (endast ett fåtal blommor) inokuleras i en vätska som består av *Agrobacterium tumefaciens*, 5%

sukros och 500 mikroliter per liter av surfactant Silwet L-77¹. Transformationen kommer att göras på den transformationsplattform som finns på SLU i Uppsala där tillstånd för innesluten användning av genmodifierade växter finns².

Efter transformationen kommer vi att erhålla en heterozygot med avseende på transgenen. I nästa generation kommer transgenen att segregeras bort och homozygota linjer som bär på CRISPR Cas9 kommer att väljas ut. Vid fältförsök med de tre mutanterna kommer transgenen på så sätt att vara bortsegregerad.

Frågor som vi vill få besvarade:

- Faller plantor som modifierats med hjälp av CRISPR Cas9 in under den rådande GMO-lagstiftningen?
- Måste vi söka tillstånd för att genomföra fältförsök med dessa plantor?
- Enligt vilket regelverk skall vi i sådana fall utforma våra fältförsök? Om regelverk saknas kan vi ändå utföra fältförsöken?

Detektion av de mutationer som skapas kommer för de två enkla mutanterna med 1-2 respektive 3-4 nukleotiders deletion att kunna ske genom en multiplex PCR-reaktion; en metod som är speciellt lämplig för detektion av mutationer och deletion av gener³. Detektionen av mutationerna som skapats i de två enkla mutanterna kommer dock inte kunna särskiljas från spontana mutationer. Den stora deletionen om 30 kb. kommer både att kunna detekteras med multiplex PCR och kommer även i princip att kunna särskiljas då denna typ av stora deletioner sker med mycket låg frekvens.

Vidare kommer vi att komplementera mutant #3, vilket innebär att vi kommer återställa denna till vildtyp. Detta kommer att göras genom transformation med *Agrobacterium*.

Fråga som vi vill få besvarad:

Behöver vi söka tillstånd för GMO för att få genomföra fältförsök med den återställda vildtypen?



Jens Sundström (Lektor/Projektledare)



Ida Eriksson (Agronomstudent)

¹ Clough, S. J., & Bent, A. F. (1998). Floral dip: a simplified method for *Agrobacterium*-

² The Uppsala Transgenic Arabidopsis Facility. http://lcpu.se/?page_id=763

³ Henegariu, O., Heerema, N. A., Dlouhy, S. R., Vance, G. H., & Vogt, P. H. (1997).

Multiplex PCR: critical parameters and step-by-step protocol. *Biotechniques*, 23(3), 504–511.



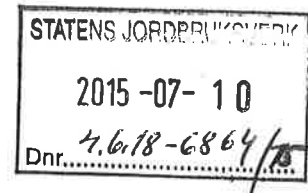
Appendix C

Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Department of Plant Biology
Associate Professor Jens Sundström

2015-07-08

Jordbruksverket
Regelenheten
Staffan Eklöf



1(3)

Kompletterande information till förfrågan om fältförsök med CRISPR-muterad *Arabidopsis* (Dnr 4.6.18-6864/15)

Nedan finns svar på de kompletterande frågor som vi blivit ålagda att besvara inför bedömning av huruvida vi behöver ansöka om tillstånd för att genomföra fältförsök med CRISPR-muterad *Arabidopsis* eller ej.

Fråga 1. (I er text står att homozygota linjer som bär på CRISPR Cas9 kommer att väljas ut. Menar ni inte "CRISPR Cas9-inducerade mutationer", alternativt "väljas bort"?)

Svar: Efter transformationen erhålls heterozygoter med avseende på transgenen. I nästa generation väljs homozygoter som **inte** bär på CRISPR Cas9-transgenen ut.

Fråga 2. (Hur har bortkorsningen bekräftats?)

Svar: De CRISPR Cas9-mutationer som induceras uppstår redan i primärtransformanterna (T_0). Då primärtransformanter alltid är heterozygota med avseende på transgenen, kommer T-DNA:t kunna segregeras bort i nästföljande generation (T_1). Att bortsegrering av T-DNA:t verkligen skett kommer att kontrolleras genom att genomföra följande:

- i) Riktad PCR mot gRNA:t och Bastagenen som ingår i DNA:t (se tabell 1 för primersekvenser). Förväntat resultat av en sådan analys kommer vara att transgenen går att detektera i $\frac{3}{4}$ av T_1 -plantorna medan $\frac{1}{4}$ visar negativa resultat.
- ii) Delar av avkomman till den $\frac{1}{4}$ -del av T_1 -plantorna som visat negativa resultat i PCR-screenen kommer att odlas på Basta-medium (50 mg/l) för att kontrollera att transgenen inte finns kvar.
- iii) För att säkerställa att endast vildtypsallelen av det lokus som bär på transgenen finns med i de linjer som används vid fältförsöken kommer positionen av transgenen i genomet att fastställas med hjälp av TAIL-PCR (http://bit.ly/Liu_Chen_2007). Utifrån

resultatet av TAIL-PCR försöken kommer primrar för genotypning som särskiljer ett vildtypslokus från ett lokus som bär på transgenen att konstrueras. Dessa primrar kommer därefter att användas för att genotypa plantor där T-DNA:t segregeras bort.

Tabell 1. Primersekvenser för riktad PCR mot gRNA:t och Bastagenen.

Primertyp	Primernamn	Sekvens
Basta	Basta (fwd)	GCACCATCGTCAACCACTAC (fwd)
	Basta (rev)	GTCATCAGATTTTCGGTGACG (rev)
gRNA	U6_F1 type 1 (fwd)	GCGGAATTCAGAAATCTCAAATTCGG
	U6_R2 type 1 (rev)	GCGTCTAGATAATGCCAACTTTGTACA
gRNA	U6_F1 type 2 (fwd)	GCGTCTAGAAGAAATCTCAAATTCGG
	U6_R2 type 2 (rev)	GCGGCGATCGCTAATGCCAACTTTGTACA

Fråga 3. (Har off-target-mutationer analyserats i växterna?)

Svar: Vid design av de gRNA-sekvenser som klonats in i vektorn pFGC-pcoCas9 i detta försök, har sekvenser med så låg off-target-frekvens som möjligt medvetet valts ut. Information om exakta positioner för var off-target-mutationer potentiellt kan ske finns sparad tillsammans med gRNA-sekvenserna.

Primärt har vi inte tänkt analysera eventuella off-target-mutationer eftersom detta inte är ett krav vid traditionell mutationsförädling. I *Arabidopsis* finns flera mutantlinjer framtagna med Ethyl methanesulfonate (EMS) som inte räknas som genetiskt modifierade organismer men som ofta bär på ett stort antal off-target mutationer. Dessutom kommer slumpmässiga mutationer naturligt att ske och skilja mellan generationerna, vilket gör att det är svårt att säkerställa huruvida en detekterad mutation är ett resultat av off-target eller om den uppstått spontant.

Om sekvensering ändå krävs kan denna göras på två olika nivåer:

- i) Sekvensering av de regioner i genomet som bedöms ha störst risk för att bära på off-target-mutationer.
- ii) Sekvensering av hela genomet.

Då sekvensering är ett kostsamt alternativ, speciellt vid sekvensering av hela genomet, vill vi om möjligt undvika det.

Vid konventionella mutationsförsök där man använder sig av EMS eller radioaktiv strålning och hundratals mutationer introduceras i ett slag brukar man göra en serie återkorsningar för att minska på antalet off-targets. Av hävd brukar man göra ca sex återkorsningar. Ett alternativ eller komplement till att sekvensera skulle kunna vara att göra återkorsningar till vild-typ för de linjer som vi avser använda i fältförsöken.

Hälsningar

Jens Sundström och Ida Eriksson

**EUROPEAN COMMISSION**
DIRECTORATE-GENERAL FOR HEALTH AND FOOD SAFETYSafety of the Food Chain
BiotechnologyBrussels,
SANTE/E1/IC/nm sante.ddg2.e.1(2015) 270852

Dear Competent Authority for Directive 2001/18/EC,

Subject: Herbicide tolerant oilseed rape produced using oligonucleotide directed mutagenesis

We have been informed that some Member States authorised, as exempted from the GMO legislation, field trials of herbicide tolerant oilseed rape produced using an oligonucleotide directed mutagenesis technique.

You are aware that the Commission is carrying out a legal analysis of a group of new plant breeding techniques, in order to evaluate whether they lead to organisms falling under the scope of Directives 2001/18/EC on the deliberate release into the environment of genetically modified organisms and 2009/41/EC on the contained use of genetically modified micro-organisms. Among these techniques, oligonucleotide directed mutagenesis (ODM) is also considered.

As you know, the deliberate release (field trials or placing on the market) of products which are subject to the rules of the EU GMO legislation without appropriate prior authorisation, is illegal.

This is why, even if we acknowledge the interest of companies to develop and market new products within the best timeframes, we would like to ask you, in the context of the current legal uncertainty, to await, as much as possible, the outcome of the Commission legal interpretation before authorising a deliberate release of organisms obtained with new plant breeding techniques. For the products which could be currently be subject to field trials with the agreement of the competent authorities of the Member States concerned, we recommend that, until the status of the techniques is clarified, a protective approach is implemented in order to be able to adapt to the requirement of EU legislation on GMOs should it be concluded that the organisms produced by these techniques are subject to its rules.

Being aware that the current legal uncertainty is unsatisfactory, the Commission' services are committed to present their legal analysis to the Competent Authorities and stakeholders before final adoption by the Commission foreseen before the end of the year.

Yours sincerely,

Dorothée André
Head of Unit*To all Permanent Representations of Member States of EU**Sent only by email*

Department of Plant Biology
Master's thesis • 30 hec • A2E
Agricultural programme - Soil/Plant • ISSN 1651-5195, No 149
Uppsala, 2015