MINISTERIUM
FRAUEN
GESUNDHEIT

RNAi-based techniques, accelerated breeding and CRISPR-Cas: basics and application in plant breeding

Universität für Bodenkultur Wien

Impressum

Eigentümer, Herausgeber und Verleger
Bundesministerium für Gesundheit und Frauen (BMGF)
Radetzkystraße 2, 1030 Wien

Autorinnen und Autoren
Dr.in Julia Hilscher
Univ. Prof. Dr. Hermann Bürstmayr
Department für Nutzplanzenwissenschaften und Department für Agrarbiotechnologie, BOKU Wien
Univ. Prof.in Dr.in Eva Stöger
Department für Angewandte Genetik und Zellbiologie, BOKU Wien

Der Bericht steht zum Download auf der Website des BMGF unter www.bmgf.gv.at im Bereich Gentechnik zur Verfügung.

Erscheinungsdatum
März 2017

ISBN 978-3-903099-18-0

Table of Contents

1 Introduction 1

1.1 Structure of the literature-based study and questions addressed 2

1.1.1 Definitions and explanatory notes to chapters 3

1.1.1.1 Definition of terms used in this study 3

1.1.1.1.1 Intended and unintended effects and safety considerations 3

1.1.1.1.2 Intermediate organism – resulting organism 3

1.1.2 Explanatory notes to chapters 4

1.1.2.1 Intended and unintended effects 4

1.1.2.2 Safety aspects 4

1.1.2.3 Aspects relating to GMO classification 4

1.1.2.4 Detection and identification 5

1.2 Interaction with stakeholders 6

1.3 Participation at GARNet/OpenPlant CRISPR-Cas Workshop 6

1.4 Recommendations (“Handlungsempfehlungen”) 8

2 CRISPR-Cas 9

2.1 Introduction 9

2.1.1 CRISPR-Cas9 mediated genome editing: underlying processes 11

2.1.2 Production processes of CRISPR-Cas9 genome edited plants 13

2.1.3 Techniques (SDN1,2,3) 15

2.2 Application in plant breeding 17

2.2.1 Potential applications of SDN1 17

2.2.1.1.1 Elimination of unwanted compounds 17

2.2.1.1.2 Increasing production of desired compounds 18

2.2.1.1.3 Engineering pathogen resistance by targeting recessive resistance genes 19

2.2.2 Potential applications of SDN2 21

2.2.3 Potential applications of SDN3 21

2.2.4 Applications other than genome editing 21

2.3 State of research and development in plants 23

2.3.1 Transferability of the system to plant species 23
2.3.2 Techniques (SDN1, 2, 3) 24
2.3.3 Delivery methods 24
2.3.4 Types of mutations generated by SDN1 technique 25
2.3.5 Off-target activity 26
2.3.6 Limiting off-target effects 28
2.4 Intended and unintended effects of CRISPR-Cas9 in genome editing 28
2.5 Safety considerations 29
2.5.1 SDN1 technique in genome modification of plants 29
2.5.1.1 Comparison of CRISPR-Cas9 and conventional mutagenesis techniques in relation to mutational load and type of modifications 29
2.5.1.2 Safety considerations in respect to CRISPR-Cas9 transgene retention, background mutations caused by transformation procedures and the use of viral vectors 31
2.5.2 SDN2 technique in genome modification of plants 31
2.5.3 SDN3 technique in genome modification of plants 32
2.6 Detection and identification 32
2.6.1 Detection and identification of SDN1 and SDN2 genome editing 32
2.6.2 Detection and identification of SDN3 genome editing 33
2.7 Aspects of GMO classification of CRISPR-Cas9 genome edited plants 33
2.7.1 Evaluation of ZFN and related genome editing techniques by the German expert commission ZKBS 35
2.8 Tables 37
3 Accelerated breeding – rapid cycle breeding 41
3.1 Introduction 41
3.2 Potential applications in plant breeding 43
3.3 State of development 44
3.3.1 Species of interest and genes tested for precocious flower induction 44
3.3.2 Experimental systems to induce precocious flower induction 45
3.3.3 Current rapid-cycle breeding programmes 46
3.3.4 Establishing infrastructure for rapid-cycle breeding programmes 48

3.4 Intended and unintended effects 49
3.5 Safety considerations 50
3.6 Identification and detection 51
3.7 Aspects of GMO classification 51
3.7.1 Evaluation of a related breeding practise by the German expert commission ZKBS 52
3.8 Tables 54
4 Small RNA-directed techniques 60
4.1 Introduction 60
4.1.1 miRNAs 61
4.1.2 siRNAs 63
4.2 Application of RNAi approaches in plant breeding 64
4.2.1 Applications based on targeting plant endogenous genes 65
4.2.2 Applications by targeting RNA expressed by plant pathogens 67
4.3 State of development 69
4.4 Intended and unintended effects 72
4.5 Safety considerations 72
4.5.1 EFSA workshop on risk assessment considerations for RNAi-based GM plants 73
4.5.2 EFSA call on literature review to support risk assessment of RNAi-based GM plants 76
4.6 Detection and identification 77
4.7 Aspects of GMO classification 77
4.8 Table 78
5 Abbreviations 84
6 References 86
7 Appendix 101
7.1 Literature Search 101
7.1.1 CRISPR-Cas 101
7.1.2 Rapid cycle breeding 101
7.2 Definition of GMO according to EU Directive 2001/18/EC on the deliberate release into the environment of genetically modified organisms 102
7.3 Tables 104

1 Introduction

The aim of plant breeding is to develop and select plants adapted to human needs [1]: breeding objectives include abiotic and biotic stress tolerance, increased yield and/or yield stability, but also for example the development of value-added crops with increased protein content or altered fatty acid composition.

After being dependent on naturally occurring variation in plants for domestication and later for breeding, the 20th century brought techniques to support the breeding process and cultivar development. Mutation breeding is a method of artificially inducing mutations, which form the genotypic basis of differing traits. Polyploidy induction, i.e. doubling chromosome sets, may lead to cultivars with higher biomass. Other techniques facilitate re-combining (nuclear and/or cytoplasmic) genomes, like protoplast fusion. Finally, in the 1980ies transformation of plants with selected additional genetic material became possible. These and other biotechnological techniques increase genotypic variation in a given gene pool, which can be utilized directly or as basis for further breeding material.

Directive 2001/18/EC regulates the deliberate release of genetically modified organisms (GMO) and Regulation (EC) 1829/2003 the food and feed use of GMOs. GMOs falling under these regulations and exemptions are defined in Directive 2001/18/EC. Since formulation of the legal definition of a GMO, progress in research and development brought questions from stakeholders to competent authorities in European Union Member States on whether certain techniques lead to such regulated GMO´s. A working group (WG) was established in 2007 to identify and discuss so called “new plant breeding techniques” (NPBT) in relation to the definition of a GMO and in light of the most recent available scientific data [2]. The techniques under scrutiny contained (1) zinc finger nuclease (ZFN) technology, (2) oligonucleotide directed mutagenesis (ODM), (3) cisgenesis and intragenesis, (4) RNA-dependent DNA methylation (RdDM), (5) Grafting (on GM rootstock), (6) reverse breeding, (7) agro-infiltration, and (8) synthetic genomics [2].

CRISPR-Cas and accelerated breeding are covered in this report under the light of the above. Based on modules of the CRISPR-Cas system a genome editing technique was developed, the most recent addition to site directed nuclease (SDN) techniques, joining ZFNs. Accelerated breeding is a breeding strategy that uses a GMO to accelerate individual breeding cycles; the resulting plants, though, do not carry the early flowering transgene. In this sense, using a GMO intermediate in a breeding process, accelerated breeding has parallels to reverse breeding. The eight NPBT have been covered in studies conducted by AGES [3, 4]; CRISPR-Cas and accelerated breeding have come into focus very recently, thus the coverage in this report.

RNAi-based plants are plants falling under the definition of GMO in Directive 2001/18/EC. They express a transgene transcribed into an RNA molecule that downregulates a third gene and so confers the desired phenotype. RNAi-based GM plants have been among the very first commercially developed GM plants (FlavrSavr^{TM}), however, to date the majority of genetically modified plants (GMP) authorized in the EU are based on expression of one or more transgenes expressing proteins that confer the desired phenotype. RNAi-based GM plants have again come into focus for example because of their potential for engineering pathogen resistance traits. There is an ongoing process in the EU to evaluate whether the risk assessment implemented for GMPs in general may be specifically adapted to RNAi-based GM plants. RNAi-based GM plants and the ongoing activities on questions in relation to adaptation of risk assessment are covered in this report. The sub-category of RNAi-based GM plants functioning through RNA-dependent DNA methylation (RdDM) has been covered by AGES [4] and is not further covered in this study.

1.1 Structure of the literature-based study and questions addressed

CRISPR-Cas9 is a novel site directed nuclease technique and accelerated breeding a relatively novel concept integrated in conventional breeding strategies. For these two, a literature search was undertaken to collect available primary research publications. Details on the search strategy in scientific literature databases can be found in Appendix 7.1.

Basic research into RNAi based pathways goes back to the early 1990ies. Description of the RNAi-based techniques, current state of application and development were guided by the most recent peer reviewed secondary literature present, and where informative to the focus of this study, expanded by data from primary research publications. Furthermore, publicly available documents by EFSA informing on ongoing developments on risk assessment evaluation are included.

Literature search ended March 2016.

Following a general description, (i) applications in plant breeding, (ii) the state of development in plant systems, (iii) intended and unintended effects upon application, (iv) safety considerations, (v) detection and identification, and (vi) aspects of GMO classification, are addressed for each of the techniques.

1.1.1 Definitions and explanatory notes to chapters

1.1.1.1 Definition of terms used in this study

1.1.1.1.1 Intended and unintended effects and safety considerations

GM risk assessment is focused on identifying and characterizing potential adverse effects on human and animal health and on the environment, both of intended and possible unintended effects caused by GM-based plants.

The term intended and unintended effects was defined in the "Scientific Opinion on Guidance on the environmental risk assessment of genetically modified plants", EFSA Journal 2010 [5]:

"Intended effects are those that are designed to occur and which fulfil the original objectives of the genetic modification. Alterations in the phenotype may be identified through a comparative analysis of growth performance, yield, pest and disease resistance, etc. Intended alterations in the composition of a GM plant compared to its appropriate comparator, may be identified by measurements of single compounds.

Unintended effects of the genetic modification are considered to be consistent (non-transient) differences between the GM plant and its appropriate comparator, which go beyond the primary intended effect(s) of introducing the transgene(s). […] these unintended effects are event-specific, applicants must supply data on the specific event. Sources of data that may reveal such effects are: 1. Molecular characterization […]. 2. Compositional analysis […]. 3. Agronomic and phenotypic characterization […]. 4. GM plant-environment interactions […].". [5]

In this report, intended and potential unintended effects on the plant genome and derived safety considerations are specified and discussed for the application of the particular techniques, based on the current state of the science.

1.1.1.1.2 Intermediate organism – resulting organism

The terms intermediate and resulting organism are used in this study in the chapters covering CRISPR-Cas and accelerated breeding. In this report the following terms are used as defined in the NTWG (New Techniques Working Group) final report of 2011; the report was never officially published but can be accessed via a link in [6].

Resulting organism was therein defined as

"… an organism that results after having gone through all the steps of the particular technique. This could be a plant or seed intended for deliberate release or placing on the market or a microorganism intended for contained use."

and intermediate organism as

"…any organism that is generated in the steps leading to the resulting organism."

The NTWG was composed of national experts nominated by the Competent Authorities of EU Member States in 2008. Their objective was to analyse whether specific biotechnological methods, including ZFN and related techniques, lead to resulting organisms falling under the definition of a GMO Directive 2001/18/EC [2].

1.1.2 Explanatory notes to chapters

1.1.2.1 Intended and unintended effects

Intended and potential unintended effects on the plant genome are specified and discussed due to the application of the particular techniques based on the current state of the science.

1.1.2.2 Safety aspects

Directive 2001/18/EC explicitly excludes plants generated by conventional mutagenesis breeding and plants generated by cell or protoplast fusion, as well as does not consider plants generated by polyploidy induction falling under the GMO definition; plants generated by these techniques are exempted from the risk assessment and regulatory procedure established by Directive 2001/18/EC that – based on the precautionary principle – has the objective to protect human health and environment.

This is based on the grounds of considerations given in recital 18 of Directive 2001/18/EC which reads that the "Directive should not apply to organisms obtained through certain techniques of genetic modification which have conventionally been used in a number of applications and have a long safety record."

The Directive therefore implicitly states that the risks associated arising from intended and unintended mutations by the exempted techniques, mutagenesis breeding, cell culture methods and bringing together related genomes or multiplication of genomes, are considered to be manageable outside the regulatory procedure of Directive 2001/18/EC, that is, by the breeding practices implemented by breeders.

Therefore, unintended effects on the genome arising due to application of these exempted techniques that may be applied during the production process of CRISPR-Cas9-based genome edited plants or during rapid-cycle breeding are treated the same in this report.

1.1.2.3 Aspects relating to GMO classification

Directive 2001/18/EC and Regulation EC/1829/2003 provide authorization procedures for deliberate release and placing on the market of genetically modified organisms (GMO) as well as for food and feed derived from GMOs. In Directive 2001/18/EC a definition of organisms falling under the

authorization procedure is given and exemptions are specified (Articles 2 and 3 and Annex IA, IB; see excerpt in Appendix 7.2).

8 NPBTs were assessed by the NTWG, for whether they generate organisms falling under the GMO definition in Directive 2001/18/EC. Similarly, the ZKBS (Zentrale Kommission für die Biologische Sicherheit), established under the scope of the German Gene Technology Act, published a position statement [7] on new plant breeding techniques.

Information in this report relating to CRISPR-Cas and to accelerated breeding may be used to interpret organisms in relation to relevant paragraphs of the legal GMO definition in Directive 2001/18/EC. In this chapter, thus, the techniques will be described in regard to the different steps involved in carrying out the techniques and the generated intermediate and resulting organisms. Where applicable, ZKBS expert opinions on analogous techniques are reported.

1.1.2.4 Detection and identification

To date, most commercialised genetically modified (GM) plants and all GM plants listed in the European Union GMO register (Regulation EC 1829/2003) are based on integration of transgenes containing one or more non-plant derived sequences, for example the cauliflower mosaic virus (CaMV) 35S promoter or bacterial herbicide tolerance conferring phosphinotricin-*N*-acetyltransferase sequences (pat, bar) [8]. Detection of (unauthorized) GMOs uses the common occurrence of these signature sequences (element and/or construct specific) in various GM plant lines; a platform (JRC-GMO-Matrix [9], storing information on known GM events) supports in deciding of an optimal screening strategy for a given sample. A first screening step detecting element and/or construct specific sequences establishes GM presence or absence (detection). In case GM presence is detected, validated analyses to identify event-specific sequences are carried out in order to unequivocally identify unique GM plant lines (identification). An event-specific detection method is an integral part of an application dossier for any GMO authorization in the European Union. Event-specific markers span the junction between the transgene insertion site and the genomic target site. Polymerase chain reaction (PCR) derived methods for detection, identification and quantification are commonly used (real time PCR using hybridization probes; see European Reference Laboratory for GM Food and Feed[1]).

In this chapter the possibility of detection and identification of intermediate and resulting organisms of the covered techniques CRISPR-Cas and accelerated breeding will be described.

[1] http://gmo-crl.jrc.ec.europa.eu/gmomethods/

1.2 Interaction with stakeholders

Information material was collected in the course of this study to be used for research education, e.g. within the program for the “Long Night of Research”, where genome editing was explained to the general public.

Information gathered within the study was also used for an article in the Austrian journal “Der Pflanzenarzt” (Neue Züchtungsmethoden: Gentechnik – oder doch keine Gentechnik?, 4/2016, p 24-27) in several talks held at meetings of breeder´s associations and other stakeholder associations.

- Vereinigung österreichischer Pflanzenzüchter, June 2015
- Klausur der Saatbau Linz, November 2015
- Saatgutgipfel der AGES; April 2016
- Interne Diskussion in der LKÖ zum Thema „Neue Methoden der Gentechnik“, April 2016

1.3 Participation at GARNet/OpenPlant CRISPR-Cas Workshop

Overall, the GARNet/OpenPlant Workshop at the John Innes Centre, UK (September 2015) provided an excellent environment to meet researchers working with CRISPR in crop species. It gave an update on state of the art of CRISPR-Cas applications in plants and made aware of where to look for current and future developments in the highly active field of CRISPR-Cas9 plant genome editing.

The Workshop gave an overview on CRISPR-Cas9 applications, reported on its current use in plant genome editing and on ongoing developments, especially in regard to optimization of efficiency and specificity. A meeting report has been published by the organisers [10].

Speakers presented data of successful genome editing by CRISPR-Cas9 in a wide variety of species, also in the crop plants maize, rice, wheat, tomato and potato. Vladimir Nekrasov (John Innes Centre, UK) described the production of a powdery mildew resistant tomato variety (cv “Moneymaker”). They used CRISPR-Cas to knock out the MILDEW RESISTANCE LOCUS O1 (Mlo1). Homozygous knockout mutants were present in the first generation of transgenic plants, and transgene free plants stably inherited the mutation. *mlo1* plants showed complete resistance against *Oidium neolycopersici*. In rice, Bing Yang (Iowa State University, USA) reported CRISPR-Cas mediated production of two independent OsSWEET13 knock-out lines which conferred resistance to *Xanthomonas oryzae*, the causal agent of rice bacterial blight.

Examples of ongoing work to further improve gene editing efficiency at various steps in the process included a database for gRNA design now also of use for diverse plant species (Edward Perello Desktop Genomics, UK) or explanation of various multiplexing strategies, like the use of synthetic tRNA-gRNA polycistronic genes (Bing Yang). The issue of specificity was for example addressed by

Oleg Raitskin (The Sainsbury Laboratory, UK) who screens variants of Cas9 nuclease and sgRNA combinations in order to find increased specificity. Holger Puchta (Kit, Germany) presented an already available strategy to decrease off-target effects by using two Cas9 nickase variants guided to adjacent positions and so resulting in a desired double strand break only if two nickases are placed in vicinity.

1.4 Recommendations ("Handlungsempfehlungen")

The present report on CRISPR-Cas and accelerated breeding applications in plant breeding provides background information on the fundamentals and the application potentials of these techniques as well as the state of development. It describes intended and unintended effects on the plant genome in relation to other plant breeding techniques and biotechnological methods.

It is intended as an information document for policy makers and stakeholders. The discussion about the so called new plant breeding techniques (NPBT) and their legal classification in the EU is now nearing a decade. In the meantime, as exemplified by the existence of this report, further techniques and breeding strategies have been developed and applied and knowledge on biotechnological methods and its impact on plant breeding have been increasing. All of the techniques hold great potential for utilization in plant breeding and development of crop cultivars. On the other hand, the legal classification of NPBTs, whether classified as falling under the GMO definition of Directive 2001/13/EC, them being exempted, or development of different regulatory procedures [11, 12], has consequences on their use and application in plant breeding.

To date, there are solid information documents available by scientific experts on the fundamentals of the different techniques and their potentials, furthermore, position statements from many stakeholder groups have been put forward; overall, a huge amount of scientific, legal and economic efforts have been carried out in regard to diverse aspects of NPBTs and related biotechnological methods. Therefore, the next step is to be done by policy makers to decide on the handling of NPBTs in order to ensure legal certainty to developers and plant breeders for their products.

Information of the public by public authorities in respect to plant breeding and biotechnological methods, their development and application in plant breeding should be an active process and guided by the current state of science and technology.

2 CRISPR-Cas

2.1 Introduction

CRISPR-Cas

CRISPR-Cas (Clustered regularly interspaced short palindromic repeats – CRISPR associated gene) is an RNA-guided DNA endonuclease complex present in bacteria and archaea. In 2012 it was recognized that it can be employed for targeted genome editing [13] and since then publication numbers have risen to develop and apply genome editing using CRISPR-Cas in various organisms, ranging from bacterial to human cells (see for example Table 1 in [14]). Fig. 2.1 illustrates publication activity for CRISPR-Cas9 in plant research.

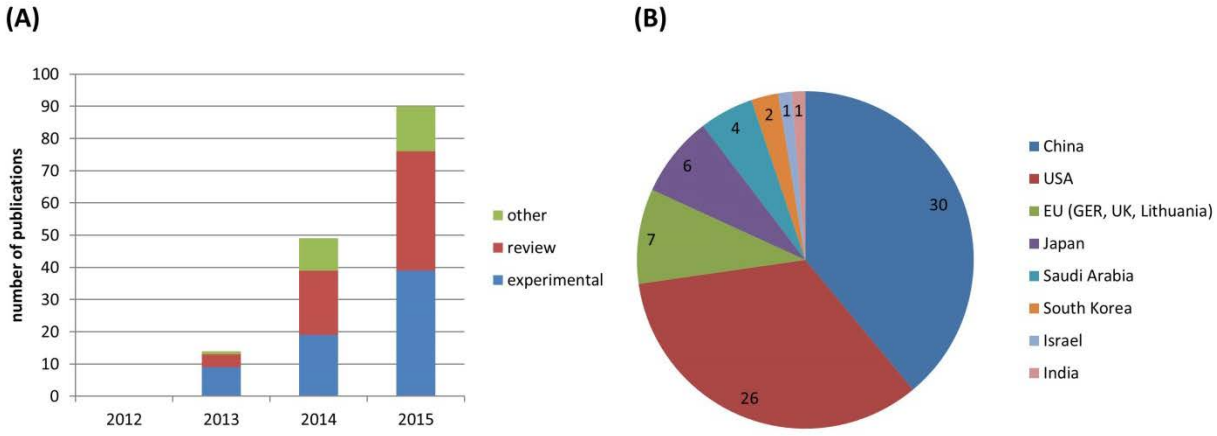


Fig. 2.1 CRISPR-Cas9 publications in plant research 2012-2015. Publications were retrieved from pubmed, Web of Science, Scopus and Ovid according to defined search criteria (see Appendix 7.1). (A) Publication numbers per year, subgrouped based on experimental (research and methodical articles), review (reviews, opinion articles and book chapters) or other publication type (meeting abstracts, publications in languages other than english, etc). (B) Country of origin of scientific papers reporting experimental data on CRISPR-Cas9 in plant research 2013 – 2015 (based on first author).

CRISPR-Cas is a recently understood adaptive “immune system” in prokaryotes against foreign DNA and RNA (reviewed for example in [15]). Present in about 90% and >40% of to date known archaeal and bacterial genomes, three main types of CRISPR-Cas systems have been identified. Fig. 2.2 outlines CRISPR-Cas function based on type II systems [16]: (1) mediated by CRISPR associated (Cas) genes, invading DNA is recognised and fragments (termed spacers) of foreign DNA are incorporated into the bacterial genome at the CRISPR locus; (2) the CRISPR locus is transcribed as precursor RNA; (3) the precursor RNA is processed into mature CRISPR RNAs (crRNAs), then hybridizes to a trans-activating CRISPR RNA (tracrRNA) and is bound by a Cas9 protein; (4) the CRISPR-Cas9 complex is

guided to specific DNA locations specified by the spacer region of crRNA component and DNA cleavage is mediated by the Cas9 protein. The RNA component of the CRISPR-Cas9 type II complex is also termed dual guide RNA (crRNA hybridized to tracrRNA).

CRISPR-Cas subtypes are classified based on the Cas genes involved, and as a consequence differing ribonucleo-protein complexes and modes of target interference. CRISPR-Cas type III has also the ability to target and cleave RNA [16].

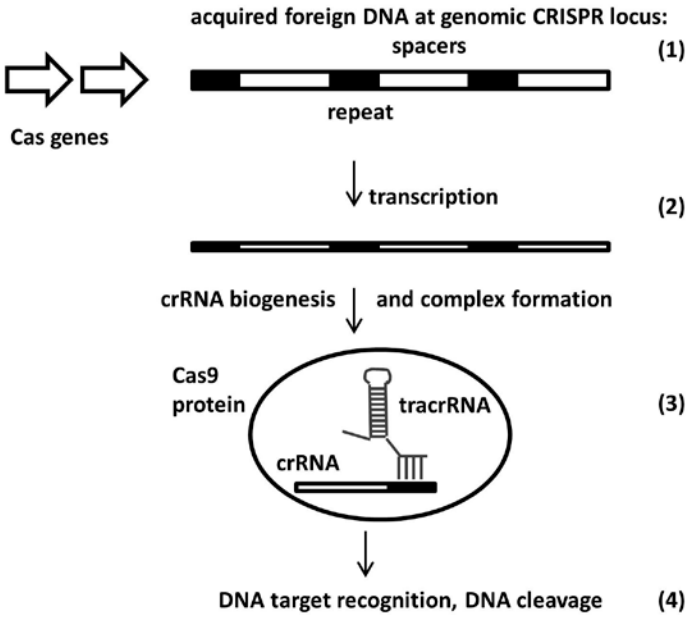


Fig. 2.2 Simplified model of CRISPR-Cas organisation, biogenesis and targeting exemplified by the type II system. (1) The CRISPR-Cas genomic locus contains the Cas protein coding genes and the CRISPR locus coding for the RNA component of the CRISPR-Cas complex. The latter is composed of acquired spacers from invading DNA and interspersed repeat sequences. (2) The Cas genes (coding for example for Cas9) and the CRISPR precursor are transcribed and (3) the CRISPR precursor RNA cleaved into crRNA moieties, which hybridized to a tracrRNA, is bound by the Cas9 protein. The crRNA and tracrRNA components together are called dual guide RNA. (4) Mature CRISPR-Cas9 complexes target DNA sequences showing complementarity to the spacer region of the crRNA and induce DNA double strand breaks. crRNA: CRISPR RNA. tracrRNA: trans-activating CRISPR RNA.

Genome editing

The technology of random mutagenesis is used to induce genetic variability in plant breeding and research. Upon exposure to, for example, radiation or chemical mutagens a large population of plants has to be screened phenotypically or genotypically to select those with desired phenotypes/genotypes. With genome editing technology it is now possible to target genomic positions to introduce variability, i.e. to generate plants with precise modifications or to insert foreign DNA at targeted genomic positions. Genome editing has been made feasible by development of several systems, all based on proteins acting as site directed nucleases (SDN), i.e. enzymes

introducing DNA double strand breaks (DSB): zinc finger nucleases (ZFN), TAL effector nucleases (TALEN), meganucleases (MN) and recently CRISPR-Cas9 (reviewed for example in [17]). These technologies share the same mechanism: they are programmable for precise typesetting of DNA double strand breaks (DSB) which are then recognised by diverse endogenous cellular repair systems. In some cases these are imperfect and incorporate errors, alternatively, DSB repair mechanisms can be tricked into modifying genomic sequences or inserting extraneous DNA by providing repair templates, all of which is exploited in genome editing.

CRISPR-Cas9 is a genome editing technique that can be used to introduce mutations at selected genomic loci. It is based on components of a naturally occurring pathway present in bacteria and archaea: the enzyme Cas9 that is able to introduce a double strand break into DNA; the associated RNA component can be easily re-programmed to target Cas9 to selected loci of eukaryotic genomes.

2.1.1 CRISPR-Cas9 mediated genome editing: underlying processes

To date, type II CRISPR-Cas9 modules are mainly used for genome editing of pro- and eukaryotes [14], including plants [18]. The Cas9 protein is mostly based on the sequence of the homolog of *Streptococcus pyogenes* (*Sp*Cas9). However, Cas9 homologs of further organisms, as well as other CRISPR-Cas subtypes are used and/or investigated for application in genome editing, or other uses. Below, CRISPR-Cas9 - target DNA interaction is explained in more detail to aid in understanding of the issue of off-target effects; it relates to the type II CRISPR-Cas9 subtype, if not indicated otherwise. Furthermore, DNA repair pathways operating in plant cells are briefly introduced.

DNA double strand break (DSB) generation by CRISPR-Cas9

To recognize target DNA sequence and execute a DNA double strand break (DSB), a natural CRISPR-Cas9 complex consists of the DNA endonuclease Cas9 protein (executing the DSB) bound to the crRNA:tracrRNA (termed dual guide RNA) (Fig. 2.2). The 5´ end of crRNA harbours the spacer, i.e. the complementary region for target recognition, the crRNA 3´end hybridizes with the tracrRNA to form a secondary structure required for Cas9 binding (Fig. 2.2).

It was discovered that engineering a chimeric guide RNA, called single guide RNA, that carries a spacer sequence of choice (depending on the desired genomic target) at the 5´end followed by a 3´end hairpin structure (mimicking tracrRNA:crRNA secondary structure) also form functional entities (Fig. 2.3) [13], which is exploited for use in genome editing.

Cas9 proteins possess two separately acting nuclease domains homologous to HNH and RuvC nucleases, cutting the complementary and non-complementary DNA strand, respectively [13]. However, Cas9 is also involved in target recognition: its PAM Interacting (PI) domain scans target

DNA for protospacer adjacent motifs (PAMs) (Fig. 2.3). PAMs are short signatures (typically 2-5 nucleotides [19]) directly downstream (type II) of protospacers (i.e. the signature sequences in the target DNA) but not incorporated into the CRISPR loci that are crucial for target recognition; they also dictate the location of the DSB executed by the Cas9 nuclease domains. Thus, if Cas9 loaded with gRNA recognises PAM sequences, the gRNA-Cas9 complex interrogates DNA directly upstream to PAMs for complementarity to the spacer sequence. In the course a RNA:DNA heteroduplex is formed, and in case of substantial complementarity target DNA is cleaved approximately three nucleotides upstream of the PAM at both strands (reviewed for example in [14]). Efficient target cleavage is dictated by near complementarity of the last 8-12 nucleotides of the spacer sequence (called seed region) to the target and the presence of the PAM nucleotides in the protospacer sequence [20] (Fig. 2.3).

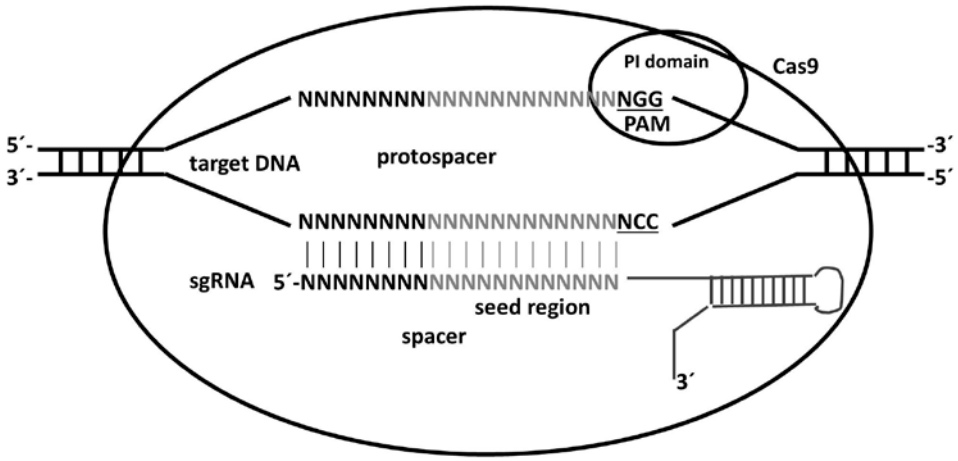


Fig. 2.3 Schematic depiction of elements involved in CRISPR-Cas9 – target DNA recognition. The PAM interacting domain (PI) of Cas9 scans DNA for PAM sequences (typically NGG in type II system of *Streptococcus pyogenes*). In case the protospacer region upstream of the PAM shows high complementarity (specially in the seed region) to the spacer region of the sgRNA, Cas9 executes a DNA DSB approximately 3 nucleotides upstream of the PAM in the target DNA. The 8 - 12 nucleotides constituting the seed region proximal to the PAM are depicted in grey. DSB: double strand break. PAM: protospacer adjacent motif; PI: PAM interacting domain; sgRNA: single guide RNA.

DNA double strand break (DSB) repair pathways exploited for genome editing

In plant cells, non-homologous end-joining (NHEJ) and homologous recombination (HR) mediated repair pathways execute repair of occurring DNA DSBs. NHEJ, the prevalent mechanism in somatic plant cells [21], is error-prone and often introduces smaller insertion or deletion mutations upon re-ligation of DNA ends. DNA ends ligated together do not need to show homology. HR mediated repair mechanisms rely on information from homologous regions. DNA ends at DSBs are processed into single-stranded 3´overhangs by 5´-3´exonuclease activity, and bound by HR-proteins (for example RAD51) which scan DNA for homologous regions. In somatic plant cells mainly two HR mediated

repair pathways seem to operate, single-strand annealing (SSA) and synthesis-dependent strand annealing (SDSA). SSA results in ligation of two annealing ssDNA strands. SDSA scans for complementary regions in duplex DNA by strand invasion and uses a detected homologous strand as repair template by initiating DNA synthesis. Synthesis finishes, and in case the now extended strand harbours again complementary sequence to the second resected single-stranded 3´overhang, the DSB can be repaired (for review see [21]).

Genome editing using CRISPR-Cas9 employs two molecular modules: it uses an engineered CRISPR-Cas9 module to execute a DNA double strand break (DSB) at a chosen site in the plant genome; in a second step, DNA DSBs, which also occur under natural conditions, are repaired by endogenous DNA DSB repair pathways. These repair pathways are error-prone, resulting in mutations; alternatively, these may be used to mediate site specific integration (at the DSB) of cis-, intra-, or transgene.

2.1.2 Production processes of CRISPR-Cas9 genome edited plants

For a given plant species, the production process of CRISPR-Cas9 genome edited plant lines depends on established transformation and, if a cell culture step is included, regeneration procedures (Fig. 2.4). They all share a step of (a) delivery of a gRNA-Cas9 module (and optionally a repair template) into plant cells and (b) screening for genome edited lines. There are several modes of gRNA-Cas9 delivery, including different vector systems, in use. CRISPR-Cas9, and in extension genome editing techniques involving site directed nucleases (SDN), introduce heritable changes *in trans*, therefore transgenic integration of a CRISPR-Cas9 gene cassette during the production process is not obligatory and if present, can be segregated out in sexually reproducing species.

Stable transformation of gRNA-Cas9 gene cassettes: gRNA-Cas9 gene cassettes including a selectable marker gene are transformed into plant cells and have become stably integrated during a selection step. gRNA-Cas9 is expressed from transgenic DNA. Transformation methods mainly used are *Agrobacterium*-mediated gene transfer and microprojectile (particle) bombardment, or electroporation and polyethylene-mediated transformation for plant protoplasts. In crop species which can be propagated by sexual reproduction genome edited progeny free of the CRISPR-Cas9 cassette including the marker gene can be selected in the next generation(s). In this case, transgenic events are present in intermediate products during the production process but are lacking in the final established plant line (resulting organism). Production processes involving stable transformation to date are the main published production processes in plants.

Transient transformation of gRNA-Cas9 gene cassettes: gRNA-Cas9 gene cassettes are transformed into plant cells and CRISPR-Cas9 is expressed from these templates. Transformation methods are as

above. The production process does not include a selection step for stable genomic integration of the gene cassette. A second strategy for transient delivery of the gRNA-Cas9 gene cassette uses viral vectors. They may either be delivered via *Agrobacterium*-mediated gene transfer, via virions or isolated viral RNA (RNA viruses). Genome editing using DNA virus (Cabbage Leaf Curl virus (CaLCuV), bean Yellow Dwarf virus (BeYDV)) systems [22, 23] and an RNA virus (Tobacco Rattle virus (TRV)) system [24, 25] have been shown to date. RNA viral vector systems were not yet shown to deliver a complete CRISPR-Cas9 gene cassette, but were shown to deliver sgRNAs into plants stably expressing the Cas9 component. However TRV virion delivery engineered to express ZFNs has been used to generate genome edited tobacco lines [26].

Delivery of pre-assembled gRNA-Cas9 ribonucleo-protein complexes: Ribonucleo-protein complexes are delivered into plant cells and directly exert their function [27]. PEG mediated delivery of particles has been carried out. This method does not involve DNA delivery into plant cells in case of SDN1 techniques (for definition of SDN1 please refer to chapter 2.1.3).

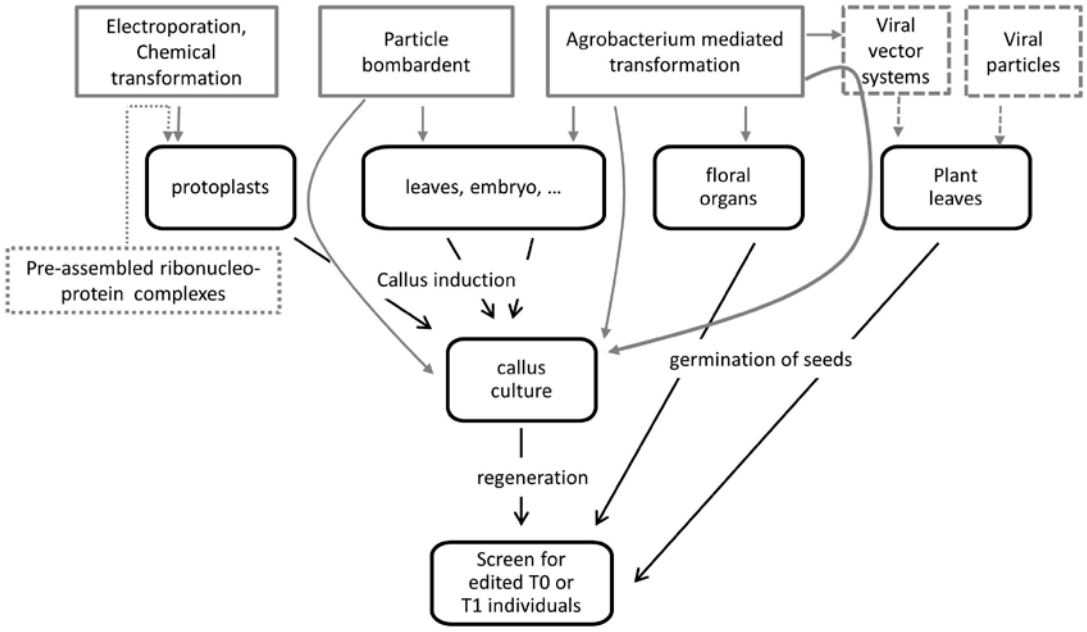


Fig. 2.4 Production processes of genome edited plants using CRISPR-Cas9 (modified after [28]). Grey boxes indicate methods to deliver CRISPR-Cas9 into cells. Delivery of CRISPR-Cas9 activity may be independent or dependent on DNA transfer into plant cells. Grey unbroken lines: DNA transfer; grey dotted lines: no DNA transfer; grey dashed lines: DNA transfer optional.

Minimal gene cassette requirements in case of recombinant DNA based transformation procedures

A gRNA-Cas9 minimal gene cassette consists of a Cas9 gene (to date mostly derived from *Streptococcus pyogenes*) fused to a nuclear localization signal (NLS) located between a polymerase II promoter and terminator to initiate and terminate transcription, respectively (Fig. 2.5). The sgRNA is driven and terminated generally by polymerase III regulatory sequences. The spacer sequence, in

plants typically 19-22 nucleotides in length, is selected based on the target of interest. For optimal guide selection and to reduce off-target potential bio-informatic tools are available (for example [29, 30]). The two components may be placed on the same or on two separate vectors [18, 20]. In case genome modification or insertion of cis-, intra-, or transgenic sequences is the goal, additionally a sequence acting as repair template is included.

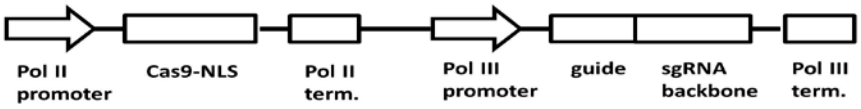


Fig. 2.5 Minimal gene cassette requirements for CRISPR-Cas9 mediated genome editing. The coding sequence giving rise to Cas9 is placed between a polymerase II promoter and terminator sequence, to initiate and stop transcription, respectively. Cas9 is fused to a nuclear localization sequence (NLS) to ensure nuclear localization. The sgRNA sequence is generally placed between a polymerase III promoter and terminator sequence.

Production processes of genome edited plants using CRISPR-Cas9 may involve generation of intermediate plants stably incorporating a gRNA-Cas9 transgene. In case of sexually reproducing crops, resulting genome edited lines without the transgene but with the intended mutation are selected.

Furthermore, genome edited lines may be established using transient transformation procedures, i.e. plants are transformed with a gRNA-Cas9 transgene, but it is not integrated into the genome. The generated mutation, but not the gRNA-Cas9 transgene, is passed on to the next generation.

Finally, gRNA-Cas9 complexes may be delivered to the cells without the involvement of DNA, as pre-assembled ribonucleoprotein complexes.

2.1.3 Techniques (SDN1,2,3)

There are different types of targeted genome modifications that can be achieved by using site directed nucleases (SDN) including CRISPR-Cas9, by placing a targeted DSB(s) and, optionally, at the same time providing a repair template (Fig. 2.6): (1a) generating gene knock outs by inducing site specific random mutations due to erroneous NHEJ repair, (1b) gene deletions by placing two DSBs leading to the loss of the genomic region within, (2) gene modification by site specific nucleotide sequence changes mediated by a repair template with homology and (3a, 3b) gene insertion by providing repair or donor templates. The NTWG (active under the request of competent authorities (CA) under Directive 2001/18/EC) subcategorized ZFN and related SDN techniques in genome editing based on their outcomes into SDN1, SDN2 and SDN3 [31] which correspond to repair pathways 1a, 2 and 3a in Fig. 2.6, respectively.

The strategic outcomes are recapitulated for CRISPR-Cas9 below:

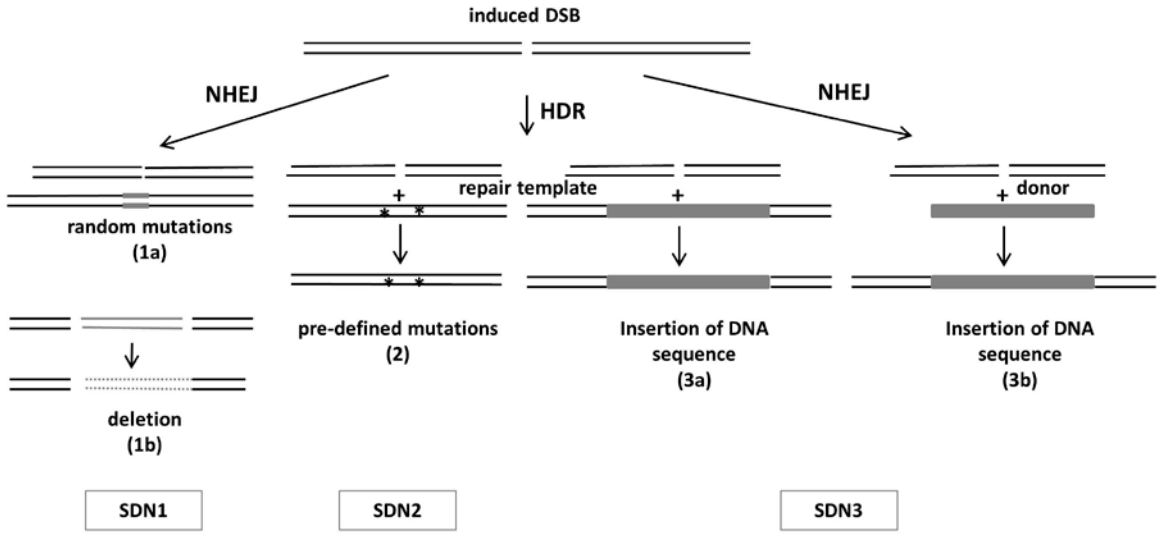


Fig. 2.6 CRISPR-Cas9 genome editing (after [14]). Targeted DSBs induced by CRISPR-Cas9 can either lead to random mutations at the DSB site (1a) or, in case two DSB are induced, to deletion of the genomic region within (1b), both mediated by NHEJ. In case a repair template with regions of homology is provided together with the CRISPR-Cas9 module, pre-defined mutations (2) or precise insertions of DNA sequences (3a) can be implemented at the DSB by HDR. Gene insertions can also be generated by providing donor molecules without homology which are inserted at the DSB by NHEJ (3b). DSB: double strand break; HDR: homology dependent repair; indel: insertion/deletion mutation; NHEJ: non-homologous end joining. SDN1, 2, 3: categories of the technique according to the definitions used in a regulatory context (site directed nuclease).

Technique SDN1: sgRNA-Cas9 activity module is delivered into cells and introduces a targeted DSB. DSBs repaired by NHEJ may lead to site specific random mutations, i.e. insertions, deletions, substitutions or a combination of these. These can be exploited in cases mutations lead to, for example, gene knock-outs by frameshift mutations when targeted to coding regions. The DSB can also be targeted to non-coding regions, for example to impair or delete regulatory elements, thereby inducing a change in gene expression (Fig. 2.6 (1a)). In extension to the original definitions by the NTWG, two DSBs can be placed by delivery of two Cas9-gRNA modules targeting different locations, resulting in deletion of the region in-between (Fig. 2.6 (1b)). Finally, placing of two DSBs has the potential to induce chromosomal re-arrangements (inversion, duplication or translocation events) which may be exploited for genome editing [32].

Technique SDN2: sgRNA-Cas9 activity together with a DNA repair template is delivered into cells. The repair template is homologous to the targeted region with exception of site specific nucleotide sequence changes (single nucleotide changes or small insertions/deletions). sgRNA-Cas9 activity induces a targeted DSB. In the course of HDR, the repair template may be used and the desired site specific nucleotide sequence changes are implemented at the genomic locus (Fig. 2.6(2)).

Technique SDN3: sgRNA-Cas9 activity together with a repair template harbouring a cis-, intra-, or transgene is delivered into cells. The repair template consists of a DNA stretch intended for insertion and is flanked on both sides by sequences homologous to the target region. sgRNA-Cas9 activity induces a targeted DSB. In the course of HDR, the DNA to be inserted is precisely inserted at the target site (Fig. 2.6 (3a)). SDN3 thus enables insertion of cis-, intra- or transgenes at specific loci. Furthermore, targeted gene insertion can also be achieved by using the repair pathway of NHEJ (Fig. 2.6 (3b)). In this case, the donor DNA harbouring the cis-, intra-, or transgene to be inserted does not need to be flanked by regions of homology to the target locus.

2.2 Application in plant breeding

In 2012 it was realized that CRISPR-Cas9 provides a valuable addition to already established systems for genome editing [13]. In the meantime further applications other than genome editing and of interest to plant breeding have been proposed. In these potential applications CRISPR-Cas9 is used as a transgenic locus to confer protection of plant virus infection [33-35]. Also, there is ongoing development of CRISPR-Cas modules for endogenous gene expression regulation [36].

2.2.1 Potential applications of SDN1

The SDN1 technique may seem of restricted use in plant breeding in comparison to transgene technology or mutation breeding since traits can mainly be altered by elimination of gene/promoter function. However, metabolic and developmental pathways function as networks and so elimination of gene function can be used to affect traits in a variety of modes, depending on the nature of the pathway and the targeted step, the eliminated gene function (positive/negative regulator) and the overall genetic architecture of the trait (redundant gene function). The SDN1 technique shares trait modification by elimination of gene function with the RNAi technology. Traits that have been engineered before using RNAi technology might now, where sensible, be implemented using SDN1 technology, and further traits beyond these will be modified using SDN1 technology.

2.2.1.1.1 Elimination of unwanted compounds

An apparent SDN1 application is elimination of unwanted compounds. Anti-nutritional compounds can be eliminated or lowered by knocking-out genes coding for enzymes in biosynthetic pathways, for example leading to phytate in maize [37], or to linamarin, a toxic compound in the staple food cassava [38]. In order to engineer food grade oil in rapeseed varieties and in other *Brassica* species, including under-utilised species like *Camelina sativa* [39] or in rapid domestication of wild species like *Thlaspi arvense* [40], low erucic acid and glucosinolate content are breeding goals (00 varieties). Potential SDN1 targets for that for example are *FATTY ACID ELONGASE 1* (*FAE1*) and the transcription factor *HIGH ALIPHATIC GLUCOSINOLATE 1* (*HAG1*), respectively [41, 42]. Low erucic acid and

glucosinolate content are quality parameters for food and feed use in Brassicas: low erucic acid varieties in general are used for production of edible oils, and low glucosinolate content allows use of the seed meal for feed purposes.

Tissue specificity, conferred in RNAi technology by promoters, can be achieved by SDN1 technology through knowledge on tissue specific gene function. In plants, paralogs, gene family members with in some cases exchangeable gene function, are often expressed in a tissue specific manner. The rapeseed (*Brassica napus*) genome encodes three functional paralogs of *FATTY ACID DESATURASE 2*, of which *FAD2-4* is expressed in root and seed tissue only, while *FAD2-1* and *FAD2-2* are expressed ubiquitously [43]. The protein derived from *FAD2* catalyses monounsaturated oleic acids into polyunsaturated fatty acids (PUFA) and for some industrial applications low PUFA content is desirable (f.e. it has higher thermal stability and a longer shelf life). Knocking out specifically *FAD2-4* might be an approach to change the fatty acid profile of rapeseed in specific organs of interest only, maintaining fatty acid metabolism in the remaining tissues. In other cases, tissue specificity may be implemented by affecting transport mechanisms: targeting homologs of *ARABIDOPSIS THALIANA GLUCOSINOLATE TRANSPORTERS 1* and *2* (*GTR1, 2*) in agronomically important *Brassicas*, may be used to lower glucosinolate content specifically in seeds while maintaining glucosinolate production and therefore biological function (protection against herbivory, for example) in source tissues, since *GTR1* and *GTR2* are required for glucosinolate transport into seeds [44]. Another problem, reducing or eliminating allergenic epitopes causing celiac disease might be challenging in the (near) future, since α-gliadin alone is encoded numerous times (at least 40 times without taking into account pseudogenes) in the wheat genome [45] and at the same time gluten is an important quality parameter of wheat. However, eliminating allergenic epitopes of less-abundant proteins eliciting strong response is a feasible breeding goal with SDN1 technology. In soybean, the p34 protein shows low abundancy, but is one of the major soybean allergens [46]. p34 is a member of the papain superfamily of Cys proteases, with as yet no reported enzymatic activity [46], and a BLAST search against the *Glycine max* genome detects few paralogous loci (3-4 loci; assembly V1.0, at EsemblPlants platform). In an RNAi approach p34 downregulated soybean lines were viable and similar in growth behavior in comparison to wild-type plants [46].

2.2.1.1.2 Increasing production of desired compounds

Besides elimination of unwanted products, knock-out of genes using SDN1 technology can also be used to change plant metabolism to enhance production of a desired metabolic product or trait. To revisit fatty acid metabolism in *Brassicas*, deletion of *FAE1* to eliminate erucic acid production at the same time leads to elevated levels of monounsaturated oleic acid content [42]. There is another breeding strategy to increase overall oil content in oil crops: reduction of fruit components (pericarp, testa) not containing oil, like for example hardened ovary tissue protecting the seed in achene fruits

of sunflowers, or for example thick testa tissue in non-oil pumpkins; but also linseed, poppy and rapeseed cultivars with reduced sclerenchymatic tissue exist [47]. The thin testa of Styrian oil pumpkin cultivars is known to be based on a recessive mutation in a major gene [47, 48], and if once mapped and based on a loss of function mutation, the locus may be an SDN1 target to generate a high oil content trait in other cultivars while at the same time maintaining the cultivars favourable genetic background (SDN2 technique might be used in case the recessive allele is a functionally recessive allele). Knock-out of inhibitors of pathways represents an additional strategy to enhance production of traits of interest using SDN1 technology. For example, non-glandular trichome production in *Brassicas* is governed by a suite of activators (certain members of WD40/bHLH/R2R3MYB genes forming a protein complex) and inhibitors (R3MYB) [49]. Knock-out of the latter increases trichome production [50]. Glandular and non-glandular trichome density has been shown to be positively correlated with protection from insect herbivory [51-53]. The genetic basis of trichome production is at least partially conserved across plant genera [54-57] and is starting to be discovered for glandular trichomes in *Cucumis sativus* [58, 59]. Trichomes, in particular glandular trichomes, are also the natural production site of a suite of specialised metabolites across plant species with commercial value (pharmaceuticals (artemisinin), fragrances/flavour (Lamiaceae plant family) or natural pesticides (involvement in pyrethrin biosynthesis) [60-62]). Therefore, a strategy to increase production of valuable trichome derived metabolites might be to increase trichome production by SDN1 targeting of trichome inhibitors, alternatively, SDN1 targeting of trichome activators might be used to generate favourable glabrous vegetable varieties [58]. WD40/bHLH/R2R3MYB complexes together with R3MYB inhibitors are also involved in regulation of the flavonoid biosynthesis pathway [55, 63-66] and production might be enhanced via targeting of the pathway specific R3MYB inhibitor by SDN1.

2.2.1.1.3 Engineering pathogen resistance by targeting recessive resistance genes

Genome editing may be used to target so called susceptibility (S) genes (or recessive resistance genes) [67] to establish lines with biotic stress tolerance (for brief introduction to S genes refer to chapter 4.2.1).

A specific example are the effector targets in rice *Xa13* and *Xa25*/*OsSWEET13* of *Xanthomonas oryzae* pv. *oryzae* (Xoo), which causes bacterial blight. *Xoo* encodes effector genes (transcription activator like effectors (TAL effectors)) that bind to effector binding sites (EBE) in promoter regions and thereby upregulate host target genes in order to promote virulence [68]. Using CRISPR-Cas9 to establish a knock out line for Xa25/OsSWEET13 in a japonica rice line otherwise susceptible to a Xoo strain transformed with a TAL effector designed to target Xa25/OsSWEET13, it could be shown that disease susceptibility was lost [68]; plants were reported to have no obvious detectable phenotype in this study. However, recessive resistance genes are endogenous plant genes with biological functions

in plants, for some of which pleiotropic effects have been reported [69]. One strategy to minimize the effects of engineering pathogen resistance of S genes, that are upregulated upon TAL effector EBE binding, has been shown using another site directed nuclease system, ironically TALEN. In this study, OsSWEET14 was not targeted by TALEN genome editing in the protein coding region to establish a knock out line, but in the EBE site of the promoter region in order to abolish Xoo TAL effector binding, and at the same time retain other OsSWEET14 functions [70]. It could be shown that genome edited rice lines with induced small deletions of 4 or 9 bp in the EBE did not induce OsSWEET14 expression after inoculation with an Xoo strain carrying the avrXa7 TAL effector protein, and displayed a resistance phenotype [70]. They mimic naturally occurring recessive resistance alleles of *Xa13* and *Xa25*/*OsSWEET13,* since these are also not null alleles, but possess polymorphisms in the EBE sequence of the promoter [71]. This study demonstrates that by introducing small, targeted mutations using genome editing valuable traits of use in plant breeding may be engineered.

SDN1 (and SDN2) genome editing techniques provide refined means to plant breeding and complement transgenic technology, traditional and mutation breeding. Conventional mutation breeding programmes offer the discovery of novel, artificially induced, trait variation; further, by TILLING (Targeting Induced Local Lesions in Genomes), mutant populations can be screened for desired variation at a locus of interest. The potential of SDN1 and SDN2 techniques in genome editing is linked to already present and increasing knowledge derived from basic and applied research on molecular variation underlying phenotypic trait expression as well as on gene function and metabolic pathways in general. By using SDN1 and SDN2 techniques, the breeder directly and specifically works with the understanding of molecular variation that has been discovered to underlie phenotypes of agronomic interest. The examples given above provide an overview on the potential use of SDN1 technology to alter traits of interest to plant breeding: removal of unwanted compounds (phytate, glucosinolates), increasing desired compounds (oleic acid, secondary metabolites) or engineering of pathogen resistance by altering recessive resistance genes.

Among other techniques, sequencing technology is generating an enormous amount of information on genomic variation within and between species (150 Tomato Genome ReSequencing Project [72], 3000 Rice Genomes Project [73], 44 sorghum line genomes [74], 302 soybean accessions [75], 115 cucumber lines [76]), which can be mined for meaningful variation of trait expression in phenotypic screens (for developments in high throughput phenotyping see for example [77, 78]; for examples of genome-wide association studies refer to [79-83]).

2.2.2 Potential applications of SDN2

The SDN2 technique can be used similar to SDN1 to translate knowledge on meaningful molecular trait variation into plant breeding programmes. The SDN2 technique, in addition to induce loss of function mutations, is of particular interest to transfer favourable functional molecular variation between cultivars or from closely related (wild) species (allele transfer) but also introducing favourable amino-acid changes deduced from methodological genetic screens into elite cultivars. A recent example is the generation of herbicide resistant maize lines by introduction of specific single nucleotide substitutions in the gene ACETOLACTATE SYNTHASE 2 (ALS2) [84].

2.2.3 Potential applications of SDN3

The SDN3 technique can be used similar to conventional cis-, intra-, or transgenesis technology to insert cis-, intra-, or transgenes into plant genomes, with the difference that the location of insertion can be determined *a priori*. To date, insertion of cis-, intra- or transgenes is largely based on insertion at random genomic loci. Several independent lines need to be screened to select suitable candidate lines which do not show undesired phenotypes because of compromised target sites, for example by gene disruption, and at the same time express the inserted gene in an adequate manner. With the SDN3 technique, insertion can be targeted at a defined locus and possibly take advantage of knowledge about regions of permissive gene expression. With this technique also gene stacking is possible, i.e. introduction of several genes in close proximity. This facilitates breeding programmes in that favourable new traits are not separated in successive breeding cycles and in turn can be easily introduced into further varieties/germplasm segregating as a single-locus trait.

The potential applications of cis- and intragenesis have been described in the study by AGES [3].

2.2.4 Applications other than genome editing

The CRISPR-Cas module provides a programmable tool to target a protein component to defined dsDNA (type II) or ssRNA (type III) sequences. The CRISPR-Cas module thus provides the potential to be used for applications other than genome editing. Both groups of potential applications below involve activity of CRISPR-Cas as a transgenic locus in plant lines. These applications have been reported recently and it remains to be seen whether they develop the potential for use in plant breeding.

CRISPR-Cas9 as a tool for conferring virus resistance in plants. Recently, three independent publications have shown in proof of principle experiments that CRISPR-Cas9 can confer protection against different types of geminiviruses in *N. tabacum* and *A. thaliana* [33-35]. The use of CRISPR-Cas9 for generation of geminivirus resistant crops offers advantages over other strategies (multiplex

targeting, fast response in targeting of newly emerging viral strains), but there are currently still questions to be addressed (off-targeting, selection pressure on virus populations) for its potential deployment as a resistance trait in plant breeding [85].

CRISPR-Cas as a tool for targeted gene expression regulation. A nuclease de-activated "dead" Cas9 (dCas9) alone or fused to effector domains is guided to loci of interest to interfere with (CRISPRi) or activate (CRISPRa) gene expression (reviewed in [36]). The mode of regulating expression is dictated by the specific dCas9 fusion protein and includes steric hindrance of transcription, mediating transcription via activation domains or epigenetic modification. In a proof of principle experiment, dCas9 guided to a reporter locus decreased gene expression in bacteria and human cells [86], and in human cells the repressive effect was enhanced by fusion of dCas9 to the chromatin modifier domain KRAB (Krüppel-associated box) [87]. The KRAB domain guided by dCas9 to HS2, a distal enhancer element of globin genes, efficiently induced histone modifications indicative of closed heterochromatin and at the same time reduced globin gene expression [88]. The feasibility of CRISPRi and CRISPRa, has been demonstrated in several systems [14] and recently also in *Nicotiana benthamiana* [89]. CRISPR-Cas subtypes (for example type III-B) may also be employed for targeted RNA interference in the future [90].

These methods offer interesting alternatives to RNAi based methods in gene expression regulation, however, it remains to be seen whether they will also be applied for plant trait development in plant breeding. They are not further considered in the remaining chapters.

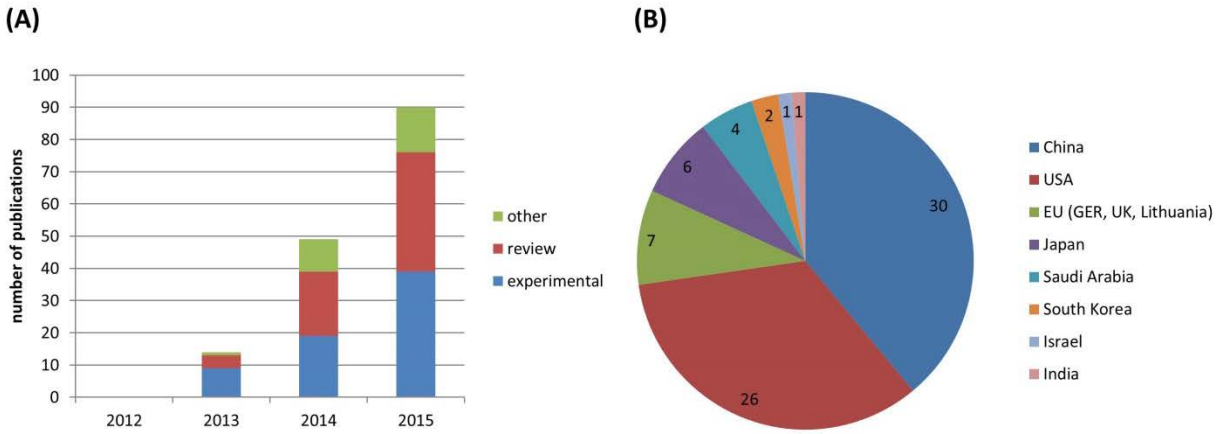
CRISPR-Cas9 has become widely applied also because it has uncoupled the technique of genome editing from know-how intensive protein engineering, as is required in TALEN- or ZFN-based genome editing. Because of that, although first applied in plants in 2013, already a large number of genome edited crop plants have been published in the scientific literature.

CRISPR-Cas9 may be applied in genome editing to introduce targeted mutations and by that engineer, for example, plants with reduced unwanted compounds, increased desired compounds or disease resistance. Traits so conferred via SDN1 technology to date are explored with great emphasis also in crop plants with prospect of applications. Genome editing resulting in accurate site directed insertion of transgenes (SDN3) is promising great strides also in basic plant research, yet still lacks ease in successful implementation. In the future, establishment of commercial cis-, intra-, and transgenic plants may benefit from developments in SDN3 technology.

Other potential applications use CRISPR-Cas9 as a trait conferring virus resistance or as a regulator (positive and negative) of gene expression. These applications would entail the insertion of foreign DNA and therefore generate transgenic plants falling under the current EU GMO legislation.

2.3 State of research and development in plants

In 2013, several research groups reported the first proof-of-principle experiments of CRISPR-Cas9 mediated genome editing in plants [91-100]. Primarily carried out in *O. sativa, N. tabacum* and *A. thaliana*, co-expression of recombinant Cas9 and gRNAs reproducibly induced targeted indels in endogenous genes in cultured cells and *in planta*. Using donor templates homologous to endogenous genes or specially designed reporter constructs it was shown that targeted DSBs carried out by CRISPR-Cas9 can result in HR-mediated repair of genomic regions. Since CRISPR-Cas9 can be easily reprogrammed via the spacer sequence of the gRNA, simultaneous delivery of multiple sgRNAs targeting different loci demonstrated the feasibility of multiplex genome editing as well as of deletion of intervening sequences in these first reports. The number of scientific publications with experimental CRISPR-Cas9 data in the plant field is increasing (Fig. 2.1).



2.3.1 Transferability of the system to plant species

CRISPR-Cas9 technology has since been shown to be transferrable to various crop plants (mostly using SDN1 technique), for example to soybean (*Glycine max*) [101], wheat (*Triticum aestivum*) [98], maize (*Zea mays*) [84], barley (*Hordeum vulgare*) [102], potato [103] and tomato [104] (*Solanum tuberosum* and *S. lycopersicum*), but also for example to tree species, like *Populus* [105] and *Citrus* [106]. By now, it has been established that genome edited sites are stably transmitted to progeny independent of CRISPR-Cas9 presence. In tomato, in a cross of a wild type plant with a bi-allelic genome edited individual, progeny lacking the CRISPR-Cas9 transgene was heterozygous with either one of the two edited alleles in combination with the wild-type allele [104]. Independent transmission has also been analysed and shown for example in *Arabidopsis*, rice, barley and *Brassica oleracea* [107-110] [102].

2.3.2 Techniques (SDN1, 2, 3)

SDN1 and, in extension, multiplexing for simultaneous editing of genes or deletion of genes are the most frequently reported genome editing techniques in the scientific literature, since it involves delivery of the nuclease component only. In wheat, for example, plant lines carrying mutations in *MILDEW RESISTANCE LOCUS* (*MLO*) *A1* have been generated [111], and in rice, knock-out of SWEET13 has proven its function in bacterial blight susceptibility [68]. Intended targeted deletions reported using two DSBs range from small deletions of for example ~50 bp in tomato [104] to ~245 kbp in rice. The latter resulted in deletion of a diterpenoid synthetic gene cluster of ten loci [110], exemplifying the potential to eliminate large genomic regions. Multiplexing ability using CRISPR-Cas9 has been shown for example in rice plants targeting up to 7 [112] and 8 [113] sites simultaneously with different gRNAs, the latter using a specially designed gRNA processing platform, or in *Arabidopsis* using a gRNA with perfect complementarity to two loci [108]. Endo *et al*., exploited the off-target activity of a gRNA targeted at CYCLIN DEPENDANT KINASE 2 (CDKB2) to generated rice lines edited at 2 further gene family members, CDKA2 and CDKB1 [114].

There are fewer reports on the SDN2 and SDN3 techniques. They, together with delivery of CRISPR-Cas9, provide templates for HR-mediated repair to either modify a locus (SDN2) or insert a cis-, intra- or transgene (SDN3). That both techniques are feasible using CRISPR-Cas9 technology in plants has been shown for example in maize, soybean and rice [84, 101, 115]. In maize, endogenous ACETOLACTATE SYNTHASE 2 (ALS2) has been converted into a sulfonylurea herbicide resistant allele by site specific modification of a proline to a serine (P165S) via SDN2; additionally, a PHOSPHINOTHRICIN ACETYLTRANSFERASE (PAT) gene driven by a constitutive promoter was site specifically inserted via SDN3. Genome edited plants transmitted the modifications into the subsequent generations. The modified ALS2 (P165S) gene conferred herbicide resistance in two tested generations [84]. In soybean, ALS1 was modified similarly (P178S) to confer herbicide tolerance via SDN2 and a hygromycin phosphotransferase (HPT) gene linked to an endogenous soybean promoter was targeted for insertion to a specific locus [101]. The SDN3 technique has further been demonstrated in *Arabidopsis* [116], and in tomato by targeting a strong promoter (CaMV 35S) in front of an anthocyanin biosynthesis gene resulting in accumulation of pigments [22]. In the study in *Arabidopsis*, one targeted insertion event has been reported with perfect repair as intended, in the study in tomato, in addition to a perfectly repaired insertion event, an event with nucleotide substitutions was recovered.

2.3.3 Delivery methods

The main delivery method of CRISPR-Cas9 reported for production of genome edited plants involves transformation of a CRISPR-Cas9 gene cassette integrated on vector systems with selectable marker

genes into cultured plant cells. The presence of marker genes allows regeneration of plants stably transformed with the CRISPR-Cas9 construct and subsequent screening of a reduced number of plants for genome edited individuals. Since the presence of the CRISPR-Cas9 transgene is not necessary and may lead to off-target effects upon retention in plant lines, it can be segregated from the intended genome modification in subsequent generations in sexually propagating crop species (for example see [107-110] [102]).

Two delivery methods of CRISPR-Cas9 into plant cells independent of DNA transfer were reported. Analogously as shown with TALEN and meganucleases [117], pre-assembled CRISPR-Cas9 ribonucleo-protein particles were delivered directly into plant cells [27]. In proof-of-principle experiments in protoplasts derived from *A. thaliana*, tobacco, lettuce and rice, genome editing was detected by this delivery method. Regenerated lettuce individuals transmitted the modified allele into the next generation [27]. A bottleneck for general application of this strategy is the ability to regenerate plants from protoplasts which is not a well-established procedure in different plant species. An alternative reported strategy uses delivery by an RNA virus [24, 25]. *Tobacco rattle virus* (TRV) has a bipartite positive strand RNA (TRV1 and TRV2) genome of which TRV2 can be modified to harbour foreign genes, which is commonly exploited in different viral systems in virus induced gene silencing (VIGS). A gRNA driven by a pea early browning virus (PEBV) promoter and targeting PDS was cloned into TRV2 and was agro-inoculated together with TRV1 into *N. benthamiana* transgenic lines stably expressing Cas9. Cas9 expression from the *N. benthamiana* genome was necessary because of limited capacity of the viral genome to harbour foreign genes and retain functionality. Gene editing at PDS took place in agro-infiltrated and in systemic *N. benthamiana* leaves [24, 25] and the edited PDS allele was transmitted into the next generation [25]. Limiting factors for a broader application using DNA free delivery by RNA viruses are the small capacity of the viral genome, the varying host range of viruses and systems to obtain virus-free genome edited progenitor plants (REF).

2.3.4 Types of mutations generated by SDN1 technique

Datasets describing the type of mutations generated by SDN1 technique are reported mainly for *Arabidopsis*, rice and soybean [107, 112, 118-125]. The mutations arise during repair by the endogenous DNA repair pathway of DSBs, mainly NHEJ in somatic cells [21].

In *Arabidopsis* and rice, based on to date available data, the most frequently detected mutations are insertions of a single adenosine or thymidine nucleotide, followed by small deletions of predominantly one nucleotide and deletions of <10 nucleotides [107, 112, 119, 120, 123-125]. Other detected mutations are nucleotide replacements and insertion of >1 nucleotides, but to a lesser extent. Based on the data available at present from *Arabidopsis* and rice, the mutation spectrum may be generalised over experimental systems, mutations detected in protoplasted cells, transgenic lines

generated by floral dip transformation (*Arabidopsis*) or somatic embryogenesis after agro-inoculation (rice). In soybean, the most frequently detected mutations were deletions <10 nucleotides [118, 121, 122]. There is the indication that dependent on the sgRNA or the targeted locus the mutation spectrum may differ in some instances; in the study of Jacobs *et al*., one sgRNA induced predominantly single nucleotide insertions, independently of the experimental system (soybean hairy root and somatic embryogenesis) [121]. Similar observations were made in rice [124, 125]. The location of the generated mutations predominantly occur starting three nucleotides off the PAM in the proto-spacer sequence (for example [118-121]).

2.3.5 Off-target activity

Recognition of the target site, the so called protospacer, by a CRISPR-Cas9 complex is guided by two different signatures, the presence of a protospacer adjacent motif (PAM) and the complementarity of the spacer sequence in the sgRNA to the protospacer sequence [13, 126] (Fig. 2.3). The PAM is present at the genomic target site directly 3´ to the protospacer and is recognized by the PAM interacting domain of Cas9 protein. *Streptococcus pyogenes* Cas9 (SpCas9) recognises the PAM sequence 5´-NGG-3´ and, with less efficiency, 5´-NAG-3´ [126]. A systematic analysis of CRISPR-Cas9 target specificity found that the 8-12 nucleotides proximal to the PAM (called seed region, (Fig. 2.3)) are on average less tolerant to mismatches than the distal region [126]. The efficiency of perfectly matched CRISPR-Cas9 modules in DSB induction was analysed over three datasets in a study by Xu *et al*., [127]. They find and confirm [128-130] that DSB induction efficiency is dependent on several features, for example nucleotide composition in the spacer region (where some nucleotide positions influence Cas9 gRNA loading) or the influence of nucleotide positions 3´downstream of the PAM (i.e. outside of the protospacer region). These, as well as for example structural features of the gRNA backbone [131] influence CRISPR-Cas9 efficiency and thus also contribute to off-target activity. There is ongoing research into specificity and efficiency which will be implemented in genome editing systems in the future, particularly in metazoan systems, since specificity and efficiency are highly critical parameters for potential therapeutic applications of CRISPR-Cas9.

Upon application of CRISPR-Cas9 in plant genome editing, characterization of off-target activity was also of interest in plant species. Table 7.1 (Appendix) lists studies which report analyses of off-target activity in plant cells. A set of 15 randomly mutant gRNAs were tested against a target locus in wheat suspension culture cells [98]. Similarly to the above studies in bacterial and human cells, mismatches at the distal region (non-seed region) were rather tolerated than mismatches proximal to the PAM, which often abolished DSB formation [98]. Two studies, in *Arabidopsis* and rice, report whole genome sequencing (WGS) data of CRISPR-Cas9 genome edited plant lines in order to survey genome-wide possible off-target effects [107, 132]. In rice, Nipponbare plant lines each transgenic

with one of 6 different gRNAs did not reveal significantly higher SNP and indels than wild type controls when compared to the rice reference sequence; in a Kasalath background a comparison was hampered by the large difference of sequencing depth between the wildtype and transgenic lines [132]. In *Arabidopsis*, comparison of 3 WGS datasets of CRISPR-Cas9 transgenic plant lines (harbouring the same gRNA) did not show an increased SNP or indel number in comparison to wild type controls when mapped to the Col-0 reference genome [107]; Off-target activity in the remaining studies was analyzed to different degrees and with different methods. While in some studies off-target sequence searches were carried out systematically by BLASTN searches, in some cases supported by available software programmes [29, 30], which search for and (CRISPR-P) rank potential off-target sites based on an experimentally derived score, other studies chose loci based on prior knowledge of sequence homology. Detection of off-target activity was either based on sequencing methods, restriction enzyme/PCR methods (PCR/RE) [96, 100] or enzyme mismatch cleavage methods [133]. In a study in soybean, 10 potential off-target sites, with varying degrees of mismatches distributed over the protospacer region, were tested and off-target effects were not detected [134]. While in this dataset a target with only three mismatches in the distal region was not targeted, another gRNA exhibited off-target activity at a locus with two mismatches in the proximal region in the same study [134]. Similar results were obtained in rice: testing 3 gRNAs on altogether 13 potential off-target sites identified one off-target locus harbouring 1 mismatch in the distal region [132] and testing 4 different gRNAs on the highest ranked potential off-target sites by CRISPR-P, off-target activity was detected at one site with one mismatch in the distal region [109]. Other potential off-target sites in these two studies harboured 3-7 and 2-4 mismatches, respectively, with varying distribution over seed and non-seed region sites [109, 132].

Based on the published data, specificity of CRISPR-Cas9 in plant cells seems to be governed by the same parameters as in other eukaryotic systems. While the majority of detected off-targets harboured mismatches in the distal region, there were also exceptions to this rule (see for example [134]).

In summary, CRISPR-Cas9 mediated genome editing has been shown to be transferable to diverse plant species. Initial proof-of-principle experiments established that all techniques (SDN1, 2, 3) are principally feasible, the highest number of publications to date report application and development of SDN1. Increasingly, there are research publications using CRISPR-Cas9 as an alternative to conventional methods in reverse genetics to analyse gene function [68, 135-137], indicating general acceptance as a validated and efficient method in plant science. Ongoing research and development is focused on establishing efficient genome editing platforms and vector systems for diverse species, and on the development of delivery modes (including those without DNA transfer). The ability to specifically modify target sites offers an alternative and site-directed mode to create variability for plant breeding. Prior knowledge of a gene

function and its physiological or phenotypic effect on plant traits can thus be implemented in a specific and efficient manner into plant breeding programmes.

2.3.6 Limiting off-target effects

There are several strategies to limit off-target effects. Hsu *et al*., postulated a set of rules to guide gRNA selection and there are several software applications to automate gRNA selection (for example [29, 30]).

Experimental strategies include the application of paired nickases or RNA-guided FokI nucleases (reviewed in [138]). Paired nickases are Cas9 proteins with introduced point mutations destroying either one of their two endonuclease domains. The resulting proteins introduce single stranded DNA breaks (nicks), and targeting two complementing paired nickases properly spaced to the same locus, a DSB is generated. At the same time, specificity is increased since two spacer sequences are needed for induction of a DSBs. Paired nickases have been used in proof-of-principle experiments in plants [119]. RNA-guided FokI nucleases confer specificity by the same principle and are based on gene fusions between dCas9 and a dimerization dependent *Flavobacterium okeanokoites* (Fok1) nuclease (reviewed in [138]).

Recently, mutant Cas9 proteins, [139] and SpCas9-HF1 [140] have been shown to confer higher specificity to the CRISPR-Cas9 complex in human cells by weakening non-specific interactions of Cas9 with its target; in the case of eSpCas9 interaction with the non-complementary target strand, in the case of SpCas9-HF1 four aa substitutions were introduced to weaken non-specific interactions of Cas9 with the target strand. Since specificity is of high importance for therapeutical applications of CRISPR-Cas, further strategies and/or mutant versions of CRISPR-Cas may be developed which may also be utilized then in plant applications.

2.4 Intended and unintended effects of CRISPR-Cas9 in genome editing

The intended effect using CRISPR-Cas9 in genome editing is the targeted site specific modification of a target locus and thereby changing expression of trait(s) modulated by that locus. Intended genetic modifications have been categorized by the NTWG (national experts nominated by the Competent Authorities of EU Member States) as site specific random mutations (SDN1), site specific modifications (SDN2), and site specific insertion of cis-, intra-, and transgenes (SDN3) [2]. Furthermore, in multiplexing, targeting of several loci or deletion of regions in between may be the intended goal.

A potential unintended effect due to application of the CRISPR-Cas9 technology in genome editing is off-target activity by placing of DSBs at loci with imperfect complementarity to the spacer sequence.

This might lead, depending on the SDN technique, to either (i) induction of random mutations at off-target loci, to (II) deletion of genomic fragments, (iii) integration of cis-, intra-, or transgenes at unintended loci or (iv) a combination of those.

Potential unintended effects by means of using transgenic CRISPR-Cas9 intermediate lines may be (i) retention of the transgene in resulting organisms and (ii) generation of background mutations due to the performed transformation process, which are passed on to resulting organisms. An unintended effect due to the use of viral vector systems is viral contamination of progeny.

2.5 Safety considerations

2.5.1 SDN1 technique in genome modification of plants

The SDN1 technique targets specific loci to introduce mutations of *a priori* unknown sequence changes. Intended changes mostly are loss of function mutations of genes or regulatory elements, since these are most likely generated using this technique. In general, the SDN1 technique introduces small insertions, small deletions or nucleotide replacement mutations at a site or sites near the PAM in the protospacer sequence. However, also larger deletions or insertions may arise. When targeting two CRISPR-Cas9 modules on the same chromosome, it is also possible to generate deletions of the genomic region in between the two DSBs. The specificity and therefore the amount of DSBs induced in the genome is determined by the spacer region.

Provided that the resulting plants do not carry a CRISPR-Cas9 module stably integrated in the established plant line, the SDN1 technique can therefore be compared to conventional physical and chemical mutagenesis techniques based on intended and unintended changes.

Genome edited, transgene-free resulting plant lines may be established, for example, (i) by selecting null-segregants of transgenic plants, (ii) in cases where a CRISPR-Cas9 DNA module had been transiently transformed and (iii) in cases where ribo-nucleoprotein complexes have been directly introduced.

2.5.1.1 Comparison of CRISPR-Cas9 and conventional mutagenesis techniques in relation to mutational load and type of modifications

Physical (for example gamma ray, X-ray) and chemical (for example ethyl methanesulfonate (EMS)) mutagenesis is used to induce variation in plants to generate mutants for conventional plant breeding. There are 3,220 mutant cultivars, in over 210 species [141], collected in the worldwide Mutant Variety Database (MVD, FAO/IAEA)[2] which have been officially and/or commercially

[2] https://mvd.iaea.org/

released. Tables 2.1 and 2.2 list studies reporting on induced genetic variation after chemical (EMS and NaN3/MNU) and physical (gamma irradiation) mutagenesis. EMS is an alkylating agent (predominantly of guanine) resulting in SNPs by changing G/C nucleotides into A/T nucleotides [142] and gamma irradiation is suggested to induce DSBs resulting in diverse mutation categories [143, 144]. The amount of induced genomewide mutational events varies, among other things, with dose and concentration of physical and chemical mutagens, respectively, but is also dependent for example on the treated tissue. Common mutation densities/effects are given in Table 2.1 and Table 2.2 for reported chemical and physical mutagenesis experiments. In TILLING datasets typical mutation densities are between 1 mutation /100 – 500 kbp (higher densities are typically present in polyploid species since they are able to buffer deleterious mutations), but may also lie outside these ranges depending on the TILLING population (see for example [142] for an overview). These mutation densities translate into several hundreds of genomewide mutations per individual in the "smaller" genomes of soybean and rice (Table 2.1), and in around 340,000 mutations per individual in wheat. Reported effects of gamma irradiation in wheat leads to estimates of around 82-110 gene deletions per individual; in a study in rice, with a high irradiation dosis, it was estimated that 9% of the genome was altered (Table 2.2). For breeding purposes, there is a trade-off in mutational density, since on the one hand the lower, the larger the population to be screened for desired genotpyic and/or phenotypic mutants, but on the other hand a large mutational load potentially affecting many other loci is undesired; depending on the species and propagation system, mutagenised individuals are either directly or indirectly (as part of breeding programmes) used for establishing commercial cultivars [145].

Whole genome sequencing of genome edited rice and *A. thaliana* lines did not suggest a genomewide elevated mutational increase when compared to control plants in the datasets of these two studies (see chapter 2.3.5; [107, 132]). In contrast to chemical and physical mutagenesis, CRISPR-Cas9 does not randomly (genomewide) induce mutagenesis events, but is restricted to the target loci and to potential off-target loci with a certain amount of sequence complementarity (see chapter 2.3.5; see selected examples of studies reporting off-target effects for soybean, rice and barley in Table 2.3). This is reproduced *in planta*, for example in the study of Endo *et al*., 2015 off-target effects were detected at two loci ranked as most likely candidates by the software CRISPR-P, however, no mutations were detected at loci ranked 3rd, 5th, 9th and 10th likely to be targeted [114]. Zhang *et al*., 2014 report similar results: while for two gRNAs off-target effects could not be detected at 5 and 3 candidate loci, one gRNA lead to off-target effects in 1 out of 5 candidate off-target loci; the effected locus showed 1 mismatch in the non-seed region in comparison to the intended target, while the remaining potential off-target loci differed at 4, 6 or 7 positions and were not targeted in this experiment [132].

The type of mutations generated by application of CRISPR-Cas9 based genome editing have been summarised in Chapter 2.3.4: small insertions (<10 nucleotides), small deletions and nucleotide replacements have predominantly been detected at sites targeted by a given gRNA. Depending on the particular mutagenic agent used, conventional mutagenesis generates for example predominantly substitutions in the case of chemical mutagenesis using EMS (Table 2.1), while for gamma irradiation substitutions, indels and copy number variations were reported for example in rice (Table 2.2). However, also in the case of EMS mutagenesis the isolated mutation used for breeding may be based, for example, on a deletion. Natural variation, including natural variation found in domesticated species, is based on the same mutation categories: for example, in a study resequencing landraces, wild progenitors and improved imbreds of *Sorghum bicolor* [74], nucleotide replacements, indels, copy number variations and larger deletions leading in some cases to gene loss can be detected.

In comparison to conventional mutation breeding techniques, the CRISPR-Cas9 SDN1 technique induces specific mutations at intended loci and potentially a smaller number of further off-target loci that can be predicted to a certain extent. This also reflects the difference in intended use of these techniques in breeding applications. Thus, the (random) unintended mutational load of CRISPR-Cas9 genome edited plants is much smaller in comparison to conventional mutation breeding methods, based on available datasets.

Generally, for plant breeding applications, CRISPR-Cas9 specificity is of importance, however, since during plant breeding practices often several generations are passed with selection based on phenotype and/or genotype and there is the possibility of backcrosses, off-target effects are tolerable and can be removed (analogous to classical mutation breeding), in contrast to therapeutic genome editing applications.

2.5.1.2 Safety considerations in respect to CRISPR-Cas9 transgene retention, background mutations caused by transformation procedures and the use of viral vectors

For safety considerations based on the above, please refer to chapter 3.4, since these are covered also in the context of rapid-cycle breeding.

2.5.2 SDN2 technique in genome modification of plants

The SDN2 technique targets specific loci to introduce mutations of *a priori* known sequence changes. For that, together with the CRISPR-Cas9 module DNA repair templates are co-transformed that are identical in sequence to the targeted locus with the exception of the intended sequence changes. In a certain proportion of cells these are used as templates by the HR repair pathway of the CRISPR-Cas9 induced DSB and thus the changes are implemented at the targeted locus.

For SDN2, the same applies in regard to unintended mutational load as for SDN1.

The repair template may be integrated as a whole at the locus with the targetd DSB for example by the NHEJ repair pathway, as well as at sites in the genome. Analysis of genome edited plant lines for ectopic integration of cisgenes can be done by standard methods (Southern Blot, PCR based methods) and plant lines without ectopic integration can be selected accordingly.

2.5.3 SDN3 technique in genome modification of plants

Safety aspects of of cis- and intragenic plants have been covered in comparison to transgenic plants in the study of AGES [3] and in a Scientific Opinion by EFSA [146]. In contrast to conventionally generated cis-, intra-, and transgenic plants, the SDN3 technique is used to insert DNA at *a priori* intended loci. Safety aspects concerning impairment of endogenous genes and creation of novel reading frames can therefore be already addressed at the development phase of the plant line.

2.6 Detection and identification

It is to be expected that genome edited plant lines free of CRISPR-Cas9 transgenes will be established, where feasible due to the production and/or the breeding process. In cases where a CRISPR-Cas9 module is present in the established genetically modified plant line detection and identification rationale follows those of conventionally transgenic plants. The CRISPR-Cas9 transgenic sequence in combination with its genomic integration location then provides a marker for GM detection and event-specific identification.

2.6.1 Detection and identification of SDN1 and SDN2 genome editing

CRISPR-Cas9 generates random site directed mutations, small insertions/deletions, larger deletions and nucleotide substitutions (SDN1), and mediates incorporation of *a priori* designed mutations, mostly small insertions/deletions and/or nucleotide substitutions (SDN2).

The quality of SDN1 and SDN2 mutations do not allow conclusions on their origin

Nucleotide changes (down to single nucleotide polymorphisms (SNPs)) are detectable by standard PCR based, hybridization based or sequencing methods [147]. The induced genomic changes cannot be distinguished from naturally occurring variation or from changes derived from conventional mutagenesis (see chapter 2.5.1.1). Therefore, the presence alone of a mutation at a genomic site cannot be causally linked to it being generated by the application of CRISPR-Cas9 technology. Circumstantial evidence based on background markers may be used for identification of a genome editing event. In case a particular mutation of a genome editing event is described in combination with marker states of the background genome of the plant line in which it was generated, these in

combination may be used to indicate the probability of the origin of a mutation (and therefore identification) in a sample. However, the use of the genome edited line in breeding programmes will break up linkage to those background markers and therefore decrease or abolish evidence of the origin of the mutation.

It is to be expected that genome editing will be targeting various loci inducing different site directed random mutations or modifications (like conventional mutagenesis). As a consequence, a general screening strategy for the detection of mutations derived from SDN1 and SDN2 techniques would have to include a combination of many tests, rather than few universal tests to collectively cover several events.

In case prior knowledge of induced mutations is absent, detection and identification is technically impracticable.

2.6.2 Detection and identification of SDN3 genome editing

Detection and identification of SDN3 genome editing follow the same principle as for conventionally generated transgenic plants. For cis- and intragenic lines the detection step, i.e. the detection of distinct sequences indicating cis- or intragenic status in a general screening step, is made more labour-intensive because of sequence homology of inserted sequences to endogenous genes (as discussed by AGES for conventionally generated cis- and intragenic plants [3]). Use of event-specific analyses (identification) provides unambiguous evidence of cis- or intragene presence or absence.

Genome modifications generated by SDN1 and SDN2 genome editing techniques can be detected, however, their presence does not provide evidence on how they originated: they cannot be distinguished from naturally occurring variation or mutations derived from conventional mutagenesis.
Genome modifications generated by SDN3 carry a cis-, intra-, or transgene, therefore, detection and identification is analogous to conventionally established cis-, intra-, or transgenic plants.

2.7 Aspects of GMO classification of CRISPR-Cas9 genome edited plants

Directive 2001/18/EC provides a definition of GMO (Annex 7.2). This report provides information on the CRISPR-Cas9 technology and its application in genome editing in plants: (i) a description of the origin and molecular mode of action of CRISPR-Cas9 (chapter 2.1.1), (ii) a description of the different types of genetic modifications possible to generate in plants (chapter 2.1.3) and (iii) an overview on production processes to obtain genome edited plants (chapter 2.1.2). By that, it covers potentially relevant aspects to classification according to Directive 2001/18/EC, which are summarised briefly for each technique below.

A functional CRISPR-Cas9 entity is a ribonucleo-protein complex that can be programmed to target certain genomic locations where it induces a DSB in the targeted DNA sequence. A DSB is repaired by endogenous cellular repair mechanisms and gives rise to an *a priori* unknown mutation. There are nucleotide changes that occur preferentially, driven by the mode of action of the endogenous repair mechanism active, and possibly dependent on the plant species and the nature of cells in question. Targeting of the CRISPR-Cas9 module is mediated by the RNA component and may lead to unintended off-target effects at genomic locations with sequence complementarity to the so called spacer sequence. If, in addition to the CRISPR-Cas9 module, DNA molecules are transformed into cells, they can be employed to achieve further modes of genome editing (SDN2, SDN3).

Genetic modifications possible to generate using ZFN technology and in extension other SDN technologies have been categorized into three classes (SDN1, SDN2 and SDN3; Fig. 2.6) by the New Techniques Working Group (NTWG) [2]. SDN1 generates site directed random mutations, SDN2 site directed intended (*a priori*) mutations and SDN3 inserts cis-, intra-, or transgenes at the targeted genomic locus. In addition, due to ease of multiplexing ability, CRISPR-Cas9 can also be used to generate small or large deletions at targeted genomic locations (subsumed under SDN1 technique based on similarity of the production process (see also Study on behalf of BAFU (Federal Office for the Environment, Switzerland) [148]). In contrast to other SDN technologies (ZFN, TALEN, MN), a functional CRISPR-Cas9 module consists of a protein and an RNA component.

SDN1, the targeted mutation of a locus with *a priori* unknown sequence change, is a form of mutagenesis using an organic particle as mutagenic substance. During the process, DNA coding for a CRISPR-Cas9 module (or, in case of multiplexing, two or more modules) may be employed for delivery of a CRISPR-Cas9 module into cells, and depending on the production process, it may be integrated as a transgenic locus in the genome. The intended heritable genetic modification is independent of the CRISPR-Cas9 module, therefore, mutant plant lines devoid of the transgenic locus can be selected in subsequent generations in sexually propagated species. The CRISPR-Cas9 transgene has thus been present in individuals (intermediate organism) during the production process, but is not present in the final established genome edited plant line (resulting organism). Alternatively, RNA coding for the sgRNA and Cas9 or ribonucleo-protein complexes can be delivered into cells as such.

SDN2, the targeted mutation of a locus with a priori intended sequence change, requires additional delivery of a DNA fragment into cells, which is used by endogenous repair pathways as a template, and by that incorporates the intended genomic modification at the targeted locus. The nucleotide sequence of the repair template is identical to the target locus with exception of a single or a small number of nucleotide sequence changes or small deletions or insertions of a few nucleotides. As for SDN1, the intended mutation is independent of the presence of the CRSIPR-Cas9 module and the

same delivery methods as for SDN1 apply for the CRISPR-Cas9 module. Similar to SDN1, it is a mutagenesis technique involving a CRISPR-Cas9 module and a repair template leading to *a priori* intended mutations.

SDN3 intends to insert a cis-, intra-, or transgene at a targeted genomic locus and requires co-delivery of the DNA sequence to be inserted. In contrast to SDN1 and SDN2 it does not aim at modification of an endogenous genomic locus, but at precise integration of extra-genomic sequences. Similar to SDN1 and SDN2, the intended sequence insertion is independent of the presence of the CRSIPR-Cas9 module and therefore, plant lines carrying the insertion but lacking a CRISPR-Cas9 module can be generated. SDN3 generated plant lines are similar to cis-, intra-, or transgenic plant lines, however the extra-genomic sequence has been inserted at a targeted genomic locus mediated by a DSB introduced by a CRISPR-Cas9 module.

2.7.1 Evaluation of ZFN and related genome editing techniques by the German expert commission ZKBS

The position statement [7] of the ZKBS (Zentrale Kommission für die Biologische Sicherheit) includes the assessment of ZFNs, and as noted in their position statement, their assessment can be extrapolated to other DSB-producing site directed endonucleases (SDN). In their statement they provide conclusions on their interpretation of the term GMO in Directive 2001/18/EC in relation to SDNs:

In relation to delivery of ZFNs, in their opinion, type B intermediate organisms (i.e. organisms with transiently present recombinant DNA which has not been chromosomally integrated) do not fall under the definition of Directive 2001/18/EC. Further, if ZFNs are delivered by isolated mRNA or as isolated proteins, in their opinion they are not covered by the GMO definition of Directive 2001/18/EC since there was no heritable genetic material involved in the production process.

In relation to resulting organisms, plants derived by ZFN1 and ZFN2 techniques are assessed by the ZKBS as not falling under the 2001/18/EC GMO definition. They remark for ZFN1 that the resulting organisms carry mutations generated with involvement of the endogenous mechanism of NHEJ and that the same mutations may be generated by natural processes as well as by conventional mutagenesis breeding.

For ZFN2, in their opinion organisms altered by the size of 20 or less nucleotides do not fall under the definition. This is based on the notion that the genetic difference between the co-delivered repair template and the endogenous to be edited gene in that case does not represent a “recombinant nucleic acid”.

Organisms resulting from ZFN3 are falling, in their opinion, under the definition of GMO given in Directive 2001/18/EC.

While SDN3 techniques generate cis-, intra-, or transgenic plants falling under the EU GMO definition (Directive 2001/18/EC), there is legal uncertainty whether genome modified plants resulting from SDN1 and SDN2 techniques do so as well.

SDN1 and SDN2 technique lead to plants with targeted introduced mutations. In the process of establishing SDN1 and SDN2 genome edited plants intermediate plants may be generated that stably integrate a CRISPR-Cas9 transgene. In sexually propagated crops, the transgene and the intended genome modification can be separated resulting in progenitors with the genome modification but not possessing any transgene. Furthermore, techniques delivering CRISPR-Cas9 into cells without transfer of heritable, genetic material are being developed.

A national expert group in Germany (ZKBS) published a position statement, in which they conclude that in their opinion resulting genome edited plants without a stably integrated transgene do not fall under the EU GMO definition (based on ZFN mediated genome editing).

Directive 2001/18/EC explicitly excludes plants generated by conventional mutagenesis breeding and plants generated by cell or protoplast fusion, as well as does not consider plants generated by polyploidy induction; plants generated by these techniques are exempted from the risk assessment and regulatory procedure established by Directive 2001/18/EC that – based on the precautionary principle – has the objective to protect human health and environment.

Directive 2001/18/EC therefore implicitly states that the risks associated with intended and unintended mutations by the exempted techniques (mutagenesis breeding, cell culture methods and bringing together related genomes or multiplication of genomes), are considered to be manageable outside the regulatory procedure of Directive 2001/18/EC, that is by the breeding practices implemented by breeders. This is based on the considerations that the Directive should not apply to techniques of genetic modification which have conventionally been used and have a long safety record (recital 18 of the Directive). From a scientific view, the mutations – intended and unintended - generated by SDNs in (cis-, intra-, and transgene free) genome edited plants are not qualitatively different from plants arising from natural mutation events or generated by breeding practises not falling under Directive 2001/18/EC. With respect to the quantity of mutations, genome editing induces a minimal number of mutation events, i.e. far less than induced by *e.g.* chemical mutagenesis breeding (typically 100s to 1000s).

2.8 Tables

Table 2.1 Reported chemical mutagenesis effects on plant genomes in TILLING projects

Species	Mutagen (concentration)	Nr of loci analysed	Size of M2 population	Mutation density (1 mut/kbp)	Predominant type of mutation	Reference	Derived average genomewide nr of mutations per individual*
Glycine max (Forrest)	EMS (40 mM)	7	529	1/140	G/C>A/T	Cooper *et al*., 2008 [149]	~ 7900
Glycine max (Williams 82)	EMS (40 mM)	7	768	1/550	G/C>A/T	Cooper *et al*., 2008 [149]	~ 2000
Oryza sativa (japonica Nipponbare)	EMS (1,5%)**	10	768	1/294	G/C>A/T	Till *et al*., 2007 [150]	~ 1300
Oryza sativa (japonica Nipponbare)	NaN3/MNU (1mM/15mM)	10	768	1/265	G/C>A/T	Till *et al*., 2007 [150]	~ 1400
Triticum aestivum	EMS (0,8%)***	3	512	1/47	G/C>A/T	Chen *et al*., 2012 [151]	~ 340,000 (Chen *et al*., 2012) [151] per individual

*based on haploid genome size of ~1115 Mb [152, 153] and ~389 Mb [154] of *Glycine max* (Williams 82) and *Oryza sativa* (*japonica* Nipponbare), respectively.
** ~145 mM
*** ~77 mM
EMS: ethyl methanesulfonate; NaN3-MNU: sodium azide-methyl nitrosourea

Table 2.2 Reported gamma irradiation effects on plant genomes

Species	Dosis (Gray)	Subject of study		Mutation type				Mutation detection	Reference	
				SNPs	Indels (1-5 bp)	Copy number variations	Presence/absence variation			
Oryza sativa (9311)	300	Red-1 (M6 inbred line)	9.19 % of genome altered	381,403*	50,116	1,279	10,026	Solexa whole genome sequencing	Cheng *et al*., 2014 [155]	
Synthetic wheat SW58	350/450	1,510 DGRH1 individuals	2 % marker loss in D genome	nd	nd	nd	nd	35 SSR marker	Kumar *et al*., 2012 [156]	
Triticum aestivum (Chara)	50	4500 M2 individuals	nd	nd	nd	nd	4 confirmed gene deletions (homozygous) of TaPFT1-D across M2 individuals 3 confirmed gene deletions (homozygous) of TaPFT1-A across M2 individuals	Hybridisation based qPCR specific for homeolog deletion detection	Fitzgerald *et al*., 2010 [157]	derived: on average 110/82 gene deletions (homozygous) per M2 individual **

*validation of SNPs by Sanger sequencing: 60/63 true

**based on the assumption of 124,000 genes (A,B,D) in *Triticum aestivum* [158]

DGRH1: D genome radiation hybrid panel; SSR: simple sequence repeat.

Table 2.3 Excerpt of Appendix Table 5.3 : off-target identification of CRISPR-Cas9. Studies/experiments with detected off-target effects are coloured green.

Target locus	Off-target candidate locus identification	Nr. of mismatches distribution	Method of detection	Off-target activity detected	Experimental system	Reference
G. max						
07g14530	BLASTn (e value threshold 5) 10 candidate loci	2-6 mm Distributed	Amplicon sequencing (n=10 biological replicates)	none detected	soybean hairy root system	[134]
DDM1 gRNA1	BLASTn (e value threshold 5) 1 candidate loci	4 mm Distributed	Amplicon sequencing (n=10 biological replicates)	none detected	soybean hairy root system	[134]
DDM1 gRNA2	BLASTn (e value threshold 5) 1 candidate locus	2 mm seed region	Amplicon sequencing (n=10 biological replicates)	Yes, in all experimental repeats	soybean hairy root system	[134]
Met1	BLASTn (e value threshold 5) 1 candidate locus	3 mm Distributed	Amplicon sequencing (n=5 biological replicates)	none detected	soybean hairy root system	[134]
miR1514	BLASTn (e value threshold 5) 2 candidate loci	6 and 2 mm Non-seed region	Amplicon sequencing (n=4 biological replicates)	yes, gRNA with 2 mm in non-seed region in all experimental repeats	soybean hairy root system	[134]
H. vulgare						
HvPM19-1	2 candidates based on homology	1 mm in seed region each	Sequencing in 93/95 T1 individuals of two independent T0 lines	Yes, gRNA with mm (further away from PAM than 2nd off-target) in seed region, 3/93 individuals	stable transformation	[102]
HvPM19-3	2 candidates based on homology	1mm in seed r. 3 mm distributed	Sequencing in 76 T1 individuals of one T0 line	None detected		[102]
O. sativa						
DERF1	Selected based on homology 5 candidates	3-5 mm 2 only in non-seed region	Sequencing at target locus in 20 GE lines (T0 and T1, all independent lines)	none detected	stable transformation	[132]
MYB1	Selected based on homology 3 candidates	3-5 mm 2 only non-seed region (5 mm)	Sequencing at target locus in 20 GE lines (T0 and T1, all independent lines)	none detected	stable transformation	[132]
YSA1	Selected based on homology 5 candidates	1-7 mm 2 only non-seed region (1 and 7 mm)	Sequencing at target locus in ~70 Cas9 positive lines (independent T0 lines)	Yes, at 1 candidate locus 7 plants with off-target activity: locus with 1 mm in non-seed region	stable transformation	[132]

Target locus	Off-target candidate locus identification	Nr. of mismatches distribution	Method of detection	Off-target activity detected	Experimental system	Reference
AOX1a	CRISPR-P Selected 2 highest ranked	3, 4 mm distributed	Sequencing of target locus	none detected (50 plants of T0 and T1)	stable transformation	[109]
AOX1b	CRISPR-P Selected 2 highest ranked	3, 4 mm distributed	Sequencing of target locus	none detected (49 plants of T0 and T1)	stable transformation	[109]
AOX1c	CRISPR-P Selected 2 highest ranked	2, 3 mm distributed	Sequencing of target locus	none detected (60 plants of T0 and T1)	stable transformation	[109]
BEL	CRISPR-P Selected 2 highest ranked	1 mm non seed r. 3 mm distributed	Sequencing of target locus	Yes, activity detected in 2 plants at locus with 1 mm (89 plants of T0 and T1)	stable transformation	[109]
CDKB2	3 candidates selected based on homology, confirmed by CRISPR-P as among possible targets (rank 1, 2, 10) Further 3 candidates ranked 3, 5, 9 by CRISPR-P	1 mm non seed r. 2 mm seed/non-seed r. 3 mm seed/non seed r.	CAPS marker, sequencing CAPS marker	Yes, activity detected (6/13 regenerated plants) Yes, activity detected (10/13 regenerated plants) none detected (0/13): mm nearest to PAM (all regenerated plants from 1 transformation event (callus); conclusion repeatable in 3 further transformation events (calli)) none detected	stable transformation	[114]

3 Accelerated breeding – rapid cycle breeding

3.1 Introduction

Accelerated breeding, also termed rapid cycle breeding, is a technique to shorten the duration of breeding programmes. Specifically of interest in species with long generation times, as in perennial, woody plants (shrubs, trees), it is achieved by establishing plant lines carrying transgenes that confer a dominant precocious flowering phenotype. These lines are used as crossing partners to shorten the individual breeding cycles. At the end of the breeding process, individuals carrying the desired trait/trait or trait/genomic background combinations but lacking the early flowering transgene are selected for further propagation (Fig. 3.1) [159].

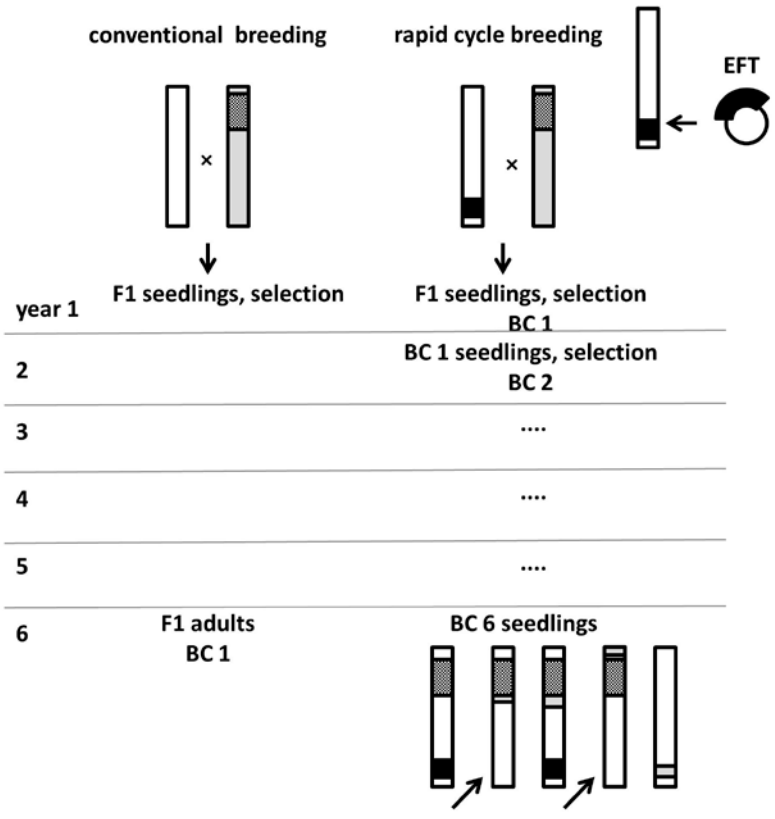


Fig. 3.1 Conventional versus rapid cycle breeding timeframes (after [159]). Conventional breeding cycles in apple may take 6-12 years. In rapid cycle breeding, first an early flowering transgenic cultivar is established and may be used for different breeding objectives, here the introgression of a desired trait from a wild relative. The transgenic line dominantly inducing early flowering is crossed with the wild relative and backcrosses of selected individuals can be carried out after shortened cycles. At a cycle where individuals carrying the trait of interest in a domestic apple background are present, individuals lacking the early flowering transgene are selected for further propagation (arrow). BC: backcross; EFT: early flowering transgene; F1: hybrid.

In addition to the above, currently it is also explored to cause precocious flowering transiently in each generation by viral induced gene expression/silencing or by grafting on transgenic rootstock [160, 161]. Naturally occurring genetic diversity (or induced by conventional mutagenesis) may be used for the same purpose, however, as of yet, there is a lack of suitable precocious flowering mutants in perennial species [159].

The juvenile phase, per definition the vegetative phase in which plants are not competent to flower independent of otherwise favourable environmental conditions, can for example, last up to 6 - 12 years for apple and pear in field conditions (see Table 1 in [162]) and so is a major determinant of generation time. The timing of flowering in plants is coordinated by an extensive gene network: it is depending on environmental and autonomous signals and is altogether suppressed in plants going through juvenile phases [163]. An increasing number of flowering time regulators are uncovered, several of which were tested for their potential in rapid cycle breeding in diverse species (Fig. 3.2). The key to successful application of rapid cycle breeding in the context of a given plant species/cultivar lies in identification of suitable candidate genes that shorten the juvenile phase and at the same time retain proper floral organ development and fertility; *Arabidopsis thaliana LEAFY* (*AtLFY*), for example, induces early flowering in a citrus hybrid (*Citrus sinensis* × *Poncirus trifoliata*) [164] but not in an apple cultivar (*Malus* × *domestica* cv 'Pinova') [165]. *BpMADS4*, a *FRUITFUL* (*FUL*) homolog from birch, is used for accelerated breeding programmes in apples [166] and poplar *FLOWERING LOCUS T1* (*PtFT1*) in plums [167].

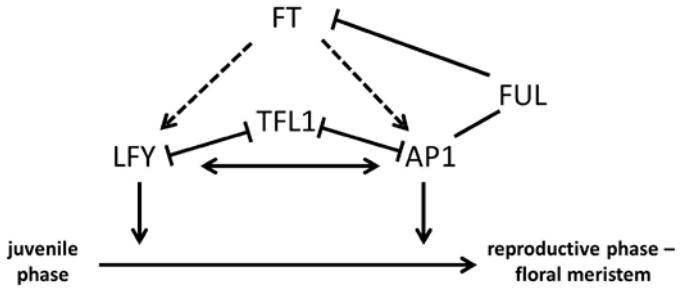


Fig. 3.2 Fraction of the gene network regulating juvenile to reproductive phase transition in *Arabidopsis thaliana* (extracted from [163]). Homologs of for example *FT* and *FUL* are used in rapid cycle breeding programmes [168, 169]. Arrow: activation; broken arrow: indirect activation; bar head: repression; line: interaction with unknown direction. *AP1*: *APETALA1*; *FT*: *FLOWERING LOCUS T*; *FUL*: *FRUITFUL*; *LFY*: *LEAFY*; *TFL1*: *TERMINAL FLOWER1.*

Synonyms used are high-speed breeding, fast breeding, FasTrack (fast track) breeding and rapid cycle breeding; it was agreed upon using the term rapid cycle breeding in the future [166].

3.2 Potential applications in plant breeding

Plant breeding in species with long juvenile phases, such as in, for example, shrubs and trees, is a time consuming process. Juvenile phases of apple and plum cultivars (or wild relatives) last between 5-12 [162, 170] and 3-7 years [167], respectively; during this time flower formation is suppressed. A central role in plant breeding play controlled crosses between varieties within a species or to related species, and depending on the breeding goal may involve several cycles of successive crosses. Rapid cycle breeding, in establishing transgenic lines with reduced juvenile phases, has the potential to reduce breeding programmes in that the crossing cycles are shortened in time.

In apple, disease resistance germplasm is also present in wild relatives [171] (fire blight: *Malus fusca [168]*, *Malus robusta* [172]; apple scab: *Malus floribunda* [173], *Malus sieversii* [174]) and markers tagging resistance genes are being developed [172, 174]. If these disease resistance gene resources from wild apple relatives are to be used by introgression into *M. domestica*, several successive cycles of pseudo-backcrosses need to be done to re-establish the *M. domestica* background genome (on average, 5 backcrosses lead to < 2% of the related species in the background genome [175]). Many of today´s scab resistant cultivars rely on Rvi6/Vf mediated resistance which was derived from the wild relative *Malus floribunda*, with initial hybridization crosses tracing back to 1914 [173, 176] and it taking several decades to establish elite cultivars carrying Vf resistance genes [177]. In apple, rapid cycle breeding programmes based on a transgenic early flowering line have been established. One breeding goal is to introgress the apple scab resistance from *Malus fusca*; generation cycles reported lasted ~ 12 months [166, 176].

Successive crosses are also needed when pyramiding genes of interest in cultivars: it is known for disease resistance that when based on a monogenic trait in combination with widespread use the possibility of resistance breakdown increases. For example, there are sporadic observations that Rvi6/Vf resistance has been overcome by a *Venturia inaequalis* strain (causative organism of apple scab), but the virulence gene has not spread through the *V. inaequalis* population due to pathogen management [173]. Therefore, breeding goals are to pyramid multiple resistance loci in a cultivar or breed for quantitative resistance, i.e. several genes underlying the resistance trait, as well as cultivars carrying resistance genes against diverse pathogens, by carrying out crosses with appropriate breeding partners. In recent years, in addition to rate breeding offspring phenotypically, also for perennial species marker assisted selection (MAS) has become feasible by establishing an increasing number of markers tagging major QTLs underlying traits of interest for breeders (markers established in apple can for example be found in [178]) and MAS applications are further being developed (RosBREED programme, USA [179], FruitBreedomics project (EU FP7 funded [180])). At the same time genome databases have begun adding genome sequences and assemblies also of perennial species,

like fruit trees (apple, pear, peach, orange [181-184]), which in the future brings the potential to integrate a large array of markers into breeding processes (for reviews and opinions please refer to [185-187]). Rapid cycle breeding, together with MAS, has the potential to support and initiate breeding programmes in perennials by reduction of time and cost of infrastructure [185-187] for current breeding goals such as disease resistance breeding, low allergenicity apples and quality traits underlying processing and fresh-cut market requirements [188].

Particularly in perennial species there is often low genetic diversity present in commercially used cultivars, because of time and costs associated with breeding; breeding is then often based on crosses between a few successful cultivars, as well as mutation breeding and spontaneous mutations ('sports') often contributing to cultivar development [186]. As a consequence for example in apple, although there is a high number of germplasm accessions and a large genepool in related species, only a small number of genotypes have been used for commercial development in the last century [189, 190]. Rapid cycle breeding may thus also contribute to increase genetic diversity in commercial species by making crossing cycles manageable in a reasonable timeframe.

Rapid cycle breeding, by shortening breeding programmes in species with long generation times, may contribute to resistance breeding in for example fruit trees, and in general may increase the number of breeding programmes. By that, it may increase genetic diversity in available germline used for breeding and establishing commercial cultivars in species with otherwise often narrow genetic breeding material.

3.3 State of development

3.3.1 Species of interest and genes tested for precocious flower induction

One of the earliest reports on precocious flower initiation induced by transgenesis is based on the *A. thaliana* gene *LEAFY* (*LFY*) and its constitutive expression in its own genomic background as well as in hybrid aspen (*Populus tremulus × tremuloides*) in 1995 [191]. To date, homologs of at least 5 genes involved in juvenile – reproductive phase transition and/or floral meristem initiation have been shown to be able to induce precocious flowering in certain woody species when overexpressed, *AP1* (*APETALA1*), *FT* (*FLOWERING LOCUS T*), *LFY* (*LEAFY*) and *MADS4* (a FRUITFUL homolog), or downregulated (via RNAi), *TFL1* (*TERMINAL FLOWER1*), depending on the regulatory function in the genetic network (see Table 3.1 and references therein; Fig. 3.1). The research focus for applications in breeding is in woody, perennial species; the most scientific publications can be found for apple and poplar, followed by citrus. Single studies can also be found for birch, eucalyptus, pear and plum in the scientific literature. However, for plum there has been set up a rapid cycle breeding program in a

collaboration of the USDA/ARS and several US based Universities (see below). Furthermore, there is one study published in soybean (*Glycine max*), an annual plant with a range of maturity groups describing the duration of the vegetative phase [192]. In this report mid to late maturity types were induced to flower after 35 – 45 days post inoculation, half the time than control plants, independent of photoperiod conditions which was discussed to be of potential interest also for soybean breeding [192].

Early flowering in most cases was induced in a certain proportion of independent transgenic lines in a given study, possibly depending, among other reasons, on the locus of transgene insertion. Furthermore, the ability of a certain transgene to induce early flowering in a species may be dependent on the respective genetic backgrounds (compatibility of origin of transgene (species, cultivar) and target genetic background (species, cultivar), as for example *Arabidopsis thaliana LEAFY* (*AtLFY*), induced early flowering in a citrus hybrid (*Citrus sinensis* × *Poncirus trifoliata*) [164] and in hybrid aspen [191] but not in transgenic apple lines (*Malus* × *domestica* cv 'Pinova') [165]. Where reported and where early flowering was successfully induced, flowers were fertile (Table x.1). Several studies report a certain extent of morphological/developmental deviations of floral organs in comparison to wild type flowers (for example in plum [193], poplar [194], apple [195]), the extent depending on the species and the transgenic strategy used to induce early flowering. In some cases this leads to reduced fertility (for example reported in relation to breeding program in apple [166]). Crosses performed (see Table 3.1 and references therein) lead to viable offspring in citrus, apple, plum, pear and poplar. Data from, for example, plum [193] and apple [195] show expected segregation pattern of progeny for the early flowering transgene (Table 3.1).

3.3.2 Experimental systems to induce precocious flower induction

In most studies stably transformed transgenic lines were generated to test the potential of transgenes to induce an early flowering phenotype and which potentially might be used as a breeding partner. The expression of the transgene conferring early flowering was mostly driven by a constitutive CaMV 35S promoter (Table 3.1). There are several studies using heat inducible promoters (heat shock promoter (HSP) from *Glycine max*) for expression of the early flowering transgenic traits in order to minimize the effect of the transgenes during plant development at times where transgenic activity is not needed [194, 196-201]. For use of the inducible system, regimens of heat treatment had to be established, in order to induce gene expression but at the same time maintain plant habitus and meristem viability. Use of the inducible system to induce early flowering was successful for example in poplar [194] and apple [200] but not in the study of Weigl *et al*., 2015 [197] due to the negative effect of heat treatment on flower formation. Transformation was mostly carried out by *Agrobacterium* mediated transformation, the tissue transformed and the cell culture

procedures of generating stable transgenic lines differ depending on established protocols in each species.

In apple, three related studies report apple latent spherical virus (ALSV) vector system to drive gene expression for precocious flower induction in apple [160, 161, 170] and one study in pear [202]. ALSV is a member of the genus Cheravirus which are bipartite (+) ssRNA viruses [203]. A vector system for ALSV is established [204]. ALSV inoculation of plants can be carried out without DNA transfer, however, ALSV has been shown to be seed transmissible [170]. In the study in pear a virus elimination procedure based on growth at high temperature has been established [202]. ALSV driven precocious flowering was also used in the study with soybean where there was a certain percentage of seed transmission [192].

Several studies (mostly in apple, one study in poplar; Table 3.1) tested whether the induced early flowering phenotype is graft transmissible, i.e. whether wild type plants (acting as scions) also are induced to flower precociously after grafting onto rootstock of transgenic early flowering lines [168, 194, 195, 200, 205]. Although in a study of transgenic poplar mRNA of AtFT could be detected in the scion [194], graft transmissibility of the phenotype, precocious flowering, based on FT, TFL1-RNAi or MADS4 could not be shown in any of the studies. FT is a compelling candidate for testing graft transmissibility since it has been shown to be part of the systemic flower inducing “florigen” signal; movement of both, FT mRNA and protein, has been implicated in florigen activity [206].

3.3.3 Current rapid-cycle breeding programmes

Based on the work cited above breeding programmes in apple and plum have been established (Table 3.2), both with involvement of Federal Research Agencies.

Example apple

In the published rapid-cycle breeding programmes in apple the goal is to pyramide disease resistance genes, both from wild apple species and domestic apple cultivars, into domestic apple to generate commercial cultivars [166, 168, 207, 208]. Federal Agencies involved are the Julius Kühn-Institut, Germany, and Agroscope, Switzerland. The breeding is built on a transgenic early flowering line (due to transformation with BpMADS4) of the cultivar ‘Pinova’ (T1190) which was crossed to *Malus fusca* to introgress fire blight resistance. Since markers are not established for the *Malus fusca* fire blight resistance, F1 individuals were screened phenotypically for resistance. Resistant individuals carrying the transgenic precocious flowering locus were then crossed to lines with (i) known scab and fire blight resistance loci (*Rvi2*, *Rvi4, FB-F7*; cv ‘Regia’) or (ii) powdery mildew resistance loci (*Pl1*, *Pl2*; germplasm 98/6-10) followed by a pseudo-backcross to ‘Golden Delicious’ to continue introgression

of the resistance loci into commercially used background germplasm (refer to breeding scheme in Fig. 5 in [168]). Individual breeding cycles were realized within a year [168].

Based on the line T1190 F1 crosses were carried out also with the ornamental apple 'Evereste' coding for a strong fire blight resistance locus (Fb_E), followed by pseudo-backrosses to various commercial *M. domestica* cultivars [207, 208]. Some BC2 individuals carried already less than 15% of background genome of the 'Evereste' while maintaining the Fb_E resistance locus [208].

The breeding programmes are made difficult by the small number of offspring as a by-product to the precocious flowering phenotype, and growth conditions were being adapted, as well as suitable crossing partners (age of wild type crossing partner) chosen [166, 168, 207, 208]. T1190 line was chosen for the breeding program because its precocious flowering phenotype is based on a single transgene insertion which was mapped to linkage group 4 (LG4) [168]. Further transgenic *M. domestica* early flowering transgenic lines to be used for breeding programmes were subsequently established which each carry the transgenic construct on different LGs in several different commercial cultivar backgrounds [166]. This ensures the presence of a diverse set of crossing partners for breeding programmes which often aim for introgression of loci present on different LGs and pyramiding of traits of interest in plant lines.

The published data show that the combination of (i) rapid cycle breeding and (ii) marker assisted selection (to optimize choosing of offspring for subsequent crosses in relation to desired trait and background genome) is a feasible breeding strategy that greatly reduces breeding time in woody species.

Example plum

A "FasTrack" breeding programme in plum is carried out in a collaboration of the United States Department of Agriculture (USDA) Agricultural Research Service (ARS) with University of California Davis, Clemson University and Pennsylvania State University, US [209-211] [3]. It is based on continually flowering transgenic plums that have been generated by stably introducing *FT1* from *P. trichocarpa* driven by the CaMV 35S promoter into diverse genomic backgrounds. A patent has been granted in plums for this system in the United States [212]. A continually flowering plum line of the cv 'Blubyrd' has been published [167]. Supported by the California Dried Plum Board (State of California), a breeding goal is to breed plum varieties suitable for dried plum production in California [210]. For that, a panel of different cultivars/germplasms have been selected for transformation with PtFT1 in order to generate FasTrack crossing partners with a range of desirable traits (for example differing harvest times, sugar content, good dried appearance and flavour). Specifically, one short

[3] http://ucanr.edu/sites/fastrack/Approach/

term goal is the introgression of the transgenic plum pox virus (PPV) resistance trait of 'Honeysweet' (a fresh market plum) into the genetic background of the dried plum cultivar 'Improved French', which is the main planted dried plum cultivar in California. The transgenic PPV resistance trait of 'Honeysweet' (event C5) is approved for cultivation and food use in the US by APHIS, FDA and EPA [211] (the application to APHIS (petition 04-264-01p)[4] contains data from experimental field tests collected in three European countries [211]). Once the PPV transgenic trait is introgressed after several cycles of backcrossing into the genome of 'Improved French', null segregants for the early flowering trait PtFT1 will be selected for potential commercial cultivation. Null segregants derived from FasTrack Breeding are not regulated by the USDA[5]. In 2013, BC1 individuals were reported to be germinated for a further cycle of backcrossing [210]. A long term goal in the breeding program is to understand, using molecular markers, high fruit sugar level, which is based on complex genetic architecture. Established markers will then be used to breed elite dried plum cultivars using the FasTrack system[1].

3.3.4 Establishing infrastructure for rapid-cycle breeding programmes

To optimize rapid-cycle breeding programmes for a given species, it is of interest to generate a panel of independent precocious flowering lines with mapped and characterized transgene locations, each carrying a single transgene on a different linkage group. Furthermore, established lines ideally maintain high fertility and exhibit a plant habitus supporting fruit growth [166]. Known insert location facilitates breeding processes because the breeder can choose suitable breeding partners depending on the breeding goal. For example, in apple if the breeding goal was to introgress a locus of interest with known linkage group location from a wild relative into domestic apple, it is of advantage to choose a breeding partner which carries the early flowering transgene on a non-homologous chromosome or as far apart as possible on the homologous chromosome [168]. If they are located on the same homologous chromosome BC1 progenies inherit both traits only in case of crossing over taking place. The closer the loci are located to each other, the smaller the number of individuals in the progeny carrying both traits. The same applies at completing the breeding process, since the early flowering transgene needs to be segregated away from the introgressed locus to generate resulting organisms which are null-segregants for the transgene.

Therefore, Weigl *et al*., 2015 [166] established several transgenic early flowering lines with transgene insertion sites at different genomic locations and in different cultivars. Initially, transformed individuals were screened for lines carrying single T-DNA insertions by Southern blotting. Insertion sites were identified by genome walking and verified by PCR assays [166]. Similarly, in the dried plum

[4] https://www.aphis.usda.gov/biotechnology/petitions_table_pending.shtml
[5] USDA/APHIS response to Letter of Enquiry by USDA/ARS

breeding program it was planned to establish several independent flowering lines in various germplasms chosen based on presence of traits of interest [6] in order to ensure a set of breeding partners.

The basis for application of rapid cycle breeding, establishing transgenic lines with precocious flowering behaviour has been achieved in several perennial species, for example apple, poplar, citrus, pear or plum. Optimal precocious flowering lines for breeding programmes selected need to retain fertility, and need to be characterised for insert number and genome location. Breeding programmes have been established in apple and plum. In apple, a major breeding goal is to combine disease resistance loci in a commercial cultivar background.
Furthermore, it is explored to induce precocious flowering using transgenic rootstock as well as transient induction using viral vectors.

3.4 Intended and unintended effects

The intended effect of using a transgenic, precocious flowering breeding partner is to shorten breeding program durations by decreasing the time between successive crosses may be carried out. To date, precocious flower initiation is induced in breeding programmes by using breeding partners carrying dominant transgenes, furthermore, applied research explores precocious flower induction by (i) grafting scions onto transgenic (i.e. harbouring precocious flower induction locus) rootstocks, and (ii) transiently expressing transgenes (conferring precocious flower induction) using viral vector systems. At the end of the breeding process, resulting individuals carrying the desired trait/trait or trait/genomic background combinations but lacking the early flowering transgene (in case of using transgenic breeding lines) are selected for further propagation (Fig. 3.1).

Unintended effects by means of using transgenic plant lines as breeding partners may be (i) retention of the transgene in resulting organisms and (ii) background mutations in the transgenic precocious flowering lines due to the performed transformation process, which are passed on to resulting organisms. An unintended effect due to application of rapid-cycle breeding in the case of using viral vector systems is viral contamination of progeny by seed transmission.

Unintended effects caused by the novel combination of different genomic backgrounds due to the breeding process are not unique to or caused by application of rapid-cycle breeding and may occur as in conventional breeding programmes.

[6] http://ucanr.edu/sites/fastrack/Approach/Obj1System/

3.5 Safety considerations

Retention of transgene in resulting organism

Resulting organisms in a rapid cycle breeding programme are selected based on the desired trait/trait or trait/genomic background combination analogously to conventional breeding programmes, additionally, resulting organisms lacking the precocious flowering phenotype conferring transgene are selected. Transgenic lines generated for rapid cycle breeding are evaluated for transgenic state (insert number, location) since it is integral to an efficient breeding programme to use well characterised transgenic lines. Presence/absence of the transgene is monitored during the rapid cycle breeding process to ensure the use of appropriate breeding partners (refer to chapter 3.3.4). Standard PCR techniques are used to map transgene integration sites and used to confirm presence/absence of the transgene; Southern blotting is routinely used to analyse transgene copy number.

For confirmation of transgene absence in resulting organisms, PCR techniques, Southern Blotting and/or genome sequencing using next generation technologies [213] may be used.

Background mutations caused by the transformation procedure elsewhere in genome

Experimental procedures during establishment of transgenic lines may lead to mutations elsewhere in the genome. In relation to partial/additional transgene copies the above considerations apply.

Background mutations may be silent as well as non-silent in regard to changes in the expression of the genome. In the latter case, mutations may have beneficial or adverse effects, or may be neutral. Unintended, unknown mutations similarly arise in conventional and mutation breeding. The transgenic line is used as an initial breeding partner to introduce the precocious early flowering transgene, and breeding programmes often involve several successive cycles of crosses. Therefore, background mutations arisen from the transformation procedure are diminished at each cycle (on average by half, with exception of mutations linked to the transgene) in the case transgenic lines are not used in successive cycles.

Viral contamination

Precocious flower formation may be induced using viral vectors (see chapter 3.3.2). Viruses may be passed on through seeds with a certain degree of transmissibility [214]. There have been several strategies of viral elimination established (heat treatment or chemical treatment, passage through tissue culture; [215]). For example, in the framework of establishing induction of precocious flower formation using the ALSV vector system in apple and pear, it has been shown that heat treatment might be an effective strategy to obtain viral free plants [202]. To control for viral contamination,

therefore, elimination procedures exist and/or may be developed for the specific virus/plant species combination and viral absence in resulting organisms can be confirmed using standard DNA or protein based methods [214].

3.6 Identification and detection

Rapid-cycle breeding uses intermediate plants with precocious flower formation to shorten the crossing cycles within breeding programmes. Intermediate plants with the precocious flower formation phenotype may be transgenic plants. In that case, the transgenic locus in combination with its genomic integration location provides a marker for GM detection and event-specific identification.

Individuals meeting the breeding goal, achieved by the conventional breeding process of crossing selected breeding partners, and at the same time being null-segregants for the precocious flowering transgene are selected for further propagation. Therefore, the resulting organism does not carry a transgene and cannot be detected or identified as being generated by rapid-cycle breeding by means of DNA marker based methods.

Similarly, in the case (i) transgenic rootstock is used to induce precocious flowering in the scion or (ii) organisms transiently expressing information of precious flower formation (VIGE or VIGS vectors that are not seed transmissible, other transient expression systems) is used to induce precocious flower formation, the resulting organism does not carry a transgene and cannot be detected or identified as being generated by that process by means of DNA marker based methods.

Intermediate plants may carry a cis-, intra-, or transgene therefore, detection and identification is analogous to conventionally established cis-, intra-, or transgenic plants.
Resulting organisms which do not carry a cis-, intra-, or transgene are not distinguishable to organisms resulting from conventional breeding programmes.

3.7 Aspects of GMO classification

Directive 2001/18/EC contains a definition of organisms falling under the authorization procedure (refer to Annex 7.2 for definition). This report provides information on rapid cycle breeding: (i) a description of the underlying principle (chapter 3) and (ii) a description of an ongoing breeding program in *Malus domestica* (chapter 3.3.3). By that, it covers potentially relevant aspects to classification according to Directive 2001/18/EC.

Rapid cycle breeding uses transgenic organisms during breeding programmes (intermediate organisms). The transgenic locus induces precocious flower formation and thereby shortens crossing

cycles within breeding programmes. Individuals of each generation segregate for the transgene in combination with genomic marker states of interest. Individuals are evaluated in terms of genotype and phenotype in each generation. Those meeting the selection criteria for desired marker states and/or phenotypes, but lacking the transgene, may be generated at certain cycles of the breeding program. The resulting organisms represent non-transgenic (null-segregant in relation to the transgene) individuals which have been passed through a breeding program using transgenic crossing partners.

Alternatively, precocious flowering may be induced by (i) grafting scion onto transgenic rootstock (which has not yet been successfully shown to induce early flowering in the scion) and (ii) by viral induced gene expression/repression. Grafting using transgenic rootstock in plant breeding in general has been covered in a report of AGES to the BMG [4]. Grafting and viral induced gene expression/repression, in case an RNA virus is used as vector, induce precocious flowering transiently in the scion and the transfected plant, respectively.

3.7.1 Evaluation of a related breeding practise by the German expert commission ZKBS

The method of rapid cycle breeding has not been analysed by the ZKBS [7].

From the aspect of the use and the state in respect to the transgene of the resulting individuals, the use of a transgene in rapid cycle breeding may be compared to that in reverse breeding. The transgene, i.e. the transgenic line, in both breeding approaches is used as a tool, not as a trait or as a breeding goal. Transgenic lines are crossing partners to shorten the individual crossing cycles (by conferring the trait of precocious flowering) in a breeding program, which follows conventional breeding goals to generate novel recombined genomic states by crossing of selected breeding partners. When achieving the breeding goal, null segregant individuals for the transgene conferring precocious flowering are selected from the breeding population.

The position statement of the ZKBS concludes on steps in reverse breeding that may be used analogously for evaluation of rapid-cycle breeding (and possibly other techniques using transgenic lines as breeding partner intermediates in the future). In rapid cycle breeding intermediate organisms are used with a precocious flower initiation phenotype that may be generated via different strategies. To date, mostly lines carrying a transgene conferring precocious flower production are used. Intermediate organisms with a stably integrated transgene are assessed by ZKBS as falling under the GMO definition of Directive 2001/18 EC by the ZKBS (here in relation to transgenic intermediates generated for suppression of meiotic recombination in reverse breeding). Furthermore, intermediate organisms exhibiting precocious flower production may be created by

transient transgene expression, i.e. no stable integration of transgenes into the genome, for example by viral induced gene expression. In the case viral vectors that are not seed transmissible are used, in an analogous situation for reverse breeding (recombinant DNA is present only transiently in the intermediary organism and is not passed to its progeny), the ZKBS assesses these intermediary organisms as not falling under the GMO definition of Directive 2001/18 EC, however they may contain a GMO (recombinant virus). Precocious flowering may also be conferred by grafting a scion onto transgenic rootstock. The ZKBS assesses progeny of these chimeras as not falling under the GMO definition of Directive 2001/18 EC.

In respect to resulting organisms, in all three (non-exhaustive) breeding strategies (precocious flower formation by transgenesis, grafting on GM rootstock and virus induced gene expression of non seed-transmissible virus) progeny is generated that does not carry recombinant DNA, i.e. the trait of precocious flowering information. In case of using transgenic lines to confer precocious flower formation, null-segregants are selected among the progeny. An analogous situation in reverse breeding is assessed as not falling under the GMO definition of Directive 2001/18 EC by the ZKBS.

Rapid cycle breeding uses transgenic intermediate plants to shorten the individual crossing cycles. At a generation yielding plants with the desired breeding goal, individuals harbouring the desired genotypes but lacking the transgene are selected.

While intermediate transgenic plants fall under the EU GMO definition (Directive 2001/18/EC), there is legal uncertainty whether plants resulting from rapid cycle breeding and lacking a transgene do so as well.

A national expert group in Germany (ZKBS) published a position statement, in which they conclude on an analogous case, transgene free plants resulting from reverse breeding, that in their opinion these do not fall under the EU GMO definition.

As covered in the chapter of CRISPR-Cas9, Directive 2001/18/EC implicitly states that the risks associated arising from intended and unintended mutations by exempted techniques of mutagenesis breeding, cell culture methods and bringing together related genomes or multiplication of genomes, are considered to be manageable outside the regulatory procedure of Directive 2001/18/EC, that is by the breeding practices implemented by breeders.

From a scientific aspect, the mutations – intended and unintended – generated or introduced in (cis-, intra-, and transgene free) plants resulting from rapid cycle breeding are not qualitatively different than to resulting plants generated by breeding practises not falling under Directive 2001/18/EC.

3.8 Tables

Table 3.1 Studies reporting on genetic engineering for precocious flowering in woody species

Species/cultivar	Precocious flower induction transgene*	Trait donor	Reference	Precocious flowering detected **	Fertility
Betula pendula (birch)				Juvenile phase under natural conditions: 10-15 years (Elo *et al*., 2007) [216]	
Betula pendula 'BPM2' (early flowering clone) 'JR1/4', 'K1898'	BpMADS4	Betula pendula	Elo *et al*., 2007 [216]	Yes (11 days versus 85 days non transgenic control) Yes (86 days after rooting)	not reported
Citrus				Juvenile phase under natural conditions: 6-20 years (Pena *et al*., 2001) [164]	
Citrange Citrus sinensis × Poncirus trifoliata	AtLFY	*A. thaliana*	Pena *et al*., 2001 [164]	Yes (6 T0 lines (out of 22) between 2 and 20 months)	fertile, F1 progeny with early flowering phenotype
Citrange Citrus sinensis × Poncirus trifoliata Sweet orange C. sinensis	AtAP1	*A. thaliana*	Pena *et al*., 2001 [164] Cervera *et al*., 2009 [217]	Yes (2 T0 lines (out of 12) after 13 and 15 months)	fertile, F1 progeny with early flowering phenotype
Poncirus trifoliata	CiFT	Citrus unshiu	Endo *et al*., 2005 [218]	Yes (T0: 12 weeks – 8 months after transfer to greenhouse F1: 2 weeks)	fertile, F1 progeny with segregating early flowering phenotype
Eucalyptus				Juvenile phase under natural conditions: 1-7 years (Klocko *et al*., 2015) [198]	
Eucalyptus grandis × urophylla	AtFT HSP::PtFT1	*A. thaliana* P.trichocarpa	Klocko *et al*., 2015 [198]	Yes (1-5 months after transplanting to glasshouse)	fertile, viable F1 generation
Malus domestica (apple)				Juvenile phase under natural conditions: 5-12 years (Yamagashi *et al*., 2014); 6-12 years (Weigl *et al*., 2014); 4-8 years (Kotoda *et al*., 2010) [166, 170, 219]	
Malus × domestica	MdTFL1	M. domestica	Kotoda *et al*.,	Yes (8 months)	fertile, seed production

Species/cultivar	Precocious flower induction transgene*	Trait donor	Reference	Precocious flowering detected **	Fertility
'Orin'	RNAi		2006 [220]		
Malus × domestica 'Pinova'	BpMADS4	B. pendula	Flachowsky *et al*., 2007 [221]	Yes (3-4 months)	fertile, seed production
Malus × domestica 'Holsteiner Cox','Gala'	MdTFL1 RNAi	M. domestica	Szankowski *et al*., 2009 [222]	Yes (6 months)	not reported
Malus × domestica 'JM2'	MdFT1	M. domestica	Kotoda *et al*., 2010 [219]	Yes (2-6 months after regeneration)	not reported
Malus × domestica 'Pinova'	MdFT2	M. domestica	Traenkner *et al*., 2010, 2011 [205, 223]	Yes (already during *in vitro* cultivation) signal not graft transmissible	not reported
Malus × domestica 'Pinova'	AtLFY	*A. thaliana*	Flachowsky *et al*., 2010 [168]	early flowering phenotype not detected (7 transgenic lines)	early flowering phenotype not detected (7 transgenic lines)
Malus × domestica	MdTFL1 RNAi viral expression system #	M. domestica	Sasaki *et al*., 2011 [160]	yes (1.5 – 2 months after virus-inoculation of seedlings)	fertile, viable seed production
Malus × domestica 'Fuji', 'Orin', 'Golden Delicious'	AtFT MdFT1 viral expression system #	*A. thaliana* M. domestica	Yamagashi *et al*., 2011 [161]	yes (1.5 – 2 months after virus inoculation of seedlings) not detected	fertile, viable seed production F1 generation virus free
Malus × domestica	AtFT & RNAi MdTFL1-1 or MdTFL2	*A. thaliana* M. domestica	Yamagashi *et al*., 2014 [170]	yes (1.5-3 months after virus inoculation of seedlings)	fertile, viable seed production

Species/cultivar	Precocious flower induction transgene*	Trait donor	Reference	Precocious flowering detected **	Fertility
	(combined) MdFT1 or MdFT2 & MdTFL1 viral expression system #	M. domestica		not detected	virus seed transmissible (detected in part of F1 lines), possibly cultivar dependent
Malus × domestica 'Holsteiner Cox', 'Gala', 'Galaxy', 'Pinova'	MdTFL1 RNAi	M. domestica	Flachowsky *et al.*, 2012 [195]	Yes (6 months; preliminary data: signal not graft-transmissible)	fertile, F1 progeny with segregating early flowering phenotype
Malus × domestica 'Pinova'	HSP::PtFT1 HSP::PtFT2	P. trichocarpa	Wenzel *et al.*, 2013 [200, 201]	Yes (6 days after 28 day heat treatment; Signal not graft-transmissible (although PtFT RNA could be detected in scion in one case))	fertile, seed production
Malus × domestica 'Pinova', 'Gala'	HSP::MdTFL 1-1,2 RNAi (same construct as in Flachowsky *et al.*, 2012)	M. domestica	Weigl *et al.*, 2015 [199]	heat treatment abolished floral organ formation	heat treatment abolished floral organ formation
Populus (poplar)				Juvenile phase under natural conditions: P. tremula 7-10 years (Hoenicka*et al.*, 2012) [197]	
P. tremula × alba (f) P. tremula × tremuloides (m)	AtLFY	*A. thaliana*	Weigel *et al.*, 1995 [191]	Yes (T0: 5 months)	not reported
P. tremula × alba female P. tremula ×	PtLFY	P. trichocarpa	Rottmann *et al.*, 2000 [224]	Yes (but only 1 line)	not reported

Species/cultivar	Precocious flower induction transgene*	Trait donor	Reference	Precocious flowering detected **	Fertility
tremuloides male					
P. tremula female P. tremula × tremuloides male	PtLFY	P. trichocarpa	Boehlenius *et al*., 2006 [225]	Yes(within 4 weeks on transformed stem segments)	not reported
P. tremula	BpMADS4	Betula pendula	Hoenicka *et al*., 2008 [226]	no	early flowering phenotype not detected
Populus tremula	MdFT2	M. domestica	Traenkner *et al*., 2010 [205]	Yes (6-10 months)	not reported
P. tremula × alba female P. tremula × tremuloides male	HSP::AtFT HSP::PtFT1, 2	*A. thaliana* P. trichocarpa	Zhang *et al*., 2010 [194]	Yes Signal not graft transmissible	fertile, seed production
P. tremula × tremuloides (male) P. tremula (male)	HSP::AtLFY 35S::AtLFY 35S::PtFT HSP::AtFT	*A. thaliana* *A. thaliana* P. trichocarpa *A. thaliana*	Hoenicka *et al*., 2012 [197] Hoenicka *et al*., 2014 [196]	not detected (heat treatment disturbed plant growth) yes, early flowering (time not indicated) yes, early flowering ((time not indicated) yes, early flowering	not reported not reported not reported fertile, viable F1 seedlings
Prunus domestica (plum)				Juvenile phase under natural conditions: 3-7 years (Srinivasan *et al*., 2012) [167]	
Prunus domestica 'Blubyrd'	PtFT1	P. trichocarpa	Srinivasan *et al*., 2012 [167] Graham *et al*., 2015 [227]	Yes (1-10 months)	fertile, F1 progeny with segregating early flowering phenotype

Species/cultivar	Precocious flower induction transgene*	Trait donor	Reference	Precocious flowering detected **	Fertility
Pyrus communis (pear)				Juvenile phase under natural conditions: 9-14 years (Freiman *et al*., 2012) [228]	
Pyrus communis 'Spadona'	PcTFL1-1, PcTFL1-2 RNAi	Pyrus communis	Freiman *et al*., 2012 [228]	Yes (already under tissue culture conditions, rooted plants 1-8 months)	fertile, F1 progeny with early flowering phenotype
Pyrus communis	AtFT & RNAi PcTFL1-1 or (combined) AtFT & RNAi MdTFL1-1 or (combined) viral expression system #	*Arabidopsis thaliana*, Pyrus communis, *Arabidopsis thaliana*, Malus domestica	Yamagishi *et al*., 2016 [202]	Yes (1-3 months after inoculation of cotyledons)	Normal flower morphology, developing fruits

AP1: APETALA1; CiFT: Citrus unshiu FLOWERING LOCUS T; HSP: heat shock promoter; LFY: LEAFY; TFL1: TERMINAL FLOWER1. VIGS: virus induced gene silencing.
Green: Studies detecting no early flowering phenotype.
*if not indicated otherwise, EFTs are transgenes which are overexpressed in the target plant under the constitutive Cauliflower Mosaic virus (CaMV) 35S promoter. RNAi denotes constructs using RNA interference to knock down genes with inhibitory effect on flower formation/juvenile phase progression
**time may vary between independent lines; earliest observed time listed
apple latent spherical virus (ALSV)

Table 3.2 Accelerated breeding programmes in apple and plum

Species/cultivar	Overexpressed transgene	Breeding goals	Trait donor	Breeding cycle duration	Status (year)	Reference
Malus domestica (apple)						
Malus × domestica 'Pinova' (T1190)	BpMADS4	Fire blight resistance Rvi2, 4 scab resistance FB-F7 fire blight resistance Pl-1, 2 powdery mildew resistance	Malus fusca cv 'Regia' clone 98/6-10	~1 year	BC1 (2011) refer to Fig.5 of [168] for breeding scheme	Flachowsky *et al*., 2011 [168]
Malus × domestica 'Pinova' (T1190)	BpMADS4	Fire blight resistance locus Fb_E	Ornamental apple cultivar 'Evereste'		BC2 (2012) Refer to Table 1 in [207] for specification of crosses	Le Roux *et al*., 2012, 2014 [207, 208]
Malus × domestica 'Pinova', 'Gala', 'Mitchgla Gala', 'Santana'	BpMADS4	Integration of early flowering transgene on various linkage groups in different cultivars for breeding as in Flachowsky *et al*., 2011 [168]	/		/	Weigl *et al*., 2015 [166]
Prunus domestica (plum)						
Prunus domestica	PtFT1	Plum pox virus resistance (transgenic trait) from fresh market plums into dried plum cultivars (f.e. 'Improved French')	P. domestica ‚Honeysweet'		BC1 individuals (2013)	Scorza *et al*., 2013 [210] Srinivasan *et al*., 2011 [212] http://ucanr.edu/sites/fastrack

BC: backcross.

4 Small RNA-directed techniques

4.1 Introduction

Small RNA directed techniques use the cellular machinery of RNA silencing pathways to downregulate gene expression of target genes. For applications in plant breeding, targets may be endogenous genes of the plant, but also of plant pathogens after interaction with the plant (feeding, viral entry, …).

In plants, RNA silencing or RNA interference (RNAi) acts through several pathways to suppress or decrease RNA abundance of, for example, endogenous genes, transposons or viral RNA, and so is involved in regulating plant development and physiology, in maintenance of genome integrity and is used by plants to battle viral attacks [229].

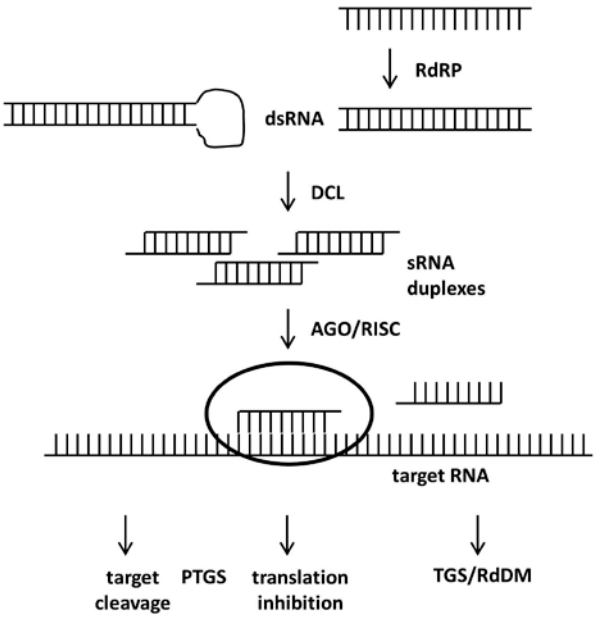


Fig. 4.1 Generalised overview of “the” RNAi pathway. dsRNA molecules are cleaved by DCL proteins into small RNAs. These are incorporated (as single stranded molecules) into the so called RISC complex, which based on sequence complementarity to the incorporated small RNA silences target RNAs by, depending on the pathway, target cleavage or translation inhibition, or, in the case of transcriptional gene silencing, by *de novo* DNA methylation at the target locus. AGO: ARGONAUT; DCL: DICER-LIKE; dsRNA: double stranded RNA; PTGS: post transcriptional gene silencing; RdDM: RNA directed DNA methylation; RdRP: RNA dependent RNA polymerase; RISC: RNA induced silencing complex; TGS: transcriptional gene silencing.

RNAi is a mechanism found in diverse eukaryotes, sharing common core components and exhibiting distinct features. Generally (see for recent reviews in plants [229-231]), central to triggering RNAi are

double stranded RNA (dsRNA) molecules of diverse sources (Fig. 4.1). They are recognised and processed by members of the Dicer family of endonucleases (DCL) into small RNA (sRNA) fragments, in plants typically ~ 21 – 25 nucleotides in length. sRNAs are loaded (as single stranded molecules) into complexes termed RISC (RNA induced silencing complex) containing at least a member of the ARGONAUTE (AGO) family of proteins. AGO proteins are the main silencing effectors and possess an RNase-H-like fold that exhibits endonuclease ("slicer") activity. Within RISC, AGO selects the sRNA guide strand, ejects the passenger strand and mediates sRNA – target RNA recognition. Depending on the particular RNAi pathway, sRNA – target recognition results in post transcriptional gene silencing (PTGS) or transcriptional gene silencing (TGS). In the former, RNA targets are cleaved or translationally repressed/destabilized, in the latter epigenetic modification is induced, RNA-directed DNA methylation (RdDM). In plants, RNAi pathways may also include the action of RNA dependent RNA polymerases (RdRP), for signal amplification or on single stranded RNAs recognized as foreign or aberrant [229-231].

RNAi pathways are further grouped based on origin and biogenesis of sRNAs and engaged members of DCL and AGO proteins into microRNA (miRNA) and small inhibitory RNA (siRNA) pathways [229-231]. sRNAs may act local or systemic; generally, in plants miRNAs act cell-autonomous or move cell-to-cell over short distances, whereas siRNAs have the potential for systemic movement [232, 233].

RNAi based methods exploit the naturally occurring cellular RNAi machinery in order to downregulate expression of target RNAs. In plants for example, a biological role of RNAi is protection from viral attacks. Double stranded (ds) RNA molecules are recognised and processed into small RNAs (sRNAs) approximately 20 nucleotides in length by Dicer proteins. They are loaded into a complex termed RISC. RISCs are targeted based on complementarity to the sRNA to to target RNAs, which are cleaved by the RISC component AGO and thereby inactivated.

4.1.1 miRNAs

miRNAs in plants have been shown to be involved in regulation of plant developmental processes and in biotic and abiotic stress responses [234]. They are encoded at *MIR* loci, non-protein coding nuclear genes, and many belong to evolutionary conserved gene families [230]. *MIR* loci preferentially encode a single miRNA *in vivo* [235] and most plants code for ≥ 100 loci [236].

MIR genes are transcribed by DNA polymerase II, their products may be spliced and give rise to imperfect self-complementary foldback precursor structures, the pri-miRNA. pri-miRNAs carry a stabilizing 5´cap structure and 3´polyadenylated tail and are processed by different progressions depending on their family affiliation. DCL1 is the main dicer activity on pri-miRNAs and finally

processes them into miRNA/miRNA* (guide/passenger strand) duplexes predominantly 21 nucleotides in length. They assemble in RISCs predominantly containing AGO1; the sorting determinant being a 5´uridine [236]. The thermodynamic stability of the miRNA/miRNA* duplex plays a role in guide strand determination and passenger strand elimination within RISC. Target sites of the miRNA in plants are frequently located in open reading frames (ORF) of mRNAs [237]. Target recognition is sequence complementarity based but perfect complementarity is not needed. Comprehensive studies identified key features in respect to thermostability, consensus sites and sequence homology important for biogenesis, strand selection and target recognition and thus effective gene silencing (summarised amongst others for plants in [238, 239]).

Target recognition of miRNAs in the RISC complex may preferentially lead to direct target cleavage (slicing) or translational inhibition/destabilization [234].

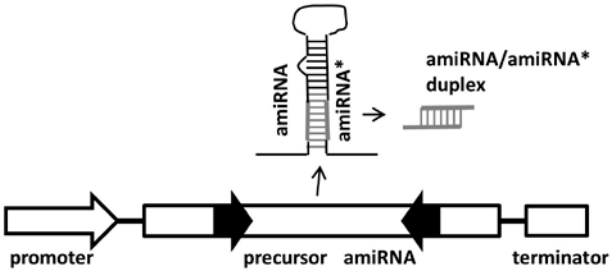


Fig. 4.2 Minimal gene cassette requirements for induction of RNAi using amiRNA constructs. The precursor amiRNA is placed between a promoter and terminator sequence, to initiate and stop transcription, respectively. The transcript gives rise to a stem-loop miRNA precursor transcript, processed primarily by DCL1 into amiRNA/amiRNA* (guide/passenger strand (see chapter 4.1)) duplexes. The guide strands are incorporated into RISC complexes and trigger downregulation of target RNAs. amiRNA: artificial miRNA. DCL1: Dicer like 1.

Methodology

Gene cassettes for induction of RNAi using miRNAs contain an artificial miRNA (amiRNA) precursor between polymerase II regulatory modules for transcription initiation (promoter) and termination (terminator) of choice (Fig. 4.2) [239]. amiRNAs carry the miRNA sequence designed to target the GOI in the context of a miRNA backbone [240]. The backbone used may be selected from a *MIR* gene from the same as well as a from a different plant species [239]. amiRNA design is guided by knowledge on binding specificity parameters, thermostability and consensus sites. Web MicroRNA Designer [239] or Plant Small RNA Maker Suite (P-SAMS) [241] are examples of programmes that integrate this knowledge and calculate and rank potential amiRNAs by sensitivity and specificity for a given target and plant species. Further, functional screens may be used to test the most efficient candidates among predicted amiRNAs [242].

Transformation methods in use to stably introduce amiRNA constructs in plants mainly are *Agrobacterium*-mediated gene transfer and microprojectile (particle) bombardment [243].

4.1.2 siRNAs

In plants, small inhibitory RNAs (siRNAs) arise mainly by DCL2, 3 and 4 activity on dsRNA derived from diverse sources, for example viral origin, transcription of natural antisense transcripts (nat-siRNAs), trans-acting siRNA (TAS) genes and transposon sequences. siRNAs derived from transposons and repeat sequences depend on plant specific DNA-dependent RNA polymerases IV and V (thus termed p4/p5-siRNAs) and ultimately mediate RNA-directed DNA methylation (RdDM). The remaining pathways function through slicing activity on target RNAs [229-231].

Biogenesis of siRNAs differs between pathways. Common to all, and as a distinctive feature to miRNA biogenesis, siRNA pathways do not depend on single siRNAs but usually dsRNA is diced into several entities. siRNA pathways in plants further may involve signal amplification steps carried out by RdRPs [229-231] which additionally to signal amplification may lead to transitive signals, i.e. secondary siRNAs different in sequence to the primary siRNAs [244].

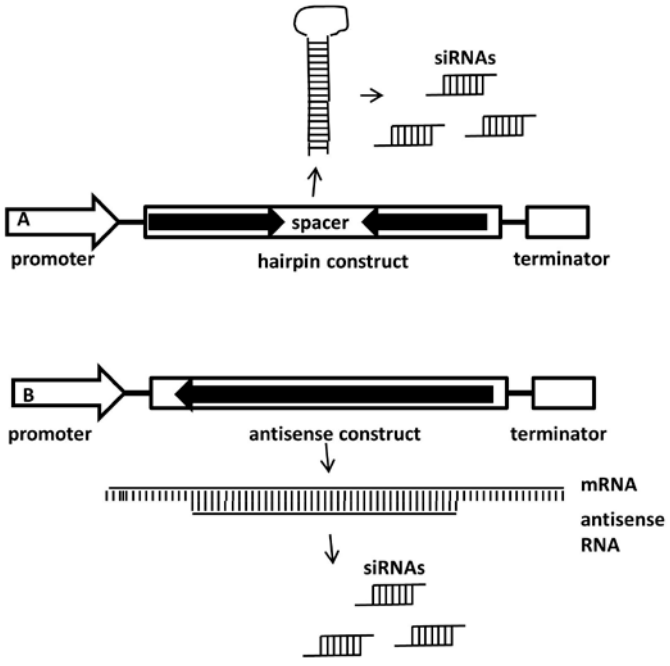


Fig. 4.3 Minimal gene cassette requirements for induction of siRNA mediated RNAi using for example (A) hairpin/inverted repeat constructs or (B) antisense constructs. The dsRNA generating constructs are placed between a promoter and terminator sequence to initiate and stop transcription, respectively. The transcript gives rise to a stem-loop structure, which is processed by members of the DCL family of endonucleases into siRNA duplexes. The guide strands are incorporated into RISC complexes and trigger downregulation of target RNAs. DCL: Dicer like.

Methodology

Gene cassettes for induction of RNAi using siRNAs usually contain a hairpin construct between polymerase II regulatory modules for transcription initiation (promoter) and termination (terminator) of choice (Fig. 4.3) [239]. A hairpin construct consists of inverted repeats complementary to the target region and separated by a spacer. Transcribed hairpin RNA folds into dsRNA and acts as RNAi trigger. Common repeat lengths are between 100 and 1000 nucleotides [245]. Alternatively, antisense and sense constructs may be expressed which trigger RNAi by base pairing to the native sense RNA and by a mechanism called co-suppression, respectively [238, 245]. Co-suppression occurs in situations where overexpression of sense transgenes leads to reduction of expression of both, the transgene and the homologous endogenous gene [238].

Transformation methods used to stably introduce RNAi constructs in plants mainly are *Agrobacterium*-mediated gene transfer and microprojectile (particle) bombardment [243].

RNAi pathways are distinguished based on origin and biogenesis of sRNAs and engaged members of DCL and AGO proteins into microRNA (miRNA) and small inhibitory RNA (siRNA) pathways. Both pathways are exploited to alter targeted traits in RNAi-based plants.

miRNAs are encoded at MIR loci which give rise to defined predominantly 21 nucleotide in length miRNAs. They have been shown to be involved in regulation of plant developmental processes and in biotic and abiotic stress responses.

siRNAs are processed from diverse double stranded RNA sources, for example viral RNA, natural antisense transcripts or transposon sequences. Common to all, and as a distinctive feature to miRNA biogenesis, siRNA pathways do not depend on single siRNAs but usually lead to a pool of differing siRNAs.

4.2 Application of RNAi approaches in plant breeding

RNAi techniques are used to study gene function by downregulation of target gene expression and have been adopted in applied plant research and development. Table 4.1 lists RNAi-based transgenic crop plants present in the scientific literature; entries are selected to exemplify potential areas of application in plant breeding (or, in case of VIRCA project, which are in development phase). Table 4.2 lists examples of RNAi-based transgenic crops which have been developed for the market and have already been evaluated by regulatory agencies; some of these are or have been placed on the market.

The RNAi-based transgene may target plant endogenous genes, and thereby affect quality or agronomical traits as well as for example affect traits involved in abiotic and biotic stress tolerance,

furthermore, it may be designed to target genes expressed in plant pathogens. The latter can be used to establish plants resistant to viral diseases, or, collectively termed host induced gene silencing (HIGS), protect against insects , nematodes (feeding on plants), fungal and bacterial diseases. In the US plants expressing transgenes (RNAi-based and proteinaceous pesticidal substances) acting against plant pests are termed plant incorporated protectants (PIPs).

4.2.1 Applications based on targeting plant endogenous genes

Most examples in the scientific literature of RNAi-based transgenic crop plants illustrating application in plant breeding are altered in respect to quality traits or in respect to abiotic stress tolerance.

Furthermore, RNAi approaches targeting so called susceptibility (S) genes (or recessive resistance genes) [67] may be exploited to establish lines with biotic stress tolerance. These are plant genes that when downregulated or present as loss of function alleles (in a homozygous state) confer (often broad-spectrum) resistance to pathogens, in turn, effectors are produced by pathogens to upregulate those genes creating a favourable cellular environment [67]. MLO (MILDEW RESISTANCE LOCUS) genes are a typical example; naturally occurring and induced MLO loss of function genotypes are used as durable resistance loci for example in plant breeding in barley [246]. The principal feasibility of using RNAi mediated downregulation of recessive resistance genes to mediate biotic stress tolerance has been shown in rice (downregulation of Os-11N3 mediates tolerance to certain *Xanthomonas oryzae* strains [247]) and in a transient expression experiment in wheat (downregulation of TAS3 mediates tolerance to *Blumeria graminis* [248]) (Table 4.1). Whether RNAi-based approaches (versus genome editing) in engineering resistance via S genes will be the method of choice remains to be seen, since the challenge will be to alter targets in respect to its response as susceptibility gene but at the same time retain function in its other cellular contexts.

Examples of how to develop abiotic stress tolerance traits are published in respect to drought tolerance, in canola, corn and potato (Table 4.1; [249-252]). In canola, an inverted repeat construct designed to downregulate farnesyl-transferase (FTA) leads to a reduced transpiration rate by enhanced stomatal closure [251, 252]. FTA is a negative regulator of abscisic acid (ABA) signaling and downregulation also leads to delayed growth and to developmental defects. To bypass these undesired effects, the inverted repeat construct targeting FTA is driven by a drought inducible, shoot-specific *Arabidopsis* promoter [252]. Under limited irrigation conditions in two field trials, seed yield was significantly higher in two transgenic lines compared to the parental line (between 10 – 20% yield increase), and, crucially, the transgenic lines did not perform worse under optimal irrigation conditions. In potato, transpiration rate was reduced by using an amiRNA construct to downregulate Abscisic Acid Hypersensitive 1 (ABH1; also known as cap binding protein 80 (CBP80)) [250]. In corn,

an enzyme involved in ethylene biosynthesis, ACC synthase 6 (ACS6) was downregulated [249]; the plant hormone ethylene is involved in diverse pathways, but it was tested as a means to engineer drought tolerance based on the finding that kernel abortion at the ear tip of corn is correlated with ethylene concentration. In several field tests over two years, two transgenic lines were detected that showed consistently a moderate but significant yield increase under drought conditions while maintaining performance under low stress environment conditions [249]. The increased yield in these lines correlated with a decreased anthesis-silking interval (ASI) under drought stress compared to wild type plants, which ensures efficient pollination of ovaries [249].

Among published crop plants with altered quality traits, there are examples with increased content of desired substances, like amylopectin (potato; [253]), amylose (wheat; [254]) or secondary metabolites (rapeseed, tomato; [255, 256]) (Table 4.1). Furthermore it is possible to reduce the amount of unwanted compounds, like phytate (shown in rice [257]) or of immunogenic epitopes. Immunogenic epitopes were shown to be reduced in transgenic apple (Mal d 1 downregulation; [258, 259]) and carrot (Dau c 1.01/ 1.02 downregulation; [260]) lines in skin prick and oral challenge tests, respectively, in humans and several wheat lines with downregulated α- and/or ω-gliadins showed impaired stimulatory capacity of gliadin reactive T-Cell clones isolated from celiac disease (CD) patients [261-263] (Table 4.1). Transgenic rice lines with reduced phytic acid content were generated based on downregulation of IPK1 (Inositol 1,3,4,5,6-pentakisphosphate 2-kinase), an enzyme involved in late stages of phytic acid biosynthesis, using a seed specific promoter [257]: Transgenic lines maintained a similar level of total phosphorus content in seeds in comparison to wild type plants, the decrease in phytate content was compensated by an increase in inorganic phosphate content. Despite these physiological changes, transgenic lines displayed normal phenotype especially assayed for agronomic parameters (grains/panicle, 1000 seeds dry weight, number of effective tillers,…), for germination behaviour, myo-inositol content and amino acid profiles of storage proteins [257]. This is in contrast to many low phytic acid (lpa) mutants which are negatively affected in seed performance and yield [264]. Plants use phytate to store minerals in seeds and a high percentage of total phosphorus in crop seeds (> 65%) is present in the form of phytic acid, however, phytic acid phosphorus and minerals complexed to phytic acid cannot be efficiently utilized by non-ruminants, and by that also contribute to waste management problems [264, 265]. Therefore, targeting IPK1 orthologs in a tissue specific manner may be of use to implement low phytic acid content in other crops important for food use of non-ruminants.

There are several examples of RNAi based transgenic crop plants with altered quality traits that have passed regulatory approval (Table 4.2). In the EU there are two soybean lines authorized under Regulation (EC) 1829/2003 on genetically modified food and feed (GMO register) altered for

increased oleic acid content. One of the first transgenic plants authorized for growth (1992) and food use (1994) was the FlavrSavr™ tomato in the US engineered for longer shelf life and with changed viscosity behavior of processed fruits (see Table 4.2 and regulatory agency reference therein). Further, recently authorized transgenic plants in the US are an alfalfa line with reduced lignin content, as well as a potato and an apple line both downregulating polyphenol oxidase genes to withstand oxidative browning after slicing or bruising (Table 4.2). The potato line additionally is engineered for purposes of processing involving heat treatment; it does not form high acrylamide content when for example fried, based on it having lower levels of reducing sugars and asparagine by downregulating enzymes involved in their synthesis (Table 4.2).

4.2.2 Applications by targeting RNA expressed by plant pathogens

Viral disease resistance

RNA silencing is used naturally by plants as a strategy of antiviral defense. Double-stranded viral RNA – either of structured genomic regions or replication intermediates of RNA viruses, or of structured transcripts of DNA viruses – present in plant cells is thought to be recognised by some members of the Dicer-like (DCL) protein family to initiate silencing and viral immunity [266]. Genetically engineered virus resistance via RNAi uses transgenes that are designed to induce siRNA formation (f.e. inverted repeat constructs) or amiRNAs aimed at viral sequences. There are several examples in the scientific literature for RNAi mediated viral resistance in crop plants, for example in barley [267], tomato [268], and wheat [269] against barley yellow dwarf virus (BYDV), cucumber mosaic virus (CMV) and wheat streak mosaic virus (WSMV), respectively (Table x.1). In cassava, an important staple crop for example in East African countries, the Virus resistant Cassava for Africa (VIRCA) project has been initiated to engineer resistance against two viral diseases [270]. In the example of wheat, a polycistronic amiRNA precursor construct giving rise to five different amiRNAs targeting WSMV genomic positions was designed using a naturally occurring miRNA precursor from rice, in order to counteract resistance breaking by rapidly evolving viruses [269].

There are at least two cases of crop plants engineered for virus resistance using RNAi with regulatory approval (Table 4.2). Plum resistant against plum pox virus (PPV) was developed by the US Agricultural research Service (ARS; [169]) and gained approval in the US around 2010. The PPV resistance trait has been shown to be stable over 15 years of field testing by natural aphid transmission and by graft inoculations; the latter showed that the virus does not spread far into the grafted wood but remains close to the graft site (reviewed in [169]). The second transgenic plant passing regulatory approval in Brazil (2011) is a bean golden mosaic virus (BGMV) resistant common bean [271, 272].

Host-induced gene silencing (HIGS) of fungi, insects and nematodes

Analogous to RNAi applications in viral protection, using HIGS RNAi constructs are expressed in plants but target RNAs in pathogenic fungi, and bacteria, insects and nematodes [273]. Targeting plant endogenous recessive resistance genes for engineering biotic resistance by RNAi needs knowledge on potential target genes and plants with recessive genotypes need to retain agronomical performance under low stress conditions, thus, engineering suitable candidate loci by RNAi is challenging. HIGS does not interfere with endogenous plant genes but requires knowledge on candidate genes in respective plant pathogens whose downregulation can be induced and which are central to the pathogen life cycle or survival. Published examples of HIGS in crop plants are listed in Table 4.1.

A recent review including the concept of *in planta* delivery of RNAi in nematode crop protection can be found in Lilley *et al*., 2012 [274]; one of the first studies in a major crop plant (soybean) using HIGS was published in 2006. Both, economically important root knot and cyst nematodes feeding on transgenic RNAi plants were shown to be amenable to HIGS Table x.1; [275-277]). In insect control, HIGS offers the potential to transgenically control also phloem feeders, such as aphids, which cannot as efficiently be controlled as chewing type insects with *Bacillus thuringiensis* (Bt) toxins [278]. In wheat, HIGS of the carboxylesterase CbE E4 of the aphid *Sitobion avenae* reduced progeny production [279]. Additionally, *in vitro* data suggest it may render *S. avenae* more sensitive to organophosphate, carbamate, and pyrethroid pesticides, since the orthologue of CbE E4 in another aphid species has been shown to mediate pesticide resistance [279]. Proof of principle studies in crop plants targeting insects started to be published around 2007, describing an engineered maize line showing resistance against the western corn root worm [280]. Recently, it has been shown that HIGS can also be exploited for fungal protection [281-283]. A specialized cell, the haustorium, formed by biotrophic fungal pathogens is used for signal exchange and nutrient uptake, and is believed to also mediate HIGS [283]. Novara *et al*, generated a barley line targeting the *Blumeria graminis* effector protein avra10 which lead to reduction in fungal development [283]. Further examples used HIGS to generate barley and wheat lines with resistance against *Fusarium graminearum* [281-283].

Recently, US-EPA issued a registration note concerning a maize line (MON-87411-9) engineered via RNAi to target an essential gene of the western corn root worm (Table 4.2). The registration is valid for 2 years for the purposes of agronomic evaluation, seed increase and production in breeding nurseries (not for commercial planting).

Traits of RNAi-based plants are modified by targeted downregulation of desired genes. Examples of RNAi-based crop plants in regard to altered quality traits (enhanced secondary metabolites, reduced allergen

potential) or abiotic stress (drought) and biotic stress tolerance (refer to Table 4.1) have been published. Furthermore, several RNAi-based GM plants have undergone successful regulatory approval (Table 4.2).

Recent interest in RNAi-based GM plants has come up in regard to engineering biotic stress resistance, however, the sRNA expressed from the transgene is targeted at viral RNA or RNAs expressed by plant pathogens coming into contact with plants; the term host induced gene silencing (HIGS) is used for this phenomenon. Proof of principle in engineering such traits has been shown for example by targeting *Fusarium* in barley, the aphid *Sitobion avenae* in wheat or nematodes in soy (Table 4.1). A transgenic maize line targeting the western corn root worm is at the moment analysed in field trials in the US, a plum and a common bean line both engineered for resistance against viral diseases have passed regulatory approval in the US and Brazil, respectively (Table 4.2).

4.3 State of development

RNAi approaches have been used in research in order to deduce the function of downregulated genes by observing the resultant phenotypes of plants. In plants, it has been the first method to interfere in a targeted manner with genes of interest in species amenable to transformation.

To date, the main strategies to engineer transgenic plants using RNAi in plant research are the use of artificial miRNAs (amiRNAs) and siRNA mediated RNAi (i.e. using constructs designed to result in longer stretches of dsRNA molecules; Fig. 4.3) to trigger silencing of target genes. Virus induced gene silencing (VIGS) is a further alternative for transient downregulation of a GOI using viral vectors for delivery incorporating fragments with complementarity to the target gene to be silenced [284, 285]; it is not covered further in this report. In the mid 1990ies reviews report on use of sense and antisense suppression techniques in plant research and designate these accepted techniques for gene expression manipulation [286-288]; in parallel and still ongoing is the functional characterisation of the diverse RNAi pathways in plants. One of the first commercial plant lines used RNAi technology, the FlavrSavr™ tomato in the US (Table 4.2; [287]). The use of an intentionally designed inverted repeat construct (also called hairpin construct) to induce silencing was reported in 1998, and targeted a GUS transgene in rice [289]. The wider use of amiRNAs in plants came after publication of the Web MicroRNA Designer (WMD) in 2006 [240], and was first applied in a monocotyledonous species, rice, in 2008 [235].

Determinants of effectiveness of RNAi approaches

The strength of target gene downregulation (expressivity) may range between partial to falling below detection limit and is determined by a combination of the properties of the RNAi construct as well as its functioning as transgene in the genomic context (e.g. location of integration) for a given

established plant line. The phenotypes of independent lines targeting a GOI may therefore form a series of hypomorphic to loss of function phenotype individuals of which suitable candidate lines can be chosen. This may be of advantage for research purposes, but potentially also for applied purposes in case of exploiting genes with severe complete loss-of-function genotypes. A further potential advantage of RNAi-based approaches in balancing negative effects of downregulation of endogenous plant genes is the use of tissue specific promoters, which allow elimination of gene function in target tissues, while gene function in remaining plant organs stays unaffected, or the potential of primarily targeting splicing isoforms (in case of the use of amiRNAs [290]).

Inverted repeat (hairpin) constructs are used now widely as RNAi-based transgenes (see also examples in Table 4.1 and Table 4.2). Early studies comparing different dsRNA constructs eliciting siRNA mediated RNAi showed that inverted repeat constructs showed a high percentage of independently transformed lines with gene silencing effects, whereas sense or antisense constructs, as well as constructs concomitantly expressing a sense and an antisense RNA from two promoters showed less penetrance [245, 291]. Inverted repeat constructs containing an intron as spacer between the inverted repeat sequences seem to be especially effective in eliciting RNAi [245].

For amiRNA design in plants, effectiveness to date is optimized by the pre-miRNA backbone chosen for a given species, as well as on consideration of empirically determined parameters in relation to for example sequence requirements and thermodynamic behaviour of miRNAs effecting their processing, their incorporation into RISC and target recognition [238].

Determinants of specificity of RNAi approaches

The sequence of the RNA component functions as a guide to target RISC complexes to its targets. However, even though plant miRNAs exhibit relatively high sequence complementarity to their targets [292], perfect complementarity is not obligatory. Other factors contribute to ensure proper functioning of RNAi pathways in the cellular context, of which, to date, there is too less information to be included into algorithms for optimization of design of RNAi constructs. Optimization of specificity, i.e. predicting and avoiding of off-targets, to date depends on the availability of transcriptome sequence information, as well as on the available understanding of requirements on specific miRNA/siRNA-target interaction [290].

For plant miRNAs sequence requirements have been studied. Information on experimentally identified miRNA-target interactions, including experiments investigating miRNA-target from non-target interactions [293], derived general patterns of miRNA-target interaction: for example, the 5´ region (~ position 2-12) of the miRNA tends to be mismatch sensitive, while the 3´region has more relaxed constraints; more than two mismatches next to each other and mismatches at the position flanking the cleavage site (10, 11) seem to be uncommon in the dataset of Schwab *et al*., [293].

Similar observations have been obtained by evaluation of experimentally proven miRNA-target duplexes for the distribution of mismatches, single-nucleotide bulges and G:U base pairs [294]. Such patterns are used in the development of scoring matrices for prediction of miRNA targets and, in turn, are also used to predict potential off-target activity of amiRNAs designed to target a gene of interest (for example, Plant Small RNA Maker Site (P-SAMS; [295]), Web MicroRNA Designer (WMD; [240])). WMD also incorporates hybridization energy in target recognition/off-target avoidance calculations [240].

Although similar factors are thought to be guiding specificity of siRNA mediated RNAi, most of the knowledge in plants is derived from studies of miRNA-target interaction (and/or transferred from metazoan studies). In contrast to amiRNA mediated RNAi, siRNA mediated RNAi leads to formation of a pool of distinct siRNAs (Fig. 4.3), each of which potentially can trigger off-target effects and production of secondary siRNAs. Furthermore, potentially, DCL proteins may cut at any site in the dsRNA to produce siRNAs, giving rise to different pools of siRNAs from different copies of the dsRNA.

It has been shown, that perfect complementarity is not needed for siRNA mediated downregulation in *N. benthamiana* using a virus induced gene silencing approach [296]. In a transgenic *A. thaliana* line expressing an antisense construct covering the coding sequence of an endogenous gene, off-target effects were shown on its paralog, as well as on two genes sharing a 23 nucleotide stretch of complete homology (however, remaining similarity of the genes to the target is not reported) [297]. Downregulation of candidates with a 21 nucleotide stretch of complete homology was not detected (again, remaining similarity of the genes to the target is not reported), as well in genes with 21 or 22 nucleotide continuous identity but one mismatch (22 candidates) [297].

In practice, sequence based considerations are integrated into the design of siRNA mediated RNAi constructs, and potential off-target candidates showing sequence similarity can be included in experimental characterisation of established transgenic RNAi-based lines. These considerations can be supported by programmes which incorporate stringency criteria derived from plant and/or metazoan studies, however, due to the high number of potentially diced siRNAs stemming from a particular dsRNA this may be challenging. A recent RNAi technique might facilitate this computational problem in the future. MIGS, miRNA-induced gene silencing, exploits the pathway of trans-acting small interfering RNAs (tasiRNAs) [298]. Certain plant miRNAs, for example miR173, target so called TAS transcripts that are converted into dsRNA by RdRP activity on the 3´ fragment and processed into a phased tasiRNA pool [298, 299]. The tasiRNA pool thus may be predictable to a certain extent.

Recent reviews detailing current considerations in design of siRNA mediated RNAi and implementation of specificity can be found in [300, 301].

RNAi based techniques are firmly established as a basic research tool. Recent interest in RNAi-based GM plants has come up with the realization of engineering plants with resistance against biotic stress by targeting gene expression in the plant pathogen (host induced stress resistance). Furthermore, already on the market are for example soy plants with modified oleic acid content or cultivars resistant against viral disease. Regulators have been increasingly contacted with respect to specific questions concerning RNAi-based GM plants. One central topic is the characterization of off-target effects of RNAi pathways, since sRNAs may also lead to downregulation of non-target RNA showing partial complementarity. Several parameters have been specified that determine specificity for plant miRNAs; at the EU level EFSA is currently collecting scientific advice to inform on potential adaptations of risk assessment of RNAi-based GM plants in the framework of Directive 2001/18/EC.

4.4 Intended and unintended effects

RNAi mediated downregulation of target genes is used in development of RNAi-based GM plants to either (i) effect plant endogenous genes or (ii) effect gene expression/RNA molecules of plant pests. The former may be used to engineer traits in respect to, among others, altered physiology, nutritional content, agronomical traits, biotic and abiotic stress tolerance, whereas the latter is used to confer biotic stress resistance to plants by targeting gene expression in plant pests or viral RNA genomes. The latter is also subsumed under the term plant incorporated protectants (PIP) in the US risk assessment framework.

A potential unintended effect which is discussed specifically in regard to RNAi-based GM plants is the potential off-target effect of the miRNAs/siRNAs, which may lead to unintended downregulation of endogenous plant genes, as well as in the case of acting as a PIP, to unintended effects in non-target organisms.

At the moment, specifics in regard to risk assessment of RNAi-based GM plants are discussed [302], at the EU level by EFSA. Chapter 4.5 covers ongoing work at EFSA.

4.5 Safety considerations

The European Food Safety Authority (EFSA) developed guidelines for risk assessment (RA) of GM plants, among other documents pertaining to food and feed use [303], to non-food/non-feed use [304], to environmental risk assessment [5] as well as supporting guiding documents for example in assessment of potential impacts on non-target organisms [305]. These documents provide guidance on the specific provisions for submission dossiers for authorization of GM plants under Regulation (EC) No. 1829/2003 on GM food and feed or under Directive 2001/18/EC on the deliberate release

into the environment. The majority of authorized GM plants internationally and in the EU are based on transgenic plants expressing one or more novel proteins, however, commercial development of RNAi-based GM plants is expected to increase due to its potential for example in engineering pest resistance or altering crop composition [306].

To date, there is an ongoing process to evaluate and refine the RA framework for the specifics of RNAi-based GM plants. The US Environmental Protection Agency (EPA) organized a Scientific Advisory Panel Meeting in 2014 on "RNAi Technology as a Pesticide: Problem Formulation for Human Health and Ecological Risk Assessment".[7] In the same year, EFSA organized the scientific workshop "Risk assessment considerations for RNAi-based GM plants" [307, 308] in order to formulate and discuss specific features of RNAi-based GM plants. Building on that, in 2015, EFSA published a call for a "Literature review of baseline information to support the risk assessment of RNAi-based GM plants" (OC/EFSA/GMO/2015/01; OC/EFSA/GMO/2015/02) "… to obtain a comprehensive literature overview on several of the risk assessment related issues identified during the EFSA´s workshop." Scientific baseline data present in the scientific literature in areas relevant to the molecular characterization, the food and feed risk assessment and environmental risk assessment will be collected and assessed. It will inform on potential future areas of research to close knowledge gaps of importance to RA of RNAi-based GM plants and/or on potential adaptations to the current framework of risk assessment of GM plants in regard to specifics of RNAi-based GM plants which may be implemented into guidance documents in the future.

Below selected topics discussed during the EFSA workshop are described (a commentary has been published by EFSA [308], the workshop documents can be found online[8] [307]), followed by the specific tasks of information retrieval identified by EFSA and subject to the call for the literature review on support for RA on RNAi-based GM plants (OC/EFSA/GMO/2015/01; OC/EFSA/GMO/2015/02).

4.5.1 EFSA workshop on risk assessment considerations for RNAi-based GM plants

During the EFSA workshop breakout sessions, the following key topics have been discussed [307, 308]:

Off-target activity in RNAi-based GM plants

RNAi-based GM plants carry either an amiRNA construct or a dsRNA construct (leading to formation of siRNAs) to downregulate a target sequence and thereby modifying the desired trait. Unintended off-target effects may arise (i) due to sufficient sequence homology to non-target genes of

[7] Meeting minutes can be found at http://www.epa.gov/sap/fifra-scientific-advisory-panel-meetings
[8] http://www.efsa.europa.eu/de/events/event/140604

amiRNAs/siRNAs as well as, especially in the case of dsRNA expressing plants, (ii) due to uncertainty of the generated pool of siRNAs, which may include secondary siRNAs.

The former problem may be addressed by bio-informatic approaches to identify possible off-target genes. The applicability and the benefit to the overall risk assessment of this approach with available knowledge to date has been discussed: (i) depending on the stringency of off-target prediction criteria applied (see chapter 4.3 for general information), specificity and sensitivity estimates vary, (ii) target prediction is also dependent on the presence and quality of genome/transcriptome sequence information of the transformed plant cultivar, which may differ to reference genomes due to natural genetic variation and/or breeding history. Taking into account that sRNA-mRNA interaction is based on a short sequence length, bio-informatic approaches to date may lead to a large variation in off-target gene candidates depending on criteria and genome sequence used, thereby may have limited additional value to the RNAi-based GM risk assessment to date. However, progress in making more reliable bio-informatic predictions of sRNA/mRNA recognition as well as the presence of suitable genome (transcriptome) sequences, in the future may provide added benefit in guiding well-informed case specific endpoint analyses, in addition to generic comparative analyses in risk assessment of GM plants.

Next generation sequencing methods may be used to characterize the sequences present in a siRNA pool in a given RNAi-based GM plant versus its comparator; a question raised was the accuracy of the methods in regard to answer questions to problems formulated during risk assessment of RNAi-base GM plants.

Food/Feed risk assessment of RNAi based GM plants

The comparative approach used to verify the intended and identify unintended effects of the established GM plant in regard to compositional, phenotypic and agronomic traits was considered to be the appropriate approach also for RNAi-based GM plants. For compositional and nutritional analyses, OECD consensus documents [309, 310] guide in selection and measurement of appropriate key compounds for a given crop species for food/feed use. Case specific analyses are guided by the intended effect of the introduced RNAi construct. As mentioned above, in the future, case specific additional analyses in risk assessment in respect to compositional, phenotypic and agronomic traits may be guided by information based on reliable bio-informatic predictions on potential off-target candidates.

The study of Zhang *et al*., [311] was debated at the EFSA workshop: the authors detected evidence of plant miRNAs in pooled sera of humans with a predominant plant based diet; in a feeding study in mice they established plant MIR168a presence in sera of mice fed a rice-based diet but not in mice fed a control diet; finally, in a feeding study in mice they report biological activity of rice MIR168a:

decrease of low-density lipoprotein receptor adapter protein 1 (LDLRAP1) protein levels in mouse plasma. A study, undertaken in collaboration with Monsanto researchers, replicating the experiments with a special emphasis on the feeding regime could not find evidence for biological activity on LDLRAP1 by dietary miRNAs [312] and postulated that compositional differences in the feeding regime between control and MIR168a administered groups may explain the differences in containing LDLRAP1 protein levels in the study by Zhang *et al*. A study examining plant dietary sRNAs in published 83 animal sRNA datasets [313] found presence of plant miRNAs in 63 datasets. The highest plant miRNA level detected was 10 times lower than that of Zhang *et al*., and datasets showed high variation (including in experimental repetitions). The authors of this study, as well as Tosar *et al*., [314] - based on analyses of publicly available human sRNA datasets and datasets from Zhang *et al*., previous to their initial finding of dietary plant miRNAs-, argue that plant sRNAs present in animal sRNA datasets may partly be due to methodological artefacts. A current review [315] summarises that the majority of work spurred by the publication of Zhang *et al*., [311] could not corroborate their finding: although there is evidence of plant miRNAs in animal tissues in some studies, levels, if detected, are low, calling into question a potential biological role. However, to find scientific consensus on the topic of dietary plant miRNAs in the framework of RNAi-based GM plant risk assessment, this topic is also reflected in the EFSA call for baseline data (see below).

Testing of RNAi molecules *per se* in oral toxicity studies was not considered relevant at the EFSA workshop [307, 308], based on (i) history of safe consumption of RNAi molecules naturally occurring in plants and (ii) information from pharmaceutical studies on bioavailability, metabolism and excretion.

Environmental risk assessment (ERA) of RNAi-based GM plants

A central topic discussed in breakout sessions were RNAi-based GM plants engineered to control insect pests (by host-induced gene silencing (HIGS)), in the US subsumed under plants expressing plant incorporated protectants (PIPs). An “area of concern” in the ERA is the “interaction of GM plants with non-target organisms (NTO), including criteria for selection of appropriate species and relevant functional groups” (Scientific Opinion on guidance for risk assessment of GM plants, EFSA, [5]). There has been issued a supporting guidance document on this particular topic by EFSA [305]. In this context, topics specific to RNAi-based GM risk assessment were discussed:

Exposure characterisation is an integral part of risk assessment which together with hazard characterization leads to risk characterization. Barriers to exposure (including, for example, degradation behaviour in soil, cellular uptake mechanisms in diverse species, sensitivity of diverse species to ingested dsRNA) were discussed since they are valuable in facilitating and refining risk assessment. It was concluded, that at present there is insufficient understanding on parameters of

specific barriers to make generalisations across taxa and to refine exposure estimates, and therefore, at the moment, most reliable conclusions are derived from non-target organism toxicity studies.

Adverse effects are tested in a tiered manner (controlled laboratory studies progressing to more realistic field conditions); for lower tier studies (laboratory conditions) there was a discussion on the appropriate composition of test diets (dsRNA, sRNAs, plant material). Not all potential non-target organisms can be tested, therefore, criteria for selection of appropriate test species have been formulated (based on for example considerations of functional groups, ecological relevance). In the future, in the presence of reliable sequence information on transcriptomes, bio-informatic analyses may be used to support the selection of NTO for adverse effect testing, by concentrating on those with genes sharing homology to the gene in the target species.

4.5.2 EFSA call on literature review to support risk assessment of RNAi-based GM plants

As mentioned above, the EFSA workshop [305] helped identify key areas to be addressed to inform on topics specific to RNAi-based GM plant risk assessment. To continue the process, a call on a literature review collecting and assessing these key areas has been issued (OC/EFSA/GMO/2015/01; OC/EFSA/GMO/2015/02).

Specifically, areas to collect and assess baseline information in the literature review to support the molecular characterization of RNAi-based GM plants identified by EFSA are: (i) characterization and distinctive features of mode-of-action of dsRNA and miRNA pathways in selected species/taxa, (ii) current knowledge on off-target effects of siRNAs and miRNAs and assessment of bio-informatic programmes available to predict off-target effects, and (iii) overview on current methodology to determine siRNA pools in plants and summary on experimental information in the scientific literature on descriptions of siRNA pools.

Areas to gather and assess data in respect to support the food/feed risk assessment of RNAi-based GM plants and derived products are: (i) data on the pharmaco-kinetics profile of RNAi molecules in humans and animals (primarily based on research and development data of RNAi molecules developed for therapeutic use and for oral administration), (ii) effects of RNAi molecules on gastrointestinal tract and annex glands on human and animals, (iii) information on barriers to absorption of RNAi molecules in gastrointestinal tract and placenta of humans and animals, and (iv) assessment of plausibility of effects of RNAi molecules on the immune system of humans and animals.

Finally, areas to be analysed to support the environmental risk assessment are the following: (i) a systematic literature search on the use of host-delivered RNAi molecules in arthropods, nematodes,

annelids and molluscs (reporting defined parameters and silencing effects) in order to assess if and under which conditions siRNA and miRNAs delivered through feeding trigger RNAi in these organisms, (ii) a review on mechanisms of dsRNA (siRNA/miRNA if relevant) uptake in arthropods, nematodes, annelids and molluscs, (iii) a review plausible routes of exposure of the biotic and abiotic environment to dsRNA (siRNA/miRNA if relevant) expressed in RNAi-based GM plants, its environmental fate and barriers of exposure, (iv) based on collected information before, a summary on information on which factors largely influence dsRNA (siRNA/miRNA if relevant) uptake in arthropods, nematodes, annelids and molluscs delivered by feeding, (v) assess plausibility and mechanisms of unintended adverse effects on arthropods, nematodes, annelids and molluscs by cultivation of RNAi-based GM plants, and (vi) an overview on species belonging to arthropods, nematodes, annelids and molluscs for which complete or partial genome data are available.

The European Food Safety Authority (EFSA) developed guidelines for risk assessment (RA) of GM plants. To date, there is an ongoing process to evaluate and refine the RA framework for the specifics of RNAi-based GM plants according to the framework given by Directive 2001/13/EC and EFSA is soliciting scientific advice. EFSA organized a scientific workshop in 2015, followed by a call for a "Literature review of baseline information to support the risk assessment of RNAi-based GM plants" in 2015. Scientific baseline data present in the scientific literature in areas relevant to the molecular characterization, the food and feed risk assessment and environmental risk assessment will be collected and assessed. It will inform on potential future areas of research to close knowledge gaps of importance to RA of RNAi-based GM plants and/or on potential adaptations to the current framework of risk assessment of GM plants in regard to specifics of RNAi-based GM plants which may be implemented into guidance documents in the future.

4.6 Detection and identification

Genomes of RNAi based GM plants contain a stably integrated transgene that in combination with its genomic integration location can be used to develop an event-specific detection method for identification. In case the transgenic construct contains elements often used in development of GMOs these can be used for screening assays for detection purposes. Examples provide the event specific identification methods for RNAi based GM plants soybean MON 87705 and soybean DP-305423-1 listed in the GMOMETHODS database [316, 317].

4.7 Aspects of GMO classification

RNAi-based GM plants fall under the legal definition of GMO given in EU Directive 2001/18/EC.

4.8 Table

Table 4.1 Examples present in the scientific literature (or in development) of RNAi based transgenic crop plants with traits of interest for potential application in plant breeding. Selected and extended from tables in Ricroch *et al*., 2015 [318], Koch *et al*., 2014 [273], Kamthan *et al*., 2015 [319], Saurabh *et al*., 2014 [320] and Tiwari *et al*., 2014 [239].

Crop	Conferred trait*	RNAi construct	References
	Quality/nutritional traits		
Potato	Enhanced amylopectin content	Antisense construct containing fragment of granule bound starch synthase (GBSS)	The EFSA Journal (2006) 324, 1-20 BASF [253]
Rapeseed	Enhanced ß-carotene, zeaxanthin, violaxanthin and lutein content in seeds	Inverted repeat construct containing fragment of lycopene synthase	Yu *et al*., 2008 [256]
Tomato	Enhanced carotenoid and flavonoid content	Inverted repeat constructs containing a partial sequence tomato DE-ETIOLATED 1 (TDET1; regulatory protein)	Davuluri *et al*. (2005) [255]
Wheat	Enhanced amylose content	Inverted repeat constructs containing fragments of Starch branching enzyme IIa and IIb (SBE IIa, IIb)	Regina *et al*. (2006) [254]
Rice	Reduced phytic acid content	Inverted repeat construct containing IPK1 (Inositol 1,3,4,5,6-pentakisphosphate 2-kinase)	Ali *et al*., 2013 [257]
Apple	Reduced allergenic potential (skin prick test, oral challenge test)	Inverted repeat construct containing fragment of apple allergen Mal d 1	Gilissen *et al*., 2005 Dubois *et al*., 2015 [258, 259]
Carrot	Reduced allergenic potential (skin prick test)	Inverted repeat construct containing fragments of carrot allergens Dau c 1.01 and 1.02	Peters *et al*., 2011 [260]
Wheat	Reduced gliadin content, retained potential for good bread baking quality	Inverted repeat construct containing fragments from α-, γ-, and ω-gladins	Gil-Humanes *et al*. 2010, Gil-Humanes *et al*. 2014 Barro *et al*., 2016 [261-263]
	Abiotic stress tolerance		
Canola	Drought tolerance in field trials (reduced transpiration rate)	Inverted repeat construct with partial sequence of farnesyl-transferase (negative regulator of abscisic acid (ABA) signaling)	Wang *et al*., 2009 Performance Plants, Inc. Canada Waltz *et al*., 2014 [251, 252]
Corn	Drought tolerance in field trials (improved kernel set at dry conditions by reduction of anthesis-silking interval (ASI))	Inverted repeat construct with partial sequence of ACC synthase 6 (ACS6; involved in ethylene biosynthesis)	Habben *et al*. 2014 Dupont/Pioneer, USA Waltz *et al*., 2014 [249, 251]
Potato	Drought tolerance in greenhouse conditions (reduced transpiration rate)	amiRNA (*Arabidopsis* miR319a backbone) construct targeting cap-binding protein 80 (CBP80; negative	Pieczynski *et al*., 2013 [250]

Crop	Conferred trait*	RNAi construct	References
		regulator of abscisic acid (ABA) signaling)	
	Biotic stress resistance: viral pathogens		
Barley	Barley yellow dwarf virus (BYDV) resistance	Inverted repeat construct containing sequence of BYDV-polymerase	Wang *et al*., 2000 [267]
Cassava	Project: Virus-Resistant Cassava for Africa (VIRCA) Cassava brown streak disease (CBSD) resistance Cassava mosaic disease (CMD) resistance	Inverted repeat constructs targeting coat protein (CP) region of CBSD virus strains and AC1,2 genes (involved in viral genome replication) in case of CMD virus strains	Taylor *et al*., 2012 [270]
Tomato	Cucumber mosaic virus (CMV) resistance	amiRNA construct (*Arabidopsis* miR159a backbone) targeting viral RdRP 2a/2b transcripts or conserved 3´UTR region of virus	Zhang *et al*., 2011 [268]
Wheat	Wheat streak mosaic virus (WSMV) resistance	amiRNA construct (rice multiplex miR395 backbone) targeting 5 viral genome locations	Fahim *et al*., 2012 [269]
	Biotic stress resistance: fungal pathogens		
Barley	*Blumeria graminis* resistance (reduced fungal development in the absence of the matching barley resistance gene *Mla10*)	HIGS of avra10 (putative *Bg* effector proteins) by inverted repeat construct	Nowara *et al*., 2010 [283]
Barley	*Fusarium graminearum* resistance	HIGS of *Fg* CYP51A, CYP51B and CYP51 by sense and antisense driven transcription of chimeric fragment harbouring partial sequences of all three genes	Koch *et al*., 2013 [282]
Wheat	*Fusarium graminearum* resistance	HIGS of *Fg* chitin synthase (Chs) 3b by inverted repeat construct	Cheng *et al*., 2015 [281]
Wheat	*Blumeria graminis f. sp. tritici* resistance (transient expression experiment)	Downregulation of endogenous recessive resistance gene *Ta*S3 (*Triticum aestivum* susceptibility 3) using a partial fragment against *Ta*S3 in an inverted repeat construct	Li *et al*., 2013 [248]
	Biotic stress resistance: bacterial pathogens		
Rice	*Xanthomonas oryzae pv oryzae* resistance (bacterial blight)	Downregulation of endogenous recessive resistance gene Os-11N3 using a partial fragment against Os-11N3 in an inverted repeat construct	Antony *et al*., 2010 [247]
	Biotic stress resistance: insects		
Maize	*Diabrotica virgifera* resistance (reduction in root damage)	HIGS of V-ATPase A using construct containing gene fragment in an inverted repeat construct	Baum *et al*., 2007 [280]
Wheat	*Sitobion avenae* resistance (reduced progeny production and reduced resistance to phoxim insecticide)	HIGS of CbE E4 (carboxylesterase) using construct containing gene fragment in an inverted repeat construct	Xu *et al*., 2014 [279]

Crop	Conferred trait*	RNAi construct	References
	Biotic stress resistance: nematodes		
Soy	*Heterodera glycines* resistance (development of soybean cyst nematode (SCN) females and number of eggs per cyst were reduced)	HIGS of MSP (major sperm protein) using inverted repeat construct	Steeves *et al*., 2006 [276]
Soy	*Heterodera glycines* resistance (decrease in the number of mature SCN females)	HIGS of HgALD (aldolase) using inverted repeat construct; hairy root system	Youssef *et al*. 2013 [277]
Soy	*Meloidogyne incognita* resistance (reduced number of egg mass and egg number; no complete resistance)	HIGS of Mi-Rpn7 (essential for the integrity of 26S proteasome) using inverted repeat construct; hairy root system	Niu *et al*., 2012 [275]

*Conferred traits were described in more detail in some listed examples; where not further defined, conferred traits may be quantitative in nature (f.e. resistance) and for the exact trait expression please refer to the cited literature. HIGS: host induced gene silencing.

Table 4.2 RNAi based transgenic crops which have been evaluated by regulatory agencies and have been approved for commercial purposes or [+]agronomic evaluation

Species	Trait	Transgene	Agency	Developer
	biotic stress resistance traits			
Plum (Event C5; `Honeysweet´)	Plum pox virus resistance (PPV)	Inverted repeat sequence of PPV coat protein driven by 35S promoter (Scorza *et al*., 2013)	USA: Determination of Non-regulated status by APHIS, USA 2007** US-FDA completed review 2009* US-EPA registration 2010 §	US Department of Agriculture (USDA) Agricultural Research Service (ARS) in cooperation with Research Institutes in Europe
Common Bean (EMBRAPA 5.1)	Bean golden mosaic virus (BGMV) resistance	Inverted repeat sequence of fragment of rep gene (AC1) of BGMV, driven by CaMV35S promoter (Aragao *et al*., 2013) (Aragao *et al*., 2009)	Brazil: Regulatory approval for food, feed and cultivation 2011*,§§	Embrapa, Brazilian Agricultural Research Corporation
Maize [+] MON-87411-9	*Diabrotica virgifera virgifera* (Western corn rootworm (WCR)) resistance	Inverted repeat sequence of fragment of the WCR Snf7 gene, driven by 35S promoter	USA: Determination of Non-regulated status by APHIS, USA 2015** US-FDA completed review 2014* US-EPA registration 2015 for agronomic evaluation (not authorised for commercial purposes) [+]	Monsanto
	quality traits			
Potato Innate™ potatoes 1st generation Events E12, E24, F10, F37, J3, J55, J78, G11, H37, H50	impaired black spot bruise development impaired asparagine (Asn1) and reducing sugar formation (pPhL, pR1) which leads to low acrylamide content upon heat treatment (frying, baking, cooking)	Chimeric construct consisting of 3'-untranslated sequence of the polyphenol oxidase-5 gene (Ppo5) and a fragment of the asparagine synthetase-1 (Asn1) gene Chimeric construct consisting consisting of fragment of promoter for the potato phosphorylase-L (pPhL) gene and a fragment of promoter for the potato R1 gene (pR1) Both designed as inverted repeat genes, each driven by two convergent S. tuberosum endogenous promoters specially active in tubers	USA: Determination of Non-regulated status by APHIS, USA, 2014** US-FDA completed review 2015*** for events in bold	J.R. Simplot Company, USA

Species	Trait	Transgene	Agency	Developer
Apple Arctic™ Apple Events GD743, GS784	impaired enzymatic browning of apple flesh after slicing or bruising	Suppression of four polyphenol oxidase genes PPO2, GPO3, APO5, pSR7 Partial sequences, expressed together in sense orientation (chimeric sense-silencing RNA) by 35S promoter	USA: Determination of Non-regulated status by APHIS, USA, 2015** US-FDA completed review 2015*** Canada: Health Canada: approved product for sale and growth as GM Food 2015 *, #	Okanagan Specialty Fruits Inc, Canada
Alfalfa KK179	Reduced lignin content which allows greater flexibility in harvest timing; high lignin content affects quality negatively	Partial sequence of caffeoyl CoA 3-*O*-methyltransferase (CCOMT) designed as inverted repeat, driven by	USA: Determination of Non-regulated status by APHIS, USA, 2014** US-FDA completed review for use in animal feed 2013***	Monsanto; Forage Genetics International, USA
Soybean MON 87705 (Vistive Gold™)	increased oleic acid and reduced linoleic acid content, which confers higher oxidative stability of the oil	Partial sequences of fatty acid desaturase (fad2-1A) and palmitoyl acyl carrier protein thioesterase (FATB1-A) genes; designed after genomic integration as chimeric inverted repeat construct, driven by a seed specific promoter from soybean	EU: Authorisation for use as/in Food and Feed 2015 ### USA: Determination of Non-regulated status by APHIS, USA, 2011** US-FDA completed review 2011***	Monsanto
Soybean DP 305423 (Plenish Soy)	increased oleic acid and reduced linoleic acid content, which confers higher oxidative stability of the oil	Partial sequence of endogenous fatty acid desaturase (fad2-1), designed to silence the expression of the endogenous fad2-1gene, driven by an endogenous soybean promoter preferentially active in seed tissue	EU: Authorisation for use as/in Food and Feed 2015 ### USA: Determination of Non-regulated status by APHIS, USA, 2010** US-FDA completed review 2009***	DuPont Pioneer
Tomato FlavrSavr™	Decreased cell wall breakdown which confers longer shelf life; processed tomatoes with higher serum viscosity	Endogenous polygalacturonase gene driven by the 35SCaMV promoter in reverse orientation	USA: Determination of Non-regulated status by APHIS, USA, 1992** US-FDA completed review 1994*	Calgene, USA

Listed RNAi plant lines may contain further transgenes to confer additional traits (for example herbicide resistance of MON87705), described are only traits based on an RNAi transgene. Listed RNAi plant lines may have gone through regulatory approval in further countries.

* Center for Environmental Risk Assessment (CERA) (http://www.cera-gmc.org)

** Petitions for Determination of Nonregulated Status Database, US Department of Agriculture (USDA) Animal and Plant Health Inspection Service (APHIS): https://www.aphis.usda.gov/biotechnology/petitions_table_pending.shtml

*** US-FDA Inventory on Biotechnology Consultations on Food from GE Plant Varieties: http://www.accessdata.fda.gov/scripts/fdcc/?set=Biocon

Health Canada, Novel Food Decisions: http://www.hc-sc.gc.ca/fn-an/gmf-agm/appro/index-eng.php

FSANZ Food Standards Code – Standard 1.5.2 – Food produced using Gene Technology https://www.comlaw.gov.au/Series/F2008B00628/Compilations

EU Register of authorised GMOs http://ec.europa.eu/food/dyna/gm_register/index_en.cfm

§ US Environmental Protection Agency (EPA) Plant Incorporated Protectant (PIP) registrations: http://www.epa.gov/regulation-biotechnology-under-tsca-and-fifra/overview-plant-incorporated-protectants

§§ ISAAA, International Service for the Acquisition of Agri-Biotech Applications, GM Approval Database: http://www.isaaa.org/gmapprovaldatabase/default.asp

5 Abbreviations

ABA abscisic acid

ALS acetolactate synthase

ALSV Apple Latent Spherical Virus

amiRNA artificial microRNA

APHIS Animal and Plant Health Inspection Service (USA)

CaMV 35S promoter Cauliflower Mosaic Virus 35S promoter

Cas CRISPR associated

CRISPR Clustered regularly interspaced short palindromic repeats

crRNA CRISPR RNA

ds double stranded

DSB double strand break

EFSA European Food Safety Authority

EMS ethyl methanesulfonate

EPA Environmental Protection Agency (USA)

ERA environmental risk assessment

FDA Food and Drug Administration (USA)

GM genetically modified

GMO genetically modified organism

GOI gene of interest

gRNA guide RNA

HDR homology directed repair

HIGS host induced gene silencing

HSP heat shock promoter

indel insertion-deletion mutation

IR inverted repeat

LG linkage group

MAS marker assisted selection

miRNA micro RNA

MN meganuclease

NHEJ non-homologous end joining

nt nucleotide

NTO non-target organisms

NTWG New Techniques Working Group

PCR polymerase chain reaction

PIP plant incorporated protectants

PTGS post-transcriptional gene silencing

PPV Plum Pox Virus

QTL quantitative trait locus

SDN site directed nuclease

RA risk assessment

RISC RNA induced silencing complex

RNAi RNA interference

sgRNA single guide RNA

siRNA small inhibitory RNA

SNP single nucleotide polymorphism

sRNA small RNA

ss single strand

TALEN transcription activator-like effector nuclease

TGS transcriptional gene silencing

TILLING Targeting Induced Local Lesions in Genomes

tracrRNA *trans*-encoded crRNA

TRV tobacco rattle virus

USDA United States Department of Agriculture

USEPA United States Environmental Protection Agency

VIGE Viral induced gene expression

VIGS Viral induced gene silencing

ZFN zinc finger nuclease

ZKBS Zentrale Kommission für Biologische Sicherheit/Central Commission for biological Safety

6 References

All links in the document were last accessed on 12.08.2016.

[1] Becker, H., *Pflanzenzüchtung*, Ulmer, Stuttgart 2011.
[2] Lusser M, Parisi C, Plan D, E, R.-C., *New plant breeding techniques*, Joint Research Centre — Institute for Prospective Technological Studies, Luxemburg 2011.
[3] AGES, *Cisgenesis - A report on the practical consequences of the application of novel techniques in plant breeding*, Bundesministerium für Gesundheit, Wien 2012.
[4] AGES, *New plant breeding techniques. RNA-dependent methylation, Reverse breeding, Grafting*, Bundesministerium für Gesundheit, Wien 2013.
[5] EFSA, Scientific Opinion on Guidance on the environmental risk assessment of genetically modified plants. *EFSA Journal* 2010, *8*.
[6] Fladung, M., Cibus' herbicide-resistant canola in European limbo. *Nat Biotech* 2016, *34*, 473-474.
[7] ZKBS, Position statement of the ZKBS on new plant breeding techniques, 2012.
[8] Broeders, S. R., De Keersmaecker, S. C., Roosens, N. H., How to deal with the upcoming challenges in GMO detection in food and feed. *Journal of biomedicine & biotechnology* 2012, *2012*, 402418.
[9] Angers-Loustau, A., Petrillo, M., Bonfini, L., Gatto, F.*, et al.*, JRC GMO-Matrix: a web application to support Genetically Modified Organisms detection strategies. *BMC bioinformatics* 2014, *15*, 417.
[10] Parry, G., Patron, N., Bastow, R., Matthewman, C., Meeting report: GARNet/OpenPlant CRISPR-Cas workshop. *Plant methods* 2016, *12*, 6.
[11] Huang, S., Weigel, D., Beachy, R. N., Li, J., A proposed regulatory framework for genome-edited crops. *Nature genetics* 2016, *48*, 109-111.
[12] Whelan, A. I., Lema, M. A., Regulatory Framework for Gene Editing and other New Breeding Techniques (NBTs) in Argentina. *GM crops & food* 2015, 0.
[13] Jinek, M., Chylinski, K., Fonfara, I., Hauer, M.*, et al.*, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (New York, N.Y.)* 2012, *337*, 816-821.
[14] Sander, J. D., Joung, J. K., CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature biotechnology* 2014, *32*, 347-355.
[15] Rath, D., Amlinger, L., Rath, A., Lundgren, M., The CRISPR-Cas immune system: Biology, mechanisms and applications. *Biochimie* 2015, *117*, 119-128.
[16] Hsu, P. D., Lander, E. S., Zhang, F., Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 2014, *157*, 1262-1278.
[17] Osakabe, Y., Osakabe, K., Genome editing with engineered nucleases in plants. *Plant & cell physiology* 2015, *56*, 389-400.
[18] Belhaj, K., Chaparro-Garcia, A., Kamoun, S., Nekrasov, V., Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. *Plant methods* 2013, *9*, 39.
[19] Sorek, R., Lawrence, C. M., Wiedenheft, B., CRISPR-mediated adaptive immune systems in bacteria and archaea. *Annual review of biochemistry* 2013, *82*, 237-266.
[20] Belhaj, K., Chaparro-Garcia, A., Kamoun, S., Patron, N. J., Nekrasov, V., Editing plant genomes with CRISPR/Cas9. *Current opinion in biotechnology* 2015, *32*, 76-84.
[21] Knoll, A., Fauser, F., Puchta, H., DNA recombination in somatic plant cells: mechanisms and evolutionary consequences. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* 2014, *22*, 191-201.
[22] Cermak, T., Baltes, N. J., Cegan, R., Zhang, Y., Voytas, D. F., High-frequency, precise modification of the tomato genome. *Genome biology* 2015, *16*, 232.
[23] Yin, K., Han, T., Liu, G., Chen, T.*, et al.*, A geminivirus-based guide RNA delivery system for CRISPR/Cas9 mediated plant genome editing. *Scientific reports* 2015, *5*, 14926.
[24] Ali, Z., Abul-faraj, A., Li, L., Ghosh, N.*, et al.*, Efficient Virus-Mediated Genome Editing in Plants Using the CRISPR/Cas9 System. *Molecular Plant* 2015, *8*, 1288-1291.

[25] Ali, Z., Abul-Faraj, A., Piatek, M., Mahfouz, M. M., Activity and specificity of TRV-mediated gene editing in plants. *Plant signaling & behavior* 2015, *10*, e1044191.
[26] Marton, I., Zuker, A., Shklarman, E., Zeevi, V*., et al.*, Nontransgenic genome modification in plant cells. *Plant physiology* 2010, *154*, 1079-1087.
[27] Woo, J. W., Kim, J., Kwon, S. I., Corvalan, C*., et al.*, DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nature biotechnology* 2015.
[28] Mahfouz, M. M., Li, L., TALE nucleases and next generation GM crops. *GM crops* 2011, *2*, 99-103.
[29] Bae, S., Park, J., Kim, J. S., Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics (Oxford, England)* 2014, *30*, 1473-1475.
[30] Lei, Y., Lu, L., Liu, H. Y., Li, S*., et al.*, CRISPR-P: a web tool for synthetic single-guide RNA design of CRISPR-system in plants. *Mol Plant* 2014, *7*, 1494-1496.
[31] Lusser, M., Parisi, C., Plan, D., Rodriguez-Cerezo, E., Deployment of new biotechnologies in plant breeding. *Nature biotechnology* 2012, *30*, 231-239.
[32] Pacher, M., Schmidt-Puchta, W., Puchta, H., Two unlinked double-strand breaks can induce reciprocal exchanges in plant genomes via homologous recombination and nonhomologous end joining. *Genetics* 2007, *175*, 21-29.
[33] Ali, Z., Abulfaraj, A., Idris, A., Ali, S*., et al.*, CRISPR/Cas9-mediated viral interference in plants. *Genome biology* 2015, *16*, 238.
[34] Baltes, N. J., Hummel, A. W., Konecna, E., Cegan, R*., et al.*, Conferring resistance to geminiviruses with the CRISPR–Cas prokaryotic immune system. *Nature Plants* 2015, *1*, 15145.
[35] Ji, X., Zhang, H., Zhang, Y., Wang, Y., Gao, C., Establishing a CRISPR–Cas-like immune system conferring DNA virus resistance in plants. *Nature Plants* 2015, *1*, 15144.
[36] La Russa, M. F., Qi, L. S., The New State of the Art: Cas9 for Gene Activation and Repression. *Molecular and cellular biology* 2015, *35*, 3800-3809.
[37] Shukla, V. K., Doyon, Y., Miller, J. C., DeKelver, R. C*., et al.*, Precise genome modification in the crop species Zea mays using zinc-finger nucleases. *Nature* 2009, *459*, 437-441.
[38] Siritunga, D., Sayre, R., Generation of cyanogen-free transgenic cassava. *Planta* 2003, *217*, 367-373.
[39] Vollmann, J., Eynck, C., Camelina as a sustainable oilseed crop: Contributions of plant breeding and genetic engineering. *Biotechnology Journal* 2015, *10*, 525-535.
[40] Sedbrook, J. C., Phippen, W. B., Marks, M. D., New approaches to facilitate rapid domestication of a wild plant to an oilseed crop: Example pennycress (Thlaspi arvense L.). *Plant Science* 2014, *227*, 122-132.
[41] Harper, A. L., Trick, M., Higgins, J., Fraser, F*., et al.*, Associative transcriptomics of traits in the polyploid crop species Brassica napus. *Nature biotechnology* 2012, *30*, 798-802.
[42] Tian, B., Wei, F., Shu, H., Zhang, Q*., et al.*, Decreasing erucic acid level by RNAi-mediated silencing of fatty acid elongase 1 (BnFAE1.1) in rapeseeds (Brassica napus L.). *African Journal of Biotechnology* 2011, *10*, 13194-13201.
[43] Lee, K.-R., In Sohn, S., Jung, J. H., Kim, S. H*., et al.*, Functional analysis and tissue-differential expression of four FAD2 genes in amphidiploid Brassica napus derived from Brassica rapa and Brassica oleracea. *Gene* 2013, *531*, 253-262.
[44] Nour-Eldin, H. H., Andersen, T. G., Burow, M., Madsen, S. R*., et al.*, NRT/PTR transporters are essential for translocation of glucosinolate defence compounds to seeds. *Nature* 2012, *488*, 531-534.
[45] Ozuna, C. V., Iehisa, J. C., Gimenez, M. J., Alvarez, J. B*., et al.*, Diversification of the celiac disease alpha-gliadin complex in wheat: a 33-mer peptide with six overlapping epitopes, evolved following polyploidization. *The Plant journal : for cell and molecular biology* 2015, *82*, 794-805.
[46] Herman, E. M., Helm, R. M., Jung, R., Kinney, A. J., Genetic modification removes an immunodominant allergen from soybean. *Plant physiology* 2003, *132*, 36-43.
[47] Vollmann, J., Rajcan, I., Oil Crop Breeding and Genetics, in: Vollmann, J., Rajcan, I. (Eds.), *Oil Crops*, Springer-Verlag New York 2010, p. 548.
[48] Zraidi, A., Pachner, M., Lelley, T., Obermayer, R., *Cucurbit Genetics Cooperative Report* 2003, pp. 57-61.
[49] Balkunde, R., Pesch, M., Hulskamp, M., Trichome patterning in *Arabidopsis thaliana* from genetic to molecular models. *Current topics in developmental biology* 2010, *91*, 299-321.

[50] Wester, K., Digiuni, S., Geier, F., Timmer, J.*, et al.*, Functional diversity of R3 single-repeat genes in trichome development. *Development (Cambridge, England)* 2009, *136*, 1487-1496.
[51] Furstenberg-Hagg, J., Zagrobelny, M., Bak, S., Plant defense against insect herbivores. *International journal of molecular sciences* 2013, *14*, 10242-10297.
[52] Kivimaki, M., Karkkainen, K., Gaudeul, M., Loe, G., Agren, J., Gene, phenotype and function: GLABROUS1 and resistance to herbivory in natural populations of *Arabidopsis* lyrata. *Molecular ecology* 2007, *16*, 453-462.
[53] Riddick, E. W., Simmons, A. M., Do plant trichomes cause more harm than good to predatory insects? *Pest management science* 2014, *70*, 1655-1665.
[54] Liu, B., Zhu, Y., Zhang, T., The R3-MYB gene GhCPC negatively regulates cotton fiber elongation. *PloS one* 2015, *10*, e0116272.
[55] Nukumizu, Y., Wada, T., Tominaga-Wada, R., Tomato (Solanum lycopersicum) homologs of TRIPTYCHON (SlTRY) and GLABRA3 (SlGL3) are involved in anthocyanin accumulation. *Plant signaling & behavior* 2013, *8*, e24575.
[56] Vendramin, E., Pea, G., Dondini, L., Pacheco, I.*, et al.*, A unique mutation in a MYB gene cosegregates with the nectarine phenotype in peach. *PloS one* 2014, *9*, e90574.
[57] Zhou, L., Zheng, K., Wang, X., Tian, H.*, et al.*, Control of trichome formation in *Arabidopsis* by poplar single-repeat R3 MYB transcription factors. *Frontiers in plant science* 2014, *5*, 262.
[58] Li, Q., Cao, C., Zhang, C., Zheng, S.*, et al.*, The identification of Cucumis sativus Glabrous 1 (CsGL1) required for the formation of trichomes uncovers a novel function for the homeodomain-leucine zipper I gene. *Journal of experimental botany* 2015, *66*, 2515-2526.
[59] Pan, Y., Bo, K., Cheng, Z., Weng, Y., The loss-of-function GLABROUS 3 mutation in cucumber is due to LTR-retrotransposon insertion in a class IV HD-ZIP transcription factor gene CsGL3 that is epistatic over CsGL1. *BMC plant biology* 2015, *15*, 302.
[60] Schilmiller, A. L., Last, R. L., Pichersky, E., Harnessing plant trichome biochemistry for the production of useful compounds. *The Plant journal : for cell and molecular biology* 2008, *54*, 702-711.
[61] Dayan, F. E., Duke, S. O., Trichomes and root hairs: natural pesticide factories. *Pesticide Outlook* 2003, *14*, 175-178.
[62] Duke, S. O., Canel, C., Rimando, A. M., Tellez, M. R.*, et al.*, Current and potential exploitation of plant glandular trichome productivity. *Advances in Botanical Research Incorporating Advances in Plant Pathology, Vol 31 2000* 2000, *31*, 121-151.
[63] Dubos, C., Le Gourrierec, J., Baudry, A., Huep, G.*, et al.*, MYBL2 is a new regulator of flavonoid biosynthesis in *Arabidopsis thaliana*. *The Plant journal : for cell and molecular biology* 2008, *55*, 940-953.
[64] Jun, J. H., Liu, C., Xiao, X., Dixon, R. A., The Transcriptional Repressor MYB2 Regulates Both Spatial and Temporal Patterns of Proanthocyandin and Anthocyanin Pigmentation in Medicago truncatula. *The Plant cell* 2015, *27*, 2860-2879.
[65] Matsui, K., Umemura, Y., Ohme-Takagi, M., AtMYBL2, a protein with a single MYB domain, acts as a negative regulator of anthocyanin biosynthesis in *Arabidopsis*. *The Plant journal : for cell and molecular biology* 2008, *55*, 954-967.
[66] Shi, M. Z., Xie, D. Y., Biosynthesis and metabolic engineering of anthocyanins in *Arabidopsis thaliana*. *Recent patents on biotechnology* 2014, *8*, 47-60.
[67] Pavan, S., Jacobsen, E., Visser, R. G. F., Bai, Y., Loss of susceptibility as a novel breeding strategy for durable and broad-spectrum resistance. *Molecular Breeding* 2010, *25*, 1-12.
[68] Zhou, J., Peng, Z., Long, J., Sosso, D.*, et al.*, Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice. *The Plant journal : for cell and molecular biology* 2015, *82*, 632-643.
[69] van Schie, C. C., Takken, F. L., Susceptibility genes 101: how to be a good host. *Annual review of phytopathology* 2014, *52*, 551-581.
[70] Li, T., Liu, B., Spalding, M. H., Weeks, D. P., Yang, B., High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nature biotechnology* 2012, *30*, 390-392.
[71] Hutin, M., Perez-Quintero, A. L., Lopez, C., Szurek, B., MorTAL Kombat: the story of defense against TAL effectors through loss-of-susceptibility. *Frontiers in plant science* 2015, *6*, 535.

[72] Aflitos, S., Schijlen, E., de Jong, H., de Ridder, D*., et al.*, Exploring genetic variation in the tomato (Solanum section Lycopersicon) clade by whole-genome sequencing. *The Plant journal : for cell and molecular biology* 2014, *80*, 136-148.
[73] Li, J. Y., Wang, J., Zeigler, R. S., The 3,000 rice genomes project: new opportunities and challenges for future rice research. *GigaScience* 2014, *3*, 8.
[74] Mace, E. S., Tai, S., Gilding, E. K., Li, Y*., et al.*, Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum. *Nature communications* 2013, *4*, 2320.
[75] Zhou, Z., Jiang, Y., Wang, Z., Gou, Z*., et al.*, Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nature biotechnology* 2015, *33*, 408-414.
[76] Qi, J., Liu, X., Shen, D., Miao, H*., et al.*, A genomic variation map provides insights into the genetic basis of cucumber domestication and diversity. *Nature genetics* 2013, *45*, 1510-1515.
[77] Araus, J. L., Cairns, J. E., Field high-throughput phenotyping: the new crop breeding frontier. *Trends in plant science* 2014, *19*, 52-61.
[78] Fiorani, F., Schurr, U., Future scenarios for plant phenotyping. *Annual review of plant biology* 2013, *64*, 267-291.
[79] Rhodes, D. H., Hoffmann, L., Jr., Rooney, W. L., Ramu, P*., et al.*, Genome-wide association study of grain polyphenol concentrations in global sorghum [Sorghum bicolor (L.) Moench] germplasm. *Journal of agricultural and food chemistry* 2014, *62*, 10916-10927.
[80] Ruggieri, V., Francese, G., Sacco, A., D'Alessandro, A*., et al.*, An association mapping approach to identify favourable alleles for tomato fruit quality breeding. *BMC plant biology* 2014, *14*, 337.
[81] Sacco, A., Ruggieri, V., Parisi, M., Festa, G*., et al.*, Exploring a Tomato Landraces Collection for Fruit-Related Traits by the Aid of a High-Throughput Genomic Platform. *PloS one* 2015, *10*, e0137139.
[82] Yang, W., Guo, Z., Huang, C., Duan, L*., et al.*, Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. *Nature communications* 2014, *5*, 5087.
[83] Yong, H. Y., Wang, C., Bancroft, I., Li, F*., et al.*, Identification of a gene controlling variation in the salt tolerance of rapeseed (Brassica napus L.). *Planta* 2015, *242*, 313-326.
[84] Svitashev, S., Young, J. K., Schwartz, C., Gao, H*., et al.*, Targeted Mutagenesis, Precise Gene Editing, and Site-Specific Gene Insertion in Maize Using Cas9 and Guide RNA. *Plant physiology* 2015, *169*, 931-945.
[85] Chaparro-Garcia, A., Kamoun, S., Nekrasov, V., Boosting plant immunity with CRISPR/Cas. *Genome biology* 2015, *16*, 254.
[86] Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A*., et al.*, Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 2013, *152*, 1173-1183.
[87] Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z*., et al.*, CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 2013, *154*, 442-451.
[88] Thakore, P. I., D'Ippolito, A. M., Song, L., Safi, A*., et al.*, Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat Meth* 2015, *advance online publication*.
[89] Piatek, A., Ali, Z., Baazim, H., Li, L*., et al.*, RNA-guided transcriptional regulation *in planta* via synthetic dCas9-based transcription factors. *Plant biotechnology journal* 2015, *13*, 578-589.
[90] Bortesi, L., Fischer, R., The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnology Advances* 2015, *33*, 41-52.
[91] Feng, Z., Zhang, B., Ding, W., Liu, X*., et al.*, Efficient genome editing in plants using a CRISPR/Cas system. *Cell Research* 2013, *23*, 1229-1232.
[92] Jiang, W., Zhou, H., Bi, H., Fromm, M*., et al.*, Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in *Arabidopsis*, tobacco, sorghum and rice. *Nucleic Acids Research* 2013, *41*.
[93] Li, J.-F., Norville, J. E., Aach, J., McCormack, M*., et al.*, Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and Nicotiana benthamiana using guide RNA and Cas9. *Nat Biotech* 2013, *31*, 688-691.
[94] Mao, Y., Zhang, H., Xu, N., Zhang, B*., et al.*, Application of the CRISPR-Cas system for efficient genome engineering in plants. *Molecular Plant* 2013, *6*, 2008-2011.
[95] Miao, J., Guo, D., Zhang, J., Huang, Q*., et al.*, Targeted mutagenesis in rice using CRISPR-Cas system. *Cell Research* 2013, *23*, 1233-1236.

[96] Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J. D. G., Kamoun, S., Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease. *Nature biotechnology* 2013, *31*, 691-693.
[97] Shan, Q., Wang, Y., Li, J., Zhang, Y*., et al.*, Targeted genome modification of crop plants using a CRISPR-Cas system. *Nature biotechnology* 2013, *31*, 686-688.
[98] Upadhyay, S. K., Kumar, J., Alok, A., Tuli, R., RNA-Guided Genome Editing for Target Gene Mutations in Wheat. *G3-Genes Genomes Genetics* 2013, *3*, 2233-2238.
[99] Xie, K. B., Yang, Y. N., RNA-guided genome editing in plants using a CRISPR-Cas system. *Molecular Plant* 2013, *6*, 1975-1983.
[100] Liang, Z., Zhang, K., Chen, K., Gao, C., Targeted Mutagenesis in Zea mays Using TALENs and the CRISPR/Cas System. *Journal of Genetics and Genomics* 2014, *41*, 63-68.
[101] Li, Z., Liu, Z.-B., Xing, A., Moon, B. P*., et al.*, Cas9-Guide RNA Directed Genome Editing in Soybean. *Plant physiology* 2015, *169*, 960-970.
[102] Lawrenson, T., Shorinola, O., Stacey, N., Li, C*., et al.*, Induction of targeted, heritable mutations in barley and Brassica oleracea using RNA-guided Cas9 nuclease. *Genome biology* 2015, *16*, 258.
[103] Wang, S., Zhang, S., Wang, W., Xiong, X*., et al.*, Efficient targeted mutagenesis in potato by the CRISPR/Cas9 system. *Plant Cell Reports* 2015, *34*, 1473-1476.
[104] Brooks, C., Nekrasov, V., Lippman, Z. B., Van Eck, J., Efficient gene editing in tomato in the first generation using the clustered regularly interspaced short palindromic repeats/CRISPR-associated9 system. *Plant physiology* 2014, *166*, 1292-1297.
[105] Fan, D., Liu, T., Li, C., Jiao, B*., et al.*, Efficient CRISPR/Cas9-mediated Targeted Mutagenesis in Populus in the First Generation. *Scientific reports* 2015, *5*.
[106] Jia, H., Wang, N., Xcc-facilitated agroinfiltration of citrus leaves: a tool for rapid functional analysis of transgenes in citrus leaves. *Plant Cell Reports* 2014, *33*, 1993-2001.
[107] Feng, Z., Mao, Y., Xu, N., Zhang, B*., et al.*, Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America* 2014, *111*, 4632-4637.
[108] Wang, Z.-P., Xing, H.-L., Dong, L., Zhang, H.-Y*., et al.*, Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in *Arabidopsis* in a single generation. *Genome biology* 2015, *16*.
[109] Xu, R.-F., Li, H., Qin, R.-Y., Li, J*., et al.*, Generation of inheritable and "transgene clean" targeted genome-modified rice in later generations using the CRISPR/Cas9 system. *Scientific reports* 2015, *5*.
[110] Zhou, H., Liu, B., Weeks, D. P., Spalding, M. H., Yang, B., Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. *Nucleic Acids Res* 2014, *42*, 10903-10914.
[111] Wang, Y., Cheng, X., Shan, Q., Zhang, Y*., et al.*, Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat Biotech* 2014, *32*, 947-951.
[112] Ma, X., Zhang, Q., Zhu, Q., Liu, W*., et al.*, A Robust CRISPR/Cas9 System for Convenient, High-Efficiency Multiplex Genome Editing in Monocot and Dicot Plants. *Molecular Plant* 2015, *8*, 1274-1284.
[113] Xie, K., Minkenberg, B., Yang, Y., Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proceedings of the National Academy of Sciences of the United States of America* 2015, *112*, 3570-3575.
[114] Endo, M., Mikami, M., Toki, S., Multigene Knockout Utilizing Off-Target Mutations of the CRISPR/Cas9 System in Rice. *Plant and Cell Physiology* 2015, *56*, 41-47.
[115] Sun, Y., Zhang, X., Wu, C., He, Y*., et al.*, Engineering Herbicide Resistant Rice Plants through CRISPR/Cas9-mediated Homologous Recombination of the Acetolactate Synthase. *Mol Plant* 2016.
[116] Schiml, S., Fauser, F., Puchta, H., The CRISPR/Cas system can be used as nuclease for *in planta* gene targeting and as paired nickases for directed mutagenesis in *Arabidopsis* resulting in heritable progeny. *Plant Journal* 2014, *80*, 1139-1150.
[117] Luo, S., Li, J., Stoddard, T. J., Baltes, N. J*., et al.*, Non-transgenic Plant Genome Editing Using Purified Sequence-Specific Nucleases. *Mol Plant* 2015, *8*, 1425-1427.
[118] Cai, Y., Chen, L., Liu, X., Sun, S*., et al.*, CRISPR/Cas9-Mediated Genome Editing in Soybean Hairy Roots. *PloS one* 2015, *10*.

[119] Fauser, F., Schiml, S., Puchta, H., Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in *Arabidopsis thaliana*. *Plant Journal* 2014, *79*, 348-359.
[120] Hyun, Y., Kim, J., Cho, S. W., Choi, Y*., et al.*, Site-directed mutagenesis in *Arabidopsis thaliana* using dividing tissue-targeted RGEN of the CRISPR/Cas system to generate heritable null alleles. *Planta* 2015, *241*, 271-284.
[121] Jacobs, T. B., LaFayette, P. R., Schmitz, R. J., Parrott, W. A., Targeted genome modifications in soybean with CRISPR/Cas9. *Bmc Biotechnology* 2015, *15*.
[122] Sun, X., Hu, Z., Chen, R., Jiang, Q*., et al.*, Targeted mutagenesis in soybean using the CRISPR-Cas9 system. *Scientific reports* 2015, *5*.
[123] Xu, R., Li, H., Qin, R., Wang, L*., et al.*, Gene targeting using the Agrobacterium tumefaciens-mediated CRISPR-Cas system in rice. *Rice* 2014, *7*.
[124] Zhang, H., Zhang, J., Wei, P., Zhang, B*., et al.*, The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. *Plant biotechnology journal* 2014, *12*, 797-807.
[125] Zhou, H. B., Liu, B., Weeks, D. P., Spalding, M. H., Yang, B., Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. *Nucleic Acids Research* 2014, *42*, 10903-10914.
[126] Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A*., et al.*, DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature biotechnology* 2013, *31*, 827-832.
[127] Xu, H., Xiao, T., Chen, C. H., Li, W*., et al.*, Sequence determinants of improved CRISPR sgRNA design. *Genome research* 2015, *25*, 1147-1157.
[128] Doench, J. G., Hartenian, E., Graham, D. B., Tothova, Z*., et al.*, Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nature biotechnology* 2014, *32*, 1262-1267.
[129] Wang, T., Wei, J. J., Sabatini, D. M., Lander, E. S., Genetic screens in human cells using the CRISPR-Cas9 system. *Science (New York, N.Y.)* 2014, *343*, 80-84.
[130] Wu, X., Scott, D. A., Kriz, A. J., Chiu, A. C*., et al.*, Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nature biotechnology* 2014, *32*, 670-676.
[131] Dang, Y., Jia, G., Choi, J., Ma, H*., et al.*, Optimizing sgRNA structure to improve CRISPR-Cas9 knockout efficiency. *Genome biology* 2015, *16*, 280.
[132] Zhang, H., Zhang, J., Wei, P., Zhang, B*., et al.*, The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. *Plant Biotechnol J.* 2014, *12*, 797 - 807.
[133] Vouillot, L., Thelie, A., Pollet, N., Comparison of T7E1 and surveyor mismatch cleavage assays to detect mutations triggered by engineered nucleases. *G3 (Bethesda, Md.)* 2015, *5*, 407-415.
[134] Jacobs, T. B., LaFayette, P. R., Schmitz, R. J., Parrott, W. A., Targeted genome modifications in soybean with CRISPR/Cas9. *BMC biotechnology* 2015, *15*, 16.
[135] Duan, Y. B., Li, J., Qin, R. Y., Xu, R. F*., et al.*, Identification of a regulatory element responsible for salt induction of rice OsRAV2 through ex situ and in situ promoter analysis. *Plant Molecular Biology* 2015.
[136] Tang, F., Yang, S., Liu, J., Zhu, H., Rj4, a Gene Controlling Nodulation Specificity in Soybeans, Encodes a Thaumatin-Like Protein, but Not the One Previously Reported. *Plant physiology* 2015.
[137] Xu, C., Liberatore, K. L., Macalister, C. A., Huang, Z*., et al.*, A cascade of arabinosyltransferases controls shoot meristem size in tomato. *Nature genetics* 2015, *47*, 784-792.
[138] Lee, J., Chung, J. H., Kim, H. M., Kim, D. W., Kim, H., Designed nucleases for targeted genome editing. *Plant biotechnology journal* 2015.
[139] Slaymaker, I. M., Gao, L., Zetsche, B., Scott, D. A*., et al.*, Rationally engineered Cas9 nucleases with improved specificity. *Science (New York, N.Y.)* 2016, *351*, 84-88.
[140] Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q*., et al.*, High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* 2016, *529*, 490-495.
[141] Bado, S., Forster, B. P., Nielen, S., Ali, A. M*., et al.*, Plant Mutation Breeding: Current Progress and Future Assessment, *Plant Breeding Reviews: Volume 39*, John Wiley & Sons, Inc. 2015, pp. 23-88.
[142] Kurowska, M., Daszkowska-Golec, A., Gruszka, D., Marzec, M*., et al.*, TILLING: a shortcut in functional genomics. *Journal of applied genetics* 2011, *52*, 371-390.
[143] Curtis, M., DNA Repair Pathways and Genes in Plant, in: Shu, Q. Y., Forster, B. P., Nakagawa, H. (Eds.), *Plant mutation breeding and biotechnology* CABI, Wallingford 2012, pp. 57-70.

[144] Bleuyard, J. Y., Gallego, M. E., White, C. I., Recent advances in understanding of the DNA double-strand break repair machinery of plants. *DNA repair* 2006, *5*, 1-12.
[145] Shu, H., Forster, B. P., Nakagawa, H., Principles and Applications of Plant Mutation Breeding, in: Shu, Q. Y., Forster, B. P., Nakagawa, H. (Eds.), *Plant mutation breeding and biotechnology* CABI, Wallingford 2012, pp. 301-325.
[146] EFSA, Scientific opinion addressing the safety assessment of plants developed through cisgenesis and intragenesis. *EFSA Journal* 2012, *10*.
[147] *Single nucleotide polymorphisms - methods and protocols*, Humana Press, New York, NY 2009.
[148] Vogel, B., *Neue Pflanzenzuchtverfahren - Grundlagen für die Klärung offener Fragen bei der rechtlichen Regulierung neuer Pflanzenzuchtverfahren*, Bundesamt für Umwelt, Bern 2012.
[149] Cooper, J. L., Till, B. J., Laport, R. G., Darlow, M. C*., et al.*, TILLING to detect induced mutations in soybean. *BMC plant biology* 2008, *8*, 9.
[150] Till, B. J., Cooper, J., Tai, T. H., Colowit, P*., et al.*, Discovery of chemically induced mutations in rice by TILLING. *BMC plant biology* 2007, *7*, 19.
[151] Chen, L., Huang, L., Min, D., Phillips, A*., et al.*, Development and characterization of a new TILLING population of common bread wheat (Triticum aestivum L.). *PloS one* 2012, *7*, e41570.
[152] Arumuganathan, K., Earle, E. D., Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter*, *9*, 208-218.
[153] Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J*., et al.*, Genome sequence of the palaeopolyploid soybean. *Nature* 2010, *463*, 178-183.
[154] Project, I. R. G. S., The map-based sequence of the rice genome. *Nature* 2005, *436*, 793-800.
[155] Cheng, Z., Lin, J., Lin, T., Xu, M*., et al.*, Genome-wide analysis of radiation-induced mutations in rice (Oryza sativa L. ssp. indica). *Molecular bioSystems* 2014, *10*, 795-805.
[156] Kumar, A., Simons, K., Iqbal, M. J., de Jimenez, M. M*., et al.*, Physical mapping resources for large plant genomes: radiation hybrids for wheat D-genome progenitor Aegilops tauschii. *BMC genomics* 2012, *13*, 597.
[157] Fitzgerald, T., Kazan, K., Li, Z., Morell, M., Manners, J., A high-throughput method for the detection of homoeologous gene deletions in hexaploid wheat. *BMC plant biology* 2010, *10*, 1-15.
[158] IWGSC, A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. *Science (New York, N.Y.)* 2014, *345*, 1251788.
[159] Flachowsky, H., Hanke, M. V., Peil, A., Strauss, S. H., Fladung, M., A review on transgenic approaches to accelerate breeding of woody plants. *Plant Breeding* 2009, *128*, 217-226.
[160] Sasaki, S., Yamagishi, N., Yoshikawa, N., Efficient virus-induced gene silencing in apple, pear and Japanese pear using Apple latent spherical virus vectors. *Plant methods* 2011, *7*, 15.
[161] Yamagishi, N., Sasaki, S., Yamagata, K., Komori, S*., et al.*, Promotion of flowering and reduction of a generation time in apple seedlings by ectopical expression of the *Arabidopsis thaliana* FT gene using the Apple latent spherical virus vector. *Plant Mol Biol* 2011, *75*, 193-204.
[162] van Nocker, S., Gardiner, S. E., Breeding better cultivars, faster: applications of new technologies for the rapid deployment of superior horticultural tree crops. *Horticulture Research* 2014, *1*, 14022.
[163] Blümel, M., Dally, N., Jung, C., Flowering time regulation in crops — what did we learn from *Arabidopsis*? *Current opinion in biotechnology* 2015, *32*, 121-129.
[164] Pena, L., Martin-Trillo, M., Juarez, J., Pina, J. A*., et al.*, Constitutive expression of *Arabidopsis* LEAFY or APETALA1 genes in citrus reduces their generation time. *Nature biotechnology* 2001, *19*, 263-267.
[165] Flachowsky, H., Hattasch, C., Hofer, M., Peil, A., Hanke, M. V., Overexpression of LEAFY in apple leads to a columnar phenotype with shorter internodes. *Planta* 2010, *231*, 251-263.
[166] Weigl, K., Wenzel, S., Flachowsky, H., Peil, A., Hanke, M. V., Integration of BpMADS4 on various linkage groups improves the utilization of the rapid cycle breeding system in apple. *Plant biotechnology journal* 2015, *13*, 246-258.
[167] Srinivasan, C., Dardick, C., Callahan, A., Scorza, R., Plum (Prunus domestica) trees transformed with poplar FT1 result in altered architecture, dormancy requirement, and continuous flowering. *PloS one* 2012, *7*, e40715.

[168] Flachowsky, H., Le Roux, P. M., Peil, A., Patocchi, A*., et al.*, Application of a high-speed breeding technology to apple (Malus x domestica) based on transgenic early flowering plants and marker-assisted selection. *New Phytologist* 2011, *192*, 364-377.
[169] Scorza, R., Callahan, A., Dardick, C., Ravelonandro, M*., et al.*, Genetic engineering of Plum pox virus resistance: 'HoneySweet' plum-from concept to product. *Plant Cell Tissue and Organ Culture* 2013, *115*, 1-12.
[170] Yamagishi, N., Kishigami, R., Yoshikawa, N., Reduced generation time of apple seedlings to within a year by means of a plant virus vector: A new plant-breeding technique with no transmission of genetic modification to the next generation. *Plant biotechnology journal* 2014, *12*, 60-68.
[171] Fischer, M., Fischer, C., Evaluation of Malus species and cultivars at the Fruit Genebank Dresden-Pillnitz and its use for apple resistance breeding. *Genetic Resources and Crop Evolution* 1999, *46*, 235-241.
[172] García-Libreros, T., *Department of Applied Genetics and Cell Biology*, University of Natural Resources and Life Sciences, Vienna 2012, p. 151.
[173] Gessler, C., Pertot, I., Vf scab resistance of Malus. *Trees - Structure and Function* 2012, *26*, 1-14.
[174] Wang, A. D., Aldwinckle, H., Forsline, P., Main, D*., et al.*, EST contig-based SSR linkage maps for Malus x domestica cv Royal Gala and an apple scab resistant accession of M. sieversii, the progenitor species of domestic apple. *Molecular Breeding* 2012, *29*, 379-397.
[175] Lewis, R. S., Kernodle, S. P., A method for accelerated trait conversion in plant breeding. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* 2009, *118*, 1499-1508.
[176] Flachowsky, H., Le Roux, P. M., Peil, A., Patocchi, A*., et al.*, Application of a high-speed breeding technology to apple (Malus×domestica) based on transgenic early flowering plants and marker-assisted selection. *New Phytologist* 2011, *192*, 364-377.
[177] Joshi, S., Soriano, J., Schaart, J., Broggini, G. A. L*., et al.*, Approaches for Development of Cisgenic Apples. *Transgenic Plant Journal* 2009, 6.
[178] Sansavini, S., Tartarini, S., *Acta Horticulturae* 2013, pp. 43-56.
[179] Iezzoni, A., RosBREED: Enabling Marker-assisted Breeding in the Rosaceae. *Hortscience* 2010, *45*, S27-S28.
[180] Baumgartner, I. O., Kellerhals, M., Costa, F., Dondini, L*., et al.*, Development of SNP-based assays for disease resistance and fruit quality traits in apple (Malus x domestica Borkh.) and validation in breeding pilot studies. *Tree Genetics & Genomes* 2016, *12*.
[181] Chagne, D., Crowhurst, R. N., Pindo, M., Thrimawithana, A*., et al.*, The draft genome sequence of European pear (Pyrus communis L. 'Bartlett'). *PloS one* 2014, *9*, e92644.
[182] Velasco, R., Zharkikh, A., Affourtit, J., Dhingra, A*., et al.*, The genome of the domesticated apple (Malus x domestica Borkh.). *Nature genetics* 2010, *42*, 833-839.
[183] Verde, I., Abbott, A. G., Scalabrin, S., Jung, S*., et al.*, The high-quality draft genome of peach (Prunus persica) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nature genetics* 2013, *45*, 487-494.
[184] Xu, Q., Chen, L. L., Ruan, X., Chen, D*., et al.*, The draft genome of sweet orange (Citrus sinensis). *Nature genetics* 2013, *45*, 59-66.
[185] Kumar, S., Bink, M. C. A. M., Volz, R. K., Bus, V. G. M., Chagné, D., Towards genomic selection in apple (Malus × domestica Borkh.) breeding programmes: Prospects, challenges and strategies. *Tree Genetics & Genomes* 2011, *8*, 1-14.
[186] McClure, K. A., Sawler, J., Gardner, K. M., Money, D., Myles, S., Genomics: a potential panacea for the perennial problem. *American Journal of Botany* 2014, *101*, 1780-1790.
[187] Myles, S., Improving fruit and wine: what does genomics have to offer? *Trends in Genetics* 2013, *29*, 190-196.
[188] Brown, S., Apple, in: Badenes, M. (Ed.), *Fruit Breeding*, Springer, New York 2012, p. 329.
[189] Bannier, H.-J., Modern Apple Breeding: Genetic Narrowing and Inbreeding Tendencies. *Erwerbs-Obstbau* 2011, *52*, 85-110.
[190] Liang, W., Dondini, L., De Franceschi, P., Paris, R*., et al.*, Genetic Diversity, Population Structure and Construction of a Core Collection of Apple Cultivars from Italian Germplasm. *Plant Molecular Biology Reporter* 2015, *33*, 458-473.

[191] Weigel, D., Nilsson, O., A developmental switch sufficient for flower initiation in diverse plants. *Nature* 1995, *377*, 495-500.
[192] Yamagishi, N., Yoshikawa, N., Expression of FLOWERING LOCUS T from *Arabidopsis thaliana* induces precocious flowering in soybean irrespective of maturity group and stem growth habit. *Planta* 2011, *233*, 561-568.
[193] Srinivasan, C., Dardick, C., Callahan, A., Scorza, R., Plum (Prunus domestica) trees transformed with poplar FT1 result in altered architecture, dormancy requirement, and continuous flowering. *PLoS ONE* 2012, *7*.
[194] Zhang, H., Harry, D. E., Ma, C., Yuceer, C*., et al.*, Precocious flowering in trees: The FLOWERING LOCUS T gene as a research and breeding tool in Populus. *Journal of experimental botany* 2010, *61*, 2549-2560.
[195] Flachowsky, H., Szankowski, I., Waidmann, S., Peil, A*., et al.*, The MdTFL1 gene of apple (Malus × domestica Borkh.) reduces vegetative growth and generation time. *Tree Physiology* 2012, *32*, 1288-1301.
[196] Hoenicka, H., Lehnhardt, D., Nilsson, O., Hanelt, D., Fladung, M., Successful crossings with early flowering transgenic poplar: interspecific crossings, but not transgenesis, promoted aberrant phenotypes in offspring. *Plant biotechnology journal* 2014, *12*, 1066-1074.
[197] Hoenicka, H., Lehnhardt, D., Polak, O., Fladung, M., Early flowering and genetic containment studies in transgenic poplar. *IForest* 2012, *5*, 138-146.
[198] Klocko, A. L., Ma, C., Robertson, S., Esfandiari, E*., et al.*, FT overexpression induces precocious flowering and normal reproductive development in Eucalyptus. *Plant biotechnology journal* 2015.
[199] Weigl, K., Flachowsky, H., Peil, A., Hanke, M.-V., Heat mediated silencing of MdTFL1 genes in apple (Malus × domestica). *Plant Cell Tiss Organ Cult* 2015, *123*, 511-521.
[200] Wenzel, S., Flachowsky, H., Hanke, M. V., The Fast-track breeding approach can be improved by heat-induced expression of the FLOWERING LOCUS T genes from poplar (Populus trichocarpa) in apple (Malus x domestica Borkh.). *Plant Cell Tissue and Organ Culture* 2013, *115*, 127-137.
[201] Wenzel, S., Flachowsky, H., Hanke, M. V., Preliminary results to establish a speed-breed program based on heat-induced precocious flowering of apple plants containing the flowering locus t gene from poplar (populus trichocarpa). *Acta Horticulturae* 2013, *976*, 471-476.
[202] Yamagishi, N., Li, C., Yoshikawa, N., Promotion of Flowering by Apple Latent Spherical Virus Vector and Virus Elimination at High Temperature Allow Accelerated Breeding of Apple and Pear. *Frontiers in plant science* 2016, *7*, 171.
[203] Tamura, A., Kato, T., Taki, A., Sone, M*., et al.*, Preventive and curative effects of Apple latent spherical virus vectors harboring part of the target virus genome against potyvirus and cucumovirus infections. *Virology* 2013, *446*, 314-324.
[204] Li, C., Sasaki, N., Isogai, M., Yoshikawa, N., Stable expression of foreign proteins in herbaceous and apple plants using Apple latent spherical virus RNA2 vectors. *Archives of virology* 2004, *149*, 1541-1558.
[205] Trankner, C., Lehmann, S., Hoenicka, H., Hanke, M. V*., et al.*, Over-expression of an FT-homologous gene of apple induces early flowering in annual and perennial plants. *Planta* 2010, *232*, 1309-1324.
[206] Jackson, S. D., Hong, Y., Systemic movement of FT mRNA and a possible role in floral induction. *Frontiers in plant science* 2012, *3*, 127.
[207] Le Roux, P. M., Flachowsky, H., Hanke, M. V., Gessler, C., Patocchi, A., Use of a transgenic early flowering approach in apple (Malus x domestica Borkh.) to introgress fire blight resistance from cultivar Evereste. *Molecular Breeding* 2012, *30*, 857-874.
[208] Le Roux, P. M., Flachowsky, H., Jänsch, M., Kellerhals, M*., et al.*, Development of apple pre-breeding genotypes highly resistant to fire blight by early flowering. *Acta Horticulturae* 2014, *1048*, 55-64.
[209] Yao, S., FasTracking Plum Breeding. *Agricultural research* 2011, *59*, 2.
[210] Scorza, R., Callahan, A. M., Dardick, C. D., Srinivasan, C*., et al.*, Biotechnological advances in the genetic improvement of prunus domestica. *Acta Horticulturae* 2013, *985*, 111-118.
[211] Scorza, R., Callahan, A., Dardick, C., Ravelonandro, M*., et al.*, Genetic engineering of Plum pox virus resistance: 'HoneySweet' plum-from concept to product. *Plant Cell, Tissue and Organ Culture* 2013, *115*, 1-12.
[212] Srinivasan, C., Scorza, R., Callahan, A., Dardick, C., *Development of Very Early Flowering and Normal Fruiting Plum With Fertile Seeds*, WIPO 2012.

[213] Park, D., Kim, D., Jang, G., Lim, J.*, et al.*, Efficiency to Discovery Transgenic Loci in GM Rice Using Next Generation Sequencing Whole Genome Re-sequencing. *Genomics & informatics* 2015, *13*, 81-85.
[214] Sastry, K. S., *Seed-borne plant virus diseases*, 2013.
[215] Panattoni, A., Luvisi, A., Triolo, E., Review. Elimination of viruses in plants: Twenty years of progress. *Spanish Journal of Agricultural Research* 2013, *11*, 173-188.
[216] Elo, A., Lemmetyinen, J., Novak, A., Keinonen, K.*, et al.*, BpMADS4 has a central role in inflorescence initiation in silver birch (Betula pendula). *Physiologia plantarum* 2007, *131*, 149-158.
[217] Cervera, M., Navarro, L., Peña, L., Gene stacking in 1-year-cycling APETALA1 citrus plants for a rapid evaluation of transgenic traits in reproductive tissues. *Journal of Biotechnology* 2009, *140*, 278-282.
[218] Endo, T., Shimada, T., Fujii, H., Nishikawa, F.*, et al.*, Development of a CiFT Co-expression system for functional analysis of genes in citrus flowers and fruit. *Journal of the Japanese Society for Horticultural Science* 2009, *78*, 74-83.
[219] Kotoda, N., Hayashi, H., Suzuki, M., Igarashi, M.*, et al.*, Molecular characterization of FLOWERING LOCUS T-like genes of apple (Malus x domestica Borkh.). *Plant & cell physiology* 2010, *51*, 561-575.
[220] Kotoda, N., Iwanami, H., Takahashi, S., Abe, K., Antisense expression of MdTFL1, a TFL1-like gene, reduces the juvenile phase in apple. *Journal of the American Society for Horticultural Science* 2006, *131*, 74-81.
[221] Flachowsky, H., Peil, A., Sopanen, T., Elo, A., Hanke, V., Overexpression of BpMADS4 from silver birch (Betula pendula Roth.) induces early-flowering in apple (Malus x domestica Borkh.). *Plant Breeding* 2007, *126*, 137-145.
[222] Szankowski, I., Waidmann, S., Omar, A. E. D. S., Flachowsky, H.*, et al.*, *Acta Horticulturae* 2009, pp. 633-636.
[223] Trankner, C., Lehmann, S., Hoenicka, H., Hanke, M. V.*, et al.*, Note added in proof to: Over-expression of an FT-homologous gene of apple induces early flowering in annual and perennial plants. *Planta* 2011, *233*, 217-218.
[224] Rottmann, W. H., Meilan, R., Sheppard, L. A., Brunner, A. M.*, et al.*, Diverse effects of overexpression of LEAFY and PTLF, a poplar (Populus) homolog of LEAFY/FLORICAULA, in transgenic poplar and *Arabidopsis*. *The Plant journal : for cell and molecular biology* 2000, *22*, 235-245.
[225] Bohlenius, H., Huang, T., Charbonnel-Campaa, L., Brunner, A. M.*, et al.*, CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. *Science (New York, N.Y.)* 2006, *312*, 1040-1043.
[226] Hoenicka, H., Nowitzki, O., Hanelt, D., Fladung, M., Heterologous overexpression of the birch FRUITFULL-like MADS-box gene BpMADS4 prevents normal senescence and winter dormancy in Populus tremula L. *Planta* 2008, *227*, 1001-1011.
[227] Graham, T., Scorza, R., Wheeler, R., Smith, B.*, et al.*, Over-Expression of FT1 in Plum (Prunus domestica) Results in Phenotypes Compatible with Spaceflight: A Potential New Candidate Crop for Bioregenerative Life Support Systems. *Gravitational and Space Research* 2015, *3*, 12.
[228] Freiman, A., Shlizerman, L., Golobovitch, S., Yablovitz, Z.*, et al.*, Development of a transgenic early flowering pear (Pyrus communis L.) genotype by RNAi silencing of PcTFL1-1 and PcTFL1-2. *Planta* 2012, *235*, 1239-1251.
[229] Martinez de Alba, A. E., Elvira-Matelot, E., Vaucheret, H., Gene silencing in plants: a diversity of pathways. *Biochimica et biophysica acta* 2013, *1829*, 1300-1308.
[230] Bologna, N. G., Voinnet, O., The diversity, biogenesis, and activities of endogenous silencing small RNAs in *Arabidopsis*. *Annual review of plant biology* 2014, *65*, 473-503.
[231] Baulcombe, D. C., VIGS, HIGS and FIGS: small RNA silencing in the interactions of viruses or filamentous organisms with their plant hosts. *Current opinion in plant biology* 2015, *26*, 141-146.
[232] Molnar, A., Melnyk, C., Baulcombe, D. C., Silencing signals in plants: a long journey for small RNAs. *Genome biology* 2011, *12*, 215.
[233] Parent, J. S., Martinez de Alba, A. E., Vaucheret, H., The origin and effect of small RNA signaling in plants. *Frontiers in plant science* 2012, *3*, 179.
[234] Xie, M., Zhang, S., Yu, B., microRNA biogenesis, degradation and activity in plants. *Cellular and molecular life sciences : CMLS* 2015, *72*, 87-99.

[235] Warthmann, N., Chen, H., Ossowski, S., Weigel, D., Herve, P., Highly specific gene silencing by artificial miRNAs in rice. *PloS one* 2008, *3*, e1829.
[236] Rogers, K., Chen, X., Biogenesis, turnover, and mode of action of plant microRNAs. *The Plant cell* 2013, *25*, 2383-2399.
[237] Liu, Q., Wang, F., Axtell, M. J., Analysis of complementarity requirements for plant microRNA targeting using a Nicotiana benthamiana quantitative transient assay. *The Plant cell* 2014, *26*, 741-753.
[238] Ossowski, S., Schwab, R., Weigel, D., Gene silencing in plants using artificial microRNAs and other small RNAs. *The Plant journal : for cell and molecular biology* 2008, *53*, 674-690.
[239] Tiwari, M., Sharma, D., Trivedi, P., Artificial microRNA mediated gene silencing in plants: progress and perspectives. *Plant Molecular Biology* 2014, *86*, 1-18.
[240] Schwab, R., Ossowski, S., Riester, M., Warthmann, N., Weigel, D., Highly specific gene silencing by artificial microRNAs in *Arabidopsis*. *The Plant cell* 2006, *18*, 1121-1133.
[241] Carbonell, A., Takeda, A., Fahlgren, N., Johnson, S. C*., et al.*, New generation of artificial MicroRNA and synthetic trans-acting small interfering RNA vectors for efficient gene silencing in *Arabidopsis*. *Plant physiology* 2014, *165*, 15-29.
[242] Li, J. F., Chung, H. S., Niu, Y., Bush, J*., et al.*, Comprehensive protein-based artificial microRNA screens for effective gene silencing in plants. *The Plant cell* 2013, *25*, 1507-1522.
[243] Barampuram, S., Zhang, Z. J., Recent advances in plant transformation. *Methods in molecular biology (Clifton, N.J.)* 2011, *701*, 1-35.
[244] Voinnet, O., Use, tolerance and avoidance of amplified RNA silencing by plants. *Trends in plant science* 2008, *13*, 317-328.
[245] Wesley, S. V., Helliwell, C. A., Smith, N. A., Wang, M. B*., et al.*, Construct design for efficient, effective and high-throughput gene silencing in plants. *The Plant journal : for cell and molecular biology* 2001, *27*, 581-590.
[246] Acevedo-Garcia, J., Kusch, S., Panstruga, R., Magical mystery tour: MLO proteins in plant immunity and beyond. *The New phytologist* 2014, *204*, 273-281.
[247] Antony, G., Zhou, J., Huang, S., Li, T*., et al.*, Rice xa13 recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene Os-11N3. *The Plant cell* 2010, *22*, 3864-3876.
[248] Li, S. H., Ji, R., Dudler, R., Yong, M. L*., et al.*, Wheat gene TaS3 contributes to powdery mildew susceptibility. *Plant Cell Reports* 2013, *32*, 1891-1901.
[249] Habben, J. E., Bao, X., Bate, N. J., DeBruin, J. L*., et al.*, Transgenic alteration of ethylene biosynthesis increases grain yield in maize under field drought-stress conditions. *Plant biotechnology journal* 2014, *12*, 685-693.
[250] Pieczynski, M., Marczewski, W., Hennig, J., Dolata, J*., et al.*, Down-regulation of CBP80 gene expression as a strategy to engineer a drought-tolerant potato. *Plant biotechnology journal* 2013, *11*, 459-469.
[251] Waltz, E., Beating the heat. *Nat Biotech* 2014, *32*, 610-613.
[252] Wang, Y., Beaith, M., Chalifoux, M., Ying, J*., et al.*, Shoot-specific down-regulation of protein farnesyltransferase (alpha-subunit) for yield protection against drought in canola. *Mol Plant* 2009, *2*, 191-200.
[253] EFSA, Opinion of the Scientific Panel on Genetically Modified Organisms on an application (Reference EFSA-GMO-UK-2005-14) for the placing on the market of genetically modified potato EH92-527-1 with altered starch composition, for production of starch and food/feed uses , under Regulation (EC) No 1829/2003 from BASF plant science. *EFSA Journal* 2006, *4*.
[254] Regina, A., Bird, A., Topping, D., Bowden, S*., et al.*, High-amylose wheat generated by RNA interference improves indices of large-bowel health in rats. *Proc Natl Acad Sci U S A* 2006, *103*, 3546-3551.
[255] Davuluri, G. R., van Tuinen, A., Fraser, P. D., Manfredonia, A*., et al.*, Fruit-specific RNAi-mediated suppression of DET1 enhances carotenoid and flavonoid content in tomatoes. *Nature biotechnology* 2005, *23*, 890-895.
[256] Yu, B., Lydiate, D. J., Young, L. W., Schafer, U. A., Hannoufa, A., Enhancing the carotenoid content of Brassica napus seeds by downregulating lycopene epsilon cyclase. *Transgenic research* 2008, *17*, 573-585.

[257] Ali, N., Paul, S., Gayen, D., Sarkar, S. N.*, et al.*, Development of low phytate rice by RNAi mediated seed-specific silencing of inositol 1,3,4,5,6-pentakisphosphate 2-kinase gene (IPK1). *PloS one* 2013, *8*, e68161.
[258] Dubois, A. E., Pagliarani, G., Brouwer, R. M., Kollen, B. J.*, et al.*, First successful reduction of clinical allergenicity of food by genetic modification: Mal d 1-silenced apples cause fewer allergy symptoms than the wild-type cultivar. *Allergy* 2015, *70*, 1406-1412.
[259] Gilissen, L. J., Bolhaar, S. T., Matos, C. I., Rouwendal, G. J.*, et al.*, Silencing the major apple allergen Mal d 1 by using the RNA interference approach. *The Journal of allergy and clinical immunology* 2005, *115*, 364-369.
[260] Peters, S., Imani, J., Mahler, V., Foetisch, K.*, et al.*, Dau c 1.01 and Dau c 1.02-silenced transgenic carrot plants show reduced allergenicity to patients with carrot allergy. *Transgenic research* 2011, *20*, 547-556.
[261] Gil-Humanes, J., Piston, F., Altamirano-Fortoul, R., Real, A.*, et al.*, Reduced-gliadin wheat bread: an alternative to the gluten-free diet for consumers suffering gluten-related pathologies. *PloS one* 2014, *9*, e90898.
[262] Gil-Humanes, J., Piston, F., Tollefsen, S., Sollid, L. M., Barro, F., Effective shutdown in the expression of celiac disease-related wheat gliadin T-cell epitopes by RNA interference. *Proc Natl Acad Sci U S A* 2010, *107*, 17023-17028.
[263] Barro, F., Iehisa, J. C., Gimenez, M. J., Garcia-Molina, M. D.*, et al.*, Targeting of prolamins by RNAi in bread wheat: effectiveness of seven silencing-fragment combinations for obtaining lines devoid of coeliac disease epitopes from highly immunogenic gliadins. *Plant biotechnology journal* 2016, *14*, 986-996.
[264] Raboy, V., Approaches and challenges to engineering seed phytate and total phosphorus. *Plant Science* 2009, *177*, 281-296.
[265] Lott, J. N. A., Ockenden, I., Raboy, V., Batten, G. D., Phytic acid and phosphorus in crop seeds and fruits: A global estimate. *Seed Science Research* 2000, *10*, 11-33.
[266] Zhang, C., Wu, Z., Li, Y., Wu, J., Biogenesis, Function, and Applications of Virus-Derived Small RNAs in Plants. *Frontiers in microbiology* 2015, *6*, 1237.
[267] Wang, M. B., Abbott, D. C., Waterhouse, P. M., A single copy of a virus-derived transgene encoding hairpin RNA gives immunity to barley yellow dwarf virus. *Molecular plant pathology* 2000, *1*, 347-356.
[268] Zhang, X., Li, H., Zhang, J., Zhang, C.*, et al.*, Expression of artificial microRNAs in tomato confers efficient and stable virus resistance in a cell-autonomous manner. *Transgenic research* 2011, *20*, 569-581.
[269] Fahim, M., Millar, A. A., Wood, C. C., Larkin, P. J., Resistance to Wheat streak mosaic virus generated by expression of an artificial polycistronic microRNA in wheat. *Plant biotechnology journal* 2012, *10*, 150-163.
[270] Taylor, N. J., Halsey, M., Gaitan-Solis, E., Anderson, P.*, et al.*, The VIRCA Project: virus resistant cassava for Africa. *GM crops & food* 2012, *3*, 93-103.
[271] Aragao, F. J., Faria, J. C., First transgenic geminivirus-resistant plant in the field. *Nature biotechnology* 2009, *27*, 1086-1088; author reply 1088-1089.
[272] Aragao, F. J., Nogueira, E. O., Tinoco, M. L., Faria, J. C., Molecular characterization of the first commercial transgenic common bean immune to the Bean golden mosaic virus. *J Biotechnol* 2013, *166*, 42-50.
[273] Koch, A., Kogel, K. H., New wind in the sails: improving the agronomic value of crop plants through RNAi-mediated gene silencing. *Plant biotechnology journal* 2014, *12*, 821-831.
[274] Lilley, C. J., Davies, L. J., Urwin, P. E., RNA interference in plant parasitic nematodes: a summary of the current status. *Parasitology* 2012, *139*, 630-640.
[275] Niu, J., Jian, H., Xu, J., Chen, C.*, et al.*, RNAi silencing of the Meloidogyne incognita Rpn7 gene reduces nematode parasitic success. *European Journal of Plant Pathology* 2012, *134*, 131-144.
[276] Steeves, R. M., Todd, T. C., Essig, J. S., Trick, H. N., Transgenic soybeans expressing siRNAs specific to a major sperm protein gene suppress Heterodera glycines reproduction. *Functional Plant Biology* 2006, *33*, 991-999.

[277] Youssef, R. M., Kim, K. H., Haroon, S. A., Matthews, B. F., Post-transcriptional gene silencing of the gene encoding aldolase from soybean cyst nematode by transformed soybean roots. *Experimental parasitology* 2013, *134*, 266-274.
[278] Zhang, H., Li, H. C., Miao, X. X., Feasibility, limitation and possible solutions of RNAi-based technology for insect pest control. *Insect science* 2013, *20*, 15-30.
[279] Xu, L., Duan, X., Lv, Y., Zhang, X*., et al.*, Silencing of an aphid carboxylesterase gene by use of plant-mediated RNAi impairs Sitobion avenae tolerance of Phoxim insecticides. *Transgenic research* 2014, *23*, 389-396.
[280] Baum, J. A., Bogaert, T., Clinton, W., Heck, G. R*., et al.*, Control of coleopteran insect pests through RNA interference. *Nature biotechnology* 2007, *25*, 1322-1326.
[281] Cheng, W., Song, X. S., Li, H. P., Cao, L. H*., et al.*, Host-induced gene silencing of an essential chitin synthase gene confers durable resistance to Fusarium head blight and seedling blight in wheat. *Plant biotechnology journal* 2015, *13*, 1335-1345.
[282] Koch, A., Kumar, N., Weber, L., Keller, H*., et al.*, Host-induced gene silencing of cytochrome P450 lanosterol C14alpha-demethylase-encoding genes confers strong resistance to Fusarium species. *Proc Natl Acad Sci U S A* 2013, *110*, 19324-19329.
[283] Nowara, D., Gay, A., Lacomme, C., Shaw, J*., et al.*, HIGS: host-induced gene silencing in the obligate biotrophic fungal pathogen Blumeria graminis. *The Plant cell* 2010, *22*, 3130-3141.
[284] Lange, M., Yellina, A. L., Orashakova, S., Becker, A., Virus-induced gene silencing (VIGS) in plants: an overview of target species and the virus-derived vector systems. *Methods in molecular biology (Clifton, N.J.)* 2013, *975*, 1-14.
[285] Pandey, P., Senthil-Kumar, M., Mysore, K. S., Advances in plant gene silencing methods. *Methods in molecular biology (Clifton, N.J.)* 2015, *1287*, 3-23.
[286] Bourque, J. E., Antisense strategies for genetic manipulations in plants. *Plant Science* 1995, *105*, 125-149.
[287] Frizzi, A., Huang, S. S., Tapping RNA silencing pathways for plant biotechnology. *Plant biotechnology journal* 2010, *8*, 655-677.
[288] Senior, I. J., Uses of plant gene silencing. *Biotechnology & genetic engineering reviews* 1998, *15*, 79-119.
[289] Waterhouse, P. M., Graham, M. W., Wang, M. B., Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. *Proc Natl Acad Sci U S A* 1998, *95*, 13959-13964.
[290] Schwab, R., Ossowski, S., Warthmann, N., Weigel, D., Directed gene silencing with artificial microRNAs. *Methods in molecular biology (Clifton, N.J.)* 2010, *592*, 71-88.
[291] Chuang, C. F., Meyerowitz, E. M., Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* 2000, *97*, 4985-4990.
[292] Jones-Rhoades, M. W., Bartel, D. P., Bartel, B., MicroRNAS and their regulatory roles in plants. *Annual review of plant biology* 2006, *57*, 19-53.
[293] Schwab, R., Palatnik, J. F., Riester, M., Schommer, C*., et al.*, Specific effects of microRNAs on the plant transcriptome. *Developmental cell* 2005, *8*, 517-527.
[294] Jones-Rhoades, M. W., Bartel, D. P., Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Molecular cell* 2004, *14*, 787-799.
[295] Fahlgren, N., Hill, S. T., Carrington, J. C., Carbonell, A., P-SAMS: a web site for plant artificial microRNA and synthetic trans-acting small interfering RNA design. *Bioinformatics (Oxford, England)* 2016, *32*, 157-158.
[296] Senthil-Kumar, M., Hema, R., Anand, A., Kang, L*., et al.*, A systematic study to determine the extent of gene silencing in Nicotiana benthamiana and other Solanaceae species when heterologous gene sequences are used for virus-induced gene silencing. *New Phytologist* 2007, *176*, 782-791.
[297] Xu, P., Zhang, Y., Kang, L., Roossinck, M. J., Mysore, K. S., Computational estimation and experimental verification of off-target silencing during posttranscriptional gene silencing in plants. *Plant physiology* 2006, *142*, 429-440.

[298] Felippes, F. F., Weigel, D., Triggering the formation of tasiRNAs in *Arabidopsis thaliana*: the role of microRNA miR173. *EMBO reports* 2009, *10*, 264-270.
[299] Montgomery, T. A., Yoo, S. J., Fahlgren, N., Gilbert, S. D.*, et al.*, AGO1-miR173 complex initiates phased siRNA formation in plants. *Proc Natl Acad Sci U S A* 2008, *105*, 20055-20062.
[300] Ahmed, F., Dai, X., Zhao, P. X., Bioinformatics tools for achieving better gene silencing in plants. *Methods in molecular biology (Clifton, N.J.)* 2015, *1287*, 43-60.
[301] Senthil-Kumar, M., Mysore, K. S., Caveat of RNAi in plants: the off-target effect. *Methods in molecular biology (Clifton, N.J.)* 2011, *744*, 13-25.
[302] Sherman, J. H., Munyikwa, T., Chan, S. Y., Petrick, J. S.*, et al.*, RNAi technologies in agricultural biotechnology: The Toxicology Forum 40th Annual Summer Meeting. *Regulatory toxicology and pharmacology : RTP* 2015, *73*, 671-680.
[303] EFSA, Scientific Opinion on Guidance for risk assessment of food and feed from genetically modified plants. *EFSA Journal* 2011, *9*.
[304] EFSA, Scientific Opinion on Guidance for the risk assessment of genetically modified plants used for non-food or non-feed purposes. *EFSA Journal* 2009, *1164*.
[305] EFSA, Scientific Opinion on the assessment of potential impacts of genetically modified plants on non-target organisms. *EFSA Journal* 2010, *8*.
[306] Parisi, C., Tillie, P., Rodriguez-Cerezo, E., The global pipeline of GM crops out to 2020. *Nature biotechnology* 2016, *34*, 31-36.
[307] EFSA, International scientific workshop 'Risk assessment considerations for RNAi-based GM plants'. 2014, *EN-705*.
[308] Ramon, M., Devos, Y., Lanzoni, A., Liu, Y.*, et al.*, RNAi-based GM plants: food for thought for risk assessors. *Plant biotechnology journal* 2014, *12*, 1271-1273.
[309] OECD, *Safety Assessment of Foods and Feeds Derived from Transgenic Crops, Volume 1*, OECD Publishing.
[310] OECD, *Safety Assessment of Foods and Feeds Derived from Transgenic Crops, Volume 2*, OECD Publishing.
[311] Zhang, L., Hou, D., Chen, X., Li, D.*, et al.*, Exogenous plant MIR168a specifically targets mammalian LDLRAP1: evidence of cross-kingdom regulation by microRNA. *Cell Res* 2012, *22*, 107-126.
[312] Dickinson, B., Zhang, Y., Petrick, J. S., Heck, G.*, et al.*, Lack of detectable oral bioavailability of plant microRNAs after feeding in mice. *Nature biotechnology* 2013, *31*, 965-967.
[313] Zhang, Y., Wiggins, B. E., Lawrence, C., Petrick, J.*, et al.*, Analysis of plant-derived miRNAs in animal small RNA datasets. *BMC genomics* 2012, *13*, 381.
[314] Tosar, J. P., Rovira, C., Naya, H., Cayota, A., Mining of public sequencing databases supports a non-dietary origin for putative foreign miRNAs: underestimated effects of contamination in NGS. *RNA (New York, N.Y.)* 2014, *20*, 754-757.
[315] Yang, J., Hirschi, K. D., Farmer, L. M., Dietary RNAs: New Stories Regarding Oral Delivery. *Nutrients* 2015, *7*, 3184-3199.
[316] Mazzara, M., Munaro, B., Grazioli, E., Savini, C.*, et al.*, *JRC validated methods, reference methods and measurements*, Joint Research Centre – Institute for Health and Consumer Protection, Luxembourg 2013.
[317] Savini, C., Mazzara, M., Delobel, C., Pinski, G., Van den Eede, G., *JRC validated methods, reference methods and measurements*, JRC- Institute for Health and Consumer Protection, Luxembourg 2012.
[318] Ricroch, A. E., Bergé, J. B., Kuntz, M., Evaluation of genetically engineered crops using transcriptomic, proteomic, and metabolomic profiling techniques. *Plant physiology* 2011, *155*, 1752-1761.
[319] Kamthan, A., Chaudhuri, A., Kamthan, M., Datta, A., Small RNAs in plants: recent development and application for crop improvement. *Frontiers in plant science* 2015, *6*, 208.
[320] Saurabh, S., Vidyarthi, A. S., Prasad, D., RNA interference: concept to reality in crop improvement. *Planta* 2014, *239*, 543-564.
[321] Jia, H., Wang, N., Targeted genome editing of sweet orange using Cas9/sgRNA. *PloS one* 2014, *9*, e93806.
[322] Gao, J., Wang, G., Ma, S., Xie, X.*, et al.*, CRISPR/Cas9-mediated targeted mutagenesis in Nicotiana tabacum. *Plant Molecular Biology* 2015, *87*, 99-110.

[323] Zhou, X., Jacobs, T. B., Xue, L. J., Harding, S. A., Tsai, C. J., Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial Populus reveals 4-coumarate:CoA ligase specificity and redundancy. *The New phytologist* 2015, *208*, 298-301.

7 Appendix

7.1 Literature Search

Literature searches were carried out using the databases PubMed, Scopus, Web of Science™ Core Collection and Ovid® (Agris, Agricola, CAB Abstracts and Food Science and Technology Abstracts). Retrieved references were combined in a library in Endnote X7 software (Thomson Reuters) and duplicates were eliminated. Remaining references were checked manually for fulfilling the intended search criteria by title and/or abstract screening. In few instances publications were included from other sources or searches (webpages, random searches).

Literature searches ended March 2016.

7.1.1 CRISPR-Cas

Database searches to find literature relating to CRISPR-Cas application in plants were carried out using the following keywords: [(plant OR plants OR plant* OR “plant breeding”) AND crispr].

7.1.2 Rapid cycle breeding

Database searches to find literature relating to accelerated breeding in plants were carried out using the following keywords: ("high speed breeding" OR "fast breeding" OR "FasTrack breeding" OR "Fast Track breeding" OR "rapid cycle breeding" OR "accelerated breeding") AND plant*.

7.2 Definition of GMO according to EU Directive 2001/18/EC on the deliberate release into the environment of genetically modified organisms

Article 2

Definitions

For the purposes of this Directive:

(1) “organism” means any biological entity capable of replication or of transferring genetic material;

(2) ”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;

Within the terms of this definition:

(a) genetic modification occurs at least through the use of the techniques listed in Annex I A, part 1;

(b) the techniques listed in Annex I A, part 2, are not considered to result in genetic modification;

Article 3

Exemptions

1. This Directive shall not apply to organisms obtained through the techniques of genetic modification listed in Annex I B.

ANNEX I A

TECHNIQUES REFERRED TO IN ARTICLE 2(2)

PART 1

Techniques of genetic modification referred to in Article 2(2)(a) are inter alia:

(1) recombinant nucleic acid techniques involving the formation of new combinations of genetic material by the insertion of nucleic acid molecules produced by whatever means outside an organism, into any virus, bacterial plasmid or other vector system and their incorporation into a host organism in which they do not naturally occur but in which they are capable of continued propagation;

(2) techniques involving the direct introduction into an organism of heritable material prepared outside the organism including micro-injection, macro-injection and micro-encapsulation;

(3) cell fusion (including protoplast fusion) or hybridisation techniques where live cells with new combinations of heritable genetic material are formed through the fusion of two or more cells by means of methods that do not occur naturally.

PART 2

Techniques referred to in Article 2(2)(b) which are not considered to result in genetic modification, on condition that they do not involve the use of recombinant nucleic acid molecules or genetically modified organisms made by techniques/methods other than those excluded by Annex I B:

(1) *in vitro* fertilisation,

(2) natural processes such as: conjugation, transduction, transformation,

(3) polyploidy induction.

ANNEX I B

TECHNIQUES REFERRED TO IN ARTICLE 3

Techniques/methods of genetic modification yielding organisms to be excluded from the Directive, on the condition that they do not involve the use of recombinant nucleic acid molecules or genetically modified organisms other than those produced by one or more of the techniques/methods listed below are:

(1) mutagenesis,

(2) cell fusion (including protoplast fusion) of plant cells of organisms which can exchange genetic material through traditional breeding methods.

7.3 Tables

Table 7.1 Summary of scientific publications in plants reporting analyses on off-target effects of CRISPR-Cas9 in genome editing (2013 – publications available November 2015).

Target locus	Off-target candidate locus identification	Nr. of mismatches distribution	Method of detection	Off-target activity detected	Experimental system	Reference
A. thaliana						
RACK1b/c	1 selected based on homology	2 mm in seed region	sequencing	none detected	transient cell culture	[93]*
GAI	2 BLASTn searches of spacer sequence against genome: complete spacer, seed region only	mm </= 2	candidate off-target sites were aligned against whole genome sequencing data of T1 (n=2) and T2 (n=1) GE lines	none detected	stable transformation germline transmission	[107]
GAI	4 selected based on homology	1-4 mm in/near seed region	60 T1 plants sequenced at each locus	none detected	stable transformation germline transmission	[107]
GAI	na	na	Are mutated target sites stable? Re-sequencing GE lines in progeny	none detected	stable transformation germline transmission	[107]
PHYB	Cas-OFFinder 3 candidates	4-5 mm distributed (fewer mm not detected)	Targeted deep sequencing	none detected	transient delivery of pre-assembled ribonucleoprotein complex into protoplasts	[27]
BRI1 gRNA1	Cas-OFFinder 6 candidates	4-5 mm Distributed (fewer mm not detected)	Targeted deep sequencing	none detected	transient delivery of pre-assembled ribonucleoprotein complex into protoplasts	[27]
BRI1 gRNA2	Cas-OFFinder 4 candidates	2-5 mm Distributed (fewer mm not detected)	Targeted deep sequencing	none detected	transient delivery of pre-assembled ribonucleoprotein complex into protoplasts	[27]
ETC2	Cas-OFFinder 3 candidates	> 4mm in all last 7 seed nt conserved	Amplicon sequencing in 2 GE lines	none detected	stable transformation germline transmission	[108]
FT	Cas-OFFinder	3-4 mm	2 chosen off-target sites with 3 mm	none detected	stable transformation	[120]

Target locus	Off-target candidate locus identification	Nr. of mismatches distribution	Method of detection	Off-target activity detected	Experimental system	Reference
gRNA1	16 candidates	mm in seed region	were amplicon sequenced (n= 48) in a T1 plant		germline transmission	
FT gRNA2	Cas-OFFinder 12 candidates	3-4 mm mm in seed region	1 chosen off-target site with 3 mm was amplicon sequenced (n= 48) in a T1 plant	none detected	stable transformation germline transmission	[120]
C. sinensis						
PDS	BLASTn search of spacer sequence against genome 46 off-target sites included	4-7 mm Distributed	8 off-targets analyzed with restriction enzyme suppressed PCR	none detected	transient Agroinfiltration of leaves	[321]*
G. max						
12g37050	1 candidate based on homology	1 mm in PAM NGG→NAG	Sequencing in 15 GE lines	Yes (1 line identified)	soybean hairy root system	[122]#
07g14530	BLASTn (e value threshold 5) 10 candidate loci	2-6 mm Distributed	Amplicon sequencing (n=10)	none detected	soybean hairy root system	[134]
DDM1 gRNA1	BLASTn (e value threshold 5) 1 candidate loci	4 mm Distributed	Amplicon sequencing (n=10)	none detected	soybean hairy root system	[134]
DDM1 gRNA2	BLASTn (e value threshold 5) 1 candidate locus	2 mm seed region	Amplicon sequencing (n=10)	Yes, in all experimental repeats	soybean hairy root system	[134]
Met1	BLASTn (e value threshold 5) 1 candidate locus	3 mm Distributed	Amplicon sequencing (n=5)	none detected	soybean hairy root system	[134]
miR1514	BLASTn (e value threshold 5) 2 candidate loci	6 and 2 mm Non-seed region	Amplicon sequencing (n=4)	yes, gRNA with 2 mm in non-seed region in all experimental repeats	soybean hairy root system	[134]
H. vulgare						
HvPM19-1	2 candidates based on homology	1 mm in seed region each	Sequencing in 93/95 T1 individuals of two independent T0 lines	Yes, gRNA with mm (further away from PAM than 2nd off-target) in seed region, 3/93	stable transformation	[102]

Target locus	Off-target candidate locus identification	Nr. of mismatches distribution	Method of detection	Off-target activity detected	Experimental system	Reference
				individuals		
HvPM19-3	2 candidates based on homology	1mm in seed r. 3 mm distributed	Sequencing in 76 T1 individuals of one T0 line	None detected		[102]
L. sativa						
BIN2	Cas-OFFinder 349 candidate loci	2-5 mm	High throughput sequencing of 92 candidate sites in 3 GE lines	none detected	transient delivery of pre-assembled ribonucleoprotein complex into protoplasts, regeneration of plants	[27]
N. benthamiana						
PDS	BLASTn 98 candidates	2-10 mm		None→ not conclusive for me	Transient Agro-infiltration of leaves	[96]
PDS	BLASTn search of spacer against genome	5-7 mm	T7EI restriction assay of 13 candidate sequences, n=?	none detected	Transient Agro-infiltration of TRV vector in stably expressing Cas9 plants	[25]
PDS	3 candidates reported by Nekrasov *et al.*, 2013	1, 3, 5 mm	Restriction enzyme suppressed PCR n=5	none detected	Transient Agro-infiltrated leaves	[103]
N. tabacum						
PDR6	BLASTn search of spacer against genome 1 candidate found	2 mm Non-seed region	Sequencing of PCR fragment in GE lines (n=?)	none detected	stable transformation	[322]*
P. tremula × alba						
4CL1	1 candidate selected based on homology	3 mm Seed region	Amplicon sequencing in 8 GE lines	none detected	stable transformation	[323]
4CL5 gRNA	4CL5 in variety with natural SNPs	1 mm seed region 1 mm PAM	Amplicon sequencing in 10 transgenic lines	none detected	stable transformation	[323]
S. tuberosum						
IAA2	BLASTn search of spacer against genome	1 mm PAM	PCR sequencing of 6 GE lines	none detected	stable transformation	[103]

Target locus	Off-target candidate locus identification	Nr. of mismatches distribution	Method of detection	Off-target activity detected	Experimental system	Reference
	1 candidate found					
O. sativa						
MPK5	BLASTn search of spacer against genome 11 candidates	3 mm distributed 3 mm distributed 5 mm distributed	RE suppressed PCR of 3 selected candidates	Yes, activity detected at off-target site with 3 mm which start furthest from PAM	Transient transformation protoplasts	[99]*
PDS	BLASTn search of spacer against genome 1 candidate	3 mm distributed	PCR – RE assay	none detected	Transient transformation protoplasts	[97]*
MPK2	BLASTn search of spacer against genome 2 candidates	1 mm non-seed 1 mm seed	PCR – RE assay PCR – RE assay/sequencing	Yes, potentially detected ## none detected	Transient transformation protoplasts	[97]*
DERF1	Selected based on homology 5 candidates	3-5 mm 2 only in non-seed region	Sequencing at target locus in 20 GE lines (T0 and T1, all independent lines)	none detected	stable transformation	[124]*
MYB1	Selected based on homology 3 candidates	3-5 mm 2 only non-seed region (5 mm)	Sequencing at target locus in 20 GE lines (T0 and T1, all independent lines)	none detected	stable transformation	[124]*
YSA1	Selected based on homology 5 candidates	1-7 mm 2 only non-seed region (1 and 7 mm)	Sequencing at target locus in ~70 Cas9 positive lines (independent T0 lines)	Yes, at 1 candidate locus 7 plants with off-target activity: locus with 1 mm in non-seed region	stable transformation	[124]*
SWEET13	Bioinformatics 6 candidates	>/= 16 identical sites	Sequencing of 7 T0 lines at 6 candidate loci	none detected	stable transformation	[125]*
BEL1	BLASTn search of spacer against genome 3 candidates detected	1 seed 3 seed/non-seed 3 seed/non-seed	Sequencing ~ 80 plants	none detected	stable transformation	[123]
AOX1a	CRISPR-P Selected 2 highest ranked	3, 4 mm distributed	Sequencing of target locus	none detected (50 plants of T0 and T1)	stable transformation	[109]
AOX1b	CRISPR-P Selected 2 highest ranked	3, 4 mm distributed	Sequencing of target locus	none detected (49 plants of T0 and T1)	stable transformation	[109]
AOX1c	CRISPR-P	2, 3 mm	Sequencing of target locus	none detected (60 plants	stable transformation	[109]

Target locus	Off-target candidate locus identification	Nr. of mismatches distribution	Method of detection	Off-target activity detected	Experimental system	Reference
	Selected 2 highest ranked	distributed		of T0 and T1)		
BEL	CRISPR-P Selected 2 highest ranked	1 mm non seed r. 3 mm distributed	Sequencing of target locus	Yes, activity detected in 2 plants at locus with 1 mm (89 plants of T0 and T1)	stable transformation	[109]
CDKB2	3 candidates selected based on homology, confirmed by CRISPR-P Further 3 candidates ranked 3, 5, 9 by CRISPR-P	1 mm non seed r. 2 mm seed/non-seed 2 mm seed/non seed	CAPS marker, sequencing CAPS marker	Yes, activity detected (6/13 regenerated plants) Yes, activity detected (10/13 regenerated plants) none detected (0/13): mm nearest to PAM (all regenerated plants from 1 transformation event (callus); result repeatable in 3 further trasnsformation events (calli)) none detected	stable transformation	[114]
T. aestivum						
INOX	Set of spacers with random mutations	1-11 mm distributed	PCR-RE analysis	Off-target activity detected in case mm are in non-seed region	Transient Protoplast cell culture	[98]*

na: not applicable; mm: mismatches
also report off-target activity with second target, however in that case both loci are 100% identical at spacer and PAM sequence
off-target site very close to target site
*taken from [90]

www.bmgf.gv.at