Background: Part 2 - Risks
Inherent risks associated with CRISPR/Cas applications

The rapid development of new genome editing techniques means there is an increasingly urgent need for responsible handling of the associated risks. Currently, the most widely used and promising procedure is the CRISPR/Cas system. The possible applications of CRISPR/Cas gene scissors are very diverse and involve a multistep process that combines different molecular biological techniques, each of which entails specific risks. This background paper provides an overview of already known and inherent risks that can occur when CRISPR/Cas and old genetic engineering techniques are applied. Introducing CRISPR/Cas into cells and the cell nucleus may cause unintended changes to the genome, RNAs or proteins.

Genome editing is a multistep process that can lead to unintended changes

CRISPR/Cas applications are multistep processes in which various unintended changes can occur that are specific to both old and new genetic engineering. The different steps of gene scissors applications are described in detail in the background paper on the technique.
In the first step, the gene scissors must be introduced into the plant cells. Currently, the most common way to do this is by introducing DNA. This DNA carries the information needed for the formation of the gene scissors in the plant cells and integrates it into the genome. This first step is carried out using old methods of genetic engineering, such as particle bombardment with a gene gun or Agrobacterium transformation. In the next step, the gene scissors are formed in the cell, the target sequences are recognized and cut. This second step is an application of new genetic engineering, i.e. once the gene scissors are in the cell, they actively screen the genome for the target sequence and then cut it.

A rice variety modified with CRISPR/Cas9 gene scissors to produce a higher yield is used here as an example with which to illustrate the risks associated with the multistep process [1]. A subsequent assessment of the genome-edited rice plants showed that various unintended changes had occurred in the genome:

1. Unintended changes from the application of old genetic engineering: the genetic information of the CRISPR/Cas9 system was introduced into the cells with the help of *Agrobacterium tumefaciens,* thereby integrating the DNA to form the gene scissors

into the genome of the rice plants. Parts of the plasmid used to transport the DNA into the plant cells were integrated into the genome during this process. The effects were subsequently detectable in the second generation after the multistep process of genome editing.

2. Unintended changes caused by the new genetic engineering techniques: so-called off-target effects were found in the genome of some of the modified rice plants, which was most likely caused by the CRISPR/Cas9 gene scissors.
3. Other unintended changes caused by the new genetic engineering techniques: in some rice plants, major rearrangements (such as deletions or insertions) of the DNA sequence occurred near the target sequence, an effect known as an on-target effect.

The following section explains in more detail the abovementioned and other unintended changes, which can occur both due to old genetic engineering methods and erroneous and imprecise functions of the gene scissors. Scientists have already described these several times.

Unintended changes due to the application of old genetic engineering techniques

Many scientific studies have already described unintended changes to the genome of plants due to the use of old genetic engineering methods. These older genetic engineering methods are currently mainly used to introduce the gene scissors DNA into plant cells. In an intermediate step, this creates a transgenic plant that carries the genetic information for the formation of the gene scissors in its genome.

Old genetic engineering methods include particle bombardment with a gene gun, transformation with the soil bacterium, *Agrobacterium tumefaciens* and chemical transformation with PEG (more information about these processes can be found in the background paper on the technique of CRISPR/Cas). When the DNA for the formation of the gene scissors is introduced into the plant cells and integrated into the plant genome, this is a random process and not targeted. The aim is to integrate the DNA to form the gene scissors completely into the plant genome just once. However, this is only a best-case scenario because the DNA can be integrated several times and sometimes even incorrectly (i.e. partially). This has been investigated in many different studies, e.g. on soybeans [2; 3], rice [1] or Arabidopsis [4]. In all these studies, Agrobacterium was used to deliver the genetic information into the plant cells. The findings show that the DNA can be inserted once, multiple times, fragmentarily and at different locations in the genome. If additional copies of the DNA unintendedly remain in the genome, the probability increases that unintended changes will occur in subsequent generations. In addition, parts of the plasmid with which the DNA is transported into the cells can also be integrated into the plant genome.

DNA sequences can also be restructured during the process of *Agrobacterium tumefaciens* transformation [5; 6]. For example, parts of the genome can be deleted (known as deletions), DNA sequences can be inserted (known as insertions), DNA sequences can be inserted at another location in the genetic material (known as translocations) or DNA sequences can be reversed in their orientation (known as inversions) [7; 4; 8-10]. Scientists have also been able to show that the composition of epigenetic markers can be changed in the genomic region where the gene scissors DNA is inserted [4]. Epigenetic markers are small appendages on DNA or proteins that regulate gene activity. These epigenetic markers can be compared to little switches that sit on genes and turn them on or off. Ultimately, old genetic engineering applications can also have an effect on the regulation of gene activity.

Unintended effects have also been described for the use of the so-called gene gun [11-13; 7]. In this procedure, DNA-coated small metal particles are shot into the cells with a gene gun. The DNA can be inserted multiple times and the delivery is untargeted [7]. This process was investigated in more detail in rice and maize with the help of whole genome sequencing methods [7]. Whole genome sequencing methods are modern DNA sequencing methods with which the exact sequence of the DNA of the entire genome can be decoded. It was shown that insertions and rearrangements of the DNA can occur during particle bombardment: physical bombardment often causes DNA damage, such as double-strand breaks in the genome. DNA fragments present in the plant cell can be inserted into the genome at these DNA double-strand breaks. These DNA fragments can be, for example, parts of the inserted DNA or even DNA from the chloroplasts of the plants. Chloroplasts are responsible for photosynthesis in the plant cell. If the gene scissors DNA is shot into the plant cell with a gene gun and the chloroplasts in the cell are hit, genes from the chloroplasts can be unintentionally incorporated into the genome in the cell nucleus of the plants.

Similar effects have been described for applications in animals and show how important it is to know the exact procedure, including all intermediate steps, when assessing risks associated with genome-edited organisms. For example, TALENs were used to modify the genome of cattle and insert a gene sequence causing ‘hornlessness’ [14]. During the procedure, a DNA template of the gene sequence for ‘hornlessness’ was introduced into the cells so that it would be integrated at a target sequence in the genome of the cattle. This is a classical SDN-3 gene scissors application, in which additional gene sequences are incorporated into the genome. Investigations subsequently found that the genome of the gene-edited cattle had not only acquired the desired gene sequence, but also gene constructs from the genetically modified bacteria used to deliver the genetic modification [15]. Complete DNA sequences conferring resistance to antibiotics were, among other things, found in the genome of the cattle. What effects this has on the health of the animals and whether the genes are biologically active has not been investigated.

Unintended changes due to incorrect or imprecise application of the gene scissors

Unintended changes in the genome of plants can also occur with new genetic engineering techniques (i.e. genome-editing methods such as CRISPR/Cas). Since CRISPR/Cas applications are so diverse, it is necessary to know exactly how the gene scissors were used in order to investigate the specific risks in more detail.
So-called off-target effects and on-target effects have already been described for established gene scissor applications in which one or more target sequences are unintentionally cut. Furthermore, unwanted DNA fragments can be additionally incorporated into the genome and new gene products can be unintentionally formed. All of the listed unintended changes can occur in SDN-1, SDN-2, and SDN-3 gene scissor applications, i.e. when DNA is cut at one or more target region(s). In SDN-2 and SDN-3, additional unintended changes can occur. For example, DNA templates may be incorporated into the genome or epigenetic markers may be rewritten. Process-specific risks may also arise in procedures that do not cut at the target sequence but use non-cutting gene scissors. All of the abovementioned unintended changes are explained below in more detail.

Off-target-effects (OTEs)

Off-target effects induced by the gene scissors have already been described for many organisms, including human, animal and plant cells. Off-target effects can result from the gene scissors cutting at unintended regions of the genome. Often these are DNA regions that are very similar to the actual target sequence [16-18]. The gene scissors confuse the off-target regions with the actual target sequence and cut there. The cuts at the off-target regions activate the cell's own repair mechanisms and can cause unintended changes, e.g. point mutations or small insertions and deletions. Thus, the same processes are set in motion that also act at the target sequence.
The off-target changes can have different effects depending on the particular genomic context. For example, the unintended changes can lead to the silencing of genes, to the alteration of a gene sequence or also to the alteration of gene activity.

The occurrence of off-target effects depends on many different factors [18]. Above all, the development of a properly designed gRNA for a specific target sequence plays a particularly important role. It was shown that the more similar two DNA sequences in the genome are, the higher the probability that the gene scissors will work imprecisely. If the target area of the gene scissors differs from off-target areas by only one or two bases, it is very likely that the gene scissors will cut at both sites [18].

The complexity of CRISPR/Cas applications also plays an important role in analyzing the risks inherent in the technology and should therefore be considered in a risk analysis:

1. CRISPR/Cas can cause changes in areas of the genome where changes are very unlikely to occur naturally. This means that unintended changes can be induced in areas of the genome that are, for example, protected by the cell's own repair mechanisms or that are located in areas of the genome that are difficult to access using conventional breeding (e.g. in the area of the centromere and other densely packed DNA areas).
2. If the gene scissors are guided to several different areas of the genome during multiplexing, OTEs can be caused at several different off-target areas. Accordingly, the genome must be analyzed in more detail.
3. The ability of CRISPR/Cas to modify all DNA regions with the same sequence should also be taken into account in risk assessment. If an off-target region is present several times in a plant, e.g. in the DNA sequence of a gene of a gene family, then several or all genes of this gene family can be unintentionally modified simultaneously. Polyploid plants are another example: CRISPR/Cas can modify an off-target region in these plants wherever gene variants (i.e., alleles) with this DNA sequence are present in the genome.

Genomics methods enable the analysis of the DNA sequence in genome-edited plants as well the investigation whether OTEs have been induced in the genome after gene scissors applications. To date, however, there have only been a few studies in which the entire genome of genome-edited plants has been examined using whole genome sequencing (WGS) methods [19-22]. All WGS studies on genome-edited plants emphasize that the design of the gRNA is crucial for the occurrence of OTEs. Compared to unintended changes that occur working in cell cultures (also called somaclonal variation), OTEs are quite a rare occurrence. The risks that may arise vary from case to case, as they are highly dependent on the particular context of where in the genome they arise.

Usually, computer programs and reference genomes are used to determine which areas of the genome are very similar to the target sequence. OTEs are more likely to occur in these areas, which is why PCR methods are mainly used to search for OTEs in these areas [19]. The rest of the genome is frequently not checked for possible errors induced by the gene scissors. Determining the off-target areas depends strongly on the computer program settings, the quality of the reference genome and the experience of the respective scientist.
Other experimental parameters can also influence the occurrence of OTEs: for optimal function, CRISPR/Cas gene scissors must be present in the cell in an amount that is precisely matched to the cell type in question. If the concentration of the gene scissors is too high, more off-target effects can occur; if it is too low, the desired effects are not achieved. Thus,

a suitable ratio must be found between the efficiency to effect a change in the target sequence and the occurrence of off-target effects.

Studies investigating the genome of genome-edited plants with WGS methods show that a comprehensive reference genome is particularly important to examine all off-target areas. In one study, a genome-edited cotton was examined for the presence of OTEs using WGS methods [21]. It was shown that the genome of the initial plants should be used as a reference genome to identify natural genetic variants and to distinguish them from newly occurring OTEs. In addition, the available reference genomes of plants often do not reflect the whole range of genetic variants present in different plant lines, so off-target regions, which the gene scissors might confuse with the target sequence, may be present in the genome. Thus, without a WGS analysis of the genome of the original plant, additional off-target regions may remain undetected.
Plants with a very complex genome and frequently repeated DNA sequences still pose a major challenge for WGS methods. Such regions often cannot be clearly sequenced. In a study on a genome-edited maize, for example, not all off-target regions could be investigated [20]. One OTE was detected at a previously calculated off-target area in a rice variety [22]. Due to the small number of studies and in part still incomplete data sets, it cannot be concluded that the OTEs caused by the gene scissors in off-target areas are negligible. For this to be the case, WGS procedures must be systematically carried out and evaluated in many different plant species, and the reference genomes and the WGS procedures must themselves be further improved.

It is often questioned that CRISPR/Cas-induced OTEs cannot be clearly distinguished from naturally occurring mutations or process-induced mutations (i.e. somaclonal variations). As already explained above, WGS analyses of the initial plants can identify genetic variants that are already present in the genome of the parental line. In addition, an existing PAM sequence in the vicinity of OTEs can provide information on whether this new change in DNA might be caused by the gene scissors. At off-target regions, the gene scissors need a PAM sequence nearby to cut the DNA sequence in front of it. This aspect has already been used to identify CRISPR/Cas-induced OTEs [21].

On-target-effects

On-target effects are unintended changes to the genome of target organisms that can occur when using the CRISPR/Cas gene scissors (see Figure 1). On-target effects are unintended rearrangements of the region at or around the target sequence as well as the unintended insertion of additional DNA fragments at the target site. Such unintended effects have already been described at target sites of CRISPR/Cas, but can also occur at unintended off-target areas.

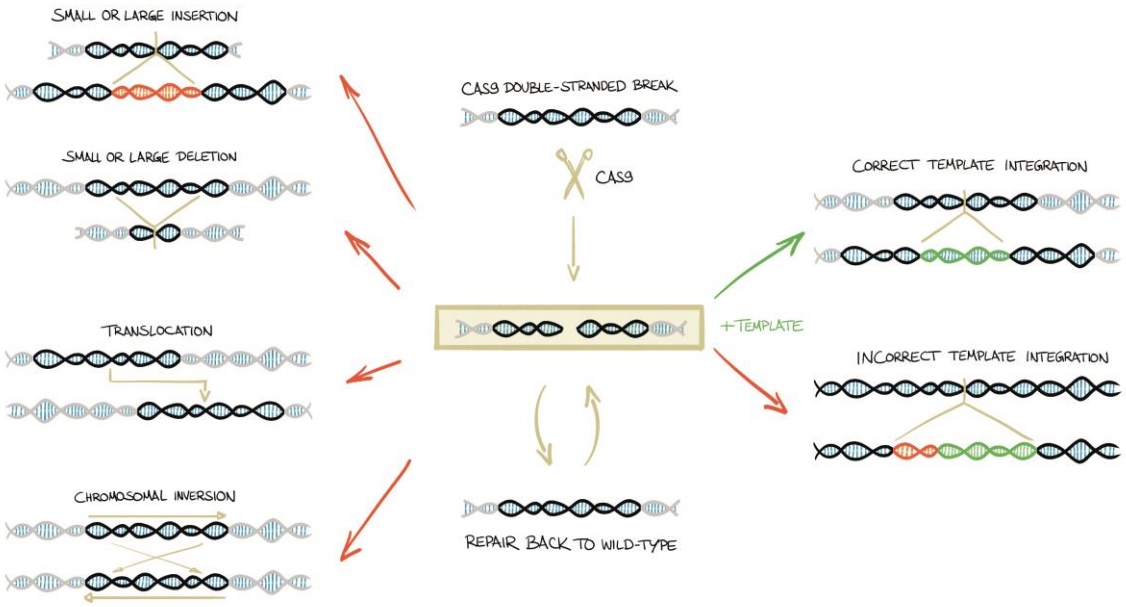

Figure 1: Schematic representation of on-target effects.
On-target effects are unintended rearrangements at or near the target sequence, such as deletions, insertions, translocations or inversions as well as the unintended insertion of additional DNA fragments at the target region, e.g. if DNA templates are used and incorrectly incorporated.

Unintended rearrangements in the genome are, on the one hand, large or small deletions (i.e. pieces of DNA can be deleted). On the other hand, large or small insertions (i.e. additional DNA sequences) can be inserted into the genome. Furthermore, the DNA region of the target sequence and adjacent regions can also be inserted at another location in the genome. This is called a translocation. For translocations to occur, several double-strand breaks must occur simultaneously in the cell and be spatially very close to each other, thus enabling an exchange of DNA regions. In addition, the orientation of the DNA target region can be reversed, this is called inversion. On-target effects have so far been studied and described mainly in human cells and mouse cells [23-26]. In plants, there are only a few studies that have examined the target sequence explicitly for on-target effects after the application of gene scissors [27; 3]. However, the methods that can be used to search for on-target effects without restrictions [28; 29] are not yet applied in plants as a standard procedure [30]. In that regard, it is important to investigate such unintended rearrangements of the DNA sequence and their consequences during risk assessment. There is already scientific work that specifically triggers such rearrangements through CRISPR/Cas in order to achieve their research goals [31-33].

Analyses of WGS data from human and mouse cells provide an overview of how often and which DNA rearrangements occur when CRISPR/Cas is applied [34; 35]: deletions at the

target sequence occur most frequently, followed by insertions and, to a lesser extent, translocations.
Large deletions are mainly caused by a repair mechanism called MMEJ (i.e. microhomology-mediated end joining), which is activated as an alternative to NHEJ repair [36].
Large insertions can occur both at the target sequence and elsewhere in the genome and are often integrations of the transport vector. A vector, also called a plasmid, is used during a CRISPR/Cas application to introduce the DNA to form the gene scissors into the cells.
Translocations occur when several double-strand DNA breaks occur simultaneously in a cell and come very close to each other spatially. Multiple double-strand breaks can occur, for example, through multiplexing, i.e. during the simultaneous modification of several different target sequences, but also when the gene scissors cut at an unintended location in the genome or through naturally occurring double-strand breaks. Most translocations are found on the same chromosome as the target sequence of the gene scissors. However, translocations can also be found on other chromosomes. In addition, translocations also occur at off-target regions [35].
Complex, unintended changes in the genome have also been reported in human cells [37; 38]. For example, parts or even whole chromosomes were deleted or massively restructured in a process called chromothripsis. These effects are triggered by the induced cut of the gene scissors at the target sequence or at off-target regions. In plants, this link between the use of CRISPR/Cas and chromothripsis has not yet been investigated.

In SDN-2 and SDN-3 applications, the DNA templates to be incorporated by HDR in the target site may either be inserted unintentionally multiple times or incorrectly in the target site. SDN-1 applications are mostly used in plants [39], as the efficiency of SDN-2 and SDN-3 applications is quite low. Therefore, it is mainly study outcomes from human and animal cells that report such unintended effects: in mouse studies, it has been shown that when CRISPR/Cas9 is applied, DNA templates are incorporated multiple times within target sites and cannot be detected by simple PCR methods [40]. In a large proportion of the offspring of the genome-edited mice, the DNA template was incorporated multiple times in succession within the target sequence. The same effect occurred repeatedly in further experiments and seems to be a general phenomenon in the application of CRISPR/Cas9. Surprisingly, it was not possible to detect unintended multiple insertions of the DNA template using normal PCR methods. Only by using a suitable combination of different methods was it possible to detect such additional insertions.

Integration of foreign DNA

In addition to the insertion of the DNA template and the insertion of the transport vector for introducing the DNA to form the gene scissors, DNA from other sources can also be inserted at the target site. Unintentionally integrated DNA sequences at the target sites were

investigated in a study on mouse cells [41]. In some cases, large pieces of DNA were incorporated. Most of this came from the vector used to introduce the DNA into the cells. Another part of the DNA fragments came from the genome of the mouse itself. A smaller proportion came from the genetic material of the bacterium *E. coli*, with which the gene scissors DNA was amplified in the laboratory before it was introduced into the cells. Surprisingly, a small proportion of the additional DNA fragments came from the genome of cattle. The explanation for this is that when working with cell cultures, some bovine serum must be added to the culture medium of the mouse cells so that the cells can divide and survive. The bovine DNA is introduced into the mouse cells via small vesicles. This raises aspects important for risk assessment in respect to CRISPR/Cas applications:
Double-strand breaks favor the incorporation of foreign DNA, which is a general problem in the basic concept of genetic scissors. Foreign DNA fragments that are introduced into the target cells as contaminations during work with cell cultures can, for example, theoretically also be incorporated at off-target areas and naturally occurring double-strand breaks in the genome.

Unintended formation of new gene products

Various changes can be induced at the target site when gene scissors are used. The target site can be altered in such a way that no more mRNAs are formed, thus preventing the formation of the corresponding protein. In addition, certain changes to the DNA sequence can lead to mRNAs being formed with an altered composition, for example, particular areas of the gene being deleted. The resulting truncated proteins can then no longer perform certain functions.

However, new mRNAs can also be formed unintentionally, which can be overlooked by the user and from which new proteins can be formed [42]. For example, the intended changes at the target site can lead to an effect called exon skipping. In exon skipping, mRNAs can be assembled differently than planned even if the intended changes at the target site are induced. This can result in the formation of shortened mRNAs. The proteins formed from such mRNAs are then also shorter but can still carry out functions in the cell. Sometimes the original proteins continue to be formed in the cell even though the underlying gene has been knocked out [43-45; 42].

In addition, the intended changes can lead to the development of frameshift mutations. Frameshift means that the reading frame of the DNA sequence from a gene shifts and the gene can be read in a completely different way. As a result, new mRNAs and proteins can be formed, which can also perform new functions in cell metabolism. Such unintended effects are often not detectable at the DNA level. Genome-edited plants need to be investigated for such effects because changes in protein composition are an important consideration when

assessing environmental and food safety. For example, allergens may arise that can trigger an immune response.

CRISPR/Cas is used in many cases to induce complex changes in the genome of plants. For example, several members of a gene family can be modified simultaneously [46; 47]. Such complex SDN-1 applications come with inherent risks: at each individual target sequence where the gene scissors induce a change, new gene products can be formed that may have an unintended effect. Such risks need to be investigated in genome-edited plants, for example, through the use of metabolomics and genomics techniques.

SDN-2 and SDN-3 applications

In SDN-2 and SDN-3 gene scissor applications, DNA templates are introduced into cells to be inserted at the respective target sequence in the genome of the target organism. Unintended changes can occur in connection with the introduction of DNA templates. As already described, the DNA template can be incorrectly incorporated into the genome several times or only parts of it might be integrated. This can happen directly at the target sequence and/or at completely different areas of the genome (e.g. at off-target areas).

In addition, studies on mouse cells show that the composition of epigenetic markers in the region of the target sequences changes when DNA templates are inserted there with CRISPR/Cas [48]. It was not investigated whether the gene activity in the genome-edited mice is influenced by the altered epigenetic markers, but it is feasible that the gene activity of genes in the immediate vicinity of the intended modification may be affected.
In plants, there has been work published on the distribution of epigenetic markers after SDN-1 gene scissor applications [49]. The results show that the distribution of epigenetic markers is not altered in these SDN-1 applications. Whether this is also the case with SDN-2 and SDN-3 applications in plants has not yet been investigated.

Unintended effects of other CRISPR/Cas variants

The background paper on the potential of CRISPR/Cas presents some examples from scientific literature on advances in gene scissor technology. The already known unwanted changes and still uninvestigated risks of these applications are described in more detail here.

Epigenome editing

The CRISPR/Cas system can be used to change epigenetic markers, i.e. gene activity. This does not involve changing the DNA, but does involve changing the composition of epigenetic

markers, which can have an influence on gene activity. These changes can lead to the activation or silencing of certain genes.
The gene scissors themselves were modified so that they were no longer able to cut the DNA (i.e. dead Cas9/dCas9), but can still bind to the DNA and recognize the target sequence. Enzymes are attached to the gene scissors to change the composition of the epigenetic markers at the target sequence [50-52].
However, it has been shown that these enzymes become non-specifically active across the entire genome and can cause epigenetic changes at unwanted sites in the genome [53-55]. Gene expression can thus be influenced at completely different genomic sites, which can fundamentally change the characteristics of the affected organisms without this being recognizable at the DNA level. Scientists are working on advancing epigenome editing to reduce off-target effects [56]. So far, epigenome editing applications have been used in human and animal cells, with only a few being used in plants.

Base editing

Base editing involves attaching enzymes to the non-cutting CRISPR/dCas9 gene scissors to convert certain base types at the target sequence of the DNA into other base types, i.e. cytosine into thymine and adenine into guanine. Unintended changes can occur at different levels of the cell during this process:
Firstly, it has been shown in rice that certain base editors can cause unintended changes in the genome [57]. The unintended changes in the DNA can occur across the entire genome and not only at specific DNA regions. Therefore, the entire genome of genome-edited plants should be assessed, using whole genome sequencing methods. The non-specific, unwanted changes at random locations in the genome show how imprecise and inaccurate base editors are.
It was also shown that base editors can unintentionally change mRNA molecules [58]. In the process, several cytosines were unintentionally converted into uracils, resulting in proteins being formed differently. These unintended C-to-U changes at the RNA level occur independently of unintended changes at the DNA level. The unintended changes in mRNA described here are caused by an additional activity of base editors. This type of unintended change at the RNA level is also found with base editors that cause a conversion from adenine to guanine [58].

Prime editing

Prime editing is a complex development of the gene scissors with various associated risks (a detailed description of this application can be found in the background paper on the possibilities of CRISPR/Cas).

Unintended changes can occur at off-target regions, i.e. at genome sites similar to the actual target sequence, but also at completely different sites of the genome. Furthermore, the translated RNA templates can be incorporated several times in one region of the genome and/or at other genomic regions.
In addition, the reverse transcriptase, the translator in prime editing, can translate other RNA molecules, such as mRNAs, which occur naturally in the cell, into DNA and these can then be unintentionally incorporated into the genome.
If the genetic information of the prime editor is introduced into the genome of the plant cells, parts of it can also be incorporated into the genome at a later time: the genetic information for the prime editor is first incorporated into the plant's genome and then expressed (i.e. formed) in the plant cells. The mRNA formed in the process is then translated into the actual prime editor (i.e. a protein complex). An intermediate step during that process can lead to unintended effects: the formed prime editor mRNA can be translated back into DNA by already formed prime editors and this DNA can then be unintentionally incorporated into the genome.

In summary, many unintended changes can occur during prime editing. Therefore, such organisms should be examined with suitable methods, such as WGS or other next-generation sequencing (i.e. NGS) methods, to check whether unintended changes have been made somewhere in the genome.

So far, only a few studies have investigated the risks of prime editing or looked for unintended changes in the genome [59; 60].
In a study on human cells, off-target effects were investigated in the whole genome after the prime editing application. According to the study, few off-target effects occur at other sites in the DNA [59]. Another study on rice looked for unintended changes in the genome using WGS after various target genes were altered with a prime editor [60]. Few off-target effects were found at sites in the genome similar to the target site. No off-target effects were found at areas of the genome that did not resemble the target sequence. No other unintended changes were found in the genome in this analysis. Further studies still need to be done in other crop species, especially those with a more complex genome than rice, to investigate the risks of prime editing in more detail.

How should genome-edited plants be examined?

CRISPR/Cas gene scissors applications are very diverse and involve a multi-step process. Different molecular biology techniques are combined in this process, each of which is associated with specific risks. The occurrence of such risks should be reduced as far as possible in applications of genome editing in plants and animals. The genome-edited

organisms must be examined in a comprehensive risk analysis to identify unwanted and unintended changes to the genome and downstream to the RNA, proteins and metabolic products. The procedures in the multi-step process that were used to modify the plants with the gene scissors should, in the best-case scenario, be known in order to select the most suitable analytical procedures. The following section sets out different analytical methods with which unintended changes can be detected.

PCR analyses for the investigation of off-target areas

Classical PCR analyses are the most frequently used applications to look for unintended off-target effects in some parts of the genome [19]. Computer programs are used to determine areas of the genome that are very similar to the actual target sequence and where the gene scissors can unintentionally cut and induce changes. These regions of the genome are then examined in the genome-edited, regenerated plants after gene scissor applications, using PCR analyses to determine the exact sequence of the DNA bases and decode possible off-target areas. In most studies, however, only a few off-target regions are examined.

Omics methods

There are a number of methods with which the various molecular levels of a cell can be analyzed. These methods are known collectively as -omics. Omics can be used to detect unintended changes that can result from the use of gene scissors as well as other methods used in the multi-step process. In addition, methods for analyzing the epigenome, i.e. the totality of all epigenetic markers that control gene activity, as well as methods that decipher the composition of an organism's microbiome, are also often counted as omics. These methods are not explained further in this background paper.

1. Analytical procedures to examine the DNA sequence of an organism, for example, by genome sequencing, are collectively known as genomics. Genomics includes whole genome sequencing (WGS) methods, with which the entire genome of an organism can be sequenced and the exact sequence of the bases of the DNA can be decoded. Comparing a reference genome and the sequenced genome of the genome-edited plant makes it is possible to see which changes CRISPR/Cas caused in the genome. The intended, the unintended and unknown changes in the DNA sequence can be determined by applying WGS. Nevertheless, WGS methods are often not detailed enough to investigate plants with a complex genome as, for example, such plants have many repetitive DNA sequences in their genome. This means that it is sometimes not possible to ensure that the entire genome is actually investigated. In addition, complete reference genomes of less well-studied plants, such as orphan crops, are often still not available.

In recent years, other methods have been developed that are also counted as genomics; these methods can be used to examine the genome specifically for off-target or on-target effects. These methods are also known as NGS or next-generation sequencing methods. One of these NGS methods is, for example, Digenome-Seq [61] and is used for genome-wide analysis of off-target effects in plants [62].

2. The generic term transcriptomics covers analytical methods that can be used to determine the composition of all RNA molecules in a cell at a specific point in time. On the one hand, RNA is the intermediate product starting from the DNA sequence for the formation of proteins. On the other, it can also have regulatory functions (e.g. miRNA, long non-coding RNA). Transcriptomics methods can be used to determine the totality of all RNAs after a cell is modified by CRISPR/Cas and to compare them with unmodified cells to detect unintended effects, such as changes in the gene expression of non-target genes or differently composed mRNAs.

3. Proteomics is the analysis of all proteins present in a cell at a certain point in time and enables the comparison of modified with unmodified cells. This can identify unwanted side effects, such as incorrectly assembled proteins due to exon skipping. Proteomics include, among others, mass spectrometry methods.

4. Metabolomics include procedures to analyze the composition of the metabolic products of a cell (i.e. the metabolome). Metabolomics methods allow molecules involved in metabolic processes to be detected within a cell and enable comparison with a corresponding reference. If a change leads to unforeseen effects in metabolism, an imbalance can arise and in plants, for example, the formation of metabolic products can be disturbed. This can have an influence on the genome-edited organism itself, but also on other organisms in the ecosystem. A metabolomics analysis is extensive and can break down the composition of precursor molecules, intermediates, end products and their derivatives in many different metabolic pathways. The more information is known about the altered metabolic pathway, the easier and more specifically the composition of the metabolites can be analyzed and, if necessary, the detection of specific metabolites can be provided. If, for example, the gene scissors interfere with the fatty acid metabolism and an enzyme for the formation of a certain fatty acid is knocked out, then the composition of all cellular fats can be investigated with the help of lipidomics methods. If it is not known which target gene and, thus, which metabolic pathway has been altered, an untargeted metabolomics analysis must be carried out that is not intended to investigate a specific metabolite type (e.g. fats, sugars, amino acids, etc.). The analysis is then broader in its scope, but less extensive in the different metabolite

species. Metabolomics has been proposed several times for the evaluation of genome-edited plants [63; 64].

Genomics and metabolomics methods, in particular, should be used for the risk assessment of genome-edited plants, as they can be used to detect many already known unintended changes.

As of June 2021

References

1.Biswas S, Tian J, Li R, Chen X, Luo Z, Chen M, Zhao X, Zhang D, Persson S, Yuan Z, Shi J (2020) Investigation of CRISPR/Cas9-induced SD1 rice mutants highlights the importance of molecular characterization in plant molecular breeding. J Genet Genomics. doi:10.1016/j.jgg.2020.04.004

2.Michno JM, Virdi K, Stec AO, Liu J, Wang X, Xiong Y, Stupar RM (2020) Integration, abundance, and transmission of mutations and transgenes in a series of CRISPR/Cas9 soybean lines. BMC Biotechnol 20 (1):10. doi:10.1186/s12896-020-00604-3

3.Li Z, Liu ZB, Xing A, Moon BP, Koellhoffer JP, Huang L, Ward RT, Clifton E, Falco SC, Cigan AM (2015) Cas9-guide RNA directed genome editing in soybean. Plant Physiol 169 (2):960-970. doi:10.1104/pp.15.00783

4.Jupe F, Rivkin AC, Michael TP, Zander M, Motley ST, Sandoval JP, Slotkin RK, Chen H, Castanon R, Nery JR, Ecker JR (2019) The complex architecture and epigenomic impact of plant T-DNA insertions. PLoS Genet 15 (1):e1007819. doi:10.1371/journal.pgen.1007819

5.Jupe M, Rivkin, Zander , Motley , Sandoval , Slotkin , Chen , Castagnon , Nery , J. R. Ecker (2018) The complex architecture of plant transgene insertions. BioRxiv (Preprint)

6.Gelvin SB (2017) Integration of *Agrobacterium* T-DNA into the plant genome. Annu Rev Genet 51:195-217. doi:10.1146/annurev-genet-120215-035320

7.Liu J, Nannas NJ, Fu F-F, Shi J, Aspinwall B, Parrott WA, Dawe RK (2019) Genome-Scale Sequence Disruption Following Biolistic Transformation in Rice and Maize. The Plant cell 31 (2):368-383. doi:10.1105/tpc.18.00613

8.Zhu C, Wu J, He C (2010) Induction of chromosomal inversion by integration of T-DNA in the rice genome. Journal of Genetics and Genomics 37 (3):189-196. doi:https://doi.org/10.1016/S1673-8527(09)60037-0

9.Clark KA, Krysan PJ (2010) Chromosomal translocations are a common phenomenon in Arabidopsis thaliana T-DNA insertion lines. Plant J 64 (6):990-1001. doi:10.1111/j.1365-313X.2010.04386.x

10.Nacry P, Camilleri C, Courtial B, Caboche M, Bouchez D (1998) Major chromosomal rearrangements induced by T-DNA transformation in Arabidopsis. Genetics 149 (2):641-650

11.Makarevitch I, Svitashev SK, Somers DA (2003) Complete sequence analysis of transgene loci from plants transformed via microprojectile bombardment. Plant Mol Biol 52 (2):421-432. doi:10.1023/a:1023968920830

12.Svitashev SK, Pawlowski WP, Makarevitch I, Plank DW, Somers DA (2002) Complex transgene locus structures implicate multiple mechanisms for plant transgene rearrangement. Plant J 32 (4):433-445. doi:10.1046/j.1365-313x.2002.01433.x

13.Pawlowski WP, Somers DA (1996) Transgene inheritance in plants genetically engineered by microprojectile bombardment. Mol Biotechnol 6 (1):17-30. doi:10.1007/bf02762320

14.Carlson DF, Lancto CA, Zang B, Kim ES, Walton M, Oldeschulte D, Seabury C, Sonstegard TS, Fahrenkrug SC (2016) Production of hornless dairy cattle from genome-edited cell lines. Nat Biotechnol 34 (5):479-481. doi:10.1038/nbt.3560

15.Norris AL, Lee SS, Greenlees KJ, Tadesse DA, Miller MF, Lombardi HA (2020) Template plasmid integration in germline genome-edited cattle. Nat Biotechnol 38 (2):163-164. doi:10.1038/s41587-019-0394-6

16.Cho SW, Kim S, Kim Y, Kweon J, Kim HS, Bae S, Kim JS (2014) Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome Res 24 (1):132-141. doi:10.1101/gr.162339.113

17.Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, Sander JD (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol 31 (9):822-826. doi:10.1038/nbt.2623
18.Modrzejewski D, Hartung F, Lehnert H, Sprink T, Kohl C, Keilwagen J, Wilhelm R (2020) Which Factors Affect the Occurrence of Off-Target Effects Caused by the Use of CRISPR/Cas: A Systematic Review in Plants. Frontiers in Plant Science 11 (1838). doi:10.3389/fpls.2020.574959
19.Modrzejewski D, Hartung, F., Sprink, T., Krause, D., Kohl, C., Wilhelm R. (2019) What is the available evidence for the range of applications of genome-editing as a new tool for plant trait modification and the potential occurrence of associated off-target effects: a systematic map. Environ Evid 8 (27). doi:10.1186/s13750-019-0171-5
20.Lee K, Zhang Y, Kleinstiver BP, Guo JA, Aryee MJ, Miller J, Malzahn A, Zarecor S, Lawrence-Dill CJ, Joung JK, Qi Y, Wang K (2019) Activities and specificities of CRISPR/Cas9 and Cas12a nucleases for targeted mutagenesis in maize. Plant Biotechnol J 17 (2):362-372. doi:10.1111/pbi.12982
21.Li J, Manghwar H, Sun L, Wang P, Wang G, Sheng H, Zhang J, Liu H, Qin L, Rui H, Li B, Lindsey K, Daniell H, Jin S, Zhang X (2019) Whole genome sequencing reveals rare off-target mutations and considerable inherent genetic or/and somaclonal variations in CRISPR/Cas9-edited cotton plants. Plant Biotechnol J 17 (5):858-868. doi:10.1111/pbi.13020
22.Tang X, Liu G, Zhou J, Ren Q, You Q, Tian L, Xin X, Zhong Z, Liu B, Zheng X, Zhang D, Malzahn A, Gong Z, Qi Y, Zhang T, Zhang Y (2018) A large-scale whole-genome sequencing analysis reveals highly specific genome editing by both Cas9 and Cpf1 (Cas12a) nucleases in rice. Genome Biol 19 (1):84. doi:10.1186/s13059-018-1458-5
23.Adikusuma F, Piltz S, Corbett MA, Turvey M, McColl SR, Helbig KJ, Beard MR, Hughes J, Pomerantz RT, Thomas PQ (2018) Large deletions induced by Cas9 cleavage. Nature 560 (7717):E8-E9. doi:10.1038/s41586-018-0380-z
24.Kosicki M, Tomberg K, Bradley A (2018) Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. Nat Biotechnol 36 (8):765-771. doi:10.1038/nbt.4192
25.Weisheit I, Kroeger JA, Malik R, Klimmt J, Crusius D, Dannert A, Dichgans M, Paquet D (2020) Detection of Deleterious On-Target Effects after HDR-Mediated CRISPR Editing. Cell Rep 31 (8):107689. doi:10.1016/j.celrep.2020.107689
26.Weisheit I, Kroeger JA, Malik R, Wefers B, Lichtner P, Wurst W, Dichgans M, Paquet D (2021) Simple and reliable detection of CRISPR-induced on-target effects by qgPCR and SNP genotyping. Nature Protocols. doi:10.1038/s41596-020-00481-2
27.Biswas S, Tian J, Li R, Chen X, Luo Z, Chen M, Zhao X, Zhang D, Persson S, Yuan Z, Shi J (2020) Investigation of CRISPR/Cas9-induced SD1 rice mutants highlights the importance of molecular characterization in plant molecular breeding. Journal of Genetics and Genomics 47 (5):273-280. doi:https://doi.org/10.1016/j.jgg.2020.04.004
28.Giraldo PA, Shinozuka H, Spangenberg GC, Smith KF, Cogan NOI (2021) Rapid and Detailed Characterization of Transgene Insertion Sites in Genetically Modified Plants via Nanopore Sequencing. Frontiers in Plant Science 11 (2292). doi:10.3389/fpls.2020.602313
29.Höijer I, Johansson J, Gudmundsson S, Chin C-S, Bunikis I, Häggqvist S, Emmanouilidou A, Wilbe M, den Hoed M, Bondeson M-L, Feuk L, Gyllensten U, Ameur A (2020) Amplification-free long-read sequencing reveals unforeseen CRISPR-Cas9 off-target activity. Genome Biology 21 (1):290. doi:10.1186/s13059-020-02206-w
30.Hahn F, Nekrasov V (2019) CRISPR/Cas precision: do we need to worry about off-targeting in plants? Plant Cell Rep 38 (4):437-441. doi:10.1007/s00299-018-2355-9
31.Huang TK, Puchta H (2021) Novel CRISPR/Cas applications in plants: from prime editing to chromosome engineering. Transgenic Res. doi:10.1007/s11248-021-00238-x
32.Beying N, Schmidt C, Pacher M, Houben A, Puchta H (2020) CRISPR-Cas9-mediated induction of heritable chromosomal translocations in Arabidopsis. Nat Plants. doi:10.1038/s41477-020-0663-x
33.Schmidt C, Fransz P, Rönspies M, Dreissig S, Fuchs J, Heckmann S, Houben A, Puchta H (2020) Changing local recombination patterns in Arabidopsis by CRISPR/Cas mediated chromosome engineering. Nature Communications 11 (1):4418. doi:10.1038/s41467-020-18277-z
34.Kosicki M, Allen F, Bradley A (2020) Cas9-induced large deletions and small indels are controlled in a convergent fashion. bioRxiv:2020.2008.2005.216739. doi:10.1101/2020.08.05.216739
35.Liu M, Zhang W, Xin C, Yin J, Shang Y, Ai C, Li J, Meng F-l, Hu J (2021) Global detection of DNA repair outcomes induced by CRISPR-Cas9. bioRxiv:2021.2002.2015.431335. doi:10.1101/2021.02.15.431335

36.Owens DDG, Caulder A, Frontera V, Harman JR, Allan AJ, Bucakci A, Greder L, Codner GF, Hublitz P, McHugh PJ, Teboul L, de Bruijn MFTR (2019) Microhomologies are prevalent at Cas9-induced larger deletions. Nucleic Acids Research 47 (14):7402-7417. doi:10.1093/nar/gkz459
37.Zuccaro MV, Xu J, Mitchell C, Marin D, Zimmerman R, Rana B, Weinstein E, King RT, Palmerola KL, Smith ME, Tsang SH, Goland R, Jasin M, Lobo R, Treff N, Egli D (2020) Allele-Specific Chromosome Removal after Cas9 Cleavage in Human Embryos. Cell. doi:https://doi.org/10.1016/j.cell.2020.10.025
38.Leibowitz ML, Papathanasiou S, Doerfler PA, Blaine LJ, Sun L, Yao Y, Zhang C-Z, Weiss MJ, Pellman D (2021) Chromothripsis as an on-target consequence of CRISPR–Cas9 genome editing. Nature Genetics. doi:10.1038/s41588-021-00838-7
39.Modrzejewski D, Hartung, F., Sprink, T., Menz, J., Kohl, C., Delventhal, R., Wilhelm, R. (2020) Aktualisierung der Übersicht über Nutz- und Zierpflanzen, die mittels neuer molekularbiologischer Techniken für die Bereiche Ernährung, Landwirtschaft und Gartenbau erzeugt wurden – marktorientierte Anwendungen. . https://www.bmel.de/SharedDocs/Downloads/DE/_Landwirtschaft/Gruene-Gentechnik/NMT_Uebersicht-Zier-Nutzpflanzen.pdf?__blob=publicationFile&v=3.
40.Skryabin BV, Kummerfeld DM, Gubar L, Seeger B, Kaiser H, Stegemann A, Roth J, Meuth SG, Pavenstadt H, Sherwood J, Pap T, Wedlich-Soldner R, Sunderkotter C, Schwartz YB, Brosius J, Rozhdestvensky TS (2020) Pervasive head-to-tail insertions of DNA templates mask desired CRISPR-Cas9-mediated genome editing events. Sci Adv 6 (7):eaax2941. doi:10.1126/sciadv.aax2941
41.Ono R, Yasuhiko Y, Aisaki KI, Kitajima S, Kanno J, Hirabayashi Y (2019) Exosome-mediated horizontal gene transfer occurs in double-strand break repair during genome editing. Commun Biol 2:57. doi:10.1038/s42003-019-0300-2
42.Tuladhar R, Yeu Y, Tyler Piazza J, Tan Z, Rene Clemenceau J, Wu X, Barrett Q, Herbert J, Mathews DH, Kim J, Hyun Hwang T, Lum L (2019) CRISPR-Cas9-based mutagenesis frequently provokes on-target mRNA misregulation. Nat Commun 10 (1):4056. doi:10.1038/s41467-019-12028-5
43.Kapahnke M, Banning A, Tikkanen R (2016) Random splicing of several exons caused by a single base change in the target exon of CRISPR/Cas9 mediated gene knockout. Cells 5 (4). doi:10.3390/cells5040045
44.Lalonde S, Stone OA, Lessard S, Lavertu A, Desjardins J, Beaudoin M, Rivas M, Stainier DYR, Lettre G (2017) Frameshift indels introduced by genome editing can lead to in-frame exon skipping. PLoS One 12 (6):e0178700. doi:10.1371/journal.pone.0178700
45.Mou H, Smith JL, Peng L, Yin H, Moore J, Zhang XO, Song CQ, Sheel A, Wu Q, Ozata DM, Li Y, Anderson DG, Emerson CP, Sontheimer EJ, Moore MJ, Weng Z, Xue W (2017) CRISPR/Cas9-mediated genome editing induces exon skipping by alternative splicing or exon deletion. Genome Biol 18 (1):108. doi:10.1186/s13059-017-1237-8
46.Sanchez-Leon S, Gil-Humanes J, Ozuna CV, Gimenez MJ, Sousa C, Voytas DF, Barro F (2018) Low-gluten, nontransgenic wheat engineered with CRISPR/Cas9. Plant Biotechnol J 16 (4):902-910. doi:10.1111/pbi.12837
47.Schachtsiek J, Stehle F (2019) Nicotine-free, nontransgenic tobacco (*Nicotiana tabacum* l.) edited by CRISPR-Cas9. Plant Biotechnol J 17 (12):2228-2230. doi:10.1111/pbi.13193
48.Farris MH, Texter PA, Mora AA, Wiles MV, Mac Garrigle EF, Klaus SA, Rosfjord K (2020) Detection of CRISPR-mediated genome modifications through altered methylation patterns of CpG islands. BMC Genomics 21 (1):856. doi:10.1186/s12864-020-07233-2
49.Lee JH, Mazarei M, Pfotenhauer AC, Dorrough AB, Poindexter MR, Hewezi T, Lenaghan SC, Graham DE, Stewart CN (2020) Epigenetic Footprints of CRISPR/Cas9-Mediated Genome Editing in Plants. Frontiers in Plant Science 10 (1720). doi:10.3389/fpls.2019.01720
50.Hilton IB, D'Ippolito AM, Vockley CM, Thakore PI, Crawford GE, Reddy TE, Gersbach CA (2015) Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nat Biotechnol 33 (5):510-517. doi:10.1038/nbt.3199
51.Huang YH, Su J, Lei Y, Brunetti L, Gundry MC, Zhang X, Jeong M, Li W, Goodell MA (2017) DNA epigenome editing using CRISPR-Cas SunTag-directed DNMT3A. Genome Biol 18 (1):176. doi:10.1186/s13059-017-1306-z
52.Fal K, Tomkova D, Vachon G, Chabouté ME, Berr A, Carles CC (2021) Chromatin Manipulation and Editing: Challenges, New Technologies and Their Use in Plants. Int J Mol Sci 22 (2). doi:10.3390/ijms22020512
53.Galonska C, Charlton J, Mattei AL, Donaghey J, Clement K, Gu H, Mohammad AW, Stamenova EK, Cacchiarelli D, Klages S, Timmermann B, Cantz T, Scholer HR, Gnirke A, Ziller MJ, Meissner A

(2018) Genome-wide tracking of dCas9-methyltransferase footprints. Nat Commun 9 (1):597. doi:10.1038/s41467-017-02708-5
54.Liu XS, Wu H, Ji X, Stelzer Y, Wu X, Czauderna S, Shu J, Dadon D, Young RA, Jaenisch R (2016) Editing DNA Methylation in the Mammalian Genome. Cell 167 (1):233-247.e217. doi:10.1016/j.cell.2016.08.056
55.Stepper P, Kungulovski G, Jurkowska RZ, Chandra T, Krueger F, Reinhardt R, Reik W, Jeltsch A, Jurkowski TP (2017) Efficient targeted DNA methylation with chimeric dCas9-Dnmt3a-Dnmt3L methyltransferase. Nucleic acids research 45 (4):1703-1713. doi:10.1093/nar/gkw1112
56.Pflueger C, Tan D, Swain T, Nguyen T, Pflueger J, Nefzger C, Polo JM, Ford E, Lister R (2018) A modular dCas9-SunTag DNMT3A epigenome editing system overcomes pervasive off-target activity of direct fusion dCas9-DNMT3A constructs. Genome Res 28 (8):1193-1206. doi:10.1101/gr.233049.117
57.Jin S, Zong Y, Gao Q, Zhu Z, Wang Y, Qin P, Liang C, Wang D, Qiu JL, Zhang F, Gao C (2019) Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. Science 364 (6437):292-295. doi:10.1126/science.aaw7166
58.Grünewald J, Zhou R, Iyer S, Lareau CA, Garcia SP, Aryee MJ, Joung JK (2019) CRISPR DNA base editors with reduced RNA off-target and self-editing activities. Nat Biotechnol 37 (9):1041-1048. doi:10.1038/s41587-019-0236-6
59.Kim DY, Moon SB, Ko J-H, Kim Y-S, Kim D (2020) Unbiased investigation of specificities of prime editing systems in human cells. Nucleic Acids Research 48 (18):10576-10589. doi:10.1093/nar/gkaa764
60.Jin S, Lin Q, Luo Y, Zhu Z, Liu G, Li Y, Chen K, Qiu J-L, Gao C (2021) Genome-wide specificity of prime editors in plants. Nature Biotechnology. doi:10.1038/s41587-021-00891-x
61.Kim D, Bae S, Park J, Kim E, Kim S, Yu HR, Hwang J, Kim J-I, Kim J-S (2015) Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. Nature Methods 12 (3):237-243. doi:10.1038/nmeth.3284
62.Xu W, Fu W, Zhu P, Li Z, Wang C, Wang C, Zhang Y, Zhu S (2019) Comprehensive Analysis of CRISPR/Cas9-Mediated Mutagenesis in Arabidopsis thaliana by Genome-wide Sequencing. Int J Mol Sci 20 (17). doi:10.3390/ijms20174125
63.Enfissi EMA, Drapal M, Perez-Fons L, Nogueira M, Berry HM, Almeida J, Fraser PD (2021) New Plant Breeding Techniques and their regulatory implications: An opportunity to advance metabolomics approaches. Journal of Plant Physiology:153378. doi:https://doi.org/10.1016/j.jplph.2021.153378
64.Fraser PD, Aharoni A, Hall RD, Huang S, Giovannoni JJ, Sonnewald U, Fernie AR (2020) Metabolomics should be deployed in the identification and characterization of gene-edited crops. Plant J. doi:10.1111/tpj.14679