TDL.ORG TEXAS DIGITAL LIBRARY

Journal of Regulatory Science

http:\\journalofregulatoryscience.org

Regulatory Science

Journal of Regulatory Science 9(1) (2021) 1–15

Regulatory Assessment of Off-Target Changes and Spurious DNA Insertions in Gene-Edited Organisms for Agri-Food Use

Martin A. Lema[a,*]

[a] *Universidad Nacional de Quilmes, Departamento de Ciencia y Tecnología & Maestría en Ciencia, Tecnología y Sociedad, Argentina*

Abstract

Worldwide, an increasing number of regulatory systems have begun to consider applications for the authorization of activities involving gene-edited organisms for agri-food use. Although a handful of countries have made advances in establishing regulatory criteria and gathering practical experience in this regard, there is still a general need for regulatory cooperation concerning capacity building and development of harmonized criteria. Consequently, many biotechnology regulators need to quickly become more acquainted with the numerous technological possibilities enclosed under the concept of "gene editing", and to incorporate criteria for their regulatory assessment. This article contains a simplified introduction to the state of the art in genome editing, described from a regulatory perspective. In particular, two issues of higher practical importance are covered in detail, namely, off-target effects and unintended DNA insertions. The detailed review of current evidence regarding those issues serves as the basis for proposing concrete regulatory criteria to address them.

Keywords: gene editing, genome editing, off-target, CRISPR-Cas, biotechnology regulation, new breeding techniques

Abbreviations: ACR, anti-Crispr proteins; CRISPR, clustered, regularly interspaced, short palindromic repeat; crRNA, CRISPR RNA; Cas, CRISPR-associated; GE, genetically engineered; GMO, genetically modified organism; LMO, living modified organism; NBT, new breeding techniques; ODM, oligonucleotide-directed mutagenesis; PAL, potentially affected locus/loci; PAM, protospacer adjacent motif; nt, nucleotide/s; r-DNA, recombinant-DNA; sgRNA, single-guide RNA; SDN, site-directed nuclease; TALEN, transcription activator-like effector nucleases; WGS, whole-genome sequencing; ZFNs, zinc-finger nucleases

1. Introduction

1.1. The Established Regulation of "Modern Biotechnology"

With the advent of recombinant-DNA (r-DNA) techniques applied to the genetic modification of organisms for agri-food uses, governments developed *ad hoc* regulatory frameworks for the so-called "modern biotechnology" [16, 7, 79]. Typically, such *ad hoc* regulation does not replace but supplement other regulations of broad applicability to organisms of agricultural use and the foodstuffs derived from them.

Concurrently, intergovernmental organizations developed international standards for "modern biotechnology". The main corpus of regulatory guidance in this regard includes several Guidelines from *Codex Alimentarius* [132], the text of the Cartagena Protocol on Biosafety [14] and many reference documents from the Organization for Economic Co-operation and Development [84].

These national regulations and intergovernmental documents were crafted from the early 1990s to the late 2000s. In that period, the only kind of products developed for agri-food use with the aid of recombinant-DNA technology consisted of transgenic organisms. Other agricultural "biotechnologies" also emerged in that period, such as plant micropropagation, marker-assisted breeding, animal cloning, microbial bioinputs, and mutagenesis from somaclonal variation. However, these other innovative biotechnologies do not use r-DNA, and for that reason, they were not included in the scope of the new regulations intended for transgenic organisms.

In general, national regulations and international guidelines are nowadays quite developed and harmonized in terms of their scientific basis and the principles to conduct a safety assessment of transgenic organisms. Basically, these frameworks begin with a comprehensive characterization of the artificial genetic construct inserted in a host organism and the resulting novel (or modified) traits. Such characterization constitutes a foundation for subsequently performing risk analyses based on a comparative approach with "conventional counterparts" (typically, wild-type organisms with a similar genetic background). The core of those analyses is the safety assessment of novel substances - in most cases meaning new proteins - and changes

*Corresponding author: Martin A. Lema, Email: mlema@unq.edu.ar

in composition (food/feed safety assessment), or biosafety assessment of changes in phenotype and the interactions with the environment.

In contrast to a high level of harmonization regarding the underlying scientific criteria, international guidance and national regulations are very diversified regarding the use of terms and definitions mandating which organisms and products are under their scope. Examples of this diversity include "genetically modified organism" (GMO), "living modified organism" (LMO), "modern biotechnology products", "genetically engineered (GE) organism", "organisms/products derived from recombinant-DNA", "organisms with novel traits", and "organisms containing pest components". Moreover, the term most frequently used by governmental regulations, which is "GMO", can have diverse operational definitions in different territories.

Nevertheless, these discrepancies were of little relevance when only transgenic organisms were presented to the regulators because, in most cases, they would be encompassed under any of the definitions for these terms. However, the lack of harmonization in this respect became very relevant during the last half of the past decade with the advent of gene (or genome) editing applied to species of agricultural use.

1.2. The Evolution in The Regulation of Gene-Edited Products

In the early 2010s, some regulators in the field of agricultural biotechnology became aware of innovations in breeding techniques using r-DNA technology [70]. These innovations, in many cases, can be used to generate changes in the host genome that do not result in transgenic organisms. They were called "new breeding techniques" (NBTs), and gene editing is their most prominent member.

Eventually, it became necessary to clarify the status of organisms and products obtained using these techniques as subjected, or not, to governmental regulations for agricultural biotechnology [51, 109]. During the first half of the decade, regulators in Canada, the United States, and some European countries took some case-by-case decisions in this regard [19, 45]. However, their decisions were based on very particular scoping definitions used only in their territory; therefore, these early decisions cannot be readily replicated in third countries having different definitions.

In 2015, Argentina issued a regulation establishing general criteria and a mechanism to define the regulatory status of these products [124, 126]. The Argentine regulation is based on definitions taken from the Cartagena Protocol on Biosafety, which are the more widely recognized. Therefore, the Argentine approach was the first that could be replicated by other countries also abiding by the Cartagena Protocol definitions. To date, another seven countries in Latin America have enacted similar regulatory approaches. At the same time, other countries in Asia and Africa have also endorsed approaches that may be compatible for most products or are officially considering to do so [22, 33, 98, 118, 120].

Finally, in 2018 the European Court of Justice issued a legal interpretation that contradicted earlier regulatory decisions of European Member States as well as official scientific reports [111]. Currently, many governments are still analyzing if (and eventually how) to regulate gene-edited agricultural organisms and the products derived from them [32, 45, 59, 88, 24].

1.3. Scope of this Article

1.3.1. Checking for off-target activity and DNA insertions in SDN products

Argentina has now accumulated significant practice in the regulatory analysis of gene-edited organisms [66, 127], and also has cooperated intensely on this subject with various governments and intergovernmental organizations. From this experience, it has become clear now that two practical issues of high regulatory relevance require harmonized approaches. These are the assessment of, (a) the so-called "off-target" activity, and (b) spurious DNA insertions. Therefore, this article proposes model criteria in this regard. These criteria have not been crafted merely on theoretical considerations; they are a result of several refinement rounds from practical experience on the regulatory assessment of different gene edited products.

To begin, we may consider the simpler scenario where gene editing is applied to generate a small number of point mutations and/or a short *indel* (insertion and/or deletion of a few nucleotides) in a specific locus. In this scenario, changes in the DNA sequence are due only to random spontaneous errors in the DNA repair process following the cut performed by a site-directed nuclease (SDN). Therefore, for this scenario the expected result should be a short stretch of nucleotide base substitutions and/or deletions and/or additions (due to DNA polymerase error), but not in the insertion of any pre-existing DNA fragment. This result/scenario is named "Site-Directed Nuclease, Type 1 (SDN1)" according to the current regulatory jargon [20]. Typically, the intent of such interventions is to knock out an endogenous gene.

1.3.2. Relevance and timeliness

The criteria presented here would be useful under any of the diverging options for regulatory approaches that are being considered worldwide. On one side, they would be applicable in situations where it may be necessary to assess if a product is (or is not) a GMO or the analogous category of regulated products. On the other hand, the criteria would be useful also in situations where it has already been decided that a product will be regulated as GMO (or analogous category). In the latter case, the criteria would be the first step in identifying those genetic alterations that would be further scrutinized by established GMO risk analysis practices.

Moreover, since many cases correspond to agri-food products that are traded internationally, these regulatory criteria, if widely adopted, could help in facilitating an unavoidable crosstalk between similar as well as dissimilar regulatory frameworks across national borders.

The specific issues and criteria discussed in this article are thus quite independent of the regulatory environment where they may be applied. In addition, they are based only on scientific considerations. They are not technically trivial, since they require considering a state of the art that is quite diversified and

continuously evolving. They tackle a regulatory challenge consisting of balancing the case-by-case approach with the principle of avoiding arbitrariness and providing equal treatment to all applicants under laws and regulations.

1.3.3. What this article is not about

This article is not intended to address any of the topics under debate with regard to the governance of genome editing in agriculture [6, 99], nor neglect their importance. These topics include whether the gene-edited products are GMOs or not [110], ethical aspects [21] and socioeconomic considerations [53, 125, 127], consumer acceptance [44] or the applicable risk assessment criteria [19, 29]. Instead, this article is anticipating that whatever the outcome of such debates, the regulatory criteria presented here would be useful for regulators when implementing any of the policy options that are being considered worldwide.

This article suggests how to harmonize the approach for searching every possible change in genetic sequences derived from a gene-editing intervention, considering that it would be a common need of regulators in different countries and for different purposes. However, there will be subsequent regulatory steps after such a search that are beyond the scope of this article and would be different for each case.

After searching, there would be a need for *characterizing* each genetic change. Characterization needs a case-by-case approach. For instance, it would be different if assessing a mutation that knocks out an endogenous gene vs. another that "resurrects" a pseudogene [74]. Also, a specific characterization would be warranted, for instance, if novel polypeptides are possibly expressed after the random insertion of foreign DNA. Finally, characterization is dependent on specific regulatory endpoints, such as analyzing if the resulting organism should be regulated as GMO or not or for the safety assessment of each change.

2. Further Considerations

2.1. What Else Can Happen at or Around a Cleaved Site?

As explained above, the activity of SDNs can lead to point mutations and indels in the cleaved site, generated during the subsequent DNA repair process. In contrast, ordinary SDNs are reported not to cause epigenetic changes [65]. Nevertheless, other kinds of changes in the host genomic sequence caused by SDN cleavage are possible, and they are briefly reviewed next.

2.1.1. Unintended DNA insertion in cleaved sites

Genome editing methods based on genomic DNA cleavage can result in unintended DNA insertion at the cut loci. Such inserted DNA may come from the host cell as well as from foreign sources [2, 46, 67, 37]. In the case of foreign DNA, the possibilities are not limited to the more obvious source of plasmids purposely introduced in the cell for expressing the SDN. The insertion of foreign DNA has been reported even in case of allegedly "DNA-free" techniques based on the intracellular delivery of nuclease proteins. In one case, for instance, the foreign DNA was identified as the *in vitro* transcription template of a Cas nuclease's RNA component; such an outcome could not be avoided even after DNAse I treatment [3]. Besides, more remote sources of minute DNA contamination have also led to foreign DNA insertions in cleaved loci. For instance, *E. coli* DNA from bacteria used to multiply plasmids, or mammalian DNA from fetal serum added to culture media [85]. For this reason, in the current state of the art no SDN technique can be claimed to be absolutely "DNA free". Developers should minimize the presence of foreign DNA and genomic DNA breakage as much as possible.

The latest improvements in sequencing and bioinformatic tools allow for a more exhaustive search of foreign DNA insertions compared with the early days of GMO safety assessment when that search could be based only on Southern Blot techniques. These improved techniques, however, given their level of detail, raise the issue of how to distinguish foreign (artificial) from endogenous (spontaneous) DNA insertions, especially for very short sequences.

2.1.2. Mutations from refilling

It has been reported [96] that the gap-refilling activity of the DNA repair mechanisms acting after SDN cleavage can lead to mutations at a short distance (up to 24 nt up/downstream) of the cleavage site. Such "satellite" mutations can happen even if there is no mutation in the cleavage site itself.

2.1.3. Local chromosomal rearrangements

It has also been reported recently [61] that large deletions and complex rearrangements (translocations, inversions, or large insertions) can result from SDN activity. Such chromosomal rearrangements can be quite distal to the cut site. Therefore, they can be missed by conventional short-range sequencing but they would still be in reach of long-range PCR or long-read sequencing. Such rearrangements also seem to be possible even if the sequence of the cut site is not modified.

2.2. CRISPR-Cas Nucleases as a Case Study

Various site-directed nucleases have been used for gene editing over the past two decades [90]. However, the advent of the CRISPR-Cas technology in 2012 [50, 92] generated an enormous increase in the number of projects and developers of genome-edited organisms [141]. CRISPR-Cas is considered easier to use by many researchers, particularly in regard to the programmability of the target sequence.

Clustered, regularly interspaced, short palindromic repeat (CRISPR) RNA-guided nucleases are highly efficient genome editing tools. These CRISPR-associated (Cas) ribonucleoprotein complexes with endonuclease activity generate a double-strand break in those DNA molecules having sequence complementarity to a specific section of its RNA component (CRISPR RNA "crRNA" or single-guide RNA "sgRNA").

2.2.1. Sequence specificity

The specificity-relevant section of the Cas RNA component consists of a "protospacer adjacent motif" (PAM), followed upstream by a so-called "spacer" sequence. A perfect pairing of the PAM with DNA is required for cleavage, while mismatches may be tolerated in the spacer.

The first nucleotides within the spacer sequence that are located immediately upstream of PAM are called the "seed region"; mismatches in this region greatly impair nuclease activity. Conversely, the middle of the spacer sequence appears to have a higher tolerance for mismatches [31, 91, 50, 52, 71, 82, 115].

Significant knowledge on the molecular basis of mismatch tolerance comes from studies of the ribonucleoprotein/DNA interaction; these studies combine molecular modeling, statistical thermodynamics, and kinetics [25, 48, 52, 129, 135]. The first critical stage of the enzymatic mechanism is PAM site recognition. The second one is the formation of a DNA/RNA bound structure (R-loop), leading to a conformational gating mechanism driven by the 14th-17th nt region of the spacer.

It has been shown that the specificity of Cas nucleases can be modified by protein mutation, leading to alterations in mismatch tolerance and the PAM canonical sequence [8, 12, 57, 58, 106, 140]. Specificity has also been modulated by creating homo- or hetero-dimeric fusion proteins [91, 119, 36].

2.2.2. Variants of Cas proteins

There are many types of Cas nucleases in nature, and just a handful have been employed so far for genome editing. The more widely used and perhaps best characterized Cas nuclease was obtained from *Streptococcus pyogenes* (SpCas9). SpCas9 PAM has the canonical sequence "NGG". The possible length of its spacer sequence ranges from 17 to 24 nt; and its seed region is reported to measure from 8 to 12 nt.

In addition to SpCas9, other natural Cas9 homologs include those obtained from *Neisseria meningitides* (NmCas9), *Treponema denticola* (TdCas9), *Staphylococcus aureus* (SaCas9), and *Streptococcus thermophilus* (StCas9) [23, 54, 112, 128].

Cas12a (a.k.a. Cpf1) is a different group claimed to have higher specificity than Cas9 nucleases [114]. The group includes the Cas12a nuclease of *Francisella novicida* (FnCas12a), *Acidaminococcus sp. BV3L6* (AsCas12a), and *Lachnospiraceae* bacterium (LbCas12a) [56, 116]. Cas12b is a related group, including representatives from *Alicyclobacillus acidophilus* (AaCas12b), *Alicyclobacillus acidoterrestris* (AacCas12b), *Bacillus thermoamylovorans* (BthCas12b), and *Bacillus hisashii* (BhCas12b) [75]. These other Cas nucleases have PAMs that are different from SpCas9 in canonical sequence and length (from 2 to 7 nt). The length of the spacer region and the relative span of the seed subregion within it can also vary.

2.2.3. Other factors affecting specificity

The standard design of Cas9 spacer sequences has a length of 20 nt. Fu et al. [31] showed that using a truncated RNA molecule having a spacer of 17-18 nt can decrease the generation of mutations in loci harboring mismatches. Conversely, spacer sequences with a high GC content (50-70 percent) might favor tolerance towards mismatches [71, 119, 134, 47].

It has been suggested that the delivery of purified Cas ribonucleoprotein complex may result in lesser cleavage of loci having mismatches compared to *in vivo* expression of the protein from a Cas gene delivered to the cell. This difference is explained using kinetic speculations based on the fast degradation rate of the protein [55]. There seems to be enough evidence from several studies that increases in cleavage efficacy (e.g., by using a more processive nuclease, increasing GC content, generating a high concentration or a longer-term presence of the nuclease) likely lead to a trade-off regarding specificity.

Finally, the current state of the art also includes techniques for modulating "tissue specificity". They include the application of anti-CRISPR (Acr) proteins [42] and photoactivatable systems [81]. However, such techniques are not expected to modify "off-target specificity" and would be applied mostly in health-related applications and basic research.

In summary, many studies show that gene editing specificity displays variability from one intervention to another. However, such variability has boundaries that can be assumed for predicting the broadest scenario with regard to loci that could be affected. Moreover, these studies also illustrate that it is possible to find and characterize every significant change generated by a particular gene editing procedure.

3. "Off-Target" Analysis

3.1. From "Off-Target" to "Potentially Affected Loci"

Numerous reports have shown that Cas nucleases are capable of performing their DNA cleavage activity in loci having some mismatches with the spacer sequence [46, 117, 134, 140, 62, 121, 101]. As a result, one of the main concerns related to technologies based in Cas and other SDNs is the possibility of "off-target" cleavage [139].

A simplistic portrayal of the "off-target" issue would be a situation where: (a) the sequence of the locus that is intended to be modified is present only once in the genome and has perfect sequence identity with the spacer sequence (i.e., one "intended" "target" loci with zero mismatches); and concurrently (b) there are other loci with imperfect sequence identity scattered along the rest of the genome that are not intended to be modified (i.e., "unintended" "off-target" loci with mismatches).

However, real-world situations are usually more complex. On one side, the developer's intention could be to modify more than one related sequence, perhaps not having a 100 percent sequence identity among them (e.g., when attempting to modify different alleles or homolog genes all at once). In such cases, the design of the spacer sequence would contemplate mismatch tolerance to reach every locus of interest. On the other hand, in addition to a locus of interest, there could be other loci with 100 percent identity with the PAM + spacer sequences (e.g., genes with repetitive sequences or random coincidences throughout the genome). In such situations, the meaning of terms like intended/unintended and target/off-target can be blurry (and ultimately irrelevant) from a regulator's perspective. Moreover, such "target/off-target" labels usually lead the

developer to study potentially affected loci using different tools and depth, on a subjective basis. Subsequently, this can lead to a situation where the developer/applicant presents the case to regulators in a way that may mislead them towards paying different attention to "loci of interests" compared with "off-target loci". Nevertheless, all potentially affected loci should receive the same treatment in regulatory issues like deciding if a product is a GMO or not, and for the safety assessment of genetic modifications.

Therefore, the term "potentially affected locus/loci" (PAL) is proposed to encircle all possibilities under a homogeneous umbrella of regulatory oversight, leaving aside the developer's subjective point of view. PAL would include every sequence of the host genome that has a significant probability of being altered by the activity of the SDN.

3.2. Ex-ante Assessment of PAL

Whole-genome studies have confirmed that Cas proteins do not affect loci lacking homology with the spacer sequence [117, 27]. Therefore, the identification of PAL should be based exclusively on a bioinformatics search of loci having a certain degree of homology with the PAM and spacer sequences.

With so many different Cas proteins and protocol options available that can affect specificity, it becomes a difficult task to establish standard consensus rules for identifying PAL. Moreover, caution should be taken when generalizing results across different reaction milieus. *In vitro* environments vs. bacterial cells vs. eukaryotic cells can display differences affecting protein function and kinetic and thermodynamic factors such as redox potential, pH, water activity, DNA repair machinery, as well as the concentration of key molecules and their degradation routes.

Developers usually base their searches for PAL on specificity rules taken from a reduced number of publications on the NnCasX protein that they are using, or a closely related one. Furthermore, in some cases developers just rely on one of the online tools available for this purpose [4, 17, 38, 39, 43]. However, these online tools do not always agree on their results, given that their algorithms were devised separately by different experts, based on different sets of publications.

This situation is particularly problematic for regulators, who have the three following options:

(i) Rely blindly on the applicant's (or the server's) criteria. This option would not be a preferred choice, since an important part of the regulator's oversight duties would be discharged to the interested party (or even worse, a third party that takes no legal responsibility on the reliability of the prediction, in the case of online tools). The applicant may select rules with bias or incomplete knowledge on the state of the art, thus missing and underreporting PAL. Besides, this would lead to different applicants arbitrarily receiving tailored regulatory stringency. In summary, there would be unacceptable regulatory slips regarding both safety assessment and fairness.

(ii) Wait until the state-of-the-art leads to repeated validation of distinctive specificity rules for each protocol (i.e., each particular combination of a NnCasX protein, host organism, and delivery method). This option could appear as the more appropriate one in theory. However, in practical terms, it would lead to a never-ending moratorium on the use of the technology, or at best to the forced use of just a couple of older and best characterized Cas proteins, thus halting innovation.

(iii) Apply a set of canonical rules based on the more characterized Cas proteins, including how to adapt these rules cautiously in case of novel Cas versions and protocols. This option seems to be the more reasonable one, after considering that there is already a significant amount of knowledge about these proteins; thus, it is further developed below.

3.2.1. General considerations

The rules for predicting PAL should be used by default in most cases, for the sake of fair treatment to all applicants and to seek harmonization between regulatory offices. They should incorporate evidence from a broad base of literature sources. The rules should be clearly made available to the public in advance, so that any potential developer can incorporate (and consult with regulators about) them during the design stage of a gene-editing procedure.

Regulators should exert a conservative criterion regarding how to identify all the genome locations potentially affected by the gene-editing procedure. Conservative in this context means that less available information on specificity may lead to more stringent criteria. Such criteria would be prone to identify more "false" PAL as a trade-off in avoiding to miss any "real" PAL (i.e., type II error over type I error) [11].

Being conservative, the rules should initially treat the latest innovations (for instance, using mutant Cas9 proteins) with the same or increased stringency. Lowering stringency shall be contemplated as an infrequent case-by-case possibility, only after strong evidence of increased selectivity is gathered and where the evidence also allows extracting clear alternative rules. "Strong evidence" in this context means a significant number of studies, from different authors, under similar conditions (NnCasX protein, host organism, delivery method) as the case presented by the applicant.

Comprehensive *in vivo* and *in vitro* empirical studies searching for effectively affected (mutated) sites have been performed in diverse species [27, 28, 30, 91, 64, 65, 75, 80, 86, 89, 95, 117, 130, 142]. These studies on specificity usually report the number of loci that have been affected (mutated) despite having mismatches against the sequence of the spacer region. Studies are diverse, but most of them report a tolerance of typically one (occasionally two) mismatch in the seed region and up to three (occasionally up to five) mismatches in the whole spacer region. In contrast, some studies report no other affected loci than those with perfect matches.

There could be many explanations for such varied results, ranging from differences in the Cas protein and the protocol used, up to the strategy used to identify affected loci. Another relevant factor could be the trueness of the genomic sequences

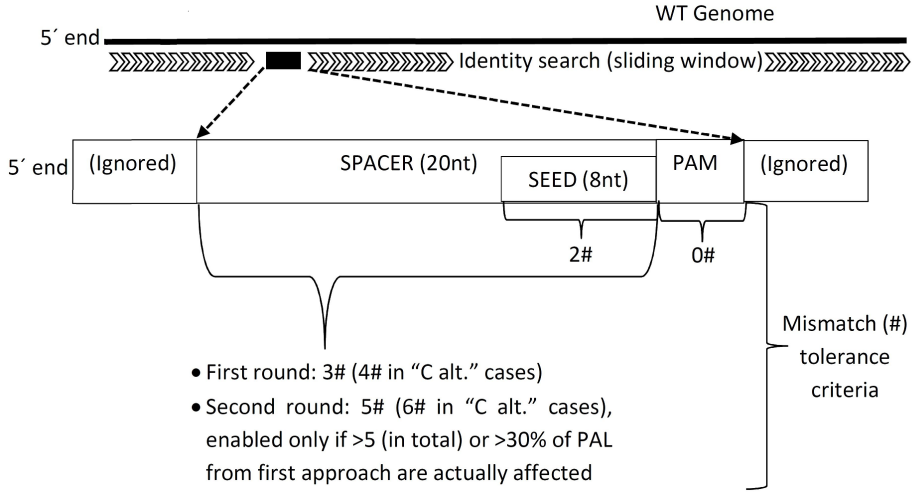

Figure 1: Proposed rules for identifying PAL

used for the *ex-ante* bioinformatics analysis compared to the actual genome of the specific strain/breed/variety to be gene-edited; for instance, when only a generic reference sequence for the species is available.

Off-target mutations appear to be reported more frequently in mammals than in plants. However, it is not clear if such difference is factual or just an artifact from more intense search in mammalian studies (given serious inferences for human health applications), vs. lesser studies in plants (where removal of mutations is feasible through backcrossing) [38].

3.2.2. Proposed rules

Considering the current state of the art, for regulatory purposes in agricultural applications, a locus should be assumed to be a PAL when **all the following** apply (see also Figure 1):

(i) It has perfect complementarity to the PAM. This assumes that the regulator acknowledges the PAM region to be sufficiently characterized for the nuclease used (such as the "NGG" motif in SpCas9). If this is not the case, only the first two nucleotides of the alleged PAM region would be assumed to be acting as PAM (i.e., not allowing mismatches), and the remainder would be considered to be part of the seed region. This approach is conservative because, in any case, it would increase the number of PAL and thus would lean towards type II errors while avoiding type I errors.

(ii) It has up to two mismatches in the seed region, assuming an 8 nt seed region for any NnCasX protein. Although the seed region has been postulated to be longer in some cases, this shorter assumption would, in any case, increase the number of PAL.

(iii) It has up to three mismatches in the whole spacer region, including any mismatch in the seed region.

Most designs use a 20 nt spacer region. However, the same limit of three mismatches should be used for other lengths, as follows:

(a) If a truncated RNA molecule is used to shorten the spacer (17-18 nt), the same limit of three mismatches would, in any case, increase the number of PAL.

(b) If the applicant claims to be using a "shorter" spacer but without truncating the RNA molecule (i.e., there is still a short stretch of 2-3 contiguous nucleotides at the tail of the molecule that could be considered part of a 20 nt spacer).

(c) If the use of a spacer *longer* than 20 nt is claimed. For regulatory analysis, the extra nucleotides should be ignored and still apply the same limit of three mismatches over a spacer region of a 20 nt. To accommodate different types of uncertainties, the limit of three mismatches in the whole spacer should be increased to four in any of the following cases:

- The sequences used for the bioinformatics analysis are not from the very same strain/breed/variety employed later in the gene editing procedure.
- Usage of a Cas nuclease that has very few studies on their specificity (regarding homologs from other microorganisms and mutant versions as different proteins). This criterion would apply especially in cases where the nuclease or the protocol is postulated to have "increased" cleavage activity (which has been shown to antagonize specificity).
- The absence of any practice that may reduce the number of loci effectively affected. This includes, where possible, designing a spacer with a low GC content (<45 percent), performing ribonucleoprotein delivery, and doing repeated backcrossing (>3 times) to bred-out undetected affected sites.

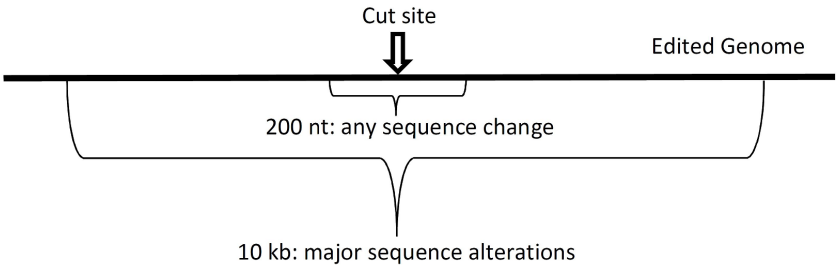

Figure 2: Bioinformatic analysis of PAL after application of the gene editing technique

The 3-4 mismatch limit is based on the assumption that very few PAL are actually found to be changed later. This assumption allows for avoiding an initial characterization of unlikely affected sites having a higher number of mismatches. Nevertheless, affected loci with up to six mismatches have been reported; this seems to be the limit imposed by the ribonucleoprotein/DNA binding process thermodynamics. Therefore, in cases where a significant number of PAL end up being affected (>5 in total or >30 percent of all PAL) using a 3-4 mismatch limit, and if the applicant is still interested in continuing, the analysis shall be repeated and expanded to loci with 5-6 mismatches.

These rules are devised on the assumption that there is a complete set of reference genome sequences available for the organism before the technique is applied, thus allowing the prior bioinformatics analysis. Nevertheless, this does not mean that deregulation shall be impossible in cases where such a reference genome is not available. A speculative possibility in this regard is the use of unassembled whole-genome sequences, which would be cheaper (S. Feingold, personal communication, November 28, 2019). Another possibility consists in the use of physicochemical methods for isolating genomic DNA fragments containing PALs; for instance, affinity chromatography using modified SDNs [131]. Finally, an alternative solution could be the "inverse approach", based on identifying affected loci after the gene editing procedure has been performed (as described later).

3.3. Ex-post Analysis of PAL

After executing a gene editing technique, in order to find which PAL have actually been affected by the SDN, developers may resort to preliminary analysis such as PCR/restriction enzyme assay [26, 48, 130, 68, 102] or T7E1 assay [134, 100]. However, for regulatory purposes these tests provide insufficient information and could miss some kinds of genomic sequence alterations. Therefore, the following should be required for regulatory purposes:

3.3.1. Whole-genome sequencing

Whole-genome sequencing (WGS), with at least 20 passes is the most straightforward way to assess if and how a PAL has been affected by the activity of the SDN. It allows exploring any PAL with any number of mismatches. It also allows any DNA insertion or satellite mutations or chromosomal rearrangements to be discarded. Moreover, this option is less prone to raise debates between applicants and regulators regarding experimental design and interpretation of results. Besides, in case of controversy regarding the parameters of the sequence analysis, it is relatively simple to perform them again.

The use of WGS does not mean that a whole-genome comparison with the reference sequence should be made (except when the "inverse approach" is used, see below). Such a comparison will always report differences in other locations not related to the PAL, which are expected due to spontaneous mutation, somaclonal variation, or pre-existing differences between the genome of the strain/variety/breed used and the reference genome [27, 117]. These kinds of changes are not a consequence of SDN activity and are not under the regulatory frameworks usually applied to agricultural biotechnology products. However, a few countries have a very particular criterion in this regard, including France (according to the latest court interpretations, see [63]), or perhaps Canada, if one of those other spurious mutations generates a novel trait [103, 93]. Therefore, only the WGS data corresponding to the PAL identified earlier should be used. The suggested approach for the bioinformatics analysis applied to every PAL should involve (see also Figure 2):

(i) Sequence alignment of a region spanning 200 nucleotides centered in the predicted cut site (100 upstream to 100 downstream). This would enable analyzing for (a) indels or change of bases at the cut site, (b) DNA insertion in the PAL, and (c) adjacent mutations from repair/refilling.

(ii) Sequence alignment of a region spanning 10 kb centered in the predicted cut site (5000 nt upstream to 5000 nt downstream). Parameters of the comparison should be adjusted for finding and reporting large deletions and complex rearrangements.

3.3.2. Alternatives to whole-genome sequencing

As mentioned earlier, WGS would be the preferred option for obtaining regulatory data on genetic changes derived from a gene editing procedure. However, this option may have a prohibitive cost in some cases. Therefore, alternative methods based on other techniques but still generating information of similar regulatory value may be used in some cases. The cost and hustle of these alternative approaches increase with

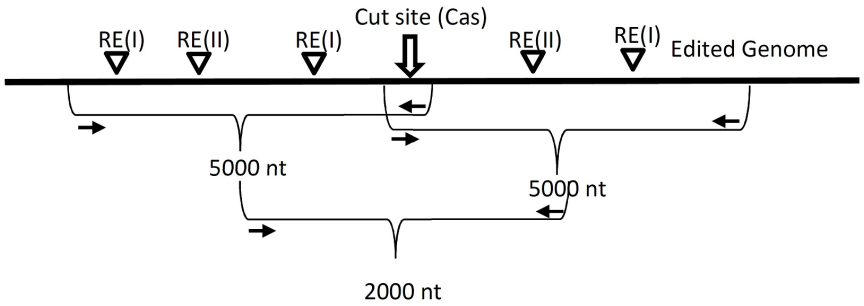

Figure 3: Alternative PCR/restriction enzyme analysis

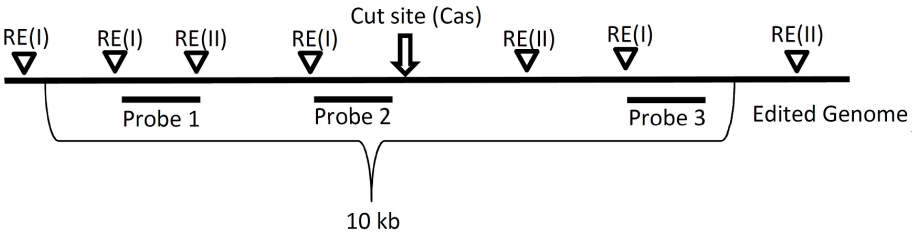

Figure 4: Alternative Southern blot analysis

the number of loci to be studied. Therefore, its relative convenience relies on the developers' capacity to find spacer sequences that minimize the number of PAL. Besides, their relative convenience will decrease over time as WGS cost gets cheaper. The suggested approach, regarding every PAL, would involve:

(i) Standard PCR, followed by sequencing and alignment, of a region spanning at least 200 nucleotides centered in the predicted cut site (100 upstream to 100 downstream). This approach would enable analyzing for (a) indels or change of bases at the cut site, (b) DNA insertion in the PAL, and (c) adjacent mutations from repair/refilling, and

(ii) a single long-range PCR spanning a 10 kb region centered in the predicted cut site (5000 bp upstream to 5000 bp downstream), followed by sequencing of the PCR fragment and alignment, where parameters of the bioinformatics comparison should be adjusted for finding and reporting large deletions and complex rearrangements, or

(iii) alternative A: Three separate long-range PCRs covering a region spanning approximately, (a) 5000 bp upstream to 50 bp upstream of the predicted cut site; (b) 2000 bp upstream to 2000 bp downstream of the predicted cut site; (c) 50 bp downstream to 5000 bp downstream of the predicted cut site. The exact location of the primers and PCR parameters should be adjusted (i.e., appropriate GC content, avoiding repetitive sequences, etc.) to avoid artifacts or misleading results. In addition, an RFLP analysis with two different restriction enzymes should be applied to assess the presence/absence and size of the amplicons. The restriction enzymes used should cut at least once on each amplicon according to the reference sequence and be used in combination (see also Figure 3), or

(iv) alternative B: Southern blot analysis of a DNA fragment spanning 10 kb centered in the predicted cut site, by using at least three probes spanning that fragment in combination with separate treatments using a different restriction enzyme each, chosen to cut in different sites spanning the fragment (see also Figure 4).

With regard to (ii), (iii Alternative A), or (iv Alternative B) above, it is relevant to note that all of them would add a significant regulatory cost and are based on currently scarce evidence (only [61]) concerning large deletions and complex rearrangements. Therefore, as more scientific evidence accumulates, this requirement might be left out in cases where the regulator has confidence that local chromosomal rearrangements are not likely to have occurred.

3.4. "Inverse Approach": Ex-post Detection of Affected Loci

An alternative approach that does not require presenting an *ex-ante* assessment of PAL begins with a comprehensive whole-genome comparison between the gene-edited organism and a closely related, wild type reference genomic sequence. As mentioned, such whole-genome comparisons will report numerous differences, most of which are not likely to be caused by the action of the SDN. Therefore, after all differences are found, the corresponding local sequences in the wild type genome should be tested with the PAL rules described previously.

This alternative approach has pros and cons. The reference (wild type) genomic sequence should be exactly the same strain/breed/variety used; if not, the parameters of the bioinformatics comparison may become quite subjective and therefore

debatable. The quantity of PAL is not examined *ex-ante*; therefore it might make more sense for techniques where the ability to redesign the tool to reduce the amount of PAL is more limited (e.g., TALEN vs. Cas). It may lead to a higher cost (two WGS determinations, on the wild type and edited organisms); however, that cost is put off until a gene-edited organism of a satisfactory phenotype is obtained.

4. Spurious DNA Insertions (Outside of PAL)

It has been shown that spurious insertions of foreign DNA in random locations (different from PAL) can occur under certain circumstances. This outcome seems to have a significant probability for techniques that introduce high quantities of DNA in the cell (e.g., those expressing the Cas and its RNA component *in vivo*). The latter is especially valid when combined with biolistic or other physical introduction methods that may increase the rate of random nuclear DNA breakage. In contrast, no random DNA insertion has been reported in protocols that use modified *Agrobacterium*-mediated methods to deliver the nuclease gene [5]. A recent article by FDA regulators [83, 137, 10] is enlightening about the possibility of spurious foreign DNA insertions, as well as insertions in PAL, to stay unnoticed by the developers because of assumptions about the "cleanliness" of an SDN technique.

Conversely, the probability of having spurious foreign DNA insertions can be dismissed as highly unlikely in cases where the presence of DNA in the milieu introduced in the cell has been greatly reduced. This is applicable for techniques where the nuclease protein is delivered into the cell instead of its gene, the Cas RNA component has been synthesized chemically, and the milieu to be inserted in the cell has been treated extensively with DNAse I, fulfilling the three conditions altogether.

In any case, further breeding, including backcrossing, as well as phenotype studies showing no changes in polygenic traits can increase regulatory confidence in many aspects, including the absence of DNA insertions in the final product that could have safety relevance.

4.1. Whole-Genome Sequencing

A bioinformatics search for foreign DNA insertion should be required in every case where the developer has performed a WGS. It should be based on the alignment of sliding windows of 100 nt having 100 percent identity. This comparison, when applicable, should be made against the following:

(a) every DNA known to have been purposely introduced in the cell; in the case of restriction fragments or PCR amplicons from a larger DNA molecule, the study should be made against the whole larger molecule;

(b) every DNA molecule used as a template in the process of producing the ribonucleoprotein complex in separate expression systems, including both the templates for the Cas protein and the RNA component;

(c) an appropriate whole reference genome of the system used for producing plasmids and/or as a separate expression system for Cas protein, typically *E. coli*; and

(d) an appropriate whole-genome reference sequence of organisms contributing with components of culture media that are presumed to contain considerable amounts of DNA (typically *Bos Taurus* in an animal cell culture medium, because of fetal calf serum).

Perhaps (c) and (d) would be an over-exaggeration for those cases described previously as having minimal possibilities of spurious DNA insertions. However, if the applicant has already invested in a WGS determination and the associated capabilities of bioinformatics analysis for assessing PAL, this additional request would not be significantly costly. Besides, the empirical experience gathered this way will enhance the regulators' ability to address different cases where other applicants base their analysis in one of the approaches described next.

4.2. Alternatives to Whole-Genome Sequencing

When the applicant is not able to provide WGS data, regulators should decide if there is a need for demanding proof of absence regarding spurious insertions of foreign DNA. This decision can be taken based on counting "flags", corresponding to those factors mentioned before that affect the possibility of having such insertions.

"Red flags" would be raised for protocols that are prone to generate random breaks in the genome and introduce significant amounts of DNA in the cell. Conversely, "green flags" can be acknowledged when repeated backcrossing has been performed and extensive phenotype information about polygenic traits is available.

If, after counting flags, such proof is required, in the absence of WGS it could be based on PCR and Southern blot analysis. For this purpose, it can be acknowledged that biotech regulators have plenty of experience with the analysis of transgenic organisms having transformation events. In that analogous situation, the applicant has to prove how many independent foreign DNA insertions from the construct and its molecular vector may be present. For years, this has been done with a combination of PCR and Southern blot analysis. Adhering to a strict scientific base as well as the principle of fairness towards all applicants, the experimental design required for the search of random DNA insertions in gene-edited and transgenic organisms should be the same. As further guidance, a suggestion of a strict design that could be equally applied to both situations is:

(i) Southern blot analysis using probes of 700 nt corresponding to at least five different, roughly equally spaced sections across the whole of each foreign DNA molecule, preferably on coding regions, or

(ii) PCR using primers targeting gross sources of foreign DNA (e.g., DNA purposely introduced in the cell, plasmids from *in vitro* transcription if the milieu was not treated with DNAse I). They should be selected to produce fragments not longer than 700 kb nor shorter than

200 kb, taking into account both the efficiency of ordinary PCR detection methods and seeking short insertions. Amplicons should span at least five different, roughly equally spaced sections across the whole target molecule, preferably on coding regions because of their higher importance on safety considerations.

Once again, regulators should make sure that whatever their criteria may be, either they are equivalent to their previously established criteria for finding the number of inserts in a transgenic organism, or they should update those earlier requirements to match the ones used for this purpose.

5. Other Gene-Editing Techniques and Site-Directed Effectors

5.1. Other Nucleases

Nowadays, dossiers reaching the regulator's desks include similar proportions of products obtained using Cas nucleases vs. products obtained using other genome editing techniques. These other techniques resort to the use of Transcription Activator-like Effector Nucleases (TALEN) [9, 78, 133, 138], zinc-finger nucleases (ZFNs) [122], and meganucleases [104, 108]. In addition, some products are obtained with oligonucleotide-directed mutagenesis (ODM) [97], a set of gene-editing techniques that do not use SDNs.

In most cases, these other products originated in projects that started before CRISPR-Cas tools were widely available, and they will likely become a minority in the near future. However, they may not disappear entirely because of intellectual property aspects that could balance their alleged technical disadvantages, and some companies that have been investing in developing proprietary technology may stick to them. In the long term, older technologies becoming off-patent sooner may be preferred by low-budget developers such as the public research sector and SMEs because they are free after entering the public domain. In principle, the overall approach presented here can be adapted to these other techniques, provided that enough information on specificity is available. For instance, it has been shown that TALEN pairs can tolerate up to 3-4 mismatches in their recognition sites [13]. However, studies characterizing the possibility of unintended genetic changes in these other techniques [15, 35, 18, 40, 73, 77, 87, 94, 107, 136] are far less in number compared with the literature on Cas nucleases. This asymmetry represents a challenge for regulators, who should adopt the same level of stringency toward different applicants, regardless of the technique used.

5.2. SDN2 and SDN3

The criteria proposed here were explained based on the simpler SDN1 scenario. However, as mentioned, there are other classes of gene editing techniques [20, 126]. One class employs an additional short DNA molecule as a template for repairing the cleaved locus (SDN2, a.k.a. "allelic repair" in some cases). Another class uses a bigger DNA molecule for its intended insertion in or replacement of the target locus (SDN3, a.k.a. "allelic replacement" in some cases).

In principle, the criteria presented here are also applicable to SDN2 and SDN3, with due consideration to the fact that these techniques always involve the deliberate introduction of specific foreign DNA in the cell in significant quantities. Hence, the additional DNA molecule has a high probability of ending up inserted, perhaps in tandem insertions, inversions, etc., in any of the PAL or elsewhere in random locations of the genome [96]. Therefore, this molecule's sequence should be taken into account in the search for spurious foreign DNA insertions, as described.

In addition, for SDN3, the short-range sequence identity analysis of those PALs where there has been an insertion should be extended. Sequence information should reach at least from 200 nt downstream of the first junction, between the inserted DNA molecule and the host genome, until 200 nt upstream of the second junction. This approach would be effective in finding unintended effects such as multiple head-to-tail insertions that can be misidentified as single insertions when only conventional PCR analysis is performed [105].

5.3. Other Site-Directed Activity

The criteria presented here perhaps could also be used as inspiration for developing regulatory criteria for SDNs-related proteins that have a different function. For instance, "prime editors" that edit single bases without double-strand breaks [34, 69], epigenome editing [41, 113, 123], transcriptional activation [60, 72], or the use of Cas13a (a.k.a. C2c2) for knocking down endogenous mRNA [1]. It has been shown that it is possible to have off-target activity in some of these techniques [49, 143].

6. Discussion and Conclusions

This article proposes a pragmatic and proportionate approach for addressing the possible existence of off-target editions and spurious DNA insertions caused by gene-editing techniques in agricultural applications. If widely adopted, there would be a harmonized approach for this important regulatory issue. Moreover, its explicit availability would help developers improve a safety aspect of their experimental design and protocols *ab initio*, lowering costs and complications during the subsequent regulatory assessment.

The value of counting with a standardized regulatory approach of the kind presented here can be compared with the allergenicity analysis of novel proteins expressed in GMOs. In 2003, the Codex Alimentarius generated guidance in this regard based on a simple bioinformatic analysis [132]. At that time, different scientific opinions were proposing alternate rules, which were similar albeit not entirely coincident. The Codex guidance was ultimately a compromise solution, agreed on the understanding that it may be a little bit prone to type II errors, but not as much as to hinder the use of GMOs for food.

The Codex guidance on allergenicity assessment was widely adopted and applied during the past two decades, and such a harmonized approach greatly facilitated that developers receive a fair and similar regulatory response in different countries. This guidance has never failed nor changed, despite updating proposals inspired by theoretical advances in the field,

since the robustness and efficacy of its original criteria are remarkable.

Similarly, the current body of knowledge on off-target effects and DNA insertions from SDN is enough for adopting criteria like the ones proposed here. Nevertheless, as more information continues to be made available, such criteria can be updated and extended to other techniques. Ongoing work on systematic reviews and information maps (such as [76]) may be very useful in this regard.

It could be argued that establishing a set of standard rules based on the specificity displayed by SDNs used nowadays would discourage the development of enhanced alternatives. Regulators may not easily relax the rules, unless an increased selectivity is considered to be fully proven, case by case; therefore, the quantity of PAL to be verified might not be affected much by innovation. However, the development of more specific SDNs or techniques will remain appealing, since they would reduce the number of loci that are actually affected. Therefore, such innovations may still reduce the overall regulatory burden.

Finally, it is important to highlight the relevance of examining phenotype to reinforce the molecular genetics approach presented here. In particular, the absence of unexpected changes in polygenic traits would act as an additional reassurance regarding the lack of genetic changes that could have gone undetected. For this purpose, polygenic traits could include agrophenotypic characterization in the case of crops, overall health indicators in the case of animals, and compositional analysis of derived foodstuff in both cases. In addition, assessing the predicted phenotypic change caused by the gene editing intervention is highly recommended. This assessment would serve as a reassurance that the biological function(s) of the target gene and the possible consequences of the intervention are understood well enough to assess their safety.

7. Declaration of Conflicting Interest

The author declares no conflicts of interest.

8. Disclaimer

The information and views are those of the author, and do not necessarily represent those of the organizations where he works.

9. Article Information

This article was received May 22, 2020, in revised form August 19, 2020, and made available online January 5, 2021.

10. References

[1] Abudayyeh, O. O., Gootenberg, J. S., Essletzbichler, P., Han, S., Joung, J., Belanto, J. J., Verdine, V., Cox, D. B. T., Kellner, M. J., Regev, A., Lander, E. S., Voytas, D. F., Ying, A. Y., & Zhang, F. (2017). RNA targeting with CRISPR–Cas13. *Nature, 550*(7675), 280-284. doi: 10.1038/nature24049

[2] Andersson, M., Turesson, H., Nicolia, A., Fält, A-S., Samuelsson, M., & Hofvander, P. (2017). Efficient targeted multiallelic mutagenesis in tetraploid potato (*Solanum tuberosum*) by transient CRISPR-Cas9 expression in protoplasts. *Plant Cell Reports, 36*(1), 117-128. doi: 10.1007/s00299-016-2062-3

[3] Andersson, M., Turesson, H., Olsson, N., Fält, A-S., Ohlsson, P., Gonzalez, M. N., Samuelsson, M., & Hofvander, P. (2018). Genome editing in potato via CRISPR-Cas9 ribonucleoprotein delivery. *Physiologia Plantarum, 164*(4), 378-384. doi: 10.1111/ppl.12731

[4] Bae, S., Park, J., & Kim, J-S. (2014). Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics, 30*(10), 1473-1475. doi: 10.1093/bioinformatics/btu048

[5] Banakar, R., Eggenberger, A. L., Lee, K., Wright, D. A., Murugan, K., Zarecor, S., Lawrence-Dill, C. J., Sashital, D. G., & Wang, K. (2019). High-frequency random DNA insertions upon co-delivery of CRISPR-Cas9 ribonucleoprotein and selectable marker plasmid in rice. *Scientific Reports, 9*, 1-13. doi: 10.1038/s41598-019-55681-y

[6] Bartkowski, B., Theesfeld, I., Pirscher, F., & Timaeus, J. (2018). Snipping around for food: Economic, ethical and policy implications of CRISPR/Cas genome editing. *Geoforum, 96*, 172-180. doi: 10.1016/j.geoforum.2018.07.017

[7] Bodiguel, L., & Cardwell, M. (2010). *The Regulation of Genetically Modified Organisms: Comparative Approaches*. Oxford University Press. doi: 10.1093/acprof:oso/9780199542482.001.0001

[8] Bratovič, M., Fonfara, I., Chylinski, K., Gálvez, E. J. C., Sullivan, T. J., Boerno, S., Timmermann, B., Boettcher, M., & Charpentier, E. (2020). Bridge helix arginines play a critical role in Cas9 sensitivity to mismatches. *Nature Chemical Biology, 16*, 587-595. doi: 10.1038/s41589-020-0490-4

[9] Carlson, D. F., Tan, W., Lillico, S. G., Stverakova, D., Proudfoot, C., Christian, M., Voytas, D. F., Long, C. R., Whitelaw, B. A., & Fahrenkrug, S. C. (2012). Efficient TALEN-mediated gene knockout in livestock. *Proceedings of the National Academy of Sciences, 109*(43), 17382-17387. doi: 10.1073/pnas.1211446109

[10] Carlson, D. F., Lancto, C. A., Zang, B., Kim, E-S., Walton, M., Oldeschulte, D., Seabury, C., Sonstegard, T. S., & Fahrenkrug, S. C. (2016). Production of hornless dairy cattle from genome-edited cell lines. *Nature Biotechnology, 34*(5), 479-481. doi: 10.1038/nbt.3560

[11] Carpenter, D., & Ting, M. M. (2005). The political logic of regulatory error. *Nature Reviews Drug Discovery, 4*(10), 819-823. doi: 10.1038/nrd1850

[12] Chen, J. S., Dagdas, Y. S., Kleinstiver, B. P., Welch, M. M., Sousa, A. A., Harrington, L. B., Sternberg, S. H., Joung, J. K., Yildiz, A., & Doudna, J. A. (2017). Enhanced proofreading governs CRISPR–Cas9 targeting accuracy. *Nature, 550*(7676), 407-410. doi: 10.1038/nature24268

[13] Clasen, B. M., Stoddard, T. J., Luo, S., Demorest, Z. L., Li, J., Cedrone, F., Tibebu, R., Davison, S., Ray, E. E., Daulhac, A., Coffman, A., Yabandith, A., Retterath, A., Haun, W., Baltes, N. J., Mathis, L., Voytas, D. F., & Zhang, F. (2016). Improving cold storage and processing traits in potato through targeted gene knockout. *Plant Biotechnology Journal, 14*(1), 169-176. doi: 10.1111/pbi.12370

[14] Convention on Biological Diversity. (2020). *Text of the Cartagena Protocol on Biosafety*. Retrieved from https://bch.cbd.int/protocol/text/

[15] Cradick, T. J., Ambrosini, G., Iseli, C., Bucher, P., & McCaffrey, A. P. (2011). ZFN-site searches genomes for zinc finger nuclease target sites and off-target sites. *BMC Bioinformatics, 12*, 152. doi: 10.1186/1471-2105-12-152

[16] Craig W., Vanga S. R., & Medaglia, J. C. (2013). Commercialisation of GM Crops: Comparison of Regulatory Frameworks. In P. Christou, R. Savin, B. A. Costa-Pierce, I. Misztal, C. B. A. Whitelaw (Eds), *Sustainable Food Production*. Springer. doi: 10.1007/978-1-4614-5797-8_837

[17] Cui, Y., Xu, J., Cheng, M., Liao, X., & Peng, S. (2018). Review of CRISPR/Cas9 sgRNA design tools. *Interdisciplinary Sciences, Computational Life Sciences, 10*(2), 455-465. doi: 10.1007/s12539-018-0298-z

[18] Du, H., Zeng, X., Zhao, M., Cui, X., Wang, Q., Yang, H., Cheng, H., & Yu, D. (2016). Efficient targeted mutagenesis in soybean by TALENs and CRISPR/Cas9. *Journal of Biotechnology, 217*, 90-97. doi: 10.1016/j.jbiotec.2015.11.005

[19] Duensing, N., Sprink, T., Parrott, W. A., Fedorova, M., Lema, M. A., Wolt, J. D., & Bartsch, D. (2018). Novel features and considerations for

ERA and regulation of crops produced by genome editing. *Frontiers in Bioengineering and Biotechnology, 6*, 79. doi: 10.3389/fbioe.2018.00079

[20] EFSA Panel on Genetically Modified Organisms. (2012). Scientific opinion addressing the safety assessment of plants developed using Zinc Finger Nuclease 3 and other Site-Directed Nucleases with similar function. *EFSA Journal, 10*(10), 2943. doi: 10.2903/j.efsa.2012.2943

[21] Eriksson, S., Jonas, E., Rydhmer, L., & Röcklinsberg, H. (2018). Invited review: Breeding and ethical perspectives on genetically modified and genome edited cattle. *Journal of Dairy Science, 101*(1), 1-17. doi: 10.3168/jds.2017-12962

[22] Eriksson, D., Kershen, D., Nepomuceno, A., Pogson, B. J., Prieto, H., Purnhagen, K., Smyth, S., Wesseler, J., & Whelan, A. (2019). A comparison of the EU regulatory approach to directed mutagenesis with that of other jurisdictions, consequences for international trade and potential steps forward. *New Phytologist, 222*(4), 1673-1684. doi: 10.1111/nph.15627

[23] Esvelt, K. M., Mali, P., Braff, J. L., Moosburner, M., Yaung, S. J., & Church, G. M. (2013). Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nature Methods, 10*(11), 1116-1121. doi: 10.1038/nmeth.2681

[24] Exec. Order No. 13874, 84 Fed. Reg. 112 (June 11, 2019).

[25] Farasat, I., & Salis, H. M. (2016). A biophysical model of CRISPR/Cas9 activity for rational design of genome editing and gene regulation. *PLoS Computational Biology, 12*(1), e1004724. doi: 10.1371/journal.pcbi.1004724

[26] Feng, Z., Zhang, B., Ding, W., Liu, X., Yang, D-L., Wei, P., Cao, F., Zhu, S., Zhang, F., Mao, Y., & Zhu, J. K. (2013). Efficient genome editing in plants using a CRISPR/Cas system. *Cell Research, 23*(10), 1229-1232. doi: 10.1038/cr.2013.114

[27] Feng, Z., Mao, Y., Xu, N., Zhang, B., Wei, P., Yang, D-L., Whang, Z., Zhang, Z., Zheng, R., Yang, L., Zeng, L., Liu, X., & Zhu, J-K. (2014). Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in *Arabidopsis*. *Proceedings of the National Academy of Sciences, 111*(12), 4632-4637. doi: 10.1073/pnas.1400822111

[28] Feng, C., Su, H., Bai, H., Wang, R., Liu, Y., Guo, X., Liu, C., Zhang, J., Yuan, J., Birchler, J. A., & Han, F. (2018). High-efficiency genome editing using a dmc1 promoter-controlled CRISPR/Cas9 system in maize. *Plant Biotechnology Journal, 16*(11), 1848-1857. doi: 10.1111/pbi.12920

[29] Friedrichs, S., Takasu, Y., Kearns, P., Dagallier, B., Oshima, R., Schofield, J., & Moreddu, C. (2019). Meeting report of the OECD conference on "Genome Editing: Applications in Agriculture - Implications for Health, Environment and Regulation". *Transgenic Research, 28*(3-4), 419. doi: 10.1007/s11248-019-00154-1

[30] Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K., & Sander, J. D. (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature Biotechnology, 31*, 822-826. doi: 10.1038/nbt.2623

[31] Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M., & Joung, J. K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nature Biotechnology, 32*, 279-284. doi: 10.1038/nbt.2808

[32] Gao, W., Xu, W-T., Huang, K-L., Guo, M-Z., & Luo, Y. B. (2018). Risk analysis for genome editing-derived food safety in China. *Food Control, 84*, 128-137. doi: 10.1016/J.FOODCONT.2017.07.032

[33] Gatica-Arias, A. (2020). The regulatory current status of plant breeding technologies in some Latin American and the Caribbean countries. *Plant Cell, Tissue and Organ Culture (PCTOC), 141*, 229-242. doi: 10.1007/s11240-020-01799-1

[34] Gaudelli, N. M., Komor, A. C., Rees, H. A., Packer, M. S., Badran, A. H., Bryson, D. I., & Liu, D. R. (2017). Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature, 551*(7681), 464-471. doi: 10.1038/nature24644

[35] Grau, J., Boch, J., & Posch, S. (2013). TALENoffer: genome-wide TALEN off-target prediction. *Bioinformatics, 29*(22), 2931-2932. doi: 10.1093/bioinformatics/btt501

[36] Guilinger, J. P., Thompson, D. B., & Liu, D. R. (2014). Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nature Biotechnology, 32*(6), 577-582. doi: 10.1038/nbt.2909

[37] Gutierrez-Triana, J. A., Tavhelidse, T., Thumberger, T., Thomas, I., Wittbrodt, B., Kellner, T., Anlas, K., Tsingos, E., & Wittbrodt, J. (2018). Efficient single-copy HDR by 5' modified long dsDNA donors. *eLife, 7*, e39468. doi: 10.7554/eLife.39468

[38] Hahn, F., & Nekrasov, V. (2019). CRISPR/Cas precision: do we need to worry about off-targeting in plants? *Plant Cell Reports, 38*(4), 437-441. doi: 10.1007/s00299-018-2355-9

[39] Hanna, R. E., & Doench, J. G. (2020). Design and analysis of CRISPR–Cas experiments. *Nature Biotechnology, 38*(7), 813-823. doi: 10.1038/s41587-020-0490-7

[40] He, Z., Proudfoot, C., Whitelaw, C. B. A., & Lillico, S. G. (2016). Comparison of CRISPR/Cas9 and TALENs on editing an integrated EGFP gene in the genome of HEK293FT cells. *SpringerPlus, 5*, 814. doi: 10.1186/s40064-016-2536-3

[41] Hilton, I. B., D'Ippolito, A. M., Vockley, C. M., Thakore, P. I., Crawford, G. E., Reddy, T. E., & Gersbach, C. A. (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nature Biotechnology, 33*(5), 510-517. doi: 10.1038/nbt.3199

[42] Hoffmann, M. D., Aschenbrenner, S., Grosse, S., Rapti, K., Domenger, C., Fakhiri, J., Mastel, M., Börner, K., Eils, R., Grimm, D., & Niopek, D. (2019). Cell-specific CRISPR–Cas9 activation by microRNA-dependent expression of anti-CRISPR proteins. *Nucleic Acids Research, 47*(13), e75. doi: 10.1093/nar/gkz271

[43] Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala V., Li, Y., Fine, E. J., Wu, X., Shalem, O., Cradick, T. J., Marraffini, L. A., Bao, G., & Zhang, F. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnology, 31*(9), 827-832. doi: 10.1038/nbt.2647

[44] Ishii, T., & Araki, M. (2016). Consumer acceptance of food crops developed by genome editing. *Plant Cell Reports, 35*(7), 1507-1518. doi: 10.1007/s00299-016-1974-2

[45] Ishii, T., & Araki, M. (2017). A future scenario of the global regulatory landscape regarding genome-edited crops. *GM Crops & Food, 8*(1), 44-56. doi: 10.1080/21645698.2016.1261787

[46] Jacobs, T. B., LaFayette, P. R., Schmitz, R. J., & Parrott, W. A. (2015). Targeted genome modifications in soybean with CRISPR/Cas9. *BMC Biotechnology, 15*, 16. doi: 10.1186/s12896-015-0131-2

[47] Jao, L-E., Wente, S. R., & Chen, W. (2013). Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proceedings of the National Academy of Sciences, 110*(34), 13904-13909. doi: 10.1073/pnas.1308335110

[48] Jiang, F., & Doudna, J. A. (2017). CRISPR–Cas9 structures and mechanisms. *Annual Review of Biophysics, 46*, 505-529. doi: 10.1146/annurev-biophys-062215-010822

[49] Jin, S., Zong, Y., Gao, Q., Zhu, Z., Wang, Y., Qin, P., Liang, C., Wang, D., Qiu, J-L., 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

[50] Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. *Science, 337*(6096), 816-821. doi: 10.1126/science.1225829

[51] Jones, H. D. (2015). Future of breeding by genome editing is in the hands of regulators. *GM Crops & Food, 6*(4), 223-232. doi: 10.1080/21645698.2015.1134405

[52] Josephs, E. A., Kocak, D. D., Fitzgibbon, C. J., McMenemy, J., Gersbach, C. A., & Marszalek, P. E. (2015). Structure and specificity of the RNA-guided endonuclease Cas9 during DNA interrogation, target binding and cleavage. *Nucleic Acids Research, 43*(18), 8924-8941. doi: 10.1093/nar/gkv892

[53] Jouanin, A., Boyd, L., Visser, R. G. F., & Smulders, M. J. M. (2018). Development of wheat with hypoimmunogenic gluten obstructed by the gene editing policy in Europe. *Frontiers in Plant Science, 9*, 1523. doi: 10.3389/fpls.2018.01523

[54] Kaya, H., Mikami, M., Endo A., & Toki, S. (2016). Highly specific targeted mutagenesis in plants using Staphylococcus aureus Cas9. *Scientific Reports, 6*, 26871. doi: 10.1038/srep2 6871

[55] Kim, S., Kim, D., Cho, S. W., Kim, J. &, Kim, J-S. (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Research, 24*(6), 1012-1019. doi: 10.1101/gr.171322.113

[56] Kim, D., Kim, J., Hur, J. K., Been, K. W., Yoon, S-H., & Kim, J-S. (2016).

Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nature Biotechnology, 34*(8), 863-868. doi: 10.1038/nbt.3609

[57] Kleinstiver, B. P., Prew, M. S., Tsai, S. Q., Topkar, V. V., Nguyen, N. T., Zheng, Z., Gonzales, A. P. W., Li, Z., Peterson, R. T., Joanna Yeh, J-R., & Aryee, M. J. (2015). Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature, 523*(7561), 481-485. doi: 10.1038/nature14592

[58] Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Zheng, Z., & Joung, J. K. (2016). High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature, 529*, 490-495. doi: 10.1038/nature16526

[59] Kleter, G. A., Kuiper, H. A., & Kok, E. J. (2019). Gene-edited crops: towards a harmonized safety assessment. *Trends in Biotechnology, 37*(5), 443-447. doi: 10.1016/j.tibtech.2018.11.014

[60] Konermann, S., Brigham, M. D., Trevino, A. E., Joung, J., Abudayyeh, O. O., Barcena, C., Hsu, P. D., Habib, N., Gootenberg, J. S., Nishimasu, H., Nureki, O., & Zhang, F. (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature, 517*(7536), 583-588. doi: 10.1038/nature14136

[61] Kosicki, M., Tomberg, K., & Bradley, A. (2018). Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. *Nature Biotechnology, 36*(8), 765-771. doi: 10.1038/nbt.4192

[62] Lawrenson, T., Shorinola, O., Stacey, N., Li, C., Østergaard, L., Patron, N., Uauy, C., & Harwood, W. (2015). Induction of targeted, heritable mutations in barley and Brassica oleracea using RNA-guided Cas9 nuclease. *Genome Biology, 16*(1), 258. doi: 10.1186/s1305 9-015-0826-7

[63] Le Page, M. (2020). A French food fight. *New Scientist 245*(3274), 20-21. doi: 10.1016/S0262-4079(20)30589-3

[64] Lee, K., Zhang, Y., Kleinstiver, B. P., Guo, J. A., Aryee, M. J., Miller, J., Malzahn, A., Zarecor, S., Lawrence-Dill, C. J., Joung, J. K., Qi, Y., & Wang, K. (2018). Activities and specificities of CRISPR/Cas9 and Cas12a nucleases for targeted mutagenesis in maize. *Plant Biotechnology Journal, 17*(2), 362–372. doi: 10.1111/pbi.12982

[65] Lee, J. H., Mazarei, M., Pfotenhauer, A. C., Dorrough, A. B., Poindexter, M. R., Hewezi, T., Lenaghan, S. C., Graham, D. E., & Stewart, C. N., Jr. (2020). Epigenetic footprints of CRISPR/Cas9-mediated genome editing in plants. *Frontiers in Plant Science, 10*, 1720. doi: 10.3389/fpls.2019.01720

[66] Lema, M. A. (2019). Regulatory aspects of gene editing in Argentina. *Transgenic Research, 28*, 147-150. doi: 10.1007/s11248-019-00145-2

[67] Li, Z., Liu, Z-B., Xing, A., Moon, B. P., Koellhoffer, J. P., Huang, L. Ward, R. T., Clifton, E., Falco, S. C., & Cigan, A. M. (2015). Cas9-guide RNA directed genome editing in soybean. *Plant Physiology, 169*(2), 960-970. doi: 10.1104/pp.15.00783

[68] Liang, Z., Zhang, K., Chen, K., & Gao, C. (2014). Targeted mutagenesis in Zea mays using TALENs and the CRISPR/Cas system. *Journal of Genetics and Genomics, 41*(2), 63-68. doi: 10.1016/j.jgg.2013.12.001

[69] Lin, Q., Zong, Y., Xue, C., Wang, S., Jin, S., Zhu, Z., Wang, Y., Anzalone, A. V., Raguram, A., Doman, J. L., Liu, D. R., & Gao, C. (2020). Prime genome editing in rice and wheat. *Nature Biotechnology, 38*, 582-585. doi: 10.1038/s41587-020-0455-x

[70] Lusser, M., Parisi, C., Plan, D., & Rodríguez-Cerezo, E. (2011). *New plant breeding techniques. State-of-the-art and prospects for commercial development*. European Union. doi: 10.2791/54761

[71] Ma, X., Zhu, Q., Chen, Y., & Liu, Y-G. (2016). CRISPR/Cas9 platforms for genome editing in plants: Developments and applications. *Molecular Plant, 9*(7), 961-974. doi: 10.1016/j.molp.2016.04.009

[72] Maeder, M. L., Linder, S. J., Cascio, V. M., Fu, Y., Ho, Q. H., & Joung, J. K. (2013). CRISPR RNA–guided activation of endogenous human genes. *Nature Methods, 10*(10), 977-979. doi: 10.1038/nmeth.2598

[73] Mahfouz, M. M., Piatek, A., & Stewart, C. N., Jr. (2014). Genome engineering via TALENs and CRISPR/Cas9 systems: challenges and perspectives. *Plant Biotechnology Journal, 12*(8), 1006-1014. doi: 10.1111/pbi.12256

[74] Merling, R. K., Kuhns, D. B., Sweeney, C. L., Wu, X., Burkett, S., Chu, J., Lee, J., Koontz, S., Di Pasquale, G., Afione, S. A., Chiorini, J. A. ,Kang, E. M., Choi, U., De Ravin, S. S., & Malech, H. L. (2017). Gene-edited pseudogene resurrection corrects p47phox-deficient chronic granulomatous disease. *Blood Advances, 1*(4), 270-278. doi: 10.1182/bloodadvances.2016001214

[75] Ming, M., Ren, Q., Pan, C., He, Y., Zhang, Y., Liu, S., Zhong, Z., Wang, J., Malzahn, A. A., Wu, J., Zheng, X., Zhang, Y., & Qi, Y. (2020). CRISPR–Cas12b enables efficient plant genome engineering. *Nature Plants, 6*(3), 202-208. doi: 10.1038/s41477-020-0614-6

[76] Modrzejewski, D., Hartung, F., Sprink, T., Krause, D., Kohl, C., Schiemann, J., & Wilhelm, R. (2018). What is the available evidence for the application of genome editing as a new tool for plant trait modification and the potential occurrence of associated off-target effects: a systematic map protocol. *Environmental Evidence, 7*, 18. doi: 10.1186/s13750-018-0130-6

[77] Morbitzer, R., Römer, P., Boch, J., & Lahaye, T. (2010). Regulation of selected genome loci using de novo-engineered transcription activator-like effector (TALE)-type transcription factors. *Proceedings of the National Academy of Sciences, 107*(50), 21617-21622. doi: 10.1073/pnas.1013133107

[78] Mussolino, C., & Cathomen, T. (2012). TALE nucleases: tailored genome engineering made easy. *Current Opinion in Biotechnology, 23*(5), 644-650. doi: 10.1016/j.copbio.2012.01.013

[79] National Academies of Sciences, Engineering, and Medicine. (2016). Regulation of Current and Future Genetically Engineered Crops. In *Genetically Engineered Crops: Experiences and Prospects*. The National Academies Press.

[80] Nekrasov, V., Wang, C., Win, J. Lanz, C., Weigel, D., & Kamoun, S. (2017). Rapid generation of a transgene-free powdery mildew resistant tomato by genome deletion. *Scientific Reports, 7*, 482. doi: 10.1038/s41598-017-00578-x

[81] Nihongaki, Y., Kawano, F., Nakajima, T., & Sato, M. (2015). Photoactivatable CRISPR-Cas9 for optogenetic genome editing. *Nature Biotechnology, 33*(7), 755-760. doi: 10.1038/nbt.3245

[82] Nishitani, C., Hirai, N., Komori, S., Wada, M., Okada, K., Osakabe, K., Yamamoto, T., & Osakabe, Y. (2016). Efficient genome editing in apple using a CRISPR/Cas9 system. *Scientific Reports, 6*, 31481. doi: 10.1038/srep31481

[83] Norris, A. L., Lee, S. S., Greenlees, K. J., Tadesse, D. A., Miller, M. F., & Lombardi, H. A. (2020). Template plasmid integration in germline genome-edited cattle. *Nature Biotechnology, 38*(2), 163-164. doi: 10.1038/s41587-019-0394-6

[84] Organisation for Economic Co-operation and Development. (2020). *Documents on Harmonisation of Regulatory Oversight in Biotechnology and the Safety of Novel Foods and Feeds*. Retrieved from https://www.oecd.org/chemicalsafety/biotrack/documentsonharmonisationofregulatoryoversightinbiotechnologyandthesafetyofnovelfoodsandfeeds.htm

[85] Ono, R., Yasuhiko, Y., Aisaki, K-I., Kitajima, S., Kanno, J., & Hirabayashi, Y. (2019). Exosome-mediated horizontal gene transfer occurs in double-strand break repair during genome editing. *Communications Biology, 2*, 1-8. doi: 10.1038/s42003-019-0300-2

[86] Pattanayak, V., Lin, S., Guilinger, J. P., Ma, E., Doudna, J. A., & Liu, D. R. (2013). High-throughput profiling of off-target DNA cleavage reveals RNA programmed Cas9 nuclease specificity. *Nature Biotechnology, 31*, 839-843. doi: 10.1038/nbt.2673

[87] Pattanayak, V., Ramirez, C. L., Joung, J. K., & Liu, D. R. (2011). Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nature Methods, 8*(9), 765-770. doi: 10.1038/nmeth.1670

[88] Pauwels, K., Podevin, N., Breyer, D., Carroll, D., & Herman, P. (2014). Engineering nucleases for gene targeting: safety and regulatory considerations. *New Biotechnology, 31*(1), 18-27. doi: 10.1016/j.nbt.2013.07.001

[89] Peterson, B. A., Haak, D. C., Nishimura, M. T., Teixeira, P. J. P. L., James, S. R., Dangl, J. L., & Nimchuk, Z. L. (2016). Genome-wide assessment of efficiency and specificity in CRISPR/Cas9 mediated multiple site targeting in Arabidopsis. *PloS one, 11*(9), e0162169. doi: 10.1371/journal.pone.0162169

[90] Podevin, N., Davies, H. V., Hartung, F., Nogué, F., & Casacuberta, J. M. (2013). Site-directed nucleases: a paradigm shift in predictable, knowledge-based plant breeding. *Trends in Biotechnology, 31*(6), 375-383. doi: 10.1016/j.tibtech.2013.03.004

[91] Ran, F. A., Hsu, P. D., Lin, C-Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., Scott, D. A., Inoue, A., Matoba, S., & Zhang, F. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell, 154*(6), 1380-1389. doi: 10.1016/j.cell.2013.08.021

[92] Riveros-Maidana, R., Méndez-Ferreira, A., Benítez-Candia, N., Nara-

Pereira, E., & Fernández-Ríos, D. (2020). CRISPR/Cas system: precision genome editing. *Memorias del Instituto de Investigaciones en Ciencias de la Salud, 18*(1), 97-107. doi: 10.18004/mem.iics/1812-9528/2020.018.01.97-107

[93] Rowland, G. G. (2009). The Effect of Plants with Novel Traits (PNT) Regulation on Mutation Breeding in Canada. In Q. Y. Shu (Ed.), *Induced Plant Mutations in the Genomics Era, Food and Agriculture Organization of the United Nations* (pp. 423-424). Food and Agriculture Organization of the United Nations, Rome.

[94] Sander, J. D., Ramirez, C. L., Linder, S. J., Pattanayak, V., Shoresh, N., Ku, M., Foden, J. A., Reyon, D., Bernstein, B. E., Liu, D. R., & Joung, J. K. (2013). *In silico* abstraction of zinc finger nuclease cleavage profiles reveals an expanded landscape of off-target sites. *Nucleic Acids Research, 41*(19), e181. doi: 10.1093/nar/gkt716

[95] Sander, J. D., & Joung, J. K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature Biotechnology, 32*, 347-355. doi: 10.1038/nbt.2842

[96] Sansbury, B. M., Hewes, A. M., & Kmiec, E. B. (2019). Understanding the diversity of genetic outcomes from CRISPR-Cas generated homology-directed repair. *Communications Biology, 2*, 1-10. doi: 10.1038/s42003-019-0705-y

[97] Sauer, N. J., Mozoruk, J., Miller, R. B., Warburg, Z. J., Walker, K. A., Beetham, P. R., Schöpke, C. R., & Gocal, G. F. W. (2016). Oligonucleotide-directed mutagenesis for precision gene editing. *Plant Biotechnology Journal, 14*(2), 496-502. doi: 10.1111/pbi.12496

[98] Schiemann, J., Dietz-Pfeilstetter, A., Hartung, F., Kohl, C., Romeis, J., & Sprink, T. (2019). Risk assessment and regulation of plants modified by modern biotechniques: current status and future challenges. *Annual Review of Plant Biology, 70*, 699-726. doi: 10.1146/annurev-arplant-050718-100025

[99] Schiemann, J., Robienski, J., Schleissing, S., Spök, A., Sprink, T., & Wilhelm, R. A. (2020). Plant Genome Editing–Policies and Governance. *Frontiers in Plant Science*, 11:284. doi: 10.3389/fpls.2020.00284

[100] Sentmanat, M. F., Peters, S. T., Florian, C. P., Connelly, J. P., & Pruett-Miller, S. M. (2018). A survey of validation strategies for CRISPR-Cas9 editing. *Scientific Reports, 8*(1), 1-8. doi: 10.1038/s41598-018-19441-8

[101] Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., Zhang, K., Liu, J., Xi, J. J., Qiu, J-L., & Gao, C. (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. *Nature Biotechnology, 31*(8), 686-688. doi: 10.1038/nbt.2650

[102] Shan, Q., Wang, Y., Li, J., & Gao, C. (2014). Genome editing in rice and wheat using the CRISPR/Cas system. *Nature Protocols, 9*(10), 2395-2410. doi: 10.1038/nprot.2014.157

[103] Shearer, H. (2014). Regulation of plants with novel traits: Canadian perspectives on the "novelty" trigger. In A. Eaglesham, & R. W. F. Hardy (Eds), *New DNA-Editing Approaches: Methods, Applications and Policy for Agriculture* (pp. 193-200). North American Agricultural Biotechnology Council, Ithaca, NY.

[104] Silva, G., Poirot, L., Galetto, R., Smith, J., Montoya, G., Duchateau, P., & Pâques, F. (2011). Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy. *Current Gene Therapy, 11*(1), 11-27. doi: 10.2174/156652311794520111

[105] Skryabin, B. V., Kummerfeld, D-M., Gubar, L., Seeger, B., Kaiser, H., Stegemann, A., Roth, J., Meuth, S. G., Pavenstädt, H., Sherwood, J., Pap, T., Wedlich-Söldner, R., Sunderkötter, C., Schwartz, Y. B., Brosius, J., & Rozhdestvensky, T. S. (2020). Pervasive head-to-tail insertions of DNA templates mask desired CRISPR-Cas9–mediated genome editing events. *Science Advances, 6*(7), eaax2941. doi: 10.1126/sciadv.aax2941

[106] Slaymaker, I. M., Gao, L., Zetsche, B., Scott, D. A., Yan, W. X., & Zhang, F. (2016). Rationally engineered Cas9 nucleases with improved specificity. *Science, 351*(6268), 84-88. doi: 10.1126/science.aad5227

[107] Smith, C., Gore, A., Yan, W., Abalde-Atristain, L., Li, Z., He, C., Wang, Y., Brodsky, R. A., Zhang, K., Cheng, L., & Ye, Z. (2014). Whole-genome sequencing analysis reveals high specificity of CRISPR/Cas9 and TALEN-based genome editing in human iPSCs. *Cell Stem Cell, 15*(1), 12-13. doi: 10.1016/j.stem.2014.06.011

[108] Sprink, T., Metje, J., & Hartung, F. (2015). Plant genome editing by novel tools: TALEN and other sequence specific nucleases. *Current Opinion in Biotechnology, 32*, 47-53. doi: 10.1016/j.copbio.2014.11.010

[109] Sprink, T., Eriksson, D., Schiemann, J., & Hartung, F. (2016a). Regulatory hurdles for genome editing: process-vs. product-based approaches in different regulatory contexts. *Plant Cell Reports, 35*(7), 1493-1506. doi: 10.1007/s00299-016-1990-2

[110] Sprink, T., Metje, J., Schiemann, J., & Hartung, F. (2016b). Plant genome editing in the European Union - to be or not to be - a GMO. *Plant Biotechnology Reports, 10*(6), 345-351. doi: 10.1007/s11816-016-0418-3

[111] Srinivas, K. R. (2018). Regulating Genome Edited Crops and European Court of Justice Ruling. *Asian Biotechnology & Development Review, 20*(1&2), 89-87.

[112] Steinert, J., Schiml, S., Fauser, F., & Puchta, H. (2015). Highly efficient heritable plant genome engineering using Cas9 orthologues from Streptococcus thermophilus and Staphylococcus aureus. *Plant Journal, 84*, 1295-1305. doi: 10.1111/tpj.13078

[113] Stepper, P., Kungulovski, G., Jurkowska, R. Z., Chandra, T., Krueger, F., Reinhardt, R., Reik, W., & Jurkowski, T. P. (2017). Efficient targeted DNA methylation with chimeric dCas9–Dnmt3a–Dnmt3L methyltransferase. *Nucleic Acids Research, 45*(4), 1703-1713. doi: 10.1093/nar/gkw1112

[114] Strohkendl, I., Saifuddin, F. A., Rybarski, J. R., Finkelstein, I. J., & Russell, R. (2018). Kinetic basis for DNA target specificity of CRISPR-Cas12a. *Molecular Cell, 71*(5), 816-824. doi: 10.1016/j.molcel.2018.06.043

[115] Svitashev, S., Young, J. K., Schwartz, C., Gao, H., Falco, S. C., & Cigan, M. (2015). Targeted mutagenesis, precise gene editing, and site-specific gene insertion in maize using Cas9 and guide RNA. *Plant Physiology, 169*(2), 931-945. doi: 10.1104/pp.15.00793

[116] Swarts, D. C., van der Oost, J., & Jinek, M. (2017). Structural basis for guide RNA processing and seed-dependent DNA targeting by CRISPR-Cas12a. *Molecular Cell, 66*(2), 221-233. doi: 10.1016/j.molcel.2017.03.016

[117] 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 Biology, 19*(1), 84. doi: 10.1186/s13059-018-1458-5

[118] Thygesen, P. (2019). Clarifying the regulation of genome editing in Australia: situation for genetically modified organisms. *Transgenic Research, 28*(2), 151-159. doi: 10.1007/s11248-019-00151-4

[119] Tsai, S. Q., Wyvekens, N., Khayter, C., Foden, J. A., Thapar, V., Reyon, D., Goodwin, M. J., Aryee, M. J., & Joung, J. K. (2014). Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nature Biotechnology, 32*(6), 569-576. doi: 10.1038/nbt.2908

[120] Tsuda, M., Watanabe, K. N., & Ohsawa, R. (2019). Regulatory Status of Genome-Edited Organisms Under the Japanese Cartagena Act. *Frontiers in Bioengineering and Biotechnology, 7*, 387. doi: 10.3389/fbioe.2019.00387

[121] Tycko, J., Myer, V. E., & Hsu, P. D. (2016). Methods for optimizing CRISPR-Cas9 genome editing specificity. *Molecular Cell, 63*(3), 355-370. doi: 10.1016/j.molcel.2016.07.004

[122] Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S., & Gregory, P. D. (2010). Genome editing with engineered zinc finger nucleases. *Nature Reviews Genetics, 11*(9), 636-646. doi: 10.1038/nrg2842

[123] Vojta, A., Dobrinić, P., Tadić, V., Bočkor, L., Korać, P., Julg, B., Klasić, M., & Zoldoš, V. (2016). Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic Acids Research, 44*(12), 5615-5628. doi: 10.1093/nar/gkw159

[124] Whelan, A. I., & Lema, M. A. (2015). Regulatory framework for gene editing and other new breeding techniques (NBTs) in Argentina. *GM Crops & Food, 6*(4), 253-265. doi: 10.1080/21645698.2015.1114698

[125] Whelan, A. I., & Lema, M. A. (2017). A research program for the socioeconomic impacts of gene editing regulation. *GM Crops & Food, 8*(1), 74-83. doi: 10.1080/21645698.2016.1271856

[126] Whelan, A. I., & Lema, M. A. (2019). Regulation of Genome Editing in Plant Biotechnology: Argentina. In H-G. Dederer, & D. Hamburger (Eds.), *Regulation of Genome Editing in Plant Biotechnology: A Comparative Analysis of Regulatory Frameworks of Selected Countries and the EU* (pp. 19-62). Springer.

[127] Whelan, A. I., Gutti, P., & Lema, M. A. (2020). Gene editing regulation and innovation economics. *Frontiers in Bioengineering and Biotechnology, 8*, 303. doi: 10.3389/fbioe.2020.00303

[128] Wolter, F., Klemm, J., & Puchta, H. (2018). Efficient in planta gene

targeting in Arabidopsis using egg cell-specific expression of the Cas9 nuclease of Staphylococcus aureus. *The Plant Journal, 94*(4), 735-746. doi: 10.1111/tpj.13893

[129] Wong, N., Liu, W., & Wang, X. (2015). WU-CRISPR: characteristics of functional guide RNAs for the CRISPR/Cas9 system. *Genome Biology, 16*(1), 218. doi: 10.1186/s13059-015-0784-0

[130] Woo, J. W., Kim, J., Kwon, S. I., Corvalán, C., Choo, S. W., Kim, H., Kim, S-G., Kim, S-T., Choe, S., & Kim, J-S. (2015). DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nature Biotechnology, 33*(11), 1162-1164. doi: 10.1038/nbt.3389

[131] Woodgate, J., Palfrey, D., Nagel, D. A., Hine, A. V., & Slater, N. K. H. (2002). Protein-mediated isolation of plasmid DNA by a zinc finger-glutathione S-transferase affinity linker. *Biotechnology and Bioengineering, 79*(4), 450-456. doi: 10.1002/bit.10296

[132] World Health Organization, & Food and Agriculture Organization of the United Nations. (n.d.). *The role of Codex in Biotechnology*. Retrieved from http://www.fao.org/fao-who-codexalimentarius/thematic-areas/biotechnology/en/

[133] Wright, D. A., Li, T., Yang, B., & Spalding, M. H. (2014). TALEN-mediated genome editing: prospects and perspectives. *Biochemical Journal, 462*(1), 15-24. doi: 10.1042/BJ20140295

[134] Xie, K., & Yang, Y. (2013). RNA-guided genome editing in plants using a CRISPR–Cas system. *Molecular Plant, 6*(6), 1975-1983. doi: 10.1093/mp/sst119

[135] Xu, X., Duan, D., & Chen, S-J. (2017). CRISPR-Cas9 cleavage efficiency correlates strongly with target-sgRNA folding stability: from physical mechanism to off-target assessment. *Scientific Reports, 7*(1), 143. doi: 10.1038/s41598-017-00180-1

[136] Yee, J. K. (2016). Off-target effects of engineered nucleases. *The FEBS journal, 283*(17), 3239-3248. doi: 10.1111/febs.13760

[137] Young, A. E., Mansour, T. A., McNabb, B. R., Owen, J. R., Trott, J. F., Brown, C. T., & Van Eenennaam, A. L. (2020). Genomic and phenotypic analyses of six offspring of a genome-edited hornless bull. *Nature Biotechnology, 38*(2), 225-232. doi: 10.1038/s41587-019-0266-0

[138] Zhang, F., Maeder, M. L., Unger-Wallace, E., Hoshaw, J. P., Reyon, D., Christian, M., Li, X., Pierick, C. J., Dobbs, D., Peterson, T., Joung, J. K., & Voytas, D. F. (2010). High frequency targeted mutagenesis in Arabidopsis thaliana using zinc finger nucleases. *Proceedings of the National Academy of Sciences of the United States of America, 107*(26), 12028-12033. doi: 10.1073/pnas.0914991107

[139] Zhang, X-H., Tee, L. Y., Wang, X-G., Huang, Q-S., & Yang, S-H. (2015). Off-target effects in CRISPR/Cas9-mediated genome engineering. *Molecular Therapy - Nucleic Acids, 4*, e264. doi: 10.1038/mtna.2015.37

[140] Zhang, Q., Xing, H-L., Wang, Z-P., Zhang, H-Y., Yang, F., Wang, X-C., & Chen, Q-J. (2018). Potential high-frequency off-target mutagenesis induced by CRISPR/Cas9 in Arabidopsis and its prevention. *Plant Molecular Biology, 96*(4-5), 445-456. doi: 10.1007/s11103-018-0709-x

[141] Zhang, Y., Pribil, M., Palmgren, M., & Gao, C. (2020). A CRISPR way for accelerating improvement of food crops. *Nature Food, 1*, 200-205. doi: 10.1038/s43016-020-0051-8

[142] Zhou, H., Liu, B., Weeks, D. P., Spalding, M. H., & Yang, B. (2014). Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. *Nucleic Acids Research, 42*(17), 10903-10914. doi: 10.1093/nar/gku806

[143] Zuo, E., Sun, Y., Wei, W., Yuan, T., Ying, W., Sun, H., Yuan, L., Steinmetz, L. M., Li, Y., & Yang, H. (2019). Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science, 364*(6437), 289-292. doi: 10.1126/science.aav9973