



Detection of genome edits in plants—from editing to seed

Raymond D. Shillito¹ · Sherry Whitt¹ · Margit Ross² · Farhad Ghavami³ · David De Vleeschauwer⁴ ·
Katelijn D'Halluin⁴ · Annelies Van Hoecke⁴ · Frank Meulewaeter⁴

Received: 28 June 2021 / Accepted: 29 June 2021 / Published online: 23 September 2021 / Editor: David Songstad
© The Author(s) 2021

Abstract

Genome editing (also known as gene editing) employs a range of tools such as Meganucleases, Zinc Finger Nucleases, TALENs, and more recently CRISPR to make defined changes in genes, regulatory sequences, untranslated regions, or intergenic regions. It is increasingly being applied in plant science research and to improve plant varieties. The benefits of having effective detection tools begin with optimization of the genome editing process itself and continue with selection and characterization of tissue cultures and/or regenerated plants. Detection tools are also used throughout the breeding process, and for preparation of regulatory dossiers when required, as well as for seed production, and may be necessary for monitoring products in the marketplace. Detection and identification of genome edits employs a wide range of analytical approaches including PCR, digital PCR, and sequencing methods. This article examines the applicability of each category of detection or identification approach, from the optimization of genome editing processes, through creation of edits, selection and characterization, and breeding. The challenges surrounding the detection of genome edits present at low levels in large seed, plant, or grain populations and of differentiating directed genome edits from conventional mutations are also explained.

Keywords Gene editing · CRISPR · Detection · PCR · Sequencing · Genome editing

Introduction

Plant breeders have traditionally relied on selection of suitable genotypes from populations, or after wide crosses to leverage the naturally occurring variability in the genome. Mutagenesis agents such as radiation or chemical mutagens have routinely been used to increase genetic variability. Genome editing tools (TALENs, CRISPR/Cas) allow enhanced precision and efficiency in creating genetic diversity and is one of the tools increasingly being applied in plant sciences research and to improve plant

varieties (Bailey-Serres *et al.* 2019; Varshney *et al.* 2020). Such technology has also been termed gene editing. However, edits can be made in genes, regulatory sequences, untranslated regions, or intergenic regions. Therefore, the term genome editing will be used throughout this article.

The need for detection tools begins with optimization of the genome editing process itself and continues with characterization of tissue cultures and/or regenerated plants. Detection tools are also used throughout the breeding process, as well as for seed production and the subsequent launch of products. The analytical techniques employed to achieve these aims vary between the different stages, but the methods often rely on the in-house knowledge of the sequences and genomes. When genome-edited crops are deployed as commercial products, then there will be a need to track valuable crops and products, to mitigate trade risks, and in some cases to achieve approval if the product is regulated.

Detection and identification of genome edits employ a wide range of analytical approaches throughout the development process. These include phenotypic characterization, PCR, digital PCR, and sequencing methods. Optimization of genome editing processes requires a range of analytical approaches as detection in tissue cultures can be challenging

✉ Raymond D. Shillito
raymond.shillito@basf.com

¹ BASF Corporation, 407 Davis Drive, Morrisville, NC 27560, USA

² Corteva Agriscience, 8325 NW 62nd Avenue, P.O. Box 7060, Johnston, IA 50131, USA

³ Eurofins BioDiagnostics Inc., 507 Highland Dr, River Falls, WI 54022, USA

⁴ BASF Belgium Coordination Center CommV, Innovation Center Gent, Technologiepark - Zwijnaarde 101, 9052 Gent, Zwijnaarde, Belgium

when a mixture of edits may be present in an unedited germplasm background, or in early phase plant populations. Applications of detection tools to single plants at early stages of the genome editing process may not be as challenging and presence/absence or zygosity assays (via PCR or digital PCR) are sufficient for this purpose. Once breeding populations are developed and seed is being readied for commercialization, then analytical sensitivity again becomes more critical. Identification of a low level of genome-edited product with a small sequence difference in a bulk seed or grain samples either for testing purity of edited populations or to ensure they are absent in non-edited materials can be challenging in many cases.

This article will cover the detection tools used in optimization of genome editing and the early stages of research and in breeding. We will also describe the general issues surrounding the detection of genome edits present at low levels in large seed, plant, or grain populations. The detection of genome edits in many cases also implies characterization of the modification.

What are genome edits

Genome editing is a term commonly used to describe site-directed mutagenesis techniques that allow the introduction of targeted changes in specific DNA sequences at a defined location in the genome of an organism (Zhan *et al.* 2021). Details of the process are described in the next section. The result of such an operation subsequently leads to “genome-edited plants” as described in other articles in this issue. Oligonucleotide-directed mutagenesis (ODM) and site-directed nucleases (SDNs) allow enhanced precision and efficiency of edits to be applied. Early gene editing applications involved the use of meganucleases (Epinat *et al.* 2003), zinc finger nucleases (ZFNs) (Porteus and Baltimore 2003), and transcription activator-like effector nucleases (TALENs) ((Bogdanove and Voytas 2011; Chen and Lin 2013). More recently, the advent of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR/Cas) SDN editing has enabled an increase in the application of genome editing tools to plant modification and breeding (Chen *et al.* 2019).

Introduction of genome edits

Genome editing is achieved by the introduction of a single-stranded nick or DNA double-strand break (DSB) at a chosen specific chromosomal location using different types of SDNs. SDNs of the CRISPR/Cas class use only a short protospacer sequence within their CRISPR RNA (crRNA) to identify the target sequence, while for Meganucleases ZFNs and TALENs, a new enzyme complex has to be generated for each target. DNA double-strand breaks induced by the SDN can be

repaired via 2 pathways, either non-homologous end joining (NHEJ) or homologous recombination (HR). Higher plants predominantly use the error-prone NHEJ pathway for DNA repair. Depending on the desired edit, nicks and breaks can be introduced at a single site or at multiple sites. The repair mechanisms for SDN1 are based on NHEJ, which due to the lack of fidelity can result in randomly introduced or deleted nucleotide(s). Conversely, repair mechanisms for SDN2 and SDN3 are based on HR between the genome target sequence and an exogenously delivered donor DNA engineered with intended edits. Genome-edited plants are usually obtained by delivering SDNs, and a donor DNA in the case of homologous recombination, into cells or explants in culture. Plants are regenerated from these cultures by standard approaches. Currently, *Agrobacterium tumefaciens* and particle bombardment remain the main methods of choice for delivery of nuclease reagents into plant cells or explants. A selectable marker can be used to select for plant cells or tissues with integrated constructs. These constructs are intended to induce the desired edits and can then be removed from the recovered genome-edited plants by segregation in the progeny. CRISPR nucleases or other effectors can also be transiently expressed in plant cells or CRISPR reagents can be delivered in a DNA-free manner using ribonucleoproteins (RNPs) (Hamada *et al.* 2018).

Protoplast systems have seen a recent revival for the introduction of CRISPR/Cas components as DNA and RNPs (Woo *et al.* 2015; Liang *et al.* 2017; Lin *et al.* 2018). Nevertheless, regeneration of plants from protoplasts remains a major bottleneck for many important crop species. Plant protoplasts also provide a robust platform for rapid validation of genome editing reagents via transient expression and for screening for optimal editing performance. An advantage is that the whole procedure from isolation of protoplasts until assessment of the activity of the nuclease reagents can be done rapidly.

When CRISPR or other effector systems need to be expressed in plants, efficient promoters and plant codon optimized versions of these nuclease genes are generally used (Zhan *et al.* 2021). RNA polymerase (Pol) II-dependent promoters with strong expression in reproductive cells or corresponding ancestor cells are good candidate promoters to express the nuclease genes. The non-coding single-guide RNAs (sgRNAs) are more suited for transcription by Pol III-dependent promoters, such as U6 and U3 promoters. Multiplex editing is possible by the delivery of a set of sgRNAs typically assembled in tandem in plant expression vectors using either identical or different Pol III-dependent promoters (Zhang *et al.* 2015a). Another possibility is to assemble the multiple sgRNAs into a single transcription unit under the control of a single Pol II promoter. In this case, the primary transcript must be processed to generate multiple mature sgRNAs by making use of self-cleaving RNAs or cleavable RNA molecules such as *csy4*, ribozyme, and tRNA sequences (Xie *et al.* 2015; Čermák *et al.* 2017).

Table 1 Genome edits classified according to the edit type and mechanism of repair.

	SDN-1	SDN-2	SDN-3
Size of edit	A few base changes	Defined point mutation to a few bases	Up to several kilobases
Mechanism	Spontaneous mutations, deletions, replacements, additions of sequence ^a	Pre-defined mutations, sequence optimization, allele replacement ^a	Addition of sequence at the target genomic location ^a
Result	Targeted but not defined deletions/mutations	Targeted and precise mutations/gene edits. No insertion of transgenic sequence	Targeted insertion of sequence. Insertion of cisgenic or transgenic sequence ^a

^a Source: <https://croplife.org/wp-content/uploads/2015/01/CLI-SDN-Definitions-Position-Paper.pdf>

To make genome editing broadly applicable, effective DNA delivery methods and efficient cell and tissue culture procedures allowing the manipulation and regeneration of large numbers of cells need to be available. For many crop species, these have still to be further developed or improved (Altpeter *et al.* 2016) as referenced by other articles in this Journal issue). Developmental genes such as *Baby Boom*, *Wuschel*, and *Grf5* or a chimeric *Grf5* protein have been shown to increase regeneration and transformation in a variety of plants, tissue types, and genotypes previously identified as recalcitrant for transformation and regeneration (Lowe *et al.* 2016; Gordon-Kamm *et al.* 2019; Debernardi *et al.* 2020; Kong *et al.* 2020). Tissue culture-free procedures for genome editing have also been described. Maher *et al.* (2020) describes a non-tissue culture method for the generation of genome edited plants through *de novo* induction of meristems using delivery of developmental regulators and gene editing reagents in somatic cells of whole plants. Other reports on “in planta” genome editing methods are based on infecting plants with an engineered virus expressing the genome editing reagents (Ma *et al.* 2020) or via bombardment of wheat embryos (Liu *et al.* 2021). Further advances in the field of tissue culture and plant regeneration, as well as tissue-culture-independent genome editing technologies combined with an increased knowledge on which gene(s) to edit in order to improve a particular trait, will gradually broaden the applications of genome editing for crop breeding.

Types of genome-editing outcomes

There are several different outcomes that can be generated using genome editing tools and have been commonly referred to as SDN-1, SDN-2, and SDN-3 edits as described in Table 1. These classifications of genome editing are based on whether the genome editing approach makes use of a repair template and on the extent to which the edited sequences differ from their predecessor sequences and from genome variations found in populations. The types of DNA alterations produced include insertions (both of cis and exogenous DNA),

deletions, combinations of the two (indels), and edits of a few or single base pairs (Table 1).

Selection and design of appropriate detection tools depends on the analytical purpose and type and size of the genome edit, the frequency of the edit in the population, *a priori* knowledge of the endogenous sequence, and the nature of the surrounding sequence.

1) Single or few base changes

Single (or few base) changes in sequence may be relatively easy or more difficult to identify depending on the surrounding sequences, especially when using traditional approaches such as PCR. There are several PCR-based methods that use a combination of cleavage enzymes or DNA mismatches. Single-nucleotide polymorphism (SNP) detection using allele specific PCR has been extensively utilized and optimized within genomics. In addition, single base detection with TaqMan^{®1} or KASP^{™2} (Kompetitive Allele Specific PCR) has become commonplace. Use of TaqMan for this purpose requires specialty primers and/or probes that incorporate enhancements such as minor groove binders (MGB; Kutuyavin *et al.* 2000; Davalieva *et al.* 2014) or locked nucleic acids (LNA; Johnson *et al.* 2004; Karkare and Bhatnagar 2006; Maertens *et al.* 2006) which increase the melting temperature of short sequence probes, allows for increased specificity, and correspondingly lowers the background signal. Detection of small changes is further discussed in the context of each detection application.

2) Deletions

Detection of a DNA deletion, for example in Waxy corn (Qi *et al.* 2020), depends upon the size of the deletion and *a priori* knowledge of the endogenous sequence prior to editing. Small deletions may occur during the SDN-1 repair process. In this case, primer design will need to be adapted to the specific change in sequence at

¹ TaqMan is a registered trademark of Roche Molecular Systems, Inc. Pleasanton, California

² KASP is a trademark of LGC Genomics, Limited. Teddington, Middlesex, England

the site of the edit. Large deletions may lead to the proximity of sequences that are not normally co-located. As with insertions, the alterations provide a target for PCR primer and probe design to differentiate the edit from the original sequence, especially if they can be positioned at or around the site of the deletion. Alternatively, PCR primers and probes designed to identify DNA sequences intended for deletion can also be used as a diagnostic for absence of the intended deletion indicating an unsuccessful attempt at editing, or to differentiate seed or grain not containing the edit within an edited population.

3) Small insertions

Small insertions (SNPs, amino acid substitutions, etc.) are typically generated via homology-directed repair using DNA donor oligos. With maximum insert lengths for use with a DNA donor oligo restricted to approximately 50 nucleotides, these changes may be relatively easy or more difficult to identify depending on the surrounding sequences, especially when using approaches such as PCR. Whether the region is AT or GC rich or repetitive can affect the ability to find good locations for primers and/or probes, as can the precise nature of the change. There is insufficient information available at present to make predictions as to whether each specific sequence change will be difficult to detect.

4) Larger insertions

Sites containing an insertion of larger sizes are relatively easy to detect using standard techniques such as PCR. Where the insertion is large enough to allow the positioning of primer and/or a probe on the inserted sequence, the procedure is the same as with transgenic events. There is not yet enough research done to determine the minimum size of the insertion that is required to develop a robust assay. However, the robustness and sensitivity of the assay may be lower when a small insertion is present than when a large insertion is present.

Should the insertion be cis-DNA (i.e., from the same species), the detection challenge is not increased as the novel juxtaposition of two sequences is the critical factor for detection, not necessarily the origin of the sequences.

A significant caveat concerning detection of DNA changes is that these modifications can also occur spontaneously in populations. De Maagd *et al.* (2020) stated that “many types of structural changes can occur spontaneously during cultivation and breeding, especially during conventional mutagenesis” and that most structural variations that arise by natural genome evolution are unnoticed. There are many similar observations, both without intervention (Gorbunova and Levy 1997) and as described in radiation induced mutation (Jo and Kim 2019), and hybridization (Bashir *et al.* 2018). In some plants such as maize, changes in the genome can occur at high frequency and in locations that vary extensively

between populations due to the presence of transposons and other active genome altering mechanisms (Bennetzen 2000).

Tools for detection of genome edits

Detection tools are an important consideration when developing genome-edited plants. To obtain the most efficient system, it is important that the success or lack thereof of a genome-editing procedure is determined as quickly as possible. A particular genome editing system can be initially tested and optimized using traits successfully described in the literature and/or having an obvious phenotype. The proof of the desired outcome can therefore be repeated and visualized. However, confirmation of the intended edit will need to be established at the nucleotide level and if a non-selection approach is used, it will be important to identify the cells or plants carrying the edits of interest within the population.

Once an edited plant is produced and selected, detection tools can be used to confirm that the plant has the desired edit, evaluate any off-target changes, and track the intended genetic change throughout the entire breeding process. Proof of the effectiveness of the edit in producing the desired phenotype is monitored through the next phases of product development.

PCR

In the past, PCR methods were traditionally carried out as end-point reactions—where the desired result is a simple presence/absence with visualization of the results routinely performed via gel electrophoresis (Lipp *et al.* 2005). However, such methods have been superseded for many purposes (Alarcon *et al.* 2019) and are limited in the sensitivity and types of genome edits that can be detected (Zischewski *et al.* 2017). We will therefore concentrate on PCR approaches that are the most applicable to this goal.

PCR is an appropriate choice for low cost and ease of use with a wide range of insertion and deletion sizes. It can be applied either as an end-point fluorescence reaction (e.g., T7E1, Surveyor, RFLP, CAPS) or as a quantitative approach, where the amplification curve is monitored. These offer the ability to illuminate guide RNA efficacy by distinguishing mutant from wild-type cells (Zischewski *et al.* 2017; Lomov *et al.* 2019). Capillary PCR fragment size analysis can also be used to detect differences as small as a few base pairs. Depending upon the size of the edit, PCR may be followed by polyacrylamide or agarose gel-based visualization (a few base pairs up to kilobases). The band intensities can provide information regarding the ratio of mutated to unmutated cells (Lomov *et al.* 2019).

Development of PCR methods requires *a priori* knowledge of the nucleotide sequence around the area to be amplified. The sequence in the immediate edited region is known to the

developers of the edit, as it is required for design of guide RNAs used in the editing process. However, this may not be publicly available, and/or extensive sequence information, as might be associated with a large insertion, may not be readily available. Sanger sequencing of an amplicon coupled with primer walking can be used to elucidate uncharacterized regions adjacent to indels. Regardless of edit type, sequencing should be applied for verification of an edit. Once the intended edit is established in a population, the shift to high-throughput and more sensitive PCR methods can be made.

Both real-time qPCR (e.g., TaqMan) and endpoint PCR (e.g., KASP) are widely applied to the detection of SNPs, indels, and insertion sites for numerous mutagenesis and other biotechnology approaches. A product of genome editing is essentially no different. The development and validation of such assays has become highly streamlined with the advent of design algorithms and proprietary chemistries. The use of such methods for SNP genotyping involves a competitive PCR. The native and edited alleles are targeted in the same reaction with two separate primers or probes and an associated fluorophore. PCR efficiency must be equivalent for each allele to avoid amplification bias. These techniques are considered qualitative or semi-quantitative and typically applied to samples taken from individual plants or single seeds. Although it is possible to detect specific single-nucleotide polymorphisms (SNPs) in breeding populations using PCR-based methods or targeted sequencing, it is usually carried out on a single plant/seed basis and with the use of multiple markers. The SNPs used in breeding for example in varietal identification are chosen due to the ease of assay design and validation. In the case of introduced small changes, the SNP and resultant assay cannot be freely chosen—it is the change that led to the characteristic expressed in the desired end product, and thus the design of the PCR-based detection method is constrained to this specific sequence.

Real-time PCR may be applied in a quantitative manner. A large insertion (or deletion) can be easily and sensitively quantified as is the practice for genetically modified organisms. PCR is also used to enumerate SNP allele frequencies (Germer *et al.* 2000). In certain polyploid crops such as canola and wheat, there may be multiple sub-genomes each containing a copy of the target. More than one series of PCR reactions can be applied in a nested fashion to increase specificity. The first series is dedicated to amplification of the relevant genome/s or genic region of interest. Subsequently, the product from the first round is used as the template to amplify smaller and more specific region of interest in a secondary PCR. This nested PCR approach is however more prone to false positive results (Wanger *et al.* 2017). If the assay is robust and reliable enough, a small edit can be quantified (as % edited content) at ratios of 1 in 1000 or even 1 in 3000. Advanced techniques and incorporation of additional chemical modifications of primers and probes (e.g., peptide nucleic

acids) can afford greater specificity and thereby enhance sensitivity (Zhou *et al.* 2018b). It is critical to use an appropriate reference and control material (e.g., known origin and sequence composition) for comparison. Relative quantification can be deduced with the use of a standard curve from previously characterized edited reference samples.

Digital PCR

Like real-time PCR, digital PCR (dPCR) technology utilizes polymerase, primers, and Taqman probes within a standard end-point PCR reaction to amplify specific targets. dPCR works, however, by partitioning nucleic acid samples into thousands of single parallel PCR reactions, each separated into small volume compartments using droplets or chambers. PCR amplification occurs simultaneously in each partition. At the end of the run, each droplet or chamber is individually assessed for the presence (positive) or absence (negative) of a specific fluorescent signal. Using Poisson statistical analysis, the ratio of positive to negative partitions yields absolute quantification of the initial number of copies of the target sequence. As such, dPCR allows for absolute quantification of a target without a reference or the need for running standard curves.

Other benefits of the dPCR technology include the high level of precision and sensitivity. By removing the amplification efficiency reliance of qPCR, error rates are strongly reduced, enabling reliable and accurate measurement of small target concentration differences among samples. The massive sample partitioning also results in an increased signal-to-noise ratio since high-copy templates and PCR inhibitors are strongly diluted, effectively enriching template concentration in target-positive partitions. Moreover, dPCR assays can be readily adapted to any target of interest and, given that reaction volumes are in the pico- to nanoliter ranges, can obtain absolute quantifications from very small amounts of DNA. Together, these features make dPCR ideally suited for the rapid and systematic quantification of genome editing outcomes at endogenous loci. A well-designed dPCR can be used to predict the number of targeted loci at the single plant level.

Over the past few years, dPCR has been used to reliably measure gene editing frequencies across a wide range of organisms (Findley *et al.* 2016; Miyaoka *et al.* 2016; Mock *et al.* 2016; Falabella *et al.* 2017). While dPCR applications in genome-edited plant cells are still limited, the methodology clearly provides novel opportunities to quantitate small genome edits or low copy number targets in bulk populations of edited cells. Recently, Jouanin *et al.* (2020) used dPCR mutation assays to detect indels (1 to 50 bp) and large deletions (>300 bp) in wheat and concluded that dPCR is suitable for high-throughput screening of copy number variation and gene editing-induced mutations in large gene families. Consonant with this are the findings of Peng *et al.* (2020)

who developed a duplexed dPCR-based method for the detection and evaluation of gene editing frequencies in rice and allotetraploid rapeseed. The authors not only showed that their dPCR-based method is sensitive to different kinds of gene editing mutations but also demonstrated its applicability to polyploid plants and processed food samples containing low initial concentrations of DNA. Moreover, compared with qPCR and NGS-based methods, the duplexed dPCR assay can yield a lower limit of detection (LOD) and was able to decipher homozygous from heterozygous mutations with superior levels of precision and sensitivity. Owing to its ease of use, reduced complexity, repeatability, and superior precision, dPCR will continue to increase in popularity (Miyaoaka *et al.* 2016) and contribute to rapid and quantitative plant genome editing workflows.

Sequencing

Rapid evolution in next (short read)- and third (long read)-generation sequencing technologies has led to sequencing emerging as a powerful tool for the detection of DNA changes. The latest advancements in methods for target enrichment, library preparation, and tools for bioinformatics analysis resulted in increased accuracy of variant detection, higher throughput, and faster turnaround times (Salk *et al.* 2018).

Applications of next-generation sequencing to detect sequence variants cover whole-genome sequencing, whole-exome sequencing, and targeted sequencing. Targeted sequencing focuses on the set of genes or targets of interest allowing higher read coverage with reduced cost and dataset size. Furthermore, a larger number of plants can be pooled in one sequencing run without affecting coverage depth.

Targeted sequencing consists of two major approaches, either PCR-based target amplification (such as amplicon-based sequencing) or target capture using biotinylated hybridization probes (such as hybrid capture sequencing; Bewicke-Copley *et al.* 2019). Amplicon-based sequencing is the most efficient of the two methods, with usually smaller target regions and higher percentage of on-target reads, while hybrid capture sequencing results in more uniform coverage of typically larger target regions. Both are powerful approaches for accurate detection of genome edits. Multiple target regions can be assessed across many samples in parallel.

Generally, using a sequencing approach for detection of genome edits have higher cost compared to PCR-based approaches. Therefore, a targeted sequencing approach is typically used when detection via PCR cannot be developed due to technical limitations and/or in cases of high throughput analysis. During early stages of development, targeted sequencing can be applied to characterize tissue cultures or plants for which the outcome of the editing is unknown or needs to be confirmed. Traditional amplicon-based

sequencing using two locus-specific primers is suitable for the characterization of single base pair variants or smaller indels. However, its application in the case of larger indels and translocations may be restricted. Newer single primer technologies based on ligation-mediated PCR have alleviated those limitations (Zheng *et al.* 2014). Since amplicons up to approx. 600 bp can be targeted for amplicon-based sequencing using short read technology, and even larger target regions in the case of hybrid capture, target enrichment approaches can be applied when primer design for PCR or dPCR methods is complicated. Reasons for complicated primer design with respect to PCR or dPCR are related to the upper limit of the amplicon size and the nature of the sequences flanking the edit. Unfavorable G-C content, repetitive sequence, or presence of polymorphisms not related to the trait in certain haplotypes may complicate the development of PCR-based detection methods. For genome edits targeted in highly repetitive regions (edit in conserved gene family, in polyploid species), long fragments may be captured and sequenced using long-read sequencing platforms to allow the level of specificity required.

Given the high read coverage across the target regions, targeted deep sequencing is more sensitive for detecting and quantifying low-frequency variants in heterogeneous samples in contrast to Sanger sequencing. The read coverage required is defined by the detection limit to be achieved. However, the lower limit of detection with respect to the sequencing technology is defined by the error rate, of which the absolute number increases along with the higher read depth. Errors accumulate throughout the different steps of the targeted sequencing protocols including damage during DNA sample preparation and introduction of errors during PCR amplification and sequencing. To differentiate low-frequency genome edits from technical artifacts, caution is needed when the variant allele frequency is near or below the limit of detection related to the targeted sequencing protocols, especially when sequence data is generated from low-quality DNA. The sensitivity of routine NGS approaches for the detection of low frequency variants is estimated to be approx. 1% (Salk *et al.* 2018).

Recent developments to increase the accuracy of next-generation sequencing protocols include computational and statistical measures to reduce the background error rate, adjustments to library preparation protocols to maintain the intactness of the DNA templates, and most effectively, single-molecule consensus sequencing. Unique molecular identifiers are attached to the target DNA fragments and remain attached throughout enrichment and sequencing. PCR duplicate reads are removed with higher accuracy which eliminates biases from variable PCR amplification and results in more precise quantification of DNA templates. In addition, sequencing errors are corrected based on the majority vote within a pool of reads originating from the same DNA fragment (Jabara *et al.* 2011; Kinde *et al.* 2011; Hong and Gresham 2017; Xu *et al.*

2017). Using duplex unique molecular identifiers, true variants present on both strands of the DNA fragment can be further differentiated from false positive variants only present on one of the two strands (Schmitt *et al.* 2012; Kennedy *et al.* 2014). As the result of accurate removal of PCR duplicates and precise error correction, target-enriched sequencing using duplex unique molecular identifiers is suitable for confident detection of small sequence edits with a frequency of approximately 0.1 to 0.2% in genetically heterogeneous samples (Peng *et al.* 2019). Base pair errors which remain unresolved with the latest targeted sequencing protocols define the limit of sensitivity of sequencing-based detection methods. Unresolved errors are typically incorporated before or during the attachment of dual unique molecular identifiers. It is recognized that the level of accuracy of sequencing is affected by different factors, including the sequencing platform, chemistry version, and sequence context, as well as by experimental variation such as the degree of DNA damage.

Digital PCR is a robust technology for detection and quantification of genome edits which are typically known *a priori* and present at low frequency in heterogeneous samples. Since the recent implementation of single-molecule consensus sequencing, next-generation sequencing may achieve an accuracy comparable to dPCR and can be applied when the resulting genome edit is unknown. Single-molecule consensus sequencing-based methods for the quantification or detection of genome-edited material present at low frequency in heterogeneous samples such as *in vitro* cultures, grain, or food have not been reported. However, the approach has been applied for reliable quantification and detection of rare variants in diverse clinical fields including cancer, aging, and metagenomics (Salk *et al.* 2018).

Despite the effective implementation of single-molecule consensus sequencing to increase the accuracy of sequencing technology, sequencing-based detection of small genome edits with low frequency requires high read coverage. The consequence is a higher cost compared to qPCR or dPCR methods, unless high throughput analysis is required (Aloisio *et al.* 2016).

Isothermal DNA detection and CRISPR-Cas-mediated edit detection

Isothermal amplification of nucleic acids is beginning to provide a rapid, sensitive, and specific diagnostic to replace the more time-consuming traditional PCR amplification methods. Much like PCR, isothermal PCR uses enzymatic amplification to amplify a nucleic acid sequence with a polymerase, but isothermal nucleic acid amplification does not require variable temperature cycling. These methods are beginning to provide the sensitivity and specificity needed for detecting single nucleotide changes (Zhou *et al.* 2018a; Shen *et al.* 2020). One

recent innovation is to couple the amplification power of isothermal DNA polymerases with CRISPR-Cas specificity (Kellner *et al.* 2019).

The CRISPR-Cas system contains programmable endonucleases. In 2017, it was reported that CRISPR-Cas editing components could also be an effective diagnostic tool for detection of specific nucleic acid changes. For example, CAS13a has a crRNA-programmed collateral activity that can detect a specific RNA target sequence. This collateral activity has been used in a Specific High sensitivity Enzymatic Reporter unLOCKing detection method named SHERLOCK (Gootenberg *et al.* 2017).

Several other platforms have been developed using this collateral cleavage activity. Such platforms are typically employed via an isothermal nucleic acid amplification step using Recombinase Polymerase Amplification (RPA), Loop-mediated isothermal AMPLification (LAMP), or Helicase Dependent Amplification (HAD). CRISPR-Cas13, Cas12a, and Csm6 (Gootenberg *et al.* 2018) as well as Cas9 (Wang *et al.* 2019) systems have been demonstrated this collateral activity. In the SHERLOCK assay, a pre-amplification of the target DNA or RNA is needed (Kellner *et al.* 2019) in addition.

Such platforms can use different indicators including a fluorescently labeled reporter, visual detection using liquid-liquid phase separation or lateral flow detection by using antigen-labeled reporters (Kellner *et al.* 2019). CRISPR-Cas13 cleaves RNA but CRISPR-Cas12 cleaves DNA, so application of the appropriate enzyme can be used for either RNA or DNA detection (Gao *et al.* 2021). Cas13 is ultrasensitive in detecting mutations but if the target is present in very low numbers (600K molecules), it cannot be used to achieve single nucleotide detection; a pre-amplification would be necessary (Kellner *et al.* 2019). Cas12a from *Acidaminococcus* sp. BV3L6 (AsCas12a) may not produce any signals from targets at lower concentrations and therefore pre-amplification of the target may be necessary.

This limitation will apply unless a very sensitive Cas system can be found. However, it has been shown that by choosing an optimal crRNA that yields a higher fluorescence level, and/or combining multiple crRNAs in one reaction, less input target is needed (Gao *et al.* 2021). Gootenberg *et al.* (2018) has combined Csm6 with Cas13 detection to increase the sensitivity and enhance the fluorescent signal. This was done to eliminate the need for pre-amplification using RPA for lateral flow detection methods.

The advantage of using a very sensitive detection method would be to eliminate the need for amplifications done through PCR or detection through sophisticated fluorophore detection instruments. Therefore, adding the isothermal pre-amplification would be a drawback to the CRISPR-Cas detection methods and limit its portability in low-resource settings (Kellner *et al.* 2019). However, the isothermal

amplification can be accomplished using handheld devices or smartphone-based detection platforms without a need for any complex instrumentation (Song *et al.* 2018; Tsaloglou *et al.* 2018). This makes the CRISPR-Cas-mediated detection method very attractive for detection single base mutations (SDN1) even considering the pre-amplification step while aiming for a method that can be used in a non-laboratory setting.

The CRISPR-Cas13a and 13b family have also been used to investigate the possibility of creating a multiplexed platform. Gootenberg *et al.* (2018) found that the activities of LwaCas13a, Cas13b from *Capnocytophaga canimorsus* Cc5 (CcaCas13b), LbaCas13a, and PsmCas13b can be combined independently and measured with the four dinucleotide reporters AU, UC, AC, and GA, respectively. Using these cleavage specificities, they could detect Zika virus using HEX, dengue virus using FAM. Later using multiplexed SHERLOCK with PsmCas13b and LwaCas13a, they could detect ZIKV and DENV RNA dilutions as well as allele-specific genotyping of human saliva samples in one reaction. This advancement allows for multiple target detection at scale and for cheaper cost.

CRISPR-Cas biosensing methods have the advantage of detecting single base variations (SDN1) occurring in femtomolar or even attomolar concentrations. The simplicity of this method and the instrument free methods of detection like lateral flows made this method very helpful in any field DNA- or RNA-based diagnostics. However, some limitation and challenges need to be resolved before making this method the mutation detection method of choice. Cas9 and Cas12 require protospacer adjacent motif (PAM) sequences adjacent to target dsDNA to be able to cleave it. The protospacer flanking site (PFS) in Cas13a creates a similar limitation since the first base following the protospacer should be a non-G base. Making CRISPR-Cas detection methods be quantitative is difficult unless they are combined with quantitative PCR or digital PCR methods. This makes them less attractive (Li *et al.* 2019). Since CRISPR-Cas systems are very new, most of these limitations may be resolved in near future.

In summary, the various tools each offer advantages and disadvantages relative to one another. Consideration of cost and the need or not for sequence information as part of the tool development are critical components. Throughput and sensitivity are tied to the stage of the edit process and therefore will have a lot of variation depending on the context. An estimate of the present situation is shown in Table 2. The cost, throughput, and sensitivity will change over time.

Specific applications of genome edit detection methods

If a population of edits are being produced as part of a research program, then each edit will need to be characterized individually at the molecular level to compare it to the desired or expected phenotype. Detection tool(s) suitable for high-throughput analyses similar to those used for varietal identification are needed for this process.

Detection tools can be used to estimate the efficiency of editing methods, to subsequently identify edits in regenerated plants and to identify off-target edits during the development and breeding process. Tools for detection of genome edits in a research context cover a wide range of platforms including PCR, dPCR, and sequencing approaches. During the early stages of development, the challenge of detection is based on the edit frequency within a given cell population.

The frequency at which edits occur includes a lot of variables, for example the edit template sequence design, the genetic location of the target, the delivery technique, and the biology of the organism of interest. Once the edit is recovered at the whole plant level, these applications only require the detection of the characteristic sequence difference in a single organism, so sensitivity is not a challenge. Detailed characterization of the intended change may also be important for and/or used in the development of intellectual property rights around the improved variety. Depending on the stage of product development, and type of edit, different technologies might be the appropriate detection tool (Table 3).

Table 2 Characterization of detection tools.

Technology	Cost	Throughput	Sensitivity	<i>A priori</i> knowledge of edit	Surrounding sequence (independency)
PCR (gel-based)	+	++	+	++	++
qPCR	+	++ to +++	+ to +++	+++	+++
ddPCR	++ to +++	++	+++	+++	+++
Hybrid capture sequencing	+++	+++	+ to ++	+ to ++	++
Amplicon-based sequencing	+++	+++	++ to +++	+ to ++	++
Isothermal PCR	+	++	+ to ++	+++	+++

Low (+) to high (+++) relative to the technologies included in this table

Table 3 Detection tools applicable to the genome editing process.

Stage	Purpose	GE type	Technologies
Set-up GE methods	Demonstrate and/or optimize GE activity	Indel	dPCR, sequencing
		Base edit	dPCR, sequencing
		HDR	Edit-specific PCR, sequencing
Select edits	Pre-select experiment with highest GE activity		dPCR, PCR
	Select plants with edited allele(s)	Indel	PCR fragment size/cleavage
		Base edit	dPCR
		HDR	Edit-specific PCR
Characterize edits	Confirm sequence of edited and unedited allele(s)	All	Sequencing
Progeny of edited plants	Confirm inheritance and absence/presence of edited alleles	All	PCR, qPCR, sequencing
Seed production	Demonstrate presence/absence of edits in population or seed lot	All	qPCR, dPCR, sequencing

Optimization of genome editing methods

The detection tools used to evaluate the performance of genome editing methods should be able to allow high-throughput assessment of editing efficiencies in large numbers of samples generated (e.g., from protoplasts, *in vitro* cultures). In the initial phase of method development, they also should be able to detect very low levels (<1%) of editing.

Large-scale screening of genome-edited events is essential to target the right edit before further regeneration or evaluation of the *in vivo* activity of the genome editing reagents (Nadakuduti *et al.* 2019). Targeted sequencing can detect SDN1 as well as SDN2 and SDN3 edits. The primary consideration for using targeted sequencing for detection of rare edits in tissue culture is the depth of sequencing to cover the necessary statistical power of detection of low frequency edits (Délye *et al.* 2015). Increasing the sequencing depth will increase the statistical power of detection of rare mutations, but also increases the cost of sequencing (Liu *et al.* 2014). Therefore, this approach may not be the best option when the sequence of the edits is known, in which case PCR would be the best tool for detecting rare event detection.

Detection of rare edits with large modification types (e.g., inserts or re-arrangements) can be done via PCR-based methods, with appropriate targeting to generate a PCR product only when the desired edit is present. In case of a sequence insertion by HDR using a repair DNA, this can be achieved using one primer located outside of the edited region and one primer within the insert. In case of low levels of editing, nested PCR is often needed to detect the edits. Special care should be taken to minimize the formation of PCR artifacts, especially recombinant products resulting from template switching, by limiting the number of PCR cycles, lowering the amount of input DNA, and using an appropriate polymerase. This is especially important when working in crops with complex genomes that typically have several homologous copies of

the target gene or when extremely high levels of (viral) repair DNA are being used.

For detection of smaller modifications (small indels, few base changes), dPCR is a convenient method that provides excellent specificity and sensitivity. Once a good dPCR assay is validated, it typically will allow for the detection and quantification of small changes below the 1% level.

Even when taking care to minimize PCR-based artifacts, final proof that the desired editing has happened requires sequencing of the edited allele. Amplicon-based sequencing may be method of choice where the exact edit is unknown or not well defined. This may be the case where the edits are imprecise, such as insertions or deletions based on NHEJ. PCR can offer a fast and efficient initial screen to identify edited lines. Then amplicon-based sequencing is used to characterize the edit sequence.

Selection of edited products

Depending on the transformation and tissue culture system used for genome editing, selection of edited products can happen at various stages, such as *in vitro* cultures or tissue explants. Analytical throughput needed to detect a sufficient number of edits depends on the tissue culture system, the stage of testing, the efficiency of the editing, and whether the edit can be selected directly or selection for a co-introduced selectable marker gene is applied. Genome edits may be introduced in systems where the edit results in ability to select the results directly. An example would be where herbicide tolerance is being introduced. In these cases, the functional edited cells and resulting plants can be identified relatively easily as the cells and/or plants will be resistant to herbicide and can be selected based on phenotype.

Genome editing is often desired in cases where the phenotype is not directly selectable for example in tissue culture and

can only be identified by the genotype. In these cases, a population of cells or large number of regenerated plants must be screened. Such situations mean that a more powerful and efficient approach is necessary and many of the same detection tools as used for optimization can be applied.

In cases where no direct selection via a co-introduced selectable marker gene is applied and/or with very low editing efficiencies, combining detection techniques could be applied to pre-screen large numbers of tissues or explants to identify rare edits (using PCR) or pools of plants could be analyzed using dPCR to identify those pools that contain the highest editing levels.

With higher editing efficiencies, screening of individual plants for the desired edits by PCR is a viable option. When creating specific edits, this can be done via an edit-specific PCR assay. When creating indels, individual plants having the desired changes can be identified by characterizing PCR products covering the target site for changes in fragment length, susceptibility to a restriction enzyme that has a recognition site at the target, by loss of susceptibility to cleavage *in vitro* by a CRISPR RNP that has the same crRNA as used for the editing. SDNs, such as Cas12a, that typically create larger deletions have the advantage that a larger fraction of the indel alleles can be detected based on fragment length.

Final proof that selected plants are edited requires targeted sequencing of the edited allele. This can be done either via cloning and Sanger sequencing, or by NGS sequencing of PCR products covering the edited allele. The level of edits detected by sequencing also gives an indication on whether the edit is chimeric or present in most of the plant. Inheritance of the edited allele to the next generation provides confirmation of the edit. Additional edits may be generated if functional SDN and gRNA are still present, and plants should be analyzed in these cases for the presence of newly formed edits.

Breeding

Although application of molecular analysis tools has increased the speed of breeding, it is still reliant upon variation that is generated via spontaneous or induced mutations (Lyzenga *et al.* 2021). Genome editing can be applied to breeding to target and produce desirable alleles. Once these changes have been introduced and regenerated into plants, there is a need for detection methods to follow the desired change through the breeding process to generate commercial varieties. A genome edit can be treated in the breeding process in much the same way as any small or large change or SNP with the caveat that it may be more or less easy to detect than a conventional SNP as explained in the “Tools for detection of genome edits” section. Such edits can usually be routinely followed using PCR analysis of large numbers of single

plants, and sequencing may be used at critical control points to confirm the identity and integrity of the edit.

Detection of gene edits in bulk seed and grain

Some jurisdictions (Dederer and Hamburger 2020) are requiring that gene-edited products be subject to a regulatory process that includes molecular characterization of edits and a way to positively identify the edits. Such detection methods typically require that one seed or grain be detectable in between 1000 and 3000 seeds. Achieving this level of sensitivity for a small sequence difference in a bulk seed or grain sample is difficult based on currently available tools. There is anecdotal information that suggests that single base changes can be detected in bulk seed or grain at single percent or tenth percent levels; there is however no published data available to date.

PCR methods to differentiate single base pair or small edits from the background sequence are highly dependent on the surrounding DNA sequence, and the specific base mismatch concerned. Therefore, no general statement can be made as to whether such methods could be developed for specific use cases (Herrera *et al.* 2021 *in press*).

Whereas marker-assisted selection routinely employs SNP analysis for use in breeding, the selected polymorphisms are typically those easiest to assay in single plants or seeds. Where the intent is to identify specific single base pair edits, there is no freedom to select the best SNP markers, which limits assay flexibility. While that is not a big challenge to using PCR detection in breeding, detection of a nucleotide single (or small) base pair change in thousands of similar copies (such as when analyzing bulk samples of 1000 to 3000 seeds) can be a challenge.

When using PCR, the reaction rate for differentiation based on mismatch at the 3' end of the primer will be dependent on the specific base change (Rejali *et al.* 2018). Various tactics can be employed to improve differentiation of the target sequence difference. These include use of PNA clamps or LNA technology (Karkare and Bhatnagar 2006), chemistries such as BHQ®-Plus³ or Taqman™ MGB probes, and altered cycling conditions such as higher annealing temperatures (Rejali *et al.* 2018).

An alternative approach can be amplicon-based sequencing. This approach requires *a priori* knowledge of the sequence around the area of the edit. Primers can be designed to amplify the edited region, and the resulting products sequenced. The amount of the edited product in the bulk can be estimated from the ratio of the sequences present.

³ BHQ is a registered trademark of Biosearch Technologies, Inc. Petaluma, California

Operational factors

Detection of off-target edits and editing components

Commercial crop production relies on several generations of backcrossing and other traditional breeding steps to attain the final varieties yielding the favorable contributions of desirable phenotypes from both parents. The introduction of gene editing is a new technique for plant breeding with the benefit that it can be targeted to specific gene(s). There has however been a great deal of discussion of the possibility of off-target editing in plants.

The specificity of CRISPR systems seems to be of minimal concern in plants as very few off-target mutations are detected (Tang *et al.* 2018; Young *et al.* 2019; Graham *et al.* 2020) and potential off-target mutations can be removed through backcrossing in plants. Any potential off-target sites can be largely avoided by designing gRNAs with high specificity.

Off-target editing, or RNA-guided endonuclease (RGEN)-induced mutations, can occur at sites with sequence similarity to on-target sequence. The nuclease that causes the double-stranded break at a specific target can potentially cause unintended double-stranded breaks at locations with significant homology in the genome. Off-target edits have been observed in clinical and therapeutic applications (Zhang *et al.* 2015b) sometimes resulting in gene function disruptions or genomic instability. These issues can create some concern for utilization of gene editing in clinical therapies as unintended mutations can pose a health risk for unregulated somatic cell proliferation. Plants, however, seem to be very resilient to somatic cell mutations (Graham *et al.* 2020). Off-target edits may be observed phenotypically in some cases as “off-types” or can be detected by molecular techniques and removed during the breeding process.

While gene editing processes are designed to minimize off-target changes, off-target mutations may be attributed to five factors (Modrzejewski *et al.* 2020): (1) number of mismatches, (2) position of mismatches, (3) G-C content of the targeted sequence, (4) altered nuclease variants, and (5) delivery method(s). Today, most major crops have comprehensive reference genomes available. Design of gene editing components and selection of unique target sites using the reference genome sequence and bioinformatic algorithms prior to the application of gene editing can help to minimize the possibility of the occurrence of off-target events (Young *et al.* 2019).

Detection of off-target edits in plants can be challenging because the number and position of off-target mutations cannot be fully predicted. Bioinformatic tools can be used to predict where edits may occur. Detection methods that can be used, provided that the specific sites are known, vary in cost, complexity, equipment, and limitations, but are the same as for on-target editing. Examples are sequencing, PCR, loss of primer binding site, mismatch cleavage assay, high-

resolution melt curve analysis, and modified sequencing approaches (Zischewski *et al.* 2017; Blondal *et al.* 2021). Approaches for identification of off-target edits continue to develop.

It has been asserted that the use of gene editing techniques may result in a genomic disruption (so-called scars) at the target cut site (Elison and Acar 2018). When relevant, such products can generally be avoided through the use of non-integrating editing components. For example, the use of a 2-step CRISPR approach avoids unintended changes by temporally separating the guide RNA from the desired product in a stepwise approach, preventing the final edit to be targeted by Cas9. In cases where the DNA coding for gene editing components has been introduced either transiently, or as an integrated gene construct, the lack of such sequences in the product must be identified. If the genes coding for the genome editing components, e.g., the site-directed nucleases, are stably integrated into the genome of the recipient, the initially regenerated plant will contain this foreign DNA. As the sequence of the components is known, PCR can be used for detection. If transient expression is used, e.g., through the use of TALEN proteins (Grohmann *et al.* 2019), no external DNA is used or expected to be in the plant, and in addition, PCR can be used to verify absence.

While off-target events are considered unacceptable in clinical therapeutic applications of genome editing (Lee *et al.* 2016), off-target events in the context of plant breeding are comparable in nature but less numerous than the variation that is introduced by conventional or mutation breeding. In addition, commercial varieties undergo continuous selection for the best phenotypes during breeding and commercial variety development. This selection process also applies to variety development that includes gene edits and can be used to remove any off-types.

Differentiation of genome edits from conventional mutations

While single base and other small changes can be made intentionally, such small changes are also constantly occurring in plant populations (Grohmann *et al.* 2019) and almost any non-deleterious mutation may be found in a commercial crop field (Sainsbury 2021). This is important from a regulatory point of view because in countries where GMO regulations are applied to gene-edited plants, the same changes selected as variants in plant breeding populations or induced by mutagenesis may not generally be regulated.

A challenge of applying detection methods to these products is that no currently available analytical tool(s) are able to determine whether a mutation occurred spontaneously in the population, or due to application of radiation or mutagens. Thus, there will always be a question as to the background level of “detection” of such changes that are not induced by

laboratory methods. Grohmann *et al.* (2019) suggest that the minimum length of a random sequence for uniqueness in plants is between 14 and 17 base pairs, depending on the genome size. As the DNA sequence may be the same whether the changes are directed or undirected, the safety of such a product will be the same.

Conclusions

Many different types of detection tools are used in the field of genome editing. Most applications focus is on detection of DNA changes, specifically PCR methods and sequencing. The preferred detection tool depends on the specific context in which it will be used—whether to optimize the process, sort through products of genome editing, or following the edited products through breeding and product development and onward to the farmer and food, feed, and fiber. Detection tools will continue to evolve as this technology develops.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alarcon CM, Shan G, Layton DT, Bell TA, Whipkey S, Shillito RD (2019) Application of DNA- and protein-based detection methods in agricultural biotechnology. *J Agric Food Chem* 67:1019–1028. <https://doi.org/10.1021/acs.jafc.8b05157>
- Aloisio M, Licastro D, Caenazzo L, Torboli V, D'Eustacchio A, Severini GM, Athanasakis E (2016) A technical application of quantitative next generation sequencing for chimerism evaluation. *Mol Med Rep* 14: 2967–2974. <https://doi.org/10.3892/mmr.2016.5593>
- Altpeter F, Springer NM, Bartley LE, Blechl AE, Brutnell TP, Citovsky V, Conrad LJ, Gelvin SB, Jackson DP, Kausch AP, Lemaux PG, Medford JI, Orozco-Cárdenas ML, Tricoli DM, VanEck J, Voytas DF, Walbot V, Wang K, Zhang ZJ, Stewart CN (2016) Advancing crop transformation in the era of genome editing. *Plant Cell* 28: 1510–1520
- Bailey-Serres J, Parker JE, Ainsworth EA, Oldroyd GED, Schroede JI (2019) Genetic strategies for improving crop yields. *Nature* 575: 109–118. <https://doi.org/10.1038/s41586-019-1679-0>
- Bashir T, Chandra Mishra R, Hasan MM, Mohanta TK, Bae H (2018) Effect of hybridization on somatic mutations and genomic rearrangements in plants. *Int J Mol Sci* 19:3758. <https://doi.org/10.3390/ijms19123758>
- Bennetzen JL (2000) Transposable element contributions to plant gene and genome evolution. *Plant Mol Biol* 42:251–269. <https://doi.org/10.1023/A:1006344508454>
- Bewicke-Copley F, Arjun Kumar E, Palladino G, Korfi K, Wang J (2019) Applications and analysis of targeted genomic sequencing in cancer studies. *Computat Struct Biotechnol J* 17:1348–1359
- Blondal T, Gamba C, Jagd LM, Su L, Demirov D, Guo S, Johnston CM, Riising EM, Wu X, Mikkelsen MJ, Szabova L, Mouritzen P (2021) Verification of CRISPR editing and finding transgenic inserts by Xdrop™ indirect sequence capture followed by short- and long-read sequencing. *bioRxiv preprint* <https://doi.org/10.1101/2020.05.28.105718>. Accessed April 15, 2021
- Bogdanove A, Voytas D (2011) TAL effectors: customizable proteins for DNA targeting. *Science* 333:1843–1846
- Čermák T, Curtin SJ, Gil-Humanes J, Čegan R, Kono TJY, Konečná E, Belanto JJ, Starker CG, Mathre JW, Greenstein RL, Voytas DF (2017) A multipurpose toolkit to enable advanced genome engineering in plants. *Plant Cell* 29:1196–1217. <https://doi.org/10.1105/tpc.16.00922>
- Chen H, Lin Y (2013) Promise and issues of genetically modified crops. *Curr Opin Plant Biol* 16:255–260. <https://doi.org/10.1016/j.pbi.2013.03.007>
- Chen K, Wang Y, Zhang R, Zhang H, Gao C (2019) CRISPR/Cas genome editing and precision plant breeding in agriculture. *Annu Rev Plant Biol* 70:667–697
- Davalieva K, Kiprijanovska S, Plaseska-Karanfilska D (2014) Fast, reliable and low cost user-developed protocol for detection, quantification and genotyping of hepatitis C virus. *J Virol Meth* 196:104–112. <https://doi.org/10.1016/j.jviromet.2013.11.002>
- de Maagd RA, van de Wiel C, Schouten HJ (2020) The plasticity of plant genomes. Causes and consequences: a survey of data on structural genome variation in plants. COGEM report 2020-04 The Netherlands Commission on Genetic Modification, Bilthoven, the Netherlands. https://cogem.net/app/uploads/2020/07/CGM-2020-04-COGEM-report_StructuralVariation.pdf. Accessed March 10, 2020
- Debernardi JM, Tricoli DM, Ercoli MF, Hayta S, Ronald P, Palatnik JF, Dubcovsky JA (2020) GRF-GIF chimeric protein improves the regeneration efficiency of transgenic plants. *Nat Biotechnol* 38:1274–1279. <https://doi.org/10.1038/s41587-020-0703-0>
- Dederer H-G, Hamburger D (2020) Regulation of genome editing in plant biotechnology. A comparative analysis of regulatory frameworks of selected countries and the EU. Springer Nature Switzerland AG., Cham, Switzerland. ISBN 978-3-030-17118-6 ISBN 978-3-030-17119-3 (eBook). <https://doi.org/10.1007/978-3-030-17119-3>
- Délye C, Causse R, Gautier V, Poncet C, Michel S (2015) Using next-generation sequencing to detect mutations endowing resistance to pesticides: application to acetolactate-synthase (ALS)-based resistance in barnyard grass, a polyploid grass weed. *Pest Manag Sci* 71:675–685. <https://doi.org/10.1002/ps.3818>
- Elison GL, Acar M (2018) Scarless genome editing: progress towards understanding genotype–phenotype relationships. *Curr Genet* 64: 1229–1238. <https://doi.org/10.1007/s00294-018-0850-8>
- Epinat J-C, Arnould S, Chames P, Rochaix P, Desfontaines D, Puzin C, Patin A, Zanghellini A, Pâques F, Lacroix E (2003) A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells. *Nucleic Acids Res* 31:2952–2962
- Falabella M, Sun L, Barr J, Pena AZ, Kershaw EE, Gingras S, Goncharova EA, Kaufman BA (2017) Single-step qPCR and dPCR detection of diverse CRISPR-Cas9 gene editing events. *In Vivo G3 (Bethesda)* 7:3533–3542. <https://doi.org/10.1534/g3.117.300123>
- Findley SD, Vincent KM, Berman JR, Postovit LM (2016) A digital PCR-based method for efficient and highly specific screening of genome edited cells. *PLoS ONE* 11:e0153901. <https://doi.org/10.1371/journal.pone.0153901>

- Gao D, Zhu X, Lu B (2021) Development and application of sensitive, specific and rapid CRISPR-Cas13-based diagnosis. *J Med Virol* 93: 4198–4204. <https://doi.org/10.1002/jmv.26889>
- Germer S, Holland M, Higuchi R (2000) High-throughput SNP allele-frequency determination in pooled DNA samples by kinetic PCR. *Genome Res* 10:258–266
- Gootenberg JS, Abudayyeh OO, Kellner MJ, Joung J, Collins JJ, Zhang F (2018) Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a, and Csm6. *Science* 360:439–444. <https://doi.org/10.1126/science.aag0179>
- Gootenberg JS, Abudayyeh OO, Lee JW, Essletzbichler P, Dy AJ, Joung J, Verdine V, Donghia N, Daringer NM, Freije CA, Myhrvold C, Bhattacharyya RP, Livny J, Regev A, Koonin EV, Hung DT, Sabeti PC, Collins JJ, Zhang F (2017) Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science* 356:438–442. <https://doi.org/10.1126/science.aam9321>
- Grubunova V, Levy AA (1997) Non-homologous DNA end joining in plant cells is associated with deletions and filler DNA insertions. *Nucleic Acids Res* 25:4650–4657. <https://doi.org/10.1093/nar/25.22.4650>
- Gordon-Kamm B, Sardesai N, Arling M, Lowe K, Hoerster G, Betts S, Jones T (2019) Using morphogenic genes to improve recovery and regeneration of transgenic plants. *Plants* 8:38. <https://doi.org/10.3390/plants8020038>
- Graham N, Patil GB, Bubeck DM, Doberst RC, Glenn KC, Gutsche AT, Kumar S, Lindbo JA, Maas L, May GD, Vega-Sanchez ME, Stupar RM, Morrell PL (2020) Plant genome editing and the relevance of off-target changes. *Plant Physiol* 183:1453–1471. <https://doi.org/10.1104/pp.19.01194>
- Grohmann L, Keilwagen J, Duensing N, Dagand E, Hartung F, Wilhelm R, Bendiek J, Sprink T (2019) Detection and identification of genome editing in plants: challenges and opportunities. *Front Plant Sci* 10:236. <https://doi.org/10.3389/fpls.2019.00236>
- Hamada H, Liu Y, Nagira Y, Miki R, Taoka N, Imai R (2018) Biolistic-delivery-based transient CRISPR/Cas9 expression enables in planta genome editing in wheat. *Sci Rep* 8:14422. <https://doi.org/10.1038/s41598-018-32714-6>
- Herrera S, Whitt S, Ghavami F, Ross M, Houchins D, Shillito R (2021) Chapter 14: detection methods for genome edited crops. In: *Sampling and Detection for Agricultural Biotechnology Products*, Eds Shillito R, Shan G, Elsevier Amsterdam (in press)
- Hong J, Gresham D (2017) Incorporation of unique molecular identifiers in TruSeq adapters improves the accuracy of quantitative sequencing. *BioTechniq* 63:221–226
- Jabara CB, Jones CD, Roach J, Anderson JA, Swanson R (2011) Accurate sampling and deep sequencing of the HIV-1 protease gene using a primer ID. *Proc Natl Acad Sci U S A* 108:20166–20171
- Jo YD, Kim JB (2019) Review frequency and spectrum of radiation-induced mutations revealed by whole-genome sequencing analyses of plants. *Quant Beam Sci* 3:7. <https://doi.org/10.3390/qubs3020007>
- Johnson MP, Haupt LM, Griffiths LR (2004) Locked nucleic acid (LNA) single nucleotide polymorphism (SNP) genotype analysis and validation using real-time PCR. *Nucleic Acids Res* 32(6):e55. <https://doi.org/10.1093/nar/gnh046>
- Jouanin A, Tenorio-Berrio R, Schaart JG, Leigh F, Visser RG, Smulders MJ (2020) Optimisation of droplet digital PCR for determining copy number variation of α -gliadin genes in mutant and gene-edited polyploid bread wheat. *J Cereal Sci* 92:102903
- Karkare S, Bhatnagar D (2006) Promising nucleic acid analogs and mimics: characteristic features and applications of PNA, LNA, and morpholino. *Appl Microbiol Biotechnol* 71:575–586
- Kellner MJ, Koob JG, Gootenberg JS, Abudayyeh OO, Zhang F (2019) SHERLOCK: nucleic acid detection with CRISPR nucleases. *Nat Protoc* 14:2986–3012. <https://doi.org/10.1038/s41596-019-0210-2> Erratum in: *Nat Protoc* 15:1311
- Kennedy SR, Schmitt MW, Fox EJ, Kohn BF, Salk JJ, Ahn EH, Prindle MJ, Kuong KJ, Shen JC, Risques RA, Loeb LA (2014) Detecting ultralow-frequency mutations by Duplex Sequencing. *Nat Protoc* 9: 2586–2606. <https://doi.org/10.1038/nprot.2014.170> Erratum in: *Nat Protoc*. 9:2903
- Kinde I, Wu J, Papadopoulos N, Kinzler KW, Vogelstein B (2011) Detection and quantification of rare mutations with massively parallel sequencing. *Proc Natl Acad Sci U S A* 108:9530–9535
- Kong J, Martin-Ortigosa S, Finer J, Orchard N, Gunadi A, Batts LA, Thakare D, Rush B, Schmitz O, Stuijver M, Olhoft P, Pacheco-Villalobos D (2020) Overexpression of the transcription factor *GROWTH-REGULATING FACTOR5* improves transformation of dicot and monocot species. *Front Plant Sci* 11:572319. <https://doi.org/10.3389/fpls.2020.572319>
- Kutyavin IV, Afonina IA, Mills A, Gorn VV, Lukhtanov EA, Belousov ES, Singer MJ, Walburger DK, Lokhov SG, Gall AA, Dempcy R, Reed MW, Meyer RB, Hedgepeth J (2000) 3'-Minor groove binder-DNA probes increase sequence specificity at PCR extension temperatures. *Nucleic Acids Res* 28:655–661. <https://doi.org/10.1093/nar/28.2.655>
- Lee CM, Cradick TJ, Fine EJ, Bao G (2016) Nuclease target site selection for maximizing on-target activity and minimizing off-target effects in genome editing. *Mol Ther* 24:475–287
- Li Y, Li S, Wang J, Liu G (2019) CRISPR/Cas systems towards next-generation biosensing. *Trends Biotechnol* 37:730–743. <https://doi.org/10.1016/j.tibtech.2018.12.005>
- Liang Z, Chen K, Li T, Zhang Y, Wang Y, Zhao Q, Liu J, Zhang H, Liu C, Ran Y, Gao C (2017) Efficient DNA-free genome editing of bread wheat using CRISPR/Cas9 ribonucleoprotein complexes. *Nat Commun*. <https://doi.org/10.1038/ncomms14261>
- Lin C-S, Hsu C-T, Yang L-H, Lee L-Y, Fu J-Y, Cheng Q-W, Wu F-H, Hsiao HC-W, Zhang Y, Zhang R, Chang W-J, Yu C-T, Wang W, Liao L-J, Gelvin SB, Shih M-C (2018) Application of protoplast technology to CRISPR/Cas9 mutagenesis: from single-cell mutation detection to mutant plant regeneration. *Plant Biotechnol J* 16:1295–1310. <https://doi.org/10.1111/pbi.12870>
- Lipp M, Shillito R, Giroux R, Spiegelhalter F, Charlton S, Pinero D, Song P (2005) Polymerase chain reaction technology as an analytical tool in agricultural biotechnology. *J AOAC Intl* 88:136–155
- Liu Y, Luo W, Linghu Q, Abe F, Hisano H, Sato K, Kamiya Y, Kawaura K, Onishi K, Endo M, Toki S, Hamada H, Nagira Y, Taoka N, Imai R (2021) *In planta* genome editing in commercial wheat varieties. *Front Plant Sci* 12:648841. <https://doi.org/10.3389/fpls.2021.648841>
- Liu Y, Zhou J, White KP (2014) RNA-seq differential expression studies: more sequence or more replication? *Bioinformatics*. 30:301–304. <https://doi.org/10.1093/bioinformatics/btt688>
- Lomov NA, Viushkova NA, Petrenko AP, Syrkinina MS, Rubtsov MA (2019) Methods of evaluating the efficiency of CRISPR/Cas genome editing. *Molecular Biology*, 53:862–875. <https://doi.org/10.1134/S0026893319060116>
- Lowe K, Wu E, Wang N, Hoerster G, Hastings C, Cho MJ, Scelongo C, Lenderts B, Chamberlin M, Cushatt J, Wang L, Ryan L, Khan T, Chow-Yiu J, Hua W, Yu M, Banh J, Bao Z, Brink K, Igo E, Rudrappa B, Shamseer PM, Bruce W, Newman L, Shen B, Zheng P, Bidney D, Falco C, Register J, Zhao ZY, Xu D, Jones T, Gordon-Kamm W (2016) Morphogenic regulators *Baby boom* and *Wuschel* improve monocot transformation. *Plant Cell* 28:1998–2015. <https://doi.org/10.1105/tpc.16.00124>
- Lyzenga WJ, Pozniak CJ, Kagale S (2021) Advanced domestication: harnessing the precision of gene editing in crop breeding. *Plant Biotechnol J* 19:660–670. <https://doi.org/10.1111/pbi.13576>
- Ma X, Zhang X, Liu H, Li Z (2020) Highly efficient DNA-free plant genome editing using virally delivered CRISPR-Cas9. *Nat Plant* 6: 773–779

- Maertens O, Legius E, Speleman F, Messiaen L, Vandesompele J (2006) Real-time quantitative allele discrimination assay using 3' locked nucleic acid primers for detection of low-percentage mosaic mutations. *Anal Biochem* 359:144–146
- Maher MF, Nasti RA, Vollbrecht M, Starker CG, Clark MD, Voytas DF (2020) Plant gene editing through de novo induction of meristems. *Nat Biotechnol* 38:84–89. <https://doi.org/10.1038/s41587-019-0337-2>
- Miyaoka Y, Berman JR, Cooper SB, Mayerl SJ, Chan AH, Zhang B, Karlin-Neumann GA, Conklin BR (2016) Systematic quantification of HDR and NHEJ reveals effects of locus, nuclease, and cell type on genome-editing. *Sci Rep* 6:23549. <https://doi.org/10.1038/srep23549>
- Mock U, Hauber I, Fehse B (2016) Digital PCR to assess gene-editing frequencies (GEF-dPCR) mediated by designer nucleases. *Nat Protoc* 11:598–615
- 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. *Front Plant Sci* 11:574959. <https://doi.org/10.3389/fpls.2020.574959>
- Nadakuditi SS, Starker CG, Ko DK, Jayakody TB, Buell CR, Voytas DF, Douches DS (2019) Evaluation of methods to assess *in vivo* activity of engineered genome-editing nucleases in protoplasts. *Front Plant Sci* 10:110. <https://doi.org/10.3389/fpls.2019.00110>
- Peng Q, Xu C, Kim D, Lewis M, DiCarlo J, Wang Y (2019) Targeted single primer enrichment sequencing with single end duplex-UMI. *Sci Rep* 9:4810. <https://doi.org/10.1038/s41598-019-41215-z>
- Peng C, Zheng M, Ding L, Chen X, Wang X, Feng X, Wang J, Xu J (2020) Accurate Detection and evaluation of the gene-editing frequency in plants using droplet digital PCR. *Front. Plant Sci* 11:1919. <https://doi.org/10.3389/fpls.2020.610790>
- Porteus M, Baltimore D (2003) Chimeric nucleases stimulate gene targeting in human cells. *Science* 300:763
- Qi X, Wu H, Jiang H, Zhu J, Huang C, Zhang X, Liu C, Cheng B (2020) Conversion of a normal maize hybrid into a waxy version using *in vivo* CRISPR/Cas9 targeted mutation activity. *Crop J* 8:440–448. <https://doi.org/10.1016/j.cj.2020.01.006>
- Rejali NA, Moric E, Wittwer CT (2018) The effect of single mismatches on primer extension. *Clin Chem* 64:801–809
- Sainsbury (2021) The Sainsbury Laboratory submission to the UK DEFRA consultation on the Regulation of genetic Technologies March 2021. <http://www.tsl.ac.uk/tsl-statement-gene-editing/> Accessed April 7th 2021
- Salk JJ, Schmitt MW, Loeb LA (2018) Enhancing the accuracy of next-generation sequencing for detecting rare and subclonal mutations. *Nat Rev Genet* 19:269–285
- Schmitt MW, Kennedy SR, Salk JJ, Fox EJ, Hiatt JB, Loeb LA (2012) Detection of ultra-rare mutations by next-generation sequencing. *Proc Natl Acad Sci U S A* 109:14508–14513. <https://doi.org/10.1073/pnas.1208715109>
- Shen H, Wen J, Liao W, Lin Q, Zhang J, Chen K, Wang S, Zhang J (2020) A sensitive, highly specific novel isothermal amplification method based on single-nucleotide polymorphism for the rapid detection of *Salmonella Pollorum*. *Front Microbiol* 11. <https://doi.org/10.3389/fmicb.2020.560791>
- Song J, Pandian V, Mauk MG, Bau HH, Cherry S, Tisi LC, Liu C (2018) Smartphone-based mobile detection platform for molecular diagnostics and spatiotemporal disease mapping. *Anal Chem* 90:4823–4831
- Tang X, Liu GQ, Zou JP, 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* 19:84–96. <https://doi.org/10.1186/s13059-018-1458-5>
- Tsaloglou MN, Nemiroski A, Camci-Unal G, Christodouleas DC, Murray LP, Connelly JT, Whitesides GM (2018) Handheld isothermal amplification and electrochemical detection of DNA in resource-limited settings. *Anal Biochem* 543:116–121. <https://doi.org/10.1016/j.ab.2017.11.025>
- Varshney RK, Sinha P, Singh VK, Kumar A, Zhang Q, Bennetzen JL (2020) 5Gs for crop genetic improvement. *Curr Opin Plant Biol* 56:190–196. <https://doi.org/10.1016/j.pbi.2019.12.004>
- Wang T, Liu Y, Sun HH, Yin BC, Ye BC (2019) An RNA-guided Cas9 nickase-based method for universal isothermal DNA amplification. *Angew Chem Int Ed* 58:165382–165386. <https://doi.org/10.1002/anie.201901292>
- Wanger A, Chavez V, Huang RSP, Wahed A, Actor J, Dasgupta A (2017) Chapter 12 – overview of molecular diagnostic principles. In: Wanger A, Chavez V, Huang R, Wahed A, Dasgupta A, Actor J (eds) *Microbiology and molecular diagnosis in pathology*. Elsevier, Amsterdam, pp 233–257. <https://doi.org/10.1016/B978-0-12-805351-5.00013-2>
- Woo JW, Kim J, Kwon SI, Corvalán C, Cho SW, Kim H, Kim SG, Kim ST, Choe S, Kim JS (2015) DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nat Biotechnol* 33:1162–1164
- Xie K, Minkenberg B, Yang Y (2015) Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proc Natl Acad Sci U S A* 112:3570–3575
- Xu C, Nezami Ranjbar MR, Wu Z, DiCarlo J, Wang Y (2017) Detecting very low allele fraction variants using targeted DNA sequencing and a novel molecular barcode-aware variant caller. *BMC Genomics* 18:5. <https://doi.org/10.1186/s12864-016-3425-4>
- Young J, Zastrow-Hayes G, Deschamps S, Svitashv S, Zaremba M, Acharya A, Paulraj S, Peterson-Burch B, Schwartz C, Djukanovic V, Lenderts B, Feigenbutz L, Wang L, Alarcon C, Siksnys V, May G, Chilcoat ND, Kumar S (2019) CRISPR-Cas9 editing in maize: systematic evaluation of off-target activity and its relevance in crop improvement. *Sci Rep* 9:6729. <https://doi.org/10.1038/s41598-019-43141-6>
- Zhan X, Lu Y, Zhu JK, Botella JR (2021) Genome editing for plant research and crop improvement. *J Integr Plant Biol* 63:3–33
- Zhang Z, Mao Y, Liu W, Botella J, Zhu J-K (2015a) A multiplex CRISPR/Cas9 platform for fast and efficient editing of multiple genes in Arabidopsis. *Plant Cell Rep* 35:1519–1533. <https://doi.org/10.1007/s00299-015-1900-z>
- Zhang XH, Tee LY, Wang XG, Huang QS, Yang SH (2015b) Off-target effects in CRISPR/Cas9-mediated genome engineering. *Mol Ther Nuc Acid* 4. <https://doi.org/10.1038/mtna.2015.37>
- Zheng Z, Liebers M, Zhelyazkova B, Cao Y, Panditi D, Lynch KD, Chen J, Robinson HE, Shim HS, Chmielecki J, Pao W, Engelman JA, Lafrate AJ, Le LP (2014) Anchored multiplex PCR for targeted next-generation sequencing. *Nat Med* 20:1479–1484. <https://doi.org/10.1038/nm.3729>
- Zhou L, Wang Y, Wittwer C (2018b) Rare allele enrichment and detection by allele-specific PCR, competitive probe blocking, and melting analysis. *Biotechniques* 50:311–318. <https://doi.org/10.2144/000113668>
- Zhou W, Hu L, Ying L, Zhao Z, Chu P, Yu X-F (2018a) A CRISPR-Cas9-triggered strand displacement amplification method for ultra-sensitive DNA detection. *Nat Commun* 9:5012. <https://doi.org/10.1038/s41467-018-07324-5>
- Zischewski J, Fischer R, Bortesi L (2017) Detection of on-target and off-target mutations generated by CRISPR/Cas9 and other sequence-specific nucleases. *Biotechnol Adv* 35:95–104. <https://doi.org/10.1016/j.biotechadv.2016.12.003>