CRISPR/Cas9 System for Crop Improvement

Subjects: Plant Sciences

Contributor: Ali Razzaq

A fascinating GE tool CRISPR/Cas9 was identified for targeted genome manipulations and to express desired genes in numerous organisms. The CRISPR/Cas9 system has emerged as the most powerful tool for GE in many species including plants. The latest ground-breaking technology of CRISPR/Cas9 is basically present as an adaptive immune system of type II prokaryotes and protects them against invading organisms during phage infection by spacer acquisition, biogenesis, and target degradation. The toolbox of CRISPR/Cas9 was adapted from bacteria as well as Archaea and in included in the toolbox of engineered nucleases. There are two main components of the CRISPR/Cas9 system: a single guide RNA (sgRNA) that identifies a specific DNA sequence and the Cas9 protein which produces DSBs at a targeted site.

Keywords: CRISPR/Cas9 ; genome editing ; crop improvement ; climate change ; abiotic stresses

1. Discovery of CRISPR/Cas9 Wonder

The discovery of the CRISPR/Cas9 system dates back to 1987 when Ishino and his colleagues first identified CRISPR while studying the *iap* gene in the genome of *E. coli*. During the cloning of the *iap* gene, they unexpectedly cloned a specific portion of CRISPR, and at the conclusion of their experiment, revealed that the bacterial genome consisted of a successive array of repeats [1]. After this discovery, an Archaea (*Haloferax mediteranii*) was also found to contain the CRISPR sequences [2]. Mojica et al. (2000) reported a similar type of regularly spaced repeats in *Haloferax mediterranei* and *Haloferax volcanii*, having interrelated functions [3]. Only prokaryotes were considered to have such repetitive sequences, which were named as CRISPR but were not present in eukaryotes and viruses [4]. These short repeats have an average length of 32 bp but are of different sizes from 21 to 47 bp in different organisms. Every repeat has a unique sequence of nucleotides that are extremely conserved in specific species [5]. It was unveiled that the short regular repeats are transcribed into small RNAs [6].

Four *Cas* genes (*Cas1–4*) were discovered in prokaryotes having CRISPR DNA sequences during that period [4]. From then on, many CRISPR/Cas sequences and multiple Cas proteins were identified [6]. In 2005, CRISPR spacers were discovered in plasmids and phages by three independent research groups by applying computational and sequencing technologies [7][8][9]. The function of CRISPR/Cas was still ambiguous before Barrangou et al. (2007) successfully demonstrated for the first time that CRISPR protected *Streptococcus thermophilus* from viral attack [10]. It was revealed that the CRISPR defense mechanism prevents the horizontal gene flow in *Staphylococci* [11]. In another study, it was observed that CRISPR RNAs regulate the CRISPR interference [12]. The presence of the CRISPR/Cas system in the bacterial genome was identified to cut specific sites in plasmid DNA and bacteriophages [13]. In 2011, the CRISPR/Cas machinery of *S. thermophilus* was exploited to confer immunity in *E. coli* [14]. Some of these important events are highlighted in **Figure 1**.

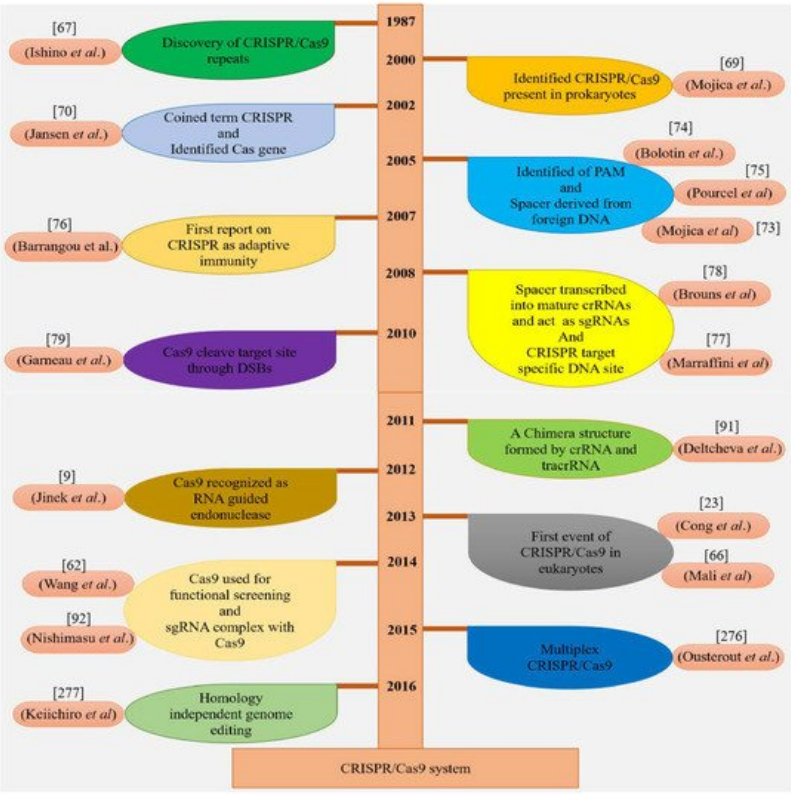

Figure 1. Historical chart illuminating key developments in the CRISPR/Cas9 system.

2. Architectural Organization of CRISPR/Cas9 System and Its Functions

Prokaryotic organisms such as bacteria and Archaea have a special type of defense machinery called CRISPR/Cas [15] in their adaptive immune network to protect them against the attack of viruses and phages [16].

All the natural CRSIPR/Cas networks are composed of many *Cas* genes, which are encoded by homologous palindromic repeated units, RNA-mediated endonucleases, and novel short RNAs, termed as “spacers” produced by the introduction of short mobile sequences called protospacers. The protospacers are derived from the unique spacers which move among the homologous palindromic sequences repeats when the cell is attacked by invaders. These spacers work as identifying units for the invaded cell and allow the CRISPR-Cas system to cut foreign DNA sequences. There are three steps to the CRISPR/Cas-mediated immune system, including adaptation, expression, and interference. The first step in this mechanism is adaptation, which is associated with the sequential layout of the CRISPR-array via a new spacer’s procurement. Precursor CRISPR-RNA (pre-crRNA) and *Cas* genes are expressed in the expression stage. Mature cr-RNAs are produced from precursor CRISPR-RNA using RNase III and Cas proteins in the interference event-specific targeted portion memorized by the combinative properties of both Cas proteins and cr-RNA [17]. The CRISPR motif, termed as protospacer adjacent motif (PAM), is connected with each protospacer and closely situated in the target portion of the sequence. The PAM was found to be a highly specific part of the foreign phage or virus genome but is not present on the CRISPR locus in bacterial genome [12].

The PAM sequence consisting of conserved dinucleotides is required upstream of the binding sites of crRNA for Cas proteins to recognize the target sequence [18]. The Cas proteins are unable to detect target DNA for effective cleavage during PAM site recognition. The PAM is also exceptionally crucial to distinguish between the bacteria’s own DNA and invader DNA sequences. Such features enable bacteria to defend their own DNA from nucleases [19]. In several kinds of CRISPR networks, PAM sequences are essential for Cas proteins functions, such as PAM sequence 5′-NNNNGATT which is targeted by Cas proteins in *Neissseria meningiditis* [20]. Similarly, Cas9 proteins target the PAM sequence 5′-NGGNG or 5′-NNAGAA in *S. thermophiles* [13][21] and 5′-NGG in *S. pyogenes* [18].

Two independent groups of scientists discovered the CRISPR/Cas9 machinery with three major kinds (type I–II–III) and two classes in host cells [22]. In 2015, Markarova and coworkers executed the comparative genomic analysis of existing data and found two further reputed types and five subtypes [23]. During the defense mechanism of CRISPR/Cas9 against invader DNA, these two classes behave differently, such as class 1 which consists of subtypes (I, III, IV) and uses many Cas proteins, while only a large Cas protein is used by the class 2 system, which has the subtypes (II, V) [24]. In the adaptation phase of the CRISPR/Cas mechanism, spacers are added by Cas1 and Cas2 proteins and pre-crRNA develop involving Cas5 or Cas6 in the type I system. The Cas6 protein is also used in the type III system for a similar process but

stimulation of 3′ end is accomplished by an uncertain element. For crRNA maturation, trans-activating crRNA (tracrRNA) and RNase are utilized in the type II mechanism [25], as shown in (**Figure 2**A). Currently, the immune system of *S. pyogenes* operates as a type II system, which is a well-established GE technique known as CRISPR. This CRISPR/Cas9 system is modified by two major units: a non-coding chimeric RNA and Cas9 endonucleases for double-stranded (dsDNA) breaks in DNA (**Figure 2**B) [18]. The Cas9 protein is directed by the guide RNA (gRNA) and Cas9 proteins recognize targeted DNA in the presence of the “seed” sequence, which is produced by spacers derived from crRNA and the *S. pyogenses* Cas9 (SpCas9) 5′ NGG ′3 sequence lying closely to the target region [18]. The crRNA and tracrRNA are complementary to each other and it directs pre-crRNA to mature crRNA by means of RNase III. After the maturation of crRNA, it guides Cas9 proteins to break specific DNA sequences [26]. In 2014, Nishimasu et al. (2014) demonstrated that the SpCas9 and gRNA DNA endonuclease has unique lobes, such as an assembly composed of a target detection lobe which is attached to the heteroduplex of sgRNA: a DNA molecule and a nuclease lobe which nicked the target DNA sequence [27], as illustrated in (**Figure 2**C).

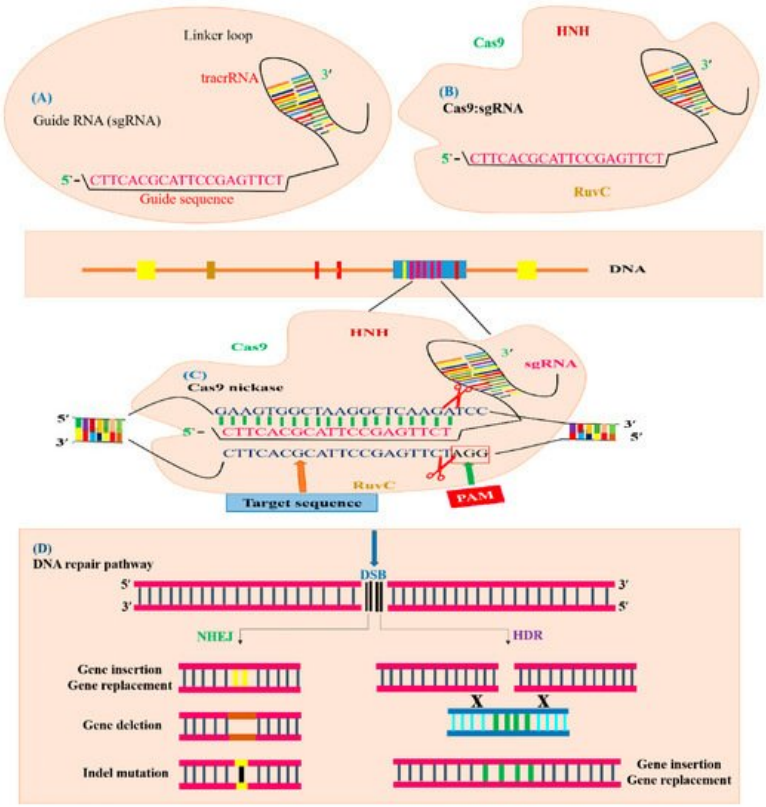

Figure 2. Illustration of CRISPR/Cas9-mediated GE. The CRISPR/Cas9 system is composed of sgRNA and Cas9. (**A**) sgRNA with a guide sequence (colored pink) is developed by the combination of protospacer with crRNA and tracrRNA. (**B**) Cas9 machinery combines with sgRNA to form a complex to trigger CRISPR/Cas9 editing. The Cas9 nuclease consists of two parts, depending on its function and structure. The recognition site identifies the target DNA and interacts with sgRNA. The nuclease site contains two domains RuvC-like and HNH which cleave the target DNA site non-complementary by the RuvC domain and complementary by the HNH domain to the gRNA. (**C**) The Cas9 nuclease detects the genomic target site (indicated with blue color) having a 20 bp target sequence that is homologous to seed or guide sequence (indicated with pink color), which is crucial for Cas9 activity and specificity. The specific PAM sequence (indicated with red color) is detected by Cas9: sgRNA complex and DSBs created by the Cas9 endonuclease three base pairs upstream of the PAM sequence. (**D**) Targeted mutagenesis of a desired gene is achieved by filling the DSB (indicated with black color) by means of the HDR or NHEJ mechanism. The NHEJ repair mechanism generally produces insertion (indicated with yellow color), deletion (indicated with brown color) or indels (indicated with black line) at the break point, generating targeted mutants. The HDR repair mechanism uses a template DNA sequence for homologous recombination to produce gene replacement or gene insertion (indicated with green color).

3. Genome Editing Mechanism of CRISPR/Cas9 System

The mode of GE is established by the healing process of the genome. After the identification of the target site, Cas9 allows sgRNA to pair with the target DNA sequence. The Cas9 endonuclease is composed of the HNH and RucV-like domain, which cuts the target DNA strands three to four bases upstream of the PAM site. The HNH domain cuts the complementary DNA strands while the RuvC domain cleaves the non-complementary to gRNA. The blunt-ended DSBs can be repaired by the HDR and NHEJ repair pathways (**Figure 2**D). The NHEJ is error prone and causes DNA insertion

or deletion at the target sequence [28]. The expression of sgRNA as pair, NHEJ mechanism came up with large deletions. The large deletions in chromosomes were attained by the NHEJ mechanism utilizing co-expressed sgRNAs. The HDR repair mechanism is only operational when a specific homologous target site is available with respect to the DSB site. In plants, through GE, many outstanding repairs were achieved via HDR, such as gene replacement, DNA correction, and targeted knock-in [29][30].

Biolistic and *Agrobacterium*-mediated transformation can be applied to transfer the sgRNA and Cas9 protein into desired cells [31]. GE by CRISPR/Cas9 is heavily dependent on the choice of sgRNA promoters and ubiquitous expression of the Cas9 enzyme. Universal *CaMV35S* RNA polymerase II promoters have been extensively used for Cas9 expression in plants. Similarly, for sgRNA expression, U3 or U6 RNA pol III promoters are applied [32]. The expression level of sgRNAs is significantly greater in endogenous promoters as compared to exogenous promoters [33]. Moreover, sgRNA expression is guided by U6 promoters which were derived from monocotyledonous or dicotyledonous varieties and can only be used in monocot or dicot plants [34]. For successful integration of CRISPR/Cas9 machinery in plant nuclei, Cas9 proteins must join with nuclear localization signals [32].

To bind the target DNA by synthetically developed short gRNA sequences of approximately 20 nucleotides, the mechanism of CRISPR/Cas editing demands the PAM 5′ NGG motif for the Cas9 enzyme to cleave 3–4 bases in the target DNA sequence after the generation of the protospacer [18]. There are two domains of Cas nucleases which have the ability to cut one strand of DNA like the HNH domain and RuvC-like domain. Simple steps involving the execution of the CRISPR mechanism are recognition of the PAM sequence; sgRNA development; cloning of sgRNA; transformation into the host cell; selection of transformed individual organisms; and edited lines confirmation, as described in **Figure 3**.

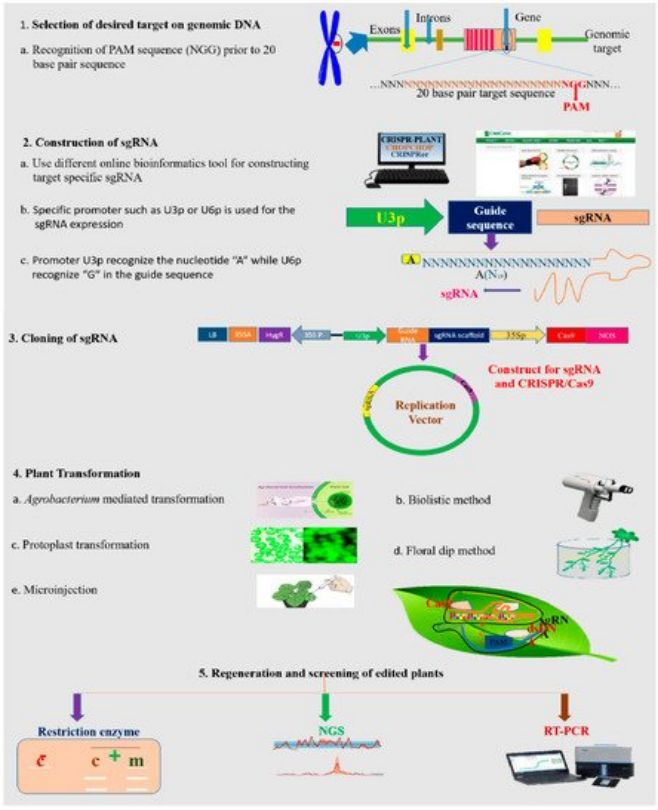

Figure 3. Basic steps in the workflow of CRISPR/Cas9-based genome editing.

References

1. Ishino, Y.; Shinagawa, H.; Makino, K.; Amemura, M.; Nakatura, A. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isoenzyme conversion in Escherichia coli, and identification of the gene product. J. Bacteriol. 1987, 169, 5429–5433.
2. Mojica, F.J.M.; Ferrer, C.; Juez, G.; Rodríguez-Valera, F. Long stretches of short tandem repeats are present in the largest replicons of the Archaea Haloferax mediterranei and Haloferax volcanii and could be involved in replicon partitioning. Mol. Microbiol. 1995, 17, 85–93.
3. Mojica, F.J.M.; Díez-Villaseñor, C.; Soria, E.; Juez, G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. Mol. Microbiol. 2000, 36, 244–246.

4. Jansen, R.; Van Embden, J.D.A.; Gaastra, W.; Schouls, L.M. Identification of genes that are associated with DNA repeats in prokaryotes. Mol. Microbiol. 2002, 43, 1565–1575.
5. Karginov, F.V.; Hannon, G.J. The CRISPR system: Small RNA-guided defense in bacteria and archaea. Mol. Cell 2010, 37, 7–19.
6. Haft, D.H.; Selengut, J.; Mongodin, E.F.; Nelson, K.E. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/cas subtypes exist in prokaryotic genomes. PLoS Comput. Biol. 2005, 1, e60.
7. Mojica, F.J.M.; Díez-Villaseñor, C.; García-Martínez, J.; Soria, E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J. Mol. Evol. 2005, 60, 174–182.
8. Bolotin, A.; Quinquis, B.; Sorokin, A.; Dusko Ehrlich, S. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology 2005, 151, 2551–2561.
9. Pourcel, C.; Salvignol, G.; Vergnaud, G. CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. Microbiology 2005, 151, 653–663.
10. Barrangou, R.; Fremaux, C.; Deveau, H.; Richards, M.; Boyaval, P.; Moineau, S.; Romero, D.A.; Horvath, P. CRISPR provides acquired resistance against viruses in prokaryotes. Science 2007, 315, 1709–1712.
11. Marraffini, L.A.; Sontheimer, E.J. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science 2008, 322, 1843–1845.
12. Brouns, S.J.J.; Jore, M.M.; Lundgren, M.; Westra, E.R.; Slijkhuis, R.J.H.; Snijders, A.P.L.; Dickman, M.J.; Makarova, K.S.; Koonin, E.V.; Oost, J. Antiviral defense in prokaryotes. Science 2008, 321, 960–964.
13. Garneau, J.E.; Dupuis, M.È.; Villion, M.; Romero, D.A.; Barrangou, R.; Boyaval, P.; Fremaux, C.; Horvath, P.; Magadán, A.H.; Moineau, S. He CRISPR/cas bacterial immune system cleaves bacteriophage and plasmid DNAT. Nature 2010, 468, 67–71.
14. Sapranauskas, R.; Gasiunas, G.; Fremaux, C.; Barrangou, R.; Horvath, P.; Siksnys, V. The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. Nucleic Acids Res. 2011, 39, 9275–9282.
15. Pawluk, A.; Davidson, A.R.; Maxwell, K.L. Anti-CRISPR: Discovery, mechanism and function. Nat. Rev. Microbiol. 2018, 16, 12–17.
16. Bhaya, D.; Davison, M.; Barrangou, R. CRISPR-Cas systems in bacteria and archaea: Versatile small RNAs for adaptive defense and regulation. Annu. Rev. Genet. 2011, 45, 273–297.
17. Bortesi, L.; Fischer, R. The CRISPR/Cas9 system for plant genome editing and beyond. Biotechnol. Adv. 2015, 33, 41–52.
18. Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J.A.; Charpentier, E. A Programmable Dual-RNA —Guided DNA endonuclease in adaptive bacterial immunity. Science 2012, 337, 816–821.
19. Mojica, F.J.M.; Díez-Villaseñor, C.; García-Martínez, J.; Almendros, C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology 2009, 155, 733–740.
20. Zhang, Y.; Heidrich, N.; Ampattu, B.J.; Gunderson, C.W.; Seifert, H.S.; Schoen, C.; Vogel, J.; Sontherimer, E.J. Processing-Independent CRISPR RNAs limit natural transformation in Neisseria meningitidis. Mol. Cell 2013, 50, 488–503.
21. Gasiunas, G.; Barrangou, R.; Horvath, P.; Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. USA 2012, 109, E2579–E2586.
22. Makarova, K.S.; Aravind, L.; Wolf, Y.I.; Koonin, E.V. Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. Biol. Direct 2011, 6, 38.
23. Makarova, K.S.; Wolf, Y.I.; Alkhnbashi, O.S.; Costa, F.; Shah, S.A.; Saunders, S.J.; Barrangou, R.; Brouns, S.J.J.; Charpentier, E.; Haft, D.H.; et al. An updated evolutionary classification of CRISPR-Cas systems. Nat. Rev. Microbiol. 2015, 13, 722–736.
24. Makarova, K.S.; Haft, D.H.; Barrangou, R.; Brouns, S.J.J.; Charpentier, E.; Horvath, P.; Moineau, S.; Mojica, F.J.M.; Wolf, Y.I.; Yakunin, A.F.; et al. Evolution and classification of the CRISPR-Cas systems. Nat. Rev. Microbiol. 2011, 9, 467–477.
25. Rath, D.; Amlinger, L.; Rath, A.; Lundgren, M. The CRISPR-Cas immune system: Biology, mechanisms and applications. Biochimie 2015, 117, 119–128.
26. Deltcheva, E.; Chylinski, K.; Sharma, C.M.; Gonzales, K.; Chao, Y.; Pirzada, Z.A.; Eckert, M.R.; Vogel, J.; Charpentier, E. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature 2011, 471, 602–607.

27. Nishimasu, H.; Ran, F.A.; Hsu, P.D.; Konermann, S.; Shehata, S.I.; Dohmae, N.; Ishitani, R.; Zhang, F.; Nureki, O. Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell 2014, 156, 935–949.
28. Cong, L.; Ran, F.A.; Cox, D.; Lin, S.; Barretto, R.; Habib, N.; Hsu, P.D.; Wu, X.; Jiang, W.; Marraffini, L.A.; et al. Multiplex genome engineering using CRISPR/Cas Systems. Science 2013, 339, 819–823.
29. Li, J.F.; Norville, J.E.; Aach, J.; McCormack, M.; Zhang, D.; Bush, J.; Church, G.M.; Sheen, J. Multiplex and homologous recombination–mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. Nat. Biotechnol. 2013, 31, 688–691.
30. Schiml, S.; Fauser, F.; Puchta, H. The CRISPR/Cas system can be used as nuclease for in planta gene targeting and as paired nickases for directed mutagenesis in Arabidopsis resulting in heritable progeny. Plant J. 2014, 80, 1139–1150.
31. Miao, J.; Guo, D.; Zhang, J.; Huang, Q.; Qin, G.; Zhang, X.; Wan, J.; Gu, H.; Qu, L.J. Targeted mutagenesis in rice using CRISPR-Cas system. Cell Res. 2013, 23, 1233–1236.
32. Belhaj, K.; Chaparro-Garcia, A.; Kamoun, S.; Nekrasov, V. Plant genome editing made easy: Targeted mutagenesis in model and crop plants using the CRISPR/Cas system. Plant Methods 2013, 9, 39.
33. Sun, X.; Hu, Z.; Chen, R.; Jiang, Q.; Song, G.; Zhang, H.; Xi, Y. Targeted mutagenesis in soybean using the CRISPR-Cas9 system. Sci. Rep. 2015, 5, 10342.
34. Mao, Y.; Botella, J.R.; Zhu, J.K. Heritability of targeted gene modifications induced by plant-optimized CRISPR systems. Cell. Mol. Life Sci. 2017, 74, 1075–1093.