Himachal Journal of Agricultural Research 45(1&2) : 96-99 (2019)

CRISPR-A powerful functional genomic tool in crop improvement

Shailja Sharma*, R.K. Mittal, V.K. Sood and Sanchit Thakur
Department of Crop Improvement
CSK Himachal Pradesh Krishi Vishvavidyalaya, Palampur-176 062, India.
*Corresponding author: shailjasharma29791@gmail.com

Manuscript Received: 27.11.2019; Accepted: 06.12.2019

Abstract

CRISPR and CRISPR Associated (Cas) Proteins (CRISPR-Cas) system is a vital defense system found in bacteria and archea. CRISPR activity requires a set of Cas genes, found adjacent to the CRISPR, that codes for protein essential for immune response. CRISPR system works in three stages to carry out a full immune response to invading foreign DNA i.e. acquisition, expression and interference stage. Recently Type II CRISPR-Cas9 system has also been adapted to perform genome engineering by inducing double-strand breaks (DSBs) in host DNA that can be repaired by either non-homologous end-joining (NHEJ) or homology-directed repair (HDR). Although the full potential of CRISPR-Cas9 has not been harnessed so far, this technology has already brought revolutionary changes in genomic research. CRISPR-Cas9 system has great potential for both research and therapeutics; however, improvements can still be made in its specificity, efficiency and off target effects. With the advancement, the use of this system is increased in model plants as well as other crop plants like *Arabidopsis*, tobacco, maize, rice, wheat, sorghum, barley, brassica, soybean etc.

Key words: CRISPR, Cas, spacer, repeat, crop improvement.

CRISPR and Cas genes first identified in *Escherichia coli* in 1987, is an array of short repeated sequences separated by spacers with unique sequences. The spacers are often derived from nucleic acid of viruses, an observation that gave rise to an idea that CRISPR are part of an anti-virus system (Guilinger *et al*. 2014). The system has also came out as a powerful genome editing tool. CRISPR containing bacteria acquire small DNA fragments of invading bacteriophages and plasmids before transcribing them into CRISPR RNAs (crRNAs) which guide cleavage of invading RNA or DNA (Wiles *et al*. 2015). The CRISPR-Cas9 system generate genome modifications by relying on the presence of a proto-spacer adjacent motif (PAM) and a twenty-nucleotide small guide RNA (sgRNA) complementary to the target DNA (Sternberg 2014).

Historically, the genetic base for plant breeding has been limited to the standing natural allelic variation within the germplasm collections or randomly induced variation generated by irradiation or chemical-based mutagenesis. The utilization of de novo site-directed variation is an attractive alternative for expanding the genetic base of a crop species, particularly for high-value traits of interest (Curtin *et al*. 2012).Creation of variation or variability in existing gene pool of crop plants is the foremost requirement in crop improvement programmes (Rani *et al*. 2016). Deciphering gene function and connecting genotype to phenotype are the primary challenges in order to utilize these resources to engineer biological systems for addressing global challenges, for example environmental cleanup, human disease treatment and clean energy production. Till date, a vast number of tools have been applied for creation of genetic modifications in several organisms. However, the demand for genetic engineering is transforming from targeting one site to targeting multiple sites in a single genome for efficient genome scale engineering (Xu *et al*. 2014).

The main focus of genome engineering is to develop methods to manipulate nucleic acids precisely. For manipulation of nucleic acids, DNA repair pathways are used to correct double stranded breaks. Two repair mechanisms are: homologous recombination (HR), in which DNA templates bearing sequence similarity to the break site are used to introduce sequence changes to the target locus, and non homologous end joining (NHEJ), in which the broken chromosomes are rejoined, often imprecisely,

thereby introducing nucleotide changes at the break site. These tools of genome engineering will soon be routinely used to repair DSB to achieve highly efficient, precise modifications of plant genomes. Two deployed to advance both basic and applied plant research (Voytas 2013).

Basically, Genome editing is the precise modification of the nucleotide sequence of the genome. Genome editing precisely modifies nucleotides (A, T, G, C) in the genetic code and is performed by using specifically engineered 'molecular scissors' to create precise breaks in the genome and repairing and editing of the genome by harnessing the DNA repair mechanisms (HR & NHEJ) of cells.

Currently, four types of engineered nucleases are used for genome editing:

1 Engineered homing endonucleases / meganucleases(EMNs)

2. Zinc finger nucleases (ZFNs)

3. Transcription activator-like effector nucleases (TALENs)

4. CRISPR (Clustered regularly interspaced short palindromic repeats)/Cas

Nowadays, CRISPR/Cas system is most commonly used because it is simple to use, highly efficient and specific as compared to earlier mutation methods; its use in plant biology research is increasing rapidly to enhance yield, biotic and abiotic stress tolerance, increased nutritional value and new trait development (Khandagale and Nadaf 2016). It was widely accepted by academics and research organizations led to CRISPR Craze. CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) is basically a family of DNA sequences found in some bacteria that has snippets of DNA from viruses that have previously attacked the bacterium. These snippets are utilized by these bacterium to detect and destroy DNA further attacks from same viruses. These sequences also play a vital role in defence system of bacteria and thus this form the basis of a genome editing technology called as CRISPR/Cas9 which permits permanent modification of genes within these organisms. It is cheaper, rapid and more precise than various other techniques of editing DNA and also has wide applications. The CRISPR/Cas system is a prokaryotic immune system that is responsible for resistance to various foreign genetic elements like those present in plasmids and phage. CRISPR/Cas9 has opened new era in biotechnology. This technology is used to understand the function of a gene or a protein, one interferes with it in a sequence-specific way and monitors its effects on the organism. In some organisms, it is difficult or impossible to perform site- specific mutagenesis, and therefore more indirect methods must be used, such as silencing the gene of interest by RNA interference (RNAi). But sometimes gene disruption by siRNA can be variable or incomplete. CRISPR/Cas9 can cut any targeted position in the genome and introduce a modification of the endogenous sequences for genes that are impossible to specifically target using conventional RNAi.

The first description of CRISPR was of Osaka University researcher named Yoshizumi Ishino in 1987, who by mistake cloned part of a CRISPR together with the iap gene. (Ishino *et al*.1987). Repeated sequences are typically arranged consecutively along DNA therefore, organization of the repeats was unusual. These interrupted clustered repeats had unknown fuction. Ishino and co-workers (1987) discovered an unusual structure of repetitive DNA downstream from the *E. coli* iap gene consisting of invariant direct repeats (29 nt) and variable spacing sequences (32 nt). These unusual sequences was named as SRSR (Short Regularly Spaced Repeats) in 2000 (Mojica *et al*. 2002). Later, Jansen and coworkers (1995) coined the term CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) and reported that CRISPRs co-localize with specific cas (CRISPR-associated proteins) genes (Jansen *et al*.1995).The CRISPR-Cas defense comprises a multistep process by which specific small fragments of foreign nucleic acids are first recognized as being non self and then incorporated into the host genome between short DNA repeats. Subsequently, these fragments or spacers, in conjunction with host Cas proteins, are used as a surveillance and adaptive immune system by which incoming foreign nucleic acids are recognized and destroyed or possibly silenced.The CRISPR-Cas9 system consists of two key molecules that introduce a change into the DNA. First key player is an enzyme called Cas9. This acts as 'molecular scissors' which cut the two strands of DNA at a definite locus in the genome so that chunks of DNA can then be removed or added. Second key player is a piece of RNA called guide RNA (gRNA). This consists of a small piece of predesigned RNA sequence (about 20 bases long) located within a longer RNA scaffold. The pre-designed sequence 'guides' Cas9 to the right part of the genome and scaffold part binds to DNA. This makes sure that the Cas9 enzyme cuts at the right point in the genome.

A practical CRISPR-Cas system has two distinguishable elements needed for activity. CRISPR locus/array set on the ordering (either body or

plasmid), that contains the hyper variable spacers non heritable from virus or cellular inclusion of DNA. The CRISPR array contains AT-rich leader sequence followed by short repeats that are separated by distinctive spacers. CRISPR repeats usually direct size from twenty eight to thirty seven base pairs (bps). Second part is various cluster of cas genes set within the neighbourhood of a CRISPR locus, that encipher proteins (commonly known as Cas proteins) needed for the multistep defense against invasive genetic components. Cas-proteins contain diagnosable domains characteristic of getting helicases, nucleases, polymerases and RNA-binding properties. There are 3 stages in CRISPR- Cas mechanism: First stage is CRISPR Spacer Acquisition, during this Cas1 and Cas2 proteins play a crucial role. Integration primarily happens at the leader finish of the CRISPR locus. First is that the recognition of the invasive macromolecule and scanning foreign deoxyribonucleic acid for potential PAMs (Protospacer adjacent motifs a terribly short stretch of preserved nucleotides within the immediate neighbourhood of the protospacer is noted because the protospacer adjacent motif (PAM), or CRISPR motif, that determine protospacers (the sequence on the infectious agent ordering that corresponds to a spacer is thought as protospacer) & the generation of a replacement repeat spacer by process of the macromolecule, and also the integration of the new CRISPR repeat spacer unit at the leader finish of the CRISPR locus. Second stage is CRISPR expression within which a primary transcript, or pre-CRISPR ribonucleic acid (pre-crRNA) is transcribed from the CRISPR locus by ribonucleic acid enzyme. Pre-crRNA isn't able to activate correct DNA cleavage. First, the pre-RNA base pairs with a complementary strand to a different ribonucleic acid molecule known as trans-activating crRNA (tracrRNA) at its 3'end. The twin advanced is then cleaved into short fragment by RNase III. This little fragmented duplex of crRNA and tracrRNA is a guide ribonucleic acid. Last stage is CRISPR interference within which crRNA-tracrRNA duplex triggers deoxyribonucleic acid cleavage activity by recruiting Cas9 nuclease. The Cas9- crRNA-tracrRNA comlex simply targets and cut the foreign deoxyribonucleic acid strand once the 5'end sequence of gRNA acknowledges and nucleotide to the protospacer region of the foreign deoxyribonucleic acid. The cleavage of the crRNA–foreign macromolecule is now advanced. On the opposite hand, if there are mismatches between the spacer and target deoxyribonucleic acid or if there are unit mutations within the PAM, then cleavage is not initiated. In 2013, the CRISPR-Cas9 system was first demonstrated to be a tool to manipulate the genome at a desired site. A chimeric single guide RNA was generated by fusing two noncoding RNAs, a crRNA and a tracrRNA. The chimeric gRNA and Cas9 endonuclease form a complex and create DSBs, which are processed to short indels at specific sites. Since then, the CRISPR-Cas9 system has gained a great attention in genetic engineering.

Emergence of CRISPR as a gene editing technique as transformed the plant biology research as it has ability to generate highly specific and efficient mutations in short time span. This technique is highly specific, rapid and cost effective, so can be used for the labour and time intensive classical plant breeding. This precise gene editing was successfully applied for functional genomics study, transcriptional regulation, disease and pest resistance and new trait development in model plants as well as cereals, vegetable and fruit crops. Shen *et al*. (2017) have successfully edited eight agronomic genes using one binary vector for each genetic transformation in rice. Kim *et al*. (2018) demonstrated CRISPER/Cas 9 system in wheat protoplasts for two abiotic stress related genes i.e. $TaDREB_2$ (Wheat ethylene responsive factor 3). Maize thermosensitive genetic male sterile 5 ($ZmTMS_5$) gene responsible to cause male sterility in maize, was targeted for genome editing by CRISPR-Cas 9 system (Li *et al*. 2017b). CRISPR/Cas 9 technology has been used in other monocot and dicot plants to improve various traits. So, it will be a prominent area to edit the plant genome in future. All the gene editing techniques have their own pros and cons, but it was seen that due to its simple, versatile nature and affordability, CRISPR/Cas9 has become a method of choice among the plant molecular biologist. Though this technique is highly specific, some degree of off-target effect has been reported but careful designing these tools and selection of target will reduce them completely. The degree of off targeting can be overcome by designing and discovering highly specific nucleases as the development in these technologies is advancing rapidly.

References

Curtin SJ, Voytas DF and Stupar RM. 2012. Genome engineering of crops with designer nucleases. The Plant Genome **5** (2): 42-50.

Guilinger JP, Thompson DB and Liu DR. 2014. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. Nature Biotechnology **32**: 577-582.

Ishino Y, Shinagawa H, Makino K, Amemura M and Nakata A. 1987. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in E coli and identification of the gene product. Journal of Bacteriology **169**: 5429-5433.

Jansen R, Embden JDAV, Gaastra W and Schouls LM. 1995. Identification of genes that are associated with DNA repeats in prokaryotes. Molecular Microbiology **43**:1565-1575.

Kim D, Alptekin B and Budak K. 2018. CRISPR/ Cas 9 genome editing in wheat. Functional and integrative Genomics **18**: 31-41.

Khandagale K and Nadaf A. 2016. Genome editing for targeted improvement of plants. Plant Biotechnology Report **4**: 51-68.

Li J, Zhang H, Si,X, Tian Y, Chen K and Liu J. 2017. Generation of thermosensistive male sterile maize by targeted knockout of ZMTMS$_5$ gene. Journal of Genetics and Genomics **44**: 465-468.

Mojica FJM, Ferrer C, Juez G and Rodrıguez-Valera F. 2002. 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. Molecular Microbiology **17**: 85–93.

Rani R, Yadav P, Barbadinkar KM, Baliyan N, Malhotra EV, Singh VK, Kumar A and Singh D. 2016. CRISPR/Cas-9: A promising way to exploit genetic variations in the plants. Biotechnological Letters **38**: 1991-2006.

Shen C, Que Z, Xia Y, Tang N, Li D & He R. 2017. Rapid generation of genetic diversity by multiplex CRISPR/ Cas 9 genome editing in rice. China life Sciences **60**: 506-515.

Sternberg SH, Redding S, Jinek M, Greene EC and Doudna JA. 2014. DNA interrogation by the CRISPR RNA guided endonuclease Cas9. Nature **507**: 62–67.

Voytas DF. 2013. Plant genome engineering with sequence specific nucleases. Annual Review Plant Biology **64**: 327-350.

Wiles MV, Qin W, Cheng AW and Wang H. 2015. CRISPR-Cas9-mediated genome editing and guide RNA design. Mammalian Genome **26**: 501-510.

Xu T, Li Y, Nostrand JDV, He Z and Zhou J. 2014. Cas-9 based tool for targeted genome editing and transcriptional control. Applied and Environmental Microbiology **80**: 1544-1552.