CRISPR-Cas 9 mediated genome editing of selected early flowering genes in Wheat

Thesis submitted in the partial fulfilment of requirements for the degree of

MASTER OF PHILOSOPHY
IN
PLANT GENOMICS & BIOTECHNOLOGY

BY

QURATULAIN

DEPARTMENT OF PLANT GENOMICS AND BIOTECHNOLOGY PARC INSTITUTE OF ADVANCED STUDIES IN AGRICULTURE NATIONAL AGRICULTURAL RESEARCH CENTRE, ISLAMABAD QUAID-I-AZAM UNIVERSITY, ISLAMABAD, PAKISTAN
2020

CRISPR-Cas 9 mediated genome editing of selected early flowering genes in Wheat

A Thesis
Submitted to Quaid-i-Azam University, Islamabad in the Partial fulfilment of the requirements for the degree of

MASTER OF PHILOSOPHY
IN
PLANT GENOMICS AND BIOTECHNOLOGY

DEPARTMENT OF PLANT GENOMICS AND BIOTECHNOLOGY PARC INSTITUTE OF ADVANCED STUDIES IN AGRICULTURE NATIONAL AGRICULTURAL RESEARCH CENTRE, ISLAMABAD QUAID-I-AZAM UNIVERSITY, ISLAMABAD, PAKISTAN
2020

CERTIFICATE

This thesis submitted by **Ms. Quratulain** from Plant genomics and biotechnology, PARC institute of advance study in Agriculture, Islamabad, Pakistan, is accepted in its current form. This thesis fulfils all the requirement for facilitating her with Degree of Master of Philosophy in **Plant Genomics and Biotechnology**.

Internal Examiner:

Dr. Ghulam Muhammad Ali
Chief Scientist-11,
NIGAB, NARC, Islamabad

External Examiner:

Dr.Jabar Zaman Khan Khattak
Associate professor
Bioinformatics & Biotechnology
International Islamic University,
Islamabad

Chairperson:

Dr. Shoukat Ali
Principle Scientific Officer,
NIGAB, NARC, Islamabad

PLAGIARISM REPORT

It is certified that **Miss Quratulain** (02361713001) has submitted her M.Phil thesis having Title "**CRISPR-Cas 9 mediated genome editing of selected early flowering genes in Wheat"** that has been checked on Turnitin for similarity index (plagiarism).

Overall plagiarism=15% that lies in the limit provided by HEC (19%)

Dr. Ghulam Muhammad Ali
Chief scientist-11,
NIGAB, NARC, Islamabad

AUTHOR'S DECLARATION

I **Quratulain** hereby declare that I have produced the work presented in thesis, **"CRISPR-Cas 9 mediated genome editing of selected early flowering genes in Wheat"** during the scheduled period of study under the supervision of **Dr. Ghulam Muhammad Ali.** I also declare that I have not taken any material from any source except referred to wherever due. If a violation of HEC rules on research has occurred in this thesis, I shall be liable to punishable action under the plagiarism rules of the HEC.

Quratulain
Reg # 02361713001

DEDICATION

I sincerely dedicate this life time achievement to my loving and very caring Husband, and My Parents whose prays and support made my path smooth and comfortable to my goals. They are my mentor and I am proud of them.

Table of Contents

LIST OF ABBREVIATIONS

BLAST	Basic Local Alignment Search Tool
DSB	DNA double strand break
cDNA	Complementary Deoxyribonucleic Acid
CDS	Coding DNA Sequence
dNTPS	Deoxy Ribonucleotide Triphosphates
HR	homologous recombination
FAOST	Food And Agriculture Organization Statistics
IWGSC	International Wheat Genome Sequencing Consortium
MEGA	Molecular Evolutionary Genetics Analysis
NARC	National Agricultural Research Centre
NIGAB	National Institute for Genomics and Advance Biotechnology
CRISPR/Cas9	Clustered Regularly Interspaced Short Palindromic Repeat/CRISPR-associated Cas9
RNA	Ribonucleic Acid

sgRNA	single guide RNA
PCR	Polymerase Chain Reaction
NHEJ	Nonhomologous End Joining
TAE	Tris- Acetate-EDTA
IWGSC	International Wheat Genome Sequencing Consortium
HDR	Homology-directed Repair
IDT	Integrated DNA Technologies
MCS	multiple cloning sites
PI	Photoperiod insensitivity
Cas9	CRISPR associated protein 9
NGS	Next Generation Sequencing
DNA	Deoxyribonucleic Acid
NHEJ	non-homologous end-joining
PAM	Protospacer Adjacent Motif
HptII	hygromycin resistance gene

ZFNs	Zinc Finger Nucleases
RNase	Ribonuclease
TALENs	Transcription Activator-like Effector Nucleases

LIST OF FIGURES

Figure No: Title: Page No

LIST OF TABLES

Table No: Title: Page No:

ACKNOWLEDGMENTS

In the name of Allah, Most Gracious, Most Merciful

First and foremost, I am very thankful to Allah Almighty for His countless blessings upon me. I would like to acknowledge my deep prayers and admiration to my university Quaid-i-Azam University Islamabad and all those who provided me the possibility to complete the research. I would like to express my sincerest gratitude to my Supervisor **Dr. Ghulam Muhammad Ali**, for his knowledge, excellent supervision, patience, continuous support and motivation during my research work. I consider myself fortunate to work under his kind supervision. Besides my advisor, I am also grateful to Chairman, DG NARC and all staff members of Plant Genomics and Advanced Biotechnology Department, for facilitating the quality of work in our department.

I wish to express my gratefulness to National Institute of Genomics and Applied Biotechnology, NARC Islamabad for valued insights, priceless assistance and for mentoring and support in the run up to completion of this work. I take this opportunity to express gratitude to all faculty members of PARC Institute of Advanced Studies In Agriculture, Islamabad for providing me with a high standard education and a memorable university experience. At last I would like thankful to my Parents and Siblings for their support throughout for which my mere expression of thanks is not enough. I am especially thankful to my Husband Muhammad Akmal for his encouragement and moral support throughout the project. Last, but no means by least, I am thankful to all lab Fellows. I acknowledge my friends Hira Hameed, and zeeshan for their warm empathy, loving support and care. Lastly, I also record my sense of gratitude to one and all who have lent their hand directly or indirectly to this undertaking.

QURATULAIN

ABSTRACT

The CRISPR (Clustered Regularly Interspaced Short Palindromic Repeat) Cas9 method is the current favorite among the available gene editing methods, as it has shown to function as editing method in many species, while following a set of simple target specific design rules. However, it may not ultimately result in the introduction of foreign DNA at the target location. Despite being robust, developed and most successful genome editing tool, it has been frequently called into question, yet it is used as most accurate available genome editing tool. It maintains the great potential of improving crop plant performance by enhancing yield, nutritional value in crop and capability to cope with biotic as well as abiotic stresses. Different technique or guidelines have been developed specifically for different plant species. Since the domestication of wheat, flowering period has been intensively modified to get high yield in varying condition and environment and also, it provided a great variety of adaptation mechanism for study and research. Hexaploid – a type of wheat, ancestrally a long day plant - may not be compatible with many environments and such environments may require the types or varieties of wheat having PI (Photoperiod Insensitivity) so that it can flower in short days. CRISPR / Cas9 mediated protocol was used for modifying genome in wheat, along with detailed record on design of RNA, construction of vector, and transformation process in plants. Mutations on A B or D genome of Ppd-1 gene causes PI and also, flowering is affected through an alternative route, most likely, by up regulating FT as a result of mutation in Ppd-1. That suggests that a strong direct quantitative relationship exists between Ppd-1 mutation and FT expression and flowering.

INTRODUCTION

Wheat (Triticum aestivum L.) - One of the most essential staple food crops - accounts for 19 per cent of human calories (FAOSTAT, 2015). Despite this socio - economic importance, progress has been slow in breeding high-yield wheat compared to other crops such as rice and maize, largely due to the complexity of their genetic systems. Expanding the agricultural region to meet the needs of the growing population worldwide is not an option given the lack of fertile land. Pakistan is facing a major challenge to deal with the explosion of population growth, there is huge demand of this staple food. The improvement in yield capacity is expected due to genetically modified local varieties. Hence, it is important to develop technologies to increase the capacity for improving genetic yield. Wheat yield capacity is anticipated to be improved by manipulating the parameters of the carbon capture, plant morphology, flowering period, biomass, and harvest index. Yield in wheat comes from parts of the flowers and their components, such as fruits and grains. Knowing the genetic basis of flower production can therefore allow us to turn inflorescences and flowers according to our wishes. There are different parameter of plant like Inflorescence / spikelet and grain size period to flower, amount and structure / arrangement are crucial criteria that decide, at least in part, where a crop plant can be grown and how many fruits or grains it can yield. To enhance the yield there is need to design wheat plant having structural properties which are radically altered. In theory, with maximum biomass it should be possible to create plants that have flowered sooner, increased the number of spikelets and grain size. Therefore, the discovery of genes associated with agronomic characteristics offers good prospects for crop development. A recent development in this respect is the characterisation of genes involved in floral transformation, spikeletal construction, grain size and regulation of biomass. Thus a model plant can be produced which has packed all the appropriate genes with optimal expression for high yield. Growth of high yielding wheat varieties would directly impact agricultural socio-economic conditions of the poor farmers, and ultimately Pakistan.

1.1. Factors which affect performance and production of wheat:

We can hardly exaggerate the importance of wheat grain as the most valuable food worldwide. In Pakistan, wheat is not only essential as staple food but also as income earner for poor farmers. Pakistan is one of the countries with the lowest production per acre (1200 k / acre in Pakistan, compared with 2500 kg / acre in America). With the passage of time, this gap in yield is widening. Major obstacle is the lack of fertile land and poor yielding varieties. The fast and feasible solution is to increase the capacity for yields by genetic manipulation. Modern research into functional genomics has shown that the following areas will increase the potential for yield. Those areas are improved biofuels, flowering, implementation of plant architecture, improved biomass, adjustment of flowering time, and performance analysis of harvest index. These are the particular areas which are being pursued for analysis for enhanced yield.(Boden et al., 2015). The progress from vegetative to regenerative development (blossoming) is related with heading time, one of the most significant trait in grain crops. AABBDD is a Triticum aestivum genome that contain 42 chromosome, commonly called bread wheat (Snape, Butterworth, Whitechurch, & Worland, 2001). The world's most widely grown crop Triticum aestivum has spike architecture of flowering, and spike architecture determines the ability to produce seed. The sensitivity of the photoperiod is determined by the main PpdA, PpdB and Ppd1 genes located on chromosome 2A, 2B and 2D (Laurie, 1997). The gene Ppd1 is called pseudo-response regulator (PRR) that contain CCT domain which is similar in structure and function to Arabidopsis PRR7 (Turner, Beales, Faure, Dunford, & Laurie, 2005). Regulation of the Ppd1 gene is used to get more spikes to increase the number of seeds. Therefore a strategy used for improving the yield potential of wheat through modified the spikelets numbers and arrangement on inflorescence architecture.

1.2. Genes involved in regulation of photoperiod in wheat:

Theoretically the gene Photoperiod Dependent (Ppd-1) regulates spikelet development in wheat. Ppd-1 is a known floral gene which regulates photoperiod, or day-long floral induction (Beales, Turner, Griffiths, Snape, & Laurie, 2007), (Shaw, Turner, & Laurie, 2012). They noticed that Ppd-1 also controls spike architecture and paired-spikelet

production in wheat, in addition to floral induction. Alternatively, Ppd-1's modulation of paired-spikelet growth is likely to occur via an FT-dependent pathway, which mediates the expression of meristem identification genes along with shaping the spike production. In addition, reduction of ZMM8-a MADS-box gene-function in transgenic maize plants has resulted in an increase in florets per spikelet (Münster et al., 1997). It is of prime importance to get early maturing varieties to improve the yield efficiency. In this way the MADS-box gene offers a wealth of information. Along with FT, other genes that can be exploited are AGL24 and MAFs etc. AGL24 belonging to MADS-box family is regulator of AP1 and promote flowering in Arabidopsis and other plants (Becker & Theißen, 2003). Result of this study indicate that growth characterises were improved in transgenic wheat plant constitutively expressing the barley HVA1 gene in response to soil water deficit. Results of the research conducted by 'Sivamani et al', shows that there were enhanced growth characteristics of transgenic plant of wheat which constitutively expressed the HVA1 barley gene in response to water deficient conditions (Sivamani et al., 2000). Enhancing the yield of wheat grain per hectare is crucial for breeders to cope with global food scarcity, rapidly growing populations and dwindling arable land in the coming years, although many obstacles remain ahead due to the complexities of this trait. Target shooting for genes responsible for higher C catch, particularly under stressful conditions or otherwise, biomass increase and harvest index optimization, in tribe Triticeae may be of considerable importance for efforts to increase unconventional yield.

1.3. Spikelet architecture in wheat:

The architecture of the wheat inflorescence (spike) is relatively simple compared with the branched inflorescences of rice (panicle) and maize (tassel and ear). It is characteristically unbranched and forms single spikelets on opposite sides of the central rachis in an alternating phyllotaxy, with a single terminal spikelet at the distal end of the spike. The spikelets of the wheat spike are indeterminate and typically produce three or four fertile florets, in contrast to the determinate spikelets of rice and maize that bear only one or two florets. (Sharman, 1944). Although the inflorescence of archetypl wheat is a relatively simple structure and variety of variants have been reported however, the

molecular mechanisms are almost unknown. Generally speaking, there are two varieties of enhanced yield wheat, one variety gives more yield through more spikelets and other variety gives more yield through grains per inflorescence. In first, spikelets form as usual on the main inflorescence, but one and sometimes more basal spikelets are replaced by long lateral branches, each forming its own spikelets and florets. In second, inflorescence branching involves a number of changes that may not include the development of two or more spikelets at a single node without a long branch being created (Sharman, 1944). Paired spikelets are distinguished by the presence of a second spikelet located close to and directly below a typical single spikelet (abaxial), analogous to the primary and secondary spikelets that make up the 'spikelet pair ' in maize and other Andropogoneae members. The second spikelet occurs in various forms, varying from simple systems not containing grain to full spikelets surrounding a variety of grains (Yen & Yang, 1992).

1.4. CRISPR/Cas9 genomic editing system in plants:

Transformation of usable genes happens in various labs but most of the time it requires either knocking down or releasing a single gene for disease resistance or development of a specific trait phenotype. But problem remained to have modification of a gene having combination of all valuable characteristics or a set of genes that alter the plant design radically. Working in the genome of plants to produce more resistant varieties is always a problem. To date, many method have been used, such as EMS mutagenesis and t-DNA insertion to generate random mutations; however, this method does not provide a solution for processing the target gene (Belhaj, Chaparro-Garcia, Kamoun, Patron, & Nekrasov, 2015). Technological advances caused the invention and version of genome modifying techniques, which includes zinc finger nucleases (ZFNs) and TAL effector nucleases (TALEN) that permit you to precisely edit interesting genes. However the development and production of genetic editing using this technology is problematic and costly, since protein engineering is required to edit the desired gene.(Gaj, Gersbach, & Barbas III, 2013);(Bortesi & Fischer, 2015). CRISPR / CAS9 (CRISPR associated Protein9) was discovered not long ago, which gained importance in no time as a precise genome editing tool and is user friendly and cost efficient in

comparison to other tools, and it has bright prospects to revolutionize the genome editing resulting in plant improvement. CRISPR/CAS9 system - versatile genome editing tool - can be used to edit multiple on the basis of small RNA's guidance (Doudna & Charpentier, 2014). It is the main component of adaptive, bacterial and archaeal, immunity, originally discovered in *Escherichia coli* in the 1980s (Ishino, Shinagawa, Makino, Amemura, & Nakata, 1987). Working of the CRISPR/CAS9 system was not known and not understood until it was determined that a bacteria called Streptococcus *thermophilus* is resistant to bacteriophage by way of incorporating a fragment of the infectious virus genome in its CRISPR locus (Barrangou et al., 2007). In CRISPR / CAS9 system the destruction of the target DNA depends on double stranded breaks guided by the cRNA transcript. (Schiml & Puchta, 2016). There are two repair mechanisms for DNA which may be activated after the DNA is damaged and those two types of mechanism are (i) NHEJ (Non-Homologous End Joining) and (ii) HDR (Homolog Directed Repair). The 'non-homologous end joining mechanism' is susceptible to errors and may not give a perfect repair but result in erroneous repairs and also interruption in gene function may result from it. However, in homologous direct repair mechanism, a perfectly repaired, new DNA is created in process. Schiml, Fauser, & Puchta, 2014); (Belhaj et al., 2015). The Non-homologous end joining regeneration tool can also produce non-specific alterations likely due to the off-target effects of engineered sgRNA in the further regions of the genome. To rise the specificity and reduce the targeting of target mutagenesis to RNA, the protein engineering method is used to transform the CAS9 nucleases and CAS9 mutant. For example, CAS9d10a (CAS9 nickase variant) is used for more accurate genomic processing (Ran et al., 2013). There are several advantages to processing the genome with the CRISPR/CAS9 system, but there are some risks that can be applied to different plant species. For example, the polypolied nature of different types of crops increases the likelihood of mutations outside the target population and reduces the specificity of genomic tissue (Peng, Lin, & Li, 2016). In addition, to edit each copy of a gene in a genome is a problem that is disputatious, especially if you have a multiple copies of a gene in a different genomic locations, is a problem. Wheat is hexaploid in nature that's the reason it makes plants a significant model for research or improving genomic editing systems (Ling et al., 2013);(Choulet et al., 2014).

1.5. CRISPR/Cas9 targeted genome of Wheat:

CRISPR / Cas9 genome editing technology allows altering 4 genes in a wheat plant at a time. Recent progress in genome editing has permitted the simultaneous silence or activation of multiple genes**.** That was made possible by the use of the CRISPR / Cas9 system's genome editing tools. The gRNA (guide RNA) is a short synthetic RNA consisting of a "scaffold" sequence needed for Cas9-binding and a user-defined "spacer" or "targeting" sequence (Protospacer Adjacent Motif, PAM) specifying the changed genomic target. Target genes can be modified very conveniently in the wheat genome due to the availability of the latest breakthrough protocol of Agrobacterium induced transformation in monocots. In turn the editing of genomes is emancipated from regulatory approvals (Lowe et al., 2016). CRISPR / Cas9 is a selective genome editing method that accelerates modern breeding by quickly and reliably and predictably altering the genome (Bortesi & Fischer, 2015). Contrary to other methods the benefit of CRISPR / Cas9 is the simplicity of multiplexing. The concurrent addition of double strand breaks (DSBs) at multiple sites can be used concurrently to modify multiple genes (Li, Zhang, & Sheen, 2014), and can be useful for knockdown or incorporation of multiple genes into a system such as folate synthesis and Zink in wheat. Several gRNAs targeting different promoters allow the creation of functional cascaded circuits through simultaneous inducible control of different genes based on either transcriptional activator or repressor (Nissim, Perli, Fridkin, Perez-Pinera, & Lu, 2014). The only attempt on silencing the gene IPK1 using ZFNs genome editing system has been made in maize (Shukla et al., 2009). Nevertheless, genome editing was attempted in wheat utilizing other genes but transiently (Zhang, Yang, Yang, Li, & Guo, 2016). Interestingly, the annotated genome sequence of wheat is available at IWGC (https:/www.wheatgenome.org/) and Ensemble genome browser.

1.6. Objectives

CRISPR-Cas9-mediated editing of local wheat with Ppd1 gene is aimed to improve spike-architecture. Availability of mutant plants of wheat having altered spikelet number and reduced flowering time for research purpose.

MATERIALS AND METHODS

2.1. Selection of genes

Three homologs of (*Ppd-, A1*, *-B1* and *D1*) and three homologs of (*FT-A1, B1 and D1)* in hexaploid common wheat (*Triticum aestivum* L.) of two genes were selected. Photoperiod-1 (Ppd-1), a pseudo-response regulator gene that controls photoperiod-dependent floral induction, has a major inhibitory effect on paired spikelet formation by regulating the expression of FLOWERING LOCUS (FT) according to (Boden et al., 2015).In this study we modulated expression of the two important flowering genes, *Ppd-1* and *FT*, which can be used to form a wheat inflorescence with a more elaborate arrangement and increased number of grain producing spikelets. Modifying the number and arrangement of spikelets on the wheat inflorescence could therefore be a strategy to improve yield potential in this important crop.

2.2. Sequences retrieval of Ppd1 and FT

The annotated Genome sequence of wheat is available at IWGC (https://www.wheatgenome.org/) and **Ensemble Plant Genome Browser**. Sequences of three homologs of (*Ppd-, A1*, *-B1* and *D1*) and three homologs (*FT-A1, B1 and D1)* of hexaploid common wheat (*Triticum aestivum* L.) were retrieved from **Ensemble Plant Genome Browser**. Homologs of (*Ppd-1 ,A1*, *-B1* and *D1*)gene located on group two chromosomes such as 2A, 2B, 2D respectively and consist of 8 coding exons, 9 introns and coding sequences contains 2013base pairs. Similarly, homologs of (*FT-A1, B1 and D1)* gene located on group seven chromosomes such as 7A, 7B, 7D, respectively and consist of 3 coding exons, 2 introns and coding sequences contains 534 base pairs. Alignment of both these genes was checked by **BioEdit** software (See Appendix A and B).

2.3. Design and selecting candidate CRISPER/Cas9 specific target site in wheat genome

Several web-based tools available for designing single guide RNAs such as CGAT, CHOPCHOP, and E-CRISP but due to unavailability of wheat genome in these

software, target sites (sgRNA) of 21 nucleotides were selected manually. We used exons of these genes for target sites selection. The potential off-target sites were detected by using BLAST in Ensemble Genome Browser.

2.4. Design and synthesize specific sgRNA oligo

The sequence of a targeted gene (PPD1, FT) of Triticum aestivum, Oryza sativa and Arabidopsis thaliana were retrieved by using ENSEMBL Plant Genome Browser. The BLAST of the retrieved sequence of both gene was performed using online ENSEMBL BLAST by keeping the default parameters. The multiple and pairwise alignment of the obtained sequences of both genes of Triticum aestivum was performed against the PPD1and FT gene. For the synthesis of 'Guide RNA', the whole genome of a specie is required. CGAT & CHOPCHOP design tools are used for the synthesis of guide RNA. These tools are capable of recognizing the possible target sites in the genome of different species. As the whole genome of wheat is partially sequenced, we designed the guide RNA (both forward and reverse) manually without using these design tools. The synthesized guide RNA were then ligated between BtgZl and Bsal restriction sites. The melting temperature and GC content of guide RNA were also checked before guide RNA synthesis. The specificity of the guide RNA was also checked by using Ensembl Plant BLAST. Nine guide RNAs were designed using the above mentioned approach. The list of the synthesized guide RNAs is attached herewith.

To design the sgRNA oligos, two complementary oligonucleotides (21 - 25 ntds) with sense Strand consisting of 4-nt overhang of TGTT at the 5' end and antisense strand consisting of extra AAAC at the 5' end for BtgZI restriction site were synthesized. Similarly, another two complementary oligonucleotides were designed with sense strand containing a 4-nt overhang of GTGT at the 5' end and antisense strand containing AAAC at the 5' end for BsaI restriction site. Finally, gRNAs were synthesized after checking different design parameters including target length, minimum percentage of GC content and melting temperature.

2.5. Construct CRISPR/Cas9 expression plasmid

To prepare the gRNA we need a genome of interest in which we target a specific site, for this procedure we used two oligonucleotides that are complementary to each other. After this we

annealed these oligos as well as ligated into the digested site called BtgZI that is located in vector pENTR4-gRNA1.

2.5.1. Annealing of the synthesized Oligos and cloning into BtgZI digested vector pENTR4-gRNA1

Restriction digestion of 1μg of vector with BtgZl was conducted. Run an agarose gel and cut out the band containing vector. The amount of 1ug PENTR4, GRNA1 vector was used for digestion in an incubator. Then PCR thermocycler was used for amplification. For digestion process 1μg pENTR4-gRNA1 is needed. BtgZI restriction site and Rnase A were added in the solution followed by addition of Cutsmart buffer. Further, water was added to the solution. We used Thermo-scientific gene get kit for purification process for digestion of pENTR4-gRNA1. The amounts of chemicals used in this reaction are listed in (table 5).

2.5.2. Annealing and phosphorylation of oligos guide RNA

Two oligos were dissolved in a duplex buffer (PNK, T4 PNK) respectively and then adenosine triphosphate was mixed in it. For oligos hybridization some salt was also added. Water in the beaker was heated until it boiled then water was cooled at room temperature. For significant results, oligos were placed in a water bath. The product is stable, double stranded and stored at 4 $℃$. After this the (dsOligo) double stranded fragment oligo cloned in the site of BtgZ1 digested spot present in scaffold vector of sgRNA. Further, these oligos were joined by ligation process and 5-phosphate was added after the PNK enzymatic phosphorylation reaction.

2.5.3. Ligation of annealed and phosphorylated oligos

In Eppendrof tube, BtgZI Digested pENTR4-gRNA1 (50 *ng*) were combined and then T4 ligase was added with ligation buffer. Ligation reaction was arranged and further, Phosphorylated and annealed dsOligo were also mixed in it. As the ligation reaction requires incubation at room temperature/25 $℃$ for 3 hours or overnight at 4 $℃$ and therefore, the reaction was carried out in such manner. After this transferred partial reaction into the rapidly growing bacteria called E. coli and competent XL-Blue cells.

2.5.4. Confirmation of positive clone

For identification of inserted clone, we need to digest DNA plasmid with digestion site called BamH1. The procedure of cloning is successful when the oligos are cloned into the vector called pENTR4-gRNA1. In result of this, the BamH1 spot at MCS disappeared. No 350bp pattern band was detected at the spot of digested BamH1 clone that indicated the possible positive clones. There are other method like colony-PCR create test and reverse primer technique, the former working via gRNA-F oligos for seed sequence (sgRNA), and the latter one using U6TR - reverse primer extracted through backbone of vector called pENTR-gRNA1, also indicate positive clone. After cloning, the sequencing can be done of candidate positive clone through using both primer forward and reverse for the confirmation of accurate seed sequence in the scaffold gRNA vector. The pgRNA1 name was assigned to this construct. Further, when the first double stranded oligo is inserted into the restriction site called BtgZI, the second double stranded oligo is ligated sequentially at restriction site Bsal.

2.6. Cloning a specific second target sequence of sgRNA

Digestion of sgRNA scaffold vector was performed that contained the first guide RNA then we target a second specific sgRNA sequence with the insertion of bsal restriction site. $1\mu g$ pgRNA1 vector is required for digestion in an incubator. Then PCR thermocycler was used for amplification process with Bsal for 3 hours at 37 $℃$ further to it, for digestion process we need pgRNA1 and addition of Bsal restriction site. Then addition of Ranse in the solution followed by adding Cutsmart buffer was performed. The H2O was added in the solution. Thermo-scientific gene get kit was used for purification process. The amounts of substances used in cloning process are listed in the (table 6)

2.6.1. Annealing and phosphorylation of oligos containing first guide RNA

The following things were combined in Eppendrof tube, $2\mu l$ Bsal Digested pgRNA1 (50 ng), and then T4 ligase with ligation buffer was added. Then set up for ligation reaction was arranged and then $5\mu l$ Phosphorylated and annealed dsOligo were added to it. As the ligation reaction requires incubation at room temperature/25 $℃$ for 3 hours or overnight at 4 $℃$ and therefore, the reaction was carried out in such manner. After this, transferred the partial reaction into the rapidly growing bacteria called E. coli and competent XL-Blue cells.

2.6.2. Confirmation of positive clone

For identification of inserted clone, we need to digest DNA plasmid with digestion site called BamH1and Bsal. The procedure of cloning is successful when the oligos are cloned into the vector called pENTR4-gRNA2. In result, the Bsal restriction spot at MCS disappeared. There was no 790bp pattern band detected at the spot of digested Bsal clone that indicated the possible positive clones. After cloning, the sequencing can be done of candidate positive clone through using both primer forward (U6PF1-b) and reverse (pENTR4-R) for the confirmation of accurate seed sequence in the both scaffold gRNA vector.

2.7. Construction of plant expression vectors

Binary Vector pOs-Cas9 was provided via Department of Genetics by Dr. Bing Yang associated with department of Development and Cell biology, at university of Iowa State Crop Bioengineering Consortium. To construct the expression vector the intermediate vector pgRNA1 was used, which is sequentially digested with BtgZI and BsaI restriction enzymes for two insertions of double-stranded oligonucleotides. The sub cloning results in two gRNA expression cassettes. The insertions was confirmed through sequencing with the primer U6.1-F. The cassettes flanked by the Gateway recombination sequences attL1 and attL2 were mobilized to the binary vector pOs-Cas9 through the Gateway recombination, which resulted in a single plasmid of Cas9/gRNA for *Agrobacterium*-mediated gene transfer into wheat callus cells.

1. pOs-Cas9 (Fig.8)
2. pETNR4-gRNA1 (Fig.8)
3. pENTR4-gRNA2 (Fig.8)

2.8. Golden gateway cloning

Golden Gateway Cloning technique was also used and by using this technique, the sgRNA was set under *Arabidopsis thaliana* promoter (U6) in Golden Gate vector at Level 1. Further, component selected marker of NPTII Inplanta, bCas9 or sgRNA together in Golden Gate vector at level 2 were assembled.

2.9. Transformation through CRISPR construct into wheat callus cells

The CRISPR/Cas9 system was introduced into LBA4404 strain of *Agrobacterium tumefaciens*. The agrobacterium mediated transformation in wheat is done by Inplanta transformation. For this method we use Agrobacterium strain LBA4404 with the wheat expression plasmid by electroporation.

2.9.1. Plant material

One wheat cultivar galaxy-2013 was grown at the National Institute of Genomics and Advanced Biotechnology, NARC. The seeds were grown in three different suitable conditions in the glass house, pots and field. We use this cultivar as an explant source for Inplanta transformation.

2.9.2. Tissue culture

Sterilization

Distilled water was used to clean glassware and autoclaved at 121o C for 20 minute to avoid contaminations.

Medium preparation

MS (Murashige & Skoog, 1962) medium for callus induction and regeneration was used. It contained different hormones. In this, 3% sucrose was added as a carbon source and ph. level was maintained at 5.8, further, to solidify the media, 6 *g/l* agar was added and maintained the ph. level of media. Then medium was autoclaved and poured into required already autoclaved glassware.

Callus induction medium

The MS media contain synthetic auxin2, 4-d with different concentration for induction of callus.

Sterilization of seed

Firstly, rinsed the seed through water 3-4 times, then immersed in 40% Clorox in 3 replications for 5-5 minute, which made it total 15 minute and then washed with distilled water till the

water after washing became clear. Seed were soaked in autoclave distilled water for 24 hours so that the embryo swell and increase in size. Turgid seeds were then plotted on a moisture filter paper in petri dishes and embryo were gently separated and amount per cultivar was transformed to sterilized MS medium.

Inoculation

In laminar flow, sterilised seeds were dispersed in callus induction medium for inoculation. Then inoculated seeds were placed at 25 ℃ with 16 hour light and 8 hour darkness photoperiod.

Regeneration medium

As embryonic and non-embryonic calli was obtained on CIM medium having 2, 4-d which used as a maintaince media for cultivar. Then flourishing calli were cut into small pieces to increase cell and shifted to regeneration media (RM). The RM contained different concentration of IAA and kinetin.

Plasmid and bacterial strain

The binary vector pOs-Cas9 is resistant to kanamycin, so it is not appropriate to use kanamycin-resistant strains of Agrobacterium tumefaciens (e.g., EHA101). Hence, Agrobacterium strain LBA4404 was used. Further, electro competent cell of Agrobacterium strain LBA4404 was prepared by following procedure. Took solid LB media and poured onto disposable plate. Streaking of E.coli on the plate was done by obtaining E.coli strain from glycerol stock and the put in incubator at 37 ℃. Then LB media was prepared and autoclaved, agar is required to be added in case of solid media.

2.9.3. Competent cell preparation:

Picked the colony from plate through tip and poured into flask, which contained 10 ml LB media liquid and then placed on shelter at 37 ℃ for culture growth. Then after 24 hour, growth appeared as solution became turbid. Again the streaking was done on plate left and placed in incubator at 37 ℃ without shaking, then after 24 hours colony appeared again and was stored at 4 ℃. Growth culture was also placed at 4 ℃ in refrigerator. Then growth cultures was added to 500 *ml* bottle of LB media and put in shaker at 37 ℃ for 3-4 hours.

LB media containing growth was poured in the six flasks, 50*ml* each and then all the six (6) flasks were placed in ice box. Then all the flasks were placed in centrifuge machine for 10 minute at 4℃ and at 8500 rpm then discarded the supernatant and then added the remaining amount of LB media in flasks. Again all the flasks were placed in centrifuge machine at the same previous calibration and discarded the supernatant again. Pallet was obtained in the flasks and then ddh20 was added in each flask and placed in centrifuge machine for 10 minute again to dissolve all the pallet. Then pour all the dissolved pallet in ddh20, in one flask and again centrifuged it for 10 minutes. Then the dissolved pallet was washed thrice after being placed in falcon tube and then 7% DMSO500 was added to it. Then almost 50μl pallet, obtained through the process explained above, was put in each eppendrof tube and stored at -80℃. These E.coli competent cells be used for integration of plasmid through electroporation.

2.9.4. Electroporation of Agrobacterium strain

First, we took LB media in two Eppendrof tubes and added 6 μl vector and poured into competent cell then dissolved fully and mixed gently by using pippete and then took all the media by tip and poured into center of cuvette. Then added all LB media and poured into the cuvette on its upper portion, then dissolved all this media gently and put into the electroporator and then take this cuvette and discard the solution. Then used the eppendrof tube that contained the competent cell, put this tube into thermo blog at 37℃ for1 hour. After incubation, the E.coli plates were used for plasmid multiplication. The plasmid extraction was carried out by using miniprep protocol (Bimboim & Doly, 1979) and then transformed into Agrobacterium strain LBA4404. Now the Agrobacterium is ready for wheat transformation.

2.9.5. Plasmid extraction

To begin with, full growth media and bacterial culture was poured in Eppendrof tubes for plasmid extraction, then centrifuged for 6 minute at 14000 *rpm* and discarded the supernatant to take the pallet. Then poured turgid culture in it and again centrifuged for 6 minute at 14000 *rpm.* For getting more pallet, absolute ethanol 100ml was added in wash buffer / ethanol buffer then added 200μl FADPL and mixed the pallet with separate tips in all Eppendrof tubes and added 2nd FADPH2 and gently mixed and then placed for 2 minute at room temperature. Then added 300μl FADPH3 and it became inert immediately. It was gently mixed and precipitation occurred. Further, it was centrifuged for 5 minute at 14000 *rpm* then poured supernatant into

purification column and then centrifuged again for 30s at 14000 *rpm.* Flow was discarded and 400μl wash buffer was added to it. Again it was centrifuged for 30s again and discarded the flow, added 600 μl wash buffer 2 and centrifuged for 30s and discarded the flow. After 3 minute then removed the collection tube from centrifuge machine and inserted purification column in new eppendrof tube and added 50μl elution buffer and leave for 2 minutes at room temperature then again centrifuged for 1 minute and removed purification column and took the extracted plasmid and run on gel .

2.10. Transformation with recombinant plasmid into wheat callus cells

The seeds were rinsed through water 3 to 4 times then immersed in 40% Clorox in 3 replications for 5-5 minute, a total of 15 minute. Then seeds were washed with distilled water till the water used for washing became completely clear. After the washing, seeds were soaked in autoclave distilled water for 24 hours so that the embryo swells and increase in size. Turgid seeds were then plotted on a moisture filter paper in petri dishes and embryos were gently separated and amount per cultivar was transformed on a sterilized medium MSD and incubated at 30 ℃ with continuous light for 14 days to start developing calli effectively. Then CRISPR / Cas9, having culture of agrobacterium size of 1.0 - 2.0 at OD600, was blended with 2ml agrobacterium cells and 20 ml medium called MSD – having 2 mm acetosyringone and pH 5.2 – was added to it. Then it was immersed in wheat calli suspension for 30 minutes and then wiped with filter paper to dry the calli, after it, calli was shifted in the medium MSD with AS plate and stored in dark for two to three days at room temperature. For the selection of transformed callus cells moved the infected calli to the selection medium (MSD supplemented with 50 mg/L of hygromycin B and 400 *mg/L* of Timentin) for 2 weeks. At this and corresponding stages, tissue culture and regeneration was performed at 30 ℃ in the growth chamber and under continuous light. To regenerate transgenic plants, single line of callus, which was resistant to hygromycin, was relocated on regeneration medium. For root activation, move the regenerated plants were moved to 1⁄2 MS medium. Finally, moved the plantlets to the soil and cultivated them in the growth chamber at a temperature of 28 ℃, total humidity of 75%, and a 12-hour photoperiod until the plants mature and seed was collected.

RESULTS

A detailed methodology has been introduced in Chapter 2. In this chapter, the results of this project are presented.

3.1. Selected target genes

The genomic sequence of wheat is available at IWGC website. The link of this website is (https://www.wheatgenome.org/) or **Ensemble Plant Genome Browser.** We retrieved the three homologous sequence of Ppd1 gene (A1, B1andD1) as well as three homologs of FT gene (A1, B1and D1) of common hexaploid *Triticum aestivum (*wheat) by using **Ensemble Plant Genome Browser** (https://plants.ensembl.org/index.html)**.** The three homologous genome of Ppd1 and FT gene are respectively located on the chromosome number two (2A, 2B and 2D) or chromosome number seven (7A, 7B and 7D). The Ppd1 gene consist of eight coding exon and nine introns or coding sequence contain2013 base pairs. Similarly the gene FT consist of three coding exons and two introns and the coding sequence contain 534 base pairs. Then we align the sequence of both gene by using default parameter by using **BioEdit** software (See Appendix1 and 2).

3.2. Designed sgRNA

THE gRNAs of both genes were synthesized after checking different design parameters including target length, minimum percentage of GC content and melting temperature. The two gRNA were selected from both genes to prepare a CRISPR/Cas9 Construct.

Here is a list of designed gRNAs.in (table1and 2).

Table 1. ***List of Ppd-1 gRNAs for Galaxy variety***

Sr. No	Primer NAME	Primer sequence 5'….3'
1	Pd1-gRNA-E2-F	TGTTTGGGAGGAGTTCCTCCACAGG
	Pd1-gRNA-E2-R	AAACCCTGTGGAGGAACTCCTCCCA
2	Pd1-gRNA-E2-F	TGTTCCGACGACTCCACCCGGCAGG
	Pd1-gRNA-E2-R	AAACCCTGCCGGGTGGAGTCGTCGG
3	Pd1-gRNA-E3-F	TGTTGACAGAGGTCTTCATGCACGG
	Pd1-gRNA-E3-R	AAACCCGTGCATGAAGACCTCTGTC
4	Pd1-gRNA-E4-F	TGTTACGAGCTTAAGAACCTTTGGG
	Pd1-gRNA-E4-R	AAACCCCAAAGGTTCTTAAGCTCGT

5	Pd1-gRNA-E5-F	TGTT AACGACGATGCCAGCATGGGG
	Pd1-gRNA-E5-R	AAACCCCCATGCTGGCATCGTCGTT
6	Pd1-gRNA-E6-F	TGTTATGCAGCAATAGATTAAGAGG
	Pd1-gRNA-E6-R	AAACCCTCTTAATCTATTGCTGCAT
7	Pd1-gRNA-E7-F	TGTTATACAGTTATGGAAAATTCGG
	Pd1-gRNA-E7-R	AAACCCGAATTTTCCATAACTGTAT

Table 2. ***List of FT gRNAs for Galaxy variety***

Sr. No	Primer NAME	Primer sequence 5'….3'
1	FT-gRNA-E1-F	TGTTACCCGCTGGTGGTTGGCAGGG
2	FT-gRNA-E1-R	AAACCCCTGCCAACCACCAGCGGGT
3	FT-gRNA-E1-F	TGTTGAACAGGACCGTGTCCAACGG
4	FT-gRNA-E1-R	AAACCCGTTGGACACGGTCCTGTTC
5	FT-gRNA-E2-F	TGTTCAAGCGATCCCAACCTTAGGG
6	FT-gRNA-E2-R	AAACCCCTAAGGTTGGGATCGCTTG
7	FT-gRNA-E3-F	TGTTAGATATCCCCGGTACAACTGG
8	FT-gRNA-E3-R	AAACCCAGTTGTACCGGGGATATCT
9	FT-gRNA-E3-F	TGTTGCCAGAACTTCAACACCAGGG
10	FT-gRNA-E3-R	AAACCCCTGGTGTTGAAGTTCTGGC

3.3. Oligos synthesized and clone into BtgZI digested vector

The $1\mu g$ vector pENTR4-gRNA1was digested with restriction site BtgZI at $60℃$ in incubator for 3 hours, listed below in (table 3)

Table 3. ***List of the chemical use for digestion.***

1.8 μl (1 μg)	pENTR4-gRNA1
1 μl	BtgZI (NEB)
0.2 μl	Rnase A
3 μl	10X Cutsmart Buffer
24 μl	H2O
30 μl	Total

3.4. Annealing and phosphorylation of oligos

The oligos were suspended in annealing buffer and mixed in equimolar concentrations. Mixed oligos were placed in a PCR tube. Then placed the tube in a thermocycler, programmed to start at 95 ℃for 2 minutes. Then, gradually cooled to 25 ℃over 45 minutes. List of chemical paste is in (table4).

Table 4.*The List of chemicals used in PCR reaction*

2 μl	Oligo 1 (10 pM)
2 μl	Oligo 2 (10 pM)
1 μl	ATP
2 μl	PNK buffer (Fisher)
1 μl	T4 PNK (Fisher)
X μl	H_2O
20 μl	Total

In a beaker, water was heated to the boiling point and then was left at room temperature for cool down.

3.5. Ligation of oligos contain BtgZl digested vector

We dilute annealed oligos with nuclease-free water and quantified the concentration. Then mixes the annealed oligos with cut vector. Chemical used in this method are listed in (table5)

Table 5. *List of chemicals used for ligation*

1 μl	H2O
1 μl	BtgZI Digested pENTR4-gRNA1 (50 ng)
5 μl	Phosphorylated and annealed dsOligo from step 3
2 μl	T4 ligase (NEB)
1 μl	10X ligation buffer
10 μl	Total

3.5.1. Confirmed positive cloned of pENTR4-gRNA1 vector

We successfully cloned the oligos into the vector pENTR4-gRNA1. The restriction site BamH1 spot at MCS was disappeared and on clone, there was no 350 bp band of digestion pattern with restriction site BamHI. Therefore, it was confirmed from non-detection of 350bp band that were positive potential clones, see in the (fig.1)

3.6. Cloned second specific target sgRNA sequence

Further, 1μg of pENTR4-gRNA2 vector was digested with Bsa1 for 3 hours at 37℃The chemicals used for this reaction are listed in (table 6).

Table 6. ***List of the chemical use for digestion.***

1.8 μl (1μg)	pENTR4-gRNA2
1 μl	Bsa1(NEB)
0.2 μl	RNAase
3 μl	10X Cutsmart Buffer
24 μl	H_2O
30 μl	Total

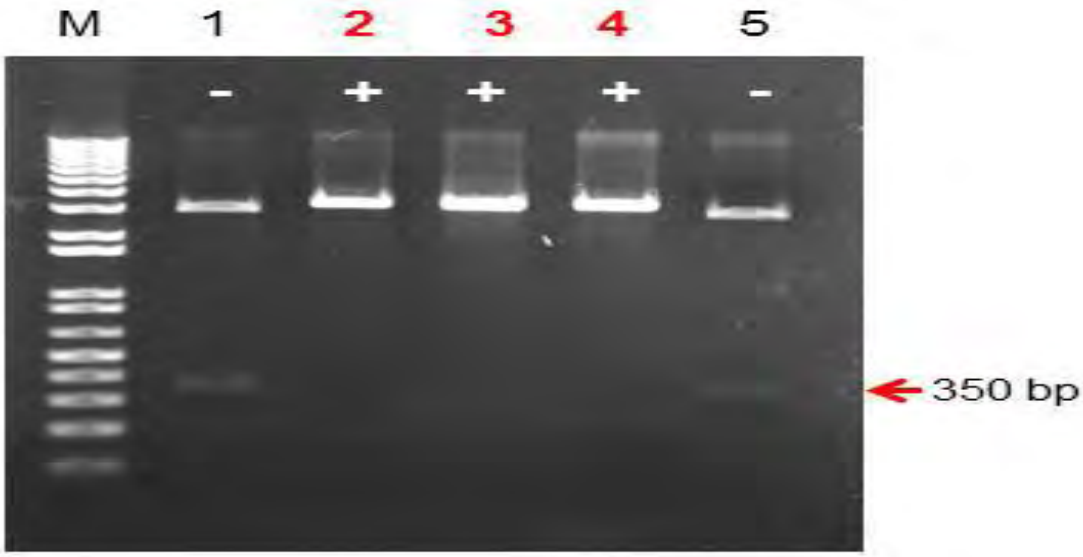


Figure 1: *Restriction and gel analysis of clones from the ligation of dsOligo at the BtgZI site of pENTR-gRNA1. The plasmid DNA of individual clone is digested with restriction enzyme BamH1 to confirm the right insertion of double-stranded oligonucleotides. The clones without the 350 bp fragment are putative positive clones.*

3.7. Accurate seed sequence in the gRNA scaffold

For the identification of positive clones, two type of tools can be used i.e. colony-PCR create test and reverse primer technique, the former working via gRNA-F oligos for seed sequence (sgRNA), and the latter one using U6TR - reverse primer extracted through backbone of vector called pENTR-gRNA1, also indicate positive clone. Therefore, both tools were used to see 560 bp positive clones and the result we can see in the (fig.2)

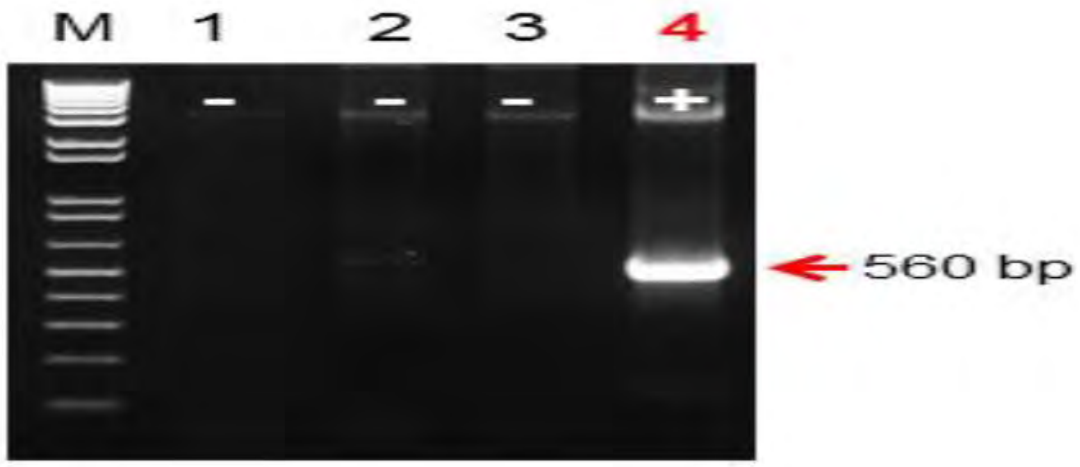


Figure 2: *Gel image of PCR reaction for identification of positive clone. The colony-PCR can be used to identify the positive clones by using gRNA-F and U6T-R primers. Clones with a detectable PCR product of 560 bp are putative positive clone.*

To confirm the accurate seed sequence in gRNA scaffold we used forward and reverse primer. Respectively, we use U6PF1-b as a forward primer and pENTR4-R as reverse primer. The pgRNA1 name was assigned to this construct. The second double stranded oligo is successfully ligated into the restriction site (Bsal), after the first double stranded oligo was inserted into restriction site (BtgZ1).

3.8. Ligation of oligos contain Bsal digested vector

We diluted, annealed oligos with nuclease-free water and quantified the concentration. Then mixed the annealed oligos with cut vector. List of chemicals used is in (table 7)

Table 7. ***List of chemicals used for ligation***

1 μl	H2O
1 μl	BsaI Digested pENTR4-gRNA2 (50 ng)
5 μl	Phosphorylated and annealed oligo duplex from step 3
2 μl	T4 ligase (NEB)
1 μl	10 ×Ligation buffer
10 μl	Total

3.8.1. Confirmed positive cloned pENTR4-gRNA2 vector

The double stranded oligos were successfully cloned into the vector, which was confirmed as the restriction site BamH1and BsaI spot at MCS was not appeared and the clone contained no 790 bp band of digestion pattern with restriction site BsaI. Therefore, the band is confirmed to contain positive potential clones, see in the (Fig.3).

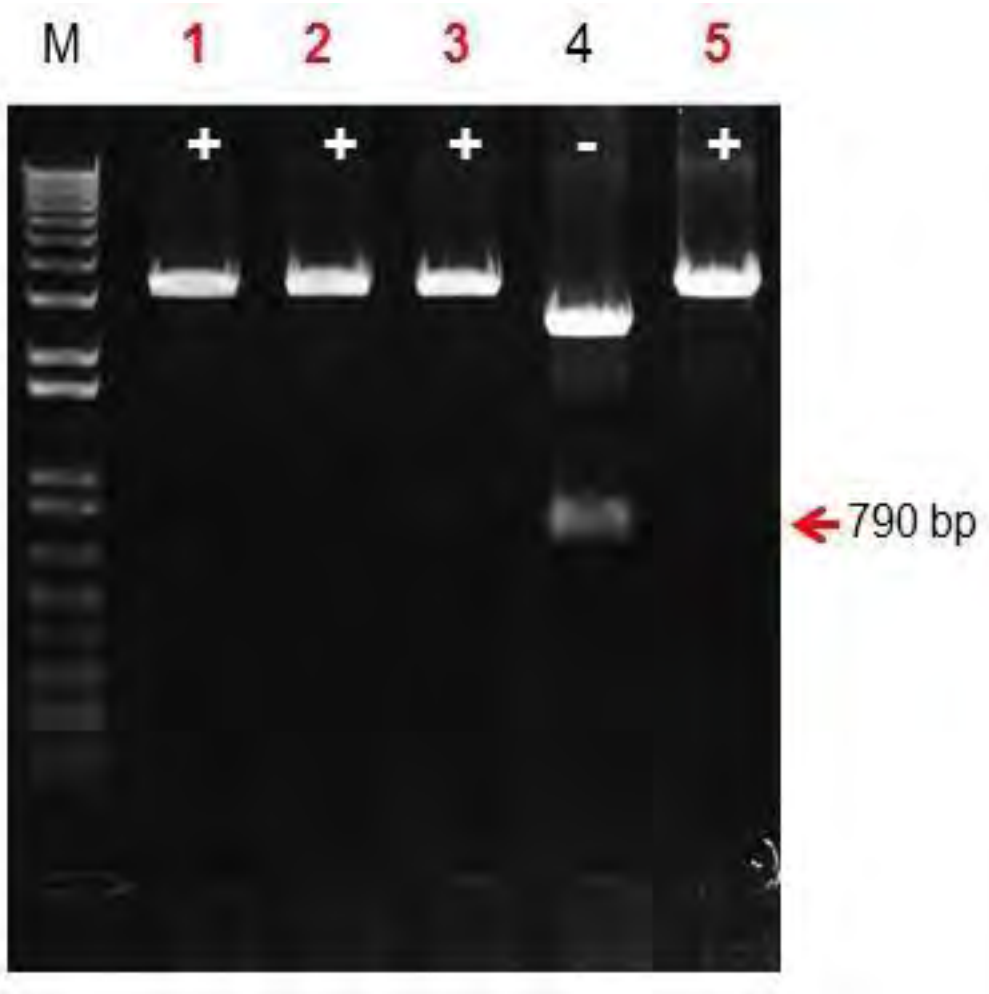


Figure 3: *Gel image of plasmids digested with BamHI and BsaI. The plasmid DNA can be digested by using two restriction enzymes BamH1 and Bsa1 to confirming the insertions of the second double-stranded oligonucleotides. The positive clones do not have any digested fragment of 790 bp below the 3kb vector band.*

3.9. In gRNA scaffold Accurate first and second seed sequence

When the cloning is done, for identification of positive clones two type of tools are used i.e. colony-PCR create test and reverse primer technique, the former working via gRNA-F oligos for seed sequence (sgRNA), and the latter one using U6TR - reverse primer extracted through backbone of vector called pENTR-gRNA1, also indicate positive clone. To confirm the accurate seed sequence in first and second seed sequence of both gRNA scaffolds, we used forward and reverse primer. Respectively, U6PF1-b was used as a forward primer and pENTR4-R as reverse primer. See in the (fig.4)

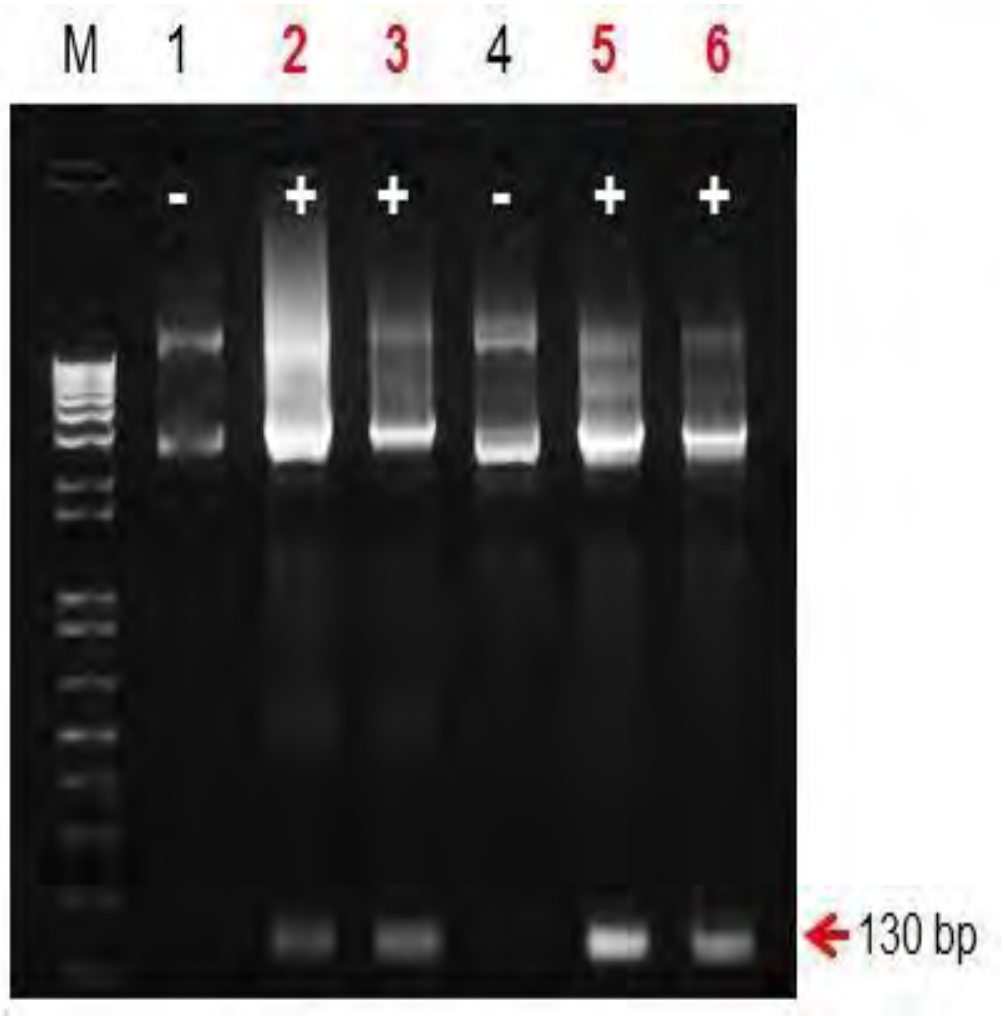


Figure 4: *Gel image of the colony-PCR products from the cloning second dsOligo. The PCR based assay can be used for confirming the positive clones by using gRNA-F2 and U6T-R primers. A PCR product around 130 bp can be detected from the putative positive clones.*

3.10. Construction of plant expression vectors

The intermediate vector was used to build up sgRNA for gene editing in wheat. The restriction enzyme AatII was linearized with intermediate vector. Linearized plasmid DNA entry can improve the efficiency of the recombination of Gateway LR. By using pENTR-gRNA1 and pENTR-gRNA2 individually the two gRNAs were constructed. As we cloned double stranded oligos into pENTR-gRNA1, likewise we cloned two double stranded oligos into the vector pENTR-gRNA2. The ingredients for vector construction see in the (table 8)

Table 8. ***The ingredients for vector construction***

Entry clone	50-150 *ng*
Destination vector	150 *ng/Ul*
TE buffer	X μl
Total	5 μl

3.11. The gRNA cassettes used for gene targeting

The pENTR4-gRNA1 gRNA cassette is sliced out using the HindIII. The gRNA cassette fragment of 1 kb is retrieved from the water, and cleaned. The retrieved fragment (1 kb) is sub-cloned into digested Hind III and subsequently CIP-treated pENTR4-gRNA2 which already contains two reference genes for RNA. We digest the DNA plasmid from the stage of cloning with the help of restriction enzyme HindIII to validate gRNA cassette insertion. By using Gateway LR Clonase (Thermo Fisher Scientific), up to 4 genes of gRNA gradually recombined in the final destination vector called pOs-Cas9 vector. The result we see in the (fig.5)

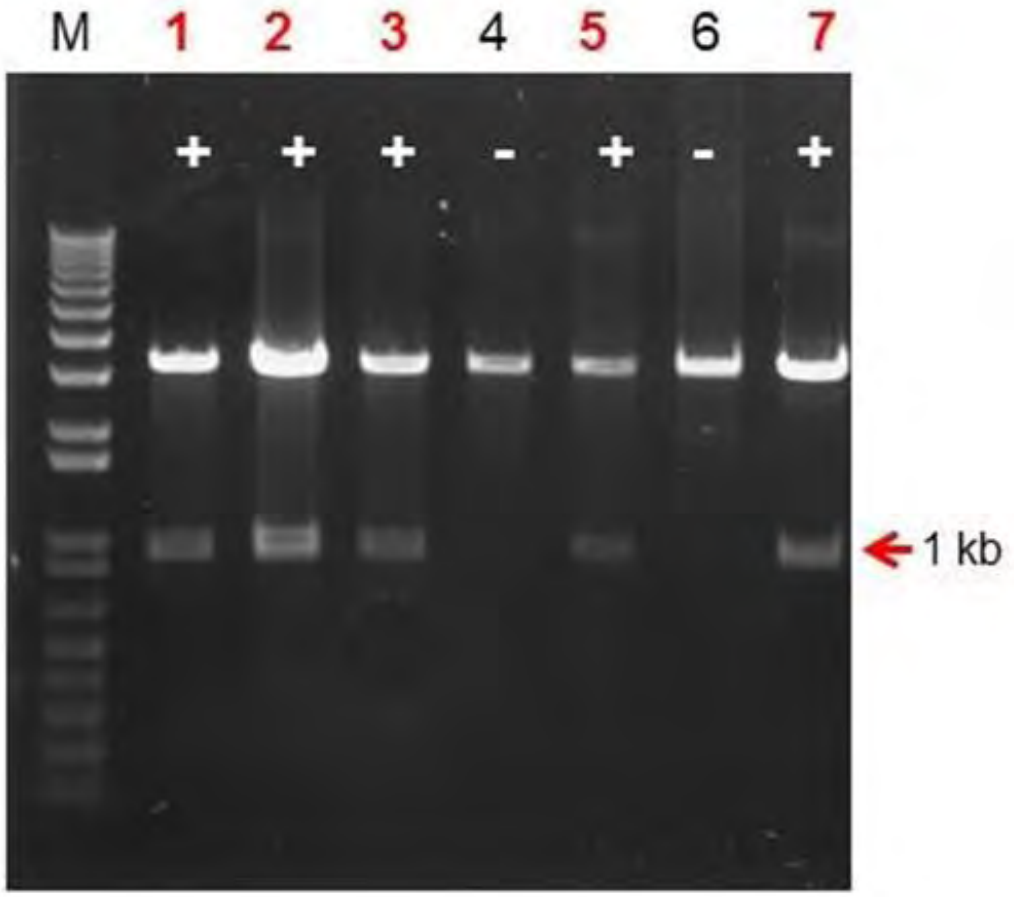


Figure 5: *Gel image of plasmids digested with HindIII. The plasmid DNA is digested with restriction enzyme HindIII to confirm the gRNA cassette from pgRNA1 is successfully subcloned into pgRNA2 that already has two guide RNA genes. The clones that contain the 1 kb fragment are positive ones.*

The DNA plasmid obtained through individual clones is digested through restriction enzyme BamH1or HindIII, and we get positive response by using gateway technique. It indicates that the intermediate vector has been successfully cloned into the final destination vector. We can see in the (fig.6)

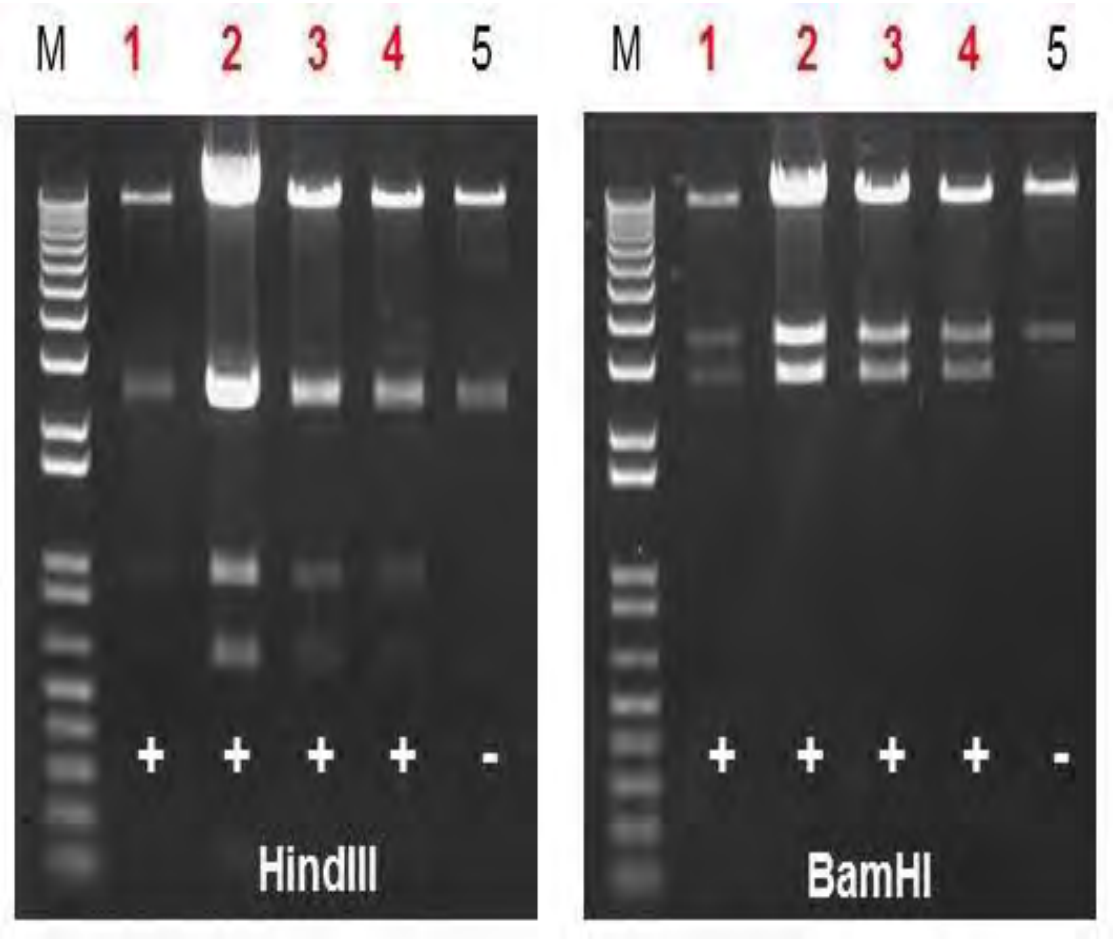
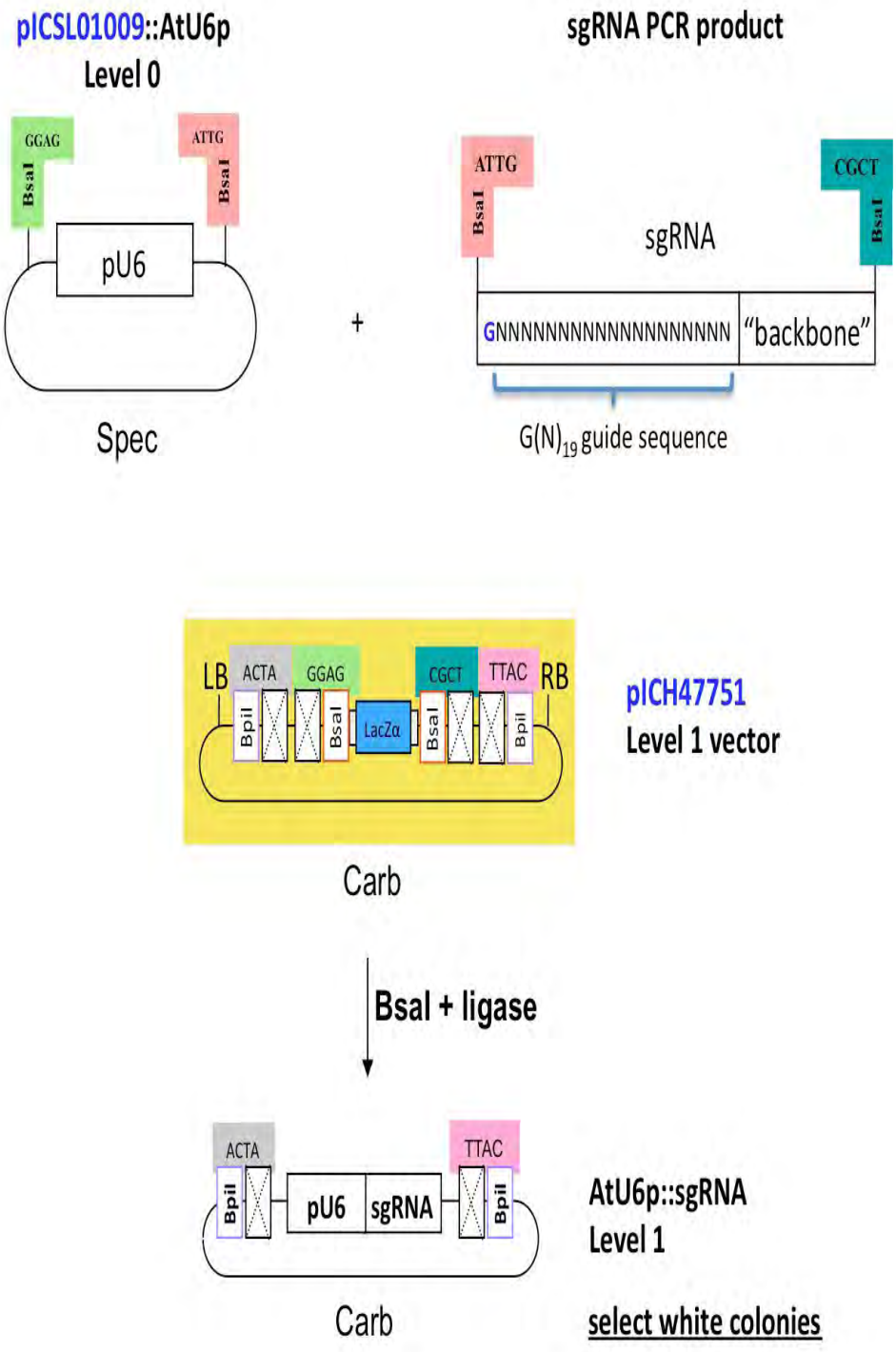


Figure 6: *Gel image of the Gateway LR recombination clones digested with HindIII and BamHI. The plasmid DNA is digested with restriction enzyme BamH1 (right gel picture) or HindIII (left gel picture). The clones with the digested pattern shown in the gel picture are the positive ones that are ready for Agrobacterium transformation.*

3.12. Cloning done through Gateway method

Both restriction sites sequences (attL1and attL2) located in the cassette flanked of gateway recombination are mobilized through Gateway recombination to the binary vector pOs-Cas9, resulting in a single Cas9/gRNA plasmid for translation of agrobacterium-mediated gene into wheat callus cells. The assembling of construct see in the (fig.7)


pICSL01009::AtU6p
Level 0
GGAG
BsaI
ATTG
BsaI
pU6
Spec
+
sgRNA PCR product
ATTG
BsaI
sgRNA
CGCT
BsaI
GNNNNNNNNNNNNNNNNNN
"backbone"
G(N)19 guide sequence
LB
ACTA
GGAG
CGCT
TTAC
RB
BpiI
BsaI
Lac2α
BsaI
BpiI
pICH47751
Level 1 vector
Carb
BsaI + ligase
ACTA
TTAC
BpiI
pU6
sgRNA
BpiI
AtU6p::sgRNA
Level 1
Carb
select white colonies

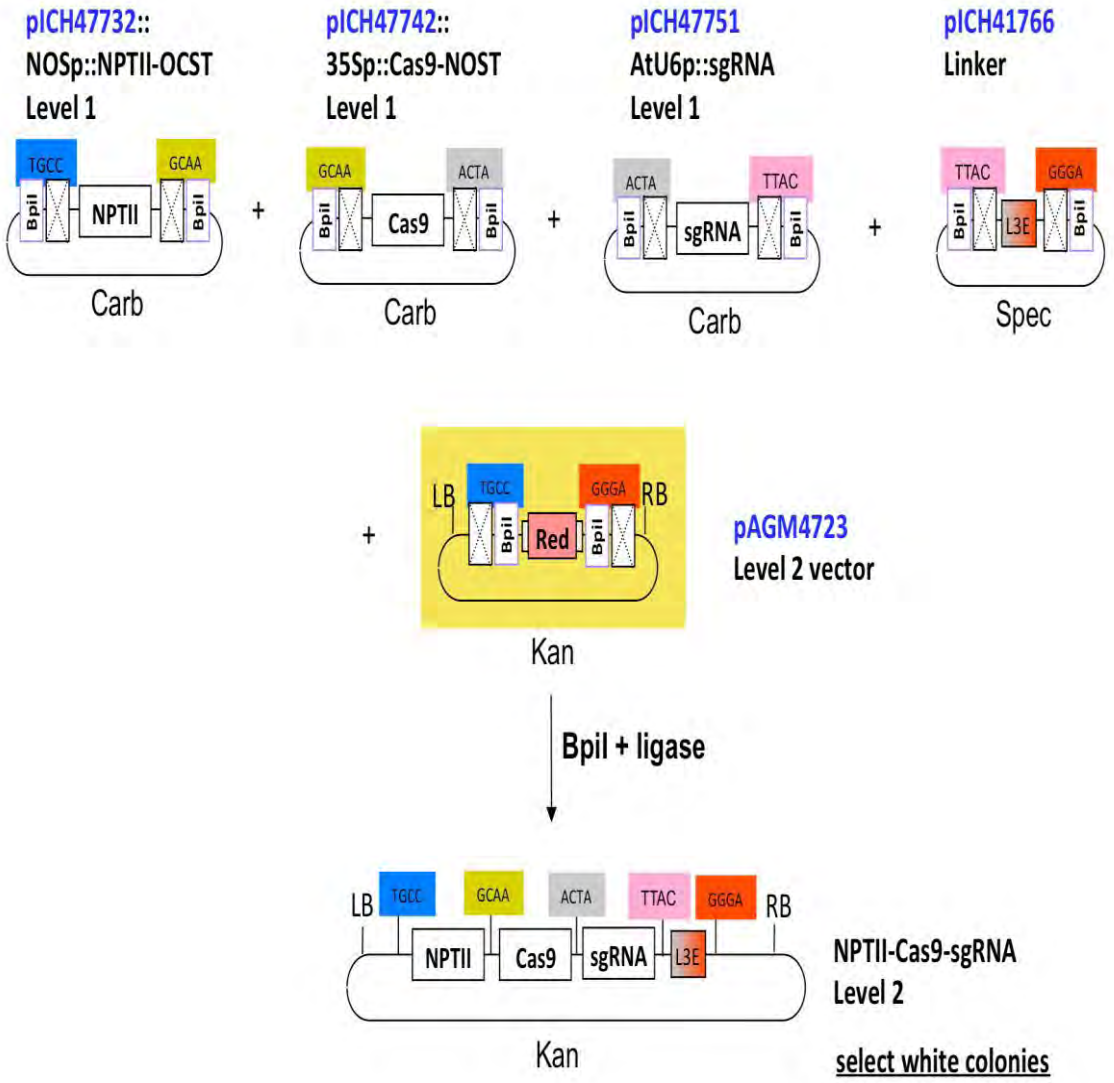


Figure 7: *Scheme illustrating the assembling of Cas9/sgRNA expressing constructs.*

The intermediate vector pgRNA1 is an enzyme that is digested sequentially with two restriction site BtgZI and BsaI by the insertion of two dsoligos. The subcloning contributes to two cassettes of the gRNA phrase. With the primer U6.1-F, the insertions can be verified by sequencing. Both restriction sites sequences (attL1and attL2) located in the cassette flanked of gateway recombination activated through the Gateway recombination to the binary vector pOs-Cas9, resulting in a single Cas9/gRNA plasmid for translation of agrobacterium-mediated gene into wheat callus cells see (fig.8)

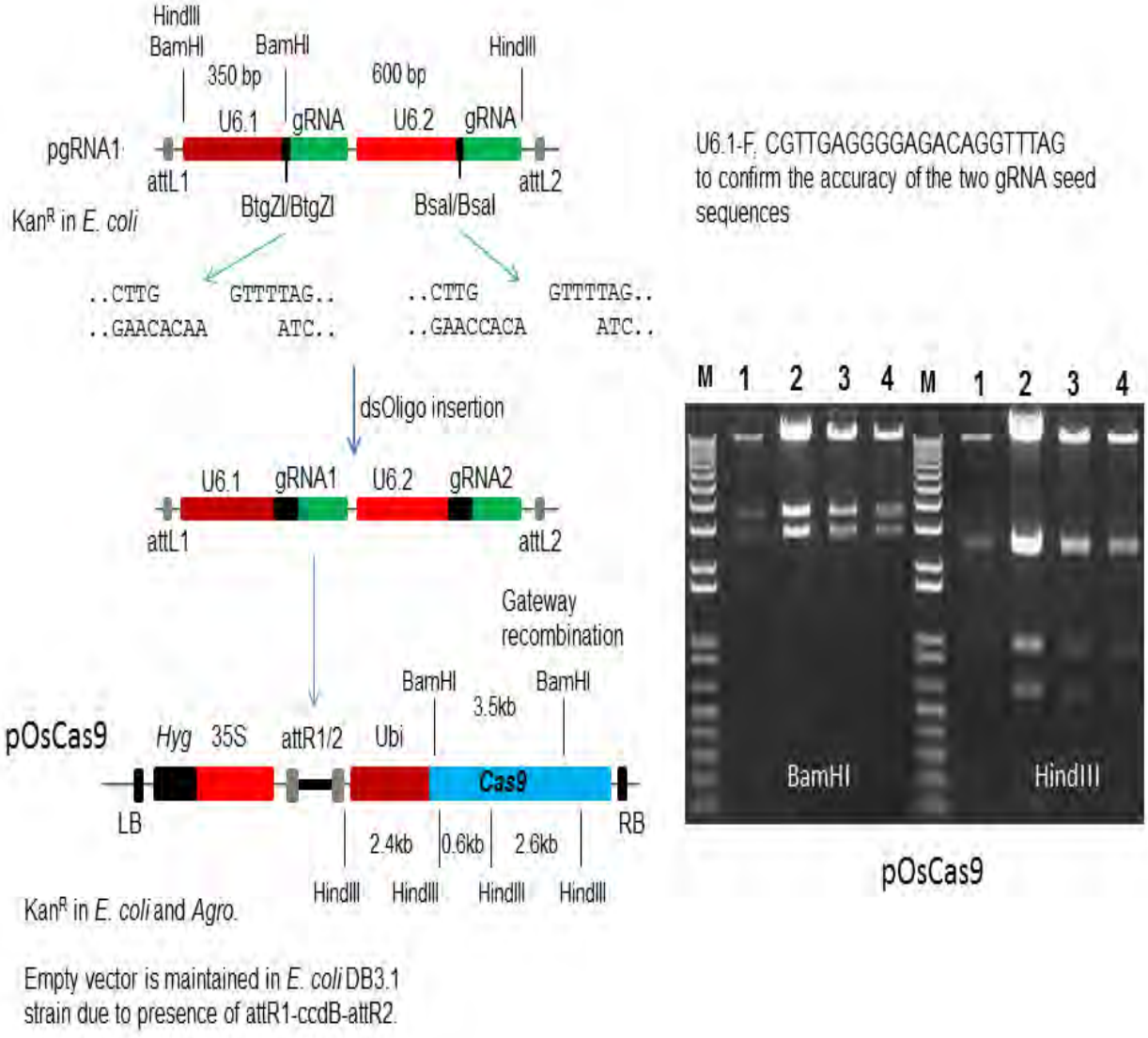


Figure 8: *Schematic diagram of wheat CRISPR/Cas9 system.*

3.13. Transformed recombinant plasmid into wheat callus cells

The CRISPR / Cas9 method has been inserted into the Agrobacterium tumefaciens strain LBA4404. Infection with Agrobacterium harbouring CRISPR / Cas9 system was attempted both in the callus and in transformation through inplanta. The MS (Murashige & Skoog, 1962) medium for shoot regeneration and selection of transformed shoots for plants edited by the genome included different plant growth regulator concentrations for effective plant material production see in the (fig.9)

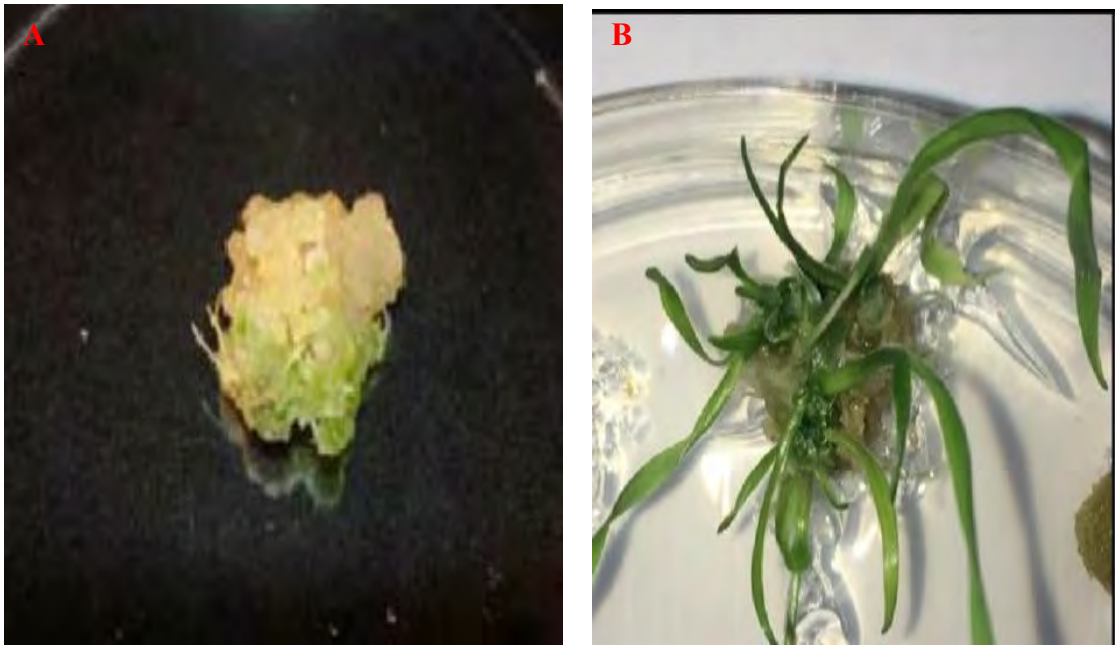


Figure 9: *Transformation of gRNA in wheat galaxy variety. A) Callus induction using Cas9 vector B) Plant regeneration from callus*

DISCUSSION

Highly conserved domains are thought to typically play a main role in performing the function of the corresponding protein. Hence, the altered functions are the result of Mutations which are located in the conserved domains. This is generally used to analyse the functions of genes that are of interest. Significant efforts were made to improve agrobacterium-mediated transformation of mature embryos and to establish direct agrobacterium-mediated transformation of germ line cells in plants to fire meristem, to develop inflorescences and immature embryos, however through this method no significant results were achieved. Alteration of numerous genetic and environmental dimensions of the process of transformation will result in high performance transformation. In this study, we have been able to effectively develop a novel approach to wheat transformation based on tissue culture through Inplanta procedure.

The Ppd-1 gene is the principal regulator of photoperiod (day length) responsiveness for flowering in temperate cereals. Expressions of a gene, where it is expressed, gives the idea that where it is functioning and with which gene it is interacting. Ppd-1 expressed itself mainly in the spikes with minor expression in leaves, apex, head and stems. The Ppd-1 action target i.e. FT, and the expression of FT is expected to be expressed in phloem companion cells, as this appears to be the only location of FT expression in Arabidopsis (Turck, Fornara, & Coupland, 2008). Due to the fact that Ppd-1 belongs to the PRR family, which is highly homologous to Arabidopsis PRR7, it was considered that Ppd-1 may operate in the circadian clock in a similar way (Nakamichi et al., 2007); (Shimada et al., 2009). In Arabidopsis, in addition to the role in the circadian clock, PRR7 plays a significant role in flowering under long days, and functions of PRR5 and PRR9 are overlapping with PRR7 (Imaizumi & Kay, 2006). All five participants of the Arabidopsis PRR family engage in the circadian clock, also FT expression is regulated by it. This indicates that Ppd-1 in wheat, given homology to PRR7, is not essential for the circadian clock feedback loops.

The sub functionalization of PRR genes in cereals is more likely to allow the photoperiodism function and regulate early flowering through mutation of the Ppd-1 gene (Higgins, Bailey, & Laurie, 2010). Level of gene expression was ascertained on the basis

of overall participations of the three homologues, with the B genome dominant for most genes in ' Paragon ' and the introgression lines. Furthermore, the B genome was the only allele to give the expression of other genomes of allele. With the introduction of Ppd-A1a or Ppd-D1a alleles in Ppd-1, in ‘Paragon’ the B gene predominated and the expression from all the other genomes increased.

For selective genome editing we used CRISPR / Cas9 method. This tool is very useful and its effectiveness has shown promising results in different species. This method is very easy and more effective as we compared this tool with other techniques for editing of genomes. However, there are some challenges in the usage of this tool for monocots plants, as mutations at off- targets sites were resulted sometime. Here, a series of experiments was carried out to demonstrate effective genome editing in wheat by utilizing CRISPR / Cas9 method. Results from experiments verified that the tool used - (CRISPR / Cas9) - is a successful method, that we used for advanced targeted editing in wheat genome.

The major effort of this project aimed at establishing and initiating optimization of the CRISPR/Cas9 system in wheat genome. Two sgRNAs were designed to target each of the genes Ppd1 and FT. For this method construct of CRISPR/Cas9 with two gene was successfully edited and inserted in wheat genome. The next step is to produce transgenic plants edited by both genes. Therefore, from results it is evident that Ppd-1 and FT genes are involved in photoperiod mechanism and produce paired spikelets in wheat. It will useful to stimulate the plants that bloom earlier with more spikelets thus, leading to increase in yield of wheat.

REFERENCES

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., . . . Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science, 315*(5819), 1709-1712.

Beales, J., Turner, A., Griffiths, S., Snape, J. W., & Laurie, D. A. (2007). A pseudo-response regulator is misexpressed in the photoperiod insensitive Ppd-D1a mutant of wheat (Triticum aestivum L.). *Theoretical and Applied Genetics, 115*(5), 721-733.

Becker, A., & Theißen, G. (2003). The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Molecular phylogenetics and evolution, 29*(3), 464-489.

Belhaj, K., Chaparro-Garcia, A., Kamoun, S., Patron, N. J., & Nekrasov, V. (2015). Editing plant genomes with CRISPR/Cas9. *Current opinion in biotechnology, 32*, 76-84.

Bimboim, H., & Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic acids research, 7*(6), 1513-1523.

Boden, S. A., Cavanagh, C., Cullis, B. R., Ramm, K., Greenwood, J., Finnegan, E. J., . . . Swain, S. M. (2015). Ppd-1 is a key regulator of inflorescence architecture and paired spikelet development in wheat. *Nature plants, 1*(2), 14016.

Bortesi, L., & Fischer, R. (2015). The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnology advances, 33*(1), 41-52.

Choulet, F., Alberti, A., Theil, S., Glover, N., Barbe, V., Daron, J., . . . Paux, E. (2014). Structural and functional partitioning of bread wheat chromosome 3B. *Science, 345*(6194), 1249721.

Doudna, J. A., & Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science, 346*(6213), 1258096.

Gaj, T., Gersbach, C. A., & Barbas III, C. F. (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in biotechnology, 31*(7), 397-405.

Higgins, J. A., Bailey, P. C., & Laurie, D. A. (2010). Comparative genomics of flowering time pathways using Brachypodium distachyon as a model for the temperate grasses. *PLoS one, 5*(4).

Imaizumi, T., & Kay, S. A. (2006). Photoperiodic control of flowering: not only by coincidence. *Trends in plant science, 11*(11), 550-558.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., & Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *Journal of bacteriology, 169*(12), 5429-5433.

Laurie, D. A. (1997). Comparative genetics of flowering time *Oryza: From Molecule to Plant* (pp. 167-177): Springer.

Li, J.-F., Zhang, D., & Sheen, J. (2014). Cas9-based genome editing in Arabidopsis and tobacco *Methods in enzymology* (Vol. 546, pp. 459-472): Elsevier.

Ling, H.-Q., Zhao, S., Liu, D., Wang, J., Sun, H., Zhang, C., . . . Tao, Y. (2013). Draft genome of the wheat A-genome progenitor Triticum urartu. *Nature, 496*(7443), 87-90.

Lowe, K., Wu, E., Wang, N., Hoerster, G., Hastings, C., Cho, M.-J., . . . Cushatt, J. (2016). Morphogenic regulators Baby boom and Wuschel improve monocot transformation. *The Plant Cell, 28*(9), 1998-2015.

Münster, T., Pahnke, J., Di Rosa, A., Kim, J. T., Martin, W., Saedler, H., & Theissen, G. (1997). Floral homeotic genes were recruited from homologous MADS-box genes preexisting in the common ancestor of ferns and seed plants. *Proceedings of the National Academy of Sciences, 94*(6), 2415-2420.

Murashige, T., & Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiologia plantarum, 15*(3), 473-497.

Nakamichi, N., Kita, M., Niinuma, K., Ito, S., Yamashino, T., Mizoguchi, T., & Mizuno, T. (2007). Arabidopsis clock-associated pseudo-response regulators PRR9, PRR7 and PRR5 coordinately and positively regulate flowering time through the canonical CONSTANS-dependent photoperiodic pathway.

Nissim, L., Perli, S. D., Fridkin, A., Perez-Pinera, P., & Lu, T. K. (2014). Multiplexed and programmable regulation of gene networks with an integrated RNA and CRISPR/Cas toolkit in human cells. *Molecular cell, 54*(4), 698-710.

Peng, R., Lin, G., & Li, J. (2016). Potential pitfalls of CRISPR/Cas9- mediated genome editing. *The FEBS journal, 283*(7), 1218-1231.

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature protocols, 8*(11), 2281.

Schiml, S., Fauser, F., & Puchta, H. (2014). The CRISPR/C as system can be used as nuclease for in planta gene targeting and as paired nickases for directed mutagenesis in A rabidopsis resulting in heritable progeny. *The Plant Journal, 80*(6), 1139-1150.

Schiml, S., & Puchta, H. (2016). Revolutionizing plant biology: multiple ways of genome engineering by CRISPR/Cas. *Plant methods, 12*(1), 8.

Sharman, B. (1944). Branched heads in wheat and wheat hybrids. *Nature, 153*(3886), 497-498.

Shaw, L. M., Turner, A. S., & Laurie, D. A. (2012). The impact of photoperiod insensitive Ppd- 1a mutations on the photoperiod pathway across the three genomes of hexaploid wheat (Triticum aestivum). *The Plant Journal, 71*(1), 71-84.

Shimada, S., Ogawa, T., Kitagawa, S., Suzuki, T., Ikari, C., Shitsukawa, N., . . . Handa, H. (2009). A genetic network of flowering- time genes in wheat leaves, in which an APETALA1/FRUITFULL- like gene, VRN1, is upstream of FLOWERING LOCUS T. *The Plant Journal, 58*(4), 668-681.

Shukla, V. K., Doyon, Y., Miller, J. C., DeKelver, R. C., Moehle, E. A., Worden, S. E., . . . Meng, X. (2009). Precise genome modification in the crop species Zea mays using zinc-finger nucleases. *Nature, 459*(7245), 437-441.

Sivamani, E., Bahieldin, A., Wraith, J. M., Al-Niemi, T., Dyer, W. E., Ho, T.-H. D., & Qu, R. (2000). Improved biomass productivity and water use efficiency under water deficit conditions in transgenic wheat constitutively expressing the barley HVA1 gene. *Plant science, 155*(1), 1-9.

Snape, J., Butterworth, K., Whitechurch, E., & Worland, A. (2001). Waiting for fine times: genetics of flowering time in wheat *Wheat in a global environment* (pp. 67-74): Springer.

Turck, F., Fornara, F., & Coupland, G. (2008). Regulation and identity of florigen: FLOWERING LOCUS T moves center stage. *Annu. Rev. Plant Biol., 59*, 573-594.

Turner, A., Beales, J., Faure, S., Dunford, R. P., & Laurie, D. A. (2005). The pseudo-response regulator Ppd-H1 provides adaptation to photoperiod in barley. *Science, 310*(5750), 1031-1034.

Yen, C., & Yang, J. (1992). THE ESSENTIAL NATURE OF ORGANS IN GRAMINEAE, MULTIPLE SECONDARY AXES THEORY--A NEW CONCEPT. *Journal of Sichuan Agricultural University*(4), 1.

Zhang, B., Yang, X., Yang, C., Li, M., & Guo, Y. (2016). Exploiting the CRISPR/Cas9 system for targeted genome mutagenesis in petunia. *Scientific reports, 6*, 20315.

APPENDICES

1. Alignment of ppd1 gene

```
Alignment: Untitled1

                ....|....| ....|....| ....|....| ....|....| ....|....|
                         10         20         30         40         50
TraesCSU02      ATGCGACCGA GGCCGCCTCC TCCTAGCGCC ATGGACCGTC ATCACCACCA
TraesCS2D0      ATGCGACCGA GGCCGCCTCC TCCTAGCGCC ATGGACCGCC ATCACCACCA
TraesCSU02      ATGCGACCGA GGCCGCCTCC TCCTAGCGCC ATGGACCGTC ATCACCACCA
Clustal Co      ********** ********** ********** ******** * **********


                ....|....| ....|....| ....|....| ....|....| ....|....|
                         60         70         80         90        100
TraesCSU02      GCAGCAGCAG CAGCCGCCGT CGCCTCAAGG GGACCATGCC GCCCAGCCGC
TraesCS2D0      GCAG------ ---CCGCCGT CGCCGCAGGG GGAGCATGCC GCCCAGCCAC
TraesCSU02      GCAGCAGCAG CAGCCGCCGT CGCCTCAAGG GGACCATGCC GCCCAGCCGC
Clustal Co      ****          ******* **** ** ** *** ****** ******** *


                ....|....| ....|....| ....|....| ....|....| ....|....|
                        110        120        130        140        150
TraesCSU02      GCTGCTGGGA GGAGTTCCTC CACAGGAAGA CCATCCGGGT CCTGCTCGTG
TraesCS2D0      GCTGCTGGGA GGAGTTCCTC CACAGGAAGA CCATCAGGGT GCTGCTCGTG
TraesCSU02      GCTGCTGGGA GGAGTTCCTC CACAGGAAGA CCATCCGGGT CCTGCTCGTG
Clustal Co      ********** ********** ********** ***** ****  *********


                ....|....| ....|....| ....|....| ....|....| ....|....|
                        160        170        180        190        200
TraesCSU02      GAGACCGACG ACTCCACCCG GCAGGTCGTC ACCGCCCTGC TCCGCCACTG
TraesCS2D0      GAGACCGACG ACTCCACCCG GCAGGTCGTC ACCGCCCTGC TCCGCCACTG
TraesCSU02      GAGACCGACG ACTCCACCCG GCAGGTCGTC ACCGCCCTGC TCCGCCACTG
Clustal Co      ********** ********** ********** ********** **********


                ....|....| ....|....| ....|....| ....|....| ....|....|
                        210        220        230        240        250
TraesCSU02      CATGTACCAA GTTATCCCTG CTGAAAACGG CCACCAGGCG TGGGCGTATC
TraesCS2D0      CATGTACCAA GTTATCCCTG CTGAAAACGG CCACCAGGCG TGGGCGTATC
TraesCSU02      CATGTACCAA GTTATCCCTG CTGAAAACGG CCACCAGGCG TGGGCGTATC
Clustal Co      ********** ********** ********** ********** **********


                ....|....| ....|....| ....|....| ....|....| ....|....|
                        260        270        280        290        300
TraesCSU02      TCCAGGACAT GCAGAGCAAC ATCGACCTTG TTCTGACAGA GGTCTTCATG
TraesCS2D0      TTCAGGACAT GCAGAGCAAC ATCGACCTTG TTCTGACAGA GGTCTTCATG
TraesCSU02      TCCAGGACAT GCAGAGCAAC ATCGACCTTG TTCTGACAGA GGTCTTCATG
Clustal Co      * ******** ********** ********** ********** **********


                ....|....| ....|....| ....|....| ....|....| ....|....|
```

```
                  310        320        330        340        350
TraesCSU02   CACGGCGGTC TCTCCGGGAT CGACCTGCTC GGCAGGATCA TGAACCACGA
TraesCS2D0   CACGGCGGTC TCTCTGGGAT CGACCTGCTC GGCAGGATCA TGAACCACGA
TraesCSU02   CACGGCGGTC TCTCCGGGAT CGACCTGCTC GGCAGGATCA TGAACCACGA
Clustal Co   ********** **** ***** ********** ********** **********

                  ....|....| ....|....| ....|....| ....|....| ....|....|
                  360        370        380        390        400
TraesCSU02   GGTCTGCAAG GACATCCCCG TCATCATGAT GTCGTCGCAC AATTCGATGG
TraesCS2D0   GGTCTGCAAG GACATCCCCG TCATCATGAT GTCGTCGCAC GATTCGATGG
TraesCSU02   GGTCTGCAAG GACATCCCCG TCATCATGAT GTCGTCGCAC AATTCGATGG
Clustal Co   ********** ********** ********** **********  *********

                  ....|....| ....|....| ....|....| ....|....| ....|....|
                  410        420        430        440        450
TraesCSU02   GCACGGTCCT CAGTTGCCTG TCAAATGGTG CTGCTGACTT CTTGGCCAAG
TraesCS2D0   GCACGGTCCT CAGCTGCCTG TCAAATGGTG CCGCCGACTT CTTGGCCAAG
TraesCSU02   GCACGGTCCT CAGTTGCCTG TCAAATGGTG CTGCTGACTT CTTGGCCAAG
Clustal Co   ********** *** ****** ********** * ** ***** **********

                  ....|....| ....|....| ....|....| ....|....| ....|....|
                  460        470        480        490        500
TraesCSU02   CCGATACGTA AGAACGAGCT TAAGAACCTT TGGGCGCATG TGTGGAGACG
TraesCS2D0   CCGATTCGTA AGAACGAGCT TAAGAACCTT TGGGCGCATG TGTGGAGACG
TraesCSU02   CCGATACGTA AGAACGAGCT TAAGAACCTT TGGGCGCATG TGTGGAGACG
Clustal Co   ***** **** ********** ********** ********** **********

                  ....|....| ....|....| ....|....| ....|....| ....|....|
                  510        520        530        540        550
TraesCSU02   CTCTCACAGC TCCAGTGGCA GTGGTAGTGG AAGTGCCATT CAGACGCAGA
TraesCS2D0   GTCTCACAGC TCCAGTGGCA GTGGCAGTGG AAGTGCCATT CAGACACAGA
TraesCSU02   CTCTCACAGC TCCAGTGGCA GTGGTAGTGG AAGTGCCATT CAGACGCAGA
Clustal Co    ********* ********** **** ***** ********** ***** ****

                  ....|....| ....|....| ....|....| ....|....| ....|....|
                  560        570        580        590        600
TraesCSU02   AGTGTACCAA ATCAAAGAGC GCTGACGATT CCAATAATAA CAGCAATAAC
TraesCS2D0   AGTGTACCAA ATCTAAGAGT GGCGATGATT CCAATAATAA CAGCAATAAT
TraesCSU02   AGTGTACCAA ATCAAAGAGC GCTGACGATT CCAATAATAA CAGCAATAAC
Clustal Co   ********** *** *****  *  ** **** ********** *********

                  ....|....| ....|....| ....|....| ....|....| ....|....|
                  610        620        630        640        650
TraesCSU02   CGCAACGACG ATGCCAGCAT GGGGCTCAAT GCAAGGGATG GCAGCGATAA
TraesCS2D0   CGCAACGACG ATGCCAGCAT GGGGCTCAAT GCAAGGGATG GCAGCGATAA
TraesCSU02   CGCAACGACG ATGCCAGCAT GGGGCTCAAT GCAAGGGATG GCAGCGATAA
Clustal Co   ********** ********** ********** ********** **********

                  ....|....| ....|....| ....|....| ....|....| ....|....|
                  660        670        680        690        700
```

```
TraesCSU02   TGGTAGTGGC ACTCAGAGCT CATGGACAAA GCGTGCCGTT GAGATCGACA
TraesCS2D0   TGGCAGTGGC ACTCAGAGCT CATGGACAAA GCGTGCCGTT GAGATCGACA
TraesCSU02   TGGTAGTGGC ACTCAGAGCT CATGGACAAA GCGTGCCGTT GAGATCGACA
Clustal Co   *** ****** ********** ********** ********** **********

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       710        720        730        740        750
TraesCSU02   GTCCACAGGA CATGTCTCCG GATCAGTCAG TTGATCCTCC TGATAGCACT
TraesCS2D0   GTCCGCAGGA CATGTCTCCA GATCAGTCAG TTGATCCGCC TGATAGCACT
TraesCSU02   GTCCACAGGA CATGTCTCCG GATCAGTCAG TTGATCCTCC TGATAGCACT
Clustal Co   **** ***** *********  ********** ******* ** **********

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       760        770        780        790        800
TraesCSU02   TGCGCGCATG TGAGCCACCT CAAGTCAGAG ATATGCAGCA ATAGATTAAG
TraesCS2D0   TGCGCGCATG TGAGCCACCT GAAGTCAGAG ATATGCAGCA ATAGATTAAG
TraesCSU02   TGCGCGCATG TGAGCCACCT CAAGTCAGAG ATATGCAGCA ATAGATTAAG
Clustal Co   ********** **********  ********* ********** **********

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       810        820        830        840        850
TraesCSU02   AGGT------ ---------- ---------- ---------- ------GATG
TraesCS2D0   AGGTACAAAT AACAAAAAAT GCCAGAAACC AAAAGAAACT AA---TGCTG
TraesCSU02   AGGTACAGAT AACAAAAAAT GCCAGAAACC AAAAGAAACT AATGGTGATG
Clustal Co   ****                                          * **

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       860        870        880        890        900
TraesCSU02   AGTTCAAAGG GAAGGAGCTG GAGATAGGTG CCCCTGGTAA TTTGAACACA
TraesCS2D0   AGTTCAAAGG GAAGGAGTTG GAGATAGGTG CCCCTGGTAA TTTGAATACA
TraesCSU02   AGTTCAAAGG GAAGGAGCTG GAGATAGGTG CCCCTGGTAA TTTGAACACA
Clustal Co   ********** ******* ** ********** ********** ****** ***

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       910        920        930        940        950
TraesCSU02   GATGATCAAT CCTCCCCGAA CGAGAGTTCG GTCAAACCAA CAGATAATGG
TraesCS2D0   GAGGATCAAT CCTCCCCGAA TGAGAGTTCG GTCAAACCAG CTGATAATGG
TraesCSU02   GATGATCAAT CCTCCCCGAA CGAGAGTTCG GTCAAACCAA CAGATAATGG
Clustal Co   ** ******* **********  ********* *********  * ********

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       960        970        980        990       1000
TraesCSU02   ACGGTGTGAG TATCTGCCAC AGAACAACTC CAACGATACA GTTATGGAAA
TraesCS2D0   ACGGTGTGAG TATCTGCCAC AGAACAATTC CAATGATACA GTTATGGAAA
TraesCSU02   ACGGTGTGAG TATCTGCCAC AGAACAACTC CAACGATACA GTTATGGAAA
Clustal Co   ********** ********** ******* ** *** ****** **********

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       1010       1020       1030       1040       1050
 TraesCSU02   ATTCGGATGA GCCAATTGTT CGAGCTGCTG ACCTAATCGG TTCGATGGCC
```

```
TraesCS2D0   ATTCGGATGA GCCAATTGTT CGAGCTGCTG ACCTAATCGG TTCGATGGCC
TraesCSU02   ATTCGGATGA GCCAATTGTT CGAGCTGCTG ACCTAATCGG TTCGATGGCC
Clustal Co   ********** ********** ********** ********** **********


                ....|....| ....|....| ....|....| ....|....| ....|....|
                       1060       1070       1080       1090       1100
TraesCSU02   AAAAACATGG ACGCCCAACA GGCAGCTAGA GCAATAGATG CTCCTAACTG
TraesCS2D0   AAAAACATGG ACGCCCAACA GGCAGCTAGA GCCATAGATG CTCCTAACTG
TraesCSU02   AAAAACATGG ACGCCCAACA GGCAGCTAGA GCAATAGATG CTCCTAACTG
Clustal Co   ********** ********** ********** ** ******* **********


                ....|....| ....|....| ....|....| ....|....| ....|....|
                       1110       1120       1130       1140       1150
TraesCSU02   CTCCTCACAA GTGCCGGAAG GGAAAGACGC CGACCGTGAG AACGCCATGC
TraesCS2D0   CTCCTCACAA GCGCCGCAAG GGAAAGACAC GGACCGTGAA AACGCCATGC
TraesCSU02   CTCCTCACAA GTGCCGGAAG GGAAAGACGC CGACCGTGAG AACGCCATGC
Clustal Co   ********** * **** *** ******** *  ********  **********


                ....|....| ....|....| ....|....| ....|....| ....|....|
                       1160       1170       1180       1190       1200
TraesCSU02   CATATCTTGA GCTGAGCCTA AAGAGGTCGA GATCGACCAC GGAGGGTGCG
TraesCS2D0   CGTATCTTGA ACTGAGCCTG AAGAGGTCGA GATCGACCGC GGACGGAGCG
TraesCSU02   CATATCTTGA GCTGAGCCTA AAGAGGTCGA GATCGACCAC GGAGGGTGCG
Clustal Co   * ********  ********  ********** ******** * *** ** ***


                ....|....| ....|....| ....|....| ....|....| ....|....|
                       1210       1220       1230       1240       1250
TraesCSU02   GATGC---GA TCCAGGAGGA ACAGAGGAAC GTCGTGAGAC GATCAGACCT
TraesCS2D0   GATGCTGCGA TCCAGGAGGA ACAGAGGAAC GTCGTGAGAC GATCGGACCT
TraesCSU02   GATGC---GA TCCAGGAGGA ACAGAGGAAC GTCGTGAGAC GATCAGACCT
Clustal Co   *****   ** ********** ********** ********** **** *****


                ....|....| ....|....| ....|....| ....|....| ....|....|
                       1260       1270       1280       1290       1300
TraesCSU02   CTCGGCATTC ACGAGGTACA ATACGTGCTC GTTCTCCAAT CAAGGCGGGG
TraesCS2D0   CTCGGCATTC ACGAGGTACA ATACGTGCGC GGTCTCCAAT CAAGGCGGTG
TraesCSU02   CTCGGCATTC ACGAGGTGCA AAGCATAA-- ---------- ----------
Clustal Co   ********** ******* ** *  * *


                ....|....| ....|....| ....|....| ....|....| ....|....|
                       1310       1320       1330       1340       1350
TraesCSU02   CAGGGTTCGT CGGGAGCTGT TCGCCCAACG GCAACAGCTC CAAGGCCGCG
TraesCS2D0   CAGGGTTCGT CGGGAGCTGC TCGCCCAACG GCAACAGCTC CGAGGCCGCG
TraesCSU02   ---------- ---------- ---------- ---------- ----------
Clustal Co


                ....|....| ....|....| ....|....| ....|....| ....|....|
                       1360       1370       1380       1390       1400
TraesCSU02   AAAACGGACG CCGCTCAGAT GAAGCAAGGC TCAAATGGCA GCAGCAACAA
TraesCS2D0   AAAACGGACG CCGCTCAGAT GAAGCAAGGC TCCAACGGCA GCAGCAACAA
```

```
TraesCSU02   ---------- ---------- ---------- ---------- ----------
Clustal Co


             ....|....| ....|....| ....|....| ....|....| ....|....|
                    1410       1420       1430       1440       1450
TraesCSU02   CAACGACATG GGCTCCACCA CCAAGAGCGT GATGACCAAG CCCGGTGGC-
TraesCS2D0   CAACGACATG GGCTCCACCA CCAAGAGCGT GGTGACCAAG CCCGCCGGCG
TraesCSU02   ---------- ---------- ---------- ---------- ----------
Clustal Co


             ....|....| ....|....| ....|....| ....|....| ....|....|
                    1460       1470       1480       1490       1500
TraesCSU02   --AATAACAA GGTTTCACCC ATCAACGGCA ACACGCACAC GTCGGCGTTC
TraesCS2D0   GAAATAATAA GGTTTCGCCC ATCAACGGCA ACACACATAC CTCGGCGTTC
TraesCSU02   ---------- ---------- ---------- ---------- ----------
Clustal Co


             ....|....| ....|....| ....|....| ....|....| ....|....|
                    1510       1520       1530       1540       1550
TraesCSU02   CATCGTGTGC AGCCATGGAC GCCGGCGACA GCAACAGGGA AAGACAAGGT
TraesCS2D0   CATCGTGTGC AGCCGTGGAC GCCGGCAACA GCAGCAGGGA AAGACAAGGC
TraesCSU02   ---------- ---------- ---------- ---------- ----------
Clustal Co


             ....|....| ....|....| ....|....| ....|....| ....|....|
                    1560       1570       1580       1590       1600
TraesCSU02   TGATGACGTG AGCAAGAAGA ATGCTGCAGC AGCAGCAGCC GCAGCAGCGA
TraesCS2D0   TGATGAAATG AGCAAGAAGA ATGCTGCAAC AGCAGCAG-- ----CGGCGA
TraesCSU02   ---------- ---------- ---------- ---------- ----------
Clustal Co


             ....|....| ....|....| ....|....| ....|....| ....|....|
                    1610       1620       1630       1640       1650
TraesCSU02   AAGACAATGG TGGTGAAGCA CAGAGCAAAC ACCCCTGTGC AGCGGCCCAT
TraesCS2D0   AAGACAATGG TGGTGAAGCA CAGAGCAAAC GCCCCTGTGC AGCGGCCCAT
TraesCSU02   ---------- ---------- ---------- ---------- ----------
Clustal Co


             ....|....| ....|....| ....|....| ....|....| ....|....|
                    1660       1670       1680       1690       1700
TraesCSU02   GATGCGAATG GTGGATC--- GGCAGGAGGC ---GCTCAGT CCAACGTGAT
TraesCS2D0   GACACGAATG GTGGATCATC GGCAGGAGGC ACTGCTCAGT CCAATGTGGT
TraesCSU02   ---------- ---------- ---------- ---------- ----------
Clustal Co


             ....|....| ....|....| ....|....| ....|....| ....|....|
                    1710       1720       1730       1740       1750
TraesCSU02   CAATCCCTCG GGTCCTGTTG AAGGTCATGC TGCCAACTAT GGTAGCAACT
TraesCS2D0   CAATCCTTCG GGTCCGGTTG AAGGTCATGC TGCCAACTAT GGTAGCAACT
TraesCSU02   ---------- ---------- ---------- ---------- ----------
```

```
Clustal Co

                .....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
                        1760       1770       1780       1790       1800
TraesCSU02      CTGGCAGTAA TAACAATACC AATAACGGTG GGAGCACCGC AGCTACTGCT
TraesCS2D0      CTGGCAGTAA TAACAACACC AACAA---TG GGAGCACCGC AGCTACTGCT
TraesCSU02      ---------- ---------- ---------- ---------- ----------
Clustal Co

                .....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
                        1810       1820       1830       1840       1850
TraesCSU02      GCTGCTGCTG CTGCTGTACA TGGTGAGACC GGTGGCATCG ACAAAAGAAG
TraesCS2D0      GCTGCTGC-- ----TGTACA TGCTGAGACC GGTGGCATCG ACAAAAGAAG
TraesCSU02      ---------- ---------- ---------- ---------- ----------
Clustal Co

                .....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
                        1860       1870       1880       1890       1900
TraesCSU02      CAACATGATG CACATGAAAC GGGAGCGCCG GGTGGCTGCC GTGAACAAAT
TraesCS2D0      CAACATGATG CACATGAAAC GGGAGCGCCG GGTGGCCGCC GTGAACAAGT
TraesCSU02      ---------- ---------- ---------- ---------- ----------
Clustal Co

                .....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
                        1910       1920       1930       1940       1950
TraesCSU02      TCAGAGAGAA GAGAAAAGAG AGGAACTTCG GGAAGAAGGT GCGTTACCAG
TraesCS2D0      TCAGAGAGAA GAGAAAAGAG AGGAACTTCG GGAAGAAGGT GCGTTACCAG
TraesCSU02      ---------- ---------- ---------- ---------- ----------
Clustal Co

                .....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
                        1960       1970       1980       1990       2000
TraesCSU02      AGCAGGAAGC GGCTGGCGGA GCAGCGCCCA CGGGTGCGCG GGCAGTTCGT
TraesCS2D0      AGCAGGAAGA GACTGGCCGA GCAGCGCCCG CGGGTGCGCG GGCAGTTCGT
TraesCSU02      ---------- ---------- ---------- ---------- ----------
Clustal Co

                .....|.....| .....|.....| .....|.....| .....|.....|
                        2010       2020       2030       2040
TraesCSU02      GCGGCAGCCG CC-ACCGCCG GCTGCCGTGG AGAGATAA--
TraesCS2D0      GCGGCTGCCG TTGAGAGATA ACCTCCCGCC ACACACCTAG
TraesCSU02      ---------- ---------- ---------- ----------
Clustal Co
```

2. Alignment of FT gene

```
Alignment: Untitled

                .....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
                        10         20         30         40         50
 TraesCS7A0     ATGGCCGGGA GGGACAGGGA CCCGCTGGTG GTTGGCAGGG TTGTGGGAGA
```

```
TraesCS7B0   ATGGCCGGTA GGGATAGGGA CCCGCTGGTG GTTGGCAGGG TTGTGGGGGA
TraesCS7D0   ATGGCCGGGA GGGACAGAGA CCCGCTGGTG GTTGGCAGGG TTGTGGGGGA
Clustal Co   ******** * **** ** ** ********** ********** ******* **


             ....|....| ....|....| ....|....| ....|....| ....|....|
                      60         70         80         90        100
TraesCS7A0   CGTGCTGGAC CCCTTTGTCC GGACCACCAA CCTCAGGGTG ACCTTCGGGA
TraesCS7B0   CGTGCTGGAC CCCTTCGTCC GGACCACCAA CCTCAGGGTG ACCTTCGGGA
TraesCS7D0   CGTGCTGGAC CCCTTCATCC GGACCACCAA CCTCAGGGTG ACCTTCGGGA
Clustal Co   ********** *****  *** ********** ********** **********


             ....|....| ....|....| ....|....| ....|....| ....|....|
                     110        120        130        140        150
TraesCS7A0   ACAGGACCGT GTCCAACGGC TGCGAGCTCA AGCCGTCCAT GGTCGCCCAG
TraesCS7B0   ACAGGACCGT GTCCAACGGC TGCGAGCTCA AGCCGTCCAT GGTCGCCCAG
TraesCS7D0   ACAGGACCGT GTCCAACGGC TGCGAGCTCA AGCCGTCCAT GGTCGCCCAG
Clustal Co   ********** ********** ********** ********** **********


             ....|....| ....|....| ....|....| ....|....| ....|....|
                     160        170        180        190        200
TraesCS7A0   CAGCCCAGGG TTGAGGTGGG CGGCAATGAG ATGAGGACCT TCTACACACT
TraesCS7B0   CAGCCCAGGG TTGAGGTGGG CGGCAATGAG ATGAGGACCT TCTACACACT
TraesCS7D0   CAGCCCAGGG TTGAGGTGGG CGGCAATGAG ATGAGGACCT TCTACACACT
Clustal Co   ********** ********** ********** ********** **********


             ....|....| ....|....| ....|....| ....|....| ....|....|
                     210        220        230        240        250
TraesCS7A0   CGTGATGGTA GACCCAGATG CTCCAAGTCC AAGCGATCCC AACCTTAGGG
TraesCS7B0   CGTGATGGTA GACCCAGATG CTCCAAGTCC AAGCGATCCC AACCTTAGGG
TraesCS7D0   CGTGATGGTA GACCCAGATG CTCCAAGTCC AAGCGATCCC AACCTTAGGG
Clustal Co   ********** ********** ********** ********** **********


             ....|....| ....|....| ....|....| ....|....| ....|....|
                     260        270        280        290        300
TraesCS7A0   AGTATCTCCA C---CTTGTG ACAGATATCC CCGGTACAAC TGGTGCCTCG
TraesCS7B0   AGTATCTCCA CTGGCTTGTG ACAGATATCC CCGGTACAAC TGGTGCGTCG
TraesCS7D0   AGTATCTCCA CTGGCTTGTG ACAGATATCC CCGGTACAAC TGGTGCATCC
Clustal Co   ********** *   ****** ********** ********** ****** **


             ....|....| ....|....| ....|....| ....|....| ....|....|
                     310        320        330        340        350
TraesCS7A0   TTCGGGCAGG AAGTGATGTG CTATGAGAGC CCTCGTCCGA CCATGGGGAT
TraesCS7B0   TTCGGGCAGG AGGTGATGTG CTACGAGAGC CCTCGTCCGA CCATGGGGAT
TraesCS7D0   TTCGGGCAGG AGGTGATGTG CTACGAGAGC CCTCGTCCGA CCATGGGGAT
Clustal Co   ********** * ******** *** ****** ********** **********


             ....|....| ....|....| ....|....| ....|....| ....|....|
                     360        370        380        390        400
TraesCS7A0   CCACCGCTTC GTGCTCGTGC TCTTCCAGCA GCTCGGCCGG CAGACGGTGT
TraesCS7B0   CCACCGCTTC GTGCTCGTAC TCTTCCAGCA GCTCGGGCGG CAGACGGTGT
TraesCS7D0   CCATCGCTTC GTGCTCGTGC TCTTCCAGCA GCTCGGCCGG CAGACCGTGT
Clustal Co   *** ****** ******** * ********** ****** *** ***** ****
```

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                        410        420        430        440        450
TraesCS7A0      ACGCCCCCGG GTGGCGCCAG AACTTCAACA CCAGGGACTT CGCCGAGCTC
TraesCS7B0      ACGCCCCCGG GTGGCGCCAG AACTTCAACA CCAGGGACTT CGCCGAGCTC
TraesCS7D0      ACGCTCCCGG GTGGCGCCAG AACTTCAACA CCAGGGACTT CGCCGAGCTC
Clustal Co      **** ***** ********** ********** ********** **********

                ....|....| ....|....| ....|....| ....|....| ....|....|
                        460        470        480        490        500
TraesCS7A0      TACAACCTTG GCCCGCCCGT CGCCGCCGTC TACTTCAACT GCCAGCGTGA
TraesCS7B0      TACAACCTCG GCCCGCCTGT CGCCGCCGTC TACTTCAACT GCCAGCGTGA
TraesCS7D0      TACAACCTCG GCCCGCCTGT CGCCGCCGTC TACTTCAACT GCCAGCGTGA
Clustal Co      ******** * ******* ** ********** ********** **********

                ....|....| ....|....| ....|....| ....
                        510        520        530
TraesCS7A0      GGCCGGCTCC GGTGGCAGGA GGATGTACAA TTGA
TraesCS7B0      GGCCGGCTCC GGCGGCAGGA GGATGTACAA TTGA
TraesCS7D0      GGCCGGCTCC GGCGGCAGGA GGATGTACAA TTGA
Clustal Co      ********** ** ******* ********** ****
```