T.R.
EGE UNIVERSITY
Graduate School of Applied and Natural Science

ESTABLISHMENT OF CAS9 ENDONUCLEASE TECHNOLOGY IN RYE

MSc THESIS

Muhammad ARSLAN

Department of Seed Science and Technology

İzmir
2022

T.R.
EGE UNIVERSITY
Graduate School of Applied and Natural Science

ESTABLISHMENT OF CAS9 ENDONUCLEASE TECHNOLOGY IN RYE

Muhammad ARSLAN

Supervisor: Assoc. Prof. Dr. Evren KOBAN BAŞTANLAR
Co-supervisor: Dr. Sameh SELIM

Department of Seed Science and Technology
Erasmus Mundus Master Program in Plant Breeding Second Cycle Programme

İzmir
2022

Here I certify that this thesis entitled "ESTABLISHMENT OF CAS9 ENDONUCLEASE TECHNOLOGY IN RYE" conducted, prepared and presented by Muhammad ARSLAN for the degree of Master of Science has been recommended for acceptance and approval for oral defence as it complies both with the "Ege University Graduate School policies" and "Ege University Graduate School of Natural and Applied Sciences" rules and regulations. This thesis for the degree of Master of Science is sufficient in scope and quality, and the Candidate was found successful and the thesis was approved by the thesis committee unanimity/majority voting on 14 / 09 / 2022.

Examining Committee Members: **Signature**

Chairman : Prof. Dr. Hülya İLBİ

Reporter : Assoc. Dr. Evren KOBAN BAŞTANLAR

Member : Prof. Dr. Eftal DÜZYAMAN

Member : Assoc. Dr. Figen ERSOY

Member : Assoc. Dr. Zelal ADIGÜZEL

EGE UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

ETHICAL CONDUCT DECLARATION

I hereby declare that this thesis study entitled "**ESTABLISHMENT OF CAS9 ENDONUCLEASE TECHNOLOGY IN RYE**" describes my own work for the degree of MSc which has been done after registration at Ege University, and does not contain any material which is already submitted for a thesis or dissertation for a degree, diploma or other qualifications. I also declare that Ege University academic rules and regulations as well as ethical conduct has been followed accordingly for all the information which is collected and presented in this document. Lastly, I declare that, I have completely cited and referenced all material and results as prescribed by these rules and conduct of EGE University, which are not pioneer to this work.

14 / 09 / 2022

Muhammad ARSLAN

ÖZET

ÇAVDARDA CAS9 ENDONÜKLEAZ TEKNOLOJİSİNİN KULLANILMASI

ARSLAN, Muhammad

Yüksek Lisans Tezi, Tohumluk Bilimi ve Teknolojisi Anabilim Dalı
Danışman: Doç. Dr. Evren KOBAN BAŞTANLAR
Eş-danışman: Dr. Sameh SELIM
Eylül 2022, 78 sayfa

Çavdar (*Secale cereale* L.), biyoteknolojik yaklaşımlara karşı oldukça dirençli olduğu kanıtlanan diploid bir kış/ilkbahar tahıl bitkisidir. Genetik ve fenotipik olarak çok yönlü olan bu bitki dünya çapında ekmek üretmek için en önemli ikinci mahsuldür ve yeşil formunda çiftlik hayvanları için de yem olarak kullanılır. Yüksek potansiyelinin aksine, haploid teknolojisi ve genetik mühendisliği uygulamaları henüz çavdar yetiştirme programlarında önemli bir rol oynamamıştır.

CRISPR-Cas9 tabanlı genom düzenleme uygulamalarını mümkün kılabilecek bir gelişme olarak, çavdar genomu yakın zamanda yayınlanmıştır (Mart ve Ekim 2021). Sürgünlerin kallus oluşumu ve rejenerasyonu, genetik mühendisliği için ön koşul olan genotip ve kültür koşullarına oldukça bağlıdır. İnbred hatlardan No5 ve Lo7 (ikincisi dizili referans genomu temsil eder), daha önce yapılmış olan deneyler kapsamında bitki rejenerasyon sisteminde en iyi performans kaydedilmiştir. Bu elde edilen ön deneysel sonuçlar ışığında; çağdaş haploid teknolojisi ve genetik mühendisliği yaklaşımları için teknolojik bir temel sağlamak amacıyla, çavdarda hedeflenmiş bölgeye özgü mutasyonlar oluşturmak için RNA güdümlü Cas9 endonükleaz tabanlı genom düzenlemesi bu tez çalışması kapsamında uygulanmıştır.

PLA1 geninde fonksiyon kaybının, mısır bitkisinde haploid indükleme kabiliyeti kazanmakla ilişkili olduğu daha önce gösterilmişti.Bu nedenle, gerçekleştirdiğimiz çalışmada Cas9 endonükleaz teknolojisini kullanarak hedeflenen mutagenez gerçekleştirmek için mısır FOSFOLİPAZ A1'in (PLA1) çavdar ortologunu ele aldık. Bu kapsamda; (Sc) PLA1 geninin 1, 2 ve 3 ekzonlarında dört hedef motif seçtik ve buna göre dört kılavuz RNA (gRNA) tasarladık. gRNA'ların yapısı ve tasarımı, çevrimiçi araçlar (RNA fold, WU-CRISPR ve Benchling) kullanılarak tasarlandı. Golden Gate klonlamasına dayalı CASCADE/Cas9 vektör dizilimi sistemi, çavdar transformasyonu için P6i binary vektöründe seçilen gRNA'ların ekspresyon kasetini klonlamak için kullanılmıştır. Cas9 ekspresyon kaseti, GFP seçim markörü ve gRNA'lar, tek bir golden gate PCR reaksiyonunda bir klonlama vektörüne bağlandı. ScPLA1 için yapılandırılmış vektörlerin fonksiyonel doğrulaması, çavdar protoplastlarının PEG aracılı transfeksiyonu kullanılarak gerçekleştirildi. İki genotip ve dört vektör protoplast transfeksiyonu için test edilmiştir. Konfokal mikroskopi altında sfGFP seçim markörü kullanılarak yapılan hesaplamada %75 transfeksiyon oranı elde edilmiştir. Mutasyona uğramış protoplastlardan DNA izole edilmiş ve hedeflenen kısmi gen fragmanları için PCR reaksiyonları yapılmıştır. Muhtemel mutasyon indüklenmesine işaret eden hedefe spesifik amplikonlar için çift bantlar içeren PCR sonuçları, dizileme analizi için saklanmıştır. Bu amplikonların dizilenmesinden elde edilecek sonuçların mutasyonları tam olarak belirlemesi beklenmektedir.

Agrobacterium yoluyla stabil transformasyon, stabil transforme edilmiş çavdar bitkilerinde mutasyonun bir göstergesi olarak kullanılabilecek amplikon dizileme analizinden sonra yapılacaktır.

Anahtar kelimeler: Çavdar genom düzenlemesi, haploid indüksiyon, Cas9 endonükleaz, amplicon

ABSTRACT

ESTABLISHMENT OF CAS9 ENDONUCLEASE TECHNOLOGY IN RYE

ARSLAN, Muhammad

Master Thesis, Department of Seed Science and Technology
Supervisor: Assoc. Prof. Dr. Evren KOBAN BAŞTANLAR
Co-supervisor: Dr. Sameh SELIM
September 2022, 78 pages

Rye (*Secale cereale* L.) is a diploid winter/spring cereal plant which proved highly recalcitrant to biotechnological approaches. It is genetically and phenotypically highly versatile and the second most important crop to produce bread worldwide, while green plants are used as feed for livestock. In contrast to its high potential, haploid technology and genetic engineering have, as yet, not been playing a significant role in rye breeding programs.

Rye genome has recently been published (March and October 2021), that made the applications of CRISPR-Cas9 based genome editing available. Callus formation and regeneration of shoots is highly dependent upon genotype and culture conditions, which are pre-requisites for genetic engineering. Inbred lines No5 and Lo7 (the latter representing the sequenced reference genome) showed the best performance in plant regeneration system in previous experiments. Taking advantage of this preliminary progress and to provide a technological basis for contemporary haploid technology and genetic engineering approaches, RNA-guided Cas9 endonuclease-based genome editing is applied to generate site-specific mutations in rye in context of this thesis study.

The loss of function of the *PLA1* gene was previously demonstrated to be sufficient to confer haploid-inducing capability in maize. Therefore, we addressed the rye orthologue of the maize PHOSPHOLIPASE A1 (*PLA1*) to perform targeted

mutagenesis using Cas9 endonuclease technology. We selected four target motifs in exons 1, 2, and 3 of (Sc) *PLA1* gene and designed four guide RNAs (gRNAs) accordingly. Structure and design of gRNAs was employed using online tools (RNA fold, WU-CRISPR and Benchling). CASCADE/Cas9 vector assembly system based on Golden Gate cloning was used to clone the expression cassette of the selected gRNAs in the P6i binary vector for rye transformation. *Cas9* expression cassette, *GFP* selection marker and gRNAs were bound to an assembly vector in a single golden gate PCR reaction. Functional validation of constructed vectors for *ScPLA1* was done using PEG-mediated transfection of rye protoplasts. Two genotypes and four vectors were tested for protoplast transfection and 75% transfection rate was observed using *sfGFP* as selection marker under confocal microscopy. DNA from mutated protoplasts was extracted and PCR reactions were performed for the targeted partial gene fragments. The PCR results containing double bands for the target specific amplicons, which indicate towards possible mutation induction, were stored for further sequencing analysis. The results to be obtained from sequencing of these amplicons are expected to exactly pinpoint the mutations.

Stable transformation through Agrobacterium will be done after the analysis of amplicon sequencing which could be used as an indication of mutation in stably transformed rye plants.

Keywords: Rye genome editing, haploid induction, Cas9 endonuclease, amplicon

PREFACE

It has been known that rye is one of the most tolerant cereal to biotic and a-biotic stresses which enables it to survive in harsh conditions. Rye has been used as a genepool to improve wheat because of its genotypic and phenotypic potential. It is also the second most important crop for bread production in Europe. The primary demand for rye breeding programs is to improve its potential towards hybrid breeding and biotechnological tools. The major hindrance in rye is recalcitrance to plant regeneration and transformation which makes it very difficult crop for biotechnological approaches. Recently, rye genome has been sequenced which made available the applications for crop improvenment especially through genome editing. Applications of CRISPR/Cas9 makes the hybrid program more facile and rapid, which overtakes the conventional breeding approaches. Due to the poor focus on rye crop and its recalcitrance to plant transformation, there is not protocol for genome editing yet. For this reason, this study was performed to establish a protocol for genome editing in rye through RNA guided Cas9 endonuclease to target *PLA1* gene. *PLA1 (PhOSPHOLIPASE A1)* is responsible for inducing haploidy in cereals and knockout of this gene leads to generating haploidy inducer lines. Haploidy inducing line has the potential to produce haploid plants when cross-pollinated with a wild-type plant of the same species. Establishment of this protocol will open new windows for rye crop improvement.

During the six-month master thesis internship in Plant Reproductive Biology research group (IPK Gatersleben), I have learned useful knowledge on haploidy induction and RNA guided Cas9 endonuclease technology. The techniques for CRISPR/Cas9 application were new to me but thanks to Dr. Diaa's support and inspiration, I have learned gene cloning and functional validation along with *in silico* planning for CRISPR experiments. I sincerely express my gratitude to my supervisor (Associate Professor Dr. Evren Koban Baştanlar) and specially co-supervisor (Dr. Diaa Eldin S. Daghma) for the kind support and guidance throughout the master thesis study.

İZMIR

14 / 09 / 2022 Muhammad ARSLAN

TABLE OF CONTENTS

Sayfa

TABLE OF CONTENTS (continued)

Sayfa

TABLE OF CONTENTS (continued)

Sayfa

TABLE OF CONTENTS (continued)

Sayfa

LIST OF FIGURES

LIST OF FIGURES (continued)

LIST OF FIGURES (continued)

Figure Page

LIST OF FIGURES (continued)

LIST OF TABLES

LIST OF TABLES (continued)

Table Page

1. INTRODUCTION

Rye (*Secale cereale* L.) is a member of the *Triticeae* tribe and is genetically close to wheat and barley grown as a multi-purpose crop; for grain production, cover-crop, and forage-crop (University of Wyoming., 2017). It has been playing crucial role in human civilization by adding agronomic, nutritional and social values (Angioloni et al., 2011). Rye grain are second most important for Bread, crisp bread, and bakery products after wheat (Németh et al., 2021). Rye has highest content of dietary fiber and bioactive compounds among cereals. Rye food has better insulin metabolism than Wheat which helps in diabetes prevention (Jonsson et al., 2018). Rye has been widely used as a gene pool for improving wheat because of its high tolerance to a wide range of biotic and abiotic stresses (Jung et al., 2014). In addition, rye has been used as cover crop to prevent soil erosion. Rye has changed its status from “forgotten crop” to “pioneer cereal” since the introduction of hybrid breeding. Rye research has widened its horizons with the advent of genomic selection-based breeding schemes, recurrent selection and commercial inbred line development (Wilde et al., 2021). Research activities were carried out to understand the rye genome to include it in breeding programs (Bauer et al., 2017). The multifaceted advantages of rye plant still contradicts with its poor amenability to the biotechnological approaches in particular genome editing and haploid production.

Genome editing based on the CRISPR/Cas (clustered regularly interspaced short palindromic repeats - CRISPR-associated protein) system has been developed and proved to be efficient and less time-consuming in plants (Kumar et al., 2015). This molecular toolbox can be applied to explore the potential of rye crop and its hidden features (Kumlehn et al., 2018). This system employs a RNA-guided Cas9 endonuclease to induce double-strand breaks (DSBs). The generated DSBs are repaired by cellular DNA repair mechanisms, either by HR (homologous recombination) or NHEJ (non-homologous end joining) (Arora et al., 2017). The NHEJ is the main repair mechanism in plant cells which can lead to a DNA sequence alteration at targeted sites (Kumlehn et al., 2018).

Rye has proved to be highly recalcitrant to plant regeneration *in vitro* as well as to genetic transformation (Popelka et al., 2003). To provide a technological basis

for contemporary haploid technology and genetic engineering approaches, the principle of adventitious shoot formation from immature embryo explants has been established in rye in the lab PRB at IPK-Gatersleben, Germany. Stable and transient genetic transformation in rye will be established involving the use of *Agrobacterium tumefaciens* and/or biolistic bombardment in immature embryos explants. Produced haploid inducer (HI) lines will be used to cross-pollinate wild type rye elite genotypes to produce haploid plants, this will provide a rapidly platform for breeding rye plants with desired traits incorporated by bypassing the lengthy procedure of repeated crossing and backcrossing used in conventional breeding (Wang et al., 2019). The establishment of this protocol will open new doors for targeting different agronomic traits by this approach.

Based on the previous studies, mutation leading to loss of function in the *PLA1 (Phospholipase A1)* gene is conferred to haploid induction in cereals especially in maize. The aim for this study is to establish a protocol for genome editing in rye using Cas9 endonuclease technology to knockout *PLA1* gene to generate haploidy inducer (HI) lines. There is no protocol devised for genome editing for haploidy induction (HI) in rye, but similar work has been done in major crops like maize, rice and wheat (Wang et al., 2019). Knocking *PLA1* gene out has the ability to generate haploidy-inducing lines in maize (Liu et al., 2017). Haploidy inducing line has the capability to produce haploid plants when cross-pollinated with a wild-type plant of the same species. As rye is recalcitrant to plant transformation, biotechnological approaches have not been established yet providing an opportunity to establish technical basis for haploid induction and genetic engineering approaches in rye. Along with establishing Cas9 endonuclease technology in rye, specifically functional validation of *PLA1* gene was performed for the Master thesis study. Following are the specific goals of the study:

1. In-silico analysis for finding *PLA1* orthologue in rye, target motif identification and designing of gRNAs for Cas9 endonuclease application.
2. Using CASCADE/Cas9 modular vector system via Golden gate cloning to assemble the binary vectors.
3. Functional validation of *PLA1* gene through PEG mediated protoplast transfection assay.

2. LITERATURE REVIEW

2.1 Rye crop

Rye *(Secale cereale L.)* is a cereal crop which belongs to Poaceae family and has extensively been grown to produce grain, often used as cover crop and fodder crop to feed livestock (Deleu et al., 2020). Rye has high resemblance to *Triticeae* tribe, mainly barley and wheat because of the genetic similarity of its grains (Wrigley et al., 2017). Rye is allogamous as its reproductive strategies promote outcrossing and is diploid (2n) in nature with 14 chromosomes (Dolezel et al., 1998). Full genome sequencing was achieved (Li et al., 2021 and Rabanus-Wallace et al., 2021) which will facilitate the rye breeding and biotechnological approaches. It has largest genome size among diploid cereals around 7.9 GB with almost 90% repetitive sequences (Baueret al., 2016). Rye has widened its horizons in research by the applications of modern breeding techniques like hybrid breeding, mutation breeding, speed breeding and most recently the precision breeding (Wilde et al., 2021). As rye is resistant to cold/frost and requires less nutrients for its growth, it has the potential to be grown in areas where wheat or barley may not survive (Bushuk et al., 2001 and Jung et al., 2014). It is grown mostly in temperate regions of the world especially in Europe and Northern America constituting about 90% of its cultivation (Figure 2.1) (Pariona et al., 2017). Rye is being grown in most parts of the world due to its endurance rapidly, it is an annual crop with leaves being linear justifying its height of approximately 2 meters. Rye has high nutritional value especially in starch, fiber and protein but has been neglected by researchers and breeders for many decades due to focus on major cereals (corn, wheat and rice).

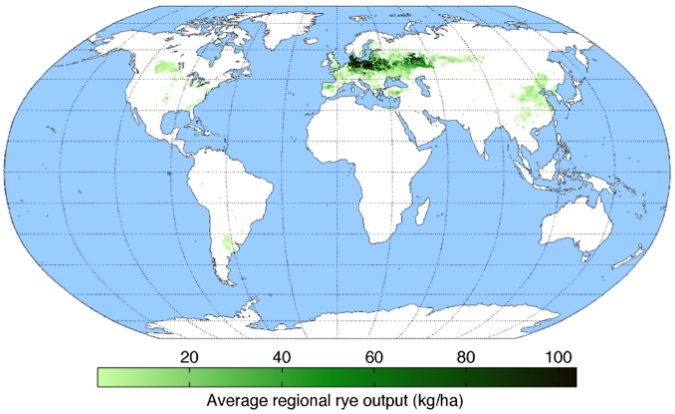

Figure 2.1. Area covered for rye production across the world, shown by the green color on world map (Monfreda et al., 2008).

2.1.1 Background and Distribution

Rye was domesticated from its wild relatives (*S. segetale* or *S. vavilovii*) which were grown as weeds, in southwestern Asia (Anatolia or Syria, the Euphrates River valley) 6600 years BC. Rye cultivation history goes back to 13000 years making it one of the oldest cereals to be grown by humans (Hillman et al., 2001). Rye was domesticated on few lands in the Asian regions and Neolithic sites (Anatolia, now Turkey) but the cultivation history was not discovered in the human history books until the 1800-1500 BC in central Europe (Zohary et al., 2012). Probably, rye traveled to west from Asia Minor thought to be wheat and later cultivation started on its own as a separate crop (McElroy et al., 2014). Ancient cultivation records of rye cultivation leads to Rhine, Danube, Ireland and Britain (Gyulai et al., 2014). Modern rye is grown extensively in Europe, Asia, and North America.

2.1.2 Nutritional value

Rye cultivation occupies almost 6 million hectares in Europe and North America thus being the valuable crop for these regions and is utilized for different purposes like making bread, beer and several products for bakery. (Németh et al., 2021). European countries (Northern and Eastern) entails majority of area for cultivation thus has largest consumption rate in making bread and beer from rye accounting more than 3 million tons of rye annually (Hansen et al., 2004). Rye contains several

components rich in nutrients which are beneficial when consumed by humans such as proteins, starch and vitamins (Figure 2.2). Rye has the highest percentage of fibers which is proven to be helpful in mediating healthy life style and comprising different health benefits like sustaining the body weight, helpful in minimizing chances of diabetes, reducing cardiac problems and helpful against cancer (El-Mahis et al., 2022). Bread made form rye is specially recommended by doctors to solve problems related to diet because of rye's richness in phytochemicals (Martinez-Villaluenga et al., 2009). Because of being rich in fibers and sugars, rye is often considered valuable as raw material for making nutritious and fresh food products among the cereals (Farag et al., 2022). Dry matter (seed hemicellulose) for rye is higher among the cereals which is responsible for surviving in harsh conditions and serving to have better storage duration compared to wheat and triticale (Alijosius et al., 2016).

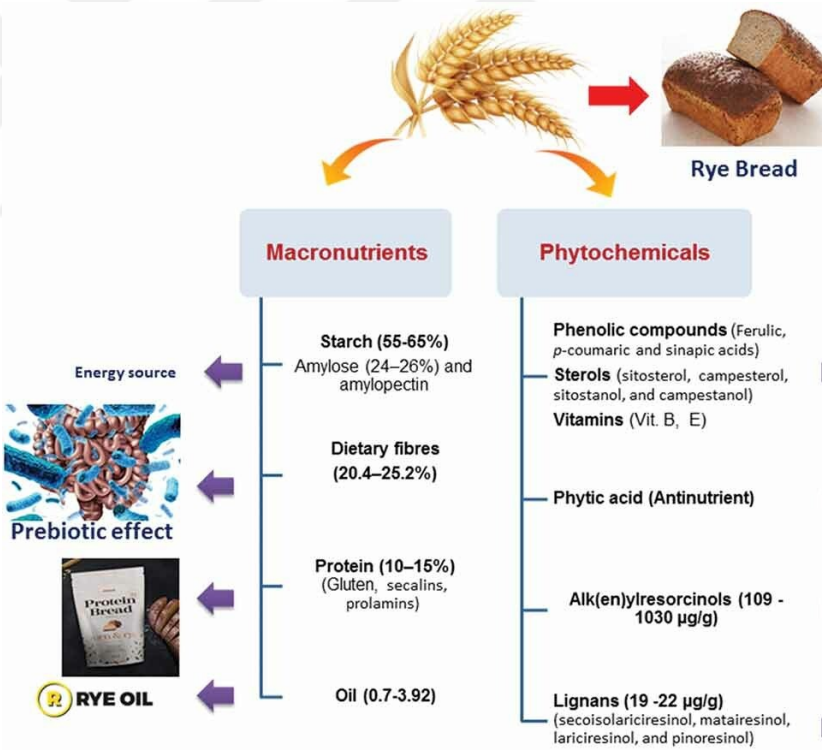

Figure 2.2. Percentage of macronutrients and phytochemicals in rye bread along with other helpful components for human health. On the left side of image, I can see the useful products for prebiotic effect of gut microbiota, protein rich products and rye oil for consumption (El-Mahis et al., 2022).

Considering the world population shifting towards more plant-based diet to avoid environmental hazards, rye has not been consumed in appropriate percentage

keeping in mind its high nutrional benefits. In recent times, accounting on usefulness of rye as food source, it has become major cereal because of its rapid growth under low key conditions and ability to yield more among other cereals in under developed areas. Along with having better production than other cereals, rye is also considered an important source of protein and fiber in rural/less developed areas where expense of growing major cereals are not affordable for people (Figure 2.3). Seeing the positive growth of rye in terms of production and nutrients, it is being considered to be grown in major/developed areas as well with being effective as food source and having micro and macro nutrients which can help prevent gut microbiota and are rich in bioactive compounds (Gomand et al., 2011). Considering the nutrient availability from rye, doctors often prescribe to add it to the diet plan. Processing of rye flour is being critically improved due to toughness of making dough compared to wheat and maize, different bakery products are being produced to enhance loaf volume and crumb structure for improved nutrional value as processing steps for bakery production hold important value for final product. Rye being exposed to all novel approaches for processing and production, scientists are always looking to make rye novel and delicious products with improved nutrients, high market value, improving rye as a crop through genetic improvement and making it more acceptable and valuable for people to consume in their daily diet (Kaur et al., 2021).

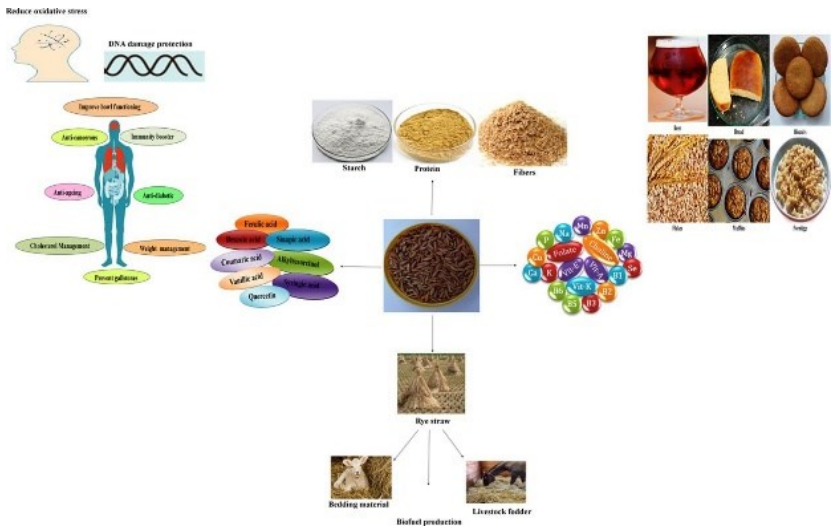

Figure 2.3. Rye importance for different aspects in daily routine showing how it can replace major cereals for various purposes (Kaur et al., 2021).

2.1.3 Technological advancements in rye

Biotechnological tools in rye are not well understood or developed. Poor focus on rye crop from scientists probably is due to its high self-incompatibility rate leading to poor breeding with available techniques and high inbreeding depression made it worse for rye to rise as a crop (Targonska et al., 2013). Prevailing breeding techniques like hybrid breeding, plant transformation approaches, haploidy technology and mutation breeding can improve the crop's aura for scientists to proceed further. However, rye has proved to show limited success in mentioned techniques due to its high recalcitrance to transformation and regeneration abilities from cultured explants *in vitro* (Ma et al., 2003). Tissue/cell culture techniques are present for other cereals like wheat and barley but similar approaches do not work for rye thus making the technological advancements in rye very difficult (Eudes et al., 2003). Additionally, mentioned methods are highly dependent on genotype in rye as tried before with different study examples (Linacero et al., 1990; Trojanowska R et al., 1995; Popelka A et al., 2001). Plant transformation and regeneration in rye is key to opening plethora of opportunities and a method for regeneration was developed through somatic embryogenesis but was genotype dependent so didn't work for a significant success in rye (Lowe et al., 2016). For hybrid breeding which is mostly used for almost all major crops, rye crop still lacks a robust and reproducible double haploid/cell culture production protocol which might open ways to improve the crop itself. To explain the reasons behind the failure of such protocol, factors affecting the culture techniques at each stage of development should be understood to address the problem and to find a long lasting solution (Dubas et al., 2015; Zimny et al., 2018). There have been efforts in the direction to improve the cell culture technologies like somatic embryogenesis but further understanding has to be done.

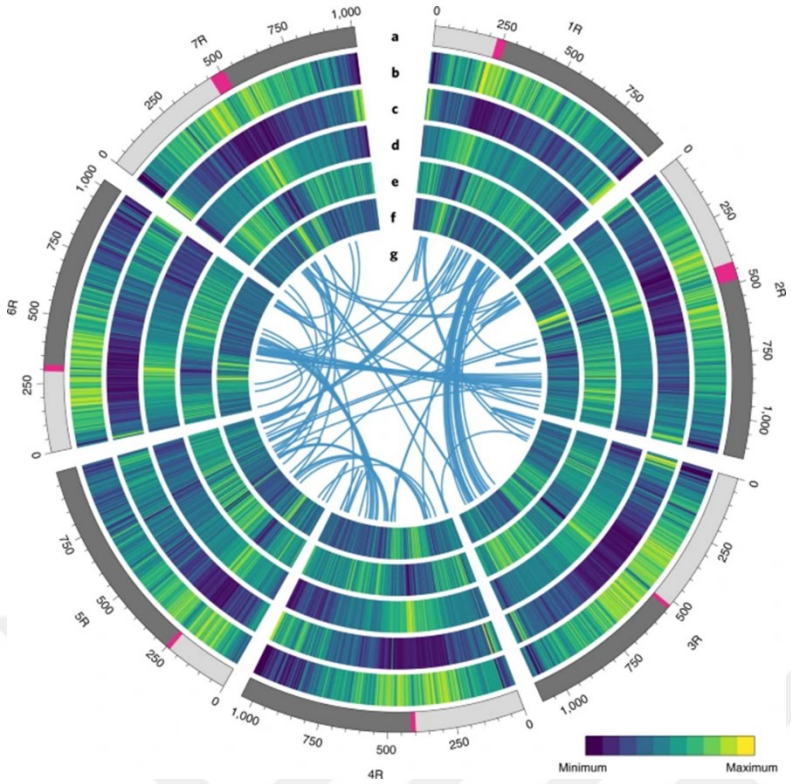

Figure 2.4. Sequenced genome display of Weining spring genotype sequences in China. Layers in the circle show the chromosome details like size and number. I can see the centromere marked in red for each chromosome displayed (Li et al., 2021).

Because of the poor understanding of the genetic basis of rye species, it was hard to achieve significant improvement in such a recalcitrance plant in approaches like plant regeneration and transformation (Miedaner et al., 2019). Though rye genome has recently been published in two genotypes in year 2021, Lo7 winter rye inbred line with genome size of 7.9 GB (Rabanus-Wallace et al., 2021) and Weining early flowering variety with genome size of 7.86 GB in China (Li et al., 2021) which opens new windows for new methods to be tried and increases the success chances for improvement. Rye has the largest genome size when compared to barley and diploid wheat species containing 90% of repeating sequences (Bartos J et al., 2008). With the sequenced genome of rye (Figure 2.4), many potential aspects uncovered like the genotypic information behind the traits, resistant genes for rye diseases, agronomic traits can be studied better, factors/reasons controlling regeneration and growth can identified and understood and much more with understanding the genome but that will take series of experimentation to validate what theoretically is correct (Rabanus-Wallace et al., 2021). There can be possible new outputs at individual genes having the coding sequences, above mentioned traits can be best

exploited with manipulation of identified genes. Sequenced genomes play an important role in understanding the genetic diversity and evolutionary outputs which become handy while improving the genotypes according to the climatic demands. With that, new ways of improvement like CRISPR-induced targeted mutagenesis, haploidy induction, hybrid breeding & high throughput MAS (marker assisted breeding) are on their way to make revolutions in this crop (Hackauf et al., 2017).

2.2. RNA Guided Customized Cas9 Endonuclease

CRISPR (Clustered regularly interspaced short palindromic repeats) is a defense mechanism that is evolved from bacteria and archaea to cope with the foreign particles/species of viruses entered to damage but made inactive by the bacterial immune system. It is a coping strategy of bacteria for invading viruses or other harmful objects, extracted/adopted the defense mechanism and being used for our desired purposes. CRISPR system of type II from bacteria has the capacity to generate site specific mutations in target DNA with the help of guide RNAs (gRNAs) which direct to target site and upon reaching, Cas9 protein performs cleavage activity and further DNA is repaired with different repair mechanisms. This type II CRISPR technology is also called RNA guided Cas9 endonuclease technology (RGEN) and is extensively used for genetic engineering purposes and was adopted from *streptococcus pyogenes* (Kumlehn et al., 2018). Since the system is modified from bacterial immune system, it can be applied to large number of plant and animal species. In RGEN system, Cas9 protein makes a complex with two types of RNAs (crRNA & tracrRNA) which holds the capacity to cleave any foreign coming particle in bacterial immune system and similar approach is employed for cleaving the target DNA *in vitro*. Each type of RNA in the RGEN system, has a defined role like crRNA (CRISPR RNA) entails the specificity to the target site by having the complementary sequences which is called a spacer and provides RGEN complex the facility to cleave at target sites (Jinek et al., 2012). These two RNAs can be replaced by a single guide RNA which later makes the gRNA/Cas9 complex that completes the tasks of targeting at the specific site to induce mutations. This RGEN system overcomes the barriers faced by zinc finger

nucleases (ZFNs) and TALENs where cleaving at target site was often not achieved due to non-specificity of RNA molecule (Cho et al., 2013).

RGEN system is comprised of two domains, REC (recognition domain) and NUC (nuclear domain) which are usually named lobes as well for this type of system (Figure 2.5) (Jinek et al., 2014). Double strand break is induced with the help of NUC which is divided further in RuvC and HNH subdomains and the later catalyzes the cleavage activity. As we know that DNA is double stranded, these two subdomains are responsible for inducing break in each of the strand, HNH for complementary strand and RuvC for non-complementary strand (Jinek et al., 2012). After the characterization of the RGEN system, gRNA/Cas9 complex generates DS-breaks three to four nucleotides upstream of the PAM (Protospacer Adjacent Motif sequence, three nucleotide sequence that are important for identification of target DNA). Due to the nature of the CRISPR/Cas9 system, the effector protein is able to target simultaneously multiple loci. There is a possibility for multiple target sites as well by having more than one gRNA at the same time.

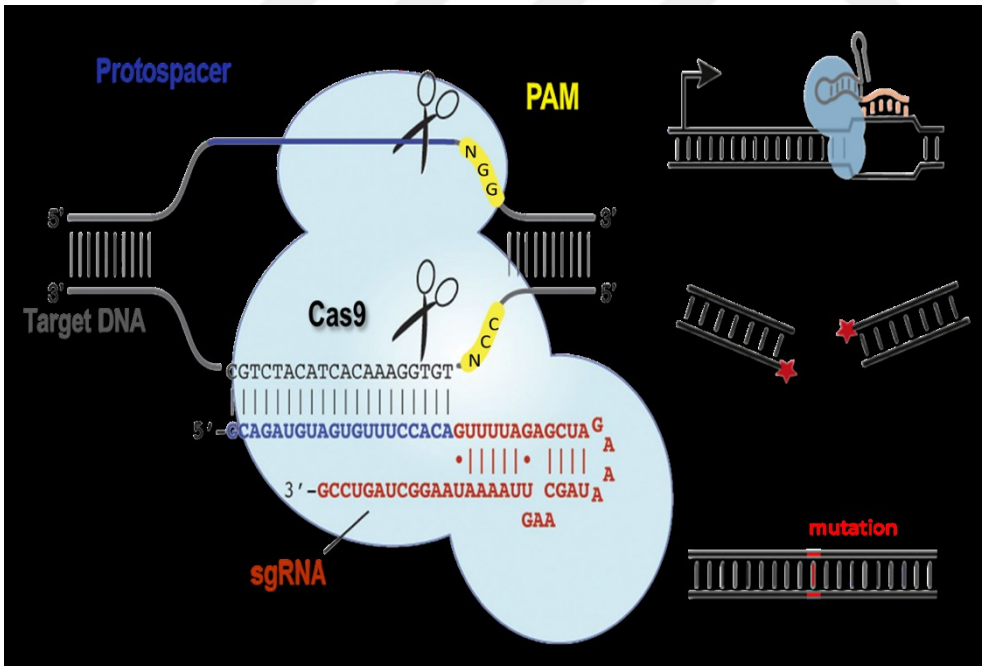

Figure 2.5. Cartoon representation of RGEN showing the activity of Cas9: gRNA complex at the target site of DNA. DNA repair occurs usually with NHEJ (adopted from Jinek et al., 2013).

2.2.1 Gene knock-out through CRISPR-Cas9

Deleting/eliminating a gene from an organism is often referred to a knockout making that genomic region non-functional. It involves perturbation of genomic DNA so that the expression of a specific gene is permanently prevented (Reski et al., 1998). Several approaches are there to make a gene knockout like RNAi, but RGEN (RNA guided Cas9 endonuclease) is being extensively used these days to generate knockouts. For generating CRISPR knockouts, 20 nucleotides specific region in target DNA is selected and often is referred target motif (TM) (Casini et al., 2018). As explained in previous section about the RGEN technology, gRNA/Cas9 complex is formed for generating site specific DNA breaks. The gRNA/Cas9 complex starts the cleavage upon annealing to target site and continues annealing in 3' to 5' direction if target motif has the similar sequences as of gRNA (Peters et al., 2019). Possibility of mismatches is always there and if mismatch happens at 3' end, cleavage will stop but it will continue if it's on 5' end (Jinek et al., 2014).

Cells try to repair DSBs to avoid cell death which otherwise results from loss of genomic integrity. After induction of double strand break, DNA repair happens and during the repair mechanisms, different types of mutation occur like Indels after the introduction of stop codon (Eid et al., 2018). Usually there are two pathways for DNA repair, homology directed repair (HDR) which needs a complementary strand for repair and non-homologous end joining (NHEJ) pathway which does not require complementary strand (Shalem et al., 2014). Indels are formed upon NHEJ repair and can lead to frameshift mutations for changing/silencing gene function (Figure 2.6). Gene knockout remains a vital approach for understanding biology. By disrupting a specific gene, scientists can discover the effects and functions of that gene (Qi et al., 2013).

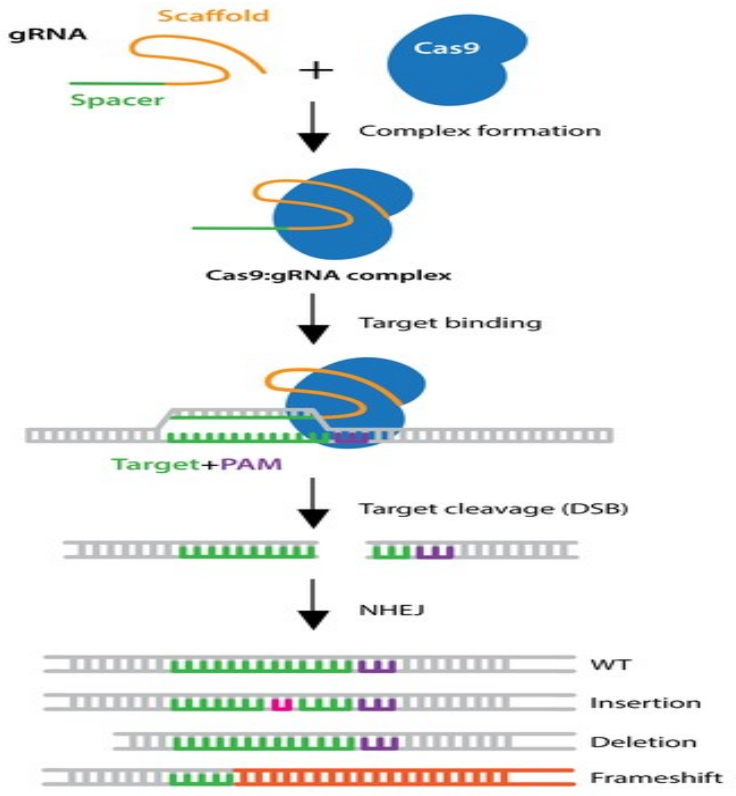

Figure 2.6. Schematic representation of Gene knockout showing the Cas9-gRNA complex formation, target binding to genomic DNA, cleavage of target site and repair strategy for different types of mutations by NHEJ (Non-homologous end joining)

There are different types of gene knockouts like insertion, deletion, inversion and frameshift mutation (Figure 2.6) (Pipiras et al., 1998). Any of the mentioned possibilities can occur that can lead to the knockout of a genetic part. To alter the phenotype of plant, frame shift mutations are considered ideal as function of the gene is altered due to the defined role of a stop codon.

2.2.2 DNA repair pathways

As DNA is hereditary material and holds the central position for the survival of a cell, maintaining its integrity becomes obligatory for the cell. There are several occasions where cell goes through a stressful time hence the DNA as well and it bears certain types of damages like strand breaks and repairing becomes indispensable (Lindahl et al., 1982). If cell wants to survive, it must have DNA repair mechanisms otherwise death is certain (Friedberg et al., 1995). Model organisms have provided us great platform to study these repair pathways and

understand the mechanisms which can occur in higher organisms. After that understanding, all genome editing platforms employ DNA repair pathways and engineer them to our desired outcome like CRISPR mediated base editing, setting the efficiency of these technologies (Figure 2.7). In CRISPR induces genome editing, different factors influence the repair pathways like DNA structure, cell cycle phases etc. (Xue et al., 2021).

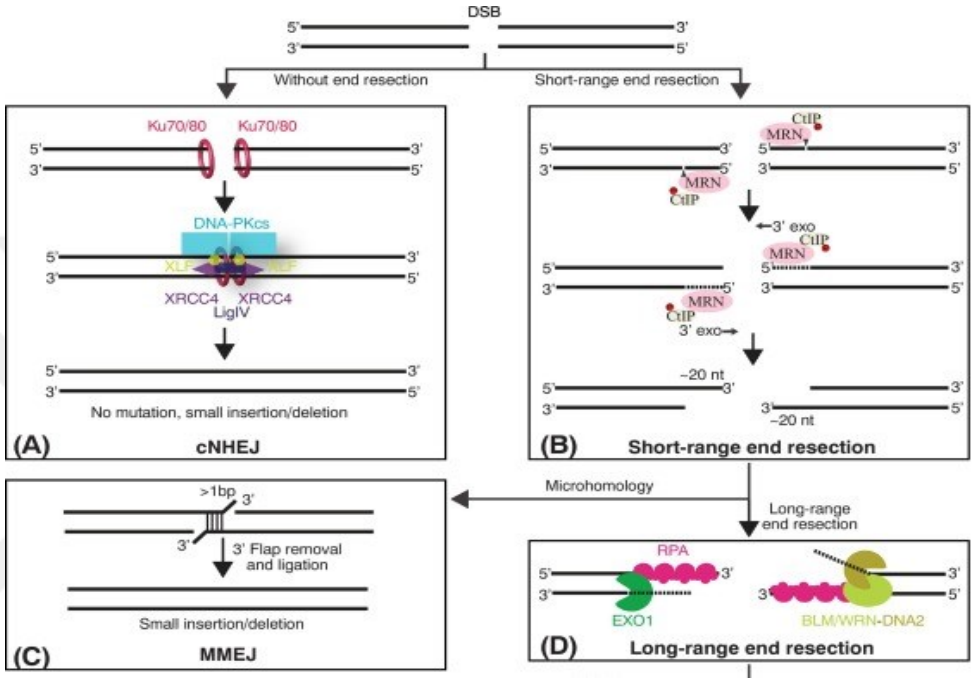

Figure 2.7. Commonly used DNA repair pathways after the induction of double strand break. Mainly there are three pathways which are employed after CRISPR technology and induces mutations. In the image, we can see all three major repair pathways with factors which influence the repair processes. (Xue et al., 2021).

2.2.2.1 Non-homologous end joining (NHEJ) pathway

NHEJ is responsible for inducing small mutations because of its inaccuracy as it involves the ligation of double stranded blunt ends which does not include large nucleotides (Chang et al., 2017). Sometimes when DNA breaks are joined automatically/naturally, other catalysts like DNA polymerases etc. are required for end processing to make sure the integrity of DNA (Bernstein et al., 2005). After cleavage, gRNA/Cas9 complex sets free the proximal end while holds on to the distal end, making free the proximal end for repairing but distal end is blocked (Clarke et al., 2018). There are few uncertainties about the functioning of NHEJ

like how Cas9 protein is released and how this affects the repair mechanisms itself after the activity of RNA polymerase to remove Cas9 from the distal end (Wang et al., 2020) and yet to be answered.

2.2.2.2 Microhomology directed repair pathway

It works on the principle of NHEJ but small sequences are required along the target DNA site for the repairing purposes and usually larger mutations are expected from this type of repair pathway. Initially the realigning of strands happens and then repair process itself starts (Sfeir et al., 2015). This repair pathway cuts any loose blunt ends and then starts the synthesis of DNA by ligating the strands with the help of microhomology sequences (Kent et al., 2015). The researchers have not yet completely understood the binding process of microhomologies but a general idea of how it works have been illustrated. Scientists have narrowed it down that about 58 percent repair happens through micro homologous sequences and the type of repair mechanism can be predicted according to the target DNA and application employed for genome editing often starting from deletion of at least two nucleotides that leads to precise edit (Shou et al., 2018).

2.2.2.3 Homology directed repair pathway (HDR)

Complementary strand is required for this type of repair and has significant role in meiotic divisions. This type of repair can prove to be lethal causing risks to genome integrity like cancer etc. if any defect occurs in the processing of the mechanism or in the helicase activity as it is mostly functioned by that. Large damages to the DNA often lead to this type of repair pathways. In the case of double strand break mediated by CRISPR, two Holliday junctions are made in DNA intermediate that initiates the repair pathway and leads to large mutations. Homology directed repair pathway comprises long range resection to form 3’ single stranded overhangs with the help of complementary DNA strand which is finally used for repair process (Chen et al., 2022). Often meiotic recombination and crossovers can lead to this type of repair pathway which usually is not applied for induced genomic mutations.

2.3 Development of Haploidy Inducer (HI) Lines

Individuals with half chromosome number are considered haploids and are utilized further to make double haploid individuals which contain the information from one parent only and this phenomenon usually takes place in plants. Genetic information in haploid plants is very obvious and all the major changes in the genome can be phenotypically seen quite well (Meng et al., 2021). Having the genetic information easily apparent on phenotypes, selection of individuals becomes very easy with great efficiency compared to the conventional lengthy breeding process (Forster et al., 2007; Geiger and Gordillo et al., 2009). Once the haploids are developed, their genome can be doubled to make double haploids (DH) by the application of colchicine. Haploidy technology is quite powerful for several approaches like mapping population, generating reverse breeding and substitution lines and doing the changes for apomixis (Wijnker et al., 2014). For the development of haploids, there are five major techniques that have been employed till date which are making haploids through tissue/cell culture (Wang et al., 2000; Chen et al., 2016), generating lines naturally (Goodsell et al., 1961; Chase et al., 1963), inducing haploidy through CENH3 (Kelliher et al., 2016), using some lines which have the natural ability to induce haploidy (Coe et al., 1959; Evans et al., 2007), and hybridization among different species (Kasha et al., 1970; Barclay et al., 1975; Laurie et al., 1988). Development of haploids through haploidy inducer lines employ two hypotheses, after fertilization chromosomes are eliminated during zygote development (Sprague et al., 1929 & 1932). Sperm cell from the male parent fertilizes the female gamete but fails to integrate its genome hence giving birth to haploid embryos. Hypothetically, chromosomes of inducer lines are deleted when zygotic divisions are taking place and in result only haploid cell with single set of chromosomes is obtained (Wedzony et al., 2009). There have been recent advents to induce haploidy that applies genome editing techniques to fasten the process and make it more efficient.

Haploid induction through genome editing has recently been established by eliminating or knocking out the genes responsible for proper fertilization and is often referred to HI-Edit. Mostly CRISPR technology is utilized for doing such acts and one maize line was produced for haploidy induction by deleting the *PLA1* gene

which was identified for improper pollen functioning and causes haploid production (Figure 2.8). After the production of haploidy inducing line, it was cross pollinated to the anther inbred line for generating haploid seeds. Colchicine was used for doubling the chromosomes to induce double haploids with the desired traits which was the ultimate goal. This technology has proved to be time efficient as we can skip the lengthy and laborious procedure of manual back-crossing and harvest maximum output in assembling the desired gene combinations in generated double haploid line. As CRISPR technology allows to edit multiple genes simultaneously, species which have multiple genes responsible for haploid induction can also be treated with the same approach and similar outcome can be expected. Now, genome editing is becoming doable for the species which are recalcitrant to plant transformation as well and scientists are achieving good efficiencies in generating such technologies (Kelliher et al., 2019; Wang et al., 2019a).

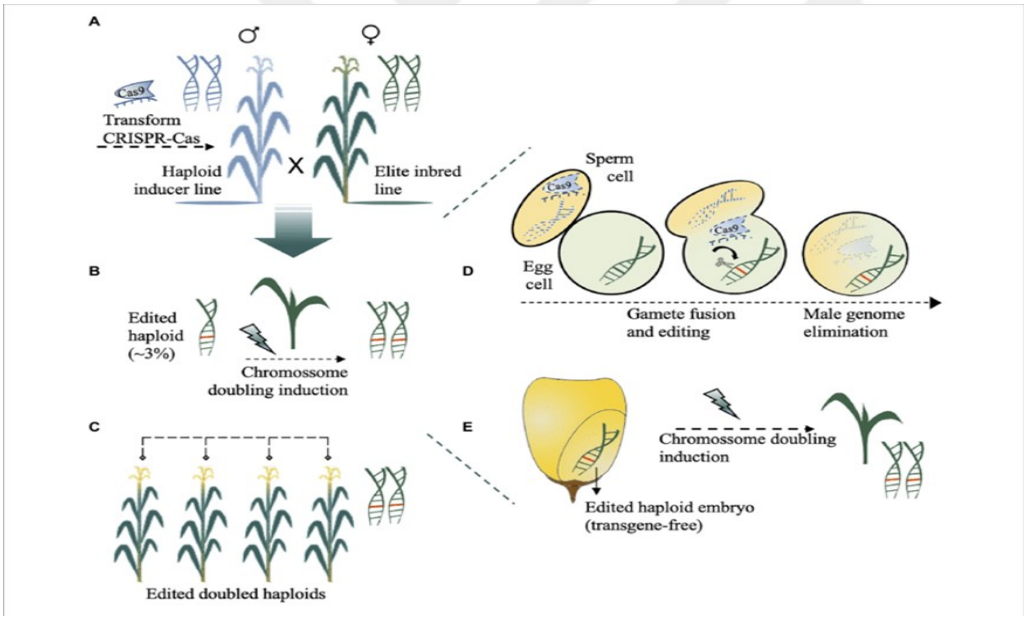

Figure 2.8. Haploid technology through CRISPR mediated genome editing is depicted in the image. Different steps are shown in the picture for complete illustration of how the technology works (Yassitepe et al., 2021).

2.3.1 *PLA1* gene in cereals

Phospholipase A1 (*PLA1*), is found to be conserved in a wide range of species, is a type of an enzyme which breaks down the phospholipids and generates simple fatty acids which play important role in lipid signaling (Aoki et al., 2007). As family of *PLAs* is conserved in many species, their role is of great importance in transferring the lipid breakdown and signaling, certain aspects of plant growth and developments and in few stress conditions. The plant phospholipases are distributed in four classes/families which are: phospholipase A1 (*PLA1*), phospholipase 2 (PLA2), phospholipase C (PLC) and phospholipase D (PLD) (Wang et al., 2012). The *PLA1* class in not defined that well in the plants. This family has been recently elaborated in plants and plays a significant role in inducing haploid technology. A recent study shows that, the maize line (Stock6) can induce haploidy in its relevant species and genotypes which is a natural haploidy inducer and was discovered about 50 years ago. With the recent advent of genomic technologies, it was found that the line has the capacity to generate 1-2 percent of haploidy induction when used as a male parent (Coe et al., 1959). With the help of genome sequencing, responsible regions for haploidy induction were characterized in the line (Stock6) and several quantitative trait loci (QTL) were identified, qhir1 was found to be the QTL as the largest contributor for inducing haploidy in the line (Stock6) (Prigge et al., 2012, Dong et al., 2013, Hu et al., 2016). Further characterization and validation illustrates the function of this loci, making the plant to abort embryos and endosperms along with distorting the segregation alignment leading to haploidy induction in the plant (Dong et al., 2013, Xu et al., 2013). Initially 13 genes were located in qhir1 loci and fine mapping studies showed that *PLA1* gene contributes the most in inducing haploidy plants making it the candidate gene. This work was done in maize with knocking out the *PLA1* gene resulting in haploidy induction of about 5-6 percent validating the role of this gene.

Recent study shows the validation for *PLA1* gene knock out starting with the characterization of qhir1 loci and narrowing down to pollen specific *PLA1* gene being the highest contributor causing 4bp frameshift mutation for haploidy induction in maize (Figure 2.9). New ways of inducing haploid plants may be found by targeting the *PLA1* gene in crop plants as this gene is conserved in wide range

of species and can help to fasten the conventional breeding process (Kelliher et al., 2017; Liu et al., 2017; Gilles et al., 2017).

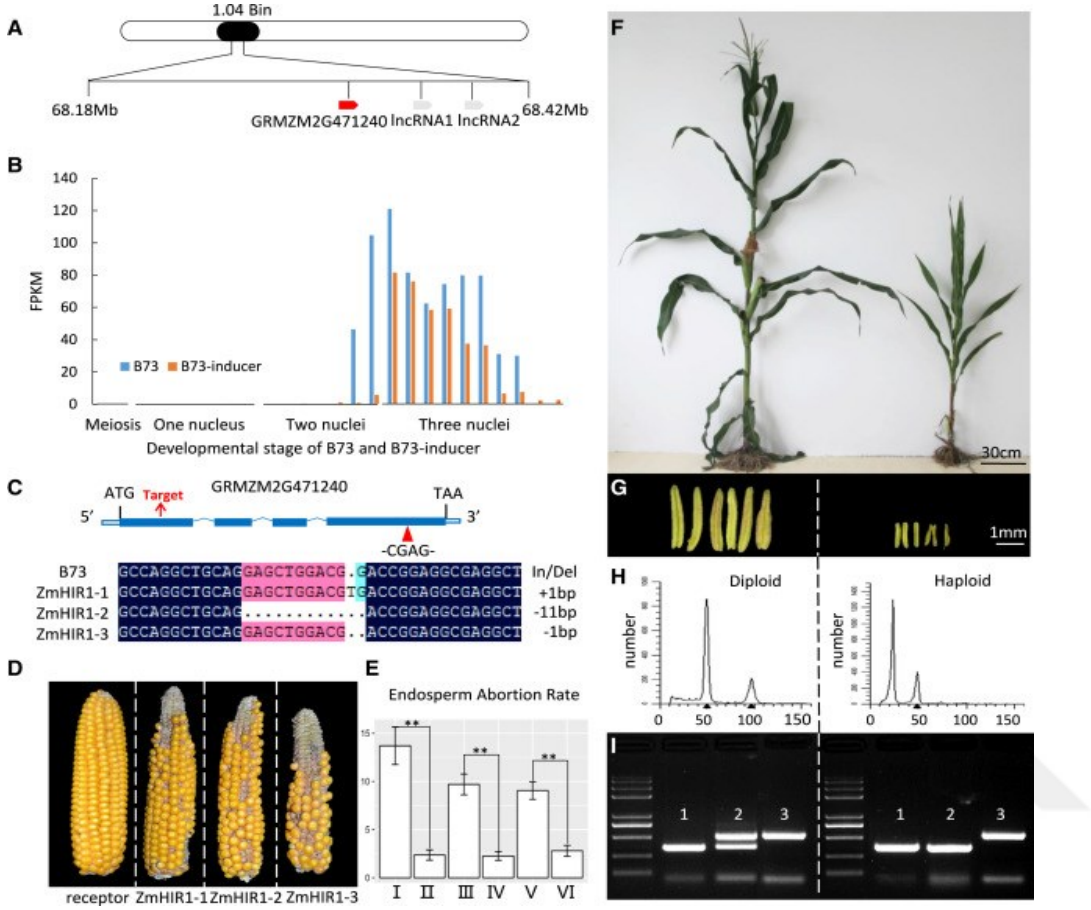

Figure 2.9 Mutation caused by knocking out the *PLA1* gene by CRISPR application in maize. The picture shows complete illustration of PLA1 gene function starting from the genetic map of responsible loci, expression of *PLA1* in different division stages, 4bp mutation in the *PLA1* gene region leading to haploidy induction. Characterization of mutation is also showed here through sequencing and PCR (Liu et al., 2017).

2.3.2 Golden gate gene cloning technique

Cloning is a toolkit which employs assembly of genetic fragments into a host organism like bacteria and utilizes host's machinery to replicate the genetic fragments inside the host. Initially the host organism is infected with a gene of interest and then put to a selection process for screening the bacterial colonies carry the gene of interest. Screened bacterial colonies are than cultured for replicating the gene of interest inside the host and once enough population of host material is acquired, extraction of gene of interest from the host takes place and used for

desired purposes. As two genetic fragments are combined in the process, cloning comprises two organisms, one would be the host DNA and the other is gene of interest (Watson et al., 2007). Once we have acquired enough DNA material from the host, this genetic fragment can be put for testing its function through different expression of proteins after being inserted into a plasmid vector. Different strategies are used for cloning purposes like golden gate, TA cloning and gateway.

Golden gate is a strategy which employs the integration of multiple gene of interests (GOIs) in a single plasmid vector with the help of PCR (Figure 2.10). Both the restriction digestion and the ligation is performed in PCR saving time and energy for doing both the procedures separately. For the integration of multiple GOIs in a single vector, overhangs are designed for each fragment in a way that ligation happens only among the compatible genetic fragments, and restriction digestion is performed with the help of type IIS restriction enzymes which helps the formation of overhangs (Engler et al., 2009). As restriction digestion and ligation are performed in a single PCR step, it makes it very efficient and cost saving strategy along with keeping intact the reaction integrity. Once the reaction is completed, overhangs created by the endonuclease enzyme activity are eliminated being just there to help the ligation of different assemblies (Casini et al., 2015). Golden gate cloning is considered error free because the recognition sites are not there anymore after the final reaction making it irreversible having no chance of mismatching due to overhangs compatibility (Werner et al., 2012). In this study, RGEN technology has employed this strategy because of its efficiency and cost effectiveness, making sure the insertion of gRNAs in plasmid vector with great success and very less chances of error. Assembly of multiple guides in a single vector is also possible with this strategy because of the overhangs for different guides with different sequence compatibility. Recognition sites for the restriction enzyme can be found in wide range of nucleotides inside the DNA, so making sure of no restriction sites presence in the fragment being used for cloning plan is indispensable for the success of this strategy. Regardless of the minor caution for this strategy, CRISPR applications employ this cloning plan for large group of experiments because of its effectiveness (Engler et al., 2014).

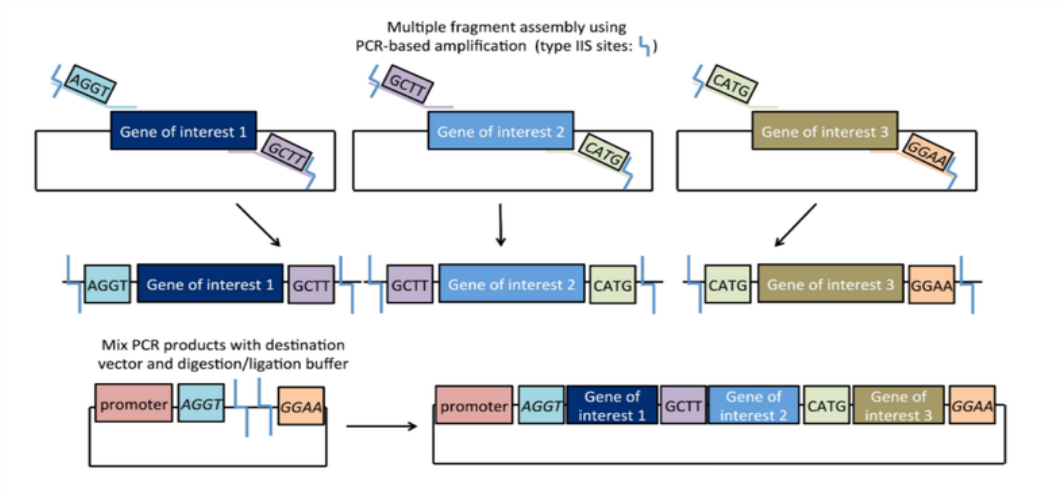

Figure 2.10. The sketch of golden gate cloning showing the integration of multiple GOIs (gene of interests) in a single plasmid vector. Single PCR reaction for restriction digestion and ligation (Gearing et al., 2015).

2.3.3 Protoplast transfection for transient gene expression

Protoplast transient expression assay is a powerful tool that has been widely used for rapid gene functional analysis, an efficient platform to validate the function of genetic material for the application of CRISPR techniques (Wedzony et al., 2002). Protoplast cells also provide an educated idea for the type of mutations that might occur in the stable mutation experiments with the help of sequencing facilities. Before conducting plant stable transformation, protoplasts provide a platform to check mutation efficiency rapidly (Zhu et al., 2020). Because the protoplast cells lack the cell wall, are considered an efficient media for checking the expression of protein or a single gene since they can take foreign particles rather easily compared to normal plant cells (Yoo et al., 2007). As these protoplast cells can bear the foreign particles for short time period only, different mechanisms/rapid tests can be made such as localization of protein, different sorts of gene and protein interactions etc. This rapid gene validation technique provides a great platform for optimizing protocols for different CRIPSR applications and helps design the different vector modules which can be helpful for efficient mutation induction (Nadakuduti et al., 2021). In each protoplast experiment to check the validity of a gene function, at least 100,000 cells are infected depending on the species, making it a complicated procedure to cultivate desired result as not all cells are mutated during the process and sequencing results show mostly chimerism in mutation

induction (Figure 2.11) (Lin et al., 2018). This method, although accurate, is expensive and time consuming. To assess the protoplast transformation efficiency, a fluorescence marker like *GFP* gene can be cloned together with gRNA/Cas9 complex in one vector. The use of a fluorescence-marker will indicate the quality and efficiency of protoplast transformation step (Yue et al., 2021).

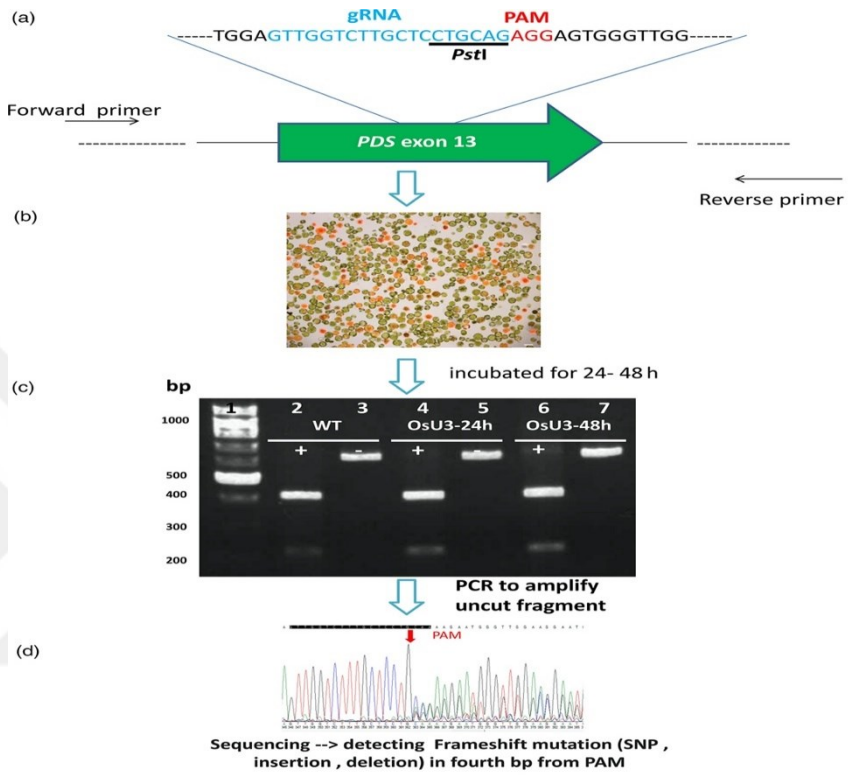

Figure 2.11. The image depicts the complete illustration of a protoplast reaction with the validation of mutations through PCR and Sequencing platforms. There are four steps shown in the picture starting from the vector design with gene of interest to test, transfected protoplasts, harvest of target specific amplicons and finally sequencing to check mutation types (Lin et al., 2017).

2.3.4 Mutation identification

Once the mutations are induced through the CRISPR systems, identification of mutation type becomes essential to validate the experiments and see how the gene has performed. There are different methods which are employed in the identification process, the very simple one is using PCR machine to detect in the genomic DNA any mutations at the target site with the help of a restriction enzyme or making the adjacent primers for target site to be mutated. Amplifying target

specific region and running the PCR product on gel electrophoresis gives an idea for mutation existence. Different sorts of gels are used to identify particular type of mutations like PAGE which helps to distinguish mutated alleles after the CRISPR applications (Li et al., 2017). However, different types of gels cannot distinguish single base changes inside the DNA fragment thus making it an initial screening step and further zooming is needed for the proper identification. Once the mutations are induced, there can be homozygous (similar) or heterozygous (chimeric) mutations at the target site which help to further categorize the screening process. There are some common mutation identification methods depending upon the repair mechanism type and the application applied to cause the mutation. For identification of mixed mutations like in protoplasts, cleavage assays are generated which identify the alleles which have been mutated and vice versa. Target specific amplicons are acquired through PCRs and restriction digestion, PCR products are run on gel electrophoresis to distinguish different fragment types (Shan et al., 2020). Further screening and confirmation is done by sequencing of target specific amplicons which address the type of mutations and help to reach the experiment goals (Wang et al., 2013). For sequencing purposes, mostly Next generation sequencing (NGS) and particularly amplicon sequencing is employed which further has many types and each help to identify different type of mutations depending on the species and method of mutation induction (Figure 2.12) (Ma et al., 2016). The type of mutations usually induced by CRISPR technology are indels either insertion or deletion of nucleotides (Shan et al., 2018). Different studies show that amplicon sequencing is a powerful tool to identify several mutation types and is becoming more and more efficient, cost effective and robust making available the sequencing technology for every organism (Fauser et al., 2014; Svitashev et al., 2015; Ma et al., 2016; Zhang et al., 2019).

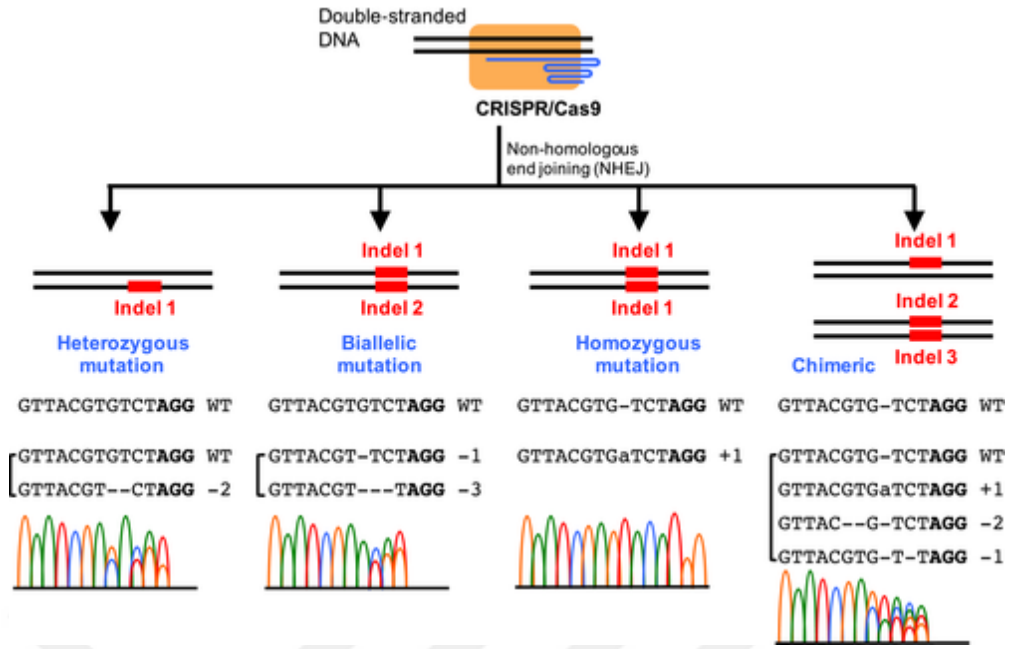

Figure 2.12. There are various mutation types shown in the image after the application of CRIPSR technology. Four major mutation types are depicted with reference to the PAM sequence shown in bold letters (Shan et al., 2020).

2.4 Research Gap ın Rye

With release of two high-quality genome assemblies (Rabanus-Wallace et al., 2021; Li et al., 2021), rye has finally reached the genomic era. As of now, we have understood that rye is highly recalcitrant to plant transformation, no genetic engineering success has been foreseen till now and that is a hindrance to making improvements in rye breeding programs despite being 2nd most important crop for bread production. In the study conducted for the Master thesis, as the sequenced genome of rye has opened new ways for crop improvement, I started to establish a protocol for genome editing in series of different experiments to try and develop genetic engineering platform in rye. Plant regeneration from the explants of rye is a major challenge leading to stable transformation and obtaining mutant plants with the improved phenotypes. With the genome sequenced, functional validation of genes has become possible that is a step forward towards applying genetic engineering approaches to gain fruitful results for improved performance in rye and that's what has been performed in this study. Once the plant reproduction from explants is achieved, it will open an era for genome-based applications in rye making an outbreak in rye breeding programs.

3. MATERIALS and METHODS

3.1 In-Silico Analysis of *PLA1* Orthologue

The phospholipase (*PLA*) gene family is present in a wide range of species including cereals where it has been shown that the knockout of the gene causes the haploidy induction such as in maize (Kelliher et al., 2017; Liu et al., 2017). *PLA1* gene hydrolyzes lipids, which plays an important role for jasmonic acid regulation, leading to the pollen abnormality and affecting the efficiency of fertilization ((Aoki et al., 2007). Sun et al. (2021) hypothesized that mutations in *PLA1* can trigger later-stage-specific reactive oxygen species that cause chromosome damage and hence the loss of damaged chromosomes during the fertilization process leading to generating haploid plants. Yet, there is no conclusive interpretations about exact function of *PLA1* in inducing haploidy in plants.

Kelliher et al. (2017) showed that knock out of *PLA1* gene in maize induced haploids up to 6.7%. First the sequence of maize "*PLA1*" gene (accession no. GRMZM2G062320) (Kelliher et al., 2017) was obtained from the database and using BLASTn (Basic local alignment search tool), it was searched against rye genome (Secale cereale Lo7 v1 pseudomolecules (2021)) on the Graingene online database (https://wheat.pw.usda.gov/jb/?data=%2Fggds%2Frye-Lo7-2021&loc=chr1R%3A290952886.436421878&tracks=&highlight=). After the confirmation of *PLA1* orthologue in rye, virtual protein translation was done and compared *in silico* against maize and durum wheat protein sequences through SnapGene software (GSL Biotech) to confirm the functionality of rye orthologue.

3.1.1 Target motifs selection

Target motifs selection was done based on important criterion to improve the efficiency the gRNA that will be designed in the following step. One of the most decisive criteria is the PAM (Protospacer adjacent motif) sequence which entails three base pairs sequence recognized by the Cas9 endonuclease and has important role in choosing the target motifs as Cas9 protein cleaves the DNA 3 to 4 nucleotides upstream of PAM sequence. The sequence of the PAM has one flexible

nucleotide followed by two Guanine (G). The extracted sequence of *PLA1* rye orthologue was tested in the Softberry website using the tool FGENESH gene-finder to predict the gene structure hence the regions of exons and introns (http://www.softberry.com/berry.phtml?topic=fgenesh&group=programs&subgroup=gfind). After gene structure prediction, online tools (Benchling, RNA fold) were used to select the target motifs and design gRNAs. The extracted coding sequence of *ScPLA1* gene was analyzed in the online open source Benchling software (https://www.benchling.com/molecular-biology) to design and select the target motifs (Figure 3.1). The following criterion were employed during target motifs selection:

- Length of target motifs (18 to 22bp, ideally 20bp),
- Design type to be a single guide to predict higher efficiency.
- PAM sequence (NGG).
- Off-Target Score based on Doench et al., 2016 that is available for NGG PAM only.
- On-Target Score based on Doench et al., 2016 that is available for *SpCas9* only.
- Target motifs starting with a G nucleotide were given higher priority to be selected.

After applying the above-mentioned criterion, target motifs with higher efficiency score were selected (Figure 3.1). In addition, few more parameters were considered to select gRNAs like percentage of GC content. There are many other parameters, which are considered useful to select the gRNAs for genome editing. Considering these parameters carefully increases possibility to higher efficiency of gRNAs.

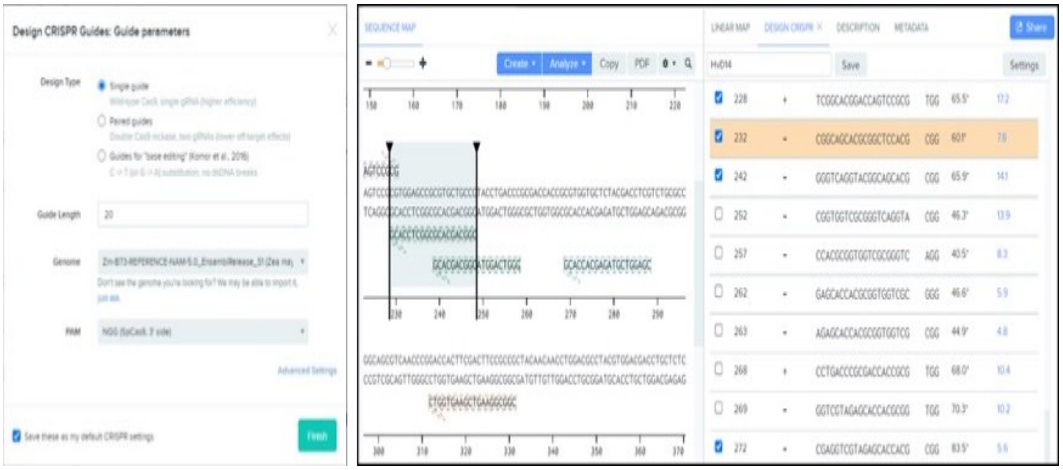

Figure 3.1. A screenshot of Benchling software representing selection criteria of target motifs. Selected sequences represent the target motifs and efficiency score is marked in blue at the right end of the picture.

3.1.2 Designing of gRNAs

After the selection of the target motifs, the next step was to design the gRNAs to forms a complex with the Cas9 protein. 5` end of gRNA is important for pairing with the sequence of the target motif of the genomic DNA. Two-dimensional (2D) secondary structures formed between gRNA and gRNA-scaffold plays significant role in base pairing of gRNA and target motif, and eventually is reflected on the activity of Cas9 endonuclease. Every candidate target motif that was selected through Benchling analysis, was tested for its 2D secondary structure features using the RNAfold webserver (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi). Online tools suggest that there should be at least 12 nucleotides, which are complementary to the target DNA and are considered enough for proper functioning of CRISPR tool (Kumlehn et al, 2018).

Importantly, three stem-loops shaped within the inalterable gRNA scaffold have been proven to be essential for proper interaction of gRNA and *Cas9* endonuclease (Figure 3.2). By contrast, any further secondary structures present in the gRNA scaffold with more or less frequency, are not considered important for cleavage function of gRNA/Cas9 complex, therefore are not taken into consideration while selecting target motifs and further in prediction of 2D structures of gRNAs.

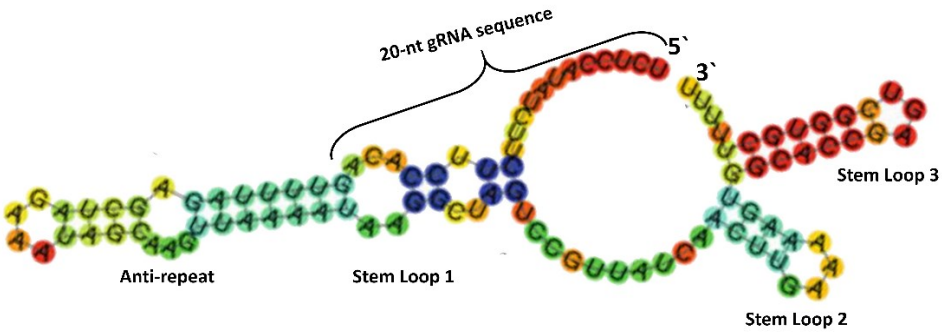

Figure 3.2. Optimal gRNA 2D secondary structure (modified from Kumlehn et al., 2018).

In this thesis study, the designed gRNAs were to be assembled together with expression cassettes of *Cas9* gene and eventually with a resistance marker gene in a binary vector using the power of an assembly modular vector system (Cascade/Cas9). Through four main steps, the CASCADE/Cas9 system in accumulative manner enables the cloning of different functional parts e.g. promoters, coding sequences and terminators.

3.2 Cloning and Construction of Modular Vectors By CASCADE/Cas9 System

During the cloning procedure, all PCR amplifications were performed using GoTaq® DNA polymerase (Promega) according to the instruction of the manufacturer. All ligation reactions were performed using T4-DNA ligase (Thermo Fisher Scientific) according to the instruction of the manufacturer. All ligated vectors were transformed into XL1-Blue heat-shock competent *E. coli* cells (Fisher Scientific) and the cloned vectors were verified via LGC Genomics (https://shop.lgcgenomics.com/) using Sanger sequencing. All planned ligation and constructed vectors were pre-designed in SnapGene® software (from Insightful Science; available at http://www.snapgene.com).

Construction of a binary vector expressing gRNA/Cas9 unit requires multiple steps to assemble various expression units for eventually performing stable plant transformation. To assemble gRNA/Cas9 vectors for transient and stable

transformation, a CASCADE modular vector system (here forth is abbreviated; CASCADE/Cas9) based on the Golden-Gate method was used. The CASCADE/Cas9 consists mainly Golden Gate methods to reduce time and efforts typically required by traditional cloning methods. The Golden Gate method relies on the type-IIS restriction enzyme for restriction digestion which has the ability to cleave DNA outside of their DNA recognition site sequence leading to a 4-nt single-stranded DNA overhang after digestion. In such a Golden Gate reaction, all DNA fragments were flanked by restriction sites for a type IIS restriction enzyme, with an inward orientation that the DNA recognition site sequences are cleaved from the DNA fragments during the digestion step. While the destination vector must contain two restriction sites for the same type IIS enzyme with an outward orientation such that the DNA recognition sites are also cleaved from the vector by the digestion step. Eventually, all single-stranded 4-nt overhangs (also called fusion sites) at the ends of the DNA fragments and on the vector were unique and complementary to allow annealing of one end to the next fragment or to the vector. The resulted circular recombinant DNA molecule was not containing restriction sites for the used type-IIS enzyme in the former step and cannot be re-digested, allowing restriction and ligation to be performed in the same reaction mix. The DNA fragments that were ligated in the initial plasmids backbone can be subjected to more cycles of digestion and ligation until being incorporated into the planned recombinant vector.

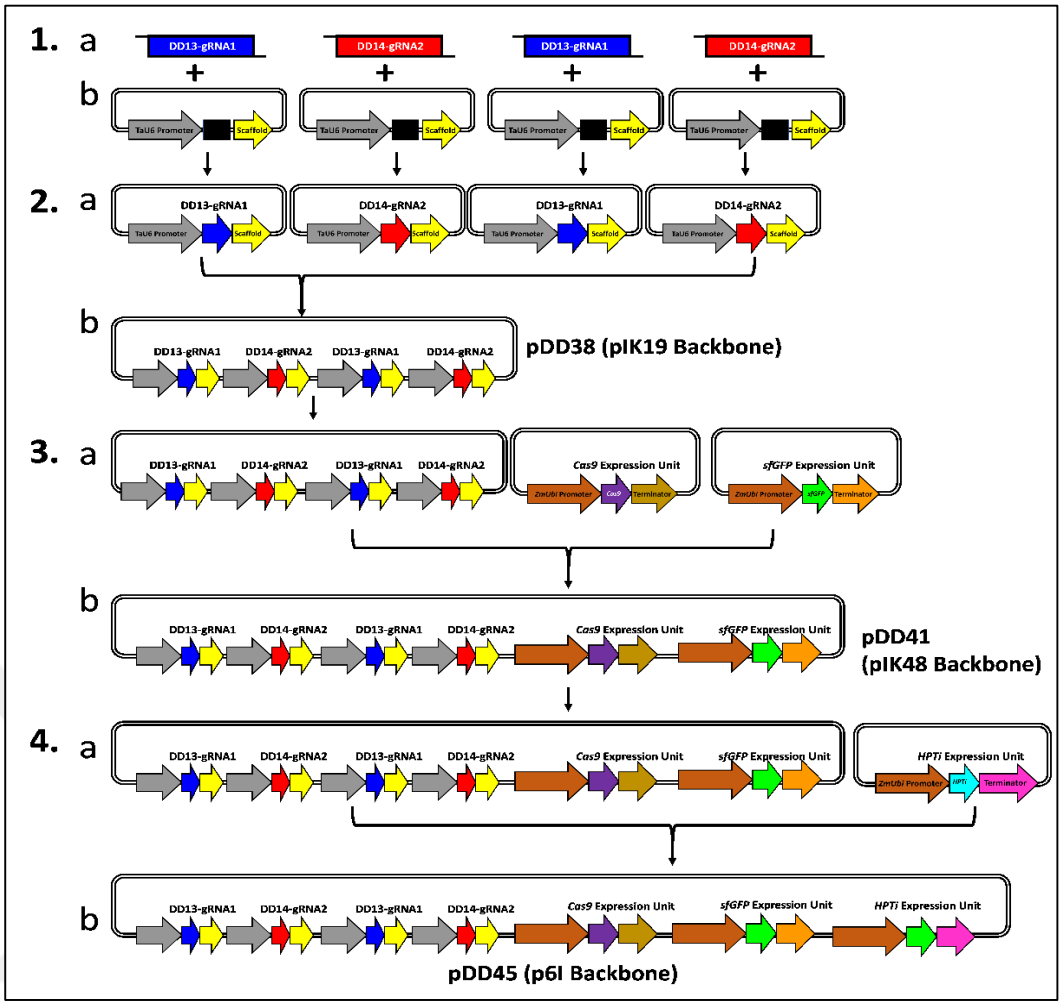

Figure 3.3. Schematic presentation of an assembled example using CASCADE/Cas9 system. There are 4 main steps indicated in the diagram. 1) Ligating the gRNA in the modular vector backbone. 2) Assembly of the gRNAs in a single vector via Golden Gate method. 3) Assembly of recombinant fragment of gRNAs together with Cas9 expression unit and sfGFP expression unit via Golden Gate method. 4) Ligating the pre-assembled recombinant fragment in binary vector (P6i) via Gibson assembly method.

In the current study, the CASCADE/Cas9 system comprised 63 modules that can be used in various combinations allowing the assembly of up to four gRNAs together with *Cas9* gene expression cassette and other expression units in a final single vector. The CASCADE/Cas9 system is designed to target either monocot or dicot plants with the advantage of optionally including *sfGFP* gene (super-fold *GFP*) as florescent marker to ease the transformation screening process. The CASCADE/Cas9 contains four main cloning steps to assemble all expression units and build the binary vector for stable plant transformation (Figure 3.3). Beside the high flexibility and variability provided by the CASCADE/Cas9 system, it shortened the cloning time and efforts significantly. The CASCADE/Cas9 was

established in the PRB-Group at IPK-Gatersleben, Germany (Hoffie I. et al., data not published yet).

3.2.1 Incorporation of gRNAs into modular vectors

CASCADE/Cas9 cloning system contains four backbone modular vectors representing four positions. All the four assembly vectors contain the TaU6 promoter, the scaffold of the gRNA and the selection gene for ampicillin resistance. In this first step, each single gRNA was incorporated in a single modular vector backbone between the TaU6 promoter and the gRNA-scaffold followed by TTTT nucleotide sequence as transcription termination signal. Eight vectors were planned to be constructed where each gRNA will be incorporated twice in two different modular vectors (Table 3.1) (Figure 3.3.1). The expected eight new vectors will be assembled in two separate vectors in the following cloning step.

Table 3.1. Vectors to be assembled to incorporate single gRNAs. Backbone of vectors and positions of gRNA shown in the table according to cloning plan.

Constructed vector	Vector backbone	Incorporated gRNA	Fragment position
pDD30	pSH215	gRNA1	Position 1
pDD31	pSH218	gRNA2	Position 2
pDD32	pSH325	gRNA1	Position 3
pDD33	pSH326	gRNA2	Position 4
pDD34	pSH215	gRNA3	Position 1
pDD35	pSH218	gRNA4	Position 2
pDD36	pSH325	gRNA3	Position 3
pDD37	pSH326	gRNA4	Position 4

All the cloning steps were performed on the bench on ice. The gRNAs oligo-nucleotides sequence (forward and reverse) were ordered as ssDNA after adding four nucleotides overhangs that were compatible with produced overhangs in the vector backbone after restriction-digesting with BsaI (type-IIS) restriction enzyme. The forward and reverse gRNA oligo-nucleotides were mixed 1:1 at concentration of 10 µl each and hybridized in thermocycler with the program in Table 3.2. The hybridized gRNAs were kept in 4 ºC until the ligation step.

Table 3.2. Thermocycler with the program for hybridized of forward and reverse gRNAs oligo-nucleotides.

95°C for 5 min	
Ramp to 85°C	1 °C/s
Ramp to 25 °C	0.1 °C/s
10 °C	∞

3.2.1.1 Restriction digestion and ligation

The backbone of the four modular vectors were digested with BsaI restriction enzyme to produce four nucleotides overhangs compatible with the hybridized gRNAs. The BsaI restriction enzyme was ordered from NEB, England (https://international.neb.com/) and used according to the manufacturer instructions. The restriction digestion protocol was done as following:

- DNA of modular vector (500 ng)
- Cut smart buffer (NEB) (2.5 µl)
- BsaI enzyme (NEB) (1 µl)
- NF/double distilled water filled to (25 µl)
- Incubation step done for 1h at 37 °C

The product of digestion was loaded on agarose gel (1%) and the band with proper size was cut from the gel. Gene-Jet Gel extraction kit from Thermo Fisher Scientific (https://www.Thermo Fisher Scientific.com/us/en/home.html) was used to extract and purify DNA from the gel. The protocol used for Gel extraction kit was according to Thermo Fisher Scientific and is as follows:

- Gel slice was excised from the gel tray by the application of clean blade and put in 1.5 ml tube. The tube was weighed after putting gel band having desired DNA fragment.
- Binding buffer was added to the weight of sliced band from gel.
- Thermoblock was used for incubation (50-60 °C for 10 minutes) and melting the gel band. Binding buffer and gel band was mixed by vortexing for a few seconds.

- The dissolved solution was put inside the purification column and centrifuged that for 1 minute. Discarded the liquid after centrifugation and saved the membranes part of column.
- Wash buffer (700 µl) was added to the column and centrifuged for 1 minutes. Discarded the liquid again.
- A round of Centrifugation was done with empty tube to eliminate any residual liquid from the column.
- Elution buffer (20 µl) was added in another clean 1.5 ml tube with membraned part of column. Centrifugation was done to extract the DNA bound to the membrane for 1 minute.
- Stored the DNA in freezer (-20 °C).

The purified DNA concentration was measured on NanoDrop™ Spectrophotometer (https://www.Thermo Fisher Scientific.com/us/en/home.html). Hybridized gRNAs and digested modular vectors were ligated according to the manufacturer of T4 ligase enzyme (NEB, England). The ligation protocol was as performed as following:

- Digested and purified vector (20ng)
- Hybridized gRNA (1.5 µl)
- T4 DNA ligase buffer (10x) 1 µl
- T4 DNA ligase enzyme 0.5 µl
- Nuclease free water filled up to 10 µl
- Incubation for 1h at room temperature

3.2.1.2 *E. coli* bacterial transformation

The ligated products were transformed into XL1-Blue *E. coli* thermo-competent cells (Thermo Fisher Scientific) as following:

- *E. coli* competent cells collected from (-80 °C) freezer.
- Mixed 5-7 µl of ligation product (did this for every vector).
- Incubated the product for half hour in ice.

- Heat shocked the product for 90 sec at 42 °C temperature without shaking in thermo block.
- After the heat shock, it was put on ice for 20-30 seconds.
- Added 500µl SOC medium to revive the cells after heat shock.
- Incubation was done at 37 °C with 550 rpm for one hour.

After transformation, *E.coli* were selected on agar LB medium including the *Ampicillin* antibiotic (100 µg/ml stock solution, 1:1000 dilution). A 100 µl of transformed *E.coli* was streaked on the agar LB medium plates under aseptic conditions. Plates were incubated overnight (18-24h) at 37 °C. Five colonies were picked from each plate and cultured overnight in glass tubes in 3 ml liquid LB media contained *Ampicillin* antibiotic at 37 °C on a shaker with 180 rpm in dark.

3.2.1.3 Plasmid DNA extraction

The plasmid DNA was extracted from *E. coli* using Gene-Jet Miniprep plasmid extraction kit (Thermo Fisher Scientific). The extraction of plasmid DNA was performed according to the manufacturer instructions as follow:

- 2ml eppendorf tube was used for centrifuging the liquid bacterial culture for 5 minutes at 13,000 rpm (same speed for all steps) and a pellet was obtained after discarding the flow-through.
- Pelleted cells Ire mixed with resuspension solution (250 µl). Transferred the product to a microcentrifuge tube (1.5 ml). Vertexing was done until the pellet was completely dissolved.
- Lysis buffer (250 µl) was added and mixed completely as the solution became clear by inverting.
- Neutralization buffer (350 µl) was added and mixed right after the addition by inverting.
- Centrifugation was done or five minutes after that.
- Supernatant products after centrifugation was put in the column tube by decanting or by the use of the pipette.
- Centrifuged for one minute. Discarded the liquid after centrifugation.
- Wash buffer (500 µl) was added. Centrifugation was done for one minute.

- Washing step was repeated.
- A round of centrifugation was done with empty tube to eliminate any residual liquid from the column.
- The column was placed on a clean 1.5 ml tube and the elution buffer (20 µl) was added on the membrane of the column. Centrifugation was done to extract the DNA bound to the membrane for 1 minute.
- Stored the DNA in freezer (-20 °C).

3.2.1.4 Validation of constructed modular vectors (Test digestion & Sanger sequencing)

The purified plasmid DNA samples were test digested with a restriction enzyme The expected fragment was detected on agarose gel (1%) and visualized using fluorescence light with wave length 470 nm in gel documentation (BLUEPAD LED Trans illuminator, FG-09). The positive samples from test digestion were sent for Sanger sequencing with the primer IK71 (Hoffie I et al. not published yet) in LGC Genomics GmbH, Germany. The sequence of the cloned vectors were aligned to the *in silico* designed vectors to confirm the incorporation of the correct gRNA. The properly ligated vectors were subjected to the following cloning step.

3.2.2 Assembly of gRNA modules via Golden Gate reaction

In this step, two vectors were constructed and each contained four inserting positions filled by two gRNAs and each gRNA is duplicated as shown in Figure 3.3 step 2b (given above). The constructed vector pDD38 comprised gRNA1 (positions 1 and 3) and gRNA2 (positions 2 and 4) gRNAs twice while vector pDD39 comprised gRNA3 (positions 1 and 3) and gRNA4 (positions 2 and 4) (Table 3.3).

Table 3.3. Assembly vectors entailing single gRNA modules. Single gRNA modules and incorporated gRNAs are shown in the table

Constructed vector	**Vector backbone**	**Assembled gRNA modules**	**Incorporated gRNA**
pDD38	pIK19	pDD 30 - 31 - 32 - 33	gRNA1 + 2 twice
pDD39	pIK19	pDD 34 - 35 - 36 - 37	gRNA3 + 4 twice

The assembly vector (pIK19) has the capacity to withhold four guides with the backbone entailing TaU6 promoter, scaffold and transcription termination signal for each guide, multiple cloning sites and antibiotic resistant gene for bacterial (*E. coli*) selection.

Assembly vectors were constructed with the help of a golden gate reaction in which restriction digestion and ligation was done in a single reaction as shown in Table 3.4. The type IIS BsmBI restriction enzyme (NEB, England) was used for restriction digestion and T4 DNA Ligase (NEB, England) according to the manufacturer. The ligation product was transformed in the XL1-blue *E.coli* thermo-competent cells (Thermo Fisher scientific) according to the protocol described in section 3.2.1.1. Validation of assembly vectors was performed through Sanger sequencing in LGC Genomics GmbH, Germany.

Table 3.4. Golden Gate reaction with Esp3I type IIS restriction enzyme to assemble vectors pDD38 and pDD39.

Golden Gate reaction [10 µl reaction]	**Golden Gate Thermocycle conditions**
75 ng of each gRNA module vector 15 ng of destination vector 0.5 µl Esp3I (BsmBI) (NEB) 0.5 µl T4 DNA ligase (NEB) 1.0 µl 10x Cut Smart buffer 0.5 µl ATP (10 mM/ NEB) ddH_20 up to 10 µl total reaction volume	**Run thermocycler** 37 °C for 5 min 22 °C for 10 min } 10x 37 °C for 30 min 75 °C for 15 min 4 °C ∞

3.2.3 Assembly of pre-assembled gRNAs, *Cas9* and *sfGFP* expression units

After the assembly of gRNA expression units in one recombinant fragment, this fragment was ligated with expression units of *Cas9* and *sfGFP* genes or a dummy position in the backbone of vector pIK48 via Golden Gate reaction. Four constructed vectors were achieved i.e. pDD40, pDD41, pDD42, and pDD43 as shown in (Table 3.5) including *sfGFP* gene to ease the transformation screening by testing transformed explant under fluorescent microscope. While including a dummy position is to provide a possibility to add a new functional expression unit such as a new gene in the future to the same vector.

Table 3.5. Assembly of gRNAs expression units together with expression units of Cas9 and sf GFP or dummy via Golden Gate method.

Constructed vector	**Vectors backbone (destination vectors)**	**Entry vectors**	**Incorporated recombinant fragments**
pDD40	pIK48	pDD38+pIk83+pIK155	gRNAs, *Cas9* and dummy
pDD41	pIK48	pDD38 pIk83+pIK290	gRNAs, *Cas9* and *sfGFP*
pDD42	pIK48	pDD39 pIk83+pIK155	gRNAs, *Cas9* and dummy
pDD43	pIK48	pDD39 pIk83+pIK290	gRNAs, *Cas9* and *sfGFP*

The entry vector pIK290 contained expression unit of *sfGFP* gene fluorescent marker, pIK83 contained Cas9 while pIK155 vector contained a dummy position.

Golden gate reaction was done according to Table 3.6:

Table 3.6: Golden Gate reaction using BsaI type IIS restriction enzyme to assemble vectors pDD40, pDD41, pDD42 and pDD43.

Golden Gate reaction [10 µl reaction]	**Golden Gate Thermocycle**
75 ng of each gRNA module vector	**Run thermocycler conditions**
15 ng of destination vector	37 °C for 5 min } 10x
0.5 µl BsaI (NEB)	22 °C for 10 min } 10x
0.5 µl T4 DNA ligase (NEB)	37 °C for 30 min
1.0 µl 10x Cut Smart buffer	75 °C for 15 min
0.5 µl ATP (10 mM/ NEB)	4 °C ∞
ddH_20 up to 10 µl total reaction volume	

The ligation product was transformed in the XL1-blue *E.coli* thermo-competent cells (Fisher scientific) according to the protocol described in section 3.2.1.1. Validation of assembly vectors was performed through Sanger sequencing in LGC Genomics GmbH, Germany.

3.2.4 Construction of binary vectors

Binary vectors construction for plant stable transformation is based on conventional cloning method. Restriction digestion and ligation are performed separately as described in (section 3.2.1.1). The p6i:d*35S* binary vector having *HPT* gene was used to incorporate the recombinant T-DNA fragments form former steps. The plan for constructing binary vectors is described in (Table 3.7).

Table 3.7. Constructed Binary vectors comprising the final assemblies/expression cassette for stable plant transformation to target ScPLA1 gene.

Constructed vector	Vectors backbone (destination vectors)	Entry vectors	Incorporated recombinant fragment
pDD44	P6i-*35S::HPT*	pDD40	gRNAs, *Cas9*, *HPT* and dummy
pDD45	P6i-*35S::HPT*	pDD41	gRNAs, *Cas9, HPT* and *sfGFP*
pDD46	P6i-*35S::HPT*	pDD42	gRNAs, *Cas9*, *HPT* and dummy
pDD47	P6i-*35S::HPT*	pDD43	gRNAs, *Cas9, HPT* and *sfGFP*

The ligation product was transformed in the XL1-blue *E.coli* thermo-competent cells (Thermo Fisher Scientific) according to the protocol described to (section 3.2.1). Validation of binary vectors was performed with the help of test digestion by using enzyme *XhoI* (Thermo Fisher Scientific) and Sanger sequencing (LGC Genomics GmbH, Germany).

The constructed vectors pDD44, pDD45, DD46 and pDD47 were transformed in agrobacterium strain LBA4404pSB1 for stable plant transformation. Transformation of agrobacterium was done by the electroporation method with following protocol:

- Agrobacterium strain (LBA4404pSB1) was collected from freezer (-80 °C) on ice
- 3 µl vector product was added in the 50 µl agrobacterium competent cells
- Electric current (2.5 kV, 200 Ω, 20 mF) was supplied to the tube
- 950 µl SOC medium (Thermo Fisher Scientific) was added immediately after the electroporation
- The transformed product was incubated for 2.5 hours at 28 °C with 550 rpm

3.3 Functional Validation Through Protoplast Transfection Assay

The constructed binary vectors (pDD44, pDD45, pDD46 and pDD47) were functionally validated through the PEG (Polyethylene glycol) mediated protoplast transfection in Lo7 and Nr5 genotypes. Three technical replicates for each genotype were used (Table 4.8). The whole experiment was repeated again to have two

biological replicates. Wild type protoplast (non-transfected) was included as a reference for healthy and well grown protoplast of both genotypes while a well-tested vector pGH215 (*Ubi::GFP*) in our lab was used as reference of transformation success. Before incubating the transfected protoplasts, three solutions were tested (W1, W5 and W1:W5) to see which solution provides with better survival rate for the transfected protoplasts.

Table 3.8. Experimental setup for protoplast transfection. Two genotypes along with four binary vectors tested.

Genotype: Lo7				**Genotype: Nr5**			
Vector	**Rep 1**	**Rep 2**	**Rep 3**	**Vector**	**Rep 1**	**Rep 2**	**Rep 3**
WT	W1	W1:W5	W5	**WT**	W1	W1:W5	W5
pGH215	W1	W1:W5	W5	**pGH215**	W1	W1:W5	W5
pDD44	W1	W1:W5	W5	**pDD44**	W1	W1:W5	W5
pDD45	W1	W1:W5	W5	**pDD45**	W1	W1:W5	W5
pDD46	W1	W1:W5	W5	**pDD46**	W1	W1:W5	W5
pDD47	W1	W1:W5	W5	**pDD47**	W1	W1:W5	W5

3.3.1 Plant material and growth conditions

Two genotypes (Lo7 and Nr5) were used in the experiment and germinated in the green house of Leibniz institute of plant genetics and crop plant research (IPK) Gatersleben, Germany. The grains were germinated for three weeks until they had 2-3 leaves around 15 cm long which were then harvested for protoplast isolation. The rye plants were grown on Jiffy7 soil in growth chambers with controlled temperature at 18 °C in dark and a 50-60% relative humidity.

3.3.2 Protoplast isolation

Protoplast isolation was done according to Yoo et al., 2007 with minor modifications. Enzyme solution to digest the cell walls was prepared as following protocol:

- Cellulose (Duchefa Biochemie) 150 mg/10ml
- Macerozyme (Duchefa Biochemie) 200 mg/10ml
- Mannitol 0.8 M 5ml/10ml
- KCL 100µl/10ml

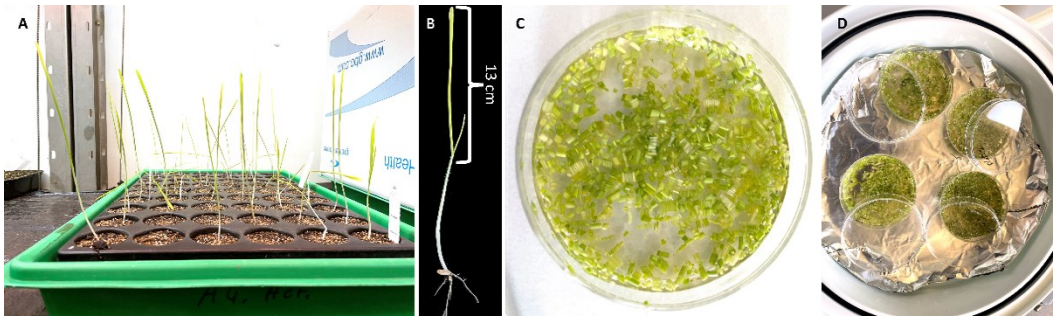

Figure 3.4 Flow of rye protoplast isolation steps. A and B) etiolated rye plants at proper stage for protoplast harvesting. C) Chopped leaves in enzyme solution. D) Chopped leaves in the exicator for vacuum treatment.

Middle-third part of the etiolated leaves from 3weeks old plants were harvested for digestion (Figure 3.4A and B). A stack of 3 to 4 leaf pieces were piled and carefully chopped using a sharp and grounded razor-blade into small stripes of approx. 1-2 mm width. It is important to hold gently the leaves without pressing to avoid any mechanical damage while with the other hand, chop of leaves. The chopped stripes were immersed rapidly into the digestion solution without giving the risk letting them dry out. Approximately, 13-15 g of chopped leaves digested in to 5 ml enzyme solution (Figure 3.4C). After chopping all the leaves, samples were vacuumed for 30 min at 600 mbar and kept in the dark to enhance infiltrating the enzyme solution through the chopped leaves. The vacuum was released very slowly to avoid mechanical damage to samples. The samples were incubated in dark at 22 ° C for 3 hours. The samples were filtered through a 75µm nylon mesh and rinsed the petri plate twice with 5 ml of W5 solution and did filtration through the nylon mesh. The filtered solution was collected in 50 ml conical-base Falcon tube and centrifuged at 60g (at very slow acceleration/deceleration speed) for 2 min and removed most the liquid solution and then added 3ml W5 solution and resuspended the protoplast by rocking very carefully to homogenate the suspension.

3.3.3 Transfection of rye protoplasts

The protoplast population was counted using Hemocytometer slide (Rogo Sampaic, France) according to the manufacturer to calculate a population of 600,000 protoplast per one transfection. While counting, rest of the protoplasts were kept on ice as protoplasts tend to settle down. After protoplast settled down, the W5 solution was carefully discarded and resuspended in MMG transfection medium with population density of $6x10^5$ ml^{-1}/100 ml MMG (means 600,000 protoplast/transfection).

For the DNA-PEG calcium mediated transfection, following protocol was employed:

- Added 100µl protoplast (in MMG) in 2 ml Eppendorf tubes.
- Added 10 µl of plasmid DNA of respected modular vector (1µg/µl).
- Added very gently on the wall of the tube 110µl PEG solution.
- Mixed very carefully the solution with just slight and gentle tap.
- Incubated 15 minutes in the dark at room temperature (ideally 22 °C)
- Then added 450 µl W5 solution and mixed the solution by tapping.
- Centrifuged at 100 g for 2 minutes and liquid was discarded carefully.
- Re-suspended the pellet in 450µl W1 solution.
- Added 250µl protoplast in the BSA pre-coated wells (Capricorn Scientific) (wells were previously coated with 300µl BSA).
- Incubated the transfected protoplasts at 22 °C in dark.
- Checked for the GFP signal after approx. 18 h with the help of fluorescent microscopy.

3.3.4 DNA isolation from transfected protoplasts

After 18 to 24 h, the protoplast was checked for the GFP signal using Zeiss Axiovert 200M Inverted Fluorescence Microscope. Then the genomic DNA from protoplasts was extracted according to the following protocol:

- Transfer the protoplasts in a 2ml Eppendorf tube and rinse the wells with 500µl of W5 buffer and transfer it to the Eppendorf tube.
- Centrifuge protoplasts at 100g for 3 min the supernatant was carefully discarded without disturbing the pellet.
- 300µl Waite-extraction buffer were added to the pellet of protoplast and vortexed vigorously and incubated at 55°C for 3 min with shaking (400-500 rpm).
- 300µl phenol: chloroform: isoamyl alcohol solution mix was added and shook it for 5 min or can invert it as well for 5 min by hand.
- The solution was centrifuged for 10 minutes at 20 000g at room temperature.
- The upper layer after centrifugation was transferred in a new 1.5ml Eppendorf tube.
- 60µl (3M) NaAc and 450µl Isopropanol was added (stored at -20°C). Mixe the solution by inverting and incubated the samples for 3h (or longer) at -20°C.
- The samples were centrifuged after incubation at 14 000g for 5 minutes and discarded the upper layer. (Note: be very careful not the remove the DNA-pellet).
- Added 800µl ethanol 70% and vortex the samples.
- The samples were centrifuged at 14 000g for 5 minutes.
- The upper layer was very carefully remove and dried the DNA pellet for 30 minutes (no ethanol should remain).
- Added 50µl TE buffer including RNAse A to suspend the DNA pellet and incubated the samples at 37°C for 1h.
- Concentrations of DNA s was measured on NanoDrop™ Spectrophotometer (Thermo Fisher Scientific) for every sample.

3.3.5 Generation of amplicons

The extracted genomic DNA was used as a template to amplify the target region. A 228 bp region including the target regions of target motifs 1 and 2 and a 234bp region including the target regions of target motifs 3 and 4 were amplified using specific primers (Table 3.9) and followed by amplicon-sequencing, which was performed by a commercial service provider on an Illumina MiSeq platform.

Table 3.9. Primers used to amplify the target region for amplicon sequencing.

Amplicon seq. Primers	Sequences	Target region
DD11_F	CAGCACGAGGGCGATG	gRNA1 and 2
DD11_R	GCTCATCACCGCCATGAT	gRNA1 and 2
DD12_F	TTCGACGTCAAGCTTCTC	gRNA3 and 4
DD12_R	CACTTCCACACCCAGG	gRNA3 and 4

Annealing temperature was set for adjacent primers beforehand with the wild-type rye genotypes. The PCR was prepared for every DNA sample along with wild type and water as control to be certain of no DNA contamination. PCR reaction volume was 35 µl for each sample. A five µl of amplicon (PCR product) were run on gel electrophoresis to confirm the amplified region size and 30 µl was purified and prepared for amplicon sequencing. PCR reaction and program were performed as in table 3.10.

Table 3.10: Primers used to amplify the target region for amplicon sequencing.

Step	PCR temp. ºC	Time	Cycles
Initial denaturation	95	2 min	1x
Denaturation	95	15 sec	40x
Annealing	62	30sec.	
Extension	72	30 sec	
Final extension	72	1 min	1x
storage	4	-	-

3.3.6 Amplicon deep sequencing samples preparation

Once the amplicon size is confirmed, the rest 30 µl of PCR product was purified with the Gene-Jet PCR purification kit (Thermo Fisher Scientific) with the following protocol:

- Binding buffer was added at equal volume to PCR product and then mixed.

- Our amplicon was less than 500bp, I added a 1:2 volume of 100% isopropanol (30 µl) and mixed.
- The solution was transferred to column tube and centrifugation was done for one minute.
- Wash buffer (700 µl) was added. Centrifugation was done for one minute.
- A round of Centrifugation was done with empty tube to eliminate any residual liquid from the column.
- Elution buffer (20 µl) was added in another clean 1.5 ml tube with membraned part of column. Centrifugation was done to extract the DNA bound to the membrane for 1 minute.
- Stored the DNA in freezer (-20 °C).

After purifying the PCR product, samples were prepared for deep sequencing of amplified regions and following steps were taken:

- Measured DNA of all the samples as I needed only 20 ng/µl for each sample and total minimum requirement for DNA concentration was 500 ng/µl per tube from the sequencing company.
- I decided to send separate tube for each genotype so in total 2 tubes.
- One tube for Lo7 genotype containing all the samples and the other one for Nr5.
- I pooled all the samples and did the average for DNA concentration of all the samples per genotype. Diluted the DNA with nuclease free water to make 20ng/µl as per the requirement from sequencing company.
- Labelled the tubes and sent them for sequencing.

4. RESULTS

4.1 Confirmation of PLA1 Orthologue ın Rye

To confirm the presence and functionality of *PLA1* orthologue in rye, the BLASTn of *ZmPLA1* sequence in the GrainGenes database resulted in an orthologue sequence with the similarity of 73% and 84% query coverage on Chr. 7

of rye genome. To further confirm the functionality of rye orthologue, the virtually translated protein sequence from resulted DNA sequence (rye orthologue) was compared *in silico* against translated PLA1 protein of maize and durum-wheat using SnapGene software (GSL Biotech) by using the algorithm MUSCLE (Multiple Sequence Comparison by Log-Expectation) to align the amino acid sequences. The Pair-wise sequence alignment of protein sequences (amino acid alignment) showed almost 85% matching of the amino acid among the three species (Figure 4.1) which indicates that it is anticipated to be the orthologue of *ZmPLA1* gene.

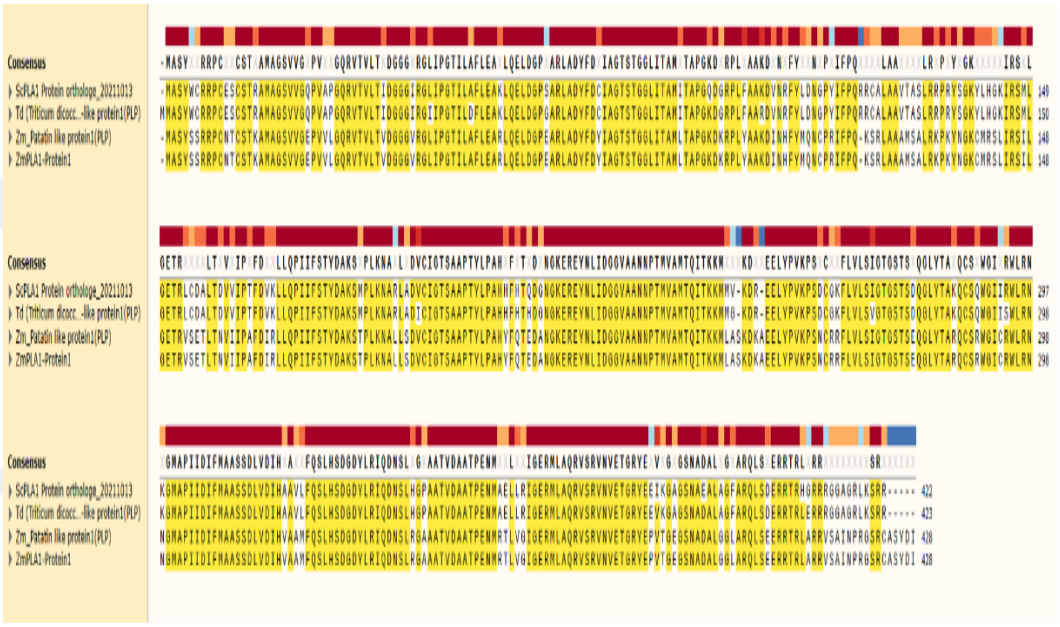

Figure 4.1 Visual example of the analytical method used to align PLA1 protein sequences. A screenshot of the PLA1 Amino acids sequence in maize, durum-wheat and rye. The highlighted yellow color indicates the similar amino acid sequences.

4.1.1 *ScPLA1* gene structure and selected target motifs

As rye Lo7 genotype was sequenced recently, the annotation of the whole genome is not complete and genes in the genome had been identified based on the coding sequences mostly but intronic regions among the genes are still not completely known. By blasting the sequence of *PLA1* orthologue in the Softberry database using the tool FGENESH for gene prediction, the tool predicted 4 exons

and 3 introns which are annotated and displayed using SnapGene software as shown in Figure 4.2. Based on the predicted coding sequence region in *ScPLA1* gene, TMs were selected to induce site directed mutagenesis in *ScPLA1* (Table 4.1).

Table 4.1. Sequence of the designed TMs for the ScPLA1 gene in rye

Name	Sequence	size	Location	Direction
TM1	5` GCCGGTGGCGCCGGGGCAGA 3`	20bp	Exon1	Sens
TM 2	5` GACTACTTCGACTGCATCGC 3`	20bp	Exon1	Sens
TM3	5` GTCGTATGTGGAGAAGATGA 3`	20bp	Exon2	Antisense
TM 4	5` GAGGCTGGATGGAGGGGCG 3`	19bp	Exon3	Antisense

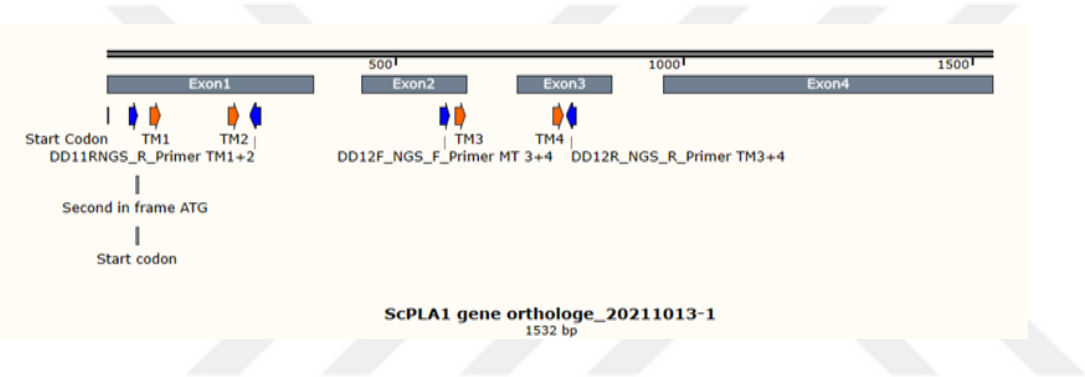

Figure 4.2. Structure of ScPLA1 gene depicting exonic region along with the target motifs pointed for knockout through CRISPR technology.

4.1.2 Designed gRNAs to target *ScPLA1* gene

Based on the sequence of TMs, gRNAs were designed according to criteria that has already been explained in Materials and Methods section 3.1.2. The prediction of gRNA secondary structure is important to predict the binding capacity of gRNA to the complementary TM on the genomic DNA of the *ScPLA1* gene. The secondary structure was predicted through RNAfold (http://rna.tbi.univie.ac.at/cgi-bin/RNAIbSuite/RNAfold.cgi) an online server which helps to see the virtual images of 2D structure of gRNAs (Figure 4.3). A good binding gRNA should have three stem loops and the gRNA nucleotides should be open to bind directly with the target motif (Kumlehn et al., 2018). Figure 4.3 shows the 2D secondary structure of the selected gRNAs. The four selected gRNAs 2D secondary structure contained

the 3 stem loops and the first nucleotides at 5`end are not bound (Figure 4.3). Table 4.2 and 4.3 indicate some important features of the designed gRNAs.

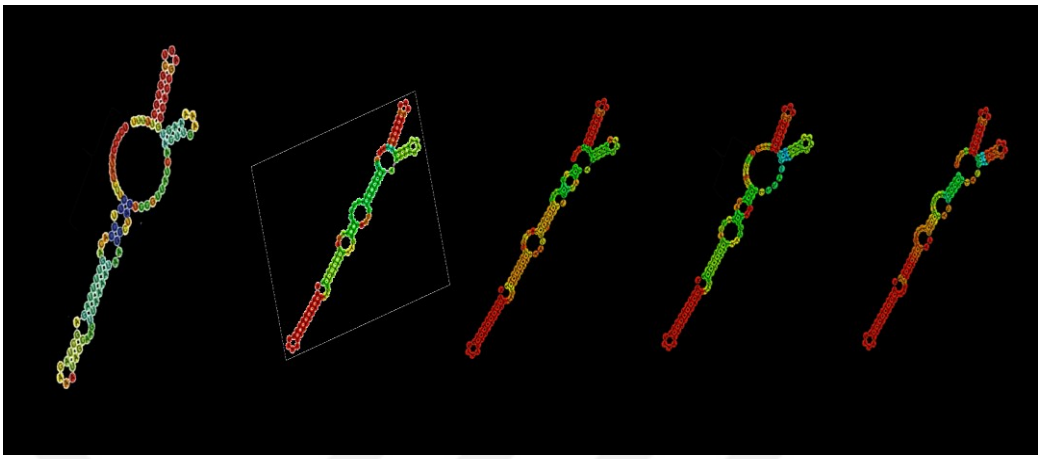

Figure 4.3 Secondary structure prediction of gRNAs from RNA fold web server. (A) Depicts the properties of a good gRNA for CRISPR applications and (B) shows the gRNA obtained for PLA1 gene of rye.

Table 4.2. gRNA sequences and efficiency score for ScPLA1 gene. Scores were predicted from online source (Benchling) for the selection of guides.

Selected guide	**Sequence**	**On-Target score (Benchling)**	**Transcription (G at 5' end)**	**PAM + six downstream bases**
gRNA1	GCCGGTGGCGCCGGGGCAGA	78	yes	GGGtgacgg
gRNA2	GACTACTTCGACTGCATCGC	92	yes	CGGcaccag
gRNA3	GTCGTATGTGGAGAAGATGA	80	yes	TGGgctgga
gRNA4	GAGGCTGGATGGAGGGGCG	88	yes	CGGggaggt

Table 4.3. Orientation and position of gRNAs in the ScPLA1 gene.

Selected guide	**Position in *PLA1***	**Cleavage**	**Length of guide sequence**	**Orientation**	**Cryptic termination (≥4 successive Us in guide sequence)**

gRNA1	75	Exon 1	20	Coding	No
gRNA2	210	Exon 1	20	Coding	No
gRNA3	604	Exon 2	20	No-coding	No
gRNA4	995	Exon 3	19	Con-coding	No

4.2 Cloning of modular vectors for rye transformation

As described in the Materials and Methods, there are four main steps to construct binary vectors harboring functional expression cassettes of the selected gRNAs, *Cas9, sfGFP, and HPT* genes. The Figure 4.4 summarizes all intermediate vectors until the construction of binary vectors.

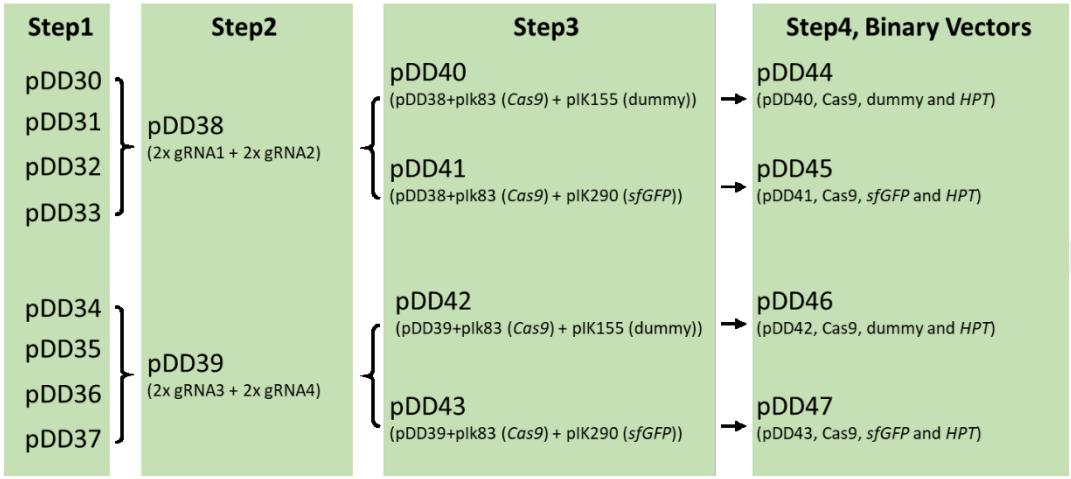

Figure 4.4. A schematic chart summarizing the cloned vectors. A) Incorporating single gRNA in a single modular vector. B) Assembly of incorporated gRNAs in a single vector via Golden Gate method. C) Assembly of the assembled gRNAs with expression cassettes *Cas9* and *sfGFP* or a dummy position. D) Generating Binary Vectors for stable plant transformation.

4.2.1 Modular vectors containing gRNAs

The cloning plan of gRNAs aimed to include 2 guides together and each is represented twice at the binary vectors. Therefore, eight modular vectors were constructed at this step (Table 4.4). gRNA1 was cloned in the backbone of modular vectors pSH215 (position1), and pSH325 (position3) while gRNA2 was cloned in modular vectors pSH218 (position2), and pSH326 (position4) which are finally the

four positions to be assembled in one assembly vector. In parallel, gRNA3 and gRNA4 were cloned in the backbone of the four modular vectors in the same manner to be assembled in another assembly vector.

Table 4.4. Assembled modular vectors for single gRNAs. Backbone of vectors and positions of gRNA are shown in the table.

Constructed vector	Vector backbone	Incorporated gRNA	Fragment position
pDD30	pSH215	gRNA1	Position 1
pDD31	pSH218	gRNA2	Position 2
pDD32	pSH325	gRNA1	Position 3
pDD33	pSH326	gRNA2	Position 4
pDD34	pSH215	gRNA3	Position 1
pDD35	pSH218	gRNA4	Position 2
pDD36	pSH325	gRNA3	Position 3
pDD37	pSH326	gRNA4	Position 4

4.2.1.1 Restriction digestion

At first step, modular vectors were digested with restriction enzyme BsaI (Thermo Scientific), the expected fragment was detected on gel electrophoresis with size of 2415bp for vectors pSH215 and 218 and 2700 bp for vectors pSH325 and 326 (Figure 4.5). In parallel, single stranded forward and reverse gRNAs were hybridized. Both hybridized and digested modular fragments were ligated according to cloning plan (section 3.2.1).

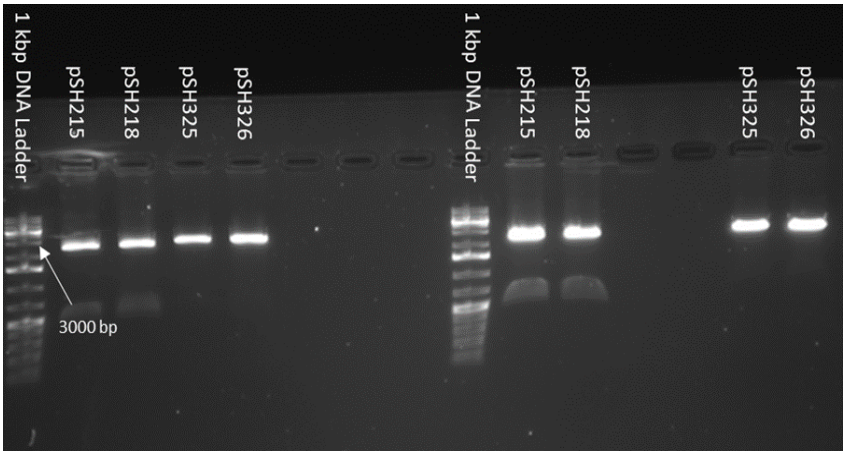

Figure 4.5. 1% agarose gel electrophoresis image of the digested fragment of single modular vectors pSH215, pSH218, pSH325, and pSH326 with BsaI restriction enzyme.

4.2.1.2 Transformed bacterial colonies

After the ligation of gRNAs with the extracted and purified fragment from the gel, the ligation product was transformed in XL1-Blue *E. coli* thermo competent cells. Once the transformation procedure (section 3.2.2.2) was done, *E.coli* cells were selected on agar LB medium having the antibiotic (*Ampicillin*, 100 µg/ml, 1:1000 dilution). A 100 µl of transformed *E.coli* was streaked on the agar LB medium plates under aseptic conditions. Figures 4.6 shows the grown *E. coli* colonies on the agar LB medium under antibiotic selection pressure and five well separated colonies from each ligated vector were randomly selected (Figure 4.6B) and grown LB liquid medium for plasmid DNA extraction and purification.

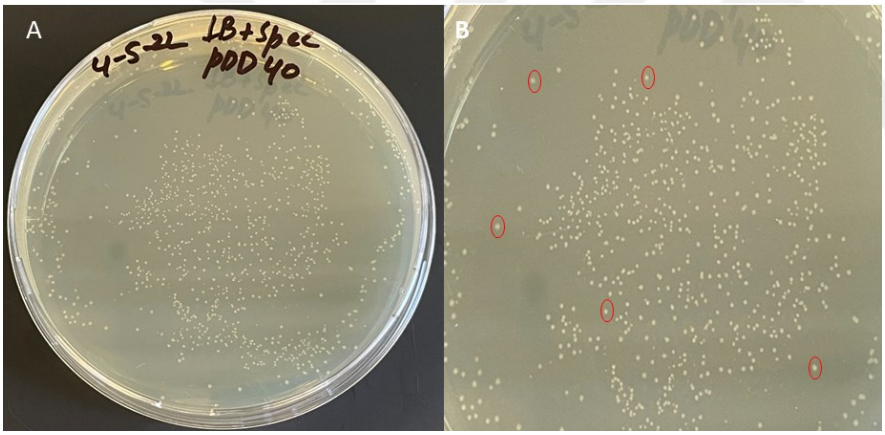

Figure 4.6. A) Transformed bacterial colonies on agar LB medium plate. B) Red circles show the randomly selected colonies.

4.2.1.3. Validation of constructed vectors

Samples were sent for Sanger sequencing with primer IK70 to sequence a region comprising TaU6, gRNAs, scaffold, and 4-Ts termination signal. The resulted sequence was aligned against the sequence of the *in-silico* pre-designed vector. Figure 4.7 showed the alignment results of selected 5 colonies of pDD30 construct. In this example, four colonies were positive displaying the correct insertion of gRNA1 expression unite while the sequence of colony 5 showed a SNP

(Figure 4.7B) in 20 nucleotides of the gRNA, highlighted by red color. Plasmid from a single positive colony was selected for further cloning steps.

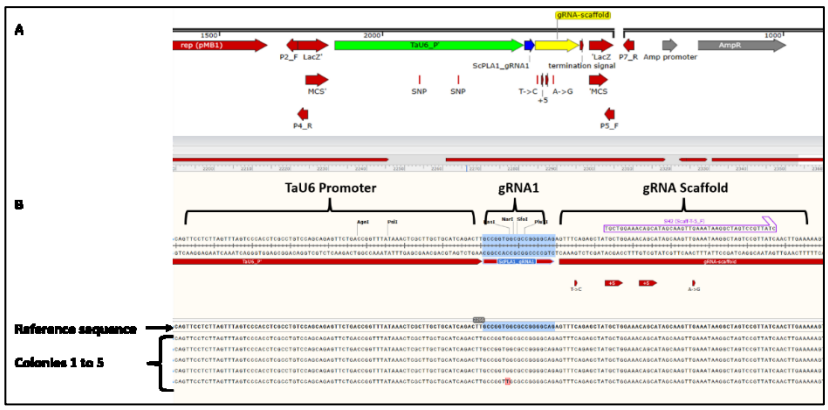

Figure 4.7. A) A screenshot of an example of a modular vector backbone (pSH218) with ligated DD13-gRNA1 indicated by blue arrow, followed by gRNA scaffold indicated by yellow, and arrow driven by TaU6 promoter indicated by green arrow. B) Sanger sequences of selected five colonies aligned against the sequence of pre-designed vector pDD30. Colonies 1 to 4 incorporated the correct sequence of gRNA1 expression unit referring to a successful cloning. Colony five displayed a SNP in the incorporated sequence where a G was replaced by a T highlighted in red.

4.2.2 Assembly of incorporated gRNA modular vectors

After construction of single modular vectors, the first Golden Gate reaction was performed with restriction digestion using Esp3I type-IIS restriction enzyme (Thermo Scientific) and ligation using T4 DNA Ligase (NEB, England) was performed to assemble the multiple gRNAs in one assembly vector via a single PCR reaction. The constructed assembly vectors contain the pIK19 backbone and gRNAs (Table 4.5). Selection of positive bacterial transformation was done according to section 4.2.1.2 and constructed vectors were validated through Sanger sequencing using primer IK70 (Figure 4.8B). The assembled vectors were subjected to the next cloning step (Figure 4.8A).

Table 4.5. Assembled vectors entailing the gRNAs. Vector backbone and incorporated modules are depicted in the table.

Constructed vector	Vector backbone	Assembled gRNA modules	Incorporated gRNA

pDD38	pIK19	pDD 30 - 31 - 32 - 33	gRNA1 + gRNA2 twice
pDD39	pIK19	pDD 34 - 35 - 36 - 37	gRNA3 + gRNA4 twice

Figure 4.8. A screenshot of assembly vector maps of pDD38 and pDD39. B) Sanger sequences of selected colonies aligned against the sequence of pre-designed vector pDD38.

4.2.3 Assembly of pre-assembled gRNAs, *Cas9* and *sfGFP* expression units

In the second Golden Gate reaction in Cascade/Cas9 modular vector system, the assembly was made using destination vector pIK48 as backbone, pIK83 vector incorporating *Cas9* and either pSH290 vector incorporating *sfGFP* or pIK155 containing a dummy position. Each assembled set of gRNAs expression units was incorporated once with *sfGFP* and another with a dummy position and in either cases was in the backbone of vector pIK48 harbouring the *Cas9* expression unit (Table 4.6). Figure 4.9 shows the vector map of the constructed vector pDD41 incorporating expression units of gRNAs 1 and 2, *Cas9*, and *sfGFP*. Bacterial transformation and colonies selection was performed according to (section 4.2.1.2) and plasmid DNA was extracted for further validation.

Table 4.6. Assembled gRNAs expression units together with expression units of Cas9 and sf GFP or dummy via Golden Gate method.

Constructed vector	Vectors backbone (destination vectors)	Entry vectors	Incorporated recombinant fragments
pDD40	pIK48	pDD38+pIk83+pIK155	gRNAs, *Cas9* and dummy
pDD41	pIK48	pDD38 pIk83+pIK290	gRNAs, *Cas9* and *sfGFP*
pDD42	pIK48	pDD39 pIk83+pIK155	gRNAs, *Cas9* and dummy
pDD43	pIK48	pDD39 pIk83+pIK290	gRNAs, *Cas9* and *sfGFP*

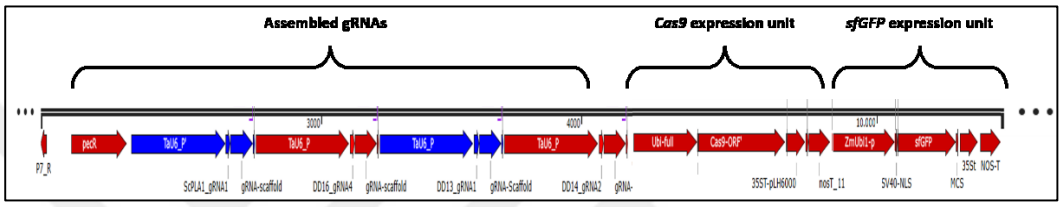

Figure 4.9. A screenshot of vector map of pDD41 as an example of constructed vectors in the second Golden Gate reaction.

The constructed vectors pDD40, 41, 42 and 43 were validated by test digestion using *EcoRV* restriction enzyme (Thermo Fisher Scientific) and the expected fragments were detected on gel electrophoresis with size of 7896 bp and 2654 bp for vectors pDD40 and pDD41 and 13759 bp for vectors pDD42 and pDD43 (Figure 4.10).

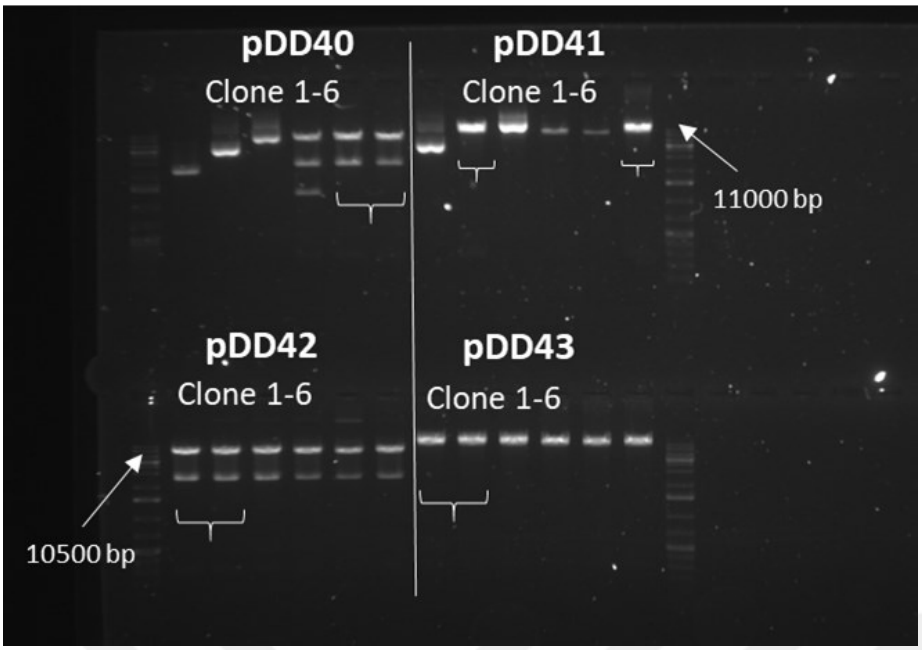

Figure 4.10. Test digestion of vectors pDD40, 41, 42 and 43 using EcoRV enzyme. Small brackets in the picture indicate the positive colonies by each vectors.

In addition, the DNA sequence of positive vector in the test-digestion test was tested via Sanger sequencing using primers IK70, IK71, and the forward sequence of last gRNA in such a vector. Each primer was used separately to sequence its corresponding fragment. Sequence alignment showed proper integration of all ligated fragments. Figure 4.11 showed an example of Cas9-ORF region in 2 colonies. An extracted plasmid DNA from single colony was selected from the positive ones to be cloned in the binary vector in the following cloning step.

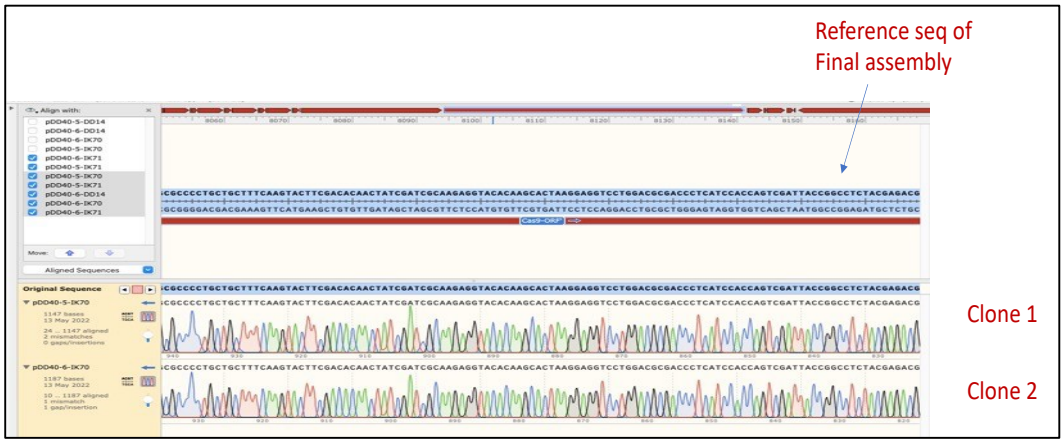

Figure 4.11. Sanger sequencing of final assembly modular vectors/expression cassette. We can see the primer set selected on the left side of the diagram and chromatogram of positive clones with reference to compared sequence.

4.2.4 Binary vectors for plant transformation

The assembled expression cassettes of gRNAs, Cas9, sfGFP or a dummy (vectors pDD40, 41, 42, and 43) were ligated in the backbone of binary vector p6i::d35S harboring HPT resistance gene to construct vectors pDD44, 45, 46, and 47 (Table 4.7). Backbone of binary vectors p6i::d35S and entry vectors pDD40, 41, 42, and 43 were digested with restriction enzyme SfiI (Thermo Scientific). The expected fragment was detected on agarose gel (1%) and visualized using fluorescence light in Gel documentation. The expected fragment size for binary vector p6i::d35S is 9988 bp, both vectors pDD40 and pDD 42 is 8916 bp, while both vectors pDD41 and pDD43 is 12123 bp (Figure 4.12).

Table 4.7. Constructed Binary vectors comprising the final assemblies/expression cassette for stable plant transformation to target ScPLA1 gene.

Constructed vector	**Vectors backbone (destination vectors)**	**Entry vectors**	**Incorporated recombinant fragment**
pDD44	P6i-*35S::HPT*	pDD40	gRNAs, *Cas9*, *HPT* and dummy
pDD45	P6i-*35S::HPT*	pDD41	gRNAs, *Cas9, HPT* and *sfGFP*
pDD46	P6i-*35S::HPT*	pDD42	gRNAs, *Cas9*, *HPT* and dummy
pDD47	P6i-*35S::HPT*	pDD43	gRNAs, *Cas9, HPT* and *sfGFP*

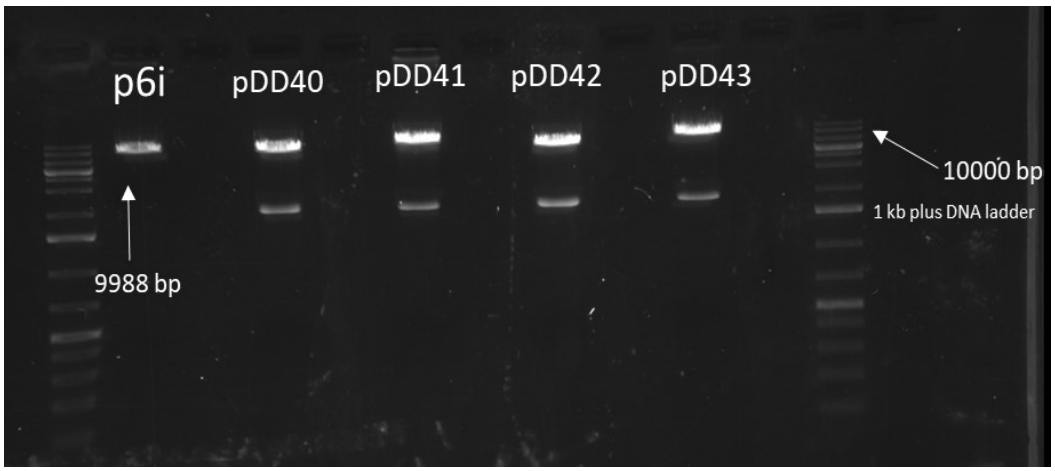

Figure 4.12. Restriction digestion of binary vector and four final assembly vectors.

After the restriction digestion, all entry fragments were subjected to ligation, E.coli transformation, and then plasmid DNA extraction according to (section 3.2.1.1). After the extraction of plasmid DNA from selected transformed bacterial colonies, the constructed four binary vectors were validated by test digestion with restriction enzyme XhoI (Thermo Scientific) and respective bands were seen on the gel electrophoresis with correct fragments size (Figure 4.14). Figure 4.13 shows the map of one of the constructed pDD45 binary vector harboring expression cassettes of gRNAs, *Cas9* protein, *sfGFP* and *HPT* cassette which is then subjected to agrobacterium transformation.

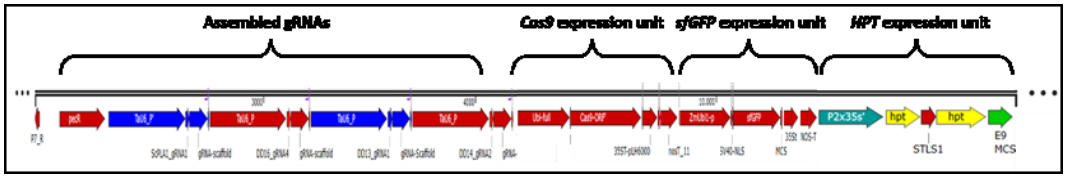

Figure 4.13. A screenshot of vector map of pDD45 as an example of constructed binary vectors plant stable transformation.

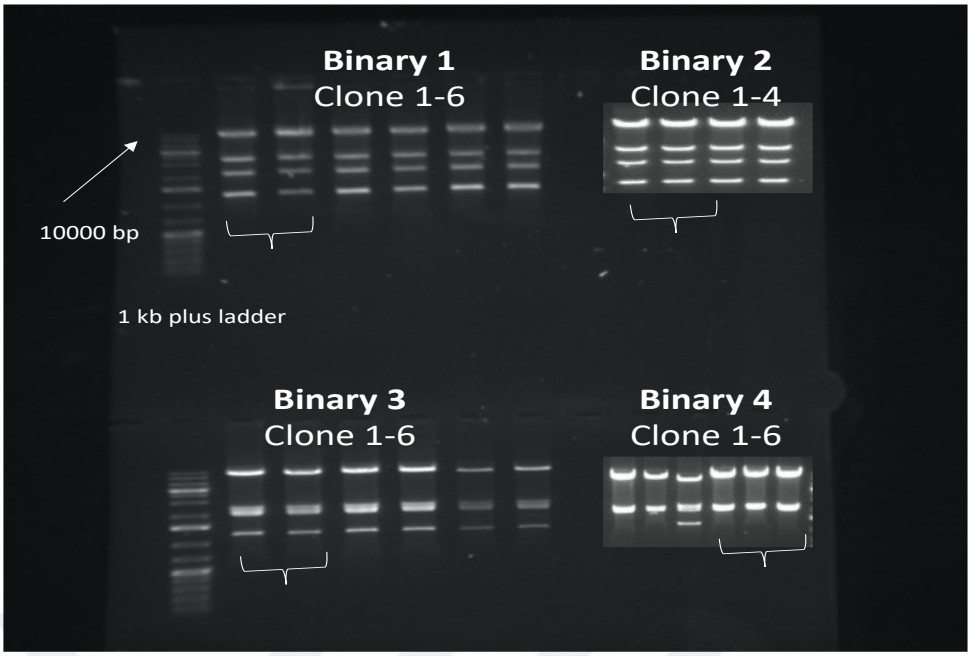

Figure 4.14. Test digestion of constructed binary vectors for rye plant transformation. Brackets in each binary vector indicate the positive clones.

4.3 Efficiency of Protoplast Transfection

The constructed binaries were employed for rye protoplast transfection to validate the vectors before performing stable transformation. Three technical replicates and two biological replicates were conducted employing three culture medium, two genotypes with four constructed binary vectors included into the experimental setup of protoplast assay. To validate the function of *ScPLA1* gene, protoplast transfected with binary vectors pDD45 and pDD47 (containing *sfGFP*) were checked under fluorescent microscopy while the four binary vector were subjected to amplicon deep sequencing. Both pDD45 and pDD47 vectors showed strong expression of *GFP* which referring to successful construction of binary vectors along with a positive sign for validation (Figure 4.15).

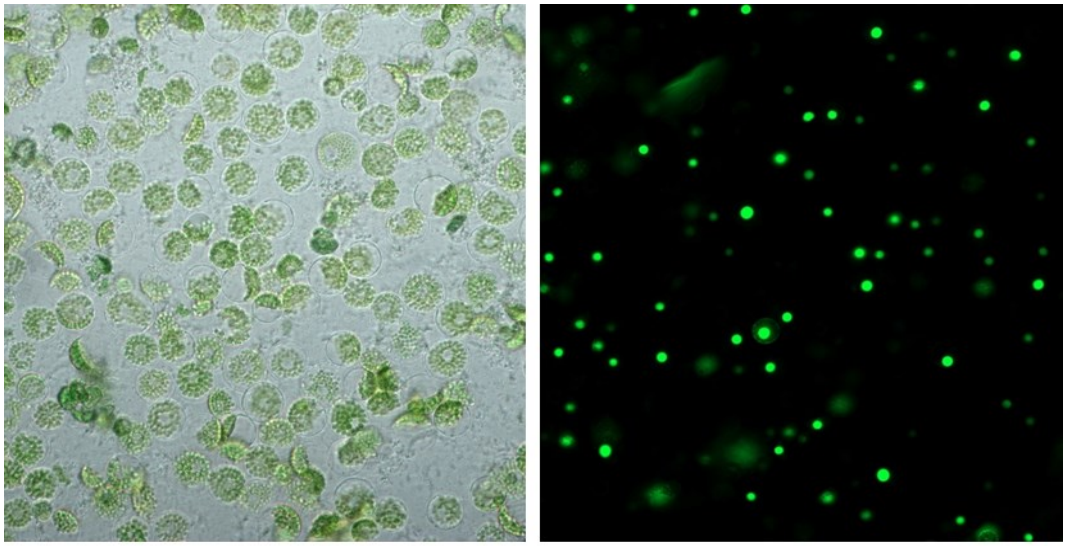

Figure 4.15. Detection of transfection efficiency using fluorescence microscopy at 20x magnification. On the left side is the bright field showing healthy and viable protoplast. On the right side is GFP-fluorescence signal. GFP-signal in the right side of image depicts the transfected protoplasts with the constructed vectors.

To calculate the transfection ratio of protoplast, the GFP-Expressing protoplast were manually counted (Figure 4.15) against the bright filter. As there were three replicates for counting with each vector (Table 4.8), the average of all three replicates was considered the transfection efficiency. The vector (pDD43) showed higher transfection rate against the vector (pDD41) in both genotypes (Table 4.8; 4.9). After 48 h of protoplast transfection, DNA was isolated and prepared for amplicon sequencing.

Table 4.8. Transfection percentage of protoplasts in Lo7 genotype.

Genotype: Lo7				
Vector ID	**Replication**	**No. protoplasts in BF** (Bright field filter)	**No. protoplasts in FL** (Fluorescence filter)	**Transfection rate**
pDD45	R1	488	279	57.1%
pDD45	R2	165	120	72.7%
pDD45	R3	587	442	75.2%
pDD47	R1	534	297	55.6%
pDD47	R2	680	426	62.6%
pDD47	R3	560	457	81.6%

Table 4.9. Transfection percentage of protoplasts in Nr5 genotype.

Genotype: Nr5				
Vector ID	**Treatment**	**No. protoplasts in BF** (Bright field filter)	**No. protoplasts in FL** (Fluorescence filter)	**Transfection rate**
pDD45	R1	519	276	53.1%
pDD45	R2	465	368	79.13%
pDD45	R3	609	473	77.6%
pDD47	R1	592	243	41.0%
pDD47	R2	581	402	69.1%
pDD47	R3	472	396	83.8%

4.3.1 Amplicons from transfected protoplast

After isolating the DNA from transfected protoplasts, PCR was run with adjacent primers to amplify target specific DNA regions. As there were two targeted regions, two primer sets were used for the amplification process. Annealing temperature for primers was set with gradient PCR beforehand using gnomic DNA of wild type Lo7 and Nr5 genotypes and optimal annealing temperature was 62 ºC.

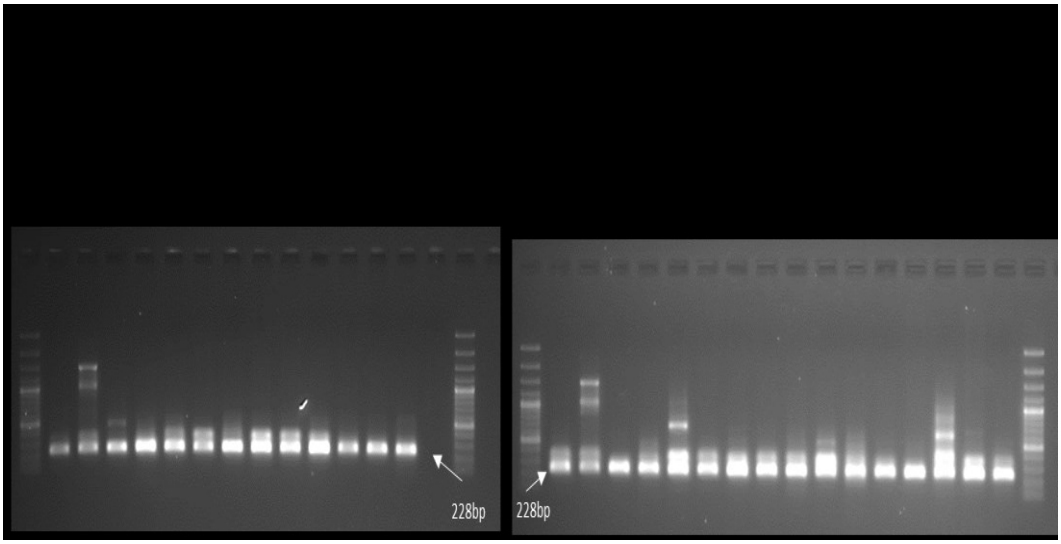

Figure 4.16. Target specific amplicons (target motif 1+2) for both genotypes Lo7 &Nr5 of mutated rye.

The amplified region for all the samples in target motif 1 and 2 (Figure 4.16) shows the same band size with some samples having double or multiple bands that might indicate the mutation induction in the targeted region. There were three replicates for each vector and correct bands (228 bp) were seen in all the replicates along with the wild type. Amplified target region including of target motif 3 and 4 did not show clear bands on the gel (Figure 4.17), except few samples which showed the correct band size (234 bp) but were not very clear on the gel.

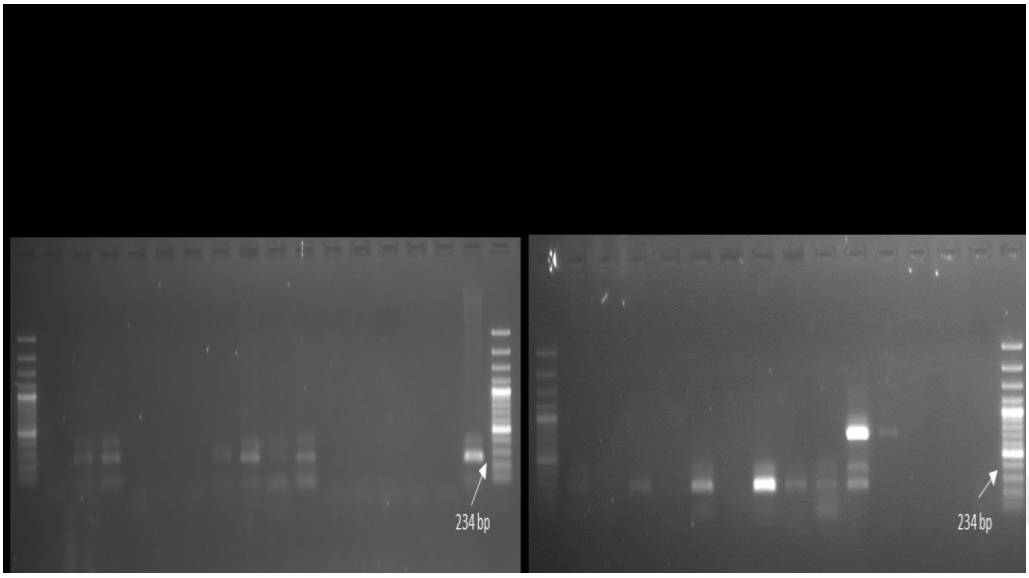

Figure 4.17. Target specific amplicons (target motif 3+4) for both the genotypes (Lo7 &Nr5) of mutated rye protoplasts for different binary vectors and the region amplified is of 234 base pairs.

After getting the amplicon results, we have sent the samples for amplicon deep sequencing to exactly pinpoint the mutations in the protoplast assay which could give us an educated guess for mutations to be induced in stable transformation. Unfortunately, the time frame for my master thesis ended and the results of amplicon deep sequencing had not arrived at the time of my thesis submission and I was not able to analyze the sequencing results.

5. SUMMARY AND DISCUSSION

Generating knockout mutants is well established and straightforward in many plants e.g. cereals through RNA guided Cas9 endonuclease technology but rye has yet to be introduced to this facile technique for plant improvement. Rye is less responsive to plant regeneration and transformation and highly genotype dependent. Due to the recalcitrance to plant transformation, there is no protocol for genome editing in rye yet as regeneration and transformation are pre-requisites for establishing successful genome editing platform. In this study, I tried to establish a protocol for genome editing in rye by knocking out *PLA1* gene to generate haploid inducing lines via RNA guided Cas9 endonuclease and validated CRSIPR/Cas9 constructs in transient expression system e.g. rye protoplasts before starting the procedure of stable transformation. In addition, genome of rye inbred line (Lo7) was sequenced recently (October, 2021) which made the applications of CRISPR systems available.

Haploid induction has been achieved in maize by 4bp frameshift mutation in *PLA1 (Phospholipase A1)* coding sequence via RNA guided Cas9 endonuclease and similar outcome was expected in rye. Orthologue of maize *PLA1* gene was addressed in rye genome and significant protein similarity (85%) was found, further *in-silico* analysis indicated that there are 4 exons and 3 introns in rye *PLA1* gene. For further confirmation and validation of *PLA1* rye orthologue, an expression analysis by RT-PCR of phospholipase1-protein in pollen is still to be performed. I designed four gRNAs for *PLA1* gene for RNA guided Cas9 endonuclease application and employed CASCADE/Cas9 modular vector system based on golden gate principle to clone the expression cassettes for mutation induction. Four binary vectors were assembled to induce mutation in *PLA1*. Because not enough information is available about Cas9 based genome editing in rye, I decided to insert each gRNA two times in each binary vector to increase the potential efficiency of inducing mutation.

Two rye genotypes for rapid transient expression system with two constructed binary vectors (pDD45 and pDD47) with *sfGFP* were used for functional validation through protoplast transfection assay. I did not observe major difference in the

expression of binary vectors in transient expression for genotypes but efficiency of mutated protoplasts for tested vectors differed from 10-15% for each replication. The modular vector (pDD45) with first two repeated guides (gRNA1 and gRNA2) showed higher efficiencies compared to the last two guides (gRNA3 and gRNA4) in vector pDD47. The expression of binary vectors in protoplasts was confirmed under the fluorescence microscopy by *sfGFP* signal. Transfection percentage was counted manually by comparing the transfected protoplasts in fluorescence filter to the bright field filter. The fact that this was the first try for establishing gene knockout platform in rye, I performed two biological replicates of the same experiment to lay the foundation for future experimentation.

After extraction of DNA from protoplasts, target specific amplicons were obtained from selected target regions in *PLA1* gene. I obtained expected band size (228 bp) on gel electrophoresis for first target region (Target motif 1 and 2) for all the replicates in both genotypes with some samples having double and multiple bands, hinting towards mutation induction. The multiple bands in the gel of amplified region might indicate insertion of a large fragment, which should be further confirmed by amplicon sequencing. In contrast to the first target region, no bands on gel electrophoresis were observed for second target region (Target motif 3 and 4) except for couple of samples where bands were not clearly visible. Samples were prepared for next generation (amplicon) sequencing to exactly pinpoint the mutated region in *PLA1* gene which could formulate a foundation before going into stable transformation.

I have verified the presence of *PLA1* gene in rye through *in-silico* analysis with the protein sequence similarity of 85% with the PLA1 protein sequences of maize and durum wheat. In addition, CASCADE/Cas9 modular vector system based on golden gate cloning showed good efficiency in cultivating expression cassettes for transient gene expression. The rye protoplast assay has given an indication towards the success of CASCADE/Cas9 modular vectors with 75% transfection efficiency, indicating towards possible mutation induction but to be confirmed by amplicon sequencing results. It could also result in higher CRISPR/Cas efficiencies in stably transformed rye plants as well but I also realize this at the same time that the results obtained using protoplasts assay may not necessarily reflect the situation in stable transgenic lines due to differences in expression patterns and the ability of the cells to repair DNA.

6. CONCLUSION

I believe that this study will provide a useful exemplification for RNA guided Cas9 endonuclease application, opening ways for haploidy technology through induced mutagenesis in rye. With the fact that it is the first try for transient gene expression in rye, this study can provide a solid base for stable transformation through CRISPR applications. I also believe that this study entails the potential for rye crop improvement, enabling scientists to target different agronomic traits with the applications of CRISPR systems.

REFERENCES

Alijosius, S., Svirmickas, G.J., Bliznikas, S., Gruzauskas, R., Sasyte, V., Raceviciute-Stupeliene, A., Kliseviciute, V.and Dauksienr, A., 2016, Grain Chemical Composition of Different Varieties of Winter Cereals. Zemdirbyste-Agriculture. 103(3), 273–280. DOI: https://doi.org/10.13080/z-a.2016.103.035.

Angioloni, A. and Collar, C., 2011, Nutritional and functional added value of oat, Kamut®, spelt, rye and buckwheat versus common wheat in breadmaking. Journal of the Science of Food and Agriculture, 91(7), 1283-1292.

Aoki, J., Inoue, A., Makide, K., Saiki, N. and Arai, H., 2007, Structure and function of extracellular phospholipase A1 belonging to the pancreatic lipase gene family. Biochimie, 89(2), 197-204.

Arora, L. and Narula, A., 2017, Gene editing and crop improvement using CRISPR-Cas9 system. Frontiers in plant science, 8, 1932.

Bartoš, J., Paux, E., Kofler, R., Havránková, M., Kopecký, D., Suchánková, P. and Doležel, J., 2008, A first survey of the rye (Secale cereale) genome composition through BAC end sequencing of the short arm of chromosome 1R. BMC Plant Biology, 8(1), 1-12.

Bauer, E., Schmutzer, T., Barilar, I., Mascher, M., Gundlach, H., Martis, M. M. and Scholz, U., 2017, Towards a whole-genome sequence for rye (Secale cereale L.). The Plant Journal, 89(5), 853-869.

Bauer, Eva., Schmutzer, Thomas., Barilar, Ivan., Mascher, Martin., Gundlach, Heidrun., Martis, Mihaela., Twardziok, Sven.,Hackauf, Bernd., Gordillo, Andrés., Wilde, Peer.,Schmidt, Malthe., Korzun, Viktor., Mayer, Klaus., Schmid, Karl., Schön, Chris-Carolin.and Scholz, Uwe., 2016, Towards a whole-genome sequence for rye (Secale cereale L.). The Plant Journal. 89. 10.1111/tpj.13436.

Bernstein, N. K., Williams, R. S., Rakovszky, M. L., Cui, D., Green, R., Karimi-Busheri, F. and Glover, J. M., 2005, The molecular architecture of the mammalian DNA repair enzyme, polynucleotide kinase. Molecular cell, 17(5), 657-670.

REFERENCES (continued)

Bushuk, W., 2001, Rye production and uses worldwide. Cereal Foods World, 46(2), 70-73.

Cai, C., Qian, L., Jiang, S., Sun, Y., Wang, Q., Ma, D. and Li, K., 2017, Loss-of-function myostatin mutation increases insulin sensitivity and browning of white fat in Meishan pigs. Oncotarget, 8(21), 34911.

Casini, A., Olivieri, M., Petris, G., Montagna, C., Reginato, G., Maule, G. and Cereseto, A., 2018, A highly specific SpCas9 variant is identified by in vivo screening in yeast. Nature biotechnology, 36(3), 265-271.

Casini, A., Storch, M., Baldwin, G. S. and Ellis, T., 2015, Bricks and blueprints: methods and standards for DNA assembly. Nature Reviews Molecular Cell Biology, 16(9), 568-576.

Chang, H. H., Pannunzio, N. R., Adachi, N. and Lieber, M. R., 2017, Non-homologous DNA end joining and alternative pathways to double-strand break repair. Nature reviews Molecular cell biology, 18(8), 495-506.

Chase, S.S., 1963, Analytic breeding in Solanum tuberosum L.–a scheme utilizing parthenotes and other diploid stocks. Canadian Journal of Genetics and Cytology, 5(4), 359-363.

Chaudhary, A., Gustafson, D. and Mathys, A., 2018, Multi-indicator sustainability assessment of global food systems. Nature communications, 9(1), 1-13.

Chen, J., Li, S., He, Y., Li, J. and Xia, L., 2022, An update on precision genome editing by homology-directed repair in plants. Plant Physiology, 188(4), 1780-1794.

Cho, S. W., Kim, S., Kim, J. M. and Kim, J. S., 2013, Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nature biotechnology, 31(3), 230-232.

REFERENCES (continued)

Clarke, R., Heler, R., MacDougall, M. S., Yeo, N. C., Chavez, A., Regan, M. and Merrill, B.J., 2018, Enhanced bacterial immunity and mammalian genome editing via RNA-polymerase-mediated dislodging of Cas9 from double-strand DNA breaks. Molecular cell, 71(1), 42-55.

Coe Jr, E. H., 1959, A line of maize with high haploid frequency. The American Naturalist, 93(873), 381-382.

Deleu, L.J., Lemmens, E., Redant, L.and Delcour, J.A., 2020**,** The major constituents of rye (SECALE CEREALE L.) flour and their role in the production of rye bread, a food product to which a multitude of health aspects are ascribed. Cereal Chem.; 97: 739– 754. https://doi.org/10.1002/cche.10306\

Dolezel, J., Greilhuber, J., Lucretti, S., Meister, A., Lysak, M.A., Nardi, L. and Obermayer, R., 1998, Plant genome size estimation by flow cytometry: Inter-laboratory comparison. Annals of Botany. ;82:17–26. doi: 10.1006/anbo.1998.0730. [CrossRef] [Google Scholar]

Dong, X., Xu, X., Miao, J., Li, L., Zhang, D., Mi, X. and Chen, S., 2013, Fine mapping of qhir1 influencing in vivo haploid induction in maize. Theoretical and applied genetics, 126(7), 1713-1720.

Eid, A., Alshareef, S. and Mahfouz, M.M., 2018, CRISPR base editors: genome editing without double-stranded breaks. Biochemical Journal, 475(11), 1955-1964.

El-Mahis, A., Baky, M.H. and Farag, M.A., 2022, How Does Rye Compare to other Cereals? A Comprehensive Review of its Potential Nutritional Value and Better Opportunities for its Processing as a Food-Based Cereal. Food Reviews International, 1-24.

Engler, C. and Marillonnet, S., 2014, Golden gate cloning. In DNA cloning and assembly methods (pp. 119-131). Humana Press, Totowa, NJ.

REFERENCES (continued)

Engler, C., Gruetzner, R., Kandzia, R. and Marillonnet, S., 2009, Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. PloS one, 4(5), e5553.

Eudes, F., Acharya, S., Laroche, A., Selinger, L. B. and Cheng, K. J., 2003, A novel method to induce direct somatic embryogenesis, secondary embryogenesis and regeneration of fertile green cereal plants. Plant Cell, Tissue and Organ Culture, 73(2), 147-157.

Evans, M. M., 2007, The indeterminate gametophyte1 gene of maize encodes a LOB domain protein required for embryo sac and leaf development. The Plant Cell, 19(1), 46-62.

Farag, M. A., Xiao, J. and Abdallah, H. M., 2022, Nutritional value of barley cereal and better opportunities for its processing as a value-added food: acomprehensive review. Critical Reviews in Food Science and Nutrition, 62(4), 1092-1104.

Forster, B. P., Heberle-Bors, E., Kasha, K. J. and Touraev, A., 2007, The resurgence of haploids in higher plants. Trends in plant science, 12(8), 368-375.

Friedberg, E. C., Bardwell, A. J., Bardwell, L., Feaver, W. J., Kornberg, R. D., Svejstrup, J. Q. and Wang, Z., 1995, Nucleotide excision repair in the yeast Saccharomyces cerevisiae: its relationship to specialized mitotic recombination and RNA polymerase II basal transcription. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 347(1319), 63-68.

Garneau, J. E., Dupuis, M. È., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P.and Moineau, S., 2010, The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature, 468(7320), 67-71.

Gearing, M. golden gate cloning. introduction to plasmids 101, 51.

Geiger, H. H. and Gordillo, G. A., 2009, Doubled haploids in hybrid maize breeding. Maydica, 54(4), 485.

REFERENCES (continued)

Gilles, L. M., Martinant, J. P., Rogowsky, P. M. and Widiez, T., 2017, Haploid induction in plants. Current Biology, 27(20), R1095-R1097.

Gomand, S. V., Verwimp, T., Goesaert, H. and Delcour, J. A., 2011, Structural and physicochemical characterisation of rye starch. Carbohydrate Research, 346(17), 2727-2735.

Gyulai, Ferenc., 2014 "Archaeobotanical overview of rye (Secale Cereale L.) in the Carpathian-basin I. from the beginning until the Roman age". Journal of Agricultural and Environmental Science

Hackauf, B., Bauer, E., Korzun, V. and Miedaner, T., 2017, Fine mapping of the restorer gene Rfp3 from an Iranian primitive rye (Secale cereale L.). Theoretical and Applied Genetics, 130(6), 1179-1189.

Hansen, H.B., Møller, B., Andersen, S.B., Jørgensen, J.R. and Hansen, A., 2004, Grain Characteristics, Chemical Composition, and Functional Properties of Rye (Secale Cereale L.) As Influenced by Genotype and Harvest Year. J. Agric. Food Chem. 52(8), 2282–2291. DOI: https://doi.org/10.1021/jf0307191.

Hillman, Gordon., Hedges, Robert., Moore, Andrew., Colledge, Susan. and Pettitt, Paul., 2001, "New evidence of Lateglacial cereal cultivation at Abu Hureyra on the Euphrates". The Holocene. 11 (4): 383–393. Bibcode: 2001Holoc...11...383H. doi:10.1191/095968301678302823. S2CID 84930632

Hu, H., Schrag, T. A., Peis, R., Unterseer, S., Schipprack, W., Chen, S. and Melchinger, A. E., 2016, The genetic basis of haploid induction in maize identified with a novel genome-wide association method. Genetics, 202(4), 1267-1276.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. and Charpentier, E., 2012, A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. science, 337(6096), 816-821.

REFERENCES (continued)

Jinek, M., Jiang, F., Taylor, D. W., Sternberg, S. H., Kaya, E., Ma, E. and Doudna, J.A., 2014, Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. Science, 343(6176), 1247997.

Jonsson, K., Andersson, R., Knudsen, K. E. B., Hallmans, G., Hanhineva, K., Katina, K. and Landberg, R., 2018, Rye and health-Where do we stand and where do we go?. Trends in Food Science & Technology, 79, 78-87.

Jung, W. J. and Seo, Y. W., 2014, Employment of wheat-rye translocation in wheat improvement and broadening its genetic basis. Journal of Crop Science and Biotechnology, 17(4), 305-313.

Kasha, K. J., Kao, K. N., 1970, High frequency haploid production in barley (Hordeum vulgare L.). Nature, 225(5235), 874-876.

Kaur, P., Sandhu, K. S., Bangar, S. P., Purewal, S. S., Kaur, M., Ilyas, R. A.and Razman, M. R., 2021, Unraveling the bioactive profile, antioxidant and dna damage protection potential of rye (Secale cereale) flour. Antioxidants, 10(8), 1214.

Kelliher, T., Starr, D., Richbourg, L., Chintamanani, S., Delzer, B., Nuccio, M. L. and Martin, B., 2017, MATRILINEAL, a sperm-specific phospholipase, triggers maize haploid induction. Nature, 542(7639), 105-109.

Kelliher, T., Starr, D., Su, X., Tang, G., Chen, Z., Carter, J. and Que, Q., 2019, One-step genome editing of elite crop germplasm during haploid induction. Nature biotechnology, 37(3), 287-292.

Kelliher, T., Starr, D., Wang, W., McCuiston, J., Zhong, H., Nuccio, M. L., & Martin, B., 2016, Maternal haploids are preferentially induced by CENH3-tailswap transgenic complementation in maize. Frontiers in Plant Science, 7, 414.

Kent, T., Chandramouly, G., McDevitt, S. M., Ozdemir, A. Y. and Pomerantz, R. T., 2015, Mechanism of microhomology-mediated end-joining promoted by human DNA polymerase θ. Nature structural & molecular biology, 22(3), 230-237.

REFERENCES (continued)

Kumar, V. and Jain, M., 2015, The CRISPR–Cas system for plant genome editing: advances and opportunities. Journal of experimental botany, 66(1), 47-57.

Kumlehn, J., Pietralla, J., Hensel, G., Pacher, M. and Puchta, H., 2018, The CRISPR/Cas revolution continues: from efficient gene editing for crop breeding to plant synthetic biology. Journal of integrative plant biology, 60(12), 1127-1153.

Laurie, D. A., Bennett, M. D., 1988, The production of haploid wheat plants from wheat x maize crosses. Theoretical and applied genetics, 76(3), 393-397.

Li, G., Wang, L., Yang, J., He, H., Jin, H., Li, X., and Wang, D., 2021, A high-quality genome assembly highlights rye genomic characteristics and agronomically important genes. Nature genetics, 53(4), 574-584.

Li, X., Meng, D., Chen, S., Luo, H., Zhang, Q., Jin, W. and Yan, J., 2017, Single nucleus sequencing reveals spermatid chromosome fragmentation as a possible cause of maize haploid induction. Nature communications, 8(1), 1-9.

Lin, C. S., Hsu, C. T., Yang, L. H., Lee, L. Y., Fu, J. Y., Cheng, Q. W. and Shih, M.C., 2018, Application of protoplast technology to CRISPR/Cas9 mutagenesis: from single-cell mutation detection to mutant plant regeneration. Plant biotechnology journal, 16(7), 1295-1310.

Linacero, R., Rueda, J., Esquivel, E., Bellido, A., Domingo, A., and Vázquez, A. M., 2011, Genetic and epigenetic relationship in rye, Secale cereale L, somaclonal variation within somatic embryo-derived plants. In Vitro Cellular & Developmental Biology-Plant, 47(5), 618-628.

Lindahl, T., 1982, DNA repair enzymes. Annual review of biochemistry, 51(1), 61-87.

Liu, C., Li, X., Meng, D., Zhong, Y., Chen, C., Dong, X. and Chen, S., 2017, A 4-bp insertion at ZmPLA1 encoding a putative phospholipase A generates haploid induction in maize. Molecular plant, 10(3), 520-522.

REFERENCES (continued)

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

Ma, R., Guo, Y. D. and Pulli, S., 2003. Somatic embryogenesis and fertile green plant regeneration from suspension cell-derived protoplasts of rye (Secale cereale L.). Plant cell reports, 22(5), 320-327.

Ma, X., Q. Zhu, Y. Chen, and Y. G. Liu. 2016. CRISPR/Cas9 platforms for genome editing in plants: Developments and applications. Molecular Plant 9: 961–974.

Ma, X., Zhu, Q., Chen, Y. and Liu, Y.G., 2016, CRISPR/Cas9 platforms for genome editing in plants: developments and applications. Molecular plant, 9(7), 961-974.

Martinez-Villaluenga, C., Michalska, A., Frías, J., Piskula, M. K., Vidal-Valverde, C. and Zieliński, H., 2009, Effect of flour extraction rate and baking on thiamine and riboflavin content and antioxidant capacity of traditional rye bread. Journal of Food Science, 74(1), C49-C55.

McElroy, J. Scott., 2014, "Vavilovian Mimicry: Nikolai Vavilov and His Little-Known Impact on Weed Science". Weed Science. 62 (2): 207–216. doi:10.1614/ws-d-13-00122.1

Meng, D., Liu, C., Chen, S., and Jin, W., 2021, Haploid induction and its application in maize breeding. Molecular Breeding, 41(3), 1-9.

Miedaner, T. and Laidig, F., 2019, Hybrid breeding in rye (Secale cereale L.). In Advances in plant breeding strategies: cereals (pp. 343-372). Springer, Cham.

Monfreda, C., Ramankutty, N., and Foley, J. A., 2008, Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000, Global Biogeochem. Cycles, 22, GB1022, doi:10.1029/2007GB002947.

REFERENCES (continued)

Nadakuduti, S. S.and Enciso-Rodríguez, F., 2021, Advances in genome editing with CRISPR systems and transformation technologies for plant DNA manipulation. Frontiers in plant science, 11, 637159.

Németh, R. and Tömösközi, S., 2021, Rye: Current state and future trends in research and applications. Acta Alimentaria, 50(4), 620-640.

Pariona, A., 2017, The top 10 rye producing countries of the world. In: Worldatlas. https://www.worldatlas.com/articles/the-top-10-rye-producing-countries-of-the-world.html

Peters, J. M., Koo, B. M., Patino, R., Heussler, G. E., Hearne, C. C., Qu, J. and Rosenberg, O. S., 2019, Enabling genetic analysis of diverse bacteria with Mobile-CRISPRi. Nature microbiology, 4(2), 244-250.

Pipiras, E., Coquelle, A., Bieth, A. and Debatisse, M., 1998, Interstitial deletions and intrachromosomal amplification initiated from a double-strand break targeted to a mammalian chromosome. The EMBO journal, 17(1), 325-333.

Popelka, J. and Altpeter, F., 2001, Interactions between genotypes and culture media components for improved in vitro response of rye (Secale cereale L.) inbred lines. Plant Cell Reports, 20(7), 575-582.

Popelka, J. C., and Altpeter, F., 2003, Evaluation of rye (Secale cereale L.) inbred lines and their crosses for tissue culture response and stable genetic transformation of homozygous rye inbred line L22 by biolistic gene transfer. Theoretical and Applied genetics, 107(4), 583-590.

Prigge, V. and Melchinger, A. E., 2012, Production of haploids and doubled haploids in maize. In Plant cell culture protocols (pp. 161-172). Humana Press, Totowa, NJ.

Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P. and Lim, W.A., 2013, Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell, 152(5), 1173-1183.

REFERENCES (continued)

Rabanus-Wallace, M. T., Hackauf, B., Mascher, M., Lux, T., Wicker, T., Gundlach, H. and Stein, N., 2021, Chromosome-scale genome assembly provides insights into rye biology, evolution and agronomic potential. Nature genetics, 53(4), 564-573.

Rakoczy-Trojanowska, M. and Malepszy, S., 1995, Genetic factors influencing the regeneration ability of rye (Secale cereale L.). II. Immature embryos. Euphytica, 83(3), 233-239.

Reski, R., 1998, Development, genetics and molecular biology of mosses. Botanica Acta, 111(1), 1-15.

Schiml, S., Fauser, F. and 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.

Sfeir, A. and Symington, L. S., 2015, Microhomology-mediated end joining: a back-up survival mechanism or dedicated pathway? Trends in biochemical sciences, 40(11), 701-714.

Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelsen, T. S. and Zhang, F., 2014, Genome-scale CRISPR-Cas9 knockout screening in human cells. Science, 343(6166), 84-87.

Shan, S., E.V. Mavrodiev, R. Li, Z. Zhang., B. A. Hauser., P. S. Soltis., D. E. Soltis., and B. Yang., 2018. Application of CRISPR/Cas9 to Tragopogon (Asteraceae), an evolutionary model for the study of polyploidy. Molecular Ecology Resources 18: 1427–1443

Shan, S., Soltis, P. S., Soltis, D. E. and Yang, B., 2020, Considerations in adapting CRISPR/Cas9 in nongenetic model plant systems. Applications in plant sciences, 8(1), e11314.

Shou, J., Li, J., Liu, Y. and Wu, Q., 2018, Precise and predictable CRISPR chromosomal rearrangements reveal principles of Cas9-mediated nucleotide insertion. Molecular cell, 71(4), 498-509.

REFERENCES (continued)

Sprague, G. F., 1929, Hetero-fertilization in maize. Science, 69(1794), 526-527.

Svitashev, S., Young, J. K., Schwartz, C., Gao, H., Falco, S. C. and Cigan, A. M., 2015, Targeted mutagenesis, precise gene editing, and site-specific gene insertion in maize using Cas9 and guide RNA. Plant physiology, 169(2), 931-945.

Targońska, M., Hromada-Judycka, A., Bolibok-Brągoszewska, H. and Rakoczy-Trojanowska, M., 2013, The specificity and genetic background of the rye (Secale cereale L.) tissue culture response. Plant cell reports, 32(1), 1-9.

Wang, B., Zhu, L., Zhao, B., Zhao, Y., Xie, Y., Zheng, Z. and Wang, H., 2019, Development of a haploid-inducer mediated genome editing system for accelerating maize breeding. Molecular plant, 12(4), 597-602.

Wang, B., Zhu, L., Zhao, B., Zhao, Y., Xie, Y., Zheng, Z. and Wang, H., 2019, Development of a haploid-inducer mediated genome editing system for accelerating maize breeding. Molecular plant, 12(4), 597-602.

Wang, G., Ryu, S. and Wang, X., 2012, Plant phospholipases: an overview. Lipases and phospholipases, 123-137.

Wang, H., Jiang, J., Chen, S., Qi, X., Peng, H., Li, P. and Chen, F., 2013, Next-generation sequencing of the Chrysanthemum nankingense (Asteraceae) transcriptome permits large-scale unigene assembly and SSR marker discovery. PloS one, 8(4), e62293.

Wang, M., van Bergen, S. and Van Duijn, B., 2000, Insights into a key developmental switch and its importance for efficient plant breeding. Plant physiology, 124(2), 523-530.

Watson, J. D., Sanderson, S., Ezersky, A., Savchenko, A., Edwards, A., Orengo, C. and Thornton, J. M., 2007, Towards fully automated structure-based function prediction in structural genomics: a case study. Journal of molecular biology, 367(5), 1511-1522.

REFERENCES (continued)

Wędzony, M., Forster, B. P., Żur, I., Golemiec, E., Szechyńska-Hebda, M., Dubas, E., and Gotębiowska, G., 2009, Progress in doubled haploid technology in higher plants. Advances in haploid production in higher plants, 1-33.

Werner, S., Engler, C., Weber, E., Gruetzner, R. and Marillonnet, S., 2012, Fast track assembly of multigene constructs using Golden Gate cloning and the MoClo system. Bioengineered, 3(1), 38-43.

Wijnker, E., Deurhof, L., Van De Belt, J., De Snoo, C. B., Blankestijn, H., Becker, F. and Keurentjes, J. J., 2014, Hybrid recreation by reverse breeding in Arabidopsis thaliana. Nature protocols, 9(4), 761-772.

Wilde, P. and Miedaner, T., 2021, Hybrid rye breeding. In The Rye Genome (pp. 13-41). Springer, Cham.

Wrigley, C. and Bushuk, W., 2017, Rye: grain-quality characteristics and management of quality requirements. In Cereal grains (pp. 153-178). Woodhead Publishing.

Wu, S., Zhu, H., Liu, J., Yang, Q., Shao, X., Bi, F. and Yi, G., 2020, Establishment of a PEG-mediated protoplast transformation system based on DNA and CRISPR/Cas9 ribonucleoprotein complexes for banana. BMC plant biology, 20(1), 1-10.

Xu, X., Li, L., Dong, X., Jin, W., Melchinger, A. E. and Chen, S., 2013, Gametophytic and zygotic selection leads to segregation distortion through in vivo induction of a maternal haploid in maize. Journal of experimental botany, 64(4), 1083-1096.

Xue, C. and Greene, E.C., 2021, DNA repair pathway choices in CRISPR-Cas9-mediated genome editing. Trends in Genetics, 37(7), 639-656.

Yassitepe, J. E. D. C. T., da Silva, V. C. H., Hernandes-Lopes, J., Dante, R. A., Gerhardt, I. R., Fernandes, F. R. andArruda, P., 2021, Maize Transformation: From Plant Material to the Release of Genetically Modified and Edited Varieties. Frontiers in Plant Science, 12.

REFERENCES (continued)

Yoo, S. D., Cho, Y. H. andSheen, J., 2007, Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. Nature protocols, 2(7), 1565-1572.

Yue, J. J., Yuan, J. L., Wu, F. H., Yuan, Y. H., Cheng, Q. W., Hsu, C. T., & Lin, C.S., 2021, Protoplasts: From isolation to CRISPR/Cas genome editing application. Frontiers in Genome Editing, 3.

Zhang, B., Wu, S., Zhang, Y. E., Xu, T., Guo, F., Tang, H. and Xue, Y., 2016, A high temperature-dependent mitochondrial lipase EXTRA GLUME1 promotes floral phenotypic robustness against temperature fluctuation in rice (Oryza sativa L.). PLoS Genetics, 12(7), e1006152.

Zhang, Z., Hua, L., Gupta, A., Tricoli, D., Edwards, K. J., Yang, B. and Li, W., 2019, Development of an Agrobacterium-delivered CRISPR/Cas9 system for wheat genome editing. Plant biotechnology journal, 17(8), 1623-1635.

Zimny, J. and Michalski, K.,2019, Development of in vitro culture techniques for advancement of rye (Secale cereale L.) breeding. Acta Biologica Cracoviensia s. Botanica, 61(1).

Zohary, D., Hopf, M. and Weiss, E., 2012, Domestication of Plants in the Old World: The origin and spread of domesticated plants in Southwest Asia, Europe, and the Mediterranean Basin. Oxford University Press.

ACKNOWLEDGEMENT

I am so grateful to the most merciful and almighty Allah without whose grace, I would never have been able to go after my higher studies in the field of plant sciences and to complete my thesis for the fulfilment of Master degree in Plant Breeding.

I would like to express my ever indebtedness, deepest sense of gratitude, sincere appreciation and profound regards to my reverend and beloved Co supervisor Dr. Diaa Eldin S. DAGHMA from Plant Reproductive Biology Research Group (IPK) Gatersleben, Germany for his scholastic supervision, uncompromising principles, valuable advice, constant inspiration and motivation. He supported me during the whole period of lab work as well as during thesis writing. The guidance helped me to find the perfect fit for my thesis research. I am immense grateful for his patience and cordial help during my difficult times. I am also grateful to Dr. Jochen KURNLEHN for insightful suggestions and providing the opportunity to work in Plant Reproductive Biology Research Group (IPK).

I would like to thank my mentor and supervisor Assoc. Prof. Dr. Evren KOBAN BAŞTANLAR from the Ege University for her kind guidance and help in perfecting my work as well as the administrative functions.

I am grateful to Leibniz Institute of Plant genetics and crop plant research (IPK) Gatersleben, Germany for providing the research facilities to complete my thesis work. I would like to thank Erasmus Mundus (emPLANT) authorities for selecting me for this program with great journey of two years at different universities in different countries.

Finally, I found no words to thank my beloved parents and family members for their unquantifiable love and continuous moral support, sacrifice, never ending affection and immense belief on my ability for higher study.

14 / 09 / 2022

Muhammad ARSLAN

CURRICULUM VITAE

Muhammad ARSLAN completed his Bachelor of Science in Agriculture (Hons.) majoring in Plant Breeding and Genetics from MNS University of Agriculture Multan, Pakistan with a CGPA 3.98 out of CGPA 4.00 grade. He was admitted to Erasmus Mundus Master Program in Plant Breeding (emPLANT) at Institute Polytechnique UniLaSalle, Beauvais, France where he completed his first year. He did his summer internship at Bayer Crop Science, Monbequi, France for 4 months. He started his second year of emPLANT program at Ege University, Izmir, Turkey. He did his master thesis work for 7 months in Leibniz Institute of Plant Genetics and Crop plant research (IPK) Gatersleben, Germany in 2022.