CRISPR/Cas9-Based Editing of Alpha-Amylase 3 (*Amy3*) Genes in Wheat

A Thesis

Presented in Partial Fulfillment of the Requirements for the

Degree of Master of Science

with a

Major in Plant Science

in the

College of Graduate Studies

University of Idaho

by

Mengmeng Lin

Major Professor: Daolin Fu, Ph.D.

Committee Members: Allan Caplan, Ph.D.; Samuel Hunter, Ph.D.

Department Administrator: Robert Tripepi, Ph.D.

May 2019

Authorization to Submit Thesis

This thesis of Mengmeng Lin, submitted for the degree of Master of Science with a Major in Plant Science and titled "CRISPR/Cas9-Based Editing of Alpha-Amylase 3 (*Amy3*) Genes in Wheat," has been reviewed in final form. Permission, as indicated by the signatures and dates below, is now granted to submit final copies to the College of Graduate Studies for approval.

Major Professor: _____ Date _____

Daolin Fu, Ph.D.

Committee Members: _____ Date _____

Allan Caplan, Ph.D.

_____ Date _____

Samuel Hunter, Ph.D.

Department Administrator: _____ Date _____

Robert Tripepi, Ph.D.

Abstract

Wheat is the second largest cereal crop in the United States. Unseasonal activation of *α*-amylase genes in developing kernels may lead to low falling numbers in wheat, which can compromise the end-use quality and causes substantial profit losses to wheat growers. In wheat, there are four subfamilies of *α*-amylases; each subfamily has multiple *α*-amylase genes. The *Amy3.1* and *Amy3.2* of *α*-amylase III (*Amy3*) are mainly expressed in developing kernels.

The goal of the present study was to test whether these two genes affect wheat falling numbers by creating mutants in each of these two groups of genes. First, the Clustered Regularly-Interspaced Short Palindromic Repeats (CRISPR) system was used to edit these two groups of genes in tetraploid wheat 'Kronos' and hexaploid wheat 'CB037'. The preliminary sequencing results showed a gene editing efficiency of 33.3%. In the preliminary tests, fourteen plants with edited versions of *Amy3.1* and/or *Amy3.2* genes were obtained. Second, using the Targeting Induced Local Lesions IN Genomes (TILLING), 21 mutants of *Amy3.1* and *Amy3.2* were identified in the tetraploid wheat 'Kronos'. The identified CRISPR and TILLING mutants of *Amy3* will be used to understand the gene's role in regulating falling numbers and other essential traits in wheat.

Acknowledgements

This thesis was completed under the guidance of my supervisor Dr. Daolin Fu, who gave me the most careful advice and warm encouragement from the first experimental design to the final writing of the paper. Dr. Fu has been my first teacher of scientific research. Under his careful guidance, I have grown from an undergraduate who knows very little about scientific research to a graduate who is familiar with various experimental skills, can think independently, raise pertinent questions and solve problems. In addition to scientific research, Dr. Fu has also been a mentor to me. I have been deeply moved by his rigorous academic attitude, broad scientific research vision, diligent and industrious professional dedication, open-minded attitude towards life, as well as his persistence and enthusiasm for research.

With Dr. Fu's encouragement and help, I came to the United States and learned many new skills and a broad spectrum of knowledge, which has broadened my horizons and made me a better person. So, I would like to express my heartfelt thanks to Dr. Fu for his concern and care for me over the years. You are a true scholar, and it has been my honor to be your student and to have had a full and wonderful three years.

I would like to thank my committee members, Dr. Samuel Hunter, and Dr. Allan Caplan, for their selfless help and support in scientific research and life. Samuel Hunter helped me a great deal in the analysis of my sequencing results. Dr. Allan Caplan gave me a lot of advice on experiments, which inspired me a lot.

I would like to express my gratitude to all the faculty, staff and students of the Department of Plant Sciences at the University of Idaho for a friendly, supportive and intellectually stimulating graduate experience. It was your genuine smiling faces and warm words that made me feel welcome when I was in a foreign country.

I would also like to thank all current and previous members in our lab. Dr. Yanmei Xiao and Miss Huifei Zhang contributed significantly on wheat transformation. Dr. Fei Ni helped me greatly in analyzing the sequence data. I would like to thank Katrina Johnson for her help in language learning. Dr. Bo Lyu also gave me a lot of practical suggestions in my experiments and helped me socially. I would also want to thank Dr. Daniel New, Dr. Jiwen Qiu, Qunqun Hao, Chaozhong Zhang, Meng Su, Matthew Fagnan at the University of Idaho, and Dr. Jiajie Wu, Meng Sun and others at the Shandong Agricultural University in China. Thank you for your

enthusiasm and help in my research and personal life. Our time together has been a precious treasure to me.

This study was supported by the Commodity grant (AP2725) from the Idaho Wheat Commission, the NACA program (2090-21000-033) from the USDA Agricultural Research Service, and the Hatch grant (IDA01587) from the USDA National Institute of Food and Agriculture. This study was also supported by the Idaho Agricultural Experiment Station, University of Idaho. Data collection and analyses performed by the IBEST Genomics Resources Core at the University of Idaho were supported in part by NIH COBRE grant P30GM103324.

Dedication

Most importantly, I would like to thank my family. I can never repay my parents, Yongzhong Lin, and Aiyun Luan, for their innumerable sacrifices they made over the years to provide me a quality education, abundant opportunities, and a loving home in which to grow up.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Cas9	CRISPR-associated protein 9 (Cas9)
CRISPR	clustered regularly interspaced short palindromic repeats
DPA	days post anthesis
DSBs	double-strand breaks
EMS	ethyl methanesulfonate
FN	falling number
HDR	homology-directed repair
LMA	late maturity alpha-amylase
NHEJ	non-homologous end-joining
PAM	protospacer adjacent motif
PCR	polymerase chain reaction
PHS	pre-harvest sprouting
pI	isoelectric point
RNP	ribonucleoprotein
sgRNA	single guide RNA
SNPs	single nucleotide polymorphisms
TALENs	transcription activator-like effector nucleases
TILLING	targeting induced local lesions in genomes

ZFNs zinc-finger nucleases

Chapter 1 Introduction

1.1 Research Background

Wheat is an important food crop worldwide. It is the main source of carbohydrates and protein in the human diet. According to the Food and Agriculture Organization of the United Nations, the world's wheat production reached 749.46 million tons in 2016, ranking behind corn as the second largest food crop. Wheat is of particular economic importance to the U.S. Pacific Northwest. Providing quality wheat for export is critical to the survival of the wheat industry. In this thesis, a portion of the genetic basis for the quality of wheat was studied in order to meet the needs of farmers, processors and wheat breeders in the future.

1.1.1 The Classification of Wheat *α*-Amylase

In wheat grain, starch accounts for more than 70% of the total dry weight and is composed of glucose-residues connected by *α*-1,4 glucosidic bonds. Starch forms branches through *α*-1,6 glucosidic bonds (Ball *et al.*, 2003). Amylase is a general term for a class of enzymes that hydrolyze these starch glucosidic bonds. According to different hydrolysis methods, amylase can be divided into *α*-amylase, *β*-amylase, isoamylase, and saccharifying amylase.

α-Amylase (EC 3.2.1.1) is one of the most essential enzymes to catalyze the hydrolysis of starch. It is an endohydrolase, belonging to the 13th family of glucosidases (Majzlová *et al.*, 2013). The *α*-1,4 glucosidic bonds in starch, glycogen, oligosaccharides or polysaccharides can be randomly cut by *α*-amylase to produce reducing sugars and dextrins. The structure of the carbon atom at the end of the product is an alpha type, so it is known as *α*-amylase (Huang *et al.*, 1996; Gale *et al.*, 2002). In the process of wheat seed formation, nutrients gradually accumulate and *α*-amylase is subsequently synthesized (Gale, 1983). During wheat germination, the activity of *α*-amylase increases sharply, the starch is hydrolyzed into small molecules, and then further hydrolyzed into sugars to promote germination (Humphreys *et al.*, 2002).

In different plants, the type and function of *α*-amylase varies. In cereal plants, rice contains three categories of *α*-amylase: AMY1, AMY2 and AMY3 from ten separate genes (Huang *et al.*, 1992). In barley, *α*-amylase can be divided into four categories: AMY1, AMY2, AMY3, and AMY4 (Radchuk *et al.*, 2009). In wheat, *α*-amylase is highly polymorphic, and three

majior isoforms, including AMY1, AMY2, and AMY3, have been reported (Barrero *et al.*, 2013). At present, people have a clearer understanding of the inheritance and function of AMY1 and AMY2 (Ainsworth *et al.*, 1985), while there are few studies on AMY3. *Amy1* is a multigene subfamily located on the long arm of the wheat homoeologous group 6 chromosomes. In the process of seed germination, the encoded *α*-amylase in the aleurone layer has a high isoelectric point (pI) (Mares *et al.*, 2008; Ral *et al.*, 2016). *Amy2* is located on the long arm of the wheat homoeologous group 7 chromosomes and this subfamily encodes a low pI *α*-amylase (Ral *et al.*, 2016). *Amy3* is located on the long arm of the wheat homoeologous group 5 chromosomes and is expressed primarily during seed development (Baulcombe *et al.*, 1987). In Whan *et al.*'s research, they studied the activity and localization of *α*-amylases in wheat at different stages of grain development (Whan *et al.*, 2014). The results showed that AMY3 was most abundant at all sampled time points from 10–30 days post anthesis (DPA), then followed by AMY2, and no AMY1 protein was detected at any time. AMY2 and AMY3 had similar localization patterns at 20 DPA, with the highest level in the pericarp (Whan *et al.*, 2014).

1.1.2 Wheat Falling Numbers and Factors Affecting Falling Numbers

The falling number (FN) test is a method for determining sprout damage and indicating the activity of *α*-amylase. The falling number reflects the level of *α*-amylase activity in flour. The higher the falling number, the lower the *α*-amylase activity. Pre-harvest sprouting (PHS) and late maturity *α*-amylase (LMA) are the main reasons for low falling values.

Pre-harvest sprouting refers to the phenomenon where the seeds have reached physiological maturity and germinate on the spike prior to harvest in a rainy or humid environment (Gubler *et al.*, 2005). Pre-harvest sprouting is a worldwide phenomenon, and global economic losses resulting from sprouting damage have amounted to approximately $1 billion per year (Black *et al.*, 2006). It is mainly caused by environmental factors including temperature and humidity. Due to differences in humidity, temperature, altitude and other factors, the harm caused by PHS varies from region to region. Genetically, pre-harvest sprouting is mainly determined by seed dormancy (Derera *et al.*, 1977). After reaching physiological maturity, wild material has a certain dormant period to endure poor environment and germinate in the appropriate season (Gubler *et al.*, 2005). In the process of

crop domestication, breeders have preferred to select seeds with weak or no dormancy requirements in order to make seeds germinate quickly and consistently, which results in many selected varieties that have lost the strong dormancy characteristics of ancient wild materials. During spike germination, a series of enzymes are activated, including various proteases and Amy1 and Amy2. This results in the breakdown of starch, protein, and other nutrients stored in grains which seriously affects the quality and subsequent processing of wheat flour. Plant hormones play an important and intermingled role in seed dormancy and germination. Gibberellic acid (GA) is synthesized in the embryo and released into the endosperm during seed formation. When gibberellin diffuses to the aleurone layer, it up-regulates*α*-amylase (Bethke *et al.*, 1997).

Late maturity *α*-amylase is a class of isoenzymes synthesized in the last stages of seed development without pre-harvest sprouting (Barrero *et al.*, 2013). Late maturity *α*-amylase is an unpredictable genetic defect in common wheat (*Triticum aestivum*). It causes higher *α*-amylase activity in mature grains and lower falling numbers (Gale, 1983). Late maturity *α*-amylase has a high isoelectric point and assubmed to be the Amy1 products. In the early stages of wheat grain development, *Amy1* is expressed only in a small number of cells in the aleurone layer stimulated by environmental factors (Mares *et al.*, 2014). Then, *Amy1* continues to be expressed in the aleurone layer during the development of seeds but doesn't affect the morphology and size of the seeds (Mares *et al.*, 2014).

The common characteristic of pre-harvest sprouting and late maturity *α*-amylase is that the *α*-amylase genes areinduced before grain harvest, which leads to a decrease of the falling number and subsequent profit losses to growers. Pre-harvest sprouting and late maturity *α*-amylases cannot be differentiated by a falling number test; both defects are treated similarly when grain is traded. Pre-harvest sprouting increases a series of metabolic activities and affects flour processing quality, but there are few studies on whether late maturity *α*-amylase affects flour processing. Recently, it has been found that overexpression of *Amy3* gene in wheat endosperm can significantly reduce the falling number, which is similar to the situation caused by PHS and LMA (Ral *et al.*, 2016). The falling number is used as an index to measure the quality of wheat flour. When the falling number is low (less than 250 s), the wheat flour is considered degraded, and its price is therefore reduced.

1.1.3 Wheat Mutant Populations and Mutant Screening

The construction of a wheat mutant population is an important basis for studying the functional genomics of wheat. Chemical mutagenesis is one of the important ways to construct mutant populations. Ethyl methane sulfonate (EMS) is a commonly used chemical mutagen which can induce high-density mutations and has the advantages of simplicity, and high point-mutation frequency (Till *et al.*, 2007). EMS has been applied to construct mutant populations in wheat (Mishra *et al.*, 2016), rice (Lee *et al.*, 2002), *Arabidopsis* (Mccallum *et al.*, 2000), maize (Till *et al.*, 2004), and other crops.

Targeting Induced Local Lesions In Genomes (TILLING) technology can effectively identify point mutations from EMS mutagenesis and has been widely used for this purpose (McCallum *et al.*, 2000). In TILLING technology, a series of point mutations are generated through chemical mutagenesis: after PCR (polymerase chain reaction) amplification, the DNA pool consisting of wild types and mutant lines will undergo one step of denaturation and renaturation. This results in DNA strands from the wild-type and mutant lines forming heteroduplexes. The endonuclease CelI, which can specifically identify the mismatched bases in the two DNA strands, is utilized for enzymatic digestion. DNA molecules can be cut at the mismatches and therefore mutants can be detected through gel electrophoresis. This technique combines the chemical mutagenesis method that induces high-frequency point mutations with PCR screening technology to find point mutations in the target region, and it is a high-throughput and low-cost reverse genetics research method (Mccallum *et al.*, 2000).

There is another technique, exon capture, that can detect mutation sites at a genomic scale, and forms a searchable, informative database. Although it requires more investment at the initial stage, this technique can provide data for many future investigations (Henry *et al.*, 2014). With the improvement of sequencing technology, the cost of sequencing has gradually decreased. However, because hexaploidy wheat has a 16-Gb haploid genome, the cost of sequencing whole genomes of thousands of individuals remains very high. Exome capture provides a method for sequencing only genomic coding regions, thus greatly reducing costs.

Using this technique to detect mutation sites in mutant populations has been used in animals and plants (Ng *et al.*, 2009; Ng *et al.*, 2010; Bolon *et al.*, 2011). Tetraploid durum wheat 'Kronos' was developed by Arizona breeders from a male sterile population (d03-21). Krasileva *et al.* sequenced exons of the whole genome of the EMS mutant 'Kronos' population and formed a database (Krasileva *et al.*, 2017). Every line of this mutant population has more than 2,000 mutations identified by multiplexed global exome capture and sequencing technology, and while the entire population had a total of more than 3 million mutations. Related information can be found on an online database (http://dubcovskylab.ucdavis.edu/home). In this study, this 'Kronos' mutant population was used to screen *Amy3* gene mutants by TILLING and blast search.

1.1.4 Genome Editing with CRISPR/Cas9 Technology

In recent years, plant genome editing technology has developed rapidly. Zinc-finger nucleases (ZFNs) (Holmes *et al.*, 2007; Wood *et al.*, 2011; Carroll, 2011), transcription activator-like effector nucleases (TALENs) (Schmidt *et al.*, 2010; Bogdanove, 2013), and today's most commonly applied approach - the clustered regularly interspaced short palindromic repeat (CRISPR) and CRISPR-associated protein 9 (Cas9) (Charpentier *et al.*, 2014; Mao *et al.*, 2008) have all been widely used for the purpose of accurately editing genomes. These three gene editing systems can induce double-strand breaks (DSBs) at a specific target site on DNA, and then the break is subsequently repaired by the nonhomologous end-joining (NHEJ) or homology-directed repair (HDR) pathways (Symington *et al.*, 2011). The NHEJ pathway is often inaccurate, introducing small insertions or deletions (InDel) (Gorbunova, 1997). InDels that occur in a coding region or in a splicing site may disrupt gene function. The HDR pathway is based on the use of homologous DNA as a template to repair the break. Sister chromatids, homologous chromosomes, or artificially supplied exogenous DNA with homologous sequences can serve as templates for HDR repair.

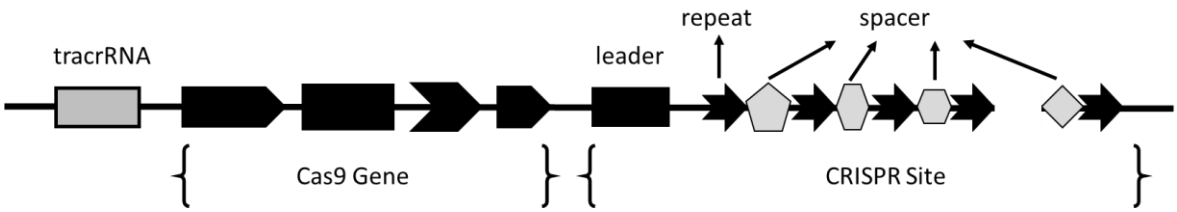


Figure 1. 1 Diagram a native CRISPR/Cas system (Wei *et al.*, 2013)

CRISPR/Cas9 is an RNA-guided genome editing system that has been developed and refined in recent years. The CRISPR sequence was first discovered by Japanese scientists in *E. coli* in 1987 (Ishino *et al.*, 1987). It is an adaptive immune system existing in many bacteria and archaea, which can help defend themselves against invading viral and plasmid DNAs (Richards *et al.*, 2007). The CRISPR/Cas system has a relatively fixed structure, consisting of tracrRNA at the 3' end of the sequence, a *Cas9* coding region and the CRISPR site (Figure 1.1). A small fraction of the invading DNA is incorporated between two repeats within the CRISPR site and becomes what is known as spacers (Bland *et al.*, 2007). When the same DNA enters the cell again, the CRISPR site is transcribed and processed to produce crRNA. crRNA binds to tracrRNA to form a tracrRNA/crRNA complex which activates and guides the Cas9 nuclease to find the sequence target that is complementary to the crRNA and then cleaves the foreign double-stranded DNA (Richards *et al.*, 2007). There are primarily six different types of CRISPR systems (I-VI) (Niewoehner *et al.*, 2014; Wardak, 2014; Sinkunas *et al.*, 2013; Terns, 2018). CRISPR/Cas9 evolved from type II of these. Sequences corresponding to the spacer sequences on phages or plasmids are usually called protospacers. In general, the 5' or 3' end of a protospacer has several bases that are very conserved; this is called the PAM (protospacer adjacent motif). The PAM has a length of 2-5 bases, 1-4 bases apart from the protospacer, and the PAM sequence is generally of the form NGG (Gasiunas *et al.*, 2012). In 2012, Jinek *et al.* found that Cas9 in the Type II CRISPR system is a nuclease, which binds two RNAs (crRNA, tracrRNA) and can cut double-stranded DNA. They constructed a new CRISPR/Cas system, in which crRNA and tracrRNA were connected by a four-base link ring, forming an amalgamation termed sgRNA (single guide RNA) (Jinek *et al.*, 2012). The first 20 nucleotides of the 5' end of sgRNA are called the target sequence, and manual modification of these 20 bases enables Cas9 to target different, specific DNA sequences. After that, CRISPR/Cas9 developed rapidly in the next several years. After manual modification of the target sequence, the CRISPR/Cas9 system

has been readily applied to gene editing of Arabidopsis (Fauser *et al.*, 2014; Jiang *et al.*, 2014), tobacco (Xia *et al.*, 2014), rice (Zhou *et al.*, 2014; Gu *et al.*, 2013), wheat (Upadhyay *et al.*, 2013), and maize (Svitashev *et al.*, 2015).

Zhu's laboratory used the CRISPR/Cas9 system to achieve the mutation of specific genes in Arabidopsis and rice (Mao *et al.*, 2008; Feng *et al.*, 2013). Seven genes of *Arabidopsis thaliana* were edited, and the mutation rates of T_1, T_2 and T_3 generation plants were 71.2%, 58.3%, and 79.4%, respectively. All homozygous mutations could be transmitted to the next generation stably. Ma *et al.* (2015) optimized the *Cas9* gene, so that it has a higher efficiency in plants. In addition, the Golden Gate Cloning method was used to construct an efficient CRISPR/Cas9 vector system for multi-gene editing in monocotyledonous plants and dicotyledonous plants. Forty-six target sites were edited in rice, with an average mutation rate of 85.4%: most of the mutations were biallelic or homozygous mutations (Ma *et al.*, 2015).

At present, it is common to use the CRISPR/Cas9 system as a method to knock-out genes in plants, but precise editing of genes is more difficult to achieve. Li *et al.* (2016) targeted the rice *OsEPSPS* gene by supplying a repair DNA template and acquired a precisely edited *OsEPSPS* gene that conferred herbicide resistance. Zong *et al.* (2017) fused the nCas9 (Cas9-D10A nickase (nCas9)) with cytosine deaminase to cause single-strand DNA breakage at the target site and replace cytosine with thymine. Li *et al.* (2017) fused cytosine deaminase and uracil glycosylase inhibitors with nCas9 to achieve the same effect. Li *et al.* (2013) designed a target point and provided an exogenous double-stranded DNA template to accurately edit the *NbPDS* gene of tobacco by the HDR pathway. Although there are some reports about the precise editing of plants with the use of CRISPR/Cas9, the actual operation efficiency is low and relevant technologies still need to be improved.

1.2 Research Objectives and Significance

During wheat grain maturation, *α*-amylases may accumulate inappropriately due to pre-harvest sprouting and/or induced LMA. This causes significant economic loss to growers and seriously affects the quality of wheat flour for processing.

This study aimed to obtain amylase-deficient plants by generating *Amy3.1* and *Amy3.2* mutants in two ways. Firstly, the CRISPR/Cas9 gene editing technology was used to knock-

out the two genes. The CRISPR/Cas9 vector utilizes a promoter from rice, so these studies can also help us to understand the gene editing characteristics and capabilities of this vector in wheat. Secondly, α-amylase gene mutants were obtained by screening an EMS-mutagenized ‘Kronos’ wheat population.

The mutant materials obtained in this study will provide an invaluable resource for in depth analysis of the effect of the *Amy3* gene on falling numbers. Gene editing data obtained by high-throughput sequencing in this study will assist in forming the theoretical foundation for the application of CRISPR/Cas9 technology in wheat.

Chapter 2 Materials and Methods

2.1 Experimental Materials

2.1.1 Plant Materials

Plant materials used in this research includes plants for wheat transformation, the 'Kronos' mutant population and transgenic plants. All the plant materials were grown in a greenhouse at the University of Idaho, Moscow, Idaho, USA. The growth condition was 16 h photoperiod with a daytime temperature range of 22-25 ℃ and a nighttime temperature range of 15-20 ℃.

The plant materials for wheat transformation are the common wheat (*Triticum aestivum* L., 2n=6x=42, AABBDD) variety 'CB037' and the tetraploid wheat variety 'Kronos' (2n=6x=28, AABB). They were grown in 16 cm pots, with 4 plants each pot. Caryopses were harvested 14 -16 days after pollination. The immature embryos, 1-2 mm in size, were taken as explants for biolistic bombardment in order to produce transgenic plants.

The 'Kronos' mutant population that was used in this experiment was established by the University of California, Davis. One thousand four hundred and ninety mutant lines of this population were grown in 6 cm cones with one plant per cone.

Transgenic plants were grown in 16 cm pots, with 3 plants per pot.

2.1.2 CRISPR/Cas9 Vectors

The sgRNA vectors pYLgRNA-*OsU3*, pYLgRNA-*OsU6a* (Figure 2.1a) and Cas9 vectors pYLCRISPR/Cas9-MB (Figure 2.1b) were provided by Dr. Yaoguang Liu at South China Agricultural University (Ma *et al.*, 2015) .

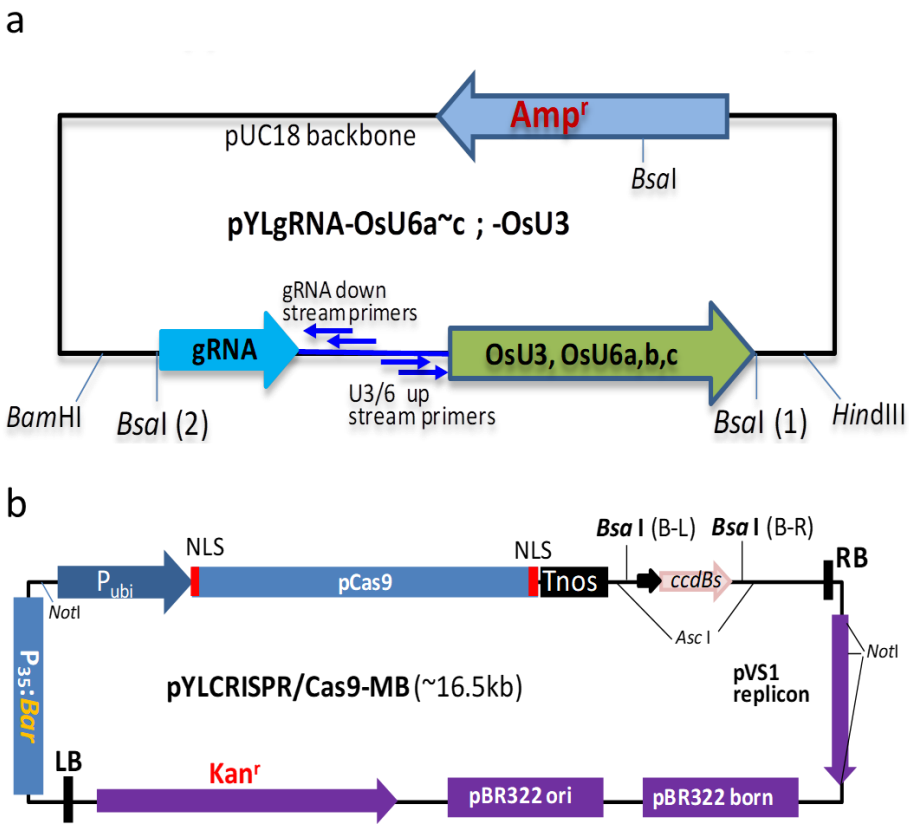


Figure 2. 1 Structures of the sgRNA vectors and a Cas9 vector

Note: *OsU3*, *OsU6a*, *OsU6b*, and *OsU6c* are small nuclear RNA promoters from rice (Ma *et al.*, 2015).

2.2 Experimental Methods

2.2.1 Gene Sequence Acquisition and Analysis

The latest wheat genome sequences and annotation (Borrill *et al.*, 2018) were downloaded from the database (https://urgi.versailles.inra.fr/download/iwgsc/IWGSC_RefSeq_Annotations/v1.1/). Predicted proteins from the IWGSC RefSeq v1.1 annotation were scanned using HMMER v3.2.1 (http://eddylab.org/software/hmmer/hmmer-3.2.1.tar.gz) with the Hidden Markov Model (HMM) model (Eddy, 1998). The profile of α-amylase obtained was from Pfam (PF00128, http://pfam.xfam.org) and was searched against the database, with a cut-off E-value 0.01 in order to filter out low confidence proteins. The cDNA sequences of these proteins were aligned using MEGA X (Kumar *et al.*, 2018), and a phylogenetic tree was constructed using the neighbor-joining method.

2.2.2 Target Selection and Primer Design

There are about 36 homologous genes of the *α*-amylase gene family in common wheat, and the third subfamily, *Amy3*, contains three groups, *Amy3.1*, *Amy3.2*, and *Amy3.3* (unlike *Amy3.1* and *Amy3.2*, this gene is located on the fourth chromosome). The target sites were designed in specific regions of *Amy3.1* or *Amy3.2*, and in the common region of all three orthologs in the A, B, and D genomes. Two target sites were chosen for each gene. According to the principle of ($N_{19/20}$)NGG (NGG: protospacer adjacent motif, PAM), nineteen or twenty nucleotides were selected as target sequences (Ma *et al.*, 2015). The CRISPRdirect website (http://crispr.dbcls.jp/) was used to retrieve potential target sites. The GC content of the target regions was more than 40% (Ma *et al.*, 2015). After concatenating the target site ($N_{19/20}$) to 5 'end of an 83-bp sgRNA (GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTT), the secondary structure of the target-sgRNA sequence was predicted by a RNA Folding Form (http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form2.3). Only target-sgRNAs that produced less than 7 bp perfect hairpins at the target site were selected in order to optimize their binding to a specific site in the *α*-amylase gene (Ma *et al.*, 2015). Finally, the target sequences were compared to the wheat genome database to exclude non-specific off-target sites.

2.2.3 CRISPR/Cas9 Vector Construction

To construct the CRISPR/Cas9 vector, we followed an established protocol (Ma *et al.*, 2015). Five CRISPR/Cas9 vectors were constructed, each vector contained two target sites.

(1) Target adaptor preparation: For each target, two target adaptor oligonucleotides (Table 2.1) were mixed resulting in 1 μM solution for each oligonucleotide, treated at 95 ℃ for 30 s, and annealed at room temperature to form a target double helix.

Table 2. 1 Target sites and the oligonucleotide sequences

Target gene	Target name	Target sequence（5'→3'）	Target adaptor	Target adaptor sequence （5'→3'）
Amy3.1	**C1**	GCTACTCGGCCGACGTCGCCAGG[1]	U6a-C1-F	GCCGCTACTCGGCCGACGTCGCC
			U6a -C1-R	AAACGGCGACGTCGGCCGAGTAG
	C2	CAAGGGCATCCTGCAGTCGGCGG	U3-C2-F	GGCACAAGGGCATCCTGCAGTCGG
			U3-C2-R	AAACCCGACTGCAGGATGCCCTTG
	C3	ATCCTGCAGTCGGCGGTGCAGGG	U3-C3-F	GGCATCCTGCAGTCGGCGGTGCA
			U3-C3-R	AAACTGCACCGCCGACTGCAGGA
Amy3.2	**C4**	TCTCGCCTCCAGCTTAGCACAGG	U3-C4-F	GGCATCTCGCCTCCAGCTTAGCAC
			U3-C4-R	AAACGTGCTAAGCTGGAGGCGAGA
	C5	ACCGACAGCTGCTCGCGAACTGG	U6a-C5-F	GCCGACCGACAGCTGCTCGCGAAC
			U6a -C5-R	AAACGTTCGCGAGCAGCTGTCGGT
	C6	CCGGCCACGGCGTTTGACTTCCC	U6a-C6-F	GCCGGGAAGTCAAACGCCGTGGC
			U6a -C6-R	AAACGCCACGGCGTTTGACTTCC

1: PAM (protospacer adjacent motif) sequence

(2) A restriction-ligation reaction: Target adaptors were then incorporated into the sgRNA vectors pYLgRNA-*OsU3*/U6a in a restriction-ligation reaction (Figure 2.2a). The reaction system was as follows:

Component	**Amount**
10×Cutsmart Buffer (NEB)	1 μL
pYLgRNA-OsU#	20 ng
Target Adaptor Pair	0.05 μM
*Bsa*I (NEB)	5 U
10×T_4 Ligation Buffer (NEB)	0.3 μL
T_4 DNA-ligase (NEB)	20 U
H_2O	Add up to 10 μL

The reactions were incubated in a thermo-cycler for five cycles: 37 ℃, 5 min; 25 ℃, 5 min.

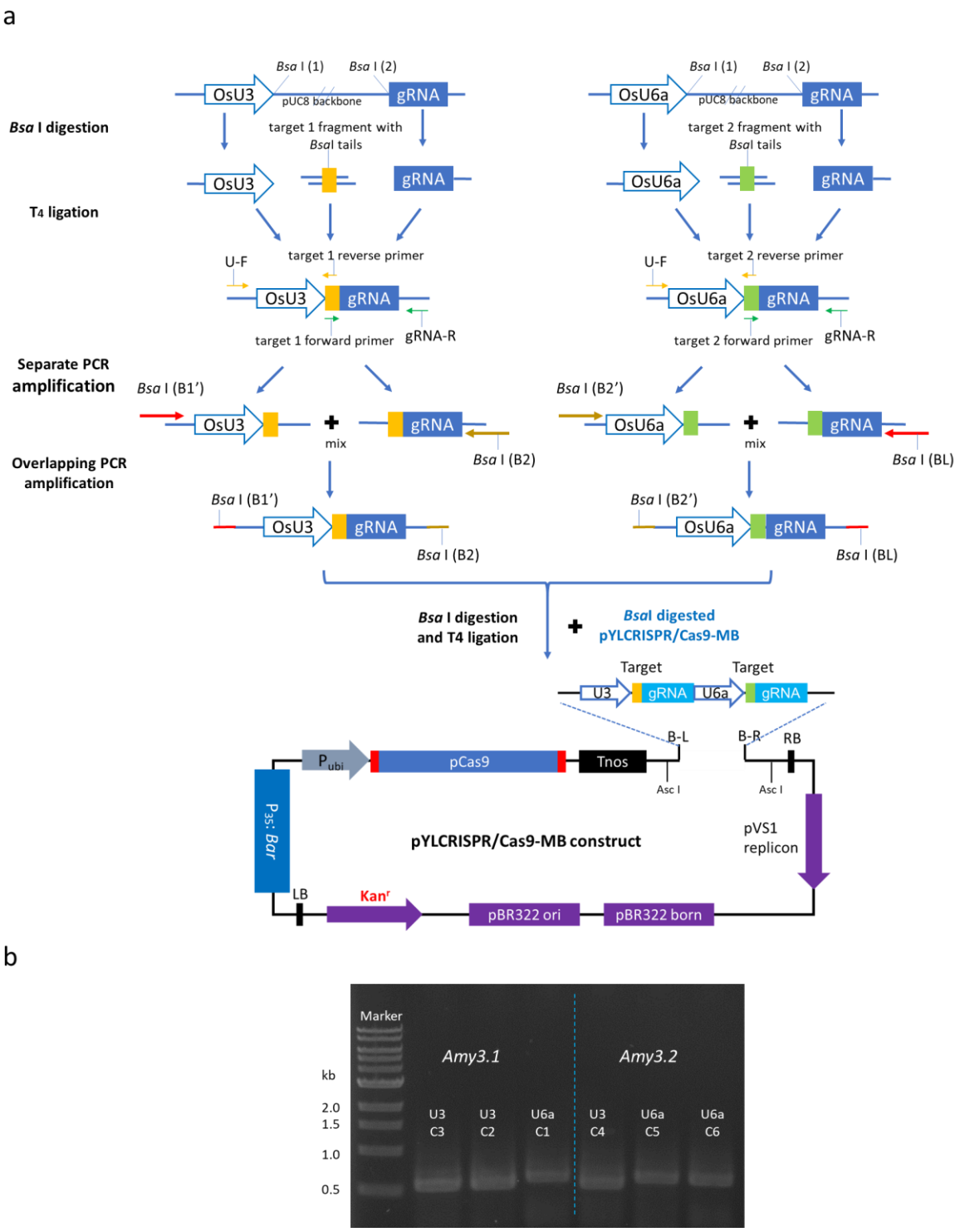


Figure 2. 2 Construction of CRISPR/Cas9 vector

(a) Steps of the CRISPR/Cas9 plasmid construction. pYLgRNA-OsU# vecters were first cut by *Bsa*I and then ligated with a designed target fragment. OsU3 (or OsU6a)-Target and Target-sgRNA were each amplified using specific primers from the ligated product. The resulting fragments were mixed as templates for overlapping PCR amplification. Different *Bsa*I tails were added on both ends of the PCR product. Finaly, the products were ligated to the *Bsa*I digested vector of

pYLCRISPR/Cas9-MB. (b) Electrophoresis of the Promoter-Target-sgRNA fragments. Target C2, C3, and C4 were under the *OsU3* promoter, about 564 bp; target C1, C5 and C6 were under the *OsU6a* promoter, about 629 bp.

(3) Two rounds of PCR were needed for each target. U-F primer and the target reverse primer were used in the first-round PCR; gRNA-R and the target forward primer were used in the second-round PCR (Tables 2.1 and 2.2; Figure 2.2a). The reaction system was as follows:

Component	Amount
2×KOD Fx Buffer	7.5 μL
KOD Fx	0.3 μL
2 mM dNTPs	3 μL
PCR Product from the Last Step	1 μL
U-F (gRNA-R) (10 μM)	0.5 μL
Target Reverse Primer (Target Forward Primer) (10 μM)	0.5 μL
H_2O	Add up to 15 μL

The reaction was repeated 28 cycles (95 ℃, 10 s; 60 ℃, 15 s; 68 ℃, 20 s), continued with a final extension at 68 ℃ for 2 min, and kept at 15 ℃.

Table 2. 2 Primer sequences for CRISPR/Cas9 vector construction

Primer name	Primer sequence (5'→3')
U-F	CTCCGTTTTACCTGTGGAATCG
gRNA-R	CGGAGGAAAATTCCATCCAC
B1'	TTCAGAGGTCTCTCTCGACTAGTGGAATCGGCAGCAAAGG
B2	AGCGTGGGTCTCGTCAGGGTCCATCCACTCCAAGCTC
B2'	TTCAGAGGTCTCTCTGACACTGGAATCGGCAGCAAAGG
BL	AGCGTGGGTCTCGACCGACGCGTCCATCCACTCCAAGCTC
SP-ML	GCGCGGTGTCATCTATGTTACT
SP-R	CCCGACATAGATGCAATAACTTC

The products of two PCR were diluted ten-fold, mixed and were used as the template for next PCR. Primer B1 'and B2 were used for the first target on each vector, and B2' and BL were used for the second target (Table 2.2; Figure 2.2a). The reaction system is as follows:

Component	Amount
2×KOD Fx Buffer	25 μL
KOD Fx	1 μL
2 mM dNTPs	2 μL
Template	2 μL
Forward Primer (10 μM)	1.5 μL
Reverse Primer (10 μM)	1.5 μL
H_2O	Add up to 50 μL

The reaction was repeated 28 cycles (95 ℃, 10 s; 60 ℃, 15 s; 68 ℃, 20 s), continued with a final extension at 68 ℃ for 2 min, and kept at 15 ℃.

Five microliters of PCR products were used for gel electrophoresis. The amplified product was either 564 bp (pYLgRNA-*OsU3*) or 629 bp (pYLgRNA-*OsU6a*) (Figure 2.2b). Each product was then isolated from the gel and purified using DNA Clean & Concentrator-25 (ZYMO Research, Irvine, CA, USA).

The purified DNA fragments were ligated into the pYLCRISPR/Cas9-MB vector (Figure 2.2a). The reaction system was as follows:

Component	Amount
10×Cutsmart Buffer (NEB)	1.5 μL
*Bsa*I (NEB)	10 U
20 ng Target 1	1 μL
20 ng Target 2	1 μL
T_4 DNA ligase (NEB)	35 U
10× T_4 Ligation Buffer (NEB)	0.5 μL
pYLCRISPR/Cas9-MB	60 -80 ng
H_2O	Add up to 15 μL

The reaction process was (37 ℃, 2 min; 10 ℃, 3 min; 20 ℃, 5 min) ×15 cycles, 37 ℃, 2 min; 15 ℃.

The ligation products were transformed into *E. coli* DH5*α* through the heat shock method. In brief, DH5*α* competent *E. coli* cells were thawed on ice. Next, 1.5 μL containing 1 pg-100 ng of plasmid DNA were added to the cell mixture and the tube was flicked 4-5 times to mix the sample. This was then placed ice for 30 minutes. Afterwards, it was put at 42 °C for 30 seconds and then returned to ice for 5 minutes. At the end of that time, 450 μL of room

temperature LB media was put into the mixture and the tube was shaken for 45 min at 37 ℃, 200 rpm min^{-1}. Finally, 150 μL of the mixture was spread onto an LB agar plate with 50 ug L^{-1} kanamycin and incubated overnight at 37 °C.

Positive clones were confirmed by colony PCR. The reaction system was as follows:

Component	Amount
5×PCR Buffer	4 μL
SP-ML (10 μM) (Table 2.3)	1 μL
SP-R (10 μM) (Table 2.3)	1 μL
2 mM dNTPs	0.5 μL
Taq DNA polymerase	0.5 μL
Bacteria	-
H_2O	Add up to 15 μL

The reaction process was: 94 ℃, 5 min, (94 ℃, 30 s; 58 ℃, 30 s; 69 ℃, 80 s) ×30 cycles, 72 ℃, 7 min; 15 ℃. The positive clones had amplified products of 1290 bp. Selected clones were sequenced to confirm both target-sgRNA cassettes in the final CRISPR/Cas9 constructs. Vectors were then prepared from them using a ZymoPURE Plasmid Mixiprep (ZYMO Research, Irvine, CA, USA). After obtaining the plasmid, adjust plasmid concentration to 2,000 ng μL^{-1}.

2.2.4 Wheat Genetic Transformation

Tissue culture and biolistic bombardment were performed as previous studies (Weeks *et al.*, 2016; Lv *et al.*, 2014). The recipient wheat genotypes were the common wheat variety ‘CB037’ and the tetraploid wheat variety ‘Kronos’. Caryopses were harvested 14-16 days after pollination. The experimental steps were as follows:

The glume was peeled off from seeds and put into a beaker. These were disinfected with 75mL disinfectant I (75% ethanol, 0.5‰ tween 20), and shaken at 200 rpm for 5 min. The liquid was poured off and the seeds were treated with 75mL disinfectant II (20% sodium hypochlorite, 0.5‰ tween 20), and shaken as before for 20 min. The disinfectant was removed, and the seeds were washed 3-4 times with sterilized water. The immature embryos were then removed under a stereoscopic microscope on a clean bench and put into

an osmotic medium for 4 h before being placed in a circle with a diameter of 1.5 cm. Embryos were bombarded 4 h later.

After bombardment, immature embryos were stored at 23 ℃, in the dark for 12-16 h, and then transferred to recovery medium (without bialaphos selection). After 2 weeks, calli were transferred to differentiation medium and given 16 h light/8 h dark. Callus was transferred to a new differentiation medium every 2 weeks until they differentiated into seedlings. Seedlings were then moved to rooting medium in a culture flask and incubated at 23 ℃, with 16 h light. After 2-3 weeks, when the length of roots is 6-10 cm, seedlings were transferred to soil.

Transgenic plants were tested in the seedling stage. Glufosinate ammonium (3 mg L^{-1}) was applied on the leaf surface with a cotton swab and observed. One week after the treatment, the smear site of positive transgenic seedlings did not change while the negative seedlings became yellow and withered.

The Mediums Used for Wheat Transformation

Recovery medium: Murashige & Skoog Salt Mixture 4.3 g L^{-1}, maltose 40 g L^{-1}, VB1 0.5 mg L^{-1}, aspartic acid 0.15 g L^{-1}, phytagel 2.5 g L^{-1}. pH was adjusted to 5.8 and autoclaved at 121 °C for 20 min. After cooling to 60 ℃, 2, 4-D (2 mg L^{-1}) and $CuSO_4$ (0.78 mg L^{-1}) was added.

Osmotic medium: Murashige & Skoog Salt Mixture 4.3 g L^{-1}, maltose 40 g L^{-1}, VB1 0.5 mg L^{-1}, aspartic acid 0.15 g L^{-1}, sucrose 171.15 g L^{-1}, phytagel 3.5g L^{-1}.

pH was adjusted to 5.8 and autoclaved at 121 °C for 20 min. After cooling to 60 ℃, 2, 4-D (2 mg L^{-1}) and $CuSO_4$ (0.78 mg L^{-1}) was added.

Differentiation medium: Murashige & Skoog Salt Mixture 4.3 g L^{-1}, maltose 40 g L^{-1}, VB1 0.5 mg L^{-1}, aspartic acid 0.15 g L^{-1}, phytagel 2.5 g L^{-1}.

pH was adjusted to 5.8 and autoclaved at 121 °C for 20 min. After cooling to 60 ℃, 2, 4-D (2 mg L^{-1}), $CuSO_4$ (0.78 mg L^{-1}), Bialaphos (3 mg L^{-1}), and 6-BA (100 mg L^{-1}). was added.

Rooting medium: Murashige & Skoog Salt Mixture 2.15 g L^{-1}, maltose 20 g L^{-1}, VB1 0.25 mg L^{-1}, aspartic acid 0.075 g L^{-1}, phytagel 3.5 g L^{-1}. pH was adjusted to 5.8 and autoclaved

at 121 °C for 20 min. After cooling to 60 ℃, 2, 4-D (2 mg L^{-1}), $CuSO_4$ (0.78 mg L^{-1}), and Bialaphos (3 mg L^{-1}) was added.

2.2.5 Samples Preparation for High-throughput Sequencing

After obtaining positive transgenic seedlings as determined by a herbicide test, high-throughput sequencing was used to identify specific nucleotide variation among transgenic plants. The amplified products required for sequencing were obtained through three rounds of PCR. Normally, researchers can obtain the amplified products required for sequencing through two rounds of PCR. However, it was difficult to design gene-specific primers around target sites in *Amy3* genes, and the sequencing length was limited to 300 bp. So, one PCR process was added in this study. In the first-round PCR, the *Amy3.1* and *Amy3.2* genes were amplified using gene-specific primers. The product of the first-round PCR was diluted and used as templates for the second-round PCR. In the second-round PCR, common primers were used, and fluidigm primers were added at two ends of the amplicon (IBEST Genomics Resources Core, University of Idaho, Moscow, Idaho). The products of the second-round PCR were then diluted and used as the template for the third-round PCR amplification. In the third-round PCR, different barcodes were added to each sample (IBEST Genomics Resources Core, University of Idaho, Moscow, Idaho).

The experimental process was as follows:

(1) The primer pairs for *Amy3.1* and *Amy3.2* genes were LM307/LM309 and LM312/LM317 (Table 2.3) respectively. The reaction system was as follows:

Component	Amount
5×PCR Buffer	4 μL
Forward Primer (10 μM)	0.5 μL
Reverse Primer (10 μM)	0.5 μL
10 mM dNTPs	0.5 μL
Taq DNA Polymerase	0.5 μL
DNA Template	150 ng
H_2O	Add up to 20 μL

The reaction process was: 94 ℃, 5 min, (94 ℃, 30 s; 60 ℃, 30 s; 72 ℃, 90 s) ×26 cycles, 72 ℃, 7 min; 15 ℃.

Table 2. 3 Primer sequences for high-throughput sequencing

Primer name	Primer sequence (5'→3')
LM307	GCGCGGTGTCATCTATGTTACT
LM309	CCCGACATAGATGCAATAACTTC
LM312	AACTATGAGCAGAGTGCTTGTCTA
LM317	GGATGGATCCCGTTCCTTGA
LM323	ACACTGACGACATGGTTCTACAATGATCTGCAGCGACGACAC
LM324	ACACTGACGACATGGTTCTACAGTACCTATAAATACAAGCGAGCTC
LM325	TACGGTAGCAGAGACTTGGTCTAGGATGTAGGCGTAGCCCAT
LM326	TACGGTAGCAGAGACTTGGTCTTGACGGTCTTCTCSGGCAT

(2) The products of the first-round PCR were diluted 370-fold and used as templates for the second-round PCR. The primer pairs for *Amy3.1* and *Amy3.2* genes were LM323/LM325 and LM324/LM326 (Table 2.3) respectively. The reaction system was as follows:

Component	Amount
5×PCR Buffer	4 μL
Forward Primer (10 μM)	0.5 μL
Reverse Primer (10 μM)	0.5 μL
10 mM dNTPs	0.5 μL
DNA Taq Polymerase	0.5 μL
DNA Template	1 μL
H_2O	Add up to 20 μL

The reaction process was: 94 ℃, 5 min, (94 ℃, 30 s; 60 ℃, 30 s; 72 ℃, 60 s) ×22 cycles, 72 ℃, 7 min; 15 ℃.

(3) The products of the second-round PCR were diluted 15-fold and used as templates for the third-round PCR. Barcode provided by IBEST Genomics Resources Core was used as the primer in the third-round PCR reaction. Each sample had a different barcode pair, and the reaction system was as follows:

Component	Volume
5×PCR Buffer	4 μL
Barcode Primer Pair (2 μM)	0.5 μL
10 mM dNTPs	0.5 μL
DNA Taq polymerase	0.5 μL
DNA Template	1 μL
H_2O	Add up to 20 μL

The reaction process was: 94 ℃, 5 min, (94 ℃, 30 s; 60 ℃, 30 s; 72 ℃, 60 s) ×15 cycles, 72 ℃, 7 min; 15 ℃. After the reaction, 5 μL product was pooled for sequencing.

The sequencing was done by IBEST Genomics Resources Core. The DNA pools were first cleaned with magnetic beads to remove primer-dimer and then were run on fragment analyzer to verify quality. Then qPCR was performed to determine quantity of sequencable library. The library was finally sequenced on 2x300 Miseq run.

2.2.6 Analysis of High-throughput Sequencing Results

The raw data of high throughput sequencing were first filtered by quality control to remove the data with poor sequencing quality. Then the data was divided into different plants according to the Barcode sequence. *Amy3.1* and *Amy3.2* were distinguished by primer sequence, and different genomes were identified by SNP on A, B and D genomes. After the sequencing data were differentiated, Geneious (11.1.5) (https://www.geneious.com/) software was used to analyze the sequences. Read 1 and Read 2 from the same plant, same gene, and same genome were imported into Geneious (11.1.5) software by ‘Paired’ method. The 'Map to Reference' option in 'Align/Assemble' was used to map the imported sequence to the corresponding gene, and the editing type of the gene was determined according to the change of the sequence in the target position.

2.2.7 DNA Extraction for Plant Materials

2.2.7.1 DNA Pools of 'Kronos' Mutants

1,440 'Kronos' mutant lines were planted in the greenhouse to screen for the *Amy3* gene mutants. Every 16 plants were arranged in a 4x4 matrix. Every four plants of one row were collected as one row (R) pool; every four plants of one column were collected as one column (C) pool. In total, 360 R pools and 360 C pools were prepared.

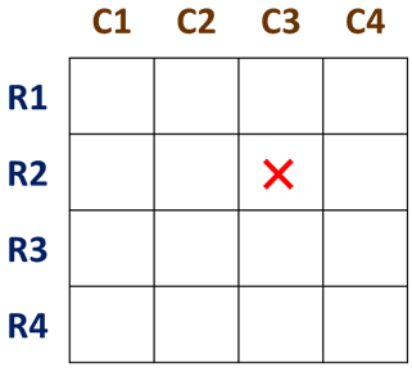


Figure 2. 3 Construction of 'Kronos' DNA pools

Each square in the figure represents one 'Kronos' mutant line, and every 16 mutant lines were arranged into a 4x4 matrix, the DNA mixture of 4 plants in the first row was labeled as R1, the DNA mixture of 4 plants in the first row was labeled as C1. If mutations could be detected in the DNA pools of R2 and C3, then the specific mutant plants could be identified as the intersection (x) between the row and column pools.

2.2.7.2 DNA Extraction Step

(1) For each plant, 4 cm fresh leaf tissue was collected (for DNA pools, leaf tissue was collected from 4 plants, 1 cm per plant, and was put into one tube), put in 2 mL tube and stored in -80 ℃. Then the tubes that contained leaf tissue were frozen in liquid nitrogen and grinded to a fine powder.

(2) 750 μL DNA extraction buffer and 5 μL RNase (10 mg mL^{-1}) were quickly added into the leaf powder. The mixtures were vigorously vortexed, and then were put in 37 ℃ water bath for 30 min. After that, 750 μL phenol: chloroform: isoamyl alcohol (25:24:1) was added into the mixtures and shaken well. The mixtures were centrifuged for 10 min at 1,2000 rpm. The liquid in the test tubes would separate into three layers.

(3) The upper phase (about 650 μL) of liquid was transferred to a new 2 mL tube. 650 μL Chloroform were added into tubes that contained the upper phase. These mixtures were

mixed gently and centrifuged for 10 min at 1,2000 rpm. 450 μL supernatant was transferred to a new 1.5 mL tube, then 45 μL 3 M sodium acetate (pH = 4.8) and 450 μL isopropanol were added into the liquid.

(4) The mixtures were mixed gently. DNA would precipitate after this step. The tubes that contained DNA and isopropanol were stored at -20 ºC for 30 min. After that, the tubes were centrifuged for 10 min at 1,2000 rpm. The supernatant was poured off and 600 μL of 70% ethanol was added into the tubes. The tubes were centrifuged for 6 min at 1,2000 rpm and the supernatant was poured off. Then the tubes were put in clean bench for 30 min to make the DNA dry. 200 μL of deionized H_2O was added to dissolve DNA. The DNA was stored at -20 ºC.

DNA Extraction Solution:

DNA extraction buffer: 100 mM Tris-HCl (pH=8.5), 100 mM NaCl, 50 mM EDTA (pH = 8.0), 2% SDS。

Phenol: Chloroform: Isoamyl alcohol: 250 ml redistilled phenol, 240 ml chloroform and 10 ml Isoamyl alcohol were mixed together. The solution could be stored at 4 °C for 1 month without light.

2.2.8 *Amy3* Gene Primers Design

Software MEGA X (Kumar *et al.*, 2018) was used for sequence alignment of all *α*-amylase genes in wheat, and specific primers of *Amy-A3.1*, *Amy-B3.1*, *Amy-A3.2*, and *Amy-B3.2* gene were designed.

Three bases at the 3' end of primers were specific for each gene. ‘Oligo Calc’ (http://biotools.nubic.northwestern.edu/OligoCalc.html)’ was used to evaluate primers. The length each primer was 18-30 bp, and GC content was between 38%-60%. The primers that were capable of forming dimers or hairpin structures were avoided. Primers were tested by gradient PCR. Six annealing temperatures were tested, from 50 °C to 68 °C. After electrophoresis, the appropriate annealing temperature and primer combinations were selected.

2.2.9 Screening 'Kronos' Mutant Library Using TILLING Technology

2.2.9.1 CelI Enzyme Extraction

CelI enzyme extraction was performed according to Yang *et al.*'s (2000) method. 500 g fresh celery was rinsed twice with distilled water, and dried. The stems were cut into pieces and crushed with a grinder (about 250 mL juice would be produced). The juice was filtered with miracloth (Millipore, Cat. 475855-1R). The filtrate was collected.

All the following steps were performed on the ice. 1M Tris-HCl and 0.1 M PMSF were added into the filtrate to form buffer A. The final concentration was 0.1 M Tris-HCl, 100 μM PMSF (For example, 449.5 mL celery filtrate was added with 50 mL 1M Tris-HCl and 0.5 mL 0.1M PMSF). The mixture was centrifuge in 4000 rpm for 20 min. After that, the supernatant was filtered with miracloth. For 1 L of liquid, 144 g $(NH_4)_2SO_4$ was added with the final concentration was 25% (wt/vol). Then the liquid was mixed in 4 ℃ for 30 min.

The mixture was centrifuge in 10,000 rpm for 20 min. Then the supernatant was collected. For 1 L of supernatant, 390 g $(NH_4)_2SO_4$ was added. The final concentration was 80% (wt/vol). Again, the liquid was mixed in 4 ℃ for 30 min.

The mixture was centrifuge in 4 ℃, 10,000 rpm for 1.5 h. The supernatant was discarded, and the sediment was collected. (The sediment could be stored in -80 ℃ for at least 2 weeks.) Next, the sediment was dissolved in 25 mL buffer B. The mixture was mix gently and the sediment was fully dissolved in buffer B.

Dialysis bag (Cellulose ester Dialysis membranes (10000 Dalton MWCO)) was pretreated for next step: it was soaked in ultrafiltration water for 15 minutes at room temperature, then rinsed in deionized.

The solution got from last step was poured into the dialysis bag to dialysis. Dialysis was performed at 4 ℃ for about 20 h. 10 L dialysate was used for every 10 mL of enzyme extraction. At first, the dialysate was changed every hour and then was changed every 2-3 hours. Finally, the enzyme solution in the dialysis bag was divided into 2 mL centrifuge tubes and stored in -80 ℃.

Preparation of Reagents

Buffer A: 0.1M Tris-HCl, pH =7.7, 100 μM PMSF

Buffer B: 0.1M Tris-HCl, pH = 7.7, 0.5 M KCl, 100 μM PMSF

Nicking Buffer: 20 mM Hepes, pH = 7.5, 10 mM KCl, 3 mM MgCl2

2.2.9.2 TILLING Procedures

(1) Gene specific primers were used for TILLING. PCR reaction system was as follows:

Component	Volume
5×PCR Buffer	4 μL
Forward Primer (10 μM)	1 μL
Reverse Primer (10 μM)	1 μL
dNTPs (2 mM)	0.5 μL
DNA Taq polymerase	0.5 μL
DNA Template	-
H_2O	Add up to 20 μL

The reaction process was: 94 ℃, 5 min, (94 ℃, 30 s; 60 ℃, 30 s; 72 ℃, 60 s) ×40 cycles, 72 ℃, 7 min; 99 ℃, 10 min; slow annealing at 72 ℃ for 10 s for 90 cycles, the annealing temperature decreased 0.3 ℃ per cycle; 15 ℃ for storage.

(2) Reaction system of CelI digestion was as follows:

Component	Volume
PCR Product	20 μL
10×Nicking Buffer	3 μL
CelI	0.66 μL
H_2O	Add up to 30 μL

The reaction process was: 45 ℃ for 30 min.

2.3 Research Technical Route

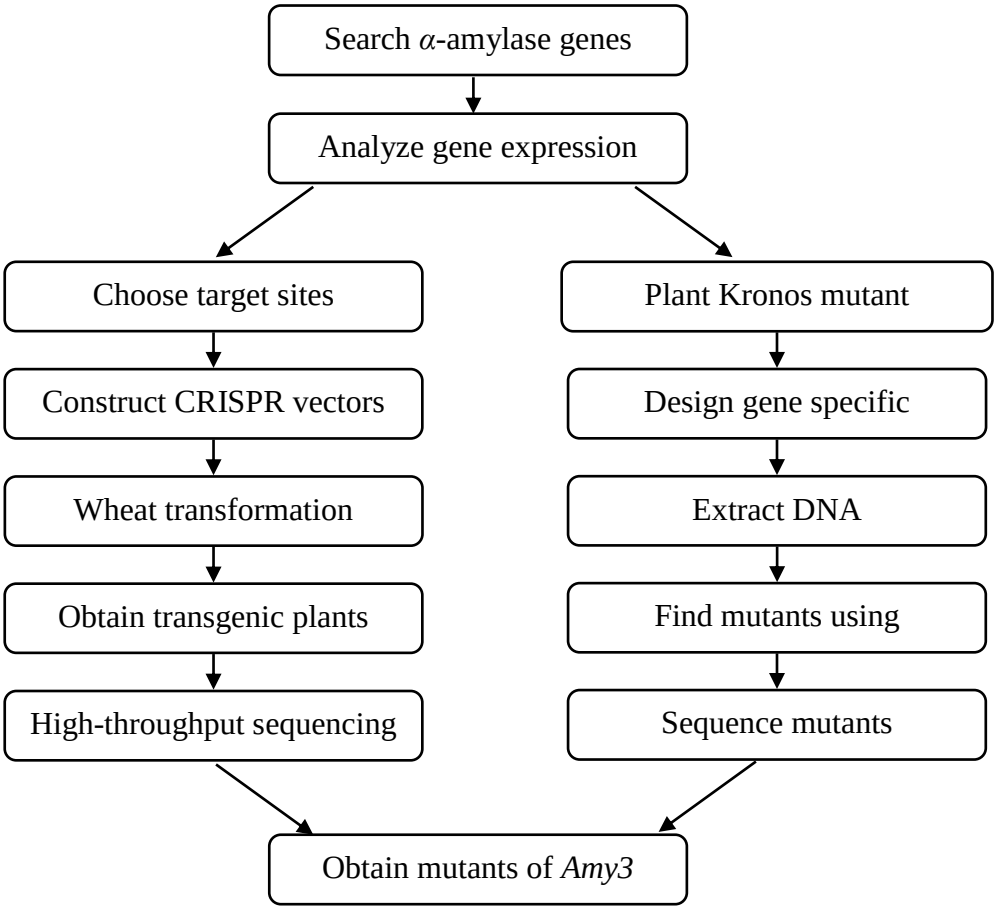


Figure 2. 4 Research technical route

Chapter 3 Gene Editing Results

3.1 *α*-amylase Genes in Wheat Genome

Per the material and methods section, predicted proteins from the IWGSC RefSeq v1.1 annotation were scanned using HMMER v3.2.1 with the Hidden Markov Model (HMM) model. The profile of *α*-amylase obtained was from Pfam and was searched against the database with a cut-off E-value of 0.01 in order to filter out low confidence proteins. After that, a total of 66 proteins were identified. The cDNA sequences of these proteins were aligned using MEGA X (Kumar *et al.*, 2018), and a phylogenetic tree was constructed (Figure 3.1).

Analysis of the phylogenetic tree revealed that 36 gene sequences could be classified as belonging to the *α*-amylase gene family. The wheat *α*-amylase gene family was divided into four subfamilies, *Amy1*, *Amy2*, *Amy3*, and *Amy4*. The first subfamily, *Amy1*, is located on the sixth chromosome, and consists of eleven different copies of the gene. The second subfamily, *Amy2*, is located on the seventh chromosome. It has nine different copies of the gene. The third subfamily, *Amy3*, is located on the fourth and fifth chromosomes. This subfamily contains ten copies of the *α*-amylase gene, and they can be divided into three groups, *Amy3.1*, *Amy3.2*, and *Amy3.3* respectively. *Amy3.1* is located on the fifth chromosome and has one copy of the gene on each of the A, B, and D genomes. *Amy3.2* is also located on the fifth chromosome. It contains one copy of the gene on both the A and D genomes and two copies of the gene on the B genome. *Amy3.3* is located on the fourth chromosome. It contains one copy of the gene on the D genome and two copies of the gene on the B genome. The fourth subfamily, *Amy4*, contains four genes, and it is located on the second and third chromosome.

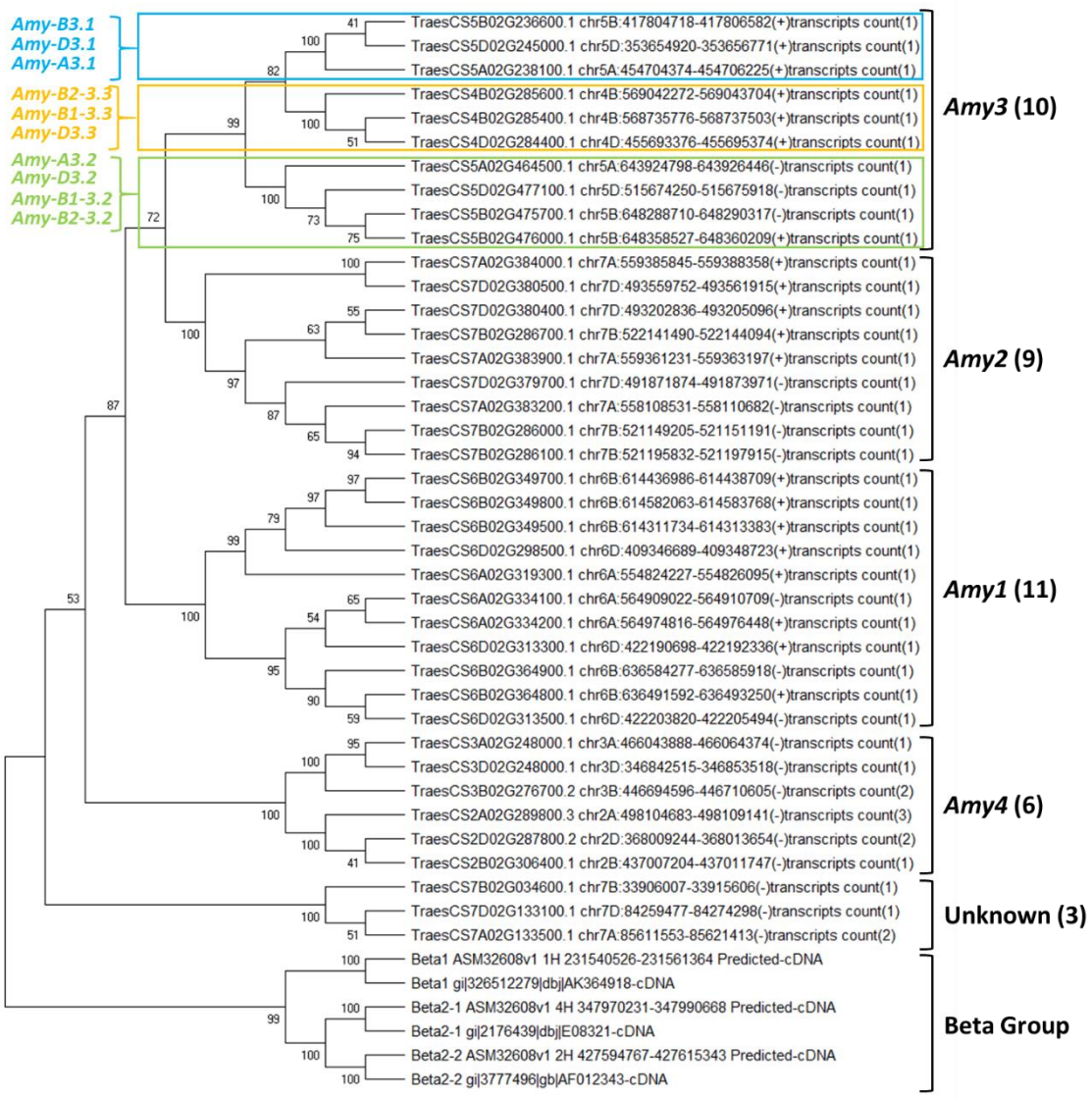


Figure 3. 1 Phylogenetic tree of wheat α-amylase gene family

The *α*-amylase gene family contains four subfamilies. The third one, *Amy3*, contains three groups: *Amy3.1* (blue box), *Amy3.2* (green box), *Amy3.3* (yellow box). The numbers in parentheses are the copy numbers of the gene. The phylogenetic tree was constructed based on the cDNA sequences of *α*-amylase genes using Neighbor-Joining method in MEGA X (Kumar *et al.*, 2018).

3.2 Expression Profiles of *α*-amylase Genes

Different *α*-amylase gene sequences were found in a gene expression assay of the WheatExp database (https://wheat.pw.usda.gov/WheatExp/). The results showed that there were significant differences in expression time and magnitude between different *α*-amylase genes.

Amy1 is expressed throughout the growth stage of wheat development and significantly increases in the spike when 50% of the florets are complete (Z65 in Zadok's scale; Zadoks *et al.*, 1974) (Figure 3.2a). *Amy2* is highly expressed in the spike at the Z65 stage and in the grain at the Z71 stage (kernel (caryopsis) watery ripe) but is rarely expressed in other tissues and other growth stages (Figure 3.2b). *Amy3* is expressed in the developing seeds but is rarely expressed in other growth stages and developing tissues (Figure 3.2c). *Amy4* gene is expressed in different tissues of wheat during the entirety of the growth period (Figure 3.2d).

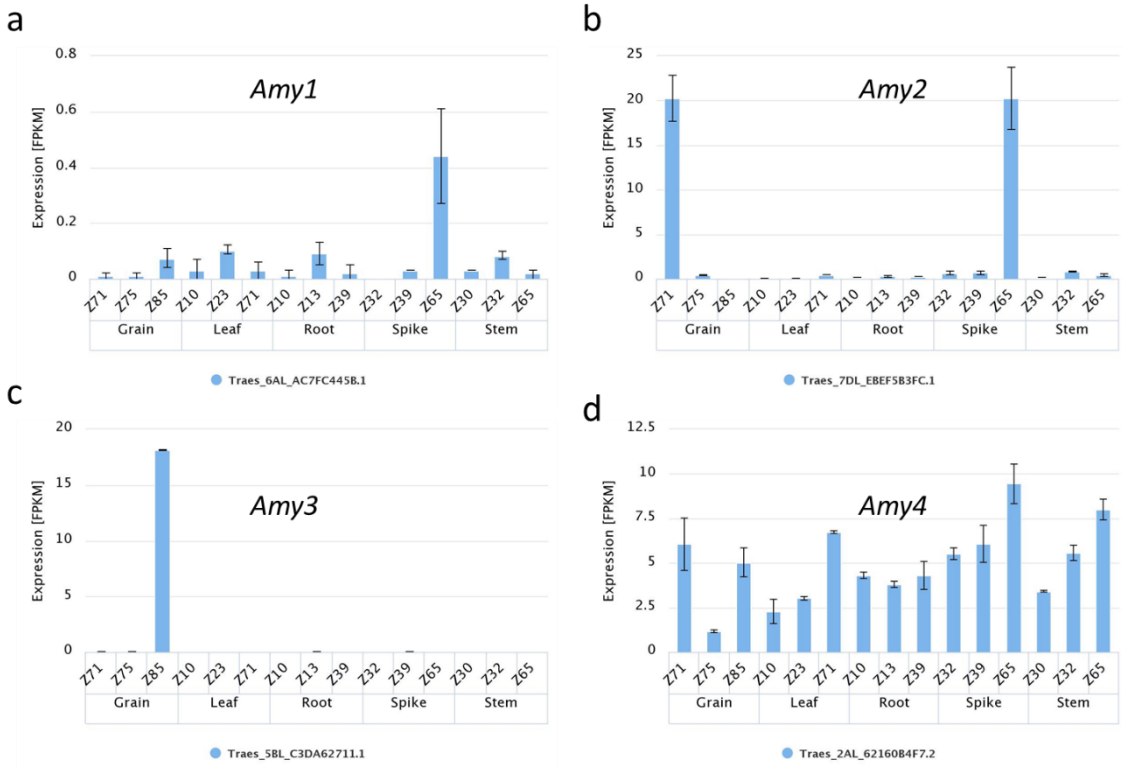


Figure 3. 2 Expression profiles of four subfamilies of *α*-amylase genes

Expression data comes from WheatExp (Pearce *et al.*, 2015). Growth stages present in Zadoks scales (Zadoks *et al.*, 1974). FPKM: Fragments per kilobase of exon per million reads mapped (Pearce *et al.*, 2015).

Next, the expression of three groups of genes in *Amy3* was analyzed. Two gene copies of *Amy3.1* in the B and D genome are expressed at the later stage of grain development (Z85, soft dough) (Figure 3.3a). Four gene copies of *Amy3.2* are expressed at the early stage of seed development (Z71) (Figure 3.3b), and *Amy-D3.2* also has a little expression in leaf tissue. *Amy-D3.3* is expressed during spike development and in the late stage of grain development (Figure 3.3c). According to the expression of those genes, *Amy3.1* and *Amy3.2* were inferred mainly play a role in wheat grains. It was thusly inferred that if the gene

function of *Amy3.1* and *Amy3.2* was destroyed, the expression of α-amylase in wheat grain could be decreased without affecting the normal growth of plants. Therefore, the following research mainly focuses on *Amy3.1* and *Amy3.2*. *Amy-B2-3.2* gene was discovered in a later research process and was not included in the current study.

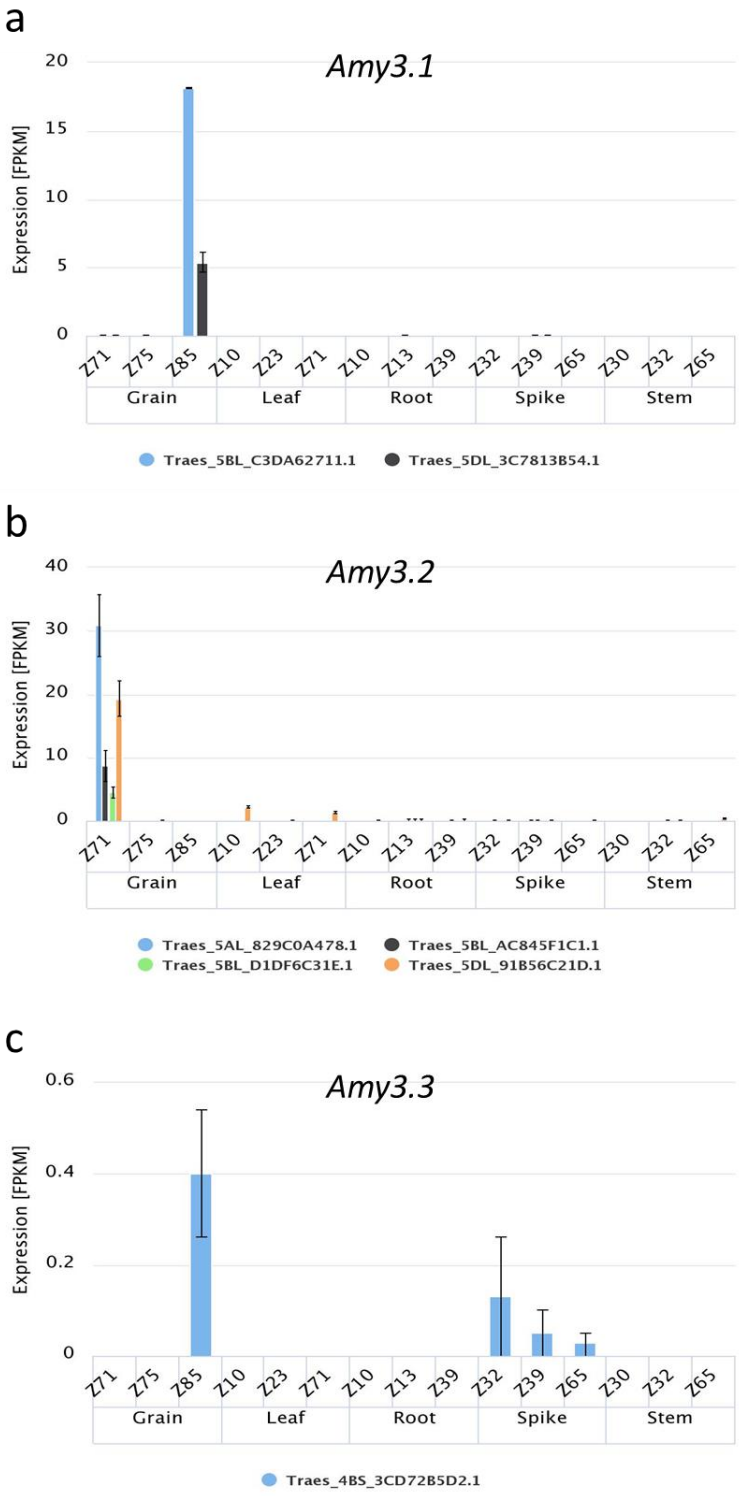


Figure 3. 3 Expression profiles of the *Amy3* genes

Expression data comes from WheatExp (Pearce *et al.*, 2015). Growth stages present in Zadoks scales (Zadoks *et al.*, 1974). FPKM: Fragments per kilobase of exon per million reads mapped (Pearce *et al.*, 2015).

3.3 CRISPR/Cas9 Plasmid Construction

3.3.1 Selection of Target Sites

CRISPR/Cas9 is a new technology developed for genome editing in recent years (Charpentier *et al.*, 2014; Mao *et al.*, 2008). This technique is very simple and convenient to operate. Appropriate target sites need to be selected according to the specific gene sequence, so that the Cas9 endonuclease can be guided to find the target site and cut double strands of DNA.

Gene structures of *Amy3* were analyzed, the results showed that genes in the *Amy3.1* and *Amy3.2* groups contain four and three exons respectively (Figure 3.4a). In this research, six target sites were chosen in the coding region of *Amy3.1* and *Amy3.2*. For *Amy3.1*, three target sites (C1, C2, C3) were chosen on the third exon. For *Amy3.2*, one target site (C1) was chosen on the first exon, and another two target sites (C2, C3) were chosen on the second exon.

There are many homologous genes of α-amylase in wheat, and the sequences in the coding region are highly homologous. When choosing the target sites, the specificity of the target had to be considered. To knockout the *Amy3.1* and *Amy3.2* genes without destroying the function of other α-amylase genes, the target sites were selected in specific regions of the *Amy3.1* and *Amy3.2* but in common regions of all three orthologs in A, B, and D genomes (Figure 3.4b). All six targets chosen in this experiment had high specificity. For C1, C2, and C4, they contain unique protospacer adjacent motif (PAM) that is absent in other homologous genes. If a target has PAM in a non-targeted gene(s), the target is designed to contain three to eleven nucleotides that are different from other α-amylase genes; this could ensure the specificity of each target and reduce off-target gene editing. In addition, the GC content at the target was also considered. The six selected targets were calculated to have the appropriate GC content (50-70%). The GC contents of the six targets C1-C6 are 75%, 65%, 65%, 60%, 65%, and 65%, respectively. Table 2.1 lists the target sequences and corresponding target adaptor sequences.

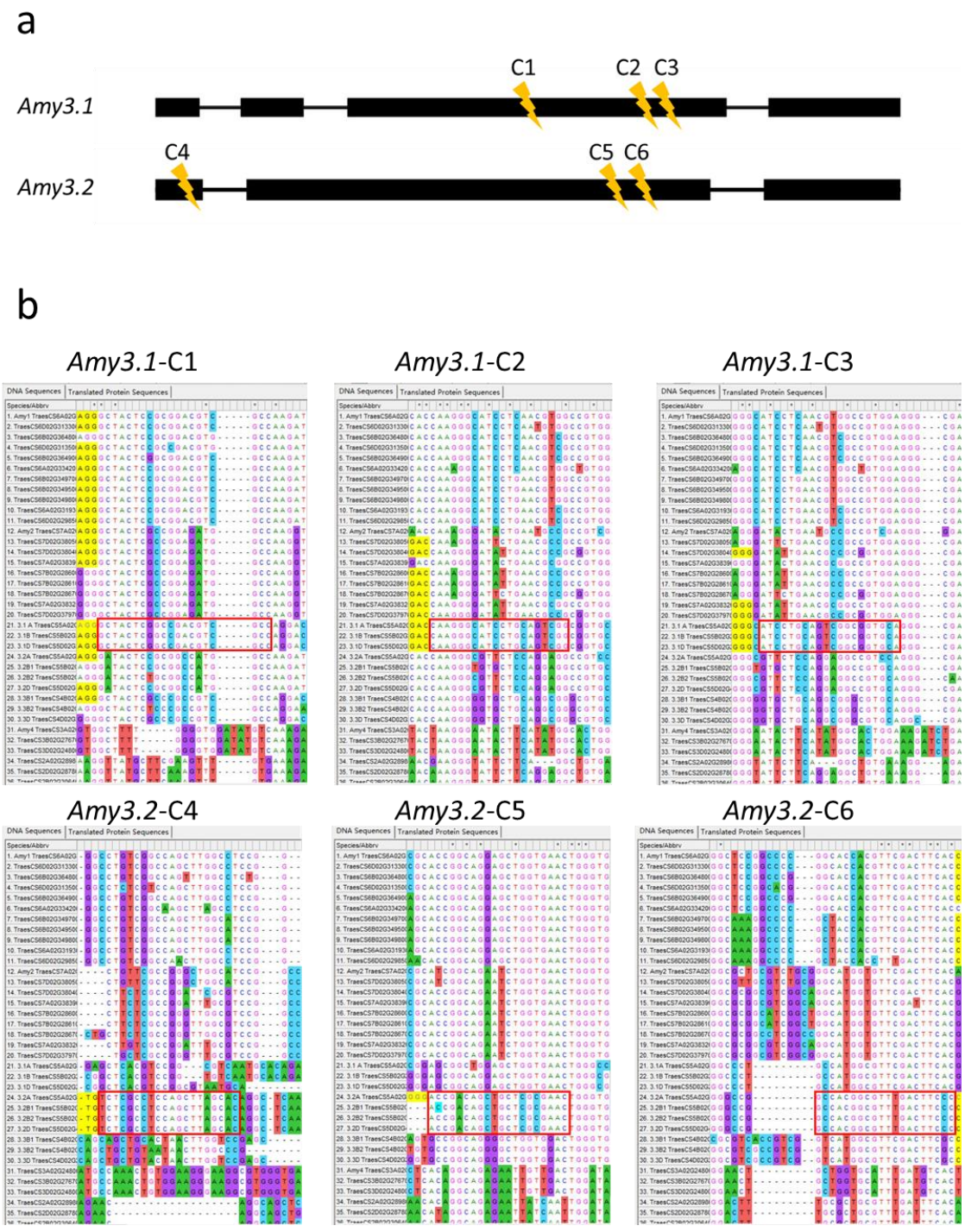


Figure 3. 4 Target sites selected for gene editing

Target sites were chosen in the specific regions of *Amy3.1* and *Amy3.2* (red box). Sequences were aligned using the MEGA X (Kumar *et al.*, 2018).

3.3.2 CRISPR/Cas9 Plasmid Construction

In order to knockout genes in *Amy3.1* and *Amy3.2* group alone, and genes in two groups simultaneously, five CRISPR/Cas9 constructs (A1-A5) were assembled in this research; each contained two target sites (Table 3.1). The first target of each vector was connected to the *OsU3* promoter, and the second target was connected to the *OsU6a* promoter. A1 and A2 vectors were constructed to knock out genes in *Amy3.1* group. The A1 vector contained the C1 and C3 sites, and the A2 vector contained the C1 and C2 sites. The A3 and A4 vectors were constructed to knockout genes in the *Amy3.2* group. A3 contained the C4 and C6 sites, and A4 contained the C4 and C5. The A5 vector contained two targets, of which C2 would target genes in the *Amy3.1* group, and C6 would target genes in the *Amy3.2* group. In this way, the A5 vector was used to knock out genes in two groups simultaneously.

Table 3. 1 CRISPR/Cas9 constructs

Target gene	Vector name	Target
Amy3.1	A1	C3 + C1
	A2	C2 + C1
Amy3.2	A3	C4 + C6
	A4	C4 + C5
Amy3.1 + Amy3.2	A5	C2 + C6

Six targets were first incorporated into the sgRNA vectors in a restriction-ligation reaction (pYLgRNA-*OsU3* for C2, C3, C4; pYLgRNA-*OsU6a* for C1, C5, C6). Then, *OsU3* (or *OsU6a*)-Target-sgRNA fragments were amplified from the sgRNA vector and incorporated tail-specific *Bsa*I sites on both ends of the PCR product (Figure 2.2a). Electrophoresis was used to determine whether the target had been properly connected to the promoter and sgRNA. Targets that connected to the *OsU3* promoter should be 564 bp, and targets that connected to the *OsU6a* promoter should be 629 bp. Results showed that all six fragments were correct in size (Figure 2.2b).

Purified PCR products were cloned into pYLCRISPR/Cas9-MB vectors, that contained two addition expression cassettes: the *Cas9* gene the *Bar* gene (Figure 2.2a). The ligation products were transformed into *E. coli* DH5*α*. Positive clones were confirmed by colony

PCR with the correct length of 1,290 bp (Figure 3.5a). Selected clones were sequenced to confirm both target-sgRNA cassettes in the final CRISPR/Cas9 constructs (Figure 3.5b). The results showed that all six constructions were correct.

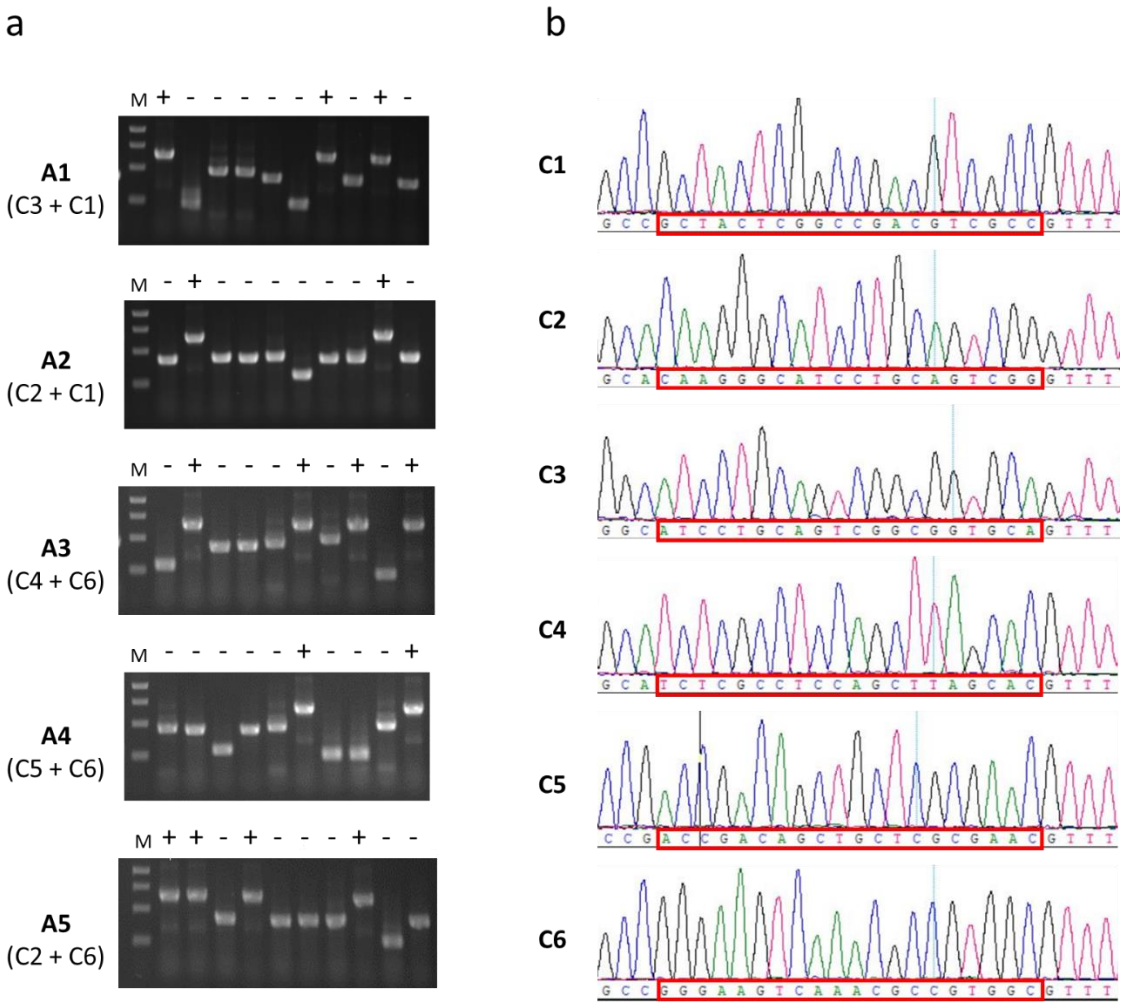


Figure 3. 5 Verification of the finished CRISPR/Cas9 constructs

(a) Positive clones were confirmed by colony PCR with the correct length of 1,290 bp. M, marker (0.5 kb, 1 kb, 1.5 kb, 2.0 kb). (b) Selected clones were sequenced to confirm the target-sgRNA cassettes in the finished CRISPR/Cas9 construct. Target sequences were marked by red box. C6 presents a reverse complementary sequence.

3.4 Wheat Transformation with CRISPR/Cas9 Constructs

Five CRISPR/Cas9 constructs were transformed into wheat through biolistic bombardment. The recipient wheat genotypes are the hexaploid wheat ‘CB037’ and the tetraploid wheat ‘Kronos’. Immature embryos were bombarded, after recovery culture, differentiation culture and rooting culture, regenerated plants were obtained from proliferating callus. In the

seedling stage, 3mg L^{-1} glufosinate ammonium was used to identify the transgenic plants. The smear sites of the positive transgenic seedlings remained healthy while the negative seedlings became yellow and withered.

The transformation material for the A1 plasmid was ‘Kronos’. From a total of 1,260 immature embryos bombarded, 129 positive transgenic plants were obtained, which contained 107 independent transformation events, giving a transformation efficiency of 8% (Table 3.2). Both plant materials were transformed using the A2 plasmid. For ‘Kronos’, 1,080 immature embryos were bombarded, 30 positive transgenic plants were obtained from 29 independent transformation events, at a transformation efficiency of 3%. For ‘CB037’, 2,100 immature embryos were bombarded, and 68 positive transgenic plants were obtained from 55 independent transformation events, at a transformation efficiency of 3%. The A3 plasmid was transformed into both wheat materials also. 1,050 immature embryos of ‘Kronos’ were bombarded to obtain 13 positive transgenic plants from 12 independent transgenic events with at a transformation efficiency of 1%. 1,560 immature embryos of ‘CB037’ were bombarded to obtain 41 positive transgenic plants from 39 independent transformation events at a transformation efficiency of 3%. The A4 plasmid was transformed into ‘Kronos’. 1,350 immature embryos were bombarded, and 117 positive transgenic plants were obtained from 110 independent transformation events, at a transformation efficiency of 8%. The A5 plasmid was transformed into both varieties. 1,470 immature embryos of ‘Kronos’ were bombarded, and 167 positive transgenic plants were obtained from 139 independent transgenic events; the transformation efficiency was 9%. 1,470 immature embryos of ‘CB037’ were bombarded, and 104 positive transgenic plants were obtained from 102 independent transformation events; the transformation efficiency was 7%. Collectively, a total of 10,290 immature embryos were transformed in this research, and 657 positive transgenic plants were obtained, of which 593 were independent transgenic events, and the overall transformation efficiency was 5%.

To knockout the *Amy3.1* gene, the A1 and A2-based transformation produced 227 positive transgenic plants, of which 159 plants were ‘Kronos’, and 68 plants were ‘CB037’. To knockout the *Amy3.2* gene, the A3 and A4-based transformation produced 171 positive transgenic plants, of which 130 plants were ‘Kronos’, and 41 plants were ‘CB037’. To knock-out the *Amy3.1* gene and the *Amy3.2* gene together, the A5-based transformation

produced 271 positive transgenic plants, of which 167 plants were ‘Kronos’ and 104 plants were ‘CB037’.

Table 3. 2 Wheat transformation of CRISPR/Cas9 vectors

Target genes	Construct name	Recipient genotype	Embryos treated	Transgenic plants	Independent transgenic events[1]	Transformation efficiency[2]
Amy3.1	**A1**	Kronos	1,260	129	107	8%
	A2	Kronos	1,080	30	29	3%
		CB037	2,100	68	55	3%
Amy3.2	**A3**	Kronos	1,050	13	12	1%
		CB037	1,560	41	39	3%
	A4	Kronos	1,350	117	110	8%
Amy3.1+ Amy3.2	**A5**	Kronos	1,470	167	139	9%
		CB037	1,470	104	102	7%
Sum			10,290	657	593	5%

[1]: Independent transgenic events refer to the number of regenerated calli. Different calli were independent, and each callus produced one or more transgenic plants. For example, A1-based transformation in ‘Kronos’ produced 129 positive transgenic plants from 107 calli.

[2]: Transformation efficiency = $\frac{\text{Independent transgenic events}}{\text{Embryo number}} \times 100\%$

3.5 Analysis of Gene Edited Plants

3.5.1 Genotype of Mutants

Tetraploid wheat ‘Kronos’ has four sets of chromosomes, each set contains seven chromosomes. Two sets of chromosomes belong to the A genome, and the other two sets of chromosomes belong to the B genome. Each gene of ‘Kronos’ has four copies, and each genome contains two copies, one on each of the sets, which is referred to as a pair of alleles. Accordingly, hexaploid wheat material ‘CB037’ has six sets of chromosomes, each set contains seven chromosomes and each gene has six copies. In this research, the targets were chosen in the common region of a gene in A, B and D genomes. When the CRISPR/Cas9 complex begins to function in the wheat nucleus, it may cut one or more copies of the gene on different genomes. For a pair of alleles, the mutation could be homozygous, heterozygous and biallelic. A homozygous mutation means that two alleles of a gene have the same mutation. A heterozygous mutation refers to the mutation of one allele of a gene while the other allele is still the wild-type. A biallelic mutation is a mutation that two alleles of a gene

have different mutations. CRISPR/Cas9 can lead to different mutations in different cells, which results in different genotypes of the same plant which is called chimeric mutation.

To understand the gene editing efficiency and characteristics of CRISPR/Cas9, transgenic plants were analyzed using high-throughput sequencing. At the time of this publishing, preliminary tests had been completed on 44 plants, and the sequencing analysis of the remaining transgenic plants is in progress. Two of the 44 plants were wild-type genotypes of Kronos and CB037, and the remaining 42 plants were transformed with different CRISPR/Cas9 constructs (A1-A5) that targeted different target sites (C1-C6). In summary, gene editing occurred in 14 of the 42 tested plants, with a frequency of 33.3%.

In this research, chimerism was very common in the T_0 generation, and was present in 4 of 14 mutant plants (Table 3.3). One of their genes, for example, *Amy-A3.1*, has more than two sequences. This may be due to different mutations caused by the CRISPR/Cas9 system in different plant cells. Overall, in 25 detected mutation sites, with 4 (16%) homozygous, 11 (44%) heterozygous, 5 (20%) biallelic, and 5 (20%) chimeric mutations.

The sequencing results revealed the gene editing results of *Amy-A3.1*, *Amy-B3.1*, *Amy-D3.1*, *Amy-A3.2*, *Amy-B1-3.2* and *Amy-D3.2* genes. The *Amy-B2*-3.2 gene was not detected, which may be caused by the low amplification efficiency of the primers. All three genomes of *Amy3.1* of plant No. 2, 5, 6 and 17 were mutant and all genomes of *Amy3.2* of plant No. 28 were mutant (‘Kronos’ does not have the D genomes) (Table 3.3). Plant No. 34 had mutations in all genomes in both *Amy3.1* and *Amy3.2*.

In addition to sequencing the specific genes for each construct, it would be interesting to sequence other homologs to detect the occurrence of possible off-target mutations. Due to the high homology of genes in *Amy3.1* and *Amy3.2*, the researcher focused on the off-target phenomenon in these two groups of genes. The results showed that the eleven mutant plants (mutant plants transformed with A1-A4 vectors) did not have off-target mutations on the non-targeted genes (Table 3.3).

Table 3. 3 Genotype of fourteen T_0 mutants

No.	Genotype	Constructs	*Amy-3.1*			*Amy-3.2*[4]		
			A	**B**	**D**	**A**	**B1**	**D**
2	Kronos	A1	biallelic	biallelic	na[1]	WT	WT	na
5	Kronos	A1	chimeric	chimeric	na	WT	WT	na
6	Kronos	A1	biallelic	chimeric	na	WT	WT	na
13	CB037	A2	het[2]	WT	WT	WT	WT	WT
14	CB037	A2	het	WT	WT	WT	WT	WT
17	CB037	A2	chimeric	hom	chimeric	WT	WT	WT
22	Kronos	A3	WT	WT	na	het	WT	na
23	Kronos	A3	WT	WT	na	het	WT	na
25	CB037	A3	WT	WT	WT	WT	het	WT
28	CB037	A3	WT	WT	WT	biallelic	hom	het
30	Kronos	A4	WT	WT	na	het	WT	-
34	Kronos	A5	het	het	na	biallelic	het	na
38	Kronos	A5	WT	WT	na	het	WT	na
42	CB037	A5	WT	WT	WT	hom	hom	WT

[1]na=not applicable; 'Kronos' does not contain the D genome.

[2] het=heterozygous mutations, hom=homozygous mutations, WT=wild-type.

[3] Specific target genes per construct were highlighted by a gray background.

[4] *Amy-B2-3.2* is not studied.

3.5.2 Gene Editing Efficiency of CRISPR/Cas9

The gene editing efficiency of different genetic background was compared. The gene editing frequency of CB037 (40.0%) was slightly higher than that of Kronos (29.6%), but the difference was not significant (Table 3.4).

Table 3. 4 The T_0 mutation rate in different genetic background

Plant materials	No. of plants examined	No. of plants with mutation	Mutation rate (%)
CB037	15	6	40
Kronos	27	8	29.6
Total	42	14	33.3

In Zhang's research, different targets had different mutation efficiencies, from 21.1% to 66.7% (Zhang *et al.*, 2014). In this research, the mutation efficiency of the six targets was also significantly different, ranging from 0% to 50%, and the average mutation efficiency was 26.2% (Table 3.5). All the targets except C5 had multiple genome editing events concurrently. The frequency of multiple genome editing ranges from 6.7% to 33.3%, with an average frequency of 13.1%.

Table 3. 5 The T_0 mutation rate among targets

Targets	No. of plants examined	No. of plants with mutations	Mutation rate (%)	Multiple genomes	
				Number	%
C1	18	6	33.3	3	16.7
C2	21	2	9.5	2	9.5
C3	6	3	50	2	33.3
C4	15	3	20	1	6.7
C5	4	0	0	0	0
C6	20	8	40	3	15
Average			26.2		13.1

In this research, the editing efficiency by two sgRNAs encoded on each construct was compared (Table 3.6). The gene editing efficiency by two sgRNAs encoded on a single construct was not affected by their order on the vector. It's a common phenomenon in this research that large fragments between two targets can be deleted due to the simultaneous action of two sgRNAs on one gene. The probability of a large fragment deletion between two targets is related to the editing efficiency of each target. When the gene editing efficiency of both targets was high, a large deletion between those two targets was likely to occur. In this preliminary test, a large deletion between two targets occurred in plants transformed with the A1-A3 vectors at frequencies of 50%, 8.3% and 9.1%, respectively. Two sgRNAs encoded in the A5 vector targeted different genes respectively, so a deletion between those two targets was highly unlikely to occur. In summary, the probability of large DNA fragment deletion between two targets was 15.2% in the T_0 plants that have been tested (5 in 33, except plants transformed with A5) or 45.5% in plants carrying targeted mutations (5 in 11, except mutant plants transformed with A5).

Table 3. 6 Site-specific mutation rates in T_0 plants

Target genes[1]	Constructs	Target 1	Target 2	No. of T_0 plants examined	No. of plants with mutations				Mutation frequency (%)			
					Site 1	Site 2	Large deletion between two targets	Both sites	Site 1	Site 2	Large deletion between two targets	Both sites
Amy3.1	**A1**	**C1**	**C3**	6	3	3	3	3	50	50	50	50
Amy3.1	**A2**	**C1**	**C2**	12	3	1	1	1	25	8.3	8.3	8.3
Amy3.2	**A3**	**C4**	**C6**	11	1	5	1	1	9.1	45.5	9.1	9.1
Amy3.2	**A4**	**C4**	**C5**	4	2	0	0	0	50	0	0	0
Amy3.1* + *Amy3.2	**A5**	**C2**	**C6**	9	1	3	0	1	11.1	33.3	0	11.1

[1]Amy3 designates a gene from hexaploid wheat and/or from tetraploid wheat.

3.5.3 Variety and Frequency of Mutations

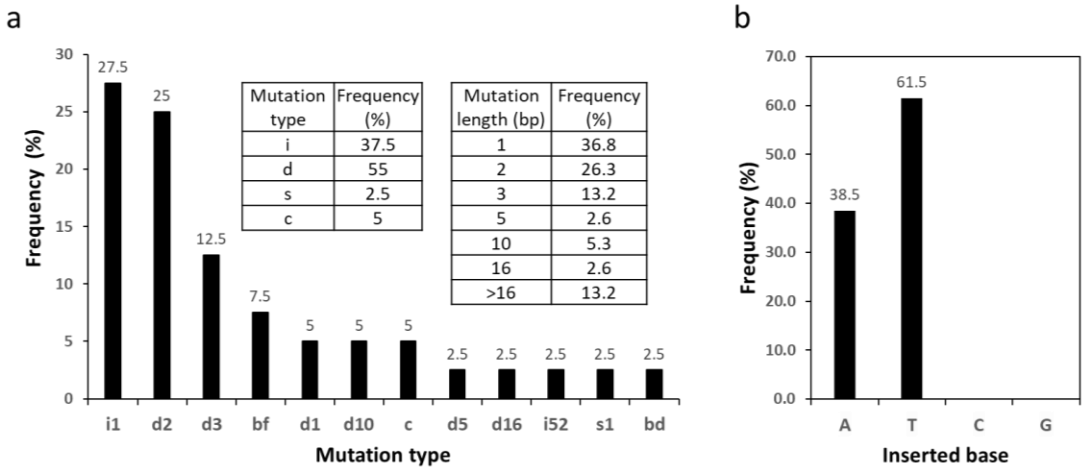


Figure 3. 6 Mutation types and frequency caused by CRISPR/Cas9

(a) The mutation types and frequency were summarized. Left inset shows the mutation types and frequency. Right inset shows the frequency of different mutation length. (b) Percentage of different bases in the 1-bp insertion mutants.

Abbreviations: i=insertion, d=deletion, s=substitution, c=multiple mutation types, # inx-axis = number of bases either inserted, deleted, or substituted, bf=big fragement change probably caused by insertion or subsitutation, bd=deletion of 16 or more bases only at one target.

In rice (Zhang *et al.*, 2014; Ma *et al.*, 2015) and Arabidopsis (Feng *et al.*, 2014), NHEJ mutations introduced by CRISPR/Cas9 were usually short insertions or deletions at the target site. Among 40 mutation types identified in rice and Arabidopsis, 1-bp insertion was the most common mutation type (Zhang *et al.* 2014; Ma *et al.* 2015; Feng *et al.* 2014), which was same as the current studies. In this research, the 1-bp insertion was followed by 2-bp (25%) and 3-bp (12.5%) deletions (Figure 3.6a). Next came the insertion or replacement of long fragments, which accounted for 7.5% of the total mutation types. Five percent of the total mutations had multiple changes in the same target region, such as an insertion and a deletion (Figure 3.6a). The insertion (37.5%) and deletion (55%) at the target sites accounted for 92.5% of the total mutation types (Left Inset, Figure 3.6a), which was also consistent with previous studies (Zhang *et al.* 2014; Feng *et al.* 2014). In terms of mutation length, single base mutations accounted for the largest proportion (36.8%), followed by two (26.3%) and three (13.2%) base mutations (Right Inset, Figure 3.6a).

A 1-bp insertion accounted for 27.5% of the total mutations, and most of the inserted bases were A (38.5%) and T (61.5%), which was consistent with the previous research results (Zhang *et al.* 2014; Ma *et al.* 2015; Feng *et al.* 2014).

Chapter 4 TILLING Results

4.1 Primers Design of *Amy3*

In hexaploid wheat, there were about 36 *α*-amylase genes on the different chromosomes, which can be divided into four gene subfamilies. Sequence alignment of all wheat *α*-amylase genes with MEGA X (Kumar *et al.*, 2018) showed that there was a high homology between different *α*-amylase genes. However, the homology of *Amy3.1* and *Amy3.2* was higher, and the gene copies in the A, B, and D chromosome were very similar, and there were only a few SNPs (single-nucleotide polymorphisms) between them in gene coding regions. Therefore, it was very difficult to design specific primers for *Amy-A3.1*, *Amy-B3.1*, *Amy-A3.2*, and *Amy-B1-3.2* in coding regions. To conduct TILLING screening, specific primers were designed for four gene copies in the UTR (untranslated region) where there was a polymorphism between the two genes (Table 4.1). The length of the amplicon was about 2 kb, and it was found that the amplification effect could not meet the requirements of CelI enzymatic digestion due to the long amplicon. So, the common primers were designed inside the *Amy3* amplicon in such a way that nested PCR could be conducted. First, the specific primers of each gene were used for the first-round PCR, and the PCR products were diluted and used as the template for the second-round PCR. The common primers were then used for the second-round PCR. By this way, the secondary PCR amplicon was more suitable for TILLING.

Table 4. 1 Primers of *Amy3* for TILLING

Genes	Primer Sequence （5’→ 3’）	Tm（℃）
Amy-A3.1	FP: CGCTACACACACGTACGGAC RP: GGGGTACAAACTTTATTGCGTG	60
Amy-B3.1	FP: CCGTCACCCTATAAATACCTTG RP: AGTGAACACAGGATTAACAACG	60
Amy-A3.2	FP: GGAGAGTAATTAAACGGCAGT RP: CTTCAGGGATAACACAATGCAGT	60
Amy-B1-3.2	FP: GGAGAGTAATTAAACGGCTCC RP: GTTTTCCTGGTTCGATTACTC	60
Amy3 common primers	FP: GTTCAACTGGGACTCGTGGA RP: AGGCCACTCTTCTCCCAGAC	59

4.2 Screening for *Amy-B3.1* Gene Mutants

The mutant population of tetraploid wheat 'Kronos' was established by the University of California, Davis. More than 2,000-point mutations have been identified in each mutant line, and the mutation information can be queried in an online database (Krasileva *et al.*, 2017). The gene sequences of *Amy-A3.1*, *Amy-B3.1*, *Amy-A3.2*, and *Amy-B1-3.2* were used as queries to search the mutants in the database, but only the mutation information of *Amy-B3.1* was recorded. The complexity of the *α*-amylase genes in wheat likely interfered the automatic computing of the exome sequencing data, and the *Amy3* gene except for *Amy-B3.1* were discarded from further analysis. However, mutations of other *TdAmy3* genes already occurred in the Kronos population.

Table 4. 2 'Kronos' mutants of the *Amy-B3.1* gene

No.	Mutant Lines	Base change	Codon change	AA change	Het/Hom	Mutation type
1	Kronos4611	C10149A	tgG/tgT	W352C	Hom	Missense variant
2	Kronos2259	C10585T	gGc/gAc	G239D	Het	Missense variant
3	Kronos3771	G10630A	cCc/cTc	P224L	Het	Missense variant
4	Kronos2852	C10658T	Gcc/Acc	A215T	Het	Missense variant
5	Kronos3964	T10721A	Aag/Tag	K194*	Het	Stop gained
6	Kronos2108	G10741A	tCc/tTc	S187F	Het	Missense variant
7	Kronos4069	G10741A	tCc/tTc	S187F	Het	Missense variant
8	Kronos798	G10741A	tCc/tTc	S187F	Het	Missense variant
9	Kronos1323	C10748T	Gag/Aag	E185K	Het	Missense variant
10	Kronos2894	G10760A	Cgc/Tgc	R181C	Het	Missense variant
11	Kronos4429	G10763A	Ccg/Tcg	P180S	Het	Missense variant
12	Kronos3474	C10813T	cGt/cAt	R163H	Het	Missense variant
13	Kronos2314	G10865A	Ccg/Tcg	P146S	Het	Missense variant
14	Kronos2364	G10880A	Ctc/Ttc	L141F	Het	Missense variant
15	Kronos4220	C10894T	gGa/gAa	G136E	Het	Missense variant
16	Kronos2317	C10918T	gGc/gAc	G128D	Het	Missense variant
17	Kronos331	C10964T	Gac/Aac	D113N	Het	Missense variant

Note: This is based on the blast search of the 'Kronos' TILLING populations database (Krasileva *et al.*, 2017).

There were 42 different mutant lines of the *Amy-B3.1* gene in the 'Kronos' mutant population, among which 19 had a missense mutation or a termination mutation. Seventeen lines of these mutants were planted in the greenhouse, of which 16 lines were missense mutants, and 1 line had a stop codon in the gene (Table 4.2). After TILLING screening, not

all lines contained mutation sites as recorded in the database. Only the amplification products of Kronos1323, Kronos2314, Kronos2364, and Kronos2894 were digested by the CelI enzyme, indicating that these four lines might contain recorded mutant bases (Figure 4.1a). Then, Singer sequencing was performed on these four lines to further verify whether the mutant information of the four lines was correct. Sequencing results showed that the mutation information of the four mutant lines was accurate. And there was a single peak at the mutation site, indicating that all the four plants were homozygous mutants (Figure 4.1b).

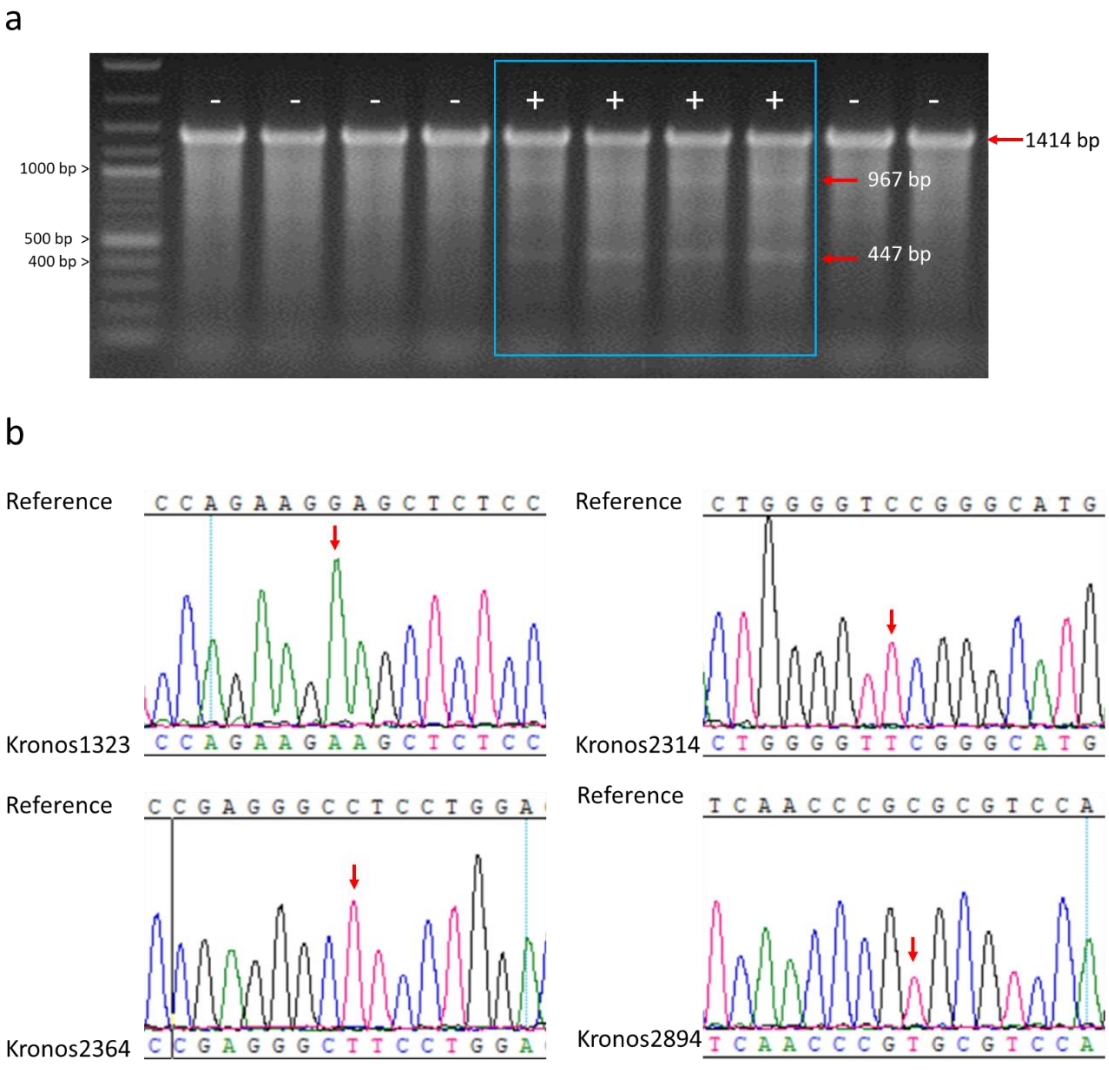


Figure 4. 1 TILLING screening of the *Amy-B3.1* mutant

(a) After PCR and CelI digestion, the PCR bands (1,414 bp) of mutant plants were digested into two small bands of 967 bp and 447 bp (blue square). (b) Mutant lines were subsequently validated by sequencing. The mutations were indicated by the red arrow in the figure. Because the mutation base was a single peak, all the four lines were homozygous mutants.

4.3 TILLING Screening of Mutations in *Amy-A3.1*, *Amy-A3.2* and *Amy-B1-3.2* Mutants

The *Amy-A3.1*, *Amy-A3.2*, and *Amy-B1-3.2* genes were not recorded in the 'Kronos' mutant database. Therefore, 1,440 'Kronos' mutant lines were planted in the greenhouse to screen for mutations in these three genes. Every 16 mutant lines in the population were arranged into a 4×4 matrix. DNA of these mutant plants was extracted in DNA pools. In the subsequent experiments, mutants were screened in the DNA R (row) pools and C (column) pools respectively. Mutant lines could thereby be found after finding the row and column pools containing mutants.

After DNA extraction, 720 row and column DNA pools were extracted in total. 360 row DNA pools were screened first using the gene specific primer (Table 4.1). Corresponding column pools were found according to the row pools; then these column pools were screened using the same method. After screening the row and column pools for mutants, the corresponding mutant lines of each gene were obtained. A total of 17 mutant lines were obtained, 13 of which were *Amy-A3.1* gene mutants and 4 of which were *Amy-A3.2* gene mutants. The mutant lines of *Amy-A3.1*, *Amy-B3.1*, and *Amy-A3.2* genes obtained from the 'Kronos' mutant population are listed in Table 4.3. The primers of *Amy-B1-3.2* gene worked poorly, so corresponding mutant lines were not found.

Sequencing results showed that five lines of the *Amy-A3.1* gene mutants contained missense mutations, among which Kronos4387 contained three mutation sites and Kronos3964 possessed a termination codon. The wild-type *Amy-A3.1* gene encodes a 468-aa α-amylase. Kronos3964 has a K194* truncation of this gene, causing a 274-aa truncation of the C terminus. Two lines of the *Amy-A3.2* gene mutants had missense mutations; another two lines had synonymous mutations.

Table 4. 3 'Kronos' mutants of *Amy3*

Gene	Mutant Lines	Base change	Codon change	AA change	Het/Hom	Consequence
Amy-A3.1	Kronos331	G/A	Gac/Aac	D113N	hom	missense_variant
	Kronos1090	G/A	Gtg/Atg	V255M	hom	missense_variant
	Kronos3771	C/T	cCc/cTc	P224L	hom	missense_variant
	Kronos4387	G/A	cGt/cAt	R163H	het	missense_variant
	Kronos4387	G/A	Gtg/Atg	V364M	hom	missense_variant
	Kronos4429	C/T	Ccg/Tcg	P180S	hom	missense_variant
	Kronos3964	A/T	Aag/Tag	K194stop	hom	stop_codon
	Kronos1159	T/C	ttT/ttC	na[1]	hom	synonymous_variant
	Kronos3863	C/T	caC/caT	na	hom	synonymous_variant
	Kronos3243	C/T	gaC/gaT	na	hom	synonymous_variant
	Kronos244	C/T	caC/caT	na	hom	synonymous_variant
	Kronos2824	G/A	ggG/ggA	na	hom	synonymous_variant
	Kronos4273	G/A	aaG/aaA	na	hom	synonymous_variant
	Kronos4387	G/C	ggG/ggC	na	hom	synonymous_variant
	Kronos4434	G/A	ggG/ggA	na	het	synonymous_variant
Amy-B3.1	Kronos1323	C/T	Gag/Aag	E185K	hom	missense_variant
	Kronos2314	G/A	Ccg/Tcg	P146S	hom	missense_variant
	Kronos2364	G/A	Ctc/Ttc	L141F	hom	missense_variant
	Kronos2894	G/A	Cgc/Tgc	R181C	hom	missense_variant
Amy-A3.2	Kronos2452	C/T	Ctc/Ttc	L259F	hom	missense_variant
	Kronos2299	G/A	agG/agA	R161S	hom	missense_variant
	Kronos2184	C/T	taC/taT	na	hom	synonymous_variant
	Kronos2242	C/T	ggC/ggT	na	het	synonymous_variant

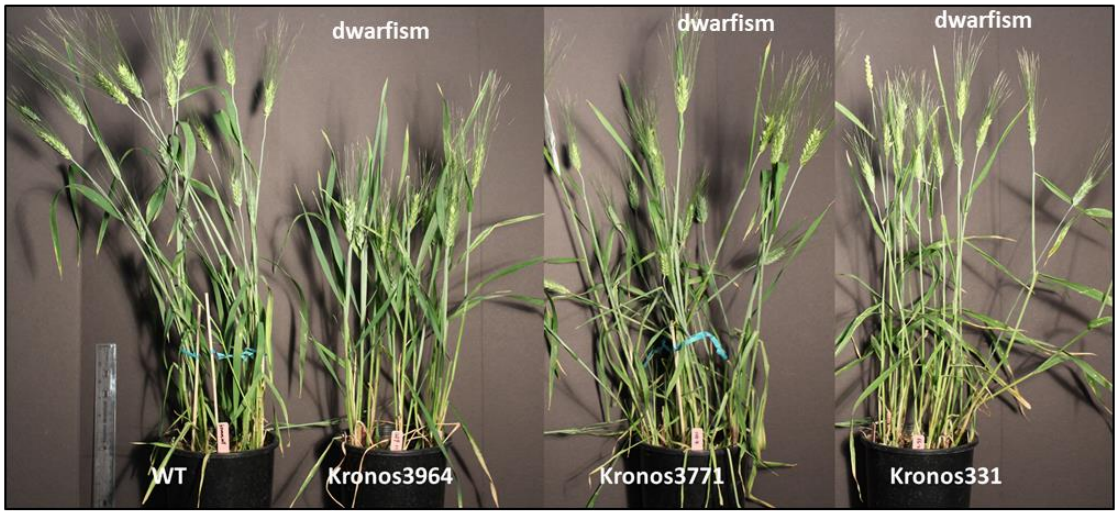
[1] na=not applicable; Synonymous variant does not change the amino acid.

4.4 Phenotype of *Amy3* Mutant

The 'Kronos' mutant plants of the α-amylase gene obtained by TILLING were planted in the greenhouse, and the initial phenotype was observed. There was no significant difference between most of the mutants and the wild-type controls. However, some missense mutants

were different from the wild-type: Kronos3964, which had a premature stop mutation of the *Amy-A3.1* gene, was weaker and had a shorter height than that of the wild-type (Figure 4.2a). Male sterility was observed in most florets, and only 2-3 stunted seeds were produced per spike (Figure 4.2b) and the seeds were misshapen. Kronos331 had a missense mutation on *Amy-A3.1*, and it had a similar phenotype as Kronos3964 (Figure 4.2a). The plant height of the *Amy-B3.1* mutant, Kronos2894, was significantly reduced, and the leaves had lesion mimic necrotic spots (Figure 4.2c). The mutant Kronos2314 had a shorter plant height, thinner stems, more tillers and a slower growth habit than the wild-type (Figure 4.2d). The *Amy-A3.2* gene mutant, Kronos2452, was weak in growth habit, thin in the stem, with only 1-2 tillers and small spikes (Figure 4.2e).

a



b

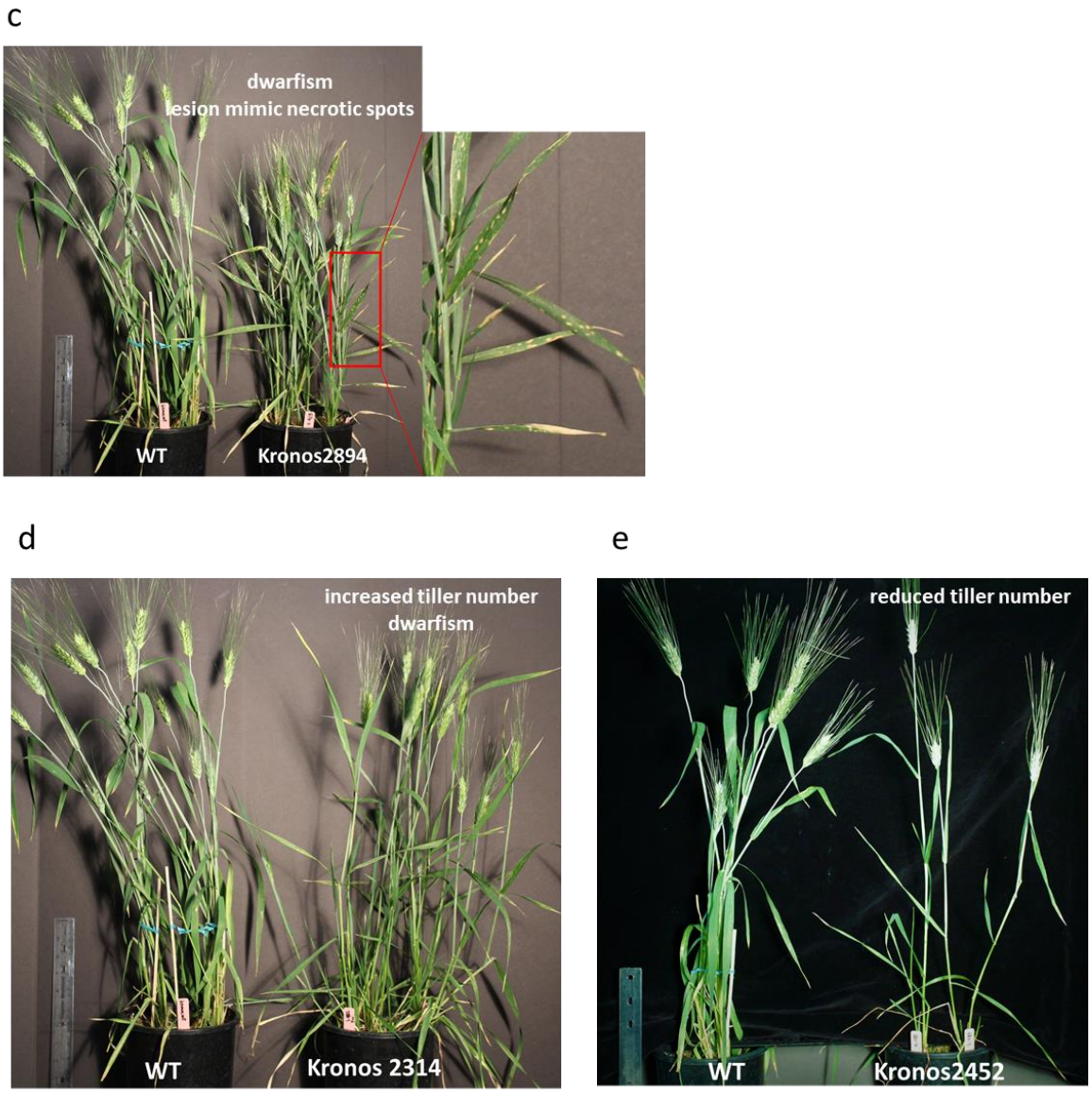


Figure 4. 2 Phenotype of 'Kronos' mutants

These lines had about 2,000 other mutations in addition to the *Amy3* gene that may affect the phenotype. To prove the effect of *Amy3* gene on plant growth and falling number, one must hybridize the mutant with the wild-type. This work still needs to be done in order purify the genetic background and thereby reduce the effect of other mutation sites on plant phenotype.

Chapter 5 Discussion, Conclusion and Future Directions

5.1 Discussion

5.1.1 Classification of α-amylase

α-amylase is the best known amylolytic enzyme (Janecek *et al.*, 2013). However, the characteristics and function of different isoforms of *α*-amylase in grains remains unclear. Rice contains ten *α*-amylase genes, which can be divided into three subfamilies (*Amy1*, *Amy2*, *Amy3*) (Huang *et al.*, 1992). In barley, *α*-amylase can be divided into four isoforms: AMY1, AMY2, AMY3, and AMY4 (Radchuk *et al.*, 2009). In wheat, three isoforms of *α*-amylase have been identified (Amy1, Amy2, Amy3), and a fourth has been suggested based on detection of expression (Barrero *et al.*, 2013).

However, due to the large size (16 Gb) of the wheat genome, many questions remain unanswered. For example, how many homologous *α*-amylase genes are there in wheat, and how many *α*-amylase genes are there in each subfamily? At present, the NCBI database is not comprehensive in its recording of the *α*-amylase genes. In August 2018, the complete genome sequence of wheat was published, and the genomic sequence data increased from 10.2 Gb to 14.5 Gb (Borrill *et al.*, 2018). Complete genome sequences and detailed annotations will enable more thorough analysis of the *α*-amylase gene family in wheat. In this study, we used the latest wheat gene annotation information to complete the analysis of the *α*-amylase gene family. After an initial search using HMMER v3.2.1, 66 *α*-amylase genes were obtained. Then, after sequence alignment and evolutionary analysis, approximately 36 *α*-amylase genes were finally identified, and their complete sequences were obtained. According to the phylogenetic tree, the *α*-amylase genes were classified into different subfamilies. Completing the identification and analysis of this family will greatly benefit future researchers.

5.1.2 Detection Method of CRISPR/Cas9 Induced Mutation

Detecting successfully edited mutants is an important part of all CRISPR/Cas9 protocols. This is not difficult for diploid organisms, but it is more difficult for plants with complex genomes, especially polyploidy wheat which has four to six genomes.

In plant genome editing, there are mainly two detection methods for mutations based on PCR and enzyme digestion. In the first method, the mutation will be detected by the PCR/restriction enzyme (PCR/RE) assay. The restriction enzyme sites within the target sequences at the Cas9 endonuclease cutting site (3-bp upstream 5'-NGG) will facilitate detection. In this way, after gene editing, specific primers were used to amplify the target sites and then the mutations were detected by a restriction enzyme digestion reaction. If the PCR products can be digested completely, then there are no mutations; if they are only partially digested, then they carry single-allelic mutations; if there are no digested products, then they carry biallelic mutations (Shan *et al.*, 2014; Li *et al.*, 2016; Yanpeng *et al.*, 2014). Use of this detection method significantly limits the choice of targets since they must contain useful restriction enzyme sites. The second method is T7EI assay (T7 endonuclease I). In this method, restriction enzyme sites do not need to be considered. PCR primers are designed to amplify 300-600 bp surrounding the genomic target site. The PCR products (a mixture of wild-type allele and mutant allele) are then denatured and renatured, forming heteroduplexes. The reaction products are digested with T7EI nuclease, and then they are analyzed by 2.0% (wt/vol) agarose gel electrophoresis (Shan *et al.*, 2014; Babon *et al.*, 2003). In this research, the CelI enzyme was used to achieve a similar effect as T7EI.

Although mutations introduced by gene editing can be detected by these two methods, the mutants still need to be sequenced to reveal the specific editing that occurred. In the procedure followed here, the PCR product of the positive transgenic plants was sequenced. If the sequencing results contained superimposed sequencing chromatograms, it indicated that editing had taken place at the beginning of overlapping peaks. Then, the mutant alleles were further characterized by subcloning and sequencing. Software programs are also available that can divide the superimposed sequencing chromatograms into two sequences (Ma *et al.*, 2015).

All of the above methods have different defects in analyzing polyploidy wheat which has a large and complex genome. Sequencing of PCR products or sequencing clones requires a lot of time, energy and money. For hexaploid wheat, an average of six clones need to be sequenced to obtain the gene editing result of one gene copy. With the development of next generation sequencing technology, the cost of high-throughput sequencing has been

dramatically reduced. CRISPR/Cas9 induced mutations can be detected in large quantities by high-throughput sequencing, which is a simple and efficient choice (Bell *et al.*, 2014).

In this research, since α-amylase has many homologous genes and they have high sequence homology, enzyme digestion detection or Sanger sequencing would cost considerable of time and money for each of the six target regions of 657 transgenic plants. Therefore, high-throughput sequencing is the preferred choice. For preliminary experiments described here, we sequenced 42 transgenic plants through MiSeq Sequencing v3 of Illumina. The reading length of each sequence was 300 bp. A total of 600 bp sequences were obtained by bi-directional sequencing. The normal amplicon coverage varied from 300 fold to 3000 fold, and this fulfilled the need for analysis purposes. With this method, the detailed mutations of each gene copy were very clear, and multiple mutant types of the same gene could be found, which is difficult to achieve by cloning and sequencing.

5.1.3 Characterization of Targeted Editing in Transgenic Wheat

CRISPR/Cas9 can create targeted double-strand breaks (DSBs) in the nuclear DNA which will trigger two independent endogenous DNA repair pathways: non-homologous end joining (NHEJ) and homology-directed repair (HDR) (Sonoda *et al.*, 2006; Symington *et al.*, 2011). In the process of plant NHEJ pathway, a remarkable feature is that large fragments are often inserted at the fracture location. The inserted fragment may come from the vector sequence or the plant genome sequence (Gorbunova, 1997). In this research, it was found that plant No.6 had a 52 bp insertion at the C3 target. It was verified that this sequence came from the promoter region of *Cas9* gene in pYLCRISPR/Cas9-MB vector, which was consistent with previous research results. This may be because the vector breaks down during the transformation or after entering the cell, and the broken fragments are randomly inserted into the broken genomic DNA during repair.

In previous studies in rice, the gene editing efficiency of different targets was different, with the efficiency ranging from 21% to 67%. Among them, 53.9% of the mutation types were a 1-bp insertion, and most of the inserted bases were A or T (Zhang *et al.*, 2014). In Ma's research, they constructed twenty CRISPR/Cas9 vectors to target 46 different locations. In the end, they got 280 different mutations, among which 177 were diallelic mutations, 81 were homozygous mutations, and 19 were heterozygous mutations. 54.1% of the total

mutations were a 1-bp insertion of A or T nucleotide (Ma *et al.*, 2015). The results of this experiment were essentially consistent with those of other studies. The gene editing efficiencies of different targets were different, ranging from 0 to 50%. One base pair of an A or T insertion (27.5%) was the most common mutation type. For the mutant genotype, heterozygous mutations account for a large proportion (44%), and the next most common class were biallelic (20%), and lastly were chimeric mutations (20%). Homozygous mutants accounted for only 16%, which was inconsistent with Ma's research (Ma *et al.*, 2015).

5.1.4 Off-target Mutation of CRISPR/Cas9

Precise gene editing technology is of great significance to biological research and crop improvement. CRISPR/Cas9 is the third-generation gene editing technology. Compared with the first-generation gene editing technology - ZFNs (Fyodor D *et al.*, 2010) and the second-generation gene editing technology - TALENs (Bedell *et al.*, 2012), the experimental design is more straightforward and cheaper. As a result, CRISPR/Cas9 has replaced the previous gene editing technologies.

One of the technical difficulties with CRISPR/Cas9 is the potential for off-target mutations. The specificity of a target sequence influences the possibility of off-target events. When choosing the target sites, the selected sgRNA coding sequences should be at different from the homologous sequences to avoid the off-target mutations (Hsu *et al.*, 2013; Jinek *et al.*, 2012).

In this experiment, due to the large number of *α*-amylase genes in wheat, the specificity of the sgRNA coding sequence should be paid careful attention when selecting the target sequence. If multiple *α*-amylase genes are knocked out due to the off-target problem, the normal growth and development of wheat may be affected. All of the six target sites in this research were specific, among which C1, C2, and C4 contain PAM sequences only in their direct target genes. For target sites associated with a PAM sequence in both the direct and indirect target genes, they are selected in a region with three to eleven different bases between the direct and indirect target genes; this could ensure the specificity of targeted gene editing. Because of their high homology, both *Amy3.1* gene and *Amy3.2* gene were sequenced in plants that had been transformed with the single-gene targeted construct. This

allowed us to detect if a single gene targeted had an off-target mutation on another gene group.

After the transformation with CRISPR/Cas9 construct, if the construct is integrated into the genome, sgRNA and Cas9 will continue to be expressed in the offspring, resulting in new editing events and increasing the possibility of off-target mutations. At present, this situation can be avoided by transforming a ribonucleoprotein (RNP) made of in vitro produced Cas9 and sgRNA (Liang *et al.*, 2017). Because only protein complexes are transferred into the plant, there is no exogenous DNA.

Off-target mutations can be detected by whole-genome sequencing. Zhang *et al.* conducted whole-genome sequencing of gene-edited rice to detect off-target mutations, and the results showed that none of the targets showed a serious off-target situation (Zhang *et al.*, 2014). In this research, no off-target mutations were found in the *Amy3.1* and *Amy3.2* gene groups of the 11 mutants.

Compared with animals, the off-target mutation is not a very serious problem in gene-edited plants. The tissue culture process itself can induce a number of mutations (Phillips *et al.*, 1994). In wheat, due to the large number of gene copies on the multiple chromosomes, the loss-of-function of one gene would be compensated for by the other copies. Also, highly specific sgRNA constructs can reduce or avoid the occurrence of off-targets mutations (Xie *et al.*, 2014). In addition, the off-target mutation can be isolated and eliminated by backcrossing or hybridization.

5.2 Conclusion and Future Directions

In the past few years, people have been striving for an understanding of wheat falling numbers. This study attempts to explore the factors that influence the falling numbers of wheat from the molecular level. In this thesis, we combined the emerging gene editing method with the traditional mutagenesis screening method to create or find α-amylase gene mutants.

I have summarized the expression levels of different α-amylase genes in wheat. Based on the expression of different α-amylase genes, *Amy3* mutants were created in two ways. On

the one hand, four single-gene and one dual-gene targeted CRISPR/Cas9 constructs of *Amy3.1* and *Amy3.2* were made. The vectors were then transformed into wheat through biolistic bombardment. The gene edited plants were screened and identified. TILLING (Targeting Induced Local Lesions In Genomes) technology was used to screen for *Amy*3 gene mutants in the 'Kronos' mutant population. Mutant plants with amino acid changes in *Amy-A3.1*, *Amy-B3.1*, and *Amy-A3.2* genes were obtained. This is expected to provide reliable mutants and experimental data for further analysis of the relationship between wheat *Amy3* genes and falling numbers. The main results are as follows:

1. Based on the latest wheat genome sequence and annotation data (IWGSC reference v1.0), 36 different copies of *α*-amylase gene in 'Chinese Spring' were found, and then divided into four gene subfamilies. The third subfamily - *Amy3*, contains three groups of genes, of which *Amy3.1* contains one copy on A, B and D genomes, *Amy3.2* contains one copy on A and D genomes, and two copies on B genome and *Amy3.3* contains two copies on B genome and one copy on D genome.

2. The gene sequences of wheat *α*-amylase were used to screen for gene expression in the WheatExp database. The results showed that the expression profiles for *α*-amylase were different between subfamilies. In all *α*-amylase genes, only genes in *Amy3.1* and *Amy3.2* were expressed in seeds, but not in other tissues. According to this result, I tried to create mutations in *Amy3.1* and *Amy3.2* to reduce their expression in the grains without affecting the normal growth of plants, thereby aiming to improve the falling numbers of wheat flour.

3. In order to create *Amy3* mutants through CRISPR/Cas9, six target sites (C1-C6) were chosen based on *Amy3*, and five CRISPR/Cas9 constructs (A1-A5) were made, each containing two target sites. A1 and A2 were used to knock out *Amy3.1*; A3 and A4 were used to knock out *Amy3.2*; and A5 could knock out both genes simultaneously.

4. Wheat genetic transformation was processed using five CRISPR/Cas9 constructs (A1-A5). A total of 10,290 immature embryos were transformed to obtain 657 positive transgenic plants among which 609 were independent transgenic events. The total transgenic efficiency of these experiments was 5%. A total of 227 transgenic plants were obtained for the A1 and A2 constructs, 171 transgenic plants were obtained for the A3 and A4 constructs, and 271 transgenic plants were obtained for the A5 construct.

5. Forty-two of the 657 transgenic plants were sequenced on the MiSeq Sequencing v3 platform. The results showed that gene editing occurred in 14 of the 42 plants, and the gene editing efficiency was 33.3%. The main type of mutation was a 1-bp insertion (27.5%), where the inserted base was mainly a T or an A. In addition, fragment deletion between two targets accounted for 45.5% of total mutations (5 in 11 mutant plants). For genotype, homozygous, heterozygous, biallelic and chimeric mutations accounted for 16%, 44%, 20% and 20%, respectively. No off-target editing phenomena were found in the 14 mutants.

6. Using TIILING technology and DNA mixing pools, 1,440 lines in a Kronos mutant population were screened, and 13 mutant lines of *Amy-A3.1*, 4 mutant lines of *Amy-B3.1*, and 4 mutant lines of *Amy-A3.2* were obtained. A preliminary phenotypic observation was made on these mutant lines, and it was found that partial mutants with amino acid changes had a reduced plant height, and some mutant lines had male sterility. More experiments are needed to prove the relationship between the *Amy3* genes and phenotype.

So far, we had obtained mutants of α-amylase in two ways. Our ultimate goal is to find out whether the mutation of *Amy3* gene will affect the growth and development of wheat, and whether it will increase the falling number of wheats. In the future, we will do more work to accomplish this goal. Firstly, the mutation induced by CRISPR/Cas9 will undergo genotype isolation in the T_1 generation. Therefore, we will determine the genotype of each plant in T_1 by next-generation sequencing. Also, the mutation in different gene copies of *Amy3* needs to be aggregated by hybridization. Detailed agronomic phenotype identification and falling numbers detection will need to be performed for the mutants of different genes.

References

AACCI Method 56-81.03 (n.d.) AACC International. Approved Methods of Analysis, 11th edn, Method 22-08.01, Determination of Falling Number. Approved 3 November 1999. St. Paul, MN, USA: AACC International, 2000; Approved Method 56-81.03.

Ainsworth C.C., Doherty P., Edwards K.G., Martienssen R.A., Gale M.D.. Allelic variation at α-Amylase loci in hexaploid wheat. Theor.appl.genet, 1985,70(4):400-6.

Babon J.J., McKenzie M., Cotton R.G.H.. The Use of Resolvases T4 Endonuclease VII and T7 Endonuclease I in Mutation Detection. Molecular Biotechnology, 2003,23(1):73-82.

Ball S.G., Morell M.K.. From bacterial glycogen to starch: understanding the biogenesis of the plant starch granule. Annual Review of Plant Biology, 2003,54(1):207-33.

Barrero J.M., Mrva K., Talbot M.J., White R.G., Taylor J., Gubler F., et al.. Genetic, Hormonal, and Physiological Analysis of Late Maturity -Amylase in Wheat. PLANT PHYSIOLOGY, 2013,161(3):1265-77.

Baulcombe D.C., Huttly A.K., Martienssen R.A., Barker R.F., Jarvis M.G.. A novel wheat alpha-amylase gene (alpha-Amy3). Molecular & General Genetics Mgg, 1987,209(1):33-40.

Bedell V.M., Wang Y., Campbell J.M., Poshusta T.L., Starker C.G., Krug R.G., et al.. In vivo Genome Editing Using High Efficiency TALENs. Nature, 2012,491(7422):114-8.

Bell C.C., Magor G.W., Gillinder K.R., Perkins A.C.. A high-throughput screening strategy for detecting CRISPR-Cas9 induced mutations using next-generation sequencing. BMC Genomics, 2014,15(1):1-7.

Bethke P.C., Schuurink R., Jones R.L.. Hormonal signalling in cereal aleurone. Journal of Experimental Botany, 1997,48(312):1337-56.

Black M., Bewley J.D., Halmer P., Black M., Bewley J.D., Halmer P.. The encyclopedia of seeds: science, technology and uses. 2006.

Bland C., Ramsey T.L., Sabree F., Lowe M., Brown K., Kyrpides N.C., et al.. CRISPR Recognition Tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. BMC Bioinformatics, 2007,8(1):209.

Bogdanove V.. TAL Effectors: Customizable Proteins for DNA Targeting. Science (New York, NY), 2013,333(6051):1-5.

Bolon Y.T., Haun W.J., Xu W.W., Grant D., Stacey M.G., Nelson R.T., et al.. Phenotypic and Genomic Analyses of a Fast Neutron Mutant Population Resource in Soybean. Plant Physiology, 2011,156(1):240-53.

Borrill P., Abrouk M., Valárik M., Poland J., Datla R., Ormanbekova D., et al.. Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science, 2018,361(6403):eaar7191.

Brown L.K., Wiersma A.T., Olson E.L.. Preharvest sprouting and α-amylase activity in soft winter wheat. Journal of Cereal Science, 2018,79:311-8.

Carroll D.. Genome engineering with zinc-finger nucleases. Genetics, 2011,188(4):773-82.

Charpentier E., Doudna J.. The new frontier of genome engineering with CRISPR-Cas9. Science, 2014,346(6213):1077-887.

Derera N.F., Bhatt G.M., Mcmaster G.J.. On the problem of pre-harvest sprouting of wheat. Euphytica, 1977,26(2):299-308.

Eddy S.R.. Profile hidden Markov models. Bioinformatics, 1998,14(9):755-63.

Fauser F., Schiml S., Puchta H.. Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in Arabidopsis thaliana. Plant Journal, 2014,79(2):348-59.

Feng Z., Mao Y., Xu N., Zhang B., Wei P., Yang D.-L., et al.. Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in Arabidopsis. Proceedings of the National Academy of Sciences, 2014,111(12):4632-7.

Feng Z., Zhang B., Ding W., Liu X., Yang D.L., Wei P., et al.. Efficient genome editing in plants using a CRISPR/Cas system. Cell Research, 2013,23(10):1229-32.

Fyodor D U., Edward J R., Michael C H., H Steve Z., Philip D G.. Genome editing with engineered zinc finger nucleases. Nature Reviews Genetics, 2010,11(9):636-46.

Gale D.M.. Alpha-amylase production in the late stages of grain development-an early sprouting damage risk period? Third International Symposium on Pre-Harvest Sprouting in Cereals, 1983:29-35.

Gale M.D., Flintham J.E., Devos K.M.. Cereal comparative genetics and preharvest sprouting. Euphytica, 2002,126(1):21-5.

Gasiunas G., Barrangou R., Horvath P., Siksnys V.. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proceedings of the National Academy of Sciences of the United States of America, 2012,109(39):15539-40.

Gorbunova V.. Non-homologous DNA end joining in plant cells is associated with deletions and filler DNA insertions. Nucleic Acids Research, 1997,25(22):4650-7.

Gu H., Qin G., Miao J., Guo D., Zhang J., Wan J., et al.. Targeted mutagenesis in rice using CRISPR-Cas system. Cell Research, 2013,23(10):1233-6.

Gubler F., Millar A.A., Jacobsen J. V.. Dormancy release, ABA and pre-harvest sprouting. Current Opinion in Plant Biology, 2005,8(2):183-7.

Henry I.M., Nagalakshmi U., Lieberman M.C., Ngo K.J., Krasileva K. V, Vasquez-Gross H., et al.. Efficient Genome-Wide Detection and Cataloging of EMS-Induced Mutations Using Exome Capture and Next-Generation Sequencing. Plant Cell, 2014,26(4):1382.

Holmes M.C., Rupniewski I., Gregory P.D., Rebar E.J., Beausejour C.M., Wang J., et al.. An improved zinc-finger nuclease architecture for highly specific genome editing. Nature Biotechnology, 2007,25(7):778-85.

Hsu P.D., Scott D.A., Weinstein J.A., Ran F.A., Konermann S., Agarwala V., et al.. DNA targeting specificity of RNA-guided Cas9 nucleases. Nature Biotechnology, 2013,31(9):827-32.

Huanbin Z., Bo L., Weeks D.P., Spalding M.H., Bing Y.. Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. Nucleic Acids Research, 2014,42(17):10903-14.

Huang N., Stebbins G.L., Rodriguez R.L.. Classification and evolution of alpha-amylase genes in plants. Proceedings of the National Academy of Sciences of the United States of America, 1992,89(16):7526-30.

Huang S., Yun S.H., Quail K., Moss R.. Establishment of flour quality guidelines for northern style Chinese steamed bread. Journal of Cereal Science, 1996,24(2):179-85.

Humphreys D.G., Noll J.. Methods for characterization of preharvest sprouting resistance in a wheat breeding program. Euphytica, 2002,126(1):61-5.

Ishino Y., Shinagawa H., Makino K., Amemura M., Nakata A.. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. Journal of Bacteriology, 1987,169(12):5429-33.

Janecek S., Svensson B., Macgregor E.A.. Alpha-amylase - an enzyme present in various sequence-based glycoside hydrolase families. 2013:

Jiang W.Z., Yang B., Weeks D.P.. Efficient CRISPR/Cas9-mediated gene editing in Arabidopsis thaliana and inheritance of modified genes in the T2 and T3 generations. PLoS ONE, 2014,9(6):e99225.

Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E.. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science, 2012,337:816-21.

Krasileva K., Vasquezgross H., Howell T., Bailey P., Paraiso F., Clissold L., et al.. Uncovering hidden variation in polyploid wheat. Proc Natl Acad Sci U S A, 2017,114(6):E913.

Kumar S., Stecher G., Li M., Knyaz C., Tamura K.. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Molecular Biology and Evolution, 2018,35(6):1547-9.

Lee J.H., Lee S.Y.. Selection of stable mutants from cultured rice anthers treated with ethyl methane sulfonic acid. Plant Cell Tissue & Organ Culture, 2002,71(2):165-71.

Li J., Liu J., Gao C., Zong Y., Li J., Meng X., et al.. Gene replacements and insertions in rice by intron targeting using CRISPR–Cas9 [En ligne]. Nature Plants. 2016. p 16139.

Li J., Sun Y., Du J., Xia L.. Generation of Targeted Point Mutations in Rice by a Modified CRISPR/Cas9 System. Molecular Plant, 2017,10(3):526-9.

Li J.-F., McCormack M., Aach J., Bush J., Church G.M., Sheen J., et al.. Multiplex and homologous recombination–mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. Nature Biotechnology. 2013. p 688-91.

Liang Z., Chen K., Li T., Zhang Y., Wang Y., Zhao Q., et al.. Efficient DNA-free genome editing of bread wheat using CRISPR/Cas9 ribonucleoprotein complexes. Nature Communications, 2017,8:1-5.

Lv B., Nitcher R., Han X., Wang S., Ni F., Li K., et al.. Characterization of Flowering Locus T1 (FT1) gene in Brachypodium and wheat. PLoS ONE, 2014,9(4):e94171.

Ma X., Chen L., Zhu Q., Chen Y., Liu Y.-G.. Rapid Decoding of Sequence-Specific Nuclease-Induced Heterozygous and Biallelic Mutations by Direct Sequencing of PCR Products. Molecular Plant, 2015,8(8):1285-7.

Ma X., Zhang Q., Zhu Q., Liu W., Chen Y., Qiu R., et al.. A Robust CRISPR/Cas9 System for Convenient, High-Efficiency Multiplex Genome Editing in Monocot and Dicot Plants. Molecular Plant, 2015,8(8):1274-84.

Majzlová K., Pukajová Z., Janeček S.. Tracing the evolution of the α-amylase subfamily GH13_36 covering the amylolytic enzymes intermediate between oligo-1,6-glucosidases and neopullulanases. Carbohydrate Research, 2013,367(367C):48-57.

Mao Y., Zhang H., Xu N., Zhang B., Gou F., Zhu J.K.. Application of the CRISPR–Cas System for Efficient Genome Engineering in Plants. Molecular Plant, 2008,6(6):2008-11.

Mares D., Mrva K.. Late-maturity -amylase: Low falling number in wheat in the absence of preharvest sprouting. Journal of Cereal Science, 2008,47(1):6-17.

Mares D.J., Mrva K.. Wheat grain preharvest sprouting and late maturity alpha-amylase. Planta, 2014,240(6):1167-78.

Mccallum C.M., Comai L., Greene E.A., Henikoff S.. Targeted screening for induced mutations. Nature Biotechnology, 2000,18(4):455-7.

McCallum C.M., Comai L., Greene E.A., Henikoff S.. Targeting induced local lesions IN genomes (TILLING) for plant functional genomics. Plant physiology, 2000,123(2):439-42.

Mishra A., Singh A., Sharma M., Kumar P., Roy J.. Development of EMS-induced mutation population for amylose and resistant starch variation in bread wheat (Triticum aestivum) and identification of candidate genes responsible for amylose variation. Bmc Plant Biology, 2016,16(1):217.

Ng S.B., Buckingham K.J., Lee C., Bigham A.W., Tabor H.K., Dent K.M., et al.. Exome sequencing identifies the cause of a mendelian disorder. Nature Genetics, 2010,42(1):30-5.

Ng S.B., Turner E.H., Robertson P.D., Flygare S.D., Bigham A.W., Lee C., et al.. Targeted capture and massively parallel sequencing of 12 human exomes. Nature, [En ligne] 2009,461. (Disponible sur : < https://www.nature.com/articles/nature08250.pdf > [Consulté le 20/02/2019]).

Niewoehner O., Jinek M., Doudna J.A.. Evolution of CRISPR RNA recognition and processing by Cas6 endonucleases. Nucleic Acids Research, 2014,42(2):1341-53.

Pearce S., Vazquez-Gross H., Herin S.Y., Hane D., Wang Y., Gu Y.Q., et al.. WheatExp: an RNA-seq expression database for polyploid wheat. Bmc Plant Biology, 2015,15(1):299.

Phillips R.L., Kaeppler S.M., Olhoft. P.. Genetic instability of plant tissue cultures: breakdown of normal controls. Proceedings of the National Academy of Sciences of the United States of America, 1994,91(12):5222-6.

Qiwei S., Yanpeng W., Jun L., Yi Z., Kunling C., Zhen L., et al.. Targeted genome modification of crop plants using a CRISPR-Cas system. Nature Biotechnology, 2013,31(8):686-8.

Radchuk V. V, Borisjuk L., Sreenivasulu N., Merx K., Mock H.P., Rolletschek H., et al.. Spatiotemporal Profiling of Starch Biosynthesis and Degradation in the Developing Barley Grain. Plant Physiology, 2009,150(1):190-204.

Ral J.P., Whan A., Larroque O., Leyne E., Pritchard J., Dielen A.S., et al.. Engineering high α-amylase levels in wheat grain lowers Falling Number but improves baking properties. Plant Biotechnology Journal. 2016. p 364-76.

Richards M., Deveau H., Romero D.A., Horvath P., Barrangou R., Moineau S., et al.. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. Science, 2007,315(5819):1709-12.

Schmidt C., Voytas D.F., Hummel A., Christian M., Cermak T., Zhang F., et al.. Targeting DNA Double-Strand Breaks with TAL Effector Nucleases. Genetics, 2010,186(2):757-61.

Shan Q., Wang Y., Chen K., Liang Z., Li J., Zhang Y., et al.. Rapid and efficient gene modification in rice and Brachypodium using TALENs. Molecular Plant, 2013,6(4):1365-8.

Shan Q., Wang Y., Li J., Gao C.. Genome editing in rice and wheat using the CRISPR/Cas system. Nature Protocols. 2014. p 2395-410.

Sinkunas T., Gasiunas G., Waghmare S.P., Dickman M.J., Barrangou R., Horvath P., et al.. In vitro reconstitution of Cascade-mediated CRISPR immunity in Streptococcus thermophilus. EMBO Journal, 2013,32(3):385-94.

Sonoda E., Hochegger H., Saberi A., Taniguchi Y., Takeda S.. Differential usage of non-homologous end-joining and homologous recombination in double strand break repair. Dna Repair, 2006,5(9):1021-9.

Svitashev S., Young J.K., Schwartz C., Gao H., Falco S.C., Cigan A.M.. Targeted Mutagenesis, Precise Gene Editing, and Site-Specific Gene Insertion in Maize Using Cas9 and Guide RNA [En ligne]. Plant Physiology. 2015. p 931-45.

Symington L.S., Gautier J.. Double-Strand Break End Resection and Repair Pathway Choice. Annual Review of Genetics, 2011,45(1):247-71.

Terns M.P.. CRISPR-Based Technologies: Impact of RNA-Targeting Systems. Molecular Cell, 2018,72(3):404-12.

Till B.J., Cooper J., Tai T.H., Colowit P., Greene E.A., Henikoff S., et al.. Discovery of chemically induced mutations in rice by TILLING. Bmc Plant Biology, 2007,7(1):19.

Till B.J., Reynolds S.H., Weil C., Springer N., Burtner C., Young K., et al.. Discovery of induced point mutations in maize genes by TILLING. Bmc Plant Biology, 2004,4(1):12.

Upadhyay S.K., Kumar J., Alok A., Tuli R.. RNA-Guided Genome Editing for Target Gene Mutations in Wheat. G3: Genes|Genomes|Genetics, 2013,3(12):2233-8.

Wardak C.. Solid contact nitrate ion-selective electrode based on ionic liquid with stable and reproducible potential. Electroanalysis, 2014,26(4):864-72.

Weeks J.T., Anderson O.D., Blechl A.E.. Rapid Production of Multiple Independent Lines of Fertile Transgenic Wheat (Triticum aestivum). Plant Physiology, 2016,102(4):1077-84.

Wei C., Liu J., Yu Z., Zhang B., Gao G., Jiao R.. TALEN or Cas9-Rapid,Efficient and Specific Choices for Genome Modifications. Journal of Genetics & Genomics, 2013,40(6):281-9.

Whan A., Dielen A.S., Mieog J., Bowerman A.F., Robinson H.M., Byrne K., et al.. Engineering α-amylase levels in wheat grain suggests a highly sophisticated level of carbohydrate regulation during development. Journal of Experimental Botany, 2014,65(18):5443-57.

Wood A.J., Lo T.W., Zeitler B., Pickle C.S., Ralston E.J., Lee A.H., et al.. Targeted genome editing across species using ZFNs and TALENs. Science, 2011,333(6040):307.

Xia Q., Zhang X., Wu Y., Ma S., Wu X., Wang G., et al.. CRISPR/Cas9-mediated targeted mutagenesis in Nicotiana tabacum. Plant Molecular Biology, 2014,87(1-2):99-110.

Xie K., Yang Y.. RNA-Guided Genome Editing in Plants Using a CRISPR–Cas System. Molecular Plant, 2013,6(6):1975-83.

Xie K., Zhang J., Yang Y.. Genome-wide prediction of highly specific guide RNA spacers for CRISPR-Cas9-mediated genome editing in model plants and major crops. Molecular Plant, 2014,7(5):923-6.

Yang B., Wen X., Kodali N.S., Oleykowski C.A., Miller C.G., Kulinski J., et al.. Purification, cloning, and characterization of the CEL I nuclease. Biochemistry, 2000,39(13):3533-41.

Yanpeng W., Xi C., Qiwei S., Yi Z., Jinxing L., Caixia G., et al.. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. Nature Biotechnology, 2014,32(9):947-51.

Zadoks J.C., Chang T.T., Konzak C.F.. A decimal code for the growth stages of cereals. Weed Research, 1974,14(6):415-21.

Zhang H., Zhang J., Wei P., Zhang B., Gou F., Feng Z., et al.. The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. Plant Biotechnology Journal. 2014. p 797-807.

Zhou H., Liu B., Weeks D.P., Spalding M.H., Yang B.. Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. Nucleic Acids Research, 2014,42(17):10903-14.

Zong Y., Zhang R., Chen K., Wang Y., Ran Y., Gao C., et al.. Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion. Nature Biotechnology, 2017,35(5):438-40.