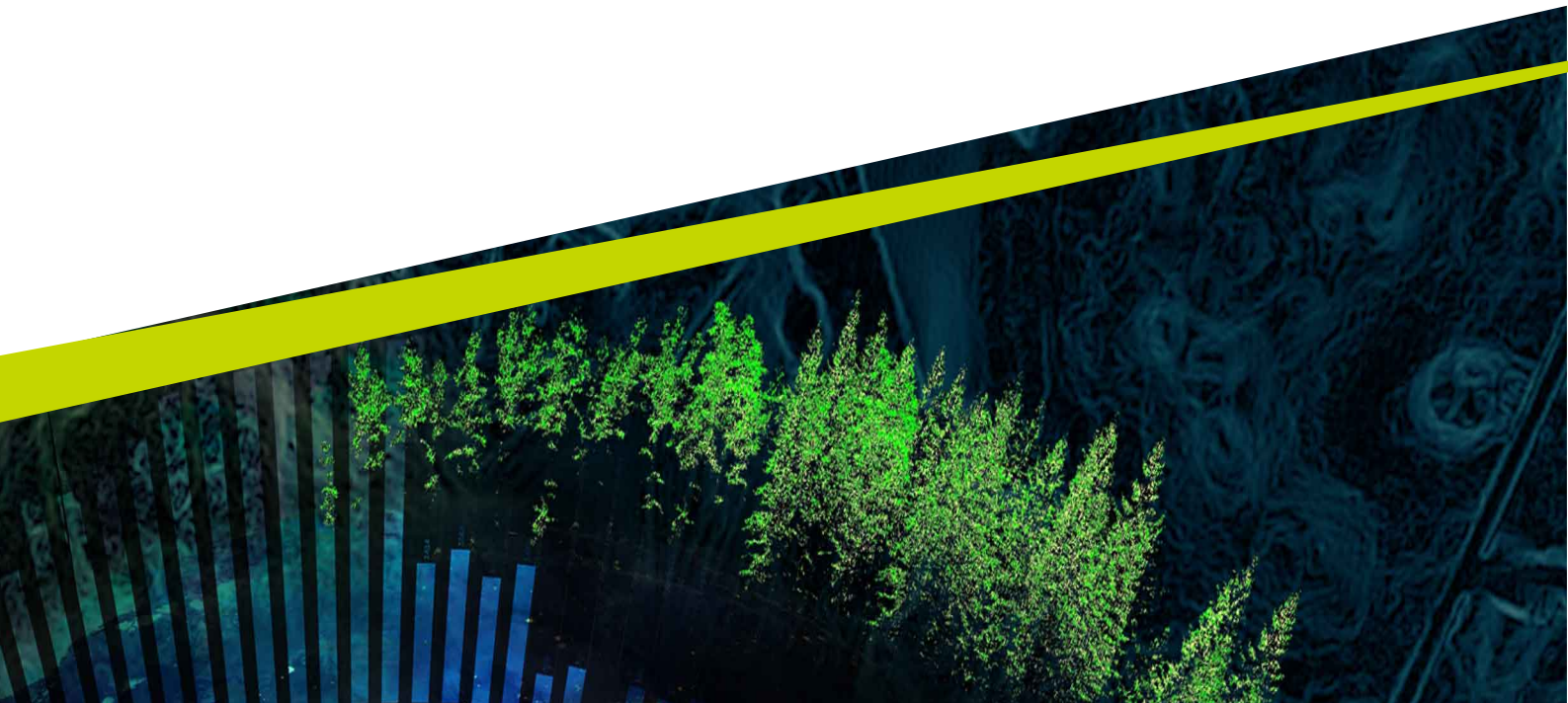
Knockout of transcription factor MYB28 by CRISPR/Cas9 for reducing glucosinolate content in rapeseed (*Brassica napus* L.)

Anusha Harshani Bogahawatta

Independent project in Horticultural Science • 30 credits
Swedish University of Agricultural Sciences, SLU
Department of Plant Breeding
Master's Programme in Horticultural Science
Alnarp, 2023



Knockout of transcription factor MYB28 by CRISPR/Cas9 for reducing glucosinolate content in rapeseed (*Brassica napus* L.)

Anusha Harshani Bogahawatta

Supervisor: Selvaraju Kanagarajan, SLU, Plant Breeding
Co-supervisor: Oliver Moss, SLU, Department of Plant Breeding
Examiner: Li-Hua Zhu, SLU, Department of Plant Breeding

Credits: 30
Level: Second cycle
Course Title: Independent Project in Horticultural Science
Course code: EX0948
Program: Horticultural Science – Master's Programme
Course coordinating dept: Department of Plant Breeding
Place of publication: Alnarp
Year of publication: 2023

Keywords: CRISPR/Cas9, Glucosinolates, *MYB28*, protoplast, rapeseed, ribonucleoprotein, transcription factor

Swedish University of Agricultural Sciences
Faculty of Landscape Architecture, Horticulture and Crop Production Sciences
Department of Plant Breeding

Abstract

Rapeseed cake is a protein-rich by-product from the rapeseed oil extraction process. It contains high levels of glucosinolates (GSLs), which possess anti-nutritional functions after being hydrolyzed. Moreover, the hydrolyzed products of GSLs produce an extremely bitter taste. These adverse properties hinder the use of seedcake for human consumption and restrict its use in livestock feed manufacturing. The GSLs, especially aliphatic GSLs, significantly contribute to importing adverse properties compared to aromatic and indole GSLs. Thus, this study focused on partially down-regulating the aliphatic GSL biosynthesis by inducing targeted mutagenesis in the *MYB28* transcription factor genes using CRISPR/Cas9.

Six out of seven *MYB28* paralogs were identified in the cv. Kumily genome of rapeseed and two sgRNAs were designed at the transcriptional regions for both MYB-type HTH domains of MYB28. The ribonucleoprotein (RNP), in which sgRNAs and Cas9 are mixed as a complex, was introduced into the rapeseed protoplasts. The transfected protoplasts were maintained for callus and shoot formation under in vitro conditions. For screening of mutation lines, genomic DNA was extracted from the true leaves of three regenerated shoots and seven randomly selected callus samples. The sgRNA target sites of five paralogs were amplified using gene-specific primers and screened for the presence of mutations. Screening results showed a single nucleotide insertion in one *MYB28* paralog, LOC106428039, in one callus. This confirms that at least one selected sgRNA was successful inducing mutation in *MYB28* using the RNA-based CRISPR-edited method.

Keywords: CRISPR/Cas9, Glucosinolates, *MYB28*, protoplast, rapeseed, ribonucleoprotein, transcription factor

Table of contents

List of Tables

List of Figures

Abbreviations

CRISPR	Clustered regularly interspaced short palindromic repeat
crRNA	CRISPR RNA
DSB	Double-strand break
GFP	Green fluorescent protein
GSL	Glucosinolate
HDR	Homology-directed repair
HTH	Helix-turn-helix
ITC	Isothiocyanates
MYB	Myeloblastosis
NHEJ	Nonhomologous end joining
ORF	Open reading frame
PAM	Protospacer adjacent motif
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
RNP	Ribonucleoprotein
RT	Room temperature
sgRNA	single guide RNA
TF	Transcription factor
tracrRNA	Trans-activating CRISPR RNA

1. Introduction

1.1 Background

The world rapeseed production accounted for about 71 million metric tons, the second largest in 2002. European Union is the major rapeseed producer (contribution- Germany 25%, France 24%, Poland 11%), followed by Canada, China, and India (USDA, 2022). The oil extraction process generates a high-quality protein by-product (up to 50% protein on a dry basis) called seed cake (from cold-pressed oil extraction) or seed meal (from solvent extraction). Nevertheless, it is a wholesome source of carbohydrates (approximately 20% (w/w)) and minerals. Additionally, it has a higher protein efficiency ratio (2.64) than the widely used soya bean cake (2.19) (Aider & Barbana, 2011). Hence, oilseed rape cake is long identified as a potential raw material for human food and animal feed production.

However, the main hurdle of using oilseed rape cake is the presence of anti-nutritional compounds such as glucosinolates (GSLs), phytic acid, sinapine, and tannins. Among them, GSLs are identified as a major factor that generates undesirable taste and/or reduces nutrient bioavailability. The hydrolyzed products of aliphatic GSLs can act as goitrogens that directly affect the iodine uptake by the thyroid gland. This could eventually cause reduced weight gain and the growth rate of different farm animals (Eskin & Przybylski, 2003). The extreme bitterness is a major obstacle to including rapeseed cake for human consumption. Only a defined quantity is used for livestock feed manufacturing (dairy and poultry). The remaining is applied directly as soil fertilizer or amendment during composting (Nega, 2018). Therefore, improving seedcake/meal quality is one of the major goals in rapeseed breeding.

Many researchers have been trying to develop lower GSL rapeseed cultivars (less than 20 μmol g^{-1} seed GSLs) using conventional and molecular breeding methods, including genome editing. Much interest was drawn to identifying and eliminating problematic GSLs in this context. According to Cartea & Velasco (2008), the aliphatic GSLs and their degradation products significantly affect the development of undesirable flavor and anti-nutritional properties. Hence, downregulating the

aliphatic GSL biosynthesis has the prospect of developing a novel low GSL cultivar, potentially eliminating the undesirable qualities of rapeseed cake/meal without significantly compromising the plant defense mechanisms.

Arabidopsis *(Arabidopsis thaliana)* is the simplest member of the Brassicaceae family and is often used as a model plant to study the *B. napus* genome (Wittstock and Halkier, 2002). The levels of aliphatic GSL in Arabidopsis are transcriptionally regulated by *AtMYB28, AtMYB29*, and *AtMYB76* of sub-group-12 of *R2R3-MYB* transcription factor (TF) family genes (Hirai et al., 2007, Gigo et al., 2009, Sønderby et al., 2010). *AtMYB28* principally regulates both long and short-chain aliphatic GSL biosynthesis, while *AtMYB29* and *AtMYB76* regulate short-chain aliphatic GSL biosynthesis (Figure 4) (Sønderby et al., 2010). Hence, a targeted mutation in the *MYB28* genes has a strong prospect of reducing the long and short-chain aliphatic GSL biosynthesis in rapeseed.

In this study, we intend to induce indels in *MYB28* paralogs through CRISPR/Cas9. The CRISPR/Cas9 technique was developed in 2012 for the purpose of gene editing (Jinek et al., 2012). This technology has the ability to insert CRISPR components into the plant cells as DNA or RNA complexes. Compared to DNA, RNA exhibits a short stay (about 48 h) inside cells; hence, the risk of unintended cleavage is minimal (Liang et al., 2015). Similarly, it is a precise and efficient method of inducing targeted mutation. However, the delivery of invitro mixed Cas9-sgRNA (ribonucleoprotein (RNP) complex) can be challenging due to the plant cell's rigid cell wall (Wang et al., 2022). An efficient delivery can be accomplished via polyethylene glycol (PEG) mediated transfection of protoplasts (living cell whose cell wall is enzymatically digested) (Woo et al., 2015). The PEG can make the plasma membrane more permeable to nucleic acids, allowing the RNP complex to enter the nucleus. The protoplast transfection has been successfully conducted by different researchers using base editors (Molla et al., 2020) or prime editing (Lin et al., 2020). However, protoplast regeneration has remained the foremost hurdle in many crops, including rapeseed. The recent availability of an efficient protocol for protoplast regeneration by Li et al., (2021) permits the application of genome editing in rapeseed. Considering the above-discussed possibilities and limitations, we intended to use the RNP-based CRISPR-editing system to generate mutants in *BnMYB28* to reduce the aliphatic GSLs in this study.

1.2 Aim and Objectives

1.2.1 Aim

Partially down-regulate the GSL biosynthesis through knockout of the *BnMYB28* genes in rapeseed using the RNA-based CRISPR editing approach.

1.2.2 Objectives

1. Identify the *MYB28* paralogs in cv. ‘Kumily’.
2. Design gRNA/s targeting all possible *MYB28* paralogs.
3. Develop transgene-free *MYB28* loss of function mutants and identify mutation events in *MYB28* paralogs in all the regenerated plants from the transfected protoplast.

1.3 Sustainability aspects of the project

The world population is projected to reach 9.7 billion by 2050, raising concerns about the imminent risk of food insecurity. Hence, the world is continuously looking for potential sources of food, methods to optimize agricultural production, and conversion of by-products into value-added products. Hence, recovering current underutilized material such as rapeseed cake into a possible food and/or feed source for the future has great importance in achieving sustainability.

Furthermore, the demand for more affordable high-protein food sources is increasing due to the world's increased maternal, neonatal, and postnatal protein malnutrition, as many people are unable to receive the recommended daily dietary allowance for protein (0.8 g kg^{-1} body weight) (Dukhi, 2020). Since rapeseed cake is identified as an affordable, high-quality plant-based protein source, it has better potential to develop as a future plant protein food for human consumption. Moreover, it can become an excellent raw material in livestock feed manufacturing once the undesirable qualities of rapeseed cake are removed.

2. Literature Review

2.1 Rapeseed (*Brassica napus*)

Rapeseed belongs to the Brassicaceae family. It is an amphidiploid (AC genome, n=19) developed through hybridization between diploid parents, *B. rapa* (AA, n=10) and *B. oleracea* (CC, n=9). Depending upon different geographical regions of cultivation, there are three ecotypes: winter, spring, and semi-winter. *B. napus* is closely related to Arabidopsis as both belong to the Brassicaceae family (Parkin et al., 2005).

2.1.1 Global rapeseed production and use

The global rapeseed production reached 88.125 million metric tons in 2022, and the major contributors to global rapeseed production are Europe Union (19.5 million MT), Canada (19.0 million MT) and China (18.5 million MT) (USDA, 2022). Rapeseed crop cultivation is much valued in temperate regions of the world as it is the only commercially viable oilseed crop that can be grown. The Europe-grown rapeseed varieties have low erucic acid (< 2%) and GSLs content (< 30 μmol g^{-1}) and are called “double-zero”, “double-low” or 00-rapeseeds (Spragg and Mailer, 2007).

The rapeseed oil is extracted by cold/ expeller pressing method or solvent extraction method. Cold or expeller-pressed rapeseed oil is called virgin oil, and the remains from the oil extraction process are called ‘rapeseed cake,’ which usually contains 5-20% residual oil (Seneviratne et al., 2010). The solvent extraction process uses chemical solvents such as hexane. It can extract more oil from seeds and the by-product ‘rapeseed meal’ only contains less than 3% residual oil (Adjonu et al., 2019).

According to recent statistics, rapeseed meal is the second largest oilseed meal produced after soybean meal (USDA, 2022). In Sweden, it is mainly used in dairy feed production. It is known as a promising source of protein (30-40%), minerals, and fiber (Table 1). It has a similar amino acid balance to the soybean, including

essential amino acids: lysine, methionine, and tryptophan. Further, it contains high levels of calcium, phosphorus, selenium, iron, and manganese (Georgeta, 2009).

Additionally, it is a potential resource that can be used in diversified ways in different industries, for instance, 1) As an additional nitrogen source and a substrate in mushroom media (Tuli et al., 1985), 2) Soil pesticide and a weed suppression agent (Boydston et al., 2008), 3) Slow-releasing soil amendment for nitrogen, phosphorus, and potassium (Ullah et al., 2008).

The isothiocyanates and other hydrolytic products of GSLs are considered health promoters in a balanced human diet that exerts protective bioactivity against diseases (Dinkova-Kostova & Kostov, 2006) and chemopreventive activity (Jiang et al., 2018). However, several research findings confirmed the potential toxicity and detrimental effects that arose due to the presence of GSLs in their hydrolyzed bioactive form in livestock fed with rapeseed cake. (Velayudhan et al., 2017; Roth-Maier et al., 2004).

2.1.2 Adverse effects of rapeseed cake/meal ingestion for livestock animals

Even though rapeseed cake has better nutrient composition and a higher protein efficiency ratio than other possible protein sources used in the feed industry, the presence of GSL-hydrolyzed products is known to limit the usage of rapeseed cake as a feed ingredient. According to several research findings, the prolonged ingestion of these hydrolyzed products can cause goiter, hemorrhagic liver, less palatability due to bitter taste, and reduced performance in major livestock animals (Schone et al., 2001; Newkirk, 2009). Furthermore, tolerance levels vary depending on the type of farm animal and its growth stage (Table 2). For instance, lactating cows can tolerate up to 24 μmol g^{-1} per diet GSLs; hence, using rapeseed cakes on lactating cow feed is non-problematic (Tripathi and Mishra, 2007). However, the total GSL concentration should not exceed 2 μmol g^{-1} for pig diets due to the risk of goitrogen that inhibits thyroid hormone production (Bell, 1993; Schone et al., 2001).

The aliphatic β-hydroxy alkenyl GSL in rapeseed cake (progoitrin and gluconapoleiferin) can hydrolyze into isothiocyanates that convert into oxazolidine-2-thiones, which act as goitrogens (Mawson et al., 1993). Goitrogens can stimulate the pituitary gland to release elevated levels of thyroid stimulating hormone that accelerate unnecessary thyroid tissue growth. Parallelly, it can inhibit the iodine uptake by thyroid follicular cells, which decreases thyroxin production (Soto-Blanco, 2022). These scenarios can eventually result in a larger thyroid gland and adversely affect the production of thyroid hormones (Bertinato, 2021). Moreover,

thiocyanates and isothiocyanates have an indirect ability to reduce thyroid hormone production by competing with thyroidal transport (Langer and Greer, 1977).

Table 1: Different protein sources used in feeds and their nutritional composition.

Feed type	Chemical composition in %									
	DM	CP	EE	CF	NFE	Ash	NDF	ADF	Ca	TP
Rapeseed cake (Feng, 2003)	88	35.7	7.4	11.4	26.3	7.2	33.3	26.0	0.59	0.96
Rapeseed meal (Feng, 2003)	88	38.6	1.4	11.8	28.9	7.3	20.7	16.8	0.65	1.02
Soybean meal (Maison et al., 2015)	89.98	47.73	1.52				8.21	5.28	0.33	0.71
Canola meal (Maison et al., 2015)	91.33	47.73	3.22				22.64	15.42	0.69	1.08

Note: DM- Dry Matter, CP- Crude Protein, EE- Eher extract, CF- Crude Fibre, NFE- Nitrogen-free extract, NDF- neutral detergent fiber, ADF- Acid detergent fiber, TP- Total phosphorous. (The variations among varieties, climatic conditions, and harvesting conditions may affect the chemical compositions of the seed meal and seed cake).

Table 2: Biological effects on glucosinolate intake on different farm animals and concentrate inclusion rate (%) guideline of oilseed rape meal for different farm animals and their growth stages. (Tripathi and Mishra, 2007; Ewing, 1998).

Farm animal	Glucosinolate levels in feed (μmol g^{-1} per diet)	Biological effect/s	% inclusion guideline
Dairy	11.7-24.3	Depressed feed intake and milk production	Calf- 5%, lactating cow- 25%
	31.0	Thyroid disturbance and depressed fertility	
Poultry	5.4-11.6	No adverse effect on intake and weight gain	Broiler and layer- 2.5%
	34.0	Severe growth depression	
Pig	1.3-2.79	Reduced feed intake and growth	Creep and weaner- 0%, Grower- 2.5%, Finisher- 5%
	7.0	Severe growth depression	
	0.16-0.78	No adverse effect during growth , pregnancy and lactation	

2.2 Glucosinolates

Intact GSLs are amino acid-derived thioglycosidic secondary metabolites (β-glucoside), abundant in Brassicales order and considered a stable, non-toxic compound in its intact form. GSLs degradative products contribute to several bioactivities, such as plant defense, flavor development (Bell et al., 2018), and chemoprevention (Becker and Juvik, 2016). Among them, the major contribution of GSLs is to protect the plant from herbivores and non-adapted pathogens. Hence, a higher concentration of GSLs is often found in the organs that contribute most to plant fitness, such as seeds - the commercially important part of the rapeseed crop (Brown et al., 2003).

2.2.1 Molecular structure

GSLs are stored in glucoside form in plants, facilitating better solubility, stability, and transport. The general chemical structure (Figure 1) comprises a β-D-glucosyl residue linked to an ester through sulfur and an amino acid-derived R group (Barba et al., 2016). Depending on the precursor amino acid in the R chain, the resulting glucosinolate can be aliphatic (methionine, alanine, leucine, isoleucine, methionine, or valine), aromatic (phenylalanine or tyrosine), or indole (tryptophan) (Redovniković et al., 2008).

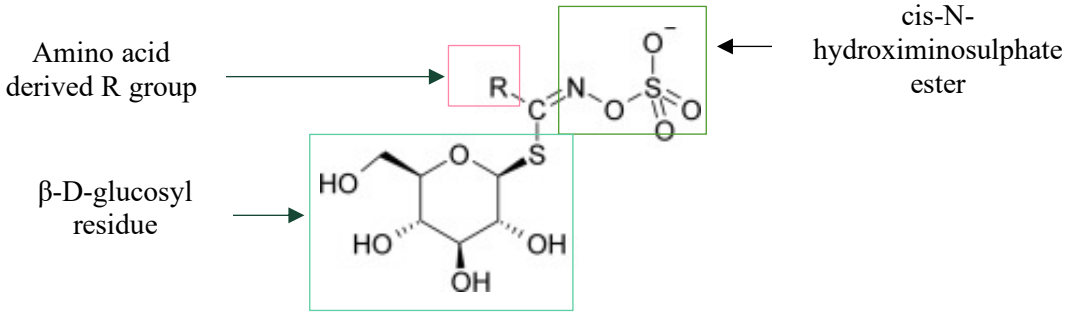


Figure 1: Generic structure diagram of a GSL (The side group R varies according to the type of GSL synthesized) (Figure source: Redovniković et al., 2008).

2.2.2 Aliphatic GSLs in rapeseed

The aliphatic GSLs are identified by their precursor amino acids. The most abundant aliphatic GSLs are derived from methionine. Depending on the length of the carbon side chain, the aliphatic GSLs can be either long-chain (6C-8C) or short-chain (3C-5C) (Table 3). Different rapeseed cultivars can express varying proportions of aliphatic GSL. Furthermore, the content and composition of aliphatic

GSL profiles of the plant can vary according to environmental conditions (Textor et al., 2007).

Table 3: A few major short-chain aliphatic GSLs identified in B. napus, their trival name, and length (Mithen et al., 1995).

Systematic name	**Trivial name**	**Length of C side chain**
1-Methylethyl	Glucoputranjivin	C3
3-Butenyl	Gluconapin	C4
2-Hydroxy-3-butenyl	Progoitrin	C4
4-Methylthiobutyl	Glucoerucin	C4
4-Methylsulfinylbutyl	Glucoraphanin	C4
4-Methylsulfinyl-3butenyl	Glucoraphenin	C4
5-Methylsulfinylpentyl	Glucoalyssin	C5
4-Pentenyl	Glucobrassicanapin	C5
2-Hydroxy-4-pentenyl	Gluconapoleiferin	C5

2.2.3 Aliphatic GSL activation and its bioactive products

Naturally, GSLs are activated by the hydrolysis of thioglucosidic bonds to bioactive aglucone forms. The process is catalyzed by myrosinase (β -thioglucosidase glucohydrolase), which is stored in myrosin cells. The myrosinase comes into contact with GSLs as a result of cell disruption. The aglucon can be converted into different products such as isothiocyanates (ITCs), thiocyanates, nitriles, epithionitriles, hydroxy nitriles, oxazolidine-2-thiones, or indoles, depending on many factors such as side chain structure of parent GSL, the hydrolysis conditions (pH, temperature), presence of Fe^{2+} ions, and additional protein factors (Rungapamestry et al., 2006) (Figure 2). These ITCs are identified as the most effective against plant pathogens and pests (Piekarska et al., 2014).

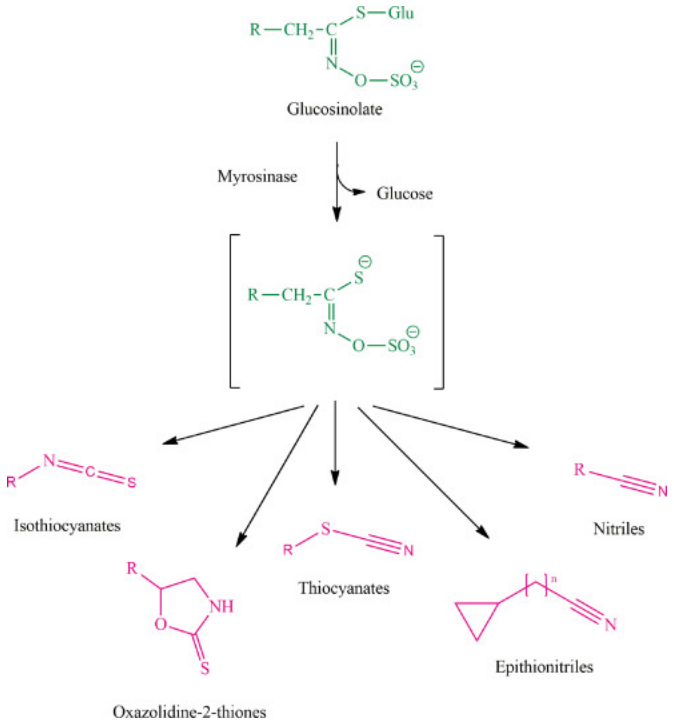


Figure 2: Hydrolysis of glucosinolate and different bioactive products (Figure source: Redovniković et al., 2008).

Oxazolidine-2-thiones Several aliphatic GSLs (progoitrin and gluconapoleiferin) consist of a hydroxylated C-side chain at carbon 2. Their hydrolyzed products can be converted into oxazolidine-2-thiones, which express goitrogenic effects on the physiology of several livestock animals (Prieto et al., 2019).

Thiocyanates and Isothiocyanates Few aliphatic GSLs, such as glucoraphanin and singrin (in cabbage), can be hydrolyzed into ITC. Allyl-, benzyl-, and 4-methylthiobutyl GSL can hydrolyze into thiocyanates in the presence of thiocyanate-forming proteins (Eisenschmidt-Bönn et al., 2019). They support metabolic enzymes in the liver and epithelial cells to balance the metabolism of carcinogenic compounds (Moreb et al., 2020).

Nitriles and Epithionitriles The nitrile-specific proteins in the plant can promote nitrile formation during GSL hydrolysis, irrespective of the side chain structure of the parent GSL. A terminally unsaturated side chain in parent GSL can allow the epithiospecific proteins to produce epithinitriles during GSL hydrolysis (Eisenschmidt-Bönn et al., 2019).

2.2.4 Impact of aliphatic GSLs on plant defense

Generally, the presence of GSLs in *Brassica* crops is known to express an inhibitory effect over many pests and pathogens. Different studies illustrated the ability to control soil pathogenic fungi and herbivore attacks by isothiocyanates- the hydrolysis products of GSLs (Ishimoto et al., 2000). Transversely, some herbivores (*Brevicoryne brassicae*, *Psylliodes chrysocephala,* and *Pieris rapae*) use GSL as attractants (Bruce, 2014). However, the overall plant defense is a rather complex scenario that does not depend solely on GSL but also on other genetic controls, environmental conditions, species involved, and agronomical conditions (Redovniković et al., 2008).

A few studies suggest that the aromatic and indole GSLs have a major role in inhibitory effects over pathogen attacks on *Brassica* crops, compared to aliphatic GSLs. According to Manici et al., (1997) and Sarwar et al., (1998), the isothiocyanates produced from aromatic GSL hydrolysis were more toxic than the aliphatic GSL hydrolysis products. The results from Doughty et al., (1991) and Li et al., (1999) also showed a similar behavior when observing the glucosinolate profile changes of pathogen-inoculated *B. napus* cultivars. Upon inoculation, the GSL profile changes were mostly due to increased aromatic and indole GSLs. Further, the exogenous application of jasmonic acid (a signaling molecule to activate plant defense) exhibited a significant increase in indole GSLs and unchanged aliphatic GSLs (Oughty et al., 1995). However, an observation from research by An Dam et al., (2004) indicated an increased level of aliphatic GSL in jasmonic acid-treated *B. oleracea* roots. On the contrary, the exact impact of aliphatic GSL on overall plant defense is complicated and yet to be discovered.

2.3 Aliphatic glucosinolate biosynthesis

Generally, the GSL biosynthesis pathway consists of three steps: 1. Side chain elongation 2. Core structure biosynthesis 3. Side chain modification. The process mainly occurs in cellular cytosol, while several steps occur in chloroplasts as enzymes such as MAMS, IPMI, and IPMDH are localized in chloroplasts (Falk, 2004). The side chain elongation is the major contributor to the production of diverse GSLs and the diversion of amino acid flux from primary and secondary metabolism (Hansen et al., 2007).

As the first step of aliphatic GSL biosynthesis, the methionine is de-aminated by aminotransferase 4 (BCAT4) for producing 2-oxo acid. The aliphatic chain of 2-oxo acid elongates by three main enzymatic catalyzers: 1. Methyl thioalkyl malate

synthase (MAMS), isopropyl malate isomerase (IPMI), and isopropyl malate dehydrogenase (IPMDH). The elongated 2-oxo acid is trans-aminated by BCAT3 and enters the core GSL assembly pathway (Knill et al., 2007).

All the elongated amino acids for all different types of GSLs undergo core structure formation in the cytosol. They convert into aldoximes by cytochrome P450 enzymes (CYP family) such as CYP79F1- all elongated methionine-derived amino acids, and CYP79F2- only long-chain methionine-derived amino acids (Glawischnig, 2006). The resulting aldoximes are oxidized to nitrile oxides or aci-nitro compounds by CYP83 cytochrome P450s. These products then conjugate with glutathione (the sulfur donor). The reaction is catalyzed by glutathione-S-transferases (GST) and produces S-alkyl-thiohydroximates. Carbon-sulfur lyase (SUR1) enzyme then transforms this S-alkyl-thiohydroximates into thiohydroximates, and it undergoes S-glycosylation catalyzed by glucosyltransferases, producing desulfoglucosinolates (Douglas Grubb et al., 2004). A set of sulfotransferases (SOTs) sulfate the desulfoglucosinolates to produce glucosinoate molecules, which may go through side chain modifications such as oxidation, hydroxylation, methoxylation, desaturation, sulfation, and glycosylation as per the requirement (Figure 3) (Kitainda & Jez, 2021).

A large number of genes involved in the aliphatic GSL biosynthetic pathway have been identified and their AGI codes are listed in Appendix 2.

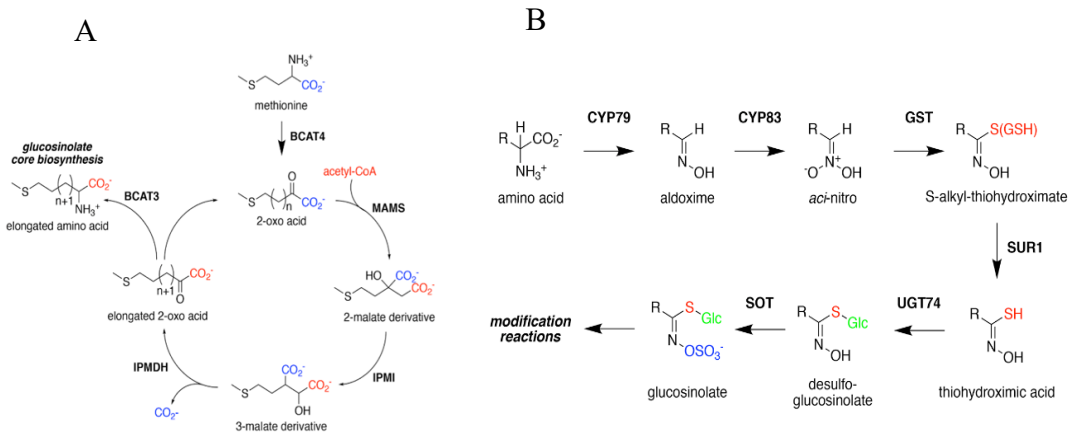


Figure 3: Schematic overview of aliphatic GSL Chain elongation (A) and GSL core structure biosynthesis pathways (B) and major enzymes involved. (Figure source: Kitainda and Jez, 2021).

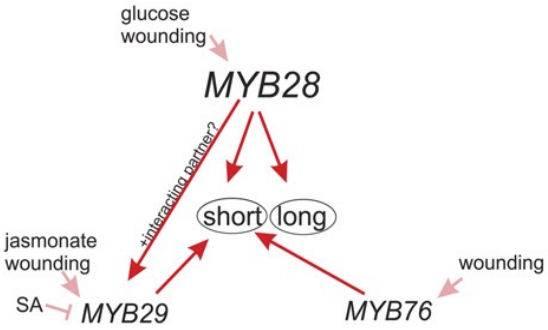
2.3.1 Transcription factor-based regulation of aliphatic GSL biosynthesis

Typically, secondary metabolite biosynthesis is considered a 'biochemically expensive' process, for instance, according to Bekaert et al., (2012), GSL biosynthesis has been estimated to increase the photosynthetic requirements by 15% in Arabidopsis. Hence, the plant generally produces GSL only for special defense requirements. This regulation occurs primarily at the transcriptional level, and several activator and/or repressor-type TFs are involved in this process (Reményi et al., 2004). These TF proteins bind to the DNA binding domain of a particular gene to up or down-regulate gene expression levels (Frerigmann, 2016). Several studies were able to identify different TFs involved in GSL biosynthesis, such as *IQD1-1* (Barda, 2022), *MYB28, MYB29* (Hirai et al., 2007), *MYB34* (Celenza et al., 2005), *MYB51, MYB76, MYB122* (Gigolashvili, Berger, et al., 2007).

The *R2R3-type MYB* TF family comprises six homologous (Table 4) that directly control the aliphatic and indole GSLs biosynthesis. Among them, *MYB28*, *MYB29*, and *MYB76* regulate the expression levels of aliphatic GSL biosynthetic genes. These three TFs form a regulatory network to control the aliphatic GSL profile (Figure 4). According to the transcriptomic analysis done by Hirai et al., (2007), the co-expression of *MYB28* was highest with aliphatic GSL biosynthetic genes in Arabidopsis. Later, Sønderby et al., (2010) documented *AtMYB28* as the principal regulator for long and short-chain aliphatic GSL biosynthesis and *AtMYB29* and *AtMYB76* as the regulators for the short-chain aliphatic GSL biosynthesis. Few studies confirm the extent of aliphatic GSL regulation by the principle regulator *MYB28* in different crops, such as *B. juncea* (Augustine et al., 2013) and *B. oleracea* (Neequaye et al., 2021). Augustine et al., (2013) achieved 89% seed GSL reduction and significant suppression of aliphatic GSL by transgene-based RNAi suppression strategy on *MYB28* in *B. juncea*. Another recent study in *B. napus* (winter rapeseed) confirms 55.3% decreased progoitrin level and 16.2 µmol g^{-1} total GSL levels from TILLING populations (mutations were induced at *MYB28* TF by chemical treatment of ethyl methanesulfonate) (Jhingan et al., 2023).

Table 4: R2R3-MYB TF *family homologous and their functions*

TF	Function	Reference
MYB34 *MYB51* *MYB122*	Regulates the tryptophan-derived indole GSL biosynthesis pathway	Frerigmann and Gigolashvili, 2014
MYB28 *MYB29* *MYB76*	Regulates the methionine-derived aliphatic GSL biosynthesis pathway	Gigolashvili et al., 2008, Sonderby et al., 2010



Single arrowheads towards circles- increasing effect on metabolites, Red single arrowheads- induction of the genes in moderate overexpression lines, Pink single arrowheads- the impact of various treatments in wild-type, Pink T-line- repression by salicylic acid.

Figure 4: Model for actual regulation of GSLs by MYB TF *in Arabidopsis (Figure source: Sønderby et al,. 2010).*

2.3.2 Structural and functional characteristics of MYB28 protein

The MYB TF protein usually contains a few (1 to 4 imperfect repeats) conserved DNA-binding sites called MYB domains. Generally, these domains can be classified into three groups: 1. MYB-type HTH domain: protein-DNA interaction, 2. SANT domain: protein-protein interaction, 3. MYB-like domain: either of the above two interactions.

Each MYB-type domain is characterized by 50-53 amino acids that encode three α-helices. The second and third helices can form the helix-turn-helix (HTH) structure, which allows the insertion of MYB TF into the target DNA's major groove. Upon successful insertion, the TF can regulate the expression of the target gene (Cao et al., 2020). According to the Uniport database, the MYB28 protein consists of 139 amino acids with two MYB-type HTH domains (Table 5).

Table 5: Major HTH structures and their positions of MYB28 *protein (Source: Uniport)*

ID position	Description	Amino acid sequence
9-65	HTH-MYB-type	GEGLKKGAWTTEEDKKLISYIHEHGEGGWRDIPQKAGLKRCGKSCRLRWTNYLKPEI
66-116	HTH-MYB-type	KRGEFSSEEEQIIIMLHASRGNKWSVIARHLPRRTDNEIKNYWNTHLKKRL

According to the TAIR database for Arabidopsis, the *MYB28 TF* is responsible for regulating several biological processes such as cellular response to sulfur starvation, defense response to fungal, bacterial, and insect attacks, induced systemic resistance, regulation of DNA-templated transcription, and regulation of glucosinolate biosynthetic process. It is expressed in many parts of a plant and abundant in mature leaf petioles and leaves, hypocotyls, and different parts of the flowers, seeds, and roots (Figure 5) during various growth and developmental stages.

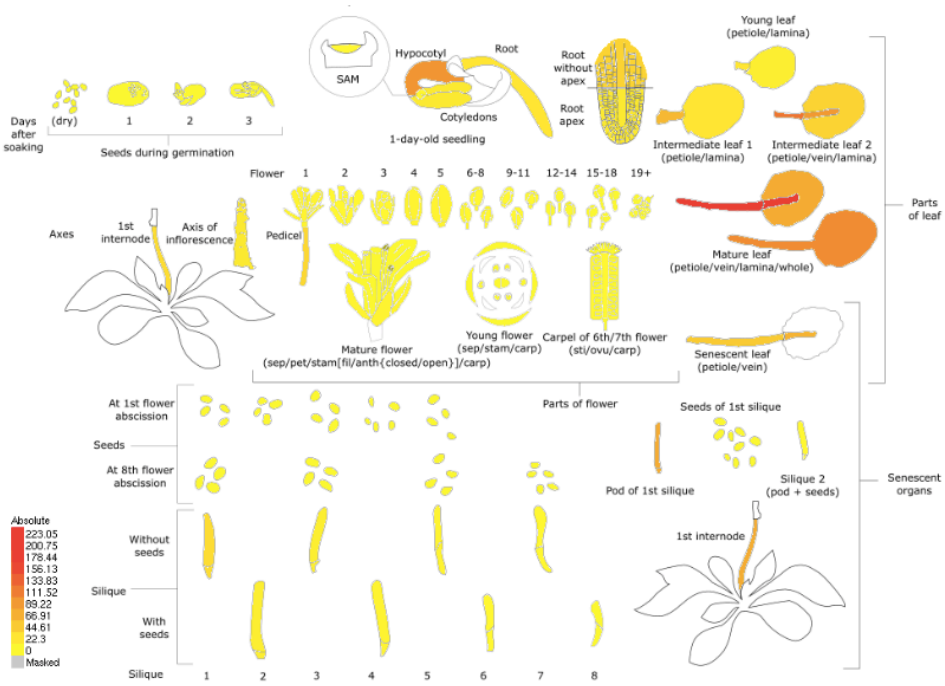


Figure 5: Transcriptome-based RNA seq profiling developed for MYB28 TF *in Arabidopsis (Figure source- TAIR).*

2.3.3 *MYB28* genes in rapeseed

The MYB28 protein translation is determined by the *MYB28* protein-coding genes in the *B. napus* genome. According to the NCBI database, seven paralogs that can transcribe *MYB28* mRNA have already been identified from cv. Da-Ae reference genome (NCBI reference seq assembly- GCF_020379485.1) (Table 6). See appendix 1, a. and b. for nucleotide similarity and protein identity of *MYB28* paralogs in *B. napus* cv. Da-Ae.

Table 6: MYB28 *paralogs, locations, and gene IDs in cv. Da-Ae, according to the NCBI database.*

Locus number	Gene ID (NCBI database)	Location (Chromo)	Genomic sequence length (bp)	Number of exons	Coding region (bp)
LOC106382746	106382746	C2	1564	3	1065
LOC106382207	106382207	C9	1399	3	1089
LOC106428039	106428039	C7	1376	3	1068
LOC125577744	125577744	A9	1324	3	1074
LOC106439923	106439923	A3	1350	3	1065
LOC106431683	106431683	Unknown	462	2	390
LOC106369912	106369912	Unknown	5306	3	759

2.4 Genome Editing using CRISPR/Cas9 technique

Compared with zinc finger nucleases, transcription activator-like effector nucleases, and RNA-guided engineered nucleases, the clustered regularly interspaced short palindromic repeat (CRISPR)-associated (Cas) system is an effective and affordable tool used to obtain precise genome editing in scientific research. Different types of CRISPR/Cas systems are available based on the distinctive composition of expression, interference, and adaptation modules of Cas proteins. However, class 2, type II Cas 9 from *Streptococcus pyogenes,* is the most widely used and studied (Makarova et al., 2022).

CRISPR/Cas gene editing can be done through DNA or RNA-based approaches. The DNA-based approach uses expression plasmids or viral vectors to deliver CRISPR components into plant cells. However, plasmids were problematic due to possible 'stable transformation events' (random integration of the plasmid T-DNA into the host genome). This may cause some undesired off-target site mutations, as reported by Cradick et al., (2013) and Fu et al., (2013). Some off-targets might be challenging to detect and more problematic due to their ability to alter host immune responses (Hemmi et al., 2000). Viral vectors are concerned with safety due to their ability to create potential immunogenic responses (Xu et al., 2019). During RNA delivery, in-vitro transcribed Cas9-mRNA and sgRNA are co-transfected and upon successful entry to the cell, the mRNA is translated to a Cas9 protein and combine with sgRNA to form the RNP complex. A pre-mixed Cas9 and sgRNA was used to transfect cells during RNP delivery (Scott et al., 2019).

The RNP delivery uses a sgRNA that is formed by hybridizing trans-activating CRISPR RNA (tracrRNA) and CRISPR RNA (crRNA). The crRNA is also called a protospacer and is composed of about 20 nucleotides complementary to the target

region of DNA. The tracrRNA binds to the Cas9 nuclease and forms the RNP complex (Garneau et al., 2010). Cas 9 protein is known as the molecular scissor. It has nuclease activity and is able to cleave DNA strands. The sgRNA guides the Cas9 protein to the target site. If the Protospacer Adjacent Motif (PAM) site is located directly next to the protospacer in an organism's genome, Cas protein can recognize the PAM sequence (Cas9 -5'-NGG, PAM site) and initiate the cleavage. Once the Cas protein binds to a PAM motif the DNA double-strand seperates and the crRNA spacer can bind to the target site if spacer sequences are matching. Cas protein is then able to cut the target DNA double strand at approximately 3bp upstream of the PAM sequence (Jinek et al., 2012).

Once a double-strand break (DSB) appears in the plant genome, the plant immediately tries to repair the damage. The repairing mechanism's nature is identified as nonhomologous end joining (NHEJ) or homology-directed repair (HDR). NHEJ usually results in indels (insertions or deletions) in the target site as the process is error-prone. These errors can terminate the gene's function and knock out the gene of interest from the plant genome. A homologous sequence to the DSB (template DNA) is used to repair the DSB in the HDR repairing mechanism (Zhang et al., 2021).

3. Methodology

3.1 Plant material and in-vitro culture conditions

The starting plant material (seeds) from spring rapeseed cv. Kumily was obtained from Lantmännen, Svalöv, Sweden. The seeds were surface-sterilized using 70% ethanol for 15 min followed by gentle shaking in kitchen bleach (20% (w/v) for 15 min, and then thorough sterile water rinse/s. Surface-sterilized seeds were dried in a sterile hood and planted on half-strength Murashige & Skoog medium, supplemented with 20 g L^{-1} sugar and 7 g L^{-1} gelrite, pH 5.7.

All the in-vitro cultures in this study were grown in a climate-controlled chamber with a temperature of 23°C/18°C (day/night) and a 16 h photoperiod with a light intensity of 40 µmol m^{-2} s^{-1}. The light levels were controlled for protoplasts and early stages of callus cultures using a cotton cloth.

3.2 Identification of *MYB28* paralogs

BnMYB28 paralogs from the NCBI database (Table 7) were amplified by PCR using the genomic DNA from cv. Kumily as the template, according to the protocol by Muthusamy et al., (2020) with minor modifications.

3.2.1 Primer Design and PCR

Gene-specific forward and reverse primers were designed for each reference gene (using Genious prime or NCBI primer blast or manually) to amplify each full-length paralog during PCR. The PCR product was visualized by gel electrophoresis (1% or 1.5% (w/v) agarose gel). The bands with the expected amplicon size were isolated and then purified using the NucleoSpin TriPrep mini kit for RNA, DNA, and protein purification.

3.2.2 Cloning and DNA Sequencing

The purified DNA segments were ligated into a linearized plasmid vector pJET1.2 (Thermo Fischer Scientific CloneJET PCR Cloning Kit #K1232) containing a lethal restriction enzyme gene that disrupts upon successful DNA fragment insertion into the cloning site.

The recombinant plasmids were introduced into the competent cells of *E. coli* HST08 strain according to the Steller ™ competent cells protocol PT5055-2. The recombinant competent cells were multiplied in 500 µl liquid SOC media and then spread-plated on a selective medium (carbendazim 5 mg ml^{-1}). Upon 12-h incubation, the single colonies were isolated, and the plasmids were purified using the GeneJET genomic DNA purification kit. The eluted DNA samples were confirmed by Sanger sequencing through Eurofins Genomics, Germany (https://eurofinsgenomics.eu).

3.3 Designing sgRNAs

The coding regions of Kumily *MYB28* paralogs were aligned to identify conserved CRISPR target sites using the Genious Prime® 2023.0.3 (http://www.geneious.com/). Two potential CRISPR target sites (20 bp) were identified from exons 1 and 2 to cover six paralogs from the core MYB protein translation regions. The off-target potential of each target site was checked over the NCBI database and by screening over an available *B. napus* whole genome sequence (accession number: JMKK00000000) through Geneious Prime. The two selected sgRNAs and Cas9 protein were purchased from Synthego, USA, and pre-assembled according to the manufacturer's instructions prior to protoplast transfections.

3.4 Protoplast transfection

The intact protoplasts were isolated from leaves and transfected with the RNP complex according to the protocol by Li et al., (2021), as elucidated below.

3.4.1 Protoplast isolation

In-vitro grown 3-4 weeks old seedlings were used for protoplast isolation. About 40 fully opened first true leaves were selected and sliced into 0.5–1 mm strips under sterile conditions and incubated in hypertonic plasmolysis solution (0.4 M mannitol at pH 5.7) for 30 min at room temperature (RT) in the dark to obtain stable osmotic environment for the protoplast cells.

Then 10 ml of enzyme solution (1.5% cellulose, 0.6% macroenzyme, 0.4 M mannitol,10 mM MES, 0.1% BSA, 1 mM $CaCl_2$, and 1 mM β-mercaptoethanol at pH 5.7) was added and incubated for 14-16 h at RT in the dark while shaking to digest the cell wall contents.

The digestion was then filtered through a 40 μm nylon cell strainer into a 50 ml Falcon tube to collect the protoplasts. The filtrate was diluted with 30 ml W5 solution and centrifuged at 100 g for 10 min. The resulting protoplast pellet underwent a washing step (the pellet was re-suspended in 10 ml W5 solution, centrifuged at 100 g for 5 min, and removed the supernatant) twice with W5 solution. Subsequently, the pellet was re-suspended in 5 ml W5 solution and ice-incubated for 30 min in the dark to collect the intact protoplasts. Finally, the supernatant was carefully discarded, and the protoplasts were diluted with 5–10 ml W5 solution.

3.4.2 Protoplast quantification

The hemocytometer was loaded with 15 μl of the protoplast suspension from 3.4.1 and counted under a light microscope. Depending on the number of protoplasts counted, the density of the original protoplast suspension was adjusted to 400,000 to 600,000 per ml using MMG (0.5 M mannitol, 750 μl $MgCl_2$, 0.2 M MES, H_2O) solution.

3.4.3 Approximation of transfection efficiency

About 120,000 protoplasts were used for transfection efficiency approximation. The protoplasts were re-suspended in 200 μl freshly prepared MMG solution (0.5 M mannitol, 15 mM $MgCl_2$, 4 mM MES) in a 2 ml Eppendorf tube. The vector pCW498-35S-GFiP-OcsT (14 743 bp) harboring the green fluorescent protein gene (*GFP*) was used as the indicator that can confirm the successful uptake of the vector DNA by the protoplasts. The solution was mixed with 40 μl vector DNA and 240 μl freshly prepared PEG-calcium solution (25% (w/v) PEG 4000, 0.5 M mannitol, and 0.1 M $CaCl_2$). After 5 min of incubation, the reaction was terminated by adding 1.5 ml W5, followed by gentle mixing. The suspension was centrifuged at 100 g for 3 min, and the supernatant was removed. The transfected protoplasts were re-suspended in 1 ml of MI medium and kept in the dark for 48 h.

Upon two-day recovery, the protoplasts were observed with a Zeiss LSM 880 Airyscan confocal laser scanning microscope using an EC-Plan-Neofluar 10x/0.30 M27 objective.

3.4.4 Assembling the RNP complex

The RNP complex was assembled by mixing 4 µl of sgRNA (0.1 nmol μl^{-1}) with 4 µl Cas9 (5 µg μl^{-1}) and 12 µl H_2O. The mixture was then incubated in the dark for at least 15 min at RT.

3.4.5 Protoplast transfection

A sample of 200 µl protoplast MMG mixture was used per transfection. The protoplast sample was mixed with 20 µl of RNP complex and 220 µl of freshly prepared PEG-calcium solution (25% (w/v)). The reaction was terminated after 6-min incubation by adding 1.5 ml W5 solution. The Eppendorf tube was inverted several times to mix the protoplast mixture with the W5 solution thoroughly. Then, the tube was centrifuged at 100 g for 3 min, and all the supernatant was carefully removed. The protoplast pellet was resuspended in 200 µl MI solution. Thereafter, 200 µl of the alginate (solidifying polysaccharide) was added and mixed gently. The mixture was carefully poured on top of the Ca-agar plate to allow polymerization and solidification of alginate for 30 min. Then, 2 ml Ca solution was poured onto the solidified alginate gel disk and incubated at RT for 1 h. Afterward, each gel disk was carefully transferred into 6 well-cell culture plates containing 3 ml of MI (0.5 mg L^{-1} 2,4-D and 0.5 mg L^{-1} NAA) solution and kept in the dark for 24 h.

3.5 Protoplast recovery and organogenesis

The medium was replaced by MII after three days of transfection. Medium was changed every 3-5 days until visible calli appeared. The gel disks with visible calli were spread on MIII agar plates to facilitate further growth. The regenerated shoots were transferred to the shooting medium. Later, the shoots were transferred to the rooting media for root induction and formation.

3.6 Identifying mutation events

The mutation events were checked in 10 samples (3 regenerated shoots and 7 callus samples). The shoot generation from calli required extra time. Considering time restrictions, the cells from randomly selected calli samples were also considered for mutation identification.

A preliminary screening was done to observe that indels had occurred in the target sites. DNA was extracted from leaf or callus tissues using a Thermo Scientific Phire

Plant Direct PCR Kit. A PCR was performed using confirmed forward and reverse primers designed to amplify the target area. The PCR products were purified and sequenced (Sanger) through Eurofins Genomics, Germany.

4. Results

4.1 Identification of the *MYB28* paralogs in cv. 'Kumily'

Full-length open reading frames (ORFs) of LOC106428039, and LOC106382207 and partial sequences for LOC106369912, and LOC106439923 were obtained through PCR using degenerative forward primer and locus-specific reverse primers.

Locus-specific primers were designed to amplify LOC106382746 (Table 7). The missing region of LOC106369912 (204bp) was identified by using GGCCACTTCCATCAAAGATA forward primer and GCTCCGACATCAATCCATTTCC reverse primer (5'-3'). The amplified PCR products were ligated into a cloning vector, purified, and sequenced to obtain the nucleotide sequence of each from cv. Kumily genome (See appendix 3 for nucleotide sequences).

Table 7: Primers used to identify nucleotide sequences from the cv. Kumily genome through sequencing.

Locus number	Gene size (bp) of the amplified region-Kumily	Nature of the amplified region	Primers used (5'-3')	Full gene size (bp)
LOC106428039	1376	Full length	F-ATGTCAAGAAARCCRTGTTGTG	1376
LOC106382207	730	Full length	R-TCATATGATTTGCTTATCGAAG R-TCAGAGGGAATCAGAATCC	730
LOC106439923	1089	Partial	R-TCATATGAGGGWATCAGACTCC	1337
LOC106369912	1438	Partial		1642
LOC106382746	1565	Full length	F-TCCTTTGGTGTTGTTTGAACCT R- ATGAGTTCTCGCCTTCTCGG	1565

The LOC106382207 nucleotides continually exhibited higher exposure to the primers used to amplify LOC125577744. Hence, the full-length ORF for LOC125577744 (644bp) and a region (partial sequence) of LOC106439923

(248bp) were identified using the database sequences for cv. Kumily whole-genome Illumina sequencing done by Oneomics Private Limited, India.

The NCBI database data for cv. Da-Ae, LOC106431683 was short (462 bp) and comprised two exons. Further, it shared 92.7% nucleotide similarity to LOC106382746. The exon 2 of LOC106431683 was 99.7% similar to LOC106382746 exon 3. Many attempts were unsuccessful in amplifying this locus from cv. Kumily genome as it resulted in no or non-specific amplification. This can occur for several reasons: 1. Sequence compatibility between two cultivars- primers designed from cv. Da-Ae might not be compatible with the target gene in the cv. Kumily, 2. Primer specificity- The non-specific binding of primers can amplify non-specific regions. Essentially, the sequence alignment results indicate that LOC106431683 lacks MYB protein coding region. Hence, the LOC106431683 was not considered for design.

The nucleotide similarity (% identity) and protein identity of cv. Kumily *MYB28* paralogs are displayed in Tables 8 and 9, respectively. It was observed that the LOC106382207 and LOC125577744 exhibited 86.7% nucleotide similarity with 100% protein identity.

4.2 Designing sgRNAs

Two sgRNAs (CRISPR target sites, with 20 bases prior to the NGG PAM site for each) with approximately 0.6 on-target activity score (Doench, 2014) were designed using Genious prime and both sgRNAs target all six paralogs (Figure 6). The nucleotide sequences for guide RNAs are as follows: 1. sgRNA from exon 1- UACAUCCACGAACAUGGAGA 2. sgRNA from exon 2- AGGCGAGUUUAGUUCAGAG. The GC content of each sgRNA was 50%. The homopolymer score for each designed sgRNA is 0.

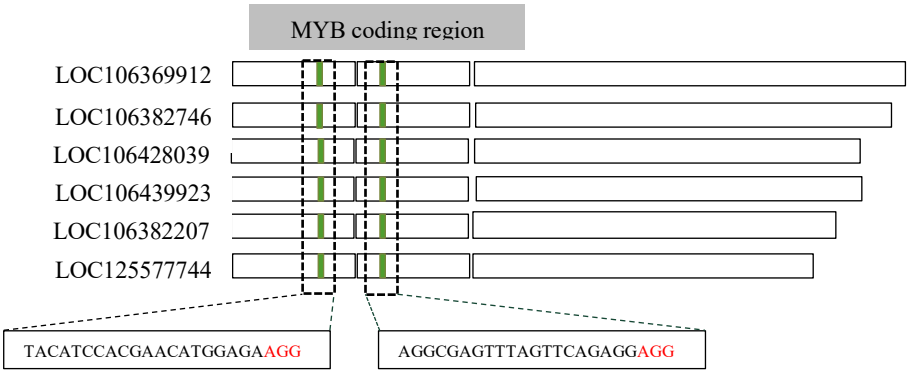


Figure 6: Schematic representation of six MYB28 *paralogs, sgRNA design, and MYB protein coding region. The PAM motifs are indicated in red.*

Table 8: Nucleotide similarity of cv. Kumily MYB28 *paralogs*

	LOC106369912	LOC106428039	LOC106382207	LOC125577744	LOC106382746	LOC106439923
LOC106369912		62.1	52.6	59.5	67.3	63.5
LOC106428039	62.1		72.0	71.6	65.3	92.5
LOC106382207	52.6	72.0		86.7	63.9	72.1
LOC125577744	59.5	71.6	86.7		59.3	72.4
LOC106382746	67.3	65.3	63.9	59.3		67.1
LOC106439923	63.5	92.5	72.1	72.4	67.1	

Table 9: Protein identity of cv. Kumily MYB28 paralogs

	LOC106369912	LOC106428039	LOC106382207	LOC125577744	LOC106382746	LOC106439923
LOC106369912		72.4	95.0	95.0	93.8	75.3
LOC106428039	72.4		91.4	91.4	68.4	94.9
LOC106382207	95.0	91.4		100	80.6	91.4
LOC125577744	95.0	91.4	100		80.6	91.4
LOC106382746	93.8	68.4	80.6	80.6		70.7
LOC106439923	75.3	94.9	91.4	91.4	70.7	

4.3 Protoplast isolation and transfection

The experiment used the first true leaves of three weeks old in vitro-grown plantlets for protoplast isolation and further gene editing. The isolated protoplasts showed adequate yield, intact plasma membrane, minimal cellular debris, and uniformity (Figure 7, B).

A plasmid vector harboring the *GFP* gene was used to transfect protoplasts to estimate the transfection efficiency. The fluorescence observation through a confocal microscope showed the successful insertion of nucleic acids into the protoplasts. The transfection efficiency was lower than 50%, and Figure 7, D observation demonstrates that about 20% of intact protoplasts can express the transgene for two days. It confirms the transfection protocol was successfully performed under the culture conditions mentioned in 2.1.

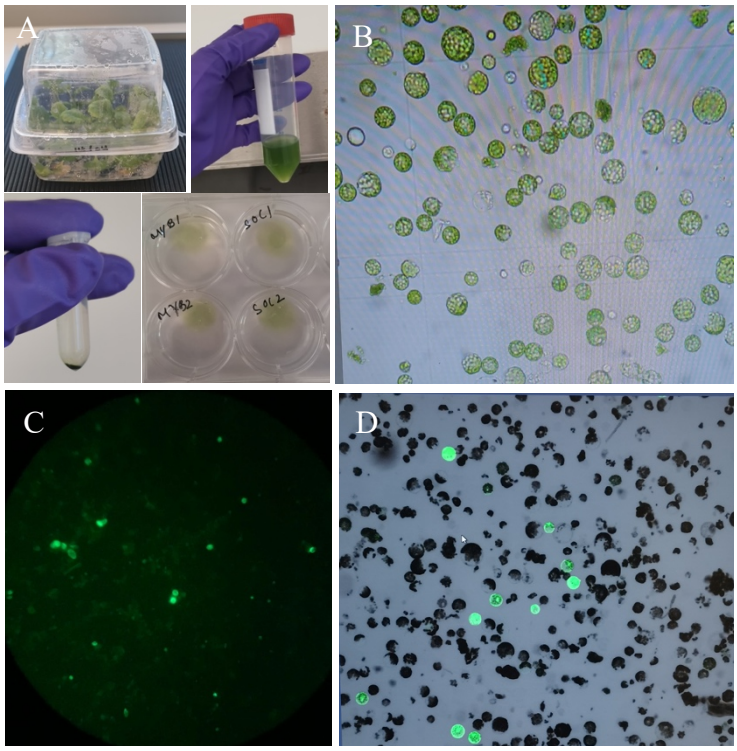


Figure 7: Different stages of protoplast isolation and transfection. A. Pictures from a transfection process; three weeks old in vitro cv. Kumily seedlings used for protoplast isolation, density-adjusted protoplast isolate (400 000 to 600 000 per ml), amount of protoplasts used per transfection, transfected protoplasts in solidified alginate gel disks. B. Isolated and purified protoplasts. C, D. *GFP* transfected protoplasts subjected to confocal fluorescence microscopy, C- direct view from the microscope (including auto-florescence), D- after adjusting auto-florescence errors.

4.4 Protoplast recovery and organogenesis

The transfected protoplasts were maintained in liquid MI medium for 3 days and MII media until the calli were visible (Figure 8, A). Li et al., (2021) observed better performance in organogenesis if the protoplasts were maintained in MII media for 15 to 20 days. However, the transfection trials in this study resulted in fewer calli development after 20 days of transfection. Hence, the protoplasts were provided with MII media for 5-10 days additionally. However, it observed accelerated growth of existing calli instead of increasing the number of calli. Out of 32 calli, only three calli regenerated into a shoot after staying one month in MIV media (Figure 8, B). However, another 9 shoots were regenerated with time and those samples are not included in to this study.

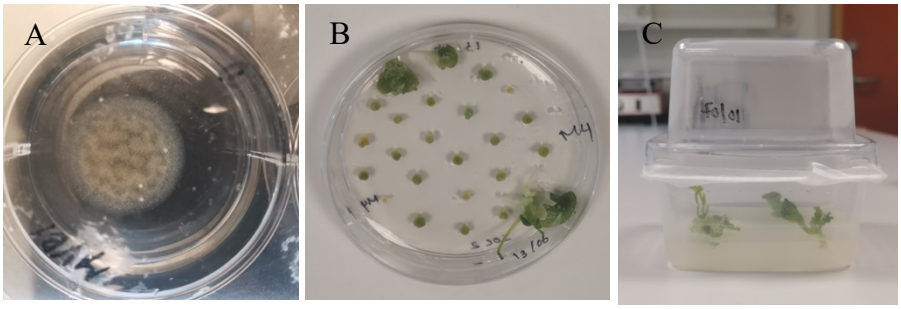


Figure 8: Different stages of protoplast culture and organogenesis. A: visible callus formation from transfected protoplasts in MII media, B: callus growth and shoot development in MIV media, C: shoot grown in shoot elongation media.

4.5 Identifying mutation events

Among six paralogs of *MYB28*, the sgRNA target sites of five paralogs were amplified using specifically designed and tested primers (Table 10). Identifying forward and reverse primers to amplify the region of interest at paralog LOC106382746 was problematic due to the frequent off-target amplification. Hence, forward and reverse primers for the paralog LOC106382746 are yet to be designed and confirmed.

Table 10: Forward and reverse primers used to amplify the target sequence region (134bp) with potential mutations in each paralog.

Locus number	Primers (5'-3')
LOC106428039	F- ATGTCAAGAAAACCGTGTTGTG R- GTAGAGCTTTTCGCGGAAGG
LOC106382207	F- ATGTCAAGAAAACCGTGTTGT R- CCAGATCTTCCGGATGGCTC
LOC106439923	F- ATGTCAAGAAAACCGTGTTGTGT R- CCTTTGTGGAGCTTTTCGCG

LOC106369912	F- ATGTCAAGAAAGCCATGTTG R- GTACGTTTATTTTTAACAAA
LOC125577744	F- TCCTTTGGTGTTGTTTGAACCTT R- TTGGAGGACCCATACACACA

The shoot initiation from calli required more time than expected. Hence, seven randomly selected calli were also included for pre-screening. Genomic DNA was isolated using 1 mg of the sample of shoots or calli was screened for CRISPR/Cas9-induced gene edits. According to the sequencing results, an insertion event was identified in LOC106428039 with the sgRNA1 of one callus tissue (Figure 9). No mutation event was observed in the other samples from the pre-screening.

	Target sequence (5'-3')
sgRNA1	TACATCCACGAACATGGAGA

Sequencing results:		
LOC106428039	**Target sequence (5'-3')**	**PAM**
Wild type	TACATCCACGAACATGGAGA	**AGG**
Mutant	TACATCCACGAACTATGGAGA	**AGG**

Figure 9: Mutation in callus sample compared with wild type of rapeseed cv. Kumily. Single base insertion at paralog LOC106428039 is highlighted in red. PAM sites are highlighted in bold text.

5. Discussion

Assessment and creation of genetic diversity in major crop plants have been recognized as significant research areas due to the influence of critical factors (increasing population, urbanization, decreasing cultivable lands, and impact of climate change on crop production) contributing to food insecurity in the future world. Nowadays, conventional breeding efforts are mainly used in plant breeding. However, inducing mutation through genetic engineering has caused some concerns about a potential increase in horizontal gene transfer. Recently, the successful application of the RNP system to induce transgene-free, precise, and time-efficient mutations in many crop varieties has drawn attention to identifying genetic engineering as a possible tool for plant breeding. Despite the advantages mentioned above, the use of the RNP system in inducing rapeseed mutations has long been obstructed due to the difficulty in protoplast regeneration. This study is among the first few investigations on rapeseed genetic engineering, which practiced the successful protoplast regeneration protocol published by researchers from the Swedish University of Agricultural Sciences, Alnarp, Sweden (Li et al., 2021).

GSL biosynthesis is a quantitative trait controlled by a highly complex network of genes and their transcriptional regulators. The expression of this trait also can alter due to environmental factors. Further, manipulating this type of quantitative trait can be challenging in any breeding effort due to the structural and expression divergence among multiple homologs of the polyploid *B. napus*. However, Hölzl et al., (2022) could successfully eliminate GSL by mutating *MYB28* and *MYB29* genes in *Camelina sativa* (hexaploid genome). This study's focus was to investigate the possibility of selectively reducing aliphatic GSL content by only inducing the mutagenesis in *MYB28*. However, due to other *MYB* family TF's (such as *MYB29* and *MYB76*) ability to contribute to the transcription of short-chain aliphatic GSL, obtaining a significant reduction of total aliphatic GSL by knocking out *MYB28* alone in the mutants can be challenging (Sønderby et al., 2010).

A recent study from Jhingan et al., (2023) confirms a significant reduction of aliphatic GSL progoitrin by random mutagenesis of *MYB28* and *BnCYP79F1* genes in *B. napus*. Since the *MYB28* TF is distinguished to positively regulate the genes involved in side chain elongation (*MAM1* and *MAM3*) and core structure formation (*CYP79F1, CYP79F2, CYP83A1, ST5b*, and *ST5c*) (Baskar & Park, 2015; Yin et

al., 2017), we predict a loss of function mutagenesis in *MYB28* would also significantly reduce the problematic aliphatic GSL such as progoitrin and gluconapoleiferin biosynthesis.

In the present study, six *MYB28* genes have been successfully identified from *B. napus* cv. Kumily. Two sgRNA sites that can mutate all six paralogs and no off-target effect were identified. Choosing suitable sgRNA is crucial to exert site-specific gene edition. Hence, several observations and principle factors were considered in designing sgRNAs. For instance, not selecting C as the closest nucleotide to the PAM site (Zhang et al., 2020). Wang et al., (2014) observed that extreme GC content could result in depleted guide RNA activity and suggested 30-70% for optimal activity. Hence, this study selected guide RNAs with 50% GC content. Moreover, the presence of homopolymers (four or more consecutive repeated nucleotides) and uracil triplet (UUU) in sgRNA design are considered detrimental to its activity (Gilbert et al., 2014). The homopolymer score for each designed gRNA was 0. However, the sgRNA designed for exon 2 contained a uracil triplet, which may adversely affect its ability to create mutations. Despite this, the sgRNA was considered for the experiment due to its zero off-target effect. However, further screening is required to confirm the above statement. The protoplast transfection of *B. napus* cv. Kumily was a meticulous process that required sharp skills and experience. The transfection efforts in this study have recovered fewer protoplasts compared to the number of protoplasts recovered by Li et al., (2021). Since the protoplast isolation and culture conditions were as described by Li et al., (2021), the lack of skills in protoplast handling could be the most likely reason for the lower recovery of protoplasts.

The involvement of GSLs in plant defense is complex (Hopkins et al., 2009). These secondary metabolites facilitate resistance from herbivores and non-adapted pathogens in rapeseed (Erb & Kliebenstein, 2020). Although the intended mutagenesis does not affect the biosynthesis of aromatic and indole GSLs, the absence of major aliphatic GSLs may reduce plant defense responses, making the plant vulnerable to pathogenic and pest attacks, especially in the root zone pathogens (Dam et al., 2004). Hence, assessing mutants' resistance/susceptibility to insect pests and diseases is imperative.

Since *B. napus* is an allotetraploid, and the trait of interest is quantitative, multiple homologs may control the same trait. Hence, it is beneficial to obtain homozygous indels in all six paralogs targeted. Therefore, the study will continue screening for all the paralogs to confirm the mutation in regenerated plants. Finally, the nature of the mutation (homozygous or heterozygous) has to be confirmed to plan the selfing or cross-pollination events under greenhouse conditions to obtain homozygous mutation lines ultimately.

6. Conclusion

The study identified six *MYB28* paralogs from rapeseed cv. Kumily. Regenerated calli or shoots were obtained from transfected protoplasts using the RNA method. Finally, one mutant callus sample sequencing on LOC106428039 proved at least one sgRNA (TACATACACACATGGAGA) was successful in inducing mutation.

7. References

Adjonu, R., Zhou, Z., Prenzler, P. D., Ayton, J., & Blanchard, C. L. (2019). Different Processing Practices and the Frying Life of Refined Canola Oil. Foods, 8(11), 527. https://doi.org/10.3390/foods8110527.

Aider, M., & Barbana, C. (2011). Canola proteins: composition, extraction, functional properties, bioactivity, applications as a food ingredient and allergenicity – A practical and critical review. *Trends in Food Science & Technology*, 22(1), 21–39. https://doi.org/10.1016/j.tifs.2010.11.002

Augustine, R., Mukhopadhyay, A., & Bisht, N. C. (2013). Targeted silencing of BjMYB28 transcription factor gene directs development of low glucosinolate lines in oilseed *Brassica juncea*. *Plant Biotechnology Journal*, 11(7), 855–866. https://doi.org/10.1111/pbi.12078.

Barba, F. J., Nikmaram, N., Roohinejad, S., Khelfa, A., Zhu, Z., & Koubaa, M. (2016). Bioavailability of Glucosinolates and Their Breakdown Products: Impact of Processing. *Frontiers in Nutrition*, 3. https://doi.org/10.3389/fnut.2016.00024.

Barda, O., & Levy, M. (2022). IQD1 Involvement in Hormonal Signaling and General Defense Responses Against Botrytis cinerea. *Frontiers in Plant Science*, 13. https://doi.org/10.3389/fpls.2022.845140

Baskar, V., & Park, S. W. (2015). Molecular characterization of *BrMYB28* and *BrMYB29* paralogs transcription factors involved in the regulation of aliphatic glucosinolate profiles in *Brassica rapa* ssp. pekinensis. *Comptes Rendus Biologies*, 338(7), 434–442. https://doi.org/10.1016/j.crvi.2015.04.001.

Becker, T., & Juvik, J. (2016). The Role of Glucosinolate Hydrolysis Products from Brassica Vegetable Consumption in Inducing Antioxidant Activity and Reducing Cancer Incidence. *Diseases*, 4(4), 22. https://doi.org/10.3390/diseases4020022.

Bekaert, M., Edger, P. P., Hudson, C. M., Pires, J. C., & Conant, G. C. (2012). Metabolic and evolutionary costs of herbivory defense: systems biology of glucosinolate synthesis. *New Phytologist*, 196(2), 596–605. https://doi.org/10.1111/j.1469-8137.2012.04302.x.

Bell, J. M. (1993). Factors affecting the nutritional value of canola meal: A review. *Canadian Journal of Animal Science*, 73(4), 689–697. https://doi.org/10.4141/cjas93-075.

Bell, L., Oloyede, O. O., Lignou, S., Wagstaff, C., & Methven, L. (2018). Taste and Flavor Perceptions of Glucosinolates, Isothiocyanates, and Related Compounds. *Molecular Nutrition & Food Research*, 62(18), 1700990. https://doi.org/10.1002/mnfr.201700990.

Bertinato, J. (2021). Iodine nutrition: Disorders, monitoring and policies. *Advances in Food and Nutrition Research*, 96, 365-415. https://doi.org/10.1016/bs.afnr.2021.01.004

Boydston, R. A., Anderson, T., & Vaughn, S. F. (2008). Mustard (Sinapis alba) Seed Meal Suppresses Weeds in Container-grown Ornamentals. *HortScience*, 43(3), 800–803. https://doi.org/10.21273/hortsci.43.3.800.

Brown, P. D., Tokuhisa, J. G., Reichelt, M., & Gershenzon, J. (2003). Variation of glucosinolate accumulation among different organs and developmental stages of Arabidopsis thaliana. *Phytochemistry*, 62(3), 471–481. https://doi.org/10.1016/s0031-9422(02)00549-6.

Bruce, T. J. A. (2014). Glucosinolates in oilseed rape: secondary metabolites that influence interactions with herbivores and their natural enemies. *Annals of Applied Biology*, 164(3), 348–353. https://doi.org/10.1111/aab.12128.

Cao, Y., Li, K., Li, Y., Zhao, X., & Wang, L. (2020). MYB Transcription Factors as Regulators of Secondary Metabolism in Plants. *Biology*, 9(3), 61. https://doi.org/10.3390/biology9030061.

Cartea, M. E., & Velasco, P. (2007). Glucosinolates in Brassica foods: bioavailability in food and significance for human health. *Phytochemistry Reviews*, 7(2), 213–229. https://doi.org/10.1007/s11101-007-9072-2.

Celenza, J. L., Quiel, J. A., Smolen, G. A., Merrikh, H., Silvestro, A. R., Normanly, J., & Bender, J. (2004). The Arabidopsis ATR1 Myb Transcription Factor Controls Indolic Glucosinolate Homeostasis. *Plant Physiology*, 137(1), 253–262. https://doi.org/10.1104/pp.104.054395.

Cradick, T. J., Fine, E. J., Antico, C. J., & Bao, G. (2013). CRISPR/Cas9 systems targeting β-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Research*, 41(20), 9584–9592. https://doi.org/10.1093/nar/gkt714.

Dinkova-Kostova, A. T., Jenkins, S. N., Fahey, J. W., Ye, L., Wehage, S. L., Liby, K. T., Stephenson, K. K., Wade, K. L., & Talalay, P. (2006). Protection against UV-light-induced skin carcinogenesis in SKH-1 high-risk mice by sulforaphane-containing broccoli sprout extracts. *Cancer Letters*, 240(2), 243–252. https://doi.org/10.1016/j.canlet.2005.09.012.

Doughty, K. J., Kiddle, G. A., Pye, B. J., Wallsgrove, R. M., & Pickett, J. A. (1995). Selective induction of glucosinolates in oilseed rape leaves by methyl jasmonate. *Phytochemistry*, 38(2), 347–350. https://doi.org/10.1016/0031-9422(94)00653-b.

Douglas Grubb, C., Zipp, B. J., Ludwig-Müller, J., Masuno, M. N., Molinski, T. F., & Abel, S. (2004). Arabidopsis glucosyltransferase UGT74B1 functions in glucosinolate biosynthesis and auxin homeostasis. The Plant Journal, 40(6), 893–908. https://doi.org/10.1111/j.1365-313x.2004.02261.x.

Dukhi, N. (2020). Global Prevalence of Malnutrition: Evidence from Literature. In *Malnutrition*. IntechOpen. https://doi.org/10.5772/intechopen.92006.

Eisenschmidt-Bönn, D., Schneegans, N., Backenköhler, A., Wittstock, U., & Brandt, W. (2019). Structural diversification during glucosinolate breakdown: mechanisms of thiocyanate, epithionitrile and simple nitrile formation. *Plant Journal*, *99*(2), 329–343. https://doi.org/10.1111/tpj.14327.

Erb, M., & Kliebenstein, D. J. (2020). Plant Secondary Metabolites as Defenses, Regulators, and Primary Metabolites: The Blurred Functional Trichotomy. *Plant Physiology*, 184(1), 39-52. https://doi.org/10.1104/pp.20.00433.

Eskin, N. A. M., & Przybylski, R. (2003). Rape Seed Oil/Canola. In *Encyclopedia of Food Sciences and Nutrition* (pp. 4911–4916). Elsevier.https://doi.org/10.1016/b0-12-227055-x/01349-3.

Ewing, W.N. 1998. The Feeds Directory: Commodity Products Guide. Context Publications, Ashby de la Zouch, UK, pp. 118. https://doi.org/10.1016/j.canlet.2005.09.012

Falk, K. (2004). Glucosinolate biosynthesis: demonstration and characterization of the condensing enzyme of the chain elongation cycle in Eruca sativa. Phytochemistry, 65(8), 1073–1084. https://doi.org/10.1016/j.phytochem.2004.02.021.

Feng D, Zuo J. (2003). Nutritional and anti-nutritional composition of rapeseed meal and its utilization as a feed ingredient for animal. In *Feed And Industrial Raw Material: Feed* (p. 265). College of Animal science, South China Agricultural University. https://www.gcirc.org/fileadmin/documents/Proceedings/IRCWuhan2007%20vol5/Pages_de_vol-5-37.pdf.

Frerigmann, H. (2016). Glucosinolate Regulation in a Complex Relationship – MYC and MYB – No One Can Act Without Each Other. In *Advances in Botanical Research* (pp. 57–97). Elsevier. https://doi.org/10.1016/bs.abr.2016.06.005.

Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M., & Joung, J. K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nature Biotechnology*, 32(3), 279–284. https://doi.org/10.1038/nbt.2808.

Garneau, J., Dupuis, MÈ., Villion, M. et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67–71 (2010). https://doi.org/10.1038/nature09523.

Georgeta C (2009) Efficiency of soybean meal replacement by rapeseed meal and/or canola seeds in commercial layer diet. In *Archiva Zootechnica* 12(1): 27-33. https://ibna.ro/arhiva/AZ%2012-1/AZ%2012-1%2003%20CiurescuG.pdf.

Gigolashvili, T., Engqvist, M., Yatusevich, R., Müller, C., & Flügge, U.-I. (2008). HAG2/MYB76 and HAG3/MYB29 exert a specific and coordinated control on the regulation of aliphatic glucosinolate biosynthesis inArabidopsis thaliana. *New Phytologist*, 177(3), 627–642. https://doi.org/10.1111/j.1469-8137.2007.02295.x.

Glawischnig, E. (2006). The role of cytochrome P450 enzymes in the biosynthesis of camalexin. Biochemical Society Transactions, 34(6), 1206–1208. https://doi.org/10.1042/bst0341206.

Hansen, B. G., Kliebenstein, D. J., & Halkier, B. A. (2007). Identification of a flavin-monooxygenase as the S-oxygenating enzyme in aliphatic glucosinolate biosynthesis in Arabidopsis. The Plant Journal, 50(5), 902–910. https://doi.org/10.1111/j.1365-313x.2007.03101.x.

Hemmi, H., Takeuchi, O., Kawai, T., Kaisho, T., Sato, S., Sanjo, H., Matsumoto, M., Hoshino, K., Wagner, H., Takeda, K., & Akira, S. (2000). A Toll-like receptor

recognizes bacterial DNA. *Nature*, 408(6813), 740–745. https://doi.org/10.1038/35047123.

Hirai, M. Y., Sugiyama, K., Sawada, Y., Tohge, T., Obayashi, T., Suzuki, A., Araki, R., Sakurai, N., Suzuki, H., Aoki, K., Goda, H., Nishizawa, O. I., Shibata, D., & Saito, K. (2007). Omics-based identification of Arabidopsis Myb transcription factors regulating aliphatic glucosinolate biosynthesis. *Proceedings of the National Academy of Sciences*, 104(15), 6478–6483. https://doi.org/10.1073/pnas.0611629104.

Hölzl, G., Rezaeva, B. R., Kumlehn, J., & Dörmann, P. (2022). Ablation of glucosinolate accumulation in the oil crop Camelina sativa by targeted mutagenesis of genes encoding the transporters GTR1 and GTR2 and regulators of biosynthesis MYB28 and MYB29. Plant Biotechnology Journal. https://doi.org/10.1111/pbi.13936.

Hopkins, R. J., van Dam, N. M., & van Loon, J. J. A. (2009). Role of Glucosinolates in Insect-Plant Relationships and Multitrophic Interactions. Annual Review of Entomology, 54(1), 57–83. https://doi.org/10.1146/annurev.ento.54.110807.090623.

Ishimoto, H., Matsumoto, A., & Tanimura, T. (2000). Molecular identification of a taste receptor gene for trehalose in Drosophila. *Science*, *289*(5476), 116–119. https://doi.org/10.1126/science.289.5476.116.

Jhingan, S., Harloff, H.-J., Abbadi, A., Welsch, C., Blümel, M., Tasdemir, D., & Jung, C. (2023). Reduced glucosinolate content in oilseed rape (*Brassica napus* L.) by random mutagenesis of *BnMYB28* and *BnCYP79F1* genes. *Scientific Reports*, *13*(1), 2344. https://doi.org/10.1038/s41598-023-28661-6.

Jiang, K., Liu, Y., Zhang, J., & Jia, J. (2017). An intracellular activation of Smoothened that is independent of Hedgehog stimulation in Drosophila. *Journal of Cell Science*, 131(1), jcs211367. https://doi.org/10.1242/jcs.211367.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, 337(6096), 816–821. https://doi.org/10.1126/science.1225829.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science, 337(6096), 816–821. https://doi.org/10.1126/science.1225829.

Kim, S., Kim, D., Cho, S. W., Kim, J., & Kim, J. S. (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Research*, 24(6), 1012–1019. https://doi.org/10.1101/gr.171322.113.

Kitainda, V., & Jez, J. M. (2021). Structural studies of aliphatic glucosinolate chain-elongation enzymes. *Antioxidants*, *10*(9). https://doi.org/10.3390/antiox10091500.

Knill, T., Schuster, J., Reichelt, M., Gershenzon, J., & Binder, S. (2007). Arabidopsis Branched-Chain Aminotransferase 3 Functions in Both Amino Acid and

Glucosinolate Biosynthesis. Plant Physiology, 146(3), 1028–1039. https://doi.org/10.1104/pp.107.111609.

Langer P, Greer MA (1977) Antithyroid Substances and Naturally Occurring Goitrogens. Letters, 240, https://doi.org/10.1159/000400572.

Li, X., Sandgrind, S., Moss, O., Guan, R., Ivarson, E., Wang, E. S., Kanagarajan, S., & Zhu, L.-H. (2021). Efficient Protoplast Regeneration Protocol and CRISPR/Cas9-Mediated Editing of Glucosinolate Transporter (GTR) Genes in Rapeseed (*Brassica napus* L.). *Frontiers in Plant Science*, 12. https://doi.org/10.3389/fpls.2021.680859.

Liang, X., Potter, J., Kumar, S., Zou, Y., Quintanilla, R., Sridharan, M., Carte, J., Chen, W., Roark, N., Ranganathan, S., Ravinder, N., & Chesnut, J. D. (2015). Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. Journal of Biotechnology, 208, 44–53. https://doi.org/10.1016/j.jbiotec.2015.04.024.

Lin, Q., Zong, Y., Xue, C., Wang, S., Jin, S., Zhu, Z., Wang, Y., Anzalone, A. V., Raguram, A., Doman, J. L., Liu, D. R., & Gao, C. (2020). Prime genome editing in rice and wheat. Nature Biotechnology, 38(5), 582–585. https://doi.org/10.1038/s41587-020-0455-x.

Maison, T., Liu, Y., & Stein, H. H. (2015). Digestibility of energy and detergent fiber and digestible and metabolizable energy values in canola meal, 00-rapeseed meal, and 00-rapeseed expellers fed to growing pigs. *Journal of Animal Science*, 93(2), 652. https://doi.org/10.2527/jas.2014-7792.

Makarova, K. S., & Koonin, E. V. (2012). Evolution and Classification of CRISPR-Cas Systems and Cas Protein Families. In *CRISPR-Cas Systems* (pp. 61–91). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-34657-6_3.

Manici, L. M., Lazzeri, L., & Palmieri, S. (1997b). In VitroFungitoxic Activity of Some Glucosinolates and Their Enzyme-Derived Products toward Plant Pathogenic Fungi. *Journal of Agricultural and Food Chemistry*, 45(7), 2768–2773. https://doi.org/10.1021/jf9608635.

Mawson, R., Heaney, R. K., Piskula, M., & Kozlowska, H. (1993). Rapeseed meal-glucosinolates and their antinutritional effects Part 1. Rapeseed production and chemistry of glucosinolates. *Food / Nahrung*, 37(2), 131–140. https://doi.org/10.1002/food.19930370206.

Mithen, R., Clarke, J., Lister, C., & Dean, C. (1995). Genetics of aliphatic glucosinolates. III. Side chain structure of aliphatic glucosinolates in Arabidopsis thaliana. *Heredity*, 74(2), 210–215. https://doi.org/10.1038/hdy.1995.29.

Molla, K. A., Shih, J., & Yang, Y. (2020). Single-nucleotide editing for zebra3 and wsl5 phenotypes in rice using CRISPR/Cas9-mediated adenine base editors. aBIOTECH, 1(2), 106–118. https://doi.org/10.1007/s42994-020-00018-x.

Moreb, N., Murphy, A., Jaiswal, S., & Jaiswal, A. K. (2020). Chapter 3 - Cabbage. In A. K. Jaiswal (Ed.), *Nutritional Composition and Antioxidant Properties of Fruits and Vegetables* (pp. 33–54). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-812780-3.00003-9.

Neequaye, M., Stavnstrup, S., Harwood, W., Lawrenson, T., Hundleby, P., Irwin, J., Troncoso-Rey, P., Saha, S., Traka, M. H., Mithen, R., & Østergaard, L. (2021). CRISPR-Cas9-Mediated Gene Editing of *MYB28* Genes Impair Glucoraphanin Accumulation of *Brassica oleracea* in the Field. *The CRISPR Journal*, 4(3), 416–426. https://doi.org/10.1089/crispr.2021.0007.

Nega, T. (2018). Review on Nutritional Limitations and Opportunities of using Rapeseed Meal and other Rape Seed by - Products in Animal Feeding. *Journal of Nutritional Health & Food Engineering*, 8(1). https://doi.org/10.15406/jnhfe.2018.08.00254.

Newkirk, R. (2011). Meal Nutrient Composition. In *Canola* (pp. 229–244). Elsevier. https://doi.org/10.1016/b978-0-9818936-5-5.50012-7.

Oilseeds: World Markets and Trade. (n.d.). USDA Foreign Agricultural Service. https://www.fas.usda.gov/data/oilseeds-world-markets-and-trade.

Palermo, G., Chen, J. S., Ricci, C. G., Rivalta, I., Jinek, M., Batista, V. S., Doudna, J. A., & McCammon, J. A. (2018). Key role of the REC lobe during CRISPR–Cas9 activation by 'sensing', 'regulating', and 'locking' the catalytic HNH domain. *Quarterly Reviews of Biophysics*, 51. https://doi.org/10.1017/s0033583518000070.

Parkin I. A. P, Gulden S. M, Sharpe A. G, Lukens L, Trick M, Osborn T. C, Lydiate D. J (2005) Segmental structure of the *Brassica napus* genome based on comparative analysis with Arabidopsis thaliana. *Genetics* 171: 765–781. https://doi.org/10.1016/b0-12-227055-x/01349-3.

Piekarska, A., Kołodziejski, D., Pilipczuk, T., Bodnar, M., Konieczka, P., Kusznierewicz, B., Hanschen, F. S., Schreiner, M., Cyprys, J., Groszewska, M., Namieśnik, J., & Bartoszek, A. (2014). The influence of selenium addition during germination of Brassica seeds on health-promoting potential of sprouts. *International Journal of Food Sciences and Nutrition*, 65(6), 692–702. https://doi.org/10.3109/09637486.2014.917148.

Porter, A. J. R., Morton, A. M., Kiddle, G., Doughty, K. J., & Wallsgrove, R. M. (1991). Variation in the glucosinolate content of oilseed rape (*Brassica napus* L.) leaves. *Annals of Applied Biology*, 118(2), 461–467. https://doi.org/10.1111/j.1744-7348.1991.tb05647.x.

Prieto, M. A., López, C. J., & Simal-Gandara, J. (2019). Glucosinolates: Molecular structure, breakdown, genetic, bioavailability, properties and healthy and adverse effects. In *Advances in Food and Nutrition Research* (pp. 305–350). Elsevier. https://doi.org/10.1016/bs.afnr.2019.02.008.

Redovniković, I. R., Glivetić, T., Delonga, K., & Vorkapić-Furač, J. (2008). Glucosinolates and their potential role in plant. *Periodicum Biologorum*, *110*(4), 297–309. https://doi.org/10.1007/978-1-59745-425-4_11.

Reményi, A., Schöler, H. R., & Wilmanns, M. (2004). Combinatorial control of gene expression. Nature Structural & Molecular Biology, 11(9), 812–815. https://doi.org/10.1038/nsmb820.

Roth-Maier, D. A., Böhmer, B. M., & Roth, F. X. (2004). Effects of feeding canola meal and sweet lupin (L. luteus, L. angustifolius) in amino acid balanced diets on

growth performance and carcass characteristics of growing-finishing pigs. Animal Research, 53(1), 21–34. https://doi.org/10.1051/animres:2003048.

Rungapamestry, V., Duncan, A. J., Fuller, Z., & Ratcliffe, B. (2006). Changes in Glucosinolate Concentrations, Myrosinase Activity, and Production of Metabolites of Glucosinolates in Cabbage (*Brassica oleracea*Var.capitata) Cooked for Different Durations. *Journal of Agricultural and Food Chemistry*, 54(20), 7628–7634. https://doi.org/10.1021/jf0607314.

Sarwar and J . A . Kirkegaard (1998). Biofumigation potential of brassicas :Effect of environment and ontogeny on glucosinolate production and implications for screening. *Plant and Soil* , Vol . 201, 91–101. https://doi.org/10.1007/s10658-009-9543-y.

Schöne, F., Leiterer, M., Hartung, H., Jahreis, G., & Tischendorf, F. (2001). Rapeseed glucosinolates and iodine in sows affect the milk iodine concentration and the iodine status of piglets. *British Journal of Nutrition*, 85(6), 659–670. https://doi.org/10.1079/bjn2001326.

Scott, T., Urak, R., Soemardy, C., & Morris, K. V. (2019). Improved Cas9 activity by specific modifications of the tracrRNA. Scientific Reports, 9(1). https://doi.org/10.1038/s41598-019-52616-5.

Seneviratne, R. W., Young, M. G., Beltranena, E., Goonewardene, L. A., Newkirk, R. W., & Zijlstra, R. T. (2010). The nutritional value of expeller-pressed canola meal for grower-finisher pigs1. *Journal of Animal Science*, 88(6), 2073–2083. https://doi.org/10.2527/jas.2009-2437.

Sønderby, I. E., Geu-Flores, F., & Halkier, B. A. (2010). Biosynthesis of glucosinolates - gene discovery and beyond. *Trends in Plant Science*, *15*(5), 283–290. https://doi.org/10.1016/j.tplants.2010.02.005.

Spragg, J. C., and R. J. Mailer. 2007. Canola meal value chain quality improvement: A final report prepared for AOF and CRC. Project code: 1B-103-0506. Available at: http://www.porkcrc.com.au/Final_Report_1B-103.pdf. Accessed Jan. 9, 2011.

Textor, S., de Kraker, J.-W., Hause, B., Gershenzon, J., & Tokuhisa, J. G. (2007). MAM3 Catalyzes the Formation of All Aliphatic Glucosinolate Chain Lengths in Arabidopsis. *Plant Physiology*, 144(1), 60–71. https://doi.org/10.1104/pp.106.091579.

Tripathi, M. K., & Mishra, A. S. (2007). Glucosinolates in animal nutrition: A review. *Animal Feed Science and Technology*, 132(1-2), 1–27. https://doi.org/10.1016/j.anifeedsci.2006.03.003.

Tuli, A., Sethi, R. P., Khanna, P. K., Marwaha, S. S., & Kennedy, J. F. (1985). Lactic acid production from whey permeate by immobilized Lactobacillus casei. *Enzyme and Microbial Technology*, 7(4), 164–168. https://doi.org/10.1016/0141-0229(85)90058-4.

Ullah, M., Islam, M., Islam, M., & Haque, T. (1970). Effects of organic manures and chemical fertilizers on the yield of brinjal and soil properties. *Journal of the Bangladesh Agricultural University*, 6(2), 271–276. https://doi.org/10.3329/jbau.v6i2.4821.

Van Dam, N. M., Witjes, L., & Svatoš, A. (2004). Interactions between aboveground and belowground induction of glucosinolates in two wild Brassica species. *New Phytologist*, 161(3), 801–810. https://doi.org/10.1111/j.1469-8137.2004.00984.x.

Velayudhan, D. E., Schuh, K., Woyengo, T. A., Sands, J. S., & Nyachoti, C. M. (2017). Effect of expeller extracted canola meal on growth performance, organ weights, and blood parameters of growing pigs. Journal of Animal Science, 95(1), 302–307. https://doi.org/10.2527/jas.2016.1046.

Wang, Y., Zhang, Y., Dong, Y., Li, D., Shi, S., Li, S., Li, L., He, Y., Li, J., Chen, H., Ge, H., & Liu, Y. (2022). A highly efficient mesophyll protoplast isolation and PEG-mediated transient expression system in eggplant. Scientia Horticulturae, 304, 111303. https://doi.org/10.1016/j.scienta.2022.111303.

Wittstock, U., & Halkier, B. A. (2002). Glucosinolate research in the Arabidopsis era. *Trends in Plant Science*, 7(6), 263–270. https://doi.org/10.1016/s1360-1385(02)02273-2.

Woo, J. W., Kim, J., Kwon, S. I., Corvalán, C., Cho, S. W., Kim, H., Kim, S.-G., Kim, S.-T., Choe, S., & Kim, J.-S. (2015). DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. Nature Biotechnology, 33(11), 1162–1164. https://doi.org/10.1038/nbt.3389.

Xu, C. L., Ruan, M. Z. C., Mahajan, V. B., & Tsang, S. H. (2019). Viral Delivery Systems for CRISPR. Viruses, 11(1), 28. https://doi.org/10.3390/v11010028.

Yin, L., Chen, H., Cao, B., Lei, J., & Chen, G. (2017). Molecular Characterization of MYB28 Involved in Aliphatic Glucosinolate Biosynthesis in Chinese Kale (*Brassica oleracea* var. alboglabra Bailey). Frontiers in Plant Science, 8. https://doi.org/10.3389/fpls.2017.01083

Zhang, S., Shen, J., Li, D., & Cheng, Y. (2021). Strategies in the delivery of Cas9 ribonucleoprotein for CRISPR/Cas9 genome editing. *Theranostics*, 11(2), 614–648. https://doi.org/10.7150/thno.47007.

Zhou, X., Zhang, H., Xie, Z., Liu, Y., Wang, P., Dai, L., Zhang, X., Wang, Z., Wang, Z., Wan, L., Yang, G., & Hong, D. (2022). Natural variation and artificial selection at the *BnaC2.MYB28* locus modulate *Brassica napus* seed glucosinolate. Plant Physiology. https://doi.org/10.1093/plphys/kiac463.

8. Popular science summary

Reap the Benefit of Rapeseed

Rapesed cake, the protein-rich by-product of rapeseed oil extraction, grasps great potential as a source of nutrition for both humans and livestock. However, its high levels of glucosinolates (GSLs) have long been a problem due to their anti-nutritional properties and extreme bitterness when hydrolyzed. These unfavorable characteristics have limited the use of rapeseed cake in both human and animal diets. Among the different types of GSLs, aliphatic GSLs have been identified as the major culprits behind these negative properties.

To address this issue, a precision plant breeding study was conducted to reduce the levels of long and short-chain aliphatic GSLs in rapeseed cake. This research work employed a cutting-edge genetic tool known as CRISPR to target the principal regulator of aliphatic GSL biosynthesis, *MYB28*. This transcription factor positively regulates several genes in different stages of GSL biosynthesis, such as side chain elongation and core structure formation. So, the absence of this transcription factor will automatically restrict the aliphatic GSL biosynthesis. They identified six out of seven *MYB28* paralogs in the cv. Kumily genome and designed two sgRNAs to target specific regions within both MYB-type HTH domains of MYB28 protein. The RNP complex that consists of sgRNA and Cas9 protein was introduced into cv. Kumily protoplasts to induce mutation. Subsequently, the transformed protoplasts were cultured under in-vitro conditions to promote organogenesis. Genomic DNA was extracted from the true leaves of three regenerated plants and seven randomly selected callus samples and screened for mutation.

In conclusion, this study has identified a successful sgRNA that can initiate CRISPR/Cas9-mediated mutagenesis at *MYB28*, as confirmed by the mutation event in LOC106428039. The research is still ongoing, with the continuation of screening the regenerated shoots from the transfected protoplasts to identify desirable homozygous mutation lines. This exciting work paves the way for a future where rapeseed cake can be used more effectively for human nutrition and improved livestock feed.

Appendix

Appendix 1

a. Nuclotide similarity (% identity) of MYB28 *paralogs in B. napus cv. Da-Ae.*

	LOC106369912	LOC106382207	LOC106428039	LOC106431683	LOC106439923	LOC125577744
LOC106369912		54.7	78.1	49.3	38.0	50.5	48.7
LOC106382207	54.7		68.4	73.0	69.9	73.8	87.4
LOC106428039	78.1	68.4		63.8	88.9	65.6	63.4
LOC106431683	49.3	73.0	63.8		70.7	92.6	73.6
LOC106439923	38.0	69.9	88.9	70.7		72.9	70.4
LOC125577744	50.5	73.8	65.6	92.6	72.9		74.0
	48.7	87.4	63.4	73.6	70.4	74.0	

b. Protein identity of MYB28 paralogs in B.napus cv. Da-Ae.

	LOC106369912	LOC106382207	LOC106428039	LOC106431683	LOC106439923	LOC125577744
LOC106369912		76.1	87.0	71.7	19.0	71.7	76.8
LOC106382207	76.1		72.7	76.4	53.2	78.0	95.5
LOC106428039	87.0	72.7		71.8	93.8	73.4	74.0
LOC106431683	71.7	76.4	71.8		59.3	94.4	77.2
LOC106439923	19.0	53.2	93.8	59.3		62.7	57.5
LOC125577744	71.7	78.0	73.4	94.4	62.7		79.6
	76.8	95.5	74.0	77.2	57.5	79.6	

Appendix 2

The gene inventory of the aliphatic GSL biosynthesis pathway (Sønderby et al., 2010)

Gene	AGI code	Gene	AGI code	Gene	AGI code
BCAT4	At3g19710	*IPMDH3***	At1g31180	*UGT74C1***	At2g3179
*BAT5**	At4g12030	*BCAT3*	At3g49680	*SOT18*	At1g74090
*MAM1****	At5g23010	*CYP79F1****	At1g16410	*SOT17*	At1g18590

MAM2	-	*CYP79F2***	At1g16400	*FMO-GSOX1*	At1g65860
*MAM3****	At5g23020	*CYP83A1****	At4g13770	*FMO-GSOX2*	At1g62540
IPMI LSU1	At4g13430	*GSTF11***	At3g03190	*FMO-GSOX3*	At1g62560
*IPMI SSU2**	At2g43100	*GSTU20***	At1g78370	*FMO-GSOX4*	At1g62570
*IPMI SSU3**	At3g58990	*GGP1**	At4g30530	*FMO-GSOX5*	At1g12140
IPMDH1	At5g14200	*SUR1*	At2g20610	*AOP3*	At4g03050
AOP2	At4g03060	*GS-OH*	At2g25450		

* = partially characterized enzyme, ** = predicted enzyme, ***= genes regulated by MYB28 TF

Appendix 3

Nucleotide sequences of each MYB28 *paralogs identified from cv. Kumily genome.*

LOC106428039

ATGTCAAGAAAACCGTGTTGTGTCGGAGAAGGGCTGAAGAAAGGGGCATGGACCACCGAAGAAGATAAGAAACTCATCTCT
TACATCCACGAACATGGAGAAGGAGGCTGGCGCGACATTCCCCAAAAAGCTGGTTAAATTTTTTTTTTTGTAACACAGCTGG
TTAAATATCTATGAAATTATACACATGCTTATTTAGTAATCGTACTTGTATGAAATTCTAATTAATTATGATTATGGTATG
TAATTAGGGTTAAAAAGGTGTGGAAAGAGTTGTAGACTGCGTTGGACTTACTACCTAAAACCTGATGTCAAAAGAGGCGAG
TTTAGTTCAGAGGAGGAACAGATTATTATTATACTTCATGCTTCTCGTGGCAACAAGTACGTTTATATGCTTAAATGTGTG
TGTATGCATGGATCCTCCAATGAATAATGAAATGCATGAAAAAGTTCCATATATTATTTTAAAGATACGATCATTATTAGT
TAATATAGTCCCTTCCGCGAAAAGCTCTACAAAGGAATCATATTAATACCGTCTTTGATATCTCAATCTCTCTTCTTGCCT
TTAGGTGGTCGGACATAGCGAGACATTTACCTAGAAGAACAGACAATGAGGTCAAGAATTACTGGAACACTCATCTTAAAA
AGCGTTTGATCGAACAGGGTATTGATCCCGTGACTCACAAGCCACTGGCTTCTAATTCCAACCCTACCGTCAACACGCCTC
CAGAGAATTTGCATTCCCTTGCTGCGCCTAGTTCCGACAAGCAACACTCCCGGTCGAGCTCATTGCCTTCCCTGTCTCGTC
TTACCAACAAAGATGGGACACCAGTTCAAGGCGGTTCCTTGAGTCACAAGAAACGTTTTAAGAGGTCGAGTTCTACATCAA
GGCTTTTGAACAAAGTTGCGGCTAAGGTCACTTCTGTAAAAGAAATATTGTCGGCTTCCATGGAAGGTAGCTTGAGCGCTA
CTACATTACCATATGCAAGCTATTCTGATGGCTTCTCTGAGCAGATTCGCAATGGAGAGGGTAGTTCCAACGCGTTCCTGA
CAAATACTCTCGCCGAGTTCGATCCCTTCTCCCAATCACCGTTGTACAGTGAGCATGAGATCAACGCTACTTCTGATCTCG
GTATGGATTACGATTTCTCACATTTTCTTGAAAAGCTTGGGAGAGATGACCACAACGAGGAGAACGATATGAATGTCGAGT
ATGGTCATGATGATCTTCTTATGTCCGATGTGTCTCAAGAAGTCTCATCAACAAGCGTTGATGATCGACACAATATATTTG
AAAATTTTGAGGGTTGGTCAAATTATCTTCTTGACCATGCGGATTTCGTATATGACACGGAGTCTGATACCCTCATATGA

LOC106382207

ATGTCAAGAAAACCGTGTTGTGTCGGAGAAGGGCTGAAGAAAGGGGCATGGACCACCGAGGAAGATAAGAAACTCATCTCT
TACATCCATGAACATGGAGAAGGAGGCTGGCGCGACATTCCTCAAAAAGCTGGTTAATATCTATTATATATATTTTGGTAA
ATTTTTAAAACATATATGTTTGTTTGGTATTTGATGTATGAAAGTTTTATATTGAATGTGGTGTTTTACTAGGATTGAAAA
GGTGTGGAAAGAGTTGTAGACTGCGATGGACTAACTACCTAAAACCTGAAATCAAAAGAGGCGAGTTTAGTTCAGAGGAGG
AACAGATTATCATCATGCTTCATGCTTCTCGTGGAAACAAGTAAGTATATATGTGTGTATGGGTCCTCCAATGATTATCAT
TCTGATTTTGTTTCTTTTCCATGTATTTGACAATTCTCTGATGCAAAATATGTGTTGGATTAGCAATTTTTGACTAAAAAT
TGCAATAGCACACATCATTTTAGTTTCAAAATAAATTTTATTAACTTTGGTCTTCCATCTTATCTCTTTCACTATTCTGGT
CTTTAGGTGGTCGGTCATAGCGAGACATTTACCTAGAAGAACAGACAATGAGATCAAGAACTACTGGAACACACATCTCAA
GAAACGTTTGATCGAACAGGGTACTGATCCCGTGACTCACAAGCCACTAGCTTCTAATACAAACCCTAGCTACTGTACCTG
A

LOC106382746

ATGTCAAGAAAGCCATGTTGTGTCGGAGAAGGGTTGAAGAAAGGTGCATGGACCACCGAGGAAGACAAGACGCTCATCTCT
TACATTCACGAACATGGAGAAGGTGGCTGGCGCGACATTCCCCAAAAAGCTGGTTAGTATATATATAGTATACACAGGCAT
GTATAATTAGTATTTATTGTATAATATTTTGATGAGAATATGGTGTTTAATTAGGGTTGAAAAGGTGTGGAAAGAGTTGTA
GACTGCGATGGACTAACTACCTAAAACCTGAGATCAAAAGAGGCGAGTTTAGTTCAGAGGAGGAACAGATTATCATCATGC
TCCATGCTGCTCGTGGCAACAAGTACGTTTATTTTTAACAAAAAGGACTATTATATGTTTTTTATGTGCGTATGAATCCTC
CGGTAATCATCATTCTAGTTGTCTCTTTTTTTTTTATTACCGCAAACAAATTTTATTAGTAAAAAATTTAAATTACAAAGT
CCATATTCAAAACACACAAAATGTTATATAATCTATATATACATCATATATTAAAAAGTACAACATGATAAATAAATTTAA
CGATGCTTCTAAAGTGAAGTTTTACTTCGAAAAAATTTATTCGTTATTTTTTCCAATGTATTTGACAATTCTCTGATGCAA
AATATGTGTTTGGTTAGCACACATCATTTAGTCTATATTCCATAAAAAAACTTCAAAATAAATTTTATTAACTGTGGTCTT
CCATCTTATCTCTTTCACTATTCTTGTCTTTAGGTGGTCGGTCATAGCGAGACATTTACCTAGAAGAACCGACAATGAGAT
CAAGAATTACTGGGAACACGCATCTCAAAAAACGTCTGATCGAACAGGGTATTGACCCCGTGACTCACAAGCCACTAGCTT
CTAATACTAGCCCTACTGCCACCACCACGCAGCCTGGGAACTTGCATTCTCTAGATGCATCTAGTTCCGACAAACGATACT
CCCGGTCAAGCTCAATGCCTTCCATGTCTCGTCCTCTTTCCTCCGGTTTCAACCCATCTTCCGAGATCACCACACCAGTTC
AGGGAGGTTCCTTGAGTTGCAAGAAAAGTATTAAGAAATCGAACTCTACATCAAGGCTTTTAAACAAAGTTGCGGCTAAGG

CCACTTCCATCAAAGATATCTTGAATGCTACTACAATACCATATGAAAGCTTTTCCAATGGCTTCCCTGAGCAGACTCGCA
ATGAAGAAGATAGTTCCAATGCTCTCGCTGACTTCGATCCCTTTTCTCAATCATTGTTGTACACTGATCACGAGATACATG
CTACTTCTGATCTCGATATGGATCAGGGTTACGATTTCTCATATTTTCTCGAGACACTCGGGAGAGATGAGCACAACATGA
ATGTCGATCAGTATGGTCATGATCTTCCTATGTCAGATGTGTCACAAGACGTCGTCTCATCAACTAGCGTTGATGATCAAG
ACAATATGACTGGAAACTTCGAGGGTTGGTCAAATTATCTTCTTGACCATGCTGGTTTGATATATGACGACACTGAATATG
ATTTCTTCGATAAGCAAATCATATGA

LOC106369912

ATGTCAAGAAAGCCGTGTTGTGTCGGAGAAGGGTTGAAGAAAGGTGCATGGACCACCGAGGAAGACAAGACGCTCATCTCT
TACATTCACGAACATGGAGAAGGTGGCTGGCGCGACATTCCCCAAAAAGCTGGTTAGTATATATATAGTATACACAGGCAT
GTATAATTAGTATTTATTGTATAATATTTTGATGAGAATATGGTGTTTAATTAGGGTTGAAAAGGTGTGGAAAGAGTTGTA
GACTGCGATGGACTAACTACCTAAAACCTGAGATCAAAAGAGGCGAGTTTAGTTCAGAGGAGGAACAGATTATCATCATGC
TCCATGCTGCTCGTGGCAACAAGTACGTTTATTTTTAACAAAAAGGACTATTATATGTTTTTTATGTGCGTATGAATCCTC
CGGTAATCATCATTCTAGTTGTCTCTTTTTTTTTTATTACCGCAAACAAATTTTATTAGTAAAAAATTTAAATTACAAAGT
CCATATTCAAAACACACAAAATGTTATATAATCTATATATACATCATATATTAAAAAGTACAACATGATAAATAAATTTAA
CGATGCTTCTAAAGTGAAGTTTTACTTCGAAAAAATTTATTCGTTATTTTTTCCAATGTATTTGACAATTCTCTGATGCAA
AATATGTGTTTGGTTAGCACACATCATTTAGTCTATATTCCATAAAAAAACTTCAAAATAAATTTTATTAACTGTGGTCTT
CCATCTTATCTCTTTCACTATTCTTGTCTTTAGGTGGTCGGTCATAGCGAGACATTTACCTAGAAGAACCGACAATGAGAT
CAAGAATTACTGGAACACGCATCTCAAAAAACGTCTGATCGAACAGGGTATTGACCCCGTGACTCACAAGCCACTAGCTTC
TAATACTAGCCCTACTGCCACCACCACGCAGCCTGGGAACTTGCATTCTCTAGATGCATCTAGTTCCGACAAACGATACTC
CCGGTCAAGCTCAATGCCTTCCATGTCTCGTCCTCTTTCCTCCGGTTTCAACCCATCTTCCGAGATCACCACACCAGTTCA
GGGAGGTTCCTTGAGTTGCAAGAAAAGTATTAAGAAATCGAACTCTACATCAAGGCTTTTAAACAAAGTTGCGGCTAAGGC
CACTTCCATCAAAGATATCTTGAATGCTACTACAATACCATATGAAAGCTTTTCCAATGGCTTCCCTGAGCAGACTCGCAA
TGAAGAAGATAGTTCCAATGCTCTCGCTGACTTCGATCCCTTTTCTCAATCATTGTTGTACACTGATCACGAGATACATGC
TACTTCTGATCTCGATATGGATCAGGGTTACGATTTCTCATATTTTCTCGAGACACTCGGGAGAGATGAGCACAACATGAA
TGTCGATCAGTATGGTCATGATCTTCCTATGTCAGATGTGTCACAAGACGTCGTCTCATCAACTAGCGTTGATGATCAAGA
CAATATGACTGGAAACTTCGAGGGTTGGTCAAATTATCTTCTTGACCATGCTGGTTTGATATATGACGACACTGAATATGA
TTTCTTCGATAAGCAAATCATATGAATCTTTTTATATTCGAACTGA

LOC106439923

ATGTCAAGAAAACCGTGTTGTGTCGGAGAAGGGCTGAAGAAAGGCGCATGGACCACCGAAGAAGATAAGAAACTCATCTCT
TACATCCACGAACATGGAGAAGGAGGCTGGCGCGACATTCCCCAAAAAGCTGGTTAAATATCTATGAAATTATACACATGC
TTATTTAGTAATCGTACTTGTATGAAATTCTAATTAATTATGATTATGGTATGTAATTAGGGTTAAAAAGGTGTGGAAAGA
GTTGTAGACTGCGTTGGACTAACTACCTAAAACCTGATGTCAAAAGAGGCGAGTTTAGCTCAGAGGAGGAACAGATTATTA
TCATGCTTCATGCTTCTCGTGGTAACAAGTACGTTTATATGCTTAAATGTGTATGTGTATGCATGGATCCTCCAATTAATA
ATGAAATGCATGAAAAAGTTTCACATATTACTTTAAAGATACGATCATTATTAGTTAATATAGTCCCTTCCGCGAAAAGCT
CCACAAAGGAATCATATTAATACCGTCTTTGATATGATCTCAATCTCCCTTCTTGTCTTTAGGTGGTCGGACATAGCGAGA
CATTTACCTAGAAGAACAGACAATGAGGTCAAGAATTACTGGAACACTCATCTTAAAAAACGTTTGATCGAACAGTGTATT
GATCCCGTGACTCACAAGCCACTGGCTTCTAATTCCAACCATACTGTCAACACGCCTCCAGAGAATTTGCATTCCCTTGCT
GCGCCTAGTTCCGACAAGCAATACTCCCGGTCGAGCTCAATGCCTTCCCTGTCTCGTCTTACCAACAAAGATGGGACACCA
GTTCAAGGCGGTGCCTTGAGTCACAAAAAACGTTTCAAGAAGTCGAGTTCTACATCAAGGCTTTTGAATAAAGTTGCGGCT
AAGGTCACTTCTGTAAAAGAAATATTGTCGGCTTCCATGGAAGGTAGCTTGAGCGCTACTACATTACCATATGCAAGCCAT
TCTAATGGCTTCTCTGAGCAGATTGGCAATGAAGAGGATAGTTCCAACGCGTTCCTGACAAATACTCTCGCCGAGTTCGAT
CCCTTCTCCCAATCACCGTTGTACAGTGAGCATGAGATCAACGCTACTTCTGATCTCGGTATGGATTACGATTTCTCACAT
TTTCTTGAAAAGCTTGGGAGAGATGACCACAACGAGGAGAACGATATGAATGTCGAGTATGGTCATGATCTTCTTATGTCT
GATGTGTCTCAAGAAGTCTCATCAACTAGCGTTGATGATCAAGACAATATGATTGAAAATTTCGAGGGTTGGTCAAATTAT
CTTCTTGACCATGCGGATTTCGTATATGACACGGAGTCTGA

LOC125577744

ATGTCAAGAAAGCCATGTTGTGTCGGAGAAGGGCTGAAGAAAGGGGCATGGACCACCGAGGAAGATAAGAAACTCATCTCT
TACATCCATGAACATGGAGAAGGAGGCTGGCGCGACATTCCTCAAAAAGCTGGTTAATATCTATTATATATATTTTGGTAA
ATTTTTAAAACATATATGTTTGTTTGGTATTTGATGTATGAAAGTTTTATATTGAATGTGGTGTTTTACTAGGATTGAAAA
GGTGTGGAAAGAGTTGTAGACTGCGATGGACTAACTACCTAAAACCTGAAATCAAAAGAGGCGAGTTTAGTTCAGAGGAGG
AACAGATTATCATCATGCTTCATGCTTCTCGTGGAAACAAGTAAGTATATATGTGTGTATGGGTCCTCCAATGATTATCAT
TCTGATTTTGTTTCTTAGTTTCAAAATAAATTTTATTAACTTTGGTCTTCCATCTTATCTCTTTCACTATTCTGGTCTTTA
GGTGGTCGGTCATAGCGAGACATTTACCTAGAAGAACAGACAATGAGATCAAGAACTACTGGAACACACATCTCAAGAAAC
GTTTGATCGAACAGGGTACTGATCCCGTGACTCACAAGCCACTAGCTTCTAATACAAACCCTAGCTACTGTACCTGA

Sequences in different color: : green - start codon; grey – exon; white – intron; red – stop codon.

Publishing and archiving

Approved students' theses at SLU are published electronically. As a student, you have the copyright to your own work and need to approve the electronic publishing. If you check the box for **YES**, the full text (pdf file) and metadata will be visible and searchable online. If you check the box for **NO**, only the metadata and the abstract will be visible and searchable online. Nevertheless, when the document is uploaded it will still be archived as a digital file. If you are more than one author, the checked box will be applied to all authors. You will find a link to SLU's publishing agreement here:

- https://libanswers.slu.se/en/faq/228318.

☒ YES, I/we hereby give permission to publish the present thesis in accordance with the SLU agreement regarding the transfer of the right to publish a work.

☐ NO, I/we do not give permission to publish the present work. The work will still be archived and its metadata and abstract will be visible and searchable.