Harnessing TILLING and TILLING by Sequencing (TbyS) - A reverse genetics approach to identify novel mutant lines in *Camelina sativa* L. with improved traits

Volume I of I

Pullanchyottu Kizhakkeveettil Anish Kumar

A thesis submitted to National University of Ireland, Galway

For the degree of Doctor of Philosophy

College of Science, School of Natural Sciences

Discipline of Botany and Plant Science

Under the supervision of Prof Charles Spillane

Head of Botany and Plant Science

January 2018

Table of Contents

DECLARATION

The thesis entitled ‘Harnessing TILLING and TILLING by Sequencing (TbyS) -A reverse genetics approach to identify novel mutant lines in *Camelina sativa* L. with improved traits’ is my own research and has not been submitted for another degree, either at National University of Ireland, Galway or elsewhere.

Signed: 

Pullanchyottu Kizhakkeveettil Anish Kumar

DEDICATION

To my dad…

ACKNOWLEDGMENTS

First of all I would like to thank GOD, who gave me the Peace of mind and the Will-power for achieving various goals in my life.

I would like to thank my supervisor Prof. Charles Spillane for his direction and guidance throughout my PhD research. He has been a kind and cooperative guide who provided me relative freedom to enhance my scientific skills and always encouraged my collaborative efforts with other researchers.

The next person whom I am very grateful is my co-supervisor and Founder and Director of Benchbio Pvt. Ltd., Dr. Manash Chatterjee, who had given me the opportunity for taking up this project. He was there with me from day one and had been very supportive for each and every instance of the progress of this project. His encouragement during each challenges during the project, had been my backup support which lead to the completion of this project.

I thank Dr. Abdel Bendahmane, Director, INRA-IPS2, who gave me the opportunity for training of TILLING and TILLING by Sequencing training. His critical interpretations of my results helped me to learn the trouble shooting of the techniques and process.

I would also like to thank Ms. Christelle Troadec, Technical officer at INRA, who had been a great teacher during the training of TILLING and TILLING by sequencing training. Her kind and supportive nature helped to achieve my results.

Now, I want to thank Dr. MSL Karuna and Mr. Srikanth Kaduri at IICT, Hyderabad, for the GC work of my camelina mutant samples. Dr. Karuna helped me to interpret the GC results and helped me to achieve a meaningful result for the project.

Special thanks to past and present members of my lab at NUIG, Ireland; Benchbio, India and INRA, France. Beginning with Kristina, Dawn, Galina, who helped with me for the admin activities at NUIG. Peter Mckneown, Post-Doc at NUIG, who always agreed and helped me to evaluate my manuscripts, reports and thesis. Thanks to my GRC members Ronan, Zoe, Peter. Special thanks and support from my lab members at NUIG – Sandesh, Simrat, Reetu, Mark, Mercy, Attia, Antoine, Kenneth, Ronan, Mohan, Xiaozhen, Alberto and especially Peter Ryder

had been immense and gave me cherishable memories during my study. And also special thanks to my colleagues at Benchbio – Vaibhav, Seema, Nikita, Riti, Rahul, Fagun, Paul, Sanjay, Bipin bhai, Jinal, Reetu, Nirali, Pooja, Tajo sir, Dr. Anjan, Mrtiyunjay, Uday, who were there every time when I was in need of some thing. I would also like to thank Lab members at INRA – Fabien, Marion, Adnane, Christine, Farhaj, Pierre, Rafael, John, Biradar, Ibrahim, Aimen, who helped me during my work and stay at INRA at Evry and Saclay in France.

I am very grateful to the College of Science, for providing me the Post-Graduate Research Scholarship at NUI, Galway, Ireland and also the financial support from Benchbio Pvt. Ltd., India.

And finally, my dear family members who were there with me at all my time of happiness and sorrow. Especially my father who is no more in this world to see this, but still his blessings are there which is still keeping me to go for the challenges in life. My mom, whose sacrifice and immense love and care helped me to reach at this stage of my life. My wife Priya, whose love and compassion helped me at every moment of my life during hard times and good times. My kids Ananya and Kuttan, who gave me the happiest moment in life. My brothers Roopesh and Kalesh and their family who helped me at all time when I am in need of their help.

PAPERS PUBLISHED

- **Kumar APK**, Troadec C, Ryder P, Karuna MSL, Kaduri S, Yelchuri V, Barot V, Vashi R, Desai R, Gondalia N, Sharma F, Boualem A, Sarkar A, Bendahmane A, Chatterjee M, Brychkova G, McKeown PC, Spillane C. (2017) C-MART: Camelina Mutant population And Reverse genetics Tool for crop improvement allows development of zero erucic acid lines of Camelina sativa. BMC Plant Biology (under review).
- **Kumar APK**, McKeown PC, Boualem A, Ryder P, Brychkova G, Bendahmane A, Sarkar A, Chatterjee M and Spillane C. (2017) TILLING BY SEQUENCING (TbyS) pipelines as high-throughput tools for functional genomics in crops. Molecular Breeding. 37:14 .
- Boualem A, Fleurier S, Troadec C, Audigier P, **Anish PK Kumar**, et al. (2014) Development of a *Cucumis sativus* TILLinG Platform for Forward and Reverse Genetics. PLoS ONE 9(5): e97963. doi:10.1371/journal.pone.0097963
- **Anish PK Kumar,** Bhattacharya A, Dutta OP and Chatterjee M. Allele Discovery Platform (ADP) in Papaya (*Carica papaya* L.). Plant Genetics and Genomics: Crops and Models. Genetics and genomics of papaya. Springer. Chapter -23, Volume 10 2014.
- **Kumar APK,** Boualem A, Bhattacharya A**,** Parikh S**,** Desai N**,** Zambelli A**,** Leon A**,** Chatterjee M and Bendahmane A. (2013) SMART – Sunflower Mutant population And Reverse genetic Tool for crop improvement. BMC Plant Biol. 13:38.
- Ryder P, **Anish PK kumar**, Chatterjee M, McKeown PC and Spillane C. (2012) Identification of variations in Camelina sativa for the generation of favourable fattyacid profiles for biofuel production. IPSAM 2012.

PRESS RELEASE

- 'New Non-GM Technology Platform for Genetic Improvement of Sunflower Oilseed Crop' May 13, 2013. Science daily. http://www.sciencedaily.com/releases/2013/05/130513123223.html

WORKSHOPS ATTENDED

- Attended Workshop under Bilateral co-operation between Germany and India on Seed Development held on 23-24th May 2016 at NASC complex, New Delhi, India.
- Presented 'TILLING: An alternative non-GM tool for Crop Improvement' at the second National workshop on TILLING on crop plants held on 11th December 2015 at Repository of Tomato Genomics Resources(RTGR),University of Hyderabad, India.
- Presented the progress report of CEFIPRA project (Indo-French) in Bangalore on 17th November 2015.

ABSTRACT

Camelina sativa L. is a neglected biofuel crop with potential for tremendous application in various industries due to its varied oil traits, which increases the scope of crop improvement. Camelina is an allo-hexaploid with a chromosome number of 40 (2n) and a genome size of 642MB. We evaluated TILLING technique to carry out Camelina crop improvement by focusing on genes controlling overall oil quality and other non-nutritive traits. TILLING by enzyme was done in *FAD2* or fatty acid desaturase 2, which controls the conversion of oleic acids (C18:1) to linoleic acid (C18:2); and *FAE1* or fatty acid elongase 1, which sequentially adds two carbon units to 18-carbon fatty acids thereby forming very long chain fatty acids (VLCFAs). A total number of 22 mutants were obtained, 11 each in *FAD2* and *FAE1* gene. TILLING in *FAD2* and *FAE1* genes showed a mutation frequency rate of 1 per 520 kb and 1 per 682 kb respectively; and an overall mutation frequency of 601 kb. A novel mutant with zero erucic acid was developed by this approach.

TILLING by sequencing (TbyS) was also done for genes affecting the lipid biosynthesis pathway (*FAD2* and *FAE1)* and the function of non-nutritive traits *MIPS1* (myo-inositol-1-phosphate synthase), and *TGG1* (thiogucoside glucohydrolase 1), we obtained 183 mutants from the 2496 M2 mutant population. Mutation frequency in each gene were as follows, for *FAD2* – 1 per 40 kb, *FAE1* – 1 per 40, *MIPS1* – 1 per 106 kb, and *TGG1* 1 per 145 kb giving an overall mutation frequency of 1 per 55 kb. Phenotyping morphologically showed an overall percentage of 9.6% various phenotypes after screening 5049 M2 families, in which a total of 1.1% of M2 families showed albino plants. The mutants generated from the present study have a potential for food and non-food applications.

1. CHAPTER ONE

INTRODUCTION

1.1 GENERAL INTRODUCTION

Camelina sativa L. (commonly known as gold-of-pleasure or false flax) is a low-input non-food oilseed crop (Kumar et al., 2012). Camelina needs very less water and nitrogen for growth as compared to other oil seed crops; it can be grown on marginal agricultural lands and does not compete with food crops (Putnam et al., 1993). It had been used as a rotation crop for wheat, to increase the health of the soil. It has been traditionally cultivated as an oilseed crop to produce vegetable oil and animal feed (http://www.renewableenergyworld.com/articles/2009/06/biofuel-could-lighten-jet-fuels-carbon-footprint-over-80-percent.html).

Camelina has exceptionally high levels (up to 45%) of omega-3 fatty acids and the oil has over 50% of polyunsaturated fatty acids (Betancor et al., 2015). The major components are alpha-linolenic acid - C18:3 (omega-3-fatty acid, about 35-45%) and linoleic acid - C18:2 (omega-6 fatty acid, about 15-20%). Camelina oil is loaded with antioxidants like tocopherols, which is resistant to oxidation and rancidity (Gugel and Falk, 2006). Other components like cholesterol and eicosenoic acid are also present which are not favorable component to be accepted as an edible oil (Waszkowiak and Rudzin´ska, 2014). Chromosome counts in Camelina species have been reported earlier as n=6 or 14 or 2n=12, 26 or 40, with 2n=40 being the most common count (Francis and Warwick, 2009; Gehringer et al., 2006; Mulligan, 2002). *Camelina sativa* genome is found to be an allohexaploid with a distribution of the three gene copies in 20 chromosomes (n=20) with around 642 MB genome size with sequenced by a hybrid solid, Illumina and Roche 454 next-generation sequencing (Galasso et al., 2010; Hutcheon et al., 2010; Kagale et al., 2014).

1.2 Morphology

Camelina sativa is an annual plant of the Brassicaceae family, which is quite vigorous weed small to medium size, sharp, pointed leaves attaining a size of about 8 cm. Stems are stout and have profuse branching pattern, which can attain maximum size of 90 to 100 cm. Leaves are green to dark green with yellow flowers with four petals and show tetracarpelary flowers (Figure 1). Seed pod size range in between 8 to 14 mm and contain about 5-10 seeds, which have the size of about 0.7 mm x 1.5 mm. Seeds mostly contain about 40% oil, as compared to other oil seed crops like soybeans, sunflower and canola (Berti et al., 2011; Putnam et al., 1993). Crop cycle differs in different species of Camelina, but *Camelina sativa* mostly exhibit about 100 to 120

days seed to seed crop cycle and short seasoned crop. Camelina is more adapted to winter climates as it does not need high temperature for seed drying (Putnam et al., 1993).

Cultivation of Camelina seeds started in Europe from Neolithic times and spread across Europe as a seed mixture in flax field. Camelina further spread across North America for the same reason and Camelina oil was used as fuel during the industrial revolution in Europe (Knorzer, 1978). But the importance was not much in the medieval period until the rediscovery of the potential of Camelina as Biofuel as a potential competitor with other conventional biofuel (Jewett, 2015).

a 

b 

Figure 1a. False flax (*Camelina sativa* L. Crantz.) morphological description with stem, leaf, flower, root and seed pods (http://wisflora.herbarium.wisc.edu/taxa/index.php?taxon=2854). b. Camelina variety BBCAM2 in flowering stage with pods.

1.3 Camelina oil profiles

Camelina oil is one having the richest source of Polyunsaturated Fatty Acids (PUFA's) which have numerous applications in various industries like nutraceutical, pharmaceuticals, cosmetics, alternative edible oil (Abedi and Sahari, 2014; Budin et al., 1995), which is mainly attributed towards the high concentrations of essential fatty acids and linoleic acid. These components are having anti ageing effects and improve skin tone that makes it an essential constituent in cosmetic industries (ZIELIŃSKA and Nowak, 2014). Biofuel production from Camelina has increased its importance as an alternative for other conventional biofuel developed from corn and sugarcane; and also Camelina can be easily adapted as an alternative for crop rotations (Keske et al., 2013). Camelina seeds are very small and can be crushed very easily and the oil can be used as biodiesel, which have similar characteristics as aviation fuels (Moser, 2010). About 40% of the remaining Camelina meal is protein rich, which could be an important by-product for poultry industry and as animal feed (Aziza et al., 2010a). The carbon footprint of Camelina-based aviation fuel's life cycle analysis demonstrated that there was a reduction of carbon dioxide emissions by 75% compared to traditional petroleum-based jet fuel (Elgowainy et al., 2012; Shonnard et al., 2010). Camelina oil is environmentally safe, which has also applications in paint and coating industries as well. There is a great potential for the use of Camelina oil as an alternative source and feed stock for biofuel for future use (Agarwal et al., 2010).

1.4 Crop improvement approaches

Crop improvement has been the major building block since humanity learned organized agriculture. The field of agriculture has undergone a massive transition with the inclusion of molecular techniques developed in recent decades. Marker assisted selection (MAS) with Random Amplification of Polymorphic DNA (RAPD), Simple Sequence Repeats (SSR), Sequence Tagged Microsatellites (STMS) etc. has become an important part of modern plant breeding. By 2050 the world population will reach about 9.7 billion, to feed these many people it will be necessary to rapidly develop new techniques for the sustainable intensification of agriculture in this century, together with other social and technological changes (Godfray and Garnett, 2014; http://www.un.org/en/development/desa/news/population/2015-report.html). The need for a 'second green revolution' has been argued to be essential to sustain human society (Jordan, 2002). With the development of new sequencing technologies, a vast pool of various

plant genome data and various phenotype databases are available through internet (Pauwels et al., 2015). Vast pool of new genomes and genes are being sequenced and characterized by various transgenic and non-transgenic techniques. Conventional plant breeding has immensely advantaged by the advent of various molecular techniques, which in turn has lead to the development of enhanced traits like yield, nutritional quality, drought and salt stress tolerance. Gene sequence information of these traits in Arabidopsis has been characterized by transgenic technologies for advancement in other crops (Pérez-de-Castro et al., 2012). There is clearly a need to generate improved germplasm, and theoretically this can come from (a) recombining existing variation in the crop (b) using existing variation in other species which can be introgressed in or moved by GM (c) transfer of favorable traits from very different organisms; (d) introduction of random novel changes in germplasm by mutagenesis; (e) introduction of targeted novel changes by Genome Editing (Hammer and Teklu, 2008; Kole et al., 2015; Telem et al., 2013). Plant mutagenesis has been an alternative for transgenic development for inducing variability. Plant mutagenesis with the help of radiation or chemical mutagens has led to the deletion or introduction of SNPs in the target gene, which is heritable to the next generation and have become alternate techniques for crop improvement (Wang et al., 2012a).

Human endeavor for continuous improvement for various traits of importance lead to the focus mostly on qualitative traits such as high yield and disease resistance (Shi et al., 2009). The discovery of induced mutation along with heterosis was important advances in the field of genetics, which lead to a major boost in the field of plant genetics (Gehringer et al., 2006). The use of high ionizing radiation frequency and application of chemicals to induce mutation has made possible to conduct genetic experiments which were not possible earlier, when natural variations were the only option available. High mutational density caused by chemical mutagenesis was competitive compared to other mutagenic tools like transposon mobilization (Bingham et al., 1981). Knowledge of genetics, including gene function, has been based upon the research based upon induced mutations (Parry et al., 2009). The use of alkylating agents, which mainly cause point mutations, have been the most important method, as the resultant modified and truncated protein has helped to map the gene of importance accurately and to know the function of the resultant protein in a better way (McCallum et al., 2000b). Mutational analysis in the field of biological research has led to a better understanding of the basic process involved in development, cell biology metabolism and even diseases (Stemple, 2004). Mutation breeding

with the help of forward genetics has become an important tool for the selection of mutant lines with improved traits.

1.5 Forward Genetics Vs Reverse Genetics

Forward genetics was used as an important tool for crop improvement, by which the plant breeding sector was benefited hugely by gaining novel traits by phenotypic screening. But the same had major drawbacks which led to the significant evolution of reverse genetics as an improved tool in plant molecular breeding, major ones being the absence of phenotype of many mutations. So, reverse genetics became a significant process by which such drawbacks were solved to get a substantial amount of innovative traits in crop improvement (McCallum et al., 2000b). Further, genome-wide screening for point mutations induced by chemical mutagens or for natural variation have been used for plant functional genomics studies (Henikoff, 2003). Further TILLING was applied on various crops to make an allelic series of a particular gene, even in case of crops with large and complex genome, where these types of allelic series were non-existent like in case of maize (Till et al., 2004). Tilling can be used for functional validation of various agriculturally important genes in several crops, which cannot be transformed easily (Boualem et al., 2014). Chemical and ionizing radiation mutagenesis have been routinely used to generate genetic variations for breeding research and genetic studies. Mutation breeding have been very successful and according to FAO/IAEA Mutant variety database there are more than 3200 officially released mutant varieties from 214 different plant species throughout the world (www-naweb.iaea.org/nafa/pbg/, 2015).

1.6 Approaches of reverse genetics - Mutagenesis

1.6.1 Physical Mutagenesis

Induced mutagenesis in plants, for agricultural or research purposes, dates back to the beginning of the 20th century. Different types of radiation such as X-rays, gamma rays, fast neutrons and Ultra-Violet radiation were used for inducing DNA deletions and aberrations in chromosomal structure in plants (Haliem et al., 2013). Fast neutron irradiation has been used in *Arabidopsis thaliana* for obtaining a genome-wide molecular profile of mutations, which are informative in nature and extent of genome-wide mutation in lines selected on the basis of mutant phenotypes

(Belfield et al., 2012). Various conditional rice mutants such as gain or loss of resistance to blast, bacterial blight etc. have also been identified by irradiation mutation (Wu et al., 2005).

1.6.2 Chemical Mutagenesis

Chemical mutagens have been widely used for creating different varieties of crops, most important of those being ethyl-methanesulfonate (EMS), Sodium azide, Ethylnitrosourea (ENU), Dimethyl sulphate (DMS), hydrogen fluoride and hydroxyl amine. EMS has a peculiar property to convert Guanine or Cytosine to Adenine or Thymine and in rare cases Thymine gets converted to cytosine (Figure 2); whereas in case of ENU it also does transversions in which A:T changes to T:A (Jansen, 1995; Skopek, 1992). EMS has been most widely used as it is most effective and easy to handle, but precautions have to be taken as it is a mutagenic, teratogenic, and carcinogenic. EMS is an organic compound with chemical formula $CH_3SO_3C_2H_5$ and produces random mutations in genetic material by nucleotide substitution; particularly by guanine alkylation leading to point mutations. Guanine alkylation occurs by the reaction of guanine with the ethyl group of EMS, which gets converted into abnormal base O-6-ethylguanine. During DNA replication, DNA polymerases that catalyze the process frequently place thymine, instead of cytosine, opposite O-6-ethylguanine, resulting in subsequent rounds of replication, the original G: C base pair can become an A: T pair (Greene et al., 2003). EMS can induce mutations at a rate of $5x10^{-4}$ to $5x10^{-2}$ per gene without substantial lethality.

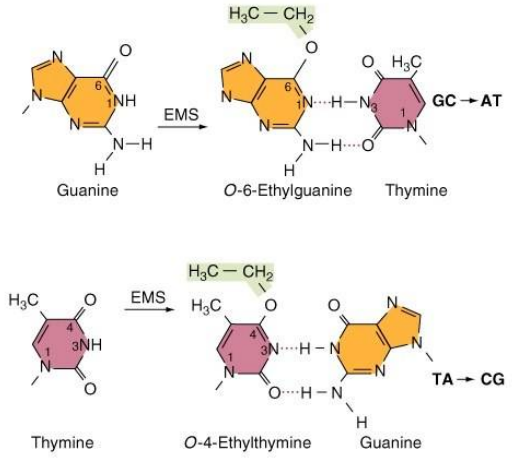


Figure 2. Action of ethylmethane sulfonate. Conversion of Guanine to 0-6-Ethylguanine and Thymine to 0-4-Ethylthyminne in the presence of EMS, results in transition of GC to AT and AT to GC respectively (http://www.mun.ca/biology/scarr/F16-17.jpg).

1.7 Targeting Induced Local Lesions IN Genomes (TILLING)

Chemically induced mutants have become a major resource for reverse genetics studies with the help of the development of TILLING (Targeting Induced Local Lesions IN Genomes) method. TILLING enables the reverse selection of single point mutations by cleavage of mismatches in heteroduplex DNA with the endonuclease CEL I or ENDO I (McCallum et al., 2000a; Triques et al., 2008). This strategy was first applied in an *Arabidopsis thaliana* mutant collection induced with EMS and later in various crops (Colbert et al., 2001; Lai et al., 2012; Tadele et al., 2010; Till et al., 2003). TILLING is a technology to detect induced and natural polymorphisms (SNP) in plants (Colbert et al., 2001). TILLING technology is a high throughput method which helps the detection of rare and recessive mutations (Colbert et al., 2001; McCallum et al., 2000b; Till et al., 2003; Till et al., 2006; Triques et al., 2007). One of the advantages of TILLING method is that it is non-transgenic and significantly cheaper than transgenic technology. A survey on the cost incurred to release a GM event in ten different countries revealed that on average it costs approximately $6-$15 million and 10 years to release a GM event (Kalaitzandonakes et al., 2007)**.** TILLING works by modifying the plants own genome and leads to creation of novel traits at a low-cost and relatively faster product development as compared to transgenic technology. Tilling has been done in model species like Arabidopsis (Lai et al., 2012), Medicago (Porceddu et al., 2008) and Lotus (Perry et al., 2009); vegetable crops like tomato (Okabe et al., 2011), Chinese cabbage (Stephenson et al., 2010), pea (Dalmais et al., 2008) and wild cabbage (Himelblau et al., 2009); oil seeds like rape seeds (Harloff et al., 2012), sunflower (Kumar et al., 2013), soybean (Cooper et al., 2008) and peanut (Knoll et al., 2011); field crops like rice (Till et al., 2007), wheat (Slade et al., 2012), barley (Caldwell et al., 2004), sorghum (Xin et al., 2008) and maize (Weil and Monde, 2007).

Alternative gene editing technologies like CRISPR/CAS9 have been used to induce SNP changes in target gene of interest, to develop transgenics with altered characters like abiotic stress repsonses (Osakabe et al., 2016). Other techniques like TALEN, have been used for targeted gene replacements earlier on tobacco (Zhang et al., 2013). These techniques have advantages as well as there own drawbacks, mainly due to the development of transgenic crops from these

techniques, which inturn have lots of procedural hurdles for approval for marketability in case of edible transgenic crops (Lucht, 2015).

1.8 TILLING in crops

Chemically induced mutants have become a major resource for reverse genetics studies with the help of the development of TILLING – Targeting Induced Local Lesions IN Genomes. TILLING enables the reverse selection of single point mutations by cleavage of mismatches in heteroduplex DNA with the endonuclease, CEL I or ENDO I (Nieto et al., 2007). This strategy was first applied in an *A. thaliana* mutant collection induced with EMS (Colbert et al., 2001; Till et al., 2003). TILLING is a technology to detect induced or natural polymorphisms (SNP) in plants (Colbert et al., 2001). There are three main steps where special care has to be taken are seed purity, mutagenesis quality and size of the mutant population. As TILLING technology is a high throughput method it can detect even rare and recessive mutations (Colbert et al., 2001; McCallum et al., 2000b; Till et al., 2003). A detailed review of TILLING in different crops has been tabulated (Table 1) (Kumar et al., 2017).

The development of a tilling platform will give us an area to explore agriculturally important genes from available genome information. INRA in collaboration with Benchbio has developed a Tilling platform for various crops (http://www-urgv.versailles.inra.fr/tilling/index.htm). Other public institutes and private companies also have their own Tilling platform e.g. UC Davis (http://tilling.ucdavis.edu/index.php/Arabidopsis_TILLING) and RevgenUK (http://revgenuk.jic.ac.uk/about.htm). More than 50 agriculturally important crops have been so far tilled with more than 100 varieties used for the same. With advent of Next Generation sequencing and RNA sequencing, the sample screening time has come down tremendously. When the sequencing technology was introduced in several genes 1 mutation per 200-500kb was observed in polyploids like wheat (Kurowska et al., 2011). Increase in dataset of natural nucleotide variations generated by NGS approaches and will add new alleles to functional genomics and will be an added advantage for detect a large number of induced mutations (Wang et al., 2012b). Next generation sequencing technology was successfully utilized in *Brassica napus* and the identification of rare mutations even polyploids (Gilchrist et al., 2013).

1.9 TILLING in Polyploids

TILLING in polyploids had been a challenge, as the function of a gene in which the mutation is induced in one copy is sufficed by other homologue of the same gene, thereby leading to limited change in the gene function (Schnable, 2015; Wells et al., 2014). Polyploidization has happened mostly by breeding events between related species or by duplication of diploid genomes (Comai, 2005). *Camelina sativa* L. exhibits a hexaploid genome and have well differentiated three gene copies, which are observed in similar fashion in case of wheat, cotton and other polyploid Brassica species (Kagale et al., 2014). TILLING in polyploid crops has given numerous traits of varied importance, which are discussed in the following sections. *Brassica rapa* has a diploid genome, but it is a palaeopolyploid i.e. triplication resulted into three copies of the same gene. Tilling in *B. rapa* var. R-o-18 was done with EMS doses of 0.3% and 0.4% and a mutant population size of 6912 M2 and 2304 M2 was raised. The mutant density was observed as 1 in 56 kb for the 0.3% population and 1 in 67 kb for the 0.4% population (Stephenson et al., 2010). More than 26 different genes were Tilled in *B. napus* (canola cv. *DH12075*) and obtained 432 unique mutations, which led to a density of 1/56 kb to 1/308 kb, with an average of 1/109 kb by utilizing next generation sequencing technology in TILLING (Gilchrist et al., 2013). *Brassica napus* commonly known as rape seed was tilled by creating a population of variety Ningyou7, 1344 M2 mutagenized with 0.3 and 0.6% of EMS (Wang et al., 2008). The gene selected for tilling was FAE1 which controls the seed erucic acid biosynthesis and the mutation frequency was found to be 1 per 41.5kb. *B. napus* was also tilled in two different varieties ie., Express 617 and YN01-429 by creating a mutant population with EMS dose of 0.5-1.2%. Mutation frequencies ranged from 1 per 12 to 1 per 22 kb for the Express 617 population and from 1 per 27kb to 1 per 60kb for the YN01-429 population. About 683 mutations were obtained in the four different sinapine genes, the sequence information of which was derived from BAC library screening and expression studies (Harloff et al., 2012).

1.10 TILLING in Camelina

Strategy for TILLING in Camelina in the present study to create novel traits of Industrial importance was taken from the polyploids which had been already TILLED for different traits. Biofuel developed from Camelina has been utilized for varied use for various industries, our

focus also is to develop Camelina mutants which shows an improved oil quality which could be utilized for applications like Biofuel development and improving edible oil qualities. Genes which were targeted in the present study were mainly *FAD2* (fatty acid desaturase2) and *FAE1* (fatty acid elongase 1), which had been very well characterized in earlier works and have been found to be potential target for improved Biofuel (Hutcheon et al., 2010). Transgenics have been developed with large amount of accumulation of docosahexanoic acid (DHA) and eicosapentanoic acid (EPA) in Camelina, which could be used as an edible option instead of fish oil. Similar works for improved Camelina oil were achieved by genetic engineering of *FAD2* and *FAE1* for improved biofuel (Napier et al., 2014). Earlier work to develop zero erucic lines have been done by targeting *fae1* gene, which was a resultant of four base deletions in some cultivars of *Brassica napus* L. (Wu et al., 2008). RNAi-mediated silencing of *FAE1* gene also was demonstrated in *Brassica napus* L. and thereby decreasing the erucic content by about 60-90 % levels (Shi et al., 2015; Tian et al., 2011). TILLING was also done *FAE1* in *Brassica napus* and lots of point mutation was tabulated and low erucic lines were identified (Wang et al., 2008).

1.11 TILLING by Sequencing in Camelina

TILLING by sequencing has facilitated targeting multiple numbers of target genes in a short span of time with increased high throughput screening of a large number of mutant population (Guo et al., 2015). Perfect pooling strategy in sequencing techniques like Next generation sequencing and exome capture, will help us to get a double check for the obtained mutants, which will decrease the number of false positives found mostly in sequencing results (Henry et al., 2014). Camelina TILLING by sequencing was done by targeting *FAD2* (fatty acid desaturase 2), *FAE1* (fatty acid elongase 1), *MIPS1* (myo-inositol phosphate synthase 1) and *TGG1* (thioglucoside glucohydrolase 1) on Camelina mutant population developed in the present study. *MIPS1* and *TGG1* genes are those which control the production of phytic acid and glucosinolates respectively, which are the non-nutritive traits mainly because of the presence of the same in the camelina meal, used as animal feed. Down regulation of these genes will lead to the use of these mutants to develop low phytate and glucosinolates camelina meal as animal feeds (Barth and Jander, 2006; Raboy, 2007). Illumina was used to screen rice and wheat mutant population by multidimensional pooling strategy for screening of various genes (Tsai et al., 2011). TbyS was done in screening Tobacco mutant population on genes controlling biomass and secondary

metabolites by 3D pooling strategy (Reddy et al., 2012). Novel pooling strategy like 2D, 3D and 4D formats will help us to attain more confidence of obtained mutants along with the sequence coverage reads and count of specific mutation will help us make great advancements to avoid false positive mutants.

1.12 Future Implications of present study on Camelina breeding

According to United Nations the world population will increase to 11.2 billion in the year 2100 and the main challenge in ahead will be mainly the struggle to meet needs for food and fuel http://www.un.org/en/development/desa/news/population/2015-report.html). The TILLING platform developed in the present study, it will help to improve oil quality of Camelina seeds for better utility as biofuel as well as an edible source by lowering the amount of the non-nutritive compounds. The genome of Camelina has been sequenced and genes annotated, which will allow the mutant population to be screened rapidly (Kagale et al., 2014). The utility of the ancient information of the utilization Camelina oil as fuel and the integration of the genome sequence information has enabled us to combine these applications together for developing new Camelina varieties (Betancor et al., 2015). Transcriptomic analysis in *Camelina sativa* for genes expressed in developing seeds will help us to get an idea to improve further the oil biosynthesis pathway (Abdullah et al., 2016). Other important genes involved in lipid biosynthesis pathway could be tilled to improve the overall oil quality of Camelina for Biofuel. Some gene like *ROD1* gene (reduce oleate desaturation 1) could be improved by TILLING, which desaturation reaction leading to the accumulation C18:2 and C18:3 could be blocked and also this gene is over expressed in plants then increase in biomass and seed yield has been reported (Haslam et al., 2016; Lu et al., 2009; Zhang et al., 2012). Other genes which are directly or indirectly related for oil biosynthesis pathway which could be a potential for future TILLING strategies are *KASIII* (3-ketoacyl-acyl-carrier protein synthase III), *FATA-1* (fatty acyl-ACP thioesterase A), *FATB* (fatty acyl thioesterase B), *SAD* (Δ9-stearoyl-acyl carrier protein desaturase) etc. TILLING could also be done on other non-nutritive gene targets like *SCPL19* (sinapoylglucose-choline O-sinapoyltransferase), UGT84A2 (UDP-glucosyltransferase 84A2), *ALDH1a* (coniferyl aldehyde dehydrogenase) etc. Further Camelina mutants developed in the present in the study could be self-pollinated and homozygous lines could be back crossed and potential hybrids could be made for various industrial applications.

Organism	Crop and Variety	Gene tilled	Function	EMS dose	Mutation frequency	Reference
Arabidopsis	*Arabidopsis thaliana* No-O	*CMT2, CMT3*	Epigenetic silencing	20mM	1/153kb	(McCallum et al., 2000b)
Arabidopsis	*Arabidopsis thaliana* Columbia	*various*	various	40mM	1/300kb	(Greene et al., 2003)
Arabidopsis	*Arabidopsis thaliana* Landsberg erecta	*various*	various	20mM to 50mM	1/89kb	(Martín et al., 2009)
Arabidopsis	*Arabidopsis thaliana* C24	*DUO1, EXO70C2, EXO70H2, APK1b*	Male germ line specific MYB transcription factor	25mM	1/345kb	(Lai et al., 2012)
Barley	*Hordeum vulgare* Optic	*Hin-a, HvFor1*	Hordoindoline-a and Floral Organ Regulator-1	20mM, 30mM	1/1000kb	(Caldwell et al., 2004)
Barley	*Hordeum vulgare* Morex	*HvCO1,Rpg1, eIF4EandNR Bmy1, GBSSI, LDA1, SSI and SSII*	*HordeumvulgareConstans-like 1, barley stem rust resistance protein gene 1, eukaryotic translation initiation factor 4E and nitrate reductase*	10mM of NaN_3	1/374kb 1/520kb	(Bovina et al., 2011; Talame et al., 2008)
Barley	*Hordeum vulgare* Barke	Various	Various	Various	1/179kb	(Gottwald et al., 2009)
Barrel Clover	*Medicago truncatula* Jemalong genotype 2HA10-9-3	*Cyt P450, MtPHY1, MsTI*	cytochrome P450 related to the absence of haemolyticsaponins, the*M. truncatula* phytase gene, and the putative orthologue of the trypsin inhibitor gene from *M. scutellata*	15mM	1/400kb	(Porceddu et al., 2008)
Barrel Clover	*Medicago truncatula* Jemalong line A17	*various*	Various	15mM, 20mM	1/485 kb	(Le Signor et al., 2009)
Camelina	*Camelina sativa*	*FAD2*	Fatty acid desaturase 2	-	1/75 to 1/120kb	(Hutcheon et al., 2009)
Chinese cabbage	*Brassica rapa* R-o-18	*BraA.RPL.a, BraA.RPL.b, RPL.c,.IND.a, MET1.a, MET1.b*	replumless, indehiscent methyltransferase1	20mM to 50mM	1/60kb	(Stephenson et al., 2010)
Lotus	*Lotus japonicus* Gifu	*Nodulation gene*	Nitrogen fixation	60mM	1/502kb	(Perry et al., 2009)
Maize	*Zea mays* B73	*various*	Various	1% 0.0625%	1/485kb 0.93/kb 2.1/kb	(Till et al., 2004) (Weil and Monde, 2007)

Organism	Crop and Variety	Gene tilled	Function	EMS dose	Mutation frequency	Reference
Melon	*Cucumis melo* PI124112, a monoecious line	*CmACS-7*	Sex determination	-	-	(Boualem et al., 2014)
Melon	*Cucumis melo* Charentais	*CmDET1*	light signaling pathway	150mM, 200mM	1/573kb	(Dahmani-Mardas et al., 2010)
		CmDHS	eIF5A activation			
		CmACO1	ethylene biosynthesis			
		CmNOR	fruit ripening process			
		CmEXP1	cell-wall modification			
		CmSP1	Inflorescence development			
		CmTCTP	cell growth			
		CmCNR	fruit ripening process			
		CmSGR	chlorophyll degradation			
		CmACS7	sex determinism			
		CmWIP1	sex determinism			
Melon	*Cucumis melo* Piel de Sapo	*Cm-eIF4E, Cm-eIF(iso)4E*	resistance toviruses	100mM	1/1.5Mb	(González et al., 2011)
		Cm-PDS	carotenoid synthesis			
		Cm-DET1	a negative regulator of light-mediated responses that affects carotenoid and flavonoid pathways in tomato and other crops			
		Cm-ACO1	conversion of ACC into ethylene			
		Cm-NOR	transcription factor related to ethylene-sensitive/insensitive phenotypes mutations of which delay fruit softening in tomato			
		Cm-DHS				
Pea	*Pisum sativum* Terese line M3T-946	*LE*	Gibberellin 3b-hydrolasea	4mM	1/401	(Triques et al., 2007)
		RAMOSUS1	Branching		1/780	
		MAX1	More Axillary Growth		1/725	
		REVOLUTA A	Meristem initiation at lateral position		1/608	
		REVOLUTA B			1/1710	
Pea	*Pisum sativum* Cameor	*PsMetI*	genome hypomethylation	20mM	1/200kb	(Dalmais et al., 2008)

Organism	Crop and Variety	Gene tilled	Function	EMS dose	Mutation frequency	Reference
Peanut	*Arachis hypogaea* L. Tifrunner	*Ara h 1, Ara h2 and AhFAD2*	Allergen gene, synthesis of linoleic acid from oleic acid	0.4%, 1.2%	1/966kb	(Knoll et al., 2011)
Rape seed	*Brassica napus* Ningyou7	*FAE1*	Seed erucic acid biosynthesis	0.3, 0.6%	1/41.5kb	(Wang et al., 2008)
Rape seed	*Brassica napus* YN01-429 Express 617	*BnaA.SGT, BnaX.REF1* *BnaA.SGT, BnaX.REF1*	ripening seeds, manipulating seed sinapine content ripening seeds, manipulating seed sinapine content	0.5 – 1.0% 0.8 – 1.2%	1/27kb to 1/60kb 1/12kb to 1/22kb	(Harloff et al., 2012)
Rice	*Oryza sativa* indica IR64	*pp2A4 and cal7*	serine/threonine protein phosphatase catalytic subunit and callose synthase	0.8-1% 1.6%	0.5/Mb 1/Mb	(Wu et al., 2005)
Rice	*Oryza sativa* Nipponbare	*various*	Various	1.5%	1/294kb	(Till et al., 2007)
Sorghum	*Sorghum bicolor*L. Moench inbred line BTx623	*COMT, MIK1,PHYA*	caffeic acid O-methyltransferase, myoinositol kinase, 1phytochrome A	10 to 60mM	1/526 kb	(Xin et al., 2008)
Sorghum	*Sorghum bicolor*L. near isogenic Moench inbred parent line	CYP79A1 and UGT85B1	Controls cyanogenicglucoside, dhurrin biosynthesis	0.15 to 0.4%	-	(Blomstedt et al., 2012)
Soybean	*Glycine max*	*various*	various	2.5mM 50mM 40mM	1/140kb 1/250kb 1/550kb	(Cooper et al., 2008)
Sunflower	*Helianthus annuus* GV-342	*kasII, kasIII, FAD2-1 AY490791 (LRR)*	fatty acid biosynthetic metabolism synthesis of linoleic from oleic *Plasmo parahalstedii* resistance	0.7%	1/475kb	(Sabetta et al., 2011)
Sunflower	*Helianthus annuus* BBS-1	*Fat-A* *SAD-III*	fatty acid biosynthetic metabolism fatty acid biosynthetic metabolism	0.6%	1/480kb	(Kumar et al., 2013)
Tomato	*Solanum lycopersicum* Red setter	*RIN* *Gr* *rab11a* *Exp1* *PG* *Lcy-b Lcy-e*	ripening-inhibitor ripening of tomato fruit tomato softening control Cellwall hydrolysis carotenoid biosynthesis	0.70%, 1.0%	1/574kb, 1/322kb	(Minoia et al., 2016)
Tomato	*Solanum lycopersicum* Shady lady (Hybrid) & NC84173(inbred line)	*PG, TBG4, LeExp1*	Polygalacturonase, tomato β-galactosidase 4, Expansin 1 - preventing softening of fruits	0.4 to 1.2%	-	(Hurst et al., 2015; McCallum et al., 2011) (Colbert et al., 2011)

Organism	Crop and Variety	Gene tilled	Function	EMS dose	Mutation frequency	Reference
Tomato	*Solanum lycopersicum* M82	*eIF4E, eIF4G*	Virus resistance	0.5%	1/574kb	(Piron et al., 2010) (Menda et al., 2004)
Tomato	*Solanum lycopersicum* Micro-Tom	*SlETR1- SlETR-6*	Ethylene receptor gene	0.5% 1.0%	1/1710kb 1/737kb	(Okabe et al., 2013; Okabe et al., 2011)
Tomato	*Solanum lycopersicum* M82	*SlDET1*	high pigment2 (hp2) tomato	0.5%	-	(Jones et al., 2012)
Wheat	*Triticum aestivum* Express (Westbred) Kronos	*GBSS1* *GBSS1*	granule-bound starch synthase I granule-bound starch synthase I	0.75-1.0% 0.75%	1/25kb 1/40kb	(Slade et al., 2005)
Wheat	*Triticum aestivum* Desert durum variety 'Kronos Hard Red Spring common wheat breeding	*SBEIIaand SBEIIb* *SBEIIa , WKS1 and WKS2*	*Starch Branching Enzyme II* genes *Starch Branching Enzyme II* genes, *Wheat Kinase Start1 and 2*	0.7 to 0.75% 0.9 to 1.0%	1/38kb 1/51kb	(Uauy et al., 2009)
Wheat	*Triticum aestivum* Express (Westbred) Kronos	*SBEIIa* *SBEIIa*	*Starch Branching Enzyme II* genes	0.75-1.0% 0.75%		(Slade et al., 2012)
Wheat	*Triticum monoccum* TA4342-96	*COMT1*, *HCT2*, and *4CL1*	lignin biosynthesis pathway	0.24%	1/92kb	(Rawat et al., 2012)
Wild cabbage	*Brassica oleracea*	*various*	various	0.4%	1/447kb	(Himelblau et al., 2009)

Table 1. List of Tilled crops with details of variety, gene tilled, function, EMS dose, mutation frequency and reference.

2. CHAPTER TWO

Crop improvement of *Camelina sativa* L. by TILLING of oil biosynthesis genes - *FATTY ACID DESASTURASE* (*FAD2*) and *FATTY ACID ELONGASE* (*FAE1*)

ABSTRACT

Camelina sativa L. is an under exploited oil crop which has a great potential for becoming an important source of oils for biofuel and other industrial applications. This Thesis focuses on improving the oil quality of *Camelina sativa* variety BBCAM2 by utilizing the molecular technique TILLING (Targeting Induced Local Lesions in Genomes). TILLING is a high throughput technique, which helps to develop mutant discovery platforms for screening large number of samples at a time. The BBCAM2 line was mutagenized by ethyl methyl sulfonate (EMS) using an optimized dose of 1.2% for 8 hours. Treatment with 12000 M1 plants resulted in 5049 M2 lines. Phenotypes of the M2 seedlings showed that the population was heavily mutagenized, with 9.6% of the seedlings showing a visible phenotype. TILLING was done on BBCAM2 mutant population for the most important genes in the oil biosynthesis pathway i.e. *FATTY ACID DESATURASE*, *CsFAD2* and *FATTY ACID ELONGASE*, *CsFAE1*. Oil extraction was carried out and total lipid concentration of seeds was determined to be 35% in the wild-type line. The fatty acid profile was analyzed by gas chromatography (GC) and was found to contain high content of poly unsaturated fatty acids (PUFAs) but variations were also observed. After sequencing the TILLING mutants, a total of 11 *FAD2* mutants were confirmed in which six were missense and five were silent mutations. Out of the six missense mutants three were predicted to be 'non-tolerated' according to SIFT (Sorting Intolerant from Tolerant) analysis. In the case of the *FAE1* mutants, sequencing identified a total of 11 mutants of which six were silent and five were missense mutants. Out of the five missense mutants, two were non-tolerated according to SIFT analysis. The M2 mutant seeds were grown and M3 seeds collected and further total lipids extractions performed. GC analysis was carried out on the total lipids obtained from the mutants of both *FAD2* and *FAE1*. The total lipid analysis confirmed that we have obtained zero erucic acid (C22:1) lines. The oil profile was profoundly changed when compared to the oil profile of the wild type Camelina BBCAM2 cultivar. A total absence of Very Long Chain Fatty acids (above C21:0) was observed in the *FAE1* mutant GC total fatty acid profile, showing the credibility of TILLING as a crop improvement tool. The *FAE1* mutant also showed an increase in α-Linolenic acid (C18:3) to 35% and Paullinic acid (C20:1) to 20% as compared to the wild-type. The mutation frequency of *FAD2* and *FAE1* TILLING was found to be 1 per 520 kb and 1 per 682 kb respectively and the total mutation frequency in the Camelina

mutant population was found to be 1 per 601 kb. The Camelina as a crop can be used as a model oil crop with the utilization of this mutant population for further crop improvement for various agronomical and industrial uses. Mutants developed from the present work could be used to improve the oil quality of Camelina, mainly for the development of biofuel. Camelina TILLING platform could be utilised by breeders for selecting and stacking the traits with desirable traits for overall crop improvement in the crop. Mutant population developed in the present study will give agronomist and plant breeders to look for both forward and reverse genetic studies for multiple characterisations of useful traits.

2.1 INTRODUCTION

Camelina sativa L., Brassicaceae (gold-of-pleasure, false flax) is a low-input non-food oilseed crop. Camelina needs little water or nitrogen for growth and can be grown on marginal agricultural lands so does not compete with food crops (Yang et al., 2015). It may also be used as a rotation crop for wheat, to increase the health of the soil. It has been traditionally cultivated as an oilseed crop to produce vegetable oil and animal feed (Gehringer et al., 2006). It has exceptionally high levels (up to 45%) of omega-3 fatty acids and the oil has over 50% of polyunsaturated fatty acids (PUFAs). The major components are alpha-linolenic acid - C18:3 (omega-3-fatty acid, about 35-45%) and linoleic acid - C18:2 (omega-6 fatty acid, about 15-20%) (Yang et al., 2016). The oil is also very rich in natural antioxidants, such as tocopherols, making this highly stable oil very resistant to oxidation and rancidity. Vitamin E content of Camelina oil is approximately 110 mg/100 g. High cholesterol and presence of eicosenoic acid (15%) poses a hurdle for its approval as cooking oil (Francis and Warwick, 2009; Mulligan, 2002).

2.1.1 Morphology and oil profiles

Camelina is an annual crop which bears yellow flowers and is tetra-carpellary and on an average the height is about 3 feet with mostly branched pattern. Leaves are mostly serrated and pointed (http://www.susoils.com/Camelina/). The seeds are small and are comparable to mustard in size with average production of 3 tons per hectare and the oil content is about 40% w/v (Vollmann and Eynck, 2015).

Camelina oil, a rich source of PUFA’s has tremendous applications in various industries like nutraceuticals, cosmetics, and toiletries. Camelina oil has excellent effects on skin due to the presence of high essential fatty acids and has been argued to act as an anti-ageing agent (http://dewolfchem.com/wp-content/uploads/2013/10/Camelina_oil_specs.pdf). Camelina has become a good option for crop rotation, which can lead to increased farm income mostly for marginal farmers with less than an acre field holding. Camelina seeds are small and oil can be easily extracted from the seeds with added advantage of producing a protein-rich Camelina meal as well (Matthaus and Zubr, 2000). The extracted oil is quite efficient and can be used as biodiesel or aviation biofuel, which have similar chemical characteristics as other fuels (Lebedevas et al., 2010). Camelina meal has high potential in animal husbandry as an animal and poultry feed (Hutcheon et al., 2010). The US Food and Drug Administration have approved the use of 10% Camelina meal for animal feed

(http://www.fda.gov/downloads/Food/IngredientsPackagingLabeling/GRAS/NoticeInventory/ucm507833.pdf). A life cycle analysis of the carbon footprint of Camelina-based bio-jet fuel concludes that the renewable fuel reduces carbon dioxide emissions by 75% compared to traditional petroleum-based jet fuel (Shonnard et al., 2010). With a variety of non-food usages of the oil such as a drying oil and in environmentally safe painting and coating applications, minimal agronomic input requirement for cultivation makes it a potential crop for use as bio-fuel without interfering with the edible oil trade and competition for available resources (Agarwal et al., 2010). A detailed agronomic characterization of Camelina is provided with detailed studies of morphological characters (Singh, 2013).

2.1.2 The C*amelina sativa* genome

Genetic mapping of the genome of *Camelina sativa* suggests a polyploid or duplicated structure (Galasso et al., 2010; Gehringer et al., 2006), and characterization of two genes in the fatty acid biosynthesis pathway further suggests a hexaploid genome (Hutcheon et al., 2010). Chromosome counts have reported the haploid number to vary from 6, 12, or 14 sometimes 20 and diploid number of 40 being the most common count (Francis and Warwick, 2009; Gehringer et al., 2006). A transcriptome study of seed has also been performed (Nguyen et al., 2013). Expression analysis of the three different homeologs of the same gene *CsFAD2* was found to not to compensate each other in site specific expression. (Kang et al., 2011). Further, a high-quality reference genome sequence for *Camelina sativa* is now available, and indicates the modern genome arose via a whole-genome triplication event (Kagale et al., 2014).

2.1.3 Targets for TILLING in Camelina

Transgenic Camelina seed oil composition has been modified for industrial oil purposes by engineering the lipid biosynthesis pathways which allowed the production of fatty acids in cotyledons (Horn et al., 2013). Camelina oil has also been enhanced for fish-oil like traits with high concentration of sciadonic acid which has pharmaceutical potential (Gonzalez-Thuillier et al., 2016; Usher et al., 2015). Interspecific hybridization between *Camelina sativa* and related wild Brassicaceae has been performed to characterize the risk of pollen-mediated gene flow from transformed Camelina and concluded the risk was very low (Julie-Galau et al., 2014). There are however still many concerns over the acceptability of products obtained from transgenic crops, especially in the case of edible products (Ishii and Araki, 2016). The main objective of our work was to generate non-GM Camelina mutant lines which are

disrupted in the functioning of lipid biosynthesis like *FAD2* gene, which converts oleic acid (C18:1) to linoleic acid (C18:2) by the introduction of double bonds in the endoplasmic reticulum. Similarly *FAE1* gene adds two carbon units to 18 carbon fatty acids thereby forming very long chain fatty acids (VLCFAs) (Figure 3). Three copies have been identified in *Camelina sativa* of *FAD2* gene i.e. *FAD2*-1, *FAD2*-2 and *FAD2*-3 and of *FAE1* gene i.e. *FAE1*-A, *FAE1*-B and *FAE1*-C (Hutcheon et al., 2010; Kang et al., 2011; Ratledge, 2004). Suppression of *FAD2* gene leads to accumulation of oleic acid and has been proved to be more beneficial for the development of biofuels (Hutcheon et al., 2010). Similarly suppression of *FAE1* gene leads to the accumulation of VLFCAs is lowered, which is also an important trait for various industrial applications. Transformation techniques like RNAi approach to down regulate *FAD2* and *FAE1*gene in *Brassica napus* was performed, leading the increase of oleic acid by 85% and reduced PUFA contents like erucic acid (Peng et al., 2010). Therefore, in the present study, we applied the TILLING method to improve the oil traits of Camelina by targeting its *FAD2* and *FAE1* genes with the aim to develop novel mutants of industrial importance.

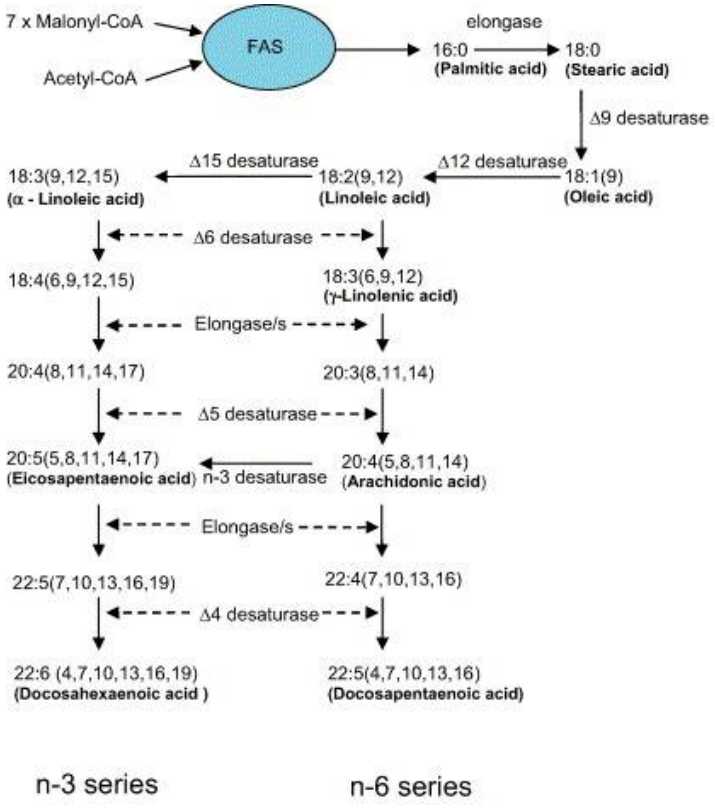


Figure 3. Lipid metabolic pathway in plants shows the function of desaturase and elongases enzymes, which results in the formation of unsaturated fatty acids and very long chain fatty acids (Ratledge, 2004)

2.2 MATERIALS AND METHODS

2.2.1 EMS dose optimization

Seeds of *Camelina sativa* accession BBCAM2, a spring variety from Europe were selected for the experiments (gift by Dr M Chatterjee). A germination test was conducted by placing 100 seeds on germination paper (Royal Crepe Paper, Ahmedabad). After obtaining the duration for initial germination i.e. radical emergence, EMS (Sigma Aldrich) dose optimization was carried out with doses of 0.5, 0.6, 0.75, 0.8, 1.0, 1.2, 1.5, 2.0 and 3.0% v/v and control 0% in 0.1M phosphate buffer at pH 7.0 (DeAngelis, 2007) for 8 hours in incubator shakers at 22°C and continuous shaking at 100 rpm. This was performed in flasks covered with aluminum foil to avoid light exposure. After the treatment, the seeds were washed in running water for 12 hours and were kept in petri plates (Himedia) for one week for observing the radical emergence. Seeds were then transferred to pot-trays for observing the emergence of seedlings (Soil mix ratio of Cocopeat,Vermicompost and Soil – 3:2:1)). The observations were taken for about 30 days.

2.2.2 Mutant population creation

Full scale treatment with EMS was done with about 12,000 seeds of *Camelina sativa* var. BBCAM2 at a dose of 1.2% EMS for 8 hours. The seeds were treated with 0.1 M phosphate buffer with continuous shaking at 100 rpm in a table top shaker. The seeds were thoroughly washed in continuous running water and were incubated for 12 hours at 22°C. The seeds were then sown in nursery beds in a polyhouse for about 30 days. The M1 seedlings were later transplanted in open field with spacing of 30 cm x 45 cm at temperate climatic condition at Kullu (hill station) in March 2012 were the temperature ranges from 20-30 °C. Proper agronomic practices were followed to grow the crop till maturity till August 2012. The M2 seeds were harvested individually from each M1 plant and dried, packeted labeled and stored in a cold room (8 °C with 40-50% relative humidity).

2.2.3 Genomic M2 DNA extraction

DNA extractions of all 5049 M2 family were performed using Qiagen plant DNA extraction kit following manufacturers instruction Ten seeds per M2 families (5049 M2) were sown in nursery pots inside a greenhouse and leaves were collected (4941 M2) from four leaf stage M2 plants. The leaf samples were collected in 96-well collection plates with two 4 mm steel beads (http://www.steelballs.co.in/) per well. The samples were frozen in liquid nitrogen and

stored at -80 ˚C prior to DNA extraction. The protocol used for DNA extraction was that from DNeasy 96-plant kit (Qiagen-Hilden, Germany). The samples were ground to powder with the help of the custom made vibrator shaker (Abraham T, 2009). The extracted genomic DNA was run on 0.8% w/v agarose gel (Invitrogen) to check the quality. The quantity of M2 DNA was quantified with Nanodrop (Implen). The M2 DNA was normalized to 50 ng/µl and 8-fold pooling plates were made for TILLING.

2.2.4 Mutation detection

Nested PCR was carried out by using target gene specific primers and 5ng of Camelina M2 genomic DNA. The first N1 PCR amplified product (1µl) was used by diluting 1:10, which served as a template for the nested PCR, using 5´end infra-red dyes labeled combinations of IRD700 and IRD800 M13 universal primers along with unlabeled N2 primers with M13 sequence and the section of N1 primer sequence (Figure 4).

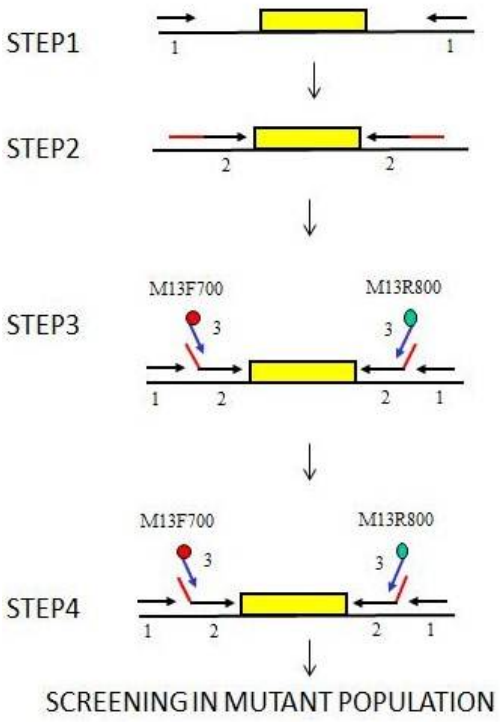


Figure 4. Schematic representation of Nested PCR. Target gene specific N1 primer (http://www.un.org/apps/news/story.asp?NewsID=45165) anneals and amplifies the required region (Step 1 and 2), N2 Primers (Ruddle et al., 2013) which have both M13 universal primer sequence and gene specific region anneals and amplifies the N1 PCR product, further the M13 IRD labeled primers (3) anneals on N2 primers M13 tail sequence, which is carried on the whole mutant population (M2 DNA) for TILLING. Modified from (Triques et al., 2008)

2.2.5 Bioinformatic analysis

The sequences of *FAD2* and *FAE1* from *Arabidopsis thaliana* were utilized to obtain the sequence of their three Camelina homeologs from the Camelina genome database (www.Camelinadb.ca). CODDLE (Codons Optimized to Discover Deleterious Lesions, http://blocks.fhcrc.org/~proweb/input/) was used to ascertain regions of the target gene in which G/C to A/T transitions, which are most likely to result in deleterious effects on the protein, had occurred. PARSESNP (Project Aligned Related Sequences and Evaluate SNPs, http://www.proweb.org/parsenp/) was used to demonstrate the distribution of mutations within the target gene, and to indicate the nature of each single mutation. To predict the impact of the mutation on the target protein sequence, SIFT (Sorting Intolerant from Tolerant, http://sift.jcvi.org/www/SIFT_seq_submit2.html) analysis and PPOPEN software analysis (http://ppopen.informatik.tu-muenchen.de/) was done. ClustalW http://www.genome.jp/tools/clustalw/) was utilized for Multiple Sequence Alignment of the gene of interest. Primers were designed according to the M13 TILLING system for each gene (Triques et al., 2008). TILLING primers for *FAD2* and *FAE1* (Table 2) was designed and further TILLING was carried out (Appendices A1).

S. No.	Target Gene	Abbreviation	Function	Camelina Sequence ID
1	*FAD2*	*Fatty Acid Desaturase -2*	Membrane bound delta-12- desaturase which converts oleic acid to linoleic acid	Csa01g013220.1 Csa15g016000.1 Csa19g016350.1
2	*FAE1*	*Fatty Acid Elongase -1*	Sequentially adds 2 C to 18 carbon fatty acid CoA conjugates, resulting in VLCFAs	Csa10g007610.1 Csa11g007400.1 Csa12g009060.1

Table 2. Target genes for TILLING with gene annotations from Camelina genome.

2.2.6 TILLING workflow

M13 TILLING methodology from INRA-URGV, France (Colbert et al., 2001; Triques et al., 2008; Triques et al., 2007) was utilized to screen the Camelina mutant population. The workflow (Figure 5) for M13 TILLING starts by designing sequence specific primers for the gene which it is intended to till. The PCR condition for the sequence specific reaction is an initial denaturation of 94 °C for 5 min, 35 cycles of denaturation at 94 °C for 10 sec, annealing at 55 °C for 15 sec, and extension at 72 °C for 1 min 30 sec, and a final extension of 72 °C for 5 min and cool down to 4 °C . The Infrared labeled M13 universal primers are multiplexed by using the PCR product of the gene specific reaction as template. The PCR condition for the nested labeled reaction is as follows, initial denaturation of 94 °C for 5 min and 10 cycles of denaturation at 94 °C for 15 sec, annealing at 60 °C for 30 sec, and extension at 72 °C for 1 min 30 sec (specific for the gene sequence), and 25 cycles of 94 °C for 15 sec, 50 °C for 30 sec and extension at 72 °C for 1 min 30 sec (specific for M13 universal primers) and a final extension of 72 °C for 5 min and cool down to 4 °C . The labeled samples are then loaded into agarose gel for observing and obtaining an idea of how much PCR product should be taken for ENDO-I digestion (Serial genetics, France). The samples are passed through a RAMP reaction in which the samples are denatured at 94 °C and gradually cooled to 8 °C at the rate of -0.1 °C per second for enhanced formation of heteroduplex. The PCR product is then used for ENDO-I digestion at 45 °C for 20 minutes and then 5 µl of 0.15 M EDTA (pH-8.0) is added to stop the reaction. The samples are then passed through the sephadex (GE, G-50, medium) for purification of the DNA samples. The Sephadex are prepared before by put equal quantities into the wells of Millipore filtration plate 94 well (Multiscreen-HV, MAHVN4550) with the help of sephadex column loader 45 µl. After transferring the sephadex into the filtration plate, about 325 µl of sterile distilled water is added to make the sephadex swell, which requires an incubation of at least one hour. And then the excess water is centrifuged out by spinning at 500 G for 2 min. Sample plate is fit onto the filtration plate with 5 µl formamide loading dye. After stopping the ENDO-I digestion the samples are transferred into the Sephadex G-50 column and then centrifuged again to obtain the purified sample onto the formamide plates. The filtrate is further dried to 5 µl at 65 °C in a vacuum centrifuge concentrator for about an hour. Further the samples are denatured at 94 °C for 2 min just before loading onto the Genetic Analyzer (Licor, 4300).

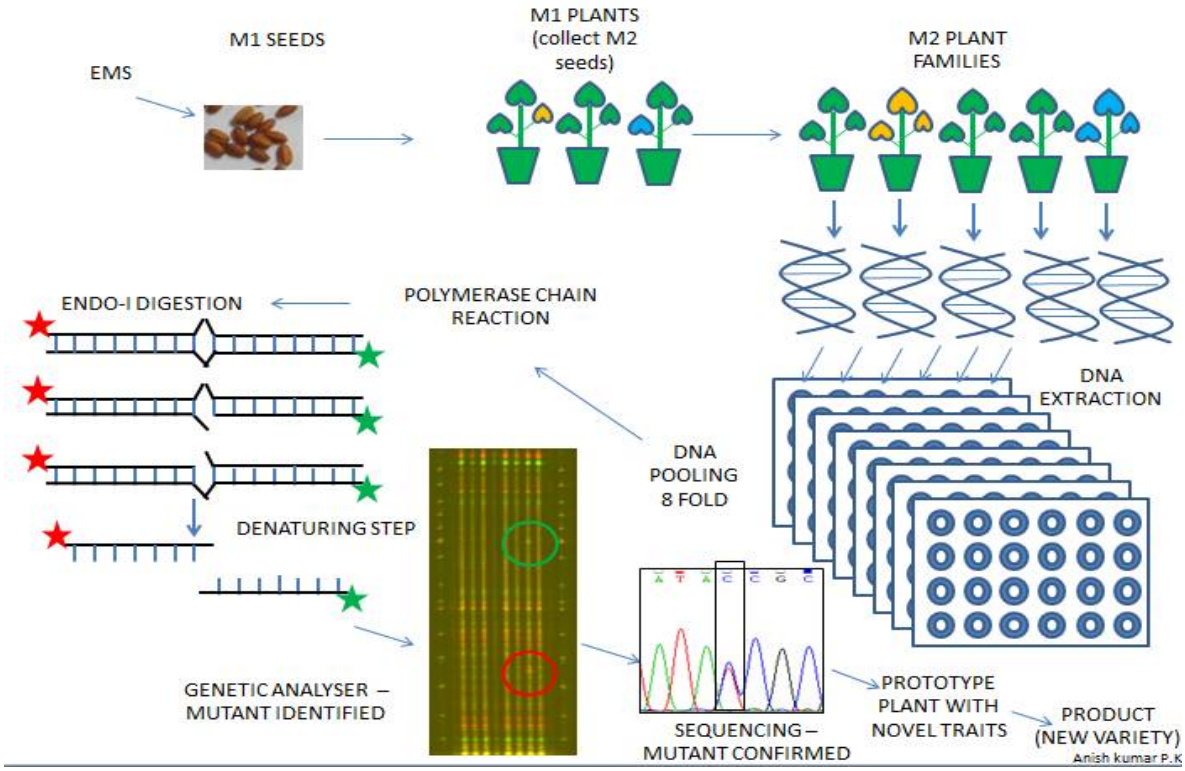


Figure 5. Schematic representation of TILLING technology: 1) EMS seed mutagenesis, 2) M1 and M2 plant production, 3) Genomic DNA bank constitution, 4) Detection of point mutations in the gene of interest. 5) Product/variety development. Modified from (Colbert et al., 2001)

2.2.7 Acrylamide gels on LICOR 4300 system

TILLING gel plates were cleaned thoroughly and the gel plates and rails were paired respectively by taking the front and back plates placed on a flat surface. The plates were sandwiched with 0.25 mm spacers. 17.5 ml of KB^+ LICOR 6.5% gel matrix was taken and 196 µl of 10% Ammonium per sulfate and 19.6 µl temed was added and mixed gently. The mix was taken up in a 20 ml syringe and the gel was squeezed into the back plate by gentle tapping on the front plate. The plastic comb (0.25 mm) was inserted in between the plates and left to polymerize for about 90 min. After the polymerization the casting combs were removed and the glass plates were cleaned thoroughly with Milli-Q water to avoid the persistence of acrylamide gel on the outer surface. Proper cleaning at the laser (camera) position is very essential. The cleaned plates were placed on to the Licor machine and the upper and lower tank was filled with 1X Tris Borate EDTA (TBE) (Licor) buffer. The notch was cleaned with a syringe and needle pre-run was started which goes on for 20 min at the following condition i.e. 1500 V, 40 W, 40 mA and 45 $^{\circ}$C . The denatured samples along with size-markers were denatured at 93 $^{\circ}$C for 3 min and both were placed on ice. Using an 8-channel pipette 1.0 µl of the sample was loaded on to the cold sample-loading tray, which should be kept on ice while loading. The samples were absorbed by inserting the membrane comb in the tray for several seconds until complete absorption the comb was left to remain for 2 min at room temperature for drying. By the mean time the notch was cleaned up with 1X TBE buffer and then the comb inserted rapidly but gently at 45-degree angle with sample one on the left. The system was closed with the tank lids and wire. The run was started with the same conditions as the one done for pre-run for initially for four minutes and then the combs removed and the notch cleaned up with 1X TBE and the run restarted. TILLING images were detected by red and green spots for IRD700 and IRD800 respectively, which are represented by black and white picture for each fluorescent image using software provided with Licor 4300 TILLING was done on *FAD2* and *FAE1* gene on all the seven super pools (SP). And further deconvolution was done on probable mutants. The TILLING mutants were confirmed by sequencing and positions of mutations were confirmed. Further the M2 seeds of mutants of *FAD2* and *FAE1* were grown to maturity and were taken to M3 generation and M3 seeds were collected. Genotyping by TILLING was done to find whether the M3 seeds are Homozygous/Heterozygous/ Wild type for the obtained mutations.

2.2.8 Lipid extraction

Oil extraction from seeds were done with some modifications of Bligh and Dyer's method (Bligh and Dyer, 1959). About 10 seeds with three replicates of *Camelina sativa* variety BBCAM2 and CELINE were lyophilized and were crushed into powder in a mortar and pestle, about 1.0 ml of chloroform: methanol (1:1 v/v) solution (solvent mix) was added to the powder and homogenized. The mixed solutions were transferred into a 5 ml glass tube with screw cap and vortexed for 30-60sec. Another 1.0ml of solvent mix was added to the motor and the wash solution was also added to the same tube to re-extract. Vortexed and centrifuged at 2000 rpm for 10 min and took the lower layer in a new glass tube, removed the top layer. Solids were re-extracted with 0.5ml of solvent mix, and the solvent mix reaction was repeated. The extract was pooled and the lower layer re-extracted. The solution was filtered with whatman filter paper 41 and rinsed with chloroform. The mixture was kept for overnight in continuous shaking at room temperature and the same was transferred to a separating funnel and the mixture was completely emptied into the separating funnel. The lower layer was taken for further overnight drying at room temperature and total lipids were gravimetrically measured.

Total Lipid extraction was also done by using hexane (Akpan et al., 2006) as the solvent mixture. About 50.0 mg of M3 seeds Camelina mutants were lyophilized with liquid nitrogen and was crushed with mortar and pestle to coarse powder and was transferred by washing the powder with hexane (1:10 weight of the seed material/ volume of hexane). The mixture was packed into a thimble in a glass soxhlet apparatus (Jensen, 2007). Hexane was removed using rotary evaporator followed by drying under reduced pressure. Oil was further measured gravimetrically.

2.2.9 Fatty acid methyl esters (FAME) preparation

Fatty acid methyl esters were prepared by trans-esterification of Camelina oil from the M3 mutants seeds obtained after TILLING. The oil was refluxed at 70 °C in 2 % sulfuric acid in methanol (Stemple, 2004). Tridecaenoic acid (C13:0) was used as an internal standard for calculating the amounts of fatty acids. The progress of the reaction was monitored by TLC using solvent system hexane/ethyl acetate (90/10vol/vol). The esters were extracted into ethyl acetate, washed with water until acid free and passed over anhydrous sodium sulfate. The ethyl acetate extracts were further concentrated using rotary evaporator to obtain fatty acid

methyl esters. The converted fatty acid methyl esters were analyzed for its fatty acid composition by gas chromatography.

2.2.10 Gas chromatographic (GC) analysis

GC analysis of the methyl ester was carried out using GC 6890 N series of Agilent and DB-225 column (30 m × 0.25 mm i.d × 0.5 μm, film thickness). The injector and flame ionization detector were at 300°C. The GC analysis was performed by injecting 1μl of FAME into the GC machine. The oven temperature was programmed to be at 160°C for 2 min and then increased to 230°C at 5°C/min with a hold time of 20 min. Nitrogen gas with a flow rate of 1.5 mL/min was used as carrier gas. The peaks were identified by flame ionization detector at 260°C. Supelco FAME mix (C4-C24) containing 36 fatty acids were used to identify the fatty acids. Statistical analysis one-way anova and histograms with error bars were made by software XLSTAT.

2.3 RESULTS

2.3.1 EMS Dose optimization of BBCAM2

EMS dose optimization or Kill-curve analysis of *Camelina sativa* variety BBCAM2 shows a variation in the radical emergence between the control and the treated samples (Figure 6, 7). The seedling germination of 0.5% EMS treated seeds was similar to control seedling emergence, so an optimization of above average was given. The high dose optimization ranging from 1.5% to 3.0% showed that the germination was remarkably delayed but not hampered even in case of doses like 3.0%. Pilot study done on optimized seedlings for field establishment showed that most of the seedlings were not growing beyond vegetative phase. Based on EMS dose optimization data 1.2% EMS dose was selected for the full scale treatment with approximately 12,000 seeds to generate the mutant population.

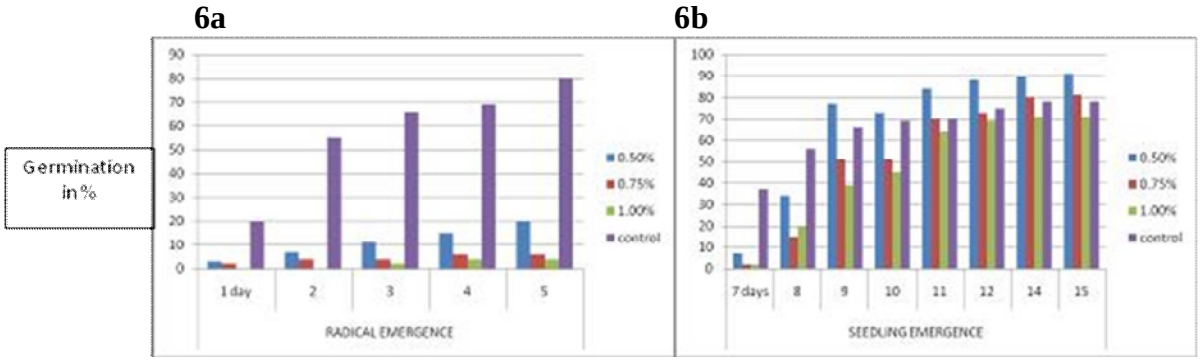


Figure 6. EMS dose optimization for radicle emergence.a)the data shows EMS effect in the emergence of radicle as compared to control and b) after transplanting in pot-trays. The seedling emergence data shows the delayed emergence in 1.0% as compared to low dose i.e. 0.5% and after 15 days, 0.5% and 0.75% germination is higher than control.

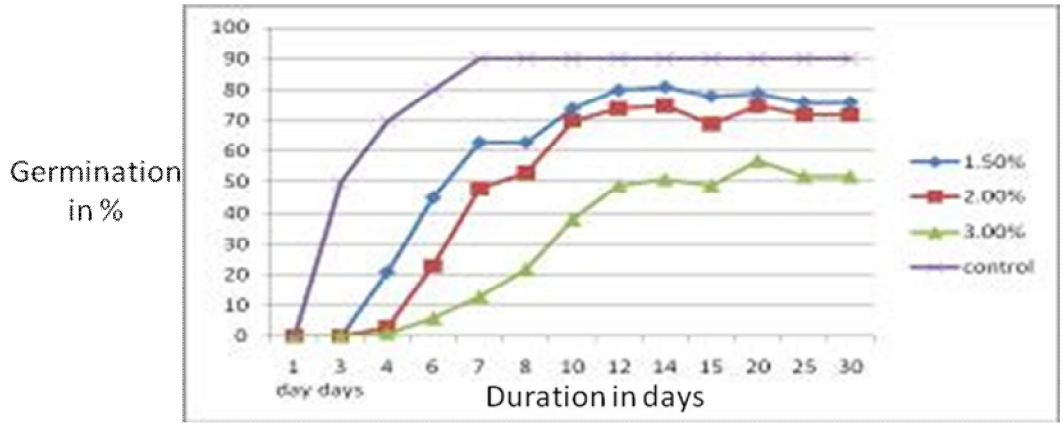


Figure 7. EMS high dose optimization for *Camelina sativa* variety BBCAM2. It shows that following treatment with a 3.0% dose, the emergence of seedlings is much delayed as compared to other doses like 1.5% and 2.0%. The data also shows the seedlings emerged reached 50% in case of 3.0% after 12 days whereas control reached 50% in 3 days

2.3.2 *Camelina sativa* L. mutant population

EMS Treatment of 12,000 seeds (M0) of *Camelina sativa* L. variety BBCAM2 were done and nursery was created inside a polyhouse. A total number of 6000 M1 seedlings germinated which were transplanted, indicating the LD-50 of the EMS dose (1.2% for 8 Hours). The disease incidence was low and showed early bolting (Figure 8a-b) and about 4% seedlings showed early flowering trait (Figure 8c). Phenotypes like albino to pale green plants were observed in the M1 generation (Figure 8d), indicating good mutation frequency in the Camelina mutant population. About 5049 M2 seed packets were made with viable seeds and the seeds were sown for DNA extraction. M2 Phenotypes like albino, variegated, dwarf, and leaf variation were observed in the M2 families (Figure 9). The percentage of M2 phenotypes observed in the mutant population was 9.6%, which shows the high mutagenic effect (Table 3) due to the EMS. This suggests that the EMS mutagenesis applied was efficient in the polyploid genome of *Camelina sativa*. In total 1.1% of M2 plants showed albino phenotypes, which is an indication of lethal mutations in chlorophyll-expressing genes.

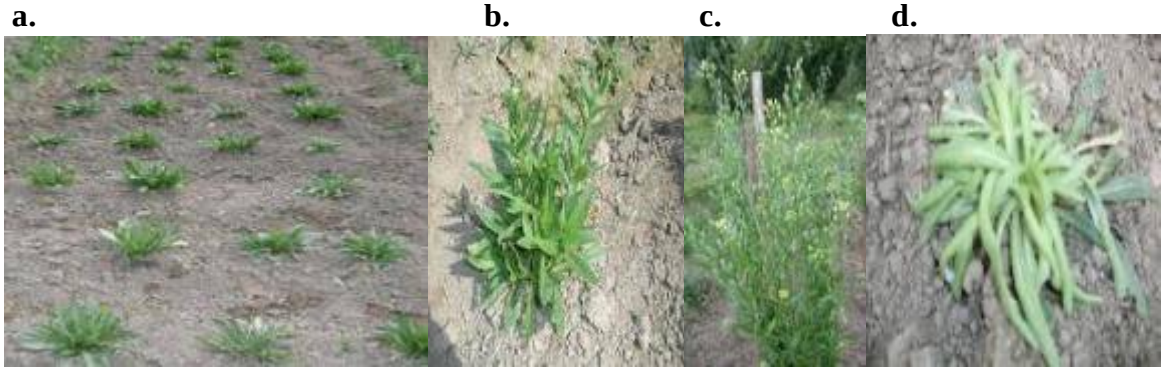


Figure 8a. 1.2% EMS treated M1 plants of *Camelina sativa* var. BBCAM2 in the field showing vegetative stage of the crop. **b**. shows early bolting in M1 plants. **c**. shows early flowering plants with pod formation as compared to other M1 plants. **d**. shows albino to pale green phenotypes which are considered a good indicator of successful mutagenesis.

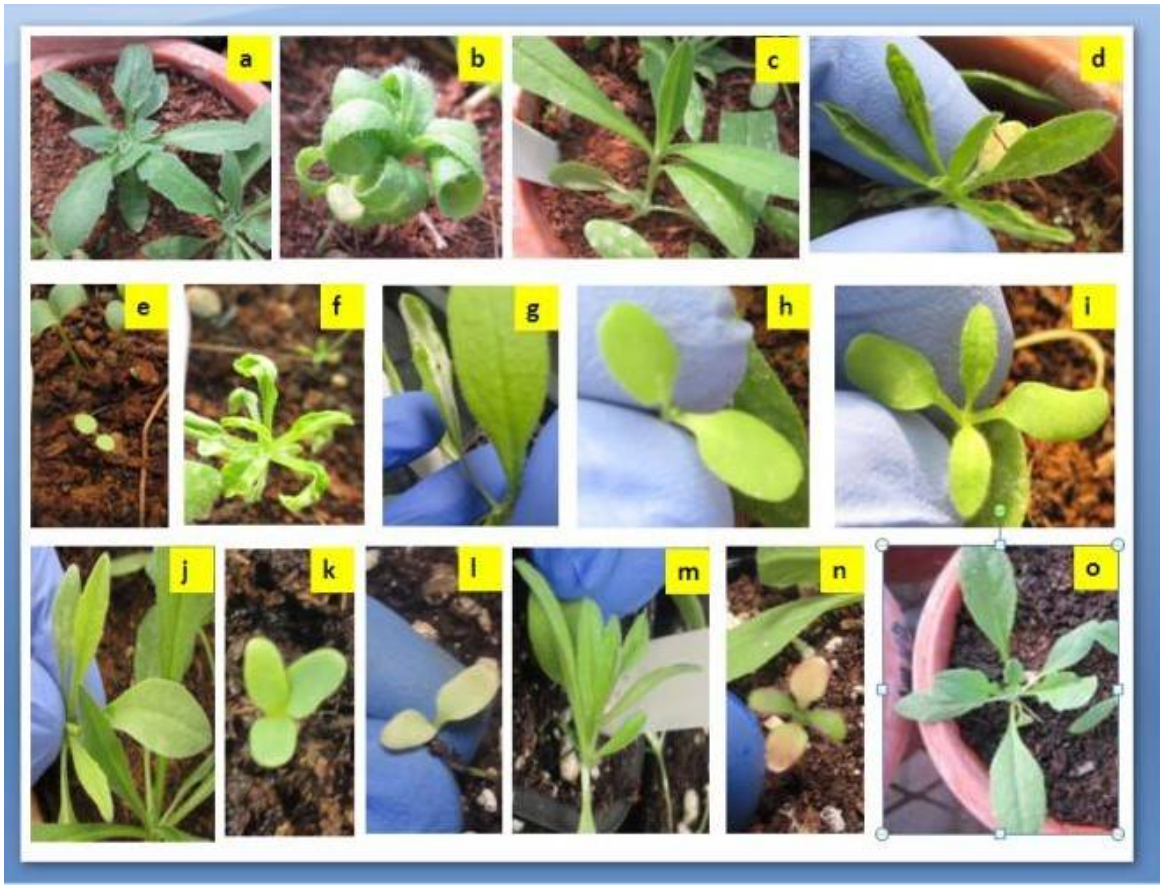


Figure 9. Phenotypes of M2 seedlings of *Camelina sativa* var. BBCAM2.. From left to right: a) Dark green leaves, b) Curly leaves, c) Leaf sheath expanded, d) Abnormal leaf; e) Pale green cotyledon, f) Irregular and curly leaf, g) Variegated leaf, h) Cotyledon arrest, i) Primary leaf arrest, no shoot apex; j) Pale green leaves, k) Triple cotyledon, l) Albino cotyledon, m) Fused leaves, n) Reddish leaves, o) Wild type.

Phenotypes	**Number of M2 lines**	**Percentage of phenotypes**
Albino	57	1.1
Pale green	7	0.1
Dwarf	7	0.1
Single cotyledon	25	0.5
Cotyledon arrest	82	1.6
No germination	108	2.1
Primary leaf arrest	66	1.3
Variegated	5	0.1
Leaf modification	78	1.5
Other phenotypes	48	1.0
Total phenotypes	**483**	**9.6**
Total M2 screened	**5049**	**100.0**

Table 3. Total number of phenotypes and percentage from the Camelina mutant population.

2.3.3 DNA extraction for TILLING platform

DNA extraction of 4941 M2 families were completed and were found to be of good quantity and quality. After DNA extraction, 5 µl each of M2 DNA from each 96-well plate a representation of 12 samples were loaded on an agarose gel to check the quality and the quantity of the same. Gel visualization showed a very good quality DNA (Figure 10). From each sample 1 µl of DNA was loaded in each well and the quantification with Nanodrop (Implen) showed an average concentration of 40ng/µl. A total number of 52, 96-well M2 DNA plates were made and rescue stock and working stock were prepared at 1:10 dilution (Figure 11a). The diluted plates were then pooled 8-fold row wise to make 7 super pool plates for screening of the population (Figure 11b). The Super pool plates were also divided into rescue stock and working stock plates. Two copies of concentrated DNA were made for long term storage for future work.

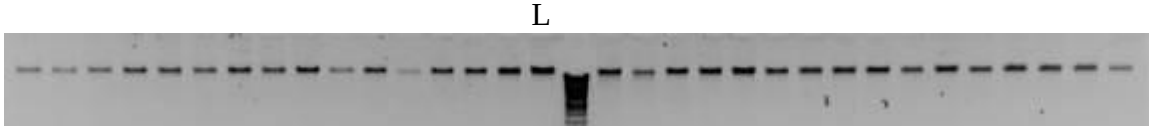
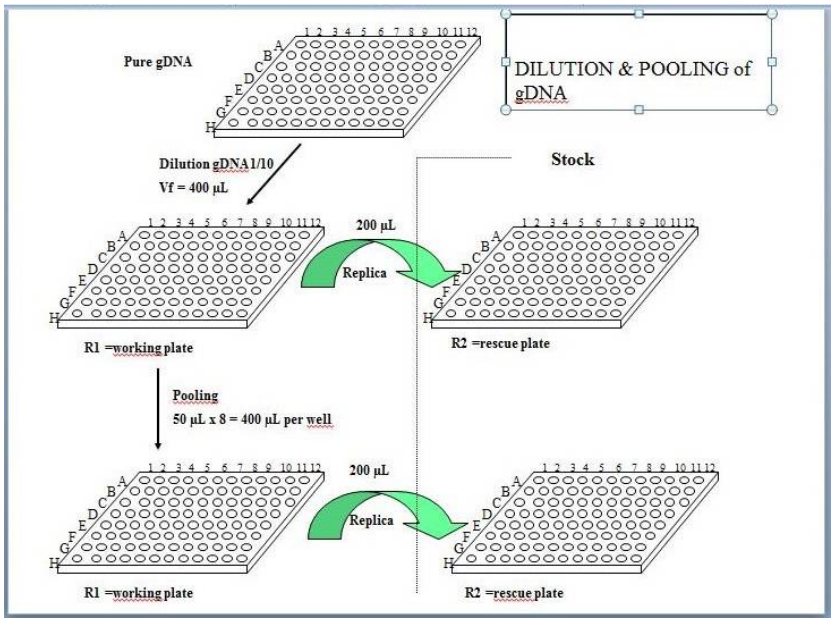
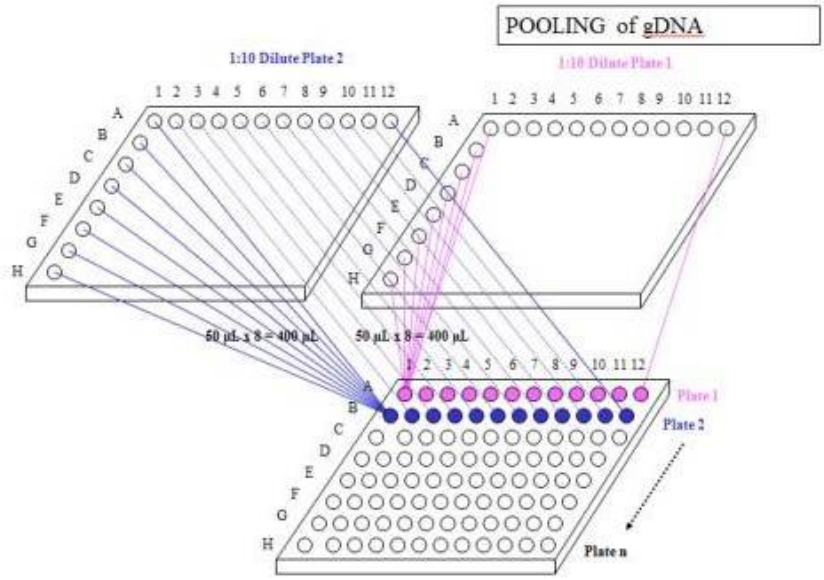


Figure 10. A representation of 5µl Camelina M2 genomic DNA profile on 0.8% agarose gel, Plate-5, 6, 7 and 8 in Ladder (3Kb – max).



a



b

Figure 11a. Dilution of different M2 DNA plates for creating working plate and rescue plate. **b.** Pooling methodology for making 8-fold superpool plates.

2.3.4 Mutant detection by TILLING in *FAD2* and *FAE1*

Validation of the Camelina mutant population was done by screening for mutation in *FAD2* and *FAE1* gene, which can be used as a proof-of concept study to decide the quality of the mutant population. TILLING will help us to estimate the mutation frequency of the population, by which it can be utilized as a source for broader improvement strategy of the crop. *FAD2* or *fatty acid desaturase -2* is a well characterized gene which encodes a membrane-bound enzyme which converts oleic acid to linoleic acid i.e. by introduction of double bonds. In case of *FAE1* or *fatty acid elongase 1,* sequential addition of 2-Carbon to the 18-C fatty acids takes place, which leads to the formation of very-Long chain fatty acids (VLCFA) (Ratledge, 2004). In the present study we have focused on these two genes for improving the oil quality for industrial use.

CODDLE (Codon Optimized to Detect Deleterious Lesion) is a web-based tool that enables the choice of gene region suitable for TILLING (Till et al., 2003). CODDLE results on all the three copies of *FAD2* shows similar version of graphs, which entails the region from 800bp to 1000bp position, is more prone for EMS induced mutations. CODDLE in *FAE1*gene results show that stretch beginning from 1000 bp to 1200 bp is GC rich and is more prone to EMS mutagenesis. The option 'TILLING w/EMS (plants)' favors an amplicon size of 1500, but the initial and final 100 bp cannot be analyzed effectively. Therefore, the new window size has been set to 1300 (Figure 12 and 13). Three isoforms of *FAE1* (*FAE1*-A, *FAE1*-B and *FAE1*-C) gene shows quite similarity as depicted from the CODDLE results.

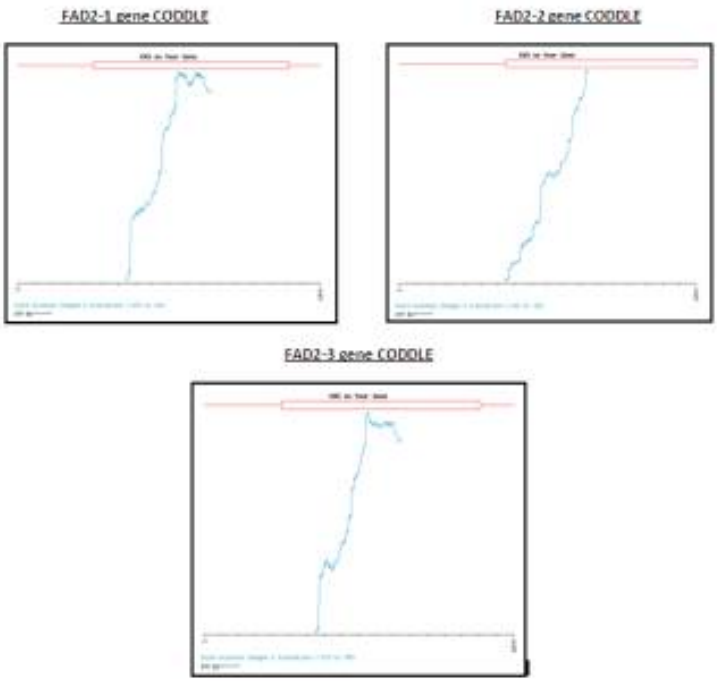


Figure 12. Coddle results for different isoforms of *FAD2* gene, as the *FAD2* is a single exon gene, most probable region for mutations are in the central part of the exon in all the three isoforms.

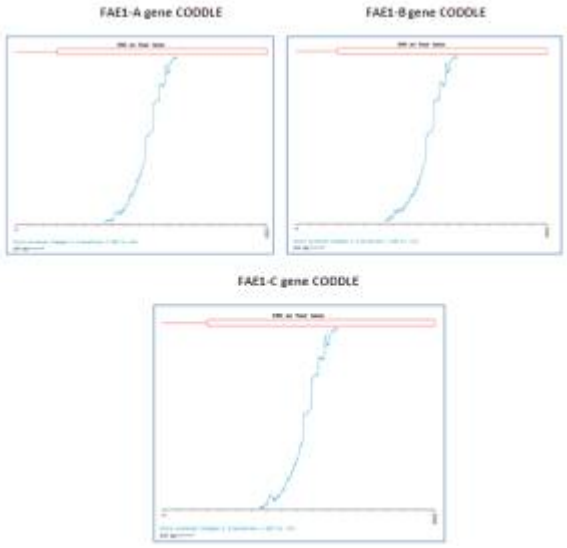


Figure 13. Coddle result for isoforms of *FAE1* gene, which depicts the probable region were maximum mutations can be found is 1000bp to 1200bp.

Both *FAD2* and *FAE1* exist as three homeologs in the Camelina genome, so the Camelina variety BBCAM2 genomic DNA was used for amplification of the same. Primers designed were conserved in all the three copies of the gene, by which all the three copies were amplified by PCR and natural SNPs were identified in the sequence after sequencing. TILLING done on pools yielded probable mutants and further were deconvoluted by taking all the eight individuals in the pool and individual mutant M2 line was identified (Figure 14). Mutations were detected from the amplified product by the heteroduplex-specific endonuclease ENDO-I in both the genes (Dalmais et al., 2008).

a. **b.** **c.**

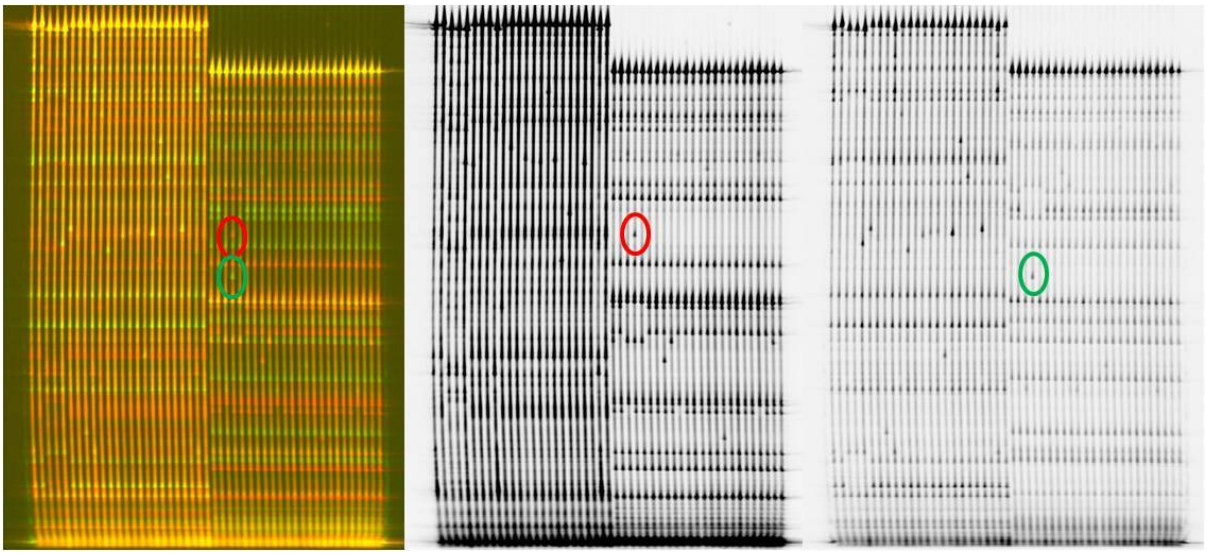
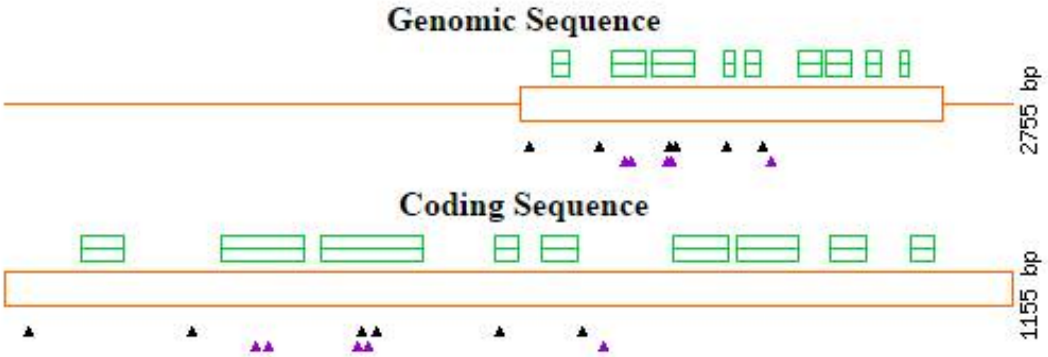


Figure 14a. TILLING poly acrylamide gel (Grossniklaus et al.) image with both IRD700 (green)and IRD800 (red)*FAD2* gene on Camelina M2 DNA pool 1; **b.** shows black and white image of IRD700; **c.** shows black and white image of IRD 800.

S No.	M2 Family Number	Nucleotide position	Codon change	Aminoacid change	SIFT RESULTS	TYPE OF MUTATION
1	CSC2-12M2(1.2) - 4960	C28T	CCT → TCT	**P10S**	TOLERATED	**MISSENSE**
2	CSC2-12M2(1.2) - 3928	C218T	ACC → ATC	**T72I**	TOLERATED	**MISSENSE**
3	CSC2-12M2(1.2) - 2397	G420A	AAG → AAA	**K139=**	-	**SILENT**
4	CSC2-12M2(1.2) - 2973	C470T	CGT → TGT	**R143C**	NOT TOLERATED	**MISSENSE**
5	CSC2-12M2(1.2) - 2277	G570A	GGG → AGG	**G190R**	NOT TOLERATED	**MISSENSE**
6	CSC2-12M2(1.2) - 790	T291C	TGT → TGC	**C96=**	-	**SILENT**
7	CSC2-12M2(1.2) - 790	T690G	TCT → TCG	**S229=**	-	**SILENT**
8	CSC2-12M2(1.2) -695	G665A	CGC → CAC	**R221H**	TOLERATED	**MISSENSE**
9	CSC2-12M2(1.2) -349	C413T	TCC TTC	**S137F**	NOT TOLERATED	**MISSENSE**
10	CSC2-12M2(1.2) -350	C408T	TAC TAT	**Y135=**	-	**SILENT**
11	CSC2-12M2(1.2) - 4068	C307T	GTC GTT	**V101=**	-	**SILENT**

Table 4. *CsFAD2* mutant details showing the M2 family number, nucleotide position with codon and amino acid change, SIFT results whether tolerated or non-tolerated (highlighted in green).

Mutant M2 line identified by TILLING gels were sequenced by Sanger sequencing and in the case of Cs*FAD2* we identified 11 point mutations, in which five were silent and six were missense mutants (Figure 15 and Table 4). Three of the missense mutants were Non-tolerated as per the SIFT results, these lines are CSC2-12M2(1.2) -2973, 2277 and 349, which were also confirmed by PARSESNP analysis.



a

b

#	View On Sequence	Nucleotide Change	Effect	Restriction Enzyme Differences from REBASE		PSSM Difference	SIFT Score	Description	Zygosity
				Gained in Variant	Lost from Reference				
1	G C	C1435T	P10S	MboII					Homo
2	G C	C1622T	T72I	BccI					Homo
3	G C	T1695C	C96=	BbvI, Fnu4HI, HgaI, TseI					Homo
4	G C	C1710T	V101=						Homo
5	G C	C1812T	Y135=						Homo
6	G C	C1817T	S137F	Hpy178III, MboII	EcoRII, PfoI, ScrFI	16.8	0.69		Homo
7	G C	G1824A	K139=		RsaI, TatI				Homo
8	G C	C1834T	R143C		Hpy99I	17.3	0.04		Homo
9	G C	G1975A	G190R	BseMII, DdeI, SauI	AvaI, SecI	20.5	0.04		Homo
10	G C	G2069A	R221H		AciI, FnuDII, McrI				Homo
11	G C	T2094G	S229=	FokI					Homo

Figure 15a. *FAD2* gene PARSESNP results showing the single exon in the genome sequence with the black triangles depicting the missense mutations and purple triangle which depicts the silent mutations. Green blocks are the various conserved domains in the exon. **b.** Table showing the various nucleotide, amino acid substitution along with various restriction enzyme site lost and introduced alongwith PSSM difference and SIFT score. Three options of substitutions have been depicted in the case of silent mutations.

TILLING of *CsFAE1* gene yielded 11 point mutants, which relates to 5 silent and 6 missense mutants (Figure 16 and Table 5). The SIFT analysis results showed the substitution as Not tolerated in case of mutant line CSC2-12M2(1.2)-1638 (Figure 17) and 2054. This line 1638 has an amino acid change at the 8th position in the protein from an aliphatic, non-polar amino acid, Leucine (Leu or L) to a cyclic non-polar amino acid Proline (Pro or P). Four of the other missense mutants were predicted to be tolerated as the amino acid substitutions didn't make significant change in the protein structure. All mutants of *FAD2* and *FAE1* gene were confirmed for respective mutations by sequencing. We have identified both conventional mutations like G/C to A/T and non-conventional EMS mutations, which shows that EMS can have both type of effects on the genome.

S No.	Packet Number	Nucleotide position (CDS)	Codon change	Aminoacid change	SIFT RESULTS	TYPE OF MUTATION
1	CSC2-12M2(1.2) - 1638	T23C	CTC ⟶ CCC	L8P	NOT TOLERATED	MIS SENSE
2	CSC2-12M2(1.2) - 1595	G453A	AAG⟶ AAA	K151=	-	SILENT
3	CSC2-12M2(1.2) - 1609	C142T	CTC ⟶ TTC	L48F	TOLERATED	MIS SENSE
4	CSC2-12M2(1.2) - 1614	C330T	ACC ⟶ ACT	T110=	-	SILENT
5	CSC2-12M2(1.2) – 77	C993T	ACC⟶ ACT	T331=	-	SILENT
6	CSC2-12M2(1.2) – 291	G1313A	GCA⟶ ACA	A438T	TOLERATED	MIS SENSE
7	CSC2-12M2(1.2) - 878	C1089T	ATC ⟶ ATT	I363=	-	SILENT
8	CSC2-12M2(1.2) - 1051	G688A	GAT⟶ AAT	D230N	TOLERATED	MIS SENSE
9	CSC2-12M2(1.2) - 1246	A1306G	ATA⟶ GTA	I436V	TOLERATED	MIS SENSE
10	CSC2-12M2(1.2) - 1424	C948T	GAC⟶ GAT	D316=	-	SILENT
11	CSC2-12M2(1.2) - 2054	C1152A	CCG⟶ ACG	P378T	NOT TOLERATED	MIS SENSE

Table 5. *CsFAE1* mutant details showing the M2 family number, nucleotide position with codon and amino acid change, SIFT results whether tolerated or non-tolerated.

Further analysis of the mutant was done with SIFT and PPOPEN to show the prediction of amino acid to be NOT Tolerated. In case of the SIFT scaled prediction, a score of 0.05 shows the substitution is NOT Tolerated. PPOPEN software gives the heat map prediction (Red color – NOT Tolerated) score of 65% that suggests similar results as the SIFT prediction (Figure 18). Thus, it can result in the protein tertiary structure being deformed and altering the oil profile of the crop.

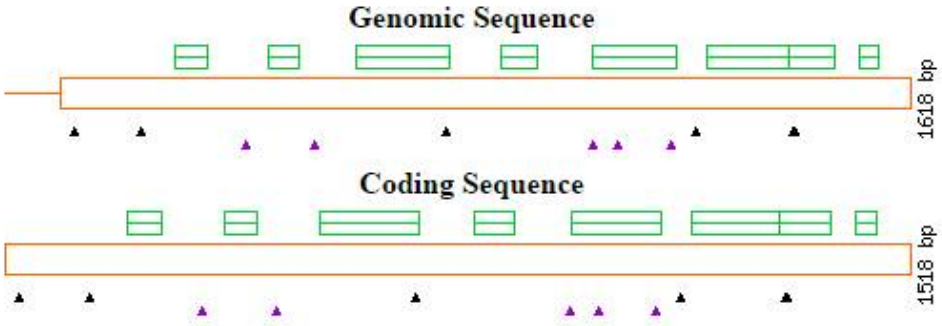
Sequences of *FAE1* and *FAD2* mutants were compared by Multiple sequence alignment with ClustalW and the mutations obtained by TILLING were found to be novel as compared to the different homeologues copies of the *FAE1* (*FAE1-A, FAE-B* and *FAE1-C*) and *FAD2* (*FAD2-1, FAD2-2* and *FAD2-3*) gene The multiple sequence alignment comparison showed that at nucleotide (CDS) position 23 ‘T’ is common in all three homologues and the substitution of ‘C’ instead in the mutant line is observed, which is an induced conversion and not a natural allele (Figure 16a).

```
CSFAE1B     CACAAACTGAGCAATGACGTCCGTTAACGCAAAGCTCCTTTACCATTACGTTCTAACCAA
CSFAE1C     CACAAACTGAGAAATGACGTCCGTTAACGCAAAGCTCCTTTACCATTACGTCCTAACCAA
CSFAE1A     CACAAACTGAGCAATGACGTCCGTTAACGCAAAGCTCCTTTACCATTACGTCCTAACCAA
CSFAE1MUT   CACAAACTGAGCAATGACGTCCGTTAACGCAAAGCTCCTTTACCATTACGTTCTAACCAA
            ********** ********************** ********** ********

CSFAE1B     CTTTTTCAACCTTTGCTTGTTTCCGTTAACGGCGTTACTTGCCGGAAAAGCCTCTAAGCT
CSFAE1C     CTTTTTCAACCTTTGCTTGTTTCCGTTAACGGCGTTACTTGCCGGAAAAGCCTCTACGCT
CSFAE1A     CTTTTTCAACCTTTGCTTGTTTCCGTTAACGGCGTTACTTGCCGGAAAAGCCTCTAGGCT
CSFAE1MUT   CTTTTTCAACCTTTGCTTGTTTCCGTTAACGGCGTTACTTGCCGGAAAAGCCTCTAAGCT
            ********************************************************* ***

CSFAE1B     TACAGCAAACGATCTCTACCACTTCTATTCCCATCTCCAACACAACCTTATAACCGTAAT
CSFAE1C     TACCACAAACGATCTCTACCACTTCTATTCCCATCTCCAACACAACCTTGTAACCGTAAT
CSFAE1A     TACCTCAAACGATCTCTACCACTTCTATTCCCATCTCCAACACAACCTTATAACCGTAAT
CSFAE1MUT   TACCTCAAACGATCTCTACCACTTCTATTCCCATCTCCAACACAACCTTATAACCGTAAT
            ***  ******************************************* **********
```

a



b

c

#	View On Sequence	Nucleotide Change	Effect	Restriction Enzyme Differences from REBASE		PSSM Difference	SIFT Score	Description	Zygosity
				Gained in Variant	Lost from Reference				
1	G C	T123C	L8P		AluI				Homo
2	G C	C242T	L48F		BccI				Homo
3	G C	C430A	T110=		MnlI				Homo
4	G C	C430G	T110=	MaeII, MaeIII, Tsp45I	MnlI				Homo
5	G C	C430T	T110=		MnlI				Homo
6	G C	G553A	K151=						Homo
7	G C	G788A	D230N	MseI, VspI	DpnI, MboI	10.4	0.11		Homo
8	G C	C1048T	D316=		BsmAI, Esp3I				Homo
9	G C	C1093A	T331=						Homo
10	G C	C1093G	T331=						Homo
11	G C	C1093T	T331=						Homo
12	G C	C1189T	I363=	BsrDI	BtgZI				Homo
13	G C	C1189A	I363=	CviJI	BtgZI				Homo
14	G C	C1232A	P378T	MaeIII, Tsp45I, TspGWI	CauII, FinI, HpaII, PfoI, ScrFI				Homo
15	G C	A1406G	I436V	BsaAI, MaeII, SnaBI		5.4	0.90		Homo
16	G C	G1412A	A438T			18.2	0.01		Homo

Figure 16a. Multiple sequence alignment (ClustalW) of *FAE1* mutant 1638 with three FAE1 homologues, with transition of T to C at nucleotide CDS position 23 (red colour). **b.** PARSESNP results showing the single exon in the genome sequence with the black triangles depicting the missense mutations and purple triangle which depicts the silent mutations. Green blocks represent the various conserved domains in the exon. **c.** Table showing the various nucleotide, amino acid substitution along with various restriction enzyme site lost and introduced alongwith PSSM difference and SIFT score. Three options of substitutions have been depicted in the case of silent mutations.

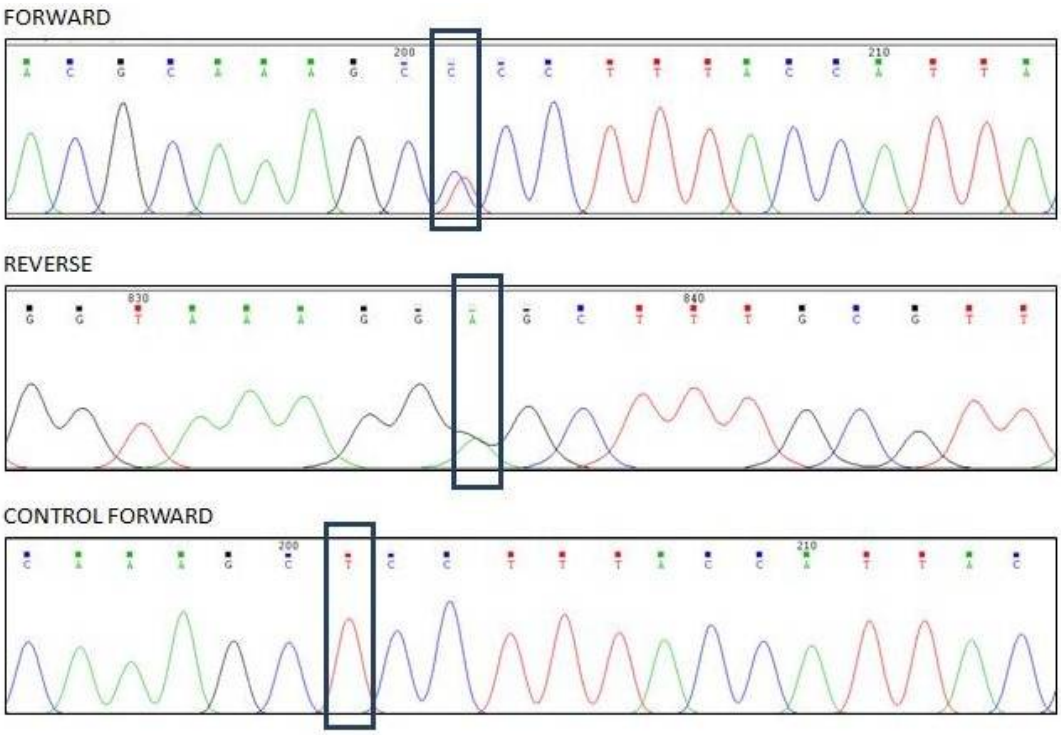


Figure 17. Chromatogram of sequencing of *FAE1* gene of CSC2-12M2(1.2) -1638 mutant shows clear double peaks in both forward and reverse sequence as compared to the control forward sequence confirming mutation of a T to C at 23 nucleotide position

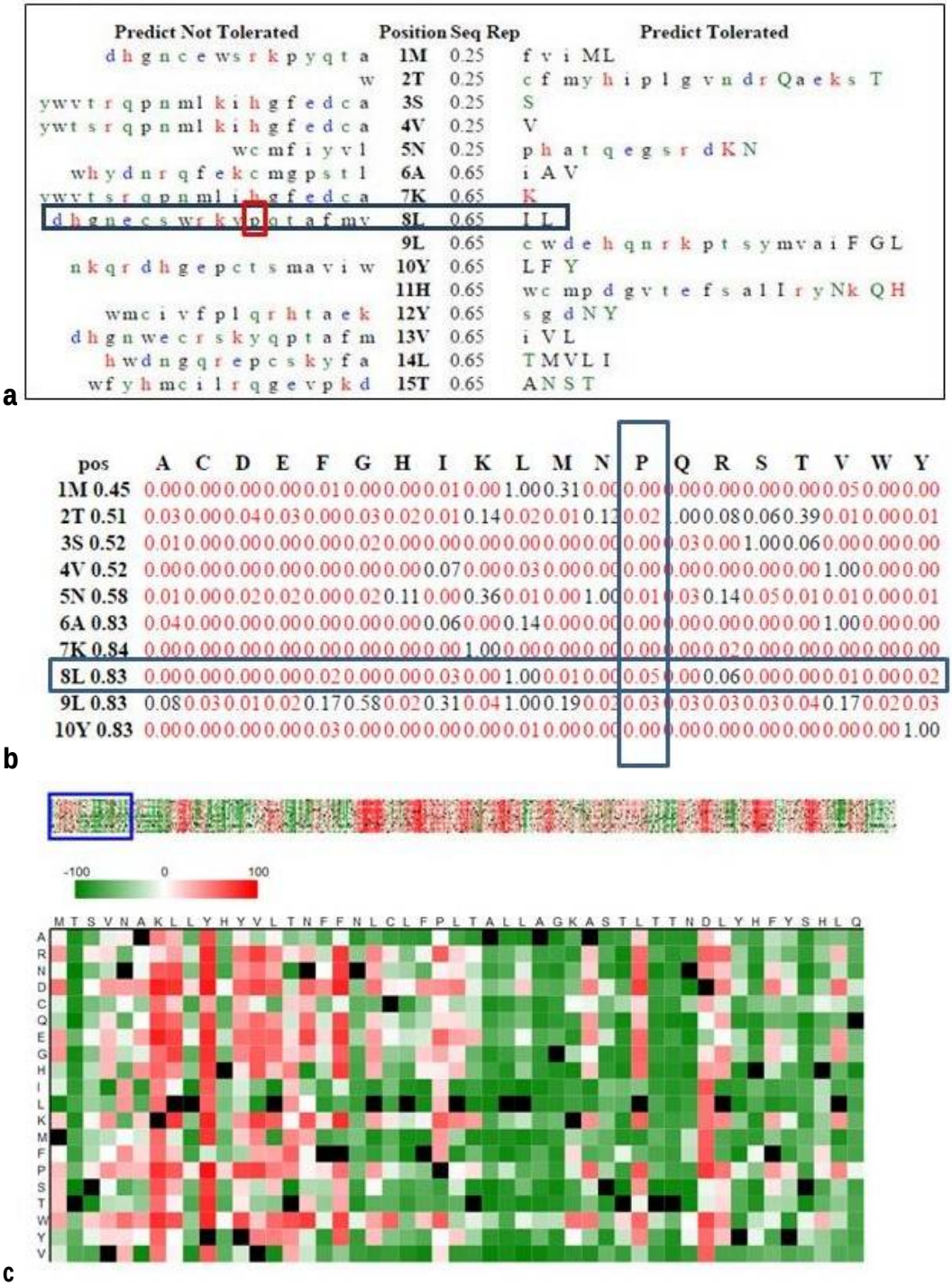


Figure 18 a. SIFT analysis score of CS*FAE1* mutant CSC2-12M2(1.2) -1638 showing prediction of substitution of 'P' instead of 'L' at the aminoacid position 8 to Not tolerated. **b.** Table shows SIFT scaled probabilities of aminoacid tolerance at the specified position i.e. 0.05, that signifies the aminoacid substitution is not tolerated. **c.** Heat map showing *CsFAE1* protein sequence tolerance with various substitutions of various amino acids, with score of 65% showing red color at the substitution of L8P, designates the Non-Tolerance nature of the mutant.

2.3.5 Oil profiling of Camelina mutants

Oil extraction by modification of Bligh and Dyer method (Bligh and Dyer, 1959) yielded a concentration of 35.21% total lipids in control Camelina BBCAM2 variety as compared to the CELINE variety which yielded 5.26% (Figure 19). The Camelina oil recovery in Celine variety has been quite low, which may be due to the method used as there are many methods optimized for each variety to get an optimized and maximum recovery of total lipids.

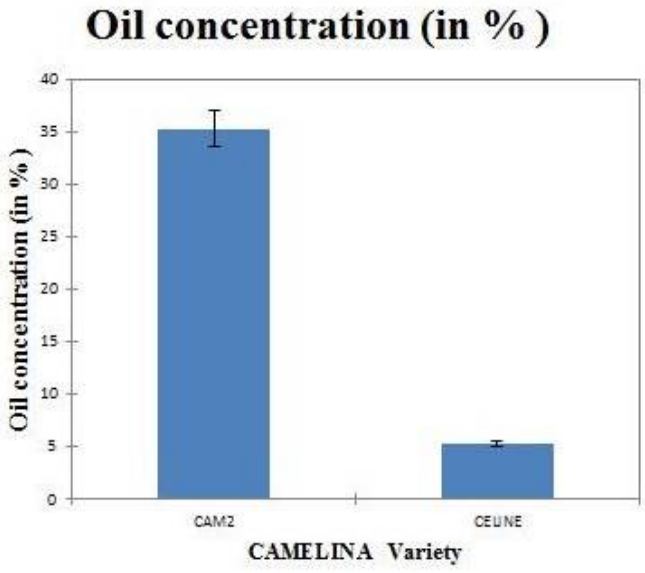


Figure 19. Comparison of total oil concentration between two different varieties of Camelina i.e. BBCAM2 and CELINE (in percentage of total weight), which shows a 35.21% and 5.26% respectively.

2.3.6 Oil analysis of *FAE1* and *FAD2* mutants

Mutant M3 seeds of *FAE1* and *FAD2* were used for GC analysis along with the BBCAM2 control seeds. Oil extracted from the M3 seed was converted to FAME before injecting into the GC. The peaks obtained were compared with an internal standard. *FAE1* mutant CSC214M3(1.2)-1638-3, which is a non-tolerant missense mutant at nucleotide CDS position T23C with an amino acid change of L8P is a heterozygous mutant confirmed after genotyping by TILLING. The oil profile obtained showed profound decrease in very long chain fatty acids as compared to the wild type oil profile. In this line the concentration of α-Linolenic acid (C18:3) has increased to 35% as compared to the control which is just 31% which indicates the lack formation of VLCFAs have lead to low erucic acid, which has good application in various industries (Table 6 and Figure 20). Variability within the M2 family was observed as in case of M3-1638-1 had shown a tremendous increase of Myristic acid (C14:0) as compared to wild type oil profile i.e. about 40% of total fatty acids.

Fatty acids	Individual M3 of Camelina *FAE1* mutant (% Fatty acids)					BBCAM 2 wild type (% of FA)
	1638-1	1638-2	1638-3	1638-4	1638-5	
Myristic acidC14:0	40.60	4.50	0	0	4.50	0
Palmitic acidC16:0	7.70	6.03	4.45	6.51	4.00	4.30
Palmitoleic acidC16:1	0	0	0	0	0	0.09
Margaric acidC17:0	0	0	0	0	0	2.23
Stearic acidC18:0	5.40	2.78	2.84	3.16	2.90	2.20
Oleic acidC18:1	16.00	14.82	15.80	18.11	16.00	13.60
Linoleic acidC18:2	2.10	18.49	17.92	20.81	18.2	19.90
α-Linolenic acidC18:3	2.40	37.39	35.31	38.75	35.70	31.92
Arachidic acidC20:0	5.60	1.98	3.65	0.80	3.20	2.48
Paullinic acidC20:1	19.90	13.91	20.03	11.86	15.00	14.20
Eicosadienoic acidC20:2	0	0	0	0	0	1.46
Heneicosylic acidC21:0	0	0	0	0	0	1.37
Behenic acidC22:0	0	0	0	0	0	0.47
Erucic acidC22:1	0	0	0	0	0	3.59
Lignoceric acidC24:0	0	0	0	0	0	0.76

Table 6. Comparative data of various fatty acids obtained by Gas Chromatographic analysis from *FAE1* mutant CSC214M3(1.2)-1638 M3 seeds (individuals of M3 Line).

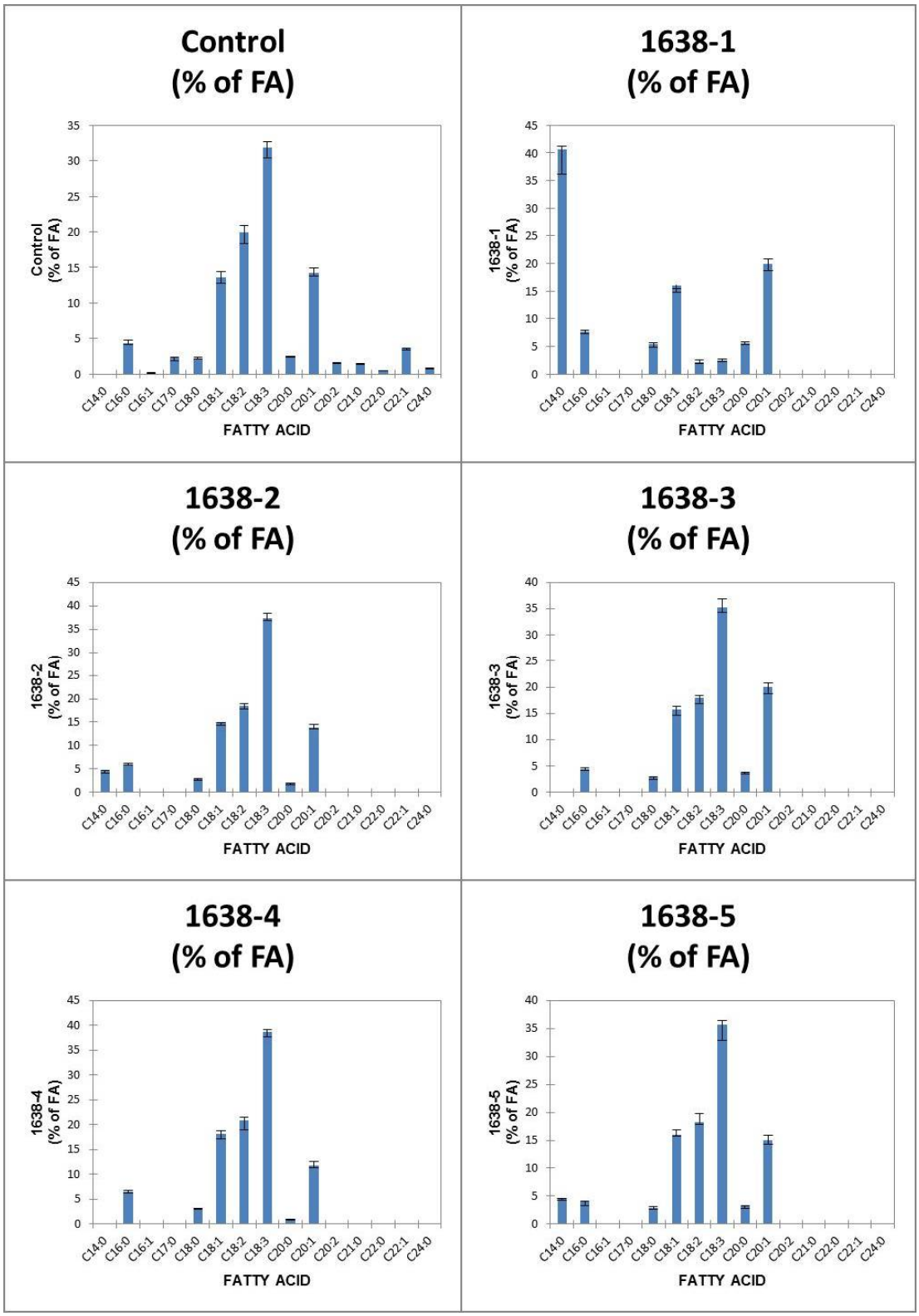


Figure 20. Histogram analysis with different fattyacid profiles in comparison to control with FAE1 M3 mutant 1638-1, 1638-2, 1638-3, 1638-4 and 1638-5. Software XLSTAT was used with error bars for making the bar diagram.

Camelina *FAD2*, non-tolerated (SIFT) missense TILLING mutants, M3- 2277-6 and M3-2973-3 oil profile shows the mutated gene is showing the phenotype in the total fatty acids composition as compared to the wild type Camelina cultivar BBCAM2. The concentration of oleic acid, linoleic and linolenic acid in total was 56.1% in the case of mutant line M3-2277-6 and 66.42% in the wild type. The decrease in this fatty acid shows that mutation of *FAD2* may be responsible for the phenotype observed. These mutant lines with improved oil composition can be bred to make better hybrids for various industrial applications (Table 7 and Figure 21).

Fatty acids	**Individual M3 of Camelina *FAD2* mutant (% Fatty acids)**		**BBCAM2 wild type (% of FA)**
	2277-6	**2973-3**	
Palmitic acid C16:0	2.1	4	4.30
Palmitoleic acid C16:1	0	0	0.09
Margaric acid C17:0	0	0	2.23
Stearic acid C18:0	1.9	2	2.20
Oleic acid C18:1	11.8	13.5	13.60
Linoleic acid C18:2	16.8	16	19.90
α-Linolenic acid C18:3	27.5	32.6	31.92
Arachidic acid C20:0	2.8	2.8	2.48
Paullinic acid C20:1	19.3	16.6	14.20
Eicosadienoic acid C20:2	2.3	1.8	1.46
Heneicosylic acid C21:0	0	0	1.37
Behenic acid C22:0	0.7	0.5	0.47
Erucic acid C22:1	6.7	4.8	3.59
Lignoceric acid C24:0	1.7	0.6	0.76

Table 7. Comparative data of various fatty acids obtained by Gas Chromatographic analysis from *FAD2* mutant CSC214M3(1.2)-2277-6 and 2973-3 M3 seeds (individuals of M3 Line).

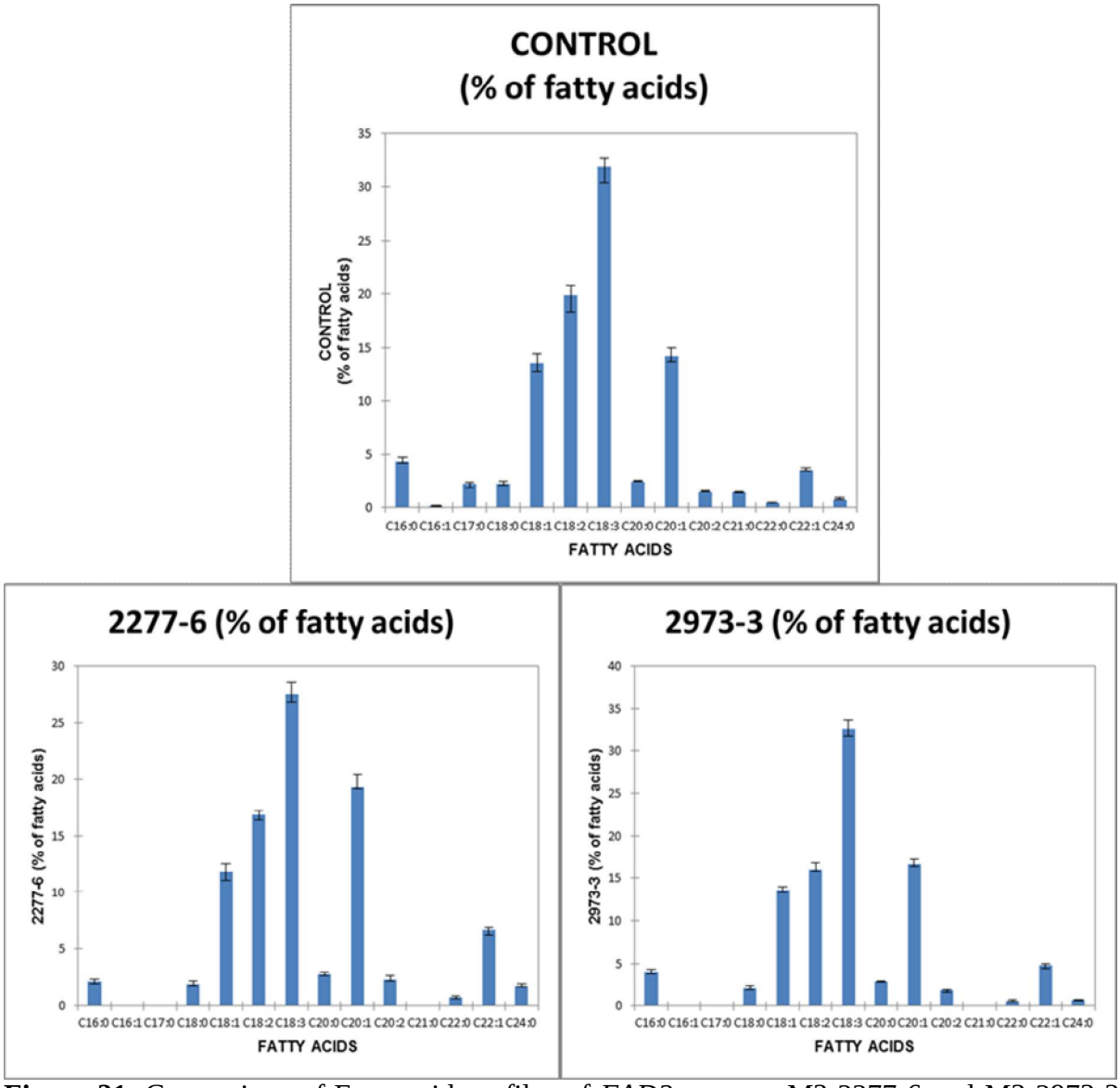


Figure 21. Comparison of Fatty acid profiles of *FAD2* mutants M3-2277-6 and M3-2973-3 with wild type BBCAM2 fatty acid, which shows marked variation in fatty acid composition as compared to the control fatty acid profile.

2.4 DISCUSSION

Camelina sativa is an under-exploited biofuel crop and recently there has been a renewal of interest in this crop for various industrial purposes, as well as for consumption. Creation of a new varieties with novel traits will help breeders to develop new varieties. *Camelina sativa* is an old Irish crop and lot of research was done earlier by Teagasc (Rice et al., 1998). Feasibility studies have been undertaken in India to assess Camelina's potential as a source of biofuel (Agarwal et al., 2010). Various research experiments at different location across the world are being carried out to tabulate and authenticate the importance of varieties from different geographical locations. Different varieties of Camelina were grown and various agronomical and fatty acid analysis were done in the Mediterranean dryland and some varieties with important traits were identified for biofuel as well as edible characteristics (Campbell et al., 2013). The Genetics and Biotechnology research lab at NUI Galway had engaged in multi-locational field trials of Camelina for studying the agronomical prominence in Ireland in association Teagasc, Ireland (Rice et al., 1998). Agronomical conditions for the Camelina variety BBCAM2 at Benchbio site Kulu, H.P., India showed better crop production as compared to the tropical site. Our results indicate that the best time for sowing in temperate climates in India is February-March (data not shown), similar to results obtained previously (Agarwal et al., 2010).

Camelina genome being an allo-hexaploid behaves like a diploid with normal disomic inheritance. *C.sativa* genome had undergone inter-specific hybridization from lower chromosome number ancestors mainly due to two polyploidy events occurred in between ancestors same as *Arabidopsis lyrata*. Expression dominance and a small protion of triplicated genes (about 22%) leads to the aspect of functional diversification. Most of the desirable agronomic traits and oil characteristics may be due to the triplicated genome of camelina (Kagale et al., 2014). Transcriptome study of camelina done at different stages of the life cycle of camelina, interactive visualization of expression data was obtained. Tissue-specific expression of various genes were observed by the RNAseq analysis and co-expression analysis tools. It was also found that the dominance of Camelina gene copy on other sub-genomes was obtained by further studying the camelina genomics (Kagale et al., 2016).

2.4.1 Camelina oil improvement by TILLING

Lipid biosynthesis leads to the formation of various fatty acids, only few fatty acids play major role in various processes. It is de novo synthesis from small precursors derived from photosynthate forms the long chain fatty acids (Harwood, 1996). There are two enzyme systems which are observed i.e., acetyl-CoA carboxylase and fatty acid synthase. Palmitic acid and stearic acid being the precursor for further biosynthesis of VLCFAs i.e. very long chain fatty acids. Mainly elongation, desaturation and other modifications are the process by which major fatty acids are produced. In this project the target genes chosen for TILLING for oil profile improvement were those encoding fatty acid desaturase 2, *FAD2* and fatty acid elongase 1, *FAE1*, yielding multiple mutant alleles of both genes.

Lipid biosynthesis pathways of poly unsaturated fatty acids (PUFA) in Camelina sativa was probed at gene expression level with the utilization of 454 pyrosequencing reads revealed about 32,759 genes which were involved in seed development, where profound changes were observed in samples from different intervals of flowering. A detailed analysis of up-regulation and down-regulation involved in the fatty acid biosynthesis pathway was also discussed and different pathway networks were matched (Wang et al., 2015a). Metabolic engineering in Camelina seed oil composition was modified by transformation of algal genes to produce higher source of n-3 LC-PUFA (Betancor et al., 2015). Camelina RNAi lines deficient in napins were obtained by targeting various genes including *FAD2* and *FAE1* to obtain high oleic acid oil lines, transcriptome data of which was used for Camelina meal and oil (Nguyen et al., 2013). Vacuum infiltration of Camelina flowers technique and floral dip transformation was demonstrated for transformation of Camelina used for improved oil composition in seeds (Liu et al., 2012; Nguyen et al., 2014). New technique of gene editing by CRISPR/CAS9 was also done to improve the oleic acid composition from 16% to 50% by targeting all the three copies of *FAD2* gene in Camelina (Jiang et al., 2016).

Fatty acid desaturases (*FAD2*) are the enzymes which lead to the introduction of double bonds in fatty acid chains. An aerobic mechanism with oxygen being reduced by four hydrogen atoms (two from substrate and two from reductant) is observed mostly in plants. All the components like the electron donors, cytochrome, varies depending on the desaturase enzyme (Harwood, 1996). The saturation of fatty acids takes place by the introduction of single (monoenoic) or multiple numbers of double bonds (polyenoic) and the number generally designates the position where the double bond has been introduced or the substrate

used. *FAD2* gene in Camelina shows high similarity with the homologs found in Arabidopsis (At*FAD2*Genebank ID:L26296) (Kang et al., 2011). *FAD2* is a membrane bound delta-12-desaturase which converts oleic acid (C18:1) to linoleic acid (C18:2) and which is located in the endoplasmic reticulum in plants (Okuley et al., 1994). Various types of mutants were obtained by mutagenesis in *Arabidopsis thaliana* with varied compositions of fatty acids especially elongation of 18:1 to 20:1 and desaturation of 18:1 to 18:2, it was also observed that the wild type allele showed incomplete dominance on the mutant allele (Lemieux et al., 1990). In *Brassica napus*, which is a tetraploid, various SNPs and QTL has been identified on A and C genes of *FAD2* and *FAD3* gene, which control the development of high oleic and low linolenic acid lines (Yang et al., 2012). *Camelina sativa* is an allohexaploid with three copies of the gene *FAD2 i.e.* A, B and C, which are expressed in developing seeds and all three copies are found to be functional (Hutcheon et al., 2010). This could be the reason why oil profile phenotype in the present study did not vary as significantly as expected. The recently developed TILLING-by-Sequencing will be explored to extensively screen for further new mutations from the population.

Fatty acid elongase (*FAE1*) in most of the Brassicaceae is the major enzyme which leads to elongation of long chain fatty acids to produce very long chain (>18C) products (both saturated and unsaturated) (Cassagne et al., 1994). These enzymes are membrane bound and used acyl-CoA substrates and malonyl-CoA directly as the source of 2C unit, which leads to formation of very long chain fatty acids (VLCFA), eiconsenoic acid (C20:1) and erucic acid (C22:1) (Fourmann et al., 1998). Modification in erucic acid content has also been due to twenty-eight base deletion at the promoter region of the *FAE1* gene resulting in differential expression of the gene leading to the variation in erucic acid composition in *Brassica rapa* L. (Yan et al., 2015). In the present study the *FAE1* mutant 1634-3 shows zero erucic acid, which is a potential line for various industrial purposes. Erucic acid has been found have harmful effects like heart ailments, if presence of the same found in edible oil (2003). Mutation in *FAE1* gene can lead to over expression of this gene, where by leading to increased formation of VLCFAs, which are the basic components of biofuel. In *Brassica carinata* the expression of one of the genes, ketoacyl-CoA synthases (KCS) through transgenic method, enhanced the concentration of major erucic acid (Taylor et al., 2010). *FAE1* has been widely studied in case of Arabidopsis especially in the expression in developing seeds (Kunst et al., 1992).

2.4.2 Protein 3D-modeling of *FAE1* mutants

Mutations obtained in EMS mutagenesis are mostly non synonymous substitution, where a nucleotide substitution alters the amino acid sequence of a protein. SWISS-MODEL is a fully automated protein structure homology-modelling server, accessible via the ExPASY web server (https://swissmodel.expasy.org/). Missense mutations predicted by SIFT analysis (not tolerated) were further analyzed with SWISS-MODEL for comparing the 3D protein structure of the wild type *FAE1* protein with the missense mutant CSC2-12M2(1.2)-1638 *FAE1* protein with amino acid substitution of L8P. Analysis of the 3D protein of both wild type and mutant protein varied 3D structures by which could affect the overall protein function (Figure 21). Similar comparative protein structure prediction was also done on tolerated mutant to observe the validity of such mutants for further phenotypic characterization. Missense mutant CSC2-12M2(1.2)-1609 which had amino acid substitution of L48F was predicted tolerated according to SIFT analysis. This mutant was further analyzed for protein structure prediction after the substitution of the amino acid phenyl alanine instead of leucine. Protein modeling predicted structural deformity (leading to function of the protein) as compared to the wild type protein by superimposing both wild type and mutant using TopMatch software (Figure 22) (https://topmatch.services.came.sbg.ac.at/), which use different techniques for protein structure alignment (Sippl and Wiederstein, 2008, 2012). SIFT prediction compares the amino acid in a protein which is conserved throughout evolution and compares with the available protein sequence data submitted in the public databases. For further confirmation on the functional effects protein modeling of the mutants is more informative in predicting the structural changes which affect the overall function of the protein and the phenotype. Phenotyping of such tolerated missense mutants (predicted by SIFT) could show whether there is any overall affect on the function of the gene.

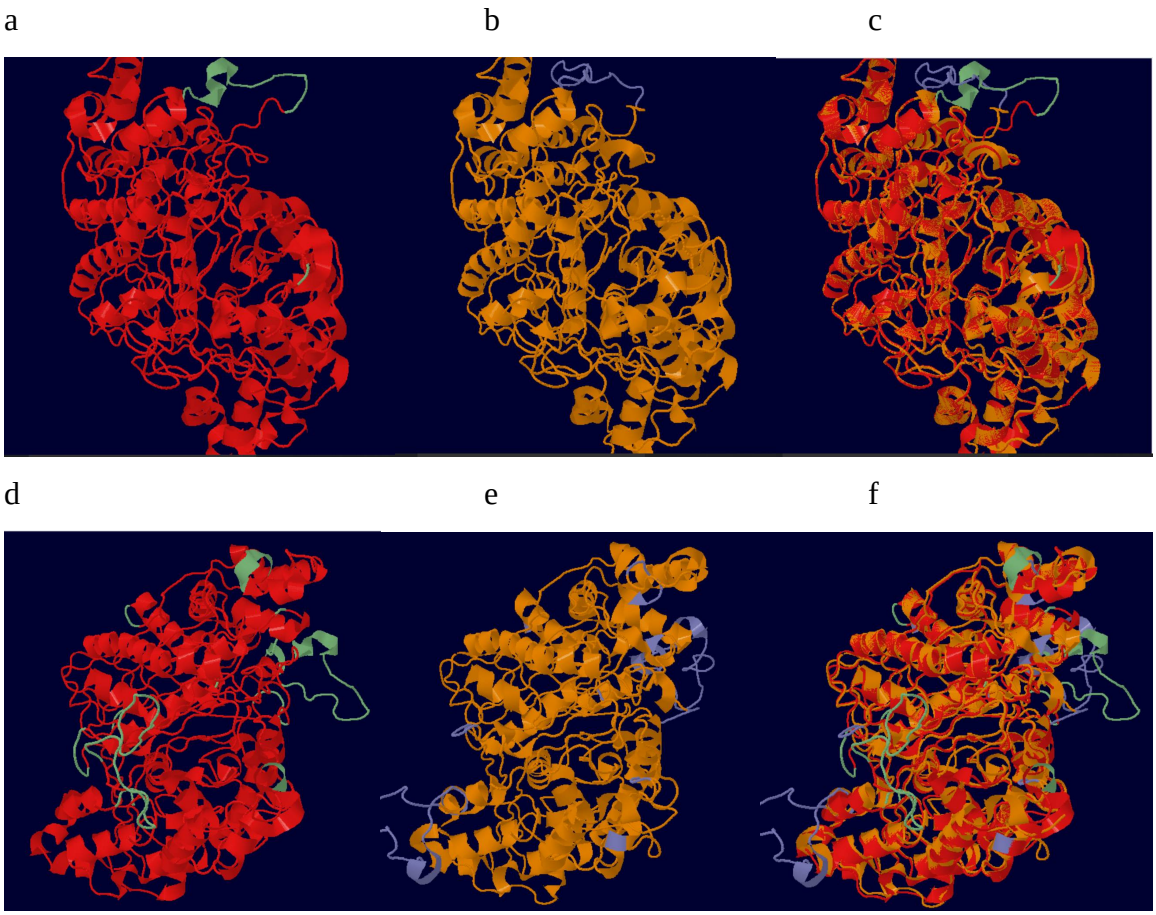


Figure 22. Camelina *FAE1* Protein 3D structure (predicted by SWISS-MODEL) a. Wild type *FAE1* protein, b. CSC2-12M2(1.2)-1638 mutant line FAE1 protein with substitution L8P, c. Superimposed protein structure (by of both a and b protein structure, d. Wild type FAE1 protein with different orientation, e. CSC2-12M2(1.2)-1609 mutant line FAE1 protein with substitution L48F, f. Non-aligned region in wild type is in green and in mutant is in Blue; Orange and red are identical residues of Mutant and Control.

2.4.3 Screening the Camelina Mutant Population

In this project, TILLING screening of Camelina mutant population demonstrates that the technology is useful in generating useful mutants as we obtained about 22 mutants. Back-crossing of potential mutants with the control parent is being planned with markers to remove the background mutations, so that plant breeders can utilize these lines for obtaining homozygous lines. Furthermore, markers can be developed as the mutations are known and marker assisted selection can be carried out. Research have shown that the mutations which results in synonymous mutations as compared to non-synonymous mutations are also of high value (Tsai et al., 2013). The mutation frequency of *FAD2* and *FAE1* in this project was 1/520kb and 1/682kb respectively and a total mutation frequency in the Camelina mutant population has been observed to be 1/601 kb (Table 8).

Gene	Fragment size (bp)	TYPES OF MUTATIONS						Mutation Frequency (1/kb)
		Non sense	Splice	Missense - non tolerated	Missense tolerated	Silent mutations	Total mutations	
FAD2	1158	0	0	3	3	5	11	1/520
FAE1	1518	0	0	2	4	5	11	1/682
	2676						**22**	**1/601**

Table 8. Total mutation frequency of BBCAM2 mutant population deciphered by TILLING Fatty acid desaturases, *FAD2* and Fatty acid elongases, *FAE1*.

The Camelina mutant population in the genetic background BBCAM2 can be used for isolation of mutant lines via reverse genetics. In this project we focused on *FAD2* as an initial gene for targeted mutagenesis via TILLING, a list of more than twenty target genes have been made which can be utilized for improvement. The mutation density has been found to be in between 1/75 kb to 1/120 kb for *FAD2* gene (Hutcheon et al., 2009). We have developed a mutant population of 5000 M2 families which has been treated with very high concentration of EMS dose of. 1.2% for 8 hours. In the M2 families we observed 483 phenotypes in our population like albino, variegated, curly leaves, expanded leaf sheath, pigmented leaves etc. (Table 8). Such phenotypes has been observed in mutant populations of oil seed crops like sunflower (Kumar et al., 2013), and rape seed (Harloff et al., 2012) and Camelina (present study). A total number of 2304 M2 seedling phenotypes has been analyzed and the number of phenotypes found were 9.6%, similar results were found in sunflower (Kumar et al., 2013). Camelina has a huge potential as a biofuel as the carbon monoxide emission is quite low, especially when used as a jet-fuel as compared to other fuels. Even the output power has been

high as compared to convention mineral oil used as a source of engine oil (Bernardo et al., 2003).

Mutants obtained in the present study should be back crossed and the desirable trait could be introduced in the parental line without any back ground mutations. EMS mutagenesis leads to random mutations at different sites in genome, which could be beneficial or deleterious for the survival of the crop. Back crossing of the mutant line with parental camelina CAM2 line till BC1 and BC2 is essential to remove the background mutations. High through put isolation of mutations by utilizing whole genome sequencing to map and detect the same without the need for prior genetic mapping of species (Zuryn et al., 2010). Mutants obtained in camelina mutant population in both *FAD2* and *FAE1* could be utilised for the back crossing and novel lines with improved oil traits and zero erucic lines could be developed.

2.5 CONCLUSIONS

Crop improvement in Camelina by TTILLING method shows that TILLING is a good method to improve crops without genetic modification. The crop has a huge potential to become a major replacement for active ingredient as a raw material for biofuel and other non-food applications. The TILLING platform can be utilized for screening other genes of importance for biofuel generation and also for non-nutritive traits for various applications on different industries. With the advent of TILLING by Sequencing, the method has become relatively easier and high throughput as it allows screening of a large number of genes from a large mutant population library yielding a large number of mutants TILLING in a relatively short period of time. The mutant population generated in the present study is of high quality as it carries high mutation frequency and is therefore an ideal resource of breeding material for developing novel traits in this species.

3. CHAPTER THREE

Review of TILLING by Sequencing (TbyS) and TbyS in *Camelina sativa* for crop improvement

ABSTRACT

Tilling by Sequencing (TbyS) refers to the application of high-throughput sequencing technologies to mutagenized TILLING populations as a tool for functional genomics. TbyS can be used to identify and characterise induced variation in genes (controlling traits of interest) within large mutant populations, and is a powerful approach for the study and harnessing of genetic variation in crop breeding programs. The extension of existing TILLING platforms by TbyS will accelerate crop functional genomics studies, in concert with the rapid increase in genome editing capabilities and the number and quality of sequenced crop plant genomes. In this chapter, I provide an overview of the growth of TbyS and its potential applications to crop molecular breeding and will provide TbyS in Camelina on selected genes i.e. *FAD2, FAE1, MIPS1* and *TGG1* genes for crop improvement. Total mutation frequency in Camelina mutant population screened by TbyS yielded 1/55 kb with mutation frequency of 1/40kb in *FAD2* gene, 1/40kb in *FAE1* gene, 1/106 in *MIPS1* gene and 1/145 in *TGG1* gene. 2-Dimensional pooling strategy utilized in TbyS of camelina led to the validation of 183 mutants from all the four genes screened, in which four non-sense, one splicing, 119 missense and 58 silent mutations were observed. These mutants could be further utilised for breeding for utilization of the specific characters required by breeders. The Camelina as a crop could be used as a model oil crop with the utilization of this mutant population for further crop improvement for various agronomical and industrial uses.

3.1 INTRODUCTION

To feed the increasing global population, it will be necessary to sustain our current growth in agricultural productivity over the course of the coming century while reducing environmental and climatic footprints (Godfray and Garnett, 2014). The availability of genome sequences for a wide range of crop plant species and their close relatives has generated a vast pool of data that needs to be harnessed for sustained and improved genetic gain (Onda and Mochida, 2016). Such genomic data provides a framework that can be used to identify genes and alleles in crop species that are causal or contributory for important traits such as yield, disease resistance, and nutritional composition (Wang et al., 2012b). However, many crop breeding programs continue to suffer from a lack of genetic diversity and/or useful traits within the primary breeding pools of many crop species (Cooper et al., 2001; Tang et al., 2010). Such constraints can arise due to genetic bottlenecks encountered during domestication and the focus on elite germplasm within twentieth century breeding programs (Gross and Olsen, 2010; Kovach and McCouch, 2008; Meyer and Purugganan, 2013; Tang et al., 2010; Till et al., 2003).

Historically, mutagenesis breeding has been as an approach to extend the low genetic diversity of many primary genepools of crops (Konzak et al., 1976). Such mutagenesis approaches have typically used radiation or chemical mutagens to introduce additional genetic polymorphisms into crop plant genomes, and thereby supplement the available natural variation within the crop primary genepool. While random mutagenesis approaches are powerful for forward genetics to identify easily scorable phenotypes within large mutagenised populations, genome-wide mutagenesis approaches on their own do not allow the identification of the underlying causal mutated genes (Murphy, 2007). TILLING was developed as a widely-applicable reverse-genetics strategy for mutagenising target genes of interest in a manner which allowed the plants carrying mutations in the target gene to be identified from within a large mutagenised population of plants (Slade and Knauf, 2005; Till et al., 2003; Wang et al., 2010; Wang et al., 2012b).

Current platforms for deploying TILLING approaches have incorporated a range of improvements to (a) the mutagenesis regime followed, (b) the pooling and deconvolution strategy, and (c) the enzymes used to identify mutations within a gene (typically based on enzymes which can pinpoint mismatches between the DNA from pools of mutagenised plants and a non-mutagenized reference genome; (Figure 21A). For example, the TILLING

workflow developed by the URGV lab, INRA, France has used a highly-efficient enzyme, ENDO-I, to identify DNA mismatches (Triques et al., 2007) which has been applied to various crop plants (Boualem et al., 2014; Dahmani-Mardas et al., 2010; Kumar et al., 2013). A range of other public research institutions have also developed their own TILLING platforms, as have a number of private plant breeding and crop genomics companies (Table 9). In some cases, the TILLING platforms are focussed on small numbers of important crop species (often with well-described and extensively annotated genomes), while others are more widely based and also facilitate the analysis of model systems or of less well-characterised crops.

TILLING programs have been used to generate and identify crop germplasm containing altered sequences in genes associated with important agronomic traits (see Table 2.2. in (McKeown et al., 2013b)). In some cases, this has provided insights into the biology of species with complex or poorly sequenced genomes. For example, novel variation in wheat genes involved in starch synthesis pathways was generated by TILLING of the genes *Sgp-1* and *Wx* (Sestili et al., 2010). Despite the potential for TILLING to induce novel variation in genes there are also limitations in such random mutagenesis approaches. Many of the allelic variants induced and identified during TILLING are considered to be phenotypically neutral, which can be due to the occurrence of synonymous mutations or due to genetic redundancy (particularly in polyploid genomes). In the accession C24 of *Arabidopsis thaliana*, analysis of mutations in four TILLED genes indicated that 69.6% of changes were missense and 29.0% sense, with 1.4% being nonsense mutations (Lai et al., 2012). Such nonsense mutations can introduce premature stop codons, and where they affect N-terminal regions are likely to generate knock-outs and are the class of mutations typically sought in many TILLING programs. However, depending on the gene target, such complete (amorphic) knockouts also carry a potential risk of being deleterious to the organism. Undoubtedly more subtle mutations (e.g. hypomorphic, hypermorphic, antimorphic and neomorphic alleles), conferring potentially valuable novel phenotypes caused by changes to the function of a gene, also exist within mutagenised populations developed during a TILLING program. Nonsense mutations in specific genes can be rare, requiring the use of larger mutagenised populations and more high-throughput screening (Wang et al., 2012b).

Category	Institution	WebLinks	Examples of TILLING work
Crop science applications	INRA, France	www-urgv.versailles.inra.fr/TILLING/index.htm	Incorporates the use of ENDO-I endonuclease (Triques et al., 2007).
	Purdue University, USA	www3.ag.purdue.edu/agry/Pages/cweil.aspx	Focussed on analysis of *Zea mays*.
	University of California, Davis, USA	http://TILLING.ucdavis.edu/index.php/ Arabidopsis_ TILLING	Provides public platform applicable to the analysis of many genomes.
	Revgen, UK	http://revgenuk.jic.ac.uk/about.htm	Focussed on Brassica TILLING
	The James Hutton Institute, Scotland, UK	http://bioinf.scri.ac.uk/barley_snpdb/	Analysis of barley by TILLING.
	University of Tsukuba, Japan	http://tomatoma.nbrp.jp/	Analysis of tomato by TILLING.
	BenchBio Pvt. Ltd., India	http://www.benchbio.com/	Focused on TILLING in various crops
Analysis of model organisms	TAIR, USA	www.arabidopsis.org/abrc/index.jsp	Provides resources for the *A. thaliana* community (Lai et al., 2012).
	University of British Columbia, Canada	www3.botany.ubc.ca/can-till/	Analysis of *A. thaliana*, but also of also includes the analysis of economically important Brassicaceae crops (*Brassica oleracea, B. napus*).

Table 9. Examples of major public and private TILLING platforms/portals and their ongoing work.

3.1.1 Enhancement of TILLING as a tool for crop improvement

TILLING by conventional method had major drawbacks like use of acrylamide gels, limitation of samples and long duration for screening, required an enhancement of this method. The widespread use of TILLING in crop improvement has been hindered by the fact that it is laborious to identify mutant lines from the high quality mutant population libraries that need to be generated and screened. Conventional TILLING approaches have typically been limited to a handful of target loci. In 2011, researchers in the Comai lab coined the term 'TILLING by sequencing' (TbyS) to describe the application of high-throughput next-generation sequencing techniques to accelerate TILLING workflows (Tsai et al., 2011). Since then, a number of different TbyS methodologies have been developed to identify point-mutations from mutagenised populations in a more high-throughput manner that avoids bottlenecks associated with conventional TILLING approaches. Further advancements of TbyS will undoubtedly emerge. It is likely that these will allow the exploration of poorly-described or unannotated genomes, and the identification of wider range of phenotypes (Dahmani-Mardas et al., 2010; Reddy et al., 2012).

3.1.2 Utilizing Next Generation Sequencing for TILLING

The difficulty of identifying suitably mutagenised lines within TILLING populations is compounded by a second key limitation facing current processes: namely, the requirement for what has been termed an '*ad hoc* investment of time and resources' in the procedure (Henry et al., 2014).This limitation will become ever-more pertinent as TILLING is applied to more genomically complex species, species with long lifecycles or complex reproductive modes, or where wider ranges of phenotypes are sought. Over recent years, TILLING approaches have been undergoing a range of modifications to allow integration with high-throughput sequencing technologies. Such changes have been accompanied by improvements to pooling strategies, usually involving use of robotics platforms. As DNA sequencing has become cheaper and faster due to the advent of so-called 'next-generation sequencing' ('TbyS') technologies, opportunities have arisen for incorporating such efficiencies into TILLING programs. TbyS platforms such as those driven by Illumina sequencing technology have allowed SNP (single-nucleotide polymorphism) analysis across multidimensional pools and have opened the door to more efficient strategies for mutant detection. The ever-advancing high-throughput sequencing platforms (e.g. Illumina, PacBio, Ion Torrent, 454, SOLiDetc) combined with consistent reductions in the cost of sequencing per base will further drive the

integration of TbyS and TILLING platforms (van Nimwegen et al., 2016). Examples already exist of efficiency gains in screening EMS-mutagenised populations of rice and tetraploid wheat (Tsai et al., 2011). The use of multidimensional pools, multiplexed into sequencing lanes, allows the simultaneous detection of larger numbers of allelic variants, and increases the likelihood of detecting rare variants within the population (Table 10). The use of multiplexing does however impose certain challenges for downstream analysis. A comparative workflow of conventional TILLING and TILLING by sequencing has been depicted in the Figure 23.

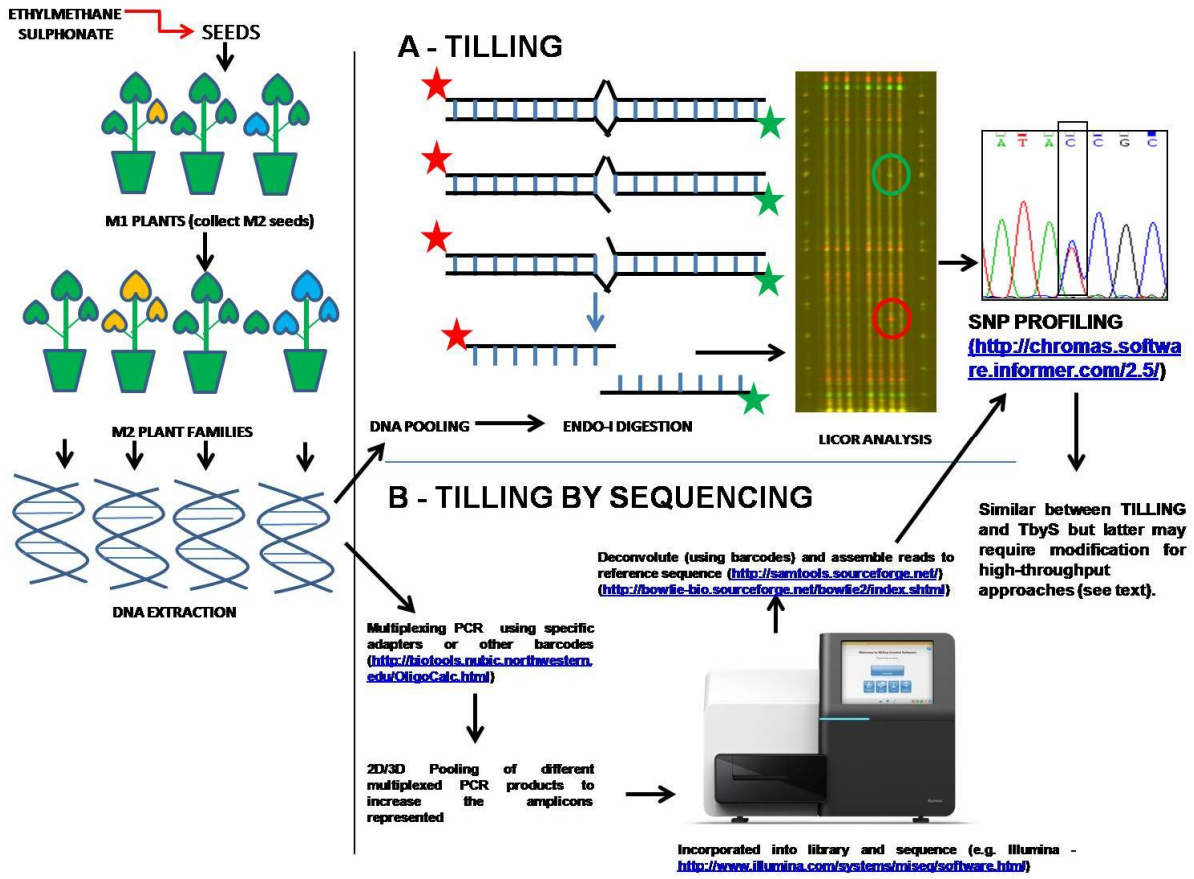


Figure 23. Schematic workflow of TILLING by Sequencing (TbyS). A schematic representation of (**A**) a traditional TILLING pathway; and (**B**) its adaptation into a TbyS pathway. The Illumina library is prepared and sequenced and undergoes software programming that leads to the assembling and SNP profiling, maximizing the coverage depending upon number of genes used for screening. Finally the mutant is detected and re-sequenced to confirm the mutant sample from the pool.

Modification	**Description and references**
Use of NGS in cereal crops	(Visendi et al., 2013); see also www.monogram.ac.uk/docs/MN2013/Monogram2013_Session1_CU.pdf.
High throughput TILLING pipelines	A high-throughput TILLING pipeline was deployed in tobacco using NGS, with a 3-D pooling methodology utilized to screen for mutants in genes for leaf biomass (*fw2.2/CNR* and *Ls*); biotic stress tolerance (*MYB12*); and altered leaf chemistry (*Geranyl Geranyl Reductase* and *Apiose Xylose Synthase*) from a total of 3162 mutations annotated from the genome (Reddy et al., 2012).
Identification of 'allelic series'	Suites of alleles have been used to study altered nodulation gene function in *Lotus japonica* (Perry et al., 2009).
Analysis of crop germplasm collections by Eco-Tilling	Morphological characterization of varieties has been performed using morphological and diversity indices based on visual phenotyping and semi-quantitative markers such as RAPD, SSR, STMS (Van Hintum et al., 2000); Eco-TILLING differs dramatically from general diversity analyses of germplasm collections as it is a reverse genetic approach which takes a candidate gene as the focus, and then uses the germplasm collection to identify cultivars, ecotypes or accessions harbouring variant versions; see (Comai et al., 2004; Cooper et al., 2013; McKeown et al., 2013b) and references therein.

Table 10. Examples of TILLING innovations in different plants and crop species.

In this chapter, we will utilize the Tilling by sequencing methodology to screen for the mutants from the *Camelina sativa* mutant population. Mainly four genes of interest in screening are *FAD2*, *FAE1*, *MIPS1* and *TGG1*. *Camelina sativa* has tremendous opportunity for improvement by various molecular techniques for various industrial applications. Camelina oil improvement by targeting genes which are involved in oil biosynthesis like *FAD2* and *FAE1* genes can lead to novel varieties for importance in the area of biofuel. To improve the oxidative stability in Camelina oils, increase in oleic acid (C18:0) and decrease in linoleic acid (C18:2) and linolenic acid (C18:3) is recommended. Oleic acid accumulation happens by the deregulation of Fattyacid desaturases enzyme (*FAD2*) in Camelina, which leads to the absence of polyunsaturated fatty acids (Kang et al., 2011). Similar beneficial traits are derived by down-regulating the enzyme, Fatty acid elongase (*FAE1*) which sequentially attach 2-C into oleic acid (C18:0) to make very long chain fatty acids (VLCFA)i.e. more than C20 (Hutcheon et al., 2010). Camelina meal is also an important by product after the removal of oil, which can be used as animal feed. The presence of some non-nutritive traits like, phytic acids and glucosinolates increases the toxicity of the meal for consumption by animals (Russo and Reggiani, 2012).Down-regulation of *MIPS1* (myo-inositol-1-phosphate synthase) which converts glusoce-6-phosphate to phytic acid will lead low phytate lines in Camelina (Colombini et al., 2014; Mitsuhashi et al., 2008). Transgenic with down regulation of the *MIPS1* gene showed 4.59 fold decrease in *MIPS1* gene expression, which lead to a significant decrease in phytate levels in rice (Ali et al., 2013).Presence of thiocynate is another important non-nutritive trait which formed from glucosinolates with the help of thioglucosideglucohydrolase enzyme (*TGG1*) (EFSA, 2008). Down regulation of *TGG1* enzyme resulted in low isothiocynate and increased glucosinolates level in arabidopsis, T-DNA insertion mutants (Barth and Jander, 2006). As *Camelina sativa* is a viable crop for oil seed production, there is an increasing demand for utilizing by-product i.e. Camelina meal or Camelina oil cake due to the high content of ω-3 fatty acids remnant as animal feed. Various animal breeding industries has evaluated the same and have shown that the meal may be cautiously used in poultry starter diet at a level of no more than 5% and may be integrated into hen finisher diet at a level of 10% without detrimental effects (Aziza et al., 2010b; Frame et al., 2008). The improvement of Camelina for the non-nutritive traits like reducing the concentration of phytic acid content and glucosinolates in seeds will have profound importance as animal feed. Present chapter deals with the details analysis of the technique of TbyS in Camelina by targeting the four genes of crop improvement.

3.2 MATERIALS AND METHODS

Tilling by sequencing in Camelina has been developed by a sequential workflow pattern described in the Figure 24. The Camelina mutant population developed in variety BBCAM2 by treating ethyl-methane-sulphonate (EMS) at 1.2% for 8 hours was utilised for the TbyS screening. DNA extraction was done with the help of DNeasy plant extraction kit (Qiagen) and same was diluted to 1:20 and further 2D1 pool was made for creation of the library.

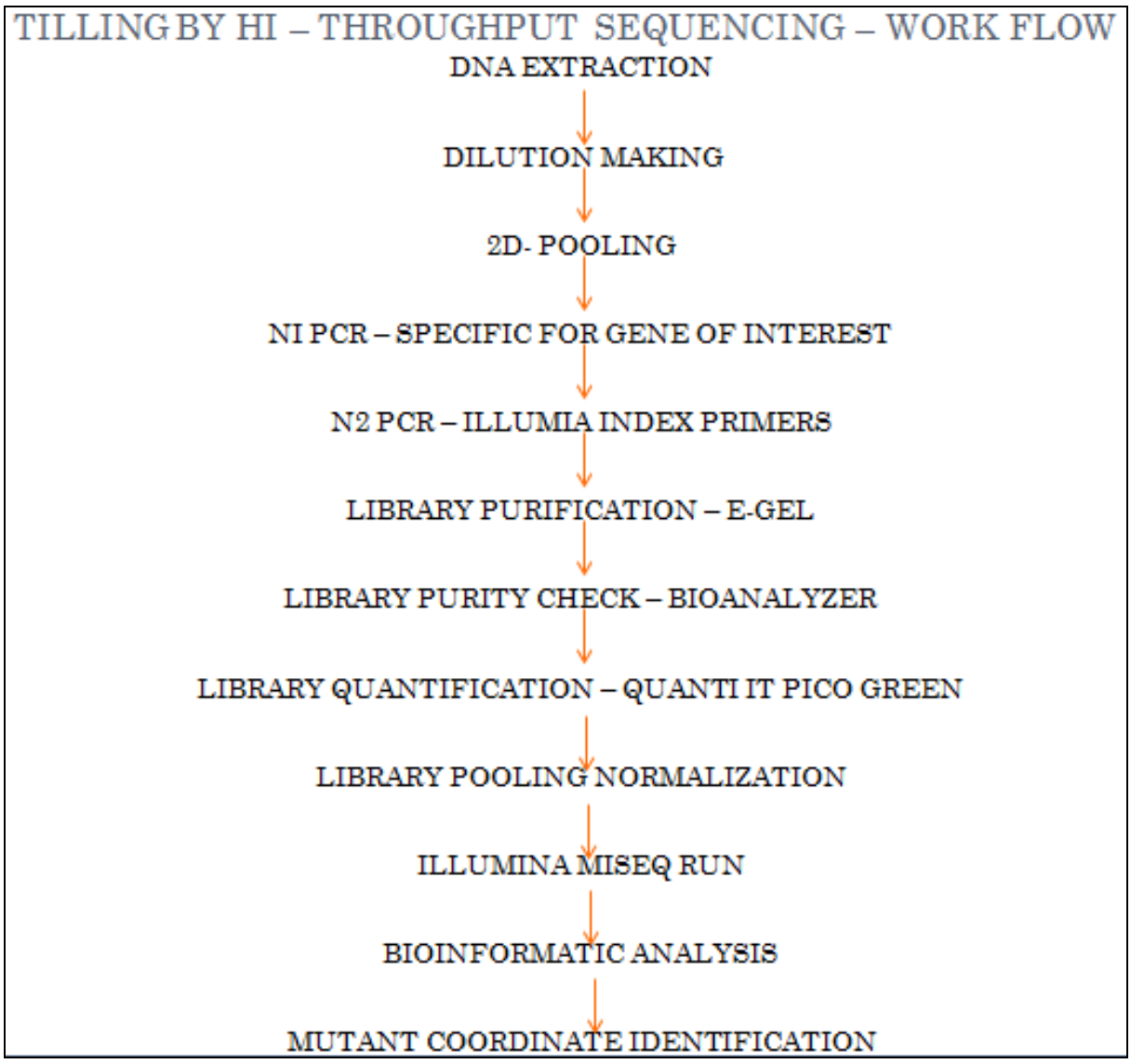


Figure 24.TILLING by High-throughput sequencing, TbyS in Camelina workflow.

3.2.1 gDNA 2-dimensional (2D) pooling

Camelina mutant population created in variety BBCAM2 was utilised for the gDNA extraction using Qiagen plant gDNA extraction kit following the user's manual. Camelina M2 DNA was pooled using a specialised 2-dimesional (2D) pooling strategy (Figure 25), where both column (C1A and C2A) and line (L1A, L2A and L3A) pooling was done. Camelina M2 gDNA plates were diluted to 1:10 dilutions and were used for making both column pool and line pool. Column pooling is done by mixing 16.0 µl each M2 gDNA from plate 1 and plate 2 to a new plate C1A, in which the well A1 will represent the gDNA from the plate 1 - A1 to H1 and plate 2 – A1 to H1. This will lead to the pooling of 16 individual M2 gDNA plate to a single C1A plate, with each having total 256.0 µl gDNA. Line pooling is done by mixing 21.0 µl each M2 gDNA from plate 1 to a new plate L1A, in which the well A1 will represent the gDNA from the plate 1- A1 to A12. This will lead to the pooling of 12 individual M2 DNA plate to a single L1A plate, with each having total of 252.0 µl gDNA. Add 250.0 µl of sterile distilled water to both column pool plate and line pool plate, to make a final dilution of 1:20. Distribute the 1:20 diluted gDNA from 4 plates i.e. C1A, L1A, C2A and L2A in the manner shown in TbyS pooling workflow (Figure 25) into 384 well plate with the help of BenchTop Pipettor (SORENSON BIOSCIENCE) (Appendices A2).

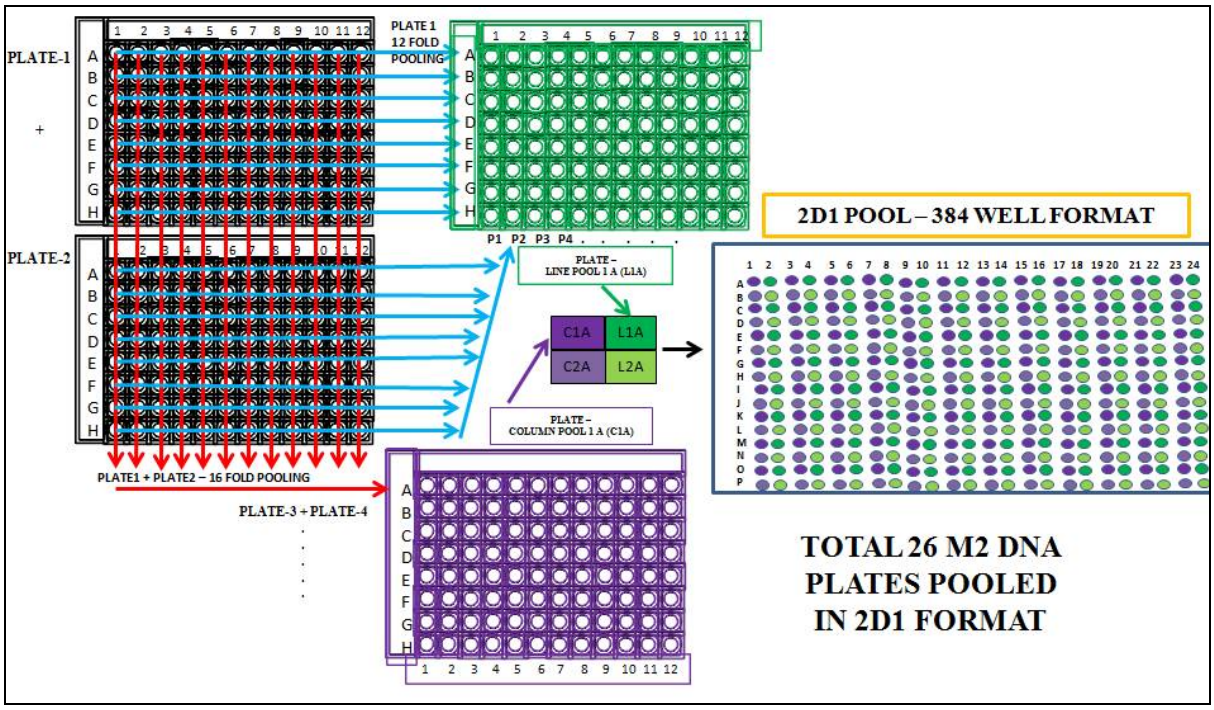


Figure 25. Tilling by high-throughput sequencing, TbyS in Camelina 2-dimensional pooling strategy.

3.2.2 Primer designing for TILLING by high throughput sequencing

Target genes for TbyS in *Camelina sativa* used were *FAD2*, *FAE1*, *MIPS1* and *TGG1* (Table 2 and 11). The maximum size of each amplicon to be amplified should not exceed 500bp, with primers having 20 bp length with 40-50% GC content. Primers were checked for hairpin and self-dimerization in oligonucleotide properties calculator software (http://www.basic.northwestern.edu/biotools/oligocalc.html). The first five nucleotides (dark red) of forward primer should have maximum diversity in all amplicons. The quality of sequencing profoundly decreases in the middle of the target region in the gene, so it is better not to centre the region of interest in the gene. Illumina adaptor is added to the 5' end of the specific region of the TbyS1 primer, forward primer 5' TTCCCTACACGACGCTCTTCCGATCT**XXXXX**xxxxxxxxxxxxxxx 3' and reverse primer 5' AGTTCAGACGTGTGCTCTTCCGATCTxxxxxxxxxxxxxxxxxxxx 3'. The maximum length of the TbyS1 primer should be not more than 46 base, the underlined part in both forward and reverse primer is the common part in the Illumina adaptor sequence and non-underlined part is the divergent part, which is part were the TbyS2 primers are hybridised in the nested PCR after TbyS1 PCR (Appendices A1).

TARGET GENES FOR NON-NUTRITIVE TRAITS				
S. No.	Target gene	Abbreviation	Function	Camelina Sequence ID
1	*MIPS1*	myo-inositol phosphate synthase	converts D-glucose 6-phosphate to 1D-myo-inositol 3-phosphate Synthesis of Inositol phosphates(Phytic acid)	Csa10g002540.1 Csa11g002930.1 Csa12g002720.1
2	*TGG1*	thioglucoside glucohydrolase 1	catalyzes the hydrolysis of glucosinolates into thiocyanate	Csa08g023710.1 Csa13g033940.1 Csa20g050340.1

Table 11. Target genes for TILLING by Sequencing in *Camelina sativa*

TbyS2 PCR primers are 384-well specific combination of a set of 16 row forward primers and 24 column reverse primers, the sequence of which are described in Table 12. Specific sets of TbyS2 Illumina primers (10 µM of concentration each - Eurofin), are made in four plates are distributed into the 384 well plates with the help of BenchTop Pipettor (Sorenson Bioscience).

Oligo_name	N2 Forward primer sequence (5' -- 3')	illumina name	tag
F_NGSi5_1	AATGATACGGCGACCACCGAGATCTACACTGAACCTTACACTCTTTCCCTACACGACGCTC	A501	TGAACCTT
F_NGSi5_2	AATGATACGGCGACCACCGAGATCTACACTGCTAAGTACACTCTTTCCCTACACGACGCTC	A502	TGCTAAGT
F_NGSi5_3	AATGATACGGCGACCACCGAGATCTACACTGTTCTCTACACTCTTTCCCTACACGACGCTC	A503	TGTTCTCT
F_NGSi5_4	AATGATACGGCGACCACCGAGATCTACACTAAGACACACACTCTTTCCCTACACGACGCTC	A504	TAAGACAC
F_NGSi5_5	AATGATACGGCGACCACCGAGATCTACACCTAATCGAACACTCTTTCCCTACACGACGCTC	A505	CTAATCGA
F_NGSi5_6	AATGATACGGCGACCACCGAGATCTACACCTAGAACAACACTCTTTCCCTACACGACGCTC	A506	CTAGAACA
F_NGSi5_7	AATGATACGGCGACCACCGAGATCTACACTAAGTTCCACACTCTTTCCCTACACGACGCTC	A507	TAAGTTCC
F_NGSi5_8	AATGATACGGCGACCACCGAGATCTACACTAGACCTAACACTCTTTCCCTACACGACGCTC	A508	TAGACCTA
F_NGSi5_9	AATGATACGGCGACCACCGAGATCTACACTATAGCCTACACTCTTTCCCTACACGACGCTC	D501	TATAGCCT
F_NGSi5_10	AATGATACGGCGACCACCGAGATCTACACATAGAGGCACACTCTTTCCCTACACGACGCTC	D502	ATAGAGGC
F_NGSi5_11	AATGATACGGCGACCACCGAGATCTACACCCTATCCTACACTCTTTCCCTACACGACGCTC	D503	CCTATCCT
F_NGSi5_12	AATGATACGGCGACCACCGAGATCTACACGGCTCTGAACACTCTTTCCCTACACGACGCTC	D504	GGCTCTGA
F_NGSi5_13	AATGATACGGCGACCACCGAGATCTACACAGGCGAAGACACTCTTTCCCTACACGACGCTC	D505	AGGCGAAG
F_NGSi5_14	AATGATACGGCGACCACCGAGATCTACACTAATCTTAACACTCTTTCCCTACACGACGCTC	D506	TAATCTTA
F_NGSi5_15	AATGATACGGCGACCACCGAGATCTACACCAGGACGTACACTCTTTCCCTACACGACGCTC	D507	CAGGACGT
F_NGSi5_16	AATGATACGGCGACCACCGAGATCTACACGTACTGACACACTCTTTCCCTACACGACGCTC	D508	GTACTGAC

Oligo_name	N2 Reverse primer sequence (5' -- 3')	illumina name	tag (5'3')
R_NGSi7_1	CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i1	ATCACG
R_NGSi7_2	CAAGCAGAAGACGGCATACGAGATACATCGGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i2	CGATGT
R_NGSi7_3	CAAGCAGAAGACGGCATACGAGATGCCTAAGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i3	TTAGGC
R_NGSi7_4	CAAGCAGAAGACGGCATACGAGATTGGTCAGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i4	TGACCA
R_NGSi7_5	CAAGCAGAAGACGGCATACGAGATCACTGTGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i5	ACAGTG
R_NGSi7_6	CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i6	GCCAAT
R_NGSi7_7	CAAGCAGAAGACGGCATACGAGATGATCTGGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i7	CAGATC
R_NGSi7_8	CAAGCAGAAGACGGCATACGAGATTCAAGTGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i8	ACTTGA
R_NGSi7_9	CAAGCAGAAGACGGCATACGAGATCTGATCGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i9	GATCAG
R_NGSi7_10	CAAGCAGAAGACGGCATACGAGATAAGCTAGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i10	TAGCTT
R_NGSi7_11	CAAGCAGAAGACGGCATACGAGATGTAGCCGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i11	GGCTAC
R_NGSi7_12	CAAGCAGAAGACGGCATACGAGATTACAAGGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i12	CTTGTA
R_NGSi7_13	CAAGCAGAAGACGGCATACGAGATTGTTGACTGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i13	AGTCAA
R_NGSi7_14	CAAGCAGAAGACGGCATACGAGATACGGAACTGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i14	AGTTCC
R_NGSi7_15	CAAGCAGAAGACGGCATACGAGATTCTGACATGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i15	ATGTCA
R_NGSi7_16	CAAGCAGAAGACGGCATACGAGATCGGGACGGGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i16	CCGTCC
R_NGSi7_17	CAAGCAGAAGACGGCATACGAGATCTCTACGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i17	GTAGAG
R_NGSi7_18	CAAGCAGAAGACGGCATACGAGATGTGCGGACGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i18	GTCCGC
R_NGSi7_19	CAAGCAGAAGACGGCATACGAGATCGTTTCACGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i19	GTGAAA
R_NGSi7_20	CAAGCAGAAGACGGCATACGAGATAAGGCCACGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i20	GTGGCC
R_NGSi7_21	CAAGCAGAAGACGGCATACGAGATTCCGAAACGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i21	GTTTCG
R_NGSi7_22	CAAGCAGAAGACGGCATACGAGATTACGTACGGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i22	CGTACG
R_NGSi7_23	CAAGCAGAAGACGGCATACGAGATATCCACTCGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i23	GAGTGG
R_NGSi7_24	CAAGCAGAAGACGGCATACGAGATTAGCTACCGTGACTGGAGTTCAGACGTGTGCTCTT	NexTill2_i24	GGTAGC

Table 12. TbyS Nested 2 PCR forward and reverse primer details from Illumina with primer name, primer sequence, Illumina primer name and specific sequence tag for primers.

3.2.3 TbyS PCR Optimization

Primer optimization is carried out based on nested-PCR in 96 well plates, with a total volume of 25.0 µl PCR mix. The TbyS N1 PCR mix contains, 10X PCR buffer, dNTPs (Invitrogen), Forward and Reverse TbyS N1 primers, 0.4 µM each (Eurofin), Taq polymerase (Sigma), 2.0 µl (10 ng) of Camelina genomic DNA, in a gradient PCR. The amplification of desired PCR product is checked in 2.0 % agarose gel. The same PCR mix further optimised in 384 well PCR 10.0 µl reaction volume with the forward and reverse primers final concentration of 0.1 µM each (Eurofin), results of which also is tested in 2.0% agarose gel. The PCR condition for the sequence specific N1 reaction is an initial denaturation of 94 °C for 5 min, 40 cycles of denaturation at 94 °C for 10 sec, annealing at 55 °C for 15 sec, and extension at 72 °C for 1min, and a final extension of 72 °C for 5 min and cool down to 4 °C. Further TbyS2 specific optimization forward and reverse primer is used for N2 PCR optimization. The PCR condition for the Illumina TbyS N2 PCR reaction is an initial denaturation of 94 °C for 5 min, 35 cycles of denaturation at 94 °C for 10 sec, annealing at 55 °C for 15 sec, and extension at 72 °C for 1min, and a final extension of 72 °C for 5 min and cool down to 4 °C. From the different combinations of primer pairs the best pair is selected after N1 and N2 optimization reactions, which is duplicated in the gDNA dilution (1:20) pool in 384 well plate. First PCR amplification is further carried out in the full 384 well TbyS 2D- pool plate with a reaction volume of 10.0 µl. After the N1 reaction, 2.0 µl of the first PCR is transferred to the TbyS N2 primer plate (as described in Table 12) with the help of BenchTop Pipettor (Sorenson Bioscience) with the help of one-touch tips (Sorenson Bioscience) and PCR is carried out with above TbyS N1 and N2 PCR conditions. Representational N1 and N2 PCR products are loaded in 2% agarose gel and confirmed the expected PCR product for the Illumina MiSeq Library preparation. About 3.0 µl of each amplified TbyS N2 PCR product is taken with a 16-channeled multichannel pipette (Thermo scientific) from all the wells of 384 well plates with help of Finntips (Thermo scientific) into a trough and pooled to a total volume of 1.0 ml of TbyS N2 PCR product into a 1.5 ml centrifuge tube.

3.2.4 TbyS library preparation and purification

3.2.4.1 E-gel library purification

The M2 Library is further taken to separate the desired target product from other unwanted PCR products by 2.0 % E-gel (Invitrogen) and the samples are purified, a total number of eight libraries are purified in one e-gel. About 20.0 µl of TbyS N2 PCR product is loaded into the upper wells and about 20.0 µl of sterile distilled water is loaded into the corresponding lower wells. A predefined 50bp ladder (Invitrogen Cat. No.: 10416-014) is loaded for checking the proper size of the target PCR product. The electrophoresis is switched-on to let the PCR product to migrate. The expected bands are further picked by pipetting out the sterile distilled water, when the PCR product reaches the lower well of the electrophoresis. Sample products are picked one by one from the same lower well until sufficient amount of the same is obtained and which is pooled into a new 1.5 ml centrifuge tube. About 100.0 µl of the samples are further concentrated in a speed-Vac system (Thermo Savant) and concentrated from a volume of 100.0 µl to 20.0 µl.

3.2.4.2 Library Quality check by BIOANALYZER

The e-gel purified Libraries were checked for purity by 2100 Bioanalyzer (Agilent). Before loading the samples, the chip priming station is set by setting the syringe by removing the plastic cap of the new syringe and the same is inserted into the clip. Further it is slided into the hole of the lucer lock adapter and screwed tightly to the priming station. The base plate and the syringe clip is adjusted according to the instructions in the user manual. Gel-dye mix is prepared and the mix is normalised to room temperature for 30 min before use. A new DNA chip was placed on the chip priming station and about 9.0 µl of gel-dye mix was pipetted into the prescribed position in the chip. The plunger is placed at 1 ml position and the chip priming station was closed. The plunger was pressed held by the clip, after 60 sec the clip was released and the plunger came back slowly to the 1 ml position. The chip priming station was opened and 9.0 µl of gel-dye mix was added in the prescribed well in user manual. About 5.0 µl of marker was added in the entire 12 sample and ladder wells and was made sure that no wells were empty. About 1.0 µl of purified library samples were loaded into the 12 sample wells and the chip was vortexed in IKA vortex mixer for 1 min at 2400 rpm. Further the chip was placed into the Bioanalyzer and the programme for analysing the sample was started.

3.2.5 Library Quantification by Quant-iT Pico Green Assay

M2 Libraries were taken for quantitation by Quant-iT Pico green dsDNA assay kit (Thermo Fisher Scientific), a total number of 40 samples fluorescence can be analysed in an Infinite M200 PRO (TECAN) spectrophotometer. The 20X TE solution (provided with the kit) was diluted to 1X TE and 100 µg/ml of standard DNA to 2 µg/ml. Aqueous working solution of the Quant-iT Pico Green reagent was made by diluting concentrated DMSO solution, 200-fold in TE. The samples are distributed in the S-Block, were five initial wells were kept for the DNA standard in varied dilutions as per the user manual (Final concentration of DNA standard is 1 µg/ml, 100ng/ml, 10ng/ml, 1ng/ml and blank). To each sample well 998.0µl of 1X TE and 2.0 µl of sample was added (TbyS library, and then finally 1.0 ml of diluted PicoGreen to each well was added. About 200.0 µl each of all the samples were pipette out (including the standards) from the S-block to a microtiter plate. As the PicoGreen is light sensitive the sample plate was protected from light by covering it with aluminium foil. The microtiter plate was placed in the TECAN spectrophotometer for quantification by giving sample coordinates in the Magellan software programme. The results were retrieved on an excel file after the analysis and standard curve showed the concentration of standard loaded along with the samples. The concentrations of different samples were obtained and the quantity to be used for loading to the Illumina MiSeq run was obtained.

3.2.6 TbyS TILLING Library Normalization and ILLUMINA MiSeq Run Loading

TbyS TILLING library which has been quantified was normalised by mixing all the samples together according to the concentration. A total number of 35 libraries can be loaded together for the MiSeq run, the required concentration of 2nM each with corresponding volume was added was made upto a total volume of 250.0 µl with sterile distilled water. The detailed descriptions of the step by step process for sample loading in the MiSeq run is available in the Illumina MiSeq system user guide. Before starting the run, the MiSeq reagent cartridge was kept in water for two hours from the deep freezer (-20^{o}C), where it is stored. PhiX library (10nM) control was used as an internal control for preparing all libraries for the MiSeq run, and the same is diluted to 8 pM which results in the cluster density of 1000-1200 K/mm^{2}.To dilute the PhiX library (Initial concentration - 10 nM) to 4 nM, 2.0 µl of the same is added with 8.0 µl of 10 mM Tris-Cl, pH 8.5 Tween 20. The 4 nM of PhiX library (10.0 µl) is then further diluted to 2 nM by adding 10.0 µl of 0.2 N NaOH. After briefly vortexing, the template solution was centrifuged at 280 xg for one minute, which is then incubated for five

minutes at room temperature to denature the PhiX library DNA into single strands. About 980.0 µl of pre-chilled HT1 solution (supplied with the Kit) was added with the 20.0 µl of the denatured PhiX library to obtain a final concentration of 20 pM PhiX library. From the diluted PhiX library (20 pM concentration) about 625.0 µl was added and a final concentration of 12.5 pM was made by mixing 375.0 µl of the pre-chilled HT1 solution. The remaining dilution of 0.2 NaOH was discarded.

The sample library was diluted the same way as described above for the PhiX library. The final dilution of the sample DNA library is diluted from 20 pM to a final concentration of 8 pM by adding 400.0 µl of the same with 600.0 µl of the pre-chilled HT1 solution. Finally the sample DNA library (Final concentration of 8 pM) and the PhiX control (Final concentration of 12.5 pM) are combined by adding 900.0 µl and 100.0 µl respectively. The samples were mixed and set aside on ice until the sample was ready to be loaded to the reagent cartridge. The reagent cartridge was taken out from water after two hours after defreezing and with a 1.0 ml pipette tip the foil over the reservoir labelled 'Load samples' was pierced. About 600.0 µl of the sample libraries was loaded into the 'Load sample' reservoir and the same was placed in the MiSeq. Flow cell (provided with the Kit) was taken out from the storage buffer and rinsed with sterile distilled water and then with ethanol, further cleaned with lint free tissue. The flow cell compartment door was raised and the flow cell latch is opened by pressing the release button to the right. The clean flow cell (label facing upwards) was placed into the flow cell stage and gently the flow cell latch was pressed down to close it over the flow cell and the flow cell compartment was closed. In the MiSeq screen 'Next' is pressed and the reagents door is opened. The reagent bottle PR2 (supplied with the kit) is loaded and the waste bottle is emptied, finally the reagent cartridge is loaded. The run parameters are reviewed by the equipment and the primer and the sample sheet are uploaded. A pre-run check is done and after all the parameters are OK then the run is initiated. During the run, monitor run progress, intensities and quality scores are displayed on the sequencing screen. When the run is complete, the review your results on the sequencing screen and proceed to the display Next, after which the post-run washes are done. For detailed protocol and troubleshooting information is described in the MiSeq (Illumina) user manual.

MiSeq run raw data is uploaded to a user interface software programme developed by IPS2-INRA, France was used for the identification of mutants. The user has to upload the exon structure of the gene, with amplicon sequence, and the CDS and protein information is automatically generated from the software.

3.3 RESULTS

Camelina M2 gDNA extraction yielded 40ng/µl DNA concentration and the dilution of 1:20 were adequate for the amplification of the target gene. N1 primers for TbyS in *FAD2*, *FAE1*, *MIPS1* and *TGG1* gene were selected. Different number of amplicons were made during the primer designing for TbyS, in case of *FAD2* – three amplicons, *FAE1* – three amplicons, *MIPS1* – two amplicons and *TGG1* – two amplicons were made to cover most of the gene (Table 13). Nested PCR for creating the library for different amplicons of the gene was done with N1 and N2 primers (Figure 26) (Appendices A1).

Gene Name	Size of the gene (CDS in bp)	Amplicons for TbyS	Size of the Amplicon (bp)
FAD2	1155	1	403
		2	460
		3	476
FAE1	1518	1	449
		2	518
		3	529
MIPS1	1533	1	457
		2	439
TGG1	1620	1	404
		2	414

Table 13.shows the total number of genes used for TILLING by Sequencing (TbyS), size of genes (CDS), number of amplicons made per gene and the size of gene.

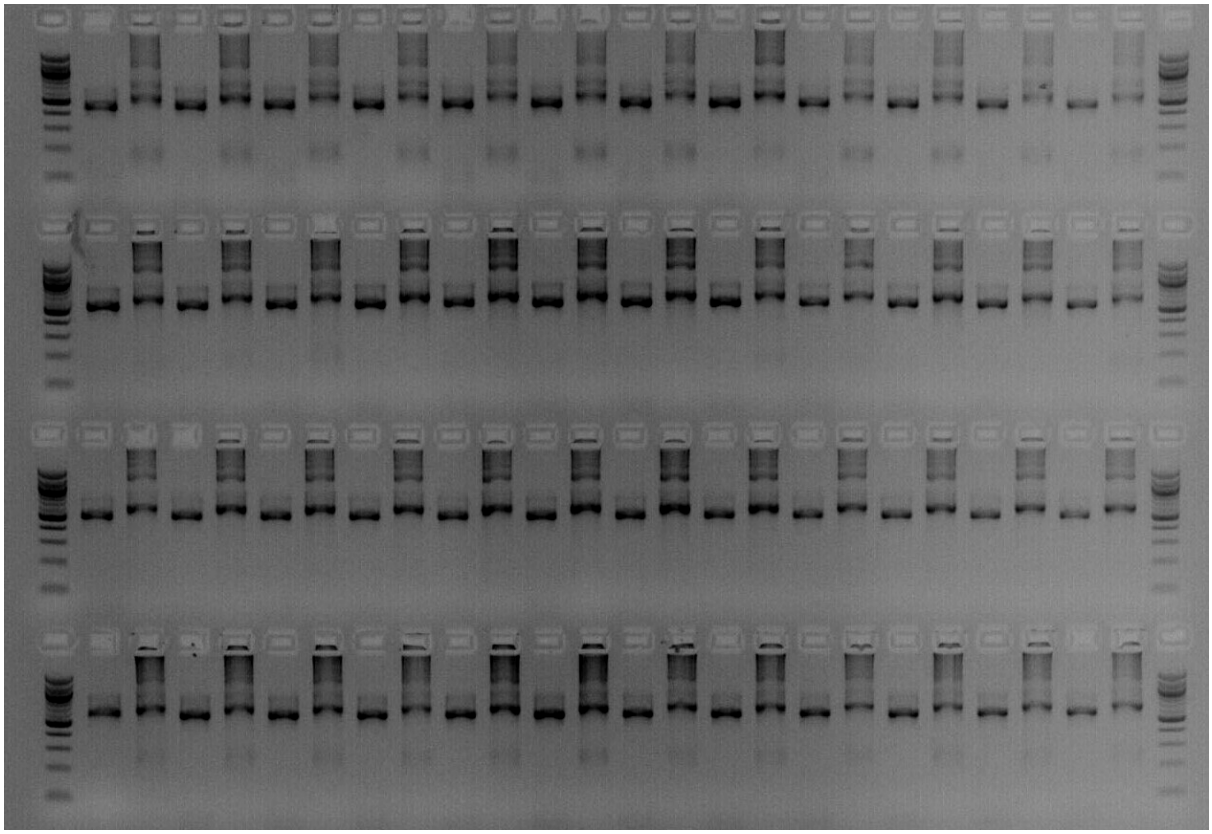


Figure 26. Alternate N1 and N2 PCR products of the *FAD2* Amplicon 1 and 3 in the first and second lane and *FAE1* Amplicon 1 and 2 in the third and fourth lane with 1Kb ladder on both ends of all four lanes.

N2 libraries of each amplicons are pooled were e-gel (Invitrogen) purified and the same was quality and purity checked by Bioanalyzer (Agilent). All the amplicons of all the four genes were e-gel purified and concentrated and tested for purity check with Bioanalyzer (Figure 27). Each library concentration and banding pattern was derived from the bioanalyzer data and was found to be of very good concentration. The peaks for each amplicons varied and had different concentrations, if the concentration was found less, the e-gel purification step was repeated to get a good concentration of each libraries. Further the MiSeq libraries were quantified by the Quant-iT pico green quantitation (Table 14), the total quantity according the average 2nM concentration of each library was used for the Library preparation of the MiSeq run.

Target libraries loaded into the Miseq undergoes four days sequencing and the raw data is uploaded to the User interface software developed by IPS2-INRA, France. The quality of the MiSeq run is established by the flow cell sample performance, base call of all the four bases during the sequencing run and the sample reads represented by heat-map figure (Figure 28). The raw data from the MiSeq run undergoes multiple alignments and coordinates are

designated with the help of the MiSeq adapter primer data uploaded by the user through the software program for the analysis.

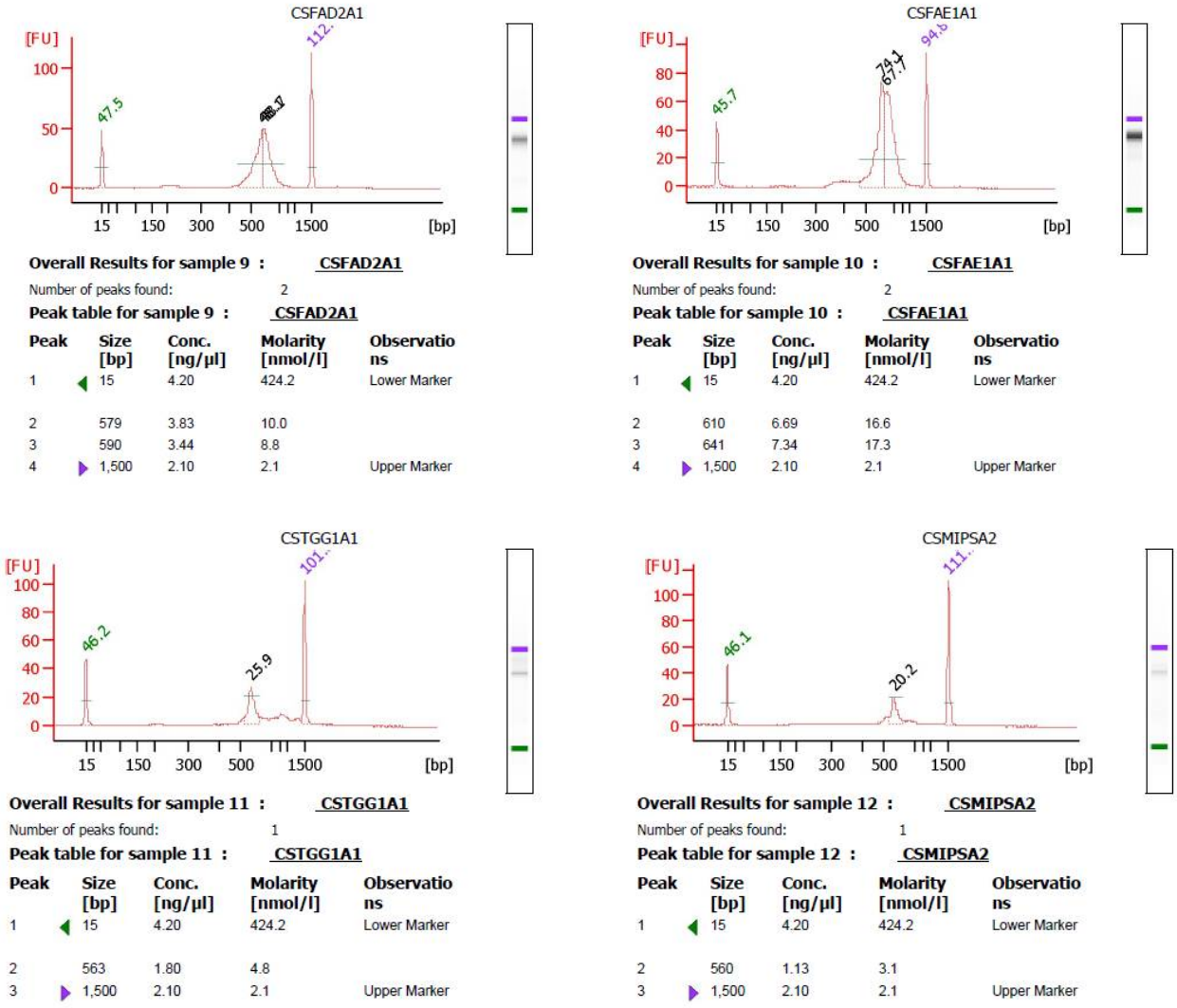


Figure 27. Bioanalyzer data for CS*FAD2* – Amplicon 1, CS*FAE1* –Amplicon 1, CS*TGG1* – Amplicon 1 and CS*MIPS1* – Amplicon 2, showing the peaks and banding pattern showing the purity and concentration of the library.

Targets	Size of the Target (bp+adapters)	Abs	Abs. (- blank)	ng/µl	MW (ng/mol)	nmol/l	Volume of targets in µl [2nM]
FAD2-A3	613	249	189	3.70	4.01E+14	9.24	1.55
FAD2-A3	613	249	189	3.70	4.01E+14	9.24	1.55
FAE1-A2	655	274	214	4.19	4.28E+14	9.79	1.46
FAE1-A2	655	274	214	4.19	4.28E+14	9.79	1.46
MIPS-1-A1	594	160	100	1.96	3.88E+14	5.05	2.83
MIPS-1-A1	594	160	100	1.96	3.88E+14	5.05	2.83
TGG1-A2	547	180	120	2.35	3.58E+14	6.57	2.17
TGG1-A2	547	180	120	2.35	3.58E+14	6.57	2.17

Table 14. Quanti-iT Pico green quantitation data for all the four gene (represented twice to increase the three copies representation in the MiSeq run) *FAD2*- Amp 3, *FAE1*- Amp2, *MIPS1* – Amp1 and *TGG1*- Amp2, showing the size of the target with the adapter, absorbance, concentration (ng/µl), molecular weight, and volume of targets required to be loaded to the MiSeq run.

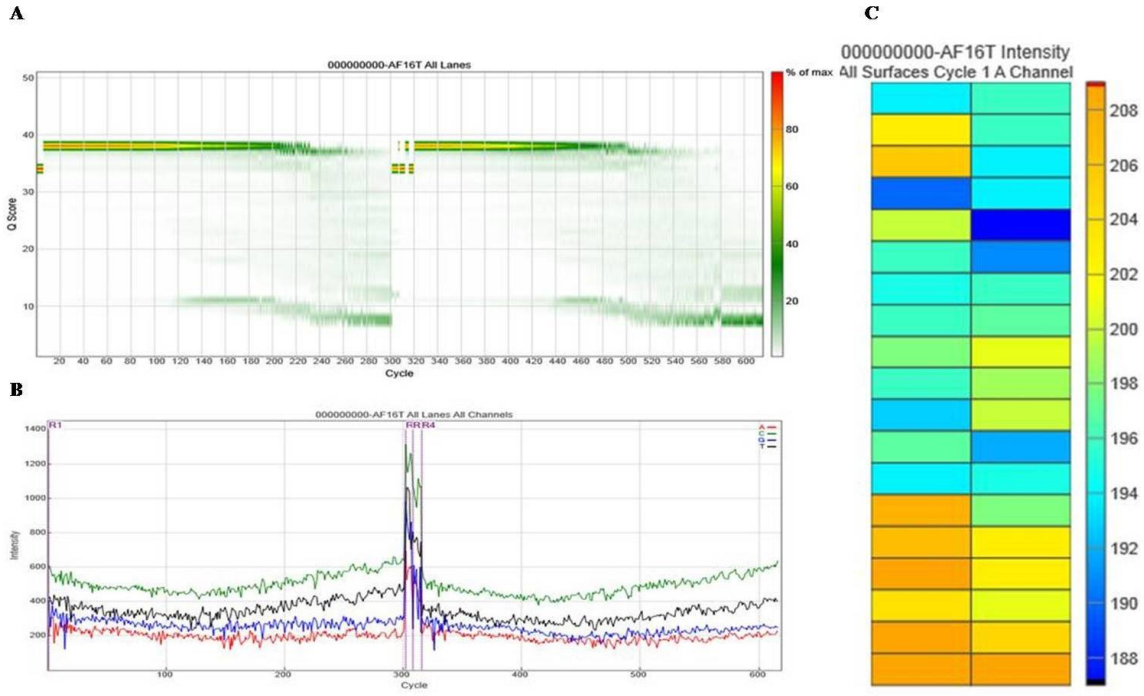


Figure 28. MiSeq run data shows A. flow cell sample performance, B. Base call representation of Adenine (green), Guanine (black), Cytosine (blue), Thiamine (red). C. Heat map representation of MiSeq Library sample reads.

Further from the analysis specific doublons (mutation detected in both column and line pool) and simplons (mutations observed in only one of the column or line pool) are obtained, which are having high frequency for each mutants. The mutants are further tabulated and coordinates of M2 family are derived from the retrieving the pooling of the NGS N2 primers and the M2 2D1 pool coordinates.

A total number of 183 mutations were observed across all the four genes which were taken for TbyS in Camelina i.e. *FAD2*, *FAE1*, *MIPS1* and *TGG1* (Appendices A3). Mutants obtained from *FAD2* were 66 in which 40 were missense (out of which 15 mutations were – Non-tolerated according to SIFT analysis), and 26 were silent mutations. Mutants obtained from *FAE1* yielded four non-sense mutations with a total of 82 mutations including missense and silent ones. In case of *MIPS1* and *TGG1* a total number of 21 and 14 mutations were obtained in different M2 families respectively. Camelina mutant population had a total mutation frequency of 1 in 55 kb by screening the four genes, in each gene the mutation frequency are described in the Table 15.

		TYPES OF MUTATIONS							
Gene	**Fragment size (bp)**	**Non sense**	**Splice**	**Missense - non tolerated**	**Missense tolerated**	**Silent mutations**	**Intronic mutation**	**Total mutations**	**Mutation Frequency (1/kb)**
FAD2	*1058*	*0*	*0*	*25*	*15*	*26*	*0*	*66*	1/40
FAE1	*1306*	*4*	*0*	*22*	*34*	*22*	*0*	*82*	1/40
MIPS1	*888*	*0*	*1*	*10*	*4*	*5*	*1*	*21*	1/106
TGG1	*814*	*0*	*0*	*4*	*5*	*5*	*0*	*14*	1/145
Total	**4066**	**4**	**1**	**61**	**58**	**58**	**1**	**183**	**1/55**

Table 15. Mutation frequency of Camelina mutant population screened by TILLING of *FAD2*, *FAE1*, *MIPS1* and *TGG1* genes with fragment screened anf types of mutations found. The mutation frequency for each amplicon is calculated as follows: [(size of the amplicon) × (total number of samples screened)] / (total number of identified mutants). The average mutation frequency was estimated to be one mutation per 55 kb by TbyS in camelina.

3.4 DISCUSSION

Approach for focusing high-throughput sequencing on regions (exons) of the genome most likely to alter protein expression (and protein-function dependent phenotypes) is to use exome capture as has been performed in wheat, rice and African rice, *Oryzaglaberrima* (Henry et al., 2014). The exome captured portions of these genomes represent the regions of the genome which are transcribed into protein-coding mRNAs. Such targeted reduced-representation approaches do not generate data on the majority "non-coding" regions of the genome (which do not code for proteins). Exome capture portions of many plant genomes can potentially be multiplexed up to 30-fold and sequenced with a standard Illumina platform. Exome capture approaches provide cost-benefit advantages for larger crop genomes which contain extensive tracts of non-coding repetitive sequences. For smaller crop genomes, such as rice, the cost-benefit of exome sequencing are less evident (Henry et al., 2014).

All TbyS requires robust bioinformatic protocols for handling the datasets produced by these high-throughput workflows. In many cases, this has been achieved by developing 'in house' bioinformatic workflows, e.g. using bespoke Python scripts. For example, the analysis of mutagenized wheat and rice populations by (Tsai et al., 2011) was performed using Coverage Aware Mutation calling using Bayesian analysis – CAMBa (Missirian et al., 2011). Development of bespoke bioinformatic pipelines allows an integrated approach to TbyS workflow development. For example, CAMBa takes account of both the pooling setup and sequencing coverage levels when calculating mutation and noise probabilities, maximising the reliability of its outputs. CAMBa also takes account of the barcoding of multiplexed samples to identify the sub-pools in which mutations of interest have occurred (Tsai et al., 2011). Similarly, other groups have employed a newly developed pipeline called *M*utations *a*nd *P*olymorphisms *S*urveyor, or MAPS (Henry et al., 2014). MAPS is of particular utility in the analysis of polyploid genomes as it uses each sample as a control against every other, allowing it to distinguish between allelic variants due to divergence between homeologues and induced mutations (Henry et al., 2014).

However, the need to develop such 'in house' solutions to bioinformatic analysis of TILLING populations also poses challenges in terms of bioinformatic expertise required to establish and sustain the bioinformatics pipeline. In the absence of community standardisation, bespoke approaches to bioinformatics pipelines run the risk of reduced

compatibility between TbyS datasets. Hence, greater standardisation of TbyS bioinformatics analysis pipelines is desirable.

The complete datasets produced by TbyS experiments from different labs should ideally be archived for use by the wider scientific community, subject to commercial interests, and according to best practice for data archiving (Sulpice and McKeown, 2015). To meet this need, some existing databases have been developed by labs in UC Davis (http://comailab.genomecenter.ucdavis.edu/index.php/TILLING_by_Sequencing) and INRA (http://www-urgv.versailles.inra.fr/TILLING/TILLING-by-sequencing.htm) for understanding and interpreting the pools of data produced by these approaches.

3.4.1 TILLING in polyploid genomes

Most crop plants have polyploid genomes (wheat, potato, Brassicaceae family, okra and many fruit crops), thus the use of high-throughput TILLING in polyploids is of particular interest in the field of crop improvement (Tsai et al., 2013). TILLING by Sequencing in *Camelina sativa* has been utilized in this chapter to generate lines with improved oil quality and other non-nutritive traits. Mainly to improve oil quality two genes were focused i.e. *FAD2* and *FAE1* gene, which were earlier used for Camelina oil improvement. Mutation frequency of *FAD2* (1/40kb)and *FAE1* (1/40kb) were comparable with the results of 1/75 kb to 1/200 kb (Hutcheon et al., 2009). The fact that polyploids have a higher gene dosage than diploids poses challenges in this regard, due to the risk of phenotypes being masked by complementation from another homeologue. However, this can also be an advantage if more subtle effects are being sought. Different types of mutations were observed in present study in case of *FAE1* gene where four STOP mutants were obtained, which can be potential lines for generating zero erucic lines. Even *FAE1* mutants with zero erucic acid lines were obtained from M13 TILLING, which are comparable to the ones obtained in *Brassica napus* (Wu et al., 2008).

A approach for TILLING autopolyploid *Arabidopsis thaliana* has been proposed as a strategy for generating higher mutational load in TILLING populations (Tsai et al., 2013). This mutant population was raised by treating the wild type Col-0 (2n=10) and autotetraploid Col-4X (4n=20) with about 30 mM and 50 mM EMS. A three-dimensional pooling method was again utilized and the analysis carried out with a modified version of CAMBa, adapted for the pooling scheme (Tsai et al., 2013). TILLING by sequencing resulted in 19.4 mutations per Mb in the 15 genes which were analyzed.

Some areas of existing and future application of TbyS are summarized in Table 16. In this context, TILLING provides new possibilities for areas such as nutritional and organoleptic improvement in crops (Elahi et al., 2015; Minoia et al., 2016), changing seed oil properties (Elahi et al., 2015; Scully et al., 2015) and improving plant biomass process ability (Scully et al., 2015).

.

Category	Aspect amenable to TbyS analysis	Rationale
Basic scientific applications	Non-sequenced genomes	Allows analysis of target genes from unsequenced and poorly-annotated genomes; it may be cost-effective to combine this with exome-capture in the case of highly-repetitive genomes.
	Polyploid genomes	TbyS may be used to screen for variation in unique regions of homeologues, increasing the likelihood of identifying phenotypes.
	Plant architecture	Novel gene variants affecting e.g. spikelet arrangement in barley (Gottwald et al., 2009) ; such approaches can be used to improve crops in accordance with the plant 'ideotype' concept.
	Promoters	Suites of point-mutations in promoter (and enhancer) regions can be used to screen for novel gene regulation motifs.
	Orphan genes	Genes with no known similarity or conserved domains are amenable for functional characterization of different regions by TbyS.
Applied agricultural outputs	Maintenance of yield under changing environmental conditions	Many crop breeding pools are limited by genetic bottlenecks during domestication and adaptation to intensive agriculture which reduces the variation available in environmental response; TbyS could identify novel alleles to help maintain yield under abiotic stress.
	Detection of genes underlying QTL	There has been interest in using TILLING to identifying genes causal for QTL (Quantitative Trait Loci) of agronomic interest; the number of genes which need to be assessed in this way depends upon marker density and the decay of linkage disequilibrium and may be large.
	Pathogen resistance	Strong R-genes generate a strong selective pressure for the evolution of resistance in the pathogen; suites of variant alleles have the potential to overcome this resistance (Nieto et al.,

		2007)
	Time to flowering	Examples of single amino acid changes altering the time to vegetative-reproductive transition have been described, suggesting that screens for induced variation could identify novel means of controlling vernalization, flowering, anthesis and grain-filling in many crops.
	Micronutrient deficiency	The enzyme pathways involved in uptake and/or biosynthesis of micronutrients lacking in many human diets are typically well known but subtle changes induced by TbyS may alter their regulation or activity and so change flux through pathways.
	Forestry	TILLING and eco-TILLING have been applied to some tree species, suggesting TbyS could be used to identify novel germplasm for forestry and tree biomass production.
Conservation science	Eco-Tilling of endangered populations	Application of TbyS to eco-TILLING could determine levels of diversity within species and identify those populations which should represent priorities for conservation efforts.

Table 16. Examples of possible future applications in the development of TILLING by Sequencing

3.4.2 TILLING-based technology and the rise of genome editing technologies

There is rapid growth underway in the development and use of genome editing (GE) technologies to directly modify nucleotides or sequence tracts within a locus, e.g. to introduce premature stop codons in the 5'-region, or other targeted changes to an open reading frame (ORF) (Quetier, 2016). Although several such methodologies have been proposed for use in crops and other plants (ZFN, TALENs), major interest has focussed on the *Streptococcus*-derived CRISPR/Cas9 system (Belhaj et al., 2013; Cermak et al., 2015; Svitashev et al., 2015; Wang et al., 2014; Zhu et al., 2016).

The major advantage of genome editing approaches is that desired specific mutation(s) or changes can be introduced to the genome, including the possibility to target multiple genes in diploid genomes, or indeed homeologs in polyploid genomes (Belhaj et al., 2015; Lowder et al., 2015; Shan et al., 2013; Upadhyay et al., 2013; Wang et al., 2014). Genome editing by CRIPSR/Cas9 is also possible in perennial species such as poplar (Fan et al., 2015) and has been applied to plant pathogens such as *Phytophthora*species (Fang and Tyler, 2015). CRISPR/Cas9 approaches were initially developed using easily identified targets genes (such as the tomato *ARGONAUTE7* (*SlAGO7*) which produces a wiry-leaved phenotype when mutated (Brooks et al., 2014), to demonstrate proof of concept prior to moving on to targets of research or agricultural interest. The use of CRISPR/Cas9 systems for genome editing is now rapidly spreading across research labs, with a wide range of applications emerging including gene replacement (Li et al., 2016), promoter swaps (Shi et al., 2016), metabolic engineering (Alagoz et al., 2016) and QTL editing (Shen et al., 2016), targeted to modify a wide range of traits (e.g. nitrogen fixation (Wang et al., 2016), plant virus resistance (Chandrasekaran et al., 2016), etc). CRISPR/Cas9 has also been proposed as a means of targeting non-protein coding genes such as lncRNA and miRNAs (Basak and Nithin, 2015), and is also subject to various improvements to accelerate the speed and reliability of locating homozygous mutants such as systems specifically expressed in germline tissue (Mao et al., 2016; Wang et al., 2015b).

To move beyond discovery science, a key challenge for genome editing (and TILLING approaches) based on candidate genes (or more precisely candidate SNPs) is some prior predictive knowledge of what phenotypic changes can be elicited by specific nucleotide changes to a particular gene or locus. For instance, genome editing or TILLING approaches may be suitable for engineering point mutations conferring single-locus heterosis, as we have

previously suggested (McKeown et al., 2013a). Both genome editing and TILLING in their current forms can (in principle) be applied to generate allelic series of the same gene to determine whether different variants of the same gene can display different phenotypic effects.

At present, an important commercial (as opposed to fundamental scientific) consideration when comparing TILLING approaches to genome editing approaches is that the regulatory status of products derived from genome editing technologies (such as CRISPr/Cas) currently remain unclear in many jurisdictions (Morris and Spillane, 2008; Ricroch et al., 2016; Sprink et al., 2016; Voytas and Gao, 2014; Waltz, 2016; Wolt et al., 2015). In contrast, in most jurisdictions the products (e.g. crop lines) derived from TILLING approaches are currently not classified as GMOs (Morris and Spillane, 2008).

3.5 CONCLUSIONS

Whole genome sequencing has led to a rapid increase in the number of TILLING populations generated for crops, where lines generated from TILLING populations are making their way into commercial breeding populations. TILLING by Sequencing provides a means for the future identification of novel variation across the genomes of crops and of model systems in a highly efficient manner. We anticipate that the generation of novel genetic variation in crops via TILLING will be facilitated over the next decades by the expected drastic reduction in the costs of TbyS, which will be a powerful complement to the rise of the new CRISPR/Cas9 genome editing within the plant science research and crop breeding communities. TbyS in Camelina will give plant breeders to work on a varied number of alleles of various industrial applications. The high mutation frequency in Camelina has proved the power of the TbyS technique. The mutant lines developed from the Camelina mutant population will have tremendous applications in the field of Biofuel development, Cosmetic industries, and animal husbandry.

4. CHAPTER FOUR

Study of *Camelina sativa* variety BBCAM2 mutant population phenotype by Mutant seed size analysis and phenotype cataloguing

ABSTRACT

Creating a mutant population has been the main source for plant breeders aiming to harness the variation induced by mutagens like irradiation or chemical treatment. The mutations which lead to small deletions or single nucleotide base changes can lead to variation in genomic as well as morphological characters, leading to profound changes in a particular trait. The *Camelina sativa* mutant population created by mutating the BBCAM2 variety by ethane methane sulphonate (EMS) has lead to the development of more than 5000 M2 families which showed different types of phenotypes, which have significant agronomical and industrial applications. In the present chapter two main phenotyping studies are presented: seed size variation and the tabulation of various phenotypes in the M2 generation, mainly focused for obtaining data on the change in flowering time in the 1000 M2 screened. Seed size variation was determined using Image J from images taken by bifocal microscopy, which showed differences in the overall size of the M2 seeds. A total number of 23 M2 family were selected which showed higher average size which were more than 1.7 mm^2, as compared to the control seed size of 1.57 mm^2. Within the mutant population smaller seeds were also detected which showed less than 1.2 mm^2, which is a common morphological change which is found in mutant populations. About five mutant families were selected with larger seed size which showed significant variation from the Camelina BBCAM2 seed size. The larger seed size mutant lines have potential to be utilised for improvement in the overall yield by weight of the total seed production, which is an important agronomical important trait. Further the creation of phenotypic database Camelina mutant population to generate early flowering lines was also done, where about ten mutant lines were detected which showed early bolting by about ten days as compared to control. These lines can be utilised to aid Camelina improvement for early flowering lines which can help lowering the duration of the crop cycle. In a snapshot, creation of phenotype database of Camelina will help plant breeders to work on this material for future crop improvement programme for various important targets.

4.1 INTRODUCTION

Phenomics is the biological study of different measurements of physical and biochemical traits of organisms, which are changed due to genetic mutations or environmental impacts (Houle et al., 2010). Creating a mutant population by artificial means like chemical mutagenesis in a crop helps us to attain enormous number of phenotypes, which can be now characterised genetically with the help of genomic sequencing. Crop genetic improvement has been also successfully demonstrated by development of various genetically modified crops by GM technology like BT-cotton, Bt-Corn although there are concerns over of the use of transgenic crops due to increased level of resistance towards *Bacillus thuringensis* (Reisig and Reay-Jones, 2015).

Phenotype database from a mutant population is an important resource for both forward and reverse genetics screen. Species like *Pisum sativum* are obstinate to transformation through *Agrobacterium* based method, so creating a mutant population can create more variability in such crop for potential phenotype studies. Both phenotypes and genotypes (sequence information) have been catalogued in various crops like garden pea (*Pisum sativum)*, flax (*Linum usitatissimum)* and is available in database, UTILLdb – http://urgv.evry.inra.fr/UTILLdb (Chantreau et al., 2013; Dalmais et al., 2008). A detailed phenotype database of *Sorghum bicolor* was created and can be accessed online through http://www.lbk.ars.usda.gov/psgd/sorghum/till/index.aspx, in which all details of the seedling phenotypes, tillers, leaf necrosis, senescence, panicle, and seed size in the M2 and M3 generations has been compared with the sorghum wild type BTx623 (Xin et al., 2008). In case of UTILLdb, we can make online search for TILLING alleles through a BLAST tool, or mutant phenotypes can be searched by keywords. A phenotype database of sunflower (*Helianthus annuus*) was also created, where detailed phenotypes like albino, sterile flowers character, double flowering heads, plant height phenotypes along with others were tabulated (Kumar et al., 2013). A number of phenotypes were observed when a soybean (*Glycine max*) mutant population was created by irradiation of fast neutrons (FN) at a dose of 32-Gy, particularly for leaf, trichome and root phenotypes (Bolon et al., 2011).

In Cotton, seed size was found to be closely related to oil content, germination and seedling emergence but there was no significant association between seed size and protein content (Pahlavani et al., 2008). In *Brassica rapa*, larger seed size lead to tolerance towards flee beetle damage, where tolerance correlated to higher seed weight which also resulted in

vigorous growth of those seeds (Elliott et al., 2007). Seed and pod characteristic in Broad bean (*Vicia faba* L.) was varied in different categories within a population but there was no significant change in seed weight (Li and Yang, 2014). A detailed review on the molecular mechanism of seed size development had been detailed in *Arabidopsis thaliana* (Orozco-Arroyo et al., 2015), where different genes controlling the seed size variation has been tabulated.

Genes like Auxin response factor 2 (*ARF2*), *APETALA2*, *DA1* and also *KLU* have been characterised for the seed size in Arabidopsis (Adamski et al., 2009). Seed size is a QTL trait and has been complex to analyze but with various high-throughput approaches for measuring the seed size and distribution, new QTL have been discovered (Herridge et al., 2011). A leucine-rich repeat kinase, *HAIKU2* gene was over-expressed in Arabidopsis and was found to have phenotypes with increased size and weight along with overall oil content (Fatihi et al., 2013).

In the present chapter we have mainly focused in tabulating the various phenotypes of mutant population and seed size analysis of the mutant seeds. This will help the breeders to correlate various traits of industrial importance and ultimately lead to Camelina crop improvement.

4.2 MATERIALS AND METHODS

Mutant population of *Camelina sativa* L. Variety BBCAM2 was created by treating 12,000 seeds were treated at a dose of 1.2% EMS. Seeds were sown in nursery at Benchbio Field, Kullu, HP, India, and after 30 days the M1 seedlings were transplanted to raise beds with spacing 30 cm x 45 cm, with temperature ranging between 20-30˚C. Plants were grown till maturity with all the phenotypes at nursery stage, vegetative stage and flowering - Seed stage were collected and tabulated. Similarly, 10 seeds each of 1000 M2 seeds were grown in the same region with same agronomic practice described above. M2 seedlings were grown till maturity and with M3 seeds collected individually from M2 plants.

4.2.1 Seed Size Measurements

Mutant population (M2 family) seed sizes were measured in triplicate by taking photographs of seeds on white paper with a bifocal microscope (Leica). Images were further analysed by Image J software, seed images are converted to solid black images and the seed size were calculated. 50 seeds were taken in triplicates from each M2 family. Statistical analysis on the seed size data were done by XL-STAT and Image J software along with Sigma plot software.

One way ANOVA with XLSTAT was done for analyzing the comparison of M2 families with larger seed size with control seeds size.

4.2.2 Phenotype catalogue

Seedlings sown for M2 DNA extraction phenotypes were looked upon for various phenotypes. A total number of 1000 M2 family (10 seeds) were planted for generating M3 seeds and phenotypes at vegetative and reproductive stage were studied. A tabulation of various flowering time study of these 1000 M2 family was also done.

4.3 RESULTS

4.3.1 Mutant seed size analysis

The Camelina mutant population created showed varied number of phenotypes at each stage. Mutant M2 seed size variation was quite profound with average size of wild type Camelina seeds variety, BBCAM2 to be 1.57 mm^2 and the seed size were distributed from range of 1.0 to 2.4 mm^2 (Figure 29). The size of the seeds were showing a sizeable difference, which was tabulated and variance plot was made in the form of Bell curve sigma plots. A comparative analysis of CSC2-12M2(1.2)-4 mutant line with control showed marked difference in the seed size with an average size of 1.896 mm^2 (Figure 30).

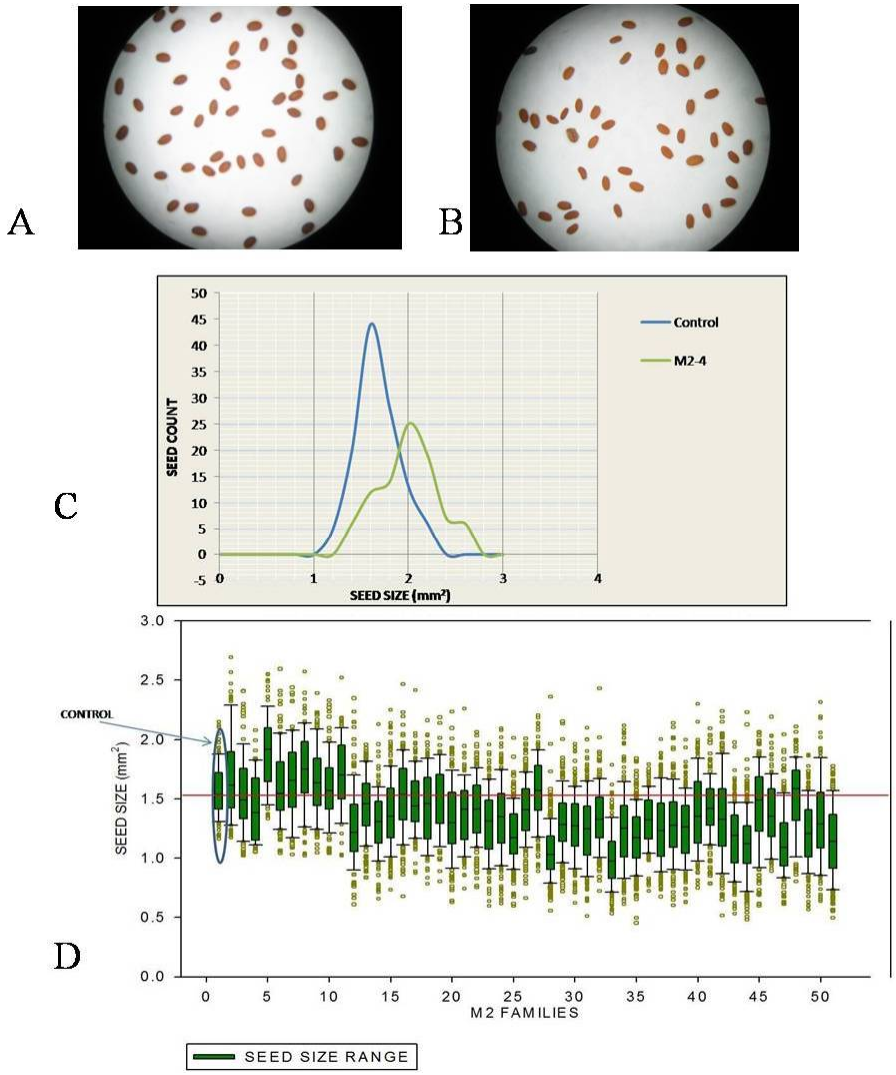


Figure 29. Seed size variation analysis with Image J software **A.** Image of control seeds, **B.** Image of mutant line M2 – 4 seeds, **C**. Histogram line plot showing distribution of the seed lots compared with the seed count **D.** Sigma plot showing the variation of seeds with control and within the M2 – 1 to M2 – 50 seeds plotted with the seed size (mm^2). The red line shows the median line of 1.57 mm^2 measured in control seed lot. Green block shows the seed size range.

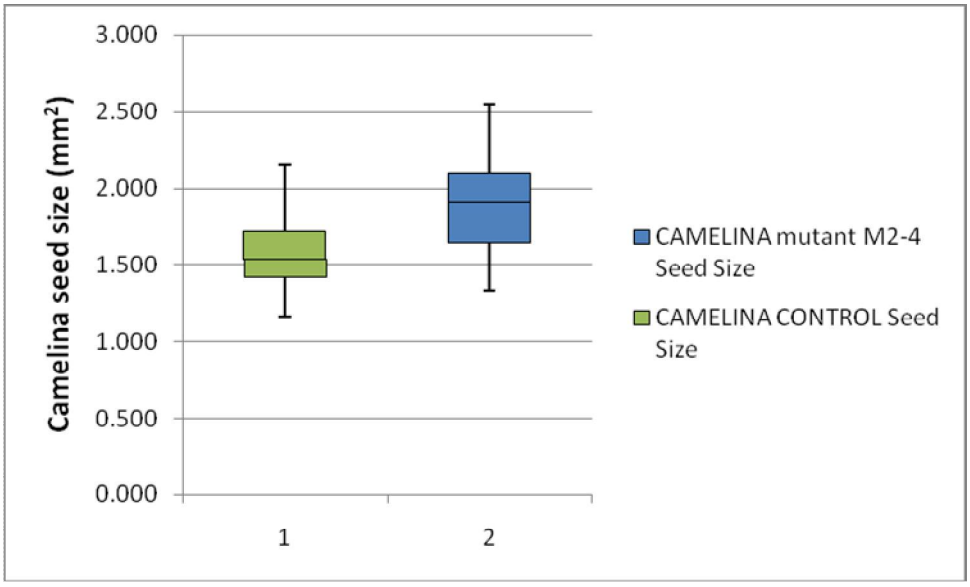


Figure 30. Box and whisker plot for comparison between control and Mutant M2 family CSC212M2(1.2)-4.

Significant variance was observed with P-value of 3.565E-16, when one way ANOVA was performed by comparing control seeds with M2-4 seed size. Further all the M2 families seed size were analysed and plotted against the control for the variance in curve. Out of the 1000 M2 families, a total number of 23 M2 families showed larger seed size, with size more than 1.7 mm^2 and about 95 M2 families showed smaller seed size, with seed size less than 1.2mm^2, as compared to the control (Figure 31). A comparative analysis of the large seed size with details and serial numbers of the M2 seed families are given in Figure 30. Camelina mutant families with larger seed sizes were significantly different from the control seed size (Figure 32). Five M2 families (M2-4, M2-168, M2-253, M2-273, M2-307) were selected for further analysis to look for significance of variance by one way ANOVA, and the P-values were well below 0.05, which shows the M2 families were showing larger seed sizes. These analysis were carried out by Tukey test and Dunnett test which compared each M2 family lines with the control to look for significant variation (Table 17 and 18).

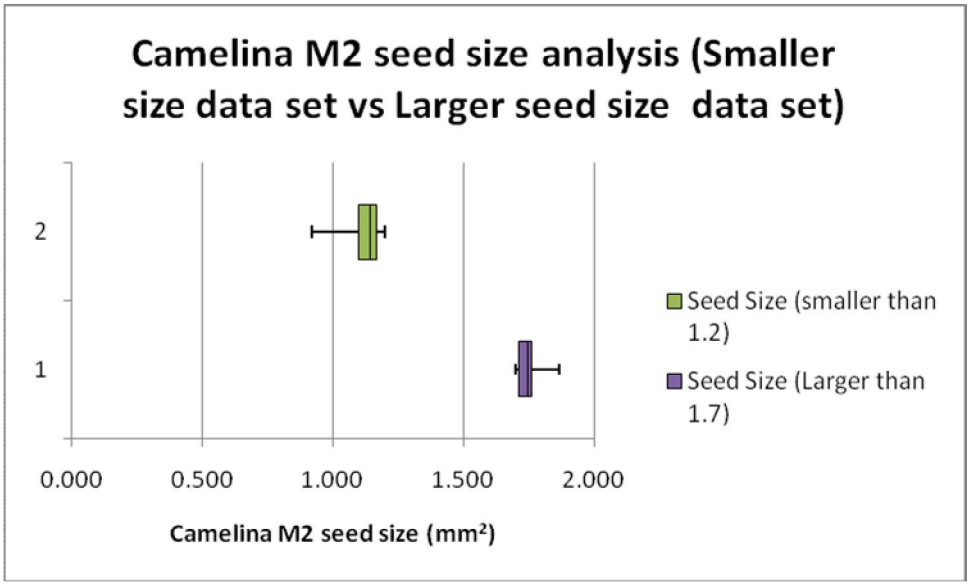


Figure 31. Box and whisker plot for the camelina M2 families seed lots with smaller seed size (green) with size below 1.2 mm^2 and larger seed size (purple) with size above 1.7mm^2 datasets.

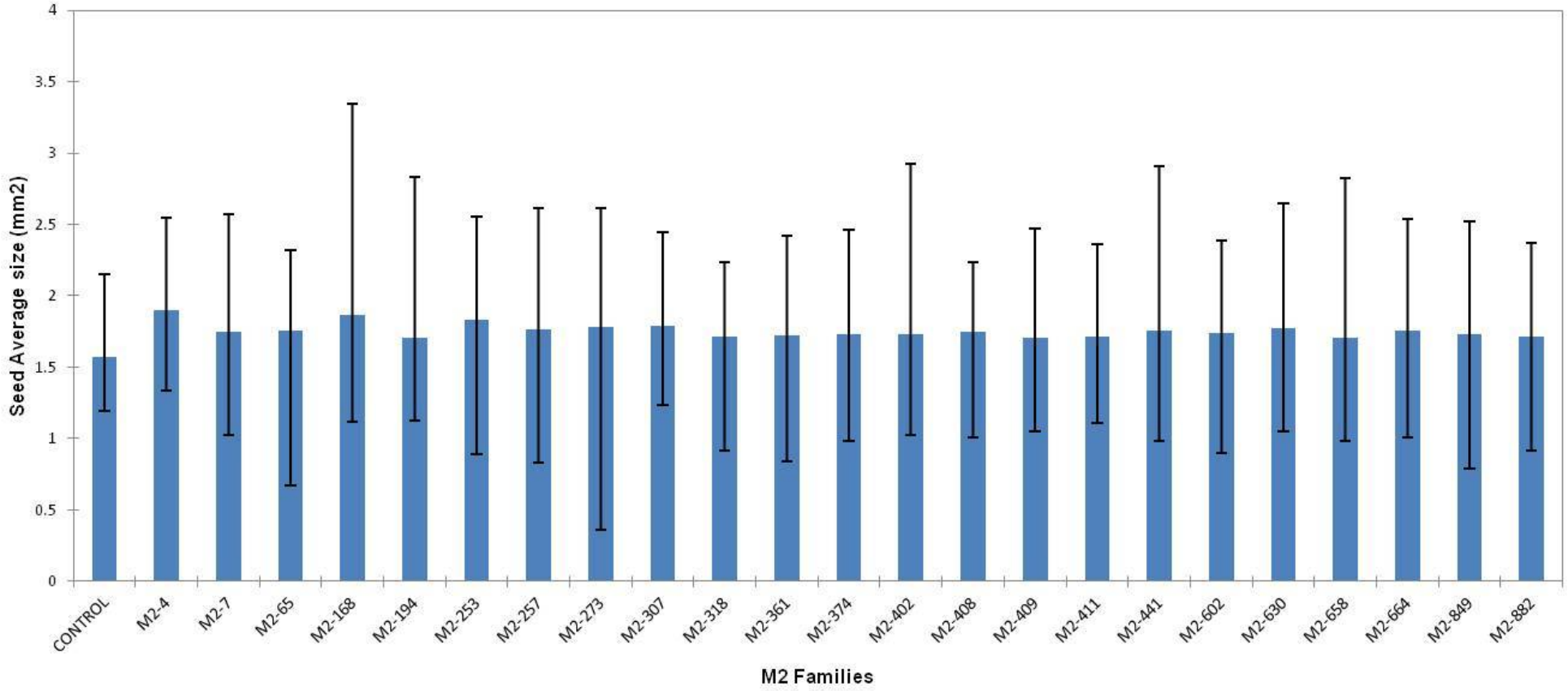


Figure 32. Comparison of average seed size of *Camelina sativa* variety BBCAM2 control seeds with 23 M2 families selected with larger seed size. Error bars shows the standard deviation of the seed lot within M2 families.

Contrast	Difference	Standardized difference	Critical value	Pr > Diff	Significant
M2-4 vs CONTROL	0.325	7.246	2.858	**< 0.0001**	Yes
M2-4 vs M2-273	0.110	2.263	2.858	0.211	No
M2-4 vs M2-307	0.104	2.040	2.858	0.321	No
M2-4 vs M2-253	0.067	1.583	2.858	0.610	No
M2-4 vs M2-168	0.029	0.652	2.858	0.987	No
M2-168 vs CONTROL	0.296	7.173	2.858	**< 0.0001**	Yes
M2-168 vs M2-273	0.081	1.789	2.858	0.473	No
M2-168 vs M2-307	0.075	1.568	2.858	0.620	No
M2-168 vs M2-253	0.038	0.987	2.858	0.922	No
M2-253 vs CONTROL	0.258	6.645	2.858	**< 0.0001**	Yes
M2-253 vs M2-273	0.043	1.006	2.858	0.916	No
M2-253 vs M2-307	0.037	0.811	2.858	0.966	No
M2-307 vs CONTROL	0.221	4.593	2.858	**< 0.0001**	Yes
M2-307 vs M2-273	0.007	0.126	2.858	1.000	No
M2-273 vs CONTROL	0.214	4.676	2.858	**< 0.0001**	Yes
Tukey's d critical value:			4.043		

Table 17. Camelina seeds / Tukey (HSD – Honest significant difference) / Analysis of the differences between the categories with a confidence interval of 95% (Seed Average size (mm^2)

Contrast	Difference	Standardized difference	Critical value	Critical difference	Pr > Diff	Significant
CONTROL vs M2-4	-0.325	-7.246	2.521	0.113	**< 0.0001**	Yes
CONTROL vs M2-168	-0.296	-7.173	2.521	0.104	**< 0.0001**	Yes
CONTROL vs M2-253	-0.258	-6.645	2.521	0.098	**< 0.0001**	Yes
CONTROL vs M2-307	-0.221	-4.593	2.521	0.121	**< 0.0001**	Yes
CONTROL vs M2-273	-0.214	-4.676	2.521	0.116	**< 0.0001**	Yes

Table 18. Camelina seeds / Dunnett (two sided) / Analysis of the differences between the control category Camelina seeds-CONTROL and the other categories with a confidence interval of 95%

4.3.2 Phenotyping of Camelina mutant population

Mutant population of Camelina generated by EMS mutagenesis showed wide phenotypic variation in the M1 generation (Details in Chapter 2). Lots of plants showed leaf variegation, pale green leaves, stunted-dwarf type and early flowering and late flowering type phenotypes were observed. Seed germination was also recorded which showed no germination in about 63 M2 families from the 1000 M2 sown (Figure 32). About 393 M2 family showed germination of less than five plants out of ten seeds sown, which are typical nature of mutant population where seed germination lethality is quite common due to mutations in essential genes. Seedling stage phenotypes of mutant population M2 families showed a number of phenotypes like albinos, cotyledon arrest, primary leaf arrest, three cotyledons, variegated leaves, early flowering and other phenotypes were observed (Details in Chapter 2). A phenotypic dataset was made to classify these morphological characters in the M2 generation. About 10 seeds were sown from the 1000 M2 seed lot and phenotypes of the seedlings were recorded at the seedling stage, vegetative stage and flowering stage (Figure 34 and 35). About 10,000 M2 individuals (1000 M2 family) were grown to maturity in field and various phenotypes were catalogued. A total of 1.1% of albino was observed from the 5000 M2 screened and a total of 9.4% of various phenotypes were observed (Chapter 2 - Table 2).

Various other phenotypes like early bolting, vigorous plants, dark green colour, were observed in the vegetative and flowering stages which have been tabulated, where phenotypes along with the flowering dates have been detailed, about 10 lines which showed about 10 days lower time for flowering (i.e. CSC2-12M2(1.2)-252, 335, 356, 366, 493, 586, 908, 979, 986 and 994), than the control line which takes about 97 days to flower after sowing. These lines could be used for characterising the genes controlling the early flowering character by positional cloning.

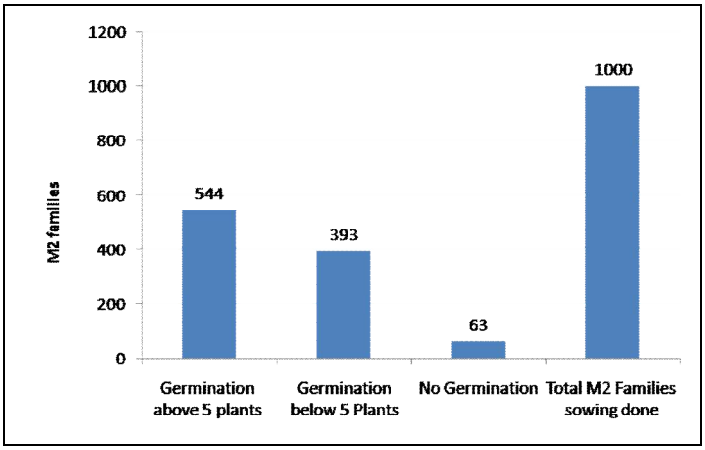


Figure 33. Germination of M2 (Mutant population second generation) family, a total number of 1000 M2 families – 544 M2 showed germination of more than five seedlings and about 393 showed less than five seedlings out of 10 seeds sown. About 63 M2 showed no germination, showing lethality in this generation out of the 1000 M2 seed lot sown.

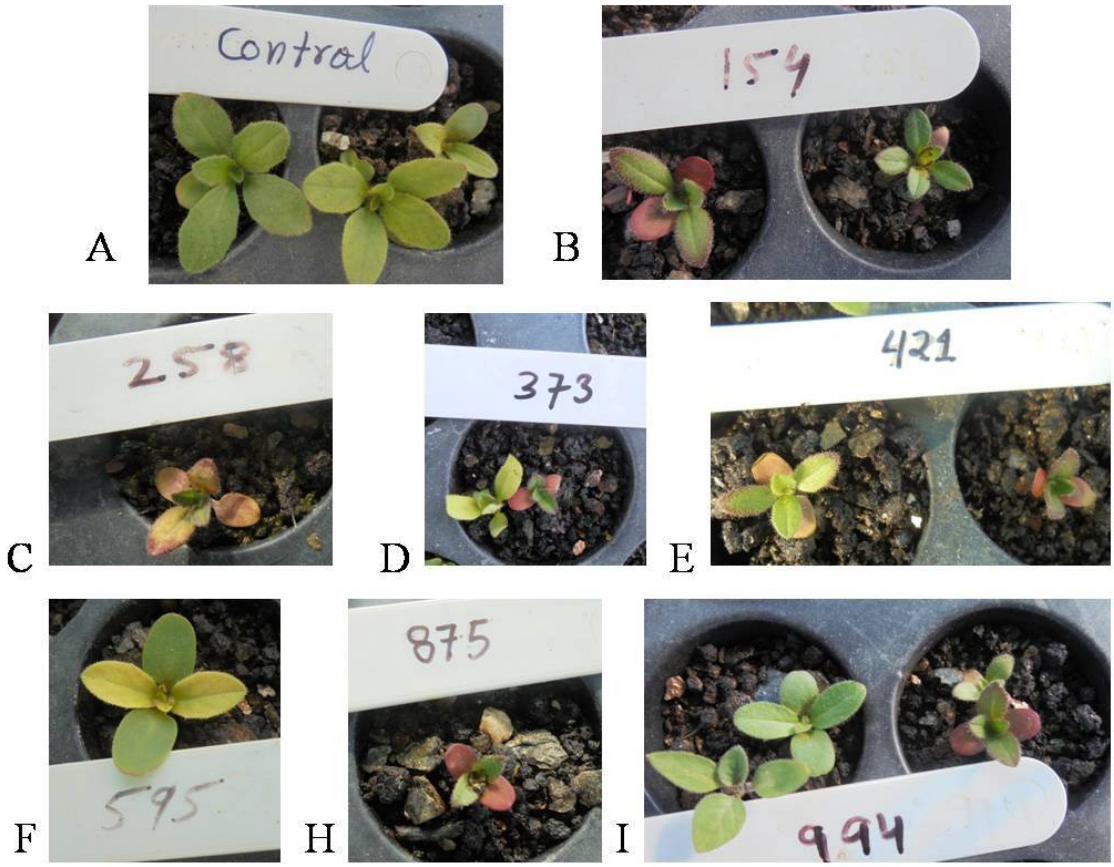


Figure 34. Phenotypes observed during 15 – 20 days seedling stage in M2 generation – A. Control, B. M2-154 with red leaves, C. M2-258 variegated red leaves, D. M2-373 with pale green leaves, E. M2-421 shows Yellow red leaves, F. M2-595 with Yellow leaves, G. M2-875 with red and green leaves, H. M2 – 994 with Red and green leaves.

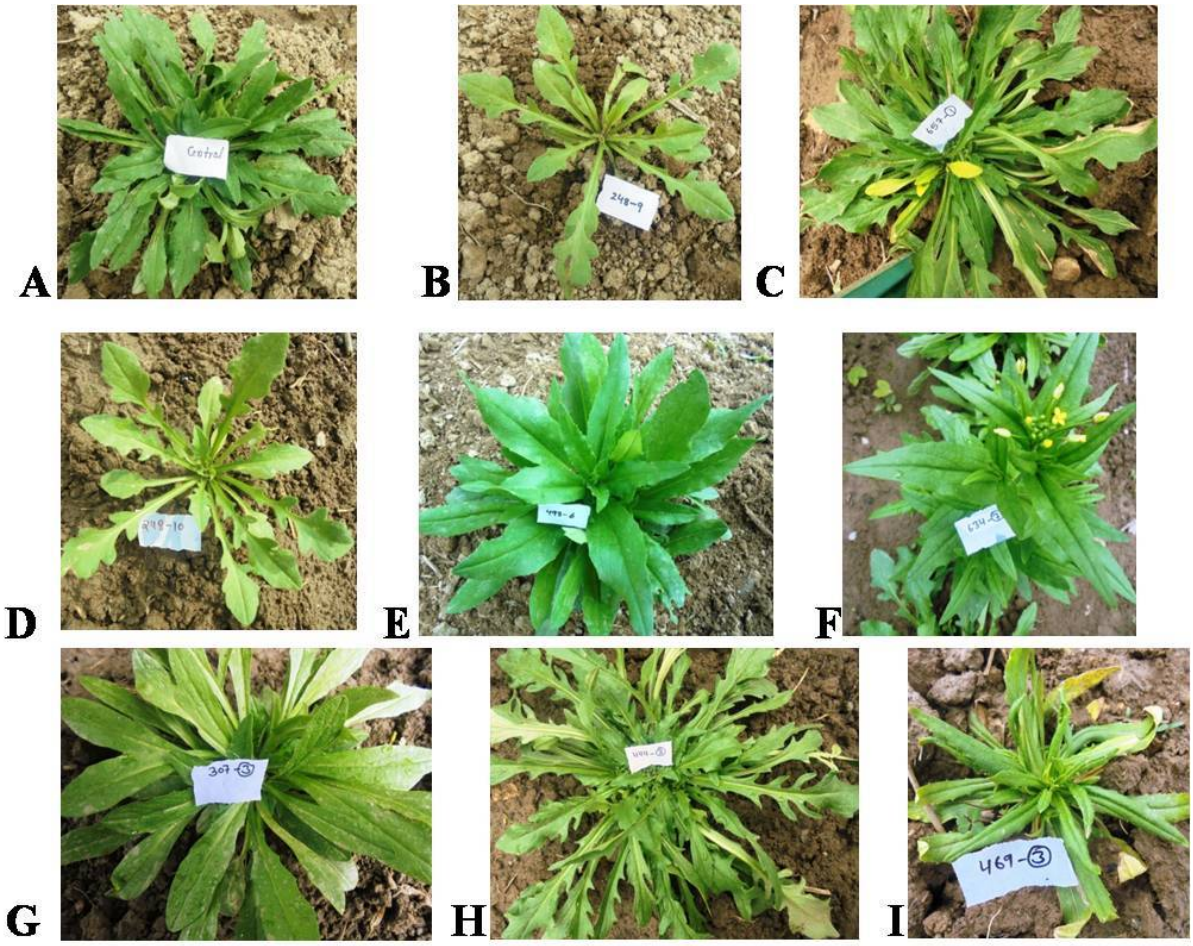


Figure 35. Various phenotypes observed in M2 families A. Control plant, B. Pale green plant, C. Variegated plant, D. Yellowish green plant, E. Dark green and erect leaf plant, F. Early flowering plant, G. Broad leaf plant, H. Serrated leaf plant, I. Pointed leaf plant.

4.4 DISCUSSIONS

Phenotyping of mutant population has been an important method by which various important agronomical important traits are characterised. Mutant population Phenotyping in tomato showed variation mostly in vegetative organ with mainly on plant size, plant habit and leaf morphology (Laskar et al., 2016) which are similar results as obtained in present work. The high mutation frequency in Camelina genome is indicated by the presence of large number of mutant phenotypes at various stages of crop cycle, which was also proved by TILLING of various genes. Similar results of varied phenotypes were attributed in Sorghum mutant population which was also proved by TILLING on myoinositol kinase (*MIK1*), phytochrome A (*PHYA*) and other genes (Xin et al., 2008). High through method of mutant population Phenotyping is being developed to increase the availability of data by including software programming and modern imaging techniques like the one developed in tomato mutant population i.e. TOMATOMA (http://tomato.nbrp.jp/indexEn.html) (Okabe et al., 2013; Okabe et al., 2011). Sol genomics network has developed a search engine for the utilization of solanaceae family crop improvement from the various mutant population repository (https://solgenomics.net/search/phenotypes) submitted in the network, similar phenotype search engines can be developed for Camelina crop improvement from the mutant population generated in the present study. Similar database are developed in brassicaceae family (https://brassica.nbi.ac.uk) and other crops, which can be utilized as a database structure on which a Camelina phenotype database can be built. Phenotypes observed in camelina M1 generation could be correlated with the M2 phenotypes obtained, but as the M1 generation is more chimaeric for mutations, the phenotype observed in the M2 generation gives better insight into the nature of mutation caused. Various phenotypes observed like dwarf type, leaf variegation, leaf serration phenotypes are significant as the variability in germplasm will lead to the development of ornamental camelina and will have better applicability in ornamental industry. Leaf variegation and serrated type of mutants could be bred to develop hybrids with both phenotypes.

4.4.1 Camelina seed size - Multiple gene family control

Seed size analysis done in the present study could be utilized for forward genetic screen of the genes controlling the expression patterns. Mutants characterized for the genes specific for the expression within seeds like expression at embryo, endosperm, seed coat and developing seeds

could be utilized as potential targets for TILLING of the selected lines (Orozco-Arroyo et al., 2015). Significant variation was observed between control and mutant M2 families which showed larger seed sizes in the present study. Seed size is a complex trait and the phenotype of seed size is controlled by a ubiquitin receptor, Arabidopsis DA1, which controls the final seed and organ size. A negative activity towards DA1 and a DA1-related (DAR) protein by the mutant protein encoded by the *da1-1* allele was observed. Increase in seed size was observed by the over-expression of *da1-1* cDNA in wild type plants, which shows that it is an important gene family which increases the seed size (Li et al., 2008). DA1 gene family in Camelina mutants which shows larger seed size, could be analysed to confirm which gene family controls the phenotype. Seed size as described earlier is a complex trait, which is controlled by the interaction of zygotic embryo (maternal) and the endosperm (parent plant) in Arabidopsis. Arabidopsis mutant, megaintegumenta (*mnt)* developed showed increased seed size and weight, which was due an extra cell division in the integuments. *mnt* is a mutant allele of AUXIN RESPONSE FATOR 2 (ARF2), which mediates gene expression in response to auxin. Further in the research on MNT/ARF2 gene in arabidopsis it was found that it is a repressor of cell division and growth (Schruff et al., 2006). As camelina is a Brassicaceae family oil seed crop similar to arabidopsis the analysis of these gene family could be correlated to the mutants obtained from camelina mutant population.

Phenotypes like seed size could vary due to various environmental factor affecting during the normal growth of a mutant population. Environmental effects on the camelina seed size could be rectified by growing and producing the seeds in closed green house conditions, where both temperature and humidity could be controlled. Further these mutant lines could be back crossed with camelina BBCAM2 wild type parent line so that the background mutations could be removed.

ANOVA (analysis of variance) test gives us the results are significant or not and it does not explain where such difference occurs. Statistical analyses carried out on the seed size analysis were mainly Tukey (HSD – Honest significant difference) and Dunnett (two sided). Tukey test does analysis of the differences between two different categories with a confidence interval of 95%. This in conjunction with ANOVA helped us to decipher the significant difference between the control camelina seeds and the mutant lines, which increases the confidence to select these

lines for further phenotypic studies. Dunnett test compares each M2 family with the control, which is also know as many to one comparison, results from which increases the confidence of the M2 seed size analysis significantly. Seed size analysis was done on camelina M3 seeds, by taking experimental triplicates to reduce the error within the seed family.

Various phytohormones like auxins, cytokinins, abscisic acid have also important functions towards the overall development of seeds. TILLING or TILLING by sequencing (TbyS) of the genes controlling such critical biosynthetic pathway will help us to understand the mutations undergone in these genes, controlling the seed size variations. Further these M3 seed families could be utilized for breeding purpose to increase the overall yield and other mutations detected in other genes could be stacked to develop camelina lines with improved quality.

Innovations in the space of phenomics have been profound by the integration of Bioinformatics and Imaging of various Phenotyping programme like the Lemnatec 3D Scanalyzer for Next generation Phenotyping, which helped in studying the Phenotyping character of wheat for drought and salinity tolerance studies (Parent et al., 2015; Rajendran et al., 2009). These types of techniques can be utilised for studying the intricate details of phenotypes in Camelina mutant population developed in the present studies.

4.5 CONCLUSIONS

Camelina mutant population developed by the mutagenesis of ethane methane sulphonate (EMS) has been proved to be one of the best mutant populations developed with the detailed cataloguing and tabulation of various phenotypes. These phenotypes of various agronomical importances have tremendous implication for plant breeders who can develop and take this information for further improvement of the germplasm. Larger seed size mutants can help the overall increase in the oil quantity and this can have major improvement in overall oil yield, as seed size is directly correlated with the increase in the overall oil concentration. Early flowering lines will help to lowering the duration of the crop cycle by which more than one crop can be raised in one year. TILLING of this mutant population with integrating more Industrial important target genes can help to add up more relevance to this work.

5. CHAPTER FIVE

Summary and Future Directions

Chapter 5 – Summary

5. Summary and Future Directions

Camelina sativa is an underutilized oil-seed crop of immense importance for various industrial applications. As it can be grown marginal lands with low nitrogen and low water inputs, it can be used as an efficient model crop for oil seed production for biofuel as well as edible purpose. We have demonstrated by TILLING and TILLING by Sequencing, the both objectives stated have been achieved by the present work.

TILLING and TILLING by sequencing in Camelina mutant population have shown the high throughput nature of these techniques. Development of mutant line has been efficiently carried out by these techniques in the present studies. With inputs in breeding aspects of the mutants developed by these techniques, like the one with zero erucic acid, could be bred for removing the background mutations by back crossing the same with the wild type parent. Further lipidomics studies could be done on these mutant lines, which will enable breeders to cross breed with a high yielding parent. The potential for such hybrids will be tremendous due to the high yielding zero erucic camelina lines in various industries. Camelina mutant seeds could be efficiently used of Biofuel production with improved oil quality, which will lead to less greenhouse gas emissions by which leading to a greener environment. More genes controlling the lipid biosynthesis pathway and other additional important pathways could be tilled to generate more number of mutants for studying the metabolomics in Camelina.

Camelina sativa mutant lines generated in the present study exhibits divergence in lipid levels and fatty acid profiles. Molecular breeders can carry out transcriptomic studies of mutants which could be grown in different geographical areas with different nutrient inputs like Nitrogen, Phosphorus etc. and even study the abiotic stress related studies on these mutants mainly focusing on drought tolerance and salt tolerance. These will lead to the development of camelina mutants which will be more stable in conditions like climate change. Camelina phenotypic database developed in the present study gives an array of morphological variations which could be utilised for various QTL studies. Genes controlling early flowering and larger seed size camelina mutants could be identified by positional cloning strategies. These traits are directly or indirectly connected to the future commercial exploitation of these mutants for greater benefits.

Chapter 5 – Summary

Whole genome study on EMS induced important mutants will help us to get an idea regarding the multiple gene activity of the three copies of different genes in the whole lipid biosynthesis pathway. Camelina, can be developed as easy crop for research, as earlier work has demonstrated easy way for creating transgenic camelina by floral-dip transformation method (Liu et al., 2012). Gene editing in *Camelina sativa* could be done by TALEN and CRISPR techniques, which will help molecular biologists to evaluate the efficiency of the mutation frequency and compare the same with the mutation frequency obtained from the Camelina mutant population created in the present study. Further as the transgenic marketability is a cause of concern in most of the country, the development of mutants from EMS induced mutant population is quite promising for future introduction of the mutants in the market. Various private companies can utilize the mutant population developed by us for creating different lines according to the need of various industries. There has been interest from private companies for the commercial utilization of Camelina with zero erucic acid mutant lines.

6. REFERENCES

(2003). ERUCIC ACID IN FOOD: A Toxicological Review and Risk Assessment TECHNICAL REPORT SERIES NO 21 4 para 1.

Abdullah, H.M., Akbari, P., Paulose, B., Schnell, D., Qi, W., Park, Y., Pareek, A., and Dhankher, O.P. (2016). Transcriptome profiling of Camelina sativa to identify genes involved in triacylglycerol biosynthesis and accumulation in the developing seeds. Biotechnology for Biofuels *9*, 136.

Abedi, E., and Sahari, M.A. (2014). Long-chain polyunsaturated fatty acid sources and evaluation of their nutritional and functional properties. Food Science & Nutrition *2*, 443-463.

Abraham T, A.P., Chatterjee M, Shroff J and Shroff V. (2009). DNA Extraction Apparatus (Gravity Assisted High throughput disruption of cellular material). BenchBio Ltd, Patent Application no: 2308/MUM/09.

Adamski, N.M., Anastasiou, E., Eriksson, S., O'Neill, C.M., and Lenhard, M. (2009). Local maternal control of seed size by KLUH/CYP78A5-dependent growth signaling. Proceedings of the National Academy of Sciences *106*, 20115-20120.

Agarwal, A., Pant, T., and Ahmed, Z. (2010). Camelina sativa: A new crop with bio-fuel potential introduced in India. Curr Science *99*, 1194-1195.

Akpan, U.G., Jimoh, A., and Mohammed, A.D. (2006). Extraction, characterization and modification of castor seed oil. Directory of open Journals.

Alagoz, Y., Gurkok, T., Zhang, B., and Unver, T. (2016). Manipulating the Biosynthesis of Bioactive Compound Alkaloids for Next-Generation Metabolic Engineering in Opium Poppy Using CRISPR-Cas 9 Genome Editing Technology. Sci Rep *6*, 30910.

Ali, N., Paul, S., Gayen, D., Sarkar, S., Datta, S., and Datta, K. (2013). RNAi mediated down regulation of myo-inositol-3-phosphate synthase to generate low phytate rice. Rice *6*, 1-12.

Aziza, A.E., Quezada, N., and Cherian, G. (2010a). Antioxidative effect of dietary Camelina meal in fresh, stored, or cooked broiler chicken meat. Poultry science *89*, 2711-2718.

Aziza, A.E., Quezada, N., and Cherian, G. (2010b). Feeding Camelinasativa meal to meat-type chickens: Effect on production performance and tissue fatty acid composition. The Journal of Applied Poultry Research *19*, 157-168.

References

Barth, C., and Jander, G. (2006). Arabidopsis myrosinases TGG1 and TGG2 have redundant function in glucosinolate breakdown and insect defense. The Plant Journal *46*, 549-562.

Basak, J., and Nithin, C. (2015). Targeting Non-Coding RNAs in Plants with the CRISPR-Cas Technology is a Challenge yet Worth Accepting. Frontiers in plant science *6*.

Belfield, E.J., Gan, X., Mithani, A., Brown, C., Jiang, C., Franklin, K., Alvey, E., Wibowo, A., Jung, M., Bailey, K., *et al.* (2012). Genome-wide analysis of mutations in mutant lineages selected following fast-neutron irradiation mutagenesis of Arabidopsis thaliana. Genome Research *22*, 1306-1315.

Belhaj, K., Chaparro-Garcia, A., Kamoun, S., and Nekrasov, V. (2013). Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. Plant Methods *9*, 39.

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

Bernardo, A., Howard-Hildige, R., O'Connell, A., Nichol, R., Ryan, J., Rice, B., Roche, E., and Leahy, J.J. (2003). Camelina oil as a fuel for diesel transport engines. Industrial Crops and Products *17*, 191-197.

Berti, M.T., Wilckens, R., Fischer, S., Solis, A., and Jonson, B. (2011). Seeding date influence on camelina seed yield, yield components and oil content in Chile S Ind Crops Prod *34*, 1358-1365.

Betancor, M.B., Sprague, M., Usher, S., Sayanova, O., Campbell, P.J., Napier, J.A., and Tocher, D.R. (2015). A nutritionally-enhanced oil from transgenic Camelina sativa effectively replaces fish oil as a source of eicosapentaenoic acid for fish. Scientific Reports *5*, 8104.

Bingham, P.M., Levis, R., and Rubin, G.M. (1981). Cloning of DNA sequences from the white locus of D. melanogaster by a novel and general method. Cell *25*, 693-704.

Bligh, E.G., and Dyer, W.J. (1959). A rapid method of total lipid extraxtion and purification. Canada Journal of Biochemistry and Physiology *37*.

Blomstedt, C.K., Gleadow, R.M., O'Donnell, N., Naur, P., Jensen, K., Laursen, T., Olsen, C.E., Stuart, P., Hamill, J.D., Moller, B.L., *et al.* (2012). A combined biochemical screen and TILLING approach identifies mutations in Sorghum bicolor L. Moench resulting in acyanogenic forage production. Plant Biotechnol J *10*, 54-66.

References

Bolon, Y.-T., Haun, W.J., Xu, W.W., Grant, D., Stacey, M.G., Nelson, R.T., Gerhardt, D.J., Jeddeloh, J.A., Stacey, G., Muehlbauer, G.J., *et al.* (2011). Phenotypic and Genomic Analyses of a Fast Neutron Mutant Population Resource in Soybean. Plant Physiology *156*, 240-253.

Boualem, A., Fleurier, S., Troadec, C., Audigier, P., Kumar, A.P.K., Chatterjee, M., Alsadon, A.A., Sadder, M.T., Wahb-Allah, M.A., Al-Doss, A.A., *et al.* (2014). Development of a Cucumis sativus TILLinG Platform for Forward and Reverse Genetics. PLoS ONE *9*, e97963.

Bovina, R., Talamè, V., Silvio, S., Sanguineti, M.C., Trost, P., Sparla, F., and Tuberosa, R. (2011). Starch metabolism mutants in barley: A TILLING approach. Plant Genetic Resources *9*, 170-173.

Brooks, C., Nekrasov, V., Lippman, Z.B., and Van Eck, J. (2014). Efficient Gene Editing in Tomato in the First Generation Using the Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-Associated9 System. Plant Physiology *166*, 1292-1297.

Budin, J.T., Breene, W.M., and Putnam, D.H. (1995). Some compositional properties of camelina (camelina sativa L. Crantz) seeds and oils. Journal of the American Oil Chemists' Society *72*, 309-315.

Caldwell, D.G., McCallum, N., Shaw, P., Muehlbauer, G.J., Marshall, D.F., and Waugh, R. (2004). A structured mutant population for forward and reverse genetics in Barley (Hordeum vulgare L.). The Plant journal : for cell and molecular biology *40*, 143-150.

Campbell, M.C., Rossi, A.F., and Erskine, W. (2013). Camelina (Camelina sativa (L.) Crantz): agronomic potential in Mediterranean environments and diversity for biofuel and food uses. Crop and Pasture Science *64*, 388-398.

Cassagne, C., Lessire, R., Bessoule, J.J., Moreau, P., Creach, A., Schneider, F., and Sturbois, B. (1994). Biosynthesis of very long chain fatty acids in higher plants. Prog Lipid Res *33*, 55-69.

Cermak, T., Baltes, N.J., Cegan, R., Zhang, Y., and Voytas, D.F. (2015). High-frequency, precise modification of the tomato genome. Genome Biol *16*, 232.

Chandrasekaran, J., Brumin, M., Wolf, D., Leibman, D., Klap, C., Pearlsman, M., Sherman, A., Arazi, T., and Gal□On, A. (2016). Development of broad virus resistance in non□transgenic cucumber using CRISPR/Cas9 technology. Molecular plant pathology.

References

Chantreau, M., Grec, S., Gutierrez, L., Dalmais, M., Pineau, C., Demailly, H., Paysant-Leroux, C., Tavernier, R., Trouvé, J.-P., Chatterjee, M., *et al.* (2013). PT-Flax (phenotyping and TILLinG of flax): development of a flax (Linum usitatissimumL.) mutant population and TILLinG platform for forward and reverse genetics. BMC plant biology *13*, 159.

Colbert, T., Till, B.J., Tompa, R., Reynolds, S., Steine, M.N., Yeung, A.T., McCallum, C.M., Comai, L., and Henikoff, S. (2001). High-throughput screening for induced point mutations. Plant Physiol *126*, 480-484.

Colbert, T.G., Hurst, S.R., and Slade, A.J. (2011). Tomatoes That Soften More Slowly Post-Harvest Due To Non-Transgenic Alterations In An Expansin Gene (Google Patents).

Colombini, S., Broderick, G.A., Galasso, I., Martinelli, T., Rapetti, L., Russo, R., and Reggiani, R. (2014). Evaluation of Camelina sativa (L.) Crantz meal as an alternative protein source in ruminant rations. Journal of the Science of Food and Agriculture *94*, 736-743.

Comai, L. (2005). The advantages and disadvantages of being polyploid. Nat Rev Genet *6*, 836-846.

Comai, L., Young, K., Till, B.J., Reynolds, S.H., Greene, E.A., Codomo, C.A., Enns, L.C., Johnson, J.E., Burtner, C., Odden, A.R., *et al.* (2004). Efficient discovery of DNA polymorphisms in natural populations by Ecotilling. The Plant Journal *37*, 778-786.

Cooper, H.D., Spillane, C., and Hodgkin, T. (2001). Broadening the genetic base of crop production (CABI).

Cooper, J.L., Henikoff, S., Comai, L., and Till, B.J. (2013). TILLING and ecotilling for rice. Methods in molecular biology (Clifton, NJ) *956*, 39-56.

Cooper, J.L., Till, B.J., Laport, R.G., Darlow, M.C., Kleffner, J.M., Jamai, A., El-Mellouki, T., Liu, S., Ritchie, R., Nielsen, N., *et al.* (2008). TILLING to detect induced mutations in soybean. BMC plant biology *8*, 9.

Dahmani-Mardas, F., Troadec, C., Boualem, A., Leveˆque, S., Alsadon, A.A., Aldoss, A.A., Dogimont, C., and Bendahmane, A. (2010). Engineering melon plants with improved fruit shelf life using the TILLING approach. PLoS One *5*, e15776.

Dalmais, M., Schmidt, J., Le Signor, C., Moussy, F., Burstin, J., Savois, V., Aubert, G., Brunaud, V., de Oliveira, Y., Guichard, C., *et al.* (2008). UTILLdb, a Pisum sativum in silico forward and reverse genetics tool. Genome Biol *9*, R43.

DeAngelis, K.M. (2007). phosphate buffer.

References

EFSA (2008). Glucosinolates as undesirable substances in animal feed Scientific Panel on Contaminants in the Food Chain The EFSA Journal.

Elahi, N., Duncan, R.W., and Stasolla, C. (2015). Decreased seed oil production in FUSCA3 Brassica napus mutant plants. Plant Physiol Biochem *96*, 222-230.

Elgowainy, A., Hanj, J., Wang, M., Carter, N., Stratton, R., Hileman, J., Malwitz, A., and Balasubramanian, S. (2012). Lifecycle analysis of Altrenative Aviation fuels in GREET

Elliott, R.H., Mann, L.W., and Olfert, O.O. (2007). Effects of seed size and seed weight on seedling establishment, seedling vigour and tolerance of summer turnip rape (Brassica rapa) to flea beetles, Phyllotreta spp. Canadian Journal of Plant Science *87*, 385-393.

Fan, D., Liu, T., Li, C., Jiao, B., Li, S., Hou, Y., and Luo, K. (2015). Efficient CRISPR/Cas9-mediated targeted mutagenesis in Populus in the first generation. Scientific reports *5*.

Fang, Y., and Tyler, B.M. (2015). Efficient disruption and replacement of an effector gene in the oomycete Phytophthora sojae using CRISPR/Cas9. Molecular Plant Pathology, 127-139.

Fatihi, A., Zbierzak, A.M., and Dörmann, P. (2013). Alterations in Seed Development Gene Expression Affect Size and Oil Content of Arabidopsis Seeds. Plant Physiology *163*, 973-985.

Fourmann, M., Barret, P., Renard, M., Pelletier, G., Delourme, R., and Brunel, D. (1998). The two genes homologous to Arabidopsis FAE1 co-segregate with the two loci governing erucic acid content in Brassica napus. Theoretical and Applied Genetics *96*, 852-858.

Frame, D.D., Palmer, M., Ward, R., and Martini, S. (2008). Feeding Camelina sativa and enhancing omega-3 fatty acid levels in market-age turkey hens. .

Francis, A., and Warwick, S.I. (2009). The biology of Canadian weeds. 142. Camelina alyssum (Mill.) Thell.; C. microcarpa Andrz. ex DC.; C. sativa (L.) Crantz. Canadian Journal of Plant Science *89*, 791-810.

Galasso, I., Manca, A., Braglia, L., Martinelli, T., Morello, L., and Breviario, D. (2010). h-TBP: an approach based on intro-length polymorphism for the rapid isolation and characterization of the multiple members of the β-tubulin gene family in Camelina sativa (L.) Crantz. . Molecular Breeding *28*, 635-645.

Gehringer, A., Friedt, W., Luhs, W., and Snowdon, R.J. (2006). Genetic mapping of agronomic traits in false flax (Camelina sativa subsp. sativa). Genome *49*, 1555-1563.

References

Gilchrist, E.J., Sidebottom, C.H.D., Koh, C.S., MacInnes, T., Sharpe, A.G., and Haughn, G.W. (2013). A Mutant <italic>Brassica napus</italic> (Canola) Population for the Identification of New Genetic Diversity via TILLING and Next Generation Sequencing. PLoS ONE *8*, e84303.

Godfray, H.C., and Garnett, T. (2014). Food security and sustainable intensification. Philos Trans R Soc Lond B Biol Sci *369*, 20120273.

Gonzalez-Thuillier, I., Napier, J., and Sayanova, O. (2016). Engineering synthetic pathways for the production of pharmaceutical sciadonic acid in transgenic oilseed Camelina sativa. New Biotechnology *33, Supplement*, S46.

González, M., Xu, M., Esteras, C., Roig, C., Monforte, A.J., Troadec, C., Pujol, M., Nuez, F., Bendahmane, A., Garcia-Mas, J., *et al.* (2011). Towards a TILLING platform for functional genomics in Piel de Sapo melons. BMC Res Notes *4*.

Gottwald, S., Bauer, P., Komatsuda, T., Lundqvist, U., and Stein, N. (2009). TILLING in the two-rowed barley cultivar "Barke" reveals preferred sites of functional diversity in the gene HvHox1. BMC Res Notes *2*.

Greene, E.A., Codomo, C.A., Taylor, N.E., Henikoff, J.G., Till, B.J., Reynolds, S.H., Enns, L.C., Burtner, C., Johnson, J.E., Odden, A.R., *et al.* (2003). Spectrum of Chemically Induced Mutations From a Large-Scale Reverse-Genetic Screen in Arabidopsis. Genet *164*.

Gross, B.L., and Olsen, K.M. (2010). Genetic perspectives on crop domestication. Trends in Plant Science *15*, 529-537.

Grossniklaus, U., Spillane, C., Page, D.R., and Köhler, C. (2001). Genomic imprinting and seed development: endosperm formation with and without sex. Current Opinion in Plant Biology *4*, 21-27.

Gugel, R.K., and Falk, K.C. (2006). Agronomic and seed quality evaluation of camelina sativa in western Canada. Can J Pl Sci *86*, 1047-1058.

Guo, Y., Abernathy, B., Zeng, Y., and Ozias-Akins, P. (2015). TILLING by sequencing to identify induced mutations in stress resistance genes of peanut (Arachis hypogaea). BMC Genomics *16*, 157.

Haliem, A.E., Abdullah, H., and AL-Huqail, A.A. (2013). Oxidative Damage and Mutagenic Potency of Fast Neutron and UV-B Radiation in Pollen Mother Cells and Seed Yield of Vicia faba L. BioMed Research International *2013*, 12.

References

Hammer, K., and Teklu, Y. (2008). Plant genetic resources: Selected issues from genetic erosion to genetic engineering. J Agric Rural Dev Trop Subtrop *109*, 15-50.

Harloff, H., Lemcke, S., Mittasch, J., Frolov, A., Wu, J., and Dreyer, F. (2012). A mutation screeing platform for rapeseed (Brassica napus L.) and the detection of sinapine biosynthesis mutants. TAG Theoretical and applied genetics Theoretische und angewandte Genetik *124*.

Harwood, J.L. (1996). Recent advancesin the biosynthesis of plant fatty acids. Biochim Biophys Acta *1301*.

Haslam, R.P., Sayanova, O., Kim, H.J., Cahoon, E.B., and Napier, J.A. (2016). Synthetic redesign of plant lipid metabolism. The Plant Journal *87*, 76-86.

Henikoff, S., and Comai, L. (2003). Single nucleotide mutations for plant functional genomics. Annual Review of Plant Biology *54*, 375-401.

Henry, I.M., Nagalakshmi, U., Lieberman, M.C., Ngo, K.J., Krasileva, K.V., Vasquez-Gross, H., Akhunova, A., Akhunov, E., Dubcovsky, J., and Tai, T.H. (2014). Efficient genome-wide detection and cataloging of EMS-induced mutationsusing exome capture and Next-Generation Sequencing. Plant Cell *26*, 1382-1397.

Herridge, R.P., Day, R.C., Baldwin, S., and Macknight, R.C. (2011). Rapid analysis of seed size in Arabidopsis for mutant and QTL discovery. Plant Methods *7*, 3.

Himelblau, E., Gilchrist, E.J., Buono, K., Bizzell, C., Mentzer, L., Vogelzang, R., Osborn, T., Amasino, R.M., Parkin, I.A., and Haughn, G.W. (2009). Forward and reverse genetics of rapid-cycling Brassica oleracea. TAG Theoretical and applied genetics Theoretische und angewandte Genetik *118*, 953-961.

Horn, P.J., Silva, J.E., Anderson, D., Fuchs, J., Borisjuk, L., Nazarenus, T.J., Shulaev, V., Cahoon, E.B., and Chapman, K.D. (2013). Imaging heterogeneity of membrane and storage lipids in transgenic Camelina sativa seeds with altered fatty acid profiles. Plant Journal *76*, 138-150.

Houle, D., Govindaraju, D.R., and Omholt, S. (2010). Phenomics: the next challenge. Nature Reviews Genetics *11*, 855-866.

http://www.mun.ca/biology/scarr/F16-17.jpg.

http://www.un.org/apps/news/story.asp?NewsID=45165 (2013).

http://www.un.org/en/development/desa/news/population/2015-report.html.

References

Hurst, S.R., Loeffler, D.L., Steine, M.N., Amen, A., and Vafeados, D. (2015). Non-transgenic tomato varieties having increased shelf life post-harvest due to alterations in β-galactosidase 4 (Google Patents).

Hutcheon, C., Ditt, R.F., Beilstein, M., Comai, L., Schroeder, J., Goldstein, E., Shewmaker, C.K., Nguyen, T., De Rocher, J., and Kiser, J. (2010). Polyploid genome of Camelina sativa revealed by isolation of fatty acid synthesis genes. BMC plant biology *10*, 233.

Hutcheon, C., Ditt, R.F., Beilstein, M., Comai, L., Shewmaker, C.K., Nguyen, T., Schroeder, J., Goldstein, E., Guillen-Portal, F., Van, A.D., *et al.* (2009). TILLING for Altered Fatty Acid Profiles in Camelina sativa. Paper presented at: Plant Biology 2009 (Honolulu, Hawaii).

Ishii, T., and Araki, M. (2016). Consumer acceptance of food crops developed by genome editing. Plant Cell Reports *35*, 1507-1518.

Jansen, J.G., Vrieling, H., van Teijlingen, C. M., Mohn, G. R., Tated A. D., and van Zeeland A. A. (1995). Marked differences in the role of O6-alkylguanine in hprt mutagenesis in T-lymphocytes of rats exposed in vivo to ethylmethanesulfonate, N-(2-hydroxyethyl)-N-nitrosourea, or N-ethyl-N-nitrosourea. Cancer Research *55*, 1875-1882.

Jensen, W.B. (2007). The Origin of the Soxhlet Extractor. Journal of Chemical Education *84*, 1913.

Jewett, F.G. (2015). Camelina sativa:For Biofuels and Biproducts Industrial Crops: Breeding for Bioenergy and Bioproducts Springer *Chapter 8*.

Jiang, W.Z., Henry, I.M., Lynagh, P.G., Comai, L., Cahoon, E.B., and Weeks, D.P. (2016). Significant enhancement of fatty acid composition in seeds of the allohexaploid, Camelina sativa, using CRISPR/Cas9 gene editing. Plant Biotechnology Journal, n/a-n/a.

Jones, M.O., Piron-Prunier, F., Marcel, F., Piednoir-Barbeau, E., Alsadon, A.A., Wahb-Allah, M.A., Al-Doss, A.A., Bowler, C., Bramley, P.M., Fraser, P.D., *et al.* (2012). Characterisation of alleles of tomato light signalling genes generated by TILLING. Phytochemistry *79*, 78-86.

Jordan, C.F. (2002). Genetic Engineering, the Farm Crisis, and World Hunger. BioScience *52*, 523-529.

Julie-Galau, S., Bellec, Y., Faure, J.D., and Tepfer, M. (2014). Evaluation of the potential for interspecific hybridization between Camelina sativa and related wild Brassicaceae in anticipation of field trials of GM camelina. Transgenic Res *23*, 67-74.

References

Kagale, S., Koh, C., Nixon, J., Bollina, V., Clarke, W.E., Tuteja, R., Spillane, C., Robinson, S.J., Links, M.G., Clarke, C., *et al.* (2014). The emerging biofuel crop Camelina sativa retains a highly undifferentiated hexaploid genome structure. Nat Commun *5*.

Kagale, S., Nixon, J., Khedikar, Y., Pasha, A., Provart, N.J., Clarke, W.E., Bollina, V., Robinson, S.J., Coutu, C., Hegedus, D.D., *et al.* (2016). The developmental transcriptome atlas of the biofuel crop Camelina sativa. The Plant journal : for cell and molecular biology *88*, 879-894.

Kalaitzandonakes, N., Alston, J.M., and Bradford, K.J. (2007). Compliance costs for regulatory approval of new biotech crops. Nat Biotechnol *25*, 509-511.

Kang, J., Snapp, A.R., and Lu, C. (2011). Identification of three genes encoding microsomal oleate desaturases (FAD2) from the oilseed crop Camelina sativa. Plant Physiol Biochem *49*, 223-229.

Keske, C.M.H., Hoag, D.L., Brandess, A., and Johnson, J.J. (2013). Is it economically feasible for farmers to grow their own fuel? A study of Camelina sativa produced in the western United States as an on-farm biofuel. biomass and bioenergy *54*, 89-99.

Knoll, J., Ramos, M., Zeng, Y., Holbrook, C., Chow, M., and Chen, S. (2011). TILLING for allergen reduction and improvement of quality traits in peanut (Arachis hypogaea L.). BMC plant biology *11*.

Knorzer, K.H. (1978). Evolution and spread of Gold of Pleasure (Camelina sativa S.L.). . Bererichte der Deutschen Botanischen Gesellschaft *91*, 187-195.

Kole, C., Muthamilarasan, M., Henry, R., Edwards, D., Sharma, R., Abberton, M., Batley, J., Bentley, A., Blakeney, M., Bryant, J., *et al.* (2015). Application of genomics-assisted breeding for generation of climate resilient crops: progress and prospects. Frontiers in plant science *6*, 563.

Konzak, C.F., Nilan, R.A., and Kleinhofs, A. (1976). Artificial mutagenesis as an aid in overcoming genetic vulnerability of crop plants. Basic Life Sci *8*, 163-177.

Kovach, M.J., and McCouch, S.R. (2008). Leveraging natural diversity: back through the bottleneck. Curr Opin Plant Biol *11*, 193-200.

Kumar, A.P., Boualem, A., Bhattacharya, A., Parikh, S., Desai, N., Zambelli, A., Leon, A., Chatterjee, M., and Bendahmane, A. (2013). SMART--Sunflower Mutant population And Reverse genetic Tool for crop improvement. BMC plant biology *13*, 38.

References

Kumar, A.P.K., McKeown, P.C., Boualem, A., Ryder, P., Brychkova, G., Bendahmane, A., Sarkar, A., Chatterjee, M., and Spillane, C. (2017). TILLING by Sequencing (TbyS) for targeted genome mutagenesis in crops. Molecular Breeding *37*, 14.

Kumar, S., Ghatty, S., Satyanarayana, J., Guha, A., Chaitanya, B., and Reddy, A.R. (2012). Paclobutrazol treatment as a potential strategy for higher seed and oil yield in field-grown camelina sativa L. Crantz. . BMC Research Notes *5*, 137.

Kunst, L., Taylor, D.C., and Underhill, E.W. (1992). Fatty acid elongation in developing seeds of Arabidopsis thaliana. Plant Physiology and Biochemistry *30*, 425-434.

Kurowska, M., Daszkowska-Golec, A., Gruszka, D., Marzec, M., Szurman, M., Szarejko, I., and Maluszynski, M. (2011). TILLING: a shortcut in functional genomics. Journal of applied genetics *52*, 371-390.

Lai, K.S., Kaothien-Nakayama, P., Iwano, M., and Takayama, S. (2012). A TILLING resource for functional genomics in *Arabidopsis thaliana* accession C24. Genes & genetic systems *87*, 291-297.

Laskar, R.A., Chaudhary, C., Khan, S., and Chandra, A. (2016). Induction of mutagenized tomato populations for investigation on agronomic traits and mutant phenotyping. Journal of the Saudi Society of Agricultural Sciences.

Le Signor, C., Savois, V., Aubert, G., Verdier, J., Nicolas, M., Pagny, G., Moussy, F., Sanchez, M., Baker, D., Clarke, J., *et al.* (2009). Optimizing TILLING populations for reverse genetics in Medicago truncatula. Plant Biotechnol J *7*.

Lebedevas, S., Lebedeva, G., Makarevičienė, V., Kazanceva, I., and Kazancev, K. (2010). ANALYSIS OF THE ECOLOGICAL PARAMETERS OF THE DIESEL ENGINE POWERED WITH BIODIESEL FUEL CONTAINING METHYL ESTERS FROM CAMELINA SATIVA OIL. Transport (16484142) *25*, 22-28.

Lemieux, B., Miquel, M., Somerville, C., and Browse, J. (1990). Mutants of Arabidopsis with alterations in seed lipid fatty acid composition. TAG Theoretical and applied genetics Theoretische und angewandte Genetik *80*, 234-240.

Li, J., Meng, X., Zong, Y., Chen, K., Zhang, H., Liu, J., Li, J., and Gao, C. (2016). Gene replacements and insertions in rice by intron targeting using CRISPR-Cas9. Nat Plants *2*, 16139.

References

Li, X., and Yang, Y. (2014). A novel perspective on seed yield of broad bean (Vicia faba L.): differences resulting from pod characteristics. Scientific Reports *4*, 6859.

Li, Y., Zheng, L., Corke, F., Smith, C., and Bevan, M.W. (2008). Control of final seed and organ size by the DA1 gene family in Arabidopsis thaliana. Genes & development *22*, 1331-1336.

Liu, X., Brost, J., Hutcheon, C., Guilfoil, R., Wilson, A.K., Leung, S., Shewmaker, C.K., Rooke, S., Nguyen, T., Kiser, J., *et al.* (2012). Transformation of the oilseed crop Camelina sativa by Agrobacterium-mediated floral dip and simple large-scale screening of transformants. In Vitro Cellular & Developmental Biology - Plant *48*, 462-468.

Lowder, L.G., Zhang, D., Baltes, N.J., Paul, J.W., 3rd, Tang, X., Zheng, X., Voytas, D.F., Hsieh, T.F., Zhang, Y., and Qi, Y. (2015). A CRISPR/Cas9 Toolbox for Multiplexed Plant Genome Editing and Transcriptional Regulation. Plant Physiol *169*, 971-985.

Lu, C., Xin, Z., Ren, Z., Miquel, M., and Browse, J. (2009). An enzyme regulating triacylglycerol composition is encoded by the ROD1 gene of Arabidopsis. Proc Natl Acad Sci U S A *106*, 18837-18842.

Lucht, J.M. (2015). Public Acceptance of Plant Biotechnology and GM Crops. Viruses *7*, 4254-4281.

Mao, Y., Zhang, Z., Feng, Z., Wei, P., Zhang, H., Botella, J.R., and Zhu, J.K. (2016). Development of germ□line□specific CRISPR□Cas9 systems to improve the production of heritable gene modifications in Arabidopsis. Plant Biotechnology Journal *14*, 519-532.

Martín, B., Ramiro, M., Martínez-Zapater, J.M., and Alonso-Blanco, C. (2009). A high-density collection of EMS-induced mutations for TILLING in Landsberg erecta genetic background of Arabidopsis. BMC plant biology *9*.

Matthaus, B., and Zubr, J. (2000). Variability of specific components in Camelina sativa oil seed cakes. Industrial crops and products *12*.

McCallum, C., Comai, L., Greene, E., and Henikoff, S. (2000a). Targeted screening for induced mutations. Nat Biotechnol *18*.

McCallum, C., Slade, A.J., Colbert, T.G., Knauf, V.C., and Hurst, S. (2011). Tomatoes having reduced polygalacturonase activity caused by non-transgenic mutations in the polygalacturonase gene (Google Patents).

References

McCallum, C.M., Comai, L., Greene, E.A., and Henikoff, S. (2000b). Targeting induced local lesions IN genomes (TILLING) for plant functional genomics. Plant Physiol *123*, 439-442.

McKeown, P.C., Fort, A., Duszynska, D., Sulpice, R., and Spillane, C. (2013a). Emerging molecular mechanisms for biotechnological harnessing of heterosis in crops. Trends in Biotechnology *31*, 549-551.

McKeown, P.C., Keshavaiah, C., Fort, A., Tuteja, R., Chatterjee, M., Varshney, R.K., and Spillane, C. (2013b). Genomics in Agriculture and Food Processing. In Biotechnology in Agriculture and Food Processing: Opportunities and Challenges, P.S. Panesar, and S.S. Marwaha, eds. (CRC Press: Taylor & Francis).

Menda, N., Semel, Y., Peled, D., Eshed, Y., and Zamir, D. (2004). In silico screening of a saturated mutation library of tomato. The Plant journal : for cell and molecular biology *38*, 861-872.

Meyer, R.S., and Purugganan, M.D. (2013). Evolution of crop species: genetics of domestication and diversification. Nature Reviews Genetics *14*, 840-852.

Minoia, S., Boualem, A., Marcel, F., Troadec, C., Quemener, B., Cellini, F., Petrozza, A., Vigouroux, J., Lahaye, M., Carriero, F., *et al.* (2016). Induced mutations in tomato SlExp1 alter cell wall metabolism and delay fruit softening. Plant science : an international journal of experimental plant biology *242*, 195-202.

Missirian, V., Comai, L., and Filkov, V. (2011). Statistical mutation calling from sequenced overlapping DNA pools in TILLING experiments. BMC bioinformatics *12*, 287.

Mitsuhashi, N., Kondo, M., Nakaune, S., Ohnishi, M., Hayashi, M., Hara-Nishimura, I., Richardson, A., Fukaki, H., Nishimura, M., and Mimura, T. (2008). Localization of myo-inositol-1-phosphate synthase to the endosperm in developing seeds of Arabidopsis. J Exp Bot *59*, 3069-3076.

Morris, S.H., and Spillane, C. (2008). GM directive deficiencies in the European Union. The current framework for regulating GM crops in the EU weakens the precautionary principle as a policy tool. EMBO reports *9*, 500-504.

Moser, B.R. (2010). Camelina (Camelina sativa L.) oil as a biofuels feedstock: Golden opportunity or false hope? lipid technology.

Mulligan, G.A. (2002). Weedy introduced mustards (Brassicaceae) of Canada. Can Field Nat *116*.

References

Murphy, D. (2007). Plant Breeding and Biotechnology (Cambridge, UK: Cambridge University Press).

Napier, J.A., Haslam, R.P., Beaudoin, F., and Cahoon, E.B. (2014). Understanding and manipulating plant lipid composition: Metabolic engineering leads the way. Current Opinion in Plant Biology *19*, 68-75.

Nguyen, H.T., Silva, J.E., Podicheti, R., Macrander, J., Yang, W., Nazarenus, T.J., Nam, J.-W., Jaworski, J.G., Lu, C., Scheffler, B.E., *et al.* (2013). Camelina seed transcriptome: a tool for meal and oil improvement and translational research. Plant Biotechnology Journal *11*, 759-769.

Nguyen, T., Liu, X., and DeRocher, J. (2014). Floral dip method for transformation of camelina (Google Patents).

Nieto, C., Piron, F., Dalmais, M., Marco, C.F., Moriones, E., Gomez-Guillamon, M.L., Truniger, V., Gomez, P., Garcia-Mas, J., Aranda, M.A., *et al.* (2007). EcoTILLING for the identification of allelic variants of melon eIF4E, a factor that controls virus susceptibility. BMC plant biology *7*, 34.

Okabe, Y., Ariizumi, T., and Ezura, H. (2013). Updating the Micro-Tom TILLING platform. Breeding Science *63*, 42-48.

Okabe, Y., Asamizu, E., Saito, T., Matsukura, C., Ariizumi, T., Bres, C., Rothan, C., Mizoguchi, T., and Ezura, H. (2011). Tomato TILLING technology: development of a reverse genetics tool for the efficient isolation of mutants from Micro-Tom mutant libraries. Plant Cell Physiol *52*, 1994-2005.

Okuley, J., Lightner, J., Feldmann, K., Yadav, N., Lark, E., and Browse, J. (1994). Arabidopsis FAD2 Gene Encodes the Enzyme That Is Essential for Polyunsaturated Lipid Synthesis. The Plant Cell Online *6*, 147-158.

Onda, Y., and Mochida, K. (2016). Exploring Genetic Diversity in Plants Using High-Throughput Sequencing Techniques. Curr Genomics *17*, 358-367.

Orozco-Arroyo, G., Paolo, D., Ezquer, I., and Colombo, L. (2015). Networks controlling seed size in Arabidopsis. Plant Reproduction *28*, 17-32.

Osakabe, Y., Watanabe, T., Sugano, S.S., Ueta, R., Ishihara, R., Shinozaki, K., and Osakabe, K. (2016). Optimization of CRISPR/Cas9 genome editing to modify abiotic stress responses in plants. Scientific Reports *6*, 26685.

References

Pahlavani, M.H., Miri, A.A., and G., K. (2008). Response of oil and protein content to seed size in cotton. International Journal of agriculture and Biology *10*, 643-647.

Parent, B., Shahinnia, F., Maphosa, L., Berger, B., Rabie, H., Chalmers, K., Kovalchuk, A., Langridge, P., and Fleury, D. (2015). Combining field performance with controlled environment plant imaging to identify the genetic control of growth and transpiration underlying yield response to water-deficit stress in wheat. Journal of Experimental Botany.

Parry, M.A., Madgwick, P.J., Bayon, C., Tearall, K., Hernandez-Lopez, A., Baudo, M., Rakszegi, M., Hamada, W., Al-Yassin, A., Ouabbou, H., *et al.* (2009). Mutation discovery for crop improvement. J Exp Bot *60*, 2817-2825.

Pauwels, K., De Keersmaecker, S.C.J., De Schrijver, A., du Jardin, P., Roosens, N.H.C., and Herman, P. (2015). Next-generation sequencing as a tool for the molecular characterisation and risk assessment of genetically modified plants: Added value or not? Trends in Food Science & Technology *45*, 319-326.

Peng, Q., Hu, Y., Wei, R., Zhang, Y., Guan, C., Ruan, Y., and Liu, C. (2010). Simultaneous silencing of FAD2 and FAE1 genes affects both oleic acid and erucic acid contents in Brassica napus seeds. Plant Cell Reports *29*, 317-325.

Pérez-de-Castro, A.M., Vilanova, S., Cañizares, J., Pascual, L., Blanca, J.M., Díez, M.J., Prohens, J., and Picó, B. (2012). Application of Genomic Tools in Plant Breeding. Current Genomics *13*, 179-195.

Perry, J., Brachmann, A., Welham, T., Binder, A., Charpentier, M., and Groth, M. (2009). TILLING in Lotus japonicus identified large allelic series for symbiosis genes and revealed a bias in functionally defective ethyl methanesulfonate alleles toward glycine replacements. Plant Physiol *151*.

Piron, F., Nicolaï, M., Minoïa, S., Piednoir, E., Moretti, A., Salgues, A., Zamir, D., Caranta, C., and Bendahmane, A. (2010). An Induced Mutation in Tomato eIF4E Leads to Immunity to Two Potyviruses. PLOS ONE *5*, e11313.

Porceddu, A., Panara, F., Calderini, O., Molinari, L., Taviani, P., Lanfaloni, L., Scotti, C., Carelli, M., Scaramelli, L., Bruschi, G., *et al.* (2008). An Italian functional genomic resource for Medicago truncatula. BMC Res Notes *1*, 129.

Putnam, D.H., Budin, J.T., Field, L.A., and Breene, W.M. (1993). Camelina: A promising low-input oilseed. In: J Janick and JE Simon (eds), New crops Wiley, New York, 314-322.

References

Quetier, F. (2016). The CRISPR-Cas9 technology: Closer to the ultimate toolkit for targeted genome editing. Plant science : an international journal of experimental plant biology *242*, 65-76.

Raboy, V. (2007). The ABCs of low-phytate crops. Nat Biotech *25*, 874-875.

Rajendran, K., Tester, M., and Roy, S.J. (2009). Quantifying the three main components of salinity tolerance in cereals. Plant, cell & environment *32*, 237-249.

Ratledge, C. (2004). Fatty acid biosynthesis in microorganisms being used for Single Cell Oil production. Biochimie *86*, 807-815.

Rawat, N., Sehgal, S.K., Joshi, A., Rothe, N., Wilson, D.L., McGraw, N., Vadlani, P.V., Li, W., and Gill, B.S. (2012). A diploid wheat TILLING resource for wheat functional genomics. BMC plant biology *12*.

Reddy, T.V., Dwivedi, S., and Sharma, N.K. (2012). Development of TILLING by sequencing platform towards enhanced leaf yield in tobacco. Industrial Crops and Products *40*, 324-335.

Reisig, D.D., and Reay-Jones, F.P.F. (2015). Inhibition of Helicoverpa zea (Lepidoptera: Noctuidae) Growth by Transgenic Corn Expressing Bt Toxins and Development of Resistance to Cry1Ab. Environmental Entomology *44*, 1275.

Rice, B., Fröhlich, A., and R, L. (1998). BIO-DIESEL PRODUCTION FROM CAMELINA OIL, WASTE COOKING OIL AND TALLOW.

Ricroch, A.E., Ammann, K., and Kuntz, M. (2016). Editing EU legislation to fit plant genome editing: The use of genome editing technologies in plant breeding requires a novel regulatory approach for new plant varieties that involves farmers. EMBO reports.

Ruddle, P., Whetten, R., Cardinal, A., Upchurch, R.G., and Miranda, L. (2013). Effect of a novel mutation in a Delta9-stearoyl-ACP-desaturase on soybean seed oil composition. TAG Theoretical and applied genetics Theoretische und angewandte Genetik *126*, 241-249.

Russo, R., and Reggiani, R. (2012). Antinutritive Compounds in Twelve <i>Camelina sativa </i>Genotypes. American Journal of Plant Sciences *Vol.03No.10*, 4.

Sabetta, W., Alba, V., and Blanco, A. (2011). Montemurro C: sunTILL: a TILLING resource for gene function analysis in sunflower. Plant Methods *7*.

Schnable, J.C. (2015). Genome Evolution in Maize: From Genomes Back to Genes. Annual Review of Plant Biology *66*, 329-343.

References

Schruff, M.C., Spielman, M., Tiwari, S., Adams, S., Fenby, N., and Scott, R.J. (2006). The AUXIN RESPONSE FACTOR 2 gene of Arabidopsis links auxin signaling, cell division, and the size of seeds and other organs. Dev *133*.

Scully, E.D., Gries, T., Funnell-Harris, D.L., Xin, Z., Kovacs, F.A., Vermerris, W., and Sattler, S.E. (2015). Characterization of novel Brown midrib 6 mutations affecting lignin biosynthesis in sorghum. J Integr Plant Biol.

Sestili, F., Botticella, E., Bedo, Z., Phillips, A., and Lafiandra, D. (2010). Production of novel allelic variation for genes involved in starch biosynthesis through mutagenesis. Molecular Breeding *25*, 145-154.

Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., Zhang, K., Liu, J., Xi, J.J., and Qiu, J.-L. (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. Nature Biotechnology *31*, 686-688.

Shen, L., Wang, C., Fu, Y., Wang, J., Liu, Q., Zhang, X., Yan, C., Qian, Q., and Wang, K. (2016). QTL editing confers opposing yield performance in different rice varieties. J Integr Plant Biol.

Shi, J., Gao, H., Wang, H., Lafitte, H.R., Archibald, R.L., Yang, M., Hakimi, S.M., Mo, H., and Habben, J.E. (2016). ARGOS8 variants generated by CRISPR-Cas9 improve maize grain yield under field drought stress conditions. Plant Biotechnol J.

Shi, J., Lang, C., Wu, X., Liu, R., Zheng, T., Zhang, D., Chen, J., and Wu, G. (2015). RNAi knockdown of fatty acid elongase1 alters fatty acid composition in Brassica napus. Biochem Biophys Res Commun *466*, 518-522.

Shi, J., Li, R., Qiu, D., Jiang, C., Long, Y., Morgan, C., Bancroft, I., Zhao, J., and Meng, J. (2009). Unraveling the Complex Trait of Crop Yield With Quantitative Trait Loci Mapping in Brassica napus. Genetics *182*, 851-861.

Shonnard, D.R., Williams, L., and Kalnes, T.N. (2010). Camelina-derived jet fuel and diesel: Sustainable advanced biofuels. Environmental Progress & Sustainable Energy *29*, 382-392.

Singh, B.P. (2013). Biofuel Crops: Production, Physiology and Genetics (CABI).

Sippl, M.J., and Wiederstein, M. (2008). A note on difficult structure alignment problems. Bioinformatics *24*, 426-427.

Sippl, M.J., and Wiederstein, M. (2012). Detection of spatial correlations in protein structures and molecular complexes. Structure (London, England : 1993) *20*, 718-728.

References

Skopek, T.R., Walker, V.E., Cochrane J.E., Cariello, N.F. (1992). Mutational spectrum at the Hprt locus in splenic T cells of B6C3F1 mice exposed to N-ethyl-N-nitrosourea. Proc Natl Acad Sci U S A *89*, 7866-7870.

Slade, A.J., Fuerstenberg, S.I., Loeffler, D., Steine, M.N., and Facciotti, D. (2005). A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. Nat Biotechnol *23*.

Slade, A.J., and Knauf, V.C. (2005). TILLING moves beyond functional genomics into crop improvement. Transgenic Research *14*, 109-115.

Slade, A.J., McGuire, C., Loeffler, D., Mullenberg, J., Skinner, W., Fazio, G., Holm, A., Brandt, K.M., Steine, M.N., Goodstal, J.F., *et al.* (2012). Development of high amylose wheat through TILLING. BMC plant biology *12*, 69.

Sprink, T., Eriksson, D., Schiemann, J., and Hartung, F. (2016). Regulatory hurdles for genome editing: process- vs. product-based approaches in different regulatory contexts. Plant Cell Rep *35*, 1493-1506.

Stemple, D.L. (2004). TILLING [mdash] a high-throughput harvest for functional genomics. Nat Rev Genet *5*, 145-150.

Stephenson, P., Baker, D., Girin, T., Perez, A., Amoah, S., King, G.J., and Ostergaard, L. (2010). A rich TILLING resource for studying gene function in Brassica rapa. BMC plant biology *10*, 62.

Sulpice, R., and McKeown, P.C. (2015). Moving Towards a Comprehensive Map of Central Plant Metabolism. Annual Review of Plant Biology *66*.

Svitashev, S., Young, J.K., Schwartz, C., Gao, H., Falco, S.C., and Cigan, A.M. (2015). Targeted Mutagenesis, Precise Gene Editing, and Site-Specific Gene Insertion in Maize Using Cas9 and Guide RNA. Plant Physiol *169*, 931-945.

Tadele, Z., Chikelu, M.B.A., and Till, B.J. (2010). TILLING for Mutations in Model Plants and Crops. In Molecular Techniques in Crop Improvement, pp. 307-332.

Talame, V., Bovina, R., Sanguineti, M.C., Tuberosa, R., Lundqvist, U., and Salvi, S. (2008). TILLMore, a resource for the discovery of chemically induced mutants in barley. Plant Biotechnol J *6*, 477-485.

Tang, H., Sezen, U., and Paterson, A.H. (2010). Domestication and plant genomes. Curr Opin Plant Biol *13*, 160-166.

References

Taylor, D.C., Falk, K.C., Palmer, C.D., Hammerlindl, J., Babic, V., Mietkiewska, E., Jadhav, A., Marillia, E.F., Francis, T., Hoffman, T., *et al.* (2010). Brassica carinata-a new molecular farming platform for delivering bio-industrial oil feedstocks: Case studies of genetic modifications to improve very long-chain fatty acid and oil content in seeds. Biofuels, Bioproducts and Biorefining *4*.

Telem, R.S., Wani, S.H., Singh, N.B., Nandini, R., Sadhukhan, R., Bhattacharya, S., and Mandal, N. (2013). Cisgenics - a sustainable approach for crop improvement. Curr Genomics *14*, 468-476.

Tian, B., Wei, F., Shu, H., Zhang, Q., Zang, X., and Lian, Y. (2011). Decreasing erucic acid level by RNAi-mediated silencing of fatty acid elongase 1 (BnFAE1.1) in rapeseeds (Brassica napus L.) African Journal of Biotechnology *10*, 13194-13201.

Till, B., Reynolds, S., Weil, C., Springer, N., Burtner, C., and Young, K. (2004). Discovery of induced point mutations in maize genes by TILLING. BMC plant biology *4*.

Till, B.J., Cooper, J., Tai, T.H., Colowit, P., Greene, E.A., Henikoff, S., and Comai, L. (2007). Discovery of chemically induced mutations in rice by TILLING. BMC plant biology *7*.

Till, B.J., Reynolds, S.H., Greene, E.A., Codomo, C.A., Enns, L.C., Johnson, J.E., Burtner, C., Odden, A.R., Young, K., Taylor, N.E., *et al.* (2003). Large-Scale Discovery of Induced Point Mutations With High-Throughput TILLING. Genome Res *13*.

Till, B.J., Zerr, T., Comai, L., and Henikoff, S. (2006). A protocol for TILLING and Ecotilling in plants and animals. Nature Protocols *1*, 2465-2477.

Triques, K., Piednoir, E., Dalmais, M., Schmidt, J., Le Signor, C., Sharkey, M., Caboche, M., Sturbois, B., and Bendahmane, A. (2008). Mutation detection using ENDO1: application to disease diagnostics in humans and TILLING and Eco-TILLING in plants. BMC Mol Biol *9*, 42.

Triques, K., Sturbois, B., Gallais, S., Dalmais, M., Chauvin, S., Clepet, C., Aubourg, S., Rameau, C., Caboche, M., and Bendahmane, A. (2007). Characterization of Arabidopsis thaliana mismatch specific endonucleases: application to mutation discovery by TILLING in pea. Plant J Cell Mol Biol *51*.

Tsai, H., Howell, T., Nitcher, R., Missirian, V., Watson, B., Ngo, K.J., Lieberman, M., Fass, J., Uauy, C., Tran, R.K., *et al.* (2011). Discovery of Rare Mutations in Populations: TILLING by Sequencing. Plant Physiol *156*.

References

Tsai, H., Missirian, V., Ngo, K.J., Tran, R.K., Chan, S.R., Sundaresan, V., and Comai, L. (2013). Production of a high-efficiency TILLING population through polyploidization. Plant Physiol *161*, 1604-1614.

Uauy, C., Paraiso, F., Colasuonno, P., Tran, R.K., Tsai, H., Berardi, S., Comai, L., and Dubcovsky, J. (2009). A modified TILLING approach to detect induced mutations in tetraploid and hexaploid wheat. BMC plant biology *9*.

Upadhyay, S.K., Kumar, J., Alok, A., and Tuli, R. (2013). RNA-guided genome editing for target gene mutations in wheat. G3: Genes| Genomes| Genetics *3*, 2233-2238.

Usher, S., Haslam, R.P., Ruiz-Lopez, N., Sayanova, O., and Napier, J.A. (2015). Field trial evaluation of the accumulation of omega-3 long chain polyunsaturated fatty acids in transgenic Camelina sativa: Making fish oil substitutes in plants. Metabolic Engineering Communications *2*, 93-98.

Van Hintum, T.J.L., Brown, A.H.D., Spillane, C., and Hodgkin, T. (2000). Core collections of Plant Genetic Resources. Paper presented at: IPGRI Technical Bulletin (IPGRI, Rome).

van Nimwegen, K.J., van Soest, R.A., Veltman, J.A., Nelen, M.R., van der Wilt, G.J., Vissers, L.E., and Grutters, J.P. (2016). Is the $1000 Genome as Near as We Think? A Cost Analysis of Next-Generation Sequencing. Clin Chem.

Visendi, P., Batley, J., and Edwards, D. (2013). Next generation characterisation of cereal genomes for marker discovery. Biology *2*, 1357-1377.

Vollmann, J., and Eynck, C. (2015). Camelina as a sustainable oilseed crop: contributions of plant breeding and genetic engineering. Biotechnol J *10*, 525-535.

Voytas, D.F., and Gao, C. (2014). Precision genome engineering and agriculture: opportunities and regulatory challenges. PLoS Biology *12*, e1001877.

Waltz, E. (2016). Gene-edited CRISPR mushroom escapes US regulation. Nature *532*, 293.

Wang, F., Chen, H., Li, X., Wang, N., Wang, T., Yang, J., Guan, L., Yao, N., Du, L., Wang, Y., *et al.* (2015a). Mining and identification of polyunsaturated fatty acid synthesis genes active during camelina seed development using 454 pyrosequencing. BMC plant biology *15*, 147.

Wang, L., Shen, W., Kazachkov, M., Chen, G., Chen, Q., Carlsson, A.S., Stymne, S., Weselake, R.J., and Zoua, J. (2012a). Metabolic Interactions between the lands cycle and the kennedy pathway of glycerolipid synthesis in Arabidopsis developing seeds. Plant Cell *24*.

References

Wang, L., Wang, L., Tan, Q., Fan, Q., Zhu, H., Hong, Z., Zhang, Z., and Duanmu, D. (2016). Efficient Inactivation of Symbiotic Nitrogen Fixation Related Genes in Lotus japonicus Using CRISPR-Cas9. Frontiers in plant science *7*, 1333.

Wang, N., Wang, Y., Tian, F., King, G.J., Zhang, C., Long, Y., Shi, L., and Meng, J. (2008). A functional genomics resource for Brassica napus: development of an EMS mutagenized population and discovery of FAE1 point mutations by TILLING. New Phytol *180*.

Wang, T., Uauy, C., Till, B., and Liu, C.M. (2010). TILLING and associated technologies. Journal of Integrative Plant Biology *52*, 1027-1030.

Wang, T.L., Uauy, C., Robson, F., and Till, B. (2012b). TILLING in extremis. Plant Biotechnology Journal *10*, 761-772.

Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C., and Qiu, J.-L. (2014). Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. Nat Biotech *32*, 947-951.

Wang, Z.-P., Xing, H.-L., Dong, L., Zhang, H.-Y., Han, C.-Y., Wang, X.-C., and Chen, Q.-J. (2015b). Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in Arabidopsis in a single generation. Genome biology *16*, 1.

Waszkowiak, K., and Rudzin´ska, M. (2014). Effect of Flaxseed Meals and Extracts on Lipid Stability in a Stored Meat Product J Am Oil Chem Soc *91*.

Weil, C.F., and Monde, R.-A. (2007). Getting the Point—Mutations in Maize Crop Sci *47*, S-60-S-67.

Wells, R., Trick, M., Soumpourou, E., Clissold, L., Morgan, C., Werner, P., Gibbard, C., Clarke, M., Jennaway, R., and Bancroft, I. (2014). The control of seed oil polyunsaturate content in the polyploid crop species Brassica napus. Molecular Breeding *33*, 349-362.

Wolt, J.D., Wang, K., and Yang, B. (2015). The Regulatory Status of Genome-edited Crops. Plant Biotechnol J.

Wu, G., Wu, Y., Xiao, L., Li, X., and Lu, C. (2008). Zero erucic acid trait of rapeseed (Brassica napus L.) results from a deletion of four base pairs in the fatty acid elongase 1 gene. TAG Theoretical and applied genetics Theoretische und angewandte Genetik *116*, 491-499.

Wu, J.-L., Wu, C., Lei, C., Baraoidan, M., Bordeos, A., Madamba, M.R., Ramos-Pamplona, M., Mauleon, R., Portugal, A., Ulat, V., *et al.* (2005). Chemical- and Irradiation-induced

References

Mutants of Indica Rice IR64 for Forward and Reverse Genetics. Plant Molecular Biology *59*, 85-97.

www-naweb.iaea.org/nafa/pbg/ (2015).

Xin, Z., Li Wang, M., Barkley, N.A., Burow, G., Franks, C., Pederson, G., and Burke, J. (2008). Applying genotyping (TILLING) and phenotyping analyses to elucidate gene function in a chemically induced sorghum mutant population. BMC plant biology *8*.

Yan, G., Li, D., Cai, M., Gao, G., Chen, B., Xu, K., Li, J., Li, F., Wang, N., Qiao, J., *et al.* (2015). Characterization of FAE1 in the zero erucic acid germplasm of Brassica rapa L. Breeding Science *65*, 257-264.

Yang, J., Astatkie, T., and He, Q.S. (2016). A comparative study on the effect of unsaturation degree of camelina and canola oils on the optimization of bio-diesel production. Energy Reports *2*, 211-217.

Yang, J., Corscadden, K., He, Q.S., and Caldwell, C. (2015). The Optimization of Alkali-Catalyzed Biodiesel Production from Camelina sativa Oil Using a Response Surface Methodology. J Bioprocess Biotech *5*, 235.

Yang, Q., Fan, C., Guo, Z., Qin, J., Wu, J., Li, Q., Fu, T., and Zhou, Y. (2012). Identification of FAD2 and FAD3 genes in Brassica napus genome and development of allele-specific markers for high oleic and low linolenic acid contents. TAG Theoretical and applied genetics Theoretische und angewandte Genetik *125*, 715-729.

Zhang, J., Liu, H., Sun, J., Li, B., Zhu, Q., Chen, S., and Zhang, H. (2012). Arabidopsis Fatty Acid Desaturase FAD2 Is Required for Salt Tolerance during Seed Germination and Early Seedling Growth. PLOS ONE *7*, e30355.

Zhang, Y., Zhang, F., Li, X., Baller, J.A., Qi, Y., Starker, C.G., Bogdanove, A.J., and Voytas, D.F. (2013). Transcription Activator-Like Effector Nucleases Enable Efficient Plant Genome Engineering. Plant Physiology *161*, 20-27.

Zhu, C., Bortesi, L., Baysal, C., Twyman, R.M., Fischer, R., Capell, T., Schillberg, S., and Christou, P. (2016). Characteristics of Genome Editing Mutations in Cereal Crops. Trends Plant Sci.

ZIELIŃSKA, A., and Nowak, I. (2014). Fatty acids in vegetable oils and their importance in cosmetic industry CHEMIK *68*, 103-110.

References

Zuryn, S., Le Gras, S., Jamet, K., and Jarriault, S. (2010). A Strategy for Direct Mapping and Identification of Mutations by Whole-Genome Sequencing. Genetics *186*, 427-430.

7. Appendices

A1. PRIMER DESIGNED FOR TILLING AND TILLING BY SEQUENCING

A1.1 M-13 TILLING PRIMERS FOR *FAD2*

>gi|315435717|gb|HQ008322.1| Camelina sativa microsomal oleate desaturase (*FAD2-3*) gene, complete cds

GAGAGAGAGGTTCTTATCTGCGGAGGAGCTTCTTCCTCGTAGGGGGTTTCTTCTTCTTCATCGTTATTAACGAAC
CATCGTTGAGTCAAATCTCCCAACCCCCTACGTCAGCTCCAGGTCCGTCCCTTCTCATTTCATATTCCGATTCAT
TTACGTCTCGTCTGATCTGTTCTTTGTTTTTATTTTTTTTTTCTTTCTCCCGCACTATCTCATTTTCGATTCTTT
TTTTTTTTAGAACCGATTTGATGATATAGATCTGGCTTATATGTCTTGCATTCAACCTTAGATCTGGTCTCGATG
CTCTGTTTTTTCTTTAGTTGAGAAATCTGATGTTGTTGTTACAATGAGTTCTTATTCATAATAATGATTACTAGT
GCCTTGGGTCATCCATGAAAACGATATGTTGTTATACTATGATTTTTTATTTGTCAAAAATGAAGTGCTGCTTTG
ACCCATTCTCTTTAGATATTTTTTTTATTATTTTTGTTGGGTTGGTAGAATAGTGAGAATCACCATAAAATTCTC
TTATCAGTTTCACGTCCTGTTTTTTTTTCAAAAAGATCCGTAATATATTCAGTTTTTTTTTATTTGTGTGTAAAA
TTTGCTTTGTATTGCCGTGTTGAATCTCTTATGGATTTGACCTATGCCTACCGGGTCTTATGGATTGATGATATA
ACTTTTTAAACAGACAAAATAAGTTTCACTGAATGATTCTCATAAACTATACAATAAAGTTGGAATTAGGGTAAT
TCAGGATCAGATGCGTAGATTCAGATGCAAATAATGAGTTGCATGACATTTTATTATTATAGATCCGTAACCGTA
AAGACATATTATGTTCTGTTTTAAATATCTGATATATGATGTCGGCAAATTTCGGTGGTTTATACATCACTTAAA
AACTGTTTCCTTTGGACTTGTTTGCCAACTTGGTGCTATTCAGGAGGATTTCTGATTTTGATTGATCCATTTACT
TTCTCTATTGTTTTTTTTTCTGGGGGTCTACTTGTTGGTATTCATAAAAGAATTTTTGATCTTGATTGAATTTAA
TTACAAGAAACTGCTGATGATAACCACAATAAAGAGATTGTGACCTGTCGTATTGAAATCTTATT**AGTAGTAGTA**
GTCGTGTTCTCAACGTCAATGGGTTTTTCTTTCTTTGGTTTCTTACTTTACGCCGCTTCTCTGCTCTTTTTATTC
CTTTTGGTCCACGCTCTCTCCTTTTGTGGCAATCCC**TCTCACAACCTCATCTCTGAATAAC**AATAATAATTACTA
GTTTGTTGATTTCATCATTACCACTCGTTTTCTAGTGCATGCAAAATTTGTCAATTAGTGATTAACAAAAGATTC
TTTTCTTGAGTTTCAGCTTTTTGATTTTTCTTTGCTCTATGTTTCTTTTGCAGAAACATGGGTGCAGGTGGAAGA
ATGCCAGTTCCTTCTTCTTCTTCCAAGAAATCTGAAACCGATGCCATAAAGCGTGTGCCCTGCGAGAAACCGCCG
TTCACGCTTGGAGAACTGAAGAAAGCAATCCCCCCGCAGTGTTTCAAACGCTCTATCCCTCGCTCTTTCTCCTAC
CTTATCACTGACATCATTGTTGCCTCCTGCTTCTACTACGTCGCCACCAATTACTTCTCTCTCCTCCCTCAGCCT
CTCTCTTACTTGGCTTGGCCTCTCTACTGGGCCTGTCAAGGCTGTGTCCTAACCGGTGTCTGGGTCATAGCTCAC
GAATGCGGTCACCACGCATTCAGCGACTACCAATGGCTTGATGACACAGTTGGTCTTATCTTCCATTCCTTCCTT
CTCGTCCCTTACTTCTCCTGGAAGTACAGTCATCGTCGTCACCATTCCAACACAGGATCTCTCGAAAGAGATGAA
GTATTTGTCCCAAAGCAGAAATCAGCTATCAAGTGGTATGGCAAATACCTCAACAACCCTCCTGGACGCATCATG
ATGTTAACCGTCCAGTTTGTCCTCGGGTGGCCCTTGTACTTGGCCTTTAACGTCTCGGGCAGACCGTACGACGGG
TTCGCTTGCCATTTCTTCCCCACAGCTCCCATCTACAACGACCGCGAACGCCTCCAGATATATCTCTCTGATGCC
GGTATCCTAGCAGTCTGTTTTGGGCTTTACCGTTACGTTGCTGCACAAGGAATGGCCTCGATGATCTGCCTCTAC
GGAGTACCGCTTCTGATAGTGAACGCGTTCCTCGTCTTGATCACTTACTTGCAGCACACTCATCCTGCGTTGCCT
CACTACGATTCATCCGAGTGGGATTGGCTTAGGGGAGCTTTGGCTACCGTTGACAGAGACTATGGAATCTTGAAC
AAGGTGTTCCACAACATCACGGACACACATGTGGCTCATCATCTGTTCTCGACAATGCCACATTATAATGCGATG
GAAGCTACAAAGGCGATAAAGCCAATACTCGGTGACTACTACCAGTTCGACGGAACACCGTGGTATGTGGCGATG
TATAGGGAGGCAAAGGAGTGTATCTATGTAGAACCGGACAGAGAAGGTGACAAGAAAGGTGTGTACTGGTACAAC
AATAAGTTATGAGGATGATGATGGTGAAAGAACACTGAAGAAATTGTCGAACTTTCTCTAGTCTGGTTCTCTTTT
GTTTAAGAAGTTATGTTTTGTTTCAATAATTTCAATGTCCATTTTGTTGTGTTATGACATTTTGGCAAATTATGT
GATGTGGGAAGTTAGTGTTTAAATGTTTTGTGTCTAAAAAAAAAAAAAAAAAAAA

SELECTED M13 TILLING PRIMER COMBINATION

Amplicon	Amplicon size	Primer	Sequence(5’-3’)	Length
Both Amp 1 and Amp 2	N1 – 1598 bp	54F	AGTAGTAGTAGTCGTGTTCTCAACG	25
		55R	AACTTCCCACATCACATAATTTGCC	25
Amp 1	N2 – 941 bp	56F	CACGACGTTGTAAAACGACTAGTTTGTTGATTTCATC	37
		57R	GGATAACAATTTCACACAGGCAAGACGAGGAACGCGT	37
Amp 2	N2 – 713 bp	58F	CACGACGTTGTAAAACGACGTCGTCACCATTCCAAC	36
		58R	GGATAACAATTTCACACAGGGTGTTCTTTCACCATCA	37

A1.2 M-13 TILLING PRIMERS FOR FAE1

>gi|306976207|gb|GU929421.1| Camelina sativa fatty acid elongase 1 (*FAE1-B*) gene, complete cds

GGTATGAATTGNCTTACACGGAAGCCAAAGGAAGAATGAGGAAAGGGAACAGAGTTTGGCAGATTGCTTT
TGGAAGCGGGTTTAAATGTAACAGCGCGGTTTGGGTGGCTCTCCGCGATGTCGAGCCCTCGTTTAAAAAT
CCTTGGGAACATTGCATCGATAGATATCCGGTTAAGATCGATCTTTGAACTCGTAAGAACGGTAGATTGG
TCTGGAAACATGGTTAGTCCTCCATGTACCAAAAAAAAAAGGTTAACTCTTATATCTTTTGTTCTTTACC
AAGGGGTCAAGAAAATCAGTGGAAGGTTAATGTATGTTTGTATATGTTGTGTATAAAGGTAACCCTTAAT
TAATGACCAAGAATGGTTATGAAGTCAAAAGATAGGACCAAAAGTGTTGACCTATTAAAATTTAAAATTT
TAGTTTTTTGGAAAAATTAAATCATTTGTGAAATAAAAATGGTGAGTTAGAAAAATCAAAACAATTTACT
TACTTACAAAGAACCTTAAAATTAAACTTAAGTGTTAAATTTAGAATAAATGAAAAAATGCTAAACTTAT
GACACCAAAACCTACGAACTCTTATAAGGACGGGGATACTAAATTTGTTTGTTTGTTTGTTTTTTTTTTT
TTTTTTGTAAATCAATCCGGTTTGTTTTCAATGTATTCAAATTAAAGTCGGGTTATGACACCAACACCAA
GCTTACGGATCTGAGGAAGACGGGATGCTAAGTCGGGTTAGTCATCAAGCGTATTCAAATAACTCTGAAG
TCTTAGAACTTTTTTTGTTAATAAAATATTTTGTATAATAAAAAAAATTTAGTTAAAAAAATTTAAAACA
ATTTTTACATAAAATATTATTTGTGAAAGAACTTTTTAAGTGTTAAATTTAGAACACTCGAAAACATGTT
AGACTTGTGACACCAAGCTTATTCCAATTTAACGCTTAAGTTTGAAAACTTTTTTTGTCAATAAAAAATA
TACTATATGAAAAGATTGGTTATAAAAACCTTAAATTTAGTTTAAAGTCTGAAAATAAAAATAAAAAATA
AAAAAAATTATTAGTTGGAAAGGAAATAAAGAAAGCAATATACCTCTGGTTGTGATGTGGAAGAAACACA
AGATTTGATCGATTCTCATCAGCTTCAGATGACTGAATGAATAGTCGGAACCATGGAAACACCACCAAAC
ACACAGGATCATGACTGGTTCACTACTCTCACTCGTCCTCGTTCCCGCATTCTCGCATCCTTGAGCCGTC
TTTGACAGATGGTTAAAACACGAACAGAGCTGAATAAGTAGAAGAAAAAAGAGAGAGTTTGAGAAGTTTG
ATCGAATGAATAAGAAGAAAGACTTTTTTTGACCTGGAATCGATCTCGATTGAACAGAGATGTTTTCCTC
TTTACCGCATTAGATCAGATCAATCCACCGTGTGAAAACGTTTCGGTGATCGTAGAGCAGATTTTGACCT
AATTAATTAAGATCGATCACGTACGATCGGATGACGATAATGTAAGAGATCAATCACGAAAAGAGATGCT
CCGATCAAATATCAAAGGGATAATAAAGTTGATGCAGAGATCATCTCTGTTCGTGGTTGTTTGAGATTAG
GGAGAAGACTTGACCTGGAATCAATCTCGATCGAATGAAGATGTTGTCACGATTAACATGGGGATGGCGA
TCAGTCAATCACGAACAGAGTCTGATATGATTCCCCCTCCGATCTTGACAGAGATCACCAACGCAAAGCT
TTGGATTTGGTTTCTCAAAGAAGAAGAAGAAGAGAGATTTGGAGAAACTTTGATGGTTTTACAAAGAAGA
AGAGAGAATGAATGAATGATTTGGAGATTTTTTAACCTAATGACTAATAATAATGAGCCGAGTGTATGAA
TGTATTTAATATAAGCAACGGCTATTTTTTCTTCTTTTCTTTTAACTTTAAGCAAATTTAGAAAAGGAAA
TTGTGATTGTTTCCTTTTTTCCCCACGAGCTCTCAAGTTGAATCCTTTTCTTATTTGAGTTGCAACCAAT
AATAAAAAAAAAAGATAGAAAAACGGAAATTTATCACAAGAGTAACAAAAAGGAGATTCGAGATCCTTTA
GAGTTTACTCAATTACATTCATATGTGCTAGTTGTGGGAGTGAGAGGAGTTTTCTCCTTCCGAAGTGATT
TATGTATGGAGGAGTTTATCACCGTTAAGAGTTCCGAATTGAAAGAGACTATTTGATTGCTAAAATTGTA
TATTACTGTCTGAGAGAAAAAATATTCGATCCCCCACAAAGTCTCCCCCCCCTTATATATTTATACAGAC
CAATTAATTACCCAATTAATGCGTAATTAATTTTACACGATTCTCGCGTCCTAATTAATTCACTTGAATG
CTAGAATCTTTGACCGAGGTCGAACTGCGAGCTGACTAATATGACGCCTCTCCGCGCTCTCTTCTTCTCT
CTGGGCCTGATGGGCCGACCAGCGCTACACCTTTGACCCGTTTAGCGTGCCCGGCCCAGGTCCTCTGCCT
GTTATCCGAGATGACGAATCGTGGGTACAACACTAGTGTATGACCTTTTGCCTTTATAGGCACTAAAAAA
CTACCCAAAAAAAATATGAAGAAGAAAAAAAGGTTAAACAATCACTGGATCAGTGTGCTCATGGTGTTCC
AGGTCCAAAGCATGTGTGTCACGAGATAAAAATAGAGAGAAACACAACTGATTATGATCACCAAGATCAC
GAATATATATAAAGGACTTATAAAAAATGAATTTGAAAGTGGTTAAACTAAGTGATTATATGTGTGATTG
CTCTTAGCCCCTTAGGTGTGGTGAATCTTATTATAGAGATGACTTATTTTAAAGATAGTTGCAATTAAAA
AAAAAAATCAGTGTCCGGATAGTTTGATGCAATTAATTAGTGTTCTATAGGGTCTGATTCTTAATATTTA
TGCAAATATTATAGTATTTCAAAGTATTTGGCCTTAACTAAACTTCCACCTGATTTATTTACTGATCTAG
TTCGGGGACAGACTTTGCGAATAAAACTCGTTCCCGAGAAACATTCATCCCATAACTGCTATTTAGTCAG
AGGCTAATCGACTATAGCCTTTCAGCCAATCAAATCTACGAACACGAATCCCCCTAAAACATCCTCAAGT
ATTTATTTAATACACATGTATCGTATTGAGCACCACTCATAAACTATTTTTTTTTTTGTTTTTAACAAAA
AAAATTTATCATACTCTTTTGTAATAATAGATGCATCAACATATTGTAGGCAACGTTGAAGAACCAGTAC
ATTCTTTTTTTTTTTGCTCCAAATTTTCAAAATTGGAAAATGAAACTTGGACGAAATAAATTTAACACTC
TGTATATATTGGCAATATAATATTGCAGAGTGGACTATTTACCTTATTTTGGCAACTTTCAGTGGACTAG
TAATTTATTTCAATGTGTATGCTTGCATGAGTGTGAATATACACATGTCTATATGCATGCTTGCAAATCG
TAACGGACCACAAAAAAGGATCCATACAAATACCTCTTAACGGCTCCTCTCTATCATACTCTCCGACACA

AACTGAGCAATGACGTCCGTTAACGCAAAGCTCCTTTACCATTACGTTCTAACCAACTTTTTCAACCTTT
GCTTGTTTCCGTTAACGGCGTTACTTGCCGGAAAAGCCTCTAAGCTTACAGCAAACGATCTCTACCACTT
CTATTCCCATCTCCAACACAACCTTATAACCGTAATTTTACTCTTTGCTTTCACCGCTTTCGGTTTGGTT
CTCTACATTGTAACCCGGCCCAAACCGGTTTACCTCGTTGACTACTCGTGCTACCTTCCACCACCGCATC
TCAAAGTTAGTGTTTCCAAGGCGATGGATATTTTCTACCAAATAAGAAAAGCTGATACCTCACGGAACGT
GGCATGCGATGATCCATCCTCGCTTGATTTCCTGAGGAAGATTCAAGAACGTTCAGGTCTAGGTGATGAA
ACGTACAGTCCCCAGGGACTCATTAACGTGCCCCCACAAAAGACCTTTGCAGCTTCACGTGAAGAGACAG
AGCAGGTAATCATCGGTGCGCTAGAAAAGCTATTCGAGAACACCAAAGTAAACCCTAGAGAGATTGGTAT
ACTTGTGGTGAACTCAAGCATGTTTAATCCAACTCCTTCGCTATCTGCGATGGTCGTTAACACTTTCAAG
CTCCGAAGCAACATCAAAAGCTTTAGTCTCGGAGGAATGGGTTGTAGTGCTGGTGTTATCGCCATTGATC
TTGCAAAGGACTTGTTGCATGTTCATAAAAACACTTATGCACTTGTGGTGAGCACTGAGAACATCACTCA
AGGCATTTATGCTGGCGAAAACAGATCCATGATGGTTAGCAATTGCTTGTTTCGTGTTGGTGGGGCAGCG
ATTTTGCTCTCCAACAAACTGGGAGATCGGAGACGGTCCAAGTACAAGCTATGTCATACTGTTCGAACGC
ATACCGGAGCTGATGACAAGTCTTTTCGATGTGTGCAACAAGGAGACGATGAGGGCGGTAAAATCGGAGT
TTGTCTGTCAAAGGACATAACCGTTGTTGCGGGGACAGCGCTTAAGAAAAACATAGCAACGTTGGGTCCG
TTGATTCTTCCTTTAAGCGAAAAGTTTCTGTTTTTAGTTACCTTCATCGCCAAGAAACTTTTGAAGGACA
AGATCAAGCACTGTTACGTCCCGGATTTCAAGCTTGCTATCGACCATTTCTGTATTCATGCGGGAGGCAG
AGCCGTGATCGATGTGCTTGAGAAGAGCTTAGGACTATCGCCAATCGATGTGGAGGCATCTAGATCAACG
TTACATAGATTTGGGAATACTTCGTCTAGCTCAATTTGGTATGAATTGGCATACATAGAAGCAAAAGGAA
GGATGAAGAAAGGGAATAGAGCTTGGCAGATTGCTTTAGGGTCAGGGTTTAAGTGTAACAGTGCGGTTTG
GGTGGCTCTATGCAATGTCAAGGCTTCGGCGAATAGTCCTTGGGAAGATTGCATCGATAGATATCCGGTT
CAAATTGATTCTGATTCATCAAAATCAGAGACTCATGTCAAAAACGGTCGGTCCTAA

SELECTED M13 TILLING PRIMER COMBINATION

Amplicon	Amplicon size	Primer	Sequence(5'-3')	Length
For both Amp 1 and Amp 2	N1 - 1743 bp	61F	GCAACGTTGAAGAACCAGTACATTC	25
		61R	TATCTATCGATGCAATCTTCCCAAG	25
Amp 1	N2 - 1018 bp	63F	CACGACGTTGTAAAACGACGCAGAGTGGACTATTTAC	36
		62R	GGATAACAATTTCACACAGGTACTTGGACCGTCTCCG	37
Amp 2	N2 - 849 bp	64F	CACGACGTTGTAAAACGACGGTAATCATCGGTGCGC	36
		65R	GGATAACAATTTCACACAGGCCAAATTGAGCTAGACG	37

A1.3 TbyS TILLING PRIMERS FOR *FAD2*

Plant: Camelina sativa

Target: FAD2
Mutation detection system:
- ✓ Miseq Desktop sequencer (Illumina®)
- ✓ MiSeq Reagent Kit v3 (2 x 300 bp)
- ✓ Pooling: 2-D, full collection, **pool A CSC2** (**2496** families)

Exon Intron Missense Mutation Silent Mutation STOP Mutation

EXON POSITIONS : ONE EXON (1408..2562)

REGION OF NGS AMPLICON 1 – 1464 TO 1866

REGION OF NGS AMPLICON 2 – 1761 TO 2220

REGION OF NGS AMPLICON 3 – 2046 TO 2521

>gi|315435717|gb|HQ008322.1| Camelina sativa microsomal oleate desaturase (*FAD2-3*) gene, complete cds

GAGAGAGAGGTTCTTATCTGCGGAGGAGCTTCTTCCTCGTAGGGGGTTTCTTCTTCTTCATCGTTATTAACGAAC
CATCGTTGAGTCAAATCTCCCAACCCCCTACGTCAGCTCCAGGTCCGTCCCTTCTCATTTCATATTCCGATTCAT
TTACGTCTCGTCTGATCTGTTCTTTGTTTTTATTTTTTTTTTCTTTCTCCCGCACTATCTCATTTTCGATTCTTT
TTTTTTTTAGAACCGATTTGATGATATAGATCTGGCTTATATGTCTTGCATTCAACCTTAGATCTGGTCTCGATG
CTCTGTTTTTTCTTTAGTTGAGAAATCTGATGTTGTTGTTACAATGAGTTCTTATTCATAATAATGATTACTAGT
GCCTTGGGTCATCCATGAAAACGATATGTTGTTATACTATGATTTTTTATTTGTCAAAAATGAAGTGCTGCTTTG
ACCCATTCTCTTTAGATATTTTTTTTATTATTTTTGTTGGGTTGGTAGAATAGTGAGAATCACCATAAAATTCTC
TTATCAGTTTCACGTCCTGTTTTTTTTTCAAAAAGATCCGTAATATATTCAGTTTTTTTTTATTTGTGTGTAAAA
TTTGCTTTGTATTGCCGTGTTGAATCTCTTATGGATTTGACCTATGCCTACCGGGTCTTATGGATTGATGATATA
ACTTTTTAAACAGACAAAATAAGTTTCACTGAATGATTCTCATAAACTATACAATAAAGTTGGAATTAGGGTAAT
TCAGGATCAGATGCGTAGATTCAGATGCAAATAATGAGTTGCATGACATTTTATTATTATAGATCCGTAACCGTA
AAGACATATTATGTTCTGTTTTAAATATCTGATATATGATGTCGGCAAATTTCGGTGGTTTATACATCACTTAAA
AACTGTTTCCTTTGGACTTGTTTGCCAACTTGGTGCTATTCAGGAGGATTTCTGATTTTGATTGATCCATTTACT
TTCTCTATTGTTTTTTTTTCTGGGGGTCTACTTGTTGGTATTCATAAAAGAATTTTTGATCTTGATTGAATTTAA
TTACAAGAAACTGCTGATGATAACCACAATAAAGAGATTGTGACCTGTCGTATTGAAATCTTATT**AGTAGTAGTA**
GTCGTGTTCTCAACGTCAATGGGTTTTTCTTTCTTTGGTTTCTTACTTTACGCCGCTTCTCTGCTCTTTTTATTC
CTTTTGGTCCACGCTCTCTCCTTTTGTGGCAATCCC**TCTCACAACCTCATCTCTGAATAAC**AATAATAATTACTA
GTTTGTTGATTTCATCATTACCACTCGTTTTCTAGTGCATGCAAAATTTGTCAATTAGTGATTAACAAAAGATTC
TTTTCTTGAGTTTCAGCTTTTTGATTTTTCTTTGCTCTATGTTTCTTTTGCAGAAACATGGGTGCAGGTGGAAGA
ATGCCAGTTCCTTCTTCTTCTTCCAAGAAATCTGAAACCGATGCCATAAAGCGTGTGCCCTGCGAGAAACCGCCG
TTCACGCTTGGAGAACTGAAGAAAGCAATCCCCCCGCAGTGTTTCAAACGCTCTATCCCTCGCTCTTTCTCCTAC
CTTATCACTGACATCATTGTTGCCTCCTGCTTCTACTACGTCGCCACCAATTACTTCTCTCTCCTCCCTCAGCCT
CTCTCTTACTTGGCTTGGCCTCTCTACTGGGCCTGTCAAGGCTGTGTCCTAACCGGTGTCTGGGTCATAGCTCAC
GAATGCGGTCACCACGCATTCAGCGACTACCAATGGCTTGATGACACAGTTGGTCTTATCTTCCATTCCTTCCTT
CTCGTCCCTTACTTCTCCTGGAAGTACAGTCATCGTCGTCACCATTCCAACACAGGATCTCTCGAAAGAGATGAA
GTATTTGTCCCAAAGCAGAAATCAGCTATCAAGTGGTATGGCAAATACCTCAACAACCCTCCTGGACGCATCATG
ATGTTAACCGTCCAGTTTGTCCTCGGGTGGCCCTTGTACTTGGCCTTTAACGTCTCGGGCAGACCGTACGACGGG
TTCGCTTGCCATTTCTTCCCCACAGCTCCCATCTACAACGACCGCGAACGCCTCCAGATATATCTCTCTGATGCC
GGTATCCTAGCAGTCTGTTTTGGGCTTTACCGTTACGTTGCTGCACAAGGAATGGCCTCGATGATCTGCCTCTAC

GGAGTACCGCTTCTGATAGTGAACGCGTTCCTCGTCTTGATCACTTACTTGCAGCACACTCATCCTGCGTTGCCT
CACTACGATTCATCCGAGTGGGATTGGCTTAGGGGAGCTTTGGCTACCGTTGACAGAGACTATGGAATCTTGAAC
AAGGTGTTCCACAACATCACGGACACACATGTGGCTCATCATCTGTTCTCGACAATGCCACATTATAATGCGATG
GAAGCTACAAAGGCGATAAAGCCAATACTCGGTGACTACTACCAGTTCGACGGAACACCGTGGTATGTGGCGATG
TATAGGGAGGCAAAGGAGTGTATCTATGTAGAACCGGACAGAGAAGGTGACAAGAAAGGTGTGTACTGGTACAAC
AATAAGTTATGAGGATGATGATGGTGAAAGAACACTGAAGAAATTGTCGAACTTTCTCTAGTCTGGTTCTCTTTT
GTTTAAGAAGTTATGTTTTGTTTCAATAATTTCAATGTCCATTTTGTTGTGTTATGACATTTTGGCAAATTATGT
GATGTGGGAAGTTAGTGTTTAAATGTTTTGTGTCTAAAAAAAAAAAAAAAAAAA

>*CSFAD2* – CDS (START CODON,STOP CODON)

ATGGGTGCAGGTGGAAGAATGCCAGTTCCTTCTTCTTCCAAGAAATCTGAAACCGATGCCATAAAGCGTGTG
CCCTGCGAGAAACCGCCGTTCACGCTTGGAGAACTGAAGAAAGCAATCCCCCCGCAGTGTTTCAAACGCTCTATC
CCTCGCTCTTTCTCCTACCTTATCACTGACATCATTGTTGCCTCCTGCTTCTACTACGTCGCCACCAATTACTTC
TCTCTCCTCCCTCAGCCTCTCTCTTACTTGGCTTGGCCTCTCTACTGGGCCTGTCAAGGCTGTGTCCTAACCGGT
GTCTGGGTCATAGCTCACGAATGCGGTCACCACGCATTCAGCGACTACCAATGGCTTGATGACACAGTTGGTCTT
ATCTTCCATTCCTTCCTTCTCGTCCCTTACTTCTCCTGGAAGTACAGTCATCGTCGTCACCATTCCAACACAGGA
TCTCTCGAAAGAGATGAAGTATTTGTCCCAAAGCAGAAATCAGCTATCAAGTGGTATGGCAAATACCTCAACAAC
CCTCCTGGACGCATCATGATGTTAACCGTCCAGTTTGTCCTCGGGTGGCCCTTGTACTTGGCCTTTAACGTCTCG
GGCAGACCGTACGACGGGTTCGCTTGCCATTTCTTCCCCAACGCTCCCATCTACAACGACCGCGAACGCCTCCAG
ATATATCTCTCTGATGCCGGTATCCTAGCAGTCTGTTTTGGGCTTTACCGTTACGTTGCTGCACAAGGAATGGCC
TCGATGATCTGCCTCTACGGAGTACCGCTTCTGATAGTGAACGCGTTCCTCGTCTTGATCACTTACTTGCAGCAC
ACTCATCCTGCGTTGCCTCACTACGATTCATCCGAGTGGGATTGGCTTAGGGGAGCTTTGGCTACCGTTGACAGA
GACTATGGAATCTTGAACAAGGTGTTCCACAACATCACGGACACACATGTGGCTCATCATCTGTTCTCGACAATG
CCACATTATAATGCGATGGAAGCTACAAAGGCGATAAAGCCAATACTCGGTGACTACTACCAGTTCGACGGAACA
CCGTGGTATGTGGCGATGTATAGGGAGGCAAAGGAGTGTATCTATGTAGAACCGGACAGAGAAGGTGACAAGAAA
GGTGTGTACTGGTACAACAATAAGTTATGA

>***CSFAD2* -PROTEIN**

MGAGGRMPVPSSSSKKSETDAIKRVPCEKPPFTLGELKKAIPPQCFKRSIPRSFSYLITDIIVASCFYYVAT
NYFSLLPQPLSYLAWPLYWACQGCVLTGVWVIAHECGHHAFSDYQWLDDTVGLIFHSFLLVPYFSWK
YSHRRHHSNTGSLERDEVFVPKQKSAIKWYGKYLNNPPGRIMMLTVQFVLGWPLYLAFNVSGRPYDG
FACHFFPNAPIYNDRERLQIYLSDAGILAVCFGLYRYVAAQGMASMICLYGVPLLIVNAFLVLITYLQHT
HPALPHYDSSEWDWLRGALATVDRDYGILNKVFHNITDTHVAHHLFSTMPHYNAMEATKAIKPILGD
YYQFDGTPWYVAMYREAKECIYVEPDREGDKKGVYWYNNKL*

Amplicon 1:

CGATGCCATAAAGCGTGTGCCCTGCGAGAAACCGCCGTTCACGCTTGGAGAACTGAAGAAAGCAATCCCCCCGCA
GTGTTTCAAACGCTCTATCCCTCGCTCTTTCTCCTACCTTATCACTGACATCATTGTTGCCTCCTGCTTCTACTA
CGTCGCCACCAATTACTTCTCTCTCCTCCCTCAGCCTCTCTCTTACTTGGCTTGGCCTCTCTACTGGGCCTGTCA
AGGCTGTGTCCTAACCGGTGTCTGGGTCATAGCTCACGAATGCGGTCACCACGCATTCAGCGACTACCAATGGCT
TGATGACACAGTTGGTCTTATCTTCCATTCCTTCCTTCTCGTCCCTTACTTCTCCTGGAAGTACAGTCATCGTCG
TCACCATTCCAACACAGGATCTCTCGAA

Amplicon 2:

GCTTGATGACACAGTTGGTCTTATCTTCCATTCCTTCCTTCTCGTCCCTTACTTCTCCTGGAAGTACAGTCATCG
TCGTCACCATTCCAACACAGGATCTCTCGAAAGAGATGAAGTATTTGTCCCAAAGCAGAAATCAGCTATCAAGTG
GTATGGCAAATACCTCAACAACCCTCCTGGACGCATCATGATGTTAACCGTCCAGTTTGTCCTCGGGTGGCCCTT
GTACTTGGCCTTTAACGTCTCGGGCAGACCGTACGACGGGTTCGCTTGCCATTTCTTCCCCAACGCTCCCATCTA
CAACGACCGCGAACGCCTCCAGATATATCTCTCTGATGCCGGTATCCTAGCAGTCTGTTTTGGGCTTTACCGTTA
CGTTGCTGCACAAGGAATGGCCTCGATGATCTGCCTCTACGGAGTACCGCTTCTGATAGTGAACGCGTTCCTCGT
CTTGATCACT

Amplicon 3:

CAACGCTCCCATCTACAACGACCGCGAACGCCTCCAGATATATCTCTCTGATGCCGGTATCCTAGCAGTCTGTTT
TGGGCTTTACCGTTACGTTGCTGCACAAGGAATGGCCTCGATGATCTGCCTCTACGGAGTACCGCTTCTGATAGT
GAACGCGTTCCTCGTCTTGATCACTTACTTGCAGCACACTCATCCTGCGTTGCCTCACTACGATTCATCCGAGTG
GGATTGGCTTAGGGGAGCTTTGGCTACCGTTGACAGAGACTATGGAATCTTGAACAAGGTGTTCCACAACATCAC
GGACACACATGTGGCTCATCATCTGTTCTCGACAATGCCACATTATAATGCGATGGAAGCTACAAAGGCGATAAA
GCCAATACTCGGTGACTACTACCAGTTCGACGGAACACCGTGGTATGTGGCGATGTATAGGGAGGCAAAGGAGTG
TATCTATGTAGAACCGGACAGAGAAG

SELECTED TbyS N1 COMBINATIONS

Amplicon	Amplicon size	Primer	Sequence	Len.	Tm	GC%
1	403	380F	TTCCCTACACGACGCTCTTCCGATCTCGATGCCATAAAGCGTGTGC	46	72.6	54
		380R	AGTTCAGACGTGTGCTCTTCCGATCTTTCGAGAGATCCTGTGTTGG	46	70.8	50
2	460	382F	TTCCCTACACGACGCTCTTCCGATCTGCTTGATGACACAGTTGGTC	46	71.7	52
		382R	AGTTCAGACGTGTGCTCTTCCGATCTAGTGATCAAGACGAGGAACG	46	70.8	50
3	476	383F	TTCCCTACACGACGCTCTTCCGATCTCAACGCTCCCATCTACAACG	46	72.6	54
		384R	AGTTCAGACGTGTGCTCTTCCGATCTCTTCTCTGTCCGGTTCTACA	46	70.8	50

PCR conditions FOR *FAD2* – AMP1 and AMP-3

Sequence Nested 1 (Tm = 57 °C)

Per reaction of 10 µl :

- 1 µl buffer PCR 10 X lab (10 mM Tris-HCl pH8, 50 mM KCl, 2.5 mM $MgCl_2$, 0.05% Igepal**)** + 0.4 µl of primer For 10 µM + 0.4 µl of primer Rev 10 µM + 0.4 µl of dNTPs 5 mM + 0.4 µl of Taq lab + 2 µl. crésol red + 3.4 µl water + 2 µl DNA D20
- 5 mn 94 °C; **40 x** (10s 94 °C, 15 s 57 °C and 1 mn 72 °C) ; 5 min 72°C ; 4°C

Sequence Nested 2 (Tm = 57 °C)

Per reaction of 10 µl :

- 1 µl buffer PCR 10 X lab (10 mM Tris-HCl pH8, 50 mM KCl, 2.5 mM $MgCl_2$, 0.05% Igepal**)** + 2 µl of primer For 0.5 µM and primer Rev 0.5 µM (TOGETHER IN ONE MIX) + 0.4 µl of dNTPs 5 mM + 0.4 µl. of Taq lab + 4.2 µl water + 2 µl DNA of PCR1
- 5 mn 94 °C; **35 x** (10s 94 °C, 15 s 57 °C and 1 mn 72 °C) ; 5 min 72°C ; 4°C

PCR conditions FOR *FAD2* – AMP-2

Sequence Nested 1 (Tm = 55 °C)

Per reaction of 10 µl :

- 1 µl buffer PCR 10 X lab (10 mM Tris-HCl pH8, 50 mM KCl, 2.5 mM $MgCl_2$, 0.05% Igepal**)** + 0.4 µl of primer For 10 µM + 0.4 µl of primer Rev 10 µM + 0.4 µl of dNTPs 5 mM + 0.4 µl of Taq lab + 2 µl. crésol red + 3.4 µl water + 2 µl DNA D20
- 5 mn 95 °C; **35 x** (30s 94 °C, 30 s 55 °C and 1 mn 72 °C) ; 5 min 72°C ; 4°C

Sequence Nested 2 (Tm = 55 °C)

Per reaction of 10 µl :

- 1 µl buffer PCR 10 X lab (10 mM Tris-HCl pH8, 50 mM KCl, 2.5 mM $MgCl_2$, 0.05% Igepal**)** + 2 µl of primer For 0.5 µM and primer Rev 0.5 µM (TOGETHER IN ONE MIX) + 0.4 µl of dNTPs 5 mM + 0.4 µl. of Taq lab + 4.2 µl water + 2 µl DNA of PCR1
- 5 mn 94 °C; **35 x** (30s 94 °C, 30 s 55 °C and 1 mn 72 °C) ; 5 min 72°C ; 4°C

A1.4 TbyS TILLING PRIMERS FOR FAE1

Plant: Camelina sativa

Target: FAE1
Mutation detection system:
- ✓ Miseq Desktop sequencer (Illumina®)
- ✓ MiSeq Reagent Kit v3 (2 x 300 bp)
- ✓ Pooling: 2-D, full collection, **pool A CSC2** (**2496** families)

Exon Intron Missense Mutation Silent Mutation STOP Mutation

EXON POSITIONS : ONE EXON (3580..5097)

REGION OF NGS AMPLICON 1 – 3689 TO 4137

REGION OF NGS AMPLICON 2 – 4049 TO 4566

REGION OF NGS AMPLICON 3 – 4466 TO 4994

>*CSFAE1* B -gDNA

GGTATGAATTGNCTTACACGGAAGCCAAAGGAAGAATGAGGAAAGGGAACAGAGTTTGGCAGATTGCTTTTGGAA
GCGGGTTTAAATGTAACAGCGCGGTTTGGGTGGCTCTCCGCGATGTCGAGCCCTCGTTTAAAAATCCTTGGGAAC
ATTGCATCGATAGATATCCGGTTAAGATCGATCTTTGAACTCGTAAGAACGGTAGATTGGTCTGGAAACATGGTT
AGTCCTCCATGTACCAAAAAAAAAAGGTTAACTCTTATATCTTTTGTTCTTTACCAAGGGGTCAAGAAAATCAGT
GGAAGGTTAATGTATGTTTGTATATGTTGTGTATAAAGGTAACCCTTAATTAATGACCAAGAATGGTTATGAAGT
CAAAAGATAGGACCAAAAGTGTTGACCTATTAAAATTTAAAATTTTAGTTTTTTGGAAAAATTAAATCATTTGTG
AAATAAAAATGGTGAGTTAGAAAAATCAAAACAATTTACTTACTTACAAAGAACCTTAAAATTAAACTTAAGTGT
TAAATTTAGAATAAATGAAAAAATGCTAAACTTATGACACCAAAACCTACGAACTCTTATAAGGACGGGGATACT
AAATTTGTTTGTTTGTTTGTTTTTTTTTTTTTTTTTGTAAATCAATCCGGTTTGTTTTCAATGTATTCAAATTAA
AGTCGGGTTATGACACCAACACCAAGCTTACGGATCTGAGGAAGACGGGATGCTAAGTCGGGTTAGTCATCAAGC
GTATTCAAATAACTCTGAAGTCTTAGAACTTTTTTTGTTAATAAAATATTTTGTATAATAAAAAAAATTTAGTTA
AAAAAATTTAAAACAATTTTTACATAAAATATTATTTGTGAAAGAACTTTTTAAGTGTTAAATTTAGAACACTCG
AAAACATGTTAGACTTGTGACACCAAGCTTATTCCAATTTAACGCTTAAGTTTGAAAACTTTTTTTGTCAATAAA
AAATATACTATATGAAAAGATTGGTTATAAAAACCTTAAATTTAGTTTAAAGTCTGAAAATAAAAATAAAAAATA
AAAAAAATTATTAGTTGGAAAGGAAATAAAGAAAGCAATATACCTCTGGTTGTGATGTGGAAGAAACACAAGATT
TGATCGATTCTCATCAGCTTCAGATGACTGAATGAATAGTCGGAACCATGGAAACACCACCAAACACACAGGATC
ATGACTGGTTCACTACTCTCACTCGTCCTCGTTCCCGCATTCTCGCATCCTTGAGCCGTCTTTGACAGATGGTTA
AAACACGAACAGAGCTGAATAAGTAGAAGAAAAAAGAGAGAGTTTGAGAAGTTTGATCGAATGAATAAGAAGAAA
GACTTTTTTTGACCTGGAATCGATCTCGATTGAACAGAGATGTTTTCCTCTTTACCGCATTAGATCAGATCAATC
CACCGTGTGAAAACGTTTCGGTGATCGTAGAGCAGATTTTGACCTAATTAATTAAGATCGATCACGTACGATCGG
ATGACGATAATGTAAGAGATCAATCACGAAAAGAGATGCTCCGATCAAATATCAAAGGGATAATAAAGTTGATGC
AGAGATCATCTCTGTTCGTGGTTGTTTGAGATTAGGGAGAAGACTTGACCTGGAATCAATCTCGATCGAATGAAG
ATGTTGTCACGATTAACATGGGGATGGCGATCAGTCAATCACGAACAGAGTCTGATATGATTCCCCCTCCGATCT
TGACAGAGATCACCAACGCAAAGCTTTGGATTTGGTTTCTCAAAGAAGAAGAAGAAGAGAGATTTGGAGAAACTT
TGATGGTTTTACAAAGAAGAAGAGAGAATGAATGAATGATTTGGAGATTTTTTAACCTAATGACTAATAATAATG
AGCCGAGTGTATGAATGTATTTAATATAAGCAACGGCTATTTTTTCTTCTTTTCTTTTAACTTTAAGCAAATTTA
GAAAAGGAAATTGTGATTGTTTCCTTTTTTCCCCACGAGCTCTCAAGTTGAATCCTTTTCTTATTTGAGTTGCAA
CCAATAATAAAAAAAAAAGATAGAAAAACGGAAATTTATCACAAGAGTAACAAAAAGGAGATTCGAGATCCTTTA
GAGTTTACTCAATTACATTCATATGTGCTAGTTGTGGGAGTGAGAGGAGTTTTCTCCTTCCGAAGTGATTTATGT
ATGGAGGAGTTTATCACCGTTAAGAGTTCCGAATTGAAAGAGACTATTTGATTGCTAAAATTGTATATTACTGTC
TGAGAGAAAAAATATTCGATCCCCCACAAAGTCTCCCCCCCCTTATATATTTATACAGACCAATTAATTACCCAA
TTAATGCGTAATTAATTTTACACGATTCTCGCGTCCTAATTAATTCACTTGAATGCTAGAATCTTTGACCGAGGT
CGAACTGCGAGCTGACTAATATGACGCCTCTCCGCGCTCTCTTCTTCTCTCTGGGCCTGATGGGCCGACCAGCGC
TACACCTTTGACCCGTTTAGCGTGCCCGGCCCAGGTCCTCTGCCTGTTATCCGAGATGACGAATCGTGGGTACAA

CACTAGTGTATGACCTTTTGCCTTTATAGGCACTAAAAAACTACCCAAAAAAAATATGAAGAAGAAAAAAAGGTT
AAACAATCACTGGATCAGTGTGCTCATGGTGTTCCAGGTCCAAAGCATGTGTGTCACGAGATAAAAATAGAGAGA
AACACAACTGATTATGATCACCAAGATCACGAATATATATAAAGGACTTATAAAAAATGAATTTGAAAGTGGTTA
AACTAAGTGATTATATGTGTGATTGCTCTTAGCCCCTTAGGTGTGGTGAATCTTATTATAGAGATGACTTATTTT
AAAGATAGTTGCAATTAAAAAAAAAAATCAGTGTCCGGATAGTTTGATGCAATTAATTAGTGTTCTATAGGGTCT
GATTCTTAATATTTATGCAAATATTATAGTATTTCAAAGTATTTGGCCTTAACTAAACTTCCACCTGATTTATTT
ACTGATCTAGTTCGGGGACAGACTTTGCGAATAAAACTCGTTCCCGAGAAACATTCATCCCATAACTGCTATTTA
GTCAGAGGCTAATCGACTATAGCCTTTCAGCCAATCAAATCTACGAACACGAATCCCCCTAAAACATCCTCAAGT
ATTTATTTAATACACATGTATCGTATTGAGCACCACTCATAAACTATTTTTTTTTTTGTTTTTAACAAAAAAAAT
TTATCATACTCTTTTGTAATAATAGATGCATCAACATATTGTAGGCAACGTTGAAGAACCAGTACATTCTTTTTT
TTTTTGCTCCAAATTTTCAAAATTGGAAAATGAAACTTGGACGAAATAAATTTAACACTCTGTATATATTGGCAA
TATAATATTGCAGAGTGGACTATTTACCTTATTTTGGCAACTTTCAGTGGACTAGTAATTTATTTCAATGTGTAT
GCTTGCATGAGTGTGAATATACACATGTCTATATGCATGCTTGCAAATCGTAACGGACCACAAAAAAGGATCCAT
ACAAATACCTCTTAACGGCTCCTCTCTATCATACTCTCCGACACAAACTGAGCAATGACGTCCGTTAACGCAAAG
CTCCTTTACCATTACGTTCTAACCAACTTTTTCAACCTTTGCTTGTTTCCGTTAACGGCGTTACTTGCCGGAAAA
GCCTCTAAGCTTACAGCAAACGATCTCTACCACTT**CTATTCCCATCTCCAACACAACCTTATAACCGTAATTTTA**
CTCTTTGCTTTCACCGCTTTCGGTTTGGTTCTCTACATTGTAACCCGGCCCAAACCGGTTTACCTCGTTGACTAC
TCGTGCTACCTTCCACCACCGCATCTCAAAGTTAGTGTTTCCAAGGCGATGGATATTTTCTACCAAATAAGAAAA
GCTGATACCTCACGGAACGTGGCATGCGATGATCCATCCTCGCTTGATTTCCTGAGGAAGATTCAAGAACGTTCA
GGTCTAGGTGATGAAACGTACAGTCCCCAGGGACTCATTAACGTGCCCCCACAAAAGACCTTTGCAGCTTCACGT
GAAGAGACAGAGCAGGTAATCATCGGTGCGCTAGAAAAGCTATTCGAGAACACCAAAGTAAACCCTA**GAGAGATT**
GGTATACTTGTGGTGAACTCAAGCATGTTTAATCCAACTCCTTCGCTATCTGCGATGGTCGTTAACACTTTCAAG
CTCCGAAGCAACATCAAAAGCTTTAGTCTCGGAGGAATGGGTTGTAGTGCTGGTGTTATCGCCATTGATCTTGCA
AAGGACTTGTTGCATGTTCATAAAAACACTTATGCACTTGTGGTGAGCACTGAGAACATCACTCAAGGCATTTAT
GCTGGCGAAAACAGATCCATGATGGTTAGCAATTGCTTGTTTCGTGTTGGTGGGGCAGCGATTTTGCTCTCCAAC
AAACTGGGAGATCGGAGACGGTCCAAGTACAAGCTATGTCATACTGTTCGAACGCATACCGGAGCTGATGACAAG
TCTTTTCGATGTGTGCAACAAGGAGACGATGAGGGCGGTAAAATCGGAGTTTGTCTGTCAAAGGAC**ATAACCGTT**
GTTGCGGGGACAGCGCTTAAGAAAAACATAGCAACGTTGGGTCCGTTGATTCTTCCTTTAAGCGAAAAGTTTCTG
TTTTTAGTTACCTTCATCGCCAAGAAACTTTTGAAGGACAAGATCAAGCACTGTTACGTCCCGGATTTCAAGCTT
GCTATCGACCATTTCTGTATTCATGCGGGAGGCAGAGCCGTGATCGATGTGCTTGAGAAGAGCTTAGGACTATCG
CCAATCGATGTGGAGGCATCTAGATCAACGTTACATAGATTTGGGAATACTTCGTCTAGCTCAATTTGGTATGAA
TTGGCATACATAGAAGCAAAAGGAAGGATGAAGAAAGGGAATAGAGCTTGGCAGATTGCTTTAGGGTCAGGGTTT
AAGTGTAACAGTGCGGTTTGGGTGGCTCTATGCAATGTCAAGGCTTCGGCGAATAGTCCTTGGGAAGATTGCATC
GATAGATATCCGGTTCAAATTGATTCTGATTCATCAAAATCAGAGACTCATGTCAAAAACGGTCGGTCCTAA

>*CSFAE1* -CDS

ATGACGTCCGTTAACGCAAAGCTCCTTTACCATTACGTTCTAACCAACTTTTTCAACCTTTGCTTGTTTCCGTTA
ACGGCGTTACTTGCCGGAAAAGCCTCTAAGCTTACAGCAAACGATCTCTACCACTTCTATTCCCATCTCCAACAC
AACCTTATAACCGTAATTTTACTCTTTGCTTTCACCGCTTTCGGTTTGGTTCTCTACATTGTAACCCGGCCCAAA
CCGGTTTACCTCGTTGACTACTCGTGCTACCTTCCACCACCGCATCTCAAAGTTAGTGTTTCCAAGGCGATGGAT
ATTTTCTACCAAATAAGAAAAGCTGATACCTCACGGAACGTGGCATGCGATGATCCATCCTCGCTTGATTTCCTG
AGGAAGATTCAAGAACGTTCAGGTCTAGGTGATGAAACGTACAGTCCCCAGGGACTCATTAACGTGCCCCCACAA
AAGACCTTTGCAGCTTCACGTGAAGAGACAGAGCAGGTAATCATCGGTGCGCTAGAAAAGCTATTCGAGAACACC
AAAGTAAACCCTAGAGAGATTGGTATACTTGTGGTGAACTCAAGCATGTTTAATCCAACTCCTTCGCTATCTGCG
ATGGTCGTTAACACTTTCAAGCTCCGAAGCAACATCAAAAGCTTTAGTCTCGGAGGAATGGGTTGTAGTGCTGGT
GTTATCGCCATTGATCTTGCAAAGGACTTGTTGCATGTTCATAAAAACACTTATGCACTTGTGGTGAGCACTGAG
AACATCACTCAAGGCATTTATGCTGGCGAAAACAGATCCATGATGGTTAGCAATTGCTTGTTTCGTGTTGGTGGG
GCAGCGATTTTGCTCTCCAACAAACTGGGAGATCGGAGACGGTCCAAGTACAAGCTATGTCATACTGTTCGAACG
CATACCGGAGCTGATGACAAGTCTTTTCGATGTGTGCAACAAGGAGACGATGAGGGCGGTAAAATCGGAGTTTGT
CTGTCAAAGGACATAACCGTTGTTGCGGGGACAGCGCTTAAGAAAAACATAGCAACGTTGGGTCCGTTGATTCTT
CCTTTAAGCGAAAAGTTTCTGTTTTTAGTTACCTTCATCGCCAAGAAACTTTTGAAGGACAAGATCAAGCACTGT
TACGTCCCGGATTTCAAGCTTGCTATCGACCATTTCTGTATTCATGCGGGAGGCAGAGCCGTGATCGATGTGCTT
GAGAAGAGCTTAGGACTATCGCCAATCGATGTGGAGGCATCTAGATCAACGTTACATAGATTTGGGAATACTTCG
TCTAGCTCAATTTGGTATGAATTGGCATACATAGAAGCAAAAGGAAGGATGAAGAAAGGGAATAGAGCTTGGCAG
ATTGCTTTAGGGTCAGGGTTTAAGTGTAACAGTGCGGTTTGGGTGGCTCTATGCAATGTCAAGGCTTCGGCGAAT
AGTCCTTGGGAAGATTGCATCGATAGATATCCGGTTCAAATTGATTCTGATTCATCAAAATCAGAGACTCATGTC
AAAAACGGTCGGTCCTAA

>*CSFAE1* – PROTEIN

MTSVNAKLLYHYVLTNFFNLCLFPLTALLAGKASKLTANDLYHFYSHLQHNLITVILLFAFTAFGLVLYIVTRPK
PVYLVDYSCYLPPPHLKVSVSKAMDIFYQIRKADTSRNVACDDPSSLDFLRKIQERSGLGDETYSPQGLINVPPQ
KTFAASREETEQVIIGALEKLFENTKVNPREIGILVVNSSMFNPTPSLSAMVVNTFKLRSNIKSFSLGGMGCSAG
VIAIDLAKDLLHVHKNTYALVVSTENITQGIYAGENRSMMVSNCLFRVGGAAILLSNKLGDRRRSKYKLCHTVRT
HTGADDKSFRCVQQGDDEGGKIGVCLSKDITVVAGTALKKNIATLGPLILPLSEKFLFLVTFIAKKLLKDKIKHC
YVPDFKLAIDHFCIHAGGRAVIDVLEKSLGLSPIDVEASRSTLHRFGNTSSSSIWYELAYIEAKGRMKKGNRAWQ
IALGSGFKCNSAVWVALCNVKASANSPWEDCIDRYPVQIDSDSSKSETHVKNGRS*

Amplicon 1:

CAGCAAACGATCTCTACCACTTCTATTCCCATCTCCAACACAACCTTATAACCGTAATTTTACTCTTTGCTTTCA
CCGCTTTCGGTTTGGTTCTCTACATTGTAACCCGGCCCAAACCGGTTTACCTCGTTGACTACTCGTGCTACCTTC
CACCACCGCATCTCAAAGTTAGTGTTTCCAAGGCGATGGATATTTTCTACCAAATAAGAAAAGCTGATACCTCAC
GGAACGTGGCATGCGATGATCCATCCTCGCTTGATTTCCTGAGGAAGATTCAAGAACGTTCAGGTCTAGGTGATG
AAACGTACAGTCCCCAGGGACTCATTAACGTGCCCCCACAAAAGACCTTTGCAGCTTCACGTGAAGAGACAGAGC
AGGTAATCATCGGTGCGCTAGAAAAGCTATTCGAGAACACCAAAGTAAACCCTAGAGAGATTGGTATACTTGTG

Amplicon 2:

GTGAAGAGACAGAGCAGGTAATCATCGGTGCGCTAGAAAAGCTATTCGAGAACACCAAAGTAAACCCTAGAGAGA
TTGGTATACTTGTGGTGAACTCAAGCATGTTTAATCCAACTCCTTCGCTATCTGCGATGGTCGTTAACACTTTCA
AGCTCCGAAGCAACATCAAAAGCTTTAGTCTCGGAGGAATGGGTTGTAGTGCTGGTGTTATCGCCATTGATCTTG
CAAAGGACTTGTTGCATGTTCATAAAAACACTTATGCACTTGTGGTGAGCACTGAGAACATCACTCAAGGCATTT
ATGCTGGCGAAAACAGATCCATGATGGTTAGCAATTGCTTGTTTCGTGTTGGTGGGGCAGCGATTTTGCTCTCCA
ACAAACTGGGAGATCGGAGACGGTCCAAGTACAAGCTATGTCATACTGTTCGAACGCATACCGGAGCTGATGACA
AGTCTTTTCGATGTGTGCAACAAGGAGACGATGAGGGCGGTAAAATCGGAGTTTGTCTGTCAAAGGAC

Amplicon 3:

ATACTGTTCGAACGCATACCGGAGCTGATGACAAGTCTTTTCGATGTGTGCAACAAGGAGACGATGAGGGCGGTA
AAATCGGAGTTTGTCTGTCAAAGGACATAACCGTTGTTGCGGGGACAGCGCTTAAGAAAAACATAGCAACGTTGG
GTCCGTTGATTCTTCCTTTAAGCGAAAAGTTTCTGTTTTTAGTTACCTTCATCGCCAAGAAACTTTTGAAGGACA
AGATCAAGCACTGTTACGTCCCGGATTTCAAGCTTGCTATCGACCATTTCTGTATTCATGCGGGAGGCAGAGCCG
TGATCGATGTGCTTGAGAAGAGCTTAGGACTATCGCCAATCGATGTGGAGGCATCTAGATCAACGTTACATAGAT
TTGGGAATACTTCGTCTAGCTCAATTTGGTATGAATTGGCATACATAGAAGCAAAAGGAAGGATGAAGAAAGGGA
ATAGAGCTTGGCAGATTGCTTTAGGGTCAGGGTTTAAGTGTAACAGTGCGGTTTGGGTGGCTCTATGCAATGTCA
AGGC

SELECTED TbyS N1 COMBINATIONS

Amp	Amp size	Primer	Sequence	Len.	Tm	GC%
1	449	398F	TTCCCTACACGACGCTCTTCCGATCTCAGCAAACGATCTCTACCAC	46	71.7	52
		397R	AGTTCAGACGTGTGCTCTTCCGATCTCACAAGTATACCAATCTCTC	46	69.0	46
2	518	399F	TTCCCTACACGACGCTCTTCCGATCTGTGAAGAGACAGAGCAGGTA	46	71.7	52
		400R	AGTTCAGACGTGTGCTCTTCCGATCTGTCCTTTGACAGACAAACTC	46	69.9	48
3	529	401F	TTCCCTACACGACGCTCTTCCGATCTATACTGTTCGAACGCATACC	46	70.8	50
		402R	AGTTCAGACGTGTGCTCTTCCGATCTGCCTTGACATTGCATAGAGC	46	70.8	50

PCR conditions FOR *FAE1* – AMP1 and AMP-2

Sequence Nested 1 (Tm = 57 °C)

Per reaction of 10 µl :

- 1 µl buffer PCR 10 X lab (10 mM Tris-HCl pH8, 50 mM KCl, 2.5 mM $MgCl_2$, 0.05% Igepal**)** + 0.4 µl of primer For 10 µM + 0.4 µl of primer Rev 10 µM + 0.4 µl of dNTPs 5 mM + 0.4 µl of Taq lab + 2 µl. crésol red + 3.4 µl water + 2 µl DNA D20
- 5 mn 94 °C; **40 x** (10s 94 °C, 15 s 57 °C and 1 mn 72 °C) ; 5 min 72°C ; 4°C

Sequence Nested 2 (Tm = 57 °C)

Per reaction of 10 µl :

- 1 µl buffer PCR 10 X lab (10 mM Tris-HCl pH8, 50 mM KCl, 2.5 mM $MgCl_2$, 0.05% Igepal**)** + 2 µl of primer For 0.5 µM and primer Rev 0.5 µM (TOGETHER IN ONE MIX) + 0.4 µl of dNTPs 5 mM + 0.4 µl. of Taq lab + 4.2 µl water + 2 µl DNA of PCR1
- 5 mn 94 °C; **35 x** (10s 94 °C, 15 s 57 °C and 1 mn 72 °C) ; 5 min 72°C ; 4°C

PCR conditions FOR *FAE1* – AMP-3

Sequence Nested 1 (Tm = 55 °C)

Per reaction of 10 µl :

- 1 µl buffer PCR 10 X lab (10 mM Tris-HCl pH8, 50 mM KCl, 2.5 mM $MgCl_2$, 0.05% Igepal**)** + 0.4 µl of primer For 10 µM + 0.4 µl of primer Rev 10 µM + 0.4 µl of dNTPs 5 mM + 0.4 µl of Taq lab + 2 µl. crésol red + 3.4 µl water + 2 µl DNA D20
- 5 mn 95 °C; **35 x** (30s 94 °C, 30 s 55 °C and 1 mn 72 °C) ; 5 min 72°C ; 4°C

Sequence Nested 2 (Tm = 55 °C)

Per reaction of 10 µl :

- 1 µl buffer PCR 10 X lab (10 mM Tris-HCl pH8, 50 mM KCl, 2.5 mM $MgCl_2$, 0.05% Igepal**)** + 2 µl of primer For 0.5 µM and primer Rev 0.5 µM (TOGETHER IN ONE MIX) + 0.4 µl of dNTPs 5 mM + 0.4 µl. of Taq lab + 4.2 µl water + 2 µl DNA of PCR1
- 5 mn 94 °C; **35 x** (30s 94 °C, 30 s 55 °C and 1 mn 72 °C) ; 5 min 72°C ; 4°C

A1.5 TbyS TILLING PRIMERS FOR MIPS1

Plant: Camelina sativa

Target: MIPS1

Mutation detection system:

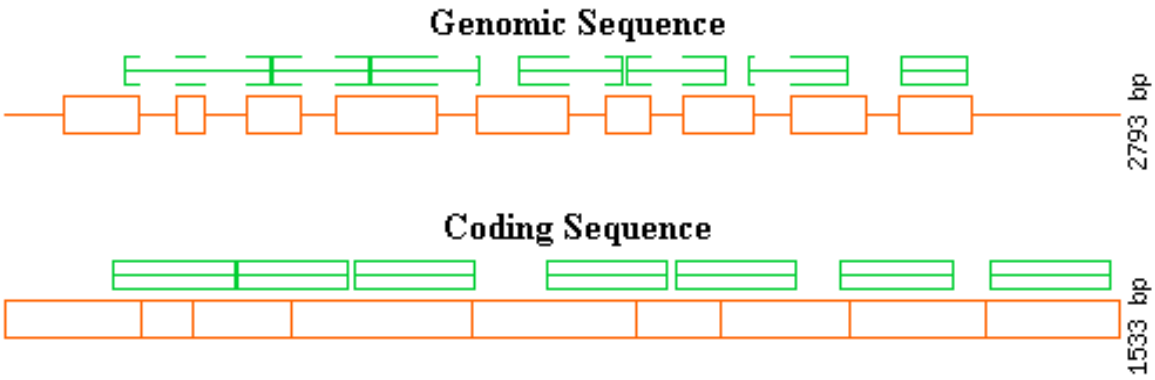
- ✓ Miseq Desktop sequencer (Illumina®)
- ✓ MiSeq Reagent Kit v3 (2 x 300 bp)
- ✓ Pooling: 2-D, full collection, **pool A CSC2** (**2496** families)

Exon Intron Missense Mutation Silent Mutation STOP Mutation

EXON POSITIONS : NINE EXON (150..337, 432..500, 608..743, 833..1080, 1183..1409, 1504..1619, 1699..1875, 1969..2157, 2241..2423)

REGION OF NGS AMPLICON 1 – 763 TO 1219

REGION OF NGS AMPLICON 2 – 77 TO 431



>*CSMIPS1*-Chr12 gDNA **(Forward sequence)**

GGACCACTAACCACTGATATGACGACACGTGTCTCTATCTCCCTCACTGTCTCGTCTATATAAACTCCCACTTCC
TCCGCTGATTAAGAGAAGCACTAAAACAAAGCGATATTACAAAACAAAGAAGCAAATTTTCTTTTATTTCGAA
A**ATGTTCATTAAAGACTTTAAGGTTGAGAGCCTGAACGTGAAGTACACTGAAAATGAGATTCACTCTGTGTA**
TGATTACGAGACGACCGAGGTTGTGCATGAGAACGTCAATGGTACTTTCCAATGGATTGTTAAGCCAAAGA
CTGTCAAATACGATTTCGTTACCGATACTCGTGTCCCCAAATTAGGGTAATTAATTACGTACTCTTTTTTTATTT
TTTTACCAAAGTTTCGATCTTCTTGATTTCTCCGTCGTTCAATTTTGATTTGTTTTTCGTTAAAG**GGTTATGCTTG**
TGGGTTTGGGAGGAAACAATGGATCAACTCTCACCGCTGGTGTTATTGCCAACAAAGAGTGAGTCTTTTTAT
TTTTTTTTGTTGGATAGAATCCTTGTAACGATTTATATATTTGATTTTGTTTTTTTCTTGTCGGAAAATATAATTG
AATTTAAAAATAAACAG**AGGAATCTCGTGGGCTACGAAGGACAAAGTGCAACAAGCGAACTACTTCGGCTC**
ACTAACTCAAGCATCATCCATTCGTGTTGGATCTTTTAACGGTGAAGAGATCTATGCTCCTTTCAAGAGTCTT
CTTCCTATGGTCTGTTTTCTATAACAAACATGATTACCTTCATCCTGAGATGGTTTATGTTATATTGAATTTTGG
GGTTTTTGTTGTTGGCAAAACAG**GTGAATCCGGATGATGTTGTGTTTGGAGGATGGGACATAAGTGATATG**
AACTTAGCAGATGCCATGGCTAGAGCCAAGGTTCTTGATATCGACTTGCAGAAACAGCTGAGGCCTTTCATG
GAGAACATTGTGCCACTCCCTGGGATCTTCGATTCTGATTTCATTGCTGCCAATCAAGGCTCACGTGCTAACC

ACGTGATCAAAGGTACCAAGAAGGAACAAATCGACCACATCATCAAGGACATGAGGTAATAATAATATAGT
GTGTTTAAAACAGAGTTTTTCTTGCTTGTTGAATTAAATCTCCTCTGTTTCGTGATTTTTATTCTTGTGATGTTTT
GGATCTTTAG**GGAGTTTAAGGAGAAGAATAAGGTGGATAGGGTTGTGGTTCTCTGGACGGCTAACACAGA**
GCGTTACAGCAATGTGGTCGTTGGGTTAAACGACACAATGGAGAATCTCATGCAGTCTGTGGAGAAGGATG
AGGCCGAGATATCTCCTTCAACGCTTTATGCTATTGCTTGTGTTCTTGAAGGCATCCCGTTCATCAATGGAAG
CCCTCAGAACACCTTTGTTCCAGGTTTCATATACTTGAAAGTTTTTAGTCATCGATCAATCAAAAATTGGGAAT
CCTTACTGTTTTTTGACTGATTTTCTTGTGGGATTTGTTGTAG**GTCTTATTGATTTGGCGATCAAAAACAATGTT**
TTGATCGGTGGAGATGATTTCAAGAGTGGTCAAACCAAGATGAAATCTGTCTTGGTTGATTTCCTCGTTGGT
GCAGGAATTAAGGTTTGTTTTAACTTGTCTATCAAGTCACACTCTTTCTTGTTAGTGAAGAAATTTGTCAAAAG
ATGTGTGTTTTATATAG**CCTACTTCAATTGTGAGCTACAATCACTTGGGGAACAACGATGGAATGAACCTCTC**
AGCTCCACAGACATTCCGATCTAAGGAGATCTCCAAAAGCAATGTCGTGGACGATATGGTTGCTAGCAACG
GTATCCTCTTTGAGCCTGGGGAACACCCGGACCATGTTGTGGTCATCAAGGTATAGCTAGAAACCATAGAGA
CAATCAAAACCTGTTTCCGCTACGACATCTTCCATTACCCATTGATCCTATTTGTGTCTTTTGTTTTGCAG**TATGT**
ACCGTATGTTGCTGATAGTAAGCGAGCGATGGACGAGTATACATCAGAGATTTTCATGGGAGGGAAGAAC
ACAATCGTGATGCACAATACCTGTGAGGACTCTCTCTTAGCTGCTCCAATCATCTTGGATCTTGTTCTCCTTGC
GGAGCTCAGCACCAGGATCCAGTTCAAATCCGAGGGAGAGGTTTGTGAATTAACAATACATTTACGGTCTTC
AATTGTTGTTAAGCAGGATTTAATTTGATGTTTTGATTCGTTTGCTGGCAG**GGAAAGTTTCATTCGTTCCATCC**
TGTTGCCACCATACTCAGCTACCTCACCAAGGCACCGCTGGTTCCGTCGGGAACGCCGGTGGTTAACGCGCT
GTCTAAGCAGCGGGCTATGCTGGAGAACATTCTTAGGGCGTGCGTTGGACTGGCCCCGGAGAATAACATGA
TCTTGGAATACAAGTGAAGAAGAAGAAGAAGAAGAAGGGATTAAAACAATCCCCTTAAATAGCCCATA
ATTTAAGTTATTAATTTGGGTCTTTGTATTGTGTTTTATTTTATGTTTAAAATTTATGAAAGCATGTGTATCTGAT
TTGATCTCTGTAGGGTGTGAAGTAACTTGTATTTGTAGTAATCCGAATGGGTTGGTATATTTTACAAGTGTTTC
TTCTATAGGTTGGAGACCAAAAAATTTCAGTATCGCTTATATGATTCGGTTTGGTATGTATGTAACCAGCTTTT
ATATATTGAGATATTTAATCATGGTTTCGGTCCCAATATTGATGATCTACGAGAGAGTGTGTATGAACAACTTT
TTTGATGATCTACGAGAA

>*CSMIPS1*-Chr12 gDNA **(original sequence-reverse orientation)**

TTCTCGTAGATCATCAAAAAAGTTGTTCATACACACTCTCTCGTAGATCATCAATATTGGGACCGAAACCATGA
TTAAATATCTCAATATATAAAAGCTGGTTACATACATACCAAACCGAATCATATAAGCGATACTGAAATTTTTT
GGTCTCCAACCTATAGAAGAAACACTTGTAAAATATACCAACCCATTCGGATTACTACAAATACAAGTTACTTC
ACACCCTACAGAGATCAAATCAGATACACATGCTTTCATAAATTTTAAACATAAAATAAAACACAATACAAAGA
CCCAAATTAATAACTTAAATTATGGGCTATTTAAGGGGATTGTTTTAATCCCTTCTTCTTCTTCTTCTTCTTC
ACTTGTATTCCAAGATCATGTTATTCTCCGGGGCCAGTCCAACGCACGCCCTAAGAATGTTCTCCAGCATAGCC
CGCTGCTTAGACAGCGCGTTAACCACCGGCGTTCCCGACGGAACCAGCGGTGCCTTGGTGAGGTAGCTGAGT
ATGGTGGCAACAGGATGGAACGAATGAAACTTTCCCTGCCAGCAAACGAATCAAAACATCAAATTAAATCCTG
CTTAACAACAATTGAAGACCGTAAATGTATTGTTAATTCACAAACCTCTCCCTCGGATTTGAACTGGATCCTGG
TGCTGAGCTCCGCAAGGAGAACAAGATCCAAGATGATTGGAGCAGCTAAGAGAGAGTCCTCACAGGTATTGT
GCATCACGATTGTGTTCTTCCCTCCCATGAAAATCTCTGATGTATACTCGTCCATCGCTCGCTTACTATCAGCAA
CATACGGTACATACTGCAAAACAAAAGACACAAATAGGATCAATGGGTAATGGAAGATGTCGTAGCGGAAAC
AGGTTTTGATTGTCTCTATGGTTTCTAGCTATACCTTGATGACCACAACATGGTCCGGGTGTTCCCCAGGCTCA
AAGAGGATACCGTTGCTAGCAACCATATCGTCCACGACATTGCTTTTGGAGATCTCCTTAGATCGGAATGTCTG
TGGAGCTGAGAGGTTCATTCCATCGTTGTTCCCCAAGTGATTGTAGCTCACAATTGAAGTAGGCTATATAAAA
CACACATCTTTTGACAAATTTCTTCACTAACAAGAAAGAGTGTGACTTGATAGACAAGTTAAAACAAACCTTAA
TTCCTGCACCAACGAGGAAATCAACCAAGACAGATTTCATCTTGGTTTGACCACTCTTGAAATCATCTCCACCG
ATCAAAACATTGTTTTTGATCGCCAAATCAATAAGACCTACAACAAATCCCACAAGAAAATCAGTCAAAAAACA

GTAAGGATTCCCAATTTTTGATTGATCGATGACTAAAAACTTTCAAGTATATGAAACCTGGAACAAAGGTGTTC
TGAGGGCTTCCATTGATGAACGGGATGCCTTCAAGAACACAAGCAATAGCATAAAGCGTTGAAGGAGATATC
TCGGCCTCATCCTTCTCCACAGACTGCATGAGATTCTCCATTGTGTCGTTTAACCCAACGACCACATTGCTGTAA
CGCTCTGTGTTAGCCGTCCAGAGAACCACAACCCTATCCACCTTATTCTTCTCCTTAAACTCCCTAAAGATCCAA
AACATCACAAGAATAAAAATCACGAAACAGAGGAGATTTAATTCAACAAGCAAGAAAAACTCTGTTTTAAACA
CACTATATTATTATTAC**CTCATGTCCTTGATGATGTGGTCGATTTGTTCCTTCTTGGTACCTTTGATCACGTGGT**
TAGCACGTGAGCCTTGATTGGCAGCAATGAAATCAGAATCGAAGATCCCAGGGAGTGGCACAATGTTCTCC
ATGAAAGGCCTCAGCTGTTTCTGCAAGTCGATATCAAGAACCTTGGCTCTAGCCATGGCATCTGCTAAGTTC
ATATCACTTATGTCCCATCCTCCAAACACAACATCATCCGGATTCACCTGTTTTGCCAACAACAAAAACCCCAA
AATTCAATATAACATAAACCATCTCAGGATGAAGGTAATCATGTTTGTTATAGAAAACAGAC**CATAGGAAGAA**
GACTCTTGAAAGGAGCATAGATCTCTTCACCGTTAAAAGATCCAACACGAATGGATGATGCTTGAGTTAGT
GAGCCGAAGTAGTTCGCTTGTTGCACTTTGTCCTTCGTAGCCCACGAGATTCCTCTGTTTATTTTTAAATTCAA
TTATATTTTCCGACAAGAAAAAAACAAAATCAAATATATAAATCGTTACAAGGATTCTATCCAACAAAAAA0AA
ATAAAAAGACTCACTCTTTGTTGGCAATAACACCAGCGGTGAGAGTTGATCCATTGTTTCCTCCCAAACCCAC
AAGCATAACCCTTTAACGAAAAACAAATCAAAATTGAACGACGGAGAAATCAAGAAGATCGAAACTTTGGTA
AAAAAATAAAAAAAGAGTACGTAATTAATTAC**CCTAATTTGGGGACACGAGTATCGGTAACGAAATCGTATT**
TGACAGTCTTTGGCTTAACAATCCATTGGAAAGTACCATTGACGTTCTCATGCACAACCTCGGTCGTCTCGTA
ATCATACACAGAGTGAATCTCATTTTCAGTGTACTTCACGTTCAGGCTCTCAACCTTAAAGTCTTTAATGAAC
ATTTTCGAAATAAAAGAAAATTTGCTTCTTTGTTTTGTAATATCGCTTTGTTTTAGTGCTTCTCTTAATCAGCGG
AGGAAGTGGGAGTTTATATAGACGAGACAGTGAGGGAGATAGAGACACGTGTCGTCATATCAGTGGTTAGT
GGTCC

>Csa12-*MIPS1* CDS
ATGTTCATTAAAGACTTTAAGGTTGAGAGCCTGAACGTGAAGTACACTGAAAATGAGATTCACTCTGTGTATGATTA
CGAGACGACCGAGGTTGTGCATGAGAACGTCAATGGTACTTTCCAATGGATTGTTAAGCCAAAGACTGTCAAATAC
GATTTCGTTACCGATACTCGTGTCCCCAAATTAGGGGTTATGCTTGTGGGTTTGGGAGGAAACAATGGATCAACTCT
CACCGCTGGTGTTATTGCCAACAAAGAAGGAATCTCGTGGGCTACGAAGGACAAAGTGCAACAAGCGAACTACTTC
GGCTCACTAACTCAAGCATCATCCATTCGTGTTGGATCTTTTAACGGTGAAGAGATCTATGCTCCTTTCAAGAGTCTT
CTTCCTATGGTGAATCCGGATGATGTTGTGTTTGGAGGATGGGACATAAGTGATATGAACTTAGCAGATGCCATGG
CTAGAGCCAAGGTTCTTGATATCGACTTGCAGAAACAGCTGAGGCCTTTCATGGAGAACATTGTGCCACTCCCTGG
GATCTTCGATTCTGATTTCATTGCTGCCAATCAAGGCTCACGTGCTAACCACGTGATCAAAGGTACCAAGAAGGAAC
AAATCGACCACATCATCAAGGACATGAGGGAGTTTAAGGAGAAGAATAAGGTGGATAGGGTTGTGGTTCTCTGGA
CGGCTAACACAGAGCGTTACAGCAATGTGGTCGTTGGGTTAAACGACACAATGGAGAATCTCATGCAGTCTGTGGA
GAAGGATGAGGCCGAGATATCTCCTTCAACGCTTTATGCTATTGCTTGTGTTCTTGAAGGCATCCCGTTCATCAATG
GAAGCCCTCAGAACACCTTTGTTCCAGGTCTTATTGATTTGGCGATCAAAAACAATGTTTTGATCGGTGGAGATGAT
TTCAAGAGTGGTCAAACCAAGATGAAATCTGTCTTGGTTGATTTCCTCGTTGGTGCAGGAATTAAGCCTACTTCAAT
TGTGAGCTACAATCACTTGGGGAACAACGATGGAATGAACCTCTCAGCTCCACAGACATTCCGATCTAAGGAGATC
TCCAAAAGCAATGTCGTGGACGATATGGTTGCTAGCAACGGTATCCTCTTTGAGCCTGGGGAACACCCGGACCATG
TTGTGGTCATCAAGTATGTACCGTATGTTGCTGATAGTAAGCGAGCGATGGACGAGTATACATCAGAGATTTTCATG
GGAGGGAAGAACACAATCGTGATGCACAATACCTGTGAGGACTCTCTCTTAGCTGCTCCAATCATCTTGGATCTTGT
TCTCCTTGCGGAGCTCAGCACCAGGATCCAGTTCAAATCCGAGGGAGAGGGAAAGTTTCATTCGTTCCATCCTGTTG
CCACCATACTCAGCTACCTCACCAAGGCACCGCTGGTTCCGTCGGGAACGCCGGTGGTTAACGCGCTGTCTAAGCA
GCGGGCTATGCTGGAGAACATTCTTAGGGCGTGCGTTGGACTGGCCCCGGAGAATAACATGATCTTGGAATACAA
GTGA

>*CSMIPS1* Amino Acid Sequence 511AA

MFIKDFKVESLNVKYTENEIHSVYDYETTEVVHENVNGTFQWIVKPKTVKYDFVTDTRVPKLGVMLVGLGGNNGS
TLTAGVIANKEGISWATKDKVQQANYFGSLTQASSIRVGSFNGEEIYAPFKSLLPMVNPDDVVFGGWDISDMNLA
DAMARAKVLDIDLQKQLRPFMENIVPLPGIFDSDFIAANQGSRANHVIKGTKKEQIDHIIKDMREFKEKNKVDRV
VVLWTANTERYSNVVVGLNDTMENLMQSVEKDEAEISPSTLYAIACVLEGIPFINGSPQNTFVPGLIDLAIKNNV
LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIVSYNHLGNNDGMNLSAPQTFRSKEISKSNVVDDMVASNGILFE
PGEHPDHVVVIKYVPYVADSKRAMDEYTSEIFMGGKNTIVMHNTCEDSLLAAPIILDLVLLAELSTRIQFKSEGE
GKFHSFHPVATILSYLTKAPLVPSGTPVVNALSKQRAMLENILRACVGLAPENNMILEYK*

AMPLICON1:

CATGATTACCTTCATCCTGAGATGGTTTATGTTATATTGAATTTTGGGGTTTTTGTTGTTGGCAAAACAG**GTGA**
ATCCGGATGATGTTGTGTTTGGAGGATGGGACATAAGTGATATGAACTTAGCAGATGCCATGGCTAGAGCC
AAGGTTCTTGATATCGACTTGCAGAAACAGCTGAGGCCTTTCATGGAGAACATTGTGCCACTCCCTGGGATC
TTCGATTCTGATTTCATTGCTGCCAATCAAGGCTCACGTGCTAACCACGTGATCAAAGGTACCAAGAAGGAA
CAAATCGACCACATCATCAAGGACATGAGGTAATAATAATATAGTGTGTTTAAAACAGAGTTTTTCTTGCTTG
TTGAATTAAATCTCCTCTGTTTCGTGATTTTTATTCTTGTGATGTTTTGGATCTTTAG**GGAGTTTAAGGAGAAG**
AATAAGGTGGATAGGGTTGTG

AMPLICON2:

CCGCTGATTAAGAGAAGCACTAAAACAAAGCGATATTACAAAACAAAGAAGCAAATTTTCTTTTATTTCGAAA
ATGTTCATTAAAGACTTTAAGGTTGAGAGCCTGAACGTGAAGTACACTGAAAATGAGATTCACTCTGTGTAT
GATTACGAGACGACCGAGGTTGTGCATGAGAACGTCAATGGTACTTTCCAATGGATTGTTAAGCCAAAGAC
TGTCAAATACGATTTCGTTACCGATACTCGTGTCCCCAAATTAGGGTAATTAATTACGTACTCTTTTTTTATTTT
TTTACCAAAGTTTCGATCTTCTTGATTTCTCCGTCGTTCAATTTTGATTTGTTTTTCGTTAAAG**GGTTATGCTTGT**
GGGTTTGGGAGGAAACAATGGATCAACTCTCACCGCTGGTGTTATTGCCAACAAAGAGTGAGTC

SELECTED TbyS N1 COMBINATIONS

Amplicon	Amplicon size	Primer	Sequence	Len.	Tm	GC%
1	457	407F	TTCCCTACACGACGCTCTTCCGATCTCACAACCCTATCCACCTTAT	46	70.8	50
		408R	AGTTCAGACGTGTGCTCTTCCGATCTCATGATTACCTTCATCCTGA	46	69	46
2	431	409F	TTCCCTACACGACGCTCTTCCGATCTGACTCACTCTTTGTTGGCAA	46	70.8	50
		409R	AGTTCAGACGTGTGCTCTTCCGATCTCCGCTGATTAAGAGAAGCAC	46	70.8	50

Sequence Nested 1 (Tm = 55 °C)

Per reaction of 10 µl :

1 µl buffer PCR 10 X lab (10 mM Tris-HCl pH8, 50 mM KCl, 2.5 mM $MgCl_2$, 0.05% Igepal) + 0.1 µl of primer For 10 µM + 0.1 µl of primer Rev 10 µM + 0.4 µl of dNTPs 5 mM + 0.4 µl of Taq lab + 0.5µl DMSO + 1 µl. crésol red + 4.5 µl water + 2 µl DNA D204mn 95 °C; 35 x (30s 95 °C, 30 s 55 °C and 1 mn 72 °C) ; 5 min 72°C ; 4°C

Sequence Nested 2 (Tm = 55 °C)

Per reaction of 10 µl :

- 1 µl buffer PCR 10 X lab (10 mM Tris-HCl pH8, 50 mM KCl, 2.5 mM $MgCl_2$, 0.05% Igepal**)** + 2 µl of primer For 0.5 µM and primer Rev 0.5 µM (TOGETHER IN ONE MIX) + 0.4 µl of dNTPs 5 mM + 0.4 µl. of Taq lab + 4.2 µl water + 2 µl DNA of PCR1
- 5 mn 94 °C; **35 x** (30s 94 °C, 30 s 55 °C and 1 mn 72 °C) ; 5 min 72°C ; 4°C

A1.6 TbyS TILLING PRIMERS FOR *TGG1*

Plant: Camelina sativa

Target: *TGG1*

Mutation detection system:

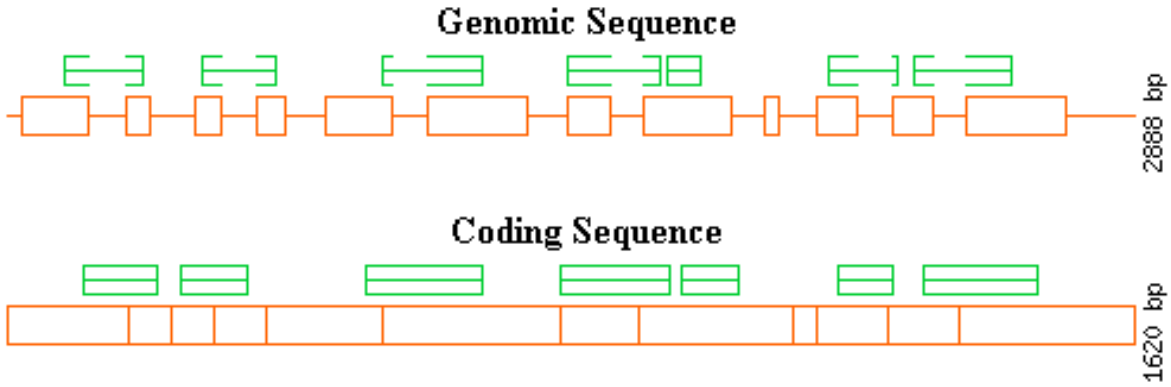
- ✓ Miseq Desktop sequencer (Illumina®)
- ✓ MiSeq Reagent Kit v3 (2 x 300 bp)
- ✓ Pooling: 2-D, full collection, **pool A CSC2** (**2496** families)

Exon Intron Missense Mutation Silent Mutation STOP Mutation

EXON POSITIONS : TWELVE EXON (36..209, 307..367, 484..545, 639..714, 817..982, 1076..1331, 1433..1545, 1632..1852, 1942..1976, 2075..2174, 2270..2372, 2459..2711)

REGION OF NGS AMPLICON 1 – 1477 TO 1880

REGION OF NGS AMPLICON 2 – 1061 TO 1470



>*TGG1* – FORWARD ORIENTATION

ATGCAACACAAAACATAACAAACTACAAATTAACC**ATGAAGTTTCATTGGTTCGCCTTAGCTTTCTTATTAGCT**
GTGGCGACTTGTAAAGGCCAGGAAGATTATACTTGCGAAGAGAACGAGCCATTCCATTGTAACCAAACTAG
TCGTTTCAATGGTAAAAGTTTCGGCGAAGATTTCATCTTCGGTGTAGCCTCCTCTGCTTACCAAGTATGTAAC
AAATCACACATTATGTTGACGTCTTGAAAGAGCATTAAAAAACATAAAAAATAATCCAACGTTGCATGGTTGTT
TCTCTTCATTTCAG**ATCGAAGGTGGTAGAGGTCGTGGACTTAACATTTGGGATGGATTCACTCACCGATACC**
CAGGTTTGTTATTTGAATGGTCTCAAACTACTATAGCTTATATAGCTATATATACGAGAAGAATGTTGTCCATA
TGCATGTTTATATATATTCATAGTGGTTCATTATTACGCTATCAG**AGAAAGCAGGAGCCGATTTGAAGAATGG**
AGACACTACTTGTGACGCATATACATATTGGCAGGTTAGTTCCTCATTATATATGCTTGCTAATGAATTATCTA
TATTAACTATACATGTGTGGATTGACATGGGATTGTATACATATAAACTGCAG**AAAGATATAGACGTGATGG**
GCGAACTCAATGCTACTGGCTACAGATTCTCCTTCGCGTGGTCAAGAATCCTTCCACGTATGTATACATTATT
GATCCATGTATATATATATATATATATAATATATAATATTCGGTTTTAGTTGTTTGGAAGAGTAAATATATGATA
TATATGAGCAG**AAGGAAAGAGGAGTAGAGGTGTGAACCAAAATGGTATTGACTACTACAACGGTCTTATA**
GACGGCCTCATAGCAAGGAATATAACGCCGTTCGTTACCCTCTTTCACTGGGACCTTCCTCAAACACTACAAG
ATGAGTATGAAGGTTTCTTGAACAGAACGATCATGTATGTCATTTATGCTACGCTGCTTATTAAGAATGACGA
AAGCATATAACGATATGATCGTACATTAATGAAATCTAACTATAACACGTGCAG**AGATG**ATTTCAAAGATTAC

GCGGATCTATGTTTCGAAAAATTTGGTGATAGGGTAAAGCACTGGATAACAATCAACCAGCTTTACACAGT
ACCTACGCGAGGATATGCAATTGGAACAGATGCACCTGGTCGATGTTCTCTTACGCATGATATAAGATGTTA
CGGCGGAAACTCGTCAACCGAACCCTATATCGTTGCACATAACCAGCTTCTTGCTCATGCCACGGTCGTGGA
TCTTTACAGGACAAAATATAAGGTGAGCTAGGTCTCCATTTAATTTGTAGAAGGTGTAGTCTAACGTAACATG
CATGTGATTCTAATTATATTATCATTGTGTATATATTTGCGTAACCACAG**GACCAAGGAGGGATGATTGGACC**
TGTGATGATAACTAGAGGGTT**TCTTCCGTTTGATGACACTC**CAGAGAGCAAAGAGGCAACTTATCGGTCTAA
AGAATTTTTCCATGGATGGTATACATATCCACGAATGAAATTGTTTATACAATATTGTATATTAGTATATATAA
AAAGTATTGAAAAGTATGCATGTATATCAG**GTTCATGGAGCCGCTAACAAAGGGTAAATACCCAGACATCAT**
GAGGAAACTTGTGGGTGAACGGCTTCCAGAGTTCACCGAGACAGAAGCCAAACTTGTAAAGGGTTCATATG
ATTTTCTTGGTCTCAACTATTACGTCACTCAGTATGCCCAAAACGATGACACAGTAGTTCCTTGGGCCAACCA
CACTGCTATGATGGACCCAAAAGCAATTCTCACATGTACCAATCTATTCCCACACTTAGTTACAACATTTCTAC
ATTTTTTCAGTTACAACATTTTGTGACTATCAATTATATGTACTTTGCAG**ATGAAAATGCAAAGGGTGAGCCCA**
TTGGTCCAATGGTATGCATGCATGGTCCACCCCTATTCATTCTTTTTTAATGGAAAATAATACCCATGTGCACTT
TTCTGATTTTTTTAATATATATTCGCTTTCTCAG**TTCAGTAAAGGAGCCTATTACTACCCAAAAGGCATTTACG**
ACGTAATGGAGTATTACAAAAACAAATATGGTGACCCTTTAATATATATCACCGAGAACGGTTAGTTCTACT
CTATTACACATTTATACAAGTTAAATGTATTTCAGAGTTGGTCCAGGAAATTTACATGAAGATTGTGTACTTGT
ACATATTAG**GAATTAGTAGCCCCGGTGATCAACCCCTTGAAGAGGCTATGGCCGATTACAAGAGGATTGATT**
ATCTATGCAGTCATCTTTGTTTTCTCCGTAAGGTCATTAAGTGAGTGGATACTACTCTTCTTCCATTTTTCTTTA
CATATGATCAGTCGTACGCCTTTAATAATTAATGGTTGTTCTTGTTTCTAG**GGAGAAGGGTGTCAACGTGAGA**
GGATACTTTGCTTGGTCTCTTGGGGATAATTACGAATTCTGCAATGGCTTTACCGTCAGATTTGGACTCAGTT
ACGTTGACTTCAACAATGTCACCGCTGATAGAGACCTCAAAGCATCTGGCAAATGGTTCCAGCAGTTCATTA
GCTCCAATGACCCTGCCGACCAAGATCTCCTCCGCTCTAGCCTCTCCTTCAAGAACCGTGATCGGAAGAGGCT
CGCAGATGCATGATAGATTTCTAACCACTTTATGCTCATCAAAATCATCTTCGTGTCTCTTCTTTCTACTTGCTCC
ATAGATAAAGGAGCTTCTTGTATTTGAATAAAAGATGATTAATGTCAATAAAGACTTTTGCTTACGACATTATA
TTATGTTGGAAGGGAAAGAACCCAGCAAACCCAATCTGCA

>*TGG1*-Chrom 20 – **ORIGINAL SEQUENCE REVERSE ORIENTATION**

TGCAGATTGGGTTTGCTGGGTTCTTTCCCTTCCAACATAATATAATGTCGTAAGCAAAAGTCTTTATTGACATTA
ATCATCTTTTATTCAAATACAAGAAGCTCCTTTATCTATGGAGCAAGTAGAAAGAAGAGACACGAAGATGATT
TTGATGAGCATAAAGTGGTTAGAAATCTATCATGCATCTGCGAGCCTCTTCCGATCACGGTTCTTGAAGGAGA
GGCTAGAGCGGAGGAGATCTTGGTCGGCAGGGTCATTGGAGCTAATGAACTGCTGGAACCATTTGCCAGATG
CTTTGAGGTCTCTATCAGCGGTGACATTGTTGAAGTCAACGTAACTGAGTCCAAATCTGACGGTAAAGCCATT
GCAGAATTCGTAATTATCCCCAAGAGACCAAGCAAAGTATCCTCTCACGTTGACACCCTTCTCCCTAGAAACAA
GAACAACCATTAATTATTAAAGGCGTACGACTGATCATATGTAAAGAAAAATGGAAGAAGAGTAGTATCCACT
CACTTAATGACCTTACGGAGAAAACAAAGATGACTGCATAGATAATCAATCCTCTTGTAATCGGCCATAGCCTC
TTCAAGGGGTTGATCACCGGGGCTACTAATTCCTAATATGTACAAGTACACAATCTTCATGTAAATTTCCTGGA
CCAACTCTGAAATACATTTAACTTGTATAAATGTGTAATAGAGTAGAACTAACCGTTCTCGGTGATATATATTA
AAGGGTCACCATATTTGTTTTTGTAATACTCCATTACGTCGTAAATGCCTTTTGGGTAGTAATAGGCTCCTTTAC
TGAACTGAGAAAGCGAATATATATTAAAAAAAATCAGAAAAGTGCACATGGGTATTATTTTCCATTAAAAAAGA
ATGAATAGGGGTGGACCATGCATGCATACCATTGGACCAATGGGCTCACCCTTTGCATTTTCATCTGCAAAGT
ACATATAATTGATAGTCACAAAATGTTGTAACTGAAAAAATGTAGAAATGTTGTAACTAAGTGTGGGAATAGA
TTGGTACATGTGAGAATTGCTTTTGGGTCCATCATAGCAGTGTGGTTGGCCCAAGGAACTACTGTGTCATCGT
TTTGGGCATACTGAGTGACGTAATAGTTGAGACCAAGAAAATCATATGAACCCTTTACAAGTTTGGCTTCTGTC
TCGGTGAACTCTGGAAGCCGTTCACCCACAAGTTTCCTCATGATGTCTGGGTATTTACCCTTTGTTAGCGGCTC
CATGAACCTGATATACATGCATACTTTTCAATACTTTTTATATATACTAATATACAATATTGTATAAACAATTTCA
TTCGTGGATATGTATACCATCCATGGAAAAATTCTTTAGACCGATAAGTTGCCTCTTTGCTCTCTGGAGTGTCA
TCAAACGGAAGAAACCCTCTAGTTATCATCACAGGTCCAATCATCCCTCCTTGGTCCTGTGGTTACGCAAATAT
ATACACAATGATAATATAATTAGAATCACATGCATGTTACGTTAGACTACACCTTCTACAAATTAAATGGAGAC
CTAGCTCACCTTATATTTTGTCCTGTAAAGATCCACGACCGTGGCATGAGCAAGAAGCTGGTTATGTGCAACG
ATATAGGGTTCGGTTGACGAGTTTCCGCCGTAACATCTTATATCATGCGTAAGAGAACATCGACCAGGTGCAT
CTGTTCCAATTGCATATCCTCGCGTAGGTACTGTGTAAAGCTGGTTGATTGTTATCCAGTGCTTTACCCTATCAC
CAAATTTTTCGAAACATAGATCCGCGTAATCTTTGAAATCATCTCTGCACGTGTTATAGTTAGATTTCATTAATG
TACGATCATATCGTTATATGCTTTCGTCATTCTTAATAAGCAGCGTAGCATAAATGACATACATGATCGTTCTGT
TCAAGAAACCTTCATACTCATCTTGTAGTGTTTGAGGAAGGTCCCAGTGAAAGAGGGTAACGAACGGCGTTAT
ATTCCTTGCTATGAGGCCGTCTATAAGACCGTTGTAGTAGTCAATACCATTTTGGTTCACACCTCTACTCCTCTT
TCCTTCTGCTCATATATATCATATATTTACTCTTCCAAACAACTAAAACCGAATATTATATATTATATATATATAT
ATATATACATGGATCAATAATGTATACATACGTGGAAGGATTCTTGACCACGCGAAGGAGAATCTGTAGCCAG
TAGCATTGAGTTCGCCCATCACGTCTATATCTTTCTGCAGTTTATATGTATACAATCCCATGTCAATCCACACAT
GTATAGTTAATATAGATAATTCATTAGCAAGCATATATAATGAGGAACTAACCTGCCAATATGTATATGCGTCA
CAAGTAGTGTCTCCATTCTTCAAATCGGCTCCTGCTTTCTCTGATAGCGTAATAATGAACCACTATGAATATATA
TAAACATGCATATGGACAACATTCTTCTCGTATATATAGCTATATAAGCTATAGTAGTTTGAGACCATTCAAAT
AACAAACCTGGGTATCGGTGAGTGAATCCATCCCAAATGTTAAGTCCACGACCTCTACCACCTTCGATCTGAAA
TGAAGAGAAACAACCATGCAACGTTGGATTATTTTTTATGTTTTTTAATGCTCTTTCAAGACGTCAACATAATGT
GTGATTTGTTACATACTTGGTAAGCAGAGGAGGCTACACCGAAGATGAAATCTTCGCCGAAACTTTTACCATT
GAAACGACTAGTTTGGTTACAATGGAATGGCTCGTTCTCTTCGCAAGTATAATCTTCCTGGCCTTTACAAGTCG
CCACAGCTAATAAGAAAGCTAAGGCGAACCAATGAAACTTCATGGTTAATTTGTAGTTTGTTATGTTTTGTGTT
GCAT

>Csa20g050340.1-*TGG1* cds
ATGAAGTTTCATTGGTTCGCCTTAGCTTTCTTATTAGCTGTGGCGACTTGTAAAGGCCAGGAAGATTATAC
TTGCGAAGAGAACGAGCCATTCCATTGTAACCAAACTAGTCGTTTCAATGGTAAAAGTTTCGGCGAAGATT
TCATCTTCGGTGTAGCCTCCTCTGCTTACCAAATCGAAGGTGGTAGAGGTCGTGGACTTAACATTTGGGAT
GGATTCACTCACCGATACCCAGAGAAAGCAGGAGCCGATTTGAAGAATGGAGACACTACTTGTGACGCATA
TACATATTGGCAGAAAGATATAGACGTGATGGGCGAACTCAATGCTACTGGCTACAGATTCTCCTTCGCGT
GGTCAAGAATCCTTCCACAAGGAAAGAGGAGTAGAGGTGTGAACCAAAATGGTATTGACTACTACAACGGT
CTTATAGACGGCCTCATAGCAAGGAATATAACGCCGTTCGTTACCCTCTTTCACTGGGACCTTCCTCAAAC
ACTACAAGATGAGTATGAAGGTTTCTTGAACAGAACGATCATAGATGATTTCAAAGATTACGCGGATCTAT
GTTTCGAAAAATTTGGTGATAGGGTAAAGCACTGGATAACAATCAACCAGCTTTACACAGTACCTACGCGA
GGATATGCAATTGGAACAGATGCACCTGGTCGATGTTCTCTTACGCATGATATAAGATGTTACGGCGGAAA
CTCGTCAACCGAACCCTATATCGTTGCACATAACCAGCTTCTTGCTCATGCCACGGTCGTGGATCTTTACA
GGACAAAATATAAGGACCAAGGAGGGATGATTGGACCTGTGATGATAACTAGAGGGTTTCTTCCGTTTGAT
GACACTCCAGAGAGCAAAGAGGCAACTTATCGGTCTAAAGAATTTTTCCATGGATGGTTCATGGAGCCGCT
AACAAAGGGTAAATACCCAGACATCATGAGGAAACTTGTGGGTGAACGGCTTCCAGAGTTCACCGAGACAG
AAGCCAAACTTGTAAAGGGTTCATATGATTTTCTTGGTCTCAACTATTACGTCACTCAGTATGCCCAAAAC
GATGACACAGTAGTTCCTTGGGCCAACCACACTGCTATGATGGACCCAAAAGCAATTCTCACATATGAAAA
TGCAAAGGGTGAGCCCATTGGTCCAATGTTCAGTAAAGGAGCCTATTACTACCCAAAAGGCATTTACGACG
TAATGGAGTATTACAAAAACAAATATGGTGACCCTTTAATATATATCACCGAGAACGGAATTAGTAGCCCC
GGTGATCAACCCCTTGAAGAGGCTATGGCCGATTACAAGAGGATTGATTATCTATGCAGTCATCTTTGTTT
TCTCCGTAAGGTCATTAAGGAGAAGGGTGTCAACGTGAGAGGATACTTTGCTTGGTCTCTTGGGGATAATT
ACGAATTCTGCAATGGCTTTACCGTCAGATTTGGACTCAGTTACGTTGACTTCAACAATGTCACCGCTGAT
AGAGACCTCAAAGCATCTGGCAAATGGTTCCAGCAGTTCATTAGCTCCAATGACCCTGCCGACCAAGATCT
CCTCCGCTCTAGCCTCTCCTTCAAGAACCGTGATCGGAAGAGGCTCGCAGATGCATGA

> *TGG1* Protein Sequence
MKFHWFALAFLLAVATCKGQEDYTCEENEPFHCNQTSRFNGKSFGEDFIFGVASSAYQIE
GGRGRGLNIWDGFTHRYPEKAGADLKNGDTTCDAYTYWQKDIDVMGELNATGYRFSFAWS
RILPQGKRSRGVNQNGIDYYNGLIDGLIARNITPFVTLFHWDLPQTLQDEYEGFLNRTII
DDFKDYADLCFEKFGDRVKHWITINQLYTVPTRGYAIGTDAPGRCSLTHDIRCYGGNSST
EPYIVAHNQLLAHATVVDLYRTKYKDQGGMIGPVMITRGFLPFDDTPESKEATYRSKEFF
HGWFMEPLTKGKYPDIMRKLVGERLPEFTETEAKLVKGSYDFLGLNYYVTQYAQNDDTVV
PWANHTAMMDPKAILTYENAKGEPIGPMFSKGAYYYPKGIYDVMEYYKNKYGDPLIYITE
NGISSPGDQPLEEAMADYKRIDYLCSHLCFLRKVIKEKGVNVRGYFAWSLGDNYEFCNGF
TVRFGLSYVDFNNVTADRDLKASGKWFQQFISSNDPADQDLLRSSLSFKNRDRKRLADA*

AMPLICON1:

TCTTCCGTTTGATGACACTCCAGAGAGCAAAGAGGCAACTTATCGGTCTAAAGAATTTTTCCATGGATGGTATACATATC
CACGAATGAAATTGTTTATACAATATTGTATATTAGTATATATAAAAAGTATTGAAAAGTATGCATGTATATCAG**GTTCAT**
GGAGCCGCTAACAAAGGGTAAATACCCAGACATCATGAGGAAACTTGTGGGTGAACGGCTTCCAGAGTTCACCGAGA
CAGAAGCCAAACTTGTAAAGGGTTCATATGATTTTCTTGGTCTCAACTATTACGTCACTCAGTATGCCCAAAACGATGAC
ACAGTAGTTCCTTGGGCCAACCACACTGCTATGATGGACCCAAAAGCAATTCTCACATGTACCAATCTATTCCCACACTTA
GTTAC

AMPLICON2:

CTATAACACGTGCAG**AGATG**ATTTCAAAGATTACGCGGATCTATGTTTCGAAAAATTTGGTGATAGGGTAAAGCACTGG
ATAACAATCAACCAGCTTTACACAGTACCTACGCGAGGATATGCAATTGGAACAGATGCACCTGGTCGATGTTCTCTTA
CGCATGATATAAGATGTTACGGCGGAAACTCGTCAACCGAACCCTATATCGTTGCACATAACCAGCTTCTTGCTCATGCC
ACGGTCGTGGATCTTTACAGGACAAAATATAAGGTGAGCTAGGTCTCCATTTAATTTGTAGAAGGTGTAGTCTAACGTAA
CATGCATGTGATTCTAATTATATTATCATTGTGTATATATTTGCGTAACCACAG**GACCAAGGAGGGATGATTGGACCTGT**
GATGATAACTAG

SELECTED TbyS N1 COMBINATIONS

Amplicon	Amplicon size (bp)	Primer	Sequence	Len.	Tm	GC%
1	404	411F	TTCCCTACACGACGCTCTTCCGATCTGTAACTAAGTGTGGGAATAG	46	69.9	48
		411R	AGTTCAGACGTGTGCTCTTCCGATCTTCTTCCGTTTGATGACACTC	46	69.9	48
2	410	414F	TTCCCTACACGACGCTCTTCCGATCTCTAGTTATCATCACAGGTCC	46	70.8	50
		413R	AGTTCAGACGTGTGCTCTTCCGATCTCTATAACACGTGCAGAGATG	46	69.9	48

Sequence Nested 1 (Tm = 55 °C)

Per reaction of 10 µl :

1 µl buffer PCR 10 X lab (10 mM Tris-HCl pH8, 50 mM KCl, 2.5 mM $MgCl_2$, 0.05% Igepal) + 0.1 µl of primer For 10 µM + 0.1 µl of primer Rev 10 µM + 0.4 µl of dNTPs 5 mM + 0.4 µl of Taq lab + 0.5µl DMSO + 1 µl. crésol red + 4.5 µl water + 2 µl DNA D204mn 95 °C; 35 x (30s 95 °C, 30 s 55 °C and 1 mn 72 °C) ; 5 min 72°C ; 4°C

Sequence Nested 2 (Tm = 55 °C)

Per reaction of 10 µl :

- 1 µl buffer PCR 10 X lab (10 mM Tris-HCl pH8, 50 mM KCl, 2.5 mM $MgCl_2$, 0.05% Igepal**)** + 2 µl of primer For 0.5 µM and primer Rev 0.5 µM (TOGETHER IN ONE MIX) + 0.4 µl of dNTPs 5 mM + 0.4 µl. of Taq lab + 4.2 µl water + 2 µl DNA of PCR1
- 5 mn 94 °C; **35 x** (30s 94 °C, 30 s 55 °C and 1 mn 72 °C) ; 5 min 72°C ; 4°C

A2. CAMELINA TbyS 2D2 POOLING STRATEGY

A2.1 Plate format of 96 well pools of 16-fold column pool and 12-fold line pool

poolC1A	1	2	3	4	5	6	7	8	9	10	11	12
A	1_2-C1	1_2-C2	1_2-C3	1_2-C4	1_2-C5	1_2-C6	1_2-C7	1_2-C8	1_2-C9	1_2-C10	1_2-C11	1_2-C12
B	3_4-C1	3_4-C2	3_4-C3	3_4-C4	3_4-C5	3_4-C6	3_4-C7	3_4-C8	3_4-C9	3_4-C10	3_4-C11	3_4-C12
C	5_6-C1	5_6-C2	5_6-C3	5_6-C4	5_6-C5	5_6-C6	5_6-C7	5_6-C8	5_6-C9	5_6-C10	5_6-C11	5_6-C12
D	7_8-C1	7_8-C2	7_8-C3	7_8-C4	7_8-C5	7_8-C6	7_8-C7	7_8-C8	7_8-C9	7_8-C10	7_8-C11	7_8-C12
E	9_10-C1	9_10-C2	9_10-C3	9_10-C4	9_10-C5	9_10-C6	9_10-C7	9_10-C8	9_10-C9	9_10-C10	9_10-C11	9_10-C12
F	11_12-C1	11_12-C2	11_12-C3	11_12-C4	11_12-C5	11_12-C6	11_12-C7	11_12-C8	11_12-C9	11_12-C10	11_12-C11	11_12-C12
G	13_14-C1	13_14-C2	13_14-C3	13_14-C4	13_14-C5	13_14-C6	13_14-C7	13_14-C8	13_14-C9	13_14-C10	13_14-C11	13_14-C12
H	17_18-C1	17_18-C2	17_18-C3	17_18-C4	17_18-C5	17_18-C6	17_18-C7	17_18-C8	17_18-C9	17_18-C10	17_18-C11	17_18-C12

poolL1A	1	2	3	4	5	6	7	8	9	10	11	12
A	1-LA	2-LA	3-LA	4-LA	5-LA	6-LA	7-LA	8-LA	9-LA	10-LA	11-LA	12-LA
B	1-LB	2-LB	3-LB	4-LB	5-LB	6-LB	7-LB	8-LB	9-LB	10-LB	11-LB	12-LB
C	1-LC	2-LC	3-LC	4-LC	5-LC	6-LC	7-LC	8-LC	9-LC	10-LC	11-LC	12-LC
D	1-LD	2-LD	3-LD	4-LD	5-LD	6-LD	7-LD	8-LD	9-LD	10-LD	11-LD	12-LD
E	1-LE	2-LE	3-LE	4-LE	5-LE	6-LE	7-LE	8-LE	9-LE	10-LE	11-LE	12-LE
F	1-LF	2-LF	3-LF	4-LF	5-LF	6-LF	7-LF	8-LF	9-LF	10-LF	11-LF	12-LF
G	1-LG	2-LG	3-LG	4-LG	5-LG	6-LG	7-LG	8-LG	9-LG	10-LG	11-LG	12-LG
H	1-LH	2-LH	3-LH	4-LH	5-LH	6-LH	7-LH	8-LH	9-LH	10-LH	11-LH	12-LH

poolC2A	1	2	3	4	5	6	7	8	9	10	11	12
A	19_20-C1	19_20-C2	19_20-C3	19_20-C4	19_20-C5	19_20-C6	19_20-C7	19_20-C8	19_20-C9	19_20-C10	19_20-C11	19_20-C12
B	21_22-C1	21_22-C2	21_22-C3	21_22-C4	21_22-C5	21_22-C6	21_22-C7	21_22-C8	21_22-C9	21_22-C10	21_22-C11	21_22-C12
C	23_24-C1	23_24-C2	23_24-C3	23_24-C4	23_24-C5	23_24-C6	23_24-C7	23_24-C8	23_24-C9	23_24-C10	23_24-C11	23_24-C12
D	25_26-C1	25_26-C2	25_26-C3	25_26-C4	25_26-C5	25_26-C6	25_26-C7	25_26-C8	25_26-C9	25_26-C10	25_26-C11	25_26-C12
E	42_43-C1	42_43-C2	42_43-C3	42_43-C4	42_43-C5	42_43-C6	42_43-C7	42_43-C8	42_43-C9	42_43-C10	42_43-C11	42_43-C12
F												
G												
H												

poolL2A	1	2	3	4	5	6	7	8	9	10	11	12
A	19-LA	20-LA	21-LA	22-LA	23-LA	24-LA	25-LA	26-LA	42-LA	43-LA	13-LA	14-LA
B	19-LB	20-LB	21-LB	22-LB	23-LB	24-LB	25-LB	26-LB	42-LB	43-LB	13-LB	14-LB
C	19-LC	20-LC	21-LC	22-LC	23-LC	24-LC	25-LC	26-LC	42-LC	43-LC	13-LC	14-LC
D	19-LD	20-LD	21-LD	22-LD	23-LD	24-LD	25-LD	26-LD	42-LD	43-LD	13-LD	14-LD
E	19-LE	20-LE	21-LE	22-LE	23-LE	24-LE	25-LE	26-LE	42-LE	43-LE	13-LE	14-LE
F	19-LF	20-LF	21-LF	22-LF	23-LF	24-LF	25-LF	26-LF	42-LF	43-LF	13-LF	14-LF
G	19-LG	20-LG	21-LG	22-LG	23-LG	24-LG	25-LG	26-LG	42-LG	43-LG	13-LG	14LG
H	19-LH	20-LH	21-LH	22-LH	23-LH	24-LH	25-LH	26-LH	42-LH	43-LH	13-LH	14-LH

poolL3A	1	2	3	4	5	6	7	8	9	10	11	12
A	17-LA	18-LA										
B	17-LB	18-LB										
C	17-LC	18-LC										
D	17-LD	18-LD										
E	17-LE	18-LE										
F	17-LF	18-LF										
G	17-LG	18-LG										
H	17-LH	18-LH										

poolC2A	1	2	3	4	5	6	7	8	9	10	11	12
A	19_20-C1	19_20-C2	19_20-C3	19_20-C4	19_20-C5	19_20-C6	19_20-C7	19_20-C8	19_20-C9	19_20-C10	19_20-C11	19_20-C12
B	21_22-C1	21_22-C2	21_22-C3	21_22-C4	21_22-C5	21_22-C6	21_22-C7	21_22-C8	21_22-C9	21_22-C10	21_22-C11	21_22-C12
C	23_24-C1	23_24-C2	23_24-C3	23_24-C4	23_24-C5	23_24-C6	23_24-C7	23_24-C8	23_24-C9	23_24-C10	23_24-C11	23_24-C12
D	25_26-C1	25_26-C2	25_26-C3	25_26-C4	25_26-C5	25_26-C6	25_26-C7	25_26-C8	25_26-C9	25_26-C10	25_26-C11	25_26-C12
E	42_43-C1	42_43-C2	42_43-C3	42_43-C4	42_43-C5	42_43-C6	42_43-C7	42_43-C8	42_43-C9	42_43-C10	42_43-C11	42_43-C12
F	17-LA	17-LB	17-LC	17-LD	17-LE	17-LF	17-LG	17-LH				
G	18-LA	18-LB	18-LC	18-LD	18-LE	18-LF	18-LG	18-LH				
H												

Transfer of the DNA from 96 wells plates to 384 wells plate with the BenchTop Pipettor to you obtain the DNA TbyS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A	1_2-C1	1-LA	1_2-C2	2-LA	1_2-C3	3-LA	1_2-C4	4-LA	1_2-C5	5-LA	1_2-C6	6-LA	1_2-C7	7-LA	1_2-C8	8-LA	1_2-C9	9-LA	1_2-C10	10-LA	1_2-C11	11-LA	1_2-C12	12-LA
B	19_20-C1	19-LA	19_20-C2	20-LA	19_20-C3	21-LA	19_20-C4	22-LA	19_20-C5	23-LA	19_20-C6	24-LA	19_20-C7	25-LA	19_20-C8	26-LA	19_20-C9	42-LA	19_20-C10	43-LA	19_20-C11	13-LA	19_20-C12	14-LA
C	3_4-C1	1-LB	3_4-C2	2-LB	3_4-C3	3-LB	3_4-C4	4-LB	3_4-C5	5-LB	3_4-C6	6-LB	3_4-C7	7-LB	3_4-C8	8-LB	3_4-C9	9-LB	3_4-C10	10-LB	3_4-C11	11-LB	3_4-C12	12-LB
D	21_22-C1	19-LB	21_22-C2	20-LB	21_22-C3	21-LB	21_22-C4	22-LB	21_22-C5	23-LB	21_22-C6	24-LB	21_22-C7	25-LB	21_22-C8	26-LB	21_22-C9	42-LB	21_22-C10	43-LB	21_22-C11	13-LB	21_22-C12	14-LB
E	5_6-C1	1-LC	5_6-C2	2-LC	5_6-C3	3-LC	5_6-C4	4-LC	5_6-C5	5-LC	5_6-C6	6-LC	5_6-C7	7-LC	5_6-C8	8-LC	5_6-C9	9-LC	5_6-C10	10-LC	5_6-C11	11-LC	5_6-C12	12-LC
F	23_24-C1	19-LC	23_24-C2	20-LC	23_24-C3	21-LC	23_24-C4	22-LC	23_24-C5	23-LC	23_24-C6	24-LC	23_24-C7	25-LC	23_24-C8	26-LC	23_24-C9	42-LC	23_24-C10	43-LC	23_24-C11	13-LC	23_24-C12	14-LC
G	7_8-C1	1-LD	7_8-C2	2-LD	7_8-C3	3-LD	7_8-C4	4-LD	7_8-C5	5-LD	7_8-C6	6-LD	7_8-C7	7-LD	7_8-C8	8-LD	7_8-C9	9-LD	7_8-C10	10-LD	7_8-C11	11-LD	7_8-C12	12-LD
H	25_26-C1	19-LD	25_26-C2	20-LD	25_26-C3	21-LD	25_26-C4	22-LD	25_26-C5	23-LD	25_26-C6	24-LD	25_26-C7	25-LD	25_26-C8	26-LD	25_26-C9	42-LD	25_26-C10	43-LD	25_26-C11	13-LD	25_26-C12	14-LD
I	9_10-C1	1-LE	9_10-C2	2-LE	9_10-C3	3-LE	9_10-C4	4-LE	9_10-C5	5-LE	9_10-C6	6-LE	9_10-C7	7-LE	9_10-C8	8-LE	9_10-C9	9-LE	9_10-C10	10-LE	9_10-C11	11-LE	9_10-C12	12-LE
J	42_43-C1	19-LE	42_43-C2	20-LE	42_43-C3	21-LE	42_43-C4	22-LE	42_43-C5	23-LE	42_43-C6	24-LE	42_43-C7	25-LE	42_43-C8	26-LE	42_43-C9	42-LE	42_43-C10	43-LE	42_43-C11	13-LE	42_43-C12	14-LE
K	11_12-C1	1-LF	11_12-C2	2-LF	11_12-C3	3-LF	11_12-C4	4-LF	11_12-C5	5-LF	11_12-C6	6-LF	11_12-C7	7-LF	11_12-C8	8-LF	11_12-C9	9-LF	11_12-C10	10-LF	11_12-C11	11-LF	11_12-C12	12-LF
L	17-LA	19-LF	17-LB	20-LF	17-LC	21-LF	17-LD	22-LF	17-LE	23-LF	17-LF	24-LF	17-LG	25-LF	17-LH	26-LF		42-LF		43-LF		13-LF		14-LF
M	13_14-C1	1-LG	13_14-C2	2-LG	13_14-C3	3-LG	13_14-C4	4-LG	13_14-C5	5-LG	13_14-C6	6-LG	13_14-C7	7-LG	13_14-C8	8-LG	13_14-C9	9-LG	13_14-C10	10-LG	13_14-C11	11-LG	13_14-C12	12-LG
N	18-LA	19-LG	18-LB	20-LG	18-LC	21-LG	18-LD	22-LG	18-LE	23-LG	18-LF	24-LG	18-LG	25-LG	18-LH	26-LG		42-LG		43-LG		13-LG		14LG
O	17_18-C1	1-LH	17_18-C2	2-LH	17_18-C3	3-LH	17_18-C4	4-LH	17_18-C5	5-LH	17_18-C6	6-LH	17_18-C7	7-LH	17_18-C8	8-LH	17_18-C9	9-LH	17_18-C10	10-LH	17_18-C11	11-LH	17_18-C12	12-LH
P		19-LH		20-LH		21-LH		22-LH		23-LH		24-LH		25-LH		26-LH		42-LH		43-LH		13-LH		14-LH

Table of 384 well plate for TbyS Miseq primer distribution

1	2
C1	L1
C2	L2
3	4

Illumina TbyS N2 primer distribution

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A	F1 / R1	F1 / R2	F1 / R3	F1 / R4	F1 / R5	F1 / R6	F1 / R7	F1 / R8	F1 / R9	F1 / R10	F1 / R11	F1 / R12	F1 / R13	F1 / R14	F1 / R15	F1 / R16	F1 / R17	F1 / R18	F1 / R19	F1 / R20	F1 / R21	F1 / R22	F1 / R23	F1 / R24
B	F2 / R1	F2 / R2	F2 / R3	F2 / R4	F2 / R5	F2 / R6	F2 / R7	F2 / R8	F2 / R9	F2 / R10	F2 / R11	F2 / R12	F2 / R13	F2 / R14	F2 / R15	F2 / R16	F2 / R17	F2 / R18	F2 / R19	F2 / R20	F2 / R21	F2 / R22	F2 / R23	F2 / R24
C	F3 / R1	F3 / R2	F3 / R3	F3 / R4	F3 / R5	F3 / R6	F3 / R7	F3 / R8	F3 / R9	F3 / R10	F3 / R11	F3 / R12	F3 / R13	F3 / R14	F3 / R15	F3 / R16	F3 / R17	F3 / R18	F3 / R19	F3 / R20	F3 / R21	F3 / R22	F3 / R23	F3 / R24
D	F4 / R1	F4 / R2	F4 / R3	F4 / R4	F4 / R5	F4 / R6	F4 / R7	F4 / R8	F4 / R9	F4 / R10	F4 / R11	F4 / R12	F4 / R13	F4 / R14	F4 / R15	F4 / R16	F4 / R17	F4 / R18	F4 / R19	F4 / R20	F4 / R21	F4 / R22	F4 / R23	F4 / R24
E	F5 / R1	F5 / R2	F5 / R3	F5 / R4	F5 / R5	F5 / R6	F5 / R7	F5 / R8	F5 / R9	F5 / R10	F5 / R11	F5 / R12	F5 / R13	F5 / R14	F5 / R15	F5 / R16	F5 / R17	F5 / R18	F5 / R19	F5 / R20	F5 / R21	F5 / R22	F5 / R23	F5 / R24
F	F6 / R1	F6 / R2	F6 / R3	F6 / R4	F6 / R5	F6 / R6	F6 / R7	F6 / R8	F6 / R9	F6 / R10	F6 / R11	F6 / R12	F6 / R13	F6 / R14	F6 / R15	F6 / R16	F6 / R17	F6 / R18	F6 / R19	F6 / R20	F6 / R21	F6 / R22	F6 / R23	F6 / R24
G	F7 / R1	F7 / R2	F7 / R3	F7 / R4	F7 / R5	F7 / R6	F7 / R7	F7 / R8	F7 / R9	F7 / R10	F7 / R11	F7 / R12	F7 / R13	F7 / R14	F7 / R15	F7 / R16	F7 / R17	F7 / R18	F7 / R19	F7 / R20	F7 / R21	F7 / R22	F7 / R23	F7 / R24
H	F8 / R1	F8 / R2	F8 / R3	F8 / R4	F8 / R5	F8 / R6	F8 / R7	F8 / R8	F8 / R9	F8 / R10	F8 / R11	F8 / R12	F8 / R13	F8 / R14	F8 / R15	F8 / R16	F8 / R17	F8 / R18	F8 / R19	F8 / R20	F8 / R21	F8 / R22	F8 / R23	F8 / R24
I	F9 / R1	F9 / R2	F9 / R3	F9 / R4	F9 / R5	F9 / R6	F9 / R7	F9 / R8	F9 / R9	F9 / R10	F9 / R11	F9 / R12	F9 / R13	F9 / R14	F9 / R15	F9 / R16	F9 / R17	F9 / R18	F9 / R19	F9 / R20	F9 / R21	F9 / R22	F9 / R23	F9 / R24
J	F10 / R1	F10 / R2	F10 / R3	F10 / R4	F10 / R5	F10 / R6	F10 / R7	F10 / R8	F10 / R9	F10 / R10	F10 / R11	F10 / R12	F10 / R13	F10 / R14	F10 / R15	F10 / R16	F10 / R17	F10 / R18	F10 / R19	F10 / R20	F10 / R21	F10 / R22	F10 / R23	F10 / R24
K	F11 / R1	F11 / R2	F11 / R3	F11 / R4	F11 / R5	F11 / R6	F11 / R7	F11 / R8	F11 / R9	F11 / R10	F11 / R11	F11 / R12	F11 / R13	F11 / R14	F11 / R15	F11 / R16	F11 / R17	F11 / R18	F11 / R19	F11 / R20	F11 / R21	F11 / R22	F11 / R23	F11 / R24
L	F12 / R1	F12 / R2	F12 / R3	F12 / R4	F12 / R5	F12 / R6	F12 / R7	F12 / R8	F12 / R9	F12 / R10	F12 / R11	F12 / R12	F12 / R13	F12 / R14	F12 / R15	F12 / R16	F12 / R17	F12 / R18	F12 / R19	F12 / R20	F12 / R21	F12 / R22	F12 / R23	F12 / R24
M	F13 / R1	F13 / R2	F13 / R3	F13 / R4	F13 / R5	F13 / R6	F13 / R7	F13 / R8	F13 / R9	F13 / R10	F13 / R11	F13 / R12	F13 / R13	F13 / R14	F13 / R15	F13 / R16	F13 / R17	F13 / R18	F13 / R19	F13 / R20	F13 / R21	F13 / R22	F13 / R23	F13 / R24
N	F14 / R1	F14 / R2	F14 / R3	F14 / R4	F14 / R5	F14 / R6	F14 / R7	F14 / R8	F14 / R9	F14 / R10	F14 / R11	F14 / R12	F14 / R13	F14 / R14	F14 / R15	F14 / R16	F14 / R17	F14 / R18	F14 / R19	F14 / R20	F14 / R21	F14 / R22	F14 / R23	F14 / R24
O	F15 / R1	F15 / R2	F15 / R3	F15 / R4	F15 / R5	F15 / R6	F15 / R7	F15 / R8	F15 / R9	F15 / R10	F15 / R11	F15 / R12	F15 / R13	F15 / R14	F15 / R15	F15 / R16	F15 / R17	F15 / R18	F15 / R19	F15 / R20	F15 / R21	F15 / R22	F15 / R23	F15 / R24
P	F16 / R1	F16 / R2	F16 / R3	F16 / R4	F16 / R5	F16 / R6	F16 / R7	F16 / R8	F16 / R9	F16 / R10	F16 / R11	F16 / R12	F16 / R13	F16 / R14	F16 / R15	F16 / R16	F16 / R17	F16 / R18	F16 / R19	F16 / R20	F16 / R21	F16 / R22	F16 / R23	F16 / R24

Table of grouping for sequence analysis after Miseq run (which corresponds to one section in column and another section for line pooling) – camelina TILLING pool 2D2

	Columns (+1)						Lines (-1)							
Groupe 1	F1 / R1	F1 / R3	F1 / R5	F1 / R7	F1 / R9	F1 / R11	F1 / R2	F3 / R2	F5 / R2	F7 / R2	F9 / R2	F11 / R2	F13 / R2	F15 / R2
	F1 / R13	F1 / R15	F1 / R17	F1 / R19	F1 / R21	F1 / R23	F1 / R4	F3 / R4	F5 / R4	F7 / R4	F9 / R4	F11 / R4	F13 / R4	F15 / R4
Groupe 2	F3 / R1	F3 / R3	F3 / R5	F3 / R7	F3 / R9	F3 / R11	F1 / R6	F3 / R6	F5 / R6	F7 / R6	F9 / R6	F11 / R6	F13 / R6	F15 / R6
	F3 / R13	F3 / R15	F3 / R17	F3 / R19	F3 / R21	F3 / R23	F1 / R8	F3 / R8	F5 / R8	F7 / R8	F9 / R8	F11 / R8	F13 / R8	F15 / R8
Groupe 3	F5 / R1	F5 / R3	F5 / R5	F5 / R7	F5 / R9	F5 / R11	F1 / R10	F3 / R10	F5 / R10	F7 / R10	F9 / R10	F11 / R10	F13 / R10	F15 / R10
	F5 / R13	F5 / R15	F5 / R17	F5 / R19	F5 / R21	F5 / R23	F1 / R12	F3 / R12	F5 / R12	F7 / R12	F9 / R12	F11 / R12	F13 / R12	F15 / R12
Groupe 4	F7 / R1	F7 / R3	F7 / R5	F7 / R7	F7 / R9	F7 / R11	F1 / R14	F3 / R14	F5 / R14	F7 / R14	F9 / R14	F11 / R14	F13 / R14	F15 / R14
	F7 / R13	F7 / R15	F7 / R17	F7 / R19	F7 / R21	F7 / R23	F1 / R16	F3 / R16	F5 / R16	F7 / R16	F9 / R16	F11 / R16	F13 / R16	F15 / R16
Groupe 5	F9 / R1	F9 / R3	F9 / R5	F9 / R7	F9 / R9	F9 / R11	F1 / R18	F3 / R18	F5 / R18	F7 / R18	F9 / R18	F11 / R18	F13 / R18	F15 / R18
	F9 / R13	F9 / R15	F9 / R17	F9 / R19	F9 / R21	F9 / R23	F1 / R20	F3 / R20	F5 / R20	F7 / R20	F9 / R20	F11 / R20	F13 / R20	F15 / R20
Groupe 6	F11 / R1	F11 / R3	F11 / R5	F11 / R7	F11 / R9	F11 / R11	F1 / R22	F3 / R22	F5 / R22	F7 / R22	F9 / R22	F11 / R22	F13 / R22	F15 / R22
	F11 / R13	F11 / R15	F11 / R17	F11 / R19	F11 / R21	F11 / R23	F1 / R24	F3 / R24	F5 / R24	F7 / R24	F9 / R24	F11 / R24	F13 / R24	F15 / R24
Groupe 7	F2 / R1	F2 / R3	F2 / R5	F2 / R7	F2 / R9	F2 / R11	F2 / R2	F4 / R2	F6 / R2	F8 / R2	F10 / R2	F12 / R2	F14 / R2	F16 / R2
	F2 / R13	F2 / R15	F2 / R17	F2 / R19	F2 / R21	F2 / R23	F2 / R4	F4 / R4	F6 / R4	F8 / R4	F10 / R4	F12 / R4	F14 / R4	F16 / R4
Groupe 8	F4 / R1	F4 / R3	F4 / R5	F4 / R7	F4 / R9	F4 / R11	F2 / R6	F4 / R6	F6 / R6	F8 / R6	F10 / R6	F12 / R6	F14 / R6	F16 / R6
	F4 / R13	F4 / R15	F4 / R17	F4 / R19	F4 / R21	F4 / R23	F2 / R8	F4 / R8	F6 / R8	F8 / R8	F10 / R8	F12 / R8	F14 / R8	F16 / R8
Groupe 9	F6 / R1	F6 / R3	F6 / R5	F6 / R7	F6 / R9	F6 / R11	F2 / R10	F4 / R10	F6 / R10	F8 / R10	F10 / R10	F12 / R10	F14 / R10	F16 / R10
	F6 / R13	F6 / R15	F6 / R17	F6 / R19	F6 / R21	F6 / R23	F2 / R12	F4 / R12	F6 / R12	F8 / R12	F10 / R12	F12 / R12	F14 / R12	F16 / R12
Groupe 10	F8 / R1	F8 / R3	F8 / R5	F8 / R7	F8 / R9	F8 / R11	F2 / R14	F4 / R14	F6 / R14	F8 / R14	F10 / R14	F12 / R14	F14 / R14	F16 / R14
	F8 / R13	F8 / R15	F8 / R17	F8 / R19	F8 / R21	F8 / R23	F2 / R16	F4 / R16	F6 / R16	F8 / R16	F10 / R16	F12 / R16	F14 / R16	F16 / R16
Groupe 11	F10 / R1	F10 / R3	F10 / R5	F10 / R7	F10 / R9	F10 / R11	F2 / R18	F4 / R18	F6 / R18	F8 / R18	F10 / R18	F12 / R18	F14 / R18	F16 / R18
	F10 / R13	F10 / R15	F10 / R17	F10 / R19	F10 / R21	F10 / R23	F2 / R20	F4 / R20	F6 / R20	F8 / R20	F10 / R20	F12 / R20	F14 / R20	F16 / R20
Groupe 12	F15 / R1	F15 / R3	F15 / R5	F15 / R7	F15 / R9	F15 / R11	F12 / R1	F12 / R3	F12 / R5	F12 / R7	F12 / R9	F12 / R11	F12 / R13	F12 / R15
	F15 / R13	F15 / R15	F15 / R17	F15 / R19	F15 / R21	F15 / R23	F14 / R1	F14 / R3	F14 / R5	F14 / R7	F14 / R9	F14 / R11	F14 / R13	F14 / R15
Groupe 13	F13 / R1	F13 / R3	F13 / R5	F13 / R7	F13 / R9	F13 / R11	F2 / R22	F4 / R22	F6 / R22	F8 / R22	F10 / R22	F12 / R22	F14 / R22	F16 / R22
	F13 / R13	F13 / R15	F13 / R17	F13 / R19	F13 / R21	F13 / R23	F2 / R24	F4 / R24	F6 / R24	F8 / R24	F10 / R24	F12 / R24	F14 / R24	F16 / R24

TABLE OF 384 WELL PLATE WITH DISTRIBUTION OF MISEQ PRIMERS IN THE TILLING POOL 2

		R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	R17	R18	R19	R20	R21	R22	R23	R24
A	F1	1	-1	1	-1	1	-2	1	-2	1	-3	1	-3	1	-4	1	-4	1	-5	1	-5	1	-6	1	-6
B	F2	7	-7	7	-7	7	-8	7	-8	7	-9	7	-9	7	-10	7	-10	7	-11	7	-11	7	-13	7	-13
C	F3	2	-1	2	-1	2	-2	2	-2	2	-3	2	-3	2	-4	2	-4	2	-5	2	-5	2	-6	2	-6
D	F4	8	-7	8	-7	8	-8	8	-8	8	-9	8	-9	8	-10	8	-10	8	-11	8	-11	8	-13	8	-13
E	F5	3	-1	3	-1	3	-2	3	-2	3	-3	3	-3	3	-4	3	-4	3	-5	3	-5	3	-6	3	-6
F	F6	9	-7	9	-7	9	-8	9	-8	9	-9	9	-9	9	-10	9	-10	9	-11	9	-11	9	-13	9	-13
G	F7	4	-1	4	-1	4	-2	4	-2	4	-3	4	-3	4	-4	4	-4	4	-5	4	-5	4	-6	4	-6
H	F8	10	-7	10	-7	10	-8	10	-8	10	-9	10	-9	10	-10	10	-10	10	-11	10	-11	10	-13	10	-13
I	F9	5	-1	5	-1	5	-2	5	-2	5	-3	5	-3	5	-4	5	-4	5	-5	5	-5	5	-6	5	-6
J	F10	11	-7	11	-7	11	-8	11	-8	11	-9	11	-9	11	-10	11	-10	11	-11	11	-11	11	-13	11	-13
K	F11	6	-1	6	-1	6	-2	6	-2	6	-3	6	-3	6	-4	6	-4	6	-5	6	-5	6	-6	6	-6
L	F12	-12	-7	-12	-7	-12	-8	-12	-8	-12	-9	-12	-9	-12	-10	-12	-10	0	-11	0	-11	0	-13	0	-13
M	F13	13	-1	13	-1	13	-2	13	-2	13	-3	13	-3	13	-4	13	-4	13	-5	13	-5	13	-6	13	-6
N	F14	-12	-7	-12	-7	-12	-8	-12	-8	-12	-9	-12	-9	-12	-10	-12	-10	0	-11	0	-11	0	-13	0	-13
O	F15	12	-1	12	-1	12	-2	12	-2	12	-3	12	-3	12	-4	12	-4	12	-5	12	-5	12	-6	12	-6
P	F16	0	-7	0	-7	0	-8	0	-8	0	-9	0	-9	0	-10	0	-10	0	-11	0	-11	0	-13	0	-13

A3. CAMELINA TbyS MUTANT DETAILS

A3.1 CAMELINA TILLING BY SEQUENCING – *FAD2* GENE MUTANT DETAILS

SR. NO	GENE	AMPLICON	Réf.	Mut.	Pos.	Type mutation	Codon	Mutated Codon	AA Change	M2 Family	EMS/NONEMS	SIMPLON/ DOUBLON
1	FAD2	AMP3	G	A	24	MISSENSE	CGT	CAT	R221H	2255	EMS	DOUBLON
2	FAD2	AMP3	G	A	68	MISSENSE	GTC	ATC	V236I	2431	EMS	DOUBLON
3	FAD2	AMP3	C	T	70	SILENT	GTC	GTT	V236V	2238	EMS	DOUBLON
4	FAD2	AMP3	C	T	80	MISSENSE	CTT	TTT	L240F	1806	EMS	DOUBLON
5	FAD2	AMP3	G	A	87	MISSENSE	CGT	CAT	R242H	2241	EMS	DOUBLON
6	FAD2	AMP3	C	T	130	SILENT	TAC	TAT	Y256Y	610	EMS	DOUBLON
7	FAD2	AMP3	C	T	154	SILENT	AAC	AAT	N264N	2017	EMS	DOUBLON
8	FAD2	AMP3	C	T	163	SILENT	CTC	CTT	L267L	526	EMS	DOUBLON
9	FAD2	AMP3	G	A	184	SILENT	CAG	CAA	Q274Q	2350	EMS	DOUBLON
10	FAD2	AMP3	C	T	219	MISSENSE	TCC	TTC	S286F	1604	EMS	DOUBLON
11	FAD2	AMP3	G	A	239	MISSENSE	GGA	AGA	G293R	1877	EMS	DOUBLON
12	FAD2	AMP3	G	A	352	SILENT	GCG	GCA	A330A	161	EMS	DOUBLON
13	FAD2	AMP3	G	A	356	MISSENSE	GAA	AAA	E332K	4130	EMS	DOUBLON
14	FAD2	AMP3	C	T	383	MISSENSE	CTC	TTC	L341F	4086	EMS	DOUBLON
15	FAD2	AMP3	G	A	387	MISSENSE	GGT	GAT	G342D	642	EMS	DOUBLON
16	FAD2	AMP3	G	A	387	MISSENSE	GGT	GAT	G342D	231	EMS	DOUBLON
17	FAD2	AMP3	G	A	422	MISSENSE	GTG	ATG	V354M	1715	EMS	DOUBLON
18	FAD2	AMP3	G	A	435	MISSENSE	AGG	AAG	R358K	3993	EMS	DOUBLON
19	FAD2	AMP3	G	A	439	SILENT	GAG	GAA	E359E	927	EMS	DOUBLON
20	FAD2	AMP3	G	A	439	SILENT	GAG	GAA	E359E	2198	EMS	DOUBLON
21	FAD2	AMP3	T	A	75	MISSENSE	TTT	TAT	F238Y	526	NO EMS	DOUBLON
22	FAD2	AMP3	G	T	79	SILENT	GGG	GGT	G239G	526	NO EMS	DOUBLON
23	FAD2	AMP3	T	A	175	SILENT	ACT	ACA	T271T	526	NO EMS	DOUBLON

SR. NO	GENE	AMPLICON	Réf.	Mut.	Pos.	Type mutation	Codon	Mutated Codon	AA Change	M2 Family	EMS/NONEMS	SIMPLON/ DOUBLON
24	FAD2	AMP3	T	C	346	SILENT	TAT	TAC	Y328Y	526	NO EMS	DOUBLON
25	FAD2	AMP1	C	T	21	MISSENSE	CCT	CTT	P26L	1795	EMS	DOUBLON
26	FAD2	AMP1	C	T	32	MISSENSE	CCG	TCG	P30S	593	EMS	DOUBLON
27	FAD2	AMP1	C	T	33	MISSENSE	CCG	CTG	P30L	563	EMS	DOUBLON
28	FAD2	AMP1	C	T	68	MISSENSE	CCA	TCA	P42S	856	EMS	DOUBLON
29	FAD2	AMP1	G	A	78	MISSENSE	TGT	TAT	C45Y	2138	EMS	DOUBLON
30	FAD2	AMP1	G	A	152	MISSENSE	GTC	ATC	V70I	90	EMS	DOUBLON
31	FAD2	AMP1	C	T	247	SILENT	GTC	GTT	V101V	4068	EMS	DOUBLON
32	FAD2	AMP1	G	A	257	MISSENSE	GCC	ACC	A105T	4123	EMS	DOUBLON
33	FAD2	AMP1	C	T	268	SILENT	TGC	TGT	C108C	1301	EMS	DOUBLON
34	FAD2	AMP1	C	T	299	MISSENSE	CTT	TTT	L119F	1595	EMS	DOUBLON
35	FAD2	AMP1	G	A	315	MISSENSE	GGT	GAT	G124D	2247	EMS	DOUBLON
36	FAD2	AMP1	G	A	315	MISSENSE	GGT	GAT	G124D	124	EMS	DOUBLON
37	FAD2	AMP1	C	T	331	SILENT	TCC	TCT	S129S	1327	EMS	DOUBLON
38	FAD2	AMP1	C	T	344	MISSENSE	CCT	TCT	P134S	1173	EMS	DOUBLON
39	FAD2	AMP1	C	T	345	MISSENSE	CCT	CTT	P134L	2232	EMS	DOUBLON
40	FAD2	AMP1	C	T	354	MISSENSE	TCC	TTC	S137F	349	EMS	DOUBLON
41	FAD2	AMP1	G	A	361	SILENT	AAG	AAA	K139K	2397	EMS	DOUBLON
42	FAD2	AMP2	C	T	34	SILENT	TCC	TCT	S129S	1327	EMS	DOUBLON
43	FAD2	AMP2	C	T	48	MISSENSE	CCT	CTT	P134L	2232	EMS	DOUBLON
44	FAD2	AMP2	C	T	52	SILENT	TAC	TAT	Y135Y	350	EMS	DOUBLON
45	FAD2	AMP2	C	T	57	MISSENSE	TCC	TTC	S137F	349	EMS	DOUBLON
46	FAD2	AMP2	G	A	64	SILENT	AAG	AAA	K139K	2397	EMS	DOUBLON
47	FAD2	AMP2	C	T	67	SILENT	TAC	TAT	Y140Y	1195	EMS	DOUBLON
48	FAD2	AMP2	C	T	71	MISSENSE	CAT	TAT	H142Y	1055	EMS	DOUBLON
49	FAD2	AMP2	C	T	163	SILENT	TAC	TAT	Y172Y	1337	EMS	DOUBLON
50	FAD2	AMP2	G	A	180	MISSENSE	GGA	GAA	G178E	746	EMS	DOUBLON

SR. NO	GENE	AMPLICON	Réf.	Mut.	Pos.	Type mutation	Codon	Mutated Codon	AA Change	M2 Family	EMS/NONEMS	SIMPLON/ DOUBLON
51	FAD2	AMP2	G	A	353	MISSENSE	GTC	ATC	V236I	2431	EMS	DOUBLON
52	FAD2	AMP2	C	T	415	SILENT	TAC	TAT	Y256Y	610	EMS	DOUBLON
53	FAD2	AMP2	T	G	358	MISSENSE	TGT	TGG	C237W	4013	NO EMS	DOUBLON
54	FAD2	AMP2	A	G	436	SILENT	GTA	GTG	V263V	1192	NO EMS	DOUBLON
55	FAD2	AMP3	G	A	24	MISSENSE	CGT	CAT	R221H	695	EMS	SIMPLON
56	FAD2	AMP3	C	T	121	SILENT	ATC	ATT	I253I	1127	EMS	SIMPLON
57	FAD2	AMP3	C	T	137	MISSENSE	CCA	TCA	P259S	2034	EMS	SIMPLON
58	FAD2	AMP3	G	A	238	SILENT	AGG	AGA	R292R	1239	EMS	SIMPLON
59	FAD2	AMP3	G	A	387	MISSENSE	GGT	GAT	G342D	1715	EMS	SIMPLON
60	FAD2	AMP3	C	T	406	SILENT	GAC	GAT	D348D	526	EMS	SIMPLON
61	FAD2	AMP1	C	T	138	MISSENSE	TCC	TTC	S65F	782	EMS	SIMPLON
62	FAD2	AMP1	G	A	152	MISSENSE	GTC	ATC	V70I	2246	EMS	SIMPLON
63	FAD2	AMP1	C	T	277	SILENT	CAC	CAT	H111H	870	EMS	SIMPLON
64	FAD2	AMP1	C	T	277	SILENT	CAC	CAT	H111H	382	EMS	SIMPLON
65	FAD2	AMP1	G	A	315	MISSENSE	GGT	GAT	G124D	437	EMS	SIMPLON
66	FAD2	AMP1	C	T	349	SILENT	TAC	TAT	Y135Y	350	EMS	SIMPLON
67	FAD2	AMP1	C	T	354	MISSENSE	TCC	TTC	S137F	1325	EMS	SIMPLON
68	FAD2	AMP2	C	T	46	SILENT	GTC	GTT	V133V	2339	EMS	SIMPLON
69	FAD2	AMP2	C	T	47	MISSENSE	CCT	TCT	P134S	1189	EMS	SIMPLON
70	FAD2	AMP2	C	T	174	MISSENSE	CCT	CTT	P176L	1311	EMS	SIMPLON
71	FAD2	AMP2	G	A	264	MISSENSE	GGG	GAG	G206E	2203	EMS	SIMPLON
72	FAD2	AMP2	G	A	305	MISSENSE	GAC	AAC	D220N	2375	EMS	SIMPLON
73	FAD2	AMP2	C	T	355	SILENT	GTC	GTT	V236V	2238	EMS	SIMPLON
74	FAD2	AMP2	C	T	439	SILENT	AAC	AAT	N264N	2352	EMS	SIMPLON

A3.2 CAMELINA TILLING BY SEQUENCING – *FAE1* GENE MUTANT DETAILS

SR. NO	GENE	AMPLICON	Réf.	Mut.	Pos.	Type mutation	Codon	Mutated Codon	AA Change	Family	EMS/ NONEMS	SIMPLON/ DOUBLON
1	FAE1	AMP2	G	A	28	MISSENSE	GGT	GAT	G166D	737	EMS	DOUBLON
2	FAE1	AMP2	C	T	65	SILENT	AAC	AAT	N178N	683	EMS	DOUBLON
3	FAE1	AMP2	G	A	100	MISSENSE	AGC	AAC	S190N	751	EMS	DOUBLON
4	FAE1	AMP2	G	A	199	MISSENSE	AGT	AAT	S223N	289	EMS	DOUBLON
5	FAE1	AMP2	G	A	205	MISSENSE	GGT	GAT	G225D	820	EMS	DOUBLON
6	FAE1	AMP2	G	A	207	MISSENSE	GTT	ATT	V226I	524	EMS	DOUBLON
7	FAE1	AMP2	G	A	303	MISSENSE	GCT	ACT	A258T	1049	EMS	DOUBLON
8	FAE1	AMP2	G	A	352	MISSENSE	GGT	GAT	G274D	3967	EMS	DOUBLON
9	FAE1	AMP2	G	A	385	MISSENSE	GGA	GAA	G285E	1989	EMS	DOUBLON
10	FAE1	AMP2	G	A	385	MISSENSE	GGA	GAA	G285E	1793	EMS	DOUBLON
11	FAE1	AMP2	G	A	387	MISSENSE	GAT	AAT	D286N	1771	EMS	DOUBLON
12	FAE1	AMP2	C	T	400	MISSENSE	TCC	TTC	S290F	1673	EMS	DOUBLON
13	FAE1	AMP2	G	A	447	MISSENSE	GAC	AAC	D306N	1053	EMS	DOUBLON
14	FAE1	AMP2	C	T	454	MISSENSE	TCT	TTT	S308F	659	EMS	DOUBLON
15	FAE1	AMP2	C	T	468	NONSENSE	CAA	TAA	Q313*	4149	EMS	DOUBLON
16	FAE1	AMP1	C	T	33	MISSENSE	CTC	TTC	L48F	1609	EMS	DOUBLON
17	FAE1	AMP1	C	T	36	NONSENSE	CAA	TAA	Q49*	2356	EMS	DOUBLON
18	FAE1	AMP1	C	T	45	MISSENSE	CTT	TTT	L52F	526	EMS	DOUBLON
19	FAE1	AMP1	C	T	63	MISSENSE	CTC	TTC	L58F	106	EMS	DOUBLON
20	FAE1	AMP1	C	T	70	MISSENSE	GCT	GTT	A60V	130	EMS	DOUBLON
21	FAE1	AMP1	C	T	76	MISSENSE	ACC	ATC	T62I	4073	EMS	DOUBLON
22	FAE1	AMP1	C	T	95	SILENT	CTC	CTT	L68L	107	EMS	DOUBLON
23	FAE1	AMP1	G	A	110	SILENT	CGG	CGA	R73R	526	EMS	DOUBLON
24	FAE1	AMP1	C	T	117	MISSENSE	CCG	TCG	P76S	1717	EMS	DOUBLON
25	FAE1	AMP1	G	A	119	SILENT	CCG	CCA	P76P	2363	EMS	DOUBLON

SR. NO	GENE	AMPLICON	Réf.	Mut.	Pos.	Type mutation	Codon	Mutated Codon	AA Change	Family	EMS/ NONEMS	SIMPLON/ DOUBLON
26	FAE1	AMP1	C	T	125	SILENT	TAC	TAT	Y78Y	2138	EMS	DOUBLON
27	FAE1	AMP1	C	T	137	SILENT	TAC	TAT	Y82Y	1124	EMS	DOUBLON
28	FAE1	AMP1	C	T	143	SILENT	TGC	TGT	C84C	2108	EMS	DOUBLON
29	FAE1	AMP1	G	A	208	MISSENSE	AGA	AAA	R106K	730	EMS	DOUBLON
30	FAE1	AMP1	G	A	208	MISSENSE	AGA	AAA	R106K	2377	EMS	DOUBLON
31	FAE1	AMP1	C	T	246	MISSENSE	CCA	TCA	P119S	858	EMS	DOUBLON
32	FAE1	AMP1	C	T	255	MISSENSE	CTT	TTT	L122F	4132	EMS	DOUBLON
33	FAE1	AMP1	G	A	269	SILENT	AGG	AGA	R126R	1327	EMS	DOUBLON
34	FAE1	AMP1	C	T	314	SILENT	CCC	CCT	P141P	1777	EMS	DOUBLON
35	FAE1	AMP1	C	T	314	SILENT	CCC	CCT	P141P	315	EMS	DOUBLON
36	FAE1	AMP1	C	T	321	MISSENSE	CTC	TTC	L144F	198	EMS	DOUBLON
37	FAE1	AMP1	C	T	333	MISSENSE	CCC	TCC	P148S	2269	EMS	DOUBLON
38	FAE1	AMP1	C	T	335	SILENT	CCC	CCT	P148P	396	EMS	DOUBLON
39	FAE1	AMP1	C	T	335	SILENT	CCC	CCT	P148P	323	EMS	DOUBLON
40	FAE1	AMP1	C	T	336	MISSENSE	CCA	TCA	P149S	2188	EMS	DOUBLON
41	FAE1	AMP1	C	T	355	MISSENSE	GCT	GTT	A155V	1712	EMS	DOUBLON
42	FAE1	AMP1	C	T	360	MISSENSE	CGT	TGT	R157C	558	EMS	DOUBLON
43	FAE1	AMP1	C	T	386	SILENT	ATC	ATT	I165I	661	EMS	DOUBLON
44	FAE1	AMP1	C	T	425	SILENT	AAC	AAT	N178N	683	EMS	DOUBLON
45	FAE1	AMP1	A	G	57	MISSENSE	ATT	GTT	I56V	526	NO EMS	DOUBLON
46	FAE1	AMP1	C	G	79	MISSENSE	GCT	GGT	A63G	526	NO EMS	DOUBLON
47	FAE1	AMP1	T	C	385	MISSENSE	ATC	ACC	I165T	1162	NO EMS	DOUBLON
48	FAE1	AMP1	G	C	401	MISSENSE	AAG	AAC	K170N	526	NO EMS	DOUBLON
49	FAE1	AMP3	G	A	116	SILENT	GCG	GCA	A334A	1278	EMS	DOUBLON
50	FAE1	AMP3	G	A	131	SILENT	AAG	AAA	K339K	1971	EMS	DOUBLON
51	FAE1	AMP3	G	A	155	SILENT	CCG	CCA	P347P	1842	EMS	DOUBLON
52	FAE1	AMP3	G	A	204	MISSENSE	GCC	ACC	A364T	2382	EMS	DOUBLON

SR. NO	GENE	AMPLICON	Réf.	Mut.	Pos.	Type mutation	Codon	Mutated Codon	AA Change	Family	EMS/ NONEMS	SIMPLON/ DOUBLON
53	FAE1	AMP3	C	T	246	MISSENSE	CCG	TCG	P378S	441	EMS	DOUBLON
54	FAE1	AMP3	C	T	247	MISSENSE	CCG	CTG	P378L	1801	EMS	DOUBLON
55	FAE1	AMP3	G	A	365	SILENT	ACG	ACA	T417T	1786	EMS	DOUBLON
56	FAE1	AMP3	G	A	527	SILENT	AAG	AAA	K471K	354	EMS	DOUBLON
57	FAE1	AMP3	C	T	246	MISSENSE	CCG	TCG	P378S	441	EMS	DOUBLON
58	FAE1	AMP3	A	G	100	MISSENSE	GAC	GGC	D329G	827	NO EMS	DOUBLON
59	FAE1	AMP3	A	G	104	MISSENSE	ATA	ATG	I330M	818	NO EMS	DOUBLON
60	FAE1	AMP3	T	C	161	SILENT	ATT	ATC	I349I	2170	NO EMS	DOUBLON
61	FAE1	AMP3	A	G	195	MISSENSE	ACC	GCC	T361A	1788	NO EMS	DOUBLON
62	FAE1	AMP3	T	A	216	MISSENSE	TTG	ATG	L368M	2275	NO EMS	DOUBLON
63	FAE1	AMP3	T	A	216	MISSENSE	TTG	ATG	L368M	1856	NO EMS	DOUBLON
64	FAE1	AMP3	C	A	234	MISSENSE	CAC	AAC	H374N	1687	NO EMS	DOUBLON
65	FAE1	AMP3	C	A	246	MISSENSE	CCG	ACG	P378T	2054	NO EMS	DOUBLON
66	FAE1	AMP3	A	G	445	MISSENSE	AAA	AGA	K444R	1800	NO EMS	DOUBLON
67	FAE1	AMP2	C	T	67	MISSENSE	CCT	CTT	P179L	344	EMS	SIMPLON
68	FAE1	AMP2	G	A	134	MISSENSE	ATG	ATA	M201I	1831	EMS	SIMPLON
69	FAE1	AMP2	G	A	303	MISSENSE	GCT	ACT	A258T	751	EMS	SIMPLON
70	FAE1	AMP2	G	A	351	MISSENSE	GGT	AGT	G274S	1826	EMS	SIMPLON
71	FAE1	AMP2	G	A	392	SILENT	CGG	CGA	R287R	1552	EMS	SIMPLON
72	FAE1	AMP2	T	C	302	SILENT	TAT	TAC	Y257Y	105	NOEMS	SIMPLON
73	FAE1	AMP2	T	G	302	NONSENSE	TAT	TAG	Y257*	4074	NOEMS	SIMPLON
74	FAE1	AMP1	C	T	76	MISSENSE	ACC	ATC	T62I	1224	EMS	SIMPLON
75	FAE1	AMP1	C	T	153	MISSENSE	CCA	TCA	P88S	1224	EMS	SIMPLON
76	FAE1	AMP1	C	T	156	MISSENSE	CCG	TCG	P89S	156	EMS	SIMPLON
77	FAE1	AMP1	C	T	334	MISSENSE	CCC	CTC	P148L	1643	EMS	SIMPLON
78	FAE1	AMP1	G	A	295	MISSENSE	GGT	GAT	G135D	2023	EMS	SIMPLON
79	FAE1	AMP1	G	A	318	MISSENSE	GGA	AGA	G143R	1239	EMS	SIMPLON

SR. NO	GENE	AMPLICON	Réf.	Mut.	Pos.	Type mutation	Codon	Mutated Codon	AA Change	Family	EMS/ NONEMS	SIMPLON/ DOUBLON
80	FAE1	AMP1	G	A	344	SILENT	AAG	AAA	K151K	1612	EMS	SIMPLON
81	FAE1	AMP1	G	A	366	MISSENSE	GAG	AAG	E159K	2241	EMS	SIMPLON
82	FAE1	AMP1	C	T	315	NONSENSE	CAG	TAG	Q142*	1554	EMS	SIMPLON
83	FAE1	AMP1	T	A	24	MISSENSE	TAT	AAT	Y45N	1141	EMS	SIMPLON
84	FAE1	AMP1	C	A	112	MISSENSE	CCC	CAC	P74H	2174	EMS	SIMPLON

A3.3 CAMELINA TILLING BY SEQUENCING – MIPS1 GENE MUTANT DETAILS

SR. NO	GENE	AMPLICON	Réf.	Mut.	Pos.	Type mutation	Codon	Mutated Codon	AA Change	M2 Family	EMS/ NONEMS	SIMPLON/ DOUBLON
1	MIPS1	AMP1	G	A	70	SPLICING				2508	EMS	DOUBLON
2	MIPS1	AMP1	G	A	79	SILENT	CCG	CCA	P134P	1111	EMS	DOUBLON
3	MIPS1	AMP1	C	T	121	SILENT	AAC	AAT	N148N	602	EMS	DOUBLON
4	MIPS1	AMP1	G	A	125	MISSENSE	GCA	ACA	A150T	330	EMS	DOUBLON
5	MIPS1	AMP1	G	A	128	MISSENSE	GAT	AAT	D151N	441	EMS	DOUBLON
6	MIPS1	AMP1	C	T	182	MISSENSE	CCT	TCT	P169S	844	EMS	DOUBLON
7	MIPS1	AMP1	G	A	202	SILENT	GTG	GTA	V175V	1160	EMS	DOUBLON
8	MIPS1	AMP1	G	A	221	MISSENSE	GAT	AAT	D182N	4148	EMS	DOUBLON
9	MIPS1	AMP1	C	T	252	MISSENSE	TCA	TTA	S192L	1052	EMS	DOUBLON
10	MIPS1	AMP1	G	A	266	MISSENSE	GTG	ATG	V197M	921	EMS	DOUBLON
11	MIPS1	AMP1	G	A	276	MISSENSE	GGT	GAT	G200D	2132	EMS	DOUBLON
12	MIPS1	AMP1	G	A	421	SILENT	AGG	AGA	R214R	1937	EMS	DOUBLON
13	MIPS1	AMP1	G	A	421	SILENT	AGG	AGA	R214R	381	EMS	DOUBLON
14	MIPS1	AMP1	G	A	431	MISSENSE	GAG	AAG	E218K	2260	EMS	DOUBLON
15	MIPS1	AMP2	G	A	70	INTRONIC				1806	EMS	DOUBLON
16	MIPS1	AMP2	C	T	159	MISSENSE	ACC	ATC	T29I	2237	EMS	DOUBLON
17	MIPS1	AMP2	G	A	261	MISSENSE	GGG	GAG	G63E	510	EMS	DOUBLON
18	MIPS1	AMP2	C	A	251	MISSENSE	CCC	ACC	P60T	412	NO EMS	DOUBLON
19	MIPS1	AMP2	C	A	252	MISSENSE	CCC	CAC	P60H	638	NO EMS	DOUBLON
20	MIPS1	AMP2	C	A	252	MISSENSE	CCC	CAC	P60H	350	NO EMS	DOUBLON
21	MIPS1	AMP2	G	T	261	MISSENSE	GGG	GTG	G63V	445	NO EMS	DOUBLON

A3.4 CAMELINA TILLING BY SEQUENCING – TGG1 GENE MUTANT DETAILS

SR. NO	GENE	AMPLICON	Réf.	Mut.	Pos.	Type mutation	Codon	Mutated Codon	AA Change	M2 Family	EMS/ NONEMS	SIMPLON/ DOUBLON
1	TGG1	AMP2	G	A	38	MISSENSE	GAT	AAT	D188N	1677	EMS	DOUBLON
2	TGG1	AMP2	C	T	49	SILENT	TTC	TTT	F191F	2140	EMS	DOUBLON
3	TGG1	AMP2	G	A	67	SILENT	AGG	AGA	R197R	2209	EMS	DOUBLON
4	TGG1	AMP2	C	T	76	SILENT	CAC	CAT	H200H	1021	EMS	DOUBLON
5	TGG1	AMP2	C	T	107	MISSENSE	CCT	TCT	P211S	601	EMS	DOUBLON
6	TGG1	AMP2	G	A	117	MISSENSE	GGA	GAA	G214E	1030	EMS	DOUBLON
7	TGG1	AMP2	G	A	117	MISSENSE	GGA	GAA	G214E	1746	EMS	DOUBLON
8	TGG1	AMP1	G	A	23	MISSENSE	GAG	AAG	E288K	4071	EMS	DOUBLON
9	TGG1	AMP1	G	A	53	MISSENSE	GAA	AAA	E298K	1579	EMS	DOUBLON
10	TGG1	AMP1	C	T	169	SILENT	CTA	TTA	L308L	921	EMS	DOUBLON
11	TGG1	AMP1	G	A	177	SILENT	AAG	AAA	K310K	217	EMS	DOUBLON
12	TGG1	AMP1	G	A	198	MISSENSE	ATG	ATA	M317I	1228	EMS	DOUBLON
13	TGG1	AMP1	C	T	224	MISSENSE	CCA	CTA	P326L	1278	EMS	DOUBLON
14	TGG1	AMP1	C	T	328	MISSENSE	CCT	TCT	P361S	2389	EMS	DOUBLON