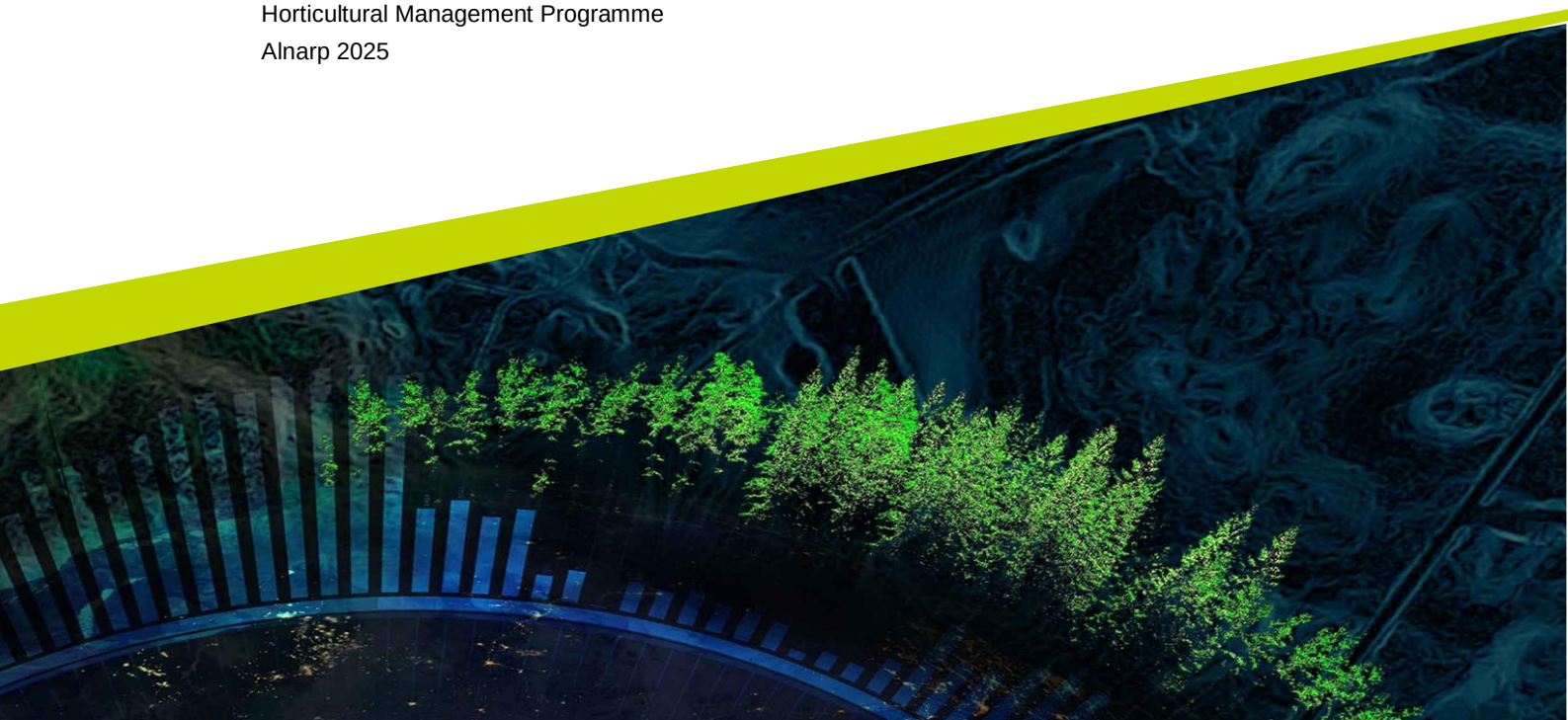
Genetic characterization of the *FAD2* gene in field cress (*Lepidium campestre*)

Lisa Hartmann

Degree project • 15 credits
Swedish University of Agricultural Sciences, SLU
Department of Plant Breeding
Horticultural Management Programme
Alnarp 2025



Genetic characterization of the FAD2 gene in field cress (*Lepidium campestre*)

Lisa Hartmann

Supervisor:	**Cecilia Hammenhag, Swedish University of Agricultural Sciences, SLU, Department of Plant Breeding**
Examiner:	**Ronja Wonneberger, Swedish University of Agricultural Sciences, SLU, Department of Plant Breeding**

Credits:	15 credits
Level:	First cycle G2E
Course title:	Independent project in Biology
Course code:	EX0855
Programme/education:	Horticultural Management Programme
Course coordinating dept:	Department of Biosystems and Technology
Place of publication:	Alnarp
Year of publication:	2025
Keywords:	Field cress, *Lepidium campestre*, *FAD2*, oil composition, polymorphism, genetic sequencing, PCR, INDEL, serine, arginine, glycine, non-synchronous mutation

Swedish University of Agricultural Sciences
Department of Plant Breeding

Abstract

Ongoing domestication efforts for field cress, *Lepidium campestre*, aim to commercialize field cress as an industrial oil crop in Sweden, potentially offering a sustainable alternative to fossil fuels and mineral oils in various industries. The identification of polymorphisms and genetic variation among different accessions is useful for identifying suitable individuals in crossbreeding. To examine genetic variation in the *FAD2* genomic coding region of field cress, plant tissue was sampled from 16 accessions of greenhouse plants grown from wild-collected field cress seeds. DNA was extracted from the accessions. The *FAD2* coding region was amplified via PCR cycling, and the purified PCR products were sequenced according to the Sanger method. Based on the alignment of the gene sequences and comparison with FAD2 gene sequences from a previous study (Lodenius 2023) and a genetic data bank, one INDEL linked to the loss of the amino acid serine and one heterozygous locus leading to the single amino acid substitution of arginine for glycine were identified. Overall, the accessions showed overwhelming homogeneity in the sequenced genetic region, despite the diverse origins of seeds within Sweden and Denmark. Future research may examine whether identified polymorphisms affect protein synthesis and whether they are linked with variation in field cress seed oil composition.

Keywords: Field cress, *Lepidium campestre*, *FAD2*, oil composition, polymorphism, genetic sequencing, PCR, INDEL, serine, arginine, glycine, non-synchronous mutation

Acknowledgement

I would like to thank my supervisor, Cecilia Hammenhag for her guidance and support in planning and executing this study. Her patience and wisdom helped me not only in my decision making throughout the research process but also gave me access to an immense amount of knowledge and experience in the field of plant breeding. I want to express my sincere appreciation for her support.

Table of contents

List of tables

List of figures

Abbreviations

Bp	Base pair
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DNA	Deoxyribonucleic Acid
dNTP	deoxynucleotide triphosphate
EA	Erucic acid (C22:1)
FAD2	*FATTY ACID DESATURASE 2*
FAD3	*FATTY ACID DESATURASE 3*
FAE1	*FATTY ACID ELONGASE 1*
INDEL	Insertion/Deletion
kb	Kilobase
KCS8	*3-KETOACYL-COA-SYNTHASE 8*
MAS	Marker Assisted Selection
NCBI	National Center for Biotechnology Information
NGT	New genomic technology
PCR	Polymerase Chain Reaction
PUFAs	Polyunsaturated fatty acids
SAF	Sustainable aviation fuel
SLU	Swedish University of Agricultural Sciences
SCAR	Sequence characterized amplified regions
SNP	Single-nucleotide polymorphism
SSR	Simple sequence repeats
STS	Sequence tagged sites
TAG1	*ACYL-COA:DIACYLGLYCEROL ACYLTRANSFERASE 1*
WRI1	*WRINKLED 1*

1. Introduction

In 2023, 81,5% of global energy consumption was sourced from fossil fuels in the form of oil, coal and natural gas (Energy Institute 2024). Due to the increasingly devastating effects of global warming, caused primarily by CO_2 emissions from the burning of fossil fuels, significantly expanding the share of renewable energy sources by 2030 was set as Sustainable Development Goal target 7.2 (UN 2015). Finding alternative sustainable energy sources is one way of meeting this goal.

Consumption of biofuels as a renewable alternative to mineral oils has increased by an average of 8,2% annually over the course of 10 years (2013-2023), currently growing at accelerating rates (Energy Institute 2024). To meet this increasing global demand for biofuel and renewable oils for industrial use, researchers are exploring new sources of plant oil, one example being field cress, *Lepidium campestre*.

Field cress, *L. campestre,* is originally a European wild species within the Brassicaceae family, which currently grows in Europe and North America (Desta et al. 2019). Field cress is under domestication as a cover crop for spring-sown grain crops in northern Sweden, complementing spring grains in seasonality and crop height (Merker et al. 2010). Covering fallow ground throughout the fall and winter months after harvest, thereby keeping weeds at bay and retaining soil nutrients, field cress as a cover crop offers a habitat for valuable insects and microbes (Merker et al. 2010). Natural characteristics of field cress such as its tolerance for hardy winter climate and the fact that it is biennial, make it a suitable cover crop for future cultivation in northern parts of Scandinavia (Gustafsson et al. 2018).

Besides its value as a cover crop, field cress has potential to become an oil crop like its relative, *Brassica napus*, known as rapeseed. Compared to commercial rapeseed cultivars, field cress has lower seed oil content and contains high levels of polyunsaturated fatty acids such as erucic acid, which makes its seed oil currently unsuitable for human nutritional consumption (Sandgrind et al. 2023). Possible industrial uses include biofuel, industrial oil for lubrication or green plastic production, and, if seed oil composition can be altered, high-quality animal fodder, also known as press cake (Hammenhag et al. 2020a). Currently, no field cress breeding lines show all desirable characteristics for large scale agricultural production, for example high oil content, retention of seeds until harvest, and the simultaneous ripening of seeds throughout the field (Hammenhag et al. 2020b).

Objective: To genetically characterize wild-collected samples of field cress, focusing on the *FATTY ACID DESATURASE 2 (FAD2)* gene, associated with oil composition in Brassicaceae plants, and to compare the genetic make-up of the

newly collected accessions with an existing collection to assess genetic variation for crop improvement.

Research hypothesis: The wild samples of field cress will exhibit genetic differences in the examined genomic regions, indicating potential beneficial traits for further improvement.

2. Background

2.1 Domestication of Field Cress

Researchers at the Swedish University for Agricultural Sciences (SLU) have been working on the domestication of field cress as a cover crop and future oil crop for northern Sweden since the 1990s (Hammenhag et al. 2020b). First trials in the process of the commercialization of field cress have now begun via a cooperation project between different actors along the value chain (SLU 2023). Current goals of domestication efforts are to refine the characteristics of field cress through plant breeding, while exploring techniques for its optimal cultivation, and building a sustainable value chain for field cress oil in Sweden (SLU 2023). Commercial uses which are currently considered for field cress oil include biofuel, sustainable aviation fuel (SAF) and industrial plant oil as an alternative to mineral oil in other industries (Hammenhag, personal communication).

Building on progress of field cress plant breeding efforts, knowledge about genetic diversity within the field cress population is useful for identifying viable accessions for crossbreeding. The complete genome of both *L. campestre* and closely related species, such as *L. heterophyllum* has been sequenced (Desta et al. 2019). *L. heterophyllum* has been found to be able to create viable hybrids with *L. campestre* (Hammenhag et al. 2020b). Hammenhag et al. (2020b) identified Quantitative Trait Loci (QTLs) in *L. campestre* for, "plant height, number of stems per plant, stem growth orientation, flowering habit, earliness, seed yield per plant, pod shattering resistance, and perenniality". Genes homologous with *Arabidopsis thaliana* have been found, based on which genetic coding regions for *L. campestre* were identified (Geleta et al. 2020). Research on genetic variation in different genetic coding regions, and the effects on amino acids and protein synthesis may assist in the further improvement of a field cress breeding line for future commercial use.

There is also ongoing research on the future potential use of field cress seed oil in nutritional products for human and animal consumption. Research focusing on improving the oil composition of field cress for nutrition has successfully used CRISPR-Cas9 to increase oleic content and decrease linoleic content of *L. campestre* seeds (Ivarson et al. 2016; Sandgrind et al. 2023). However, the use of organisms modified by new genomic technologies (NGTs), such as CRISPR, is strictly regulated according to EU law, and research is currently only possible under highly controlled conditions (Stokstad 2024).

2.2 Plant Breeding

Domestication of crops occurs when wild species are selected by humans and grown in new environments, leading to their evolutionary adaption to new environmental conditions (Purugganan 2019). Domestication is closely connected to the process of humans purposefully selecting plants based on beneficial traits (Bernal-Gallardo and de Folter 2024). Plant breeding has historically been conducted via decade long processes, starting with selection of parent plants based on observation of plant attributes in the field, followed by targeted crossbreeding of the selected parent plants, and testing of subsequent hybridized generations for the desired traits (Hasan et al. 2021). The goal of this process is the establishment of an elite cultivar that exhibits all desirable traits for human cultivation.

The use of molecular biological advances in plant genetics has led to a detailed understanding of plant genomes, shifting the plant breeding process from phenotype-based to genotype-based. Since the complete genome sequencing of the very first plant genome, *Arabidopsis thaliana*, in 2000, the genomes of 1482 plant species have been sequenced as of March 2024 (Bernal-Gallardo and de Folter 2024). These genome sequences enable researchers to map genomes, perform association studies and identify associated traits for specific genes, giving new insights into historic domestication processes and evolution (Bernal-Gallardo and de Folter 2024).

A genetic marker is "a DNA sequence with a known physical location on a chromosome" (NCI 2024). Genetic markers help identify genetic diversity among individuals of the same species, by aiding in finding, for example, the location of heterozygous alleles (Collard and Mackill 2008). When genetic markers are compared across accessions to identify polymorphisms, meaning genetic variation that can be used to track, select for or modify desirable or undesirable traits in breeding, this process is called marker assisted selection (MAS) (Martínez-Gómez 2019). MAS can accelerate breeding and novel domestication processes by changing the selection process of individuals for crossbreeding, thereby reducing the need for lengthy field studies to identify viable individuals (Jewel et al. 2019).

Different types of genetic markers include, simple sequence repeats (SSRs), sequence tagged sites (STSs), sequence characterized amplified regions (SCARs), and single nucleotide polymorphism (SNPs) (Collard and Mackill 2008). One example of an SNP could be a biallelic variation between individuals of the same species. While some individuals within a species are homozygotes with regards to a trait, carrying identical alleles in one genetic locus linked to this trait, some individuals are heterozygotes, carrying two different alleles at this genetic locus (Evert et al. 2013). Different alleles are linked to different phenotypes and may either exhibit dominance over one another or show incomplete dominance leading to intermediate phenotypes (Evert et al. 2013).

A central challenge in plant breeding is ensuring that desired genetic traits are consistently inherited and expressed in hybrid individuals and subsequent generations (Zhong et al. 2019). Especially when targeting multiple plant traits, hybridization can require several generations to stabilize the desirable genetic combination (Hammenhag et al. 2020b). Despite advances in molecular breeding and MAS crossbreeding, the goal of establishing homozygous lines still involves a degree of selection and optimization over multiple generations. Genetic traits and their expression can depend on environmental factors, such as photoperiod, temperature, precipitation and soil nutrients (Hammenhag et al. 2020b). This further increases the uncertainty of gene expressions in the field, despite targeted selection in laboratory or greenhouse conditions.

2.3 Desirable Characteristics of Field Cress

A single genetic marker can influence multiple traits (pleiotropy), and multiple genetic regions can interact to determine a single trait (epistasis) (Evert et al. 2013). Current plant breeding efforts for field cress are investigating a wide range of traits, to which pleiotropy and epistasis add complexity. Desirable characteristics in domestication and plant breeding efforts for field cress include hardiness against harsh winter climate, high seed yield, oil content/composition, reduced levels of anti-nutritional compounds, as well as pod shattering resistance, synchronous flowering, and early seed production (Hammenhag et al. 2020a).

Being developed as an oilseed crop, achieving high oil content has a high priority in field cress plant breeding efforts (Hammenhag et al. 2020b). Furthermore, oil composition is a chief characteristic, determining suitable areas of commercial use for field cress oil. Table 1 gives an overview of the fatty acid composition of wild collected field cress from different studies. Fiel cress is high in erucic acid C22:1, (EA) comprising between 23%-27% of wild field cress seed oil (Nilsson et al. 1998; Sandgrind et al. 2023; Lodenius 2023). EA is suitable for sustainable aviation fuel (SAF), because of its long carbon chain which makes it energy dense (Redda et al. 2024). Plant oils high in EA can also be used to produce sustainable plastic bags, cosmetics, and detergent (Harwood et al. 2017).

When considering field cress oil for commercial use in the food industry on the other hand, high levels of EA and polyunsaturated fatty acids (PUFAs) are not desirable. High amounts of PUFAs generally make oil prone to oxidation, thereby unstable, reducing its shelf life (Sandgrind et al. 2023). High oleic oils, meaning monounsaturated fatty acids, are less prone to oxidation, and therefore preferable for commercial use (Lee et al. 2013).

Oleic composition is one of the main traits currently targeted by plant breeding efforts for nutritional consumption of field cress (Sandgrind et al. 2023). Wild field cress has relatively low amounts of monounsaturated fatty acids, and high

amounts of PUFAs, as can be seen in Table 1. In contrast, cultivated *B. napus* lines for nutritional consumption have monounsaturated fatty acid contents of 60-70% (Lee et al. 2013).

Table 1. Fatty acid composition of Lepidium campestre seeds, measured in Nilsson et al. (1998), Sandgrind et al. (2023), and Lodenius (2023)

Fatty acid	**Percentage content in wild *L. campestre* seeds**		
	Lodenius (2023)	Sandgrind et al. (2023)	Nilsson et al. (1998).
Myristic acid (14:0)	0.1%	*	*
Palmitic acid (16:0)	4.6%	5%	4.7%
Stearic acid (18:0)	0.7%	1%	1.1%
Oleic acid (18:1)	12.9%	13.3%	14.4%
Linoleic acid (18:2)	9.9%	8.8%	9.6%
Linolenic acid (18:3)	39.1%	38.4%	35%
Arachidic acid (20:0)	0.6%	0.5%	0.6%
Eicosenoic acid (20:1)	5.5%	4.8%	5.4%
(20:2)	*	*	0.6%
Behenic acid (22:0)	0%	1%	1.4%
Erucic acid (22:1)	26.7%	24%	23.1%
(22:2)	*	*	0.3%
Lignoceric acid (24:0)	*	0.5%	0.7%
Nervonic acid (24:1)	*	3%	2.1%
Others	*	*	0.8%

*not measured

Furthermore, the toxicity of seeds is of concern for plant breeding efforts for commercial use in the food industry or as animal feed. Brassicaceae seeds, typically containing high levels of EA (C22:1), are unsuitable for human and animal consumption, as this fatty acid leads to heart disease (Lee et al. 2013). Glucosinolates, substances that are also prevalent within the Brassicaceae family, are toxic to mammals, and therefore not desirable should the biproducts of oil pressing be used as animal feed (Arefaine et al. 2019).

2.4 Genes linked to Production and Composition of Fatty Acids

There are several genes responsible for fatty acid biosynthesis in most Brassicaceae plants, among others genes steering oil production itself, such as *ACYL-COA:DIACYLGLYCEROL ACYLTRANSFERASE 1 (TAG1)* and *WRINKLED 1 (WRI1)*, and genes linked to oil composition, such as *FATTY ACID DESATURASE 2 (FAD2), FATTY ACID ELONGASE 1 (FAE1)* and *3-KETOACYL-COA-SYNTHASE 8 (KCS8)* (Gustafsson 2018).

The *FAD2* gene encodes an enzyme that catalyses the desaturation of oleic acid (C18:1) in Brassicaceae species, leading to the synthesis of linoleic acid in the endoplasmic reticulum and the storage of polyunsaturated fatty acids in cell membranes and seeds (Lee et al. 2013). While the gene *FAE1* steers oleic content by elongating monounsaturated fatty acids to erucic acids (C22:1), *FAD2* is responsible for desaturating monounsaturated fatty acids (C18:1) to polyunsaturated fatty acids (C18:2) (Sandgrind et al. 2023). *FAD3* then desaturates these fatty acids further (C18:3) (Sandgrind et al. 2023).

Silencing the *FAE1* and *FAD2* genes in different plant species has shown varying effects on oleic content, indicating that the oleic regulation process associated with the *FAE1* and *FAD2* genes is specific to each species (Ivarsson et al. 2016). The specific functions of *FAE1* in combination with *FAD2* in *L. campestre* have been confirmed by studies that successfully silenced these genes, resulting in higher oleic acid content and reduced PUFAs and EA in field cress seeds (Ivarsson et al. 2016; Sandgrind et al. 2023).

Pleiotropy and epistasis of *FAD2* and *FAE1* with other genes in *L. campestre* are not fully understood by researchers yet. Ivarsson et al. (2016) note that several of the transgenic plants with silenced *FAD2* and *FAE1* genes had poor or no seed production, and some did not produce flowers. The exact reason for this has not been established (Ivarsson et al. 2016). Another trait possibly linked with fatty acid composition is winter hardiness. A study in *A. thaliana* has shown that lower levels of PUFAs were associated with lower cold tolerance (Miquel et al. 1993). This finding could be significant for plant breeding efforts if it can be confirmed for *L. campestre*, making PUFAs essential if they are shown to contribute to cold tolerance. A future field cress variant for large scale cultivation in northern Sweden must survive harsh winter climate, and its winter hardiness is the very reason why field cress was selected as a suitable species for northern Sweden (Hammenhag et al. 2020b).

3. Methodology

3.1 Sampling and DNA Extraction

Prior to this study, plant seeds from wild field cress specimen, *L. campestre*, were collected in different locations in Sweden and Denmark (see Appendix 1, Table 2 and Figure 6, for an overview of collection sites). One of the sixteen accessions, NEH006, is of the closely related species *L. heterophyllum*, also known as Smith's cress. Leaf tissue was sampled from three-week-old plants, that were grown under greenhouse conditions. The leaf tissue samples were frozen in liquid nitrogen, homogenized with glass beads in a Retsch MM400 shaker for 60s at 30Hz, and subjected to DNA extraction using a kit (GeneJET Plant Genomic DNA Purification kit, according to protocol, "Purify DNA from liquified, polyphenol-rich plant tissue or rapeseed"). DNA quality was tested with 1.8% Agarose gel electrophoresis, and the quality and concentration of the extracted DNA was tested on a ds-11 FX DeNovix Nanodrop Spectrometer (see Appendix 2 for detailed gel electrophoresis method). One DNA template per accession was selected for Polymerase Chain Reaction (PCR), using the template with the highest DNA concentration and quality per accession. Criteria set for the selection of DNA was DNA concentration >15ng/µl and the 260/280 ratio between 1.80-2.00, measured on a Nanodrop Spectrometer.

3.2 PCR and Sequencing

Using the extracted DNA of 16 accessions, two specific genomic regions in the *FAD2* gene were amplified using PCR, with a reaction volume of 10µl, composed of 6µl H_2O, 2µl HF x5 Phusion Buffer, 0.2µl of 10mM dNTP, 0.15µl of 10µM forward Primer, 0.15µl of 10µM reverse Primer, 0.1µl Polymerase, and 1µl DNA. Two primer pairs designed by Lodenius (2023), LepiFAD2-1a and LepiFAD2-2a, were used to amplify the entire *FAD2* coding region for each accession, where each primer amplified one of two adjacent segments of 726bp and 794bp, respectively (see Appendix 3, Table 5 for primer sequences). These two primers were used. PCR cycling conditions were identical for both primers, with a cycle of denaturation (98°C), annealing (60°C) and extension (72°C), repeated 40 times, as illustrated in Figure 1. The quality and concentration of the amplified PCR products were examined using agarose gel electrophoresis and a Nanodrop Spectrometer (see Appendix 3 for details on the method of gel electrophoresis used).

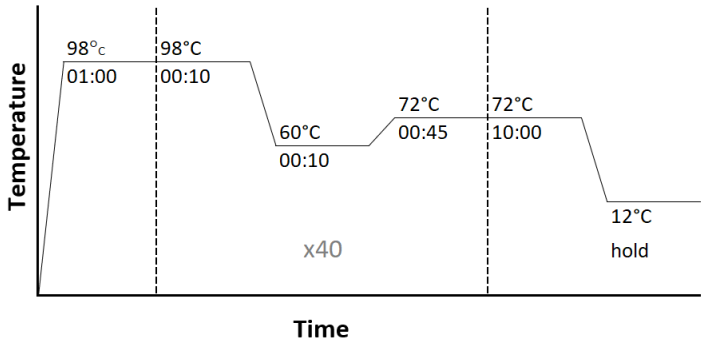


Figure 1. Polymerase Chain Reaction (PCR) Cycle. The cycling conditions depicted were used for both primer pairs. Depicted between dotted lines, the cycle of denaturation at 98 ℃, annealing at 60 ℃ and extension at 72 ℃, was repeated x40.

PCR products were purified using a kit (Thermal Scientific GeneJET PCR Purification, Protocol A). The quality of these purified PCR products was assessed based on their DNA concentration on the Nanodrop Spectrometer. Purified PCR products with a DNA concentration <5ng/µl were not selected for sequencing, and the PCR cycle for these accessions was repeated with double the PCR reaction volume of 20 µl to yield higher volumes of DNA. Purified PCR products from 16 field cress accessions were sent to Eurofins Genomics, Germany for Sanger sequencing of the two gene segments comprising the *FAD2* coding region. Each DNA sample was sequenced in the forward and reverse directions, using the same primers as in the PCR reaction.

3.3 Sequence Analysis

Only sequences which passed Eurofins' quality criteria with an average quality score of above 30 were analysed. For sequences, which did not meet the quality criteria, the PCR process was repeated. The obtained sequences were aligned with the known gene sequence for the *FAD2* gene of *L. campestre*, available in the online data bank of the National Center for Biotechnology Information (NCBI), to confirm that the correct genomic region had been amplified. Bioinformatics software ClustalX and BioEdit were used to align genetic sequences and evaluate marker diversity, assessing polymorphisms within the amplified genetic region. To assess genetic diversity and determine whether the newly collected samples deviate from the existing collection, the *FAD2* gene sequences were compared with the genomic sequences of 9 accessions of *L. campestre* from a previous study by Lodenius (2023), giving a total number of genetic sequences of 25.

4. Result

4.1 DNA Extraction and PCR Results

The majority of the extracted DNA templates selected for PCR cycling met the quality criteria set in this study (15ng/µl DNA concentration and 1.80-2.00 260/280 ratio). The templates with the relatively highest quality for each accession were selected for PCR (see Appendix 2, Table 3 for Nanodrop Spectrometer results of the selected DNA templates). Despite three of the templates not meeting the quality criteria, with a 260/280 ratio outside the 1.80-2.00 span, PCR was attempted with these templates to ensure that the FAD2 gene could be amplified in all accessions. The 260/280 ratios outside the 1.80-2.00 span for three DNA templates indicates impurities in the DNA template.

Agarose UV images after electrophoresis of PCR products showed a band for all PCR products except one, CVW004, for which the PCR process was repeated (see Appendix 3, Figure 8 for a UV image of agarose gel for electrophoresis of PCR products). An error likely occurred while mixing the PCR reaction mix for CVW004 which showed no band, for example, the DNA template may not have been added to the PCR mix by mistake.

For all other PCR products, agarose gel electrophoresis resulted in a single band in the expected position at 726bp for primer LepiFAD2-1a and 794bp for primer LepiFAD2-2a. However, despite the expected bands showing after electrophoresis, most of the samples had unspecific bands that showed relatively broad in the electrophoresis result (for example, see TRU001 and TRU002 in Appendix 3, Figure 8). These unspecific bands could either indicate lower DNA quality or overloading of the agarose gel pockets.

However, because the expected band was present for all PCR products and the Nanodrop Spectrometer measured high DNA concentrations between 105-605 ng/µl as well as very good values for purity and quality of the DNA, with 260/280 ratios close to 1.80, the PCR products were purified and sent for sequencing, despite the broadness of the electrophoresis bands.

4.2 Sequence Alignment and Genetic Variation in *FAD2*

When aligning the forward and reverse sequences of two accessions (NEH002 and TRU001) with the confirmed *FAD2* gene sequence for *L. campestre* via the NCBI BLAST tool, the E-value was 0.0, indicating a very high likelihood that the

sequences match the standard *FAD2* sequence. Alignment scores were above 200 for each subject confirming the high congruency between the accession sequences and the *FAD2* sequence for *L. campestre*. For each accession, between 96% and 100% of the subject nucleotides matched those in the query sequence. Since these values indicate a high congruency between the accession sequences and the reference *FAD2* gene sequence for *L. campestre*, the findings confirm that the targeted gene region was successfully amplified in the PCR. These values also show that the accession sequences match existing *L. campestre* gene sequences, which confirms that the DNA extraction from sampled plant tissue was successful.

Query coverage for the accession sequences, generated using the primer pair LepiFAD2-1a, was 48% (NEH002) and 45% (TRU001), when compared to the *FAD2* sequence for *L. campestre*, available at NCBI (accession number FJ907546). Coverage for the primer pair LepiFAD2-2a was 61% and 62% respectively. When the two sequences resulting from the different primers were combined for each of these two accessions, query cover was 100% and 98% respectively, confirming that, together, the two gene segments amplified by the primers represent the complete coding region of the *FAD2* gene.

Complete alignment of the forward and reverse complement sequences of all sixteen accessions revealed two polymorphisms within the *FAD2* gene, firstly an INDEL composed of three nucleotides in two accessions, and secondly a heterozygous locus in two accessions. Figure 2 shows the location of the polymorphisms found in the coding region of the *FAD2* gene in the sixteen *L. campestre* accessions, compared with polymorphisms in this region found by Lodenius (2023). The two accessions (NEH002 and CVW005) which were found to have an ambiguous base at one locus in the second segment of the *FAD2* coding region, do not have the trinucleotide deletion, and vice versa.

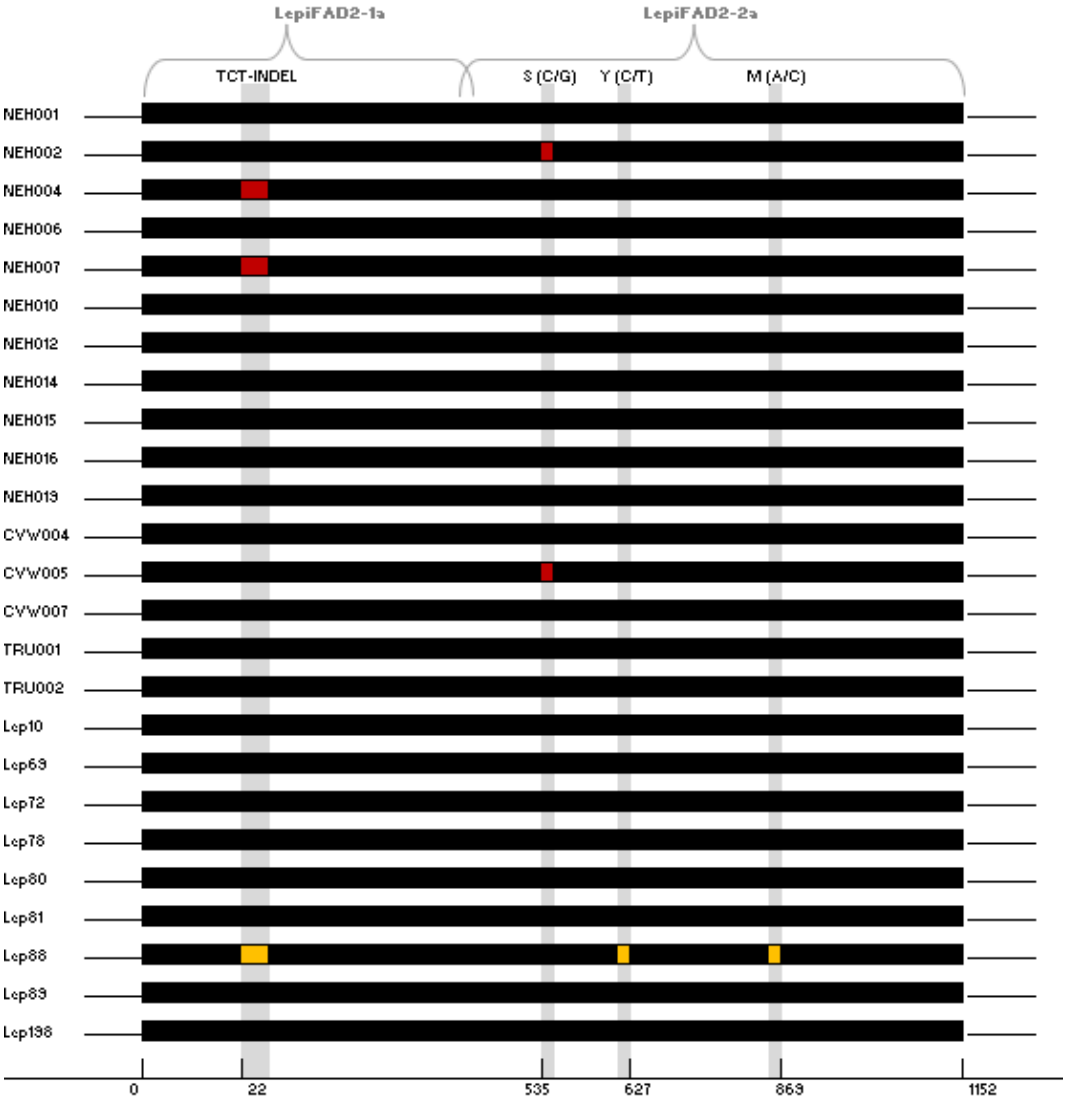


Figure 2. Illustration of the locations of the polymorphisms in the FAD2 coding region found in this study in the 16 accessions of L. campestre. Samples labelled "Lep" are from Lodenius (2023). Black lines represent homogenous loci, red lines represent heterogeneous loci in the 16 accessions in this study, and yellow lines represent heterogenous loci in Lodenius' (2023) accessions. In the second segment of the FAD2 coding region, heterozygous loci with ambiguous bases S, M and Y have been identified.

4.3 TCT Trinucleotide INDEL

In the gene region amplified by the primer pair LepiFAD2-1a, a trinucleotide INDEL on the sequences generated by both the forward and reverse strands, was found in two samples (NEH004 and NEH007). Figure 3 shows an excerpt of the alignment of gene sequences, where it is visible that for both accessions which have the trinucleotide deletion, codons TCT are absent at the same locus in the forward and reverse strands. Both accessions which show this variation were

collected at different locations, Halmstad and Haverdal in Sweden, 17km apart (see Appendix 1, Figure 6 for a map of collection sites).

When aligning the gene sequences with *L. campestre* sequences from Lodenius (2023), the TCT-INDEL was the only congruent finding among the total of 25 accessions. Figure 3 shows one of the nine accessions from Lodenius (2023) has this polymorphism at the same locus as two accessions sequenced in this study. Lodenius' (2023) accession, Lep88, was collected in Eisenach, Germany (see Appendix 1, Table 2 for collection sites). Three out of 25 accessions (12%) aligned in this study have a TCT-deletion.

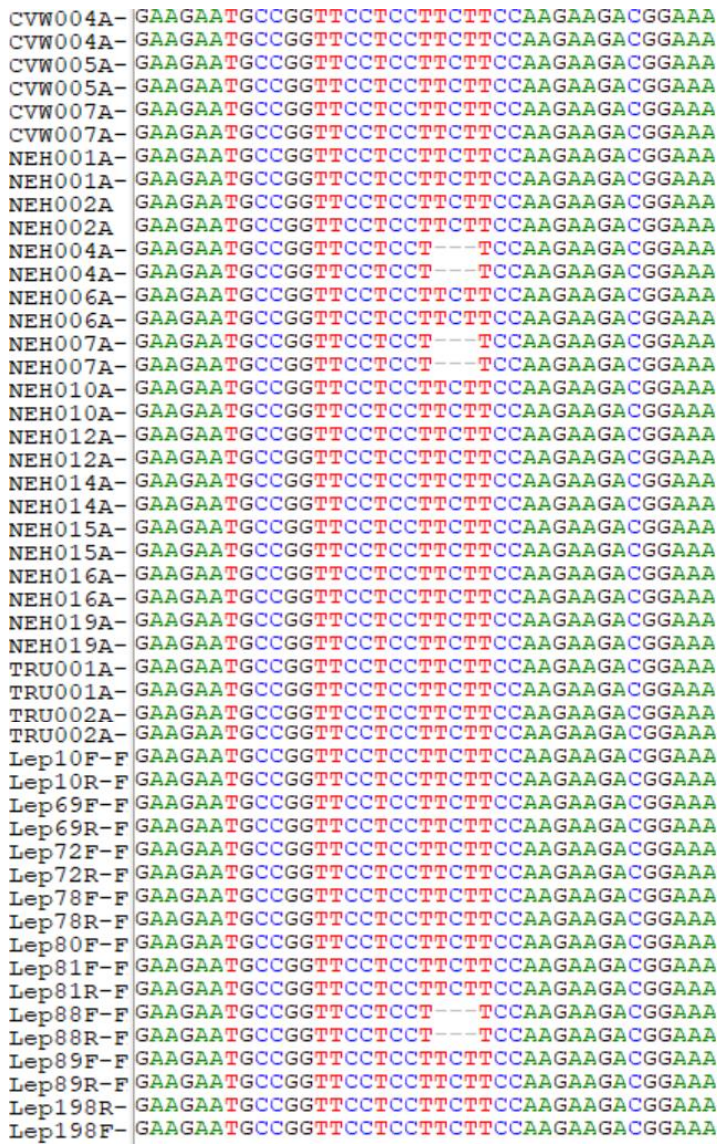


Figure 3. Excerpt of the sequence alignment for the FAD2 gene L. campestre. The sequence labels which end in "F" and "R" represent the forward and reverse strands of the same DNA template. All sequences aligned above are in the '5 to '3 direction. The sequences named "Lep" are L. campestre sequences from Lodenius (2023).

In Figure 4, excerpts of the gene sequence chromatograms show the locus of the TCT-INDEL. The chromatograms depict the similarity in the curves' magnitudes between the two gene sequences with the TCT-deletion (Figure 4a-b).

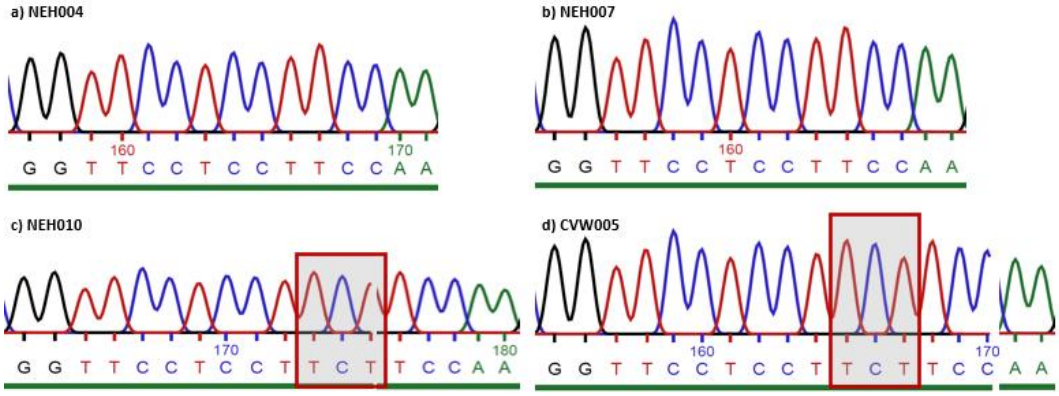


Figure 4. Sanger sequencing chromatograms showing a trinucleotide INDEL (TCT) in four accessions. (a) NEH004 and (b) NEH007 exhibit the deletion, as indicated by the absence of the TCT peak. (c) NEH010 and (d) CVW005 represent all other accessions sequenced in this study, where the TCT trinucleotide is present.

4.4 Heterozygous locus

In two samples (CVW005 and NEH002) a biallelic polymorphism was identified in the form of a heterozygous locus in the second segment of the *FAD2* gene. The chromatograms in Figure 5 show the heterozygous locus with an ambiguous base S (nucleotide code representing C or G), present only in the forward gene sequences of the two accessions. Despite the variation only being present in the forward sequence, it was identified in two separate accessions and is therefore less likely to be a sequencing error than a heterozygous locus. A possible reason for why the reverse sequence does not show the ambiguous base could be that the reverse primer might not be able to bind to both alleles.

The chromatograms show that the curves at the heterozygous locus are of similar magnitude, indicating that the variation is unlikely to be caused by background disturbances. All other samples have an unambiguous nucleotide base C at the same locus. The samples, which have this variation were collected at different sites, 250km apart, one in Denmark and one in Sweden (see Appendix 1, Figure 6 for a map of collection sites).

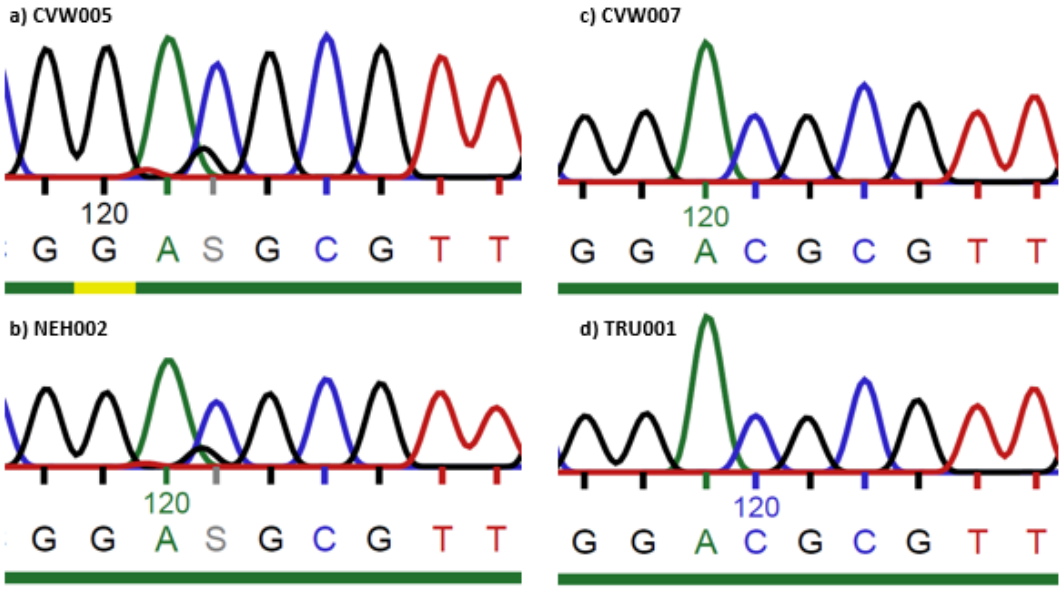


Figure 5. Heterozygote identified in the FAD2 coding region of two accessions. Curves a and b depict the two accessions that have this biallelic polymorphism with base S (C or G). Curves c and d represent all other samples, which have an unambiguous base C at this locus. The labelling of the base pair number is slightly shifted for each sample due to trimming of low-quality results at the beginning of some sequences.

The trinucleotide INDEL was the only congruent finding identified when aligning the total of 25 accessions, including nine from Lodenius (2023). The sixteen accessions sequenced in this study were homogenous at the loci of two biallelic polymorphisms found by Lodenius (2023). Neither in the alignment of BIOEdit were any ambiguous bases found at the loci specified by Lodenius (2023), nor did the chromatographs of each *FAD2* sequence for this study give any indications for heterozygous alleles at the given loci. Conversely, at the locus of the biallelic locus identified in this study in the *FAD2* coding region, the nine sequences from Lodenius (2023) were homogenous.

In the forward sequence of four accessions of this study (CVW005, TRU001A, NEH002, NEH015), a second curve on the chromatograph is visible, adjacent to the locus of a polymorphism identified by Lodenius (2023). The base at this locus was not identified as ambiguous in any of the samples, so that the sequence alignment shows no variation, and the second curve therefore is likely not indicative of a heterozygous locus. The second chromatogram curve has a much lower magnitude than the curve of the identified base A (see Appendix 4, Figure 9). Though a low magnitude second curve is present at the same locus in four samples, it is likely a disturbance, either due to an adjacent nucleotide G, which shows a high magnitude curve, or due to errors in the PCR process, such as polymerase slippage. Such a disturbance could also be caused by impurities in the DNA templates.

5. Discussion

Genetic differences were found among the sixteen accessions, thereby the research hypothesis is accepted. Through sequence alignment of a total of 25 accessions of *L. campestre* of the genetic region representing the *FAD2* coding region, one INDEL and one heterozygous locus were identified. Other than the TCT-INDEL, no other SNPs identified in previous studies could be confirmed in the 16 accessions sequenced in this study.

Overall, the findings indicate that the *FAD2* coding region of *L. campestre*, even when compared with one sequence of *L. heterophyllum*, is highly homogenous, despite the diversity of seed collection sites. As the genetic region amplified in this study is a genetic coding region, little genetic variation is to be expected between different individuals of the same species. Gustafsson et al. (2018) also found low genetic variation in the *FAD2* gene of *L. campestre* and the closely related species *L. heterophyllum* and *L. hirtum*. From a plant breeding perspective, the high degree of genetic homogeneity may pose difficulties in identifying suitable individuals for hybridization when targeting fatty acid composition.

L. campestre has a high tendency toward self-fertilization (Desta et al. 2019). Self-fertilization often leads to a higher frequency of homozygotes and a lower frequency of heterozygotes in a population (Evert et al. 2013). The 25 gene sequences of *L. campestre* aligned in this study reflect the low frequency of heterozygotes. However, with the increase in homozygotes, the frequency of alleles in the population does not change (Evert et al. 2013). Therefore, larger samples might simply be necessary for identifying heterozygotes and polymorphisms in *L. campestre*, due to the high homogeneity in the *FAD2* gene.

The evolutionary development of *A. thaliana*, a self-fertilizing plant with which *L. campestre* shares a common evolutionary ancestor, is dominated by purifying selection, where polymorphisms that lead to undesirable traits are passed down through generations at a lower frequency (Foxe et al. 2008). Therefore, there is a high likelihood that traits not furthering species survival will not be passed on to subsequent generations. A gene with a central function for plant metabolism, such as *FAD2*, likely shows high homogeneity as a result of evolutionary de-selection of traits that disturb fatty acid synthesis.

Furthermore, from and evolutionary standpoint, fatty acid composition may play an important role for the survival of plants that have high lipid content. A study by Miquel et al. (1993) observed mutants of *A. thaliana* that show low levels of PUFAs in leaf and root tissue compared to the average wild type. With 16% 18:2 plus 18:3 in leaf tissue and 7-9% in root tissue, stem elongation of plants at 12°C failed to occur, and at 6°C the low PUFAs plants were inviable (Miquel et al. 1993). Similar links between PUFAs and cold sensitivity were

found in other species. Nampoothiri et al. (2021) found a link between the overexpression of genes linked to fatty acid synthesis and temperature sensitivity in bacteria. Nieminen et al. (2013) observed an increase in PUFAs in the ectoparasite deer ked (*Lipoptena cervi*) between the summer and autumn seasons, indicating that cells underwent a process of cold-induced desaturation. The generally high levels of PUFAs in Brassicaceae wild types (Lee et al. 2013) might have cold resistance as an evolutionary purpose.

Despite the overall lack of genetic diversity found in this study, two polymorphisms were identified in the sixteen accession sequences. The trinucleotide INDEL found in two accessions, has been previously identified by Gustafsson et al. (2018) and Lodenius (2023), located in the first segment of the *FAD2* gene. Lodenius (2023) found this genetic variation to be polymorphic in *L. campestre*. According to Gustafsson (2018), the deletion results in the loss of a serine amino acid. By examining the codon sequence of the *FAD2* gene sequence for *L. campestre*, it was confirmed that the INDEL is congruent with the locus of the codon TCT, one of the six codons coding for serine. What the absence of one serine amino acid means for enzyme function of these field cress accessions is uncertain and could be explored in further studies of that accessions which have a TCT deletion.

Genetically diverse individuals for crossbreeding are necessary for conventional plant breeding methods. The presence of one heterozygous locus in two of the accessions in this study, despite the overall homogeneity, therefore, might be useful if its dominant expression and function can be determined. Yet the fact that the biallelic variation found in this study has not been identified in other studies might indicate that these variations are rare and difficult to find. The large distance between the collection sites of the seeds of individuals showing the same genetic variation, however, indicates that the genetic variations are unlikely to be isolated spontaneous mutations.

Since the heterozygous locus was found in a coding region, there is a possibility that it impacts protein synthesis and fatty acid synthesis. The ambiguous base S (nucleotide code for C and G) causes a non-synonymous mutation as it was identified in the first position of the codon CGC, which codes for arginine. The second variant of this codon GGC codes for glycine.

A single amino acid substitution can decrease, amplify or disrupt the functions of a protein (NCBI 2025). Examples of effective single amino acid substitutions have been identified in *A. thaliana*, where certain single amino acid substitutions have been linked to both temperature dependent protein activation (Xiong et al. 2004) and cell growth defects (Xiong et al. 2013). However, amino acid substitution with functionally equivalent substitutes might not have any effect on protein synthesis (NCBI 2025). Prediction of the functional effect of a single

amino acid substitution on an organism can be done by investigating other mutations at the same site, given the mutations' effects are known (Zabin et al. 1991).

Determining whether the arginine amino acid and the substituting glycine in the two heterozygous accessions are functionally equivalent in field cress requires further research. Furthermore, whether the heterozygous locus identified in the two accessions is linked to differing traits and whether one trait exhibits dominant expression over the other is unknown. Growing plant specimen from the seeds of the accessions that have the heterozygous locus, and observing the traits of subsequent generations could provide an answer. One would expect a recessive trait to appear in one out of four individuals of the second filial generation in an environment of controlled self-pollination (Evert et al. 2013).

Both types of polymorphisms found were present in two accessions, which reduces the likelihood that the variations result from a faulty PCR amplification or sequencing error. Furthermore, the samples with the trinucleotide deletion do not show any heterozygous alleles and vice versa, which also reduces the likelihood that a mistake in the PCR process led to the variations in gene sequences. PCR errors would have likely resulted in a more random distribution of both types of variations across all samples, rather than this clear separation.

However, this study did not verify whether the samples showing genetic variation were truly distinct or if they might be genetically identical due to handling errors, such as accidental seed mixing during sowing or mislabelling of plants in the greenhouse. DNA samples could also have been mislabelled in the transferring processes between tubes. Replication studies are needed to confirm whether the heterozygous alleles can be found in other individuals of *L. campestre*.

6. Conclusion

Low genetic variation in the *FAD2* coding region of *Lepidium campestre* can pose a challenge to plant breeders when targeting the fatty acid composition of field cress seeds. Few individuals show polymorphisms and the impact on protein synthesis and oil composition of the genetic variation found is unknown. Polymorphisms in other genes involved in oil composition and synthesis could be examined for the 16 accessions used in this study, searching for possible links between polymorphisms and gaining insight into associated plant characteristics. Identified in accessions in this study, the TCT-INDEL and the biallelic polymorphism and their effect on protein synthesis could be further investigated, for example via a vector in a yeast expression system testing its ability to transform oleic acid (18:1) to linoleic acid (18:2). Future research could also investigate the oil composition of seeds of the accessions in this study to find possible indications for variations linked to the identified polymorphisms.

Gaining knowledge on the genetic diversity within the field cress species is an important part of catalysing the commercialization of this promising future oil crop. Plant oils exemplify the principles of a circular economy, with renewable fuel or nutritional oil as a primary product, and high protein biproducts to the oil pressing process as animal fodder. Field cress as a new oil crop has the potential to help meet SDG 7, expanding the global share of renewable energy sources.

References

Arefaine, H., Rydhmer, L., Andersson, R. & Ivarsson, E. (2019). Lepidium cake as a feedstuff for pigs. *Livestock Science*, 225, 47–52. https://doi.org/10.1016/j.livsci.2019.04.022

Bernal-Gallardo, J.J. & de Folter, S. (2024). Plant genome information facilitates plant functional genomics. *Planta*, 259 (5), 117–117. https://doi.org/10.1007/s00425-024-04397-z

Collard, B.C.Y. & Mackill, D.J. (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philosophical Transactions of the Royal Society of London. Series B. Biological Sciences*, 363 (1491), 557–572. https://doi.org/10.1098/rstb.2007.2170

Desta, Z.A., Kolano, B., Shamim, Z., Armstrong, S.J., Rewers, M., Sliwinska, E., Kushwaha, S.K., Parkin, I.A.P., Ortiz, R. & de Koning, D.-J. (2019). Field cress genome mapping: Integrating linkage and comparative maps with cytogenetic analysis for rDNA carrying chromosomes. *Scientific Reports*, 9 (1), 17028. https://doi.org/10.1038/s41598-019-53320-0

Energy Institute (2024), "Statistical Review of World Energy", 73rd edition. Accessed: 01/03/2025, https://www.energyinst.org/statistical-review

Evert, R.F., Eichhorn, S.E. & Raven, P.H. (2013). Raven Biology of Plants. Eighth edition, [International edition]. *W.H. Freeman*.

Foxe, J.P., Dar, V.-N., Zheng, H., Nordborg, M., Gaut, B.S. & Wright, S.I. (2008). Selection on Amino Acid Substitutions in *Arabidopsis*. *Molecular Biology and Evolution*, 25 (7), 1375–1383. https://doi.org/10.1093/molbev/msn079

Geleta, M., Gustafsson, C., Glaubitz, J.C. & Ortiz, R. (2020). High-Density Genetic Linkage Mapping of *Lepidium* Based on Genotyping-by-Sequencing SNPs and Segregating Contig Tag Haplotypes. *Frontiers in Plant Science*, 11, 448–448. https://doi.org/10.3389/fpls.2020.00448

Gustafsson, C., Willforss, J., Lopes-Pinto, F., Ortiz, R. & Geleta, M. (2018). Identification of genes regulating traits targeted for domestication of field cress (*Lepidium campestre*) as a biennial and perennial oilseed crop. *BMC Genetics*, 19 (1), 36–36. https://doi.org/10.1186/s12863-018-0624-9

Hammenhag, C., Sandin, P., Zhu, L. H., Stymne, S., Geleta, M. (2020a). Fältkrassing - en framtida oljegröda för Norrland. *Sveriges Utsädesförenings Tidskrift*, 2, 6–15. https://www.sverigesutsadesforening.se/data/pdf/sut_2020_2.pdf

Hammenhag, C., Saripella, G.V., Ortiz, R. & Geleta, M. (2020b). Qtl mapping for domestication-related characteristics in field cress (*Lepidium campestre*)— a novel oil crop for the subarctic region. *Genes*, 11 (10), 1–20. https://doi.org/10.3390/genes11101223

Harwood, J.L., Woodfield, H.K., Chen, G., Weselake, R.J. & Ahmad, M.U. (2017). Modification of Oil Crops to Produce Fatty Acids for Industrial Applications. *In:*

Fatty Acids. AOCS Press. 1–2. https://doi.org/10.1016/B978-0-12-809521-8.00005-2

Hasan, N., Choudhary, S., Naaz, N., Sharma, N. & Laskar, R.A. (2021). Recent advancements in molecular marker-assisted selection and applications in plant breeding programmes. *Journal of Genetic Engineering and Biotechnology*, 19 (1), 128–26. https://doi.org/10.1186/s43141-021-00231-1

Ivarson, E., Ahlman, A., Lager, I. & Zhu, L.-H. (2016). Significant increase of oleic acid level in the wild species *Lepidium campestre* through direct gene silencing. *Plant Cell Reports*, 35 (10), 2055–2063. https://doi.org/10.1007/s00299-016-2016-9

Jewel, Z.A., Ali, J., Mahender, A., Hernandez, J., Pang, Y. & Li, Z. (2019). Identification of quantitative trait loci associated with nutrient use efficiency traits, using SNP markers in an early backcross population of rice (*Oryza sativa* L.). *International Journal of Molecular Sciences*, 20 (4), 900-. https://doi.org/10.3390/ijms20040900

Lee, K.-R., In Sohn, S., Jung, J.H., Kim, S.H., Roh, K.H., Kim, J.-B., Suh, M.C. & Kim, H.U. (2013). Functional analysis and tissue-differential expression of four *FAD2* genes in amphidiploid *Brassica napus* derived from *Brassica rapa* and *Brassica oleracea*. *Gene*, 531 (2), 253–262. https://doi.org/10.1016/j.gene.2013.08.095

Lodenius, N. (2023). Identifying molecular markers for breeding a future oil crop, *Lepidium campestre*. *Swedish University of Agricultural Sciences*.

Martínez-Gómez, P. (2019). Editorial for special issue "plant genetics and molecular breeding." *International Journal of Molecular Sciences*, 20 (11), 2659-. https://doi.org/10.3390/ijms20112659

Merker, A., Eriksson, D. & Bertholdsson, N.-O. (2010). Barley yield increases with undersown *Lepidium campestre*. *Acta Agriculturae Scandinavica*. Section B, Soil and Plant Science, 60 (3), 269–273. https://doi.org/10.1080/09064710902903747

Miquel, M., James, D.J., Dooner, H. & Browse, J. (1993). *Arabidopsis* requires polyunsaturated lipids for low-temperature survival. *Proceedings of the National Academy of Sciences* - PNAS, 90 (13), 6208–6212. https://doi.org/10.1073/pnas.90.13.6208

Nampoothiri, K.M., Hoischen, C., Bathe, B., Mockel, B., Pfefferle, W., Krumbach, K., Sahm, H. & Eggeling, L. (2002). Expression of genes of lipid synthesis and altered lipid composition modulates L-glutamate efflux of *Corynebacterium glutamicum*. *Applied Microbiology and Biotechnology*, 58 (1), 89–96. https://doi.org/10.1007/s00253-001-0861-z

NCBI (2025), "Amino Acid Substitution", *National Institutes of Health*. Accessed: 28/02/2025, https://www.ncbi.nlm.nih.gov/mesh?Db=mesh&Cmd=DetailsSearch&Term=%22Amino+Acid+Substitution%22%5BMeSH+Terms%5D#:~:text=The%20naturally%20occurring%20or%20experimentally,enhance%2C%20or%20eliminate%20protein%20function

NCI (2024), “Genetic Marker”, *National Cancer Institute*. Accessed: 03/03/2025, https://www.cancer.gov/publications/dictionaries/genetics-dictionary/def/genetic-marker

Nieminen, P., Käkelä, R., Paakkonen, T., Halonen, T. & Mustonen, A.-M. (2013). Fatty acid modifications during autumnal cold-hardening in an obligatory ectoparasite, the deer ked (*Lipoptena cervi*). *Journal of Insect Physiology*, 59 (6), 631–637. https://doi.org/10.1016/j.jinsphys.2013.04.001

Nilsson, P., Johansson, S.-A. & Merker, A. (1998). Variation in seed oil composition of species from the genera *Barbarea* and *Lepidium*. *Acta Agriculturae Scandinavica*. Section B, Soil and Plant Science, 48 (3), 159–164. https://doi.org/10.1080/09064719809362493

Purugganan, M.D. (2019). Evolutionary Insights into the Nature of Plant Domestication. *Current Biology*, 29 (14), R705–R714. https://doi.org/10.1016/j.cub.2019.05.053

Redda, Z.T., Laß-Seyoum, A., Yimam, A., Barz, M. & Jabasingh, S.A. (2024). Solvent extraction and characterization of *Brassica carinata* oils as promising alternative feedstock for bio-jet fuel production. *Biomass Conversion and Biorefinery*, 14 (11), 12207–12226. https://doi.org/10.1007/s13399-022-03343-x

Sandgrind, S., Li, X., Ivarson, E., Wang, E.S., Guan, R., Kanagarajan, S. & Zhu, L.H. (2023). Improved fatty acid composition of field cress (*Lepidium campestre*) by CRISPR/Cas9-mediated genome editing. *Frontiers in Plant Science*, 14, 1076704–1076704. https://doi.org/10.3389/fpls.2023.1076704

SLU, (2023), “Cultivating sustainable energy and empowering Swedish agriculture through a novel oil crop”, *Swedish University of Agricultural Sciences*. Updated: 21/12/2023, Accessed: 27/02/2025, https://www.slu.se/en/faculties/ltv/forskningsprojekt/vaxtforadling---forskningsprojekt/odling-av-hallbar-energi-och-ett-starkt-svenskt-jordbruk-genom-en-ny-oljegroda/

Stokstad, E. (2024) “European Parliament votes to ease regulation of gene-edited crops”, *Science.org*, Published 07/02/2024. Accessed 25/02/2025, https://www.science.org/content/article/european-parliament-votes-ease-regulation-gene-edited-crops

UN (2015), “Sustainable Development Goals”, *United Nations Department of Economic and Social Affairs (UNDESA) and the Division for Sustainable Development Goals (DSDG).* Accessed on: 01/03/2025, at https://sdgs.un.org/goals

Xiong, L., Lee, H., Huang, R. & Zhu, J.K. (2004). single amino acid substitution in the *Arabidopsis* FIERY1/HOS2 protein confers cold signaling specificity and lithium tolerance. *The Plant Journal*: for cell and molecular biology, 40 (4), 536–545. https://doi.org/10.1111/j.1365-313X.2004.02225.x

Xiong, X., Xu, D., Yang, Z., Huang, H. & Cui, X. (2013). A Single Amino-Acid Substitution at Lysine 40 of an *Arabidopsis thaliana* α-tubulin Causes Extensive Cell Proliferation and Expansion Defects. *Journal of Integrative Plant Biology*, 55 (3), 209–220. https://doi.org/10.1111/jipb.12003

Zabin, H.B., Horvath, M.P. & Terwilliger, T.C. (1991). Approaches to predicting effects of single amino acid substitutions on the function of a protein. *Biochemistry* (Easton), 30 (25), 6230–6240. https://doi.org/10.1021/bi00239a022

Zhong, X., Zhou, Q., Cui, N., Cai, D. & Tang, G. (2019). BvcZR3 and BvHs1pro-1 Genes Pyramiding Enhanced Beet Cyst Nematode (*Heterodera schachtii* Schm.) Resistance in Oilseed Rape (*Brassica napus* L.). *International Journal of Molecular Sciences*, 20 (7), 1740-. https://doi.org/10.3390/ijms20071740

Appendix 1 - Seed Collection Sites

Table 2. Collection sites of accessions in this study. Accessions labelled "Lep" were sequenced by Lodenius (2023).

Species	Accession	Collection site	Closest Village
L. campestre	NEH001	Cow pasture	SE, Arlöv
L. campestre	NEH002	Gravel pit	SE, Staffanstorp Laholms kommun
L. campestre	NEH004	Gravelhill	SE, Halmstad
L. heterophyllum	NEH006	Close to gravel pit	SE, Vapnö
L. campestre	NEH007	Dock	SE, Haverdal
L. campestre	NEH010	Rocky road side	SE, Kärradal
L. campestre	NEH012	Hill next to road	SE, Sörred
L. campestre	NEH014	Parking lot	SE, Vallhamns industriby
L. campestre	NEH015	Parking lot	SE, Vallhamns industriby
L. campestre	NEH016	Parking lot	SE, Vallhamns industriby
L. campestre	NEH019	Hill next to road and train tracks	SE, Jönköping
L. campestre	TRU001	Varbergs castle	SE, Varberg
L. campestre	TRU002	Varbergs castle	SE, Varberg
L. campestre	CVW004		DK, Ore strand, Vording-borg
L. campestre	CVW005		DK, Langeskov 1A, Od-ense
L. campestre	CVW007		DK, Gjerrild myrstack, Gjerrild
L. campestre	Lep10		SE, Arrie
L. campestre	Lep69		Breeding line (dev. at SLU)
L. campestre	Lep72		Breeding line (dev. at SLU)
L. campestre	Lep78		SE, Västerås
L. campestre	Lep80		Breeding line (dev. at SLU)
L. campestre	Lep81		Breeding line (dev. at SLU)
L. campestre	Lep88		DE, Eisenach, Thüringen
L. campestre	Lep89		FR
L. campestre	Lep198		DE, Marburg

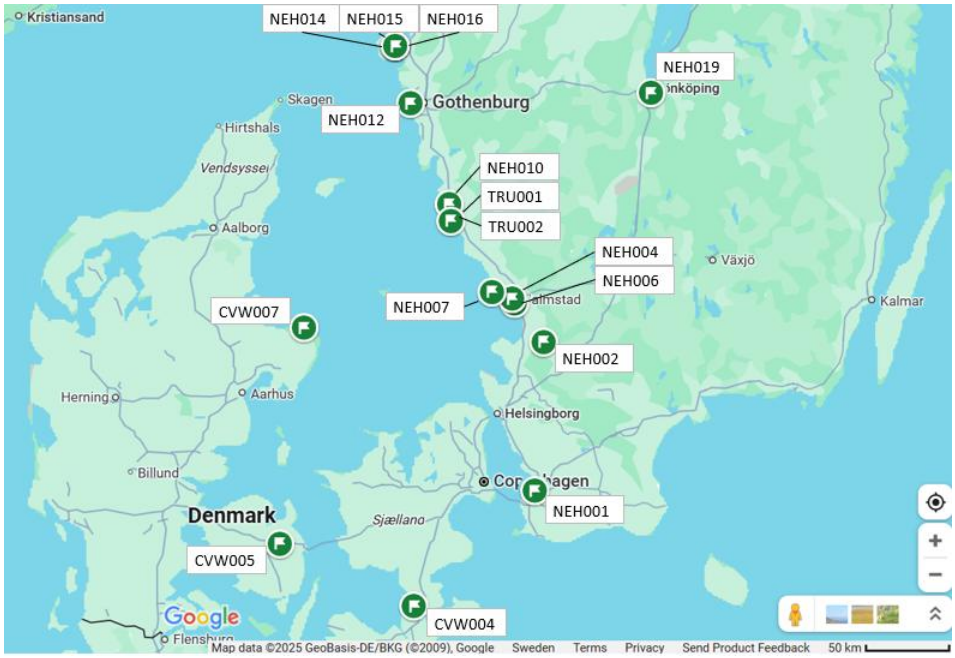


Figure 6. Map showing seed collection sites of the 16 accessions sequenced in this study. Sites were marked on Google Maps using the exact coordinates of collection sites.

Appendix 2 – Band Size, Quality and Concentration of DNA after Extraction

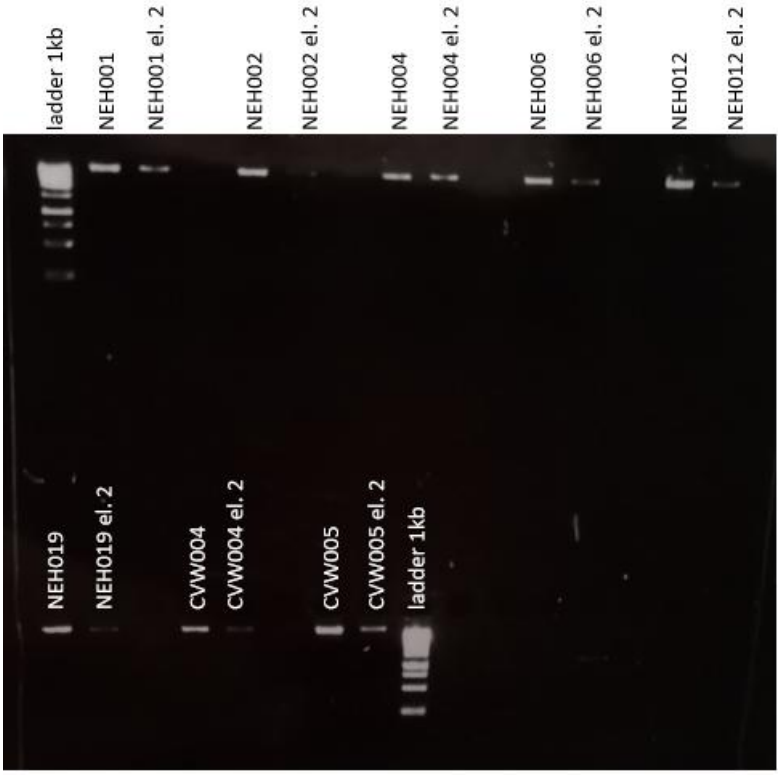


Figure 7. Example of agarose gel electrophoresis result for extracted DNA templates. Method: The loading solutions were mixed as follows: ladder = 1µl 1kb DNA ladder + 9µl H2O + 2µl loading dye; DNA = 1µl DNA template + 9µl H2O + 2µl loading dye. 3µl of the DNA ladder solution and 12µl of the DNA template solution were loaded in pockets of 1.8% agarose gel. A current of 80V was run for 30min on the EC Apparatus. The second band for each accession ("el. 2") was of a weaker DNA concentration, which resulted from a second elution during the DNA extraction process.

Table 3. Nanodrop result after DNA extraction. DNA templates were analysed on deNovix Nanodrop Spectrophotometer. Red values are outside the quality threshold set in this study for PCR cycling (>15ng/µl and 260/280 between 1.80-2.00).

Accession	DNA concentration (in ng/µl)	260/280	Accession	DNA concentration (in ng/µl)	260/280
NEH001	15.74	2.02	**NEH015**	16.209	1.91
NEH002	15.952	1.87	**NEH016**	19.159	2.26
NEH004	14.889	1.80	**NEH019**	17.917	1.86
NEH006	6.289	1.71	**CVW004**	16.029	1.92
NEH007	22.572	1.80	**CVW005**	15.891	1.78
NEH010	31.556	1.96	**CVW007**	52.053	1.84
NEH012	14.574	1.68	**TRU001**	17.027	2.23
NEH014	12.452	1.88	**TRU002**	23.515	2.02

Appendix 3 – PCR Reaction

Table 4. Composition of PCR reaction mix for both primer pairs.

	10µl-reaction volume for 22 of the PCR reactions (in µl)	**20µl-reaction volume** for 10 of the PCR reactions (in µl)
H_2O	6.4	13.8
Phusion Buffer	2.0	4.0
dNTP	0.2	0.4
Primer F	0.15	0.3
Primer R	0.15	0.3
Polymerase	0.1	0.2
DNA template	**1.0**	**1.0**

Table 5. Primer sequences, designed in a study by Lodenius (2023)

LepiFAD2-1a-F	AACGCACTTTCCATTTTTGG
LepiFAD2-1a-R	GGAAGAAATGGCTAGCGAAC
LepiFAD2-2a-F	CCTTCCTCCTCGTCCCTTAC
LepiFAD2-2a-R	TTCGCTATTCCTTCTCAATCG

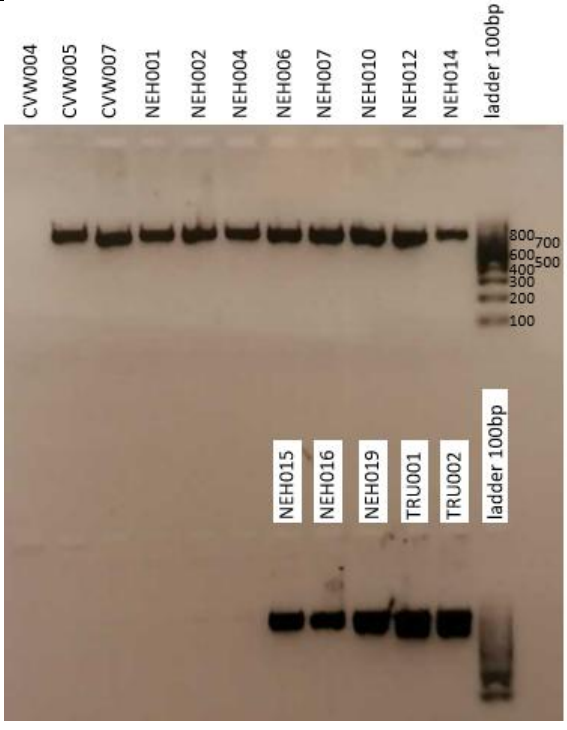


Figure 8. Agarose gel electrophoresis result for LepiFAD2-1a after PCR reaction. The expected band was 726bp for primer LepiFAD2-1a and 794bp for LepiFAD2-2a (Lodenius 2023). Method: The loading solutions were mixed as follows: ladder = 1µl 100bp DNA ladder + 9µl H2O + 2µl loading dye; DNA = 3µl DNA template + 7µl H2O + 2µl loading dye. 3µl of the DNA ladder solution and 12µl of the DNA template solution were loaded in pockets of 1.8% agarose gel. A current of 90V was run for 30min on the EC Apparatus.

Table 6. Nanodrop Spectrometer result after PCR reaction with Primer LepiFAD2-1a. PCR products were analysed on deNovix Nanodrop Spectrophotometer. 1µl of the PCR product was tested for each accession.

Accession	**DNA concentration** (in ng/µl)	**260/280**	**Accession**	**DNA concentration** (in ng/µl)	**260/280**
NEH001	228.866	1.86	**NEH015**	566.949	1.83
NEH002	431.152	1.82	**NEH016**	298.972	1.87
NEH004	106.840	1.89	**NEH019**	558.613	1.84
NEH006	370.802	1.82	**CVW004**	602.120	1.81
NEH007	577.578	1.82	**CVW005**	527.564	1.83
NEH010	523.554	1.83	**CVW007**	581.400	1.81
NEH012	261.721	1.82	**TRU001**	578.668	1.83
NEH014	555.252	1.83	**TRU002**	346.033	1.87

Appendix 4 – Chromatograph Excepts

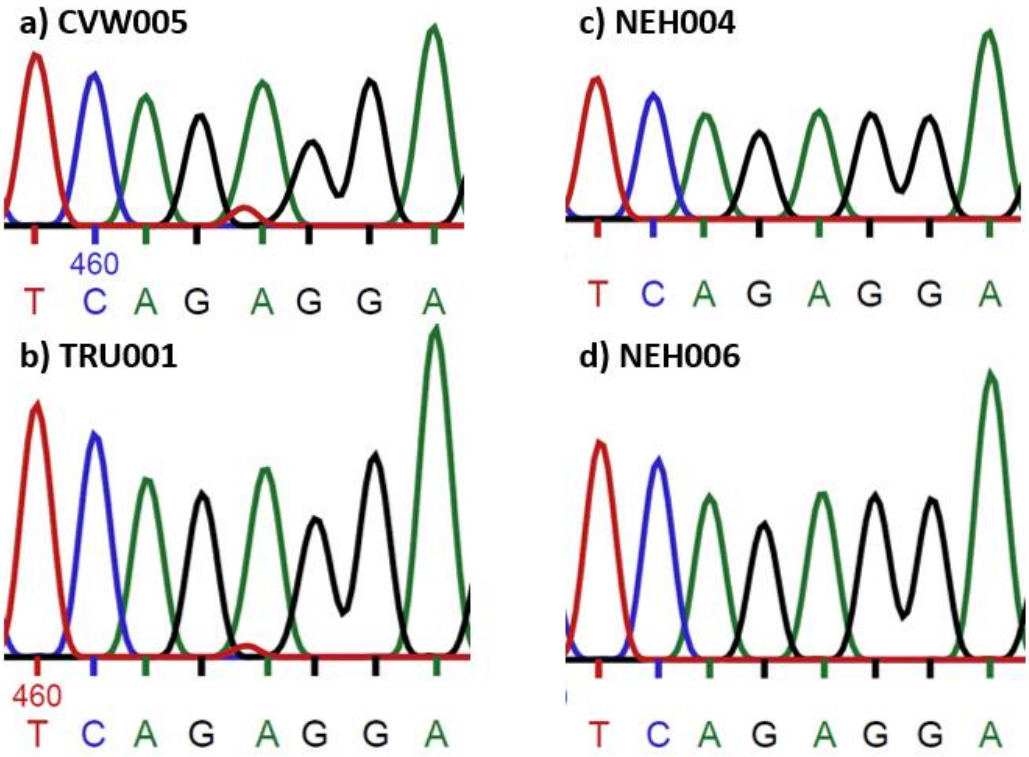


Figure 9. Chromatograph Excerpt showing disturbance found in forward sequence of four accessions. The labelling of the base pair number is slightly shifted for each sample due to trimming of low-quality results at the beginning of some sequences.

Publishing and archiving

Approved students' theses at SLU can be published online. As a student you own the copyright to your work and in such cases, you need to approve the publication. In connection with your approval of publication, SLU will process your personal data (name) to make the work searchable on the internet. You can revoke your consent at any time by contacting the library.

Even if you choose not to publish the work or if you revoke your approval, the thesis will be archived digitally according to archive legislation.

You will find links to SLU's publication agreement and SLU's processing of personal data and your rights on this page:

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

x YES, I, Lisa Hartmann, have read and agree to the agreement for publication and the personal data processing that takes place in connection with this.

☐ NO, I/we do not give my/our permission to publish the full text of this work. However, the work will be uploaded for archiving and the metadata and summary will be visible and searchable.