Vicia faba determinate and indeterminate inflorescence genotypes

– comparison of genetic variation at *TFL1* locus

Vicia faba determinanta och icke-determinanta genotyper - jämförelse av genetisk variation vid TFL1 locus

Jenny Östberg

Degree project/Independent project • 30 hp
Swedish University of Agricultural Sciences, SLU
Department of Plant Breeding
Horticultural Science Programme
Alnarp 2021

Vicia faba determinate and indeterminate inflorescence genotypes – comparison of genetic variation at *TFL1* locus

Vicia faba determinanta och icke-determinanta genotyper – jämförelse av genetisk variation vid TFL1 locus

Jenny Östberg

Supervisor:	Åsa Grimberg, Swedish University of Agricultural Sciences, Department of Plant Breeding
Assistant supervisor:	Per Hofvander, Swedish University of Agricultural Sciences, Department of Plant Breeding
Examiner:	Ida Lager, Swedish University of Agricultural Sciences, Department of Plant Breeding

Credits:	30 hp
Level:	Second cycle
Course title:	Independent Project in Biology, A2E
Course code:	EX0856
Program:	Horticultural Science programme
Course coordinating dept:	Department of Plant Breeding
Place of publication:	Alnarp
Year of publication:	2021
Keywords:	**Faba bean, growth habit, *TERMINAL FLOWER1,* genetic marker, legume, mutant**

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

Abstract

Faba bean (*Vicia faba* L.) plants usually express an indeterminate inflorescence growth habit regulated by the gene *TERMINAL FLOWER1* (*TFL1*) which allows the apical meristem to continuously produce new vegetative and reproductive plant parts. In turn, this prolongs the growth and flowering period and causes an uneven ripening of the pods. The opposite growth habit is determinate inflorescence where the apical meristem evolves into a terminal inflorescence which can give a more homogenous ripening of pods, a trait desired for its agronomical benefits especially in a northern country like Sweden where the growing season is relatively short. In this study, the *TFL1* gene and part of its upstream region was amplified, sequenced and analysed from genomic DNA of eleven different faba bean varieties with previously described growth habit, indeterminate or determinate, to explore the genetic variation between the varieties. The aim was to find a genetic molecular marker which could be used in plant breeding of new cultivars better adapted for short growing seasons. Several SNPs were identified in the upstream region of *TFL1* and were shown to be shared by varieties of the same inflorescence growth habit, but no observed sequence difference was clearly separating all the determinate varieties from the indeterminate. One of the mutant varieties showed an SSR which distinguished it from all other varieties in this study. These findings propose that there might be more than one genetic variation in the *TFL1* locus which could cause the determinate inflorescence growth habit in determinate faba bean varieties.

Sammanfattning

Åkerböna (*Vicia faba* L.) har vanligtvis ett icke-determinant växtsätt på grund av genen *TERMINAL FLOWER1*(*TFL1*) som gör att det apikala meristemet kontinuerligt producerar både vegetativa och reproduktiva växtdelar. Detta växtsätt innebär en lång tillväxt- och blomningsperiod med ojämn mognad av baljorna. Det motsatta växtsättet är determinant där det apikala meristemet utvecklar en terminal och avslutande blomställning som kan ge en mer homogen mognad av baljorna. Denna egenskap i växtsätt är efterfrågad då det innebär flera agronomiska fördelar speciellt i ett nordiskt land som Sverige där odlingssäsongen är relativt kort. I denna studie har *TFL1*-genen samt dess uppströmsregion amplifierats, sekvenserats och analyserats från genomiskt DNA av elva olika genotyper av åkerböna med olika växtsätt, determinant eller icke-determinant för att studera den genetiska variationen. Målet var att hitta en genetisk molekylär markör som kunde urskilja de två växtsätten och som kan användas inom växtförädling av nya åkerbönsorter som är bättre anpassade till korta odlingssäsonger. Skillnader i sekvensen mellan de olika genotyperna kunde identifieras i uppströmsregionen av *TFL1,* men ingen SNP som tydligt skiljde de determinanta från de icke-determinanta genotyperna hittades. En av mutant-genotyperna visade en SSR vilket urskiljde denna genotyp mot de

resterande i studien. Resultaten tyder på att det eventuellt finns mer än en genetisk variation som kan orsaka det determinanta växtsättet i determinanta åkerbönsgenotyper.

Keywords: Faba bean, growth habit, *TERMINAL FLOWER1,* genetic marker, legume, mutant

Acknowledgements

Thank you Åsa Grimberg and Per Hofvander for being great supervisors. I truly appreciate all the time, the wise words of knowledge and the feedback you have given me during this project!

I also wish to thank the kind co-workers in the lipid-lab and Horticum building for always being helpful and inspiring.

Lastly, I thank The Nordic Genetic Resource Center (NordGen) for providing the seeds of the mutant genotype Ö38.

This work was part of the SLU Grogrund project ‘Future faba beans for food and feed’.

Table of contents

Abbreviations

ANFs	Antinutritional Factors
SNP	Single Nucleotide Polymorphism
SSR	Simple Sequence Repeat
SAM	Shoot apical meristem
TFL1	TERMINAL FLOWER1
PCR	Polymerase Chain Reaction
LFY	LEAFY
AP1	APETELATA1
FT	FLOWERING LOCUS T
PIM/PEAM4	PROLIFERATING INFLORESCENCE MERISTEM
UNI	UNIFOLIATA
DET	DETERMINATE
VEG1	VEGETATIVE1

1. Introduction

1.1. The role of legume crops

The Leguminosae (Fabaceae) plant family includes several important agricultural crops which provides the world with both food and feed, and additionally plays an important role for a more sustainable agricultural system (Magrini et al, 2018). Some of the commonly cultivated legume crops are soybean (*Glycine max*), pea (*Pisum sativum*), faba bean (*Vicia faba*), lentil (*Lens culinaris*) and chickpea (*Cicer arientinum*). The species in the Leguminosae plant family is generally characterized by the legume fruit which is a pod filled with seed/seeds (Allen, 2013). Agricultural legume crops can be cultivated for many different purposes such as for the mature dry seeds, the immature seeds and pods, or for the extraction of oils or starch, depending on the legume species. Many of the grain legume crops have a high nutritional value containing protein, fiber, starch, micro- and macronutrients and vitamins (Watson et al, 2017). Cultivation of Leguminosae species in agriculture can also provide several environmental benefits. One benefit is through the symbiosis with *Rhizobium* bacteria, which makes the legume plants able to fix atmospheric nitrogen and so reduce the need of synthetic nitrogen fertilizers for the legume crop itself and, in many cases, also the subsequent crop (Watson et al, 2017). Another benefit is that by adding legume plants in the agricultural landscape and the crop rotation the agrobiodiversity increases which can reduce pest and pathogen outbreaks and also positively affect the soil structure and soil fertility (Peoples et al, 2019).

In Sweden, as well as in Europe, a lot of legumes, typically soy bean, are imported both for food and feed purposes. As much as 63 % of the plant protein use in the EU during 2016/2017 was imported, and for soy protein the self-succifiency rate was only 5 % (European commission, 2018). Of the total plant protein use in the EU 93% was used for animal feed and only 7 % for food. In Sweden, approximately 2 % of the arable land has been cultivated with legumes in the past 5 years (Olsson, 2020) and the

percentage in Europe has been around 3 % (European commission, 2018). The two most cultivated legume crops in Sweden are pea and faba bean (Olsson, 2020). An increasing interest and demand for plant protein from consumers and food producers has been seen in Sweden and in the EU for products such as fresh and frozen produce, flour or other ingredient for different food products (Grimberg, 2019; European commission, 2018). By adding more legume crops in the agricultural landscape the self-sufficiency rate of both food and feed in Sweden could increase and the Swedish food value chain could get more sustainable and competitive, which is a goal set by the Swedish Government (Näringsdepartimentet, 2016).

1.2. Plant breeding targets for faba bean in Sweden

In this study, the focus has been on the legume crop faba bean. Faba bean is a crop cultivated for the seeds which are used for both food and feed. The seeds are rich in protein (27-32% of seed dry matter (Crépon et al, 2010)) and provide a good source of carbohydrate, fibers, carotenoids, vitamins and minerals. It should also be mentioned that the seeds contain antinutritional factors (ANFs) such as tannins which can reduce the digestion and uptake of nutrients, and also vicine and convicin, which can cause a medical syndrome called favism to individuals with the inherited disorder (Ray and Georges, 2010). Through plant breeding of faba bean many of today's faba bean cultivars are low in ANFs. For example, there are white-flowering cultivars that have lower levels of seeds coat tannins, as compared to the variegated (i.e. black and white) flowered cultivars. Other important breeding targets for faba bean are higher yield, earliness (shorter growth period), tolerance for abiotic and biotic stresses, and shape and size of the seeds (Maalouf et al, 2018). Faba bean plants have a relatively long growth period and in a northern country like Sweden it can be difficult for the crop to reach maturity in time for harvest. With new cultivars better adapted to shorter growing seasons it could be possible to grow faba bean in bigger parts of Sweden, including more northern regions.

1.3. Genetic markers as tools in plant breeding

Selection of individual plants or lines expressing the desired target traits for use in breeding programs can be performed by using so called genetic or molecular markers. A molecular marker is detected at subcellular level as a DNA sequence with a known approximate location on a chromosome which is associated with a specific gene or trait

(George, 2012). Molecular markers can be applied during an early stage of the development of a plant and allows plant breeders to efficiently monitor if a specific trait of interest is present or not in the genotype, and then perform selection. With molecular markers plant breeders can save both time and resources in the development of new cultivars, as well as gaining a better understanding of the breeding material and specific traits. Examples of molecular markers are Simple Sequence Repeats (SSRs) which are repetitive DNA sequences of shorter motifs, and Single Nucleotide Polymorphism (SNP) which is a single nucleotide difference at a specific site in the DNA sequence between different genotypes (George, 2012). SNPs are highly interesting markers since a single nucleotide difference in the coding DNA sequence can change the amino acid sequence for a protein and furthermore affect its function.

1.4. Determinate and indeterminate inflorescence growth habit in faba bean

As mentioned previously, one important target trait is to develop faba bean cultivars that are early in development which gives a reduced growth period. Such cultivars could extend the regions in which faba bean could be grown within, including more northern regions in Sweden. One specific trait of interest which affect the growth period of faba bean is whether the plant develops a determinate or indeterminate inflorescence. The indeterminate inflorescence growth habit is the most common behavior in faba bean and is characterized by the continuous production of new vegetative and reproductive plant parts until a maximum number of pods have been produced or the growth environment is no longer favorable (Sjödin, 1971). This growth habit prolongs the reproductive growth and therefore extends the time to ripening of the pods on the whole plant, i.e. some pods are mature while others are immature. Through induced mutations in faba bean, mutants with determinate inflorescence growth habit were identified (Sjödin, 1971). In contrast to the indeterminate plants, these determinate plants reached a terminal vegetative phase and the inflorescence transformed into a terminal floral meristem. Examples of faba bean plants expressing the two different inflorescence growth habits can be seen in Figure 1. The determinate growth habit is of interest from a plant breeder's and faba bean grower's point of view since it allows a more concentrated ripening, and therefore a shorter growth period and an earlier harvest. Faba bean cultivars with traits allowing for earlier harvest is of high interest in the Swedish climate, where the long growth period of this crop is often a challenge. Another advantage with determinate inflorescence growth habit in faba bean is that

some varieties showed reduced plant height and thickening of the stem (Sjödin, 1971). These traits reduce the risk of lodging and simplifies culture management and mechanical harvesting.

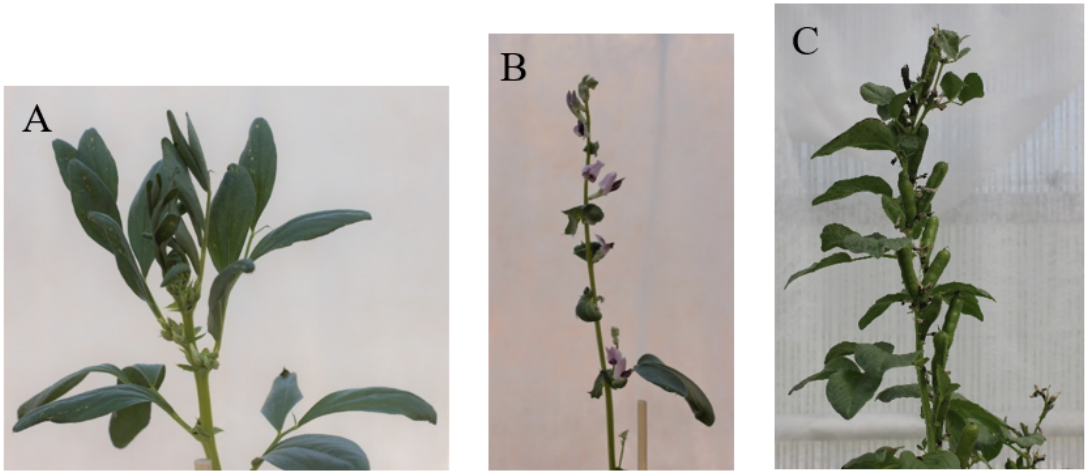


Figure 1. The apical meristem of A) an indeterminate variety, B) and C) determinate varieties. *(Photos Jenny Östberg and Åsa Grimberg)*

1.5. Genetic control of flowering and growth habit in plants

In studies of the genetic control of the inflorescence growth habit and the onset of the floral transition in plants, the most well studied species is *Arabidopsis thaliana,* which also like faba bean usually express an indeterminate inflorescence growth habit. The maintenance of the indeterminate inflorescence growth habit with the continuous development of new leaves and flowers occurs in the shot apical meristem (SAM) at the tip of the plant shoot. Three different genes are known to be mainly responsible for the regulation and development of the inflorescence architecture in *A. thaliana*. These are *TERMINAL FLOWER1 (TFL1),* a gene promoting inflorescence meristem identity, *LEAFY (LFY)* and *APETELATA1 (AP1),* which both are floral meristem identity genes, promoting floral transition in the SAM (Benlloch et al, 2015). Together they maintain the balance between inflorescence and floral meristem identity at the SAM. TFL1 is a protein acting as a repressor of flowering by repressing the gene expression of *LFY* and *AP1,* and so maintaining the indeterminate growth of the inflorescence. In the same but opposite way, the gene expression of *TFL1* is repressed by the presence of LFY and

AP1, which allows the formation of floral meristem (Figure 2). The balance of inflorescence and floral meristem development is maintained through regulated expression levels of the three genes during different phases of plant development. Mutation in the *TFL1* gene has shown to increase the expression of *LFY* and *AP1* and resulted in determinate growth of the inflorescence, meaning development of a terminal floral meristem instead of an inflorescence meristem (Benlloch et al, 2015). Another important key regulatory protein which acts as a florigen, a signal of floral stimulus, is the FLOWERING LOCUS T (FT). FT is a leaf-derived photoperiodically regulated floral stimulus, which is transported through the phloem companion cells from the leaf to the shoot apical meristem (SAM) where it promotes the onset of floral transition by activating the floral identity genes *AP1* and *LFY* (Taiz et al, 2018). Altogether, TFL1 and FT can both be regarded as key regulators of the inflorescence development and flowering time. The two proteins have opposite functions as a repressor and an activator, respectively. Interestingly, the primary sequences of these two proteins are highly similar and a single amino acid substitution can convert the function of FT into TFL1, and vice versa (Zhu et al, 2020).

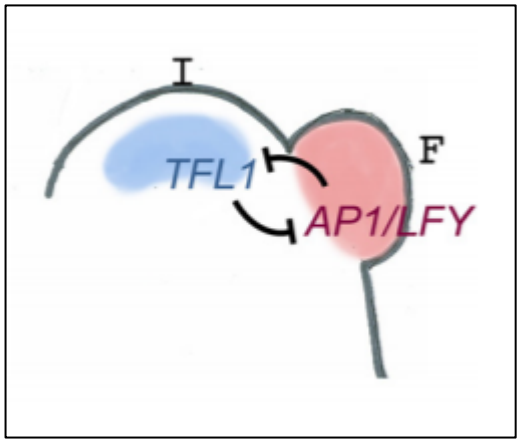


Figure 2. Meristem identity genes in Arabidopsis thaliana (Benlloch et al, 2015). *Illustration shows the mutual repression of the meristem identity genes TFL1 and AP1/LFY, which controls the maintenance of the indeterminate inflorescence growth habit in A. thaliana.*

Among legume crops the genetic control of flowering and growth habit has been most studied in pea, *Pisum sativum*. Genes in pea that are homologous (which means genes that share a common ancestor) to the genes controlling meristem development and identity in *A. thaliana* have been described and studied for pea. The regulation of the meristem identity in pea seems to be in many ways similar to the regulation in *A. thaliana* but there is one important difference between the two species, which makes the regulation in pea more complex. *A. thaliana* develops a simple inflorescence

architecture, where flowers are directly developed from the main inflorescence (I1 in Figure 4), while pea (and legumes in general) produce a more complex compound inflorescence, where the main inflorescence produces secondary inflorescences (I2 in Figure 4) which in turn develop flowers (Figure 3)(Benlloch et al, 2015). The known homologous genes of *A.thaliana LFY* and *AP1* are known as *PROLIFERATING INFLORESCENCE MERISTEM* (*PIM/PEAM4*) and *UNIFOLIATA* (*UNI*), respectively, in pea. The inflorescence meristem identity gene *TFL1* has a homologous gene in pea known as *DETERMINATE* (*DET*) or *PsTFL1a* which regulates the primary inflorescence (I1) meristem identity and the indeterminate growth habit (Benlloch et al, 2015). As mentioned earlier, legumes have a more complex compound inflorescence architecture and therefore an additional gene which controls the secondary inflorescence (I2) meristem identity is also known to be involved called *VEGETATIVE1* (*VEG1*). *DET*/*PsTFL1a* and *VEG1* repress each other's expression and regulates the balance of indeterminate growth and the formation of secondary inflorescence, while the floral meristem identity in the secondary inflorescence (I2) is controlled by *PIM/PEAM4* which represses *VEG1* (Figure 4) (Benlloch et al, 2015).

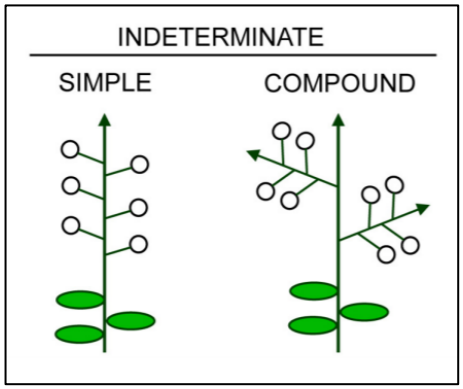


Figure 3. Different types of inflorescence architecture in plants (Benlloch et al, 2015). *The illustration shows the different types of inflorescence architecture in A.thaliana (left) and legumes (right). Open circles represent flowers and arrows represent indeterminate shoots.*

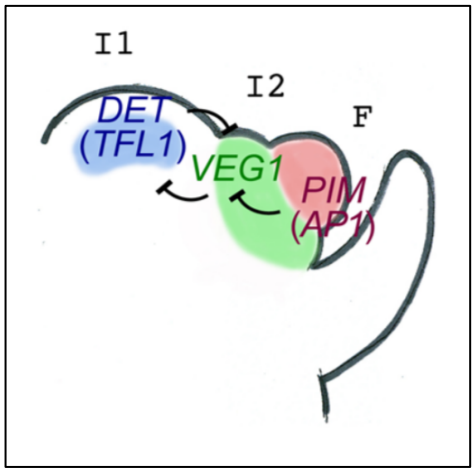


Figure 4. Meristem identity genes in pea (Pisum sativum). *Model for specification of meristem identity in the compound pea inflorescence. In the pea inflorescence apex, DET expression in the primary inflorescence meristem (I2), VEG1 in the secondary inflorescence meristem (I2) and PIM in the floral meristem (F) are required for these meristems to acquire their identity. Expression of these genes in their correct domains is maintained by a network of mutual repressive interactions (Benlloch et al, 2015).*

The genetic knowledge about the regulatory mechanisms of inflorescence growth habit available for *A. thaliana* and *P. sativum* can be used for better understanding of the same regulation in faba bean, *V. faba*.

In previous studies, a molecular marker for the determinate growth habit in faba bean was developed (Avila et al, 2006; Avila et al, 2007). The marker was located on the *TFL1* sequence and was represented by a single nucleotide substitution, a SNP. However, the preciseness of this molecular marker was shown to be limited, i.e. it could only differentiate between a limited number of varieties with determinate and indeterminate growth habit. Therefore, it is of interest to develop new more precise markers for determinate growth habit in faba bean that can be used in breeding programs in the future. *TFL1* is one of the highly interesting target genes for such marker development.

1.6. Aim of thesis

In this study, eleven varieties of faba bean with different inflorescence growth habit, determinate or indeterminate, were selected for comparison of the *TFL1* locus, including part of the upstream promotor region of the gene, to explore potential

polymorphism in the DNA sequence. The upstream promotor region of the gene is of high interest since it is containing sites to which transcription factors and other regulatory elements can bind to which affect the transcription and expression of the gene. With a better understanding of the regulation of the inflorescence growth habit in faba bean, and with more precise molecular markers, new cultivars with a growth habit better adapted for shorter growing seasons could be developed.

2. Material and methods

Selection of faba bean varieties

In this study, eleven varieties with known growth habit, either determinate or indeterminate were selected. The selected varieties were the following: Mikko, Pirhonen, Fanfare, Aurora, Taifun, Tina, Tinova, Bruno, TicolH, TicolL, and Ö38. Of these varieties the first five ones were indeterminate, and the other six were determinate.

Seed surface sterilization and plant growth conditions

Seeds from the selected varieties were surface sterilized using 1% hypochlorite, 0,25 ‰ Silwet and sterile water, and treated under agitation for 5 minutes. The seeds were then treated with a solution of 70 % ethanol, 0,25 ‰ Silwet and sterile water, under agitation for another 5 minutes, and then washed with sterile water in a sterile bench. Seeds were sown in germination media consisting of ½ strength MS (2.2 g/L) (Murashige-Skoogh medium, Dushefa Biochemie), sucrose (5g/L) (Duchefa Biochemie) and bacto agar (7g/L) (Saveen Werner). Three seeds of each variety were sown, except for the variety Ö38 where only two seeds were available (Figure 2). The seeds were placed in a growth chamber (22°C ± 2°C, 16 h photoperiod (at 60–70 µmol m–2 s–1) and leaves were harvested as they appeared and immediately snap-freezed in liquid nitrogen. All samples were milled to a fine powder using metal beakers and beads chilled with liquid nitrogen (Retsch® MM400 Mixer Mill). The milling settings were 30 Hz for 40 seconds.

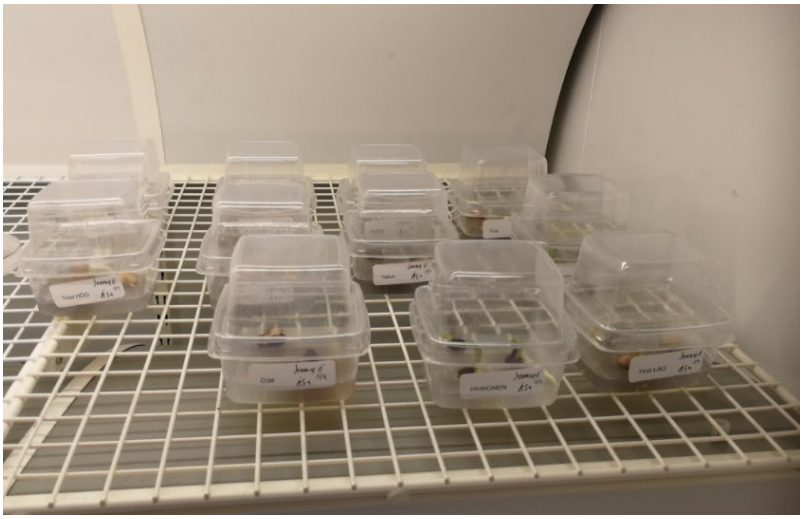


Figure 2. Boxes with seeds sown in germination media and placed in growth chamber*. (Photo Jenny Östberg)*

Genomic DNA extraction

Genomic DNA from the milled plant material of each variety was extracted using Nucleospin® plant II (MACHERY-NAGEL). In brief, genomic DNA was extracted using lysis buffer (PL1) and the crude lysates were cleared by the Nucleospin® plant filters and centrifugation which removed polysaccharides, contaminations, and residual cellular debris, leaving a clear flow through. The clear flow through, containing the genomic DNA, was mixed with the binding buffer so that the genomic DNA could bind to the silica membrane in a second plant filter. The filter with the genomic DNA was then washed using washing buffers to remove contaminants. Finally, the genomic DNA was eluted once using low salt elution buffer into 50 µL and the concentration was measured with Nanodrop (Saveen Werner) or DNA chip (Xpose).

Amplification of *TFL1* gene in faba bean

The whole *TFL1* gene was amplified from faba bean varieties Ö38, TicolL and Tinova, using Phusion U Multiplex PCR Master Mix (Thermo Scientific), with the primer set *TFL1* C (ATGGCAAGAATGGCTCAAGAAC) and B (GCGTCTTCTTGCAGCGGTT), available from Avila (2006) & Avila (2007). Three different PCR (Polymerase Chain reaction) settings were applied with an increasing annealing temperature to test which gave most of the desired product (Table 1). The Tm calculator from Thermo Fisher was used to calculate the optimal annealing temperature for the primers, and the PCR machine used was a Unocycler (VWR).

***Table 1. PCR settings for amplification of the TFL1 gene in faba bean**. Three different annealing temperatures were tested to see which gave most of the desired PCR product.*

Step	Temperature, °C	Time
Heating lid	105	
Initial	98	30 sec
Denaturation	98	10 sec
Annealing	**58.1 / 61.6 / 64.4**	30 sec
Extension	72	30 sec
Final	72	3 min
Storage	12	infinite

Amplification and cloning of upstream region of the *TFL1* gene in faba bean

The upstream region of *TFL1* was amplified using PCR with two different combinations of primers (1 and 2, see Figure 3) and the Phusion U Multiplex PCR Master Mix (Thermo Scientific).

Combination 1:
Forward primer Vfab_F1 (GTGTACGTGTAGGGGARAG) and
Reverse primer: Vfab _R1 (ATCTCATGTCAGCACCATC)

Combination 2:
Forward primer: Vfab_F1 (GTGTACGTGTAGGGGARAG) and
Reverse primer: Vfab_R2 (AGGATCTCATGTCAGCACCATC)

Figure 3. Primer sequences and combinations for amplification of upstream region of TFL1.

The two primer combinations were expected to yield a PCR product covering approximately 300 base pairs upstream of *TFL1* gene and around 180 base pairs of the *TFL1* gene. The expected size of the obtained fragment was approximately 500 base pairs. The PCR machine Unocycler (VWR) was used for all PCR reactions and the PCR cycle settings were as follows in Table 2. Two different annealing temperatures (56°C and 58°C) were tested to see which gave most specific fragment of the expected size (i.e. 500 base pairs). The Tm calculator from Thermo Fisher was used to calculate the optimal annealing temperature for the primers.

Table 2. PCR settings for the PCR reactions for the amplification of the upstream region of TFL1. *Two different annealing temperatures were tested to see which gave most specific fragment of expected size (i.e. 500 bp).*

Step	Temperature, °C	Time
Heating lid	105	
Initial	98	30 sec
Denaturation	98	10 sec
Annealing	**56/58**	30 sec
Extension	72	30 sec
Final	72	3 min
Storage	12	infinite

The PCR products from both amplifications were analysed on agarose gel using electrophoresis to see if the fragments of the expected sizes (around 1000 bp for whole *TFL1* gene, and around 500 bp for upstream region of *TFL1*) were present, and to examine if the primer combinations were successful. Selected PCR products were ligated into plasmids using the CloneJET PCR Cloning kit (Thermo Scientific) and following the protocol for blunt-end cloning. The ligations were then transformed with chemically competent cells (One Shot ®, Thermo Fisher) heated in a water bath, allowing the pores of the cell membrane to open and the plasmid to enter. The cells were then spread onto LB agar plates and placed at 37 °C overnight. Colonies from the plates were harvested individually and inoculated into liquid LB media with antibiotic (ampicillin) and grown at 37 °C with shaking at 225 rpm, overnight. The plasmid DNA was purified using the GeneJET Plasmid Miniprep Kit (Thermo Scientific) and DNA concentration was measured using DNA chip (Xpose). To ensure the samples contained the target fragment digestion of the DNA was performed using the Fast Digestion of DNA protocol (Thermo Scientific) with the restriction enzyme BglII. 10X FastDigest®Green buffer was used. Samples were then run on gel with a 100 base pair ladder and a 1 kb ladder. Samples showing target fragments on gel were selected and sent for sequencing. The samples were sent to Eurofins Genomics or LGC Biosearch Technologies, both using Sanger sequencing. Each variety was represented with four individual replicates to ensure coverage of possible allelic differences of *TFL1* between the genomes, since faba bean is a diploid.

Bioinformatic analysis of sequences

Data from the sequencing results were uploaded and analysed using the software CLC Main workbench 20.0.4. All sequences were trimmed to remove parts belonging to the plasmid. The sequencing data was aligned both per variety with the four different replicates and aligned with one representative sequence per variety. The presence of Single Nucleotide Polymorphism (SNP), Simple Sequence Repeats (SSRs) and other sequence differences were investigated.

Binding site analysis of upstream region of the *TFL1* gene

To further explore the upstream region sequence of *TFL1,* a binding site analysis was performed, using the database PLACE (Higo et al, 1999).

3. Results

Eleven different varieties with known growth habit, determinate or indeterminate inflorescence, were grown in sterile conditions and leaf tissue was harvested using liquid nitrogen for snap freeze. One variety, Taifun, did not germinate and was excluded. DNA was extracted from the remaining ten varieties and PCR reactions for amplification of both the *TFL1* gene and the upstream region of *TFL1* was performed (Table 3). The PCR products were ligated into plasmids and transformed into chemically competent bacterial cells for cloning. Plasmid DNA that was purified from bacterial cells were sent for sequencing and analyzed to identify sequence differences between the varieties.

***Table 3**. The eleven different varieties, their growth habit and the amplifications that were performed for each variety.*

Variety	Growth habit	Amplification of upstream region of *TFL1*	Amplification of whole *TFL1* gene
Tina	Determinate	✓	
Tinova	Determinate	✓	✓
Bruno	Determinate	✓	
TicolH	Determinate	✓	
TicolL	Determinate	✓	✓
Ö38	Determinate	✓	✓
Mikko	Indeterminate	✓	
Pirhonen	Indeterminate	✓	
Fanfare	Indeterminate	✓	
Aurora	Indeterminate	✓	
Taifun	Indeterminate		

3.1. Amplification of the *TFL1* gene and the upstream region of *TFL1*

Amplification of the *TFL1* gene in faba bean

A PCR reaction was performed to amplify the whole *TFL1* gene using genomic DNA from three chosen varieties, but only two of them, Ö38 and TicolL showed successful amplification of the gene with clear fragment around the expected size 1000 bp (Figure 5). The PCR reaction based on DNA from Tinova showed several weak bands of different sizes and amplification of *TFL1* in this variety was therefore not regarded successful. Three different PCR settings (i.e. annealing temperature was 58.1°C, 61.6°C, or 64.4°C) were performed to test which gave results with most specific fragment of the expected size. The selected PCR products from Ö38 and TicolL belonged to samples amplified using the annealing temperature 61.6 °C.

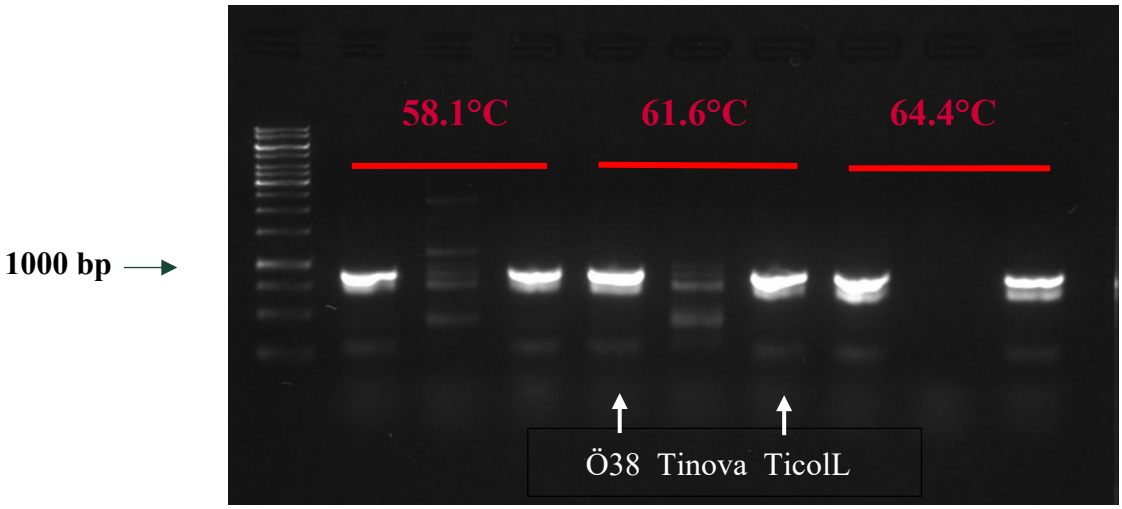


Figure 5. Gel showing PCR products from amplification of TFL1 gene in faba bean derived from three different annealing temperatures and three different varieties*. Arrows point out which products were selected for cloning.*

Amplification of the upstream region of *TFL1*

For each variety, two different combinations of primers were applied and run in the PCR, primers Vfab_F1 with Vfab_R2, or Vfab_F1 with Vfab_R2. Primers Vfab_R1 and Vfab_R2 were located at the beginning of the known sequence of the *TFL1* gene in V. faba (Figure 6). The forward primer Vfab_F1 was designed by using previous knowledge of genomic partial sequences upstream of the *TFL1* gene in *Pisum sativum*

and *Medicago truncatula,* two model species within legume genomics. A region of the upstream sequence with high homology between *P. sativum* and *M. truncatula* was selected as a basis for the design of Vfab_F1. Primer Vfab_F1 included an R in the sequence (see Figure 3) which means it could be either an A or a G nucleotide. This was included because that was the only position where there was a deviation between the sequences of the two species.

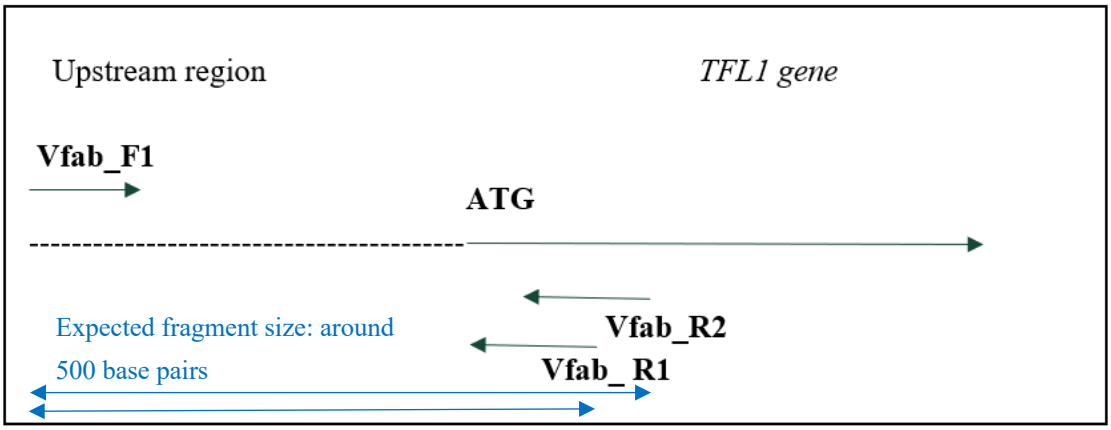


Figure 6. The Expected size of the obtained PCR fragment*. The primers used covered both upstream region and beginning of TFL1 gene. Expected size of fragment was around 500 base pairs. ATG marks start of the TFL1 gene.*

Two different annealing temperatures in the PCR reaction were tested. In the first run, the annealing temperature was 56 °C, and in all the following runs the annealing temperature was 58°C. A clearly defined fragment of around the target size 500 base pairs was visible in both gels with both primer combinations (Figure 7). However, less fragments of other sizes were visible in the gel with the PCR annealing temperature at 58°C. PCR products chosen for cloning were selected based on which samples showed the most specific desired product on the gel (Table 4).

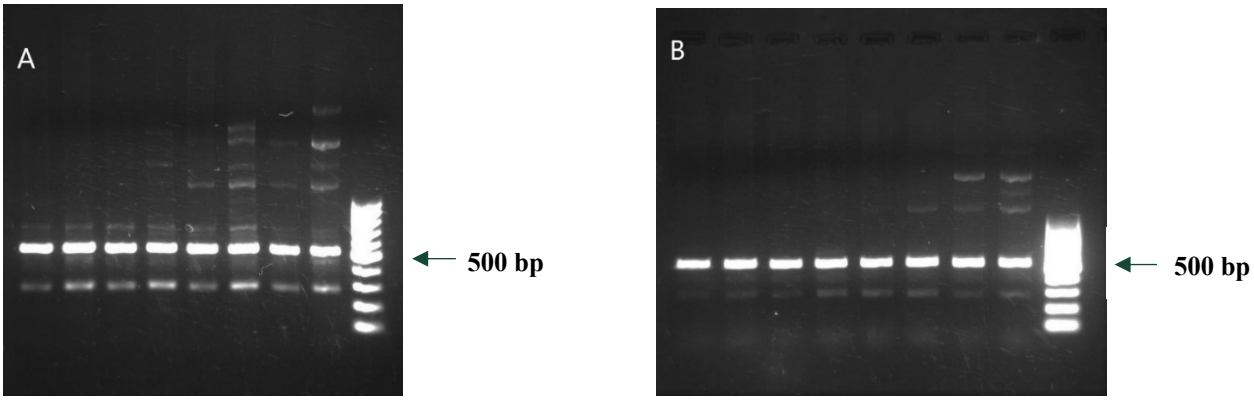


Figure 7. Examples of gels showing PCR products from amplification of the TFL1 upstream region at different annealing temperatures. ***A)*** *Annealing temperature was 56°C.* ***B)*** *Annealing temperature was 58°C.*

Table 4. Settings for the PCR products from amplification of the upstream sequence of TFL1 in faba bean chosen for cloning.

Variety	PCR annealing temperature	Primer combination
TicolL	58°C	Vfab_F1+ Vfab_R1
TicolH	56°C	Vfab_F1+ Vfab_R1
Bruno	58°C	Vfab_F1+ Vfab_R1
Aurora	56°C	Vfab_F1+ Vfab_R1
Tina	58°C	Vfab_F1+ Vfab_R1
Tinova	58°C	Vfab_F1+ Vfab_R1
Mikko	58°C	Vfab_F1+ Vfab_R2
Pirhonen	58°C	Vfab_F1+ Vfab_R1
Ö38	58°C	Vfab_F1+ Vfab_R1
Fanfare	58°C	Vfab_F1+ Vfab_R1

Selection of samples sent for sequencing

PCR products from the two amplifications were successfully ligated into plasmids and then transformed into chemically competent cells. The purified plasmid DNA samples showed concentration values between 76-152 ng/µL. Using the restriction enzyme BglII for digestion of the plasmid DNA, clear fragments of the desired sizes (1000 bp for the TFL1 gene, and 500 bp for the upstream region of *TFL1*) were visible in several samples after separation on agarose gel (Figure 8 and figure 9). Those samples that showed clear fragments of desired sizes were selected for sequencing. Four samples

from each variety that showed the desired fragment sizes were sequenced to ensure coverage of possible allelic differences.

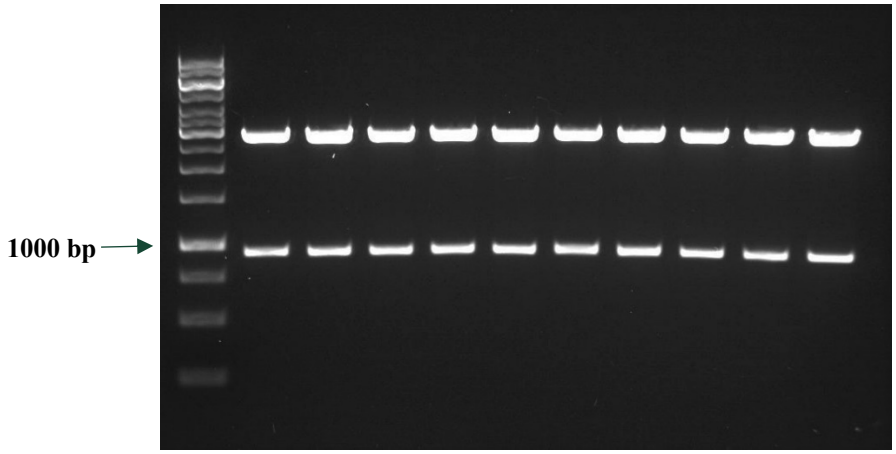


Figure 8. Example of gel picture showing digested plasmid-DNA products for the whole TFL1 gene using restriction enzyme BglII. *Gel shows clear fragment around the target size of 1000 base pairs for all samples.*

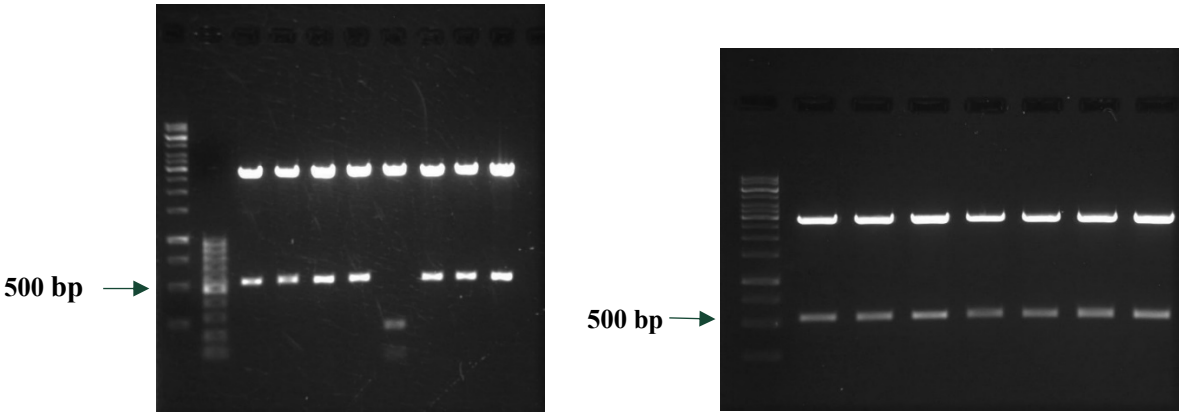


Figure 9. Examples of gel pictures showing digested plasmid-DNA products for the upstream region of TFL1 using restriction enzyme BglII. *The gels show clear fragment around the target size of 500 base pairs for all but one sample after digestion by the restriction enzyme BglII.*

3.2. Variety differences in sequence of faba bean *TFL1* gene and the upstream region

DNA-fragments containing the whole *TFL1* gene, and fragments containing upstream region of *TFL1* were sequenced using Sanger sequencing and then trimmed and aligned to identify sequence differences between the varieties.

Genetic variation in the *TFL1* gene

Three faba bean varieties were selected for amplification of the *TFL1* gene, two of them showed successful products for sequencing (i.e. Ö38 and TicolL). These sequences were aligned with previous results of *TFL1* available from Avila et al (2007) to evaluate the marker developed in their study which should distinguish determinate varieties from indeterminate. Both Ö38 and TicolL have previously been described as having determinate inflorescence growth habit. However, in this study, only TicolL showed the SNP described for determinate inflorescence growth habit while Ö38 unexpectedly showed the SNP described for indeterminate inflorescence growth habit (Figure 10). The complete alignment of the *TFL1* sequences can be found in appendix, Figure S1.

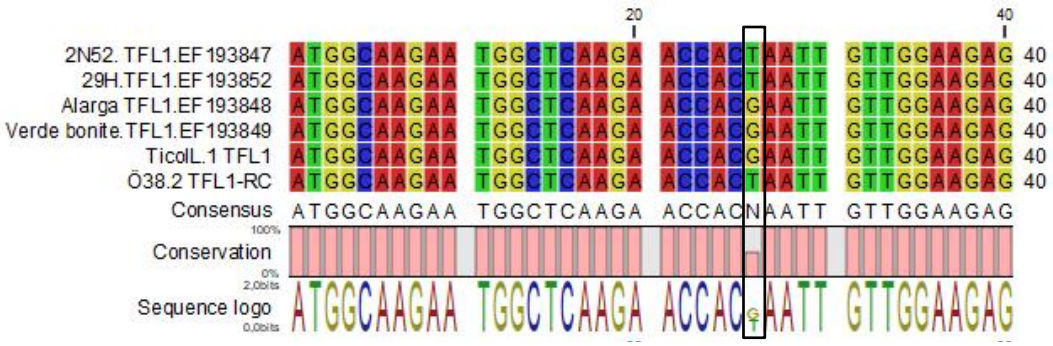
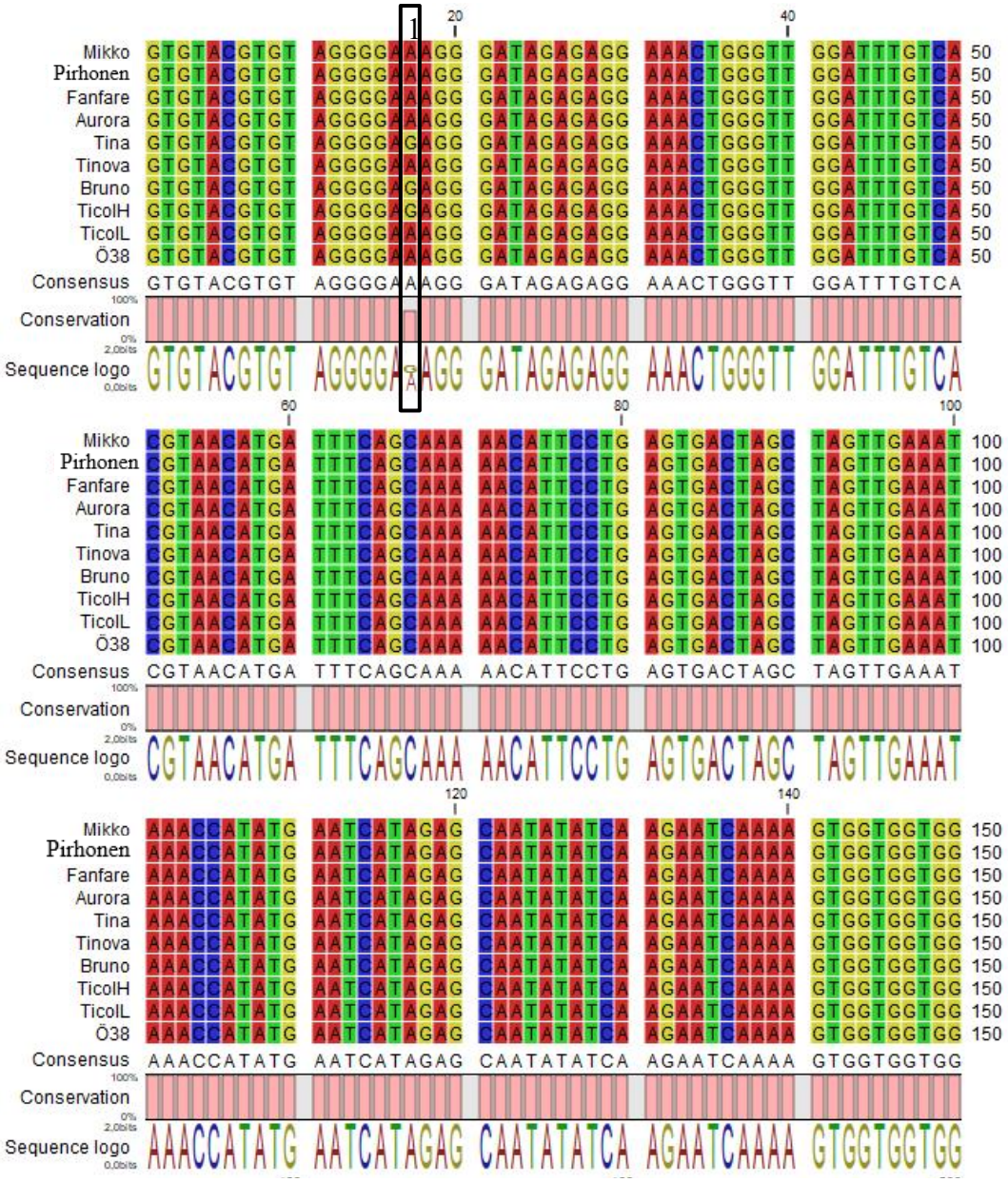


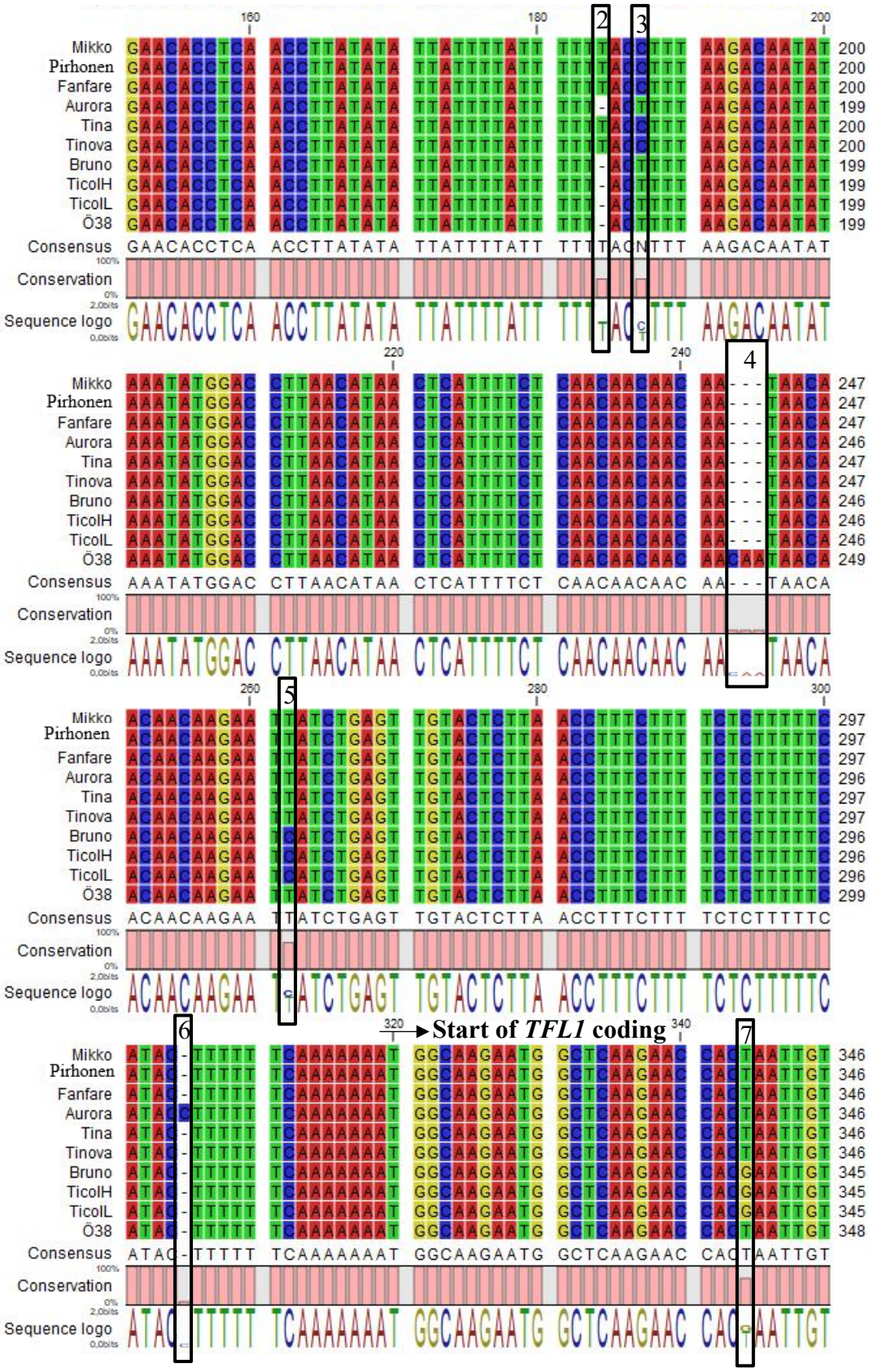
Figure 10. Alignment of the first 40 bp of the TFL1 sequence*. The box indicates the location of the SNP marker identified by Avila et al (2007). The varieties 2N52 and 29H are previously described as having indeterminate inflorescence growth habit. Alarga, Verde bonite, TicolL and Ö38 are previously described as having determinate inflorescence growth habit. Ö38 shows the SNP for indeterminate inflorescence growth habit, while TicolL shows the SNP for the determinate inflorescence growth habit.*

Genetic variation in the upstream region of *TFL1*

The upstream region of *TFL1* was amplified and sequenced for ten faba bean varieties. These sequences were aligned to explore SNPs and SSRs, which are highlighted in boxes in Figure 11. In total, seven different locations of variety differences were found in the sequences. In Table 5 these differences, their location on the obtained sequence, and which varieties that share the same difference is presented. I should be noted that primer Vfab_F1 included an R in the sequence where there was a variation between *P. sativum* and *M. truncatula,* which means it contained two different primer sequences with either an A or a G nucleotide. The sequencing results show that both primer variations could be used to obtain successful PCR products for almost all products, only a few showed a clear preference (Table 5). Due to this, variety differences at this location (box 1, Figure 11) are therefore not discussed further as 'true' SNP.

In summary, no single identified SNP or SSR clearly separated the determinate varieties from the indeterminate varieties, but several SNPs were shared by varieties of the same inflorescence growth habit, but not all SNPs (Table 5). For example, Bruno, TicolL and TicolH showed the same differences throughout the whole sequence, and so did Mikko, Pirhonen, Fanfare, Tina and Tinova. Ö38 showed an SSR of a three nucleotide addition in the sequence (box 4, Figure 11) which differed from all other varieties. In appendix the four replicates of each variety can be found (Figure S2-S11). The results showed that in most varieties the sequences from the four replicates were identical.


Mikko
Pirhonen
Fanfare
Aurora
Tina
Tinova
Bruno
TicolH
TicolL
Ö38
Consensus
Conservation
Sequence logo
GTGTACGTGT AGGGGAAAGG GATAGAGAGG AAACTGGGTT GGATTTGTCA
CGTAACATGA TTTCAGCAAA AACATTCCTG AGTGACTAGC TAGTTGAAAT
AAACCATATG AATCATAGAG CAATATATCA AGAATCAAAA GTGGTGGTGG


Mikko
Pirhonen
Fanfare
Aurora
Tina
Tinova
Bruno
TicolH
TicolL
Ö38
Consensus
Conservation
Sequence logo
2
3
4
5
6
7
Start of TFL1 coding

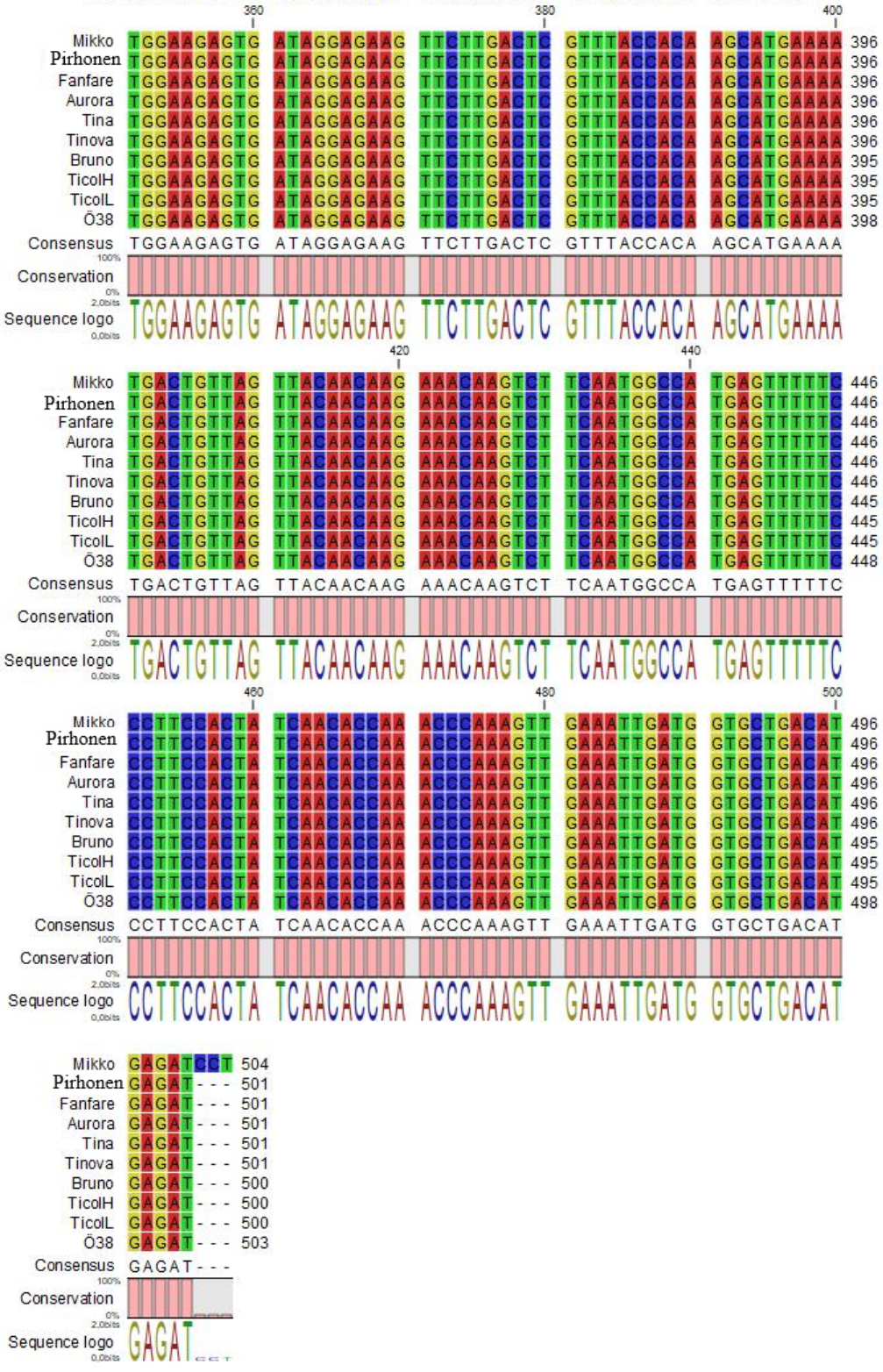


Figure 11. Alignment of the upstream region of faba bean TFL1. *Boxes marks out variety differences in sequence and is explained in Table 5.*

Table 5. Identified differences in the sequence of TFL1 upstream region and part of the TFL1 gene. *The differences are numbered according to the boxes in Figure 11 and the location on sequence is referring to the numbers of bases from start in the sequence obtained in this study. $^{2}/_{4}$ A means that two out of four sequencing replicated showed an A in the sequence, while two replicated showed a G (see text for further explanation).*

Box nr and difference	Location on sequence	Variety
1. Primer variation	17	Mikko $^{4}/_{4}$ **A** Pirhonen $^{2}/_{4}$ **A** Fanfare $^{4}/_{4}$ **A** Aurora ¾ **A** Tina ¾ **G** Tinova $^{4}/_{4}$ **A** Bruno $^{2}/_{4}$ **A** TicolH $^{2}/_{4}$ **A** TicolL ¾ **A** Ö38 $^{4}/_{4}$ **A**
2. Deletion (no T)	184	Bruno TicolH TicolL Ö38 Aurora
3. T instead of C	187	Bruno TicolH TicolL Ö38 Aurora
4. Addition of three nucleotides (SSR)	243	Ö38
5. C instead of T	262	Bruno TicolH TicolL
6. Addition of C	305	Aurora
7. G instead of T	344 (on *TFL1* coding sequence)	Bruno TicolH TicolL

Binding site motifs in the upstream region of *TFL1* gene

To further explore the differences in the upstream region sequence of *TFL1,* a prediction of binding site motifs was performed, using the database PLACE (Higo et al, 1999). Genomic sequence from *Pisum sativum* and *Medicago truncatula* was included in the analysis to explore possible similarities and differences between the different species. The sequences from *P.sativum* and *M.truncatula* were derived from NCBI and starts where the primer Vfab_F1 begins and ends at translational start of *TFL1* gene. The result showed that *P.sativum, M.truncatula* and all *V.faba* varieties included in this study, possibly shared several binding site motifs. Some were predicted to be present in *V.faba* and *P.sativum* only, and others were instead predicted to be present in *V.faba* and *M.truncatula.* Predicted differences were also found between *V.faba* varieties, for example an IBOXCORE motif which was shown to be present in all studied *V.faba* sequences (and in the sequences for *P.sativum,* and *M.truncatula*) except for in the sequence from the variety Bruno. Another predicted difference was shown in the sequence of the variety Ö38. As mentioned above, Ö38 had an additional three nucleotides in position 243 in the upstream region of *TFL1* (Figure 11). This addition resulted in an extra RAV1AAT motif. All other varieties contained five repeated RAV1AAT motifs, while Ö38 contained six (see Table S1). Another difference among the sequences shown in the prediction was the presence of TAAAGSTKST1 and SEF4MOTIFGM7S motifs. These motifs were shown to be present in only a few varieties. In Table 6, the predicted binding site motifs found in the upstream region of *TFL1* is presented. Motifs specific for *P.sativum* and *M.truncatula* are not included in Table 6, but can be found in a more comprehensive table in appendix (Table S1), which also includes repeats of the motifs. A sequence alignment of the upstream region of *TFL1* in *P.sativum, M.truncatula,* and *V.faba* cultivar Mikko, can be found in appendix (Figure S12) where the similarities in sequence motifs also can be seen.

Table 6. Predicted similarities and differences in binding site motifs in the upstream region of TFL1 for Vicia faba varieties, Pisum sativum and Medicago truncatula.

Present in:	Binding site Motif:
All: *Vicia faba* varieties (Mikko, Pirhonen, Fanfare, Aurora, Tina, Tinova, Bruno, TicolH, TicolL, Ö38), *Pisum sativum, Medicago truncatula*	ABRELATERD1 ACGTATERD1 ACGTOSGLUB1 ARR1AT BIHD1OS CAATBOX1 CACTFTPPCA1 CAREOSREP1 CURECORECR DOFCOREZM EBOXBNNAPA GATABOX GT1CONSENSUS GT1GMSCAM4 GTGANTG10 INRNTPSADB MYB1AT MYCCONSENSUSAT NODCON2GM OSE2ROOTNODULE POLLEN1LELAT52 RAV1AAT ROOTMOTIFTAPOX1 SEF1MOTIF TATABOX2 TATABOX5 WRKY71OS
V. faba and *P. sativum*	CATATGGMSAUR MARTBOX MYBPZM POLASIG1 POLASIG3 SEF3MOTIFGM TATABOX4
V.faba and *M. truncatula*	GT1CORE SEBFCONSSTPR10A WBOXHVISO1 WBOXNTERF3
V.faba varieties only	MYBST1 PREATPRODH RBCSCONSENSUS RHERPATEXPA7
All but Bruno, TicolL and TicolH	IBOXCORE
Aurora	PYRIMIDINEBOXOSRAMY1A
All but Tina and Tinova	SEF4MOTIFGM7S
All but Aurora, Bruno, TicolL, TicolH, Ö38 and *M.truncatula*	TAAAGSTKST1

4. Discussion

In this study, the polymorphism in the *TFL1* locus of determinate and indeterminate varieties of faba bean was analyzed. A previously developed genetic marker in the *TFL1* gene for distinguishing determinate from indeterminate growth habit in faba bean varieties (Avila et al, 2007) was tested, but showed limited effectiveness. This result show that new genetic markers with higher precision needs to be developed for the determinate growth trait, which potentially can be located outside of the coding gene sequence. Therefore, the upstream region of *TFL1* was amplified and sequenced for ten different faba bean varieties to explore potential differences that could affect the expression of the *TFL1* gene. The varieties used in this study were Bruno, TicolH, TicolL, Ö38, Tina, Tinova which have determinate inflorescence growth habit, and Mikko, Pirhonen, Fanfare and Aurora which have indeterminate inflorescence growth habit. The growth habit of all these varieties except for Ö38 have been confirmed by previous phenotype analyses under greenhouse conditions (Hanna Ohm, Alnarp, 2019). Ö38 is a mutant from Jan Sjödin's collection described with determinate growth habit, donated to the Nordic Genetic Resource Center.

In the upstream region of faba bean *TFL1*, four SNPs were observed, and an addition of three nucleotides in one of the varieties (Ö38). No observed difference in the sequences was clearly separating the determinate varieties from the indeterminate, but several SNPs were shared by varieties of the same inflorescence growth habit, proposing that there could be more than one genetic variation that could cause the determinate inflorescence growth habit in faba bean. In fact, observations of plants in greenhouse might indicate that the phenotype of determinate inflorescence growth habit could vary among determinate varieties. Examples of the phenotype of the shoot apical meristem of two determinate and one indeterminate variety are shown in Figure 1. The images might possibly illustrate an example of variation of the determinate phenotype, which would be interesting to look further into in the future.

The genetic differences between the studied varieties were all found in the upstream region at 180-350 base pair from start of the sequenced upstream fragment. This could indicate that this region in the upstream sequence is important for the regulation of *TFL1* transcription (i.e. might contain binding motifs for transcription factors). Studying the similar region upstream of *TFL1* in other closely related species of faba

bean such as *Pisum sativum* (pea) and *Medicago truncatula* (barrel clover) might lead to more interesting findings.

To further explore the polymorphism found in the upstream region of *TFL1*, a binding site analysis was performed using available software. The prediction showed that some of the identified SNPs from the comparative study of faba bean varieties resulted in changes of the predicted binding site motifs. For example, all varieties but Bruno, TicolL and TicolH had a cis-regulatory element (an element located on the upstream region of a gene which can regulate the gene's expression) called IBOXCORE which is a conserved sequence upstream of light-regulated genes of both monocotyledons and dicotyledons. This element is one of many that are involved in plant response to light (Junhua et al, 2014). Since light is an environmental factor that affect several aspects of growth and development of a plant, including the onset of flowering, it can be possible that a SNP that affects the IBOXCORE in the upstream region of *TFL1* could affect the growth habit. Another interesting finding is the addition of an extra RAV1AAT motif in Ö38, found at the three nucleotide addition site in the upstream region of *TFL1.* RAV1AAT is an AP2/EREBP-like protein which belongs to the AP2/EREBP family of transcription factors which can regulate several developmental processes in plants including flower development (Riechmann et al, 1998). The results from the binding site analysis are predictions made from a software and would have to be verified experimentally for confirmation.

The conclusion from this study was that sequence differences between determinate and indeterminate faba bean varieties could be found within the upstream region of *TFL1,* which in turn affected the structure of the predicted binding sites for regulatory motifs. None of the identified sequence differences or similarities between the studied varieties could be used as a marker for the growth habit type, determinate or indeterminate, but several interesting findings propose that there might be more than one genetic cause for the determinate inflorescence growth habit in faba bean. Further studies of the effect of the identified SNPs in the upstream region of the *TFL1* gene, as well as the addition of the three nucleotides in the determinate variety Ö38, is necessary for better understanding and development of a more precise marker, or markers, for the determinate inflorescence growth habit in faba bean. With genetic markers plant breeders can save both time and resources in the development of new high-yielding faba bean cultivars with agronomic traits better adapted for a northern country like Sweden.

References

Allen L.H. (2013). Legumes. Caballero, Benjamin (editor) *Encyclopedia of Human Nutrition.* Elsevier. URL: https://www.sciencedirect.com/science/article/pii/B9780123750839001707#s0015 [2021-01-11]

Avila, C.M., Atienza, S.G., Moreno, M.T. & Torres, A.M. (2007). Development of a new diagnostic marker for growth habit selection in faba bean (Vicia faba L.) breeding. *Theoretical and Applied Genetics*, vol. 115 (8), pp. 1075–1082. DOI: https://doi.org/10.1007/s00122-007-0633-y

Avila, C.M., Nadal, S., Moreno, M.T. & Torres, A.M. (2006). Development of a simple PCR-based marker for the determination of growth habit in Vicia faba L. using a candidate gene approach. *Molecular Breeding*, vol. 17 (3), pp. 185–190. DOI: https://doi.org/10.1007/s11032-005-4075-4

Benlloch R, Berbel A, Ali L, Gohari G, Millán T and Madueno F (2015). Genetic control of inflorescence architecture in legumes. *Frontiers in Plant Science.* 6:543. Doi:10.3389/fpls.2015.00543

Crépon, K. Marget, P. Peyronnet, C. Carrouée, B. Arese, P. Duc, G. (2010) Nutritional value of faba bean (*Vicia faba* L.) seeds for food and feed. *Field Crops Research*, Volume 115, Issue 3, pages 329–339

European commission (2018). *Report from the commission to the council and the European parliament on the development of plant proteins in the European Union.* Brussel: European Commission. URL: https://ec.europa.eu/info/sites/info/files/food-farming fisheries/plants_and_plant_products/documents/report-plant-proteins-com2018-757-final_en.pdf [2021-01-11]

George A. (2012). *Principles of Plant Genetics and Breeding*, Second Edition. John Wiley & Sons, Ltd.

Goretti D., Silvestre M., Collani S., Langenecker T., Méndez C., Madueño F., Schmid M. (2020) TERMINAL FLOWER1 Functions as a Mobile Transcriptional Cofactor in the Shoot Apical Meristem. *Plant Physiology.* DOI: 10.1104/pp.19.00867

Grimberg, Åsa (2019) The faba bean back in the field and on the plate. *Sveriges Utsädesförenings Tidsskrift* 1-2019

Higo, K., Ugawa, Y., Iwamoto, M. and Korenaga, T. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Research*, Volume 27, Issue 1, 1999, Pages 297-300.

Heather Ray & Fawzy Georges (2010). A genomic approach to nutritional, pharmacological and genetic issues of faba bean (Vicia faba): *Prospects for genetic modifications, GM Crops*, 1:2, 99-106

Junhua Li, Jinhong Yuan, Mingjun Li (2014). Characterization of Putative cis-Regulatory Elements in Genes Preferentially Expressed in *Arabidopsis* Male Meiocytes, *BioMed Research International*. https://doi.org/10.1155/2014/708364

Magrini Marie-Benoit, Anton Marc, Chardigny Jean-Michel, Duc Gerard, Duru Michel, Jeuffroy Marie-Helene, Meynard Jean-Marc, Micard Valerie, Walrand Stephane (2018). Pulses for Sustainability: Breaking Agriculture and Food Sectors Out of Lock-In. *Frontiers in Sustainable Food Systems*. Vol. 2. Page 64. DOI: 10.3389/fsufs.2018.00064. URL: https://www.frontiersin.org/article/10.3389/fsufs.2018.00064 [2021-01-11]

Maalouf F., Hu J., O'Sullivan D., Zong X., Hamwieh A., Kumar S., Baum M. (2018). Breeding and genomics status in faba bean (Vicia faba). *Plant Breeding*. 138. 10.1111/pbr.12644. URL: file:///C:/Users/Namn/Downloads/Breeding_and_genomics_status_in_faba_bean_Vicia_fa.pdf [2021-01-13]

Näringsdepartimentet (2016) En livsmedelsstrategi för Sverige – fler jobb och hållbar tillväxt i hela landet. *Regeringens proposition* 2016/17:104.

Olsson Ylva, 2020, Jordbruksmarkens användning 2020 Preliminär statistik, Jordbruksverket, Serie JO – Jordbruk, skogsbruk och fiske, JO 10 SM 2001, ISSN 1654-4102, URL: https://jordbruksverket.se/download/18.5b7c91b9172c01731758cd12/1592494021096/JO10SM2001.pdf [2021-01-13]

Riechmann JL, Meyerowitz EM (1998). The AP2/EREBP family of plant transcription factors. *Biol Chem.* 379(6):633-46. doi:10.1515/bchm.1998.379.6.633. PMID: 9687012.

Sjödin J. 1971. Induced morphological variation in Vicia faba. *Hereditas* 67: 155–180.

Stagnari, F., Maggio, A., Galieni, A. et al. Multiple benefits of legumes for agriculture sustainability: an overview. *Chem. Biol. Technol. Agric*. 4, 2 (2017). https://doi.org/10.1186/s40538-016-0085-1

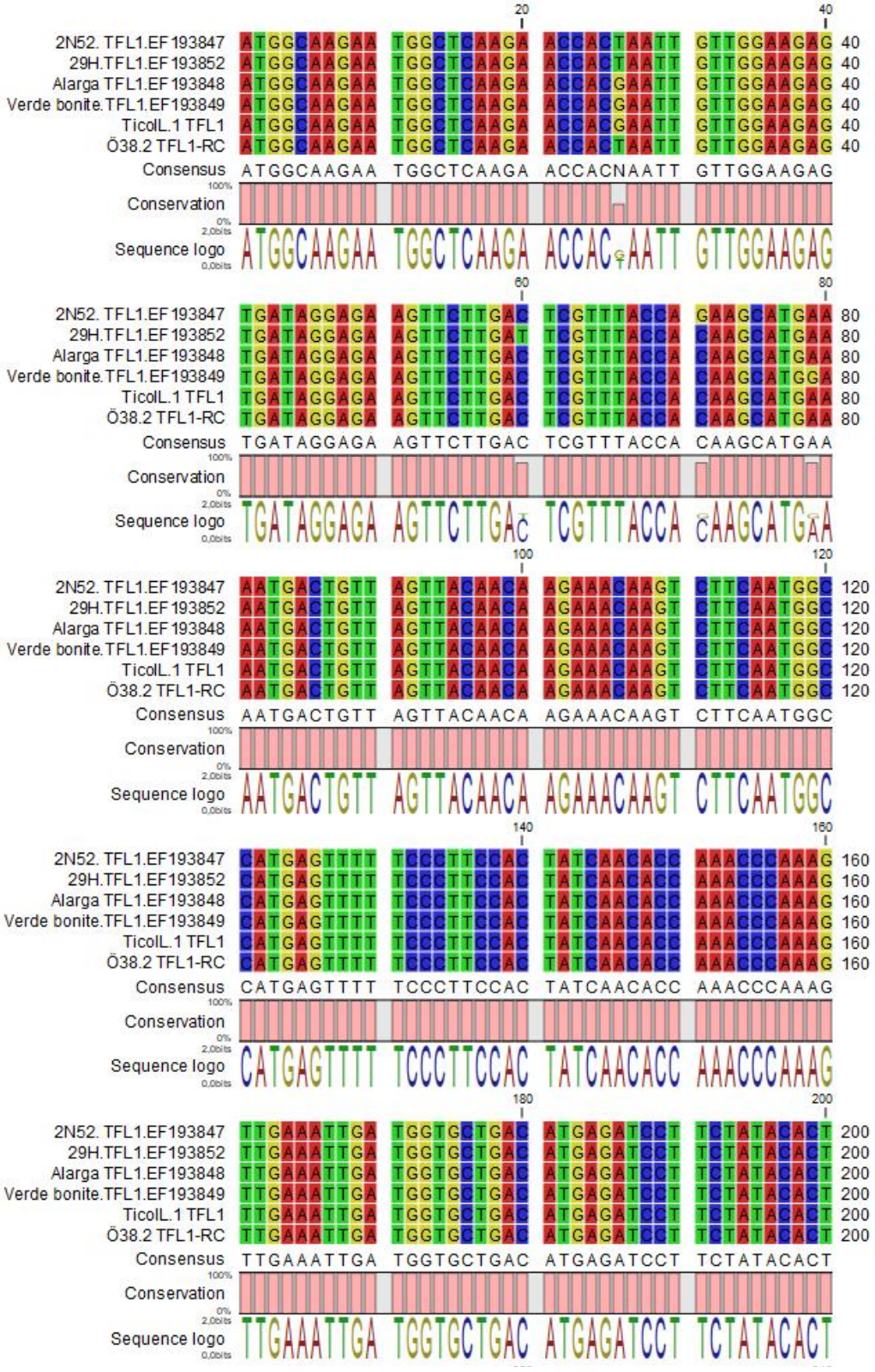
Taiz L., Zeiger E., Moller M., Murphy A. (2018). The Control of Flowering and Floral Development. *Plant Physiology and Development.* Sixth edition, New York: Oxford University Press

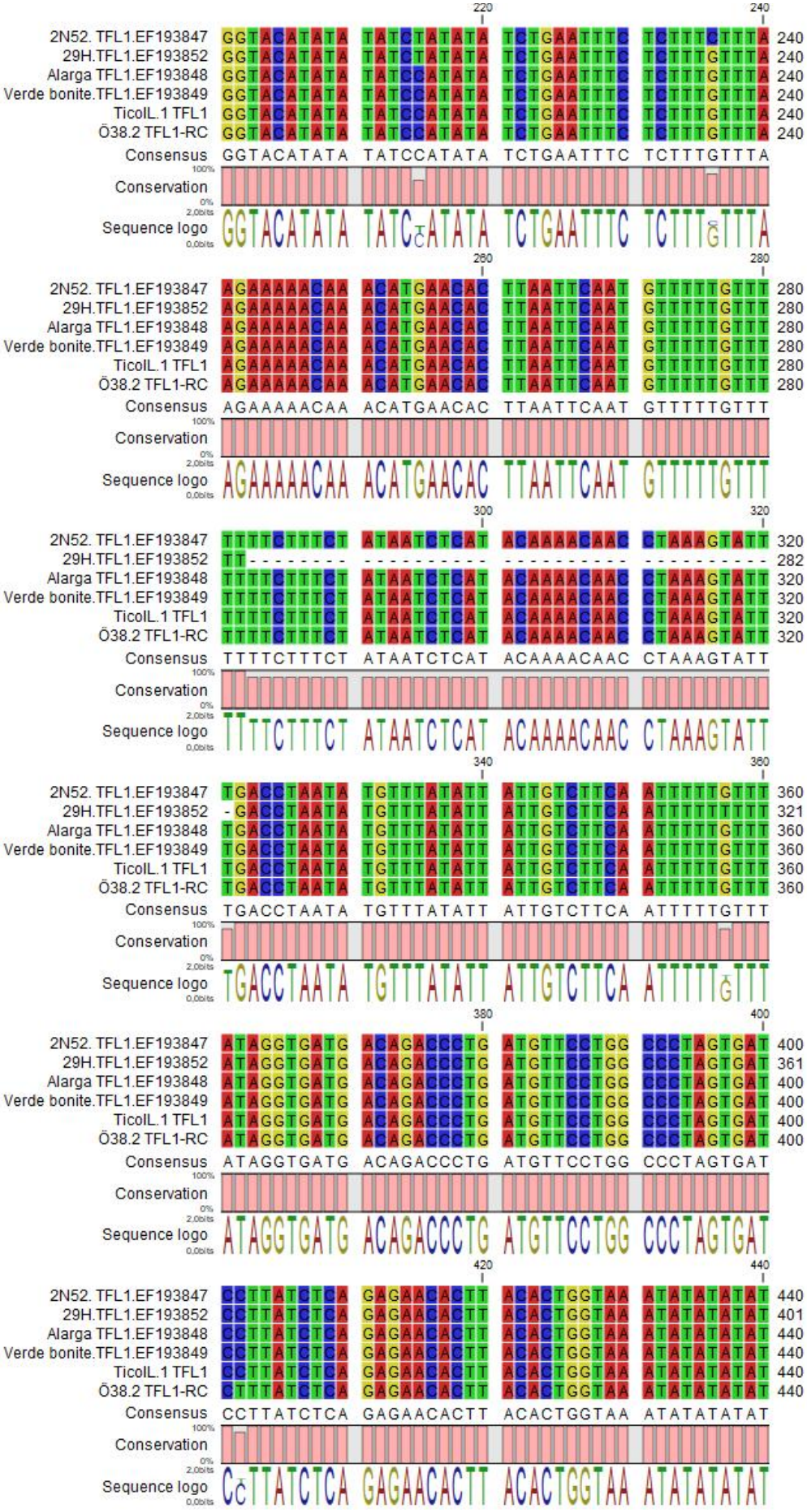
Terzaghi WB, Cashmore AR. Light-regulated transcription (1995). Annu. Rev. *Plant Mol Biol*.46:445–474.

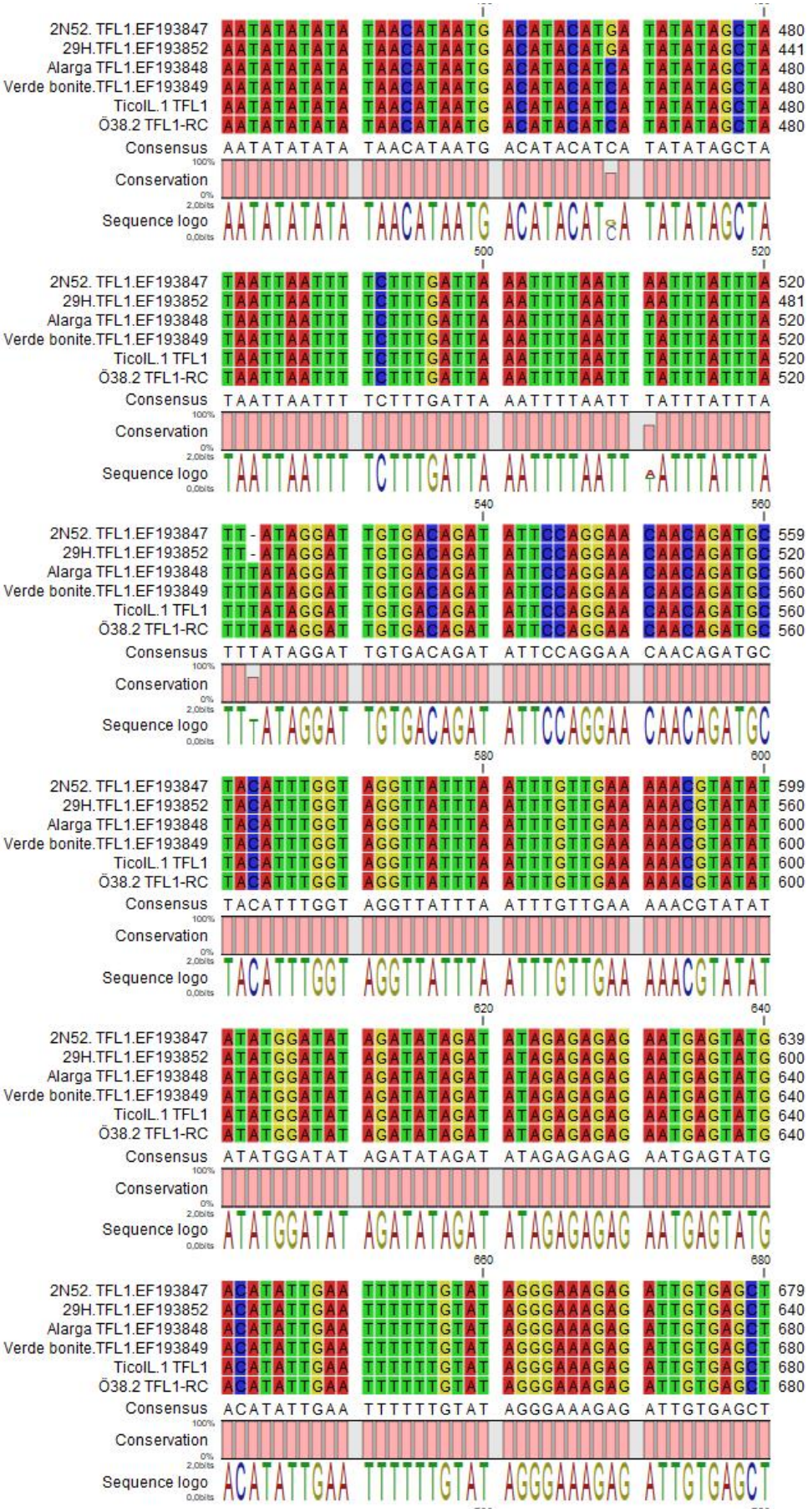
Watson, Christine & Reckling, Moritz & Preissel, Sara & Bachinger, Johann & Bergkvist, Göran & Kuhlman, Tom & Lindström, Kristina & Nemecek, Thomas & Topp, Kairsty & Vanhatalo, Aila & Zander, Peter & Murphy-Bokern, Donal & Stoddard, Frederick. (2017). Grain Legume Production and Use in European Agricultural Systems. *Advances in Agronomy*. 144. 235-303. 10.1016/bs.agron.2017.03.003.

Zhu Y, Klasfeld S, Jeong C.W, Jin R, Goto K, Yamaguchi N, Wagner D (2020). TERMINAL FLOWER 1-FD complex target genes and competition with FLOWERING LOCUS T. *Nature Communications* 11, 5118 (2020). https://doi.org/10.1038/s41467-020-18782-1

Appendix


2N52. TFL1.EF193847
29H.TFL1.EF193852
Alarga TFL1.EF193848
Verde bonite.TFL1.EF193849
TicolL.1 TFL1
Ö38.2 TFL1-RC
Consensus
Conservation
Sequence logo


220
240
2N52. TFL1.EF193847 GGTACATATA TATCTATATA TCTGAATTTC TCTTTCTTTA 240
29H.TFL1.EF193852 GGTACATATA TATCTATATA TCTGAATTTC TCTTTGTTTA 240
Alarga TFL1.EF193848 GGTACATATA TATCCATATA TCTGAATTTC TCTTTGTTTA 240
Verde bonite.TFL1.EF193849 GGTACATATA TATCCATATA TCTGAATTTC TCTTTGTTTA 240
TicoIL.1 TFL1 GGTACATATA TATCCATATA TCTGAATTTC TCTTTGTTTA 240
Õ38.2 TFL1-RC GGTACATATA TATCCATATA TCTGAATTTC TCTTTGTTTA 240
Consensus GGTACATATA TATCCATATA TCTGAATTTC TCTTTGTTTA
100%
Conservation
0%
2,0bits
Sequence logo
0,0bits
260
280
2N52. TFL1.EF193847 AGAAAAACAA ACATGAACAC TTAATTCAAT GTTTTTGTTT 280
29H.TFL1.EF193852 AGAAAAACAA ACATGAACAC TTAATTCAAT GTTTTTGTTT 280
Alarga TFL1.EF193848 AGAAAAACAA ACATGAACAC TTAATTCAAT GTTTTTGTTT 280
Verde bonite.TFL1.EF193849 AGAAAAACAA ACATGAACAC TTAATTCAAT GTTTTTGTTT 280
TicoIL.1 TFL1 AGAAAAACAA ACATGAACAC TTAATTCAAT GTTTTTGTTT 280
Õ38.2 TFL1-RC AGAAAAACAA ACATGAACAC TTAATTCAAT GTTTTTGTTT 280
Consensus AGAAAAACAA ACATGAACAC TTAATTCAAT GTTTTTGTTT
100%
Conservation
0%
2,0bits
Sequence logo
0,0bits
300
320
2N52. TFL1.EF193847 TTTTCTTTCT ATAATCTCAT ACAAAACAAC CTAAAGTATT 320
29H.TFL1.EF193852 TT--------- ---------- ---------- ---------- 282
Alarga TFL1.EF193848 TTTTCTTTCT ATAATCTCAT ACAAAACAAC CTAAAGTATT 320
Verde bonite.TFL1.EF193849 TTTTCTTTCT ATAATCTCAT ACAAAACAAC CTAAAGTATT 320
TicoIL.1 TFL1 TTTTCTTTCT ATAATCTCAT ACAAAACAAC CTAAAGTATT 320
Õ38.2 TFL1-RC TTTTCTTTCT ATAATCTCAT ACAAAACAAC CTAAAGTATT 320
Consensus TTTTCTTTCT ATAATCTCAT ACAAAACAAC CTAAAGTATT
100%
Conservation
0%
2,0bits
Sequence logo
0,0bits
340
360
2N52. TFL1.EF193847 TGACCTAATA TGTTTATATT ATTGTCTTCA ATTTTTTGTTT 360
29H.TFL1.EF193852 -GACCTAATA TGTTTATATT ATTGTCTTCA ATTTTTTTTTT 321
Alarga TFL1.EF193848 TGACCTAATA TGTTTATATT ATTGTCTTCA ATTTTTTGTTT 360
Verde bonite.TFL1.EF193849 TGACCTAATA TGTTTATATT ATTGTCTTCA ATTTTTTGTTT 360
TicoIL.1 TFL1 TGACCTAATA TGTTTATATT ATTGTCTTCA ATTTTTTGTTT 360
Õ38.2 TFL1-RC TGACCTAATA TGTTTATATT ATTGTCTTCA ATTTTTTGTTT 360
Consensus TGACCTAATA TGTTTATATT ATTGTCTTCA ATTTTTTGTTT
100%
Conservation
0%
2,0bits
Sequence logo
0,0bits
380
400
2N52. TFL1.EF193847 ATAGGTGATG ACAGACCCTG ATGTTCCTGG CCCTAGTGAT 400
29H.TFL1.EF193852 ATAGGTGATG ACAGACCCTG ATGTTCCTGG CCCTAGTGAT 361
Alarga TFL1.EF193848 ATAGGTGATG ACAGACCCTG ATGTTCCTGG CCCTAGTGAT 400
Verde bonite.TFL1.EF193849 ATAGGTGATG ACAGACCCTG ATGTTCCTGG CCCTAGTGAT 400
TicoIL.1 TFL1 ATAGGTGATG ACAGACCCTG ATGTTCCTGG CCCTAGTGAT 400
Õ38.2 TFL1-RC ATAGGTGATG ACAGACCCTG ATGTTCCTGG CCCTAGTGAT 400
Consensus ATAGGTGATG ACAGACCCTG ATGTTCCTGG CCCTAGTGAT
100%
Conservation
0%
2,0bits
Sequence logo
0,0bits
420
440
2N52. TFL1.EF193847 CCTTATCTCA GAGAACACTT ACACTGGTAA ATATATATAT 440
29H.TFL1.EF193852 CCTTATCTCA GAGAACACTT ACACTGGTAA ATATATATAT 401
Alarga TFL1.EF193848 CCTTATCTCA GAGAACACTT ACACTGGTAA ATATATATAT 440
Verde bonite.TFL1.EF193849 CCTTATCTCA GAGAACACTT ACACTGGTAA ATATATATAT 440
TicoIL.1 TFL1 CCTTATCTCA GAGAACACTT ACACTGGTAA ATATATATAT 440
Õ38.2 TFL1-RC CTTTATCTCA GAGAACACTT ACACTGGTAA ATATATATAT 440
Consensus CCTTATCTCA GAGAACACTT ACACTGGTAA ATATATATAT
100%
Conservation
0%
2,0bits
Sequence logo
0,0bits


2N52. TFL1.EF193847 480
29H.TFL1.EF193852 441
Alarga TFL1.EF193848 480
Verde bonite.TFL1.EF193849 480
TicolL.1 TFL1 480
Ö38.2 TFL1-RC 480
Consensus AATATATATA TAACATAATG ACATACATCA TATATAGCTA
Conservation
Sequence logo
500
520
2N52. TFL1.EF193847 520
29H.TFL1.EF193852 481
Alarga TFL1.EF193848 520
Verde bonite.TFL1.EF193849 520
TicolL.1 TFL1 520
Ö38.2 TFL1-RC 520
Consensus TAATTAATTT TCTTTGATTA AATTTTAATT TATTTATTTA
Conservation
Sequence logo
540
560
2N52. TFL1.EF193847 559
29H.TFL1.EF193852 520
Alarga TFL1.EF193848 560
Verde bonite.TFL1.EF193849 560
TicolL.1 TFL1 560
Ö38.2 TFL1-RC 560
Consensus TTTATAGGAT TGTGACAGAT ATTCCAGGAA CAACAGATGC
Conservation
Sequence logo
580
600
2N52. TFL1.EF193847 599
29H.TFL1.EF193852 560
Alarga TFL1.EF193848 600
Verde bonite.TFL1.EF193849 600
TicolL.1 TFL1 600
Ö38.2 TFL1-RC 600
Consensus TACATTTGGT AGGTTATTTA ATTTGTTGAA AAACGTATAT
Conservation
Sequence logo
620
640
2N52. TFL1.EF193847 639
29H.TFL1.EF193852 600
Alarga TFL1.EF193848 640
Verde bonite.TFL1.EF193849 640
TicolL.1 TFL1 640
Ö38.2 TFL1-RC 640
Consensus ATATGGATAT AGATATAGAT ATAGAGAGAG AATGAGTATG
Conservation
Sequence logo
660
680
2N52. TFL1.EF193847 679
29H.TFL1.EF193852 640
Alarga TFL1.EF193848 680
Verde bonite.TFL1.EF193849 680
TicolL.1 TFL1 680
Ö38.2 TFL1-RC 680
Consensus ACATATTGAA TTTTTTGTAT AGGGAAAGAG ATTGTGAGCT
Conservation
Sequence logo

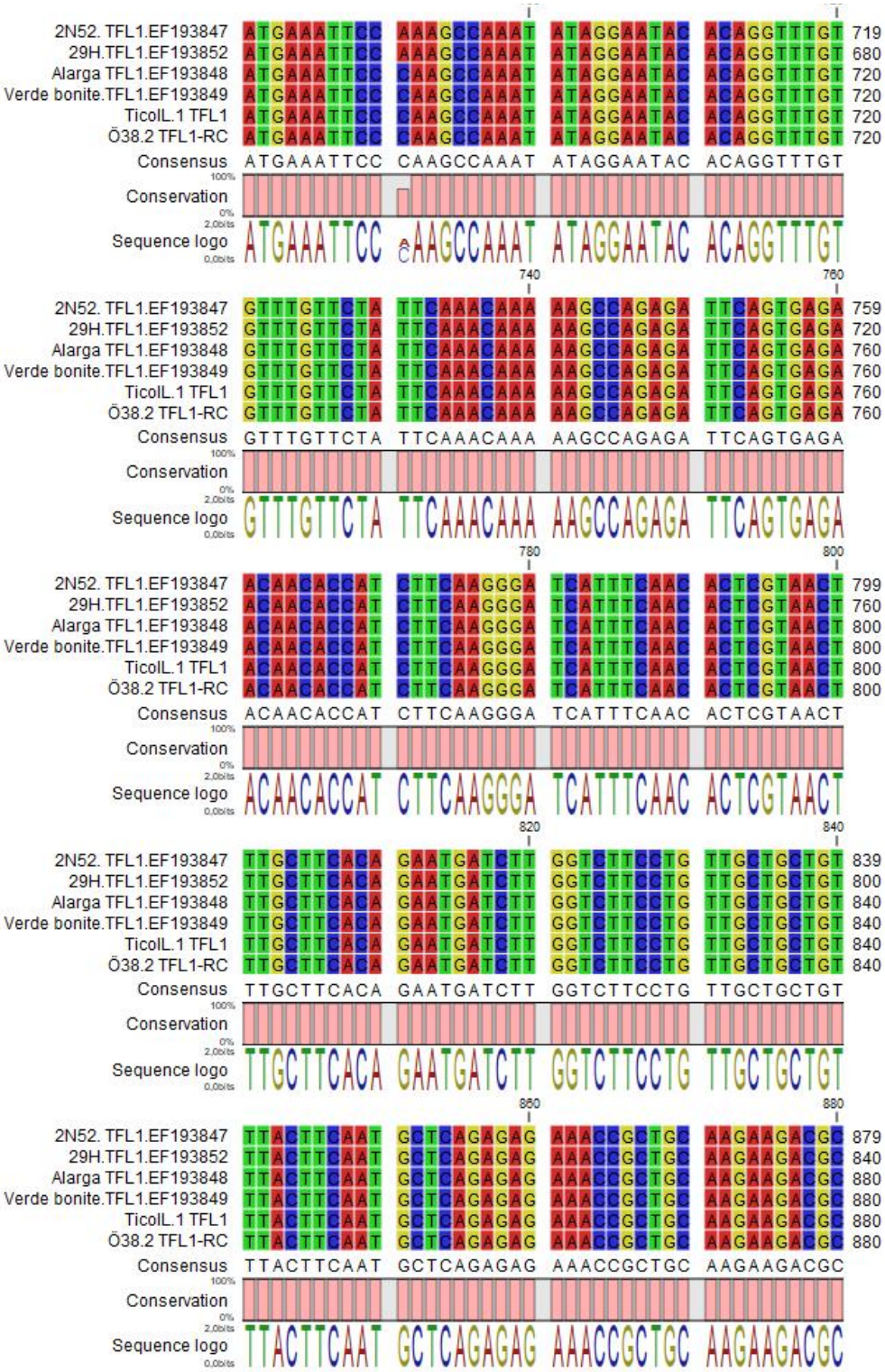


Figure S1. Alignment of whole TFL1 sequence, with Avila's et al (2007) sequences, and Ö38 and TicolL sequences from this study.

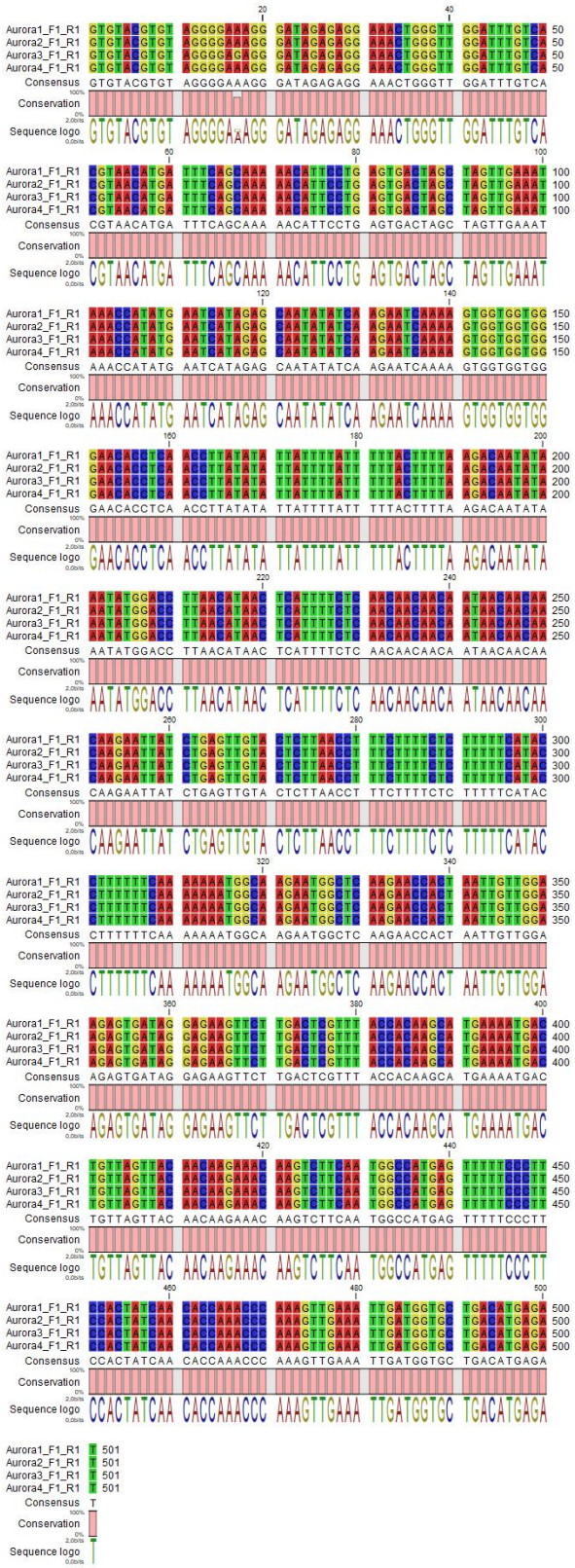


Figure S2. The four replicates of the variety Aurora. Aurora 1 was chosen as representative.

Figure S3. The four replicates of the variety Tina. Tina 2 was chosen as representative.

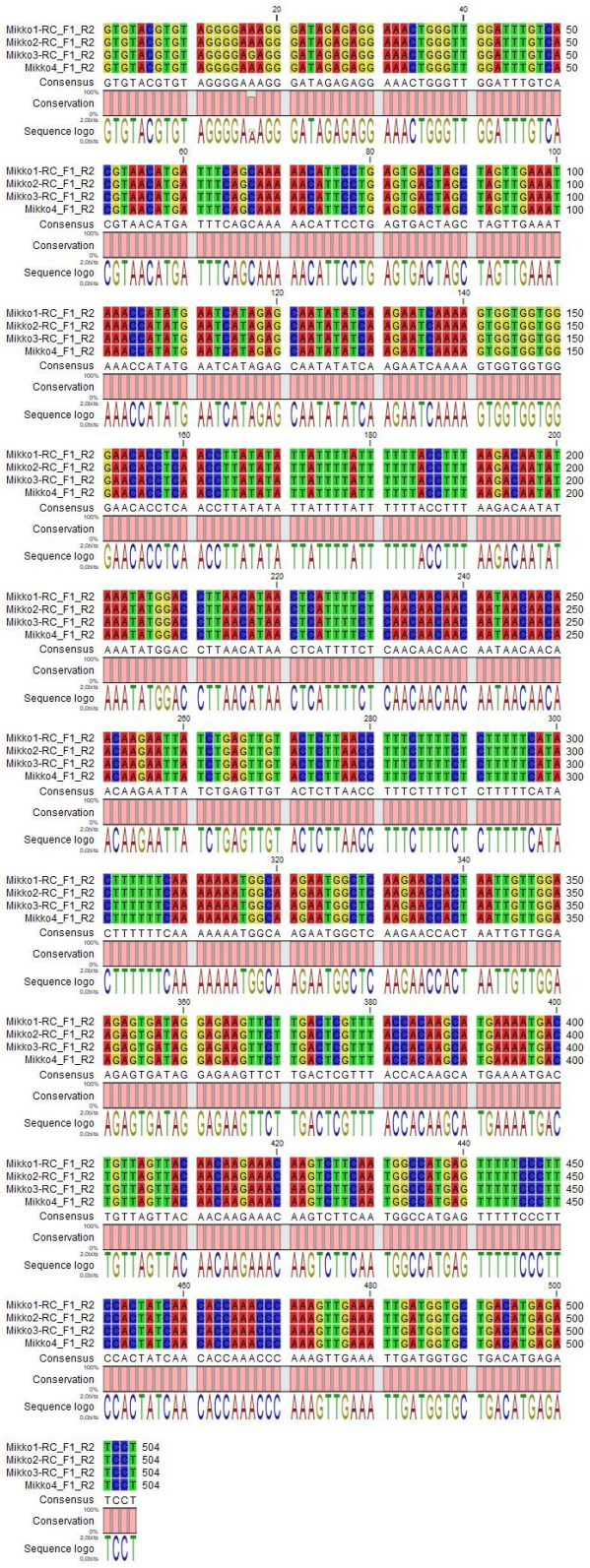


Figure S4. The four replicates for the variety Mikko. Mikko 1 was chosen as representative.

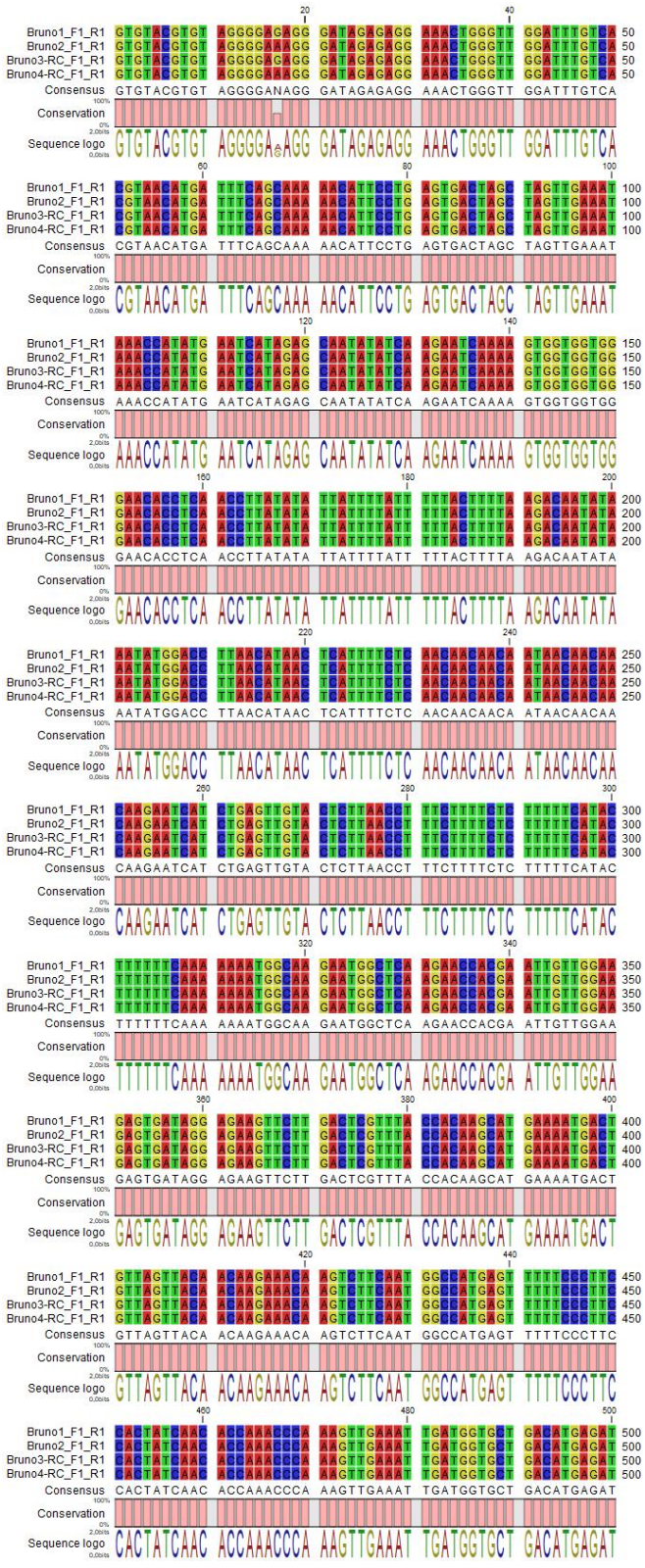


Figure S5. The four replicates of the variety Bruno. Bruno 1 was chosen as representative.

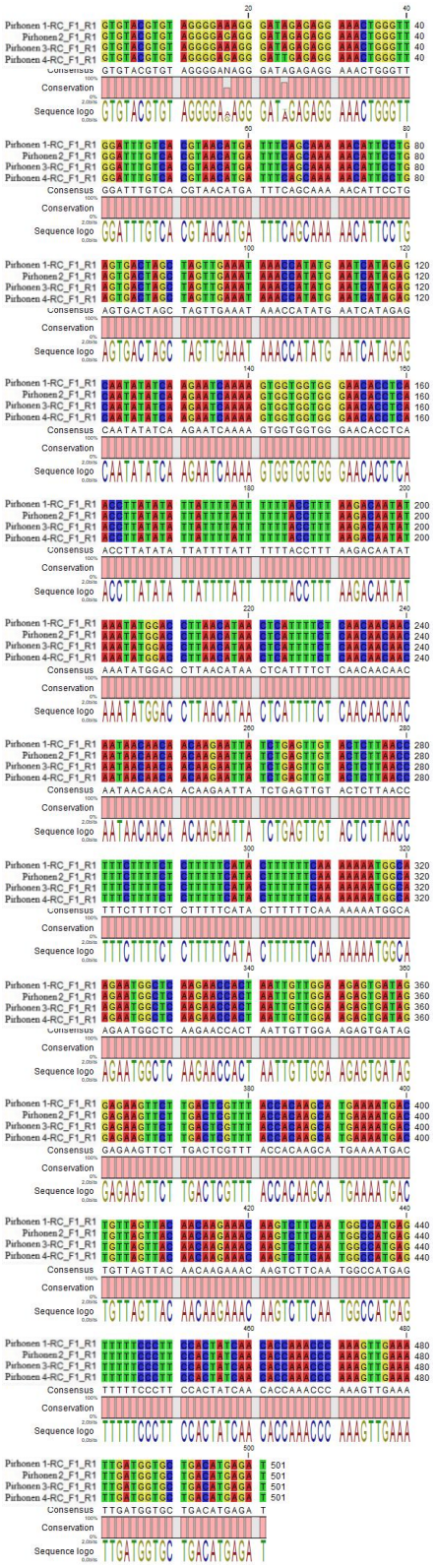


Figure S6. The four replicates of the variety Pirhonen. Pirhonen 2 was chosen as representative.

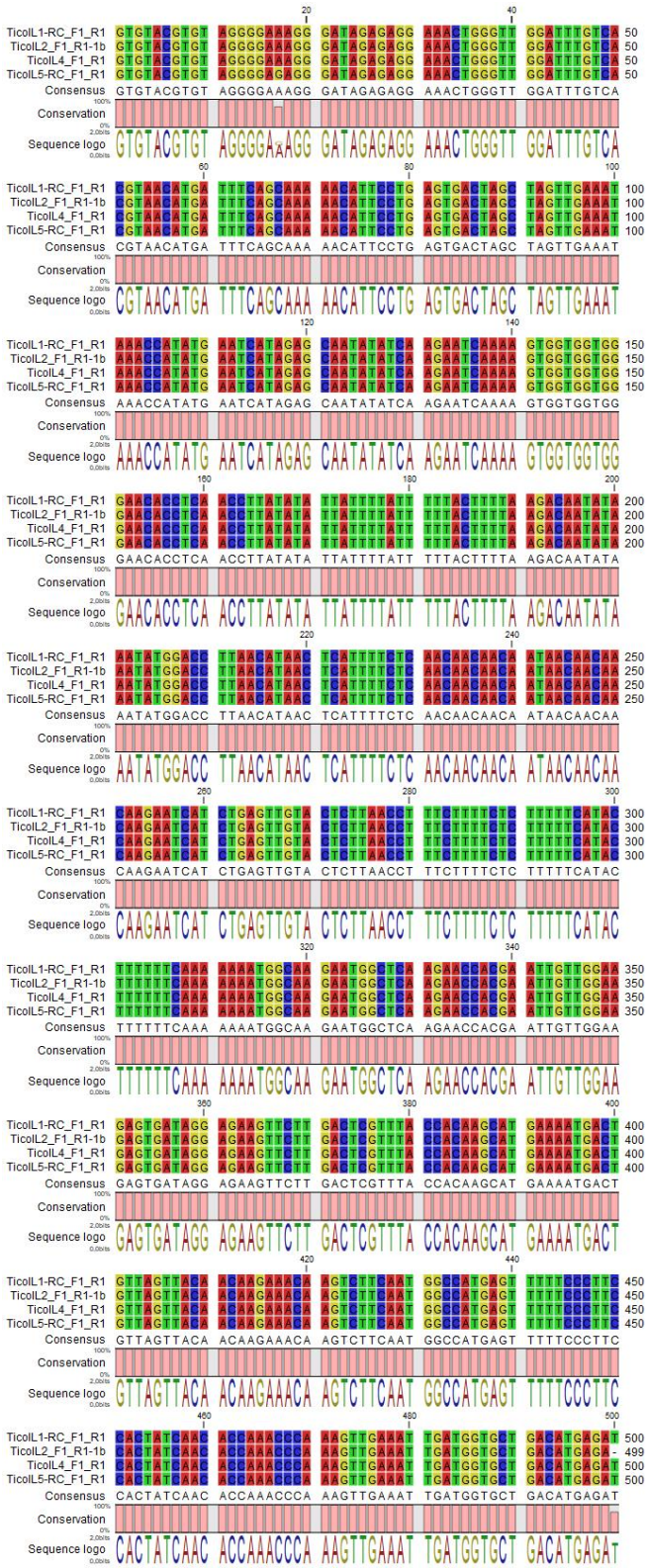


Figure S7. The four replicates of the variety TicolL. TicolL 1 was chosen as representative

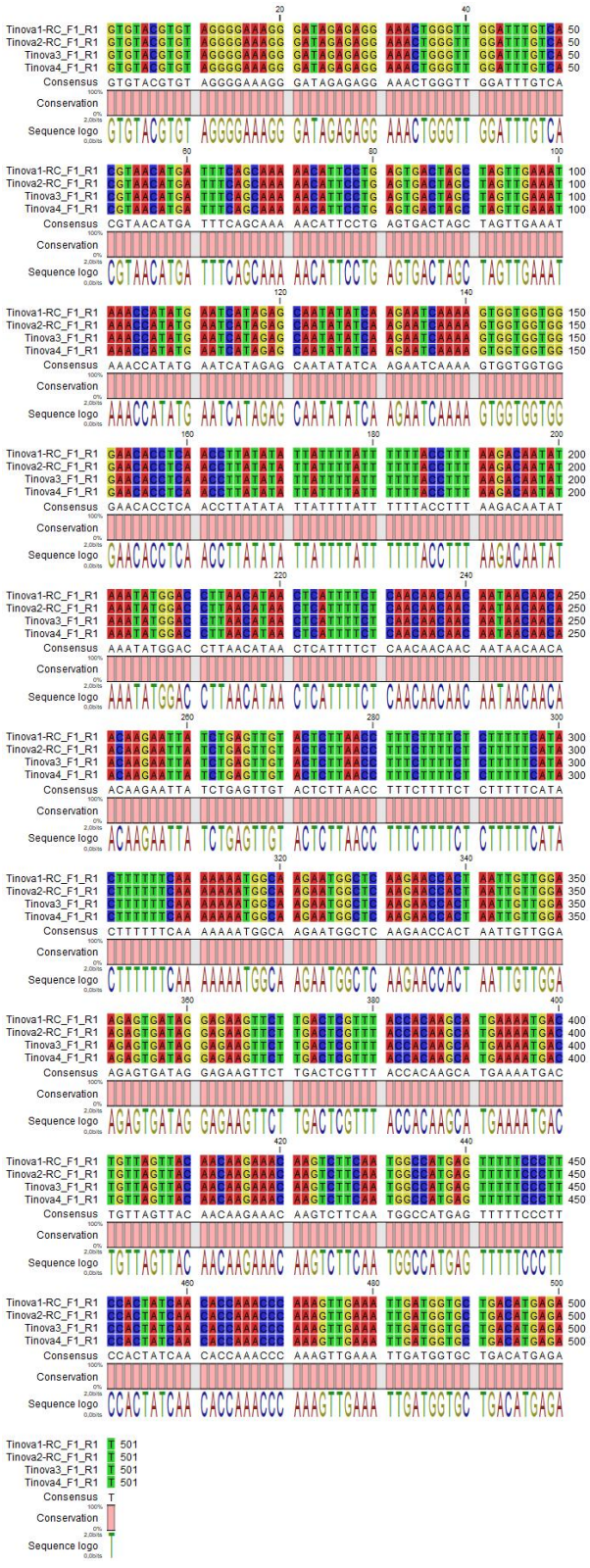


Figure S8. The four replicates of the variety Tinova. Tinova 1 was chosen as representative.

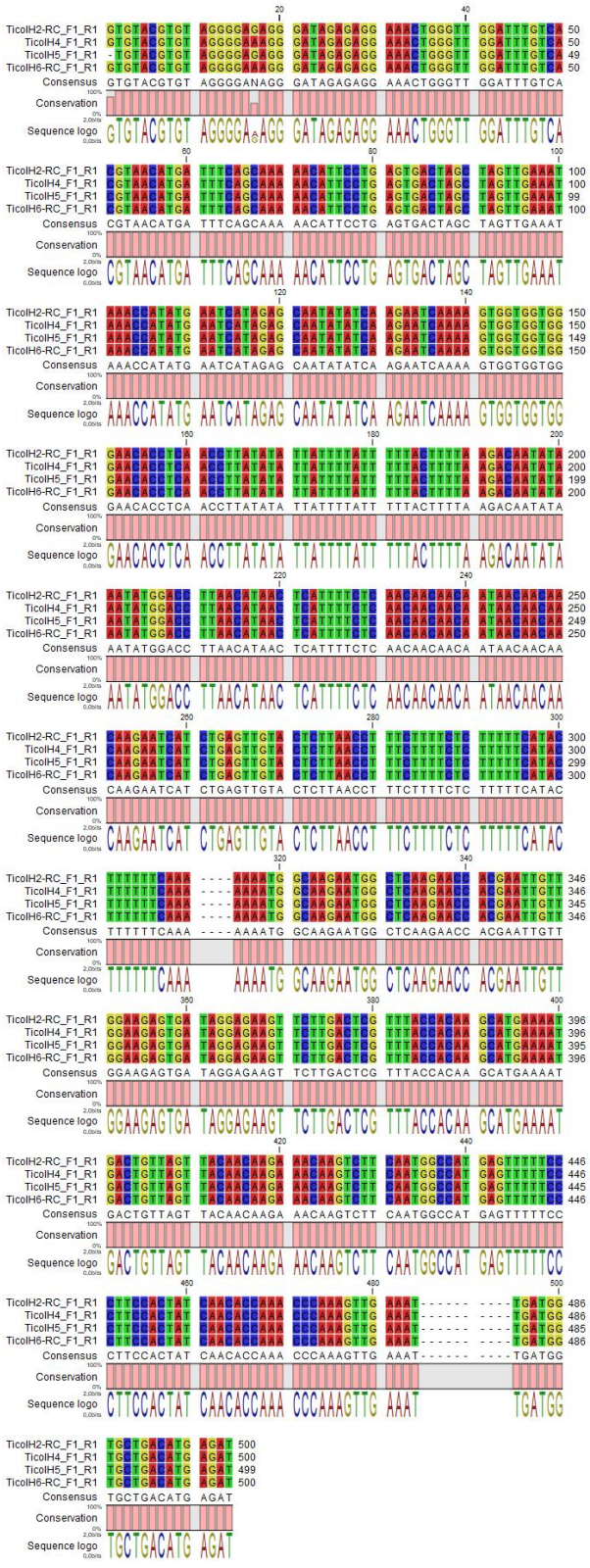


Figure S9. The four replicates of the variety TicolH. TicolH 2 was chosen as representative.

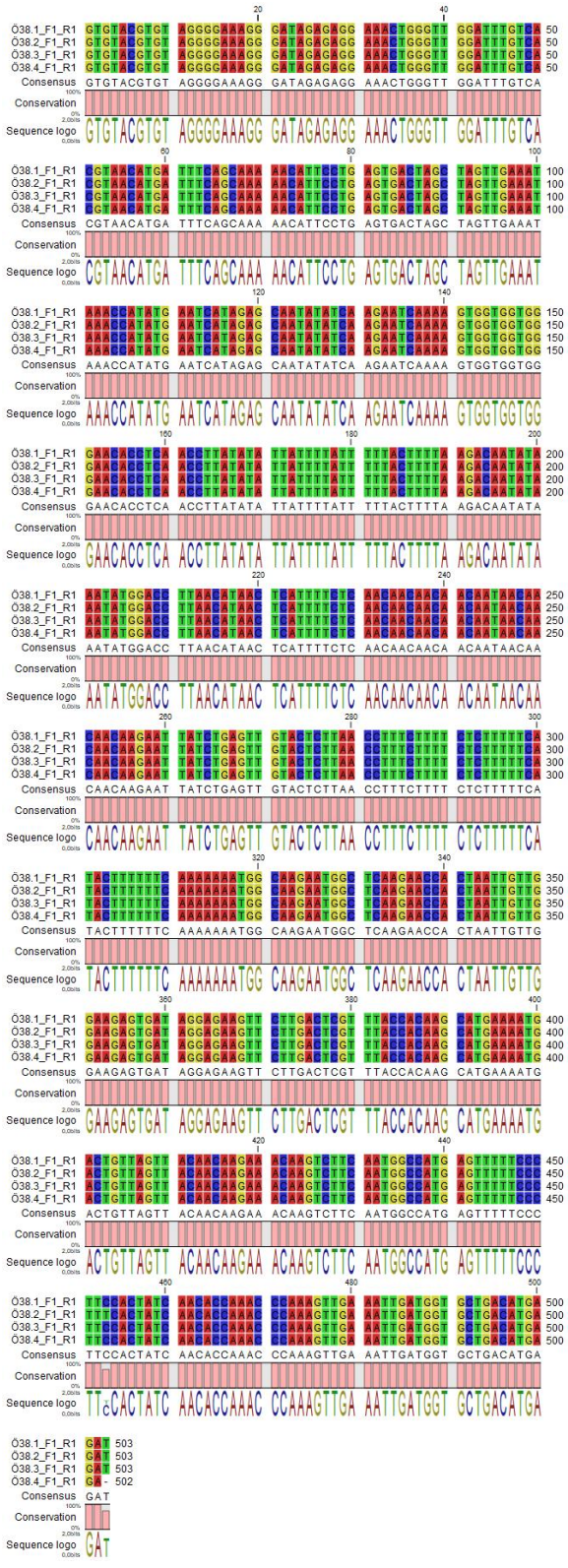


Figure S10. The four replicates of the variety Ö38. Ö38 1 was chosen as representative.

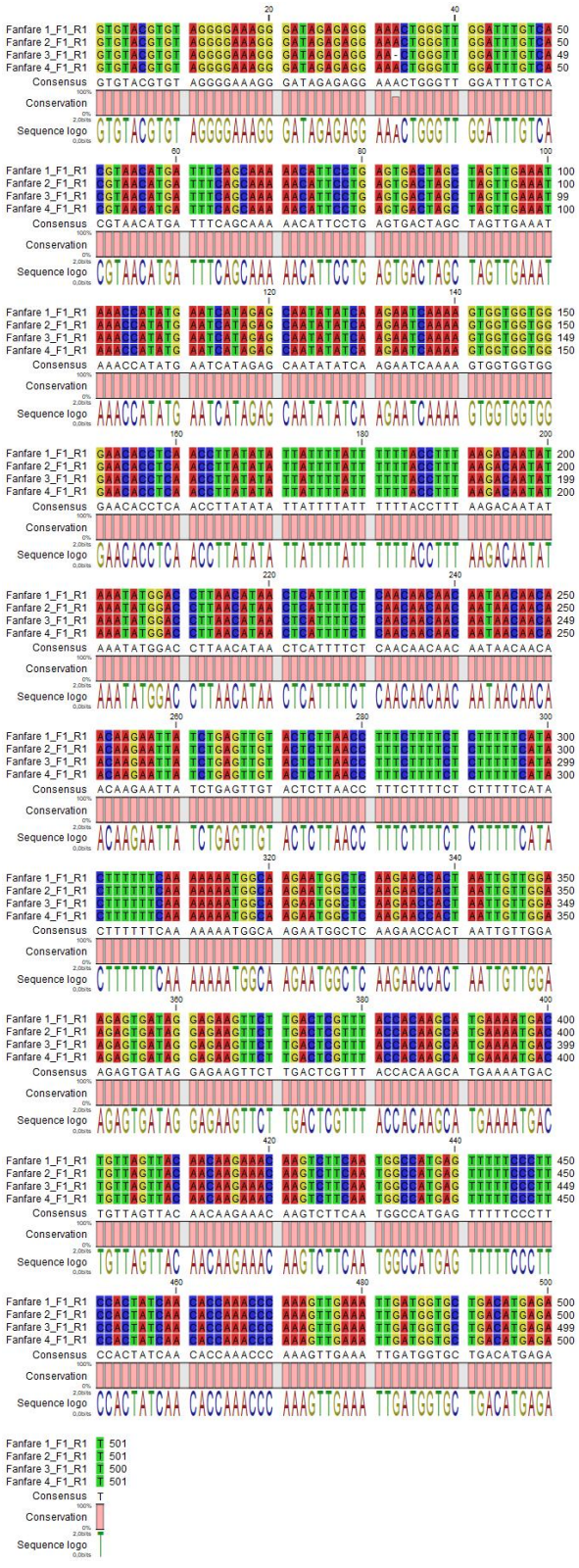


Figure S11. The four replicates of the variety Fanfare. Fanfare 1 was chosen as representative.

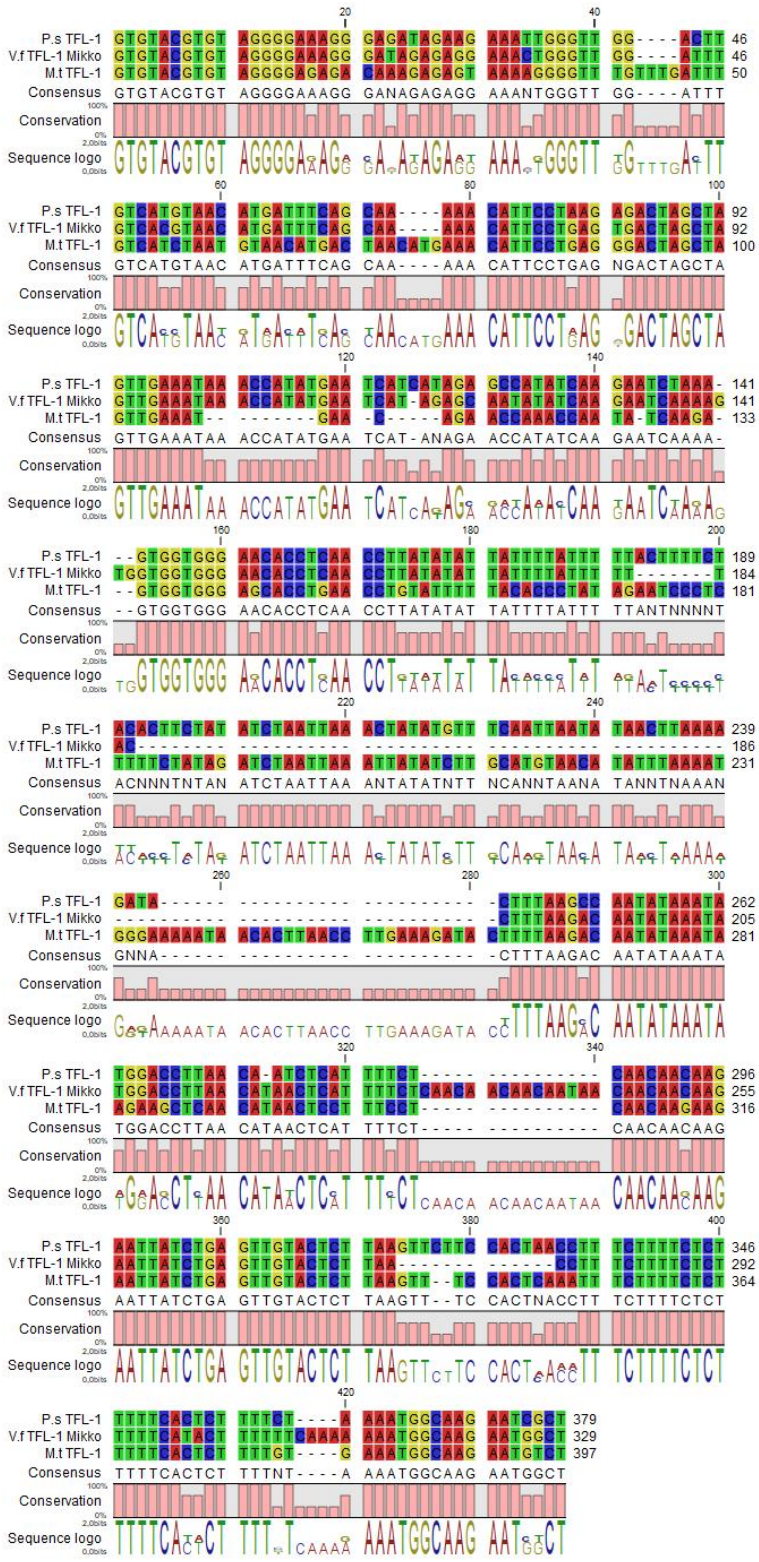


Figure S12. Upstream sequence alignment of the TFL1 homologous genes in Pisum sativum, Vicia faba cultivar Mikko, and Medicago truncatula.

Table S1. Predicted binding site motifs for the upstream region of TFL1. *Binding sites are sorted alphabetically and in color code to visualize similarities and differences.*

Colour code	*Pisum Sativum*	*Medicago truncatula*	*Vicia faba*, Mikko, Pirhonen and Fanfare	*Vicia faba,* Bruno, TicolL and TicolH	*Vicia faba,* Ö38
Green – Present in all	-300ELEMENT	-300ELEMENT	-300ELEMENT	-300ELEMENT	-300ELEMENT
	ABRELATERD1	ABRELATERD1	ABRELATERD1	ABRELATERD1	ABRELATERD1
	ACGTATERD1	ACGTATERD1	ACGTATERD1	ABRELATERD1	ABRELATERD1
	ACGTATERD1	ACGTATERD1	ACGTATERD1	ACGTATERD1	ACGTATERD1
	ACGTOSGLUB1	ACGTOSGLUB1	ACGTATERD1	ACGTATERD1	ACGTATERD1
	ARR1AT	ARR1AT	ACGTATERD1	ACGTATERD1	ACGTATERD1
	ARR1AT	ARR1AT	ACGTOSGLUB1	ACGTATERD1	ACGTATERD1
	ARR1AT	BIHD1OS	ARR1AT	ACGTOSGLUB1	ACGTOSGLUB1
	ARR1AT	CAATBOX1	ARR1AT	ARR1AT	ARR1AT
	ARR1AT	CAATBOX1	ARR1AT	ARR1AT	ARR1AT
	BIHD1OS	CACTFTPPCA1	ARR1AT	ARR1AT	ARR1AT
	CAATBOX1	CACTFTPPCA1	BIHD1OS	ARR1AT	ARR1AT
	CAATBOX1	CACTFTPPCA1	CAATBOX1	ARR1AT	BIHD1OS
	CAATBOX1	CACTFTPPCA1	CAATBOX1	BIHD1OS	CAATBOX1
	CAATBOX1	CACTFTPPCA1	CAATBOX1	CAATBOX1	CAATBOX1
	CACTFTPPCA1	CACTFTPPCA1	CACTFTPPCA1	CAATBOX1	CAATBOX1
	CACTFTPPCA1	CACTFTPPCA1	CACTFTPPCA1	CAATBOX1	CACTFTPPCA1
	CACTFTPPCA1	CAREOSREP1	CACTFTPPCA1	CACTFTPPCA1	CACTFTPPCA1
	CACTFTPPCA1	CURECORECR	CACTFTPPCA1	CACTFTPPCA1	CACTFTPPCA1
	CACTFTPPCA1	CURECORECR	CAREOSREP1	CACTFTPPCA1	CACTFTPPCA1
	CACTFTPPCA1	CURECORECR	CURECORECR	CACTFTPPCA1	CACTFTPPCA1
	CACTFTPPCA1	CURECORECR	CURECORECR	CACTFTPPCA1	CAREOSREP1
	CAREOSREP1	DOFCOREZM	CURECORECR	CAREOSREP1	CURECORECR
	CURECORECR	DOFCOREZM	CURECORECR	CURECORECR	CURECORECR
	CURECORECR	DOFCOREZM	DOFCOREZM	CURECORECR	CURECORECR
	CURECORECR	DOFCOREZM	DOFCOREZM	CURECORECR	CURECORECR
	CURECORECR	DOFCOREZM	DOFCOREZM	CURECORECR	DOFCOREZM
	DOFCOREZM	DOFCOREZM	DOFCOREZM	DOFCOREZM	DOFCOREZM
	DOFCOREZM	DOFCOREZM	DOFCOREZM	DOFCOREZM	DOFCOREZM
	DOFCOREZM	DOFCOREZM	DOFCOREZM	DOFCOREZM	DOFCOREZM

	DOFCOREZM	DOFCOREZM	DOFCOREZM	DOFCOREZM	DOFCOREZM
	DOFCOREZM	EBOXBNNAPA	EBOXBNNAPA	DOFCOREZM	DOFCOREZM
	DOFCOREZM	EBOXBNNAPA	EBOXBNNAPA	DOFCOREZM	DOFCOREZM
	DOFCOREZM	GATABOX	GATABOX	EBOXBNNAPA	EBOXBNNAPA
	DOFCOREZM	GATABOX	GATABOX	EBOXBNNAPA	EBOXBNNAPA
	DOFCOREZM	GATABOX	GATABOX	EBOXBNNAPA	GATABOX
	EBOXBNNAPA	GATABOX	GT1CONSENSUS	EBOXBNNAPA	GATABOX
	EBOXBNNAPA	GT1CONSENSUS	GT1CONSENSUS	GATABOX	GATABOX
	GATABOX	GT1CONSENSUS	GT1CONSENSUS	GATABOX	GT1CONSENSUS
	GATABOX	GT1CONSENSUS	GT1CONSENSUS	GT1CONSENSUS	GT1CONSENSUS
	GATABOX	GT1CONSENSUS	GT1CONSENSUS	GT1CONSENSUS	GT1CONSENSUS
	GATABOX	GT1GMSCAM4	GT1GMSCAM4	GT1CONSENSUS	GT1CONSENSUS
	GATABOX	GT1GMSCAM4	GT1GMSCAM4	GT1GMSCAM4	GT1GMSCAM4
	GT1CONSENSUS	GTGANTG10	GTGANTG10	GT1GMSCAM4	GT1GMSCAM4
	GT1CONSENSUS	GTGANTG10	GTGANTG10	GTGANTG10	GTGANTG10
	GT1CONSENSUS	INRNTPSADB	INRNTPSADB	GTGANTG10	GTGANTG10
	GT1GMSCAM4	INRNTPSADB	MYB1AT	INRNTPSADB	INRNTPSADB
	GTGANTG10	MYB1AT	MYCCONSENSUSAT	MYB1AT	MYB1AT
	INRNTPSADB	MYCCONSENSUSAT	MYCCONSENSUSAT	MYCCONSENSUSAT	MYCCONSENSUSAT
	INRNTPSADB	MYCCONSENSUSAT	NODCON2GM	MYCCONSENSUSAT	MYCCONSENSUSAT
	MYB1AT	NODCON2GM	NODCON2GM	MYCCONSENSUSAT	NODCON2GM
	MYCCONSENSUSAT	NODCON2GM	OSE2ROOTNODULE	MYCCONSENSUSAT	NODCON2GM
	MYCCONSENSUSAT	NODCON2GM	OSE2ROOTNODULE	NODCON2GM	OSE2ROOTNODULE
	NODCON2GM	NODCON2GM	POLLEN1LELAT52	NODCON2GM	OSE2ROOTNODULE
	NODCON2GM	NODCON2GM	POLLEN1LELAT52	OSE2ROOTNODULE	POLLEN1LELAT52
	NODCON2GM	NODCON2GM	POLLEN1LELAT52	OSE2ROOTNODULE	POLLEN1LELAT52
	NODCON2GM	OSE2ROOTNODULE	RAV1AAT	POLLEN1LELAT52	POLLEN1LELAT52
	OSE2ROOTNODULE	OSE2ROOTNODULE	RAV1AAT	POLLEN1LELAT52	RAV1AAT
	OSE2ROOTNODULE	OSE2ROOTNODULE	RAV1AAT	POLLEN1LELAT52	RAV1AAT
	OSE2ROOTNODULE	OSE2ROOTNODULE	RAV1AAT	RAV1AAT	RAV1AAT
	OSE2ROOTNODULE	OSE2ROOTNODULE	RAV1AAT	RAV1AAT	RAV1AAT
	POLLEN1LELAT52	OSE2ROOTNODULE	ROOTMOTIFTAPOX1	RAV1AAT	RAV1AAT
	POLLEN1LELAT52	POLLEN1LELAT52	ROOTMOTIFTAPOX1	RAV1AAT	RAV1AAT
	POLLEN1LELAT52	POLLEN1LELAT52	ROOTMOTIFTAPOX1	RAV1AAT	ROOTMOTIFTAPOX1
	POLLEN1LELAT52	POLLEN1LELAT52	ROOTMOTIFTAPOX1	ROOTMOTIFTAPOX1	ROOTMOTIFTAPOX1
	POLLEN1LELAT52	RAV1AAT	ABRELATERD1	ROOTMOTIFTAPOX1	ROOTMOTIFTAPOX1
	POLLEN1LELAT52	RAV1AAT	SEF1MOTIF	ROOTMOTIFTAPOX1	ROOTMOTIFTAPOX1
	RAV1AAT	ROOTMOTIFTAPOX1	TATABOX2	ROOTMOTIFTAPOX1	SEF1MOTIF

	RAV1AAT ROOTMOTIFTAPOX1 ROOTMOTIFTAPOX1 ROOTMOTIFTAPOX1 ROOTMOTIFTAPOX1 SEF1MOTIF TATABOX2 TATABOX5 TATABOX5 TATABOX5 WRKY71OS	ROOTMOTIFTAPOX1 ROOTMOTIFTAPOX1 SEF1MOTIF TATABOX2 TATABOX5 TATABOX5 WRKY71OS WRKY71OS	TATABOX5 TATABOX5 TATABOX5 WRKY71OS WRKY71OS	SEF1MOTIF TATABOX2 TATABOX5 TATABOX5 TATABOX5 WRKY71OS WRKY71OS	TATABOX2 TATABOX5 TATABOX5 TATABOX5 WRKY71OS WRKY71OS
Light blue- Present in *Vicia faba* varieties and *Pisum sativum*	CATATGGMSAUR CATATGGMSAUR MARTBOX MARTBOX MYBPZM POLASIG1 POLASIG1 SEF3MOTIFGM TATABOX4 IBOXCORE TAAAGSTKST1 TAAAGSTKST1 SEF4MOTIFGM7S SEF4MOTIFGM7S BOXIINTPATPB BOXIINTPATPB CCA1ATLHCB1 CCAATBOX1 CCAATBOX1 GT1MOTIFPSRBCS NODCON1GM NTBBF1ARROLB NTBBF1ARROLB OSE1ROOTNODULE POLASIG2 POLASIG3		MARTBOX MARTBOX MYBPZM POLASIG1 POLASIG1 POLASIG3 SEF3MOTIFGM TATABOX4 CATATGGMSAUR CATATGGMSAUR	MARTBOX MARTBOX MYBPZM POLASIG1 POLASIG1 POLASIG3 SEF3MOTIFGM TATABOX4 CATATGGMSAUR CATATGGMSAUR	MARTBOX MARTBOX MYBPZM POLASIG1 POLASIG1 POLASIG3 SEF3MOTIFGM TATABOX4 CATATGGMSAUR CATATGGMSAUR

	SURECOREATSULTR11 TATABOX3 CATATGGMSAUR CATATGGMSAUR MARTBOX MARTBOX MYBPZM POLASIG1 POLASIG1 SEF3MOTIFGM TATABOX4				
Yellow- Present in *Vicia faba* and *Medicago truncatula*		GT1CORE SEBFCONSSTPR10A WBOXHVISO1 WBOXNTERF3	GT1CORE SEBFCONSSTPR10A WBOXHVISO1 WBOXNTERF3	GT1CORE SEBFCONSSTPR10A WBOXHVISO1 WBOXNTERF3	GT1CORE SEBFCONSSTPR10A WBOXHVISO1 WBOXNTERF3
Orange- Present only in *Vicia faba* varieties			MYBST1 PREATPRODH RBCSCONSENSUS RHERPATEXPA7	MYBST1 PREATPRODH RBCSCONSENSUS RHERPATEXPA7	MYBST1 PREATPRODH RBCSCONSENSUS RHERPATEXPA7
Red- Present only in Mikko, Pirhonen, Fanfare Tina, Tinova, Aurora, Ö38, *P,sativum* and *M.truncatula*	IBOXCORE	IBOXCORE	IBOXCORE		IBOXCORE
Blue - Present in *P.sativum*,	TAAAGSTKST1 TAAAGSTKST1		TAAAGSTKST1		

Mikko, Pirhonen, Fanfare, Tina and Tinova					
Grey- Present in all but Tina and Tinova	SEF4MOTIFGM7S SEF4MOTIFGM7S	SEF4MOTIFGM7S	SEF4MOTIFGM7S	SEF4MOTIFGM7S SEF4MOTIFGM7S	SEF4MOTIFGM7S SEF4MOTIFGM7S
Present only in *Pisum Sativum*	BOXIINTPATPB BOXIINTPATPB CCA1ATLHCB1 CCAATBOX1 CCAATBOX1 GT1MOTIFPSRBCS NODCON1GM NTBBF1ARROLB NTBBF1ARROLB OSE1ROOTNODULE POLASIG2 POLASIG3 SURECOREATSULTR11 TATABOX3				
Present only in *Medicago truncatula*		ANAERO1CONSENSUS ARFAT BOXCPSAS1 BOXIINTPATPB BOXIINTPATPB CCAATBOX1 GT1MOTIFPSRBCS HSELIKENTACIDICPR1 HSELIKENTACIDICPR1 MYBPLANT NODCON1GM OSE1ROOTNODULE POLASIG2 PYRIMIDINEBOXOSRAMY1A			

		RAV1BAT REALPHALGLHCB21 REALPHALGLHCB21 RYREPEATBNNAPA SURECOREATSULTR11 TATABOXOSPAL AACACOREOSGLUB1			

Colour code	*Vicia faba*. Tina and Tinova	*Vicia faba,* Aurora
Green – Present in all	-300ELEMENT ABRELATERD1 ACGTATERD1 ACGTATERD1 ACGTATERD1 ACGTATERD1 ACGTOSGLUB1 ARR1AT ARR1AT ARR1AT ARR1AT BIHD1OS CAATBOX1 CAATBOX1 CAATBOX1 CACTFTPPCA1 CACTFTPPCA1 CACTFTPPCA1 CACTFTPPCA1 CAREOSREP1 CURECORECR CURECORECR	-300ELEMENT ABRELATERD1 ABRELATERD1 ACGTATERD1 ACGTATERD1 ACGTATERD1 ACGTATERD1 ACGTOSGLUB1 ARR1AT ARR1AT ARR1AT ARR1AT BIHD1OS CAATBOX1 CAATBOX1 CAATBOX1 CACTFTPPCA1 CACTFTPPCA1 CACTFTPPCA1 CACTFTPPCA1 CAREOSREP1 CURECORECR

	CURECORECR	CURECORECR
	CURECORECR	CURECORECR
	DOFCOREZM	CURECORECR
	DOFCOREZM	DOFCOREZM
	DOFCOREZM	DOFCOREZM
	DOFCOREZM	DOFCOREZM
	DOFCOREZM	DOFCOREZM
	DOFCOREZM	DOFCOREZM
	DOFCOREZM	DOFCOREZM
	EBOXBNNAPA	DOFCOREZM
	EBOXBNNAPA	EBOXBNNAPA
	GATABOX	EBOXBNNAPA
	GATABOX	GATABOX
	GATABOX	GATABOX
	GT1CONSENSUS	GATABOX
	GT1CONSENSUS	GT1CONSENSUS
	GT1CONSENSUS	GT1CONSENSUS
	GT1CONSENSUS	GT1CONSENSUS
	GT1GMSCAM4	GT1CONSENSUS
	GTGANTG10	GT1GMSCAM4
	GTGANTG10	GT1GMSCAM4
	INRNTPSADB	GTGANTG10
	MYB1AT	GTGANTG10
	MYCCONSENSUSAT	INRNTPSADB
	MYCCONSENSUSAT	MYB1AT
	NODCON2GM	MYCCONSENSUSAT
	NODCON2GM	MYCCONSENSUSAT
	OSE2ROOTNODULE	NODCON2GM
	OSE2ROOTNODULE	NODCON2GM
	POLLEN1LELAT52	OSE2ROOTNODULE
	POLLEN1LELAT52	OSE2ROOTNODULE
	POLLEN1LELAT52	POLLEN1LELAT52
	RAV1AAT	POLLEN1LELAT52
	RAV1AAT	POLLEN1LELAT52
	RAV1AAT	RAV1AAT
	RAV1AAT	RAV1AAT
	RAV1AAT	RAV1AAT
	ROOTMOTIFTAPOX1	RAV1AAT

	ROOTMOTIFTAPOX1 ROOTMOTIFTAPOX1 ROOTMOTIFTAPOX1 SEF1MOTIF ABRELATERD1 TATABOX2 TATABOX5 TATABOX5 TATABOX5 WRKY71OS WRKY71OS	RAV1AAT ROOTMOTIFTAPOX1 ROOTMOTIFTAPOX1 ROOTMOTIFTAPOX1 ROOTMOTIFTAPOX1 SEF1MOTIF TATABOX2 TATABOX5 TATABOX5 TATABOX5 WRKY71OS WRKY71OS
Light blue- Present in *Vicia faba* varieties and *Pisum sativum*	CATATGGMSAUR CATATGGMSAUR MARTBOX MARTBOX MYBPZM SEF3MOTIFGM TATABOX4 POLASIG1 POLASIG1 POLASIG3	CATATGGMSAUR CATATGGMSAUR MARTBOX MARTBOX MYBPZM POLASIG1 POLASIG1 POLASIG3 SEF3MOTIFGM TATABOX4
Yellow- Present in *Vicia faba* and *Medicago truncatula*	GT1CORE SEBFCONSSTPR10A WBOXHVISO1 WBOXNTERF3	GT1CORE SEBFCONSSTPR10A WBOXHVISO1 WBOXNTERF3
Orange- Present only in *Vicia faba* varieties	MYBST1 PREATPRODH RBCSCONSENSUS RHERPATEXPA7	MYBST1 PREATPRODH RBCSCONSENSUS RHERPATEXPA7
Red- Present only in Mikko, Pirhonen, Fanfare, Tina, Tinova, Aurora, Ö38, *P,sativum* and *M.truncatula*	IBOXCORE	IBOXCORE

Blue - Present in *P.sativum*, Mikko, Pirhonen, Fanfare, Tina and Tinova	TAAAGSTKST1	
Grey- Present in all but Tina and Tinova		SEF4MOTIFGM7S SEF4MOTIFGM7S
Black- Present only in Aurora		PYRIMIDINEBOXOSRAMY1A